How Does the Market Use Citation Data? The Hirsch Index in Economics[†]

By Glenn Ellison*

A large literature following Hirsch (2005) has proposed citation-based indexes of individuals' research output. This paper views Hirsch's index as one member of a larger class and examines how well different indexes align with labor market outcomes for young, tenured economists at 50 US departments. Variants that emphasize smaller numbers of highly-cited papers are more aligned with labor market outcomes than is Hirsch's original index. It also examines how the market assesses jointly authored work, and how indexes can be adjusted for differences in citations across fields and years of experience. (JEL A14, C43)

There is great interest in "objective" numerical indexes that can be used to quantify researchers' output. They are a critical tool enabling studies that examine topics like the determinants of research productivity and whether women suffer from discrimination in academic labor markets. And they are used by market participants. A recent paper by Hirsch (2005) proposing that scientists could be ranked using a simple index h computed from citation data has spurred a rapidly growing literature that now includes scores of papers proposing alternate indexes. In this paper, I propose that a useful criterion for assessing citation indexes is to examine whether they are consistent with labor market outcomes. I carry out such an exercise for the economics profession using data on the citation records of young, tenured economists at 50 US departments. Among the findings are that Hirsch-like indexes are strongly correlated with labor market outcomes in economics (when subfield and age controls are used), and that indexes that focus on a smaller number of more highly cited papers are preferable to Hirsch's original index.

Analyses of the scientific research process often require quantitative measures of a researcher's output. The literature on citation indexes is motivated by the thought that earlier metrics of research output—usually based on quality-weighted publication or page counts—could be improved upon by focusing instead on citations; papers in second-tier journals that go on to be more influential than top-tier publications can be given more credit. This immediately, however, brings up an index problem: how should the vector containing the number of citations to each paper

a researcher has written be converted to a scalar that can be used as a variable in a regression (or compared across the candidates for a job or research grant)? Simply summing citations across all papers is seen as unattractive because the resulting number would be a noisy reflection of an underlying research production process and might change dramatically depending on how one allocated credit across authors on one or two highly cited, jointly authored papers.

Hirsch's h index is a simple clever construction. The index is defined to be the largest number h such that the researcher has at least h papers with h or more citations. It thereby de-emphasizes the number of citations to a researcher's most-cited paper. It is also seen as appealing that the focus of the index shifts when comparing researchers at different levels. When comparing young researchers, it emphasizes whether they have written a few papers that have had some impact; whereas when comparing distinguished senior researchers, it ignores minor papers and considers only papers that have a substantial number of citations. Although the h index has attracted a great deal of attention, the index is unappealing on its face when applied to economists. Economists write fewer papers than do physicists, and individual papers get many citations. As a result, the h index is uncomfortably like a publication count. An illustrative example is that Roger Myerson's h index is 44, a number that has already been surpassed by a number of economists in my sample with PhDs from the 1990s.[1] The h index is poorly aligned with the profession's view of Myerson because the profession cares much more about the several tremendously important papers Myerson has written than it does about whether his forty-fourth most-cited paper has as many citations as someone else's 44th most-cited paper.

Hirsch's original paper argued that an attractive feature of the h index is that it is not "arbitrary" as would be, for example, counting the number of papers with at least 100 citations. In this paper, however, I will view the Hirsch index as an arbitrary choice from a larger class; for any positive numbers a and b, I define the Hirsch-like index $h_{(a, b)}$ to be the largest number h such that a researcher has at least h papers with ah^b citations each. This arbitrariness can be thought of as an attractive feature of Hirsch-like indexes; indexes using different values of a and b could potentially be applied in different disciplines to better reflect how records of contributions would be regarded in each field.

Section III of this paper presents its first main empirical exercise—an assessment of which Hirsch-like indexes $h_{(a, b)}$ are most aligned with how the economics labor market appears to evaluate economists. I do the analysis in two different ways, but each is motivated similarly. I write down a simple model in which an economist's place of employment would be a noisy signal of how that economist would be evaluated by an unbiased, informed "market;" and then I note that in the model one can estimate which Hirsch-like indexes $h_{(a, b)}$ are more aligned with the idealized market assessment by fitting a model predicting where each economist will be employed. The analyses are carried out on a dataset that contains the citation records of almost all tenured economists at 50 top US departments who received their PhDs in 1987–2004. The motivation for restricting the analysis to relatively young,

---

[1] The index above and all others in this paper were computed using citation counts from Google Scholar, which tend to be substantially larger than citation counts from ISI, RePEc, and other sources. RePEc lists Myerson as having an h index of 18, which places him in a tie for 327th on their list.

tenured economists is that it seems most plausible that employment outcomes are aligned with "market" assessments in this group. The first analysis is an ordered probit estimation in which the 1995 NRC rank of each department is used as a proxy for each department's quality. The second analysis is a logit model that uses no a priori information about department qualities and instead views them as additional parameters to be estimated. The two estimates can be seen as relying on different sources of information. The first looks for an index formula that aligns individuals' citation indexes with the NRC ranks of the departments in which they are employed; whereas the second looks for an index formula that makes departments appear to be groups of economists with similar indexes.

Rather than focusing on which index $h_{(a, b)}$ provides the best fit I report how well the model fits for a range of values of a and b. The most important observation is that fits are substantially improved by departing from Hirsch's original index in the direction of increasing both a and b. The best fit is obtained using an index that in practice assesses tenured economists who are 20 years post-PhD roughly on their 5 to 10 most influential papers. But it is also true that a range of indexes can perform nearly as well if one incorporates appropriate corrections for differences across fields and years of experience. Indeed, a motivation for presenting many likelihoods rather than estimating a single best fit (a, b) is that the likelihood is sufficiently flat so that the "best" index would be sensitive to details of the sample selection.

Section IV further explores the performance of Hirsch-like indexes in two directions. First, it repeats the estimation on a number of different samples to explore whether conclusions about which Hirsch-like indexes are most aligned with labor market outcomes change when one examines younger economists, older economists, economists at the most highly ranked departments, etc. Here, I find that there are indexes that perform well and could be applied in a variety of situations. Second, I examine how well Hirsch-like indexes perform relative to old-fasioned citation counts. Here, I find that Hirsch's original index is inferior to simple citation counting, but an appropriate Hirsch-like index is an improvement.

Section V examines how the "market" credits jointly authored work. In earlier sections I maintain the assumption that each author receives $1/n$ credit for an n-authored paper. This is appealing because it reflects contributions to aggregate output. And in some applications, e.g., designing research assessment exercises, it may be necessary to prevent gaming. But whether the market does in practice credit joint work this way is far from obvious. And in other applications, e.g., in a study of discrimination in academic labor markets, a researcher would want to select an index that treated joint work however the market did. The analysis in Section V considers a generalized $h_{(a, b, c)}$ class of indexes that nests the $1/n$ credit model $(c = 1)$, giving full credit to each coauthor $(c = 0)$, and in between cases. In the full sample of young, tenured economists the best fit is obtained with a model that is closer to the full credit model than the $1/n$ model. But it is not clear if this is a broadly applicable conclusion or whether it reflects that the authors in this sample tend to be better known than their coauthors.

Section VI presents additional results designed to help researchers who wish to apply Hirsch-like indexes and also to provide a better sense of the goodness-of-fit obtained with these indexes. One conclusion of the earlier analysis was that citation indexes are more aligned with labor market outcomes if one controls for

differences across subfields. And correcting for years of experience is obviously extremely important. The analysis involves estimating nonlinear least squares models with $h_{(a, b, c)}$ indexes as the dependent variable. The estimation is carried out on an expanded dataset that also includes almost all faculty at the 50 departments from first-year assistant professors to the longest serving (nonemeritus) full professors. The estimated field dummies and experience effects from these regressions can serve as correction factors when applying Hirsch-like indexes. Estimated department dummies and estimates of the residual variance provide a sense of the magnitude of within- and across-department variation. Tables of adjustment factors appropriate for two different indexes—one maintaining the $1/n$ credit model and one not—are presented in the text and tables for additional indexes are presented in the Appendix.

This paper contributes to a rapidly growing literature. A number of papers following Hirsch (2005) proposed variants on his index including several that belong to the class studied in this paper.[2] The most closely related paper in terms of relating the citation indexes to labor market outcomes is Jensen, Rouquier, and Croissant (2009), which examines whether the h-index and other measures predict which CNRS researchers are promoted in a sample containing 586 candidates from many disciplines. They report that the h index outperforms the other measures considered, but that the predictive power of the regressions is low. Other papers have attempted to validate the h-index by showing that it is correlated with peer assessments and decisions of grantmaking bodies.[3] I have not seen other papers calculating adjustment factors that could be applied to improve comparisons between subfields of any other field, but Iglesias and Pecharromán (2007) propose that one could compare h-indexes across fields, e.g., between chemistry and mathematics, by multiplying the raw h-index by a field specific correction factor.[4]

Other noteworthy papers in the literature include Hirsch (2007), which compares the h-index and other measures (including variants that place more weight on highly cited papers) in terms of their ability to predict future success (measured with the h-index and other ways), and Lehmann, Jackson, and Lautrup (2006), which argues that one can derive an upper bound on the ability to distinguish between scientists using any index by examining how much the index varies when one resamples from the distribution of a scientists' citations, and that the h index fares worse in this test than do some other measures.

At least two papers have previously computed h-indices for economists. Ruane and Tol (2008) compute individual and successive h indexes to rank economics departments in Ireland. Tol (2009) computes h indices for the 100 economists with the largest number of papers listed in IDEAS/REPEC and finds that the h-index is highly correlated with several other citation and publication indexes.

---

[2] Kosmulski (2006) proposed what I refer to as the $h_{(1, 2)}$ index. Wu (2010) proposed the $h_{(10, 1)}$ version. Lehmann, Jackson, and Lautrup (2006) is mostly critical as opposed to advocating variants on Hirsch's index, but mentions that Hirsch's index is arbitrary in that $h_{(a, b)}$ indexes are also possible. Egghe (2008) and Schreiber (2008) proposed the fractional counting of coauthored papers I maintain through most of the paper.

[3] van Raan (2006) examined the correlation between the h index and peer ratings (on a three point scale) of 147 chemistry research groups in the Netherlands. Bornmann, Wallon, and Ledin (2008) examine grants to biomedical researchers.

[4] They discuss what these factors should be if citations in different fields follow different power laws and present correction factors for 21 fields derived from an analysis of citation distributions.

Hamermesh and Pfann (2012) does not compute Hirsch-like indexes, but is more closely related. Using a dataset containing information on tenured economists at 88 economics departments, it examines how the rank of the department in which an economist works is related to the economist's total citations, the number of citations to his or her most cited paper, and to the number of papers he/she has published. It finds that total citations are a significant determinant of job outcomes, but that the other two variables are not. It also examines other dependent variables including salaries and whether economists have won honors, such as being elected as a fellow of the Econometric Society.

I. The Hirsch Index

Hirsch's h index is defined as the largest number h such that the researcher has at least h papers with h or more citations each. It can be thought of as a count of the number of "good" papers that a researcher has written, with the clever addition that what "good" means becomes more demanding when comparing more accomplished researchers. It can also be thought of as similar to a citation count, but de-emphasizing the most-cited papers.[5] De-emphasizing outlier papers is seen as attractive for two reasons. First, the distribution of citations per paper has been reported to have a power law distribution with an infinite variance, which would make the sum of citations a noisy estimate the rate at which a researcher produces citations in expectation.[6] Second, it is sometimes not clear how much credit an author should receive for a joint paper, so it is attractive to reduce the impact that any one paper can have. Despite these attractive features, the h index is unappealing when applied to economics. Economists write fewer papers than do physicists, and individual papers get many citations. As a result, the h index is unappealingly similar to a publication count.

