

IZA DP No. 8344

**And the Winners Are...
An Axiomatic Approach to Selection from a Set**

Daniele Checchi
Gianni De Fraja
Stefano Verzillo

July 2014

And the Winners Are... An Axiomatic Approach to Selection from a Set

Daniele Checchi

*University of Milan
and IZA*

Gianni De Fraja

*University of Nottingham,
Università di Roma "Tor Vergata" and CEPR*

Stefano Verzillo

University of Milan

Discussion Paper No. 8344

July 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

And the Winners Are... An Axiomatic Approach to Selection from a Set^{*}

We study here the problem of evaluating whether the selection from a set is close to the ranking determined by a measurable criterion. We propose a number of natural axioms and show that an index exists that satisfies all these axioms. The paper ends with an application of the index to Italian academia.

JEL Classification: I23, I30

Keywords: selection, merit, ranking

Corresponding author:

Daniele Checchi
University of Milan
Department of Economics
via Conservatorio 7
20138 Milano
Italy
E-mail: daniele.checchi@unimi.it

^{*} We would like to thank Eyal Winter for helpful suggestions.

1 Introduction

Selections are typically made according to a varying blend of objective measures and subjective judgements. A sport coach might pick a team using both the recent performance of individuals in her squad (batting and bowling/pitching averages, tennis rankings, trial times for track and field, and so on), and her sense of who is the best person for a given role in the team, given the expected conditions. Many universities do not strictly follow SATs results and school exam grades, but take into account a student's social background and his potential contribution to desirable characteristics of the student body, like diversity. Applicants for academic jobs might be ranked according to bibliometric measures, but the appointment panel's judgement often leads to decision that do not map precisely into the ranking.¹ Large and complex procurement contracts often demand the subtle evaluation of complex qualitative elements, and lowest price is seldom the only criterion used to award these contracts.² And so forth and so on. Notice that selection is a broader concept than full ranking: choosing the all-time best 20 in a list of 1000 footballers is different and simpler than ranking the best 20.

Observers and decision makers might be interested in some means of comparing the choices of different selectors. If, for want of a better term, the property of following the measurable dimension is labelled "metric-basedness" someone might want to compare different selections, and determine which is more "metric based". For example, a cricket fan may want to know whether Australia's selection for the Ashes team is more metric-based than England's. Or whether it is more metric-based than it was sixty years ago. A university whose admission policies are under scrutiny in court may want to argue that

¹An example from recent implementation of policy which some readers will be familiar with is the extent by which bibliometric criteria should be used in the evaluation of university research departments. Unlike in Italy, in the UK funding body was persuaded to allow panels not to adhere strictly to bibliometric measures of departmental output, but allow the latitude afforded by peer review. We reprise this theme in Section 4.

²As, for instance, the 1991 auctions for the 16 regional television franchises in the UK, when only half the franchises were awarded to the highest bidder; see Cabizza and De Fraja (1998), especially Table 1, pp 11-12.

its selections are as based on SAT's as those of comparable institutions. Or a government minister concerned about corruption in awarding procurement contracts or personnel hiring may want to compare the "metric basedness" of various commissioning boards or hiring panels. Finally, the punchline of the popular book and film *Moneyball* (Lewis 2004) is that by being obsessively "metric based" in its selection, a baseball team with relatively scarce financial resources can systematically outperform its much wealthier rivals; Hakes and Sauer (2006) confirm econometrically the book's intuition.

Comparing selections is straightforward only in the starkest cases. Sure, the selection of the best ranked is unquestionably more metric-based than one that selects a different element from the pool. But is a university which, from its 100 applicants, admits as students the second, eighth, ninth, twenty-second, and thirtieth-ranked more metric-based than one that chooses the first five, and those ranked between 77 and 81 out of 200 applicants? Or in an even simpler example, is picking the second ranked out of ten candidates for a job more metric-based than selecting the third ranked out of twenty?

This paper suggests an axiomatic approach to comparing selections. In the next section a set of minimal axioms is proposed: they are all variations on the idea of dominance in the comparison of sets (Barberà et al 2004): in practice, each of these axioms takes a binary comparison between selections where the answer is unambiguous to the question as to which of the two selections is more metric based, and determines in accordance their relative metric-basedness. In Section 3 we provide operational content to these axioms, by constructing an index that satisfies them all. Because these axioms are intended to be minimal, they necessarily define an *incomplete* ordering: there are pairs of selection that cannot be compared on the basis of the axioms, and so the index we propose is not the only possible one: different indices satisfy all the axioms, and give a different answers to the comparison of the "metric-basedness" of two given selections from sets. This is due to the different dimensionality of the set of possible selections from ranked sets: mathematically, the problem is to map selections taken from ordinally ranked sets into an interval of the real line, which clearly cannot be done bijectively. This is deliberate, the aim of this paper being the statement of the minimal

axioms and to show that even these minimal, unobjectionable requirements have some bite, in the sense that they exclude alternative indices that violate at least one of them.

