Local Bibliometric Indicators and the International Mobility of Researchers

prof. dr. Elizabeth Amann
Ghent University

Abstract

In recent years, a number of countries and regions have developed local bibliometric indicators to take into account research not included in standard international metrics. This paper argues that these initiatives could have a negative impact on the international mobility of researchers. A comparison of the indicators of five countries/regions demonstrates very low levels of overlap among them. As a result of these discrepancies, researchers moving from one academic system to another may find that their publications are devalued or not recognized in the new context. This paper advocates the creation of a standard transnational classification of publication venues as well as a common system of weighting research (i.e. determining the relative value of books and articles and of single-authored versus co-written publications).

In recent years, a growing body of studies has examined the problem of 'academic inbreeding' and the obstacles to the international mobility of researchers. Although there is currently no consensus about the causes of inbreeding, a number of factors have been identified as contributing to immobility in academia. These include: 1. language barriers; 2. uncompetitive salaries; 3. interpersonal relationships (bonds of loyalty among faculty members, which lead them to privilege internal over foreign candidates); 4. local traditions (and a reluctance to hire those unfamiliar with them); 5. the privileging of teaching over research needs in evaluating candidates; 6. selection committees made up entirely of internal or local professors; 7. the insufficient dissemination of job announcements and 8. differences in the structures of academic careers (for example, privileging or requiring local qualifications such as an *agrégation* or *Habilitation*).[1] Less attention, however, has been given to the role that selection criteria, which often vary considerably from one academic context to another, play in hiring decisions.

In the past two decades, quantitative measures (e.g. the number of Web of Science articles, the impact factor of a journal or the citation rate of a publication) have become increasingly important in hiring and promotion decisions. In some instances, this shift toward numerical indicators may favour internationalization by reducing the weight given to more subjective criteria, which are more easily manipulated in favour of local candidates. Nevertheless, it would be a mistake to conclude that this new culture of counting transcends borders. For the reality of modern academia is that people count in different ways in different places. Committees vary in the weight they give to different types of publications (the relative value of a monograph, a journal article, a book chapter and a proceedings paper), in the way they assess publication venues and in the way they assign points to co-publications (the relative value of co-written and single-authored articles). These differences in academic cultures shape not only the ways in which researchers evaluate one another but also the ways in which they conduct their careers and tye type of scholarship they produce.

This essay will draw attention to a development that illustrates the discrepancies among the 'cultures of counting' of different academic systems and that may in some instances have a negative impact on international mobility: the use of local bibliometric indicators. In recent years, a number of countries and regions have generated their own lists of publishing venues in order to take into account forms of research that are not registered in standard international metrics. Examples of these local initiatives include the Scholarly Publishers Indicators (SPI) in Spain, the Flemish Academic Bibliographic Database (VABB-

SHW), the Bibliometric Research Indicator (BRI) in Denmark, the JUFO Publication Forum in Finland, the Current Research Information System (CRISTIN) in Norway, the AERES list of humanities and social sciences journals in France and the ANVUR classification of journals in Italy.

These local initiatives are generally well-intentioned and thoughtfully constructed. Often they attempt to correct omissions or perceptions of bias in standard international indicators. Some are designed to account for book monographs and collective volumes, which are not included in the Web of Science. Some attempt to give more space to disciplines underrepresented in other lists such as the Social Sciences and Humanities (SSH). Often they seek to remedy the perceived overrepresentation of Anglophone venues in international indicators and to make room for publications in other languages or from other places.

Nevertheless, as I will argue in this paper, the use of these local indicators is problematic for two reasons: (1) the relatively low level of overlap among them and (2) their potential negative impact on the international mobility of researchers. The first section of this paper will compare and contrast the rankings of academic book publishers in five regions in order to assess the degree of overlap or discrepancy among them. The second section will consider the uses to which these measures are put in different countries and their possible impact on the experience of internationally mobile researchers. Finally, the third section will consider the advantages and disadvantages of this type of indicator and will offer a series of policy recommendations.

1. Rates of overlap and discrepancy among local indicators:

This study compares and contrasts the top-level book publishers in the local indicators of five countries or regions: Denmark, Finland, Norway, Flanders and Spain. I have chosen to focus in this paper on classifications of book publishers (rather than journals) because of the high stakes involved in these rankings.[2] An academic monograph typically takes a number of years to write, edit and pass through peer review and thus represents a considerable time investment for the author(s).[3] Criteria that lead to the recognition or non-recognition of a book publication, therefore, have a potentially large impact on the evaluation of a researcher.