Through most of this paper I consider indexes that depart from Hirsch's in two ways. First, I define the index $h_{(a, b)}$ to be largest h for which a researcher has at least h papers with at least ah^b citations each.[7] Intuitively, one way in which the h index can be made less like a publication count is by raising the standard for what "good" means, so that the index focuses on a smaller number of papers. This can be achieved by increasing a or b. For example, when using both the $h_{(9, 1)}$ and $h_{(1, 3)}$ indexes a researcher achieves an index of 3 when he or she has 3 papers with at least 27 citations each, whereas the original h index reaches 3 as soon as a researcher has 3 papers with at least 3 citations each. The motivation for adding two parameters, a and b, instead of just one is that it also gives the index flexibility to capture how the profession shifts its focus when evaluating researchers at different levels. Intuitively, indexes with larger values of b increase the number of papers considered more slowly. For example, the $h_{(9, 1)}$ index counts the tenth paper if it has at least 90 citations, whereas the $h_{(1, 3)}$ index will ignore a researcher's tenth most cited paper unless it has at least 1,000 citations.

---

[5] Hirsch (2005) reports that h is approximately proportional to the square root of the total number of citations.

[6] See Price (1965) for an early study.

[7] For example, if a researcher has sole-authored papers with 85, 60, 35, 15, 10, and 8 citations, then the researcher's (10, 1) index would be 3 because he or she has 3 papers with at least 30 citations each. The researcher's (1, 2) index would be $\sqrt{15}$ because he or she has at least $\sqrt{15}$ papers with $\sqrt{15}^2$ citations each.

A second way in which I depart from Hirsch's index is in the treatment of jointly authored papers. Hirsch (2005) proposed giving a researcher full credit for each paper on which he is listed as a coauthor even if the paper has dozens of authors. Usually, I will consider coauthorship-adjusted versions of the index in which an author only gets $1/n$ credit for a paper with n authors.[8] But later in the paper I will also consider a broader class that nests both Hirsch's full-credit model and the $1/n$ credit model. I define the $h_{(a, b, c)}$ index to be the largest h for which a researcher has at least h papers with at least ah^b citations each when one gives $1/n^c$ for a paper with n authors. Here, $c = 1$ corresponds to the $1/n$ credit model and $c = 0$ corresponds to giving full credit for each paper regardless of the number of authors. Early in the paper, however, I often omit the c parameter for readability. When omitted it should be understood to be 1.

II. Data

The dataset for this paper includes citation records of almost all faculty at the top 50 economics departments in the 1995 NRC ranking. Faculty lists were taken from the departmental websites in the fall of 2011. Information on rank and year of PhD were collected from departmental and individual websites. Citation data for each economist were collected from Google Scholar. The data include the the number of citations and the number of coauthors for up to 100 papers by each researcher.[9] Economists were classified into 1 or more of 15 subfields primarily by mapping keywords appearing in descriptions of research interests on departmental or individual websites. Slightly less than half of economists are classified as working in a single field. The remainder are classified as working in multiple fields with the most common split being classified as two-thirds in one field (the one mentioned first in the description) and one-third in another. Two traditional fields (International and Econometrics) were divided into subfields because my impression from previous analyses of citations I have carried out at the paper-level suggest that there might be substantial differences.[10]

Table 1 presents summary statistics for two samples: the "young-tenured" sample, which includes 513 tenured economists who earned their PhDs in 1987–2004; and the "full" sample, which includes 1,486 faculty members. The analyses in the first part examining which citation indexes are most aligned with labor market outcomes will mostly use the "young-tenured" sample, whereas the analysis in the final section will analyze the full sample. The average number of papers per author is capped at 100 by the data collection procedure. It is larger than what one would get from counting lines on authors' CVs in part because Google sometimes fails

---

[8]For example, if a researcher has coauthored papers with 85, 60, 35, 25, 15, 10, and 8 citations, then the researcher's (10, 1) index would be 2 because he or she has 2 "papers" (the first four each given one-half credit) with at least 20 citations.

[9]More precisely, a page containing up to 100 references was downloaded by typing "au:"Firstname Lastname"" into the search box with advanced search parameters restricting the search to the "business, administration, finance, and economics" and "social sciences, arts, and humanities" subject areas. Some searches for which the first-name search seemed to miss papers were broadened to use just the first initial and some searches for which the first-name search included papers by another author were narrowed to also include a middle initial. A few authors for whom separating papers was difficult were dropped from the data. Most pages were downloaded on April 28–29, 2012.

[10]Some results from these anaylses appear in Ellison (2002b, 2011).

Table 1—Summary Statistics for Author Database

Variable	Mean	SD	Min.	Max.
Panel A. Young, tenured sample: 513 economists				
YearsPostPhD	16.7	5.2	8.0	25.0
NumPapers	72.0	25.3	8.0	100.0
AvgAuthors	2.2	0.3	1.2	3.1
TotalCites(full)	3,247	4,803	40.0	52,343
TotalCites($1/n$)	1,793	2,603	20.5	28,551
Count > 10	17.1	12.5	1.0	60.0
Count > 100	4.1	5.3	0.0	41.5
h	20.6	11.7	3.0	79.0
$h_{(1, 1, 1)}$	14.0	8.0	2.0	54.3
$h_{(15, 2, 1)}$	2.7	1.0	0.9	7.5
$h_{(90, 3, 1)}$	1.3	0.4	0.5	3.0
Panel B. Full sample: 1,470 economists				
YearsPostPhD	21.5	13.9	1.0	62.0
NumPapers	66.7	33.4	1.0	100.0
AvgAuthors	2.1	0.4	1.0	3.6
TotalCites(full)	4,550	8,699	0.0	120,561
TotalCites($1/n$)	2,684	5,399	0.0	60,420
Count > 10	20.3	19.4	0.0	85.9
Count > 100	5.7	9.6	0.0	85.9
h	22.0	18.0	1.0	100.0
$h_{(1, 1, 1)}$	15.6	12.9	1.0	85.9
$h_{(15, 2, 1)}$	2.6	1.5	0.3	9.8
$h_{(90, 3, 1)}$	1.2	0.5	0.2	3.5

to match different versions of a paper.[11] The average number of authors per paper understates the actual number of coauthors because in the period when we collected our data Google Scholar's display format often did not include all authors on papers with more than three authors.

The table also lists several citation related measures. The total citations variable is computed both giving full credit for each citation and giving $1/n$ credit for a citation to an n-authored paper. The variables giving the number of papers with at least 10 and 100 citations are computed under the $1/n$ counting system. The table also presents summary statistics for several Hirsch-like indexes. h is Hirsch's original index, which gives full credit for joint work. It has a mean of 20.6 in the young-tenured sample and 22.0 in the full sample.[12] The mean of about 20 illustrates the earlier informal comment that the h index becomes similar to a publication count. It is as if economists are judged on their twentieth most-cited papers. The $h_{(1, 1, 1)}$ index is similar, but gives $1/n$ credit for joint work. In practice the coauthorship adjustment tends to make it about two-thirds as large. Also listed are two Hirsch-like indexes that count many fewer papers for most authors and turn out to be more aligned with labor-market outcomes. The $h_{(15, 2, 1)}$ index has a mean of 2.7 and a standard deviation of 1.0 in our

[11] Google is much better at matching different versions than other sources. Counts are also inflated because Google often includes items that appear not to be academic papers. Hirsch-like indexes, fortunately, ignore this tail of little-cited items.

[12] The minimum value of 1 in the full sample reflects that I depart slightly from the stated formula for authors with zero citations and instead treat them as having one paper with a single citation. This is done to avoid dropping such authors from "full-sample" calculations. Such authors are, however, always dropped in analyses that compare the performance of different Hirsch-like indexes.

young-tenured sample. Hence, one can think of applying this index roughly as evaluating most economists in the young-tenured sample on the basis of their three to eight most-cited papers. The $h_{(90, 3, 1)}$ index has a focus that changes much less: in practice it assesses many young economists in terms of the citations to their second most cited paper and many senior economists on their third most cited paper.

III. Which Hirsch-Like Indices Predict Employment Patterns?

In this section, I present estimates from two models examining which Hirsch-like models are most predictive of labor market outcomes. The first model uses a traditional ranking of an economist's department as the dependent variable. The second does not impose any a priori ranking; each department's quality is estimated as part of the model. I find that Hirsch-like indexes that count smaller numbers of papers are more aligned with labor market outcomes than the original Hirsch index. But there is no clear "winner," and a range of somewhat different Hirsch-like indexes perform approximately as well.

A. *Estimates of an Ordered Probit Model Using NRC Rankings*

My first analysis uses the 1995 NRC rankings of the department at which each economist works as a proxy for the market assessment of an economist's quality. Note that the NRC rankings were conducted before most economists in my sample were at their current positions. Hence, the motivation for the use of the NRC rankings is that they may reflect persistent differences in departmental prestige/resources that may have allowed the higher-ranked departments to attract or retain highly regarded faculty.

The analyses in this section use various Hirsch-like indexes to try to explain the NRC rank of the department in which an economist works. An economic motivation for the model would be to assume that economics departments share a "market" assessment q_i^* of economist i and to assume that

$$q_i^* = q_i + \epsilon_i,$$

where q_i is the true quality of the economist and ϵ_i is an error term that could reflect biases or common errors in assessment, e.g., due to an economist having attended a less prestigious graduate school, having a powerful advocate, having turned out to be less promising than was thought at the time of his hiring, having written outstanding papers that were nonetheless rejected at top journals, etc. Suppose also that the true quality q_i is related to some Hirsch-like index $h_{(a, b)}$ by

$$q_i = \log\left(h_{(a, b)i}e^{\mathbf{X}_i \Gamma}\right) + u_i,$$

where $h_{(a, b)i}$ is economist i's Hirsch-like index; $\mathbf{X}_i$ is a vector of controls; and u_i is another error term.

Suppose also that the labor market for economists is such that matching is perfectly assortative; the N_1 economists with the highest market assessment q_i^* are

in department 1, the next N_2 are in department 2, and so on.[13] If the error terms u_i and ϵ_i are independent of the $\mathbf{X}$'s and $h_{(a, b)}$ with their sum being normally distributed, then this model would correspond to the classic ordered probit model: economist i would work in department j if $y_i^* \in [\mu_j, \mu_{j-1})$, where

$$y_i^* = \log(h_{(a, b)i}) + \mathbf{X}_i \mathbf{\Gamma} + u_i + \epsilon_i$$

is the unobserved attractiveness of the economist, and $\mathbf{\Gamma}$ and the thresholds μ_j are parameters that can be estimated. In this model a motivation for estimating a and b and using the $h_{(a, b)}$ index as a proxy for quality (rather than simply using the rank of the employing department) is that the Hirsch-like index is correlated with q_i and orthogonal to any systematic biases in the how the market assesses economists, i.e in the ϵ's, and, hence, might be appropriate for use in studies examining issues like discrimination.

