# Scaling Rules in the Science System: Influence of Field-Specific Citation Characteristics on the Impact of Individual Researchers

**Rodrigo Costas and Maria Bordons**
*IEDCYT–Centre for Humanities and Social Sciences (CCHS), Spanish Council for Scientific Research (CSIC), Madrid, Spain. E-mail: {rodrigo.costas,mbordons}@cindoc.csic.es*

**Thed N. van Leeuwen and Anthony F.J. van Raan**
*Centre for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands. E-mail: leeuwen@fsw.leidenuniv.nl,vanraan@cwts.leidenuniv.nl*

**The representation of science as a citation density landscape and the study of scaling rules with the field-specific citation density as a main topological property was previously analyzed at the level of research groups. Here, the focus is on the individual researcher. In this new analysis, the size dependence of several main bibliometric indicators for a large set of individual researchers is explored. Similar results as those previously observed for research groups are described for individual researchers. The total number of citations received by scientists increases in a cumulatively advantageous way as a function of size (in terms of number of publications) for researchers in three areas: Natural Resources, Biology & Biomedicine, and Materials Science. This effect is stronger for researchers in low citation density fields. Differences found among thematic areas with different citation densities are discussed.**

## Introduction

Science can be considered as a complex system of highly interconnected entities (e.g., researchers, research groups, universities) that produce and transfer knowledge. There is much recent work on the study of networks in science, such as those focused on the study of links among authors, publications, and citations (Albert & Barabási, 2002; Dorogovtsev & Mendes, 2002; Leicht, Clarkson, Shedden, & Newman 2007). But there is little work on the study of bibliometric indicators and their statistical properties in the context of science as an interconnected system. Particularly important in large networked systems (Caldarelli, Erzan, & Vespignani, 2004) are the relations between large-scale attributes (e.g., in science, the citation characteristics of fields) and local patterns (e.g., the performance in terms of citation-based impact of individual researchers).

The scaling relationships between number of citations and number of publications have been analyzed across countries, research fields, and institutes (Katz, 1999, 2000, 2005) as well as across research groups (van Raan, 2008a). An important finding was that citations increase in a power law relationship with the size (in terms of number of publications) of the groups, institutions, or nations, and a cumulative advantage effect in the scientific community was observed; that is, a size-dependent Matthew effect (Merton, 1988). Mechanisms for generating power laws and the methods to detect them were discussed by Newman (2005).

In a series of studies on the statistical properties of bibliometric characteristics of research groups (van Raan, 2006a, 2006b, 2008a), the size-dependent nature of impact was analyzed focusing on the differences between top-performance and lower performance groups. The crucial finding was that particularly the lower performance groups have a size-dependent *cumulative* advantage[1] for receiving citations. Two different underlying factors interact: First, the fraction of not-cited publications, which for lower performance groups

[1]By "cumulative advantage," we mean that the dependent variable (e.g., number of citations of a Group C) scales in a disproportional, nonlinear way (power law) with the independent variable (e.g., the "size" of a research group, in terms of number of publications, $P$). Thus, larger groups (in terms of $P$) do not just receive more citations (as can be expected), but they do so increasingly more "advantageously:" Groups that are twice as large as other groups receive, for instance, 2.4 times more citations. For a detailed discussion, please refer to our previous article (van Raan, 2006b). For a general discussion of cumulative advantage in science, please refer to Merton (1988) and Price (1976).

Received March 7, 2008; revised November 20, 2008; accepted November 20, 2008

 • Published online 12 January 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.21017

decreases considerably with size (van Raan, 2006a),[2] and second, the *citation density*[3] of the fields since research groups working in low field-citation-density regions tend to benefit the most from a higher number of publications (van Raan, 2008a). In the latter publication, the scaling behavior in relation to the size dependency of the main bibliometric indicators applied to the study of research groups for different levels of field-specific citation densities was analyzed.

Following this line of research, we continue in this article our exploration of these interdependencies of the science system as a landscape characterized by field-specific citation densities, including a new level of analysis: the individual researcher. We wondered whether the scaling behavior identified at the research-group level also could be observed at the individual level. Different questions emerge. Do the scaling rules described for the research groups also apply for individual researchers? Is research performance of researchers influenced by the characteristics of fields, and if so, how? The structure of this article is as follows. First, we discuss the data material, the application of the method, and the calculation of indicators. Second, we present the results of our data analysis for "external" (i.e., non-self-) citations, and finally, we discuss the main outcomes of this study in the framework of the landscape model.

## Data, Indicators, Citation-Density Landscape

Data for this study come from the analysis of scientific activity of a total of 1,064 researchers working as scientific staff at the Spanish CSIC[4] in 2005, which represents 45% of total researchers of the institution staff. These researchers are grouped according to the thematic orientation of their institutes in three main scientific areas, namely Natural Resources (349 researchers), Biology & Biomedicine (388 researchers), and Materials Science (327 researchers); 1,038 researchers have at least one *Web of Science* (*WoS*) publication[5] in the period under study, 1994 to 2004. In total, the analysis covers about 25,000 publications and 222,300 citations (excluding self-citations). Obtaining bibliometric indicators at the individual level is laborious due to the lack of normalization of author names in the publications. A careful analysis of author names and addresses in publications was carried out to properly identify the scientific production of researchers (Costas & Bordons, 2005).

The indicators are calculated on the basis of a total time-period analysis. This means that publications are counted for the entire 11-year period (1994–2004) and that citations are counted up to and including 2004 (e.g., for publications from 1994, citations are counted from 1994–2004; for publications from 2004, citations are counted only in 2004).[6] We applied the Centre for Science and Technology Studies (CWTS) standard bibliometric indicators. Here, only "external" citations (i.e., citations corrected for self-citations[7]) are taken into account. Next, we present the standard bibliometric indicators, each with a short description. For a detailed discussion, please refer to van Raan (2004). For the analysis, only researchers with at least five publications were considered.

The standard bibliometric indicators are:

- Number of publications ***P*** in CI-covered journals of a researcher in the specified period;
- Number of citations ***C*** received by ***P*** during the specified period, without self-citations;
- Average number of citations per publication, without self-citations (***CPP***);
- Journal-based worldwide average impact as an international reference level for a researcher (***JCS***, journal citation score, which is our journal impact indicator), without self-citations (on a world-wide scale!); in the case of more than one journal, we use the (weighted) average ***JCSm***; for the calculation of ***JCSm***, the same publication and citation counting procedure, time windows, and article types are used as in the case of ***CPP***;
- Field-based[8] world-wide average impact as an international reference level for a researcher (***FCS***, field citation score), without self-citations (on a world-wide scale!); in the case of more than one field (as almost always), we use the (weighted) average ***FCSm***; for the calculation of ***FCSm***, the same publication and citation counting procedure, time windows, and article types are used as in the case of ***CPP***; we refer in this article to the ***FCSm*** indicator as the "field-specific citation density;"
- Comparison of the ***CPP*** of a researcher with the world-wide average based on ***JCSm*** as a standard, without self-citations, indicator ***CPP/JCSm***; it allows us to observe whether the impact of a researcher is above or below the international average in his or her publication journals.
- Comparison of the ***CPP*** of a researcher with the world-wide average based on ***FCSm*** as a standard, without self-citations, indicator ***CPP/FCSm***; it allows us to observe whether the

[2]In this context, the role of self-citation as impact-reinforcing mechanism is discussed in van Raan (2008b).

[3]By "citation density," we refer to the mean number of citations received by publications in a certain field, as measured by the *FCSm* indicator (see description of indicators in the text).

[4]Consejo Superior de Investigaciones Científicas (Spanish Council for Scientific Research).

[5]Thomson Reuters, formerly the Institute for Scientific Information (ISI) in Philadelphia, is the producer and publisher of the *Web of Science* (*WoS*) that covers the *Science Citation Index* (-Extended), the *Social Science Citation Index*, and the *Arts & Humanities Citation Index*. Throughout this article, we use the term *WoS* for the aforementioned set of databases.

[6]All researchers were given an 11-year period for publications and citations, including publications before entering CSIC for those who joined the institution during the period of analysis.

[7]A citation is a self-citation if any of the authors of the citing paper also are an author of the cited paper. It must be taken into account that self-citations have been removed from the publications of the individual researchers and also from all publications used as an international reference. Thus, only "external citations" (i.e., citations given by authors different from the coauthors of the original paper) have been considered for the calculation of all indicators.

[8]Here, we use the definition of fields based on a classification of scientific journals into *categories* developed by Thomson Reuters. Although this classification is not perfect, it provides a clear and "fixed" consistent field definition suitable for automated procedures within our data system.


DOI: 10.1002/asi

impact of a researcher is above or below the international average in his or her field.
- Ratio ***JCSm/FCSm*** is the relative, field-normalized journal impact indicator. It indicates if the researcher publishes in journals with high or low impact within the field.