The paper is simply organised: the axioms are stated in the next section, and Section 3 presents the index and demonstrates that it satisfies the axioms proposed. The paper ends with Section 4, which showing how the index can be used to compare the evaluation mechanisms for promotion to professorship in Italian universities.

2 Axioms of “metric-basedness”

Consider a set \mathcal{N} with $N > 1$ elements, and a binary relation $R \subseteq \mathcal{N} \times \mathcal{N}$, which we interpret as “ranking according to the metric considered”: that is if $j_1, j_2 \in R$, which we write compactly as $j_1 R j_2$, then j_1 is ranked at least as well³ as j_2 . R is assumed to be reflexive, complete, transitive and antisymmetric:⁴ thus R is a linear ordering (Barberà et al 2004, p 903).

We next define a selection \mathcal{K} as a proper and non-empty subset of \mathcal{N} . Let K be the number of elements of \mathcal{K} , $K \in \{1, \dots, N - 1\}$. We define the pair $(\mathcal{N}, \mathcal{K})$ a “selected set”.⁵ The set of all selected sets of \mathcal{N} , \mathfrak{S}_N , is the power set of the set \mathcal{N} minus the empty set and the set \mathcal{N} itself. \mathfrak{S}_N has $2^N - 2$ elements. Finally, we let $\mathfrak{S} = \prod_{N \in \mathbb{N} \setminus \{1\}} \mathfrak{S}_N$ be the set of all selected sets.

The question we ask is the following: given two selected sets, $(\mathcal{N}_A, \mathcal{K}_A)$ and $(\mathcal{N}_B, \mathcal{K}_B)$ in \mathfrak{S} , which is more metric-based, in the sense of being a selection “closer” to the actual ranking of that set?

To formalise this question we define a binary relation $M \subseteq \mathfrak{S} \times \mathfrak{S}$, which

³As a matter of terminology, we use best and worst ranked element, rather than highest and lowest, given the potential linguistic ambiguity that the lowest/highest terminology could determine, where the lowest number is associated with the highest ranked element.

⁴Or “strict”: given any $j_1, j_2 \in \mathcal{N}$, either $j_1 R j_2$ or $j_2 R j_1$, but not both. It is straightforward to extend the theory to rankings where some elements of the set are ranked equally. This would complicate the presentation and add little of substance, and given the unexplored topic of this paper, we feel that clarity should take precedence over generality.

⁵Mathematically, selected-sets are the proper K -subsets of \mathcal{N} (Mazur, 2010 p 5).

we interpret as “metric-basedness”. $(\mathcal{N}_A, \mathcal{K}_A), (\mathcal{N}_B, \mathcal{K}_B) \in M$ if and only if $(\mathcal{N}_A, \mathcal{K}_A)$ is at least as metric based as $(\mathcal{N}_B, \mathcal{K}_B)$. For simplicity we write $(\mathcal{N}_A, \mathcal{K}_A) \succeq_M (\mathcal{N}_B, \mathcal{K}_B)$, if and only if $(\mathcal{N}_A, \mathcal{K}_A), (\mathcal{N}_B, \mathcal{K}_B) \in M$. “Strict metric-basedness”, \succ_M , is naturally defined: $(\mathcal{N}_A, \mathcal{K}_A)$ is strictly more metric based than $(\mathcal{N}_B, \mathcal{K}_B)$ if $(\mathcal{N}_A, \mathcal{K}_A) \succeq_M (\mathcal{N}_B, \mathcal{K}_B)$ and not $(\mathcal{N}_B, \mathcal{K}_B) \succeq_M (\mathcal{N}_A, \mathcal{K}_A)$.

We require that the metric-basedness relation satisfies a number of axioms, which are all specifications in this more complex set-up of the *dominance axiom*, which “requires that adding an element which is better (worse) than all elements in a given set \mathcal{N} according to R leads to a set that is better (worse) than the original set” (Barberà et al 2004, p 905).

