Majorization by L^p-Norms

Tapan Mitra*

and

Efe A. Ok†

February, 2001

Abstract

Some submajorization type integral inequalities are proved in arbitrary measure spaces. The main results provide sufficient conditions for a nonnegative measurable and rearrangable function to have a larger L^p-norm than another such function, regardless of which particular L^p-norm is used in the comparison.

AMS Subject Classification: 26D15

Keywords: majorization, monotonic rearrangements, entropy

*Department of Economics, Cornell University, Ithaca NY 14853. E-mail: tm19@cornell.edu.

†Department of Economics, New York University, 269 Mercer St., New York NY 10003. E-mail: efe.ok@nyu.edu.

1. INTRODUCTION

Let $(\Omega, \mathcal{F}, \mu)$ be any measure space, and take any two (rearrangable) nonnegative $\mathcal{F}$-measurable functions f, g on Ω. In this note we ask and provide some answers to the following question: under what conditions is the L^p-norm of f at least as large as that of g, regardless of which L^p-norm is used in the comparison? Put differently, we wish to identify easy-to-check conditions in terms of f and g such that $\|f\|_p \geqslant \|g\|_p$ holds for all $p \geqslant 1$ (read g is L^p-*norm dominated* by f). The nature of this question suggests that f and g must be related by some sort of a submajorization inequality. To formalize this intuition we shall extend here the classical Tomic-Weil submajorization theorem to the class of all $\mathcal{F}$-measurable and rearrangable nonnegative real functions (Theorem 3.1). The resulting theorem entails readily that if f submajorizes g (Remark 3.1), then $\|f\|_p \geqslant \|g\|_p$ holds for all $p \geqslant 1$ (Proposition 4.2). The converse turns out to be false. In particular, the entropy supermajorization turns out to lie strictly between the submajorization and the L^p-norm preorders (Theorem 5.1 and Example 6.1). Obtaining a majorization type characterization of the L^p-norm preorder is presently open.

Being closely related to the theory of mean inequalities, the investigation of the L^p-norm preorder is interesting on its own right. What is more, this preorder arises naturally in the context of measurement of economic welfare. To see this, fix any population size $n \in \mathbb{Z}_{++}$, and define

$$m_p(x) := \left(\frac{1}{n} \sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

for any $x \in \mathbb{R}^n_+$. If x_i is interpreted as the difference between person i's income and the mean income level, then m_p is called a *compromise inequality measure* (Ebert, 1988). If x_i is the difference between person i's initial and final period incomes, then m_p corresponds to an *absolute income mobility index* (Mitra and Ok, 1998). Finally, if x_i is interpreted as person i's income shortfall from the poverty line (with $x_i = 0$ if person i's income is above the poverty line), then m_p is called the P_p*-index of poverty* (Foster, Greer and Thorbecke, 1984). Whichever interpretation is chosen, there is no compelling reason to choose a specific $p \geqslant 1$ value in making use of the m_p function in empirical applications. Moreover, while an income distribution can be declared more unequal (or less mobile, etc.) than another by an m_r function, the opposite conclusion may well ensue according to an m_s function for $r \neq s$. It is thus natural to investigate the circumstances under which one can *unanimously* rank income distributions via m_p for *all* $p \geqslant 1$. The present results, when specialized to $\mathbb{R}^n_+$, provide sufficient conditions for such a unanimous ranking to take place.

2. MONOTONIC REARRANGEMENTS

Fix any measure space $(\Omega, \mathcal{F}, \mu)$ and take an $\mathcal{F}$-measurable mapping $f : \Omega \to \mathbb{R}_+$. We say that f is *rearrangable* if $\lim_{s \nearrow \infty} \mu\{\omega : f(\omega) > s\} = 0$, and denote the set of all $\mathcal{F}$-measurable and rearrangable nonnegative functions on Ω by $\mathcal{L}_+(\Omega, \mathcal{F}, \mu)$. The *decreasing* and *increasing rearrangements* of an $f \in \mathcal{L}_+(\Omega, \mathcal{F}, \mu)$ are defined on $(0, \mu(\Omega))$ as

$$f_\downarrow(t) := \sup\{s : \mu\{\omega : f(\omega) > s\} > t\} \quad \text{and} \quad f_\uparrow(t) := \sup\{s : \mu\{\omega : f(\omega) < s\} < t\},$$

respectively. The rearrangability of f guarantees that $f_\downarrow$ is real-valued. Furthermore, $f_\downarrow$ is nonnegative, left-continuous and decreasing.[1] We have $f_\downarrow(0+) = f_\uparrow(\mu(\Omega)-) = \|f\|_\infty$ provided that f is essentially bounded. Moreover, the measures $\mu \circ f^{-1}$ and $\ell \circ f_\downarrow^{-1}$ defined on the Borel σ-field of $\mathbb{R}$ (where ℓ is the Lebesgue measure) coincide on all open intervals in $\mathbb{R}$. (This fact is sometimes expressed by saying that f and $f_\downarrow$ are *equimeasurable*.) A useful consequence of this fact is the following equality which is obtained simply by changing the variables:

$$\int_\Omega f \, d\mu = \int_{f(\Omega)} t \; \mu \circ f^{-1}(dt) = \int_{f(\Omega)} t \; \ell \circ f_\downarrow^{-1}(dt) = \int_{f_\downarrow^{-1}(f(\Omega))} f_\downarrow = \int_0^{\mu(\Omega)} f_\downarrow. \tag{1}$$

(Here and throughout we omit giving explicit reference to Lebesgue measure in integration.)

Decreasing and increasing rearrangements are studied extensively within the context of majorization theory, see, for instance, Hardy, Littlewood and Polya (1929), Ryff (1963, 1970), Chong (1974) and Joe (1987). Mitrinović, Pečarić and Fink (1991, Chapter 10) provide a comprehensive survey of inequalities involving rearrangements and their derivatives.

