Comparing and ranking universities when quantity and quality both matter[1]

Nicolas Carayol [‡,◊,2] and Agenor Lahatte[‡]

[◊] GREThA, Université de Bordeaux - CNRS
[‡] Observatoire des Sciences et Techniques, HCERES

February 1, 2017

[1]This article has benefited from many discussions with colleagues. We would like, in particular, to thank P. David, G. Demange, G. Filliatreau, S. Goyal, M. Jackson, M. Lubrano, T. Marchant, P. Moyes, P. Reny, M. Zitt, and the participants in the DIME conference in Paris, the ENID conference in Oslo, the BRICK conference in Turin, the LAGV conference in Marseille, the Bordeaux conference on the economics of science, the EARIE conference in Milan, the Barcelona GSE Summer Forum and seminars at CORE, GREQAM and OST. An earlier version of this paper circulated under the title "Dominance relations and ranking when both quantity and quality matter: Applications to US universities and econ. departments worldwide". The editor and the two anonymous referees should also be strongly acknowledged.

[2]Corresponding author. Nicolas Carayol, Université de Bordeaux, GREThA - CNRS, Avenue Leon Duguit, F-33608 Pessac Cedex. Tel: +33-556844051. Email: nicolas.carayol@u-bordeaux.fr

Abstract

We use and extend the concept of stochastic dominance to compare the scientific outcome of research institutions. Value judgments are explicitly captured by imposing unanimous comparisons for given classes of article value functions. It is shown how this theory can be applied by comparing top U.S. research universities in all fields of science. We further argue that this approach can simultaneously take quality and quantity into consideration as the ranking obtained from the dominance relations focusing on excellence highly correlates with well-know size-dependent ranking counting top cited articles.

Keywords: Stochastic Dominance, Ranking, Tournaments, Citations.

JEL codes: D63, I23.

1 Introduction

During the last decade, we have witnessed the emergence of a growing number of global university and department rankings (ARWU, Leiden, THE, to name but a few). In a context of intensifying globalization of higher education, these rankings are widely used by increasingly mobile students and their families when choosing a university. Because students' country of study will not necessarily be that of their future residence, one of their goals, if not their main one, is to obtain a diploma that will constitute a consistent signal worldwide. They thus want to know, ex ante, which institutions can provide them with such a signal. In turn, university leaders and governments scrutinize these rankings within a context of global competition for good students and funds. However, the existing rankings do not rely on theoretical basis that would ground them on explicit and sound value judgments. Instead, those rankings often mix several indicators, following *ad hoc* recipes.[1] In this article, we introduce a generic theory that provides explicit foundations for comparing universities (or any other research institution), and we document its application to the comparison of top-US universities, by discussing the rankings it leads to, and by confronting them to one well known ranking.

Our basic idea is that to provide such worldwide signals to the public, students, or governments, universities need to combine to some extent a large size and high quality. The universities that succeed doing so are also often called "world class universities" in the popular press. But size and quality are not two dimensions to account for separately. When focusing on the scientific outcome of universities, as most university rankings do (at least in part), size typically refers to the quantity of articles produced, and quality is mainly associated to the scientific impact of the papers. The scientific impact of an article is usually (not exclusively though) approximated by citation counts. In fact, the h-index (Hirsch, 2005) is perhaps the most well known measurement built on such information taking into account both quality and quantity, in a very specific way. The strong specificities of this index may make sense with respect to its given goals, but often render it far less relevant for other purposes, especially at department or university level. Yet, a series of articles, which have provided axiomatics of bibliometric rankings based on the h-index and other well-known indicators, clearly showed some of its undesirable properties (Woeginger 2008; Marchant, 2009; Albaran et al. 2011; Bouyssou and Marchant, 2014).[2] Some articles already considered the quality

[1] The Leiden ranking would be an exception as, in fact, the CWTS releases several rankings, each one being based on a single-indicator. It therefore offers connections with the scolarly literature which to understand the their premisses (see Albarrán et al., 2011 and Bouyssou and Marchant, 2014).

[2] An author's h is the maximum number of articles written by her/him which have received at least h citations. This index has the specific characteristic that it neither takes into account citations received by articles with fewer than h citations, nor considers citations received by papers above the threshold of the first h citations. The citations that are not considered are said to fall outside the h-core. Many contributions

dimension using the traditional stochastic dominance concepts (Lubrano and Protopopescu, 2004, Ravaillon and Wagstaff, 2011, Bazen and Moyes, 2014). To our knowledge, our paper is the first academic article to derive comparisons and ranking from explicit value judgments that take into consideration both quantity and impact.

How do we proceed to take both quantity and quality of scientific articles into account on explicit value judgments? Everything works as if any bilateral comparison should be obtained by consensus among a large number of neutral evaluators. We model the total value of a given university's research as being additively separable in the value of each article. Evaluators unanimity is captured by imposing non-contradictory comparisons of scientific productions value, for given classes of the article value function. Evaluators are assumed to share the generic view that the value of an article is a non-negative function of its scientific impact. This captures the idea that quantity is important. They also agree that publishing a higher impact paper can never decrease the total value of any scientific production, and this translates evaluators' interest in article quality. Assumptions about the second derivative of the value function of each paper are more debatable. However, it appears, implicitly or explicitly, that convexity is a widely accepted assumption in the academic sphere. Typically, convexity implies that the value of two articles of impact x is never larger than the value of one article of impact $2x$. Most university representatives consider it extremely important for their institution to produce highly cited papers. The convexity of the value function, together with the assumption that a zero quality item (or a nearly zero quality item) has value zero, is the way excellence is captured in this model.

Our main theoretical result consists in providing a condition, which can be empirically verified, that stands if and only if the total value of a university publication-impact list is greater than the value of an other one, for any positive and convex individual article value function which is null (as well as its first derivative) when impact is zero. This result (Theorem 3) is new as we had to depart significantly from the standard applications of stochastic dominance. In the theory of choice under uncertainty, quantity does not make sense since, when comparing lotteries, the probabilities of occurrence of each possible state of the world always sum to unity. In the context of income distribution comparisons, judgments could, in principle, take quantity into account (population size, in this case), but are designed not to do so. Moreover, we also depart from these applications in that our value function is assumed to be convex with article quality, while utility function in the theory of choice under uncertainty or in the study of income distributions are always concave.

have criticized this index (Egghe, 2006, Anderson et al., 2008, Woeginger, 2008, Marchant, 2009; Albaran et al. 2011; Bouyssou and Marchant, 2014). In particular, the h-index is shown to violate a number of desirable axioms, such as the Independence axiom, while some of the axioms characterizing it are not appealing (see the discussion in Bouyssou and Marchant, 2014).

When a dominance relation can be established between two institutions, ordering them is obvious. However, a dominance relation does not necessarily exist within each pair. At this point, we are thus left with a partial ordering. There is a lot that can be said with such incomplete ranking. However, we would like to compare our results with the ones obtained by well known university rankings. Therefore, a methodology has been developed to infer a ranking (a complete order) on clear and consistent principles, using the information on bilateral dominance relations (the partial order). Technically speaking, we are here close to the tournament literature and to the literature on centrality in networks.[3] We propose an index (the Importance index) that is consistent with a set of desirable conditions, and on which a ranking can be built. Actually, we show that only two axioms (Symmetry and Indirect Dominance Homogeneity) characterize the Importance index.

This allows us to compare our approach with a well-known ranking of universities which focuses on research excellence at world level: the CWTS Leiden ranking, based on the number of top-10% cited articles, which is a size-dependent and quality-oriented indicator. This ranking is well known by researchers in the field and by many practitioners, and has a solid reputation of being professionally built. Moreover, it shares many commonalities with our approach in terms of data so that it is possible to avoid the interference of any differences in data sources, treatment and normalization by recalculating the index on our data. Further, the underlying assumptions of this indicator can be interpreted within our theoretical framework. On the set of top U.S. research universities,[4] we find that our ranking correlates strongly with the Leiden ranking based the number of articles in the top-10%. Our ranking exhibits a certain capacity to balance the advantages of larger but still excellent universities against the advantages of smaller institutions of outstanding quality.

The rest of the article is organized as follows. Our basic theory of extended dominance relations is developed in the following section. In the third section, we show how this theory can be used to compare the scientific production of research institutions. The fourth section applies it to US research universities. The fourth section presents how unilateral pairwise dominance relations can be turned into rankings, so that the the fifth section can propose new rankings and compare our approach with the Leiden ranking. The last section concludes.

[3]This literature on tournaments was initiated by Wei (1952), Kendall (1955) or Daniels (1969). It has also been applied to majority voting theory (cf. Laslier, 1997). See Langville and Meyer (2012) for an up-to-date and comprehensive review of ranking methods. This literature has been recently revisited using an axiomatic approach (Palacios-Huerta and Volij, 2004; Slutzky and Volij, 2006; Demange, 2013). Concerning the literature on networks centrality, see in particular Katz (1953) and Jackson (2010) for a review.

[4]Only a limited number of results is presented here due to space constraints. More detailed results (in particular at disciplinary level) are available on a companion website: http://ncarayol.u-bordeaux4.fr/ranking.html. More applied papers are also available (Carayol et al., 2012, 2015). The present article is the lead paper with respect to the theoretical foundations.

2 The extended theory of dominance relations

2.1 Notation

Let us define a set I whose elements are to be compared by some evaluators. Elements of the set are called agents, and can denote either individuals or institutions. Let $\Im$ denote the collection of possible sets. Each agent is characterized by a list of items. In the application developed here, the items are articles published in scientific journals, which have been produced by the employees of each considered research institution (the agent). Items are heterogeneous in quality so that each one is characterized by an associated quality $s \in S$, with $S = [0, \bar{s}[\subset \mathbb{R}^+$, the bounded set of all possible values of quality. Even if one would first think of the number of citations received (which is discrete) as characterizing quality, the quality set is assumed to be continuous for two main reasons. First, alternative raw measures of scientific impact, which would be continuous, could be considered in practice. Second, citations (as most other raw measures of scientific impact) can not be directly assumed to measure quality. Some corrections need to be made to more accurately assess quality because there are several external factors that affect impact but which are not related to quality. Such corrections are likely to transform discrete numbers in positive real numbers. Numerous details on these aspects come in the following sections.

An agent $i \in I$ is fully described by the vector $x^i = (x^i_1, x^i_2, ..., x^i_a, ..., x^i_{n(x^i)})$ of size $1 \times n(x^i)$, of which a^{th} entry, x^i_a, denotes the quality of the a^{th} item associated to i. $n(x^i)$ is the number of articles of i. Let us define $q\left(s, x^i\right)$, the function that gives the number of items in a vector x^i, that have a quality level strictly greater than s. Formally, it is given by:

$$q\left(s, x^i\right) = \sum_{a=1,...,n(x^i)} 1_{\{x^i_a > s\}}, \tag{1}$$

with $1_{\{.\}}$ the indicator function, which is equal to one if the condition in brackets is satisfied and zero otherwise. Obviously function $q\left(s, x^i\right)$ is a continuous and increasing function of s, which goes from zero to $n(x^i)$. We now introduce $f\left(s, x^i\right)$ that stands for the (piece-wise) variations of $q\left(s, x^i\right)$: $f\left(s, x^i\right) = \left(q\left(s', x^i\right) - q\left(s'', x^i\right)\right)/2$ where s' and s'' are either entries of x^i, or 0, or $\bar{s}$, such that $s' > s \geq s''$, and so that there is no other entry of x^i that would take a value $\widetilde{s} \neq s$ such that either $s' > \widetilde{s} > s$ or $s > \widetilde{s} \geq s''$. We further assume that vector x^i is large enough and populated with quality values distributed all over S, so that function q is assumed to be differentiable with respect to s, and that thus function f can be seen as the fist derivative of q for all $s \in S$. This assumption allows us to use integrals that will prove to be convenient. In particular, we write:

$$q\left(s, x^i\right) = \int_s^{\bar{s}} f\left(t, x^i\right) dt. \tag{2}$$

Note that the lower bound of the integral is here, by convention, excluded. The function $f(s, x^i)$ can be interpreted as the number of items in a vector x^i that have exactly quality level s.

