Scientometrics, Vol. 8. Nos 5–6 (1985) 315–320

THE MONEY VALUE OF CITATIONS TO SINGLE-AUTHORED AND MULTIPLE-AUTHORED ARTICLES

A. M. DIAMOND, Jr.

Department of Economics,
The Ohio State University,
Hagerty Hall, 1775 College Road,
Columbus, Ohio 43210-1309 (USA)

(Received September 17, 1984)

This note presents evidence for the surprising conclusion that a citation to a multiple-authored article is worth more to its author than a citation to a single-authored article.

Nudelman and *Landers*[1] have presented evidence that the total credit given to all authors of a jointly authored paper is greater than the credit given to the author of a singly authored article. For the case of a three author article they found that the first author received 75% of the credit of a singly authored article while the second author received 62% and the third 58%. The weight to be given to multiple-authored articles has also been much discussed elsewhere in the literature.[2]

Longitudinal data for the Berkeley Mathematics Department on salaries and citations by authorship category can provide important new evidence on the weight employers attach to various categories. 'Longitudinal' means that up to 15 years of data are available for each mathematician in the sample. The University of California at Berkeley was chosen from among the universities with highly ranked departments because it, as a state supported school, is required by law to make faculty salary data publicly available. The salary data was obtained from the Office of Academic Personnel after considerable delay and resistance from the Mathematics Department and various officials.

The basic sample was obtained from faculty listings in a catalog from the late 1970's.[3] Since these listings underrepresented those who were nearing the end of their careers in the early years of the *Science Citation Index*[4] (i.e., the 1960's) the sample was augmented by the addition of all those full and emeriti professors listed in a catalog from the middle 1960's[5] who were not listed in the catalog from the late 1970's. From these samples, any scientist was dropped for whom biographical

Elsevier, Amsterdam–Oxford–New York
Akadémiai Kiadó, Budapest

information was not available from any of the editions of *Cattell's American Men and Women of Science.*[6] Occasionally a scientist was also omitted from the sample if his name was identical to another scientist's as listed in the *Science Citation Index* since it would have been too costly to distinguish citations to his work from those to the work of the other scientist with the same name.

Volumes of the *Science Citation Index* have been published annually since 1961, but because the coverage of math journals in the first four years was very limited, only the years 1965–1979 were used to obtain mathematics citation counts. A well-known defect of the *Science Citation Index* is that it only lists citations under the first author of a multiple-authored article. A total citation count that distinguishes between various categories of multiple authorship is thus much more costly than a first-author citation count because the researcher must first find, using some source other than the *Science Citation Index,* an authoritative list of all of the scientist's multiple-authored publications and then the researcher must separately look up in the *Science Citation Index* each non-first-authored article under the first author's name.

All previous studies of the money value of a citation have used first-author citation counts.[7] To learn the value of citations to multiple authored articles and to test for any bias introduced by the omission of non-first-authored articles, citation counts for the Berkeley mathematics department were constructed that included citations to co-authored articles of which the mathematician was not the first author. Since multiple authorship in mathematics is considerably less common than in the physical sciences, total citation counts for mathematics are less costly to obtain, but also perhaps less informative.

Biographical data, salaries and total number of first-author citations per year were also collected for physics and economics at Berkeley and mathematics, physics and chemistry at the University of Illinois at Urbana. Those departments are not discussed here,[8] however, because the citation data was not collected according to multiple authorship category.

The effect of citations on earnings was estimated using ordinary-least-squares regressions with the natural log of salary as the dependent variable and various independent variables that have traditionally been thought to influence earnings. Descriptive statistics for the Berkeley mathematics sample used in the salary regressions are reported in Table 1. The natural log of salary is used as the dependent variable following the usual practice of economists which they justify on both theoretical and empirical grounds. Years since receipt of Ph.D. measures experience and is expected to be positively related to salary. The period dummy variables are intended to control for general changes over time in citation practices or in the salaries of mathematicians. A cohort is defined as a group of mathematicians who all received their Ph.D.'s within the same period of time. The cohort dummy variables are intended to control for