Example 2.1. Let $n \in \mathbb{Z}_{++}$, and suppose $\Omega = \{1, ..., n\}$, $\mathcal{F} = 2^\Omega$ and μ is the counting measure. Then, for any $x \in \mathbb{R}^n_+$ with $f(i) = x_i$, we have

$$f_\downarrow(t) = x_{[\lfloor t \rfloor]}, \quad 0 < t < n,$$

where $x_{[\cdot]}$ is a permutation of x such that $x_{[1]} \geqslant \cdots \geqslant x_{[n]}$. (Here $\lfloor t \rfloor$ denotes the smallest integer at least as large as t.) We shall make extensive use of this example in what follows.

For future reference, we state the following elementary properties of decreasing rearrangements without proof. Of course, analogous properties hold for increasing rearrangements as well.

[1] Throughout this paper, the terms decreasing and increasing will be used to mean non-increasing and non-decreasing, respectively.

Lemma 2.1. *Let $f \in \mathcal{L}_+(\Omega, \mathcal{F}, \mu)$, and let $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ be continuous. Then*

$$\int_\Omega \varphi \circ f \; d\mu = \int_0^{\mu(\Omega)} \varphi \circ f_\downarrow \quad \textit{and} \quad \mu\{f = 0\} = 0 \;\; \textit{implies} \;\; f_\downarrow > 0.$$

Moreover, if φ is strictly increasing and surjective, then $(\varphi \circ f)_\downarrow = \varphi \circ f_\downarrow$.

Lemma 2.2. *For any $f \in \mathcal{L}_+(\Omega, \mathcal{F}, \mu)$ and any $q \in (0, \mu(\Omega))$, we have $\mu\{f > f_\downarrow(q)\} \leqslant q$. Moreover,*

$$\mu\{f > f_\downarrow(q)\} < t \leqslant q \;\; \textit{implies} \;\; f_\downarrow(t) = f_\downarrow(q).$$

3. A SUBMAJORIZATION THEOREM

To study the L^p-norm preorder at a suitably general level, we need to extend the well-known (Tomic-Weil) submajorization theorem (which was originally proved for finite measure spaces) to the case of arbitrary nonnegative rearrangable and measurable real functions.[2] Similar generalizations have occurred in the literature. In particular, Chong (1974) proves a result for nonnegative integrable (rearrangable) functions.[3] Being based on the stochastic dominance approach, our method of proof is slightly more direct than that of Chong's, and yields a more general theorem. In Remark 3.1, however, we shall show that our general submajorization theorem can also be obtained as a consequence of Chong's theorem.

Fix any $a \in \overline{\mathbb{R}}_{++}$, and let ν_1 and ν_2 be any two Borel measures on $[0, a)$. We shall write $\nu_1 \rhd \nu_2$ whenever

$$\int_{[\alpha,a)} (t - \alpha) \; \nu_1(dt) \geqslant \int_{[\alpha,a)} (t - \alpha) \; \nu_2(dt) \quad \text{for all } \alpha \in [0, a).$$

Clearly, $\rhd$ defines a preorder on the class of all Borel measures on $[0, a)$. Define next the class

$$\mathcal{C} := \{\varphi : \mathbb{R}_+ \to \mathbb{R}_+ : \varphi \text{ is increasing and convex}\}$$

and observe that any member φ of this class is right-differentiable (with $\varphi' := \varphi'_+ < \infty$) on its entire domain.

The main result of this section provides a characterization of the classical submajorization ordering in terms of the preorder $\rhd$.

[2]The original Tomic-Weil theorem is, in fact, an easy consequence of Theorem 10 of Hardy, Littlewood and Polya (1929).

[3]There are some relatively minor differences in our definitions of monotonic rearrangements, but these are largely inconsequential.

Theorem 3.1. *For any $f, g \in \mathcal{L}_+(\Omega, \mathcal{F}, \mu)$, the following statements are equivalent:*

(a) $\int_0^q f_\downarrow \geqslant \int_0^q g_\downarrow$ for all $0 < q \leqslant \mu(\Omega)$.

(b) $\int_\Omega \varphi(f)\, d\mu \geqslant \int_\Omega \varphi(g)\, d\mu$ for all $\varphi \in \mathcal{C}$.

(c) $\mu \circ f^{-1} \rhd \mu \circ g^{-1}$.

Remark 3.1. When part (a) of the above theorem holds, we say that f (weakly) *submajorizes* g. It is easily seen that this particular notion of submajorization generalizes the notions of submajorization of finite and infinite sequences (Marshall and Olkin, 1979, p. 10 and p. 17), and that of continuous submajorization (Ryff, 1963, and Marshall and Olkin, 1979, p. 16). If $\int_0^q f_\uparrow \geqslant \int_0^q g_\uparrow$ holds for all $q \in (0, \mu(\Omega))$, we say that f (weakly) *supermajorizes* g. □

The main argument for the proof of Theorem 3.1 is contained in the following lemma which is substantially more general than what is actually needed for the theorem. But this lemma reduces to the well-known stochastic variability theorem in the case where ν_1 and ν_2 are probability measures, and hence it is of independent interest.

Lemma 3.1. *For any $a \in \overline{\mathbb{R}}_{++}$ and any Borel measures ν_1 and ν_2 on $[0, a)$, we have $\nu_1 \rhd \nu_2$ if and only if*

$$\int_{[0,a)} \varphi\, d\nu_1 \geqslant \int_{[0,a)} \varphi\, d\nu_2 \tag{2}$$

for all $\varphi \in \mathcal{C}$ with $\varphi(0) = 0$. Moreover, provided that $\nu_1[0, a) \geqslant \nu_2[0, a)$, we have $\nu_1 \rhd \nu_2$ if and only if (2) holds for all $\varphi \in \mathcal{C}$.