Relative indicators of impact are especially useful because they compare the activity of scientists with an international reference (e.g., see one of the first works on relative indicators by Moed, Burger, Frankfort & van Raan, 1985; also see Aksnes & Taxt, 2004; Schubert & Braun, 1986; Vinkler, 1986). Among these indicators, the internationally standardized (field-normalized) ***CPP/FCSm*** indicator is considered by CWTS as the "crown indicator" since it enables us to observe whether the performance of a unit of analysis is significantly far below (indicator value < 0.5), below (0.5–0.8), around (0.8–1.2), above (1.2–1.5), or far above (>1.5) the international impact standard of the field. Particularly with a ***CPP/FCSm*** value above 1.5, units of analysis can be considered as scientifically strong. A value above 2 indicates a very strong unit, and units with values above 3 generally can be considered as excellent and comparable to top units in the best U.S. universities (van Raan, 2004). A good correlation of ***CPP/FCSm*** and quality judgment of peers has been described elsewhere (Rinia, van Leeuwen, van Vuren, & van Raan 1998, 2001).

In this work, the relationship between variables is studied through correlation analyses to detect "advantages" and "disadvantages" of the increasing size (i.e., number of publications) on the impact of research. Since differences are expected according to the citation density of the subfields (as measured through the ***FCSm*** indicator), the characteristics of *high* and *low field-citation-density* researchers are compared; high and low field-citation-density researchers are those in the top-25% and bottom-25% of the distribution of ***FCSm***, respectively (i.e., Percentiles 25 and 75 of the ***FCSm*** distribution). A research area may comprise subfields with different citation patterns; however, in this study, we distinguish between researchers in low and high citation-density fields. This largely reduces the risk of inadequate comparison among scientists from subfields with different citation characteristics.

In the correlation analysis, due to the high variability of data, a low $R^2$ (determination coefficient) was obtained in some cases, and complementary tests were used to support the results. Researchers were grouped into four categories (P1–P4) according to their productivity (i.e., their "scientific size" in terms of number of publications). Size-dependent differences of a number of impact measures were explored for high and low field-citation-density researchers with the help of ANOVA,[9] after normalization of the variables by the logarithm. For the classification of researchers into four productivity (i.e., size) classes, the percentiles of the distribution of the number of publications within each thematic area are used (see Appendix, Table A1).

[9]Analysis of variance (ANOVA) is a statistical procedure in which the observed variance is partitioned into components due to different explanatory variables.

## Results and Discussion

### *Influence of Field-Specific Citation Density and Journal Impact*

In Figure 1, we present the distribution of publications by scientific fields[6] to show the thematic composition of each of the three main CSIC areas. Thus, in this study, we consider an area as a higher, interdisciplinary aggregate of several fields.

In this sense, there are clear scientific orientations in each area. Biology & Biomedicine researchers present a higher percentage of their publications in the fields Biochemistry & Molecular Biology and Neurosciences. Materials Science researchers show an orientation towards Condensed Matter Physics, Materials Science, Physical Chemistry, Polymer Sciences, and so on. Natural Resources is the most interdisciplinary area, with publications in a wide range of fields such as Marine and Freshwater Biology, Ecology, Oceanography, Zoology, and so on.

Differences in field-citation density and research performance of researchers in the three areas also can be observed (Table 1). The high field-citation density of Biology & Biomedicine (***FCSm***) is remarkable; it is far above the densities in the other two areas. Although Biology & Biomedicine researchers are not the most productive ones in terms of ***P***, they obtain the highest number of ***CPP***, which is, of course, related to the high ***FCSm*** value. In the three areas, researchers tend to publish in journals with an impact above the average in their field (***JCS/FCSm*** >1), although they do not obtain as many citations as their journals (*CPP/JCSm* <1). The number of citations received is below the average of their field (***CPP/FCSm*** <1) for Materials Science and Natural Resources, and not statistically different from the world-wide average in the case of Biology & Biomedicine. Note that 48% of researchers in Biology & Biomedicine show a ***CPP/FCS*** higher than 1 while this percentage is around 37% in the other two areas.

In Figure 2, the correlation of the number of citations (***C***) with number of publications (***P***) for each of the three areas is presented. Researchers are classified according to *high* and *low* field-citation densities (i.e., the top-25% and bottom-25%, respectively) of the ***FCSm*** distribution. Figure 2 shows that there is a cumulative "advantage" effect in the three fields (i.e., the power law exponents of the correlation functions are >1); but it is higher for the *low* field-citation-density researchers (i.e., ***C*** increases with ***P*** more for bottom-25% researchers (i.e., higher power law exponents). These results are consistent with those obtained at the research-group level in a previous article (van Raan, 2008a).

Thus, as ***P*** increases, the difference in number of citations between high and low field-citation-density researchers will become smaller. This trend also can be observed in Figure 3, where we show the relationship between ***CPP*** and ***P*** for the high and the low field-citation-density researchers. Individual


DOI: 10.1002/asi

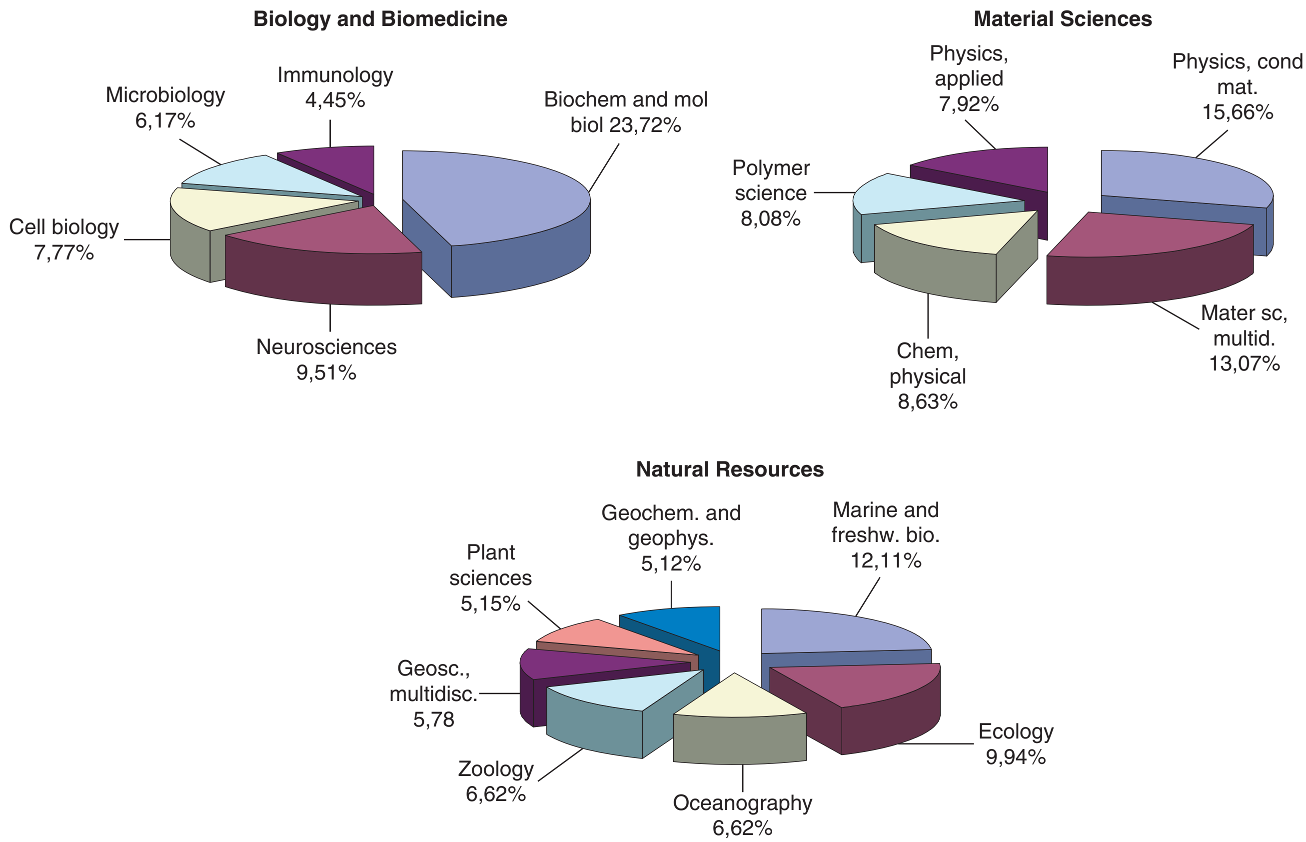


FIG. 1. Distribution of publications within the three areas by field (Largest fields account for 50% of publications).

