Benchmarking

Christopher P. Chambers* and Alan D. Miller†‡

December 7, 2015

Abstract

We introduce a theory of ranking sets of accomplishments in the presence of objectively incomparable marginal contributions (apples and oranges). Our theory recommends *benchmarking*, a method under which an individual is deemed more accomplished than another if and only if she has achieved more benchmarks, or important accomplishments. We show that benchmark rules are characterized by four axioms: transitivity, monotonicity, incomparability of marginal gains, and incomparability of marginal losses. The latter two properties are local, and hence we endogenously derive a concept of a global, objective good, which we call a benchmark.

Introduction

Employers making hiring decisions commonly follow a two-stage process. The pile of candidates is first winnowed according to objective criteria. Difficult cases that remain are then resolved by an executive decision. While the latter choice may be made on the basis of gut feeling, the initial stage tends to follow a more formal process. This may be because objective criteria make it easier to delegate the winnowing decision to administrative staff, or because an executive prefers to rely on more than mere instinct.

There is a general problem in that two candidates' credentials may not be directly comparable. Two economists, Alice and Bob, may have basically the same CV with the following exception. Alice has one more top publication than Bob, and Bob has one more year of excellent teaching

*Department of Economics, University of California, San Diego, 9500 Gilman Drive, #0508, La Jolla, CA 92093-0508. Email: cpchambers@uscd.edu Web: http://econ.ucsd.edu/∼cpchambers/

†Faculty of Law and Department of Economics, University of Haifa, Mount Carmel, Haifa, 31905, Israel. Email: admiller@econ.haifa.ac.il Web: http://econ.haifa.ac.il/∼admiller/

‡Acknowledgments will come in a final version.

evaluations. An employer may not want the administrative staffer charged with winnowing the pile of candidates to make a judgment about the relative value of these marginal accomplishments. What should the staffer be instructed to do when comparing these "apples and oranges?"

We provide a general theory of ranking in the presence of objects whose marginal values are incomparable. Because of the inherent incommensurability, we necessarily seek a ranking that is incomplete. One can view this ranking as an objective "meta-ranking," in which all interested parties agree. We do not model a formal procedure that results in this meta-ranking; rather, we take it as a primitive. The meta-ranking is intended to be an object that can be "completed" in later stages by specific parties.[1] Incompleteness is therefore a normative property, which affords flexibility of completing the ranking in different ways. Our theory is based on two primary axioms that require the ranking to respect the general marginal incommensurability of apples and oranges. These axioms are used to characterize a class of methods that staffers may use when ranking candidates.

To understand our properties, consider two kinds of activities pursued by academic economists: (a) publishing papers and (b) teaching classes. These accomplishments are not comparable in any objective sense, and their marginal value will depend on the context. Given a CV, a research university will place high value on an additional publication but may consider extra teaching experience to be of much less importance. A small teaching college, on the other hand, will value the teaching experience much more highly than the extra publication.[2]

If one must choose between the two, which is more desirable? Our first axiom, *incomparability of marginal gains*, requires that we not make this choice: when both an additional publication and an additional teaching experience add value, these marginal contributions must not themselves be comparable. We generalize the statement to set of accomplishments: if two disjoint sets of accomplishments are added to a given CV, and each result in a strict marginal gain, then these marginal contributions should not be comparable.[3] Different institutions are free to complete this incompleteness in their own way.

Our second axiom is similar in nature and motivation to the first, except in that it is concerned with deletions from (and not additions to) the candidate's résumé. Suppose that both the removal of a publication and the removal of a teaching experience both weaken the candidate. Which

[1]So, the ranking may be viewed as a "Pareto" order over candidates; the ranking over which all interested parties agree. We make this statement precise later in the paper.

[2]At the margin, however, each of these institutions views each of the contributions as net positives.

[3]With non-disjoint sets of accomplishments, we can never be sure that the marginal gains are not due to a common accomplishment. Hence, we would not want to force incomparability in this case.

deletion should weaken the candidate more? *Incomparability of marginal losses* requires that these two marginal losses not be comparable.

To these axioms we add two ancillary conditions. First, *monotonicity* requires that all accomplishments be "goods," in a weak sense. An accomplishment may add value, or not, perhaps depending on what else the candidate has achieved, but it will never harm the candidate. Second, *transitivity* ensures that the method of comparison we derive is a ranking.

The general incommensurability of marginal apples and oranges places a strong restriction on the set of possible ranking methods. Any method consistent with our axioms consists of a fixed list of relevant accomplishments, which we call benchmarks. One candidate is stronger than another if and only if the former has achieved all of the latter's benchmarks. Every set of benchmarks gives us a different ranking method; conversely, any method that respects the incommensurability of apples and oranges must be a benchmarking method. Benchmarks can be interpreted as strict "goods," in an absolute and objective sense. Hence, we have established that any accomplishment which is ever viewed as a marginal good in the presence of other accomplishments must always be a marginal good. Thus, the local properties of incomparability of marginals have global implications about an absolute notion of "good."

To grasp what exactly our axioms are ruling out, consider three candidates, Alice, Bob, and Charles, who are candidates on the academic job market. Alice has taken all requisite courses but has not yet defended her dissertation. Bob has defended his dissertation but has not yet finished his coursework. Charles has both finished all courses and successfully defended his dissertation. Charles qualifies for a PhD, and therefore might be viewed as objectively better than either Alice or Bob. But, from a pragmatic perspective, neither Alice nor Bob can be hired by an academic department, and hence, they may be considered objectively the same as a candidate with a blank CV. Hence, the coursework and the dissertation function as "extreme complements," neither of them have any value on their own to anybody, but together they are valued by everyone.