Proof. For any $\alpha, \beta \geqslant 0$, we define $\psi_{\alpha,\beta} \in \mathcal{C}$ by $\psi_{\alpha,\beta}(t) := (\beta t - \alpha)^+$. "If" part of both of the claims follow evaluating (2) for each $\psi_{\alpha,1}$. To prove the converse, let $\mathcal{A}$ be the class of all piecewise linear convex and increasing functions ϕ in $\mathcal{C}$ with $\phi(0) = 0$. Observe that any $\phi \in \mathcal{A}$ can be written as a finite positive linear combination of maps of the form $\psi_{\alpha,\beta}$. Thus, since $\nu_1 \rhd \nu_2$ implies that (2) holds for any $\psi_{\alpha,\beta}$, it follows readily that (2) holds for any $\varphi \in \mathcal{A}$. But it is easy to see that for any $\varphi \in \mathcal{C}$ with $\varphi(0) = 0$, there exists a sequence of maps $\phi_m \in \mathcal{A}$ with $\phi_m \searrow \varphi$. By the monotone convergence theorem, therefore, (2) must hold for all $\varphi \in \mathcal{C}$ with $\varphi(0) = 0$. As for the second claim, let $\nu_1[0, a) \geqslant \nu_2[0, a)$, pick any $\varphi \in \mathcal{C}$ with $\varphi(0) > 0$, and define $\psi(t) := \varphi(t) - \varphi(0)$ for all $t \geqslant 0$. Clearly, $\int_{[0,a)} \varphi\, d\nu_i = \int_{[0,a)} \psi\, d\nu_i + \varphi(0)\nu_i[0, a)$, $i = 1, 2$, and hence (2) follows from the first part of the lemma. □

Proof of Theorem 3.1. (a) ⇒ (b). Since $\mu \circ f^{-1}[0, \infty) = \mu(\Omega) = \mu \circ g^{-1}[0, \infty)$, if we can show that $\mu \circ f^{-1} \rhd \mu \circ g^{-1}$, Lemma 3.1 (with $a = \infty$) will settle this case. To this end, fix an arbitrary $\alpha > 0$, and notice that choosing $q = \mu\{g \geqslant \alpha\}$ in (a) we obtain

$$\int_{\{f_\downarrow \geqslant \alpha\}} (f_\downarrow - \alpha) \geqslant \int_0^{\mu\{g\geqslant\alpha\}} (f_\downarrow - \alpha) \geqslant \int_0^{\mu\{g\geqslant\alpha\}} (g_\downarrow - \alpha) = \int_{\{g_\downarrow \geqslant \alpha\}} (g_\downarrow - \alpha) \tag{3}$$

by definition of a decreasing rearrangement. By changing the variables,

$$\int_{\{f_\downarrow \geqslant \alpha\}} (f_\downarrow - \alpha) = \int_{f_\downarrow^{-1}[\alpha,\infty)} (f_\downarrow - \alpha) = \int_{f_\downarrow(f_\downarrow^{-1}[\alpha,\infty))} (t - \alpha)\, \ell \circ f_\downarrow^{-1}(dt) = \int_{[\alpha,\infty)} (t - \alpha)\, \ell \circ f_\downarrow^{-1}(dt)$$

where the last equality follows from the fact that $f_\downarrow^{-1}([\alpha,\infty)\backslash f_\downarrow(f_\downarrow^{-1}[\alpha,\infty))) = \emptyset$. The same equality also obtains for $g_\downarrow$, and since α is arbitrary in this discussion, we may use (3) to conclude that $\ell \circ f_\downarrow^{-1} \rhd \ell \circ g_\downarrow^{-1}$. By equimeasurability, then, $\mu \circ f^{-1} \rhd \mu \circ g^{-1}$.

(b) $\Rightarrow$ (c). Apply Lemma 3.1 and change the variables.

(c) $\Rightarrow$ (a). By using equimeasurability and changing the variables, we see that (c) implies $\int_{f_\downarrow^{-1}[\alpha,\infty)} (f_\downarrow - \alpha) \geqslant \int_{g_\downarrow^{-1}[\alpha,\infty)} (g_\downarrow - \alpha)$, that is, $\int_0^{\mu\{f>\alpha\}} (f_\downarrow - \alpha) \geq \int_0^{\mu\{g>\alpha\}} (g_\downarrow - \alpha)$ for all $\alpha \geqslant 0$. Now take any $q \in (0, \mu(\Omega))$, and use this observation (with $\alpha = f_\downarrow(q)$) along with Lemma 2.2 to get

$$\int_0^{\mu\{f>f_\downarrow(q)\}} (f_\downarrow - f_\downarrow(q)) \geqslant \int_0^{\mu\{g>f_\downarrow(q)\}} (g_\downarrow - f_\downarrow(q)) \geqslant \int_0^{q} (g_\downarrow - f_\downarrow(q)).$$

Adding $\int_0^q f_\downarrow(q)$ to both sides, and recalling that $f_\downarrow$ is decreasing, we get $\int_0^q f_\downarrow \geqslant \int_0^q g_\downarrow$. $\square$

Remark 3.2. The equivalence of (a) and (b) is proved in the case of nonnegative integrable functions by Chong (1974, Theorem 2.1). While our Theorem 3.1 is more general at face value than Chong's theorem, we should note that the equivalence of (a) and (b) for nonnegative measurable functions can be obtained as a direct corollary of Chong's theorem. This is because if f is such a function, and $\int_\Omega \varphi(f)\, d\mu < \infty$ for some $\varphi \in \mathcal{C}\backslash\{0\}$, then f must be integrable. (Notice that, for any $\varphi \in \mathcal{C}$ such that $\int_\Omega \varphi(f)\, d\mu = \infty$ or $\varphi := 0$, the inequality in (b) holds trivially.) If this was not the case, then $f_\downarrow$ would not be Lebesgue integrable, and hence for $n \in \mathbb{N}$ with $1/n < \mu(\Omega)$, we would have

$$\lim_{n\nearrow\infty} \int_{1/n}^{\min\{n,\mu(\Omega)\}} f_\downarrow = \infty. \tag{4}$$

But by Lemma 2.1 and Jensen's inequality, for all such n,

$$\infty > \int_\Omega \varphi(f)\, d\mu = \int_0^{\mu(\Omega)} \varphi(f_\downarrow) \geqslant \int_{1/n}^{\min\{n,\mu(\Omega)\}} \varphi(f_\downarrow) \geqslant \varphi\left(\int_{1/n}^{\min\{n,\mu(\Omega)\}} f_\downarrow\right)$$

so that, by (4), $\lim_{u\nearrow\infty} \varphi(u) < \infty$. Yet $\varphi \in \mathcal{C}$ guarantees that $\varphi(nu) \geqslant n\varphi(u)$ for all $n \in \mathbb{N}$ and $u \geqslant 0$, and hence we get the contradiction $\lim_{u\nearrow\infty} \varphi(u) = \infty$, given that $\varphi(v) > 0$ for some $v > 0$.

