Another characterization of the Hirsch index

Antonio Quesada[†]

Departament d'Economia, Universitat Rovira i Virgili, Avinguda de la Universitat 1, 43204 Reus, Spain

9th November 2008

149.1

Abstract
The Hirsch index is a number that synthesizes a researcher's output. It is defined as the maximum number *h* such that the researcher has *h* papers with at least *h* citations each. This index is axiomatically characterized using three axioms. One of them sets upper and lower bounds to the index. The other two require the index to satisfy conditions of convexity and strict subadditivity.

Keywords: Hirsch index, publications, citations, research quality.

JEL Classification: C43, A11, D80, D70

[†] E-mail address: aqa@urv.cat. Financial support from the Spanish *Ministerio de Educación y Ciencia* under research project SEJ2007-67580-C02-01 and from the *Departament d'Universitats, Recerca i Societat de la Informació* (*Generalitat de Catalunya*) under research project 2005SGR-00949 is gratefully acknowledged.

1. Introduction

Research has often to do with the aggregation of quantitative magnitudes. An empirical phenomenon in general involves many variables and the theoretical analysis of that phenomenon typically requires the aggregation of many, if not all, of those variables. For instance, several heterogeneous commodities are produced in an economy but macroeconomic models presume that some function can be used to map any n-dimensional vector, whose components indicate the production of each commodity, into a number representing "total output".

The problem of aggregating heterogeneous magnitudes has a bearing on researchers themselves. After all, a researcher produces an output that can be measured in terms of papers and citations of each paper. And, conceptually, that output is also a heterogeneous collection of magnitudes lacking an indisputable standard of measurement. For an economy, it is not obvious whether there is more output in the production vector (1, 2) consisting of 1 unit of commodity 1 and 2 units of commodity 2 or in the vector (2, 1) consisting of 2 units of commodity 1 and 1 unit of commodity 2. Similarly, it is not obvious whether it is more valuable to have one paper cited ten times or ten papers cited once.

The Hirsch (2005) index is a measure that summarizes a researcher's output by reducing this output to a number. Assuming that a researcher i is characterized by a set of papers and, for each paper, the number of citations of that paper, the Hirsch index of researcher i is the maximum number h of papers by i having at least h citations each; see Wikipedia (2008) for a discussion of advantages and criticisms. For instance, if x is the one paper output in which the paper is cited ten times and y is the output consisting of ten papers each of which is cited once then x and y have the same Hirsch index: 1.

Woeginger (2008) has axiomatically characterized the Hirsch index using five axioms. The first axiom attributes index zero to having no paper or no citation. The second axiom is monotonicity: more citations cannot lower the index (see MON in Section 2). The other three axioms appear to be too circumscribed to the particular interpretation of the magnitudes involved (papers and citations) and, as Woeginger (2008, p. 227) himself stresses, should be interpreted within the context of monotonicity. These features of Woeginger's axiomatization leave a space for a characterization requiring fewer axioms, not hinging on monotonicity and based on conditions susceptible to be applied to a general aggregation problem.

This note suggests one such characterization. It relies on axioms A1, A2 and A3 in Section 2. A1 sets the upper and the lower bounds of the index: it is never higher than the highest number of citations and never smaller than the smallest number between the number of cited papers and the smallest number of citations. A2 is a property of convexity: loosely speaking, if two outputs have the same index then any combination of them has also the same index. Lastly, A3 is a property of strict subadditivity: it identifies a situation in which the sum of two outputs must be higher than the index of the output obtained by merging the two outputs.

