the review of income and wealth

Review of Income and Wealth
Series 65, Number 3, September 2019
DOI: 10.1111/roiw.12402

LEXICAL MEASURES OF SOCIAL INEQUALITY: FROM PIGOU-DALTON TO HAMMOND

BY KUI OU-YANG*

School of Economics and Management, Northwest University, China

A measure of social inequality is essentially a rational ordering over a space of social distributions. However, different measures, including the most popular ones, may provide very different rankings over the same set of typical distributions. We thus propose an axiomatic approach to inequality measurement mainly based on the Hammond principle, a natural generalization of the Pigou-Dalton principle, attempting to clarify the true nature of social inequality: the rich get richer and the poor get poorer. Under the standard assumptions of anonymity and scale independence, we show that a social inequality ordering is the leximinimax measure if and only if it satisfies the first Hammond principle, and it is the leximaximin measure if and only if it satisfies the second Hammond principle.

JEL Codes: D63

Keywords: inequality measurement, the Pigou-Dalton principle, the Hammond principle, leximinimax, leximaximin

1. INTRODUCTION

A measure of social inequality is essentially a rational ordering over a space of social distributions. However, different measures, including the most popular ones, may provide very different rankings over the same set of typical distributions. For example, consider three income distributions ***a*** = (10, 40, 50), ***b*** = (20, 20, 60), and ***c*** = (100/3, 100/3, 100/3). The three famous measures of inequality, i.e., the Gini index (Gini, 1912), the Theil index (Theil, 1967), and the coefficient of variation, are calculated in Table 1.

While all indices consider ***c*** = (100/3, 100/3, 100/3) the lest unequal, the Gini index considers ***a*** = (10, 40, 50) as unequal as ***b*** = (20, 20, 60), the Theil index considers ***b*** less unequal than ***a***, and the coefficient of variation considers ***a*** less unequal than ***b***. Which index should we believe in the end? Or perhaps naively, what is inequality and how to measure it?

As pointed out by Foster (1985), there are two common approaches to deal with the paradox of measuring inequality. The first approach is to find a partial ranking that some large class of measures all would be consistent with; for

Note: The author would like to thank the editor and two anonymous referees for their helpful suggestions and comments. The support from the Key Research Foundation of Humanities and Social Sciences of the Ministry of Education of China (16JJD790046) is gratefully acknowledged. Any remaining errors are the responsibility of the author.

*Correspondence to: Kui Ou-Yang, School of Economics and Management, Northwest University, Chang'an District, Xi'an, Shaanxi, China (byloyk@gmail.com).

TABLE 1
DIFFICULTY IN MEASURING INEQUALITY

Measures	Formulas	***a***	***b***	***c***	Rankings
Gini index	$\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\lvert x_i - x_j\rvert}{2n(n-1)\mu}$	.4000	.4000	0	$c \prec \boldsymbol{a} \sim \boldsymbol{b}$
Theil index	$\frac{1}{n\ln n}\sum_{i=1}^{n}\frac{x_i}{\mu}\ln\frac{x_i}{\mu}$	.1413	.1350	0	$c \prec \mathbf{b} \prec \boldsymbol{a}$
Coefficient of variation	$\sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}\left(1-\frac{x_i}{\mu}\right)^2}$	.3606	.4000	0	$c \prec \boldsymbol{a} \prec \boldsymbol{b}$

example, the Lorenz criterion, an incomplete ranking, is generally accepted as an unambiguous principle for inequality comparisons and satisfied by most types of reasonable inequality measures, including the Gini index, the Theil index, and the coefficient of variation. The second approach is to axiomatically characterize a reasonable inequality measure by a relevant set of axioms elucidating the nature of measuring inequality. Those axioms, of course, should better be intuitive, ethically justifiable, mathematically tractable, and empirically implementable (Cowell, 2000, 2016). The following three principles are fairly standard in the theory of inequality measurement.

Anonymity: Permutation of income distribution should not change the degree of inequality, i.e. inequality measures should be symmetrical between social members. For example, $\boldsymbol{a}$ = (10, 40, 50) should be as unequal as both (40, 50, 10) and (50, 10, 40).

Scale Independence: Equally proportionate changes in individual incomes should not change the degree of inequality. Thus $\boldsymbol{d}$ = (6, 24, 30) should be as unequal as $\boldsymbol{a}$ = (10, 40, 50), and $\boldsymbol{e}$ = (10, 10, 30) should be as unequal as $\boldsymbol{b}$ = (20, 20, 60).

Pigou-Dalton Principle (Pigou, 1912; Dalton, 1920): A mean-preserving transfer of income from a person to a richer one should increase the degree of inequality. Hence (10, 30, 60) should be more unequal than both $\boldsymbol{a}$ = (10, 40, 50) and $\boldsymbol{b}$ = (20, 20, 60).

The Pigou-Dalton principle elucidates the true nature of inequality: it gets more unequal when the rich get richer and the poor get poorer; but it only applies when two distributions have the same mean. When two distributions have unequal means, the Pigou-Dalton principle can be extended into the following principle.

Hammond Principle (Hammond, 1976): Increasing the income of a richer person and decreasing the income of a poorer person should increase the degree of inequality. Therefore, $\boldsymbol{e}$ = (10, 10, 30) must be less unequal than $\boldsymbol{d}$ = (6, 24, 30), which cannot be derived from the Pigou-Dalton principle.

Now, we claim that $\boldsymbol{b}$ = (20, 20, 60) is less unequal than $\boldsymbol{a}$ = (10, 40, 50). By the Hammond principle, $\boldsymbol{e}$ = (10, 10, 30) must be less unequal than $\boldsymbol{d}$ = (6, 24, 30). By scale independence, $\boldsymbol{d}$ is as unequal as $\boldsymbol{a}$, and $\boldsymbol{e}$ is as unequal as $\boldsymbol{b}$; that is, $\boldsymbol{e}$ is less unequal than $\boldsymbol{d}$ if and only if $\boldsymbol{b}$ is less unequal than $\boldsymbol{a}$. Therefore, $\boldsymbol{b}$ must be less unequal than $\boldsymbol{a}$, which means that both the Gini index and the coefficient of variation violate the Hammond principle. In fact, the Theil index must also violate the Hammond principle, since, as we will show, there is no continuous inequality measure satisfying the Hammond principle.

Furthermore, it is somewhat surprising that under the standard assumptions of anonymity and scale independence, the Hammond principle with some slight extensions is enough to capture the full nature of inequality comparisons: we can always make a rational judgement confidently about the inequality relation between any two social distributions, as long as we have some sympathy with those three assumptions, and no more assumptions are needed at all. The Hammond principle can be slightly extended to the following two principles.

First Hammond Principle: While the rich do not get poorer, decreasing the income of a poorer person (i.e. the poor get poorer) should increase the degree of inequality. Therefore, (20, 30, 50) must be less unequal than (10, 30, 50), which cannot be directly derived from the Hammond principle.

Second Hammond Principle: While the poor do not get richer, increasing the income of a richer person (i.e. the rich get richer) should increase the degree of inequality. Therefore, (20, 30, 50) must be less unequal than (20, 30, 60), which also cannot be directly derived from the Hammond principle.

Specifically, we can show that a social inequality ordering is i) the leximinimax rule if and only if it satisfies anonymity, scale independence, and the first Hammond principle; and ii) it is the leximaximin rule if and only if it satisfies anonymity, scale independence, and the second Hammond principle. In a fixed society, the Lorenz criterion can be fully characterized by anonymity, scale independence, and the Pigou-Dalton principle (Foster, 1985), and thus we can show that a social inequality ordering is i) the leximinimax rule if and only if it satisfies the Lorenz criterion and the first Hammond principle; and ii) it is the leximaximin rule if and only if it satisfies the Lorenz criterion and the second Hammond principle.

The Lorenz criterion has been generally recognized as the fundamental principle in the theory of inequality measurement (Foster and Ok, 1999). When the Lorenz curves intersect, however, the ethical foundations of popular synthetic indices, such as the Gini index, the Theil index, and the coefficient of variation, are somewhat vague; that is, when the comparison of two income distributions cannot be completed by the principle of Lorenz dominance, we cannot clearly explain the fundamental ideas about the inner nature of income comparisons. On the other hand, the Atkinson-Kolm-Sen approach (Dalton, 1920; Kolm, 1969; Atkinson, 1970; Sen, 1973; Blackorby et al., 1999; Dutta, 2002), connecting inequality measurement with social welfare, makes the measurement of inequality more controversial, since one can easily obtain different inequality measures by different choices of social welfare functions. In particular, Arrow's impossibility theorem (Arrow, 1951, 1963, 1951, 1963) makes the ideal choice of social welfare functions more problematic.