The five lists under consideration are organized in different ways. The Danish (BRI) and Norwegian (CRISTIN) indicators divide publishers into two categories: Level 1, the most common designation, and Level 2, which represents the leading publishers and may not include more than 20% of the total production worldwide in the field. The Danish list currently ranks 89 presses as Level 2, and the Norwegian list includes 87.[4] The Finnish list (JUFO) distinguishes among three levels, defining Level 1 as 'basic', Level 2 as 'leading' and Level 3 as 'top'. Level 2, which is limited to 10% of book publishers, includes 93 presses, and Level 3, which is limited to 25% of Level 2, includes 15 presses. For research assessment purposes, however, no distinction is made between Levels 2 and 3 (book publications on these tiers are granted the same number of points).[5] Unlike the Scandinavian lists, the Spanish SPI does not distinguish between levels but rather ranks presses in two separate lists, one for national and the other for foreign publishers (Cambridge University Press is ranked number 1; Oxford, number 2; Routledge, number 3; and so on).[6] Finally, the Flemish VABB-SHW database features a list of 135 publishers whose publications are automatically recognized.[7] Flemish academics, however, may present books published by other presses for consideration; generally these works are included in the database if the authors can prove that the work has passed through an adequate peer-review process. There is no distinction among levels in this indicator.[8] Both the Spanish and Flemish list are specifically for the Humanities and Social Sciences. The Nordic lists, include non-SSH publishers, although relatively few.

As is clear from this overview, the five indicators differ considerably both in size and methodology. The comparison is also complicated by the volatility and complexity of the publishing world (some indicators list imprints or subsidiary companies separately whereas others do not).[9] It is important, however, to make this comparison because researchers who move from one system to another are potentially subject to multiple indicators.

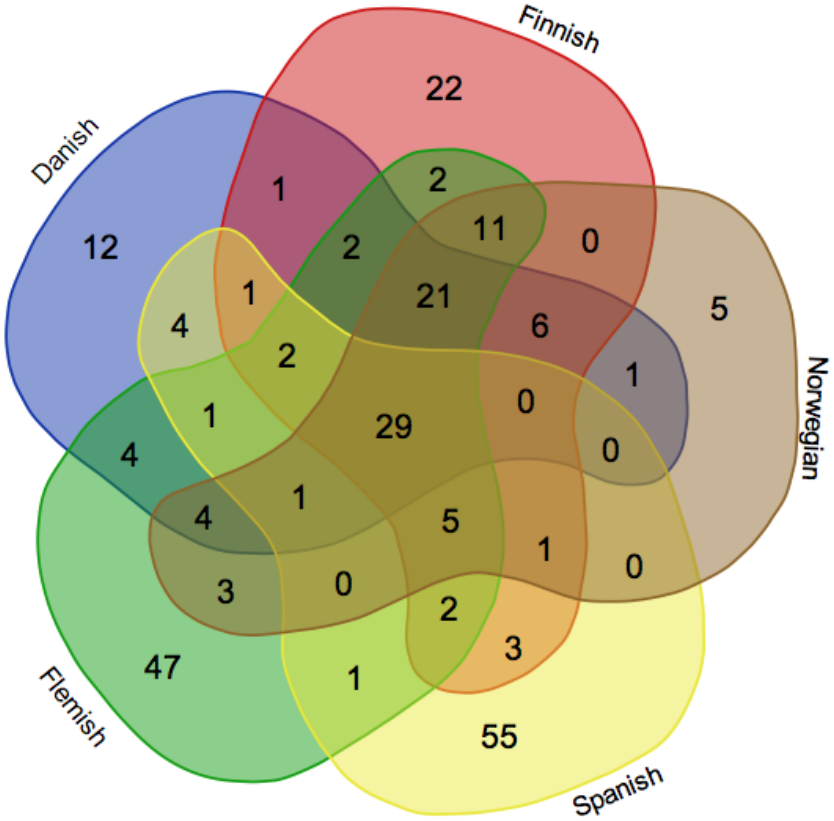
Previous attempts to calculate the overlap among these sorts of lists have generally worked with percentages. This approach, however, is problematic when the lists differ in size.[10] The discussion below, therefore, will use the Jaccard similarity coefficient, a formula for measuring the similarity of two sets. It is calculated by dividing the number of overlapping items by the total number of overlapping and non-overlapping items on the two lists. The Jaccard index can range from zero (when the two sets have no overlap) to one (when the two sets overlap completely).

The Finnish list is unique in that it distinguishes an extremely limited tier of élite publishers (Level 3). As this category does not exist in the Danish, Finnish and Flemish lists, the only possible comparison is with the top fifteen foreign presses on the Spanish SPI ranking. Only 6 presses coincide. The rate of overlap is thus 40%, and the Jaccard similarity coefficient is .25. Moreover, three of the presses on the Spanish list (Peter Lang, the Fondo de Cultura Económica and Thomas Reuters) do not appear among the leading presses of any of the other indicators (they are not among the recognized publishers in Flanders or on the Level 2 lists of the Nordic countries).