We estimate this model on the sample of young, tenured economists. The vector $\mathbf{X}_i$ of controls includes a quadratic in years-post-PhD and estimated field effects for 15 subfields of economics.

To give a sense of how well different Hirsch-like indexes align with employment outcomes, the top panel of Figure 1 graphs the maximized per observation log likelihood of this model when the quality measure is taken to be the Hirsch-like indexes $h_{(a, b)}$ for $a = 0.5, 1, 2, 5, 10, 15, \ldots, 100$ and $b = 1, 2, 3$. The decision to focus on this set of (a, b) reflects in part that practitioners seem to prefer indexes involving round numbers. The results below also suggest that there would be little point in trying to instead derive precise estimates of a and b. A variety of somewhat different indexes provide similar fits, so the "best" a and b could vary greatly depending on the details of the sample chosen. The dashed line with circular marks graphs the likelihoods obtained with the $h_{(a, 1)}$ indexes. The second data point from the left corresponds to the $h_{(1, 1)}$ index.[14] The fit of the model using this index is substantially worse than the fit of models that focus on fewer papers by using larger values of a. The likelihoods increase monotonically until $a = 15$. Beyond this the likelihood is fairly flat. The best fit is obtained with $a = 75$, but there are local maxima around 15 and 55, and differences in fit are such that 15, 20, 25 and all values from 55 to 90 would not be rejected at the 5 percent level in a likelihood ratio test. The similar fits are obtained with fairly different indexes: the $h_{(15, 1)}$ index has a mean of 4.7, so, given the coauthorship adjustment, one can think of it as perhaps reflecting the citations of an economist's eighth or tenth most-cited paper; whereas the $h_{(90, 1)}$ index has a mean of 1.8, and therefore may reflect citations to the third most-cited paper.

The solid line with square markers graphs the fit of models using $h_{(a, 2)}$ indexes. They outperform the $h_{(a, 1)}$ indexes throughout most of the range. The $h_{(15, 2)}$ provides the best fit, but again we see that very different indexes also fit well with the $h_{(95, 2)}$, providing the third best fit among those graphed. The thin line with triangle markers graphs the fit of models using $h_{(a, 3)}$ index. The $h_{(60, 3)}$ index provides the

---

[13] A variety of models with complementarities between worker quality and firm quality can produce such an outcome.

[14] The index is not exactly Hirsch's because it gives partial credit for jointly authored papers. In practice, most papers are coauthored so Hirsch's index is more similar to the $h_{(2, 1)}$ index. In this dataset, a model using Hirsch's index would have a maximized per observation log likelihood of -3.623.

Likelihood of ordered probit model for various $h_{(a, b)}$

Likelihood of model with estimated dept. quality for various $h_{(a, b)}$

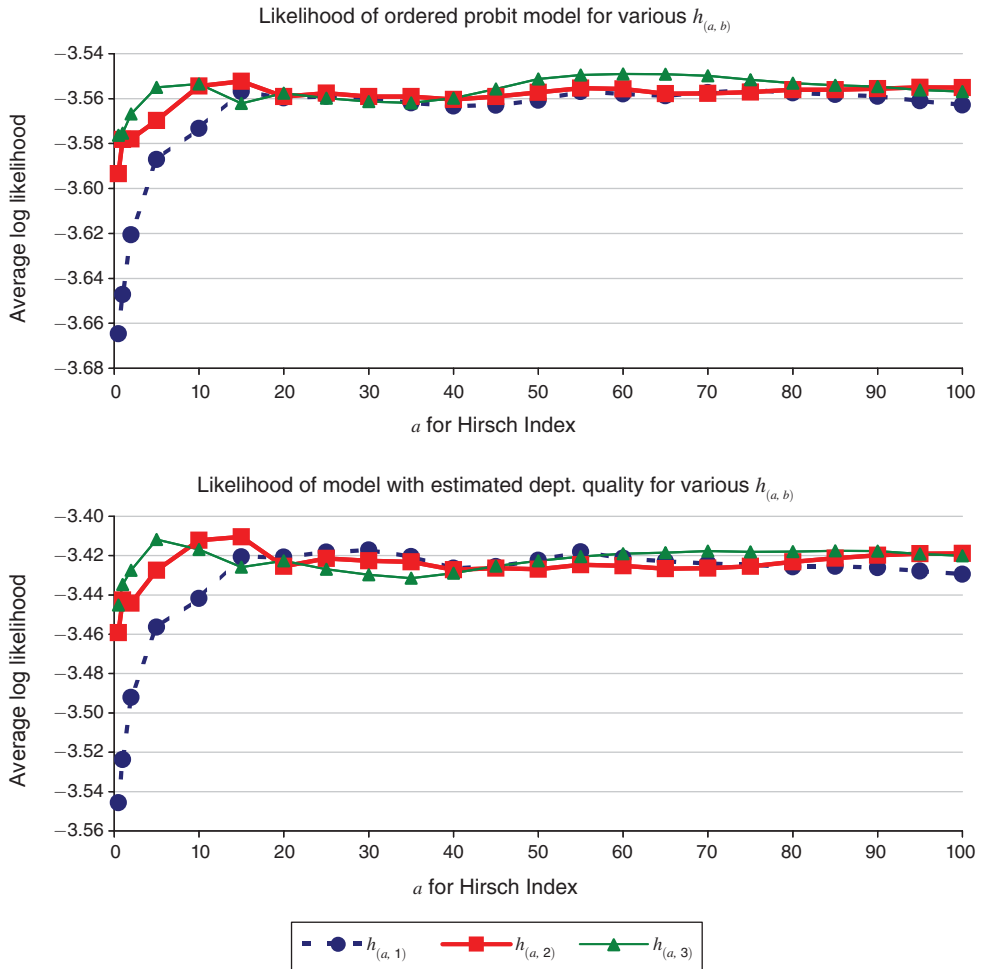$h_{(a, 1)}$ $h_{(a, 2)}$ $h_{(a, 3)}$

FIGURE 1. AVERAGE LOG LIKELIHOOD OF MODELS USING DIFFERENT $h_{(a, b)}$

best fit of any index considered and the difference between its likelihood and those of the $h_{(a, 1)}$ index lets one reject the $b = 1$ restriction at the 1 percent level. This index has a mean of 1.4.

To summarize, one main conclusion is that Hirsch-like indexes with larger a and/ or b are substantially more aligned with employment outcomes than the original Hirsch index. A second conclusion is that similar fits can be obtained using a fairly broad range of indexes.

An obvious limitation of the method of this section is that it assumes that the 1995 NRC rankings of departments are perfectly aligned with the market's assessment of researchers' contributions. These ratings are quite old by now, even in 1995 they were not designed solely to reflect research contributions, and in a world where economists and schools have idiosyncratic preferences, it is unrealistic to assume that any ranking of departments is perfectly aligned with how the market would assess economists.

B. *Estimates with Estimated Department Quality*

My second analysis does not employ any a priori information about rankings of different economics departments. Instead, it treats each department's quality as unobserved and uses a discrete choice model to simultaneously estimate which index the market is using and how departments should be ranked.

To motivate the model suppose that the utility that economist i would receive if employed at school j is

$$u_{ij} = \alpha_{0j} + \alpha_1 w_{ij} + \epsilon_{ij},$$

where α_{0j} is a preference common across economists; w_{ij} is the attractiveness of the wage (and nonwage) offer department j would make to i; and ϵ_{ij} is an idiosyncratic preference. Suppose also that the attractiveness of the offer that school j would make to economist i is

$$w_{ij} = \overline{w}_j + \delta_j \log(q_i) + \eta_{ij},$$

where $\overline{w}_j$ reflects the average offer by department j; q_i is the market's assessment of the quality of economist i; δ_j reflects both department j's monetary willingness-to-pay for quality and indirect benefits that it can provide to an economist of quality q_i due to department characteristics like visibility, student quality, etc.; and η_{ij} is another idiosyncratic preference. Finally, suppose that the market assessment of economist i is given by a Hirsch-like index adjusted for age and field differences:

$$q_i = h_{(a, b)i} e^{\mathbf{X}_i \Gamma}. \text{[15]}$$

In this model, departments have idiosyncratic preferences across economists and economists have idiosyncratic preferences across schools. Knowing where an economist works will not reveal the economist's quality for two reasons. Economists are not perfectly sorted into departments by quality, and no ranking of departments is assumed to be known. But there will be partial sorting with high q economists being relatively more likely to take jobs in high δ departments and low q economists being relatively more likely to take jobs in low δ departments.

If one assumes that the weighted sum $\epsilon_{ij} + \alpha_1 \eta_{ij}$ of the two idiosyncratic taste shocks is independent of the $\mathbf{X}$'s and of the Hirsch index, and that it has a type 1 extreme value distribution, then the model's prediction for the probability that economist i will work in department j takes on a simple and intuitive form:

$$\text{Prob}\{y_{ij} = 1 | \mathbf{X}_i, h_{(a, b)i}\} = \frac{e^{\beta_{0j} - \beta_1 d_{ij}}}{\sum_{j'} e^{\beta_{0j'} - \beta_1 d_{ij'}}},$$

---

[15] Note (unlike in the ordered probit model) I assume here that the market assessment is exactly a Hirsch-like index rather than being the sum of the Hirsch index and a noise term. A noise term could be added, but the probabilities would then no longer have a simple logit form.

where d_{ij} is a measure that can be thought of as reflecting the mismatch between economist i and department j,

$$d_{ij} = \left(\delta_j - (\log (h_{(a, b)i}) + \mathbf{X}_i \mathbf{\Gamma}) \right)^2;$$

and β_{0j} and β_1 are constants.[16] In this model, a motivation for estimating a and b and using $h_{(a, b)}$ is that the $h_{(\hat{a}, \hat{b}), i}$ is an estimate of the labor market's assessment of a researcher that can be constructed without having any a priori ranking of departments, and that can be superior to using rank-of-employing-department as a proxy because employment also reflects idiosyncratic preferences.

I again estimate this model by maximum likelihood on the sample of 513 young, tenured economists. The parameters to be estimated are the quality-related department dummies δ_j for each of the 50 departments, controls $\mathbf{\Gamma}$ consisting of a quadratic in years-post-PhD and field dummies for 15 fields, and the βs. To limit the number of parameters to be estimated I assume that $\beta_{0j} = \beta_0 \log(N_j)$, where N_j is the number of economists in the dataset who are observed to work in department j and estimate a single parameter β_0, rather than estimating 50 additional dummies.[17]

The dashed line with circular marks in the lower panel of Figure 1 graphs the maximized per observation log likelihood obtained by estimating the model using various $h_{(a, 1)}$. The pattern is similar to that obtained from the ordered probit model. The $h_{(1, 1)}$ index (which again is the second point from the left) performs much worse than indexes with larger a.[18] The best fit $h_{(a, 1)}$ model is obtained with $a = 30$, and likelihood is not quite as flat as that of the ordered probit model with only a's from 15 to 35, and 55 being close enough to avoid rejection at the 5 percent level.