The function $v(\cdot) : S \to \mathbb{R}$, gives a unique "value" of any unit item as a function of its quality. This function is assumed to be continuous and twice differentiable on S. Assuming that the value of the whole vector of agent i, W_v^i, is the sum of the value of each element, itself obtained by function v, this writes as follows:

$$W_v^i = W_v(x^i) = \int_0^{\bar{s}} v(s) f\left(s, x^i\right) ds. \quad (3)$$

2.2 Dominance relations

The generic dominance relation is noted $\succcurlyeq$. It is a binary relation which compares the elements of any I in class $\Im$. We introduce three specific dominance relations: strong dominance, dominance and weak dominance. Each one requires unanimity within a given particular category of judgment which is defined by a class of admissible $v(\cdot)$ functions. Definitions 1 to 3 require that the total value of an institution's production be superior to that of another institution for any function $v(\cdot)$, within clearly defined classes, for it to be dominant. Theorems 1 to 3 establish the necessary and sufficient conditions for each dominance relation to hold.

Let us first define the notion of strong dominance over the set of agents I. This relation only requires that function $v(\cdot)$ be non-negative, i.e. no item will contribute negatively to the performance of the agent. This very weak condition implies that quality plays almost no role. Therefore, we say that strong dominance relations capture a size or volume dimension.

Definition 1 *Agent i (characterized by vector x^i) strongly dominates agent j (characterized by vector x^j) noted $i \blacktriangleright j$, if, for any non-negative function $v(\cdot)$ over set S: $W_v^i \geq W_v^j$.*

Theorem 1 below simply means that the necessary and sufficient conditions for there to be a strong dominance of one agent over another is that it does not perform less for any possible level of quality. This condition is intuitive, since strong dominance requires unanimity of judgment for any non-negative value function, which may arbitrarily increase the value of any positive level of quality.[5]

Theorem 1 *$i \blacktriangleright j$ if and only if $f(s, x^i) - f(s, x^j) \geq 0, \forall s \in S$.*

[5] All the proofs of the theorems of this section are exposed in Appendix A.

We now introduce the notion of dominance, which requires unanimity among all non-negative and now also non-decreasing functions $v(\cdot)$; that is to say that articles of a higher quality should never have a lower value. Therefore, that dominance relation captures a focus on quality.

Definition 2 *Agent i (characterized by vector x^i) dominates agent j (characterized by vector x^j), and is noted $i \rhd j$, if, for any-non negative and non-decreasing function $v(\cdot)$ over set S: $W_v^i \geq W_v^j$.*

Theorem 2 introduces necessary and sufficient (empirical) conditions for a dominance relation to hold. Interestingly, this condition is equivalent to the monotonicity condition introduced by Bouyssou and Marchant (2014) which, according to them, any bibliometric ranking should satisfy.

Theorem 2 *$i \rhd j$ if and only if $q(s, x^i) - q(s, x^j) \geq 0, \forall s \in S$.*

Additional assumptions can be introduced relative to the second derivative of the value function. Definition 3 introduces the notion of weak dominance which requires that $v(\cdot)$ be convex and that the value of a null quality item should be zero, as well as its first difference at this point. Woeginger (2008) clearly argues that a null value of a zero-citation article should be a "condition sine qua non for measuring the scientific impact: If one's scientific research fails to generate citations (all-zero vector x), it has no impact" (p. 225). We further impose a null first derivative of the value function when quality is zero. This simply captures the idea that no significant change in value is to be expected as quality slightly goes away from zero. Note that if $v(0) = 0$ and if $v(\cdot)$ is positive and convex over S, then it is also increasing. Thus, the notion of weak dominance imposes new conditions that add up to the ones associated to dominance relations. These additional assumptions imply that the value function gives proportional or more than proportional weight to the best quality items. Thus, weak dominance reflects a focus on excellence, that is here an accentuated focus on the highest quality items.

Definition 3 *Agent i (characterized by vector x^i) weakly dominates agent j (characterized by vector x^j) noted $i \trianglerighteq j$, if, for any non-negative and weakly convex function $v(\cdot)$ over set S, such that $v(0) = v'(0) = 0$: $W_v^i \geq W_v^j$.*

We now have the following and main theorem which gives a necessary and sufficient condition for weak dominance to hold.

Theorem 3 *$i \trianglerighteq j$ if and only if $\int_s^{\bar{s}} [q(t, x^i) - q(t, x^j)]\, dt \geq 0, \forall s \in S$.*

Let us consider the following example, which illustrates how weak dominance differentiates from the other forms of dominance.

Example 4 *Let* $x^i = (2,5); x^j = (1,3,3), S = [0,5[$. *Clearly, strong dominance and dominance notions do not allow us to compare the two productions: neither* i *strongly dominates or dominates* j*, nor* j *strongly dominates or dominates* i*. However, if* $v(\cdot)$ *is convex and such that* $v(0) = v'(0) = 0$*, then the following inequalities are obviously true:* $2v(2)+v(5) \geq 3v(3)$*;* $v(2) \geq 2v(1)$ *and* $v(3) \geq \frac{3}{2}v(2)$*. Combining these inequalities yields:*

$$v(2) + v(5) \geq 3v(3) - v(2) = 2v(3) + v(1) + (v(3) - v(1) - v(2)).$$

However, we also have $v(3) \geq \frac{3}{2}v(2) \geq v(2) + v(1)$*, and therefore, we can conclude that:* $v(2) + v(5) \geq 2v(3) + v(1)$*, that is* $W_v(x^i) \geq W_v(x^j)$*, and thus infer that* i *weakly dominates* j*. Table 2 shows that the necessary and sufficient empirical conditions in Theorem 3 are indeed verified in this example.*

Each one of the three notions of dominance requires the unanimity of the judgments associated with any value function belonging to a specific class. The results provided by the three axioms are important because they make it possible to compute the dominance relations, without having to further specify the functional forms of the various dominance relations.

2.3 Some basic properties of dominance relations

We present here some simple properties of the dominance relations that will prove useful in the next sections. Before doing so, we need to define a principle of comparison of dominance relations. A dominance relation $\succcurlyeq$ is said to be "stronger" than any other dominance relation $\succcurlyeq'$, noted $\succcurlyeq' \subseteq \succcurlyeq$, if $\forall I$ in $\Im$, and $\forall i,j \in I$, $i \succcurlyeq j$ implies $i \succcurlyeq' j$. The symbols $\succcurlyeq$ and $\succcurlyeq'$ account for any one of the dominance relations introduced above ($\succcurlyeq, \succcurlyeq' \in \{\blacktriangleright, \rhd, \unrhd\}$). In other words, a dominance relation $\succcurlyeq$ is stronger than $\succcurlyeq'$ if a dominance of the first type of any given agent over any another implies the dominance of the second type.

A first lemma establishes connections between dominance relations. This states that the weaker the dominance relation, the greater the number of dominance relations it is possible to establish between the agents of any given set of agents I in $\Im$. The proofs are derived directly from the definitions of the different forms of dominance, which have conditions of increasing strength.

Lemma 5 $\unrhd \subseteq \rhd \subseteq \blacktriangleright$.

The second lemma below focuses on some properties of the dominance relations. Part i) simply establishes that all the dominance relations introduced are transitive. Part ii) states the basic reflexivity property derived from the definitions of the different forms of dominance in which inequalities are weak. Altogether, i) and ii) imply that the three dominance relations are in fact preorders. The last two statements of the lemma state the necessary and sufficient conditions for having reciprocal dominance relations between two agents. The production performances of these two agents should be identical for all possible quality levels, excluding the zero case for weak dominance only. If one can not find any pair $(i, j) \in I^2$, for any I of $\Im$, such that the conditions in iii) or in iv) are verified, then the associated relations become antisymmetric for all I in $\Im$. The dominance relations are then always partial orders.

Lemma 6 *The following statements hold:*

i) if $i \succcurlyeq j$ *and* $j \succcurlyeq h$, *then* $i \succcurlyeq h, \forall \succcurlyeq \in \{\blacktriangleright, \triangleright, \trianglerighteq\}$; *(transitivity)*

ii) $i \succcurlyeq i$, $\forall i \in I$, $\forall \succcurlyeq \in \{\blacktriangleright, \triangleright, \trianglerighteq\}$; *(reflexivity)*

iii) $\forall \succcurlyeq \in \{\blacktriangleright, \triangleright\}$ *[*$i \succcurlyeq j$ *and* $j \succcurlyeq i$ *if and only if* $f\left(t, x^i\right) = f\left(t, x^j\right), \forall s \in S$*].*

iv) $i \trianglerighteq j$ *and* $j \trianglerighteq i$ *if and only if* $f\left(t, x^i\right) = f\left(t, x^j\right), \forall s \in S \setminus \{0\}$.

2.4 Dominance networks

As it is often useful to provide graphs associated with dominance relations, we introduce the notion of (directed) dominance network that can be simply built from any dominance relation $\succcurlyeq$ as follows: a link from i to j exists if and only if $i \succcurlyeq i$. In the graphical representations, for the sake of clarity, we remove self-dominance and redundant dominance relations, which are uninformative, since reflexivity and transitivity always hold. Such a network is called the adjusted dominance network associated to dominance relation $\succcurlyeq$.

3 Quality, quantity and value

In this article we want to show how our theory goes to real data. For this purpose, we first need, in this section, to introduce some clarifications on the computation of scientific production and of impact, which will proxy quantity and quality in this context. We do not try to extensively justify the choices here as we did in the previous section. We rather present some standard measures used in the field and show how they can be manipulated in the framework of our theory. It should be clear that our theory is not constrained by the specific measurements of quantity or quality proposed here. The appropriate assumptions for the value function are subsequently discussed.

3.1 Fractional counts of scientific production: quantity

We here briefly present the publication counting method known in the scientometric literature as fractional count. Again, in this article we are using that method, but the theory presented above could perfectly apply have we used the full counting method or other methods. Fractional counting is however, in our eyes, more precise and sophisticated. It also conveniently avoids double counting down the road. Fractional counting between institutions basically splits each paper between institutions and fields, given that using bibliometric data implies some constraints. Let index μ specifically denote an article in A, the set of all articles, and let $p_\mu^i \in [0,1]$ account for the fact that, in practice, most articles are attributed to several universities (about 60% in our data) and should, therefore, be divided between them. Either one author is employed by several institutions or several authors are employed by different institutions (or both). In practice, it is often impossible to know the precise affiliations of each author, and therefore one can only count the number of times an institution is referred to in the list of addresses of the authors. An article μ provides institution i with a gross volume of academic production of:

$$p_\mu^i = \frac{\Delta(\mu, i)}{\Delta(\mu)}, \tag{4}$$

where $\Delta(\mu)$ is the number of address lines of the authors of μ and $\Delta(\mu, i)$ the number of address lines of μ which actually correspond to institution i. p_μ^i thus gives the weight of institution i in article μ. For instance, if μ has three authors two of them mention i as their affiliated institution and the third mentions another institution, then the ratio will be equal to 2/3.

Let us also introduce the term $q_\mu^k \in [0,1]$, which is intended to account for the fact that not all papers are associated with field $k \in K$, while those that are, are not necessarily exclusively associated with that field. Typically, in scientometric databases, the information on articles' fields comes from the association of journals to scientific fields. We abstract from this here to simply compute the weight of field k in article μ as follows:

$$d_\mu^k = \frac{1_{\{k \in d(\mu)\}}}{\# d\left(\mu\right)}, \tag{5}$$

where $d(\mu)$ is be the set of fields with which article μ is associated and $\# d\left(\mu\right)$ the cardinal of that set. Thus, d_μ^k serves as a filter for selecting the articles related to field k. It also gives the weight of each field when the journal (and thus the article) is related to several fields.