Given this observation, it is enough to prove (a) ⇒ (b) in Theorem 3.1 by assuming that f is integrable, for otherwise (b) holds trivially. If f is integrable, however, (a) guarantees that so is g, and thus Theorem 2.1 of Chong (1974) establishes the claim. Integrability of the functions does not play a role in establishing the converse implication.

4. THE L^p-NORM PREORDER

For any $p \geqslant 1$, we refer to the functional $\|\cdot\|_p : \mathcal{L}_+(\Omega, \mathcal{F}, \mu) \to \overline{\mathbb{R}}_+$ defined by $\|f\|_p = (\int_\Omega f^p d\mu)^{1/p}$ as the *L^p-norm*. In turn, the *L^p-norm preorder* $\succsim \subset \mathcal{L}_+(\Omega, \mathcal{F}, \mu)^2$ is defined as

$$f \succsim g \;\Leftrightarrow\; \|f\|_p \geqslant \|g\|_p \quad \text{for all } p \geqslant 1.$$

As usual, the asymmetric factor of $\succsim$ is denoted by $\succ$, that is, $f \succ g$ iff $f \succsim g$ but not $g \succsim f$.

Remark 4.1. In the definitions above, we do not require the L^p-norm to be finite-valued. However, if we concentrate on the restriction of the ordering $\succsim$ to $L^{1,\infty}_+ \times L^{1,\infty}_+$, where $L^{1,\infty}_+ := L^1_+(\Omega, \mathcal{F}, \mu) \cap L^\infty_+(\Omega, \mathcal{F}, \mu)$, then we may define $\|\cdot\|_p$ as a real-valued functional. This is because $f \in L^{1,\infty}_+$ implies that $\int_\Omega f^p \; d\mu < \infty$ for all $p > 1$. Indeed, since essential boundedness of f guarantees that $f^{p-1} \in L^\infty_+$, by Hölder's inequality, we have $\int_\Omega f^p \; d\mu = \int_\Omega f f^{p-1} \; d\mu \leqslant \|f\|_1 \|f^{p-1}\|_\infty < \infty$ for any $f \in L^{1,\infty}_+$ and any $p > 1$.

Clearly, whenever $f \succsim g$ holds, we understand that f has a larger norm than g, regardless of which particular L^p-norm underlies the comparison. This preorder concept ties in closely with the theory of mean inequalities in the case of finite measure spaces. To see this, we recall that the pth *power mean* of $f \in L^1_+$ is defined as

$$M_p(f) := \left(\frac{1}{\mu(\Omega)} \int_\Omega f^p \; d\mu \right)^{1/p}, \quad p \geqslant 1,$$

provided that $\mu(\Omega) < \infty$. We define the *generalized power mean difference operator* $\mathcal{D} : [1, \infty)^2 \times L^1_+ \times L^1_+ \to \mathbb{R}$ as $\mathcal{D}_{r,s}(f, g) := M_r(f) - M_s(g)$. Two restrictions of the operator $\mathcal{D}$ are worth considering: $D^1_f(r, s) := \mathcal{D}_{r,s}(f, f)$ and $D^2_p(f, g) := \mathcal{D}_{p,p}(f, g)$. Now, it is readily verified that $r \geqslant s$ implies $D^1_f(r, s) \geqslant 0$ for all $f \in L^1_+$. Indeed, by Hölder's inequality,

$$M_s(f) = \left(\frac{1}{\mu(\Omega)} \int_\Omega (f^r)^{s/r} \, 1^{1-s/r} \; d\mu \right)^{1/s} \leqslant \left(\frac{1}{\mu(\Omega)} \left(\int_\Omega f^r \; d\mu \right)^{s/r} \mu(\Omega)^{1-s/r} \right)^{1/s} = M_r(f)$$

whenever $r \geqslant s$. In fact, considerable work has been devoted to finding upper bounds for $D^1_f(r, s)$ in the special case where L^1_+ is a finite Euclidean space; see, for instance, Shisha and Mond (1967), and Bullen, Mitrinović and Vasić (1988), Section III.5.2. On the other hand, the analogous question for D^2 has not yet been satisfactorily answered: what sort of relations between f and g guarantee that $D^2_p(f,g) \geqslant 0$ for all $p \geqslant 1$? (A related question is asked and partially answered for the ratio of $M_p(f)$ and $M_p(g)$ by Marshall, Olkin and Proschan (1967).) Since the purpose of this note is to understand the basic structure of the preorder $\succsim$, and since it is obvious that, for any $f, g \in L^1_+$, $f \succsim g$ holds iff $D^2_p(f,g) \geqslant 0$ for all $p \geqslant 1$, our main results will provide partial answers to this particular question.

The next proposition shows that it is sometimes possible to transform the problem of determining whether or not $f \succsim g$ holds to a simple optimization problem.

Proposition 4.1. *Let $f, g \in L^\infty_+(\Omega, \mathcal{F}, \mu)$ satisfy $\inf_{r \geqslant 1} \|f\|_r > \|g\|_\infty > 0$. Then $f \succsim g$ holds if, and only if,*

$$\min_{1 \leqslant p \leqslant K(f,g)} \int_\Omega (f^p - g^p)\, d\mu \geqslant 0 \quad \textit{where} \quad K(f,g) := \max\left\{1, \frac{\log \mu(\Omega)}{\log \inf_{r \geqslant 1} \left(\|f\|_r / \|g\|_\infty\right)}\right\}. \quad (5)$$

Proof. Necessity is obvious. To see sufficiency, note that $p > K(f,g)$ implies $\inf_{r \geqslant 1} \|f\|^p_r > \|g\|^p_\infty \mu(\Omega)$. Thus,

$$\int_\Omega f^p\, d\mu = \|f\|^p_p > \|g\|^p_\infty \mu(\Omega) \geqslant \int_\Omega g^p\, d\mu, \quad p > K(f,g).$$