The solid line with square markers again graphs the likelihoods of models using $h_{(a, 2)}$ indexes. The $h_{(15, 2)}$ provides the best fit of any of the indexes in the graph. The improvement in likelihood relative to the best-fit $h_{(a, 1)}$ index is sufficient to reject the $b = 1$ restriction at the 5 percent level. The $h_{(15, 2)}$ index has a mean of 2.7 in this sample and a standard deviation of 1, so one can think of it as roughly similar to judging young, tenured economists on the quality of their third to eighth most cited paper. Another observation from the ordered probit model that carries over to this model is that a fairly broad range of models work comparably well. The $h_{(90, 3)}$ index provides the seventh-highest likelihood of the indexes considered here. Recall that this index has a mean of 1.26 and a standard deviation of 0.36 so it is evaluating most economists in the sample on their second- or third-most-cited paper.

The main reason why ostensibly very different indexes can both be comparably aligned with labor market outcomes is that they are very highly correlated. The indexes could in theory be very different, but in practice economists who have a few extremely highly cited papers tend to also have a number of fairly highly cited papers, and vice versa. In the young-tenured sample, the correlation between the $h_{(15, 2)}$ and $h_{(90, 3)}$ indexes is 0.94. If we restrict to the much more homogeneous

---

[16] In terms of the primitives of the model, the constants are $\beta_{0j} = \alpha_{0j} + \alpha_1 \bar{w}_j + \frac{1}{2} \alpha_1 \delta_j^2$ and $\beta_1 = \alpha_1 / 2$.

[17] By construction, the estimated β_0 will be very close to one.

[18] Hirsch's original index without a coauthorship adjustment is better than the $h_{(1, 1)}$ index, but a little worse than the $h_{(2, 1)}$ index with a per observation log likelihood of -3.500.

sample of 61 tenured economists at top ten schools with 1987–1995 PhDs, the correlation is still 0.89.

The maximized likelihood of this model is greater than that of the ordered Probit model. This reflects that some departments with low NRC rankings have relatively high average Hirsch indexes, and vice versa. But estimated department qualities do end up being highly aligned with NRC rankings. For example, in the model using the $h_{(15, 2)}$ index, the ten departments with the highest estimated $\hat{\delta}_j$ are, in order, Harvard University, Massachusetts Institute of Technology, Princeton University, Stanford University, University of California at Berkeley, University of Chicago, Northwestern University, Columbia University, University of Minnesota, and Johns Hopkins University, and the correlation between the 50 estimated $\hat{\delta}_j$ and the 1995 NRC ratings is 0.80.

IV. Performance of Hirsch-Like Indexes

The Hirsch index has attracted an extraordinary amount of attention. For example, Prathap (2010) says

> *The h index (*Hirsch *2005) has rapidly captured the imagination of sciento-metricians and bibliometricians to such an extent that one can now divide the history of the subject virtually into a pre-Hirsch and a post-Hirsch period.*

Presumably, this excitement reflects a hope that the Hirsch-index is a sufficient advance on existing techniques so as to allow many new analyses of scientific research and of academia. In this section, I explore two questions related to the question of whether the Hirsch index is a great advance.

A. *Is the Hirsch Index a One-Size-Fits-All Index?*

Part of the appeal of the Hirsch index is that it adjusts its focus in a seemingly natural way when comparing researchers at different levels or different stages of their career. For example, a young researcher's $h_{(15, 2)}$ index advances from 0.5 to 1.0 when he or she goes from having some paper with four citations to having 1 sole-authored or 2 coauthored papers with at least 15 citations each. But when applied to distinguished senior researchers, who typically have indexes around 3 or 4, the index essentially ignores all papers with fewer than 100 citations, and instead focuses on the researcher's 5 to 10 most important papers. An index that is appropriate in a range of situations is appealing because it could facilitate broad-ranging studies examining researchers at different stages of their careers and across a wide range of institutions.

Although the way in which Hirsch-like indexes change when applied to researchers at different levels seems natural, whether any single Hirsch-index can actually capture how the market assesses researchers at different levels is an empirical question. In this section, I present related evidence by examining whether different Hirsch-like indexes work better on different subsamples.

First, I reestimated the model of Section IIIB (the model with estimated department qualities) on two subsamples defined by the NRC rank of the departments: one subsample is departments ranked 1 to 25 and the other is departments ranked 26 to 50. To avoid cluttered figures, I will present results here just for indexes with
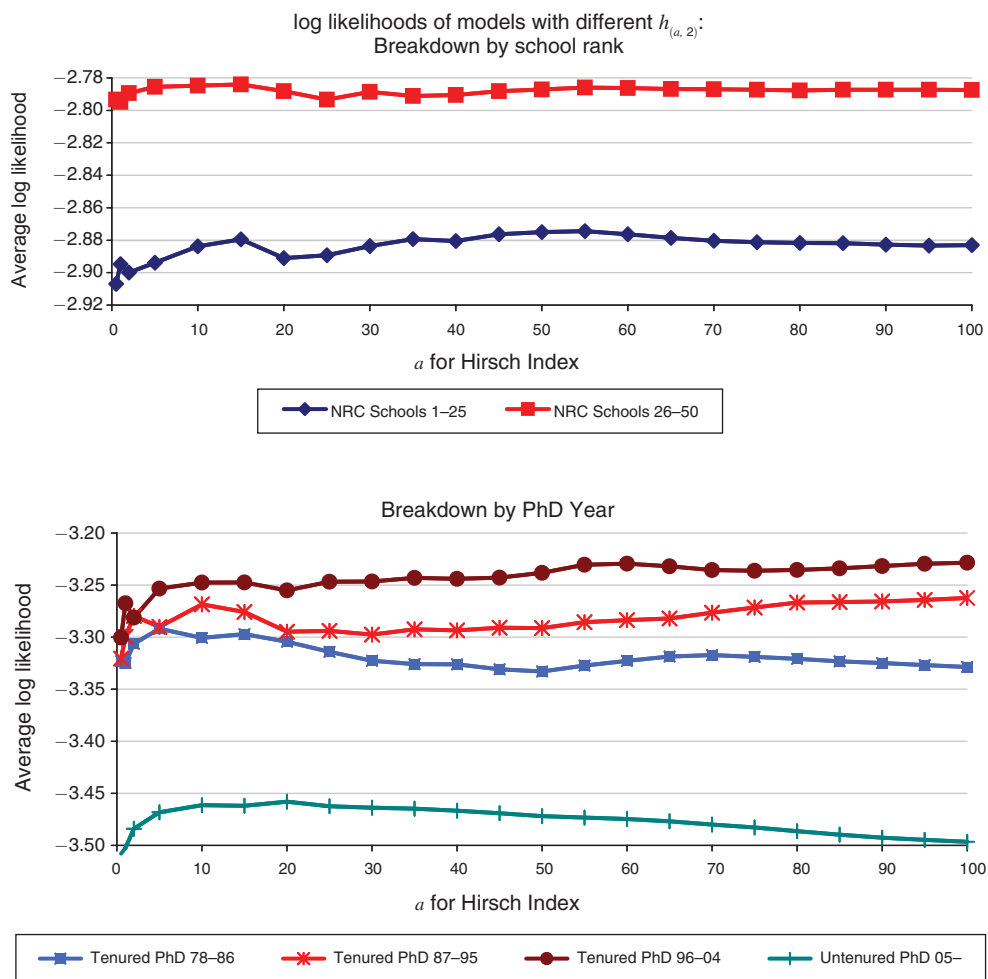
FIGURE 2. AVERAGE LOG LIKELIHOOD OF MODELS USING $h_{(a, 2)}$ INDEXES ON DIFFERENT SUBSAMPLES

$b = 2$. The top panel of Figure 2 graphs the estimated log likelihoods from models using $h_{(a, 2)}$ indexes as a ranges from 0.5 to 100. (Recall that the $h_{(15, 2)}$-based model provided the best fit in the full sample.) The lower curve, estimated on the sample of economists at NRC 1–25 departments, is maximized at $a = 55$. But the improvement in fit going from the $h_{(15, 2)}$ index to the $h_{(55, 2)}$ index is not significant at the 5 percent level. The upper curve, estimated using data on economists at NRC 26–50 departments, is maximized at $a = 15$. Hence, it appears that the $h_{(15, 2)}$ could be regarded as a reasonable measure to use in both sets of schools. It should be noted that differences between the two sets of schools are not so great. The mean value of the $h_{(15, 2)}$ index is 3.0 in the NRC 1–25 sample and 2.2 in the NRC 26–50 sample.

Second, I reestimated the model on four samples of economists at different career stages. The top two lines in the lower panel of Figure 2 present estimated likelihoods for the younger and older halves of the original young senior sample: the top line uses the 255 tenured economists with PhDs from 1996–2004; and the second uses the 258 tenured economists with PhDs from 1987–1995. A somewhat odd result is

that within each subsample the best fit is obtained using the largest value of *a,* and in each case, the increase in fit relative to $h_{(15, 2)}$ is significant at the 5 percent level. This is not inconsistent with the fact that the $h_{(15, 2)}$ provides the best fit in the full sample because the school dummies and field effects are separately estimated on each subsample and the two sets of estimates are fairly different; the correlation between the two sets of estimated school dummies is just 0.65 and the correlation between the two sets of field effects is just 0.38.

The figure also graphs comparable log-likelihoods obtained by reestimating the model on older and younger cohorts than those analyzed so far. The assumptions used to justify the model are more problematic in these samples—the model assumes that job offers reflect the market's assessment of each economist and that market assessments are perfectly aligned with the current Hirsch-index—but it still may be useful to see which Hirsch-like indexes most align with employment outcomes. The third line from the top in the graph, marked by squares, corresponds to the next older cohort: 302 tenured economists who received their PhDs from 1978–1986. The lowest line reports log likelihoods from estimating the model on younger econ-omists: 315 untenured economists at these departments with PhDs from 2005 or later. Citation levels in the two populations are quite different: the mean $h_{(15, 2)}$ index is 3.5 in the older sample and 1.0 in the younger sample. In the older sample the best fit is obtained with the $h_{(5, 2)}$ index. The $h_{(15, 2)}$ index provides the second-best fit among the indexes shown, and the difference in likelihood is not significant at the 5 percent level. One would expect it to be more difficult to use citations to explain employment outcomes for very young economists—most have very few citations—and indeed the fit of this model is substantially worse. The best fit is obtained using the $h_{(20, 2)}$ index. Again, the fit is not significantly better than that of the $h_{(15, 2)}$ index.

In summary, a single Hirsch-like index can perform reasonably well across a range of departments and age cohorts, although one might want to choose a somewhat dif-ferent index for some applications.

B. *Is the Hirsch Index a Revolutionary Advance in Scientometrics?*

Informal arguments for the Hirsch index often emphasize that an advantage rela-tive to traditional citation counts is that the latter can be heavily influenced by a researcher's most cited paper and can therefore become a very noisy measure of quality. But whether the way the Hirsch index avoids this problem actually makes it a more powerful tool than old-fashioned citation counting is again an empiri-cal question. To examine this issue I reestimated the model of Section IIIB using $1 + TotalCites$ in place of the $h_{(a, b)}$ index.[19]

The first row of Table 2 reports estimated log likelihoods from estimating three models with estimated department qualities on our standard sample of 513 young, tenured economists: one using $1 + TotalCites$, one using Hirsch's (2005) original *h*, and one using the $h_{(15, 2)}$ index. A comparison of the first two columns indicates that using Hirsch's index instead of a citation count would not be a revolutionary

---

[19] This count gives $1/n$ credit for citations to *n* authored papers. Due to the data collection procedure it includes at most the 100 most-cited papers of each author.