The comparison becomes more comprehensive when it is expanded to include the next tier: the Level 2 lists in the Danish and Norwegian indicators, the combined Level 2 and 3 lists in Finland and the automatically recognized presses of the Flemish VABB-SHW. As the Spanish SPI is based on a ranking rather than a differentiation of levels, the average length of the other four lists (105) was used to determine the number of publishers to be included from this ranking.

Taken together, these five lists consist of 246 different publishers. The overlap among the lists is represented in the Venn diagram in Chart 1. Of the 246 publishers, 29 appear on all five lists; complete overlap thus exists for 33% of the Danish list, 27% of the Finnish list, 33% of the Norwegian list, 28% of the Spanish list and 21% of the Flemish list. 58 publications appear on at least four of the five lists. These publications account for 60% of the Danish list, 53% of the Finnish list, 64% of the Norwegian list, 35% of the Spanish list and 43% of the Flemish list. Troublingly, however, more than half of the 246 presses—141 publishers—appear on only one of the five lists. These outliers made up 13% of the Danish list, 20% of the Finnish list, 6% of the Norwegian list, 35% of the Flemish list and 52% of the Spanish list.

Chart 1



As is suggested by this last statistic, the greatest level of divergence occurs with the Spanish list, which has the lowest Jaccard similarity coefficients (from .21 to .25) (see Chart 2). The highest Jaccard coefficients occur among the Nordic lists (Danish-Finnish: .46; Danish-Norwegian: .54; Norwegian-Finnish: .60). The Flemish list falls somewhere between the Spanish and Scandinavian ones. Its Jaccard coefficients with the Nordic lists range from .40 to .50. The lowest rate of overlap of all the possible combinations of lists is between the Flemish and the Spanish list (Jaccard index of .21).[11]

Chart 2

	Jaccard similarity coefficient
Danish Level 2 + Finnish Level 2/3	.46
Danish Level 2 + Norwegian Level 2	.54
Danish Level 2 + Spanish top 105	.24
Danish Level 2 + Flemish recognized presses	.40
Finnish Level 2/3 + Norwegian Level 2	.60
Finnish Level 2/3 + Spanish top 105	.25
Finnish Level 2/3 + Flemish recognized presses	.44
Norwegian Level 2 + Spanish top 105	.23
Norwegian Level 2 + Flemish recognized presses	.50
Spanish top 105 + Flemish recognized presses	.21

The discrepancies among the lists may be attributed to several factors, the most evident being the different methods and criteria used to create them. The Norwegian, Danish and Finnish lists were created by committees of experts for each discipline who were charged with distinguishing among several levels of quality. The Danish and Finnish lists were based on the Norwegian one, which may account in part for the high degree of overlap among the Nordic indicators.[12]

The initial version of the Flemish list was also based on the Norwegian one but it has since been expanded to include many new publishers. It is overseen by a single committee, which focuses primarily on the existence and rigor of the peer review process and whether the press is considered international. Another label (GPRC) is granted to Flemish publishers that meet the VABB peer review criteria. In many ways, this is the most objective of the five lists but it is also the least selective. The fact that the indicator is limited to the social sciences and humanities may also account for some of the discrepancy between the Flemish and the Finnish list, which includes about ten presses that focus on the sciences.

The Spanish SPI, which claims to measure the 'perceived prestige' of a venue, is based not on discussions—the consensus of a panel—but rather on surveys of Spanish researchers who were asked to identify the ten best publishers in their field. This methodology is perhaps the most questionable, for Spanish academics, who are evaluated by these rankings, have a strong incentive to list the venues in which they themselves publish.[13] It is striking, for example, that Peter Lang, generally considered a second-tier press elsewhere, is ranked sixth in the Spanish ranking above publishers such as Harvard, Yale or Princeton. It is also noteworthy that some of the publishers that appear in the top 15 of the Spanish list but not on the Finnish Level 3 are presses that are open to publishing in Spanish: John Benjamins, the Fondo de Cultura Económica, Peter Lang and Wolters Kluwer. The fact that these surveys are filled out privately and anonymously and not subject to committee discussion may make them less reliable. The Spanish list, however, does something that none of the others does: in addition to a general ranking, it generates separate lists for each discipline.

