



Co-authorship and the Measurement of Individual Productivity

by

Karol Flores-Szwagrzak and Rafael Treibich

Discussion Papers on Business and Economics
No. 17/2015

FURTHER INFORMATION
Department of Business and Economics
Faculty of Business and Social Sciences
University of Southern Denmark
Campusvej 55, DK-5230 Odense M
Denmark

E-mail: lho@sam.sdu.dk / <http://www.sdu.dk/ivoe>

Co-authorship and the Measurement of Individual Productivity

Karol Flores-Szwagrzak

Rafael Treibich

ABSTRACT. Consider a database of academic papers where each paper has a scientific worth and a group of authors. We propose a new way of measuring individual academic productivity by evaluating *authorship*, the extent of an author's contribution to each paper. Our method, *CoScore*, uses the varying levels of success of all academic partnerships to infer, simultaneously, overall individual productivity and authorship: the worth of a paper is distributed proportionally to each co-author's productivity, defined as the sum of her contributions to all papers. The *CoScores* of all authors are determined endogenously via the solution of a fixed point problem. We show that *CoScore* is well-defined and that it is uniquely characterized by three properties: *consistency*, *invariance to merging papers*, and *invariance to merging scholars*. We illustrate *CoScore* for the two thousand most cited papers in economics.

Keywords: Authorship, Co-authorship network, Ranking methods, PageRank.

JEL classification: D70, D71, D89.

1. INTRODUCTION

In economics, as well as in other fields, authorship is listed alphabetically. A researcher's contribution to a publication is thus not explicit; it can neither be observed nor verified objectively. Not surprisingly, the existing measures of individual academic productivity either ignore the collaborative nature of research, assigning full authorship to every co-author or attempt to correct for it ad hoc, for example, by dividing a paper's citations equally among its co-authors. We argue that individual *authorship*, the extent of an individual's contribution to collaborative papers, can be approximated by systematically observing the varying levels of success of all academic partnerships in a field. We propose a new measure of individual productivity, the co-author score or *CoScore*, reflecting this inferred authorship.

Date: December 21, 2015.

karolszw@sam.sdu.dk, University of Southern Denmark.

rtr@sam.sdu.dk, University of Southern Denmark.

The information we rely on is freely available in bibliographic databases that can be accessed through the Internet. It consists of a collection of papers specifying the authors as well as the scientific worth or value of each paper. Depending on the application, different indices may be used to evaluate a paper’s scientific value: its number of citations, its number of *American Economic Review*-equivalent papers (Kalaitzidakis et al., 2003; Conley and Önder, 2014), the impact factor of the journal in which the paper was published, etc. The citation analysis literature is extensive and growing, providing various well-founded metrics for the worth of articles.¹ Thus, we take the scientific value of articles as exogenous parameters and study the methods assigning individual credit for co-authored papers. To the best of our knowledge, no systematic analysis of such methods exists.

To evaluate the contribution or authorship of the researchers involved in a given paper, we account for the track record of *all* researchers in the field, not just those co-authoring the paper. The main idea is that stronger authors usually contribute more than their weaker co-authors and should therefore be given, *as a first approximation*, a bigger credit for their joint papers. Which researcher is relatively stronger in turn depends on the authorship and worth of the papers that researcher has contributed to. In other words, the way authorship of a paper should be assigned is inherently connected to the overall strength of its authors, which is itself determined by the distribution of authorship on possibly all papers in the database.

Our measure, *CoScore*, naturally captures the relationship between a researcher’s contribution to a paper and her individual productivity, as quantified by her *score*, which is determined endogenously. The worth of each paper is distributed proportionally to each of its co-authors’ scores, where the score of an author is defined as the sum of her contributions to all of her papers. Crucially, the scores of all authors are determined endogenously and simultaneously as the solution of a fixed point problem. *CoScore* is well defined since the fixed point always exists and is unique. The endogeneity behind *CoScore* is also an essential feature of Google’s PageRank algorithm (Page et al., 1998), the invariant (Pinski and Narin, 1976; Palacios-Huerta and Volij, 2004) and handicap (Demange, 2014) journal ranking methods, and various measures of network centrality (Jackson, 2008).

By exploiting all of the information in the database, *CoScore* aims at improving the assessment of individual contributions in co-authored papers, and thereby the measurement of *individual* productivity. *CoScore* can be used to rank scholars, either as an alternative or as a complement to the existing h (Hirsch, 2005), step-based (Chambers and Miller, 2014), or Euclidean (Perry and Reny, 2015) indices. In contrast to these rankings, *CoScore* reflects the whole co-authoring network and the complete records of all scholars, not just the publication record of the author being ranked.

¹For a survey, see Palacios-Huerta and Volij (2014).

We illustrate *CoScore* for the 1888 most cited papers in economics. Our results show that it differs substantially from the “egalitarian score” obtained by assigning all co-authors an equal number of citations for each paper: *CoScore* concentrates authorship among those it identifies as stronger. Typically, these are the authors who consistently publish highly cited works either individually or in partnership with multiple groups of co-authors. This is striking for Andrei Shleifer, who has 33 papers with a total of 17 different co-authors in the database. Many of these co-authors only contribute to papers where he is also a co-author. As a result, Andrei Shleifer goes from being ranked eighth according to the egalitarian score to being ranked first with *CoScore*. In contrast, authors who tend to write papers on their own, such as Robert E. Lucas, do not experience a significant change.

Finally, we provide axiomatic foundations for *CoScore*, showing that the associated method to allocate credit for papers is uniquely characterized by three properties. The first property, *consistency*, requires the distribution of authorship to remain invariant when an author is taken away from the problem, and the value of every paper she has contributed to is reduced by the amount she was previously allocated. Similar consistency properties have been used in the literature on the measurement of intellectual influence (Palacios-Huerta and Volij, 2004) and, more extensively, in resource allocation and cooperative game theory (Thomson, 2011). The second property, *invariance to merging papers*, requires that individuals do not benefit or suffer from merging papers with the same co-authors. It can be justified both on informational and strategic grounds. The third property, *invariance to merging scholars*, requires that two authors who contribute to the same joint papers do not benefit or suffer from merging into one single author. The three axioms are uniquely satisfied by our co-author method. Furthermore, they are logically independent.

Although we focus here on academic authorship, analogous problems arise in other settings: What is the contribution of a football player to her team? What is the contribution of a manager to her firm’s profits? What is the contribution of an actor to the box-office revenue of a movie? The common thread in these questions is that groups of individuals produce observable output jointly but their individual contributions cannot be perfectly and objectively observed. The insights developed in this paper can similarly be applied to these different environments, where they provide new ways of measuring individual productivity.

The remainder of this paper is organized as follows: Section 2 introduces the model and provides examples of its various applications. Section 3 introduces *CoScore* and shows that it is well defined. Section 4 illustrates *CoScore* for the 1888 most cited papers in economics. Section 5 discusses formal properties of authorship measurement and provides an axiomatic rationale for the co-author method and hence for *CoScore*. Section 6 concludes. All of the proofs are included in the appendix.

2. MODEL

A problem or database is described by a collection of papers where each paper is described by a set of authors and its scientific worth (such as the number of citations). Authors are drawn from a countable set of potential agents which we identify with the natural numbers, \mathbb{N} . Let \mathcal{N} denote the collection of finite subsets of \mathbb{N} . Papers are indexed by a countable set denoted by \mathbb{C} . Let \mathcal{C} denote the collection of finite subsets of \mathbb{C} .

For each $N \in \mathcal{N}$, a problem involving N is a triple $P \equiv (C, w, S)$ where

$$C \in \mathcal{C}, w : C \rightarrow \mathbb{R}_+, S : C \rightarrow 2^N.$$

Thus, problem P is described by a finite collection of papers C and, for each paper $p \in C$, a weight $w(p) \in \mathbb{R}_+$ and a collection of authors $S(p) \subseteq N$. For each $i \in N$, let C_i denote the papers in C involving i as an author, let C_{ii} denote the papers involving i as the sole author, and let w_{ii} denote the aggregate solo contribution of individual i , $w_{ii} = \sum_{p \in C_{ii}} w(p)$. For simplicity, we assume throughout that w_{ii} is positive.² For each $N \in \mathcal{N}$, let \mathcal{P}^N denote all problems involving the authors in N .

Our objective is to infer a measure of individual productivity from any database. A score is a systematic procedure which associates to every problem a profile of individual productivity scores. Formally, a **score**, s , is a function such that, for each $P \in \mathcal{P}^N$,

$$s(P) \in \Delta(N) \equiv \left\{ x \in \mathbb{R}_+^N : \sum_N x_i = 1 \right\}.$$

We refer to the i th coordinate of $s(P)$, denoted by $s_i(P)$, as the score of i under $s(P)$ or, simply, as the score of author i when there is no room for confusion. Note that a score is normalized so that the individual scores add up to one.

The model is formulated in the context of academic authorship. However, it may similarly be applied in many other environments where groups of individuals engage in joint production. Examples include:

Measuring the contribution of actors to the success of a movie. A problem may now be interpreted as a collection C of movies where each movie $p \in C$ is described by a cast of actors $S(p)$ and a measure of success $w(p)$ that is both observable and quantifiable (box-office revenue, number of Oscars, IMDb rating, etc.). The group of authors N may be extended to account for other key players in movie production such as movie directors, screenwriters or producers. The measure of success $w(p)$ may also be corrected to control for other determinants to the success of a movie such as genre, year, country, etc.

