DISTRIBUTION OF SCIENTIFIC PRODUCTIVITY:
AMBIGUITIES IN THE ASSIGNMENT OF AUTHOR RANK‡

NEVENKA PRAVDIC, VESNA OLUIC-VUKOVIC

Institute for Information Sciences (formerly: Referral Centre of the University of Zagreb),
P.O.Box 327, 41001 Zagreb, Croatia (Yugoslavia)

(Received December 15, 1989)

Methodological implications of four accounting procedures applied in multiple authorship
treatment relating to author productivity distribution were investigated. The emphasis was
given to the individual author rank and inequality pattern of data. It was found that similar
pattern of inequality holds in three of the four analysed cases, in spite of the fact that
significant changes were observed on the individual level. By introducing the concept of dual
approach a plausible interpretation of that phenomenon was obtained.

Introduction

Although fundamental regularities of the scientific productivity distribution were
ingeniously and clearly put down by Yablonsky almost ten years ago, there are still
many open questions that stimulate further investigation on theoretical and empirical
grounds. One of these questions, concerning the distribution of publications over the
set of their authors, i.e. author productivity distribution, is the reliability of
conclusions derived from quantitative studies.

The real motive for such a doubt was found in the comprehensive review by Vlachy. As
early as 1970, Vlachy pointed out that in accomplishing the
requirements for the investigation of productivity distributions, several characteristics
of data, such as the size of the sample, time interval, and multiple authorships, should
be included. In recent contributions to this field, considerable attention was actually
paid to the first two requirements, while the fulfilment of the third one was
considered only fragmentarily, referring to comparison of full and fractional
authorships. This, however, is not surprising since the treatment of multi-authored
papers is not solved in general. The agreement on the unit of publication to be
credited to each author has not been reached, confirming the long-life actuality of the
statement formulated by Price: "for modern data there must be much uncertainty
because we have as yet no adequate model or theory for the attribution of credit in
the case of multi-authored collaborative paper". Furthermore, consensus about the

‡ Dedicated to the memory of Michael J. Moravcsik
possible impact of the accounting procedure on analytical results is missing as well.\textsuperscript{5-8} It might be supposed that the problem in quantitative analysis emerges from the limitations of the approaches applied in author productivity distribution. Irrespective of whether the frequency-size or rank-frequency distribution is applied, the one-dimensional description of the relation author - paper is required, although in fact this relation is by no means straightforward, but mutually determined. Each author can be associated with a set of papers, and \textit{vice versa}, each paper can be associated with a set of its authors. From quantitative point of view these two relations are not the same since each of them employs different units of analysis. When the authors are units of analysis and papers which they published alone or in collaboration are the dependent variable, the methodology chosen for attribution of credit becomes the stipulating factor, giving rise to significant differences in the value of author productivity indicator.

Multiple-authorship doubtfulness on the one side and one-dimensionality of the distribution approach on the other, constitute major shortcomings of the methods applied in the author productivity distribution studies. The aim of the present empirical examination was therefore:

- to evaluate the possible implications of different procedures for attribution of credit in multi-authored papers on the assignment of the author rank, and
- to apply the concept of dual approach\textsuperscript{9} as a methodological improvement in author productivity distribution analysis.

In this paper we will report and comment on: (a) the application of four different methodologies to multi-authored papers; (b) the exploration of differences in the rank distribution of authors performed on two levels: (1) for the author corpus on the whole, and (2) for individuals; (c) quantitative evaluation of the impact of different accounting procedures; and (d) the application of dual approach.

Four accounting procedures

The author productivity indicator is influenced not only by the number of papers to which an author contributed, but also by the number of coauthors. To quantify an author's credit in a multi-authored paper a number of weighting procedures of different objectivity are available. In the course of this study, the following were applied:

(A) \textit{Normal Count} procedure meaning that equal credit is given to all contributors; one full unit is assigned for each appearance of the author's name in the by-line regardless of the number of coauthors present. This procedure leads to the so