The divergences among the lists may also be attributed to differences in academic cultures. It is noteworthy that the Spanish ranking gives greater weight than the other indicators to French and Italian presses and to venues that publish in Spanish.[14] The Northern lists, in contrast, tend to be dominated by presses based in Anglophone regions: whereas these publishers make up 64% of the Danish list, 61% of the Norwegian list, 60% of the Finnish list and 58% of the Flemish, they account for only 38% of the Spanish list. The lists also vary significantly in the relative weight given to university presses, which are only 17% of the Spanish top publishers as opposed to 41% of the Flemish one.

2. Uses of local indicators and impact on internationally mobile researchers:

The five indicators vary not only in their content but also in the uses to which they are put in their respective academic systems. In some cases, the lists are used only for aggregate assessments of output; they serve to determine the distribution of funding among universities and other research institutions. In the Finnish system, for example, the indicator is designed for 'the evaluation of large publication quantities, such as the entire production of universities or research institutions but not for the evaluation of individual researchers'.[15]

In other cases, however, these tools are used for both aggregate and individual assessments, even though they were not intended for the latter. The background report on the Norwegian model states that it cannot be used to assess individual researchers or publications. Nevertheless, as Aagaard et al. report, it is commonly employed to determine salary, awards and promotions: 66% of department heads use it 'to assess individual researchers' publishing performance'.[16] Similarly, the Flemish VABB-SHW list was designed to take SSH publications more into account in the calculation of the *BOF-sleutel*, the formula

by which funding is divided among the Flemish universities. An official document explaining the database indicates that it is 'NOT appropriate for comparing researchers individually, in groups or by disciplines and certainly not in a cross-disciplinary manner'.[17] Nevertheless, one of the most important Flemish funding institutions, the Research Foundation-Flanders (FWO), regularly asks applicants for doctoral or post-doctoral mandates and research projects to list publications included in the VABB-SHW. The list is also used in decisions about promotions at individual universities. Finally, the Danish indicator has been used 'to supplement citation-based evaluations of applications for higher academic positions or for determining if an academic should receive an increase in salary'.[18]

The systems also differ in the relative weight they assign to publications at the various levels. Most of the regions under consideration give different numbers of points to journal articles, books and book chapters published at the different levels (in the Norwegian system, for example, a Level 2 journal article is worth three times as much as a Level 1 journal article). For the sake of this comparison I will use the journal article published at the lowest recognized level as my basic unit of measurement. In the Danish and Norwegian systems, a Level 1 book counts as 5 units, whereas a Level 2 book is worth 8 units. In the Finnish system, a Level 1 book is worth 6 units, and a Level 2 or 3 book is worth 12 units. Finally, in the Flemish system, books recognized by the VABB-SHW are worth 4 units in the *BOF-sleutel* whereas others are given no points at all.[19]

As these point systems demonstrate, the inclusion of a publisher on the top level of these lists has at times significant consequences for the career of a scholar. The difference between Level 1 and Level 2/3 in Finland is the difference between writing one book and writing two. And in Flanders, VABB-SHW recognition is in some cases the difference between writing one book and writing none at all.

Given the weight given to these lists, the discrepancies among them are particularly troubling from the standpoint of mobility. Whereas applicants who have trained in the local context are familiar with the regional indicator from the very beginning of their careers and strategically submit research to the privileged venues throughout their doctoral and post-doctoral periods, foreign candidates discover its existence only when they apply for a job and are asked to classify their publications according to the local criteria. Let us imagine the case of a Finnish scholar who has published a book in a press that is included in the Finnish Level 2 list but not on the Flemish VABB-SHW. In applying for a professorial position or a post-doctoral fellowship in Flanders, this candidate's publication could count significantly less than a book by an internal candidate that has been recognized by VABB-SHW.

Indeed, not only do Flemish candidates have the advantage of prior awareness of the VABB-SHW list but they are also the only ones who can request recognition for publications that are not automatically accepted into it. In the Flemish system, no publication venue is incorporated into the list until a researcher affiliated with a Flemish institution has published in it.[20] This gives internal candidates for Flemish job openings a clear advantage over foreign ones who cannot submit their work for consideration for VABB-SHW inclusion prior to applying. Local researchers who have spent part of their careers abroad may also find many of their earlier publications excluded from the database, which is at times used for decisions about promotions and project funding. In short, when local bibliometric indicators are used for individual assessments, foreign researchers often find themselves evaluated by criteria of which they have been hitherto unaware and which they have had less of a role in shaping than local researchers.