Having defined these axioms, we propose an index, which maps every possible selected set $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ into the unit interval of the real line, in such a way that if $(\mathcal{N}_A, \mathcal{K}_A) \succeq_M (\mathcal{N}_B, \mathcal{K}_B)$, then the value of the index associated to $(\mathcal{N}_A, \mathcal{K}_A)$ is higher than the value associated to selected set $(\mathcal{N}_B, \mathcal{K}_B)$: if the selection $(\mathcal{N}_A, \mathcal{K}_A)$ is more metric-based than the selection $(\mathcal{N}_B, \mathcal{K}_B)$ the value of the index at $(\mathcal{N}_A, \mathcal{K}_A)$ is higher than the value of the index at $(\mathcal{N}_B, \mathcal{K}_B)$.

To assist with the interpretation of the axioms, we introduce a less formal description of the comparison among selected sets. We begin by defining changes to the set \mathcal{N} and the selection \mathcal{K} . We consider two types of changes: the first type maintains the size of the set \mathcal{N} , N , unchanged, but modifies the selection \mathcal{K} .

Definition 1 *Given a selected set $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$, an “IN-SWITCH” is the inclusion in the selection of an element of \mathcal{N} not in \mathcal{K} . An “OUT-SWITCH” is the removal from the selection \mathcal{K} of an element of \mathcal{N} .*

The second type of change modifies the size of the set \mathcal{N} .

Definition 2 *Given a selected set $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$, an “ADD” is the addition a new element of \mathcal{N} . An “ADD” is an “ADD-IN” if the new element is in \mathcal{K} , an “ADD-OUT” if it is not in \mathcal{K} .*

We note that given any two selected sets $(\mathcal{N}_A, \mathcal{K}_A)$ and $(\mathcal{N}_B, \mathcal{K}_B)$ in \mathfrak{S} , it is possible to repeatedly apply the two changes described in Definitions 1 and 2 to selected set $(\mathcal{N}_A, \mathcal{K}_A)$ and transform it into selected set $(\mathcal{N}_B, \mathcal{K}_B)$.⁶ Before proceeding, we single out the situation where the selection follows exactly the ranking.

Definition 3 *Given a selected set $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$, the selection \mathcal{K} is “PERFECT” if no selected element has a rank worse than a non-selected element. The selection \mathcal{K} is “ANTIPERFECT” if every selected element has a worse rank than every non-selected element.*

We are now ready to make explicit the concept of dominance; each of the axioms below is a “natural” requirement of the comparison between selections, when both the selection and the set from which the selection is made can change in composition. The first two consider an increase in the size of the selection, K , keeping the size of the set, N , constant.

Axiom 1 *For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \in \mathcal{N} \setminus \mathcal{K}$ such that yRx for all $y \in \mathcal{N} \setminus \mathcal{K}$, $(\mathcal{N}, \mathcal{K}) \succsim_M (\mathcal{N}, \mathcal{K} \cup \{x\})$.*

That is, the in-switch of the worst ranked non-selected element in a selected set $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ makes the selection less metric-based; the axiom can be naturally extended to require that this is strictly unless the selection is antiperfect.

Axiom 2 *For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \in \mathcal{N} \setminus \mathcal{K}$ such that xRy for all $y \in \mathcal{N} \setminus \mathcal{K}$, $(\mathcal{N}, \mathcal{K} \cup \{x\}) \succsim_M (\mathcal{N}, \mathcal{K})$.*

In words, the in-switch of the best ranked non-selected element in a selected set $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ makes the selection more metric-based; again, this can be strictly so unless the selection is perfect.

⁶Clearly, the switch and add operations are not defined when the result of the operation does not belong to \mathfrak{S} (for example, it is not possible to add-in to the selected set $(\mathcal{N}, \mathcal{K})$ if \mathcal{K} has $N - 1$ elements). As there is no danger of confusion, to avoid unnecessary burdensome terminology, we leave this detail implicit.

In Axioms 1 and 2, the size of the set is not changed, while the number of selected elements increases. In the next two axioms, the converse is considered: fewer elements are selected while still keeping constant the size of the set \mathcal{N} .

Axiom 3 For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \in \mathcal{K}$ such that yRx for all $y \in \mathcal{K}$, $(\mathcal{N}, \mathcal{K} \setminus \{x\}) \succsim_M (\mathcal{N}, \mathcal{K})$.