Measuring the share of the profits of a firm that can be attributed to each of the members of its board of directors. A board of directors is a body of elected members who jointly oversee the

²This mild technical assumption can be justified by assuming that every author has made an independent contribution, however small, such as writing her PhD dissertation.

activities of a company. The board of directors may change several times over a fixed period of time, which means a firm has to be treated separately every time its board is modified. A problem would thus consist of a collection C of firm-period pairs (f, t) where the board of director for firm f is maintained identical over period t . Each element $p = (f, t) \in C$ is defined by the corresponding board of directors $S(p)$ and the profit $w(p)$ generated by the company during period t , typically the tenure of board $S(p)$.

Measuring the contributions of managers to the profitability of their company. Managers or executives within a given company participate in several team projects. The composition of a team is observable and its success can be measured explicitly. The information provided by the whole database of projects can thus be used to infer the relative contribution of each manager in a given team. The problem now consists of a collection of team projects C , each project being defined by a team of managers $S(p)$ and the revenue $w(p)$ generated by the project.

Measuring the value of individual players in team sports. A problem consists of a collection C of games where each game $p \in C$ is characterized by a winning team $W(p) \subset N$ and a losing team $L(p) \subset N$. The worth of a game for the winning team $w(p)$ is defined by the value of the losing team, as measured endogenously by a pre-determined tournament solution (Laslier, 1997), or by some exogenous criteria for the selected period such as the number of victories, rank in championship, etc.

3. THE CO-AUTHOR SCORE

We now introduce the “co-author score” or “*CoScore*” which is central in our analysis. Here, the score of each author is equal to her aggregate contribution where, critically, her contribution to a paper is itself assigned in proportion to her fraction of the scores of all the authors involved in that paper. The score of an author is thus determined endogenously and simultaneously with the scores of all other authors.

Formally, the co-author score or **CoScore**, \check{s} , is the value such that, for each $P = (C, w, S) \in \mathcal{P}^N$,

$$(1) \quad \check{s}_i(P) = \frac{1}{\sum_{p \in C} w(p)} \sum_{p \in C_i} w(p) \frac{\check{s}_i(P)}{\sum_{j \in S(p)} \check{s}_j(P)} \quad \text{for each } i \in N.$$

In contrast, the score commonly used to discount for co-authorship, the **egalitarian score**, \tilde{s} , divides the worth of each paper equally among its co-authors: for each $P = (C, w, S) \in \mathcal{P}^N$,

$$\tilde{s}_i(P) = \frac{1}{\sum_{p \in C} w(p)} \sum_{p \in C_i} \frac{w(p)}{|S(p)|} \quad \text{for each } i \in N.$$

A central result of this paper is that *CoScore* is well defined. That is, the system of equations (1) defining the co-author score yields a unique solution. Existence follows from Brouwer’s fixed point theorem which implies that the operator $\phi : \Delta(N) \rightarrow \Delta(N)$ defined by

$$\phi_i(x) = \frac{1}{\sum_{p \in C} w(p)} \sum_{p \in C_i} w(p) \frac{x_i}{\sum_{j \in S(p)} x_j} \quad \text{for each } x \in \Delta(N) \text{ and each } i \in N.$$

has a fixed point. Moreover, each such fixed point satisfies (1). Uniqueness, follows from the fact that each solution of system (1) maximizes $\sum_{p \in C} w(p) \ln \sum_{j \in S(p)} x_j$ over $x \in \Delta(N)$. Since this amounts to maximizing a strictly concave function over a compact and convex set, the solution is in fact unique. See the appendix for the formal proof.

Theorem 1. *CoScore is well defined.*

CoScore is implicitly defined by a system of equations that, in general, has no closed form solution. However, there is a simple class of problems where we can immediately express *CoScore* in terms of the worth of papers. We say that a problem P is **symmetric** if each co-authored paper in P includes all authors. This implies that, in a symmetric problem, authors only differ in their solo contributions. In the domain of symmetric problems, *CoScore* is such that the credit allocated to each author for a given paper is simply proportional to her solo contribution relative to that of her co-authors. Accordingly, *CoScore* can be expressed as follows.

Proposition 1. *For each symmetric problem $P = (C, w, S) \in \mathcal{P}^N$, the *CoScore* is such that*

$$\check{s}_i(P) = \frac{1}{\sum_{p \in C} w(p)} \sum_{p \in C_i} \frac{w_{ii}}{\sum_{j \in N} w_{jj}} w(p) \quad \text{for each } i \in N.$$

In general, the pattern of co-authorship is much more complicated than the one of symmetric problems and the above formula does not apply. However, the *CoScore* of each author can easily be obtained by computing the fixed point of function ϕ .

4. ILLUSTRATIONS FROM THE ECONOMIC LITERATURE

We illustrate *CoScore* for the one thousandth most cited papers in economics, as listed on the IDEAS website.³ We compute both the *CoScore* and the egalitarian score for the 1737 authors included in the database using citations as the measure of scientific worth of a

³The data was retrieved on Nov. 10th 2015, at <https://ideas.repec.org/top/top.item.nbcites.html>. The citation data used by the IDEAS website comes from the analysis of the Citec project, which is currently used by several RePEc services, including Socionet, EconPapers and IDEAS. The database consists of 1888 papers and provides a consolidated number of citations for each paper, taking into account all of its registered versions.

paper.⁴ The results are reported in Figure 1 and Table 1. The two scores differ significantly both in values and in their associated author rankings.

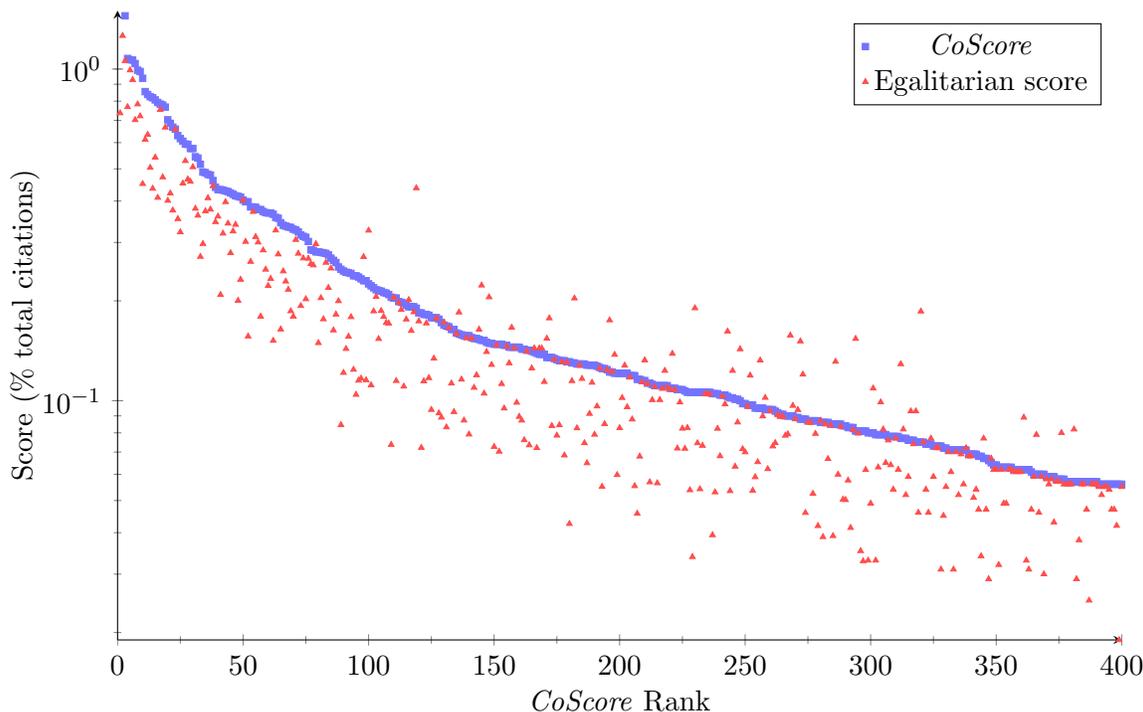


FIGURE 1. Distribution of *CoScore* and the Egalitarian score of the 400 highest *CoScore* ranked economists.

Compared to the egalitarian score, *CoScore* concentrates authorship among the authors it identifies as stronger. Typically, these are the authors who have established their standing by publishing highly cited works either individually or in partnership with multiple groups of co-authors. As a consequence, *CoScore* decreases the ranking of co-authors who may have written highly cited papers but only in collaboration with stronger scholars. While the forty economists with the highest egalitarian score make for roughly 22.2% of the total number of citations, the forty economists with the highest *CoScore* account for more than 31% of that total.

The largest increases when going from the egalitarian score to the *CoScore* typically occur for researchers who co-author with many and less productive researchers. Conversely, the largest decreases typically occur for researchers who co-author with few and highly productive researchers. This effect is striking for Andrei Shleifer, who has 33 papers with a total of 17 different co-authors, many of them only contributing to papers where he is also a co-author. As a result, Andrei Shleifer goes from being ranked 8th to being ranked first, and his score jumps from 0.73% to more than 2% of the total number of citations in the database. On the other hand, authors who tend to write single-authored papers do not experience significant

⁴We choose here to give the equivalent of 10 citations to every author in the database for their PhD dissertation.