3. Discussion and policy recommendations

Bibliometric criteria such as Web of Science are often criticized for falling into a metonymical logic: confusing a container with its contents or a thing with its location. The underlying assumption of these indicators is that there is a correspondence between the

journal and the articles inside it or between a book and the press in which it is published. It is simply not the case, however, that all articles in a given journal or all books published by a given press are of the same quality. As Claes Vesterager Pedersen points out, the Danish BFI 'can give the impression that it is concerned with measuring quality' but this is not what it does.[21] For this reason some assessment exercises such as the Research Excellence Framework (REF) in the United Kingdom explicitly prohibit the use of 'journal impact factors, rankings, lists or the perceived standing of publishers in assessing the quality of research outputs'.[22]

Supporters of these indicators generally acknowledge that one can assess the quality of a publication only by reading it but point to the time, labour and cost of peer-review evaluations. In appointment procedures one must often assess dozens of candidates with long publication lists in a short period of time. It is simply not possible to read the entire output. Researchers inevitably resort to short-cuts in these situations.

In the Anglo-American academic world, which has a long tradition of associating presses with prestige, a common short-cut is to consider publication venue. The hierarchy of presses in the Anglophone world tends to privilege university over non-university publishers (as well as Anglophone over non-Anglophone ones) and is loosely correlated with the hierarchy of universities, which is sharply differentiated and accorded great importance in these cultures. Indeed, many institutions of higher education established presses in the late nineteenth and early twentieth century in order to assert their importance within an academic world increasingly focused on research.[23] In continental Europe, the hierarchy among universities has generally been less pronounced (people tend to study where they live), and researchers often publish in more local venues. It is not a coincidence that the focus on bibliometric criteria in Europe has coincided with the rise of university rankings such as the Shanghai and Times Higher Education lists in the past decades.

Publication venue, however, is not the only short-cut used in research assessments. In the German academic world in which edited volumes are a privileged mode of scholarly communication, the reputations of the compiling editor(s) and of the researchers included in a collection can be as important as the press. This system is also based on a metoymnical logic ('birds of a feather flock together') but is in some ways more intuitive (it replaces fetishism of imprint with trust in a fellow academic). Ultimately, however, it is more suited to a small academic community than to a larger, international one. A German academic may know or know of the other German researchers in his or her field but it is difficult to have the same familiarity with scholars working abroad. Moreover, this system may give an unfair advantage to local candidates, who are likely to have greater access to the recognized networks.

Yet another short-cut is the writing sample, which could be considered a type of synecdoche, i.e. taking the part for the whole. This is a common practice in appointment procedures and is also the principle of the REF evaluations in the United Kingdom in which researchers submit several key publications, which are then read and evaluated by expert panels. In many ways, this is the most intuitive short-cut as the judgment is based on the researcher's work rather than an associated attribute (journal, press or editor). This method, however, requires a greater investment of time and labour and is at times perceived as subjective. Often, indeed, committees are reluctant to give the qualitative evaluation of the writing samples the same weight as quantitative indicators, which may feel more objective or seem more relevant to the financing model of a particular academic system.

Given that the use of short-cuts is in some situations inevitable, there is something to be said for a standardization of these methods. In the Anglo-American system, the hierarchy of academic book publishers is often an unwritten code that is transmitted orally from one generation of researchers to another. These values, however, are not entirely homogenous, and a young scholar who is not properly informed can make poor choices,

which can have serious repercussions for his or her career prospects. Given that journals and academic publishers vary considerably in their selection and editorial procedures, the idea of a standardized list that distinguishes vanity presses from peer-reviewed venues and that guides researchers in deciding where to publish is not without merit. It is, however, unhelpful and indeed counter-productive to have a proliferation of highly divergent lists, which may lead researchers to believe that a press is well-regarded when it is only recognized in a specific region.

The most compelling argument for compiling lists at the regional or national level is that local scholars are more likely to recognize quality publishers in the local language/context. It is striking, however, how few regional presses appear in the indicators under consideration. The Danish list includes no Danish presses, and the Norwegian indicator, no Norwegian ones. There are only four Finnish presses on the Finnish list and five Belgian one on the Flemish VABB-SHW. The Spanish indicator places foreign and Spanish presses in separate lists, and the Flemish have a separate label for local publishers. The justification for the existence of five separate lists is somewhat tenuous. That said, it is important to recognize that these lists also have a positive side: they are useful short-term correctives for the omissions and oversights of international indicators, and they may encourage researchers to be more ambitious in their choice of publication venue. Eventually, they may also be a useful starting point for establishing a common list. Nevertheless, given the high level of discrepancy among them, their use should be strictly limited to aggregate assessments.

In the long term it would be better to develop a single, transparent list of publication venues (both books and journals) at a supranational level. A standard list would not only level the playing field for local and foreign scholars but also allow for more meaningful cross-national comparisons of research output. The EU would be an obvious candidate for such a project (it has already embarked on a similar venture with the European Reference Index for the Humanities), but given the recent decision of the United Kingdom to leave the union, the creation of a European list could lead to a double system in which Anglo-American scholars adhere to one set of values, and EU researchers to another. A better option might be a transnational organization with a focus on research policy and experience in cross-national comparisons such as the OECD. Such an institution would give the indicator greater stability and legitimacy.