The share of paper μ that goes to institution i in field k is thus simply given by the product of the two terms: $p_\mu^i \cdot d_\mu^k$. Actually, the previous section considered only unitary papers for simplicity. As we have assumed that total value is additive in the value of each item, considering fractions of articles as we do in this section changes nothing.

3.2 Impact and quality

The most basic way of computing scientific impact simply counts the number of citations received by each article in a given time window after publication. This is computed as follows for any article μ:

$$c_\mu = \#\left\{u : t_u \in w(\mu) \text{ and } \mu \in r\left(u\right)\right\}, \tag{6}$$

with $\#\left\{\cdot\right\}$ the cardinal of the set defined into brackets, t_u the year of publication of article u, $w(\mu)$ the citation window of article μ, and $r\left(u\right)$ the reference list of article u. Therefore $\mu \in r\left(u\right)$ means that μ is cited by u, and $t_u \in w(\mu)$ imposes that the publication year of the citation source falls in the time window. The number of citations is very attractive because it is a direct measure of impact. Less direct measures of quality are often used in the literature such as the impact factor of the journal in which the article was published. Since Pinski and Narin (1976), several authors have also suggested using the whole information contained in the citation matrix between journals to calculate more relevant measures of journal influence (Liebowitz and Palmer, 1984; Laband and Piette, 1994; Palacios-Huerta and Volij, 2004; Bergstrom et al., 2008; Demange, 2013). These journal-based indicators indicators could easily be implemented here. We however prefer using direct measures of impact as there is a skew distribution of impact within-journals (Seglen, 1992). Even at the the article level, we could in principle use the whole citation matrix between articles to better appreciate the scientific impact of each article (to grasp indirect impact). As this would however require huge computational capacities we will here limit ourselves to citation counts defined in Equation (6).

The simplest and naive way of proceeding would be to assume that the chosen impact measure (here citation counts) is per se the appropriate proxy of quality. There are, however, good reasons not to make this assumption. The main one is that impact varies dramatically among fields, simply because citation practices vary across fields. For instance, the average size of reference lists in chemistry is much greater than in mathematics and thus the average impact is higher, so raw impact calculations can not be reliable measures of quality. Other differences between scientific fields concern the timely arrival of citations, the coverage of the databases (use of external sources), the distribution of citations, the collaboration patterns... We therefore propose to normalize impact in each field through its position in the distribution of articles within its corresponding field to compute its quality. The quality of a given article μ in field k is equal to the larger s, such that its raw impact is higher than that of s percent of the articles published in field k. In other words, the quality of article μ in field k is equal to the probability that a randomly drawn article in field k would have a lower impact than μ. The normalized quality of a paper of (non-normalized) impact t in field k is thus given by:

$$s = \Lambda(c,k) = \Pr\left(X^k < c\right), \tag{7}$$

with X^k the raw impact of a randomly drawn paper in field k. In principle, that probability should depend on time as well. From now on, we will assume in this section that this field-normalized (appropriately chosen) impact accounts for article quality. Notice that this definition (with strict inequality) ensures that a zero raw impact item is always of zero quality, and that it is upper bounded by one: $S = [0,1[$. We need now to redefine slightly q to consider this problem of aggregation across fields. Let $q^k\left(s,x^i\right)$ denote the production of institution i in field k, of quality greater than s, computed as follows:

$$q^k\left(s,x^i\right) = \sum_{\mu\in A} 1_{\{s>\Lambda(x(\mu),k)\}} \cdot p^i_\mu \cdot d^k_\mu, \tag{8}$$

where $x(\mu)$ denotes the impact of article μ. Note that we do not need to restrict the sum to articles published by institution i because p^i_μ is null otherwise. The expression can be aggregated across fields for any quality levels s: $q\left(s,x^i\right) = \sum_{k\in K} q^k\left(s,x^i\right)$. Similar notations stand for function f: $f\left(s,x^i\right) = \sum_{k\in K} f^k\left(s,x^i\right)$. This way of proceeding avoids double counting thanks to the fractional counts methods used. It is also very conveniently offering a solution to deal with the multiplicity of impact "norms" across fields. For instance, when an article is associated to two different fields, each half of it is benchmarked according to a different "impact standard", and may then be associated to a different level of quality.

3.3 Value

The function $v(\cdot)$ gives here the "value" of any unit of scientific production in its associated field. Then, the value of the whole production performance of agent i in field k is simply:

$$W^{i,k}_v = \int_S v(s) f^k\left(s,x^i\right) ds. \tag{9}$$

The value of the whole publication performance of institution i can be computed either directly, or by aggregating the values over all fields:

$$W^i_v = \int_S v(s) f\left(s,x^i\right) ds = \sum_k W^{i,k}_v. \tag{10}$$

The establishment of dominance relations between universities is, therefore, a natural extension of the general theory presented in Section 2. If one focuses on comparisons within a given field k, the publication data are the associated $f^k\left(s,x^i\right)$, and the corresponding values are the W^i_v. When one focuses on comparisons across fields, the publication data are the $f\left(s,x^i\right)$, and the corresponding values are the W^i_v.

4 Comparing top US research universities

In this section, the general theory introduced above is applied to the comparison of the scientific production of top US research universities. The data are first presented before the results are exposed.

4.1 Data

A set of the top US universities was selected on the basis of their rank in the Academic Ranking of World Universities (ARWU) produced by Shanghai Jiao Tong University. This ranking is well known to be "research oriented", a specificity which, though based on very different premisses to ours, fits well with them. As our goal was to restrict our analysis to research universities, the best-ranked universities, representing about 30% of all Ph.D. granting universities in the US, was selected, i.e. a total of 112 universities.

Publications by these institutions[6] and the citations they have received have been collected in the Thomson-Reuters-Web of Science (WoS) database.[7] Since the publication data are available only from 2003 onwards and the citations data are only available up to the year 2007, this analysis was carried out using a set of publication data from 2003 to 2005, with a 3-year citation window for each of these publication years. Over the period of observation and for the citation window selected, the scientific production of the 112 universities/institutions considered in this experiment amounted to $329,910$ articles published in the journals referenced in the WoS database, articles which received $2,316,576$ citations. The citation scores achieved by these papers were between 0 and $1,292$.

The assignment of the papers to fields was based on the association of the journals with nine field categories according to the disciplinary nomenclature of the Observatoire des Sciences et Techniques (OST, Paris). The first eight correspond to clear broad disciplinary lines of inquiry, whereas the ninth, labeled Multidisciplinary Sciences, groups together journals that have a truly multidisciplinary focus, as well as some large multidisciplinary journals that publish articles pertaining to several fields. In the disciplinary based comparisons, excluding papers published in such large journals would introduce a significant bias since it would eliminate a significant percentage of the best articles across several fields. Therefore, the articles published in the most influential of these multidisciplinary journals (namely Proceedings of the National Academy of Sciences USA, Science and Nature) were reallocated to

[6] Only research articles are considered here (letters, review... are not considered). The lexical tokens used to collect publications have been kindly provided by Cheng and Zitt (2009).

[7] These data are imported and maintained by the Observatoire des Sciences et Techniques (OST) for national evaluation purposes and research and thus all computations (citations, impact factors, etc.) are performed in-house.

their parent field according to the procedure of Thomson Reuters re-assignation. Therefore, the set K we are working with contains eight well defined broad disciplines or fields (see Table 3). Because the database provides an imperfect coverage of social sciences and humanities, we decided not to consider these disciplines.

As mentioned above, the impact of publications by universities was considered through the direct citations the articles received The scientific production curves of each institution in each field ($f^k(\cdot, x^i)$) were linearized in one hundred points, positioned at equal intervals between zero and the maximum (normalized) impact reached.[8]

4.2 Completeness

The first result proposed concerns the extent to which the various dominance relations allow us to compare universities. For this purpose, we introduce the notion of rate of completeness. It is computed as follows:

$$C(I, \succcurlyeq) = \frac{\#\{\succcurlyeq, I\} - \#I}{\#I(\#I-1)/2}, \tag{11}$$

for dominance relation $\succcurlyeq$ over a given set I, of cardinal $\#I$, where $\#\{\succcurlyeq, I\}$ stands for the cardinal of $\succcurlyeq$ when applied to set I. The numerator is thus equal to the total number of dominance relations that can be inferred on I, having excluded the ones that are always obtained thanks to the reflexivity property. Equation (11) simply gives the percentage of pairs of (distinct) institutions in I for which one dominance relation of type $\succcurlyeq$ can be established. In other words it is the share of bilateral comparisons that are directly inferred by the associated dominance relation. Table 4 presents the rates of completeness for each one of the eight fields, and for all fields together, associated with dominance relations $\blacktriangleright$, $\triangleright$, and $\trianglerighteq$. The results show that completeness varies across domains and depends on the type of dominance and the field. The completeness rates are all strictly below unity and, therefore, we cannot establish a complete ranking from the different types of dominance relations discussed here. Weak dominance relations achieve significantly higher rates than other types of dominance, regardless of the domain, with completeness rates that are generally between eighty and ninety percent. We will therefore focus on weak dominance comparisons.

[8] Different numbers of points were tried. In principle, institutions which produce the highest impact papers in their field should be favored as the number of points increases. However, it also reduces the proportion of bilateral comparisons that can be directly assessed. It turns out that the results change only marginally between twenty and one hundred. All the results presented in this article are obtained with one hundred points.

4.3 Comparisons

Figure 1 presents the adjusted dominance network associated with weak dominance ($\trianglerighteq$) among the top institutions, proxying impact with the number of direct citations. We observe that, just below Harvard, the dominance structure is more sophisticated than expected. In fact, no dominance relation can be found between the University of Michigan at Ann Arbor, the University of Washington (Seattle), the University of California at Los Angeles (UCLA) and Stanford University. In the Appendix, we provide similar weak dominance network for the field of Medicine only (Figure 3).