Rank	Scholars	# Papers	<i>CoScore</i> (%)	Egalitarian Score (%)
1	Shleifer, Andrei	33	2.218	0.737
2	Barro, Robert	22	1.553	1.260
3	Fama, Eugene F	21	1.447	1.058
4	Heckman, James	16	1.076	0.769
5	Lucas, Robert E	13	1.067	0.992
6	Johansen, Soren	7	1.064	0.928
7	Rogoff, Kenneth	20	1.039	0.702
8	Becker, Gary S	20	0.992	0.783
9	Engle, Robert	14	0.983	0.721
10	Stock, James H	16	0.938	0.451
11	Bollerslev, Tim	13	0.855	0.613
12	Jensen, Michael C	8	0.838	0.634
13	Levine, Ross	11	0.825	0.504
14	Stiglitz, Joseph E	8	0.819	0.436
15	Myers, Stewart C	6	0.807	0.541
16	Kahneman, Daniel	11	0.793	0.410
17	Romer, Paul M	5	0.784	0.754
18	Gertler, Mark	15	0.779	0.472
19	Granger, Clive W J	13	0.768	0.667
20	Prescott, Edward C	8	0.703	0.401
21	Blanchard, Olivier J	13	0.686	0.422
22	Pesaran, M Hashem	8	0.668	0.375
23	Merton, Robert C	8	0.659	0.658
24	Acemoglu, Daron	13	0.630	0.353
25	Alesina, Alberto	17	0.617	0.322
26	Campbell, John	16	0.606	0.452
27	Krugman, Paul	9	0.594	0.529
28	Bernanke, Ben	16	0.593	0.465
29	Perron, Pierre	9	0.576	0.458
30	Bond, Stephen	5	0.576	0.506
31	Reinhart, Carmen	11	0.544	0.380
32	Tirole, Jean	11	0.539	0.362
33	Aghion, Philippe	7	0.516	0.272
34	Milgrom, Paul	12	0.489	0.297
35	Arellano, Manuel	2	0.487	0.373
36	Sims, Christopher A	9	0.481	0.408
37	Hall, Robert E	7	0.479	0.377
38	White, Halbert	4	0.461	0.444
39	Diamond, Douglas W	6	0.440	0.346
40	Hansen, Lars Peter	6	0.433	0.360

TABLE 1. The forty highest *CoScore* ranked economists according to *CoScore*.

changes. For instance, Robert E. Lucas, with only three different co-authors for a total of 13 articles in the database, goes from 0.99% to 1.07%.

The results reported here should be taken with caution since (i) the database only accounts for a very small share of the papers in economics, and (ii) citation numbers have not been

re-scaled to account for across-field differences.⁵ The main purpose of this exercise is simply to illustrate *CoScore* for a real-life database, contrasting it with the egalitarian score.

5. AXIOMATIC FOUNDATIONS

5.1. Methods. Our premise is that the score of an author should reflect her total contribution to all the papers in the database. Accordingly, a problem closely related to assigning scores is that of dividing the authorship of a given paper among its co-authors. A *method* is a systematic procedure which distributes the worth of each paper among its co-authors. The close relationship between methods and scores is a salient feature of our analysis and motivates our discussion of methods here. Formally, a **method**, m , is a function such that, for each $N \in \mathcal{N}$ and each $P = (C, w, S) \in \mathcal{P}^N$,

$$m(P) \in Z(P) \equiv \prod_{p \in C} \left\{ x \in \mathbb{R}_+^{S(p)} : \sum_{i \in S(p)} x_i = w(p) \right\}.$$

For each $p \in C$ and each $i \in S(p)$, $m_i^p(P)$ denotes the authorship attributed to individual i in paper p . Thus, $\sum_{i \in S(p)} m_i^p(P) = w(p)$.

To each method, m , we can associate the score s^m such that, for each $N \in \mathcal{N}$, and each $P = (C, w, S) \in \mathcal{P}^N$,

$$s_i^m(P) = \frac{1}{\sum_{q \in C} w(q)} \sum_{p \in C_i} m_i^p(P) \quad \text{for each } i \in N.$$

In words, the score of an author is the sum of her contributions to each of the papers she co-authors which we normalize dividing by the total worth of the papers in the database.

The **co-author method**, \check{m} , allocates the value of each paper proportionally to the individual worth of every coauthor, as (endogenously) measured by the *CoScore*: for each $N \in \mathcal{N}$, each $P = (C, w, S) \in \mathcal{P}^N$, each $p \in C$, and each $i \in S(p)$,

$$\check{m}_i^p(P) = w(p) \frac{\check{s}_i(P)}{\sum_{j \in S(p)} \check{s}_j(P)}.$$

Note that \check{m} is well defined because \check{s} is by Theorem 1.

Remark 1. *The co-author method can equivalently be defined as follows: for each $P = (C, w, S) \in \mathcal{P}^N$, each $p \in C$, and each $i \in S(p)$,*

⁵Perry and Reny (2015) offer a thorough discussion of why such re-scaling is desirable.

$$(2) \quad \check{m}_i^p(P) = w(p) \frac{\sum_{q \in C_i} \check{m}_i^q(P)}{\sum_{j \in S(p)} \sum_{q \in C_j} \check{m}_j^q(P)}.$$

The co-author method is our main proposal, but other alternative methods may also be considered. For example, the **egalitarian method**, $\tilde{\mathbf{m}}$, allocates the value of each paper equally among all its co-authors: for each $N \in \mathcal{N}$, each $P = (C, w, S) \in \mathcal{P}^N$, each $p \in C$, and each $i \in S(p)$,

$$\tilde{m}_i^p(P) = \frac{w(p)}{|S(p)|}.$$

When assigning the authorship of a paper, the egalitarian method disregards any information about the publication records of its co-authors. Any two co-authors are deemed equally deserving of the paper's authorship.

The **proportional method**, $\hat{\mathbf{m}}$, allocates the value of each paper proportionally to the value of the (aggregate) solo contribution: for each $N \in \mathcal{N}$, each $P = (C, w, S) \in \mathcal{P}^N$, each $p \in C$, and each $i \in S(p)$,

$$\hat{m}_i^p(P) = \frac{w_{ii}}{\sum_{j \in S(p)} w_{jj}} w(p).$$

When assigning the authorship of a paper, the proportional method makes inferences on authorship based only on the worth of the individual contributions of its co-authors. For example, if researchers Alice and Bob have similar individual contributions, they will receive similar shares of their joint work, regardless of the fact that Alice also co-authored several other important papers while Bob didn't.

5.2. Axiomatic Characterization. We provide two axiomatic characterizations of the co-author method on the basis of equity, strategic, and informational simplicity properties. The first property, consistency, requires that, if the assignment of authorship is considered desirable for a group of authors, then it should still be considered desirable if an author is taken away from the problem and the worth of every paper is reduced by the amount she was previously assigned. Formally, let $P = (C, w, S) \in \mathcal{P}^N$, $x \in Z(P)$, and $k \in N$. The problem obtained from P upon the departure of author k following authorship assignment x , denoted by $c_k^x(P)$, is $\tilde{P} = (\tilde{C}, \tilde{w}, \tilde{S}) \in \mathcal{P}^{N \setminus \{k\}}$ such that

- i. the papers in \tilde{P} are the papers in P that are not single-authored by k , i.e., $\tilde{C} = C \setminus C_{kk}$;
- ii. the worth of each paper in \tilde{P} is revised down by the authorship of the paper attributed to k , i.e., for each $p \in \tilde{C}$ such that $p \in C_k$, $\tilde{w}(p) = w(p) - x_k^p$ and, for each $p \in \tilde{C}$ such that $p \notin C_k$, $\tilde{w}(p) = w(p)$;

- iii. each paper in \tilde{P} is co-authored by the same authors as in P , excluding k , i.e., for each $p \in \tilde{C}$, $\tilde{S}(p) = S(p) \setminus \{k\}$.

Consistency: For each $P \in \mathcal{P}^N$, each $k \in N$, and each $i \in N \setminus \{k\}$,

$$m_i(P) = m_i(c_k^{m(P)}(P)).$$

Thus, a method, m , is *consistent* if the assignment of authorship among the researchers in $N \setminus \{k\}$ is the same in problems P and $c_k^{m(P)}(P)$. A related consistency condition is used in the axiomatic characterization of the “invariant” journal ranking of Palacios-Huerta and Volij (2004). This ranking relies on the same ideas at the core of PageRank (Page et al., 1998), the procedure used by Google to rank web pages. Other conceptually related conditions have also been central in the analysis of resource allocation and cooperative game theory. See Thomson (2011) for a survey and Thomson (2012) for a discussion of the ethical content of consistency-type requirements. *Consistency* is satisfied by the egalitarian and the co-author methods but not by the proportional method.