2. Definitions and axioms

Let $\mathbb{N}$ be the set of non-negative integers. Members of $\mathbb{N}$ represent both the number of papers of a given researcher and the number of citations that a paper can receive. Define X to be the set of all vectors $x = (x_1, x_2, \ldots, x_n)$ such that $n \in \mathbb{N}\backslash\{0\}$ and $x_1 \geq x_2 \geq \ldots \geq x_n$. For $x \in X$: (i) d_x is the number of components (or dimension) of vector x; (ii) c_x is the number of components of vector x different from 0; (iii) for $i \in \{1, \ldots, d_x\}$, x_i is the ith component of vector x and stands for the total number of citations of paper i; and (iv) $x^- = min\{x_1, \ldots, x_{d_x}\}$ and $x^+ = max\{x_1, \ldots, x_{d_x}\}$. With $\varnothing$ designating the empty vector (the no paper case), a researcher's output will be represented by a member of $D = X \cup \{\varnothing\}$. When $x = \varnothing$, the convention is that $c_x = d_x = x^- = x^+ = 0$.

Definition 2.1. A research output index (or index, for short) is a mapping $f : D \to \mathbb{N}$.

This definition of index, which is Woeginger's (2008, p. 225), is restrictive because it excludes reasonable indices like the average citation index. Since papers and citations are measured by members of $\mathbb{N}$, the presumption that the units of measure of the index are papers or citations may justify forcing the index to yield values in $\mathbb{N}$. For instance, for output x, $f(x)$ can be interpreted as the number of "valuable" papers in x or as the number of representative citations in x, with valuable papers or representative citations being a standard of measurement implicitly defined. It is worth noticing that the suggested characterization (Proposition 3.3) does not hold for indices taking values in the set of rational numbers; see Example 3.9. This probably testifies to the weakness of the axioms. By contrast, Woeginger's (2008, pp. 229-230) proof of his characterization remains valid for real-valued indices.

Definition 2.2. The Hirsch index is the research output index h such that $f(\varnothing) = 0$ and, for all $x \in X$, $h(x) = max\{n \in \{0, 1, \ldots, d_x\} : x_n \geq n\}$.

A1. For all $x \in D$, $min\{x^-, c_x\} \leq f(x) \leq x^+$.

A1 is just a condition setting upper and lower bounds to the index. On the one hand, A1 states that the index of output x cannot be greater than the maximum number of citations that a paper receives in x. On the other, A1 also requires the index to be at least equal to the smallest of two values: the number of cited papers and the minimum number of citations that a paper receives in x. Define the binary relation $\geq$ on D as follows: for all $x \in D$ and $y \in D$, $x \geq y$ if, and only if: (i) $d_x \geq d_y$; and (ii) for all $i \in \{1, \ldots, d_y\}$, $x_i \geq y_i$.

A2. For all $x \in X$, $y \in X$ and $z \in X$, if $x \geq z \geq y$ and $f(x) = f(y)$ then $f(z) = f(y)$.

A2 is a convexity property: if output z lies "between" two outputs x and y having the same index then z has the same index as x and y. A2 expresses as well some sort of monotonicity condition: if the changes from y to $x \geq y$ do not alter the index then smaller changes are also incapable of modifying the index. A2 is weaker than Woeginger's (2008, p. 225) monotonicity axiom MON.

MON. For all $x \in D$ and $y \in D$, $x \geq y$ implies $f(x) \geq f(y)$.

If f satisfies MON then $x \geq z \geq y$ implies $f(x) \geq f(z) \geq f(y)$. In view of this, $f(x) = f(y)$ implies $f(z) = f(y)$. This proves that MON implies A2. The index f such that, for all $x \in D$, $f(x) = h(x)$ if $h \geq 2$, $f(x) = 1$ if $h(x) = 0$ and $f(x) = 0$ if $h(x) = 1$ satisfies A2 but not MON. As a second example, $f(x) = 1 / (1 + h(x))$ satisfies A2 but not MON.

For $x \in X$ and $y \in X$ with $d_x \geq d_y$, let $x + y$ and $y + x$ both designate the member z of X such that: (i) for all $i \in \{1, \ldots, d_y\}$, $z_i = x_i + y_i$; and (ii) for all $i \in \{d_y + 1, \ldots, d_x\}$, $z_i = x_i$. For $x \in X$, define x^{-1} to be the member of D obtained from x by deleting the first component x_1 of x (the most cited paper in x). Formally: for $x \in D_1$, $x^{-1} = \varnothing$; and, for $n \geq 2$ and $x \in D_n$, x^{-1} is the member y of D_{n-1} such that, for all $i \in \{1, \ldots, n-1\}$, $y_i = x_{i+1}$.