This type of ranking is ruled out by incomparability of marginal losses. Specifically, removing Charles' coursework results in a strict loss. Removing his dissertation also results in a strict loss. Incomparability of marginal losses requires that these losses are incomparable; specifically, Alice and Bob's CV's must be incomparable. But each of Alice and Bob are equally worthy: they are the same as an individual with no qualifications whatsoever.

A method of resolving this issue would be to reconsider what is deemed an "accomplishment." It seems natural in this case to consider "being qualified for a PhD" to be the accomplishment. Neither completing the coursework or defending the dissertation on its own would be considered an accomplishment.

On the other hand, consider another extreme case. Alice gets a PhD in economics from Alpha University, and Bob gets a PhD in economics from Beta University. Charles gets a PhD in economics from each of those institutions. Each of Alice and Bob are deemed strictly better than a candidate with a blank CV. However, the marginal contribution of Charles' second PhD to either of Alice's or Bob's is deemed irrelevant; he is ranked the same as each of Alice and Bob. As a consequence, Alice and Bob must be ranked, contradicting incomparability of marginal gains.

To resolve this issue and remain consistent with the ranking, we could reconsider the notion of "accomplishment" in a different way: the accomplishment is earning a PhD in economics (irrespective of the institution).

Benchmarking is used in a variety of settings. Universities may use benchmarks in admissions. Investment firms use benchmarks when deciding between projects. Consumers use benchmarks when buying computers. Governments frequently use benchmark rules in assigning priorities in procurement contracts.[4] Schools may be compared among among benchmarks, such as the scores that their students have received in math, science, history, and literature. While this practice is controversial, its' study is important; large sums of federal money are allocated to schools according to these metrics.[5]

A common benchmark used in governmental hiring is the veterans' preference, according to which a military veteran is deemed superior to a non-veteran when they would otherwise be equivalent. Benchmarking may not provide a complete ranking; Alice may be a veteran without work experience, while Bob may be a non-veteran with work experience. But the benchmark rule nonetheless provides valuable comparisons that can be used by decision makers.

Federal courts evaluating the decisions of administrative agencies often use benchmark rules in determining whether to uphold the agency decision. For example, the Federal Communications Commission was held to have acted unreasonably by awarding a television station to a broadcaster who was weaker than another on all relevant criteria used by the commission.[6] Administrative agencies themselves may choose whether particular criteria are important enough to qualify as benchmarks.[7]

[4]For the criteria used in determining eligibility for federal contracts, see 48 C.F.R. Chapter 1.

[5]In particular, the "adequate yearly progress" requirements of the No Child Left Behind act (see 20 U.S.C. 6311) can function like a benchmark rule under which schools are compared to themselves in prior years. Over $14 billion was allocated for these grants in 2014.

[6]*Central Florida Enterprises, Inc. v. Federal Communications Commission*, 598 F.2d 37 (1979).

[7]For an example see *Baltimore Gas & Electric Co. v. Natural Resources Defense*

Benchmarking may be used when comparing scholars by their citation profiles, as in Chambers and Miller (2014b). Here, an accomplishment is a pair of two numbers (x, y) where the individual has x publications with at least y citations each. The step-based indices characterized by Chambers and Miller (2014b) are all benchmark rules, but the reverse is not true. For example, the h-index (Hirsch, 2005) and the i-10 index are two popular measures of scholarly accomplishment; each is a step-based index. However, many believe that multiple such indices should be used in practice. A method of comparison that determines Alice to be better than Bob if she is better according to both measures would not be a step-based index, but would be a benchmark rule. Benchmark rules are more versatile in that they can be applied to a wider array of problems than can be the step-based indices.

The relationship between the step-based indices and the benchmark rules is not a coincidence, but can be seen in the axioms as well. Our axioms imply two properties, intersection and union dominance, which are weaker forms of the lattice-theoretic notions of meet and join homomorphisms used in Chambers and Miller (2014b). These properties were first studied in economics by Kreps (1979) and have since been studied in a wide variety of settings. For example, Hougaard and Keiding (1998), Christensen et al. (1999), and Chambers and Miller (2014a,b) study these axioms in the context of measurement, while Miller (2008), Chambers and Miller (2011), Dimitrov et al. (2012), Leclerc (2011), and Leclerc and Monjardet (2013) study them in the context of aggregation.

Our work is also related to previous literature on incomplete preferences. It is a relatively easy corollary of our main result that $\succeq$ satisfies our axioms if and only if there is a family $\mathcal{R}$ of *complete* relations satisfying our axioms such that for all x, y, $x \succeq y$ if and only if for all $\succeq^* \in \mathcal{R}$, $x \succeq^* y$. This explains our claim that $\succeq$ serves as a Pareto relation for interested parties. Results of this type were pioneered by Dubra et al. (2004) for the expected-utility case. Other such results include Duggan (1999), for the case of general binary relations, Donaldson and Weymark (1998), Dushnik and Miller (1941), and Szpilrajn (1930).[8]

Council, 462 U.S. 87 (1983), which upheld the decision of the Nuclear Regulatory Commission to ignore potential harm from the accidental release of spent nuclear fuel from long term storage.

[8]See also the work of Ok (2002) on incomplete preferences in economic environments, and a choice theoretic foundation for incomplete preferences (Eliaz and Ok, 2006).