Those two complete measures, i.e. leximinimax and leximaximin, are proposed to compensate the incompleteness of the Lorenz criterion and to establish an objective ground for measuring inequality; the ethical foundation is mainly established by the Hammond principle, not by any subjective criterion of social welfare. Since the Lorenz criterion is hardly controversial, if there are any doubts about the reasonability of the lexical measures, one actually has some doubts about the Hammond principle. It seems to us, however, that it is a little hard to raise any serious objections to the Hammond principle if one really agrees that the true nature of social inequality is "the rich get richer and the poor get poorer." At first glance, one usually cannot accept the lexical measures, as their intuitive understanding of inequality have already been shaped stubbornly by the traditional inequality measures such as the Gini index, which generates "considerable cultural inertia in the field of inequality analysis" (Cowell, 2000). Of course, we do not consider the Hammond principle to be uncontroversial, and indeed almost every axiom about inequality can be challenged more or less (Kolm, 1999), but when the classic measures violate the Hammond principle, there must be some reasonable explanations for this failure.

The rest of this paper is organized as follows. Section 2 presents the basic definitions about social inequality orderings and provides an elementary proof of Foster's characterization of the Lorenz criterion. Section 3 proposes and characterizes the leximinimax and leximaximin measures and shows the inconsistency between continuity and the Hammond principle. As an illustrative example for the application of the lexical measures, income inequality in the United States (1967-2016) is computed in Section 4. We conclude in Section 5.

2. Social Inequality Orderings

The society is a finite set $N = \{1, 2, \cdots, n\}$, $n \geq 2$. The space of social distributions is denoted by $X = \{\boldsymbol{x} \in \mathbb{R}^n \mid \forall i \in N, x_i > 0\}$, where a vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ can be interpreted as a list of income levels. The space of (rational, i.e. complete and transitive) social inequality orderings over X is denoted by $O(X)$. A typical social inequality ordering in $O(X)$ is denoted by $\precsim$. As usual, $\prec$ denotes the asymmetric part of $\precsim$, and $\sim$ the symmetric part. For all $\boldsymbol{x}, \boldsymbol{y} \in X$, $\boldsymbol{x} \precsim \boldsymbol{y}$ means that $\boldsymbol{x}$ is at least as unequal as $\boldsymbol{y}$; $\boldsymbol{x} \prec \boldsymbol{y}$ means that $\boldsymbol{x}$ is strictly less unequal than $\boldsymbol{y}$; $\boldsymbol{x} \sim \boldsymbol{y}$ means that $\boldsymbol{x}$ is as unequal as $\boldsymbol{y}$.

Anonymity: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, if there is a bijection $\pi: N \rightarrow N$ such that for each $i \in N$, $x_i = y_{\pi(i)}$, then $\boldsymbol{x} \sim \boldsymbol{y}$.

Anonymity says that relabeling social members in social distributions should not change social inequality, i.e. social inequality measures should be symmetrical between social members.

Scale Independence: For each $\boldsymbol{x} \in X$, for each $t > 0$, $\boldsymbol{x} \sim t\boldsymbol{x}$.

Scale independence implies that equally proportionate change in social distributions should not change the level of inequality. Scale independence makes us consider only relative inequality measures;[1] that is, only income shares matter in the measurement of inequality. The distributions of incomes essentially represent the distributions of real social resources, and equiproportional increase of incomes or measuring incomes in pounds instead of in dollars cannot change the distribution of real social resources.

In the theory of inequality measurement, scale independence is sometimes replaced with the following similar but weaker axiom.

Scale Invariance: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, for each $t > 0$, $\boldsymbol{x} \precsim \boldsymbol{y}$ if and only if $t\boldsymbol{x} \precsim t\boldsymbol{y}$.

[1]There are, of course, absolute measures of inequality, requiring translation independence or translation invariance. In this paper, we confine our main discussion to relative measures.

Scale invariance means that equally proportionate changes in any two social distributions should not change the inequality comparison of those two distributions. As we will see, however, scale invariance is not enough to characterize the Lorenz criterion formally defined later.

Pigou-Dalton Principle: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, for each $i, j \in N$, for each $\varepsilon > 0$, if $y_i - \varepsilon = x_i \geq x_j = y_j + \varepsilon$, and for each $k \in N \setminus \{i, j\}$, $x_k = y_k$, then $x \prec y$.

The Pigou-Dalton principle states that a mean-preserving transfer from a poorer person to a richer person should increase social inequality. This is indeed the central axiom that attempts to directly describe the true nature of social inequality, since neither anonymity nor scale independence can specify any situations in which one distribution should be strictly less unequal than another. Of course, the Pigou-Dalton principle is far from a complete measure of social inequality. We can strengthen it in two directions: first, anonymity and scale independence can be incorporated into the Pigou-Dalton principle, and then the Lorenz criterion is obtained; second, we simply get rid of the mean-preserving requirement and extend the Pigou-Dalton principle to the Hammond principle.

For each $\boldsymbol{x} \in X$, there is a bijection σ_x: $N \rightarrow N$ such that (i) $x_{\sigma_x(1)} \leq x_{\sigma_x(2)} \leq \cdots \leq x_{\sigma_x(n)}$ and (ii) for each $i, j \in N$, if $i < j$ and $x_i = x_j$, then $\sigma_x^{-1}(i) < \sigma_x^{-1}(j)$. For convenience, let $x_{(i)} = x_{\sigma_x(i)}$; that is, $x_{(i)}$ denotes the income of the ith worst-off.

For any income distribution $\boldsymbol{x} \in X$, let $\varphi_k(\boldsymbol{x}) = \sum_{i=1}^{k} x_{(i)}$ and $L(\boldsymbol{x}, k) = \frac{\varphi_k(\boldsymbol{x})}{\varphi_n(\boldsymbol{x})}$ for each $k = 1, 2, \cdots, n$; that is, $L(\boldsymbol{x}, k)$ denotes the percentage of income accruing to the k worst-off individuals in $\boldsymbol{x}$. For all $\boldsymbol{x}, \boldsymbol{y} \in X$, if for all $k \in N$, $L(\boldsymbol{x}, k) \geq L(\boldsymbol{y}, k)$, then we say that $\boldsymbol{x}$ *weakly Lorenz-dominates* $\boldsymbol{y}$; if for all $k \in N$, $L(\boldsymbol{x}, k) \geq L(\boldsymbol{y}, k)$, and for some $m \in N$, $L(\boldsymbol{x}, m) > L(\boldsymbol{y}, m)$, then we say that $\boldsymbol{x}$ *strongly Lorenz-dominates* $\boldsymbol{y}$.

Lorenz Criterion: For all $\boldsymbol{x}, \boldsymbol{y} \in X$, if $\boldsymbol{x}$ weakly Lorenz-dominates $\boldsymbol{y}$ then $\boldsymbol{x} \precsim \boldsymbol{y}$, and if $\boldsymbol{x}$ strongly Lorenz-dominates $\boldsymbol{y}$ then $\boldsymbol{x} \prec \boldsymbol{y}$.

Lemma 1: (Foster, 1985): *A social inequality ordering satisfies the Lorenz criterion if and only if it satisfies anonymity*, *scale independence*, *and the Pigou-Dalton principle*.

Proof: (Only if part) It is easy to see that the Lorenz criterion implies anonymity and scale independence. To see the Pigou-Dalton principle, consider any $\boldsymbol{x}, \boldsymbol{y} \in X$ such that for some $i, j \in N$ and some $\varepsilon > 0$ if $y_i - \varepsilon = x_i \geq x_j = y_j + \varepsilon$, and for each $t \in N \setminus \{i, j\}$, $x_t = y_t$. Since the Lorenz criterion implies anonymity, we can suppose that $x_1 \leq x_2 \leq \cdots \leq x_n$ and $y_1 \leq y_2 \leq \cdots \leq y_n$, without any loss of generality. Then we have $\varphi_k(\boldsymbol{x}) = \varphi_k(\boldsymbol{y})$ for each $k < j$, $\varphi_l(\boldsymbol{x}) > \varphi_l(\boldsymbol{y})$ for each $l \in \{j, j+1, \cdots, i-1\}$, and $\varphi_m(\boldsymbol{x}) = \varphi_m(\boldsymbol{y})$ for each $m \geq i$. Since $\varphi_n(\boldsymbol{x}) = \varphi_n(\boldsymbol{y})$, $\boldsymbol{x}$ must strongly Lorenz-dominate $\boldsymbol{y}$, i.e., $\boldsymbol{x} \prec \boldsymbol{y}$.

(If part) Suppose that anonymity, scale independence, and the Pigou-Dalton principle are satisfied. Consider any $\boldsymbol{x}, \boldsymbol{y} \in X$ such that for each $m \in N$, $L(\boldsymbol{x}, m) = L(\boldsymbol{y}, m)$. Then by scale independence we have $\boldsymbol{x} \sim \boldsymbol{x}/\varphi_n(\boldsymbol{x})$ and $\boldsymbol{y} \sim \boldsymbol{y}/\varphi_n(\boldsymbol{y})$, and by anonymity we must have $\boldsymbol{x}/\varphi_n(\boldsymbol{x}) \sim \boldsymbol{y}/\varphi_n(\boldsymbol{y})$ and thus $\boldsymbol{x} \sim \boldsymbol{y}$.

Now, consider any $\boldsymbol{x}, \boldsymbol{y} \in X$ such that for each $m \in N$, $L(\boldsymbol{x}, m) \geq L(\boldsymbol{y}, m)$, and for some $t \in N$, $L(\boldsymbol{x}, t) > L(\boldsymbol{y}, t)$.