TABLE 1. Performance of Individual Researchers by Areas.

| | ***N*** | ***P*** | ***C*** | ***CPP*** | ***JCSm*** | ***FCSm*** | ***JCS/FCSm*** | ***CPP/JCSm*** | ***CPP/FCSm*** |
|---|---|---|---|---|---|---|---|---|---|
| Biology & Biomedicine | 371 | 24 | 334 | 12.38 | 17.29 | 12.38 | 1.35 | 0.74 | 0.97 |
| Materials Science | 302 | 42 | 149 | 3.47 | 5.41 | 4.65 | 1.22 | 0.73 | 0.85 |
| Natural Resources | 304 | 23 | 116 | 4.97 | 5.40 | 5.64 | 1.00 | 0.84 | 0.86 |

*Note.* $N$ = number of researchers. All other figures (***P***, ***C***, etc.) represent the median of the distribution of the indicator values for all researchers per area.

researchers are represented in the figure on the left while researchers are grouped into four categories according to their production (scientific size, using the percentiles of ***P***: P1–P4) in the figure on the right. In Materials Science and Natural Resources, ***CPP*** tends to increase with ***P***, as observed by the positive power law exponent $\alpha$ as well as by the fact that high-productive researchers (P4) obtain a significant higher ***CPP*** than do low-productive researchers (P1). Interestingly, the increase of ***CPP*** as a function of ***P*** is higher for *low* field-citation-density researchers (higher power law exponent $\alpha$), which is consistent with previous results for research groups (van Raan, 2008a). However, in that previous study, a slight downward trend of ***CPP*** for increasing values of ***P*** was described for high field-citation-density regions.

In Figure 4, the same data as in Figure 3 are presented, but now we distinguish within the high/low field-citation-density researchers between top- and bottom-performance researchers (i.e., the top-50% and the bottom-50% of the ***CPP/FCSm*** distribution, respectively). According to this, four classes can be considered: Top-Top (*high* field-citation density and *top* performance), Top-Bottom (*high* field-citation density and *bottom* performance), Bottom-Top (*low* field-citation density and *top* performance), and Bottom-Bottom (*low* field-citation density and *bottom* performance). These data should be analyzed with caution since we have a low number of researchers in some classes. As a general result, observed in all three scientific areas, the Bottom-Bottom researchers benefit more from a larger number of publications, as indicated by their higher power law exponent and the fact that P4 researchers are more productive than are P1 researchers ($p < .05$ in Biology & Biomedicine and Materials Science).

Behavior of the field-citation density itself as a function of the number of publications for both the high as well as the low field-citation-density regions has been investigated. The results are shown in Figure 5, and similar properties to those previously described for research groups can be observed (van Raan, 2008a).


DOI: 10.1002/asi

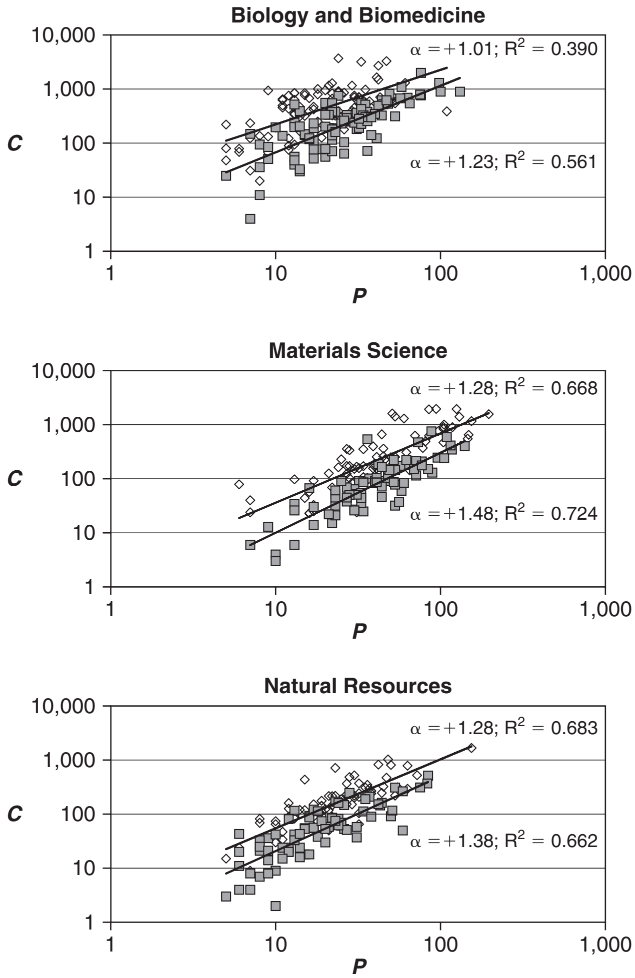


FIG. 2. Correlation of the number of citations (***C***) with the number of publications (***P***) for researchers in the top-25% (diamonds) and in the bottom-25% (squares) of the field-citation-density (***FCSm***) distribution.

For the *high* field-citation-density researchers, the ***FCSm*** tends to decrease very slightly or remains stable with increasing ***P***. Statistically significant results are found only in Biology & Biomedicine, in which the most productive researchers show a lower ***FCSm*** than do the least ones ($p < .01$). This means that for researchers operating in *high* field-citation-density regions in this area, a larger number of publications mostly implies extension towards regions with a somewhat lower field-citation density.

However, for the *low* field-citation-density researchers, there is a slight upward trend in ***FCSm*** as ***P*** increases. This finding is supported by the observation of a higher ***FCSm*** for the most productive researchers as compared to the least ones (significant differences in Biology & Biomedicine and in Materials Science). Thus, for researchers operating in the *low* field-citation-density regions, a larger number of publications appears to correspond with an "expansion" into regions with higher field-citation density. The difference between top- and lower performance researchers is shown in Figure 6. Clearly, particularly in the right-hand figures, there is no significant difference between top and low performance, as also was found in the case of groups (van Raan 2008a).

How does the average journal-citation impact of a researcher relate to the field-citation density? The answer is given by Figure 7. For the *low* field-citation-density researchers, a larger production (i.e., size) implies a somewhat higher average ***JCSm*** value in two areas. More specifically, the differences in ***JCSm*** of researchers in relation to size levels (P1–P4) are significant in the cases of Materials Science and Natural Resources ($p < .05$; figures on the right-hand side of Figure 7). Concerning high field-citation-density regions, a larger number of publications do not significantly change the average journal-citation impact, and even the trend is slightly negative in some cases. This means that "expanding in size" may take place within the same field-citation-density region, publishing in journals with the same or lower impact.

Figure 8 shows that the lower performers in *low* field-citation-density fields (Bottom-Bottom) are the ones who benefit the most: For them, a larger number of publications implies higher ***JCSm*** scores in the three areas analyzed. In fact, significant differences in ***JCSm*** by size classes were found for Bottom-Bottom researchers in Biology & Biomedicine and in Materials Science ($p < .05$ and $p < .01$, respectively).

Thus, for researchers operating in *low* citation-density regions, a larger number of publications can be seen as an "expansion" into regions with *higher* field-citation density, as we saw earlier, as well as an expansion towards journals with a higher average impact. The interrelation between field-citation density and journal impact, and its influence on the total number of citations of a researcher, also needs to be studied. We take the results presented in Figure 2 and make a breakdown for both the high (top-25% of the ***FCSm*** distribution) as well as the low (bottom-25% of ***FCSm***) into the higher (top-50% of the ***JCSm*** distribution) and the lower (bottom-50% of ***JCSm***) journal impact (see Figure 9).

Clearly, one can observe that researchers in low citation-density fields who publish in *low* impact journals (Bottom-Bottom) benefit the most from the increase in number of publications in the three areas (Power law exponent $\alpha$ is between $+1.42$ and $+1.51$.)