The second property prevents authors from increasing their authorship by either fragmenting or consolidating papers with the same group of authors. The idea here is for the content of papers to be presented in the most natural way possible. For example, no author should be given incentives to split one of her solo works, say an *Econometrica* article, into two *Journal of Economic Theory* articles. More generally, no group of authors, should gain by splitting a single paper jointly authored by all of them into lesser papers with a total value equal to that of the original paper and, again, co-authored by the whole group. In the context of ranking authors based only on their citations list (Hirsch, 2005), Perry and Reny (2015) propose a related property, “depth relevance”, which requires that an author’s rank does not increase upon splitting an article in her publication list into two articles with the same total number of citations. The property can also be motivated on grounds of informational simplicity since joint papers with the same group of authors can now be replaced by a single representative paper of total equal value.

Formally, consider a problem $P = (C, w, S) \in \mathcal{P}^N$ and let $D \subseteq C$ consist of papers with the same set of authors. The problem obtained from P by merging the papers in D into a single paper r , denoted by $c^{D \rightarrow r}(P)$, is $\tilde{P} = (\tilde{C}, \tilde{w}, \tilde{S}) \in \mathcal{P}^N$ such that

- i. the papers in D are all merged into a new paper, r , so that $\tilde{C} = [C \setminus D] \cup \{r\}$;
- ii. the worth of paper r in \tilde{P} is the sum of the worth of the papers in D from problem P while the worth of all other papers is the same as in P , i.e., $\tilde{w}(r) = \sum_{q \in D} w(q)$ and, for each $q \in C \setminus D$, $\tilde{w}(q) = w(q)$;
- iii. the authors of a paper in \tilde{P} that is also in P are the same while the authors of r are the former authors of the papers in D , i.e., for each $p \in D$, $\tilde{S}(r) = S(p)$ and, for each $p \in C \setminus D$, $\tilde{S}(p) = S(p)$.

Invariance to merging papers (IMP): For each $P = (C, w, S) \in \mathcal{P}^N$, each paper subset $D \subseteq C$ consisting of papers with the same set of authors, and each paper $r \notin C \setminus D$,

$$m^r(c^{D \rightarrow r}(P)) = \sum_{q \in D} m^q(P) \quad \text{and, for each } q \in C \setminus D, \quad m^q(c^{D \rightarrow r}(P)) = m^q(P).$$

Thus, a method, m , satisfies *Invariance to merging papers* if the credit allocated to any author for paper r in the new problem $c^{D \rightarrow r}(P)$ is equal to the the total credit that used to be allocated to her for all papers in D in problem P .

Invariance to merging papers is satisfied by the egalitarian, the proportional, and the co-author methods.

The third property prevents authors from increasing their authorship by replicating their identity or multiplying their affiliations. A conceptually related property proposed for journal rankings, “invariance to splitting of journals” (Palacios-Huerta and Volij, 2004), requires that upon splitting journals into smaller replica journals, each with the same aggregate number of citations as the original journal, the ratio between the value of two journals is the same as that of their corresponding replica journals.

Formally, consider a problem $P = (C, w, S) \in \mathcal{P}^N$ and a group of authors $I \subseteq N$. Suppose that each author in I has the same list of co-authored papers. The problem obtained by merging authors in I into a new author k , denoted by $c_{I \rightarrow k}(P)$, is the problem $\tilde{P} = (\tilde{C}, \tilde{w}, \tilde{S})$ with authors $N \setminus I$ and k such that $\tilde{C} = C$, $\tilde{w} = w$, and each paper involving the authors I in P involves author k instead in \tilde{P} , and every other paper has the same authors as in P , i.e., for each $p \in \tilde{C}$, if $I \subseteq S(p)$, $\tilde{S}(p) = [S(p) \setminus I] \cup \{k\}$, and, if $I \not\subseteq S(p)$, $\tilde{S}(p) = S(p)$.

Invariance to merging scholars (IMS): For each $P = (C, w, S) \in \mathcal{P}^N$, each author set $I \subseteq N$ where each author has the same co-authored papers, each $k \notin N \setminus I$, and each $p \in C$,

$$m_k^p(c_{I \rightarrow k}(P)) = \sum_{i \in I} m_i^p(P) \quad \text{and} \quad m_{N \setminus I}^p(c_{I \rightarrow k}(P)) = m_{N \setminus I}^p(P).$$

Thus, a method, m , satisfies *invariance to merging scholars* if the credit allocated to author k for a paper p in the new problem $c_{I \rightarrow k}(P)$ is equal to the the total credit that used to be allocated for paper p to all individuals in I in problem P . *Invariance to merging scholars* is satisfied by the proportional and the co-author methods, but is violated by the egalitarian method.

Our first characterization identifies the only method satisfying the three properties above, namely the co-author method.

Theorem 2. *The only method satisfying consistency, invariance to merging papers, and invariance to merging scholars is the co-author method.*

We introduce a further property supporting the co-author method. It is weak in that it only applies to a basic class of problems, two-author problems, where each author has a single-authored paper and a co-authored paper. The main issue of how the authorship of the

joint paper ought to be assigned is already present here. The question is precisely that of inferring authorship from the worth of the individual contributions. In fact, the proportional and co-author methods both assign authorship proportionally to the individual contributions within this class of two-author problems. That is, they both satisfy:

Pairwise proportionality (PP): For each pair $i, j \in \mathbb{N}$ and each $P = (C, w, S) \in \mathcal{P}^{\{i,j\}}$ where $C = \{p_i, p_j, p\}$, $S(p_i) = \{i\}$, $S(p_j) = \{j\}$, and $S(p) = \{i, j\}$,

$$\frac{m_i^p(P)}{w(p_i)} = \frac{m_j^p(P)}{w(p_j)}.$$

In contrast to the co-author and proportional methods, the egalitarian method does not satisfy *pairwise proportionality*.

Our second characterization establishes that the co-author method is also the only *consistent* method satisfying *invariance to merging papers* and *pairwise proportionality*.

Theorem 3. *The only method satisfying consistency, invariance to merging papers, and pairwise proportionality is the co-author method.*

The properties in Theorems 2 and 3 are logically independent. See appendix E.

6. CONCLUSION

Over half of the one hundred most cited articles in economics are co-authored.⁶ Research is collaborative and yet no ranking has been able to ascertain the *individual* productivity of a scholar, separating her contribution from that of her collaborators. This is not surprising given that the authorship of a publication, the degree to which it can be attributed to any one of its co-authors, is not observable. However, much can be inferred about authorship by exploiting all of the information in the bibliographic database. *CoScore* uncovers individual productivity by observing the varying levels of success of each author in all of her academic partnerships.

It is worthwhile mentioning that *CoScore* is intended to judge and compare individual research records *as they stand*, using only observable information. It is not meant to predict the future productivity of individuals, as discussed more generally by Perry and Reny (2015). However, it may be argued that *CoScore* underestimates the contributions of younger authors. If inferences on the strength of researchers, as quantified by their *CoScores*, are to be based on their publication records, young researchers will receive unreasonably low authorship for their joint work with senior researchers.

CoScore can naturally be extended to address this issue. For instance, suppose that author i has had a career spanning y_i years after she received her PhD. An “age-corrected *CoScore*”

⁶See <https://ideas.repec.org/top/top.item.nbcites.html>.

can be defined as follows: for each $P = (C, w, S)$, each $p \in C$, and each author i ,

$$s_i^y(P) = \frac{1}{\sum_{p \in C} w(p)} \sum_{p \in C_i} w(p) \frac{\frac{s_i^y(P)}{y_i}}{\sum_{j \in S(p)} \frac{s_j^y(P)}{y_j}}.$$

Note that the score of an author is now weighted downwards by her seniority. Our proof techniques can readily establish that the above score and its associated method are well defined.

We stress that *CoScore* can be used to *complement* the existing citation indices, not only as an alternative. The h (Hirsch, 2005), step-based (Chambers and Miller, 2014), and Euclidean (Perry and Reny, 2015) indices rely exclusively on the citation count of each scholar. *CoScore* and its associated co-author method can be used to extract a refined citation count that corrects for co-authorship: an author’s solo papers would maintain their current citations while, for each co-authored paper, citations would be reduced as recommended by the co-author method. The resulting “purely individual” citation count can then be used to compute the above mentioned indices.

To conclude, our analysis seeks to measure the academic value of individuals in a world where research is increasingly collaborative. This requires quantifying authorship by determining the contribution of an individual to a publication. Of course, such estimates are bound to ignore essential elements in the production of joint research and may not always coincide with the actual contributions. However, such an approximation is necessary in assessing individual productivity. As the physicist Lord Kelvin argued,

When you can measure something that you are speaking about, express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of the meager and unsatisfactory kind (Thomson, 1889).

Jacob Viner’s reaction to this dictum was that “even when we can measure a thing, our knowledge will be meager and unsatisfactory” (Merton et al., 1984). However, formalizing the problem of measuring authorship and analyzing it according to objective criteria is the first step in progressing towards sharper measures of scholarly output, thus making our knowledge less meager and more satisfactory.

APPENDIX A. PROOF OF THEOREM 1

Proof. Let $N \in \mathcal{N}$ and $P = (C, w, S) \in \mathcal{P}^N$. Let $\phi : \Delta(N) \rightarrow \Delta(N)$ denote the function such that

$$(3) \quad \text{for each } x \in \Delta(N) \text{ and each } i \in N, \quad \phi_i(x) = \frac{1}{\sum_{p \in C} w(p)} \sum_{p \in C_i} w(p) \frac{x_i}{\sum_{j \in S(p)} x_j}.$$

The co-author score is well defined if and only if ϕ has a unique fixed point.