What shape should this indicator take? Given the difficulty of distinguishing among presses in an objective way, a system of levels seems preferable to a ranking. The Norwegians opted for a two-tiered model after observing Australia's experiment with performance-based research funding in the 1990s, in which publications were simply counted with no weighting system. After a number of years, the total number of Australian publications increased but the impact of the research as measured by citation rates decreased. Scholars had adapted to the call for productivity by opting for lower-impact publishing venues. Eventually, the system was discontinued.[24] In contrast, the Norwegian system of levels, which was the first to be created in this way, seems to have had no adverse effect on impact so far.[25]

The question of the number of levels is complicated. Currently, the Finnish system is the only one to distinguish a restricted top tier of presses. It does not, however, reward Level 3 publications with extra points. This ambivalence is understandable. Most researchers would probably acknowledge that not all the Level 2 presses on the Danish and Norwegian lists or all presses on the VABB-SHW list are equal. Nevertheless, it is important to point out that not all presses publish in all disciplines. Some scholars, thus, could be excluded from a highly restricted Level 3 simply because of the field in which they work. The idea of creating lists for each discipline might be a way of dealing with this but disciplinary boundaries are notoriously hard to define and can be problematic for cross-discliplinary research.

Many suggestions have been put forward as to how book publishers should be evaluated.[26] Some have suggested the use of quantitative measures such as citation frequency.[27] This method, however, not only requires a longer time lapse (several years must pass before citations can be collected) but could also penalize presses (and by extension scholars) publishing in smaller fields in which there are fewer researchers and by extension fewer citations. Moreover, non-Anglophone and SSH publications are generally considered to be inadequately represented in existing citation indices.[28]

Others have proposed focusing on the international dissemination of the publications of a press by considering the library collections in which they are included.[29] In some countries, however, library collections are developed not by acquisitions librarians but by professors who at times order books not in a comprehensive way but rather in relation to their own research areas and needs. As with citation analysis, moreover, this system could disadvantage presses that publish in smaller and more specialized fields or that lack the financial resources to promote their publications abroad. Another proposal is the use of book reviews.[30] But given that fewer and fewer books are reviewed (due in part to the non-recognition of book reviews in evaluation procedures), this method could also work against publishers specializing in smaller fields.

Still others have proposed surveys of experts such as scholars or collection development librarians.[31] Specialized librarians often have a more panoramic view of a field than individual scholars and have less of a stake in the evaluations game than researchers. The training and degree of specialization of librarians, however, varies greatly from one academic context to another (not all countries or institutions have discipline-specific acquisitions librarians).

Given the relatively greater coherence of the Nordic lists, it seems clear that a system of expert panels whose members meet regularly and consult both existing precedents and with one another, is more likely to lead to a stable and authoritative consensus than a system of large-scale voting or surveys. Ideally, the panels would consist of a mix of researchers and specialized librarians from a wide range of countries and would work with criteria that are generally accepted in the disciplines that they represent. Humanities committees, thus, should not be asked to use citation metrics, which are highly controversial in their fields. In all cases, however, the panels should consider the 'publisher's good practice', factors such as international visibility, external peer review and whether authors pay to publish.[32]

Finally, it is important to define a homogenous system of weights and equivalencies. There should be international standards as to the points assigned to a book versus an article and to the weighting of co-authored publications. To facilitate mobility in the academic world, it is essential to establish a common evaluation framework that is transparent, stable and internationally accepted.

NOTES

[1] On academic inbreeding and the obstacles to mobility in different academic contexts, see Maria Yudkevich, Philip G. Altbach and Laura E. Rumbley, eds., *Academic Inbreeding and Mobility in Higher Education: Global Perspectives* (New York: Palgrave, 2015).

[2] For a comparison of journal databases in the humanities and social sciences (WOS, Scopus, ERIH, AERES), see Michèle Dassa, Christine Kosmopoulos and Denise Pumain, 'JournalBase—A Comparative International Study of Scientific Journal Databases in the Social Sciences and Humanities (SSH)', *Cybergeo: European Journal of Geography* (online), document 484, http://cybergeo.revues.org/22862; DOI: 10.4000/cybergeo.22862.

[3] On the importance and prestige of book publications in the Humanities and Social Sciences, see Elea Giménez-Toledo, Carlos Tejada-Artigas and Jorge Mañana-Rodríguez, 'Evaluation of scientific books' publishers in social sciences and humanities: Results of a survey', *Research Evaluation* (2012), 1-14; and Elea Giménez-Toledo, Jorge Mañana-Rodríguez, Tim C. E. Engels, Peter Ingwersen, Janne Pölönen, Gunnar Silvertsen, Frederik T. Verleysen and Alesia A. Zuccala, 'Taking Scholarly Books into Account: Current Developments in Five European Countries', *Scientometrics* 107:2(2016), 685-699.