Combining this with (5) yields $f \succsim g$. □

Example 4.1. Consider the setting described in Example 2.1. Take any $x, y \in \mathbb{R}^n_{++}$ with $f(i) = y_i$ and $g(i) = x_i$. Assume that $y_{[1]} > x_{[1]}$. It is easily verified that $\inf_{r \geqslant 1} \|f\|_r = y_{[1]}$ and $K(f,g) = (\log n)\left(\log \frac{y_{[1]}}{x_{[1]}}\right)^{-1}$. By Proposition 3.1, therefore, $\sum_{i=1}^n y_i^p \geqslant \sum_{i=1}^n x_i^p$ holds for all $p \geqslant 1$, if, and only if,

$$\min\left\{\sum_{i=1}^n (y_i^p - x_i^p) : \frac{\log n}{\log y_{[1]} - \log x_{[1]}} \geqslant p \geqslant 1\right\} \geqslant 0.$$

Thus, identifying whether or not $\sum_{i=1}^n y_i^p \geqslant \sum_{i=1}^n x_i^p$ holds for all $p \geqslant 1$ can in fact be thought of as solving a particular optimization problem on a compact interval.

In what follows, we will identify certain interesting subrelations of $\succsim$. Our aim is to develop methods that will help identify cases in which two functions in $\mathcal{L}_+$ are ordered by $\succsim$. One important observation in this regard (immediate from Theorem 3.1) is that the submajorization preorder (Remark 3.1) is a proper subrelation of $\succsim$.

Proposition 4.2. *For any $f, g \in \mathcal{L}_+(\Omega, \mathcal{F}, \mu)$, if $\int_0^q f_\downarrow \geqslant \int_0^q g_\downarrow$ holds for all $q \in (0, \mu(\Omega))$, then we have $f \succsim g$.*

Remark 4.2. The converse of Proposition 4.2 does not hold. That is, $\succsim$ is strictly coarser than the submajorization relation. To see this, consider the setting described in Example 2.1 with $n = 3$. Let $x = (12, 12, 2)$, $y = (20, 3, 3)$, and let $f(i) = y_i$ and $g(i) = x_i$. It is immediately observed that $\int_0^2 f_\downarrow = y_1 + y_2 < x_1 + x_2 = \int_0^2 g_\downarrow$. However, as we shall show in Remark 5.1(a), we in fact have $f \succ g$.

5. SUPERMAJORIZATION IN ENTROPY

Theorem 5.1. *Let $f, g \in \mathcal{L}_+(\Omega, \mathcal{F}, \mu)$ satisfy $f_\downarrow \wedge g_\downarrow \geqslant 1$. We have $f \succsim g$, provided that $\|f\| \geqslant \|g\|$ and*

$$\int_0^q f_\downarrow \log f_\downarrow \geqslant \int_0^q g_\downarrow \log g_\downarrow, \quad 0 < q < \mu(\Omega). \tag{6}$$

In words, given that $f_\downarrow$ and $g_\downarrow$ are bounded by 1 from below and given that $\|f\| \geqslant \|g\|$, if the entropy of $f_\downarrow$ (defined as $-f_\downarrow \log f_\downarrow$) supermajorizes the entropy of $g_\downarrow$, then $\|f\|_p \geqslant \|g\|_p$ must hold for all $p \geqslant 1$.

Remark 5.1. (a) Consider the example given in Remark 4.2, where we have claimed that $f \succ g$ holds. To prove this, we note that $\sum_{i=1}^q y_i \log y_i \geqslant \sum_{i=1}^q x_i \log x_i$ for all $q = 1, 2, 3$, and then apply Theorem 5.1 to conclude that $f \succsim g$. Moreover, $g \succsim f$ does not hold since $\|f\|_\infty > \|g\|_\infty$.

(b) Under the domain restriction $f_\downarrow \wedge g_\downarrow \geqslant 1$, Theorem 5.1 is, in fact, strictly stronger than Proposition 4.2. To see this, assume that f is integrable (otherwise $f \succsim g$ follows trivially; see the first paragraph of the proof of Theorem 5.1) and let $\int_0^q f_\downarrow \geqslant \int_0^q g_\downarrow$ for all $0 < q < \mu(\Omega)$. Now fix any $q \in (0, \mu(\Omega))$ and observe that $\int_0^r f_\downarrow \geqslant \int_0^r g_\downarrow$ for all $0 < r \leqslant q$, and $f_\downarrow, g_\downarrow \in L_+^1((0, q], \mathcal{B}, \ell)$. Define next $\varphi(u) = (u \log u)^+$ for all $u \geqslant 0$, and notice that $\varphi \in \mathcal{C}$. Therefore, by Theorem 3.1, we have $\int_0^q f_\downarrow \log f_\downarrow \geqslant \int_0^q g_\downarrow \log g_\downarrow$. Since q is arbitrary in this finding, we may conclude that (6) holds. That is, *if f submajorizes g, then the entropy of f supermajorizes the entropy of g.* (The discrete version of this result follows from Theorem 5.A.2. of Marshall and Olkin, 1979.)