Axiom 4 For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \in \mathcal{K}$ such that xRy for all $y \in \mathcal{K}$, $(\mathcal{N}, \mathcal{K}) \succsim_M (\mathcal{N}, \mathcal{K} \setminus \{x\})$.

That is, the out-switch of the worst (best) ranked selected element in a selected set $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ makes the selection more (less) metric-based; strictly unless the selection is perfect (antiperfect).

To see these four axioms “in action”, take a typical selected set:

$$11000011000010001000 \tag{1}$$

where elements are ranked from best to the left, to worse to the right, and a “1” in the j -th position indicates that the j -th element is selected. The four axioms are illustrated by the following changes in the selected set (the element different is in bold type):

Axiom	Change in the selected set (1)	New selected set
1	in-switch of the worst non-selected element	1100001100001000100 1
2	in-switch of the best non-selected element	11 100011000010001000
3	out-switch of the worst selected element	1100001100001000 0 000
4	out-switch of the best selected element	0 1000011000010001000

Like Axioms 1-4, Axiom 5 alters the selection, without changing the size of the set, but unlike them, which apply only to the best and to the worse elements of the selection \mathcal{K} and its complement $\mathcal{N} \setminus \mathcal{K}$, Axiom 5 does not change the size of the selection and it applies to any change in the selection.

Axiom 5 For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \in \mathcal{N} \setminus \mathcal{K}$ and $y \in \mathcal{K}$, xRy if and only if $(\mathcal{N}, \mathcal{K} \cup \{x\} \setminus \{y\}) \succsim_M (\mathcal{N}, \mathcal{K})$.

In words, an in-switch accompanied by an out-switch makes the selected set strictly more (less) metric-based if the rank of the in-switched element is better (worse) than the rank of the out-switched element. Less formally, swapping a selected element with a non-selected one improves the metric-basedness of the selection if and only if the newly selected element is better ranked than the newly excluded one.⁷

So far we have kept constant the size of the set \mathcal{N} ; the last four axioms add a new element to the set.

Axiom 6 For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \notin \mathcal{N}$ such that xRy for all $y \in \mathcal{K}$, $(\mathcal{N} \cup \{x\}, \mathcal{K} \cup \{x\}) \succeq_M (\mathcal{N}, \mathcal{K})$.

Axiom 7 For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \notin \mathcal{N}$ such that xRy for all $y \in \mathcal{K}$, $(\mathcal{N}, \mathcal{K}) \succeq_M (\mathcal{N} \cup \{x\}, \mathcal{K})$.

Axiom 8 For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \notin \mathcal{N}$ such that yRx for all $y \in \mathcal{K}$, $(\mathcal{N}, \mathcal{K}) \succeq_M (\mathcal{N} \cup \{x\}, \mathcal{K} \cup \{x\})$.

Axiom 9 For all $(\mathcal{N}, \mathcal{K}) \in \mathfrak{S}$ and $x \notin \mathcal{N}$ such that yRx for all $y \in \mathcal{K}$, $(\mathcal{N} \cup \{x\}, \mathcal{K}) \succeq_M (\mathcal{N}, \mathcal{K})$.

Again these are natural requirements. In words, an add-in with a rank better (worse) than every selected element makes the selected set more (less) metric-based; strictly unless the final selection is antiperfect (perfect), Axioms 6 and 8. And an add-out with a better (worse) rank than every selected element makes the selected set less (more) metric-based; strictly unless the initial selection is antiperfect (perfect), Axioms 7 and 9. Put differently, adding to the set \mathcal{N} and to the selection \mathcal{K} a new element which is better (worse) than all the selected elements makes the selection more (less) metric-based. Conversely, adding to the set \mathcal{N} but not to the selection \mathcal{K} a new element which is better (worse) than all the selected elements makes the selection less (more) metric-based.

⁷One could make an analogy with the Dalton-Pigou principle (Dalton, p 351); a transfer of a resource (being selected in our case, or income in Dalton's) from a worse ranked/richer to a better ranked/poorer element/person, so long as that transfer does not reverse the ranking of the two, will result in greater metric-basedness/equity.