5 From bilateral dominance relations to rankings

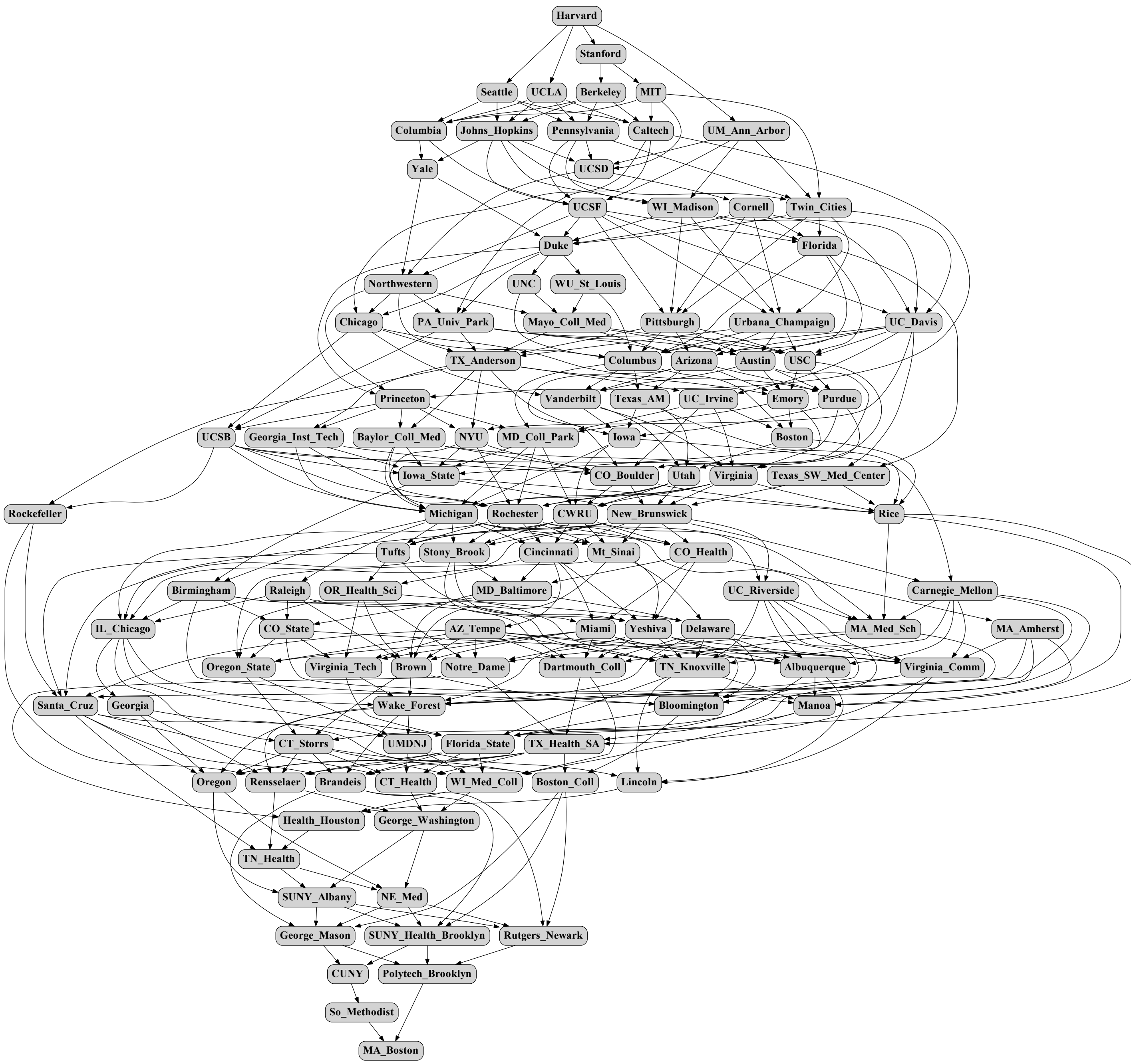
Section 2 showed how to construct a partial ordering in such a way that any bilateral comparison should be based on the unanimity among all judgments that are consistent with some well-defined conditions concerning the shape of the value function. The following section has showed how this theory can be applied for comparing the scientific productions of research institutions. The application to US top research universities was then delivered in the previous section. Now, we would like to compare our approach with existing rankings. One way of proceeding for this purpose, is to build a ranking, that is a complete order, on the basis of the information contained in the bilateral comparisons only.

For instance, in Figure 1, we see that Stanford University and UCLA can not be directly compared according to weak dominance. However, there is some information in the way these two institutions compare to other institutions which could be used to rank them. As they are both dominated by Harvard only, they tie looking at their upward weak dominance relations. But, looking downward, we see that Stanford University only dominates Berkeley and the MIT. That should lead Stanford University to be better ranked than UCLA. How can we precisely use all the information available on dominance relations to compare institutions that are not directly comparable? This is the purpose of his section. To do so, we first clarify our problem, before some indexes are proposed and compared.

5.1 Definitions

Let us first consider any agent set I such that the dominance comparisons between the production curves of its elements i, $f\left(\cdot, x^i\right)$, are necessarily antisymmetric: $\forall i, j \in I, i \neq j$, if $i \succcurlyeq j$ then $j \not\succcurlyeq i$, for any $\succcurlyeq \in \{\blacktriangleright, \triangleright, \trianglerighteq\}$, *i.e.* there is no reciprocal dominance for any pair of distinct agents in I. Parts *iii* and *iv* of Lemma 6 imply that reciprocal dominance of distinct agents would only arise when two agents have exactly the same production, for all quality levels when $\succcurlyeq \in \{\blacktriangleright, \triangleright\}$ and for all non-null quality levels when $\succcurlyeq = \trianglerighteq$. Is this an

Figure 1: The adjusted dominance network (without self-dominance and dominance relations that could be inferred by transitivity) among the top US research universities associated with weak dominance, when impact is measured with citations and for all fields.



acceptable assumption to make? From an empirical point of view, the answer is yes as, in the empirical applications, this assumption is never contradicted if the number of quality points are sufficiently numerous (more than ten points while we use a hundred points in practice). We impose a second restriction as there should be no "separate leagues" in I for any dominance relation. In words, separate leagues would manifest as disconnected components in the dominance network.[9] Formally, $\forall i, j \in I$, there is always a finite sequence $i_0, i_1, ..., i_T$ with $i_0 = i$ and $i_T = j$ such that either $i_{t-1} \succcurlyeq i_t$ or $i_t \succcurlyeq i_{t-1}$, for any $\succcurlyeq \in \{\blacktriangleright, \triangleright, \trianglerighteq\}$. From now on, we assume that $\Im$ is the collection of all sets I whose elements' production functions are such that these two conditions (antisymmetry and no separate leagues) are respected.

A ranking problem is a pair $(\succcurlyeq, I)$ with $\succcurlyeq \in \{\blacktriangleright, \triangleright, \trianglerighteq\}$ and $I \in \Im$. We denote by $\Re$ the set of all such ranking problems. A ranking method (or an index)[10] is defined as a function $\phi : \Re \to \mathbb{R}^N$, with N the set of all finite subsets of $\mathbb{N}$. For any I and dominance relation $\succcurlyeq$, that function returns a vector $1 \times \#I$ of scores . This vector is noted (with a slight abuse of notation) $\phi(\succcurlyeq)$, of which the ith entry $\phi_i(\succcurlyeq)$ is the evaluation of agent i in I with respect to $\succcurlyeq$.

5.2 Indexes

The structure of our problem is reminiscent of the tournament literature initiated by Wei (1952), Kendall (1955), David (1963) and Daniels (1969). We will therefore present the indexes they have introduced. However, the type of tournaments we have here has specificities which will lead us to introduce a new index: the Importance Index. The first characteristic of our tournaments is that there is at most one comparison between two distinct agents. Secondly, since the dominance relations are reflexive, self-dominance always applies. Third, transitivity holds in any dominance matrix and, therefore, the presence of inconsistencies that Wei (1952), Kendall (1955) and Daniels (1969) try to deal with is not the problem here. As antisymmetry is imposed, there is no cycle in the dominance network (when self-dominance is not considered) and, consequently, a complete order can be built avoiding any contradiction with the dominance relations. Instead, incompleteness, when present, is clearly the problem that we are trying to solve so as to obtain a complete order. When two agents can not be compared (there is no dominance relation between them), we are willing

[9] As comparisons between institutions located in separate components would then become adventurous, comparisons and rankings should rather only be made separately for each component. We have however never encountered that case in our data.

[10] We here define a ranking method as an index which can be modified by any linear positive relation. That is to say, any two vectors $\phi(\succcurlyeq)$ and $\phi'(\succcurlyeq)$, such that $\phi'(\succcurlyeq) = b\phi(\succcurlyeq)$ for a given strictly positive b, are considered as equivalent. This is a slight abuse of denomination as the ranking is inferred by the index. See Bouyssou and Marchant (2014) for a precise discussion of the distinction between a ranking and an index.

to use the information on how others relate to both in order to infer some sound form of differentiation. This kind of argument will prompt us to break any form of irrelevance of independent alternatives axiom that Rubinstein (1980) proposed (in the context of complete tournaments) and which led him to the row sum ranking method.

Due to space constraints, only a certain number of the existing methods can be examined. The first one, proposed by David (1962), gives each agent a score equal to the sum of all outgoing paths of length one and two, minus the sum of all incoming paths of the same lengths. This can be written as follows:

$$da_i(\succcurlyeq) = \sum_{j \in I} 1_{\{i \succcurlyeq j\}} \cdot (r_j(\succcurlyeq) - c_j(\succcurlyeq)), \tag{12}$$

where $r_j(\succcurlyeq) = \#\{k : k \in I, j \succcurlyeq k\}$ and $c_j(\succcurlyeq) = \#\{k : k \in I, k \succcurlyeq j\}$. The first ranking method intended to weight each win by the score of the outranked agent was developed by Wei (1952) and Kendall (1955), and is accordingly called the Wei-Kendall method. The most common way of writing this score is as follows:

$$wk_i(\succcurlyeq) = \lambda \sum_{j \in I \setminus i} 1_{\{i \succcurlyeq j\}} \cdot wk_j(\succcurlyeq). \tag{13}$$

This equation states that the score of each agent is proportional to the sum of the scores of all the agents it dominates.[11] Moon and Pullman (1970) and Daniels (1969) have proposed two improvements to this scoring function, which are better known as the Invariant Method and the Fair Bets method. The Invariant Method, noted $im_i(\succcurlyeq)$, is calculated in a very similar way as in Equation (13) but the contribution of each agent j to the score of i is to be divided by the number of distinct agents dominating j ($\#\{k : k \in I \setminus j, k \succcurlyeq j\}$). The Fair Bets method $fb_i(\succcurlyeq)$ is very similar to the Invariant Method, but relies on a different normalization of the contribution of each agent j to the score of the focal agent i: instead of dividing it by the number of agents that dominate j, it considers the number of agents that dominate i ($\#\{k : k \in I \setminus i, k \succcurlyeq i\}$).

We now propose an evaluation function, Influence, which, like the last three ones, has a fixed point inspiration, and which can be traced back to Katz (1953):

$$\alpha_i(\succcurlyeq) = \varepsilon \sum_{j \in I \setminus i} 1_{\{i \succcurlyeq j\}} \cdot \alpha_j(\succcurlyeq) + \delta, \tag{14}$$

with δ and ε two strictly positive parameters. This function increases proportionally with the sum of the score of the agents that the considered agent dominates, plus some exogenous

[11] As long as there is no cycle in the network of dominance relations excluding self relations, existence is guarantied for all positive values of λ.

parameter. In fact, this is very similar to the "real" Page Rank, that is when one also considers the perturbations of the matrices introduced by Brin and Page (1998) who transform them into irreducible ones.[12] Moreover, in the spirit of Ramanujacharyulu (1964), we would like to capture not only the dominance structure "below" agents but also the structure "over" agents. In other words, we also intend to account for the inability to free oneself from the dominance of agents that are themselves dominated. In fact, this sort of "dependence" is exactly what we would obtain with an index $\alpha\left(\succcurlyeq^T\right)$, where $\succcurlyeq^T$ would be defined as the dominance relation obtained from $\succcurlyeq$ by inverting all dominance relations. To take into account both upward and downward dimensions, we can construct a synthetic index called *Importance* that combines the two, by dividing the former by the latter. This is given by:

$$\gamma_i\left(\succcurlyeq\right) = \alpha_i\left(\succcurlyeq\right)/\alpha_i\left(\succcurlyeq^T\right) \tag{15}$$

The division captures the idea that the (downward) influence of each agent is to be benchmarked by its dependence (upward).

Is the Importance index well-defined so that we have one and only one value for each agent in any given ranking problem? In a general problem, ε cannot be arbitrarily large because, at some point, the system may diverge. However, as transitivity always holds (see Lemma 6-i) and as we have assumed antisymmetry (no reciprocal dominance between distinct agents), there is thus no cycle in the directed network of dominance relations, ones self-dominance relations have been dropped as in Equation (14). Therefore $\alpha\left(\succcurlyeq\right)$ and $\alpha\left(\succcurlyeq^T\right)$ can always be calculated, and as $\alpha\left(\succcurlyeq^T\right) > 0$, thus $\gamma_i\left(\succcurlyeq\right)$ exists and is unique for some δ and ε.

In Appendix B, we study whether the indexes introduced above respect a series of desirable axioms. It turns out that the Wei-Kendall, the Invariant Method and the Fair Bets perform very badly because they do not deal with the "dead end" issue: all dominance paths necessarily end at one agent who dominates no other and, therefore, all scores are null.[13] In comparison, most simple indexes as David's are considerably more robust. Further, we show that the Importance index with $\varepsilon = 1$ is the only one among all those introduced above which satisfies a series of desirable axioms. Moreover, it is actually the only measurement which satisfies two of the axioms introduced: Symmetry and Indirect Dominance Homogeneity. This index will thus be used, in the next section, to rank top US universities and to compare our approach with other indexes.