The Model

Let A be a set of accomplishments, and let $\leq$ be a partial order on A.[9] A subset $B \subseteq A$ is *comprehensive* if for all $b \in B$ and $a \in A$, $a \leq b$ implies that $a \in B$. Let X be the set of all finite comprehensive subsets of A. The set X forms a lattice when ordered by set inclusion. A **benchmarking rule** is a binary relation $\succeq$ on X for which there exists a set of benchmarks $B \subseteq A$ such that, for all $x, y \in X$, $x \succeq y$ if and only if $B \cap x \supseteq B \cap y$. For two sets $x, y \in X$, we write $x \parallel y$ to denote that x and y are not comparable with respect to $\succeq$; *i.e.* neither $x \succeq y$ nor $y \succeq x$.[10]

For a binary relation $\succeq$ on X, we define the following properties:

Transitivity: For all $x, y, z \in X$, if $x \succeq y$ and $y \succeq z$, then $x \succeq z$.

Monotonicity: For all $x, y \in X$, if $x \supseteq y$, then $x \succeq y$.

Incomparability of Marginal Gains:
For all $x, y \in X$, if $x \succ (x \cap y)$ and $y \succ (x \cap y)$, then $x \parallel y$.

Incomparability of Marginal Losses:
For all $x, y \in X$, if $(x \cup y) \succ x$ and $(x \cup y) \succ y$, then $x \parallel y$.

Transitivity and monotonicity are self-explanatory. Incomparability of marginal gains was described in the introduction; here, imagine we start from a "baseline" set of accomplishments, $(x \cap y)$. Adding the marginal set of accomplishments $(x \backslash y)$ results in x, which is deemed better than $(x \cap y)$. Adding the marginal $(y \backslash x)$ results in y, which is also deemed better than $(x \cap y)$. Since $(x \backslash y) \cap (y \backslash x) = \varnothing$, there are no common accomplishments in the marginals. So, we require that these marginal contributions are unranked, which implies that x and y must be unranked.

Incomparability of marginal losses is similar. Imagine starting from the baseline set of accomplishments $(x \cup y)$. Removing the marginal $(y \backslash x)$ results in x, which is deemed worse than $(x \cup y)$. Similarly, removing the marginal $(x \backslash y)$ results in y, which is deemed worse than $(x \cup y)$. Since $(x \backslash y)$ and $(y \backslash x)$ have no common elements, we require that x and y be unranked.

We use these axioms to prove our main result.

Theorem 1. *A binary relation $\succeq$ on X satisfies transitivity, monotonicity, incomparability of marginal gains, and incomparability of marginal losses if and only if it is a benchmarking rule.*

---

[9] A partial order is a binary relation that is *reflexive*, *transitive*, and *antisymmetric*.

[10] To avoid confusion, we emphasize that there are three important binary relations in this paper: $\leq$, $\succeq$, and $\subseteq$. Incomparability of $\succeq$ is denoted $\parallel$. We need no notation for incomparability of $\leq$ or $\subseteq$.

The proof is given the appendix.

Complete benchmarking rules enjoy additional structure. Formally, a $\leq$-*chain* $C \subseteq A$ is a set for which for all $a, b \in C$, either $a \leq b$ or $b \leq a$.

Completeness: For all $x, y \in X$, either $x \succeq y$ or $y \succeq x$.

Corollary 1. *A benchmarking rule* $\succeq$ *is complete if and only if its associated set of benchmarks* B *is a* $\leq$*-chain.*

Corollary 1 generalizes the earlier results of Chambers and Miller (2014b), which effectively assumes that $\leq$ is the standard order on $\mathbb{Z}^2_+$.[11]

Finally, we present the corollary affording the interpretation of benchmarking rules as Pareto relations of interested parties.

Corollary 2. *For any benchmarking rule* $\succeq$*, there is a family* $\mathcal{R}$ *of complete benchmarking rules for which for all* $x, y \in X$*,* $x \succeq y$ *if and only if for all* $\succeq^* \in \mathcal{R}$*,* $x \succeq^* y$*.*

Examples

We provide three examples of sets that can be compared according to benchmark rules. These sets vary according to the specification of A and $\leq$, and of course, to the interpretation given to them.

Ranking Scholars

Academic institutions often use influence measures to compare scholars in terms of citations to their scientific publications. Popular influence measures include the h-index (Hirsch, 2005), the largest number h such that the scholar has at least h publications with at least h citations each, the $i10$-index, the number of publications with at least ten citations each, and the citation count, the combined number of citations to all of the author's publications.[12] Chambers and Miller (2014b) study a model of influence measures and characterize the family of step-based indices.

Influence measures can be studied in our framework. Let $A \equiv \mathbb{N} \times \mathbb{Z}_+$, the set of pairs of integers (m, n) where m is positive and n is non-negative, and let $\leq$ be the natural order, where $(m, n) \leq (m', n')$ if and only if $m \leq m'$ and $n \leq n'$. The set X of all finite comprehensive subsets of A is equivalent

[11] It is also conceptually related to earlier literature on efficiency measurement; *e.g.* (Hougaard and Keiding, 1998; Christensen et al., 1999; Chambers and Miller, 2014a). This earlier literature works on a space in which there are no nontrivial finite comprehensive sets. Generalizing our results to such environments requires the imposition of continuity properties; we leave this to future research.

[12] These particular measures are widely used, in part, due to their inclusion in the internet service "Google Scholar Profiles", available at http://scholar.google.com/.

to the set of scholars. Benchmark rules are related to the step-based indices characterized in that paper: a relation $\succeq$ is a benchmark rule if and only if there is a collection $\{\succeq_i\}$ of step-based indices for which $x \succeq y$ if and only if $x \succeq_i y$ for all i. A benchmark rule is complete if and only if it can be represented by a single step-based index.