3 An index of “metric-basedness”

The axioms proposed in the above Section aim to provide a minimum set of principles which are not subject to plausible immediate objections, and so are not intended to define a complete relation. In this section we take the opposite view, and we propose, through the construction of an index of “metric-basedness”, a way to compare *any two* selected sets. Relatively simple examples show that it is in principle difficult to construct a complete ordering. Consider the following two selected sets:

01001110010000010000000100010000000
001001010110000000000

In the first selection, there are some selected elements among the best ranked, but there are also some below the median. The second, smaller, selection picks only elements above the median, but quite close to it: different people might well have different views as to which of the two above selections is more metric-based, implying a degree of arbitrariness in any index. Nevertheless, any index which violates one of the Axioms 1-9 proposed above would give the wrong answer at least in some cases, and therefore should not be considered a measure of metric-basedness. Formally, given any selected set $(\mathcal{N}, \mathcal{K})$, in this section we assign to this set a real number

$$M_{(\mathcal{N}, \mathcal{K})},$$

in such a way that if selected set $(\mathcal{N}_A, \mathcal{K}_A)$ is more metric-based than selected set $(\mathcal{N}_B, \mathcal{K}_B)$, then the index assigned to $(\mathcal{N}_A, \mathcal{K}_A)$ is higher than that assigned to $(\mathcal{N}_B, \mathcal{K}_B)$: $M_{(\mathcal{N}_A, \mathcal{K}_A)} > M_{(\mathcal{N}_B, \mathcal{K}_B)}$.

The following normalisation is convenient, though not logically necessary.

Criterion 1 *The index $M_{(\mathcal{N}, \mathcal{K})}$ is non-negative, and takes value 1 if and only if the selection from the set $(\mathcal{N}, \mathcal{K})$ is perfect, and takes value 0 if and only if the selection from the set $(\mathcal{N}, \mathcal{K})$ is antiperfect.*

The main hinge around which the index we propose is built is the sum of the ranks of the selected elements. Given a selected set $(\mathcal{N}, \mathcal{K})$, define an

indicator function $\delta_{(\mathcal{N}, \mathcal{K})} : \{1, \dots, N\} \rightarrow \{0, 1\}$ in such a way that for every $j = 1, \dots, N$, $\delta_{(\mathcal{N}, \mathcal{K})}(j) = 1$ if and only if $j \in \mathcal{K}$, that is if and only if the j -th ranked element of \mathcal{N} is selected. The sum of the ranks of the selected elements is given by:

$$r = \sum_{j \in \mathcal{K}} j = \sum_{j=1}^N \delta_{(\mathcal{N}, \mathcal{K})}(j) j \quad (2)$$

Notice that the minimum (maximum) value of r occurs when all the selected elements are the best (worst) ranked elements, that is when the selection is perfect (antiperfect). Formally, if the minimum and the maximum value of r are denoted as r_{\min} and r_{\max} , then

$$r_{\min} = \sum_{j=1}^K j = \frac{K(K+1)}{2}$$

$$r_{\max} = \sum_{j=N-K+1}^N j = \sum_{j=1}^N j - \sum_{j=1}^{N-K} j = \frac{K(2N-K+1)}{2}$$

Here is the promised index.

$$M_{(\mathcal{N}, \mathcal{K})} = 1 - \frac{r - r_{\min}}{r_{\max} - r_{\min}} = \frac{r_{\max} - r}{r_{\max} - r_{\min}}$$

$$M_{(\mathcal{N}, \mathcal{K})} = \frac{\frac{K(2N-K+1)}{2} - \sum_{j=1}^N \delta_{(\mathcal{N}, \mathcal{K})}(j) j}{K(N-K)} \quad (3)$$

We can now establish that the index satisfies the Axioms given above.

Proposition 1 *The index $M_{(\mathcal{N}, \mathcal{K})}$ satisfies Axioms 1-9 and Criterion 1.*

Proof. Criterion 1 is obvious: when r takes the lowest possible value, that is, when only the K best ranked elements of \mathcal{N} are selected, then $M_{(\mathcal{N}, \mathcal{K})} = 1$: the selection is fully metric-based. Conversely, when only the K worst ranked elements are selected, then we have $r = r_{\max}$ and so $M_{(\mathcal{N}, \mathcal{K})} = 0$: the selection is the exact opposite of being metric-based.