[12]It can easily be shown that, after some normalization, adding perturbations amounts to adding some exogenous value to each agent. On this point, one may refer to Newman (2010).

[13]These indexes are more adapted to ranking scientific journals for instance, for which no such "dead-ends" are found or can fairly be excluded (see Palacios-Huerta and Volij, 2004; Demange, 2013).

6 Ranking US Universities

In this section, we first apply the ranking principles introduces in the previous section to our set of top US research universities. The rankings are based on the importance index, using the different forms of dominance. Next, we compare those rankings with the ranking based on the P(TOP 10%) index, a size-dependent indicator focused on scientific excellence, which is used by the CWTS to produce one of the Leiden rankings.

6.1 The Importance ranking based on the dominance relations

We focus on the rankings based on strong dominance, dominance and weak dominance relations obtained using the Importance scoring function γ, defined in the preceding section. The rankings of the top-40 institutions in the weak dominance ranking are are presented in the first three columns of Table 5. Though the rankings are correlated (see the two first rows and columns of Table 6), some institutions, however, have very different rankings depending on the associated dominance relation. For instance, MIT is twenty-ninth when ranking is based on strong dominance, but seventh in the weak dominance ranking. This result should be interpreted bearing in mind the relative size of this institution. Similar statements can be made for the California Institute of Technology, Princeton University, or the University of California at San-Francisco. The notion of weak dominance provides an opportunity for excellent but smaller institutions to remain at the top of the ranking.

It is interesting not just to confine the investigation to the ranking results, but to simultaneously rely on the associated (adjusted) dominance networks that we have introduced in Section 4 (Figure 1). In fact, no dominance relation can be found between the University of Michigan at Ann Arbor, the University of Washington (Seattle), the University of California at Los Angeles (UCLA) and Stanford University. Stanford is, however, more important and thus better ranked than the other three institutions, because it dominates Berkeley and MIT while the others do not. The University of Washington and UCLA rank better than Berkeley because the latter is dominated by Stanford University, while the other two are not. Although the University of Michigan is not dominated by Stanford University, Berkeley is still better ranked because the University of Michigan does not dominate John Hopkins, Columbia and Pennsylvania Universities, whereas Berkeley does. The Importance index rewards less the capacity of being free from the domination of Stanford than the capacity to dominate John Hopkins, Columbia and Pennsylvania. This is because all these institutions are very high in the dominance network hierarchy.

6.2 Comparisons with Leiden ranking

We now compare our ranking based on weak dominance with the well known Leiden ranking. The CWTS releases each year a series of different rankings based on different indexes. We focus on the ranking built on the P(TOP 10%) indicator, which is the closest to our weak dominance notion. The P(TOP 10%) indicator is equal to the number of articles that are among the top-10% most cited compared with other publications in the same field and in the same year. This indicator is being used as excellence indicator in the literature , most notably for evaluating an research institute (Bornmann 2013). As this indicator, our dominance relations are build on the percentile rank approach for measuring the impact of publications within a reference domain. Actually, the Leiden P(TOP 10%) indicator can be reinterpreted in our framework: it assumes any article in the top-10% has a unitary value, and a null value otherwise. Its implicit value function of papers with respect to s as defined in Equation (10) is thus: $v(s) = 1_{\{s \geq .9\}}$. This function is non-negative and non-decreasing, and is therefore one of the functions considered for the dominance relation (see Definition 2). This function however does not respect the convexity requirement of the weak dominance. As it does not make any difference among the top-cited articles and as it just ignores the 90% less cited items, this indicator thus shares most of the drawbacks of the h-index.

As we do, the Leiden ranking is based on the WoS data, concern all science fields and use the fractional counts and the percentile rank approach for measuring the impact of publications within a reference domain. There are however some differences in the data and their treatment. Just to mention a few, the two approaches use different normalization choices (a classical issue in bibliometrics raised by Zitt et al. 2005a,b among others), use different document types (CWTS considers articles, letters and reviews whereas we only consider articles), and select different lists of publication outlets (the Leiden ranking relies exclusively on publications in the "core journals" of the WoS, while we do not discriminate among publication outlets within the WoS).

Thus, directly comparing our results with the 2012 release of the Leiden ranking, which best matches our 2003-2005 publication period, would raise many issues. We would not be able to identify whether observed dissimilarities between the rankings would be due to the ranking methods or to the data sources and data treatments. Therefore, we recompute the P(TOP 10%) indicator using our dataset and our normalization procedures. This way, we are completely sure the observed differences in ranking are only due to the ranking principles.[14] The last two column of Table 5 present the rankings of our top universities (in the weak dominance ranking based on citations) according to the original size-dependent 2011-12 Leiden Ranking (P(TOP 10%)), and according to our own computation of the same index

[14]We would like to thank an anonymous referee for suggesting us to proceed this way.

on our data (OST #10%). The three first columns correspond to our ranking based on the importance index based on strong dominance, dominance and weak dominance.

Table 6 presents the Spearman correlations between these rankings. It turns out that the ranking based on weak dominance highly correlates with the OST computation of the size-dependent TOP 10% (at 97%). Correlation is slightly less important with the original Leiden Ranking (.96) due to differences in the data and their treatment at CWTS and OST. An examination of the detailed rankings exposed in Table 5 reveals that the ranking based on weak dominance exhibits a surprising capacity to order alternatively, on the one hand institutions that are of more limited size but which have produced outstanding research and, on the other hand, institutions that are both larger and yet excellent. For instance, it gives the California Institute of Technology an intermediate and reasonable seventeenth position, between the twenty-eighth position given by the original P(TOP 10%) Leiden ranking and the thirteenth position according in our own calculation of this indicator.

7 Conclusion

This article introduces a new theory for ranking institutions when both quantity and quality matter: it extends the well-known stochastic dominance theory and proposes a new ranking method based on the unanimous comparisons that is characterized axiomatically. We have applied this theory to compare and rank the scientific production of US research universities. We should emphasize, in conclusion, that this theory provides an original solution for the size problems that most rankings face. Although our tool is not size-independent (simply because it is not a desired implicit assumption), it does, however, give those smaller institutions that perform well in terms of quality the opportunity to compete with larger institutions, in particular when excellence is the focus of the underlying dominance relations.

We also believe that this theory has great application potential because there are numerous contexts in which the quantity and quality of each item both matter. For instance, research departments want to hire big stars and welcome more members. Schools care about the numbers of students they train and their future wages. Social clubs care about both the number of members and their social status. Museums value both the number of artistic items and their importance in the history of arts (assuming that such a quality can be unambiguously assessed).

The final note addresses the issue of social wealth of the method we have introduced. Since our approach helps to better understand and discuss their premisses which, more often than not, are implicit, such comparisons and rankings may become truly useful to the users and to the evaluated institutions.

8 References

Albarrán, P., Ortuño, I., Ruiz-Castillo, J., 2011, The measurement of low- and high-impact in citation distributions: Technical results, Journal of Informetrics 5, 48-63.

Anderson, T.R., Hankin, R.K.S., Killworth, P.D., 2008, Beyond the Durfee square: Enhancing the h-index to score total publication output, Scientometrics 76, 577-588.

Bazen, S., Moyes, P., 2012, Elitism and stochastic dominance, Social Choice and Welfare 39, 207-251.

Bergstrom, C. T., West, J. D., Wiseman, M. A.; 2008, The Eigenfactor™ Metrics, Journal of Neuroscience 28 (45), 11433–11434.

Bouyssou, D., Marchant, T. 2014, An axiomatic approach to bibliometric rankings and indices, Journal of Informetrics 8, 449-477.

Bornmann, L., 2013, How to analyse percentile citation impact data meaningfully in bibliometrics: The statistical analysis of distributions, percentile rank classes and top-cited papers, Journal of the American Society of Information Science and Technology 64 (3), 587-595.

Brin, S., Page, L. 1998, Anatomy of a large-scale hypertextual Web search engine, Computer Networks and ISDN Systems 30, 107-117.

Carayol, N., Filliatreau, G., Lahatte, A., 2015, Impact, dominance et classements des universités, Revue Economique 66, 173-194.

Carayol, N., Filliatreau, G., Lahatte, A., 2012, Reference classes: A tool for benchmarking universities research, Scientometrics 93 (2), 351-371.

Cheng, Y., Zitt M., 2009, Towards a global ranking by size-independent bibliometric measures: An experiment on selected US, French and Chinese Universities, mimeo OST.

Daniels, H.E., 1969, Round-Robin tournament scores, Biometrika 56, 295-299.

David, H.A., 1963, The method of paired comparisons, Griffin, London.

Demange, G., 2013, A ranking method based on handicaps, Theoretical Economics, forthcoming.

Egghe, L., 2006, Theory and practice of the g-index, Scientometrics 69, 131-152.

Hirsch, J.E., 2005, An index to quantify an individual's scientific research output, Proceedings of the National Academy of Sciences of the USA 102, 16569-16572.

Jackson, M.O., 2010, Social and Economic Networks, Princeton University Press.

Katz, L., 1953, A new index derived from sociometric data analysis, Psychometrika 18, 39-43.

Kendall, M.G., 1955, Further contributions to the theory of paired comparisons, Biometrics 11, 43-62.

Laband, D.N., Piette, M.J., 1994, The relative impacts of economics journals: 1970-1990, Journal of Economic Literature 32(2), 640-666.

Langville, A.N., Meyer, C.D., 2012, Who's #1. The Science of Rating and Ranking, Princeton University Press, Princeton and Oxford.

Laslier, J-F., 1997, Tournament Solutions and Majority Voting, Springer-Verlag, Berlin Heidelberg.

Liebowitz, S.J., Palmer, J.P., 1984, Assessing the relative impacts of economics journals, Journal of Economic Literature 22(1), 77-88.

Lindsay, D., 1978, The scientific publication system in social science, San Francisco, CA: Jossey-Bass.

Lubrano, M., Protopopescu, C., 2004, Density inference for ranking European research systems in the field of economics, Journal of Econometrics 123, 345-369.

Marchant, T. 2009, An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors, Scientometrics 80, 325-342.

Pinski, G., Narin, F., 1976, Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics, Information Processing and Management 12(5), 297-312.

Newman, M.E. J., 2010, Networks: An Introduction, Oxford University Press.

Palacios-Huerta, I., Volij, O., 2004, The measurement of intellectual influence, Econometrica, 963-977.

Ramanujacharyulu, C., 1964, Analysis of preferential experiments, Psychometrika 29, 257-261.

Ravaillon, M., Wagstaff, A., 2011, On measuring scholarly influence by citations, Scientometrics 88, 321-337.

Rubinstein, A., 1980, Ranking the participants in a tournament, Journal of the Society of Industrial and Applied Mathematics 38, 108-111.

Seglen, P.0., 1992, The skewness of science, Journal of the American Society for Information Science, 43(9), 628-638.

Slutzki, G., Volij, O., 2006, Scoring of Web Pages and Tournaments - Axiomatizations, Social Choice and Welfare 26(1), 75-92.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J.W., Van Eck, N.J., Van Leeuwen, T.N., Van Raan, A.F.J., Visser, M.S., & Wouters, P., 2012, The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation, Journal of the American Society for Information Science and Technology 63(12), 2419-2432.

Wei, T.H., 1952, The algebraic foundations of ranking theory, Doctoral Thesis, Cambridge University.

Woeginger, G.J., 2008, An axiomatic characterization of the Hirsch-index, Mathematical Social Sciences 56, 224-232.

Zitt, M., Ramanana-Rahary, S., Bassecoulard, E., 2005, Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation, Scientometrics 63, 373-401.

Zitt, M., 2005, Facing Diversity of Science: A Challenge for Bibliometric Indicators, Measurement 3(1), 38-49.