First note that, for each $x \in \Delta(N)$ and each $i \in N$,

$$\phi_i(x) \geq \frac{w_{ii}}{\sum_{q \in C} w(q)} > 0.$$

Thus, letting

$$K \equiv \left\{ x \in \Delta(N) : \forall i \in N, x_i \geq \frac{w_{ii}}{\sum_{q \in C} w(q)} \right\},$$

we can redefine ϕ as a function mapping K into K . Note that K is compact and convex. Furthermore, ϕ is continuous on K . Thus, by Brouwer's fixed point theorem, ϕ has at least one fixed point. Let x^* denote any one of such fixed points. By (3),

$$(4) \quad \text{for each } i \in N, \quad \sum_{q \in C} w(q) = \sum_{p \in C_i} w(p) \frac{1}{\sum_{j \in S(p)} x_j^*}, \quad \text{and} \quad 1 = \sum_{j \in N} x_j^*.$$

We now prove that (4) admits a unique fixed point. Let $f : \mathbb{R}_{++}^N \rightarrow \mathbb{R}$ denote the function such that

$$\text{for each } x \in \mathbb{R}_{++}^N, \quad f(x) \equiv \sum_{p \in C} w(p) \ln \sum_{j \in S(p)} x_j.$$

Let $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$ denote the function that coincides with f on \mathbb{R}_{++}^N and that takes the value $-\infty$ elsewhere. Note that F is upper-semi-continuous. Thus, since $\Delta(N)$ is compact, $\max_{x \in \Delta(N)} F(x)$ is well defined. By the strict concavity of f ,⁷ there is a unique element in $\arg \max_{x \in \Delta(N)} F(x)$ which we denote by x^{**} . Moreover, since x^{**} is in the relative interior of $\Delta(N)$, for each $i \in N$, $x_i^{**} > 0$. The Lagrangian of the maximization problem is

$$L = F(x) + \lambda \left[\sum_{i \in N} x_i - 1 \right] + \sum_{i \in N} \mu_i x_i$$

where λ is the Lagrange multiplier and corresponding to the constraint $\sum_N x_i = 1$ and, for each $i \in N$, μ_i is Kuhn-Tucker multiplier corresponding to the constraint $x_i \geq 0$. By Corollary 28.2.2 in Rockafellar (1970) these Lagrange and Kuhn-Tucker multipliers exist. By

⁷To see that f is strictly concave, let $g : \mathbb{R}_{++}^N \rightarrow \mathbb{R}$ be defined, for each $x \in \mathbb{R}_{++}^N$, by $g(x) = \sum_{i \in N} w_{ii} \ln x_i$. Function g is strictly concave since its Hessian is negative definite and, for each $i \in N$, $w_{ii} > 0$. Let $D \subseteq C$ consist of all of the papers in C with at least two co-authors: $p \in D$ if and only if $|S(p)| \geq 2$. Let $h : \mathbb{R}_{++}^N \rightarrow \mathbb{R}$ be defined, for each $x \in \mathbb{R}_{++}^N$, by $h(x) = \sum_{p \in D} w(p) \ln \sum_{j \in S(p)} x_j$. As the sum and composition of concave functions, h is concave. Finally, note that $f = g + h$. Since the sum of a strictly concave function and a concave function is strictly concave, f is strictly concave, as desired.

complementary slackness, for each $i \in N$, since $x_i^{**} > 0$, $\mu_i = 0$. The first order conditions are thus,

$$(5) \quad \text{for each } i \in N, \quad \lambda = \sum_{p \in C_i} w(p) \frac{1}{\sum_{j \in S(p)} x_j}, \quad \text{and} \quad 1 = \sum_{j \in N} x_j.$$

By Theorem 28.3 in Rockafellar (1970), if x satisfies (5), $x \in \arg \max_{x \in \Delta(N)} F(x)$, and thus, since the maximizer is unique, $x = x^{**}$. Moreover, by (5),

$$\lambda = \sum_{i \in N} x_i \lambda = \sum_{i \in N} x_i \sum_{p \in C_i} w(p) \frac{1}{\sum_{j \in S(p)} x_j} = \sum_{p \in C} w(p) \frac{\sum_{j \in S(p)} x_j}{\sum_{j \in S(p)} x_j} = \sum_{p \in C} w(p).$$

Thus, by (4) and (5), $x^* = x^{**}$. Recall that x^* is an arbitrary fixed point of ϕ . Thus, if x^* is a fixed point of ϕ , then $x^* = x^{**}$, establishing uniqueness. \blacksquare

APPENDIX B. PROPERTIES OF THE CO-AUTHOR METHOD

Lemma 1. *The co-author method satisfies consistency, invariance to merging papers, invariance to merging scholars, and pairwise proportionality.*

Proof. Let $N \in \mathcal{N}$ and $P = (C, w, S) \in \mathcal{P}^N$.

Consistency: Let $k \in N$, $x \equiv \check{m}(P)$, and $y \equiv \check{m}(c_k^x(P))$. Recall that $c_k^x(P) = (\tilde{C}, \tilde{w}, \tilde{S})$ is such that:

- $\tilde{C} = C \setminus C_{kk}$
- For each $p \in \tilde{C}$ such that $p \in C_k$, $\tilde{w}(p) = w(p) - x_k^p$.
- For each $p \in \tilde{C}$ such that $p \notin C_k$, $\tilde{w}(p) = w(p)$.
- For each $p \in \tilde{C}$, $\tilde{S}(p) = S(p) \setminus \{k\}$.

Then, for each $p \in C_k \setminus C_{kk}$, and each $i \in S(p) \setminus \{k\}$,

$$(6) \quad \begin{aligned} y_i^p &= [w(p) - x_i^p] \frac{\sum_{q \in C_i} y_i^q}{\sum_{j \in S(p) \setminus \{k\}} \sum_{q \in C_j} y_j^q} \\ &= w(p) \left[1 - \frac{\sum_{q \in C_i} x_i^q}{\sum_{j \in S(p)} \sum_{q \in C_j} x_j^q} \right] \frac{\sum_{q \in C_i} y_i^q}{\sum_{j \in S(p) \setminus \{k\}} \sum_{q \in C_j} y_j^q} \\ &= w(p) \left[\frac{\sum_{j \in S(p) \setminus \{k\}} \sum_{q \in C_j} x_j^q}{\sum_{j \in S(p)} \sum_{q \in C_j} x_j^q} \right] \frac{\sum_{q \in C_i} y_i^q}{\sum_{j \in S(p) \setminus \{k\}} \sum_{q \in C_j} y_j^q} \end{aligned}$$

while for each $p \in C \setminus C_k$, and each $i \in S(p)$,

$$(7) \quad y_i^p = w(p) \frac{\sum_{q \in C_i} y_i^q}{\sum_{j \in S(p)} \sum_{q \in C_j} y_j^q}.$$

On the other hand, for each $p \in C_k \setminus C_{kk}$ and each $i \in S(p) \setminus \{k\}$,

$$(8) \quad x_i^p = w(p) \left[\frac{\sum_{q \in C_i} x_i^q}{\sum_{j \in S(p)} \sum_{q \in C_j} x_j^q} \right] = w(p) \left[\frac{\sum_{j \in S(p) \setminus \{k\}} \sum_{q \in C_j} x_j^q}{\sum_{j \in S(p)} \sum_{q \in C_j} x_j^q} \right] \frac{\sum_{q \in C_i} x_i^q}{\sum_{j \in S(p) \setminus \{k\}} \sum_{q \in C_j} x_j^q}.$$

while for each $p \in C \setminus C_k$, and each $i \in S(p)$,

$$(9) \quad x_i^p = w(p) \frac{\sum_{q \in C_i} x_i^q}{\sum_{j \in S(p)} \sum_{q \in C_j} x_j^q}.$$

Let $z \in Z(r_k^x(P))$ be such that, for each $p \in \tilde{C}$ and each $i \in N \setminus \{k\}$, $z_i^p = x_i^p$. Note that, by (8) and (9), z satisfies the system of equations in (6) and (7). By Remark 1, (6) and (7) uniquely define y . Thus, $z = y$. Thus, for each $i \in N \setminus \{k\}$, $\check{m}_i(c_k^x(P)) = \check{m}_i(P)$, so \check{m} satisfies *consistency*.

Invariance to merging papers: Let $D \subseteq C$ be such that, for each pair $p, q \in D$, $S(p) = S(q)$. Recall that $c^{D \rightarrow r}(P) = (\tilde{C}, \tilde{w}, \tilde{S}) \in \mathcal{P}^N$ is such that:

- $\tilde{C} = [C \setminus D] \cup \{r\}$ where $r \notin C \setminus D$.
- $\tilde{w}(r) = \sum_{q \in D} w(q)$ and, for each $q \in C \setminus D$, $\tilde{w}(q) = w(q)$.
- For each $q \in \tilde{C}$, $S(q) = \tilde{S}(q)$.