### *Analyzing the Observations in the Framework of the Science Landscape*

In a previous article (van Raan, 2008a), the size-dependence nature of several bibliometric indicators was analyzed for a large set of research groups. This work is relevant to the understanding of complex, networked systems, particularly the interdependencies of variables in both natural as well as artificial systems (Newman, 2005). One of the main observations was that groups in low density-citation fields have a size-dependent cumulative advantage for receiving citations, and they benefit most from an increase in the number of publications. Does this behavior also apply to a lower aggregation level; namely, individual researchers? Moreover, since differences between thematic areas can be expected,

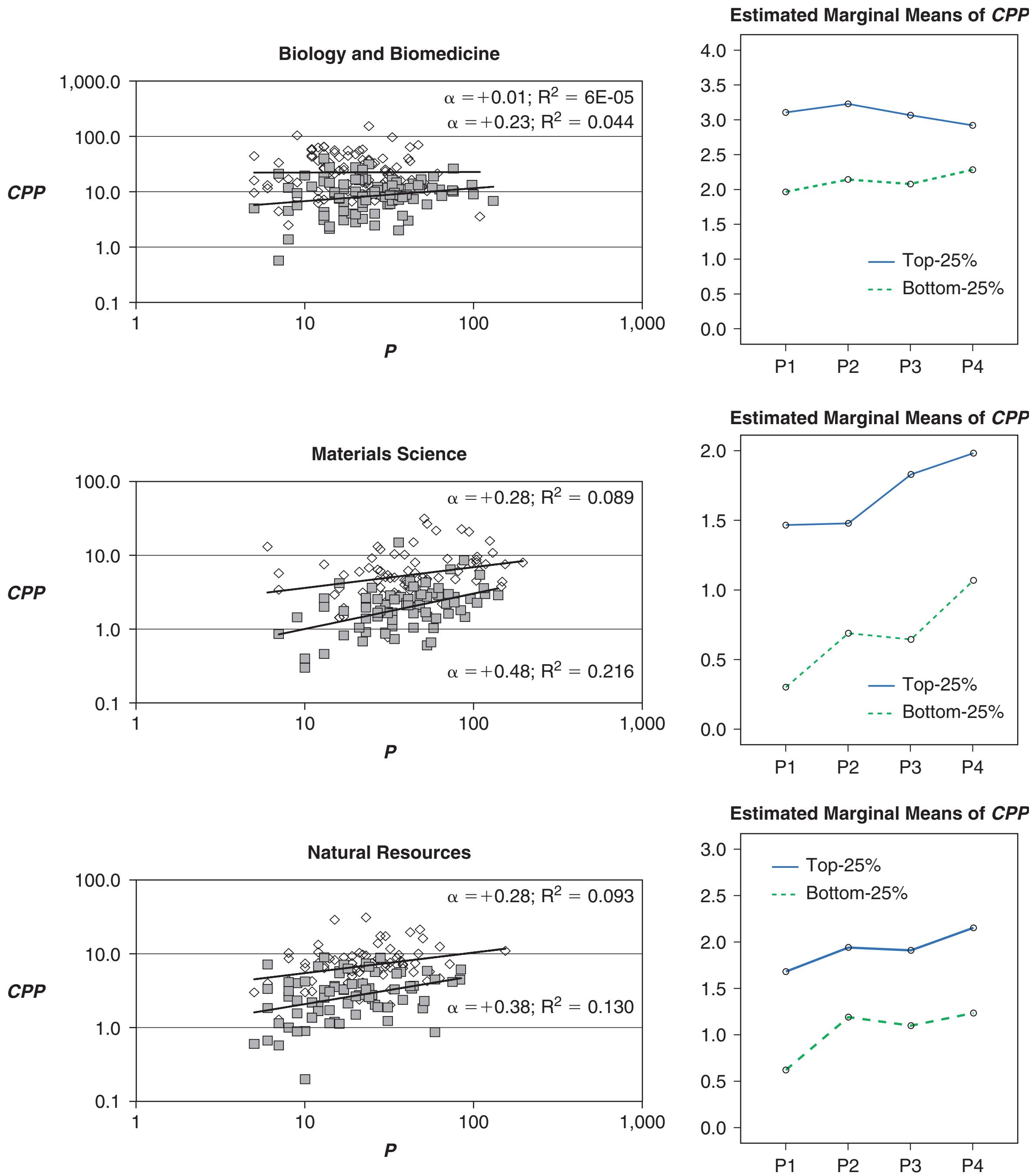


FIG. 3. Correlation of citations-per-publication (***CPP***) with the number of publications (***P***) for researchers in the top-25% (diamonds) and in the bottom-25% (squares) of the field-citation-density (***FCSm***) distribution. On the right side, we show the results of the percentile-based ANOVA analysis as discussed in the text, with the size (***P***) percentile-class on the horizontal axis and the natural logarithm (ln) of ***CPP*** on the vertical axis.

the performance of individual researchers in three different areas was studied in this article.

When comparing the present results with previous ones, several facts need to be considered. First, note that different thematic areas are analyzed—Chemistry in the previous study versus Biology & Biomedicine, Materials Science, and Natural Resources in this article—and even differences within these areas might exist. Thus, area-specific features in research performance should be taken into account in the understanding of the results. Second, differences in the organization of research in The Netherlands (previous article) and Spain (this article) also could have influenced the results. Concerning methodological issues, the fact that the same source data and indicators were used in both studies makes comparisons possible.

We are especially interested in the aggregation level-specific differences between groups (previous study) and individual researchers (this study). Some features of the study


DOI: 10.1002/asi

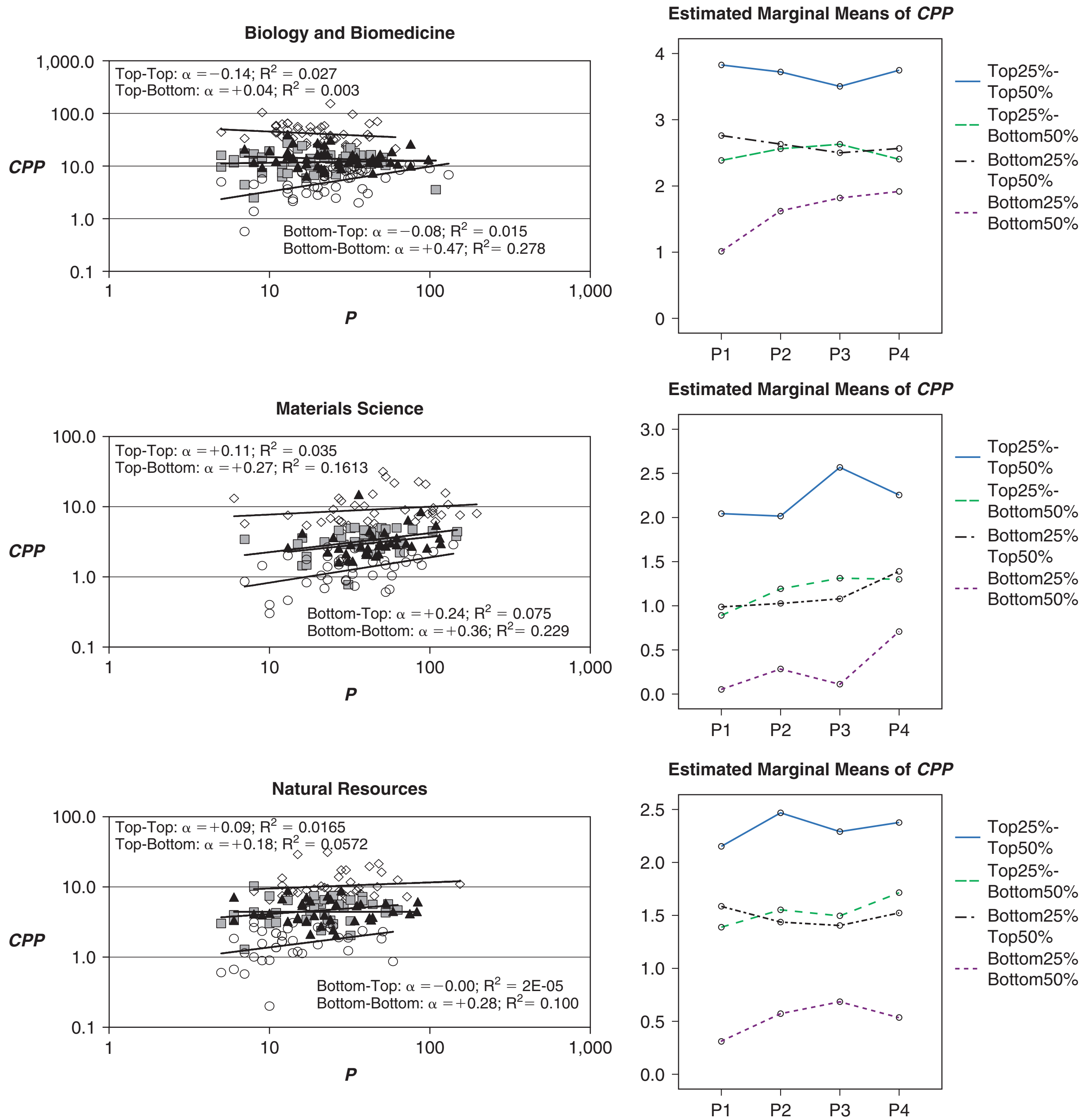


FIG. 4. Correlation of citations-per-publication (***CPP***) with number of publications (***P***) for high field-citation-density researchers (top-25% of ***FCSm***), divided in top-performance (top-50% of ***CPP/FCSm***, diamonds) and lower performance (bottom-50% of ***CPP/FCSm***, squares), and for low field-citation-density researchers (bottom-25% of ***FCSm***), again divided in top-performance (top-50% of ***CPP/FCSm***) (triangles) and lower performance (bottom-50% of ***CPP/FCSm***, circles). For an explanation of the figures on the right-hand side, refer to Figure 3.

at the individual level need to be taken into account, such as the higher variability of data and the fact that different researchers from the same group (probably sharing similar bibliometric features) are playing in the analysis, while at the aggregation level of groups, only one value per group appears in the analysis. Comparison of the statistical properties of the bibliometric performance indicator values of researchers and groups within the same population should be addressed in future work to answer these questions.