Appendix A. Proofs of the theorems introduced in Section 2

Proof of Theorem 1 - Strong Dominance

We prove $i \blacktriangleright j \Leftrightarrow f(s, x^i) - f(s, x^j) \geq 0, \forall s \in S$. To do so we begin by the right-to-left implication and then prove the left-to-right implication.

i) Let $i \blacktriangleright j$, and assume $f(t, x^i) - f(t, x^j) < 0$ for some given t. Then let us consider the positive function $v_0(s) = 0$ if $s \neq t$ and $v_0(s) = 1$ if $s = t$. Then: $W_{v_0}(x^i) < W_{v_0}(x^j)$ which contradicts $i \blacktriangleright j$ for an appropriately chosen continuous approximation of v_0.

ii) If $f(s, x^i) - f(s, x^j) \geq 0$ then $v(s)n(s, x^i) - v(s)n(s, x^j) \geq 0$, for all s (since $v(s)$ is positive), and if we sum on all s, we have $W_v(x^i) \geq W_v(x^j)$. $\square$

Proofs of Theorem 2 - Dominance

We prove $i \rhd j \Leftrightarrow \int_s^{\bar{s}} [f(t, x^i) - f(t, x^j)]\, dt \geq 0, \forall s \in S$. Again we begin by the left-to-right implication and then prove the right-to-left implication.

i) Assume $i \rhd j$ and suppose $f(s_0, x^i) - f(s_0, x^j) < 0$ for some t and $f(s, x^i) = f(s, x^j)$, $\forall s > s_0$ so that $\int_{s_0}^{\bar{s}} [f(t, x^i) - f(t, x^j)]\, dt < 0$ and $\int_s^{\bar{s}} [f(t, x^i) - f(t, x^j)]\, dt = 0, \forall s > s_0$. Now consider the increasing positive function $v_1(s) = 0$ if $s < s_0$ and $v_1(s) = 1$ if $s \geq s_0$. Then: $W_{v_1}(x^i) < W_{v_1}(x^j)$ which contradicts $i \rhd j$ and this is also true for an appropriately chosen continuous approximation of v_1.

ii) Assume $\int_s^{\bar{s}} [f(t, x^i) - f(t, x^j)]\, dt \geq 0, \forall s \in S$. A special case would be $f(s_0 + \varepsilon, x^i) - f(s_0 + \varepsilon, x^j) = a > 0$, $f(s_0, x^i) - f(s_0, x^j) = -a$ with $\{s_0, s_0 + \varepsilon\} \subset S$, and $\varepsilon > 0$ and $f(s, x^i) = f(s, x^j)$, $\forall s \neq s_0, s_0 + \varepsilon$; so that $\int_{s_0}^{\bar{s}} [f(t, x^i) - f(t, x^j)]\, dt > 0$ and $\int_s^{\bar{s}} [f(t, x^i) - f(t, x^j)]\, dt = 0, \forall s \neq s_0$. Now assume there is a positive and increasing function v such that $\int_0^{\bar{s}} v(s)\left(f(s, y^i) - f(s, y^j)\right) ds < 0$. In the example chosen, the left part of this inequality becomes

$$v(s_0)\left(f\left(s_0, x^i\right) - f\left(s_0, x^j\right)\right) + v(s_0 + \varepsilon)\left(f\left(s_0 + \varepsilon, x^i\right) - f\left(s_0 + \varepsilon, x^j\right)\right)$$

$$= a\left(-v(s_0) + v(s_0 + \varepsilon)\right).$$

Since v is positive and increasing, replacing this expression in the inequality yields $a < 0$, which contradicts the initial statement. This conclusion extends to any appropriately chosen continuous approximation of v. $\square$

Proofs of Theorem 3 - Weak Dominance

Theorem 3 states that $i \trianglerighteq j \Leftrightarrow \int_s^{\bar{s}} [q(t, x^i) - q(t, x^j)]\, dt \geq 0, \forall u \in S$. Again we begin by sufficiency before proving necessity.

i) Assume $i \trianglerighteq j$ and suppose $\exists u \in S$ such that $\int_s^{\bar{s}} [q(t, x^i) - q(t, x^j)]\, dt < 0$. For instance if $f(t, x^i) = f(t, x^j)$ when $t < s$,$f(t, x^i) < f(t, x^j)$ when $t = s$ but $f(t, x^i) = f(t, x^j)$ when $t > s$. Now consider the increasing positive function $v_2(t) = 0$ if $t \leq s - \varepsilon$ and $v_2(t) = t - s$ if $t > s - \varepsilon$. Then:

$$W_{v_2}(x^i) - W_{v_2}(x^j) = \left(f\left(s, x^i\right) - f\left(s, x^j\right)\right) v_2(s),$$

which is of the same sign as $(f(s, x^i) - f(s, x^j)) < 0$, which thus contradicts $i \trianglerighteq j$. This result is preserved for an appropriately chosen continuous approximation of v_1.

ii) To prove necessity, we need first to consider specifically the situation in which the two compared institutions may have produced a different number of items. We show that this situation reduces to a situation in which the two institutions have produced the same number of items, and then prove necessity.

Let first define $m = max_{i \in I} n(x^i)$ and the y^i vectors which are build from the vectors x^i as follows: $y^i := (y_1^i, y_2^i, .., y_m^i) = (0, ..., 0, x_1^i, x_2^i, ..., x_a^i, ..., x_{n(x^i)}^i)$. In other words each entry of the $1 \times m$ vector is defined as follows: $y_t^i = 0$ if $t \leq m - n(x^i)$ and $y_t^i = x_{t-(m-n)}^i$ if $m - n(x^i) < t \leq m$. We should now note that:

$$\int_s^{\bar{s}} \left[q\left(t, x^i\right) - q\left(t, x^j\right)\right] dt = \int_s^{\bar{s}} \left[q\left(t, y^i\right) - q\left(t, y^j\right)\right] dt, \forall s \in S,$$

and thus (empirically) reasoning with the x^i and the y^i are similar here because the zero quality items are never taken into account in the $q(t, y^i)$ (strict inequality in the definition). Moreover, since $v(0) = 0$, we also have $W_v(x^i) = W_v(y^i)$ and then $W_v(x^i) - W_v(x^j) \geq 0$ is equivalent to $W_v(y^i) - W_v(y^j) \geq 0$. Therefore we can rewrite Theorem 3 as follows:

$$i \trianglerighteq j \Leftrightarrow \int_s^{\bar{s}} \left[q\left(t, y^i\right) - q\left(t, y^j\right)\right] dt \geq 0, \forall u \in S. \tag{16}$$

We now prove the necessity implication of this statement. The left part of equivalence statement 16 writes as: $\int_0^{\bar{s}} v(s) \left(f(s, y^i) - f(s, y^j)\right) ds \geq 0$, for all continuous and twice differentiable v functions respecting the conditions exposed in Theorem 3. If we integrate this expression by parts, we obtain:

$$\left[v(s) \int_s \left(f\left(t, y^i\right) - f\left(t, y^j\right)\right) dt\right]_0^{\bar{s}} - \int_0^{\bar{s}} v'(s) \int_0^s \left(f\left(t, y^i\right) - f\left(t, y^j\right)\right) dt ds \geq 0.$$

We now define:

$$p\left(s, y^{i}\right)=\int_{s} f\left(t, y^{i}\right) d t=\int_{0}^{s} f\left(t, y^{i}\right) d t$$

$$=\sum_{a=1, \ldots, n(x)} 1_{\left\{x_{a} \leq s\right\}}=n\left(y^{i}\right)-q\left(s, y^{i}\right)=m-q\left(s, y^{i}\right).$$

The inequality thus becomes:

$$\left[v(s)\left(p\left(s, y^{i}\right)-p\left(s, y^{j}\right)\right)\right]_{0}^{\bar{s}}-\int_{0}^{\bar{s}} v^{\prime}(s)\left(p\left(s, y^{i}\right)-p\left(s, y^{j}\right)\right) d s \geq 0.$$

Notice that $p\left(\bar{s}, y^{i}\right)=m-q\left(\bar{s}, y^{i}\right)=m$. Using this, we obtain after some simplifications the following inequality:

$$-\int_{0}^{\bar{s}} v^{\prime}(s)\left(p(s, y)-p\left(s, y^{j}\right)\right) d s \geq 0$$

Now, we again integrate by parts, and obtain:

$$-v^{\prime}(\bar{s}) \int_{0}^{\bar{s}}\left(p\left(s, y^{i}\right)-p\left(s, y^{j}\right)\right) d s+v^{\prime}(0) \int_{0}^{0}\left(p\left(t, y^{i}\right)-p\left(t, y^{j}\right)\right) d t$$
$$+\int_{0}^{\bar{s}} v^{\prime \prime}(s) \int_{0}^{s}\left(p\left(t, y^{i}\right)-p\left(t, y^{j}\right)\right) d t d s \geq 0$$

Since by assumption $v^{\prime}(0)=0$, this expression reduces to

$$-v^{\prime}(\bar{s}) \int_{0}^{\bar{s}}\left(p\left(s, y^{i}\right)-p\left(s, y^{j}\right)\right) d s+\int_{0}^{\bar{s}} v^{\prime \prime}(s) \int_{0}^{s}\left(p\left(t, y^{i}\right)-p\left(t, y^{j}\right)\right) d t d s \geq 0.$$

Knowing that $p\left(s, y^{i}\right)=m-q\left(s, y^{i}\right)$, it comes:

$$v^{\prime}(\bar{s}) \int_{0}^{\bar{s}}\left(q\left(s, y^{i}\right)-q\left(s, y^{j}\right)\right) d s-\int_{0}^{\bar{s}} v^{\prime \prime}(s) \int_{0}^{s}\left(q\left(t, y^{i}\right)-q\left(t, y^{j}\right)\right) d t d s \geq 0$$

Note that:

$$\int_{0}^{s}\left(q\left(t, y^{i}\right)-q\left(t, y^{j}\right)\right) d t=\int_{0}^{\bar{s}}\left(q\left(t, y^{i}\right)-q\left(t, y^{j}\right)\right) d t-\int_{s}^{\bar{s}}\left(q\left(t, y^{i}\right)-q\left(t, y^{j}\right)\right) d t.$$

Replacing the right part of this equality in the preceding expression, it comes:

$$v^{\prime}(\bar{s}) \int_{0}^{\bar{s}}\left(q\left(s, y^{i}\right)-q\left(s, y^{j}\right)\right) d s$$
$$-\int_{0}^{\bar{s}} v^{\prime \prime}(s)\left(\int_{0}^{\bar{s}}\left(q\left(t, y^{i}\right)-q\left(t, y^{j}\right)\right) d t-\int_{s}^{\bar{s}}\left(q\left(t, y^{i}\right)-q\left(t, y^{j}\right)\right) d t\right) d s \geq 0.$$

After some simplifications and recombinations, we obtain:

$$\int_0^{\bar{s}} \left(q\left(s, y^i\right) - q\left(s, y^j\right)\right) ds \left(v'(\bar{s}) - \int_0^{\bar{s}} v''(s) ds\right)$$
$$+ \int_0^{\bar{s}} v''(s) \int_s^{\bar{s}} \left(q\left(t, y^i\right) - q\left(t, y^j\right)\right) dt ds \geq 0,$$

which ultimately rewrites as:

$$\int_0^{\bar{s}} v''(s) \int_s^{\bar{s}} \left(q\left(t, y^i\right) - q\left(t, y^j\right)\right) dt ds \geq 0.$$

Since v is weakly convex, this last expression is clearly always verified when

$$\int_s^{\bar{s}} \left(q\left(t, y^i\right) - q\left(t, y^j\right)\right) dt \geq 0, \forall s \epsilon S. \square$$

Appendix B. Axioms and indexes characterization

As we want to make use of matrix computations in this appendix, we thus introduce, for any agent set I and dominance relation $\succcurlyeq$, the matrix $G = (g_{ij})_{i,j=1,\ldots,\#I}$ which captures all the relevant information. It is built as follows: $g_{ij} = 1$ if $i \succcurlyeq j$ and zero otherwise. We call such a matrix the dominance matrix of set I associated with dominance relation $\succcurlyeq$.[15] The set of admitted matrices for any $I \in \Im$ is noted Γ. A ranking problem can thus be rewritten as (G, I) and a ranking method as $\phi(G)$.