TABLE 2—COMPARISON OF FITS OBTAINED WITH TRADITIONAL AND
HIRSCH-LIKE CITATION MEASURES

	Maximized log likelihood using each citation index		
Sample	$1 + TotalCites$	h	$h_{(15, 2)}$
Standard young tenured	−3.43	−3.50	−3.41
NRC 1–25	−2.88	−2.91	−2.88
NRC 26–50	−2.80	−2.79	−2.78
Tenured PhD 1978–1986	−3.33	−3.39	−3.30
Tenured PhD 1987–1995	−3.28	−3.37	−3.28
Tenured PhD 1996–2004	−3.25	−3.32	−3.25
Untenured PhD 2005–	−3.50	−3.51	−3.46

advance; it would be taking a substantial step backward. The next column, however, illustrates that Hirsch's index does advance the field in the sense a suitable Hirsch-like index does have more explanatory power. The result does not hinge on a precise fine-tuning of the Hirsch-like index. Models using the $h_{(a, 2)}$ index outperform the *TotalCites* model for all but one of the a's from 10 to 100 that I have tried. Note also that even if the advantage is not so large in magnitude, this does not mean that the Hirsch-like index is not a powerful tool. It may be that traditional citation counting was underappreciated.

To investigate the source of advantage/disadvantage further, the remaining rows of Table 2 report log likelihoods on the subsamples discussed above. The breakdown by the NRC rank of the department indicates that the $h_{(15, 2)}$ index is not superior to the old-fashioned measure when applied to young, tenured economists in top 25 departments.[20] The $h_{(15, 2)}$ index does perform better than *TotalCites* on the schools with NRC ranks from 26 to 50. The Hirsch-like index's reduced sensitivity to outliers may be a more desirable feature in this sample. Indeed, even the original Hirsch index outperforms the *TotalCites* measure in this sample.

The next four rows show breakdowns by PhD year. Here, the $h_{(15, 2)}$ index outperforms the the *TotalCites* measure in all age groups. The advantage is nonmonotonic, being largest in the oldest and youngest cohorts considered and very small in the middle two cohorts.

V. How Does the Market Treat Jointly Authored Work?

In the models estimated so far I have maintained the assumption that authors get $1/n$ credit for n-authored papers. In some applications using indexes that credit work in other ways seems unwise, e.g., if a research assessment exercise gave greater than $1/n$ credit to n-authored papers, then departments could boost their ranking by encouraging faculty to add each other as coauthors. But it is far from clear that the market does give $1/n$ credit to n-authored papers. How the market treats jointly authored work is a practically important question because it should impact researchers' choices of projects and thereby affect the aggregate productivity of the profession. In addition to this academic motivation, examining how the market treats

[20] In this subsample the *TotalCites* model outperforms the $h_{(15, 2)}$ index model (−2.875 versus −2.879) and also outperforms the almost all of the $h_{(a, b)}$ models I have estimated.

TABLE 3—COMPARISON OF FITS OBTAINED WITH DIFFERENT COAUTHORSHIP ADJUSTMENTS

	Maximized log likelihood using best $h_{(a, b, c)}$ index				
Sample	$c = 0$	$c = 0.25$	$c = 0.50$	$c = 0.75$	$c = 1$
Standard	−3.413	−3.406	−3.409	−3.407	−3.410
NRC 1–25	−2.852	−2.854	−2.856	−2.865	−2.874
NRC 26–50	−2.782	−2.780	−2.780	−2.781	−2.783
Tenured PhD 1978–1986	−3.300	−3.296	−3.295	−3.291	−3.292
Tenured PhD 1987–1995	−3.251	−3.251	−3.256	−3.262	−3.262
Tenured PhD 1996–2004	−3.215	−3.216	−3.216	−3.212	−3.223
Untenured PhD 2005–	−3.477	−3.466	−3.453	−3.446	−3.448

coauthored work shares the practical motivations of our other analyses. In some applications, such as in studies in which a citation index is being used as a proxy for the market's assessment of a researcher, one would want to weight multiauthored papers differently if the market did so.

To investigate how the market treats multi-authored papers I consider here a further generalization of the Hirsch index. I define $h_{(a, b, c)}$ to be the largest number h such that a researcher has at least h papers with at least ah^b citations each when n-authored papers are counted as $1/n^c$ of a paper. Note that this index with $c = 1$ corresponds to the $1/n$-credit model that we have used so far, and that $c = 0$ corresponds to giving each author full credit for a multiauthored paper. Values of c between 0 and 1 give researchers more than $1/n$ but less than full credit for jointly authored papers. For example, when $c = 0.5$ a researcher is given 0.71 credit for a coauthored paper and 0.58 credit for a three-authored paper.

Table 3 lists the highest log likelihood of the model of Section IIIB over the same set of a and b we have been using for various values of c. The best fit is obtained for $c = 0.25$, using the $h_{(10, 3, 0.25)}$ index. The improvement in likelihood relative to the best index with $c = 1$ is again significant at the 5 percent level. This provides evidence that the "market" gives researchers more than $1/n$ credit for jointly authored papers. An index with $c = 0.25$ gives each author 0.84 credit for a coauthored paper and 0.76 credit for a three-authored paper. If the market does assess researchers in this manner it would provide a strong incentive for coauthoring.

The next two rows examine a breakdown by NRC rank. In the NRC 1–25 subsample the full credit ($c = 0$) model provides the best fit and the improvement is significant at the 1 percent level. In the NRC 26–50 subsample a model with $c = 0.5$ provides the best fit, but neither $c = 0$ nor $c = 1$ can be rejected at the 5 percent level. The bottom four rows of the table examine a breakdown by experience cohorts. In three of the four cohorts the $c = 0.75$ model provides the best fit. The strongest evidence is obtained in the untenured cohort; the full-credit model performs much worse than the $1/n$ credit model. The full-credit model outperforms the $1/n$ model in two of the three tenured cohorts.

An extension of this analysis that seems natural and might potentially help explain some differences across cohorts would be to examine whether the market credits joint papers differently depending on the identity of the coauthors. For example, the market might give a researcher nearly full credit if she is better known than her coauthors, and otherwise divide credit more evenly. This could explain why the

full-credit model does not work as well for young economists—they could be getting less credit on joint papers with advisors and senior colleagues. Our data, unfortunately, is not well suited to investigate this further. When we collected our data Google Scholar would replace many coauthors' names with an ellipsis. And we are also limited by the fact that many coauthors are not in our dataset.

VI. Using Hirsch-Like Indexes: Patterns across Fields, Time, and Schools

People using citation indexes typically want to do one of two things: compare the accomplishments of a set of researchers, or assess whether a given researcher seems appropriate for a particular position. For the former, one needs to be able to correct for differences across fields and years of experience. For the latter, one also needs some idea of what would be typical for the department in question. This section presents estimates that facilitate such comparisons. The results will also give more of a sense of the degree to which labor market outcomes and citation indexes are aligned.

As noted above a variety of Hirsch-like citation indexes are similarly aligned with labor market outcomes and could reasonably be adopted for a citation analysis. One would, however, need to know how to make age adjustments and field adjustments appropriate for the chosen index. In this section I will present results using the $h_{(15, 2, 1)}$ and $h_{(10, 3, 0.25)}$ indexes. The $h_{(10, 3, 0.25)}$ index provided the best fit in our standard sample and therefore might be a good choice if one were interested, for example, in studying if there appeared be discrimination against some group in the labor market. The $h_{(15, 2, 1)}$ index was the best fitting $c = 1$ index in the standard sample and also performed well in most of the subsamples. It would be a reasonable index to apply if a researcher or market participant felt that authors "should" receive $1/n$ credit for jointly authored work or that the improvement in fit from indexes that give more credit for jointly authored work is not sufficiently large or consistent across subsamples to warrant departing from the $1/n$ benchmark. The Appendix reports adjustment factors for two other indexes as well: one that might be chosen if one wanted to give each author full credit on jointly authored papers, and one that might be chosen if one wanted to stick with the simpler $h_{(a, 1, 1)}$ class.

In order to provide as broad a view as possible, the analyses of this section will use the full dataset of 1,486 economists. Relative to the standard dataset of the previous sections this adds 360 untenured faculty with PhDs from 2001 or later, 604 full professors who earned PhDs prior to 1987, and 9 tenured faculty with post-2004 PhDs.[21] The main focus will be on a simple model of how the $h_{(a, b, c)}$ indexes vary across school, fields, and age cohorts estimated by weighted nonlinear least squares:

$$h_{(a, b, c)i} = \left(\sum_s \delta_s School_{is}\right)e^{\mathbf{X}_i\mathbf{\Gamma}} + \epsilon_i.$$

---

[21] The sample consists of almost all faculty at the 50 departments considered. The largest omitted group is researchers for whom we had difficulty separating their citations from citations to others with similar names. The sample construction also omits assistant professors who earned their PhDs prior to 2002 and associate professors who earned their PhDs prior to 1987. Assistant professors with zero citations, who were not included in the previous analyses comparing different Hirsch-like indexes, are included here and treated as having one paper with a single citation.

The vector of controls $\mathbf{X}_i$ will include both controls for different fields and a flexible function of the number of years post-PhD. There is substantial heteroskedasticity in the ϵ_i in the above equation. Hirsch-like indexes are close to zero for first-year assistant professors, whereas there is a great deal of variation among older economists. Accordingly, I estimate this using a feasible GLS-style procedure weighting the observations by the inverse of a predicted residual variance.[22] Alternatively, I also estimate the parameters via the related least-squares regression:

$$\log(h_{(a,\,b,\,c)i}) = \left(\sum_s \log(\delta_s) School_{is}\right) + \mathbf{X}_i\mathbf{\Gamma} + \epsilon_i.$$

Here, there is also systematic heteroskedasticity—there is more noise among very young and very old economists than among those in the middle—so I again use a weighted least squares procedure.[23]

A. *Variation in Citation Indexes across Fields*

There is substantial variation in citations across fields of economics. Citation indexes are more aligned with employment outcomes if one adjusts for differences across fields, so researchers comparing economists in different fields may want to make such adustments. The first two panels of Table 4 report the estimated field effects from our base NLLS regression of the $h_{(15,\,2,\,1)}$ and $h_{(10,\,3,\,0.25)}$ indexes on field dummies, a flexible function of years Post-PhD., and school dummies. The omitted field is macroeconomics, so the coefficient estimates can be thought of as measuring the extent to which economists in the field in question tend to have a higher or lower citation index than macroeconomists in the same department.

In each case, the three largest estimates turn out to be those behavioral/experimental, time series econometrics, and international trade. Researchers in these fields, on average, have $h_{(15,\,2,\,1)}$ indexes that are about 10 percent higher than those of their macro colleagues although the differences fall short of being significant at the 5 percent level. At the other extreme are economic historians who are estimated to have $h_{(15,\,2,\,1)}$ indexes that are more than 30 percent below those of their macro colleagues. Economic theorists have indexes about 20 percent below those of their macro colleagues, and researchers in industrial organization, cross-section econometrics, public finance, and development are 10 percent to 15 percent below. Each of these estimates is statistically significant at the 5 percent level. The field corrections for the $h_{(10,\,3,\,0.25)}$ index are similar, but usually smaller in magnitude. This is primarily due to the difference in the b parameter; a cubic ($b = 3$) index will be less affected by a change in citations than a quadratic ($b = 2$) index.