First, suppose that $\varphi_n(\boldsymbol{x}) = \varphi_n(\boldsymbol{y})$. By anonymity, we can suppose that $x_1 \leq x_2 \leq \cdots \leq x_n$ and $y_1 \leq y_2 \leq \cdots \leq y_n$, without any loss of generality. There must be some

$s(1) < t(1) \in N$ and $M^1 = \{s(1), s(1)+1, \cdots, t(1)\} \subseteq N$ such that $x_{s(1)} > y_{s(1)}$, $x_{t(1)} < y_{t(1)}$, $x_l = y_l$ for all $l < s(1)$, and $x_m = y_m$ for all $m > t(1)$. Let $\boldsymbol{a}^0 = \boldsymbol{y}$. There must be some $\boldsymbol{y}^1 \in X$ such that

(i) $y^1_{s(1)} = x_{s(1)}$ and $y^1_{t(1)} = a^0_{t(1)} - (x_{s(1)} - a^0_{s(1)})$ if $x_{s(1)} - a^0_{s(1)} < a^0_{t(1)} - x_{t(1)}$;
(ii) $y^1_{t(1)} = x_{t(1)}$ and $y^1_{s(1)} = a^0_{s(1)} + (a^0_{t(1)} - x_{t(1)})$ if $x_{s(1)} - a^0_{s(1)} > a^0_{t(1)} - x_{t(1)}$;
(iii) $y^1_{s(1)} = x_{s(1)}$ and $y^1_{t(1)} = x_{t(1)}$ if $x_{s(1)} - a^0_{s(1)} = a^0_{t(1)} - x_{t(1)}$;
(iv) $y^1_i = a^0_i$ for all $i \in N \setminus \{s(1), t(1)\}$.

By the Pigou-Dalton principle, $\boldsymbol{y}^1 \prec \boldsymbol{a}^0$. Let $\boldsymbol{a}^1$ be the permutation of $\boldsymbol{y}^1$ such that $a^1_1 \leq a^1_2 \leq \cdots \leq a^1_n$; by anonymity, $\boldsymbol{a}^1 \sim \boldsymbol{y}^1$.

Then there must be some $s(2) < t(2) \in M^1$ and $M^2 = \{s(2), s(2)+1, \cdots, t(2)\} \subseteq M^1$ such that $x_{s(2)} > a^1_{s(2)}$, $x_{t(2)} < a^1_{t(2)}$, $x_l = a^1_l$ for all $l < s(2)$, and $x_m = a^1_m$ for all $m > t(2)$. There must be some $\boldsymbol{y}^2 \in X$ such that

(i) $y^2_{s(2)} = x_{s(2)}$ and $y^2_{t(2)} = a^1_{t(2)} - (x_{s(2)} - a^1_{s(2)})$ if $x_{s(2)} - a^1_{s(2)} < a^1_{t(2)} - x_{t(2)}$;
(ii) $y^2_{t(2)} = x_{t(2)}$ and $y^2_{s(2)} = a^1_{s(2)} + (a^1_{t(2)} - x_{t(2)})$ if $x_{s(2)} - a^1_{s(2)} > a^1_{t(2)} - x_{t(2)}$;
(iii) $y^2_{s(2)} = x_{s(2)}$ and $y^2_{t(2)} = x_{t(2)}$ if $x_{s(2)} - a^1_{s(2)} = a^1_{t(2)} - x_{t(2)}$;
(iv) $y^2_i = a^1_i$ for all $i \in N \setminus \{s(2), t(2)\}$.

By the Pigou-Dalton principle, $\boldsymbol{y}^2 \prec \boldsymbol{a}^1$. Let $\boldsymbol{a}^2$ be the permutation of $\boldsymbol{y}^2$ such that $a^2_1 \leq a^2_2 \leq \cdots \leq a^2_n$; by anonymity, $\boldsymbol{a}^2 \sim \boldsymbol{y}^2$.

Similarly, we can construct $\boldsymbol{a}^0, \boldsymbol{y}^1, \boldsymbol{a}^1, \boldsymbol{y}^2, \boldsymbol{a}^2, \cdots, \boldsymbol{y}^k, \boldsymbol{a}^k \in X$ and $M^1, M^2, \cdots, M^k \subseteq N$ for some finite integer k such that $\boldsymbol{x} = \boldsymbol{a}^k \sim \boldsymbol{y}^k \prec \boldsymbol{a}^{k-1} \sim \boldsymbol{y}^{k-1} \prec \cdots \prec \boldsymbol{a}^1 \sim \boldsymbol{y}^1 \prec \boldsymbol{a}^0 = \boldsymbol{y}$ and $M_k \subseteq M_{k-1} \subseteq \cdots \subseteq M_2 \subseteq M_1$, which means $\boldsymbol{x} \prec \boldsymbol{y}$.

Second, if $\varphi_n(\boldsymbol{x}) \neq \varphi_n(\boldsymbol{y})$, then let $\boldsymbol{s} = \boldsymbol{x}/\varphi_n(\boldsymbol{x})$ and $\boldsymbol{t} = \boldsymbol{y}/\varphi_n(\boldsymbol{y})$. Since $\varphi_n(\boldsymbol{s}) = 1 = \varphi_n(\boldsymbol{t})$, we have $\boldsymbol{s} \prec \boldsymbol{t}$ by the result of the first step. By scale independence, $\boldsymbol{s} \sim \boldsymbol{x}$ and $\boldsymbol{t} \sim \boldsymbol{y}$. Therefore, $\boldsymbol{x} \prec \boldsymbol{y}$.

Q.E.D.

Lemma 1 is a fundamental result in the theory of relative measurement of social inequality. Forster's proof of Lemma 1 (Foster, 1985) is based on a result of Hardy et al. (1934, 1952, 1934, 1952), originally stated in terms of majorization. Our proof of Lemma 1 is elementary and actually provides a direct procedure for approaching the Lorenz criterion through anonymity, scale independence, and the Pigou-Dalton principle.[2]

[2]In fact, Foster (1985) utilizes Dalton's population principle (Dalton, 1920) and shows that in a variable society, an inequality measure satisfies the Lorenz criterion if and only if it satisfies anonymity, scale independence, the Pigou-Dalton principle, and the Population principle, and then Lemma 1 can be considered as a corollary. In our proof of Lemma 1, we do not utilize the population principle since we only consider a fixed society. But together with the population principle, our proof can be easily extended to the case of variable society.

Scale independence cannot be replaced with scale invariance in Lemma 1, since there are social orderings, such as the leximin and leximax rules, satisfying anonymity, scale invariance, and the Pigou-Dalton principle, but not scale independence.

Leximin: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, $\boldsymbol{x} \prec \boldsymbol{y}$ if and only if $x_{(i)} > y_{(i)}$ for some $i \in N$ and $x_{(j)} = y_{(j)}$ for all $j < i$.

Leximax: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, $\boldsymbol{x} \prec \boldsymbol{y}$ if and only if $x_{(i)} < y_{(i)}$ for some $i \in N$ and $x_{(j)} = y_{(j)}$ for all $j > i$.[3]

The leximin rule says that for any two income distributions $\boldsymbol{x}$ and $\boldsymbol{y}$, $\boldsymbol{x}$ is less unequal than $\boldsymbol{y}$ if the income of the poorest is larger in $\boldsymbol{x}$ than in $\boldsymbol{y}$, or if the incomes of the poorest in $\boldsymbol{x}$ and in $\boldsymbol{y}$ are equal and the income of the second poorest is larger in $\boldsymbol{x}$ than in $\boldsymbol{y}$, and so on. On the other hand, the leximax rule says that for any $\boldsymbol{x}$ and $\boldsymbol{y}$, $\boldsymbol{x}$ is less unequal than $\boldsymbol{y}$ if the income of the richest is smaller in $\boldsymbol{x}$ than in $\boldsymbol{y}$, or if the incomes of the richest in $\boldsymbol{x}$ and in $\boldsymbol{y}$ are equal and the income of the second richest is smaller in $\boldsymbol{x}$ than in $\boldsymbol{y}$, and so on. While failing to satisfy scale independence, the leximin and leximax social inequality orderings both satisfy Hammond's equity axiom, which is similar to but stronger than the Pigou-Dalton principle.

Hammond Principle: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, for each $i, j \in N$, if $y_i > x_i \geq x_j > y_j$ and $x_k = y_k$ for each $k \in N \setminus \{i, j\}$, then $\boldsymbol{x} \prec \boldsymbol{y}$.

First Hammond Principle: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, for each $i, j \in N$, if $y_i \geq x_i \geq x_j > y_j$ and $x_k = y_k$ for each $k \in N \setminus \{i, j\}$, then $\boldsymbol{x} \prec \boldsymbol{y}$.

Second Hammond Principle: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, for each $i, j \in N$, if $y_i > x_i \geq x_j \geq y_j$ and $x_k = y_k$ for each $k \in N \setminus \{i, j\}$, then $\boldsymbol{x} \prec \boldsymbol{y}$.