For example, the h-index and the $i6$-index (the number of publications with at least six citations each) are step-based indices and benchmark rules. The benchmarks for these rules are depicted in Figures 1(a) and 1(b) respectively. One can see that, according to the h-index, scholars B and C are equivalent and that each dominates scholar A. According to the $i6$-index, on the other hand, scholar C dominates scholar A, and each of them dominates scholar B. From these two indices we can construct a composite benchmark rule, under which a scholar is at least as good as another if the former scholar is at least as good according to both the h-index and the $i6$-index. This is shown in Figure 1(c). Under this composite rule, scholars A and B are incomparable, but both are dominated by scholar C. In the context of ranking scholars, every collection of step-based indices can generate a benchmark rule in this manner, and every benchmark rule can be generated by some collection of step-based indices.

Evaluating Schools

Under the No Child Left Behind act, schools are compared on the basis of their performance on a set of exams. Let E be a set of exams, and define $A \equiv E \times \mathbb{N}$, where an element (e, n) corresponds to a score of n on exam e. Define $\leq$ so that $(e, n) \leq (e', n')$ if and only if $e = e'$ and $n \leq n'$. A school is a finite comprehensive subset of A; that is, it consists of the schools' score on each exam, and all lesser scores. For example, if a school has received a scores of 2, 0, and 3 on exams e_1, e_2, and e_3, respectively, then the school is represented by the set $\{(e_1, 1), (e_1, 2), (e_3, 1), (e_3, 2), (e_3, 3)\}$. Note that in this example, the score of 2 on exam e_3 is imputed to the school because it received a higher score of 3.

In Figure 2, each exam is represented by a separate vertical line containing the possible scores on that exam. A school takes the form of a finite comprehensive subset, which implies that the school can be represented by its' upper boundary. In the example, the school has achieved scores of 4, 3, and 5 on the respective exams.

Of course, this representation of schools can be applied to any case in which alternatives are compared among multiple dimensions. Academics, for example, may be compared according to their accomplishments in different areas, such as the numbers of papers published, citations received, seminars given, and conferences attended. In this framework, the citation count can satisfy the axioms, although it would not satisfy the axioms in

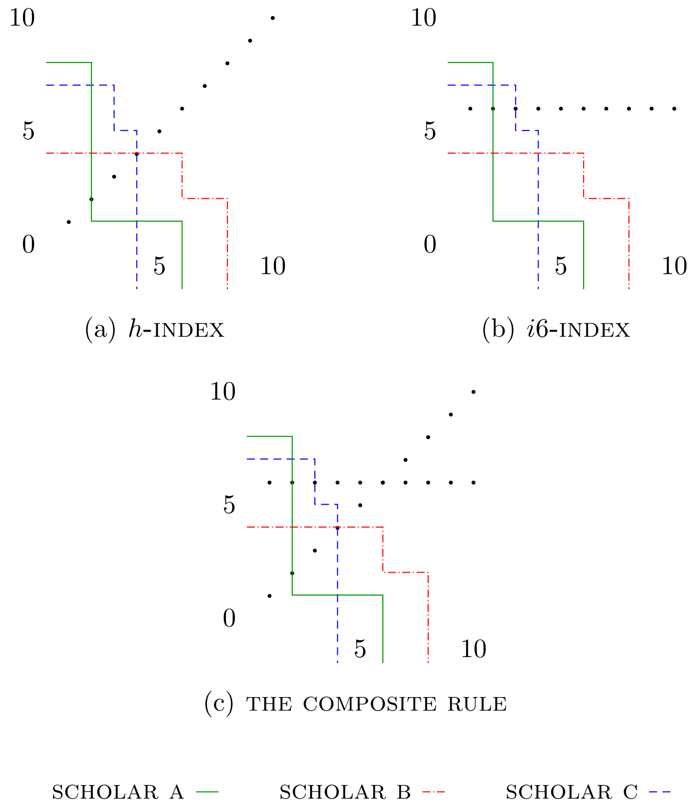

Figure 1: Ranking Scholars.

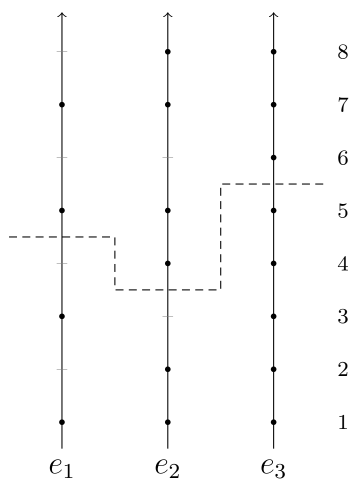

Figure 2: Ranking Schools.

the case of ranking scholars, above. The dimensions may also represent criteria used in multicriterial decision-making.[13]

Unordered Accomplishments

We can also compare all finite subsets of any set A, regardless of comprehensiveness. To do this, we define $\leq$ such that no two distinct elements are comparable (*i.e.*, so that $a \leq b$ implies that $a = b$). A set of benchmarks is pictured in Figure 3. Here, scholars A and B are incomparable, because each contains benchmarks not contained by the other. Similarly, scholars A and C are also not comparable. However, one can see that scholar B dominates scholar C.