---

[22] More precisely, the model was first estimated via a standard NLLS regression. A second NLLS regression was then used to estimate the relationship between the squared residuals and a constant, the log of the NRC rank of the department, and the same flexible function of years post-PhD, assuming a relationship of the form $\hat{\epsilon}_i^2 = e^{\mathbf{X}_i\beta} + u_i$. The square root of the inverse of the predicted values from this regression were then used as weights in a weighted NLLS estimation of the model.

[23] In this case, the regression of the squared residuals on the NRC rank and a function of years post-PhD is done via an OLS regression.

TABLE 4—ESTIMATED FIELD DUMMIES FROM NLLS MODELS AND FROM A
PAPER-LEVEL CITATION REGRESSION

| | Field Effects in NLLS Models Using $h_{(a, b, c)}$ | | | | Field effects in a citation regression | |
| | $h_{(15, 2, 1)}$ | | $h_{(10, 3, 0.25)}$ | | | |
Field	Coef. est.	SE	Coef. est.	SE	Coef. est.	SE
Behavioral/exper.	0.10	0.06	0.08	0.04	−0.21	0.11
Development	−0.11	0.05	−0.06	0.04	0.14	0.19
Finance	0.02	0.06	0.03	0.04	−0.13	0.09
History	−0.38	0.06	−0.32	0.04	−0.33	0.22
Industrial org.	−0.14	0.04	−0.10	0.03	−0.25	0.10
International finance	−0.09	0.06	−0.04	0.04	0.32	0.18
International trade	0.09	0.06	0.04	0.04	−0.10	0.12
Labor	−0.05	0.04	−0.04	0.03	−0.05	0.09
Metrics-cross sec.	−0.14	0.04	−0.10	0.03	−0.28	0.15
Metrics–time ser.	0.10	0.07	0.05	0.05	0.12	0.15
Micro theory	−0.22	0.03	−0.18	0.02	−0.28	0.07
Public finance	−0.12	0.04	−0.11	0.03	−0.19	0.10
Political econ.	−0.09	0.07	−0.11	0.04	−0.35	0.14
Other	−0.06	0.07	−0.09	0.06	−0.20	0.11

The estimated field effects are of sufficient magnitude to make them practically important. For example, a time series econometrician at the twentieth-ranked school would be expected to have a higher $h_{(15, 2, 1)}$ index than a macroeconomist at the tenth-ranked school, whereas an economic historian at the twentieth-ranked school would be expected to have a lower $h_{(15, 2, 1)}$ index than a macroeconomist at 46 of the 50 schools in the sample (and likely below many schools not in the sample as well).

The final two columns of the table present quantitatively noncomparable estimates of field effects from a different regression/data source. The dataset is one from Ellison (2002b). It contains information on 1,393 papers published in the top five general interest journals in 1990–1998. The regression has the log of the total citations to each paper as the dependent variable and includes a number of controls in addition to field fixed effects as explanatory variables.[24] There are a number of similarities across columns. The fields in which papers are estimated to have the fewest citations—political economy, history, micro theory, cross-section econometrics, and industrial organization—are all fields where we estimated that researchers have relatively low citation indexes. But the lists are not perfectly aligned. The most notable exceptions are behavioral/experimental and international finance.

B. *Variation in Citation Indexes with Academic Age*

An economist's Hirsch index will increase over time for two reasons: he or she will write more papers; and existing papers will accumulate more citations. Hirsch's (2005) original paper included a model of these two processes under which the Hirsch index would be directly proportional to the number of years for which a researcher has been active.

---

[24] Citations were collected from ISI in February of 2000. Other regressors include journal dummies, journal-specific time trends, several author characteristics, the length of the paper, and the order in which the paper appeared in its journal issue.
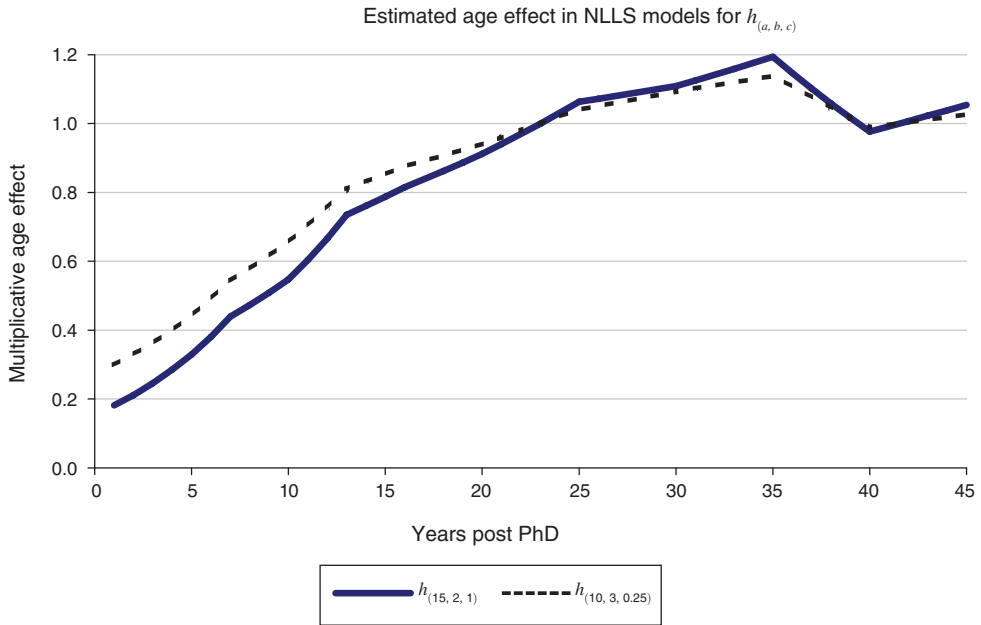
FIGURE 3. ESTIMATED EXPERIENCE EFFECTS FOR THE $h_{(15, 2, 1)}$ AND $h_{(0, 3, 0.25)}$ INDEXES

The NLLS model includes a flexible multiplicative experience effect. The Hirsch-like index of a researcher who is t years post-PhD is modeled as differing from the base level by a multiplicative factor $e^{-f(t)}$. The function $f(t)$ was assumed to be piece-wise linear with the slope allowed to change at 1, 4, 7, 10, 13, 16, 20, 25, 30, 35, and 40 years post-PhD. The estimated slope coefficients for the $h_{(15, 2, 1)}$ and $h_{(10, 3, 0.25)}$ are reported in the Appendix. Most are fairly precisely estimated and growth rates are substantially higher in the early years than in the later ones.

Figure 3 provides graphs the estimated $e^{-\hat{f}(t)}$ versus years post-PhD. (The models are normalized by setting this multiplicative age effect to 1 for researchers who are 23 years post PhD.) The solid line is that for the $h_{(15, 2, 1)}$ index. The most striking pattern is that $h_{(15, 2)}$ index appears to grow approximately linearly from years 1 through 25. The slope is perhaps a bit steeper at a of couple points, but there is no dramatic jump up around the time of the tenure decision. The index grows somewhat more slowly from year 25 to a peak at 35 years post-PhD. Researchers who are five years post PhD tend to have citation indexes that are a little more than one-fourth as large as those of researchers at the peak age. Researchers who are 11 years post-PhD have indexes about one-half as large. Any individual researcher's Hirsch-index must increase monotonically over time. Accordingly, the lower average $h_{(15, 2, 1)}$ indexes for older researchers must be due to some combination of other factors: more of the papers citing their work may never have made it onto the web and into Google Scholar's database; citations have been increasing over time and current researchers may be more likely to cite research performed by a younger generation; etc. The dashed line in the figure graphs the corresponding pattern for the $h_{(10, 3, 0.25)}$ index. It is quite similar. The most notable difference is that the $h_{(10, 3, 0.25)}$ index is relatively larger for the youngest researchers and grows more slowly.

TABLE 5—ESTIMATED SCHOOL DUMMIES FROM NLLS MODELS FOR $h_{(15, 2, 1)}$ AND $h_{(10, 3, 0.25)}$

School (NRC rank)	$h_{(15, 2, 1)}$		$h_{(10, 3, 0.25)}$	School (NRC rank)	$h_{(15, 2, 1)}$		$h_{(10, 3, 0.25)}$
	Coef.	SE	Coef.		Coef.	SE	Coef.
Harvard (1)	5.98	0.27	4.47	Penn State (45)	3.04	0.19	2.78
MIT (3)	5.31	0.28	4.04	Texas (31)	3.03	0.24	2.68
Chicago (2)	4.97	0.38	3.81	Virginia (24)	3.01	0.18	2.66
UC-Berkeley (7)	4.71	0.23	3.76	Purdue (50)	2.97	0.29	2.65
Princeton (5)	4.71	0.21	3.88	George Mason (47)	2.94	0.17	2.64
Stanford (4)	4.66	0.26	3.70	Wash-St. Louis (29)	2.92	0.21	2.77
Columbia (12)	4.25	0.23	3.53	UC-S. Barbara (49)	2.90	0.18	2.64
Northwestern (9)	4.18	0.23	3.42	Vanderbilt (48)	2.89	0.19	2.55
Yale (6)	4.09	0.20	3.39	Arizona (37)	2.88	0.23	2.62
UC-San Diego (16)	3.91	0.23	3.22	Ohio State (35)	2.87	0.17	2.56
UCLA (11)	3.90	0.20	3.27	Rochester (14)	2.87	0.17	2.69
Penn (8)	3.89	0.19	3.31	Southern Cal (40)	2.77	0.22	2.47
NYU (17)	3.88	0.19	3.21	Michigan State (27)	2.76	0.18	2.54
Brown (23)	3.84	0.29	3.21	Illinois (28)	2.74	0.16	2.51
Boston U (21)	3.77	0.21	3.19	Washington (26)	2.64	0.18	2.52
Wisconsin (15)	3.71	0.22	3.09	NC State (42)	2.64	0.16	2.46
Maryland (20)	3.65	0.17	3.20	Iowa (30)	2.59	0.15	2.47
Cornell (18)	3.63	0.18	3.11	North Carolina (25)	2.55	0.20	2.41
Johns Hopkins (32)	3.62	0.29	2.94	Pittsburgh (34)	2.50	0.22	2.48
Michigan (13)	3.58	0.15	3.06	Rice (46)	2.48	0.20	2.30
Cal Tech (19)	3.54	0.32	3.00	Texas A&M (33)	2.47	0.18	2.44
Duke (22)	3.43	0.24	2.97	Indiana (44)	2.42	0.19	2.35
Minnesota (10)	3.42	0.19	3.02	SUNY-SB (39)	2.26	0.14	2.27
UC-Davis (38)	3.33	0.29	2.89	Iowa State (36)	2.24	0.16	2.23
Boston College (43)	3.12	0.21	2.75	Florida (41)	2.21	0.16	2.29

The Appendix reports estimates from two other indexes and provides a large table of adjustment factors.