Axioms

We first introduce a series of axioms that we would like the measurement introduced in Section 4 to respect. The first axiom we want any ranking method to satisfy is the standard Anonymity Axiom.

Axiom 7 ***[Anonymity]*** *Let $\kappa : I \to I$ be a permutation function, and for any $G \in \Gamma$, let $G_\kappa = \left(g_{\kappa(i)\kappa(j)}\right) \in \Gamma$. A ranking method ϕ is anonymous if, for all ranking problems (G, I) in $\Re$, all $i \in I$ and any permutation function κ: $\phi_i(G) = \phi_{\kappa(i)}(G_\kappa)$.*

The following dominance consistency axiom is simple but important, in the sense that it posits that the ranking of any two agents should be strictly consistent with their direct dominance relation (when it exists).

Axiom 8 ***[Dominance Consistency]*** *A ranking method ϕ satisfies Dominance Consistency if, for all ranking problems (G, I) in $\Re$, and $\forall i, j \in I$: if $g_{ij} = 1$ then $\phi_i(G) > \phi_j(G)$.*

Up to this point, the proposed axioms are based on the initial partial order directly obtained with the dominance relation used.

A Symmetry axiom is also in order, since there is no reason why upward dominance relations should be treated non-symmetrically from downward relations. This idea was first expressed by Ramanujacharyulu (1964). According to him, the score an agent has in a tournament should be the inverse of what that agent would have if all wins and loses were reversed, that is here when the transpose of the dominance matrix (noted G^T) is used. The following axiom precisely states this point.

Axiom 9 ***[Symmetry]*** *A ranking method ϕ satisfies Symmetry, if, for all ranking problems (G, I) in $\Re$, and $\forall i \in I : \phi_i(G)\, \phi_i\left(G^T\right) = 1$.*

[15] As noted by a referee, matrix G does not bring any additional information than its associated dominance relation $\succcurlyeq$. We however use this notation here because it conveniently allows us to use matrix notations and calculations.

We could have chosen other involutions to capture the idea of symmetry. For instance the difference could be an avenue, in particular if we had in mind some sort of transfers. The Symmetry Axiom introduced here rather aims to capture the idea that the "power" or the "influence" an agent gains from its domiance relations over the others is to be benchmarked by the dominance the others excert on that agent. Note that this version of the axiom imposes $\phi_i(G)$ and $\phi_i(G^T)$ to be of the same sign.

We now introduce the condition that a ranking method should correspond to the row sum ranking in the subset of complete tournaments in Γ. Let $\Gamma^1 \subset \Gamma$ be the subset of all dominance matrices such that $g_{ij} + g_{ji} = 1$ for distinct i and j, and Γ^2 the complement of Γ^1 in Γ. The set Γ^1 is the subset of all complete tournaments in Γ, and Γ^2 the subset of incomplete tournaments. The row sum is defined as follows: $r_i(G) = \sum_{j \in I} g_{ij}$.

Axiom 10 ***[Row Sum Correspondence]*** *A ranking method ϕ satisfies the condition of Row Sum Correspondence if, $\forall G \in \Gamma^1$ and $\forall i, j \in I$: $\phi_i(G) \geq \phi_j(G)$ if and only if $r_i(G) \geq r_j(G)$*

The Symmetry axiom has led us to consider both upward dominance and downward dominance, and we therefore also introduce the column sum given by $c_i(G) = \sum_{j \in I} g_{ji}$. However it is not necessary to define a column sum correspondence axiom since, given that $c_i(G) = r_i(G^T)$, it is easy to show that Symmetry and Row Sum Correspondence imply column sum correspondence, that is $\forall G \in \Gamma^1$ and $\forall i, j \in I$: $\phi_i(G) \geq \phi_j(G)$ if and only if $c_i(G) \leq c_j(G)$.

We now investigate the idea that indirect dominance should be considered positively. The next axiom builds directly on the ideas developed by David (1963) who argued that the row sum of the defeated agents should account for the value of each win. We propose that one more dominance relation of i or of any of the agents that i beats as compared to some other agent j (who has an otherwise structurally identical position than i), should allow us to rank i strictly better than j. This is stated in the Indirect Dominance Consistency axiom below.

Axiom 11 ***[Indirect Dominance Consistency]*** *A ranking method ϕ respects the condition of Indirect Dominance Consistency if, for all ranking problems (G, I) in $\Re$, for which there exists a permutation function κ, and a pair of agents (t, w) and an agent i: such that $g_{a,b} = g_{\kappa(a)\kappa(b)}$ for all $(a, b) \in I^2 \setminus \{(t, w), \ (\kappa^{-1}(t), \kappa^{-1}(w))\}$, such that $g_{it} = 1 - g_{\kappa(i)\kappa(t)} = 1$ and $g_{tw} = 1 = 1 - g_{\kappa(t)\kappa(w)}$, then $\phi_i(G) > \phi_{\kappa(i)}(G)$.*

Similarly, one more dominance relation of i or of an agent that dominates i, as compared to some j who would otherwise be in a structurally similar position, should allow us to rank j better than i because i is then defeated by a less performing agent. However it is not necessary to state this in a new axiom since the Symmetry axiom will naturally bring in this property, provided Indirect Dominance Consistency is verified. The following very simple examples should clarify most aspects of what this axiom is intended to capture.

Example 12 *Consider a dominance matrix G whose associated adjusted dominance network over is depicted in Figure 2 (dominance relations that can be induced by transitivity have been removed). Dominance Consistency brings in a better ranking of i over u and v, as well as of j over v. But we can not yet order i and j on the one hand, and u and v on the other hand. However, a ranking should reflect a preference for the position of i as compared to the position of j, and of u as compared to v. Indirect Dominance consistency will bring these in. Indeed, there is a permutation function κ_0, defined by: $\kappa_0(i) = j; \kappa_0(j) = i; \kappa_0(u) = u; \kappa_0(v) = v$, that respects $g_{iv} = 1 = g_{\kappa_0(i)\kappa_0(v)}$; $g_{jv} = 1 = g_{\kappa_0(j)\kappa_0(v)}$; $g_{ij} = 0 = g_{\kappa_0(i)\kappa_0(j)}$; $g_{uv} = g_{\kappa_0(u)\kappa_0(v)} = 0$ and there is a pair of agents (i, u), such that $g_{ii} = g_{\kappa_0(i)\kappa_0(i)} = 1$ and $g_{iu} = 1 - g_{\kappa_0(i)\kappa_0(u)} = 1$. Therefore a ranking method ϕ that respects the Indirect Dominance Consistency Axiom will exhibit $\phi_i(G) > \phi_{\kappa_0(i)}(G) = \phi_j(G)$. It is also easy to see that if ϕ also respects Symmetry then $\phi_u(G) > \phi_v(G)$ since the same reasoning as before would apply to v and u if all dominance relations would be reversed.*

Figure 2: The adjusted dominance network (without self-dominance and dominance relations that could be inferred by transitivity) of Example 12 and Example 13.

Example 13 *Let us now consider a dominance matrix (G) whose associated adjusted dominance network is depicted in Figure 2, right graph (again, the dominance relations that can be induced by transitivity are not pictured). We see there that i and j have identical structural positions, but one agent (u), that i dominates, has one more dominance relation than his structural equivalent for j (x). Indirect Dominance Consistency axiom captures the*

idea that agent i's position is preferable to j's position because i dominates an agent (u) that has a greater ability (than x) to dominate other agents. Indeed, there is a permutation function κ_1, defined by: $\kappa_1(i) = j; \kappa_1(j) = i; \kappa_1(u) = x; \kappa_1(v) = y; \kappa_1(a) = a$, that respects the required conditions and a pair of agents (u, v), such that $g_{iu} = g_{\kappa_1(i)\kappa_1(u)} = 1$ and $g_{uv} = 1 - g_{\kappa_1(u)\kappa_1(v)} = 1$. Therefore any ranking method ϕ that respects the Indirect Dominance Consistency Axiom will be such that $\phi_i(G) > \phi_j(G)$.

Indirect Dominance Consistency amounts to considering the direct dominance relations and the indirect dominance relations of length equal to two. However, as suggested by Wei (1952) and Kendall (1955), there is no reason to restrict ourselves to direct dominance and indirect dominance of length two, and we should also consider dominance paths of all lengths. Moreover, in its simplest form, indirect dominance relations may be considered as votes of equal weight. Each dominance path could be considered as a vote and, thus, the scores should be proportional to the number of outgoing dominance paths to others. Before presenting the next axiom which builds directly on this idea, we first introduce some notations. Let the matrix G_i^1 be built from G, replacing by zero all entries that concern, in column, agents that i does not dominate, and self-dominance relations: $G_i^1 = \left(g_{ju} \cdot g_{ij} \cdot 1_{\{i \neq j\}}\right)_{j,u=1,\ldots,\#I}$. In matrix form, we have $G_i^1 = (G - II) \times II_{G,i}$, where $II_{G,i}$ is the diagonal matrix which diagonal is formed with the i^{th} line of matrix G. The matrix $G_i^2 = G - II - G_i^1$, captures the remaining information. Moreover, let's define a function that precisely counts the number of outgoing paths from agent i in G:

$$\sigma_i(G) = \sigma_i\left(G_i^1\right) = \sum_{j \in I} \sum_{k=0,\ldots\infty} h_{ij}^{(k)}. \tag{17}$$

The term $h_{ij}^{(k)}$ here denotes the i^{th} line and j^{th} column of matrix $(G - II)^{(k)}$ (with II the identity matrix) which is the k^{th} power of matrix $(G - II)$.[16] It is clear that the computation of all the $h_{ij}^{(k)}$ only uses the non null information contained in G_i^1: it is also the i^{th} line and j^{th} column of matrix $(G_i^1 - II)^{(k)}$.

Axiom 14 ***[Indirect Dominance Homogeneity]*** *A ranking method ϕ respects the condition of indirect dominance consistency, if, for all ranking problems (G, I) in $\Re$, and $\forall i \in I$: it is of the form: $\phi_i(G) = \psi_i(G_i^2) \cdot \sigma_i(G_i^1)$ with $\sigma_i(\cdot)$ defined in Equation 17 and with $\psi_i(\cdot)$, any given strictly positive function.*

[16]Matrix $(G - II)$ is considered here instead of matrix G, because it makes no sense to include the others' self-dominance as votes for the considered agent. One's self-dominance is however included since, by convention, $H^{(0)} = II$. It is well known from graph theory that the i^{th} line and j^{th} column entry of the k^{th} power of the adjacency matrix of some digraph gives the number of paths of length k from i to j. Therefore, $\sum_{k=0,\ldots\infty} h_{ij}^{(k)}$ gives the number of paths (irrespective of their length) from i to j on $(G - II)$.

The axiom of Indirect Dominance Homogeneity simply requires the score of agent i to be proportional to the number of paths originating from i, not excluding that it could also be affected by some other positive function of the dominance relations, but which would not use the information contained in G_i^1.

Measurements' characterization

We now investigate how the proposed indexes relate to the different axioms introduced previously. Table 1 synthesizes all the information. What may surprise the reader is the very weak performance of the Wei-Kendall, the Invariant Method and the Fair Bets, that only respect the Anonymity Axiom. However, a little reflection leads to the conclusion that these indexes are not well designed to the class of matrices considered here. In particular, they do not even satisfy the Dominance Consistency Axiom. That is because, by transitivity, all dominance paths necessarily end at one agent who dominates no other and, therefore, all scores are null. Consequently these indexes always yield a vector of zeros and are not even able to differentiate two agents when one dominates the other.[17] In comparison, David's index is considerably more robust since it satisfies four of the six axioms.