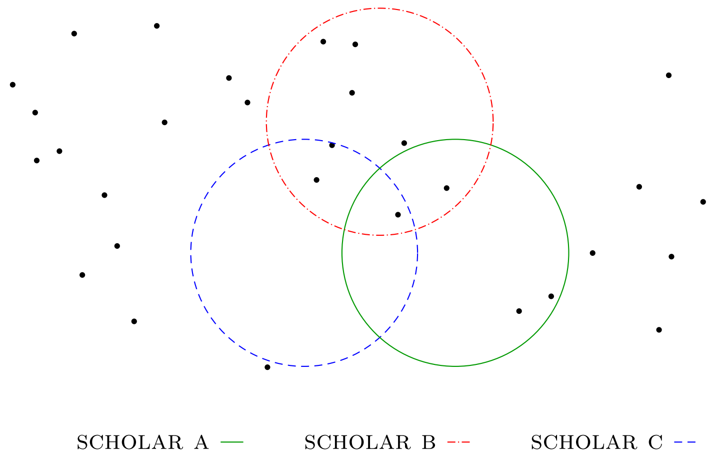

Figure 3: Unordered Accomplishments.

Conclusion

We have described a method of comparison that we term "benchmarking." Benchmark rules are characterized by four axioms, transitivity, monotonicity, incomparability of marginal gains, and incomparability of marginal losses.

Benchmark rules are not necessarily complete, and can be used in cases where completeness is considered undesirable or is otherwise not required. One can see that a benchmark rule will satisfy completeness if and only if the benchmarks are totally ordered. In the case of ranking scholars, this

[13] See Spitzer (1979) and Katz (2011) for applications of multicriterial decision-making in law.

implies that they must form a step-based index. In the case of evaluating schools, it implies that only a single test will matter. In the unordered case, it implies that there can be only one benchmark.

In some cases it may seem as if, in practice, the benchmark rule is simply a comparison according to set-inclusion. This does not necessarily mean that all potential accomplishments are benchmarks. Alternatively, it may be that only benchmarks are included on résumés. This would be expected if the rule were to be known with certainty.

In other cases, however, résumés often include accomplishments that are not benchmarks. One might ask why this would occur. As we noted, this might come across as a result of uncertainty about the rule. However, there are other possibilities. One is that the applicant finds it efficient to use the same résumé for multiple employers, or in multiple markets, where different rules may be applied. Alternatively, as we explained in the introduction, the benchmark rule might be used only as a first step in sorting applicants; the otherwise extraneous information might still be relevant when making an executive decision.

Appendix

We first consider the following two axioms:

Intersection dominance: For all $x, y \in X$, if $x \succeq y$, then $(x \cap y) \succeq y$.

Union dominance: For all $x, y \in X$, if $x \succeq y$, then $x \succeq (x \cup y)$.

Lemma 1. *If $\succeq$ satisfies transitivity, monotonicity, incomparability of marginal gains, and incomparability of marginal losses, then it satisfies both intersection dominance and union dominance.*

Proof of Lemma 1. We shall establish that intersection dominance is satisfied. Union dominance follows from a dual argument. To this end, let $x, y \in X$, and suppose that $x \succeq y$. If either $x \subseteq y$ or $y \subseteq x$, the implication is trivial, so suppose that neither of these are satisfied.

Suppose by means of contradiction that $(x \cap y) \succeq y$ is false. By monotonicity, since $y \succeq (x \cap y)$, it follows that $y \succ (x \cap y)$. By transitivity, we know that $x \succeq y \succ (x \cap y)$, so that $x \succ (x \cap y)$.

Now, we have that both $x \succ (x \cap y)$ and $y \succ (x \cap y)$. Conclude by intersection dominance that $x \parallel y$, a contradiction. $\square$

The proof of Theorem 1 makes use of the concept of an *interior operator*,[14] a function $i : X \to X$ for which the following three properties are satisfied:

[14]An interior operator is a dual concept to that of a closure operator (Ward, 1942).

Contractivity: for all $x \in X$, $i(x) \subseteq x$.

Monotonicity: for all $x, y \in X$, $x \subseteq y$ implies $i(x) \subseteq i(y)$.

Idempotence: for all $x \in X$, $i(i(x)) = i(x)$.

Proof of Theorem 1. **If:** Let $\succeq$ be a benchmarking rule and with benchmarks B. The transitivity of $\succeq$ follows from the transitivity of $\supseteq$. To see this, let $x \succeq y$ and $y \succeq z$. It follows that $B \cap x \supseteq B \cap y$ and $B \cap y \supseteq B \cap z$. Hence $B \cap x \supseteq B \cap z$ and therefore $x \succeq z$. To see that $\succeq$ is monotonic, note that $x \supseteq y$ implies that $B \cap x \supseteq B \cap y$ and therefore that $x \succeq y$. To see that incomparability of marginal gains is satisfied, suppose that $x \succ (x \cap y)$ and $y \succ (x \cap y)$. Since $x \succ (x \cap y)$, there is $b \in B \cap (x \backslash y)$, and since $y \succ (x \cap y)$, there is $b \in B \cap (y \backslash x)$. Hence $x \parallel y$. To see that incomparability of marginal losses is satisfied, suppose that $(x \cup y) \succ x$ and $(x \cup y) \succ y$. Since $(x \cup y) \succ x$, there is $b \in B \cap (y \backslash x)$ and since $(x \cup y) \succ y$, there is $b \in B \cap (x \backslash y)$. Conclude that $x \parallel y$.

Only if: Let $\succeq$ satisfy the four axioms transitivity, monotonicity, intersection dominance, and union dominance. Define, for every $x \in X$, the function

$$i(x) \equiv \bigcap \{z \subseteq x : z \succeq x\}.$$

Notice that because $|\{z \in X : z \subseteq x\}| < +\infty$ for all x, it follows that $0 < |\{z \subseteq x : z \succeq x\}| < +\infty$, so $i(x) \in X$. Observe also that $i(x) = \bigcap \{z : z \succeq x\}$, since for every z such that $z \succeq x$, we know that $x \cap z \succeq x$ (by intersection dominance) and that $x \cap z \subseteq x$. We claim that $i(\cdot)$ is an interior operator satisfying, for all $x, y \in X$, (1) $x \succeq y$ if and only if $i(x) \supseteq i(y)$ and (2) $i(x) \cup i(y) = i(x \cup y)$.