Table 1
Descriptive statistics on Berkeley mathematics sample*

Variable	Mean	Standard deviation
Citations to single authored articles	5.0	8.7
Citations to articles where first of more than one author	1.4	3.2
Citations to articles where second, third or fourth of more than one author	1.6	3.8
Year of birth	1927.0	10.2
Year of Ph.D.	1952.9	10.6
Salary in 1967 $	18235.9	4706.3

*An observation represents data on a given mathematician in a given year. So if 10 years of data are available for a mathematician, then that mathematician will account for 10 observations in the sample.

differences in the quality and citation practices of mathematicians who received their Ph.D.'s under different conditions.

Heckman and *Robb*[9] have shown that even for longitudinal data a regression that incorporates age, period and cohort effects is underidentified. They suggest reprlacing either period or cohort variables with more sharply focused behavioral variables. Unfortunately, for our data set no good behavioral variables exist that pick up the kinds of effects intended by the period and cohort variables. To proceed with estimation of the effect of age either the cohort effect must be normalized to zero (following *Johnson* and *Stafford*[10]), or the period effect must be normalized to zero (following *Weiss* and *Lillard*[11]). We estimated the basic regression using both normalizations and found that the qualitative results were robust.

Although some studies have shown unobserved person effects to be statistically significant in determining earnings, we do not control for them here. The justification is that, given our data, the tractability of a fixed effects model is greatly reduced due to differences in the number of observations for each scientist. The lack of control for person effects will not bias the estimates of the observed variables if the standard assumption is true that the observed variables are uncorrelated with the unobserved person effects.[12]

Table 2
Multiple authorship log-salary regression
for Berkeley mathematics sample*

Variable	Regression #	
	1	2
Years since Ph.D.	0.041 (15.678)	0.018 (5.976)
Years since Ph.D. squared	−0.0005 (−9.126)	−0.0004 (−6.305)
Citations to single authored articles	0.0019 (2.259)	0.0056 (7.324)
Citations to articles where first of more than one author	0.0084 (3.670)	0.0107 (5.350)
Citations to articles where second, third or fourth of more than one author	0.0023 (1.184)	0.0081 (4.719)
Period 1968–1971	0.120 (6.758)	–
Period 1972–1975	0.063 (3.651)	–
Cohort 1941–1950	–	−0.239 (−7.861)
Cohort 1951–1960	–	−0.385 (−10.282)
Cohort 1961–1970	–	−0.683 (−15.088)
Constant	9.166 (319.985)	9.962 (195.490)
Number of observations	564	564
Number of mathematicians	45	45
R^2	0.60	0.70

*t-statistics are reported in parentheses. The dependent variable was the natural log of salary. The omitted period is 1965–1967 and the omitted cohort consists of those who received their Ph.D.'s before 1941.

An initially attractive specification for the regression would include separate independent variables for citations to articles where the mathematician was: the single author, first of two, second of two, first of three and so on. Unfortunately, the paucity of multiple-authored articles in mathematics did not permit the precise estimation of a regression with such refined distinctions. Instead, citations were

Table 3
The marginal dollar value of citations to single-authored and multiple-authored articles

Marginal dollar value of:	Regression # used to calculate values	
	1	2
A citation to a single-authored article	92.30	272.05
A citation to an article where the author is the first of more than one author	408.07	519.81
A citation to an article where the author is the second, third or fourth of more than one author	111.73	393.50

divided into three categories: citations per year to articles of which the mathematician was the single author; citations per year to any multiple-authored article of which the mathematician was the first author; and citations per year to any multiple-authored article of which the mathematician was a second, third or fourth author. The coefficients estimated in this regression are reported in Table 2. Citations to multiple-authored papers are worth more than citations to singly authored work no matter what the order of the author's name.

To present the results of Table 2 in more concrete form, the marginal value of a citation of various types was calculated from both regression 1 and regression 2. The 'marginal' value simply means the value of an additional citation. The results are reported in Table 3. The surprising conclusion to be drawn from Table 3 is that a citation to a single-authored article is worth less to its author than a citation to a multiple-authored article. The difference is greater when the author was the first among the multiple authors but holds even when the author was a 'secondary' author.