Despite these limitations, it is fascinating to see that some of the patterns previously observed at the group level also emerge at the level of the individual researcher. We show in this article that the total number of citations received by researchers increases in a cumulatively advantageous way as a function of number of publications, with a higher benefit for researchers publishing in fields of *low* citation density. The number of ***CPP*** also tends to increase with the number of ***P*** in two areas, and this increment is higher for *low*


DOI: 10.1002/asi

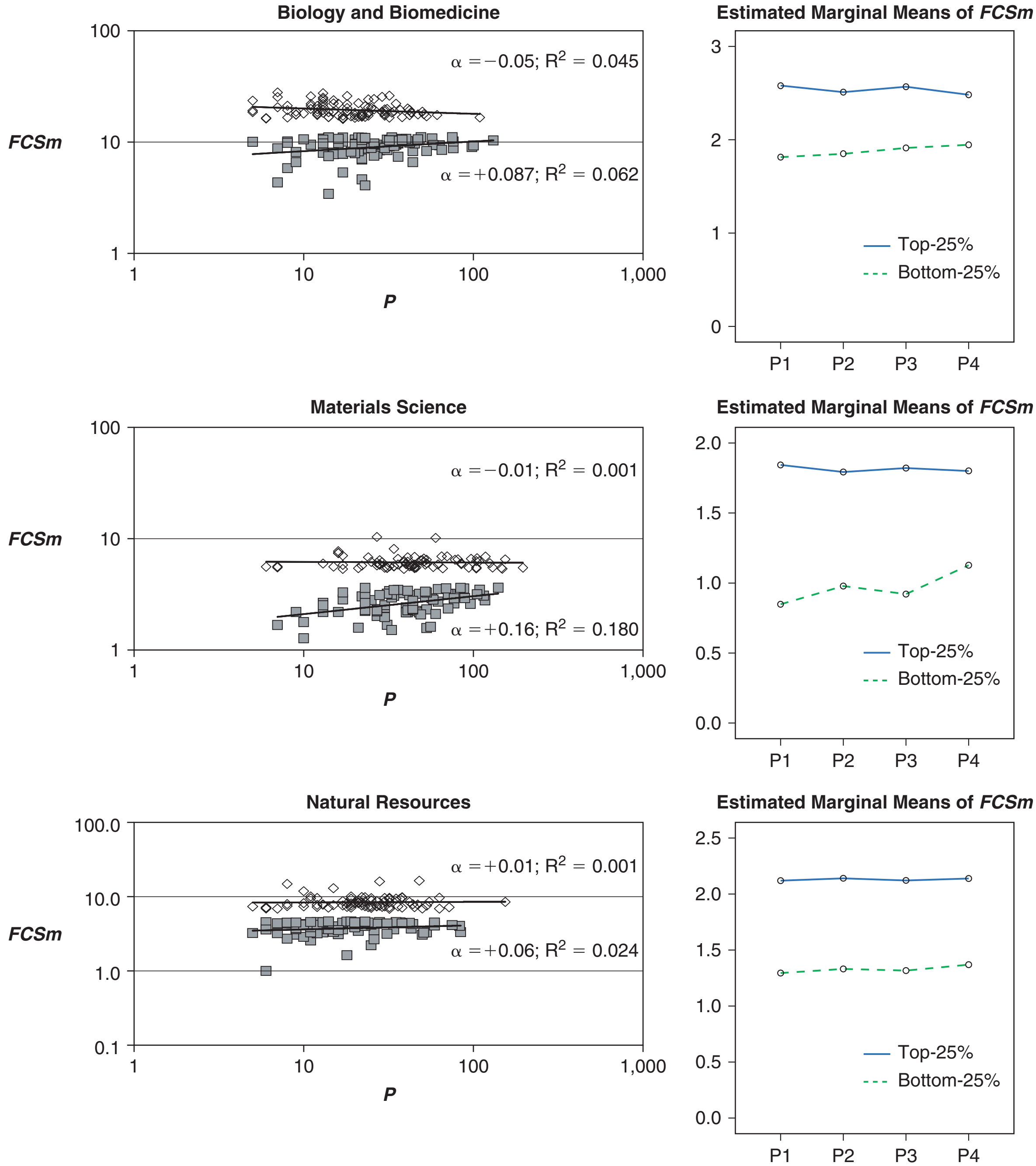


FIG. 5. Correlation of field citation density (***FCSm***) with size (***P***) for researchers in fields with a high (diamonds) and a low (squares) field citation density (***FCSm***). On the right side, we show the results of the percentile-based ANOVA analysis, as discussed in the text, with the size (***P***) percentile-class on the horizontal axis and the natural logarithm (ln) of ***FCSm*** on the vertical axis.

field-citation-density researchers, but there is no cumulative advantage (power law exponent $<1$).

As the production (i.e., size) of researchers increases, a trend to publish in higher impact journals (***JCSm***) is more significant for *low* field-citation-density researchers as compared to *high* field-citation-density researchers. The trend to publish in fields with a higher density of citations (***FCSm***) as a function of the increasing number of publications is weak and was only found for researchers in *low* citation-density fields. Thus, for researchers in the latter fields, a larger number of publications implies a higher probability of expansion into higher citation densities. These trends are consistent with results shown at the research-group level (van Raan, 2008a), although specific differences in the power law exponent depending on the area can be observed.

According to our results, Materials Science researchers working in *low* citation-density fields have a stronger advantage with size since they show the highest power law exponent in the correlation of the different indicators