Remark 5.2. (*An application to the theory of probability inequalities*) We now briefly present a probabilistic reading of Theorem 5.1. For any $S \subseteq \mathbb{R}_+$, let $\mathbf{1}_S : \mathbb{R}_+ \to \{0, 1\}$ stand for the indicator function of S. For any $q > 0$, and any nonnegative random variable X with $X \geqslant 1$ a.s., we define the *q-truncated entropy of X* as the random variable

$$e_q(X) := (-X \log X)\mathbf{1}_{(0,q]},$$

where we adopt the convention of setting $e_q(X) = 0$ at all sample points at which X vanishes. Where F is the distribution function of X, let $F_* := 1 - F$, and define the *pseudoinverse* of

F_* as the function $F_*^{-1} : [0,1] \to [1,\infty)$ given by $F_*^{-1}(s) := \inf\{t \geqslant 1 : s \geqslant F_*(t)\}$. One can easily check that F_* is right-continuous and decreasing, and satisfies $-\int_0^q \varphi dF_* = \int_0^q \varphi \circ F_*^{-1}$ for any $q > 0$ and any continuous $\varphi : \mathbb{R}_+ \to \mathbb{R}$. We are now ready to prove the following

Fact. *Let X and Y be two random variables such that $X, Y \geqslant 1$ a.s., and $\mathbf{E}(X) \geqslant \mathbf{E}(Y)$. If $\mathbf{E}(e_q(Y)) \geqslant \mathbf{E}(e_q(X))$ for all $q > 0$, then $\mathbf{E}(X^p) \geqslant \mathbf{E}(Y^p)$ for all $p \geqslant 1$.*

To see this, let F and G be the distribution functions of X and Y, respectively. Observe that

$$\int_0^\infty t dF \geqslant \int_0^\infty t dG \text{ iff } -\int_0^\infty t dF_* \geqslant -\int_0^\infty t dG_* \text{ iff } \int_0^\infty F_*^{-1} \geqslant \int_0^\infty G_*^{-1}$$

so that $\mathbf{E}(X) \geqslant \mathbf{E}(Y)$ implies that $\int_0^\infty F_*^{-1} \geqslant \int_0^\infty G_*^{-1}$. It is similarly observed that $\mathbf{E}(e_q(Y)) \geqslant \mathbf{E}(e_q(X))$ implies $\int_0^q F_*^{-1} \log F_*^{-1} \geqslant \int_0^q G_*^{-1} \log G_*^{-1}$ for each $q > 0$. Therefore, letting $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}_+, \mathcal{B}, \ell)$, $f := F_*^{-1}$ and $g := G_*^{-1}$, and applying Theorem 5.1 yield $\int_0^\infty (F_*^{-1})^p \geqslant \int_0^\infty (G_*^{-1})^p$ for all $p \geqslant 1$. But

$$\int_0^\infty (F_*^{-1})^p \geqslant \int_0^\infty (G_*^{-1})^p \text{ iff } -\int_0^\infty t^p dF_* \geqslant -\int_0^\infty t^p dG_* \text{ iff } \int_0^\infty t^p dF \geqslant \int_0^\infty t^p dG$$

so we get $\mathbf{E}(X^p) \geqslant \mathbf{E}(Y^p)$ for all $p \geqslant 1$.

Proof of Theorem 5.1. Notice first that $f_\downarrow \geqslant 1$ guarantees that $f \geqslant 1$ μ-a.e.. Thus, if $\int_\Omega f\, d\mu = \infty$, then we have $\|f\|_p = \infty$ for all $p \geqslant 1$, and hence $f \succsim g$ obtains trivially. We then assume in what follows that f is integrable. Since $\|f\| \geqslant \|g\|$, g must then also be integrable.

We proceed by means of three easy steps.

(*Step 1*) Define $G : [1,\infty) \to [0,\infty)$ by $G(a) := a \log a$. G is invertible with a strictly increasing inverse, G^{-1}, defined on $[0,\infty)$. For any $p \geq 1$, define next the function

$$\varphi_p(u) := G\left(\left[G^{-1}(u)\right]^p\right), \quad u \geq 0.$$

Clearly, φ_p is a strictly increasing function for any $p \geq 1$. Moreover, by using the inverse function theorem, we find $\varphi_p'(u) = p\left[G^{-1}(u)\right]^{p-1} \xi_p(u)$ where

$$\xi_p(u) := \frac{G'\left(\left[G^{-1}(u)\right]^p\right)}{G'\left(G^{-1}(u)\right)} = \frac{1 + p \log G^{-1}(u)}{1 + \log G^{-1}(u)} = 1 + (p-1)\frac{\log G^{-1}(u)}{1 + \log G^{-1}(u)}$$

for any $u > 0$. Since $t \mapsto t/(1+t)$ is a strictly increasing mapping on $\mathbb{R}_+$ and G^{-1} is strictly increasing, it is evident that, for any $p > 1$, ξ_p is strictly increasing on $(0,\infty)$. It follows that φ_p is strictly convex on $(0,\infty)$. Since φ_p is continuous, we may conclude that $\varphi_p \in \mathcal{C}$ for any $p > 1$.

Now, given that $f_\downarrow \wedge g_\downarrow \geqslant 1$ and G is a continuous and strictly increasing bijection on $[1, \infty)$, by Lemma 2.1 and (6), we have

$$\int_0^q (G \circ f_\downarrow)_\downarrow = \int_0^q G(f_\downarrow) \geqslant \int_0^q G(g_\downarrow) = \int_0^q (G \circ g_\downarrow)_\downarrow ,$$

for all $0 < q < \mu(\Omega)$, that is, $G \circ f_\downarrow$ submajorizes $G \circ g_\downarrow$. But then by Lemma 2.1, we have $\int_0^{\mu(\Omega)} \varphi_p(G(f_\downarrow)) \geqslant \int_0^{\mu(\Omega)} \varphi_p(G(g_\downarrow))$ for all $p \geq 1$. Equivalently,

$$\int_0^{\mu(\Omega)} G(f_\downarrow^p) \geqslant \int_0^{\mu(\Omega)} G(g_\downarrow^p), \quad p \geq 1. \tag{7}$$