C. *Average Values of the $h_{(15, 2, 1)}$ and $h_{(10, 3, 0.25)}$ Index by School*

The NLLS model also includes estimated school dummies. Table 5 presents these estimates. The coefficients can be thought of as reflecting the expected values of the $h_{(15, 2, 1)}$ and $h_{(10, 3, 0.25)}$ indexes for a macroeconomist working at the school who received his or her PhD in 1989. They can also provide a benchmark for evaluating the citation records of other economists by applying the field adjustments and age adjustments reported above. The estimates turn out to be quite aligned with traditional reputational rankings—the correlation of the school dummies from the $h_{(15, 2, 1)}$ model with the 1995 NRC rating is 0.86. The highest estimate (for NRC #1 Harvard) is 5.98. The next 5 schools have indexes ranging from 4.66 to 5.31. The tenth largest estimated coefficient is 3.91, the twentieth is 3.58, the thirtieth is 2.94, and the fortieth is 2.64. Most of the standard errors are around 0.2. The estimated school dummies for the $h_{(10, 3, 0.25)}$ model have an 0.994 correlation with the estimates from the $h_{(15, 2, 1)}$ model (and an even higher 0.88 correlation with the 1995 NRC ratings). They range from 2.23 to 4.47.

D. *Within-School Variation*

In an additive sense, there is almost no variation in the $h_{(15, 2, 1)}$ or $h_{(10, 3, 0.25)}$ indexes among new assistant professors, almost all have indexes very close to 0, and
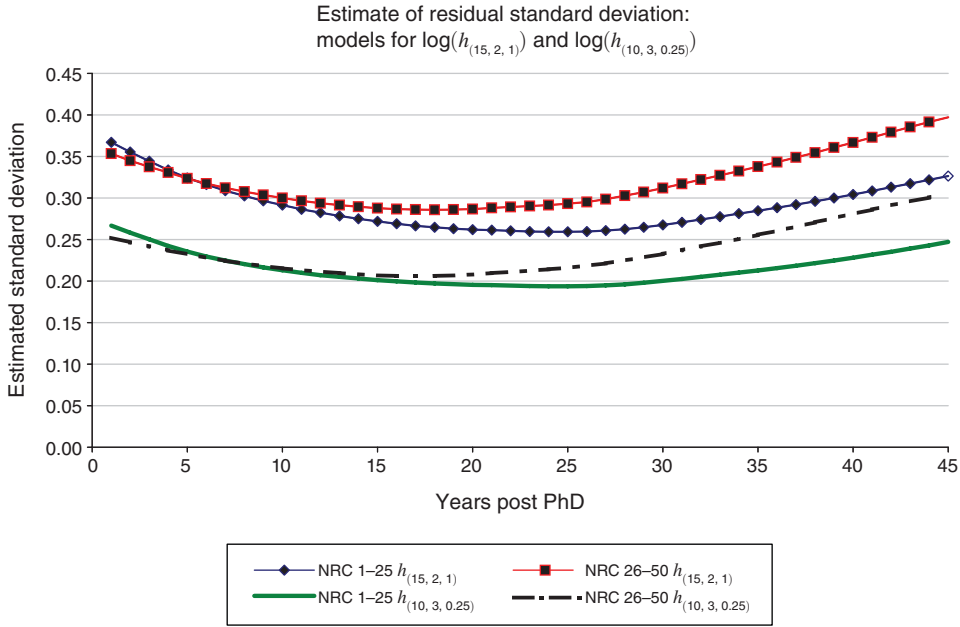
Estimate of residual standard deviation:
models for $\log(h_{(15, 2, 1)})$ and $\log(h_{(10, 3, 0.25)})$

FIGURE 4. ESTIMATE OF IDIOSYNCRATIC WITHIN-SCHOOL VARIATION IN $h_{(15, 2, 1)}$ AND $h_{(0, 3, 0.25)}$ INDEXES

much more variation among older economists. To provide a view of the variation that is informative across a broad range of situations, Figure 4 graphs estimated variances for the error terms in the regression

$$\log(h_{(a, b, c)i}) = \left(\sum_s \log(\delta_s) School_{is} \right) + \mathbf{X}_i \mathbf{\Gamma} + \epsilon_i.$$

Specifically, I first estimated this model using a weighted least squares procedure, then performed a local linear regression to estimate the expected value of the square of the residual as a function of years post-PhD (using a bandwidth of 10), and then graph the square root of this estimated variance. The local linear regression was done separately for economists at schools with NRC ranks from 1–25 and for economists at schools with NRC ranks from 26–50.

The lower line with markers in Figure 4 is that for economists at schools in the NRC top 25. The estimated standard deviation reaches a minimum of about 0.26. One can think of this informally as saying that the within school variation in the $h_{(15, 2, 1)}$ index tends to be about 26 percent of the mean. Relative to the estimated school dummies one can think of this as implying that an economist at the sixth-ranked department (Stanford) would be one standard deviation above average for that department if his or her index was slightly above the Harvard mean and one standard deviation below average if his or her index was around the mean of the twentieth highest department (Michigan).

There is more variation in this log sense for younger and older economists, but the standard deviation remains between 0.26 and 0.30 for most tenured cohorts (those between 9 and 38 years post Ph.D.). The upper line with markers is that for economists at schools with NRC ranks from 26–50. There is somewhat more idiosyncratic

variation in this sample and the minimum occurs earlier in the career: the minimum is about 0.29 and the estimate is below 0.3 for those from 10 and 27 years post PhD.

The lines without markers give corresponding estimates for the $h_{(10, 3, 0.25)}$ index. The residual standard deviation is lower here. It drops below 0.20 for economists in the NRC 1–25 group who are 16–29 years post-PhD, and is below 0.21 for economists in the NRC 26–50 group who are 13–21 years post PhD. The lower variance is due to the nature of a $b = 3$ index. Outliers are less extreme because the cube root means that highly cited researchers end up with lower indexes and little-cited researchers have higher indexes.

E. *Some Highly Cited Economist Lists*

In this section, I present a few lists of highly cited economists as an illustration of the performance of citation indexes (and because the outliers may be informative).

First, the left half of Table 6 presents lists consisting of the five economists in several cohorts who have the highest field- and age- adjusted $h_{(15, 2, 1)}$ indexes. The table lists both the $h_{(15, 2, 1)}$ index and an adjusted index normalized to a macroeconomist with a 1989 PhD. Unsurprisingly, the list contains many highly regarded economists. Four of the five in the oldest cohort have won Nobel prizes.

The right half lists the economists in each cohort who have the highest field-, age- and department- adjusted $h_{(15, 2, 1)}$ indexes. Formally, this is the economists for whom the ratio of the actual to model-predicted index is largest. It is a mix of people at top schools with extremely high indexes and people with very good records who might be regarded as underplaced. Consistent with the earlier results on age-related heteroskedasticity, there are larger outliers in the youngest and oldest cohorts than in the middle cohorts.

VII. Conclusion

In this paper, I have examined the degree to which Hirsch-like citation indexes are aligned with market outcomes for economists. The original Hirsch index is poorly suited to the economics profession. But variants of the Hirsch index that focus on more highly cited papers do appear to be an advance on traditional citation counting techniques.

Models that adjust the Hirsch index in this direction and correct for differences across fields can do a fairly good job of accounting for labor market outcomes. Measures of productivity are central to many questions one would like to ask in studying academic or scientific research. For example, studies of the extent to which scientific progress is predictable, studies of the effects of grants or other resources on subsequent research success, studies of the effects of incentive schemes, and studies of discrimination. I hope that citation indexes like those discussed here will spur more interesting work in these areas. The adjustment factors presented here obviously apply only to economics, but I hope that similar studies will be carried out by others with expertise in other disciplines.[25]

---

[25] See Ellison (2012) for a study of computer science.

TABLE 6—LISTS OF HIGHLY CITED ECONOMISTS WITH AND WITHOUT CONTROLS

Highest $h_{(15, 2, 1)}$ indexes relative to field and age			Highest $h_{(15, 2, 1)}$ indexes relative to field, age, school		
Name	$h_{(15, 2, 1)}$	Fld and age adj. $h_{(15, 2, 1)}$	Name	$h_{(15, 2, 1)}$	Actual/ predicted
Panel A. Economists with −1977 PhDs					
Robert Barro	9.66	9.59	Joseph Stiglitz	8.05	2.11
Amartya Sen	9.35	9.51	Peter Schmidt	4.55	1.89
Robert Lucas	9.81	9.21	Robert Lucas	9.81	1.85
James Heckman	8.00	9.00	James Heckman	8.00	1.81
Joseph Stiglitz	8.05	8.96	Herve Moulin	4.00	1.70
Panel B. Economists with 1978–1986 PhDs					
Andrei Shleifer	9.00	8.65	Jeffrey Wooldridge	4.83	1.88
Paul Milgrom	7.83	8.46	Tim Bollerslev	7.22	1.82
John Campbell	8.10	7.32	Paul Milgrom	7.83	1.82
Kenneth Rogoff	7.51	7.20	James N. Brown	4.38	1.62
Bengt Holmstrom	7.33	7.19	Jennifer Reinganum	4.70	1.57
Panel C. Economists with 1987–1995 PhDs					
Daron Acemoglu	7.08	8.29	Ross Levine	7.21	1.78
Alan Krueger	7.50	7.39	David Hummels	4.49	1.65
Edward Glaeser	6.28	7.28	Joon Park	3.66	1.64
Ross Levine	7.21	6.83	David A. Hennessy	3.00	1.62
Guido Imbens	5.08	6.23	Alan Krueger	7.50	1.57
Panel D. Economists with 1996–2004 PhDs					
Ricardo Reis	3.63	7.68	Sandra Black	4.20	1.85
Esther Duflo	5.03	7.61	Ricardo Reis	3.63	1.81
Enrico Moretti	4.50	7.10	Steven Puller	2.34	1.80
Benjamin Olken	3.00	6.99	Andrew J. Patton	3.41	1.69
Emmanuel Saez	4.30	6.61	Brian Jacob	3.44	1.67
Panel E. Economists with 2005− PhDs					
Flavio Cunha	2.42	7.68	Jonathan Meer	1.13	2.10
Hunt Allcott	1.67	7.50	Antonio Galvao	1.13	2.04
Parag Pathak	2.08	7.26	Flavio Cunha	2.42	1.97
Cynthia Kinnan	1.34	7.06	Ryan Michaels	1.17	1.95
Kalina Manova	2.54	7.05	Hunt Allcott	1.67	1.94

Given that much of the demand for citation analyses comes from market participants I would like to be careful to mention a couple limitations. First, it should be kept in mind that the association between citation indexes and employment indexes identified here is not necessarily causal—economists may be highly cited in part because they are at top departments. Second, the estimates presented here cannot be regarded as indicating how researchers in different fields *should* be treated. The field and experience adjustments will reflect both differences in citations holding "quality" fixed and differences in the "quality" thresholds that schools apply when hiring and promoting researchers in different fields and at different levels of experience. The field adjustments are informative about what the economics labor market is doing, not what one should do if one wanted to maximize some objective or treat researchers equitably.

Although I believe that Hirsch-like indexes are useful and have tried to make their adoption as easy as possible, I think that the search for better indexes should remain an active area. The indexes used here are convenient in that they only use data that can be collected from Google Scholar in a matter of seconds, but by doing so they are ignoring a great deal of potentially valuable information. Just as Google's search

engine revolutionized web search by paying attention to which websites were citing which other websites, I imagine that citation indexes that pay attention to where papers are being cited could be much more powerful.[26] There are many potentially informative characteristics of each citation: the citing journal, the citing author, citations to the citing paper, and perhaps some day text-derived estimates of how central the cited paper is to the citing paper. The potential seems vast and this could be an exciting area for many years to come.