A3. For all $x \in X$, if $f(x^{-1}) < f(x)$ then, for all $y \in X$ such that $y^+ \neq 0$, $f(x + y) < f(x) + f(y)$.

An index f is subadditive when, for all $x \in X$ and $y \in X$, $f(x + y) \leq f(x) + f(y)$. Subadditivity means that output evaluation is subject to diseconomies of scale: the index associated with the addition of two outputs is never higher than the addition of the indices of the two outputs evaluated independently. Subadditivity suggests that the effort (measured in citations) necessary to increase the index increases itself with the index. The Hirsch index, for instance, satisfies subadditivity. A3 identifies a condition under which the index is strictly subadditive: if the most cited paper in x is decisive (losing that paper reduces the index) then the index of any expansion $z = x + y$ of x is smaller than the sum of the index of x and the index of the additional output y transforming x into z, provided y does not consist of papers without citations.

3. Result

Remark 3.1. The Hirsch index satisfies A1, A2 and A3.

With respect to A1, $h(x)$ means that some paper receives $h(x)$ citations and, in view of this, it cannot be that $h(x) > x^+$. Hence, $h(x) \leq x^+$. If $x^- \geq c_x$ then each of the c_x papers having some citation receives at least c_x citations, so $h(x) = c_x = min\{x^-, c_x\}$. Finally, if $x^- < c_x$ then each of the c_x papers having some citation receives at least x^- citations, for which reason $h(x) \geq x^- = min\{x^-, c_x\}$.

As for A2, let $x \geq z \geq y$ and $h(x) = h(y) = h$. Given that y has h papers with at least h citations each, $z \geq y$ implies $h(z) \geq h$. On the other hand, x has h papers with at least h citations each and does not have h + 1 papers with at least h +1 citations each. Consequently, $x \geq z$ implies $h(z) \leq h$. As a result, $h(z) = h$.

Concerning A3, suppose $h(x^{-1}) < h(x)$ and $y^+ \neq 0$. It follows from $h(x^{-1}) < h(x)$ that $x_{h+1} < h(x)$, so $x_{h+1} \leq h(x) - 1$. Case 1: $h(y) \leq h(x)$. In this case, paper $h + 1$ in $x + y$ receives at most $h(y) + h(x) - 1$ citations. Given this, $h(x + y) < h(x) + h(y)$. Case 2: $h(y) > h(x)$. Then $x + y$ is such that: (i) the first $h(x)$ papers have at least $h(x) + h(y)$ citations each; (ii) those after paper $h(x)$ and up to paper $h(y)$ have at least $h(y)$ each; and (iii) the rest of papers have at most $h(x) - 1 + h(y)$ each. In view of this, $h(x + y) \leq h(x) + h(y) - 1 < h(x) + h(y)$.

Lemma 3.2. If an index f satisfies A1 and A2 then, for all $x \in D$, $f(x) \geq h(x)$.

Proof. For $n \in \mathbb{N}$, define $D_n = \{x \in D\text{: } d_x = n\}$ to be the set of outputs having exactly n papers. The proof is by induction on the sets D_n. Step 1: for all $x \in D_0$, $f(x) \geq h(x)$. Given that $D_0 = \{\varnothing\}$, if $x \in D_0$ then, by A1, $f(x) \geq min\{x^-, c_x\} = 0 = h(x)$. Step 2: for all $x \in D_1$, $f(x) \geq h(x)$. Let $x \in D_1$. If $x_1 = 0$ then, by A1, $f(x) \geq min\{x^-, c_x\} = 0 = h(x)$. If $x_1 \geq 1$ then, by A1, $f(x) \geq min\{x^-, c_x\} = 1 = h(x)$. Step 3: for all $n \in \mathbb{N}\backslash\{0, 1\}$ and $x \in D_n$, $f(x) \geq h(x)$. By steps 1 and 2, choose $n \in \mathbb{N}\backslash\{0, 1\}$ and, arguing inductively, suppose that, for all $k \in \{0, 1, \dots, n-1\}$ and $x \in D_k$, $f(x) \geq h(x)$. To show that, for all $x \in D_n$, $f(x) \geq$