The Hammond principle asserts that a transfer (not necessarily mean-preserving) from a person to a richer one should increase social inequality. Intuitively, the Hammond principle attempts to formalize the idea that "the rich get richer and the poor get poorer," while the first Hammond principle lays more emphasis on "the poor get poorer" and the second on "the rich get richer."

It is worth noting that the Hammond principle is originally proposed as an equity axiom for characterizing social welfare orderings (Hammond, 1976). But we consider the Hammond principle more natural as an income inequality axiom rather than a welfare axiom.[4] For example, the distribution (41,41,60) is less unequal than the distribution (40,80,60) according to the Hammond principle, but it is more controversial to say that (41,41,60) is socially better than (40,80,60), even though the Hammond principle as a welfare axiom says so. To understand how to measure social inequality, one should always keep in mind that "$\boldsymbol{x}$ is less unequal than $\boldsymbol{y}$" does not necessarily mean "$\boldsymbol{x}$ is socially better than $\boldsymbol{y}$", although they may be closely related in some way.

[3]Note that the leximax social inequality ordering is different from the leximax social welfare inequality ordering. The latter is usually defined such that for each $\boldsymbol{x}, \boldsymbol{y} \in X$, $\boldsymbol{x}$ is socially better than $\boldsymbol{y}$ if and only if $x_{(i)} > y_{(i)}$ (not $x_{(i)} < y_{(i)}$) for some $i \in N$ and $x_{(j)} = y_{(j)}$ for all $j > i$. The leximax social welfare ordering satisfies the Pareto principle, but the leximax social inequality ordering does not. For example, by the leximax social inequality ordering, (1, 2, 3) is less unequal than (1, 2, 5), while by the leximax social welfare ordering, (1, 2, 5) is socially better than (1, 2, 3). On the other hand, the leximin social inequality ordering and the leximin social welfare ordering are actually the same rule and both satisfy the Pareto principle.

Obviously, both the first and second Hammond principles imply the Hammond principle, and the latter implies the Pigou-Dalton principle. Unfortunately, the first and second Hammond principles are inconsistent.

Lemma 2: *When $n \geq 3$, there is no social inequality ordering satisfying both the first and the second Hammond principles.*

Proof: Let $\boldsymbol{x} = (1, 3, 4, \cdots, 4)$ and $\boldsymbol{y} = (1, 2, 4, \cdots, 4)$. By the first Hammond principle, we have $(1, 3, 4, \cdots, 4) \prec (1, 2, 4, \cdots, 4)$, since $y_3 \geq x_3 > x_2 > y_2$ and for all $i \in N \setminus \{2, 3\}$, $x_i = y_i$. But by the second Hammond principle, we have $(1, 2, 4, \cdots, 4) \prec (1, 3, 4, \cdots, 4)$, since $x_2 > y_2 > y_1 \geq x_1$ and for all $j \in N \setminus \{1, 2\}$, $x_j = y_j$, and we thus get a contradiction.

Q.E.D.

3. Leximinimax and Leximaximin

It is easy to verify that (i) the leximin and leximax rules both satisfy anonymity, scale invariance, and the Hammond principle, but not scale independence; (ii) the leximin rule satisfies the first Hammond principle but not the second; (iii) the leximax rule satisfies the second Hammond principle but not the first. Inspired by those observations, we can further propose two other similar rules that inherit the very spirit of the leximin and leximax rules and do satisfy scale independence.

Let $x_{i|j} = \frac{x_{(i)}}{x_{(j)}}$, i.e., $x_{i|j}$ denotes the ratio of the income of the *i*th worst-off to the income of the *j*th worst-off.

Leximinimax: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, $\boldsymbol{x} \prec \boldsymbol{y}$ if and only if $x_{i|n} > y_{i|n}$ for some $i \in N$ and $x_{j|n} = y_{j|n}$ for each $j < i$.

Leximaximin: For each $\boldsymbol{x}, \boldsymbol{y} \in X$, $\boldsymbol{x} \prec \boldsymbol{y}$ if and only if $x_{i|1} < y_{i|1}$ for some $i \in N$ and $x_{j|1} = y_{j|1}$ for each $j > i$.

The leximinimax measure says that for any two income distributions $\boldsymbol{x}$ and $\boldsymbol{y}$, $\boldsymbol{x}$ is less unequal than $\boldsymbol{y}$ if the income ratio of the poorest to the richest is larger in $\boldsymbol{x}$ than in $\boldsymbol{y}$, or if the income ratio of the poorest to the richest in $\boldsymbol{x}$ and in $\boldsymbol{y}$ are equal and the income ratio of the second poorest to the richest is larger in $\boldsymbol{x}$ than in $\boldsymbol{y}$, and so on. The interpretation of the leximaximin measure is similar.

We will show that under the assumptions of anonymity and scale independence, the lexminimax and leximaximin measures can be fully characterized by the first and second Hammond principles respectively. Before this, we need to prove several preparatory lemmas. For convenience, let $D = \{\boldsymbol{x} \in X \mid x_1 \leq x_2 \leq \cdots \leq x_n\}$.

Lemma 3: *If a social inequality ordering $\precsim$ satisfies scale independence and the first Hammond principle, then for each $\boldsymbol{x}, \boldsymbol{y} \in D$, $\boldsymbol{x} \prec \boldsymbol{y}$ whenever $x_{1|n} > y_{1|n}$.*

[4]In the social choice literature, some authors attempt to consider the Hammond principle as an axiom for measuring inequality between ordinal utilities, without imposing the assumption of scale independence (Gravel et al. 2017).

Proof: If $n = 2$, then by the first Hammond principle, we have $(x_{1|n}, 1) \prec (y_{1|n}, 1)$, and by scale independence, we have $\boldsymbol{x} \sim (x_{1|2}, 1) \prec (y_{1|2}, 1) \sim \boldsymbol{y}$.

Let $n \geq 3$. By scale independence, there must be some $\boldsymbol{a} \in D$ such that $\boldsymbol{a} = \frac{y_n}{x_n}\boldsymbol{x} \sim \boldsymbol{x}$. Hence $a_1 = \frac{y_n}{x_n}x_1 > y_1$ and $a_n = y_n$. Then there must be some sufficiently small $\varepsilon > 0$ and some $\boldsymbol{a}^1, \boldsymbol{a}^2, \cdots, \boldsymbol{a}^{n-1} \in D$ such that

(i) $\boldsymbol{a}^1 = \boldsymbol{a}$;

(ii) for each $k \in \{2, 3, \cdots, n-1\}$, $a_1^k = a_1^{k-1} - \varepsilon$, $a_{n-k+1}^k = a_n$, and for each $j \in N \setminus \{1, n-k+1\}$, $a_j^k = a_j^{k-1}$;

(iii) $a_1^{n-1} = a_1 - (n-2)\varepsilon > y_1$.

By the first Hammond principle, we must have $\boldsymbol{a} = \boldsymbol{a}^1 \precsim \boldsymbol{a}^2 \precsim \cdots \precsim \boldsymbol{a}^{n-1}$.

There must also be some $\boldsymbol{b}^1, \boldsymbol{b}^2, \cdots, \boldsymbol{b}^{n-1} \in D$ such that $\boldsymbol{b}^1 = \boldsymbol{y}$, $b_{n-k+1}^k = y_n$ for each $k \in \{2, 3, \cdots, n-1\}$, and $b_j^k = b_j^{k-1}$ for each $j \in N \setminus \{n-k+1\}$.

Again, by the first Hammond principle, we must have $\boldsymbol{b}^{n-1} \precsim \boldsymbol{b}^{n-2} \precsim \cdots \precsim \boldsymbol{b}^1 = \boldsymbol{y}$.

Finally, by the first Hammond principle, we have $\boldsymbol{a}^{n-1} = (a_1 - (n-2)\varepsilon, a_n, \cdots, a_n) \prec (y_1, y_n, \cdots, y_n) = \boldsymbol{b}^{n-1}$. Therefore, we must have $\boldsymbol{a} \prec \boldsymbol{y}$, which means $\boldsymbol{x} \prec \boldsymbol{y}$.

Q.E.D.

Lemma 4: *If a social inequality ordering* $\precsim$ *satisfies scale independence and the first Hammond principle, then for each* $\boldsymbol{x}, \boldsymbol{y} \in D$, $\boldsymbol{x} \prec \boldsymbol{y}$ *whenever* $x_{i|n} > y_{i|n}$ *for some* $i \in N$ *and* $x_{j|n} = y_{j|n}$ *for each* $j < i$.

Proof: If $n = 2$, then we must have $i = 1$, and thus by scale independence and the first Hammond principle, we have $\boldsymbol{x} \sim (x_{1|2}, 1) \prec (y_{1|2}, 1) \sim \boldsymbol{y}$.