[4] The Authority Lists of the Danish BFI may be consulted at: http://ufm.dk/forskning-og-innovation/statistik-og-analyser/den-bibliometriske-forskningsindikator/autoritetslister. The Norwegian database is found at: https://dbh.nsd.uib.no/publiseringskanaler/KanalTreffliste.action?sok.uhrNivaa=2&sok.avansert=true&tidsskriftTreffside=1&forlagTreffside=3&treffliste.vis=true&bibsys=false.

[5] The Finnish database may be consulted at: https://www.tsv.fi/julkaisufoorumi/haku.php?lang=en.

[6] Given that there are no Spanish presses on any of the lists under consideration, this separation may not be that important for this comparison. The SPI rankings may be consulted at: http://ilia.cchs.csic.es/SPI/rankings.html. The 2014 ranking was used in this study.

[7] The Flemish list may be downloaded from this site: https://www.ecoom.be/en/vabb. The fifth edition of the list was used in this study.

[8] On the Danish indicator, see Claus Vesterager Pedersen, 'The Danish Bibliometric Research Indicator—BFI: Research publications, research assessment, university funding', *ScieCom Info* 4 (2010), 1-4. On the Finnish indicator, see Otto Auranen and Janne Pölönen, 'Classification of scientific publication channels: Final report of the Publication Forum project (2010-2012)', Federation of Finnish Learned Societies: http://www.julkaisufoorumi.fi/sites/julkaisufoorumi.fi/files/publication_forum_project_final_report_0.pdf. On the Flemish indicator, see Frederik Verleysen, Pol Ghesquière and Tim Engels, 'The Objectives, Design and Selection Process of the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB-SHW)' in W. Blockmans et al., *Bibliometrics: Use and Abuse in the Review of Research Performance* (London: Portland Press, 2014), 115-125. On the Norwegian system, see Jesper W. Schneider, 'An Outline of the Bibliometric Indicator Used for Performance-Based Funding of Research Institutions in Norway', *European Political Science* 8 (2009), 364-378; and Kaare Aagaard, Carter Bloch and Jesper W. Schneider, 'Impacts of Performance-Based Research Funding Systems: The Case of the Norwegian Publication Indicator', *Research Evaluation* 24 (2015), 106-117. For an overview of the five indicators, see Elea Giménez-Toledo et al., 2016.

[9] Whereas the Flemish VABB-SHW distinguishes among the series/imprints of a press by listing ISBN numbers, the other indicators simply list the publishers' names.

[10] In the 2014 essay cited above, Verleysen, Ghesquière and Engels compare the Flemish VABB-SHW list to the Danish, Finnish and Norwegian lists (they work with slightly earlier versions). Unlike the present study, however, they consider the list of automatically recognized presses in Flanders in relation to *both* the Level 1 and Level 2 lists in the Nordic countries. This results in a somewhat optimistic figures about the overlap among the lists (83.0% of the Flemish publishers are on the Danish list; 86.1% on the Finnish and 86.8% on the Norwegian). It is important to point out, however, that the Nordic systems assign very different weights to Level 1 and Level 2 publications. A Flemish researcher who publishes in a VABB press that is recognized as Level 1 (Basic) on the Norwegian list would thus find the publication devalued in Norway. Moreover, the authors do not sufficiently consider the degree of overlap in the opposite direction. Although they mention that the VABB does not

include 188 of the Nordic Level 1 presses and 19 of the Nordic Level 2 ones (i.e. a total of 207 publishers), they do not calculate what percentage these presses represent on the Danish, Norwegian and Finnish lists. If one runs a calculation based on the authors' numbers (112 on the VABB-SHW and 207 excluded), it would seem that only 35.1% of the Nordic Level 1 and 2 presses are included on the Flemish list. This figure hardly supports the conclusion that there is a 'general agreement' as to the 'international layer' of publishers (126).

[11] Dassa, Christine Kosmopoulos and Denise Pumain's work on journal databases points to a similar lack of overlap in international SSH indicators. Of the 3453 journals in Scopus 'Social Sciences' (which includes Humanities) and the 3178 SSH journals in the Web of Science (WOS-AHCI and WOS-SSCI), they find 2175 in common. That makes a Jaccard similarity coefficient of .328.