$h(x)$, choose $x \in D_n$. With $h = h(x)$ and $k = f(x)$, suppose $k < h$. The aim is to reach a contradiction. Since $k \geq 0$, this implies that $h \geq 1$. Let $z \in D_h$ satisfy, for all $i \in \{1, \ldots , h\}$, $z_i = h$. Case 1: $h < n$. By the induction hypothesis,

$$f(z) \geq h(z) = h. \tag{1}$$

As $d_x = n > h = d_z$ and, for all $i \in \{1, \ldots , d_z\}$, $x_i \geq z_i = h$ it follows that $x \geq z$. Let $y \in D_k$ satisfy, for all $i \in \{1, \ldots , k\}$, $y_i = k$. By the induction hypothesis, $f(y) \geq h(y) = k$ and, by A1, $f(y) \leq y^+ = k$. Therefore, $f(y) = k$. Since $d_z = h > k = d_y$ and, for all $i \in \{1, \ldots , d_y\}$, $z_i = y_i$ it follows that $z \geq y$. By A2, $x \geq z \geq y$ and $f(x) = f(y) = k$ imply $f(z) = k$, which contradicts (1). Case 2: $h = n$. This implies that, for all $i \in \{1, \ldots , h\}$, $x_i \geq h$. By A1, $f(x) \geq min\{x^-, n\} = min\{h, n\} = h$, which contradicts the assumption that $f(x) < h$.■

Proposition 3.3. An index f satisfies A1, A2 and A3 if, and only if, f is the Hirsch index.

Proof. "⇐" Remark 3.1. "⇒" Let f be an index that satisfies A1, A2 and A3. For $n \in \mathbb{N}$, $C_n = \{x \in D\colon x^+ = n\}$ is the set of outputs in which n is the maximum number of citations that a paper receives. The proof is by induction on the sets C_n.

Step 1: f agrees with the Hirsch index on C_0. Let $x \in C_0$. By A1, $f(x) \leq x^+ = 0$. By Lemma 3.2, $f(x) \geq h(x) = 0$. Consequently, $f(x) = 0 = h(x)$.

Step 2: f agrees with the Hirsch index on C_1. Let $x \in C_1$. By A1, $f(x) \leq x^+ = 1$. By Lemma 3.2, $f(x) \geq h(x) = 1$. Therefore, $f(x) = 1 = h(x)$.

Step 3: for $n \in \mathbb{N}\backslash\{0, 1\}$, f agrees with the Hirsch index on C_n. By steps 1 and 2, choose $n \in \mathbb{N}\backslash\{0, 1\}$ and, arguing inductively, suppose that, for all $k \in \{0, 1, \ldots , n - 1\}$, f agrees with the Hirsch index on C_k. To show that f agrees with the Hirsch index on C_n, choose $x \in C_n$. With $h = h(x)$, by Lemma 3.2, $f(x) \geq h$. Since the proof amounts to obtaining a contradiction from $f(x) > h$, assume $f(x) > h$. If $x^+ < n$ then, by the induction hypothesis, $f(x) = h$. Accordingly, $x^+ = n$.