Consider the Axioms next, beginning with Axiom 1. Let the worst ranked non-selected element be ranked z . In-switching it, increase r by z and K by 1. So the difference in the value of the index is

$$\frac{\frac{(K+1)(2N-(K+1)+1)}{2} - (r+z)}{(K+1)(N-(K+1))} - \frac{\frac{K(2N-K+1)}{2} - r}{K(N-K)},$$

which can be written as

$$\frac{(N+1)\frac{(K+1)}{2} + (N-1-2K)\frac{r}{K} - (N-K)z}{(K+1)(N-K-1)(N-K)};$$

we want to establish that the above is negative. The denominator is positive, and so we need

$$(N+1)\frac{(K+1)}{2} + (N-1-2K)\frac{r}{K} - (N-K)z < 0. \quad (4)$$

If K is less than $\frac{N-1}{2}$, then the highest possible value for the LHS occurs when the selection is antiperfect: r reaches its maximum value, r_{\max} , and z reaches its minimum value, $N-K$. In this case the RHS in (4) becomes 0, and so except in this extreme case it is strictly negative. If $K = \frac{N-1}{2}$ (4) becomes $\frac{1}{4}(N+1)(N-2z+1) < 0$, which is true as $z > N-K = \frac{N+1}{2}$ (unless the selection is antiperfect). Consider next $K > \frac{N-1}{2}$. In this case, the worst case occurs when the selection is perfect with $r = r_{\min}$ and so $z = N$, and (4) becomes $-(N-K)(N-K-1) < 0$, which holds.

Consider Axiom 2 next. When the best ranked non-selected element is in-switched, K increases by 1 and r also increases. If the selection is perfect, then $r = r_{\min}$ and $z = K+1$, and $M_{(\mathcal{N}, \mathcal{K})} = M_{(\mathcal{N}, \mathcal{K} \cup \{K+1\})} = 1$, and so their difference is 0. If the selection is not perfect, then r is higher and z is lower, increasing the difference, as long as K is lower than $\frac{N-1}{2}$. Suppose therefore that K exceeds $\frac{N-1}{2}$. In this case, the worse value of r is r_{\max} , and the worse value of z is $N-K$: note that these values are not compatible with each other: if one reaches its maximum, the other cannot. But even in this theoretical worse scenario, the difference is 0.

Axiom 3 next. The out-switch of the worst ranked selected element decreases K by 1 and decreases r by z , and so the converse argument applies which was used to establish Axiom 2. And similarly for Axiom 4, which mirrors Axiom 1.

We consider Axiom 5 next. This is straightforward: an in-switch accompanied by an out-switch changes neither K nor N . It only changes r , and so clearly the index M increases if r decreases, that is if a better ranked element takes the place of a worse ranked one in the selection.

Axiom 6 next. The value of M changes from

$$\frac{\frac{K(2N-K+1)}{2} - r}{K(N-K)}$$

to

$$\frac{\frac{(K+1)(2(N+1)-(K+1)+1)}{2} - r - w}{(K+1)((N+1)-(K+1))}$$

where w is the rank of the new selected element. The difference is

$$\frac{\frac{1+K}{2} + \frac{r}{K} - w}{-(N-K)(K+1)} \quad (5)$$

which is negative if w is less than every previously selected element, $\frac{r}{K}$, and is positive.

For the next axiom, N increases by 1, and r by K (because each selected element's rank increases by 1). The difference between the new index and the previous value is $\frac{1}{-2K} \frac{K(2N+1-K)-2r}{(N+1-K)(N-K)}$; the maximum value of r is r_{\max} , and in this case the difference is 0. For lower values of r it is negative, as required. Next Axiom 8: adding a new non-selected element worse than all the selected elements changes N by 1, does not change K , and does not change r the index goes $\frac{\frac{K(2N-K+1)}{2} - r}{K(N-K)}$ to $\frac{\frac{K(2(N+1)-K+1)}{2} - r}{K((N+1)-K)}$ and the difference is $\frac{\frac{K(K+1)}{2} - r}{(N+1-K)(N-K)K}$, which is negative if $r > r_{\min}$. That the index $M_{(\mathcal{N}, \mathcal{K})}$ satisfies Axiom and 9 is established simply by differentiating the expression $\frac{\frac{K(2N-K+1)}{2} - r}{K(N-K)}$ with respect to N , given that neither K nor r change.

$$\frac{d}{dN} \left(\frac{\frac{K(2N-K+1)}{2} - r}{K(N-K)} \right) = \frac{r - r_{\min}}{K(N-K)^2}$$

which is true, unless the selection is perfect. ■

While we have not been able to establish it formally, extensive simulations suggest that when the selection is completely random, then the index M tends to take value $\frac{1}{2}$, for every N and for every K . That is, given random selection, the distribution of the values of the index M for a large number of repetition tends to a symmetric β distribution with mean $\frac{1}{2}$.