DOI: 10.1002/asi

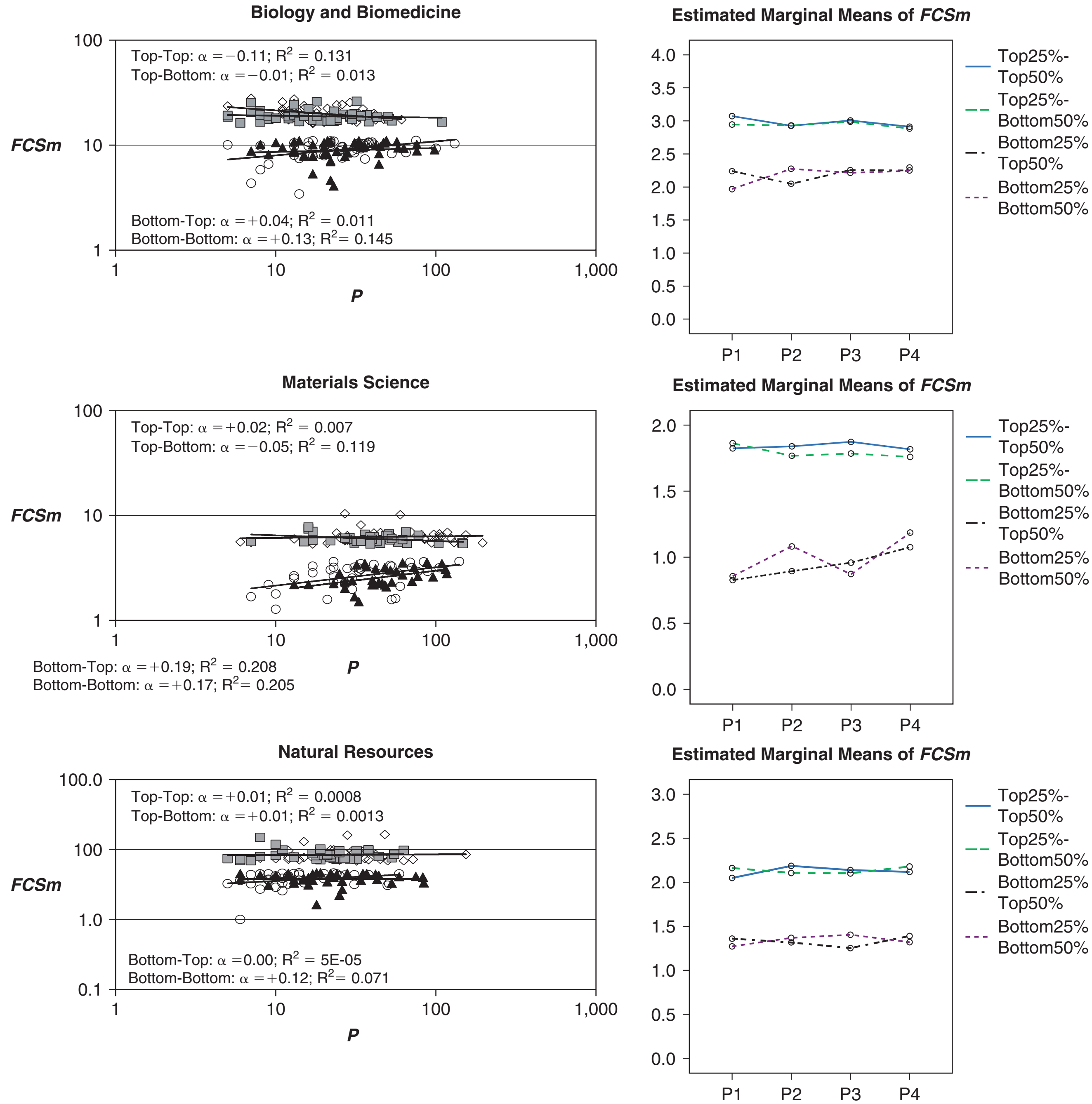


FIG. 6. Correlation of field citation density (***FCSm***) with number of publications (***P***) for high field-citation-density researchers (top-25% of ***FCSm***), divided in top-performance (top-50% of ***CPP/FCSm***, diamonds) and lower performance (bottom-50% of ***CPP/FCSm***, squares), and for low field-citation-density researchers (bottom-25% of ***FCSm***), again divided in top-performance (top-50% of ***CPP/FCSm***) (triangles) and lower performance (bottom-50% of ***CPP/FCSm***, circles). For an explanation of the figures on the right-hand side, refer to Figure 5.

(***C***, ***CPP***, ***FCSm***, ***JCSm***) with ***P***, and the most productive scientists always present significantly higher values than do the less productive (see Appendix, Table A2). The other side of the spectrum is Biology & Biomedicine, which shows for most of the indicators the lowest exponents, and no significant differences in ***CPP*** and ***JCSm*** by size classes were observed. For Materials Science researchers active in *low* citation-density regions, citations and ***CPP*** tend to increase with the number of publications, and they tend to publish in higher impact factor journals and even expand to higher field-citation-density regions with increasing production. For Natural Resources researchers working in *low* citation-density regions, citations and ***CPP*** also tend to increase with ***P***, and better journals within the same field are used as scientific production increases. However, in Biology & Biomedicine, only researchers working in *low* citation-density regions and publishing in below-average impact journals show an increase in ***CPP*** for larger ***P***. The


DOI: 10.1002/asi

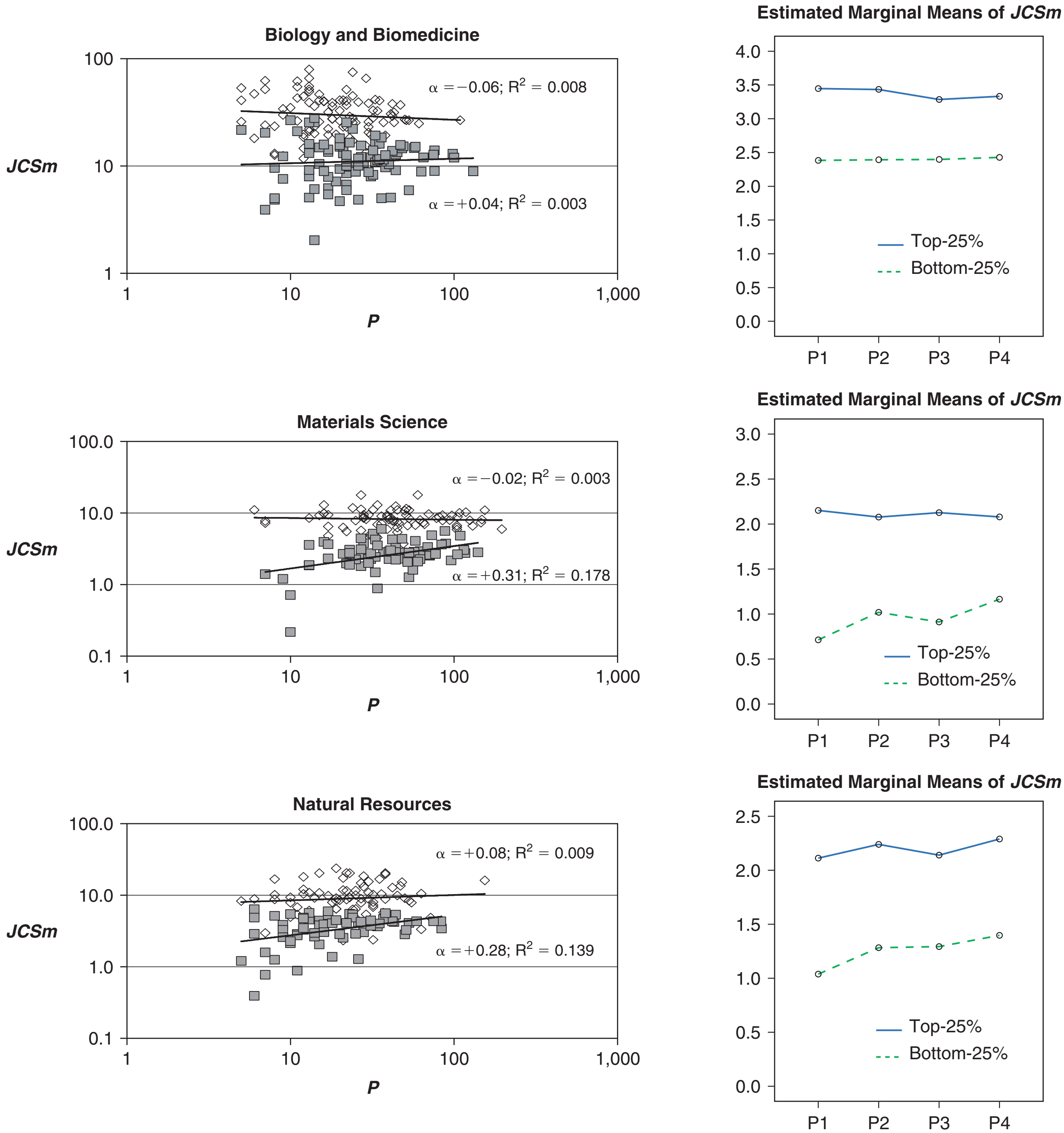


FIG. 7. Correlation of journal impact (***JCSm***) with size (***P***) for researchers in fields with a high (diamonds) and a low (squares) field citation density (***FCSm***). On the right side, we show the results of the percentile-based ANOVA analysis as discussed in the text with the size (***P***) percentile-class on the horizontal axis and the natural logarithm (ln) of ***JCSm*** on the vertical axis.

fact that Biology & Biomedicine shows the highest density of citations and ***CPP/FCSm*** score (Table 1) might explain its peculiarities. It is a very competitive area, and researchers are oriented towards high-impact journals within their subfields.

In summary, we show that researchers in *low* citation-density fields benefit most from increasing number of publications, as previously observed at the level of research groups; however, the difference in advantage between *low* and *high* field-citation-density researchers is smaller here than in the study of groups. In fact, a negative effect of the number of publications on ***CPP*** or ***JCS*** was observed at the level of groups in high field-citation-density regions (van Raan, 2008a), so "expanding in size" could be counterproductive in ***CPP*** scores for these groups. For individual researchers, a slight negative trend was observed in ***JCSm*** in two areas, but not in ***CPP***. Since CSIC researchers have, on average, impact scores below the international level (***CPP/FCSm*** $<1$) (Table 1), we hypothesize that these researchers still have "room" for improvement (especially in Natural Resources and in Materials Science). The fact that