Thus, using the definition of the co-author score in (1), $\check{v}(c^{D \rightarrow r}(P)) = \check{v}(P)$. By the definition of the co-author score, for each $p \in D$, and each $i \in S(p)$,

$$\sum_{q \in D} \check{m}_i^q(P) = \sum_{q \in D} w(q) \frac{\check{s}_i(P)}{\sum_{j \in S(p)} \check{s}_j(P)} = \tilde{w}(r) \frac{\check{s}_i(c^{D \rightarrow r}(P))}{\sum_{j \in S(p)} \check{s}_j(c^{D \rightarrow r}(P))} = \check{m}_i^r(c^{D \rightarrow r}(P)).$$

while for each paper $\hat{p} \in C \setminus D$, and each $i \in S(\hat{p})$,

$$\check{m}_i^{\hat{p}}(P) = w(\hat{p}) \frac{\check{s}_i(P)}{\sum_{j \in S(p)} \check{s}_j(P)} = \tilde{w}(\hat{p}) \frac{\check{s}_i(c^{D \rightarrow r}(P))}{\sum_{j \in S(p)} \check{s}_j(c^{D \rightarrow r}(P))} = \check{m}_i^{\hat{p}}(c^{D \rightarrow r}(P)).$$

Thus, \check{m} satisfies *IMP*.

Invariance to merging scholars: Let $I \subseteq N$ be such that, for each pair $i, j \in I$, $C_i \setminus C_{ii} = C_j \setminus C_{jj}$. Let $k \in N \setminus I$ and recall that $c_{I \rightarrow k}(P) = (\tilde{C}, \tilde{w}, \tilde{S}) \in \mathcal{P}^{N \setminus I \cup \{k\}}$ is such that:

- $\tilde{C} = C$.
- For each $p \in \tilde{C}$, $\tilde{w}(p) = w(p)$.
- For each $p \in \tilde{C}$, if $I \subseteq S(p)$, then $\tilde{S}(p) = S(p) \setminus I \cup \{k\}$, and, if $I \not\subseteq S(p)$, $\tilde{S}(p) = S(p)$.

By Remark 1, for each $p \in C$ and each $r \in S(p)$:

$$(10) \quad \check{m}_r^p(P) = w(p) \frac{\sum_{q \in C_r} \check{m}_r^q(P)}{\sum_{t \in S(p)} \sum_{q \in C_t} \check{m}_t^q(P)}.$$

Let $z \in Z(c_{I \rightarrow k}(P))$ be such that, for each $p \in \tilde{C}$,

$$z_k^p = \sum_{i \in I} \check{m}_i^p(P) \quad \text{and, for each } r \in S(p) \setminus \{k\}, \quad z_r^p = \check{m}_r^p(P).$$

Then, by (10), for each $p \in \tilde{C}$, and each $r \in \tilde{S}(p)$,

$$z_r^p = \tilde{w}(p) \frac{\sum_{q \in \tilde{C}_r} z_r^q}{\sum_{t \in \tilde{S}(p)} \sum_{q \in \tilde{C}_t} z_t^q}$$

which means z satisfies the system of equations (2) for problem $c_{I \rightarrow k}(P)$. By Remark 1 and Theorem 1, $\check{m}(c_{I \rightarrow k}(P)) = z$. Thus, \check{m} satisfies *IMS*.

Pairwise proportionality: Follows immediately from Proposition 1. ■

APPENDIX C. PROOF OF THEOREMS 2 AND 3

We introduce two additional properties that will be useful in the proofs of Theorems 2 and 3. The first of these reflects the requirement that the indexing of papers is irrelevant to the authorship assignment. Only the worth of papers and the co-authorship relations are taken into consideration:

Neutrality: For each $N \in \mathcal{N}$ and each pair $P = (C, w, S), P' = (C', w', S') \in \mathcal{P}^N$, if there is a bijection $\sigma : C \rightarrow C'$ such that, for each $p \in C$, $w'(\sigma(p)) = w(p)$ and $S'(\sigma(p)) = S(p)$, then, for each $p \in C$ and each $i \in S'(\sigma(p))$, $m_i^{\sigma(p)}(P') = m_i^p(P)$.

The second property specifies that the name of authors bears no influence on the assignment of authorship. Again, only the worth of papers and the co-authorship pattern are taken into consideration:

Anonymity: For each pair $N, N' \in \mathcal{N}$, each $P = (C, w, S) \in \mathcal{P}^N$, and each $P' = (C', w', S') \in \mathcal{P}^{N'}$, if $C = C'$, $w = w'$, and there is a bijection $\pi : N \rightarrow N'$ such that, for each $p \in C$, $S'(p) = \pi(S(p))$, then, for each $p \in C$ and each $i \in S(p)$, $m_i^p(P) = m_{\pi(i)}^p(P')$.

Remark 2. *Invariance to merging papers implies neutrality and invariance to merging scholars implies anonymity.*

Proof of Theorem 3. Let m denote a method satisfying consistency, IMP, and pairwise proportionality. By Lemma 1, it suffices to prove that m is the co-author method, \check{m} .

For each $N \in \mathcal{N}$, each $P = (C, w, S) \in \mathcal{P}^N$, and each $T \subseteq N$, let $I_P(T)$ denote the set of all papers with the same set of authors T , i.e., $I_P(T) = \{p \in C : S(p) = T\}$.

Step 1: For each $N \in \mathcal{N}$ such that $|N| = 2$ and each $P \in \mathcal{P}^N$, $m(P) = \check{m}(P)$.

Let $N = \{a, b\} \in \mathcal{N}$ and $P = (C, w, S) \in \mathcal{P}^N$. Let $K(P)$ denote the total value of the joint papers in P , $K(P) = \sum_{p \in I_P(N)} w(p)$. Let $\mathcal{Q}(P) \subseteq \mathcal{P}^N$ consist of all problems $Q = (\tilde{C}, \tilde{w}, \tilde{S}) \in \mathcal{P}^N$ such that:

- $\tilde{C}_{aa} = C_{aa}$ and $\tilde{C}_{bb} = C_{bb}$.
- for each $p \in \tilde{C}_{aa} \cup \tilde{C}_{bb}$, $\tilde{w}(p) = w(p)$.
- $\sum_{p \in I_Q(N)} w(p) = K(P)$.

Note that $P \in \mathcal{Q}(P)$.

Step 1A: For each $P' = (C', w', S')$, $P'' = (C'', w'', S'') \in \mathcal{Q}(P)$, each $p' \in I_{P'}(N)$, and each $p'' \in I_{P''}(N)$,

$$w'(p') = w''(p'') \quad \Rightarrow \quad m^{p'}(P') = m^{p''}(P'').$$

Let all of the notation be as in the statement of Step 1A. Without loss of generality, suppose that $|I_{P'}(N)| \geq 2$ and $|I_{P''}(N)| \geq 2$.⁸ Let $\hat{P} = (\hat{C}, \hat{w}, \hat{S}) \in \mathcal{Q}(P)$ be such that $I_{\hat{P}}(N) = \{p', \hat{p}\}$, $\hat{w}(p) = w'(p')$, and $\hat{w}(\hat{p}) = K(P) - w'(p')$. By IMP, merging all the papers in $I_{P'}(N) \setminus \{p'\}$ into \hat{p} , $m^{p'}(P') = m^{p'}(\hat{P})$. Similarly, let $\tilde{P} = (\tilde{C}, \tilde{w}, \tilde{S}) \in \mathcal{Q}(P)$ be such that $I_{\tilde{P}}(N) = \{p'', \tilde{p}\}$, $\tilde{w}(p) = w''(p'')$, and $\tilde{w}(\tilde{p}) = K(P) - w''(p'')$. By IMP, merging all the papers in $I_{P''}(N) \setminus \{p''\}$ into \tilde{p} , $m^{p''}(P'') = m^{p''}(\tilde{P})$. By neutrality, which is implied by IMP (Remark 2), since the problems \hat{P} and \tilde{P} differ only in the indexing of papers, $m^{p'}(P') = m^{p'}(\hat{P}) = m^{p''}(\tilde{P}) = m^{p''}(P'')$.

Step 1B: For each pair $p, q \in C$,

$$\frac{m_a^p(P)}{m_a^q(P)} = \frac{w(p)}{w(q)}.$$

⁸Otherwise, we necessarily have $|I_{P'}(N)| = |I_{P''}(N)| = 1$ and the result simply follows from neutrality which, by Remark 2, is implied by IMP.

Let $f : [0, K(P)] \rightarrow \mathbb{R}$ be such that

$$f(x) = m_a^p(Q) \quad \forall x \in [0, K(P)]$$

where $Q = (\mathring{C}, \mathring{w}, \mathring{S}) \in \mathcal{Q}(P)$ and p is any paper in \mathring{C} such that $\mathring{w}(p) = x$ and $\mathring{S}(p) = N$. By Step 1A, f is well defined. Furthermore, by IMP,

$$f(x+y) = f(x) + f(y) \quad \forall x, y \in [0, K(P)] \text{ such that } x+y \leq K(P).$$

Thus, f satisfies the Cauchy functional equation and there is a constant $c_{K(P)}$ such that, for each $x \in [0, K(P)]$, $f(x) = c_{K(P)}x$ (see Theorem 3 in page 48 of Aczél, 2006). Since $P \in \mathcal{Q}(P)$, for any $p, q \in C$, $m_a^p(P) = f(w(p)) = c_{K(P)}w(p)$ and $m_a^q(P) = f(w(q)) = c_{K(P)}w(q)$. This establishes Step 1B.