(*Step 2*) Fix any $h \in L^1_+(\Omega, \mathcal{F}, \mu)$ with $h_\downarrow(t) \geqslant 1$ for all $t \in (0, \mu(\Omega))$, and define $F_h : [1, \infty) \to \mathbb{R}_+$ as

$$F_h(p) := \int_0^{\mu(\Omega)} h_\downarrow^p.$$

We shall show that F_h is differentiable on $(1, \infty)$ and that its derivative is obtained by differentiation under the integral sign. To see this, fix any p_0 and α such that $\alpha > p_0 > 1$. Define $V : (0, \mu(\Omega)) \times (1, \alpha) \to \mathbb{R}_{++}$ as $V(t,p) = h_\downarrow(t)^p$. Since $h_\downarrow$ is integrable, $V(\cdot, p)$ is also integrable on $(0, \mu(\Omega))$ for any $p > 1$. Moreover, $\frac{\partial}{\partial p} V(t,p)$ obviously exists on $(0, \mu(\Omega)) \times (1, \alpha)$ and equals $h_\downarrow(t)^p \log h_\downarrow(t)$. Finally, we need to show that there exists an integrable function $H : (0, \mu(\Omega)) \to \mathbb{R}_+$ such that $H(t) \geqslant \left| \frac{\partial}{\partial p} V(t,p) \right|$ for all $0 < t < \mu(\Omega)$ and $1 < p < \alpha$. But this is readily obtained by choosing $H := h_\downarrow^\alpha \log h_\downarrow$. Consequently, we can apply the theorem of differentiation under the integral sign (Apostol (1974, Theorem 10.39)) to conclude that F_h is differentiable on $(1, \alpha)$. In particular, F_h is differentiable at p_0 and

$$F_h'(p_0) = \int_0^{\mu(\Omega)} \frac{\partial}{\partial p} V(t, p_0) \, dt.$$

Since p_0 is arbitrary in this demonstration, we can thus conclude that $F_h'(p) = \int_0^{\mu(\Omega)} h_\downarrow^p \log h_\downarrow$ for all $p > 1$.

(*Step 3*) Define $A(p) := \int_0^{\mu(\Omega)} (f_\downarrow^p - g_\downarrow^p)$ for any $p \geqslant 1$. The argument sketched in Step 2 shows that A is differentiable on $(1, \infty)$. By (7) and the definition of G, we have

$$pA'(p) = \int_0^{\mu(\Omega)} \left(f_\downarrow^p \log f_\downarrow^p - g_\downarrow^p \log g_\downarrow^p \right) \geqslant 0$$

so that $A'(p) \geqslant 0$ for all $p > 1$. Given that $A(1) \geqslant 0$, therefore, we find $A(p) \geqslant 0$ for all $p \geqslant 1$. We may thus conclude by using (1) that $\int_\Omega (f^p - g^p) \; d\mu \geqslant 0$, that is, $\|f\|_p \geqslant \|g\|_p$ for all $p \geqslant 1$. □

One shortcoming of Theorem 5.1 is that it only works under a certain domain restriction, namely, it helps one rank functions bounded below by 1 almost everywhere. Whether this restriction can be completely removed from the statement of Theorem 5.1 is at present an open problem. The following two propositions, however, show that one can assume this restriction away at least partially.

Corollary 5.1. *Let* $(\Omega, \mathcal{F}, \mu)$ *be any measure space,* $f, g \in \mathcal{L}_+(\Omega, \mathcal{F}, \mu)$ *satisfy* $\min\{\inf f_\downarrow, \inf g_\downarrow\} > 0$ *and let* $\eta = \max\{1, 1/\inf f_\downarrow, 1/\inf g_\downarrow\}$. *We have* $f \succsim g$, *provided that* $\|f\| \geqslant \|g\|$ *and*

$$\int_0^q f_\downarrow \log[\eta f_\downarrow] \geqslant \int_0^q g_\downarrow \log[\eta g_\downarrow], \quad 0 < q < \mu(\Omega).$$

Proof. Apply Theorem 5.1 to $\eta f_\downarrow$ and $\eta g_\downarrow$. □

Before stating our final result, we need to introduce one last bit of notation. We define, for any $h \in L_+^\infty(\Omega, \mathcal{F}, \mu)$,

$$\sigma(h) := \begin{cases} \sup\{t \in (0, \mu(\Omega)] : h_\downarrow(t) \geqslant 1\}, & \text{if } h_\downarrow(0+) \geqslant 1 \\ 0, & \text{otherwise.} \end{cases}$$

The following result generalizes Theorem 5.1 in the case of finite measure spaces and integrable functions.

Corollary 5.2. *Let* $(\Omega, \mathcal{F}, \mu)$ *be any finite measure space, and let* $f, g \in L_+^1(\Omega, \mathcal{F}, \mu)$ *be positive-valued* μ*-a.e.. We have* $f \succsim g$, *provided that,* $\sigma(f) \geqslant \sigma(g) > 0$,

$$\int_0^q f_\downarrow \log f_\downarrow \geqslant \int_0^q g_\downarrow \log g_\downarrow, \quad 0 < q < \sigma(g),$$

and[4]

$$\int_0^{\sigma(g)} f_\downarrow \geqslant \int_0^{\sigma(g)} g_\downarrow + \mu(\Omega) - \sigma(f). \tag{8}$$

Proof. First notice that if $\sigma(g) = \mu(\Omega)$, we must have $\|f\| \geqslant \|g_\downarrow\| = \|g\|$. By Theorem 5.1, then, we obtain $f \succsim g$. If, on the other hand, $0 < \sigma(g) < \mu(\Omega)$, we define, for any $p \geqslant 1$

$$B(p) := \int_0^{\sigma(g)} \left(f_\downarrow^p - g_\downarrow^p\right) \quad \text{and} \quad C(p) := \int_{\sigma(f)}^{\mu(\Omega)} \left(f_\downarrow^p - g_\downarrow^p\right).$$

Since $\sigma(f) \geqslant \sigma(g)$ guarantees that $\min\{f_\downarrow(\sigma(g)), g_\downarrow(\sigma(g))\} \geqslant 1$, by Theorem 5.1, we have $B(p) \geqslant 0$ for all $p \geqslant 1$. Moreover, the arguments given in Steps 2 and 3 of the proof of

[4]By Lemma 2.1, $\log f_\downarrow$ and $\log g_\downarrow$ are well-defined on $(0, \mu(\Omega))$.