4 An application: metric-basedness in Italian universities

The paper ends with a concrete application of the proposed index, inspired by two recent assessment exercises carried out in the Italian university system, one assessing the research activities of every Italian university department, and the other determining which Italian assistant and associate professors were qualified for promotion. In these exercises, each of the 371 scientific sectors (SSDs) which pigeon hole all Italian professors was classified as either “bibliometric” or “non bibliometric”,⁸ with different assessment rules for the two types: assessment in the bibliometric scientific sectors had to utilise explicitly quantitative measures of publications in journals indexed in the Web of Knowledge and Scopus databases. As far as we are aware, however, whether a given scientific sector was bibliometric or non bibliometric was decided by the government agency subjectively rather than following some objective rational. It might however be of interest to know whether the ministerial classification reflects differences in the criteria used in the past by promotion and appointment panels in the different scientific sectors, and the index M is ideally suited to determine these criteria. To do so, we use the dataset constructed by Checchi et al (2014), which contains the Web of Knowledge publications by anyone holding a post in an Italian university from 1990 to 2011 (details of the data collection, cleaning and organisation are in Checchi et al 2014). A person’s productivity is constructed from this as a principal component of the number of publications weighted by the number of co-authors and the real h-index (Guns and Rousseau 2009, p 67).

As is explained in some detail in Checchi et al (2014), the promotion process in Italy is highly centralized and firmly separated along scientific sectors: all the members of the appointment panel for a post in a given scientific sector must hold a post in the same sector, and while individuals

⁸Mathematics, Physics, Chemistry, Biology, Medicine, Engineering, Psychology, Agriculture and Veterinary Sciences were classified as “bibliometric”, and Humanities, Law, Sociology, Political Sciences, Architecture, History, Philosophy, Economics and Statistics as “non-bibliometric”.

Table 1: Descriptive Statistics of the M-index.

Period	Non-Bibliometric			Bibliometric			T-test
	mean	st.dev	n	mean	st.dev	n	p-value
1990-1994	0.508	(0.090)	144	0.524	(0.129)	190	0.88
1995-1998	0.546	(0.116)	128	0.657	(0.169)	172	1.00
1999-2002	0.523	(0.083)	159	0.681	(0.111)	192	1.00
2003-2006	0.520	(0.074)	154	0.670	(0.127)	195	1.00
2007-2011	0.543	(0.113)	136	0.651	(0.180)	190	1.00
1990-2011	0.526	(0.094)	877	0.615	(0.159)	1132	1.00

Note: Metric-basedness of appointments and promotions to associate professor of bibliometric and non-bibliometric disciplines in Italian universities. Mean, standard deviation and number of scientific subsectors by time-period. P-values of mean difference tests between non-bibliometric and bibliometric in the last column.

may apply for promotions in areas different from the one in which they are currently allocated, this occurs rarely: less than 1% of promotion are simultaneous with a change of scientific sector. It therefore makes sense to measure the metric-basedness of each scientific sector, and we do so using information about recent productivity to rank the potential appointees, that is the elements of the set \mathcal{N} , and the knowledge of who was promoted, namely the selected set \mathcal{K} in each scientific sector. Schematically, we imagine that each appointment panel is aware of the ranking of the candidates according to the publications and citations measures, and we apply the index of metric-basedness (3) to the aggregate decisions taken by all the appointment panels in each scientific sector.

We have calculated the index M for the promotions to associate professor in each scientific sector, in five separate periods, to reflect the frequency of appointment and promotion rounds, and Table 1 reports the mean by scientific sector, split between bibliometric and non-bibliometric scientific sectors. As Table 1 shows, there are substantial differences in the role of ranking as determined by publication in international journals, which play on average a greater role in scientific sectors classified as bibliometric by the