Let $n \geq 3$. If $i = 1$, then by Lemma 3, we have $\boldsymbol{x} \prec \boldsymbol{y}$. If $i \in \{2, 3, \cdots, n-1\}$, then by scale independence, there must be some $\boldsymbol{a} \in D$ such that $\boldsymbol{a} = \frac{y_n}{x_n}\boldsymbol{x} \sim \boldsymbol{x}$. Hence $a_1 = \frac{y_n}{x_n}x_1 = y_1$, $a_2 = \frac{y_n}{x_n}x_2 = y_2, \cdots, a_{i-1} = \frac{y_n}{x_n}x_{i-1} = y_{i-1}$, $a_i = \frac{y_n}{x_n}x_i > y_i$, and $a_n = \frac{y_n}{x_n}x_n = y_n$. Then there must be some sufficiently small $\varepsilon > 0$ and some $\boldsymbol{a}^1, \boldsymbol{a}^2, \cdots, \boldsymbol{a}^{n-i} \in X$ such that

(i) $\boldsymbol{a}^1 = \boldsymbol{a}$;

(ii) for each $k \in \{2, 3, \cdots, n-i\}$, $a_i^k = a_i^{k-1} - \varepsilon$, $a_{n-k+1}^k = a_n$, and for each $j \in N \setminus \{i, n-k+1\}$, $a_j^k = a_j^{k-1}$;

(iii) $a_i^{n-1} = a_i - (n-i-1)\varepsilon > y_i$.

By the first Hammond principle, we have $\boldsymbol{a} = \boldsymbol{a}^1 \precsim \boldsymbol{a}^2 \precsim \cdots \precsim \boldsymbol{a}^{n-i}$.

There must also be some $\boldsymbol{b}^1, \boldsymbol{b}^2, \cdots, \boldsymbol{b}^{n-i} \in D$ such that $\boldsymbol{b}^1 = \boldsymbol{y}$, $b^k_{n-k+1} = y_n$ for each $k \in \{2, 3, \cdots, n-i\}$, and $b^k_j = b^{k-1}_j$ for each $j \in N \setminus \{n-k+1\}$.

Again, by the first Hammond principle, we must have $\boldsymbol{b}^{n-i} \precsim \boldsymbol{b}^{n-i-1} \precsim \cdots \precsim \boldsymbol{b}^1 = \boldsymbol{y}$.

Finally, by the first Hammond principle, we have $\boldsymbol{a}^{n-i} = (a_1, a_2, \cdots, a_{i-1}, a_i-(n-i-1)\varepsilon, a_n, a_n, \cdots, a_n) \prec (y_1, y_2, \cdots, y_{i-1}, y_i, y_n, y_n, \cdots, y_n) = \boldsymbol{b}^{n-i}$. Therefore, we have $\boldsymbol{a} \prec \boldsymbol{y}$, which means $\boldsymbol{x} \prec \boldsymbol{y}$.

Q.E.D.

Theorem 1: *A social inequality ordering is the leximinimax measure if and only if it satisfies anonymity, scale independence, and the first Hammond principle.*

Proof: The only if part is obvious, and we now prove the if part. Suppose that a social inequality ordering $\precsim$ satisfies anonymity, scale independence, and the first Hammond principle. Consider any $\boldsymbol{x}, \boldsymbol{y} \in X$ such that for some $i \in N$, $x_{i|n} > y_{i|n}$, and for all $j < i$, $x_{j|n} = y_{j|n}$. There must be some $\boldsymbol{a}, \boldsymbol{b} \in D$ such that for each $k \in N$, $a_k = x_{(k)}$ and $b_k = y_{(k)}$, and hence we $a_{k|n} = x_{k|n}$ and $b_{k|n} = y_{k|n}$ for each $k \in N$. We also have $a_{i|n} > b_{i|n}$ for some $i \in N$ and $a_{j|n} = b_{j|n}$ for all $j < i$. By anonymity, we have $[\boldsymbol{x} \prec \boldsymbol{y}] \Longleftrightarrow [\boldsymbol{a} \prec \boldsymbol{b}]$. By Lemma 4, we have $\boldsymbol{a} \prec \boldsymbol{b}$, and hence we must have $\boldsymbol{x} \prec \boldsymbol{y}$, as required.

Q.E.D.

Corollary 1: *A social inequality ordering is the leximinimax measure if and only if it satisfies the Lorenz criterion and the first Hammond principle.*

Proof: It follows immediately from Lemma 1 and Theorem 1.

Q.E.D.

The characterization process of the leximaximin measure will have many similarities to that of the leximinimax measure, but for the purpose of being self-contained, we still present the proofs in detail.

Lemma 5: *If a social inequality ordering $\precsim$ satisfies scale independence and the second Hammond principle, then for each $\boldsymbol{x}, \boldsymbol{y} \in D$, $\boldsymbol{x} \prec \boldsymbol{y}$ whenever $x_{n|1} < y_{n|1}$.*

Proof: If $n = 2$, then by the second Hammond principle, we have $(1, x_{2|1}) \prec (1, y_{2|1})$, and by scale independence, we have $\boldsymbol{x} \sim (1, x_{2|1}) \prec (1, y_{2|1}) \sim \boldsymbol{y}$.

Let $n \geq 3$. By scale independence, there must be some $\boldsymbol{a} \in D$ such that $\boldsymbol{a} = \frac{y_1}{x_1}\boldsymbol{x} \sim \boldsymbol{x}$. Hence $a_n = \frac{y_1}{x_1}x_n < y_n$ and $a_1 = y_1$. Then there must be some sufficiently small $\varepsilon > 0$ and some $\boldsymbol{a}^1, \boldsymbol{a}^2, \cdots, \boldsymbol{a}^{n-1} \in D$ such that

(i) $\boldsymbol{a}^1 = \boldsymbol{a}$;

(ii) for $k = 2, 3, \cdots, n-1$, $a_k^k = a_1, a_n^k = a_n^{k-1} + \varepsilon$, and for each $j \in N \setminus \{k, n\}$, $a_j^k = a_j^{k-1}$;

(iii) $a_n^{n-1} = a_n + (n-2)\varepsilon < y_n$.

By the second Hammond principle, we must have $\boldsymbol{a} = \boldsymbol{a}^1 \precsim \boldsymbol{a}^2 \precsim \cdots \precsim \boldsymbol{a}^{n-1}$.

There must also be some $\boldsymbol{b}^1, \boldsymbol{b}^2, \cdots, \boldsymbol{b}^{n-1} \in D$ such that $\boldsymbol{b}^1 = \boldsymbol{y}$, $b_k^k = y_1$ for each $k \in \{2, 3, \cdots, n-1\}$, and $b_j^k = b_j^{k-1}$ for each $j \in N \setminus \{k\}$.

By the second Hammond principle, we must have $\boldsymbol{b}^{n-1} \precsim \boldsymbol{b}^{n-2} \precsim \cdots \precsim \boldsymbol{b}^1 = \boldsymbol{y}$.

Again, by the second Hammond principle, $\boldsymbol{a}^{n-1} = (a_1, a_1, \cdots, a_1, a_n + (n-2)\varepsilon) \prec (y_1, y_1, \cdots, y_1, y_n) = \boldsymbol{b}^{n-1}$. Therefore, we must have $\boldsymbol{a} \prec \boldsymbol{y}$, which means $\boldsymbol{x} \prec \boldsymbol{y}$.

Q.E.D.

Lemma 6: *If a social inequality ordering* $\precsim$ *satisfies scale independence and the second Hammond principle, then for each* $\boldsymbol{x}, \boldsymbol{y} \in D$, $\boldsymbol{x} \prec \boldsymbol{y}$ *whenever* $x_{i|1} < y_{i|1}$ *for some* $i \in N$ *and* $x_{j|1} = y_{j|1}$ *for each* $j > i$.

Proof: If $n = 2$, then we must have $i = 2$, and thus by scale independence and the second Hammond principle, we have $\boldsymbol{x} \sim (1, x_{2|1}) \prec (1, y_{2|1}) \sim \boldsymbol{y}$.

Let $n \geq 3$. If $i = n$, then by Lemma 5, we have $\boldsymbol{x} \prec \boldsymbol{y}$. If $i \in \{2, 3, \cdots, n-1\}$, then by scale independence, there must be some $\boldsymbol{a} \in D$ such that $\boldsymbol{a} = \frac{y_1}{x_1}\boldsymbol{x} \sim \boldsymbol{x}$. Hence $a_n = \frac{y_1}{x_1}x_n = y_n$, $a_{n-1} = \frac{y_1}{x_1}x_{n-1} = y_{n-1}, \cdots, a_{i+1} = \frac{y_1}{x_1}x_{i+1} = y_{i+1}$, $a_i = \frac{y_1}{x_1}x_i < y_i$, and $a_1 = \frac{y_1}{x_1}x_1 = y_1$. Then there must be some sufficiently small $\varepsilon > 0$ and some $\boldsymbol{a}^1, \boldsymbol{a}^2, \cdots, \boldsymbol{a}^{n-i} \in X$ such that

(i) $\boldsymbol{a}^1 = \boldsymbol{a}$;

(ii) for each $k \in \{2, 3, \cdots, n-i\}$, $a_k^k = a_1, a_i^k = a_i^{k-1} + \varepsilon$, and for each $j \in N \setminus \{k, i\}$, $a_j^k = a_j^{k-1}$;

(iii) $a_i^{n-i} = a_i + (n-i-1)\varepsilon < y_i$.