DOI: 10.1002/asi

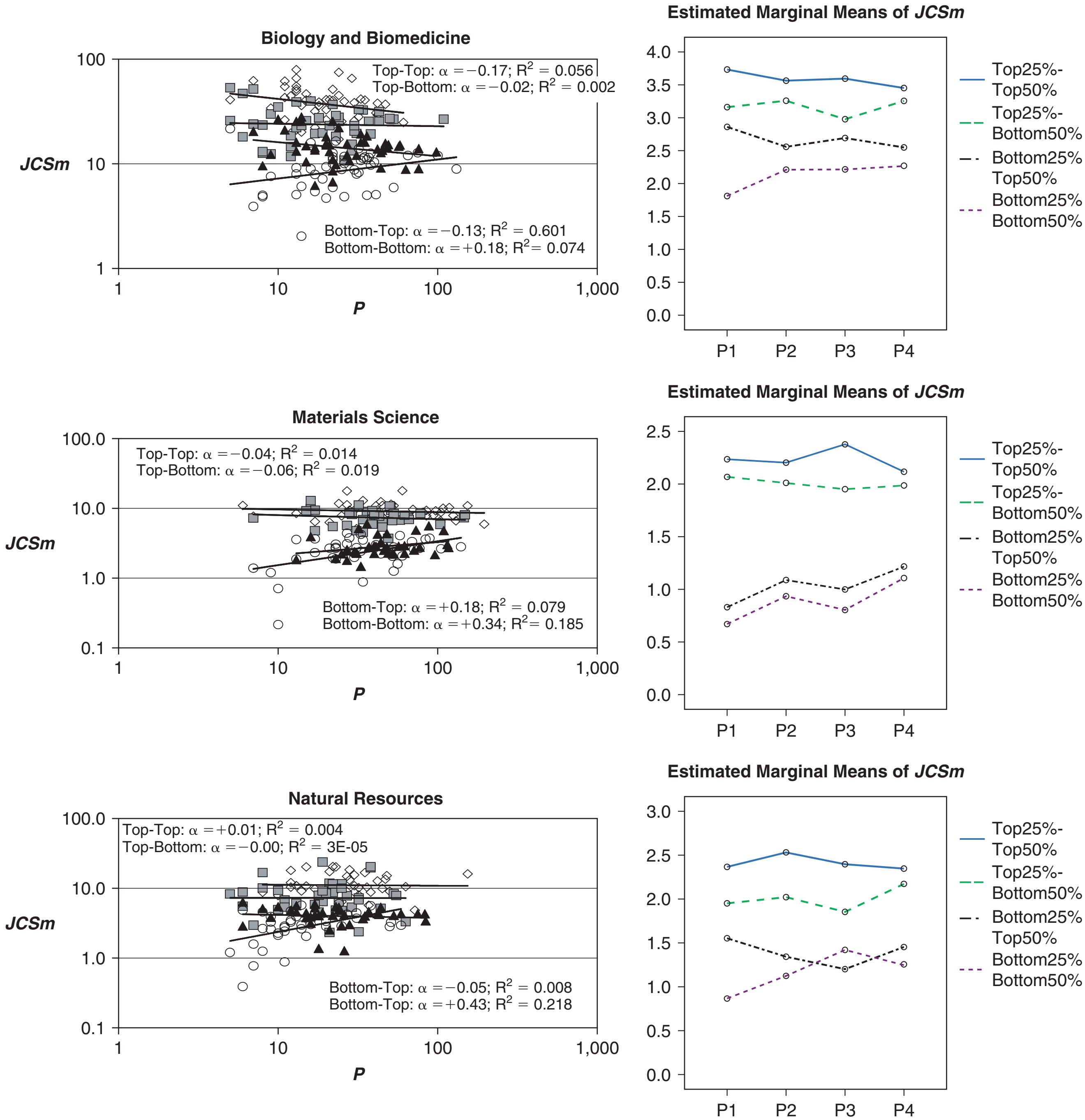


FIG. 8. Correlation of field citation density (***JCSm***) with number of publications (***P***) for high field-citation-density researchers (top-25% of ***FCSm***), divided in top-performance (top-50% of ***CPP/FCSm***, diamonds) and lower performance (bottom-50% of ***CPP/FCSm***, squares), and for low field-citation-density researchers (bottom-25% of ***FCSm***), again divided in top-performance (top-50% of ***CPP/FCSm***) (triangles) and lower performance (bottom-50% of ***CPP/FCSm***, circles). For an explanation of the figures on the right-hand side, refer to Figure 7.

Biology & Biomedicine researchers are working in higher field-citation-density regions (***FCSm*** = 12.38 vs. values <5 for the other two areas) but manage to publish in high-impact journals within the field (***JCS/FCSm*** = 1.35) suggests that for them, improvement is more difficult. In fact, the hypothesis that researchers with ***CPP/FCSm*** >1 are less likely to benefit in their ***CPP*** when increasing their number of publications is supported by data in the Appendix, Table A3.

Our results show the existence of a size-dependent cumulative advantage for receiving citations, which has been previously described at the country, institution, and group levels, and also was observed at the individual level in this study. Researchers in low field-citation-density regions and those whose impact is below world class tend to benefit the most from an increase in number of publications. Interarea differences can be explained by different factors such as the


DOI: 10.1002/asi

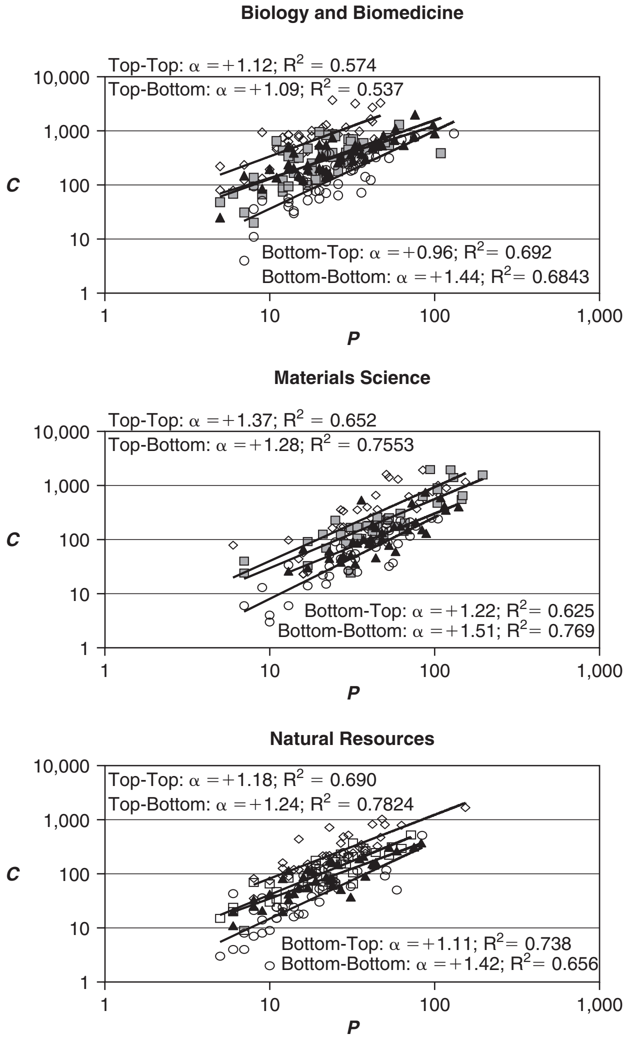


FIG. 9. Correlation the number of citations (***C***) with the number of publications (***P***) for high field-citation-density researchers (top-25% of ***FCSm***), divided in high journal impact (top-50% of ***JCSm***, diamonds) and low journal impact (bottom-50% ***JCSm***, squares), and for low field-citation-density researchers (bottom-25% of ***FCSm***), again divided in high journal impact (top-50% of ***JCSm***) (triangles) and low journal impact (bottom-50% ***JCSm***, circles).

field citation density, the level of development of the area in the country of analysis, and the distribution of high/low performance of researchers as compared to the world average. These results can be useful for policy makers and research managers as well as for researchers themselves. For scientists with high performance in high-citation-density fields, it is increasingly difficult to maintain good scores in impact as productivity increases, an issue that also was suggested by Egghe and Rousseau (1990); so for them, it is especially important to select appropriate publication strategies focusing on quality rather than on quantity of publications. This article focuses on scaling relationships between major bibliometric indicators in three different areas to gain insight into the dynamics of the research process at the individual level. Further research is needed to analyze whether the trends described in this article also occur in other fields. Moreover, the study of the influence of size on bibliometric indicators is important for an adequate interpretation of these indicators in research-performance assessment.

## Acknowledgments

This work was possible because of an I3P fellowship from the CSIC and a research stay of Rodrigo Costas at the CWTS (Leiden University). We also thank Laura Barrios (CSIC) for her advice on statistical issues as well as the referees whose comments substantially improved this article.

## References


Aksnes, D.W., & Taxt, R.E. (2004). Peer reviews and bibliometric indicators: A comparative study at a Norwegian university. Research Evaluation, 13(1), 33–41.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. Reviews of Modern Physics, 74, 47–97.

Caldarelli, G., Erzan, A., & Vespignani, A. (Eds.). (2004). Virtual round table on ten leading questions for network research. European Physical Journal, B38, 143–145.

Costas, R., & Bordons, M. (2005). Bibliometric indicators at the micro-level: Some results in the area of natural resources at the Spanish CSIC. Research Evaluation, 14(2), 110–120.

Dorogovtsev, S.N., & Mendes, J.F.F. (2002). Evolution of networks. Advances in Physics, 51, 1079–1187.