[12] According to the website of the Norwegian list, Denmark, Finland, Norway, Sweden, Iceland and Faroe Island plan to collaborate on a 'Nordic list': https://dbh.nsd.uib.no/publiseringskanaler/Forside.action;jsessionid=XpoP5bUbrU+s2q8+DAI49CE4.undefined?request_locale=en

[13] CNEAI (Comisión Nacional Evaluadora de la Actividad Investigadora) allows the use of the SPI in the assessment of individual researchers. See, for example, 'Resolución de 26 de noviembre de 2015, de la Comisión Nacional Evaluadora de la Actividad Investigadora...' in the *Boletín Oficial del Estado*, no. 286, 30 November 2015, p. 113075.

[14] The Spanish list includes the following French, Italian or Spanish-language publishers, which do not appear on other lists: Armand Colin, Bulzoni, CNRS, Cedam, Dalloz, École Française de Rome, Éditions de Minuit, Electa, Flammarion, Fondo de Cultura Económica, Francis Lefebvre, Giappichelli, Giuffre, Giulio Einaudi, Hachette, Il Mulino, Klincksieck, L'Erma di Bretschneider, L'Harmattan, La Découverte, Laterza, Leo S. Olschki, Les Belles Lettres, Masson, Mondadori, Presses de l'université de Paris Sorbonne, Sismel. Edizioni del Galluzzo and UNAM.

[15] Auranen & Pölönen, n.p.

[16] Aagaard et al., 113-14.

[17] See the 'Begeleidende nota VABB-SHW—vijfde versie', emphasis in original, translation mine: https://www.ecoom.be/sites/ecoom.be/files/150617%20GP%20begeleidende%20nota%20VABB%20V.pdf.

[18] Giménez-Toledo et al., 2015, 689.

[19] Giménez-Toledo et al., 2016. According to this article, the SPI is used 'only as a reference' in evaluations (688), and SSH book publications in Spain are considered the equivalent of two papers (694).

[20] 'The Authoritative Panel ... will only assess a journal in the year following the one in which for the first time a researcher associated with a Flemish university (college) has published in the relevant journal'. https://www.ecoom.be/en/services/vabb/faq

[21] Vesterager Pedersen, 1.

[22] http://www.ref.ac.uk/about/guidance/faq/researchoutputsref2/

[23] Steven E. Gump, 'Prestige and the University Press', *Journal of Scholarly Publishing* 37:2 (2006), 69-85.

[24] On the Australian experience, see Linda Butler, 'What Happens when Funding is linked to Publication Counts?' in H. F. Moed et al., eds., *Handbook of Quantitative Science and Technology Research* (Netherlands: Kluwer, 2004), 389-405.

[25] Schneider, 374; Aagaard et al., 111.

[26] For an overview of the options, see Elea Giménez-Toledo and Adelaida Román-Román, 'Assessment of Humanities and Social Sciences Monographs through their Publishers: A Review and a Study towards a Model of Evaluation', *Research Evaluation* 18:3 (2009), 201-213.
[27] See Alesia Zuccala, Raf Guns, Roberto Cornacchia and Rens Bod, 'Can we rank scholarly book publishers? A bibliometric experiment with the field of history', *Journal of the Association of Information Science and Technology* 66:7 (2015), 1333-1347.
[28] On the disadvantages of citation metrics, see Schneider, 368.
[29] On this idea, see Daniel Torres Salinas and Henk F. Moed, 'Library catalog analysis as a useful tool in studies of social sciences and humanities' in *Excellence and Emergence: a New Challenge of Quantitative and Qualitative Approaches* (Vienna: Austrian Research Centers GmbH-ARC, 2008), 246-250.
[30] Matthew L. Jordy, Eileen L. McGrath, and John B. Rutledge, 'Book Reviews as a Tool for Assessing Publisher Reputation', *College & Research Libraries* 60:2 (1999), 132-142.
[31] For examples of surveys of researchers, see Larry P. Goodson, Bradford Dillman and Anil Hira, 'Ranking the Presses: Political Scientists' Evaluations of Publisher Quality', *PS: Political Science and Politics* 32:2 (1999), 257-262; and E. Giménez-Toledo, C. Tejada-Artigas and J. Mañana-Rodríguez, 'Evaluation of Scientific Books' Publishers in Social Sciences and Humanities: Results of a Survey', *Research Evaluation* 22:1 (2013), 64-77. On the consultation of collection development librarians, see Janice S. Lewis, 'An Assessment of Publisher Quality by Political Science Librarians', *College & Research Libraries* 61:4 (2000), 313-323; and Paul Metz and John Stemmer, 'A Reputational Study of Academic Publishers', *College & Research Libraries* 57:3 (1996), 234-247.
[32] On the importance of examining 'the internal processes of publishers', see Giménez-Toledo et al., 2009.