By the second Hammond principle, we must have $\boldsymbol{a} = \boldsymbol{a}^1 \precsim \boldsymbol{a}^2 \precsim \cdots \precsim \boldsymbol{a}^{n-i}$.

There must also be some $\boldsymbol{b}^1, \boldsymbol{b}^2, \cdots, \boldsymbol{b}^{n-i} \in D$ such that $\boldsymbol{b}^1 = \boldsymbol{y}$, $b_k^k = y_1$ for each $k \in \{2, 3, \cdots, n-i\}$, and $b_j^k = b_j^{k-1}$ for all $j \in N \setminus \{k\}$. By the second Hammond principle, we must have $\boldsymbol{b}^{n-1} \precsim \boldsymbol{b}^{n-2} \precsim \cdots \precsim \boldsymbol{b}^1 = \boldsymbol{y}$.

By the second Hammond principle, $\boldsymbol{a}^{n-i} = (a_1, a_1, \cdots, a_1, a_i + (n-i-1)\varepsilon, a_{i+1}, a_{i+2}, \cdots, a_n) \prec (y_1, y_1, \cdots, y_1, y_i, y_{i+1}, y_{i+2}, \cdots, y_n) = \boldsymbol{b}^{n-i}$. Therefore, we must have $\boldsymbol{a} \prec \boldsymbol{y}$, which means $\boldsymbol{x} \prec \boldsymbol{y}$.

Q.E.D.

Theorem 2: *A social inequality ordering is the leximaximin measure if and only if it satisfies anonymity, scale independence, and the second Hammond principle.*

TABLE 2
INCOME INEQUALITY IN THE UNITED STATES, 1967–2016[a]

Year	$x_{1\|5}$	$x_{2\|5}$	$x_{3\|5}$	$x_{4\|5}$	Year	$x_{1\|5}$	$x_{2\|5}$	$x_{3\|5}$	$x_{4\|5}$
1967	.0898	.2488	.3971	.5557	1992	.0796	.1996	.3362	.5161
1968	.0970	.2601	.4126	.5755	1993	.0727	.1843	.3089	.4800
1969	.0942	.2542	.4062	.5690	1994	.0728	.1815	.3057	.4757
1970	.0919	.2489	.4007	.5649	1995	.0763	.1864	.3117	.4792
1971	.0926	.2448	.3970	.5644	1996	.0744	.1826	.3072	.4755
1972	.0920	.2377	.3880	.5570	1997	.0720	.1800	.3028	.4691
1973	.0942	.2369	.3885	.5589	1998	.0723	.1826	.3056	.4726
1974	.0976	.2425	.3908	.5644	1999	.0733	.1800	.3013	.4689
1975	.0967	.2376	.3898	.5664	2000	.0714	.1783	.2969	.4615
1976	.0966	.2366	.3891	.5650	2001	.0694	.1745	.2920	.4579
1977	.0940	.2314	.3827	.5608	2002	.0695	.1767	.2978	.4684
1978	.0943	.2316	.3824	.5601	2003	.0680	.1746	.2964	.4691
1979	.0926	.2303	.3797	.5572	2004	.0676	.1731	.2933	.4624
1980	.0927	.2307	.3807	.5609	2005	.0668	.1714	.2901	.4564
1981	.0910	.2267	.3756	.5598	2006	.0675	.1711	.2868	.4539
1982	.0875	.2215	.3666	.5452	2007	.0688	.1753	.2975	.4710
1983	.0874	.2195	.3630	.5447	2008	.0681	.1726	.2931	.4663
1984	.0875	.2180	.3612	.5439	2009	.0676	.1712	.2899	.4606
1985	.0846	.2144	.3559	.5354	2010	.0649	.1684	.2903	.4657
1986	.0815	.2096	.3507	.5283	2011	.0631	.1641	.2800	.4498
1987	.0819	.2081	.3479	.5258	2012	.0632	.1633	.2814	.4513
1988	.0821	.2072	.3465	.5238	2013	.0600	.1594	.2779	.4472
1989	.0818	.2035	.3382	.5116	2014	.0602	.1602	.2785	.4526
1990	.0822	.2069	.3418	.5153	2015	.0616	.1612	.2808	.4548
1991	.0820	.2060	.3421	.5215	2016	.0605	.1613	.2765	.4449

[a]According to the figures in Table 2, one can also compute the leximaximin measures: $x_{5|1} = 1/x_{1|5}$, $x_{4|1} = x_{4|5}/x_{1|5}$, $x_{3|1} = x_{3|5}/x_{1|5}$, and $x_{2|1} = x_{2|5}/x_{1|5}$.

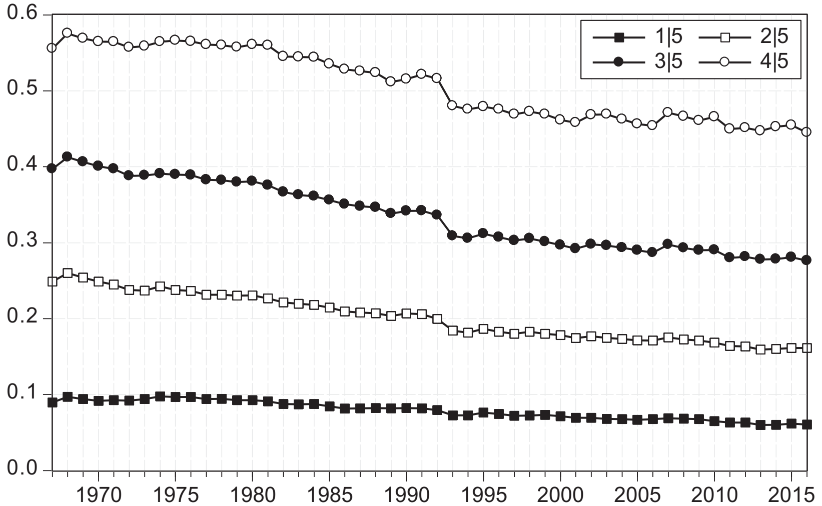

Figure 1. Income Inequality in the United States, 1967–2016

Proof: The only if part is obvious, and we now prove the if part. Suppose that a social inequality ordering $\precsim$ satisfies anonymity, scale independence, and the second Hammond principle. Consider any $\boldsymbol{x}, \boldsymbol{y} \in X$ such that for some $i \in N$, $x_{i|I} < y_{i|I}$, and for all $j > i$, $x_{j|I} = y_{j|I}$. There must be some $\boldsymbol{a}, \boldsymbol{b} \in D$ such that for each $k \in N$, $a_k = x_{(k)}$ and $b_k = y_{(k)}$, and hence we have $a_{k|I} = x_{k|I}$ and $b_{k|I} = y_{k|I}$ for each $k \in N$, and we also have $a_{i|I} > b_{i|I}$ for some $i \in N$ and $a_{j|I} = b_{j|I}$ for all $j > i$. By anonymity, we have $[\boldsymbol{x} \prec \boldsymbol{y}] \Longleftrightarrow [\boldsymbol{a} \prec \boldsymbol{b}]$. By Lemma 6, we have $\boldsymbol{a} \prec \boldsymbol{b}$, and hence we must have $\boldsymbol{x} \prec \boldsymbol{y}$, as required.

Q.E.D.

Corollary 2: *A social inequality ordering is the leximaximin measure if and only if it satisfies the Lorenz criterion and the second Hammond principle.*

Proof: It follows immediately from Lemma 1 and Theorem 2.

Q.E.D.

It is easy to see that those lexical measures are not continuous. One may wonder whether there are some continuous measures satisfying anonymity, scale independence, and the Hammond principle. Unfortunately, we can show that the Hammond principle is incompatible with the assumption of continuity, even without the assumptions of anonymity and scale independence.

Continuity: For any sequences $(\boldsymbol{x}^k)_{k=1,2,\dots}$ and $(\boldsymbol{y}^k)_{k=1,2,\dots}$ in X such that $\boldsymbol{x}^k \precsim \boldsymbol{y}^k$ for all k, if $\lim_{k\to\infty} \boldsymbol{x}^k = \boldsymbol{x}$ and $\lim_{k\to\infty} \boldsymbol{y}^k = \boldsymbol{y}$, then we have $\boldsymbol{x} \precsim \boldsymbol{y}$.

Theorem 3: *If* $n \geq 4$, *then there is no continuous social inequality ordering satisfying the Hammond principle*; *if* $n \geq 3$, *then there is no continuous social inequality ordering satisfying the first or the second Hammond principle*.