Step 1C: For each $q \in C$,

$$m_a^q(P) = \frac{w_{aa}}{w_{aa} + w_{bb}}w(q).$$

Let $P^r = (C^r, w^r, S^r) \in \mathcal{P}^N$ be such that

$$\begin{aligned} C^r &= \{p_a, p_b, p_{ab}\} \\ S^r(p_a) &= \{a\}, \quad S^r(p_b) = \{b\}, \quad S^r(p_{ab}) = N, \\ w^r(p_a) &= w_{aa}, \quad w^r(p_b) = w_{bb}, \quad w^r(p_{ab}) = \sum_{p \in I_P(N)} w(p). \end{aligned}$$

By IMP,

$$(11) \quad m_a^{p_{ab}}(P^r) = \sum_{p \in I_P(N)} m_a^p(P).$$

By PP,

$$(12) \quad m_a^{p_{ab}}(P^r) = \frac{w^r(p_a)}{w^r(p_a) + w^r(p_b)}w^r(p_{ab}).$$

By Step 1B,

$$\text{for each pair } p, q \in C, \quad m_a^p(P) = \frac{w(p)}{w(q)}m_a^q(P)$$

Thus, by (11),

$$m_a^{p_{ab}}(P^r) = \sum_{p \in I_P(N)} \frac{w(p)}{w(q)}m_a^q(P) = \frac{m_a^q(P)}{w(q)} \sum_{p \in I_P(N)} w(p) = \frac{m_a^q(P)}{w(q)}w^r(p_{ab}).$$

Thus, by (12),

$$\frac{w^r(p_a)}{w^r(p_a) + w^r(p_b)} = \frac{m_a^q(P)}{w(q)}.$$

Thus,

$$\text{for each } q \in C, \quad m_a^q(P) = \frac{w^r(p_a)}{w^r(p_a) + w^r(p_b)} w(q) = \frac{w_{aa}}{w_{aa} + w_{bb}} w(q) = \check{m}_a^q(P).$$

Step 2: For each $N \in \mathcal{N}$ and each $P \in \mathcal{P}^N$, $m(P) = \check{m}(P)$.

Induction hypothesis: Let $n \in \mathbb{N}$ and suppose that, for each $N \in \mathcal{N}$ such that $|N| \leq n$ and each $P \in \mathcal{P}^N$, $m(P) = \check{m}(P)$.

Let $N \in \mathcal{N}$ be such that $|N| = n + 1$, $P = (C, w, S) \in \mathcal{P}^N$, and $x \equiv m(P)$. For each $k \in N$, let $N_k = N \setminus \{k\}$. Let $k \in N$ and consider problem $c_k^x(P)$. By *consistency*,

$$(13) \quad \text{for each } p \in C \setminus C_{kk} \text{ and each } i \in S(p) \setminus \{k\}, \quad m_i^p(c_k^x(P)) = x_i^p.$$

By the induction hypothesis, since $|N_k| = n$,

$$\text{for each } p \in C \setminus C_{kk} \text{ and each } i \in S(p) \setminus \{k\}, \quad m_i^p(c_k^x(P)) = \check{m}_i^p(c_k^x(P)).$$

Thus, by the definition of \check{m} in Remark 1, for each $p \in C \setminus C_{kk}$ and each $i \in S(p) \setminus \{k\}$,

$$m_i^p(c_k^x(P)) = (w(p) - x_k^p) \frac{\sum_{q \in C_i} m_i^q(c_k^x(P))}{\sum_{j \in S(p) \setminus \{k\}} \sum_{q \in C_j} m_j^q(c_k^x(P))}$$

where, abusing notation, if $p \in C_k$, we let $x_k^p = 0$. Thus, by (13),

$$\text{for each } p \in C \setminus C_k \text{ and each } i \in S(p) \setminus \{k\}, \quad x_i^p = (w(p) - x_k^p) \frac{\sum_{q \in C_i} x_i^q}{\sum_{j \in S(p) \setminus \{k\}} \sum_{q \in C_j} x_j^q}$$

where, again, abusing notation, if $p \in C_k$, we let $x_k^p = 0$. Thus,

$$\text{for each } p \in C \setminus C_{kk} \text{ and each pair } i, j \in S(p) \setminus \{k\}, \quad \frac{x_j^p}{x_i^p} = \frac{\sum_{q \in C_j} x_j^q}{\sum_{q \in C_i} x_i^q}.$$

Repeating the same argument for each $k \in N$,

$$\text{for each } p \in C \text{ and each pair } i, j \in S(p), \quad x_j^p = x_i^p \cdot \frac{\sum_{q \in C_j} x_j^q}{\sum_{q \in C_i} x_i^q}.$$

Thus, for each $p \in C$ and each $i \in S(p)$, summing over $j \in S(p)$,

$$w(p) = \sum_{j \in S(p)} x_j^p = x_i^p \cdot \frac{\sum_{j \in S(p)} \sum_{q \in C_j} x_j^q}{\sum_{q \in C_i} x_i^q}.$$

Thus,

$$\text{for each } p \in C \text{ and each } i \in S(p), \quad x_i^p = w(p) \frac{\sum_{q \in C_i} x_i^q}{\sum_{j \in S(p)} \sum_{q \in C_j} x_j^q}.$$

Hence, x satisfies the system of equations (2) defining $\check{m}(P)$ in Remark 1. However, $\check{m}(P)$ is the only solution to the system of equations (2). Thus, in fact, $\check{m}(P) = x = m(P)$, as desired. \blacksquare

Lemma 2. *If m satisfies neutrality and invariance to merging scholars, then it satisfies pairwise proportionality.*

Proof. Suppose that m satisfies *neutrality* and *IMS*. By Remark 2, m also satisfies *anonymity*. For any $N \in \mathcal{N}$, let $\mathcal{Q}^N \subseteq \mathcal{P}^N$ consist of all problems such that each author, $i \in N$, has exactly one single-authored paper, denoted by p_i , and all authors participate in writing a joint paper, denoted by p_0 . Let $i, j \in \mathbb{N}$ and $P = (C, w, S) \in \mathcal{Q}^{\{i, j\}}$. For each $a \in [0, w(p_i) + w(p_j)]$, let $P_a = (\tilde{C}, \tilde{w}, \tilde{S}) \in \mathcal{Q}^{\{i, j\}}$ be such that

$$\tilde{C} = C, \quad \tilde{w}(p_i) = a \quad \text{and} \quad \tilde{w}(p_j) = w(p_i) + w(p_j) - a.$$

Note that $P = P_{w(p_i)}$. Let $a', a'' \in [0, w(p_i) + w(p_j)]$ be such that $a' + a'' = a$. Let $i', i'' \in \mathbb{N} \setminus \{i, j\}$ and $P'_a = (C', S', w') \in \mathcal{Q}^{\{i', i'', j\}}$ be such that

$$w'(p_{i'}) = a', \quad w'(p_{i''}) = a'', \quad w'(p_j) = w(p_j), \quad w'(p_0) = w(p_0).$$

By *IMS*,

$$m_i^{p_0}(P_a) = m_{i'}^{p_0}(P'_a) + m_{i''}^{p_0}(P'_a).$$

Similarly, let $j', j'' \in \mathbb{N} \setminus \{i, j\}$ and $P''_a = (C'', S'', w'') \in \mathcal{Q}^{\{i, j', j''\}}$ be such that

$$w''(p_i) = a', \quad w''(p_{j'}) = a'', \quad w''(p_j) = w(p_j), \quad w''(p_0) = w(p_0).$$

By *IMS*,

$$m_i^{p_0}(P'_a) = m_i^{p_0}(P''_a).$$

By *anonymity* and *neutrality*,

$$m_i^{p_0}(P''_a) = m_{i'}^{p_0}(P'_a).$$

so that $m_i^{p_0}(P'_a) = m_{i'}^{p_0}(P'_a)$. Similarly, $m_i^{p_0}(P''_a) = m_{i''}^{p_0}(P'_a)$. We conclude that

$$m_i^{p_0}(P_a) = m_i^{p_0}(P'_a) + m_i^{p_0}(P''_a).$$

Define $f : [0, w(p_i) + w(p_j)] \rightarrow [0, w(p_0)]$ such that, for each $a \in [0, w(p_i) + w(p_j)]$, $f(a) = m_i^{p_0}(P_a)$. Since, $f(a' + a'') = f(a') + f(a'')$, f satisfies the Cauchy functional equation and there is a constant c_P such that, for each $a \in [0, w(p_i) + w(p_j)]$, $f(a) = c_P a$ (see Theorem 3 in page 48 of Aczél, 2006). Since $P = P_{w(p_i)}$, $m_i^{p_0}(P) = c_P w(p_i)$. Furthermore, by *anonymity* and *neutrality*, $m_j^{p_0}(P) = m_i^{p_0}(P_{w(p_j)}) = c_P w(p_j)$. We conclude that

$$\frac{m_i^{p_0}(P)}{w(p_i)} = \frac{m_j^{p_0}(P)}{w(p_j)}.$$

That is, m satisfies *PP*. ■

Proof of Theorem 2. Let m denote a method satisfying the properties in Theorem 2. By Remark 2, m satisfies *neutrality*. Thus, by Lemma 2, m satisfies *PP*. Thus, by Theorem 3, m is the co-author method. ■