Appendix

This Appendix presents more detail on the age and field effects discussed earlier and presents estimates for two other Hirsch-like indexes some researchers might choose to apply. The $h_{(40, 1, 0)}$ index gives the highest likelihood among indexes giving full credit for coauthored papers. The $h_{(30, 1, 1)}$ index gives the highest likelihood among indexes that give $1/n$ credit on jointly authored papers and have $b = 1$.

Field effects for these two additional indexes are reported in Table A1. One would multiply the index of an economist in field j by $e^{-\hat{\gamma}_j}$ to make it comparable to that of a macroeconomist of the same age. The correction factors for the $h_{(40, 1, 0)}$ and $h_{(30, 1, 1)}$ indexes are often much larger in magnitude than those for the $h_{(15, 2, 1)}$ and $h_{(10, 3, 0.25)}$ indexes presented in Table 4. This reflects that an index with $b = 1$ increases more rapidly when citations increase than does an index with $b = 2$ or $b = 3$. I have not reported standard errors, which are mostly between 0.07 and 0.10.

Experience effects for the four Hirsch-like indexes are reported in Table A2. For example, to compare an economist who is 12 years post PhD with one who is 23 years post PhD, one would multiply the $h_{(15, 2, 1)}$ index of the younger economist by $e^{(1 \times 0.099) + (3 \times 0.034) + (4 \times 0.028) + (3 \times 0.031)}$. (Such comparison, of course, become more dubious as the difference in ages becomes larger.) Again, a comparison of the magnitudes across columns reflects that a linear citation index increases more rapidly and a cubic index increases more slowly. To save space I only report standard errors for the estimates from the model using $h_{(15, 2, 1)}$. The standard errors for the second model are mostly 0.01. The standard errors for the third model are 0.05 for the earliest period, 0.04 for the next four, and 0.02 or 0.01 for the others. The standard errors for the fourth model are similar to those for the third.

To make applying the $h_{(15, 2, 1)}$ and $h_{(10, 3, 0.25)}$ indexes as easy as possible, Table A3 reports the factors that an economist's index would be multiplied by in order to make his or her index comparable to that of someone 23 years post-PhD. These factors can be computed from the numbers in Table A2, but the more explicit table makes it easier. The archived data materials for this paper include a spreadsheet that automates both field and age corrections for these indexes.

---

[26] See Liebowitz and Palmer (1984); Palacios-Huerta and Volij (2004); and West, Bergstrom, and Bergstrom (2010) for discussions of ranking journals by such approaches.

TABLE A1—ESTIMATED FIELD DUMMIES FROM NLLS MODELS
USING TWO $h_{(a, b, c)}$ INDEXES

Field	Estimated field effects	
	$h_{(40, 1, 0)}$	$h_{(30, 1, 1)}$
Behavioral/experimental	0.14	0.16
Development	−0.23	−0.15
Finance	0.04	0.06
History	−0.72	−0.54
Industrial organization	−0.30	−0.26
International finance	−0.12	−0.12
International trade	0.10	0.14
Labor	−0.14	−0.07
Metrics-cross section	−0.30	−0.25
Metrics-time series	0.09	0.12
Micro theory	−0.41	−0.34
Public finance	−0.29	−0.21
Political economy	−0.16	−0.09
Other	−0.05	−0.01

TABLE A2—ESTIMATED ACADEMIC AGE EFFECTS FROM NLLS MODELS
USING SEVERAL $h_{(a, b, c)}$

	Estimated time effects from NLLS Models Using $h_{(a, b, c)}$				
	$h_{(15, 2, 1)}$		$h_{(10, 3, 0.25)}$	$h_{(40, 1, 1)}$	$h_{(30, 1, 1)}$
Field	Coef.	SE	Coef.	Coef.	Coef.
Years 1–4	0.149	0.02	0.096	0.237	0.254
Years 4–7	0.144	0.02	0.103	0.181	0.227
Years 7–10	0.073	0.02	0.062	0.124	0.098
Years 10–13	0.099	0.02	0.071	0.170	0.146
Years 13–16	0.034	0.02	0.026	0.048	0.068
Years 16–20	0.028	0.01	0.018	0.048	0.041
Years 20–25	0.031	0.01	0.020	0.045	0.047
Years 25–30	0.008	0.01	0.010	0.024	0.023
Years 30–35	0.015	0.01	0.008	0.014	0.016
Years 35–40	−0.040	0.01	−0.028	−0.068	−0.052
Years 40+	0.015	0.01	0.008	0.016	0.022

TABLE A3—MULTIPLICATIVE EXPERIENCE ADJUSTMENTS FOR $h_{(15, 2, 1)}$ AND $h_{(10, 3, 0.25)}$

	Factor for $h_{(a, b, c)}$			Factor for $h_{(a, b, c)}$			Factor for $h_{(a, b, c)}$	
Year	(15,2,1)	(10,3,0.25)	Year	(15,2,1)	(10,3,0.25)	Year	(15,2,1)	(10,3,0.25)
1	5.49	3.35	16	1.23	1.14	31	0.89	0.91
2	4.73	3.04	17	1.19	1.12	32	0.88	0.90
3	4.07	2.76	18	1.16	1.10	33	0.86	0.89
4	3.51	2.51	19	1.13	1.08	34	0.85	0.89
5	3.03	2.26	20	1.10	1.06	35	0.84	0.88
6	2.63	2.04	21	1.06	1.04	36	0.87	0.90
7	2.27	1.84	22	1.03	1.02	37	0.91	0.93
8	2.11	1.73	23	1.00	1.00	38	0.94	0.96
9	1.97	1.62	24	0.97	0.98	39	0.98	0.98
10	1.83	1.53	25	0.94	0.96	40	1.02	1.01
11	1.66	1.42	26	0.93	0.95	41	1.01	1.00
12	1.50	1.33	27	0.92	0.94	42	0.99	1.00
13	1.36	1.23	28	0.92	0.93	43	0.98	0.99
14	1.31	1.20	29	0.91	0.92	44	0.96	0.98
15	1.27	1.17	30	0.90	0.91	45	0.95	0.97

REFERENCES

Bornmann, Lutz, Gerlind Wallon, and Anna Ledin. 2008. "Is the *h* Index Related to (Standard) Bibliometric Measures and to the Assessments by Peers? An Investigation of the *h* Index by Using Molecular Life Sciences Data." *Research Evaluation* 17 (2): 149–56.

Egghe, Leo. 2008. "Mathematical Theory of the *h*- and *g*-Index in Case of Fractional Counting of Authorship." *Journal of the American Society of Information Science and Technology* 59 (10): 1608–16.

Ellison, Glenn. 2002a. "Evolving Standards for Academic Publishing: A *q-r* Theory." *Journal of Political Economy* 110 (5): 994–1034.

Ellison, Glenn. 2002b. "The Slowdown of the Economics Publishing Process." *Journal of Political Economy* 110 (5): 947–93.

Ellison, Glenn. 2011. "Is Peer Review in Decline?" *Economic Inquiry* 49 (3): 635–57.

Ellison, Glenn. 2012. "Assessing Computer Scientists Using Citation Data." http://economics.mit.edu/files/7784.

Ellison, Glenn. 2013. "How Does the Market Use Citation Data? The Hirsch Index in Economics: Dataset." *American Economic Journal: Applied Economics*. http://dx.doi.org/10.1257.app.5.3.63.

Hamermesh, Daniel S., and Gerard A. Pfann. 2012. "Reputation and Earnings: The Roles of Quality and Quantity in Academe." *Economic Inquiry* 50 (1): 1–16.

Hirsch, J. E. 2005. "An Index to Quantify an Individual's Scientific Research Output." *Proceedings of the National Academy of Sciences* 102 (46): 16569–72.

Hirsch, J. E. 2007. "Does the *h* Index Have Predictive Power?" *Proceedings of the National Academy of Sciences* 104 (49): 19193–98.

Iglesias, Juan E., and Carlos Pecharromán. 2007. "Scaling the *h*-Index for Different Scientific ISI Fields." *Scientometrics* 73 (3): 303–20.

Jensen, Pablo, Jean-Baptiste Rouquier, and Yves Croissant. 2009. "Testing Bibliometric Indicators by Their Prediction of Scientists Promotions." *Scientometrics* 78 (3): 467–79.

Kosmulski, Marek. 2006. "A New Hirsch-Type Index Saves Time and Works Equally Well as the Original h-Index." *ISSI Newsletter* 2 (3): 4–6.

Lehmann, Sune, Andrew D. Jackson, and Benny E. Lautrup. 2006. "Measures for Measures." *Nature* 444 (1): 1003–04.

Liebowitz, S. J., and J. P. Palmer. 1984. "Assessing the Relative Impacts of Economics Journals." *Journal of Economic Literature* 22 (1): 77–88.

Palacios-Huerta, Iganacio, and Oscar Volij. 2004. "The Measurement of Intellectual Influence." *Econometrica* 72 (3): 963–77.

Prathap, Gangan. 2010. "Is There a Place for a Mock *h*-Index?" *Scientometrics* 84 (1): 153–65.

Price, Derek J. de Solla. 1965. "Networks of Scientific Papers." *Science* 149 (3683): 510–15.

Ruane, Frances, and Richard S. J. Tol. 2008. "Rational (Successive) *h*-Indices: An Application to Economics in the Republic of Ireland." *Scientometrics* 75 (2): 395–405.

Schreiber, Michael. 2008. "A Modification of the *h*-Index: The *hm*-Index Accounts for Multi-Authored Manuscripts." *Journal of Informetrics* 2 (3): 211–16.

Tol, Richard S. J. 2009. "The *h*-Index and Its Alternatives: An Application to the 100 Most Prolific Economists." *Scientometrics* 80 (2): 317–24.

van Raan, Anthony F. J. 2006. "Comparison of the Hirsch-index with Standard Bibliometric Indicators and with Peer Judgment for 147 Chemistry Research Groups." *Scientometrics* 67 (3): 491–502.

West, Jevin D., Theodore C. Bergstrom, and Carl T. Bergstrom. 2010. "The Eigenfactor Metrics: A Network Approach to Assessing Scholarly Journals." *College & Research Libraries* 71 (3): 236–44.

Wu, Qiang. 2010. "The w-Index: A Measure to Assess Scientific Impact by Focusing on Widely Cited Papers." *Journal of the American Society for Information Science and Technology* 61 (3): 609–14.

This article has been cited by:

1. Motty Perry, Philip J. Reny. 2016. How To Count Citations If You Must. *American Economic Review* **106**:9, 2722-2741. [Abstract] [View PDF article] [PDF with links]
2. Tyler Cowen, Alex Tabarrok. 2016. A Skeptical View of the National Science Foundation's Role in Economic Research. *Journal of Economic Perspectives* **30**:3, 235-248. [Abstract] [View PDF article] [PDF with links]
3. Les Oxley. 2016. Elites and Secret Handshakes Versus Metrics and Rule-Based Acclamation: A Comment on "Measuring the Unmeasurable". *Econometric Reviews* **35**:1, 44-49. [CrossRef]
4. JOHN GIBSON, DAVID L. ANDERSON, JOHN TRESSLER. 2014. WHICH JOURNAL RANKINGS BEST EXPLAIN ACADEMIC SALARIES? EVIDENCE FROM THE UNIVERSITY OF CALIFORNIA. *Economic Inquiry* **52**:4, 1322-1340. [CrossRef]