Egghe, L., & Rousseau, R. (1990). Introduction to informetrics: Quantitative methods in library, documentation and information science. Amsterdam: Elsevier.

Katz, J.S. (1999). The self-similar science system. Research Policy, 28, 501–517.

Katz, J.S. (2000). Scale independent indicators and research assessment. Science and Public Policy, 27(1), 23–26.

Katz, J.S. (2005). Scale-independent bibliometric indicators. Measurement, 3(1), 24–28.

Leicht, E.A., Clarkson, G., Shedden, K., & Newman, M.E.J. (2007). Large-scale structure of time evolving citation networks. European Physical Journal B, 59, 75–83.

Merton, R.K. (1988). The Matthew Effect in science: II. Cumulative advantage and the symbolism of intellectual property. Isis, 79, 606–623.

Moed, H.F., Burger, W.J.M., Frankfort, J.G., & van Raan, A.F.J. (1985). The use of bibliometric data for the measurement of university research performance. Research Policy, 14, 131–149.

Newman, M.E.J. (2005). Power laws, Pareto distributions and Zipf's law. Contemporary Physics, 46, 323–351.

Price, D. de Solla. (1976). A general theory of bibliometric and other cumulative advantage processes. Journal of the American Society for Information Science, 27(5), 292–306.

Rinia, E.J., van Leeuwen, Th.N., van Vuren, H.G., & van Raan, A.F.J. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria: Evaluation of condensed matter physics in The Netherlands. Research Policy, 27, 95–107.

Rinia, E.J., van Leeuwen, Th.N., van Vuren, H.G., & van Raan, A.F.J. (2001). Influence of interdisciplinarity on peer-review and bibliometric evaluations. Research Policy, 30, 357–361.

Schubert, A., & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. Scientometrics, 9(5–6), 281–291.

van Raan, A.F.J. (2004). Measuring science. Capita selecta of current main issues. In H.F. Moed, W. Glänzel, and U. Schmoch (Eds.), Handbook of quantitative science and technology research (pp. 19–50). Dordrecht, The Netherlands: Kluwer.

van Raan, A.F.J. (2006a). Statistical properties of bibliometric indicators: Research group indicator distributions and correlations. Journal of



DOI: 10.1002/asi

the American Society for Information Science and Technology, 57(3), 408–430.

van Raan, A.F.J. (2006b). Performance-related differences of bibliometric statistical properties of research researchers: Cumulative advantages and hierarchically layered networks. Journal of the American Society for Information Science and Technology, 57(14), 1919–1935.

van Raan, A.F.J. (2008a). Scaling rules in the science system: Influence of field-specific citation characteristics on the impact of research groups. Journal of the American Society for Information Science and Technology, 59(4), 565–576.

van Raan, A.F.J. (2008b). Self-citation as an impact-reinforcing mechanism in the science system. Journal of the American Society for Information Science and Technology, 59(10), 1631–1643.

Vinkler, P. (1986). Evaluation of some methods for the relative assessment of scientific publications. Scientometrics, 10(3–4), 157–177.

## Appendix

TABLE A1. Size classes (no. of publications/researcher) by area.

| | Biology & Biomedicine | Materials Science | Natural Resources |
|---|---|---|---|
| P1 | 5–15 | 5–27 | 5–14 |
| P2 | 16–24 | 28–42 | 15–23 |
| P3 | 26–36 | 43–62 | 23–34 |
| P4 | >36 | >62 | >34 |

TABLE A2. Statistical data on the analysis of the size-dependence advantage of different bibliometric indicators: (a) correlation between bibliometric indicators and number of publications and (b) comparison of means by size classes.

| | | | Correlation with ***P*** | ANOVA by size classes | |
|---|---|---|---|---|---|
| Indicator | Area | ***FCSm*** class | $\alpha$ | $p$ | Significance: post hoc test |
| ***C*** | All | Top | +0.98 | .000 | All significant |
| | | Bottom | +1.32 | .000 | All significant |
| | Biology & Biomedicine | Top | +1.01 | .000 | All significant |
| | | Bottom | +1.23 | .000 | All significant |
| | Material Science | Top | +1.28 | .000 | All significant |
| | | Bottom | +1.48 | .000 | All significant |
| | Natural Resources | Top | +1.28 | .000 | All significant |
| | | Bottom | +1.38 | .000 | All significant |
| ***CPP*** | All | Top | −0.02 | n.s. | – |
| | | Bottom | +0.32 | .000 | P1 vs. P2, P3,P4 |
| | Biology & Biomedicine | Top | +0.01 | n.s. | – |
| | | Bottom | +0.23 | n.s. | – |
| | Material Science | Top | +0.28 | .05 | P4 vs. P1, P2 |
| | | Bottom | +0.48 | .01 | P1 vs. P2, P4 |
| | Natural Resources | Top | +0.28 | .01 | P1 vs. P4 |
| | | Bottom | +0.38 | .01 | P1 vs. P2, P3, P4 |
| ***FCSm*** | All | Top | −0.05 | .000 | P1 vs. P2, P3, P4 |
| | | Bottom | +0.02 | n.s. | – |
| | Biology & Biomedicine | Top | −0.05 | .01 | P1 vs. P2, P4 |
| | | Bottom | +0.09 | .05 | P1 vs. P4 |
| | Material Science | Top | −0.01 | n.s. | – |
| | | Bottom | +0.16 | .01 | P4 vs. P1, P3 |
| | Natural Resources | Top | +0.01 | n.s. | – |
| | | Bottom | +0.06 | n.s. | – |
| ***JCSm*** | All | Top | −0.17 | .000 | P1 vs P2, P3, P4; P2 vs P4 |
| | | Bottom | +0.18 | .001 | P1 vs. P2, P3, P4 |
| | Biology & Biomedicine | Top | −0.06 | n.s. | – |
| | | Bottom | +0.04 | n.s. | – |
| | Material Science | Top | −0.02 | n.s. | – |
| | | Bottom | +0.31 | .05 | P1 vs. P2, P4 |
| | Natural Resources | Top | +0.08 | n.s. | – |
| | | Bottom | +0.28 | .05 | P1 vs. P4 |

***C*** = number of citations; ***CPP*** = citation per publication; ***FCSm*** = field citation score ($M$); ***JCSm*** = journal citation score ($M$). "Top" refers to researchers in the top-25% of the ***FCSm*** distribution. "Bottom" refers to researchers in the bottom-25% of the ***FCSm*** distribution.


DOI: 10.1002/asi

TABLE A3. ***CPP*** by the Crown Indicator (***CPP/FCSm***) classification.

| Scientific field | Crown Indicator | Size class | *N* | *M* ***CPP*** | *p* |
|---|---|---|---|---|---|
| Biology & Biomedicine | ***CPP/FCSm*** > 1 | P1 | 46 | 3.33 | .001 |
| | | P2 | 49 | 3.15 | |
| | | P3 | 44 | 2.96 | |
| | | P4 | 41 | 2.93 | |
| | | Total | 180 | 3.1 | |
| | ***CPP/FCSm*** < 1 | P1 | 48 | 1.76 | .001 |
| | | P2 | 50 | 1.97 | |
| | | P3 | 43 | 2.06 | |
| | | P4 | 50 | 2.16 | |
| | | Total | 191 | 1.99 | |
| Materials Science | ***CPP/FCSm*** > 1 | P1 | 27 | 1.81 | n.s. |
| | | P2 | 18 | 1.89 | |
| | | P3 | 30 | 2.06 | |
| | | P4 | 39 | 1.95 | |
| | | Total | 114 | 1.93 | |
| | ***CPP/FCSm*** < 1 | P1 | 51 | 0.7 | .000 |
| | | P2 | 61 | 0.95 | |
| | | P3 | 45 | 0.91 | |
| | | P4 | 31 | 1.08 | |
| | | Total | 188 | 0.89 | |
| Natural Resources | ***CPP/FCSm*** > 1 | P1 | 20 | 2.02 | n.s. |
| | | P2 | 30 | 2.04 | |
| | | P3 | 26 | 2.1 | |
| | | P4 | 33 | 2.03 | |
| | | Total | 109 | 2.05 | |
| | ***CPP/FCSm*** < 1 | P1 | 62 | 0.85 | .000 |
| | | P2 | 52 | 1.26 | |
| | | P3 | 39 | 1.31 | |
| | | P4 | 42 | 1.42 | |
| | | Total | 195 | 1.18 | |

For researchers with a ***CPP/FCSm*** < 1, ***CPP*** tends to increase with size (***P***), so "expanding in size" means higher ***CPP*** scores (differences statistically significant in all scientific fields). For researchers with ***CPP/FCSm*** > 1, no statistical differences in ***CPP*** by size class were found, except in Biology & Biomedicine, where ***CPP*** tends to decrease for the most productive researchers.


DOI: 10.1002/asi