First, the fact that $x \in \{z : z \succeq x\}$ implies that $i(x) = \bigcap \{z : z \succeq x\} \subseteq x$, so $i(\cdot)$ is contractive. Next, we claim that $x \sim i(x)$. To see this, by contractivity and monotonicity of $\succeq$ we know that $x \succeq i(x)$. Furthermore, by intersection dominance, $z \succeq x$ implies $x \cap z \succeq x$, and thus by finite induction it follows that $i(x) = \bigcap \{z \subseteq x : z \succeq x\} \succeq x$.

We prove that $x \succeq y$ if only if $i(x) \supseteq i(y)$. First, let $x \succeq y$. It follows that $\{z : z \succeq y\} \supseteq \{z : z \succeq x\}$ and therefore that $i(x) \supseteq i(y)$. To show the converse, let $i(x) \supseteq i(y)$. By monotonicity of $\succeq$ it follows that $x \sim i(x) \succeq i(y) \sim y$, and thus by transitivity, $x \succeq y$.

To see that $i(\cdot)$ satisfies monotonicity, let $x \supseteq y$. By monotonicity of $\succeq$ it follows that $x \succeq y$, and therefore that $i(x) \supseteq i(y)$. To see that $i(\cdot)$ satisfies idempotence, note that because $x \sim i(x)$ it follows that the sets $\{z : z \succeq z\}$ and $\{z : z \succeq i(x)\}$ coincide.

These objects are intimately connected with the theory of Galois connections and residuated mappings; see, e.g. Blyth and Janowitz (1972). Galois connections have recently found economic application in Nöldeke and Samuelson (2015).

Finally, let $x, y \in X$. Recall that $x \sim i(x)$ and $y \sim i(y)$. By transitivity and monotonicity of $\succeq$, we know that $i(x) \cup i(y) \succeq x$ and $i(x) \cup i(y) \succeq y$. By union dominance and monotonicity of $\succeq$, $y \cup i(x) \cup i(y) \succeq i(x) \cup i(y)$. By transitivity and monotonicity of $\succeq$, we know that $x \succeq y \cup i(x) \cup i(y)$, thus by union dominance and monotonicity of $\succeq$, $x \cup y \cup i(x) \cup i(y) \succeq y \cup i(x) \cup i(y)$. By transitivity, $x \cup y \cup i(x) \cup i(y) \succeq i(x) \cup i(y)$. By contractivity, $x \cup y \cup i(x) \cup i(y) = x \cup y$, hence $x \cup y \succeq i(x) \cup i(y)$. Using the fact that for arbitrary a, $i(a) = \bigcap\{z : z \succeq a\}$, it follows that $i(x \cup y) \subseteq i(x) \cup i(y)$. Lastly $i(x) \subseteq i(x \cup y)$ and $i(y) \subseteq i(x \cup y)$ (from the monotonicity of c), and thus $i(x) \cup i(y) \subseteq i(x \cup y)$. Therefore $i(x) \cup i(y) = i(x \cup y)$.

For a subset $C \subseteq A$, we define $\mathcal{K}(C)$ as its *comprehensive hull*; the smallest comprehensive set containing C. To simplify notation, for an element $a \in A$, we write $\mathcal{K}(a)$ in place of $\mathcal{K}(\{a\})$.

Define $B \equiv \{a \in A : \mathcal{K}(a) = i(\mathcal{K}(a))\}$. We claim that $i(x) = \mathcal{K}(B \cap x)$. To see this, let $a \in \mathcal{K}(B \cap x)$. It follows that there exists $b \in B \cap x$ for which $a \leq b$. Because $b \in x$ it follows that $\mathcal{K}(b) \subseteq x$, and by monotonicity of $i(\cdot)$, $i(\mathcal{K}(b)) \subseteq i(x)$. Because $b \in B$ it follows that $\mathcal{K}(b) = i(\mathcal{K}(b))$ and therefore $b \in i(x)$. Because $i(x)$ is comprehensive and $a \leq b$ it follows that $a \in i(x)$. Conversely, let $a \in i(x)$. Let $b \in i(x)$ such that $a \leq b$ and b on in the comprehensive frontier of $i(x)$, so that $i(x) \setminus \{b\} \in X$. Then $i(x) = \mathcal{K}(b) \cup \mathcal{K}(i(x) \setminus \{b\})$. By idempotence and the join homomorphism property, $i(x) = i(i(x)) = i(\mathcal{K}(b)) \cup i(i(x) \setminus \{b\})$. Because $b \in i(x)$ and $b \notin i(x) \setminus \{b\}$ it follows that $b \in i(\mathcal{K}(b))$ and hence $\mathcal{K}(b) \subseteq i(\mathcal{K}(b))$. By contractivity it follows that $\mathcal{K}(b) = i(\mathcal{K}(b))$ and hence $b \in B$. Because $b \in i(x)$ it follows from contractivity that $b \in x$ and therefore that $b \in B \cap x$. Because $a \leq b$ it follows that $a \in \mathcal{K}(B \cap x)$ and hence $i(x) = \mathcal{K}(B \cap x)$.