APPENDIX D. PROOF OF PROPOSITION 1

Proof. Let $N \in \mathcal{N}$ and $P = (C, w, S)$ denote a symmetric problem in \mathcal{P}^N . We show that \check{m} coincides with \hat{m} for problem P . Note that, for each pair $i, j \in N$, $C \setminus C_{ii} = C \setminus C_{jj}$. For each joint paper $p \in C$ and each $i \in N$ we have:

$$\begin{aligned} \frac{\sum_{q \in C_i} \hat{m}_i^q(P)}{\sum_{j \in N} \sum_{q \in C_j} \hat{m}_j^q(P)} w(p) &= \frac{w_{ii} + \sum_{q \in C \setminus C_{ii}} \frac{w_{ii}}{\sum_{k \in N} w_{kk}} w(q)}{\sum_{j \in N} w_{jj} + \sum_{q \in C \setminus C_{jj}} \frac{w_{jj}}{\sum_{k \in N} w_{kk}} w(q)} w(p) \\ &= \frac{w_{ii} \left(1 + \frac{\sum_{q \in C \setminus C_{ii}} w(q)}{\sum_{k \in N} w_{kk}} \right)}{\sum_{j \in N} w_{jj} \left(1 + \frac{\sum_{q \in C \setminus C_{jj}} w(q)}{\sum_{k \in N} w_{kk}} \right)} w(p) \\ &= \frac{w_{ii}}{\sum_{j \in N} w_{jj}} w(p) \\ &= \hat{m}_i^p(P). \end{aligned}$$

Thus, $\hat{m}(P)$ satisfies equation system (2). By Remark 1 and Theorem 1, $\check{m}(P) = \hat{m}(P)$, as desired. ■

Proposition 2. *Consistency, invariance to merging papers, and pairwise proportionality are logically independent. Consistency, invariance to merging papers, and invariance to merging scholars are logically independent.*

Proof. The egalitarian rule satisfies *Consistency* and *IMP* but not *PP*. The proportional rule satisfies *IMP* and *PP* but not *consistency*. We now exhibit a method that satisfies *consistency* and *PP* but not *IMP*.

Take an author $i \in \mathbb{N}$ and a function $x : \mathbb{C} \rightarrow \{0, 1\}$. Let $N \in \mathcal{N}$ and $P = (C, W, S) \in \mathcal{P}^N$. If $i \in N$, let $z(P) \in Z(P)$ be such that, for each $p \in C$, $z_i^p(P) = x(p)w(p)$. Define the method \mathring{m} as follows:

1st Case: i appears in at most 2 papers in C . Then, for each paper $p \in C$,

$$m^p(P) = \check{m}^p(P)$$

2nd Case: i appears in at least 3 papers in C . Then, for each paper $p \in C$,

$$\begin{aligned} m_i^p(P) &= x(p)w(p) \\ m_j^p(P) &= \check{m}_j^p(c_i^{z(P)}(P)) \quad \text{for each } j \neq i \end{aligned}$$

First note that *PP* only applies in Case 1. Therefore, by Lemma 1, \mathring{m} satisfies *PP*. Consider any problem $P = (C, w, S) \in \mathcal{P}^N$, where $i \in N$, such that there exists two identical papers p and q in C_i : $i \in S(p) = S(q)$ and $w(p) = w(q)$. Then, since $x(\cdot)$ may be chosen in an arbitrary way, we can write:

$$m_i^p(P) = x(p)w(p) \neq x(q)w(q) = m_i^q(P)$$

thus violating *IMP* (which implies that two identical papers are shared in the same way).

Finally, we show that \mathring{m} satisfies *consistency*. Let $N \in \mathcal{N}$ and $P = (C, w, S) \in \mathcal{P}^N$. Let $k \in N$ and consider the problem $c_k^{\mathring{m}(P)}(P) = (\tilde{C}, \tilde{w}, \tilde{S})$.

1st Case: i appears in at most 2 papers in C . Then, whether $k = i$ (i is taken out) or $k \neq i$, i still appears in at most 2 papers in \tilde{C} . Therefore, for each $p \in C$ and each $j \in N \setminus \{k\}$:

$$\mathring{m}_j^p(P) = \check{m}_j^p(P) = \check{m}_j^p(c_k^{\mathring{m}(P)}(P)) = \mathring{m}_j^p(c_k^{\mathring{m}(P)}(P))$$

since the co-author method \check{m} satisfies *consistency*. We conclude that \mathring{m} satisfies *consistency*.

2nd Case: i appears in at least 3 papers in C .

2.1: $k = i$. Then, since i doesn't appear in any paper in \tilde{C} (i has been taken out), we have $\mathring{m}_j^p(c_k^{\mathring{m}(P)}(P)) = \check{m}_j^p(c_k^{\mathring{m}(P)}(P))$. Furthermore, since $c_k^{\mathring{m}(P)}(P) = c_k^{z(P)}(P)$, we also have:

$\hat{m}_j^p(P) = \check{m}_j^p(c_i^z(P)) = \check{m}_j^p(c_k^{\hat{m}(P)}(P))$ so that $\hat{m}_j^p(P) = \hat{m}_j^p(c_k^{\hat{m}(P)}(P))$. We conclude that \hat{m} satisfies *consistency*.

2.2: $k \neq i$. Then i still appears in at least 3 papers in \tilde{C} . Therefore, for each $p \in C$,

$$\hat{m}_i^p(c_k^{\hat{m}(P)}(P)) = x(p)(w(p) - \hat{m}_k^p(P)) = x(p)w(p) = \hat{m}_i^p(P)$$

For each $j \in N \setminus \{i, k\}$,

$$\hat{m}_j^p(c_k^{\hat{m}(P)}(P)) = \check{m}_j^p(c_i^{z(c_k^{\hat{m}(P)}(P))}(c_k^{\hat{m}(P)}(P)))$$

Furthermore, since i is allocated the same share of paper p in problems $c_k^{\hat{m}(P)}$ and P :

$$c_i^{z(r_k^{\hat{m}(P)}(P))}(c_k^{\hat{m}(P)}(P)) = c_i^{z(P)}(c_k^{\hat{m}(P)}(P))$$

and since (a) reducing P first by taking out k (according to $\hat{m}(P)$) then by taking out i (according to $z(P)$) is equivalent to (b) reducing P first by taking out i (according to $z(P)$), then by taking out k (according to $\hat{m}(c_i^{z(P)}(P))$), which is also equivalent to (c) reducing P first by taking out i (according to $z(P)$), then by taking out k (according to $\check{m}(c_i^{z(P)}(P))$):

$$c_i^{z(P)}(c_k^{\hat{m}(P)}(P)) = c_k^{\hat{m}(c_i^{z(P)}(P))}(c_i^{z(P)}(P)) = c_k^{\check{m}(c_i^{z(P)}(P))}(c_i^{z(P)}(P))$$

We conclude that

$$\hat{m}_j^p(c_k^{\hat{m}(P)}(P)) = \check{m}_j^p(c_k^{\check{m}(r_i^{z(P)}(P))}(c_i^{z(P)}(P))) = \check{m}_j^p(c_i^{z(P)}(P)) = \hat{m}_j^p(P),$$

using the fact that \check{m} satisfies *consistency*. Thus, \hat{m} satisfies *consistency*. ■

REFERENCES

- Aczél, J. (2006). *Lectures on Functional Equations and their Applications*. Dover.
- Chambers, C. and A. Miller (2014). Scholarly influence. *Journal of Economic Theory* 151.
- Conley, J. and A. S. Önder (2014). The research productivity of new PhDs in economics: the surprisingly high non-success of the successful. *Journal of Economic Perspectives* 28(3), 205–216.
- Demange, G. (2014). A ranking method based on handicaps. *Theoretical Economics* 9, 915–942.
- Hirsch, J. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Science* 102(46), 16569–16572.
- Jackson, M. (2008). *Social and economic networks*. Princeton University Press.
- Kalaitzidakis, P., T. Mamuneas, and T. Stengos (2003). Ranking of academic journals and institutions in economics. *Journal of the European Economic Association* 1, 1346–1366.
- Laslier, J. F. (1997). *Tournament solutions and majority voting*. Berlin, Heidelberg, New York: Springer.
- Merton, R., D. Sills, and S. Stigler (1984). The Kelvin dictum and social science: an excursion into the history of an idea. *Journal of the History of the Behavioral Sciences* 20.

- Page, L., S. Brin, R. Motawani, and T. Winograd (1998). The PageRank citation ranking: Bringing order to the web. *Working Paper, Stanford University*.
- Palacios-Huerta, I. and O. Volij (2004). The measurement of intellectual influence. *Econometrica* 72(3), 963–977.
- Palacios-Huerta, I. and O. Volij (2014). Axiomatic measures of intellectual influence. *International Journal of Industrial Organization* 34.
- Perry, M. and P. Reny (2015). How to count citations if you must. *working paper*.
- Pinski, G. and F. Narin (1976). Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics. *Information Processing and Management* 12.
- Rockafellar, T. (1970). *Convex Analysis*. Princeton University Press.
- Thomson, W. (1889). *Popular lectures and addresses*. London Macmillan.
- Thomson, W. (2011). Consistency and its converse: an introduction. *Review of Economic Design* 15(4), 257 – 291.
- Thomson, W. (2012). On the axiomatics of resource allocation: Interpreting the consistency principle. *Economics and Philosophy* 28(3), 385–421.