Table 1: Ranking methods and axioms.

	A	DC	S	RSC	IDC	IDH
r, $1/c$	√	√	-	√	-	-
da	√	√	-	√	√	-
wk, im, fb	√	-	-	-	-	-
α	√	√	-	√	-	-
γ	√	√	√	√	√	√

In fact, it turns out that the two axioms of Symmetry and Indirect Dominance Homogeneity are also sufficient to characterize Importance as stated in the following theorem.

Theorem 15 *A ranking method ϕ satisfies the axioms of Symmetry and Indirect Dominance Homogeneity if and only if it is the Importance evaluation function γ with $\varepsilon = 1$.*

Proof.

[17]These indexes are more adapted to ranking scientific journals for instance, for which no such "dead-ends" are found (see Palacios-Huerta and Volij, 2004; Demange, 2013).

We start with the "if" part of the proof. Applying the recursivity of the right-hand side of Equation 14, one obtains:

$$\alpha_i\left(G\right)=\varepsilon\sum_{j\in I\backslash i}g_{ij}\times\left(\varepsilon\sum_{k\in I\backslash j}g_{jk}\left(\varepsilon\sum_{y\in I\backslash k}g_{ky}\left(\ \ldots\right)+\delta\right)+\delta\right)+\delta.$$

This expression reduces to:

$$\alpha_i\left(G\right)=\delta\sum_{k=1\ldots,n-1}\varepsilon^k\sum_{j\in I}h_{ij}^{(k)}+\delta,$$

where $h_{ij}^{(k)}$ is the i^{th} line and j^{th} column entry of the k^{th} power of matrix $G-II$. It can be shown easily that $h_{ij}^{(k)}$ is equal to the number of paths of length k that start from i and end at j. We can stop the sum at $k=n-1$ (recalling that $n=\#I$) because there is no cycle in the network associated with $G-II$ and, thus, the longest possible path is composed of $n-1$ intermediary links. Letting $\varepsilon=1$, one obtains:

$$\alpha\left(G\right)=\delta\left(\sum_{k=0,\ldots,n-1}\sum_{j\in I}h_{ij}^{(k)}\right),$$

because by convention, for any matrix A, $A^{(0)}=II$ and thus we have $h_{ii}^{(0)}=1$.

Let us now write $h_{ij}^{T(k)}$ the i^{th} line and j^{th} column entry of matrix $\left(G^T-II\right)^{(k)}$. Similar computations can be introduced for $\alpha_i\left(G^T\right)$ which leads to:

$$\alpha_i\left(G^T\right)=\delta\left(\sum_{k=0,\ldots,\#I-1}\sum_{j\in I}\left(h_{ij}^T\right)^{(k)}\right),$$

when $\varepsilon=1$. We can thus obtain $\gamma_i\left(G\right)$ using previous calculations:

$$\gamma_i\left(G\right)=\alpha_i\left(G\right)/\alpha_i\left(G^T\right)=\frac{\sum_{k=0,\ldots,\#I-1}\sum_{j\in I}h_{ij}^{(k)}}{\sum_{k=0,\ldots,\#I-1}\sum_{j\in I}\left(h_{ij}^T\right)^{(k)}}.$$

Since $h_{ij}^{T(k)}$ is also equal to the i^{th} line and j^{th} column entry of matrix $\left(\left(G_i^2\right)^T-II\right)^{(k)}$, $\gamma_i\left(G\right)$ is of the form:

$$\gamma_i\left(G\right)=\psi_i\left(G_i^2\right)\cdot\sigma_i\left(G_i^1\right),$$

with $\sigma_i\left(G_i^1\right)=\sum_{j\in I}\sum_{k=0,\ldots\infty}h_{ij}^{(k)}$ (which is also equal to $\sum_{j\in I}\sum_{k=0,\ldots\#I-1}h_{ij}^{(k)}$ since there is no cycle in the network) and $\psi_i\left(G_i^2\right)=1/\sum_{k=0,\ldots,\#I-1}\sum_{j\in I}\left(h_{ij}^T\right)^{(k)}$. Therefore, the Importance ranking method γ with $\varepsilon=1$ respects the Indirect Dominance Homogeneity axiom.

Since $\gamma_i\left(G^T\right)=\alpha_i\left(G^T\right)/\alpha_i\left(G\right)$, then obviously $\gamma_i\left(G\right)\gamma_i\left(G^T\right)=1$ and γ thus respects the Symmetry axiom.

We have proved that the Importance ranking method respects the two axioms of Indirect Dominance Consistency and Symmetry. To show the converse, let us assume that a given ranking method ϕ respects the conditions of Indirect Dominance Consistency and of Symmetry. Thus, for all ranking problems (G, I) in $\Re$ and $\forall i \in I$, we have:

$$\phi_i(G) = \psi_i\left(G_i^2\right) \cdot \sigma_i\left(G_i^1\right), \tag{18}$$

and

$$\phi_i(G)\,\phi_i\left(G^T\right) = 1. \tag{19}$$

If we introduce the former equation into the second, we obtain:

$$\psi_i\left(G_i^2\right) \cdot \sigma_i\left(G_i^1\right) \cdot \psi_i\left(\left(G^T\right)_i^2\right) \cdot \sigma_i\left(\left(G^T\right)_i^1\right) = 1. \tag{20}$$

As we have shown in the first part of this proof, we know that $\sigma_i\left(G_i^1\right) = \sum_{j\in I}\sum_{k=0,\ldots\infty} h_{ij}^{(k)} = \sum_{j\in I}\sum_{k=0,\ldots\#I-1} h_{ij}^{(k)} = \alpha_i(G)$ since there is no cycle in the network. Similarly, we have $\sigma_i\left(\left(G^T\right)_i^1\right) = \alpha_i\left(G^T\right)$. Therefore, Equation 20 thus becomes:

$$\psi_i\left(G_i^2\right) \cdot \alpha_i(G) \cdot \psi_i\left(\left(G^T\right)_i^2\right) \cdot \alpha_i\left(G^T\right) = 1. \tag{21}$$

We thus have three possible definitions for ψ_i so that this equation is verified for all G and all i:

1. $\psi_i\left(G_i^2\right) = 1/\alpha_i(G)$ and $\psi_i\left(\left(G^T\right)_i^2\right) = 1/\alpha_i\left(G^T\right)$,
2. $\psi_i\left(G_i^2\right) = 1/\psi_i\left(\left(G^T\right)_i^2\right)$, and then $\alpha_i(G) = 1/\alpha_i\left(G^T\right)$,
3. $\psi_i\left(G_i^2\right) = 1/\alpha_i\left(G^T\right)$ and $\psi_i\left(\left(G^T\right)_i^2\right) = 1/\alpha_i(G)$.

The first case implies that $\phi_i(G) = \phi_i\left(G^T\right) = 1$ for all G and i, that is the ranking would be always flat. The second possibility can not be verified since the equality $\alpha_i(G) = 1/\alpha_i\left(G^T\right)$ is clearly not verified in general. Therefore only the last possibility applies. Now, replacing $\psi_i\left(G_i^2\right)$ by its expression in the indirect dominance homogeneity condition, we obtain $\phi_i(G) = \psi_i\left(G_i^2\right) \cdot \sigma_i\left(G_i^1\right) = \alpha_i(G)/\alpha_i\left(G^T\right)$ which is precisely equal to $\gamma_i(G)$. This completes the proof. $\square$

Appendix C. Tables and Figures

Table 2: Empirical condition of Theorem 3 in Example 4.

s	$q(s, x^i)$	$q(s, x^j)$	$\int_s^5 q(t, x^i)\, dt$	$\int_s^5 q(t, x^j)\, dt$
$1-\varepsilon$	2	3	7	7
$2-\varepsilon$	2	2	5	4
$3-\varepsilon$	1	2	3	2
$4-\varepsilon$	1	0	2	0
$5-\varepsilon$	1	0	1	0

Table 3: The scientific fields (disciplines).

k	Fields
1	Fundamental Biology
2	Medicine
3	Applied Biology/Ecology
4	Chemistry
5	Physics
6	Science of the Universe
7	Engineering Sciences
8	Mathematics

Table 4: The rate of completeness of a series of dominance relations in the set of 112 US higher education and research institutions.

	Citations		
Dominance relation	$\blacktriangleright$	$\triangleright$	$\trianglerighteq$
1 Fundamental Biology	.540	.786	.829
2 Medicine	.705	.853	.880
3 Applied Biology/Ecology	.523	.764	.802
4 Chemistry	.518	.786	.828
5 Physics	.605	.830	.857
6 Science of the Universe	.599	.788	.821
7 Engineering Sciences	.629	.807	.843
8 Mathematics	.432	.667	.716
All fields	.162	.826	.863

Figure 3: The adjusted dominance network (without self-dominance and dominance relations that could be inferred by transitivity) among the top US research universities associated with weak dominance, when impact is measured with citations in the field of medicine.

Table 5: Top 40 rankings of 112 US higher education and research institutions in all fields according to the Importance index, built upon the three dominance relations, the top-40 according to the number of articles in the top-10% most cited articles of its discipline, the original size-dependent 2011-12 Leiden Ranking P(TOP #10%).

	▶	▷	⊵	OST #10%	Leiden P(top10%)
Harvard University	1	1	1	1	1
Stanford University	3	2	2	2	2
University of California, Los Angeles	6	3	3	5	4
University of Washington, Seattle	10	3	3	6	5
University of California, Berkeley	9	5	5	4	9
University of Michigan - Ann Arbor	4	6	6	7	3
Massachusetts Institute of Technology	29	9	7	3	8
The Johns Hopkins University	15	7	8	9	6
University of Pennsylvania	18	8	9	8	7
Columbia University	14	10	10	10	10
University of California San Diego	11	12	11	14	11
University of Wisconsin Madison	2	11	12	15	12
University of Minnesota, Twin Cities	5	13	13	17	16
Yale University	16	16	14	12	15
University of California, San Francisco	38	15	15	13	13
Cornell University	8	13	16	16	14
California Institute of Technology	41	17	17	11	28
Duke University	18	18	18	18	17
Northwestern University	22	20	19	19	19
University of Illinois at Urbana-Champaign	12	22	20	20	24
University of Pittsburgh	21	23	20	21	18
University of North Carolina at Chapel Hill	23	24	22	22	23
Washington University in St. Louis	23	26	22	23	22
University of California, Davis	7	21	24	28	21
University of Florida	17	19	25	32	26
Pennsylvania State University - Univ Park	23	25	26	26	20
University of Chicago	28	28	27	24	29
Mayo Medical School	65	27	28	25	-
Princeton University	46	34	29	27	31
University of Texas Anderson Cancer Ctr	83	39	30	30	-
Ohio State University - Columbus	20	29	31	33	25
University of Arizona	26	32	32	35	44
University of Texas at Austin	31	30	33	31	27
University of Southern California	37	31	34	37	32
University of California, Santa Barbara	39	36	35	29	39
University of California, Irvine	26	35	36	36	34
Emory University	65	43	37	41	36
Baylor College of Medicine	75	40	38	34	49
Texas A&M Univ - College Stn	13	33	39	44	40
Vanderbilt University	43	38	40	40	35

Table 6: Spearman rank correlation between the top-40 universities according to the Importance index based on strong dominance, dominance, weak dominance ranking, the number of top-10% articles and the original size-dependent P(TOP-10%) ranking in the 2011-12 Leiden Ranking.

	$\blacktriangleright$	$\rhd$	$\unrhd$	OST #10%
$\rhd$	0.77			
$\unrhd$	0.70	0.98		
OST #10%	0.58	0.93	0.97	
Leiden P(TOP 10%)	0.72	0.96	0.96	0.92