Proof: Suppose that $\precsim$ is continuous. First, let $n \geq 4$ and suppose $\precsim$ satisfies the Hammond principle. Then let $\boldsymbol{x} = (1, 4, 7, 10, x_5, \cdots, x_n)$, $\boldsymbol{y} = (1, 5, 6, 10, x_5, \cdots, x_n)$, $\boldsymbol{z} = (1, 3, 8, 10, x_5, \cdots, x_n)$, and $\boldsymbol{z}^k = (1 + 1/k, 3, 8, 10 - 1/k, x_5, \cdots, x_n)$ for each $k = 1, 2, \cdots$. By the Hammond principle, we have $\boldsymbol{y} \prec \boldsymbol{x}$, $\boldsymbol{x} \prec \boldsymbol{z}$, and for each $k = 1, 2, \cdots$, $\boldsymbol{z}^k \prec (1 + 1/k, 3, 6, 10, x_5, \cdots, x_n) \prec \boldsymbol{y} \prec \boldsymbol{x}$, i.e., $\boldsymbol{z}^k \prec \boldsymbol{x}$. Hence by continuity, we have $\boldsymbol{z} \precsim \boldsymbol{x}$, a contradiction.

Second, if $n \geq 4$, then by the first step, there must be no continuous social inequality ordering satisfying the first or the second Hammond principle since both the first and the second Hammond principles imply the Hammond principle. Now, let $n = 3$. Then let $\boldsymbol{x} = (1,4,10)$, $\boldsymbol{y} = (1,5,10)$, $\boldsymbol{z} = (1,3,10)$, and $\boldsymbol{z}^k = (1 + 1/k, 4, 10 - 1/k)$ for each $k = 1, 2, \cdots$. If the first Hammond principle is satisfied, then we have $\boldsymbol{y} \prec \boldsymbol{x}$ and $\boldsymbol{x}^k \prec (1 + 1/k, 3, 10) \prec \boldsymbol{y}$ for each $k = 1, 2, \cdots$. Thus by continuity, we have $\boldsymbol{x} \precsim \boldsymbol{y}$, a contradiction. On the other hand, if the second Hammond principle is satisfied, then we have $\boldsymbol{z} \prec \boldsymbol{x}$ and $\boldsymbol{x}^k \prec (1,5,10 - 1/k) \prec \boldsymbol{z}$. Hence by continuity, we have $\boldsymbol{x} \precsim \boldsymbol{z}$, a contradiction.

Q.E.D.

When $n = 3$, however, there is a continuous social inequality ordering satisfying anonymity, scale independence, and the Hammond principle: for any $\boldsymbol{x}, \boldsymbol{y} \in X$, $[\boldsymbol{x} \prec \boldsymbol{z}] \Longleftrightarrow \frac{\min_{i\in N}\{x_i\}}{\max_{i\in N}\{x_i\}} > \frac{\min_{i\in N}\{y_i\}}{\max_{i\in N}\{y_i\}} \Longleftrightarrow \frac{\max_{i\in N}\{x_i\}}{\min_{i\in N}\{x_i\}} < \frac{\max_{i\in N}\{y_i\}}{\min_{i\in N}\{y_i\}}$.

Continuity is a useful technical assumption in the characterizations of the generalized entropy measures (Toyoda, 1975; Cowell, 1977; Shorrocks, 1980; Cowell and Kuga, 1981a, 1981b), and some characterizations even use the stronger assumption of differentiability (Chakravarty, 1999). Intuitively, continuity ensures that a "small" change in the income distribution should lead to only a "small" change in measured inequality. The lexical measures are not continuous, and Theorem 3 shows that there are indeed no continuous measures satisfying the Hammond principle. We thus have to make a choice between continuity and the Hammond principle. We believe that in the characterizations of social inequality measures, ethical justifiability must be prior to mathematical convenience, and any technical assumptions however elegant must be rejected if they unavoidably conflict with the requirements for defining the inner nature of social inequality.

4. Income Inequality in the United States: an Illustrative Example

We now present an example to illustrate the application of the leximinimax measure (the application of the leximaximin measure is similar and thus omitted). The leximinimax measure is simple and intuitive, but it may usually exhibit unsatisfactory statistical stability when n is large. In practice, a society can be divided into $m < n$ equal sized classes according to income level, and then we can measure the corresponding inequality between classes. A traditional way is to let $m = 5$, and then the society can be divided into the following five classes: the lower class (the lowest quintile), the lower middle class (the second quintile), the middle class (the third quintile), the upper middle class (the fourth quintile), and the upper class (the highest quintile).

Based on the quintile classification, we use data of mean household income of quintiles from the U.S. Census Bureau to evaluate income inequality in the United States for the period 1967-2016 (Semega et al., 2017, Appendix A-2). From Table 2 and Figure 1, it is easy to see that each of the four income ratios (i.e., $x_{1|5}$, $x_{2|5}$, $x_{3|5}$, $x_{4|5}$) has experienced a long-term decrease: the ratio of the lower class to the upper class is about 0.09 in 1967 and 0.06 in 2016; the ratio of the lower middle class to the upper class is about 0.25 in 1967 and 0.16 in 2016; the ratio of the middle class to the upper class is about 0.40 in 1967 and 0.28 in 2016; the ratio of the upper middle class to the upper class is about 0.56 in 1967 and 0.45 in 2016. According to the leximinimax measure, we can conclude that the long-term decrease of the ratio of the lower class to the upper class clearly suggests a long-term increase in income inequality in the United States from 1967 to 2016.

The calculation result of the leximinimax measure of income inequality in the United States is not surprising, since other inequality measures on the whole suggest similar increasing trends of inequality. For instance, from the Current Population Reports, the Gini index of income inequality in the United States is about 0.397 in 1967 and 0.481 in 2016, and the Theil index is about 0.287 in 1967 and 0.426 in 2016. A minor advantage of the leximinimax measure is that one can not only see the trend of inequality, but also understand that the trend of inequality is actually the change of the income ratio of the lower class to the upper class, while the changes of the Gini index or the Theil index cannot be easily explained to anyone who has not learned enough about inequality measurement.

In addition, it is interesting to note that in practice the bottom/top ratios under comparison are usually different at different times, and thus the lexical comparisons are usually irrelevant. Therefore, leximinimax and leximaximin can be approximated respectively by $\frac{\min_{i\in N}\{x_i\}}{\max_{i\in N}\{x_i\}}$ and $\frac{\max_{i\in N}\{x_i\}}{\min_{i\in N}\{x_i\}}$, which actually represent the same ordering.

5. Conclusion

This paper proposes two lexical measures of social inequality. We show that a social inequality ordering is the leximinimax measure if and only if it satisfies anonymity, scale independence, and the first Hammond principle, if and only if it satisfies the Lorenz criterion and the first Hammond principle; similarly, a social inequality ordering is the leximaximin measure if and only if it satisfies anonymity, scale independence, and the second Hammond principle, if and only if it satisfies the Lorenz criterion and the second Hammond principle. In addition, we also show that there are no continuous measures satisfying the Hammond principle.

The lexical measures of inequality are simple and intuitive. Their ethical foundations are established mainly by the Hammond principle, an attempt to formalize the fundamental idea about the true nature of income inequality: the rich get richer and the poor get poorer. They are also empirically implementable, as shown by the example presented in the previous section. We thus propose the lexical measures as reasonable alternatives to the traditional measures of inequality. The lexical measures, of course, are not perfect; as any other existing measures of inequality, they do have some inevitable drawbacks of their own. But we still hope that the lexical measures may shed some light on the basic issues in the theory of inequality measurement.

First, the lexical measures take a descriptive approach to the measurement of inequality rather than a normative approach. According to the Atkinson-Kolm-Sen approach, we can first choose a social welfare ordering $\succsim_W$ on X and then define a social inequality ordering $\precsim$ on X such that for any $\boldsymbol{x}, \boldsymbol{y} \in X$, if $\sum_{i\in N} x_i = \sum_{i\in N} y_i$ then $[\boldsymbol{x} \precsim \boldsymbol{y}] \Longleftrightarrow [\boldsymbol{x} \succsim_W \boldsymbol{y}]$. As a result, the choice of social inequality measures totally depends on the choice of social welfare criterions. This is definitely a subjective approach to measurement of inequality. Unfortunately, given any social inequality ordering $\precsim$, we can always choose a utilitarian social welfare ordering $\succsim_W$ such that for any $\boldsymbol{x}, \boldsymbol{y} \in X$, if $\sum_{i\in N} x_i > \sum_{i\in N} y_i$ then $\boldsymbol{x} \succ_W \boldsymbol{y}$, and if $\sum_{i\in N} x_i = \sum_{i\in N} y_i$ then $[\boldsymbol{x} \precsim \boldsymbol{y}] \Longleftrightarrow [\boldsymbol{x} \succsim_W \boldsymbol{y}]$. This makes the choice of social inequality orderings somewhat arbitrary: whichever inequality measure you choose, there is always a utilitarian interpretation for it.

Without doubt, social inequality and social welfare are closely related. The increase or decrease of social inequality usually affects the judgements about social welfare. On the other hand, social inequality should depend less on social welfare. Rather than first choosing a social welfare ordering and then defining a social inequality ordering associated with it, one should first choose a social inequality measure and then incorporate it into the definition of social welfare criterion. In some sense, social inequality can be considered only an objective description about the social distribution of the real world, and the measure of social inequality does not necessarily involve any questions like "what is socially better," unless it is totally hopeless to find any satisfactory measures of inequality. The lexical measures are thus proposed as an attempt to establish an objective ground for measuring social inequality, especially when the Lorenz curves intersect.