It follows then that $x \succeq y$ if and only if $\mathcal{K}(B \cap x) \supseteq \mathcal{K}(B \cap y)$. It remains to be shown that $\mathcal{K}(B \cap x) \supseteq \mathcal{K}(B \cap y)$ if and only if $B \cap x \supseteq B \cap y$. One direction (if) is trivial. To see the converse, let $\mathcal{K}(B \cap x) \supseteq \mathcal{K}(B \cap y)$ and let $a \in B \cap y$. Then $a \in \mathcal{K}(B \cap x)$ and thus there exists $b \in B \cap x$ such that $a \leq b$. By construction $a \in B$, and $b \in x$, and $a \leq b$ implies that $a \in x$ and hence $a \in B \cap x$.

Independence of the Axioms: We describe four rules: each satisfies three axioms while violating the fourth.

Transitivity: Let $\succeq_1, \succeq_2, \succeq_3$ be distinct complete benchmark rules.

Let $\succeq$ be the rule where $x \succeq y$ if and only if $|\{i : x \succeq_i y\}| \geq 2$. Since each of $\succeq_i$ are complete, so is $\succeq$, and it follows that $x \succ y$ if and only if $|\{i : x \succ_i y\}| > 2$.

To see that this rule is monotonic, let $x \supseteq y$. Because the benchmark rules are monotone, it follows that $x \succeq_i y$ for all i and thus $x \succeq y$.

The rule satisfies incomparability of marginal gains because there does not exist $x, y \in X$ such that $x \succ (x \cap y)$ and $y \succ (x \cap y)$. To see this, suppose contrariwise that there does exist $x, y \in X$ such that $x \succ (x \cap y)$

and $y \succ (x \cap y)$. It follows that $|\{i : x \succ (x \cap y)\}| \geq 2$ and $|\{i : y \succ (x \cap y)\}| \geq 2$. Hence, there must exist i for which $x \succ_i (x \cap y)$ and $y \succ_i (x \cap y)$. Because $\succeq_i$ satisfies incomparability of marginal gains it follows that $x \parallel_i y$, contradicting the fact that $\succeq_i$ is complete.

Incomparability of marginal losses follows from a similar argument.

To see that the rule violates transitivity in general, consider $A = \{1, 2, 3\}$, and let $\succeq_i$ be the benchmark rule with set of benchmarks $\{i\}$. Then it is easy to see that $\{1\} \succeq \{1, 2\} \succeq \{1, 2, 3\} \succ \{1\}$.

Monotonicity: Let $\succeq$ be the rule where $x \succeq y$ if and only if $x = y$. Transitivity of $\succeq$ follows from transitivity of $=$. The relation $\succeq$ is reflexive, and thus trivially satisfies incomparability of marginal gains and incomparability of marginal loses. To show that $\succeq$ is not monotone, let $x \subsetneq y$. Then $y \not\succeq x$, a violation of monotonicity.

Incomparability of Marginal Gains: Let $\succeq^*$ be a weak order (complete and transitive) over A (the set of accomplishments). For each $x, y \in X$, let $x \succeq y$ if and only if for each $c \in y$, there is $a \in x$ for which $a \succeq^* c$. Clearly, $\succeq$ is transitive and monotonic. It satisfies incomparability of marginal losses, as there is no $x, y \in X$ for which $(x \cup y) \succ x$ and $(x \cup y) \succ y$. To see this, let $x, y \in X$ and suppose that $(x \cup y) \succ x$ and $(x \cup y) \succ y$. Since $(x \cup y) \succ x$, there is $a \in (y \backslash x)$ for which $a \succ^* c$ for all $c \in x$. And since $(x \cup y) \succ y$, there is $d \in (x \backslash y)$ for which $d \succ^* c$ for all $c \in y$. Then we have $a \succ^* d$ and $d \succ^* a$, a contradiction.

Incomparability of marginal gains is violated in general: Let $A = \{1, 2\}$ where $\leq$ is defined so that no two distinct items are comparable. Suppose that $1 \sim^* 2$. Observe that $\{1\} \succ \varnothing$ and that $\{2\} \succ \varnothing$, yet $\{1\} \sim \{2\}$.

Incomparability of Marginal Losses: Let $\mathcal{B}$ be a family of nested elements of X (a $\subseteq$-chain), so that for any $B, B' \in \mathcal{B}$ either $B \subseteq B'$ or $B' \subseteq B$. Now, define $\succeq$ on X by $x \succeq y$ if and only if for all $B \in \mathcal{B}$, $B \subseteq y$ implies $B \subseteq x$. Observe that $\succeq$ is complete. Further it satisfies monotonicity and transitivity. Incomparability of marginal gains is satisfied because there is no $x, y \in X$ such that $x \succ (x \cap y)$ and $y \succ (x \cap y)$. To see this, let $x, y \in X$ and suppose that $x \succ (x \cap y)$ and $y \succ (x \cap y)$. There is therefore $B_x \in \mathcal{B}$ for which $B_x \subset x$, but B_x is not a subset of $(x \cap y)$. Likewise, there is $B_y \in \mathcal{B}$ for which $B_y \subseteq y$, but B_y is not a subset of $(x \cap y)$. Either $B_x \subseteq B_y$ or $B_y \subseteq B_x$. Without loss of generality, suppose the former. Then we have $B_x \subseteq B_y \subseteq y$, and $B_x \subseteq x$, so that $B_x \subseteq (x \cap y)$, a contradiction.

Incomparability of marginal losses is violated in general: Let $A = \{1, 2\}$ where $\leq$ is defined so that no two distinct items are comparable. Let the family $\mathcal{B}$ consist of $\{\{1, 2\}\}$. Observe that $\{1, 2\} \succ \{1\}$ and $\{1, 2\} \succ \{2\}$, but that $\{1\} \sim \{2\}$.