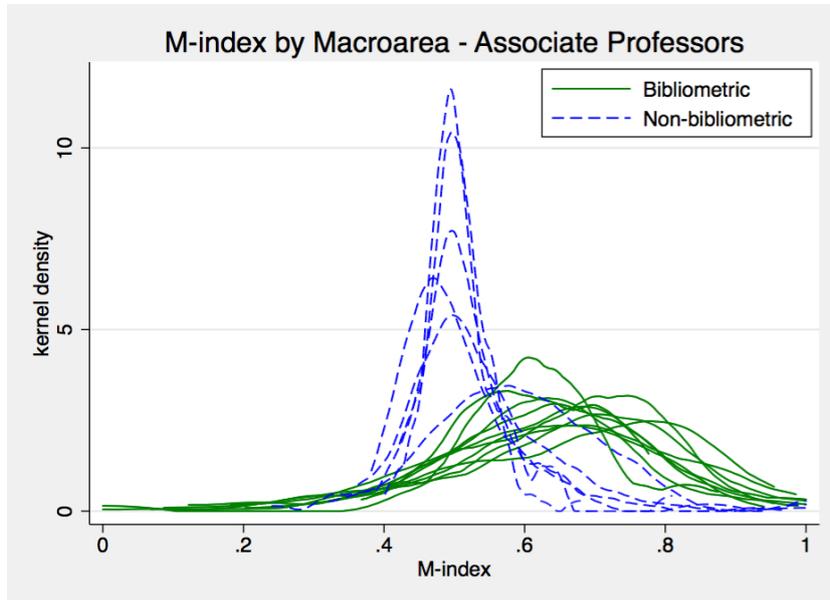


Figure 1: Kernel density estimation of M index by subject.

government. The same message is conveyed by Figure 1, which gives the kernel density of the metric-basedness computed for each scientific sector within each macroarea, and indicates considerable variation within macroareas, suggesting that, although the bibliometric/non-bibliometric divide is valid on average, there is sufficient variation within macroareas to cast doubts on the validity of the classification for a substantial minority of the scientific sectors.

5 Concluding remarks

We have been unable to find in the literature a way to assess, in situations where there is an objective measure of performance, how close a selection is to the ranking established by the measure. Availability of such a measure would assist the evaluation, among others, of medical services (Iverson 1998, Bickman 2012). We propose an index which can be used to this aim. The index is built to satisfy some natural requirements, or axioms; this is a minimal set of axioms, and further extension should attempt to add more natural

ones. The paper ends with an application to the Italian university system, but we are confident that it will prove to have potentially wider applications, including, for example, quantifying up Benabou’s seminal ideas on the role of meritocracy⁹ on growth (Benabou 2000).¹⁰

References

- Barber, Salvador, Walter Bossert, and Prasanta K Pattanaik**, “Ranking Sets of Objects,” in Salvador Barber, Peter J Hammond, and Christian Seidl, eds., *Handbook of Utility Theory*, Springer US, 2004, pp. 893–977.
- Benabou, Roland**, “Meritocracy, Redistribution and the Size of the Pie,” in Kenneth Arrow, Samuel Bowles, and Steve Durlauf, eds., *Meritocracy and Economic Inequality*, Princeton NJ: Princeton University Press, 2000, pp. 317–340.
- Bickman, Leonard**, “Why Can’t Mental Health Services be More Like Modern Baseball?,” *Administration and Policy in Mental Health and Mental Health Services Research*, 2012, *39*, 1–2.
- Cabizza, Mariella and Gianni De Fraja**, “Quality Considerations in Auctions for Television Franchises,” *Information Economics and Policy*, 1998, *10*, 9–22.
- Cecchi, Daniele, Gianni De Fraja and Stefano Verzillo**, “Publish or Perish: An Analysis of the Academic Job Market in Italy, 1991-2011,” mimeo, Milan, Italy, 2014.
- Dalton, Hugh**, “The Measurement of the Inequality of Incomes,” *The Economic Journal*, 1920, *30*, 348–461.

⁹We have deliberately avoided the term “meritocracy”, whose ideological connotations we prefer to eschew: in some environments, such as where corruption or nepotism are rife, being metric-based might well be a natural way to ensure meritocracy, that is that those most deserving are selected, but in other the mechanical application of a metric might prove inefficient and counterproductive.

¹⁰The paper by Kobus and Milos (2008), has an axiomatic approach to inequality in the presence of meritocracy, and takes therefore a different departure from Benabou’s ideas.

Hakes, Jahn K and Raymond D Sauer, “An Economic Evaluation of the *Moneyball* Hypothesis,” *The Journal of Economic Perspectives*, 2006, 20/3, 173–185.

Iverson, Donald J, “Meritocracy in Graduate Medical Education? Some Suggestions for Creating a Report Card,” *Academic Medicine*, 1998, 73, 1223–1225.

Kobus, Martyna and Piotr Milos, “On Meritocratic Inequality Indices,” Munich Personal RePEc Archive 2008.

Lewis, Michael, *Moneyball: The Art of Winning an Unfair Game*, New York: Norton, 2004.

Mazur, David R, *Combinatorics: A Guided Tour*, Mathematical Association of America, 2010.