Before much effort is spent trying to explain this finding, we would do well to encourage the creation of data sets that would permit tests of the robustness of the finding for other disciplines (as well as for different samples of mathematicians). If the finding turns out to be robust one explanation worth considering would be that citations to multiple-authored articles are a proxy for the trait of collegiality that is rewarded by departments in the determination of salaries.

*

The research was in part conducted while the author was a postdoctoral fellow supported by the Sloan Foundation. The Walgreen Foundation provided funds for the collection and processing of the data Able research assistance was provided by *Gregory Armotrading, Patricia Campana, Blair Gifford, Bolivar Ramos, James Rasulo, Myriam Rasulo, John Robb,* and *Kwang-Shik Shin.*

References

1. A. E. NUDELMAN, C. E. LANDERS, The failure of 100 divided by 3 to equal 33–1/3, *The American Sociologist,* 7 (Nov. 1972) 9.
2. See, e. g.: H. A. ZUCKERMAN, Patterns of name ordering among authors of scientific papers: A study of social symbolism and its ambiguity, *The American Journal of Sociology*, 74 (Nov. 1968) 276.; E. GARFIELD, Where the action is, was, and will be-for first *and* secondary authors, as reprinted in *Essays of an Information Scientist,* ISI Press, Philadelphia, Vol. 1, 1977, p. 281, M. D. GORDON, A critical reassessment of inferred relations between multiple authorship, scientific collaboration, the production of papers and their acceptance for publication, *Scientometrics,* 3 (1980) 193. S. PRESSER, Collaboration and quality of research, *Social Studies of Science,* 10 (1980) 95, D. LINDSEY, Production and citation measures in the sociology of science: The problem of multiple authorship, *Social Studies of Science,* 10 (1980) 145; J. S. LONG, R. McGINNIS, P. D. ALLISON, The problem of junior-authored papers in constructing citation counts, *Social Studies of Science,* 10 (1980) 127; J. S. LONG, R. McGINNIS, On adjusting productivity measures for multiple authorship, *Scientometrics,* 5 (1982) 379; D. LINDSEY, Futher evidence for adjusting for multiple authorship, *Scientometrics,* 5 (1982) 389; R. RUSTUM, N. R. ROY, G. G. JOHNSON, Jr., Approximating total citation counts from first author counts and from total papers, *Scientometrics,* 2 (1983) 117.
3. *University of California, Berkeley 1979/80 General Catalog,* Vol. 73, No. 11 (August 1979).
4. E. GARFIEDL (chairman) *Science Citation Index 1979,* Institute for Scientific Information, Inc., Philadelphia; 1980.
5. *University of California, Berkeley General Catalogue* 1966–1967, Vol. 60, No. 15 (July 15, 1966).
6. J. CATELL PRESS (Ed)., *American Men and Women of Science,* 14th ed., R. R. Bowker Co., New York; 1979.
7. For references, see: A. M. DIAMOND, Jr., What is a citation orth? Xerox draft, The Ohio State University, 1984.
8. For an analysis of this data, see the reference in Ref.[7] and also: A. M. DIAMOND, Jr., New evidence on the life-cycle productivity of scientists, Xerox draft, The Ohio State University, 1984.
9. J. HECKMAN, R. ROBB, *Using Longitudinal Data to Estimate Age, Period and Cohort Effects in Earnings Equations,* Economics Research Center/NORC Discussion Paper Series, 83–9, 1983.
10. G. JOHNSON, F. STAFFORD, Lifetime earnings in a professional labor market: Academic economists, *Journal of Political Economy*, 82 (1974) 549.
11. Y. WEISS, L. LILLARD, Experience, vintage and time effects in the growth of earnings: American scientists, 1960–1970, *Journal of Political Economy,* 86 (1978) 427.
12. See: S. ROSEN, P. TAUBMAN, Changes in life-cycle earnings: What do social security data show? *The Journal of Human Resources,* 17 (Summer 1982) 329.