Second, the interrelation between social welfare and social inequality also causes some confusion about the choices between relative and absolute measures of inequality. In the example given by Kolm (1999), income distribution (0.01, 1) is simply transformed into (0.1, 10), and then social inequality may not change for at least two reasons: first, the distribution of real social resources represented by the income distribution may remain unchanged; second, even if real social resources have increased and thus social welfare is higher (whether from the viewpoint of utilitarianism or Rawlsianism), equiproportional changes in individual incomes do not change the resource shares represented by the income shares, i.e., we actually define a social inequality ordering on the space $\{\boldsymbol{x} \in (0, 1)^n \mid \sum_{i \in N} x_i = 1\}$. That is partly why we are interested in relative measures. Of course, relative measures can be inappropriate in some cases. For instance, when directly measuring inequality of nonmonetary social resources (such as consumption), equiproportional changes in incomes may change the shares of real social sources, and thus relative measures may become less appealing. Then translation invariance or other "intermediate" versions of invariance may be specified (Kolm, 1969, 1976; Bossert and Pfingsten, 1990).

Furthermore, the lexical measures of inequality might be easily criticized at first glance because they are too sensitive to income differences between the poorest and the richest, and the middle members usually play no role except as tie breakers. Note that Rawls's indifference principle and Sen's leximin rule (Rawls, 1971; Sen, 1970) could also be criticized for the same reason. This criticism, however, is actually a prejudice since it is not directly based on any fundamental axioms about inequality. This prejudice may largely arise from the traditional synthetic indices such as the Gini index. Consider two income distributions $\boldsymbol{x} = (1, 11, 11, 11, 11, 11, 11, 11, 11, 11)$ and $\boldsymbol{y} = (7, 8, 9, 10, 10, 10, 10, 11, 12, 13)$ in a society with ten members. According to the formula in Table 1, the Gini index of $\boldsymbol{x}$ is 0.1, the Gini index of $\boldsymbol{y}$ is 0.10222⋯, and hence $\boldsymbol{x}$ is less unequal than $\boldsymbol{y}$. But by the leximinimax and leximaximin measures, $\boldsymbol{y}$ is obviously less unequal than $\boldsymbol{x}$, which can also be simply derived from scale independence and the Hammond principle: for some sufficiently small $\varepsilon > 0$,

$$
\begin{aligned}
\boldsymbol{y} &\sim (3.5, 4, 4.5, 5, 5, 5, 5, 5.5, 6, 6.5) \\
&\prec (3.5 - \varepsilon, 4, 4.5, 5, 5, 5, 5, 5.5, 6, 11) \\
&\prec (3.5 - 2\varepsilon, 4, 4.5, 5, 5, 5, 5, 5.5, 11, 11) \\
&\qquad \cdots\cdots \\
&\prec (3.5 - 8\varepsilon, 4, 11, 11, 11, 11, 11, 11, 11, 11) \prec \boldsymbol{x}.
\end{aligned}
$$

The above procedure is very general for the applications of the lexical measures and thus makes the explanation of all inequality comparisons of incomes extremely simple and transparent. On the contrary, there are always some mysteries in inequality comparisons when applying synthetic indices such as the Gini index.

Finally, we can easily apply the lexical measures for populations of different sizes if we can utilize Dalton's population principle (Dalton, 1920), which states that an income distribution must be as unequal as any distributions simply formed by replications of it. Therefore, for any two distributions $\boldsymbol{x}$ and $\boldsymbol{y}$ with different sizes n and m, we can say that $\boldsymbol{x} \precsim \boldsymbol{y}$ if and only if $\boldsymbol{x}' \precsim \boldsymbol{y}'$, where $\boldsymbol{x}' \in \mathbb{R}^{\mathrm{n}\times\mathrm{m}}$ is an m-replication of $\boldsymbol{x}$ and $\boldsymbol{y}' \in \mathbb{R}^{\mathrm{n}\times\mathrm{m}}$ is an n-replication of $\boldsymbol{y}$. In addition, we cannot approximate the leximinimax and leximaximin measures simply by $\frac{\min_{i\in N}\{x_i\}}{\max_{i\in N}\{x_i\}}$ and $\frac{\max_{i\in N}\{x_i\}}{\min_{i\in N}\{x_i\}}$ any more when dealing with populations of different sizes. For example, consider three income distributions (1, 2), (1, 1, 2, 2), and (1, 1, 1, 2). By the population principle, (1, 2) is as unequal as (1, 1, 2, 2). But by the leximinimax measure, (1, 1, 2, 2) must be less unequal than (1, 1, 1, 2), hence (1, 2) must be less unequal than (1, 1, 1, 2); on the contrary, by the leximaximin measure, (1, 1, 1, 2) must be less unequal than (1, 1, 2, 2), and hence (1, 1, 1, 2) must be less unequal than (1, 2).

References

Arrow, K. J., *Social Choice and Individual Values*, Wiley, New York, 1951, 1963.

Atkinson, A. B., "On the Measurement of Income Inequality," *Journal of Economic Theory*, 2, 244–63, 1970.

Blackorby, C., W. Bossert, and D. Donaldson, "Income Inequality Measurement: The Normative Approach," in J. Silber (ed.), *Handbook of Income Inequality Measurement*, Kluwer, Dordrecht, *133–61*, 1999.

Bossert, W. and A. Pfingsten, "Intermediate Inequality: Concepts, Indices and Welfare Implications," *Mathematical Social Sciences*, 19, 117–34, 1990.

Chakravarty, S. R., "Measuring Inequality: The Axiomatic Approach," in J. Silber (ed.), *Handbook of Income Inequality Measurement*, Kluwer, Dordrecht, *163–86*, 1999.

Cowell, F. A., *Measuring Inequality*, Phillip Allan, Oxford, 1977.

______, "Measurement of Inequality," in A. B. Atkinson and F. Bourguignon (eds), *Handbook of Income Distribution, Volume I*. Elsevier, North-Holland, *87–166*, 2000.

______, "Inequality and Poverty Measures," in M.D. Adler and M. Fleurbaey (eds), *The Oxford Handbook of Well-being and Public Policy*, Oxford University Press, Oxford, 82–125, 2016.

______, and K. Kuga, "Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement," *Journal of Economic Theory*, 25, 131–43, 1981a.

______, and K. Kuga, "Inequality Measurement: An Axiomatic Approach," *European Economic Review*, 15, 287–305, 1981b.

Dalton, H., "The Measurement of the Inequality of Incomes," *Economic Journal*, 30, 348–61, 1920.

Dutta, B., "Inequality, Poverty and Welfare," in K. J. Arrow, A. K. Sen, and K. Suzumura (eds), *Handbook of Social Choice and Welfare, Volume I*, Elsevier, North-Holland, 597–633, 2002.

Foster, J. E., "Inequality measurement," In H. P. Young (ed.), *Fair Allocation*, American Mathematical Society, Providence RI, *38–61*, 1985.

Foster, J. E. and E. A. Ok, "Lorenz Dominance and the Variance of Logarithms," *Econometrica*, 67, 901–7, 1999.

Gini, C., "Variabilità e Mutuabilità," *Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche*, C. Cuppini, Bologna, 1912.

Gravel, N., B. Magdalou, and P. Moyes, "Hammond's Equity Principle and the Measurement of Ordinal Inequalities," *Working Papers 17–03*. LAMETA, University of Montpellier, 2017.

Hammond, P. J., "Equity, Arrow's Conditions, and Rawls' Difference Principle," *Econometrica*, 44, 793–804, 1976.
Hardy, G. H., J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, London, 1934, 1952.
Kolm, S.-C., "The Optimal Production of Social Justice," in J. Margolis and H. Guitton (eds), *Public Economics*, 145–200. Macmillan, London, 1969.
______, "Unequal Inequalities I, II," *Journal of Economic Theory*, 12, 416–42, 13, 82-111, 1976.
______, "The Rational Foundations of Income Inequality Measurement," in J. Silber (ed.), *Handbook of Income Inequality Measurement*, Kluwer, Dordrecht, 19–100, 1999.
Pigou, A. C., *Wealth and Welfare*. Macmillan, London, 1912.
Rawls, J., *A Theory of Justice*. Harvard University Press, Cambridge, MA, 1971.
Sen, A. K., *Collective Choice and Social Welfare*. North-Holland, Amsterdam, 1970.
______, *On Economic Inequality*. Clarendon Press, Oxford, 1973.
Semega, J. L., K. R. Fontenot, and M. A. Kollar, "*Income and Poverty in the United States: 2016*," *Current Population Reports P60–259*, U.S. Census Bureau, U.S. Government Printing Office, Washington, DC, 2017.
Shorrocks, A. F., "The Class of Additively Decomposable Inequality Measures," *Econometrica*, 48, 613–25, 1980.
Theil, H., *Economics and Information Theory*. North Holland, Amsterdam, 1967.
Toyoda, T., "Inequalities of Income Distributions: Their Comparisons and Inequality Measures," *Kokumin Keizai*, 134, 15–41, Japanese, 1975.