Case 1: $x_h > h$. For $x \in X$, define α_x to be the member y of X such that, for all $i \in \{1, \ldots , c_x\}$, $y_i = x_i - 1$ and, for all $i \in \{c_x + 1, \ldots , d_x\}$, $y_i = 0$. It follows from $x_h > h$ that $h(\alpha_x) = h(x)$. When $n = h$, it is plain that $h(\alpha_x^{-1}) < h(\alpha_x)$. When $n > h$, $h(x) = h$ implies $x_{h+1} \leq h$. In view of this, paper h in α_x (which is paper $h + 1$ in x) has at most $h - 1$ citations in α_x. Consequently, $h(\alpha_x^{-1}) \leq h - 1 < h = h(\alpha_x)$. To sum up, $h(\alpha_x^{-1}) < h(\alpha_x)$. By the induction hypothesis, $f(\alpha_x^{-1}) = h(\alpha_x^{-1})$ and $f(\alpha_x) = h(\alpha_x)$. Hence, $f(\alpha_x^{-1}) < f(\alpha_x)$. By A3, $f(x) < f(\alpha_x) + f(y)$, where $y \in X$ satisfies: (i) $d_y = d_x$; (ii) for all $i \in \{1, \dots, c_x\}$, $y_i = 1$; and (iii) for all $i \in \{c_x + 1, \dots, d_x\}$, $y_i = 0$. By step 2, $f(y) = 1$. Therefore, $f(\alpha_x) > f(x) - 1$. By the assumption that $f(x) \in \mathbb{N}$, $f(x) > h$ implies $f(x) \geq h + 1$. Accordingly, $f(\alpha_x) > f(x) - 1 \geq (h + 1) - 1 = h = h(\alpha_x)$: contradiction.

Case 2: $x_h = h$. Let $z \in X$ be such that: (i) $d_z = d_x$; (ii) for all $i \in \{1, \dots, h\}$ with $x_i = h$, $z_i = h + 1$; and (iii) otherwise, $z_i = x_i$. Output z is obtained from x by adding one citation to paper h and, if necessary, to those before paper h having the same number of citations as paper h. By Lemma 3.2, $f(z) \geq h(z)$. Since $h(x) = h$ and z comes from x by increasing the number of citations of a subset of the first h papers, $h(z) = h(x)$. Case 2a: $f(z) = h$. Let $v \in X$ satisfy: (i) $d_v = h$; and (ii) for all $i \in \{1, \dots, h\}$, $v_i = h$. By A1, $f(v) = h$. As $z \geq x \geq v$ and $f(z) = f(v) = h$, by A2, $f(x) = h$: contradiction. Case 2b: $f(z) > h$. By construction, $z_h > h$. In addition, $h(z) = h$. This leads to case 1, which (by replacing x with z) shows how to reach a contradiction from $f(z) > h$.■

Remark 3.4. Examples 3.5, 3.6 and 3.7 prove that no axiom in Proposition 3.3 is redundant.

Example 3.5. Let f be the index such that, for all $x \in D$, $f(x) = 0$. Then f satisfies A2 and A3; does not satisfy A1; and is not the Hirsch index.

Example 3.6. Let f be the index such that, for all $x \in D_0 \cup D_1$, $f(x) = h(x)$ and, for all $n \in \mathbb{N}\backslash\{0, 1\}$ and $x \in D_n$, $f(x) = x_2$. Then f satisfies A1 and A3; does not satisfy A2; and is not the Hirsch index.

Example 3.7. Let f be the index such that, for all $x \in D$, $f(x) = x^+$. Then f satisfies A1 and A2; does not satisfy A3; and is not the Hirsch index.

Remark 3.8. Example 3.9 proves that Proposition 3.3 does not hold if the index can take values in the set of non-negative rational numbers.

Example 3.9. With $D^* = \{x \in D: h(x) = 2 \text{ and } x_1 \geq 3\}$, let f be the mapping defined on D such that, for all $x \in D^*$, $f(x) = 2.1$ and, for all $x \in D\backslash D^*$, $f(x) = h(x)$. Then f satisfies A1, A2 and A3 but is not the Hirsch index.

References

Hirsch, J. E. (2005): "An index to quantify an individual's scientific research output", Proceedings of the National Academy of Sciences 102(46), 16569–16572.

Wikipedia (2008): http://en.wikipedia.org/wiki/Hirsch_index, accessed the 9th of November, 2008.

Woeginger, G. J. (2008): "An axiomatic characterization of the Hirsch-index", Mathematical Social Sciences 56(2), 224–232.