□

Proof of Corollary 1. Suppose that $\succeq$ is a benchmarking rule, with associated benchmarks B.

First, suppose that B is a $\leq$-chain. We will show that $\succeq$ is complete. Let $x, y \in X$. If $x \sim y$, we are done, so suppose that $(B \cap x) \neq (B \cap y)$. Without loss of generality, let $b \in (B \cap x)$, and suppose that $b \notin (B \cap y)$. Let $c \in (B \cap y)$. Now either $b \leq c$ or $c \leq b$. If the former, then since y is comprehensive, $b \in (B \cap y)$, which is false. Hence $c \leq b$. As x is comprehensive, it follows that $c \in x$. Conclude that $(B \cap y) \subset (B \cap x)$, so that $x \succeq y$.

Conversely, suppose that $\succeq$ is complete, and let $b, c \in B$. Either $\mathcal{K}(b) \succeq \mathcal{K}(c)$, or conversely. Suppose without loss of generality that $\mathcal{K}(b) \succeq \mathcal{K}(c)$. Conclude that $c \in (B \cap \mathcal{K}(c)) \subseteq (B \cap \mathcal{K}(b))$. By definition of $\mathcal{K}(b)$, $c \in \mathcal{K}(b)$ implies that $c \leq b$. $\square$

Proof of Corollary 2. Suppose that $\succeq$ is a benchmarking rule, and let B be the associated set of benchmarks. Let $\mathcal{C}$ be the family of all $\leq$-chains C for which $C \subseteq B$. Each $C \in \mathcal{C}$ induces a complete benchmarking rule $\succeq_C$. Let $\mathcal{R} = \cup\{\succeq_C : C \in \mathcal{C}\}$.

Now, let $x, y \in X$. Suppose that $x \succeq y$. Then $(B \cap y) \subseteq (B \cap x)$. In particular, for all subsets $C \subseteq B$ (chains or not), we have $(C \cap y) \subseteq (C \cap x)$; conclude that $x \succeq_C y$.

Conversely, suppose that for all $C \in \mathcal{C}$, we have $x \succeq_C y$. Let $b \in (B \cap y)$. Consider $\{b\} \in \mathcal{C}$. Clearly $x \succeq_{\{b\}} y$; conclude that $b \in x$. Hence $(B \cap y) \subseteq (B \cap x)$, so that $x \succeq y$. $\square$

References

Blyth, T. and M. Janowitz (1972): *Residuation Theory*, Pergamon Press.

Chambers, C. P. and A. D. Miller (2011): "Rules for Aggregating Information," *Social Choice and Welfare*, 36, 75–82.

——— (2014a): "Inefficiency Measurement," *American Economic Journal: Microeconomics*, 6, 79—92.

——— (2014b): "Scholarly Influence," *Journal of Economic Theory*, 151, 571–583.

Christensen, F., J. L. Hougaard, and H. Keiding (1999): "An axiomatic characterization of efficiency indices," *Economics Letters*, 63, 33–37.

Dimitrov, D., T. Marchant, and D. Mishra (2012): "Separability and Aggregation of Equivalence Relations," *Economic Theory*, 51, 191–212.

Donaldson, D. and J. A. Weymark (1998): "A Quasiordering Is the Intersection of Orderings," *Journal of Economic Theory*, 78, 382 – 387.

Dubra, J., F. Maccheroni, and E. A. Ok (2004): "Expected utility theory without the completeness axiom," *Journal of Economic Theory*, 115, 118–133.

Duggan, J. (1999): "A General Extension Theorem for Binary Relations," *Journal of Economic Theory*, 86, 1 – 16.

Dushnik, B. and E. W. Miller (1941): "Partially Ordered Sets," *American Journal of Mathematics*, 63, 600–610.

Eliaz, K. and E. A. Ok (2006): "Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences," *Games and Economic Behavior*, 56, 61 – 86.

Hirsch, J. E. (2005): "An index to quantify an individual's scientific research output," *Proceedings of the National Academy of Sciences of the United States of America*, 102, 16569–16572.

Hougaard, J. L. and H. Keiding (1998): "On the Functional Form of an Efficiency Index," *Journal of Productivity Analysis*, 9, 103–111.

Katz, L. (2011): *Why the Law Is So Perverse*, University of Chicago Press.

Kreps, D. M. (1979): "A Representation Theorem for "Preference for Flexibility"," *Econometrica*, 47, 565–577.

Leclerc, B. (2011): "Galois Connections in Axiomatic Aggregation," in *Formal Concept Analysis*, ed. by P. Valtchev and R. Jäschke, Springer Berlin Heidelberg, vol. 6628 of *Lecture Notes in Computer Science*, 24–25.

Leclerc, B. and B. Monjardet (2013): "Aggregation and Residuation," *Order*, 30, 261–268, 10.1007/s11083-011-9242-0.

Miller, A. D. (2008): "Group Identification," *Games and Economic Behavior*, 63, 188–202.

Nöldeke, G. and L. Samuelson (2015): "The Implementation Duality," Cowles Foundation Discussion Paper No. 1993.

Ok, E. A. (2002): "Utility Representation of an Incomplete Preference Relation," *Journal of Economic Theory*, 104, 429 – 449.

Spitzer, M. L. (1979): "Multicriteria Choice Processes: An Application of Public Choice Theory to Bakke, the FCC, and the Courts," *The Yale Law Journal*, 88, 717–779.

Szpilrajn, E. (1930): "Sur l'extension de l'ordre partiel," *Fundamenta Mathematicae*, 16, 386—389.

Ward, M. (1942): "The closure operators of a lattice," *Annals of Mathematics*, 43, 191–196.