Theorem 5.1 show that B is differentiable on $(1,\infty)$ and $B'(p) \geqslant 0$ for all $p > 1$. Therefore, given that

$$|C(p)| \leqslant \int_{\sigma(f)}^{\mu(\Omega)} \left|f_\downarrow^p - \ g_\downarrow^p\right| \leqslant \mu(\Omega) - \sigma(f)$$

we have, by (8),

$$B(p) \geqslant B(1) \geqslant |C(p)| \quad \text{ for all } p \geqslant 1,$$

which ensures that $B(p) + C(p) \geqslant 0$ for all $p \geqslant 1$. Since $f_\downarrow(t) > 1 > g_\downarrow(t)$ for all $t \in (\sigma(g), \sigma(f))$, we thus find

$$\int_0^{\mu(\Omega)} \left(f_\downarrow^p - g_\downarrow^p\right) \geqslant B(p) + C(p) + \int_{\sigma(g)}^{\sigma(f)} \left(f_\downarrow^p - g_\downarrow^p\right) \geqslant 0, \quad p \geqslant 1.$$

Example 5.1. Consider the setting described in Example 2.1 with $n = 4$. Let $x = (65, 60, 4, 1)$, $y = (100, 20, 10, 1/2)$, and let $f(i) = y_i$ and $g(i) = x_i$. It can be checked that Proposition 4.2, Theorem 5.1 and Corollary 5.1 are not helpful in determining whether or not $f \succsim g$ holds. However, f and g satisfy all of the conditions of Corollary 5.2, and hence we have $f \succsim g$. It is then trivial to establish that $f \succ g$.

6. AN OPEN PROBLEM

Submajorization theorems are found to be very useful in applications, for they provide easy-to-check conditions which identify precisely when the inequality $\int_\Omega \varphi(f)\ d\mu \geqslant \int_\Omega \varphi(g)\ d\mu$ holds for a large class of real functions φ. For instance, Theorem 3.1 provides such conditions for the set of all convex and increasing φ (with two simple boundary conditions that could be dropped in the case of finite measure spaces). It is of interest to see if a similar characterization can be obtained when we require that the said inequality holds for all φ which is of the exponential form: $\varphi(t) = t^p$, $p \geqslant 1$. As the following example illustrates, our present results fall short of providing such a theorem.

Example 6.1. Consider the setting described in Example 2.1 with $n = 3$. Let $x = (9, 7, 3)$ and $y = (10, 5, 5)$, and let $f(i) = y_i$ and $g(i) = x_i$. One can easily check that none of the results reported above is helpful in verifying that $f \succ g$. To see that $f \succ g$ indeed holds, define $K(w, p) := w_1^p + w_2^p + w_3^p$ for all $w \in \mathbb{R}_+^3$ and $p \geqslant 1$. Since it is readily checked that $\sum_{i=1}^h y_i^3 \geqslant \sum_{i=1}^h x_i^3$ for $h = 1, 2, 3$, Theorem 3.1 yields that $K(y, p) > K(x, p)$ for all $p \geqslant 3$. It remains to establish that $K(y, p) > K(x, p)$ for all $p \in [1, 3)$. To this end, define $w = (10, 6, 3)$ and notice that $K(y, p) > K(w, p)$ iff $5^p + 5^p > 6^p + 3^p$ for all $p \in [1, 3)$. But since $(5^3, 5^3)$ strictly supermajorizes $(6^3, 3^3)$, by the dual of the submajorization theorem, we must have $2\varphi(5^3) > \varphi(6^3) + \varphi(3^3)$ for all increasing and concave $\varphi : \mathbb{R}_+ \to \mathbb{R}$. For any $p \in [1, 3)$, by choosing $\varphi(t) := t^{p/3}$, we thus find $5^p+5^p = 2(5^3)^{p/3} > (6^3)^{p/3}+(3^3)^{p/3} = 6^p+3^p$. Consequently, $K(y, p) > K(w, p)$ holds for all $p \in [1, 3)$. But since $\sum_{i=1}^h w_i \geqslant \sum_{i=1}^h x_i$ for $h = 1, 2, 3$, Theorem 3.1 entails $K(w, p) > K(x, p)$ for all $p \geqslant 1$. Combining these two observations, we conclude that $K(y, p) > K(x, p)$ for all $p \in [1, 3)$.

The characterization of $\succsim$ in a way analogous to that of the submajorization relation remains an open problem.

REFERENCES

Apostol, T. M., 1974, *Mathematical Analysis*, Addison-Wesley.

Bullen, P. S., D. S. Mitrinović, and P. M. Vasić, 1988, *Means and their Inequalities*, Kluwer Academic Publishers.

Chong, K-M., 1974, Some extensions of a theorem of Hardy, Littlewood and Polya and their applications, *Canad. J. Math.*, 26, 1321-1340.

Ebert, U., 1988, A family of aggregative compromise inequality measures, *Int. Econom. Rev.,* 29, 363-376.

Hardy, G., J. E. Littlewood and G. Polya, 1929, Some simple inequalities satisfied by convex functions, *Messenger Math.*, 58, 145-152.

Joe, H., 1987, Majorization, randomness and dependence for multivariate distributions, *Ann. Probab.,* 15, 1217-1225.

Foster, J., J. Greer and E. Thorbecke, 1984, A class of decomposable poverty measures, *Econometrica,* 52, 761-766.

Marshall, A. W. and I. Olkin, 1979, *Inequalities: Theory of Majorization and its Applications*, Academic Press.

Marshall, A. W., I. Olkin and F. Proschan, 1967, Monotonicity of ratios of means and other applications of majorization, in *Inequalities*, O. Shisha ed., Academic Press, 177-190.

Mitra, T. and E. A. Ok, 1998, The measurement of income mobility: a partial ordering approach, *Econ. Theory,* 12, 77-102.

Mitrinović, D. S., J. E., Pečarić and A. M. Fink, 1991, *Inequalities Involving Functions and Their Integrals and Derivatives*, Kluwer Academic Publishers.

Ryff, J. V., 1963, On the representation of doubly stochastic operators, *Pacific J. Math.*, 13, 1379-1386.

Ryff, J. V., 1970, Measure preserving transformations and rearrangements, *J. Math. Anal. Appl.*, 31, 449-458.

Shisha, O. and B. Mond, 1967, Differences of means, *Bull. Amer. Math. Soc.*, 73, 328-333.

Tong, Y. L., 1994, Some recent developments on majorization inequalities in probability and statistics, *Linear Algebra Appl.*, 199, 69-90.