

Discussion Paper Series

RIEB

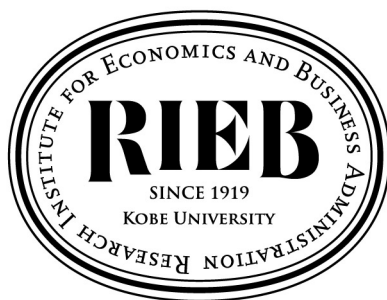
Kobe University

DP2019-J08

人文社会科学の研究評価について

浜口 伸明

2019年10月30日



神戸大学 経済経営研究所

〒657-8501 神戸市灘区六甲台町 2-1

On the Research Evaluation of Social Science and Humanities

1. はじめに

文部科学省は、平成 27 年 6 月 8 日付「国立大学法人等の組織及び業務全般の見直しについて」という文書の中で、「特に教員養成系学部・大学院、人文社会科学系学部・大学院については、18 歳人口の減少や人材需要、教育研究水準の確保、国立大学としての役割等を踏まえた組織見直し計画を策定し、組織の廃止や社会的要請の高い分野への転換に積極的に取り組むよう努めることとする。」という考えを表明した。この文書は第 3 期中期目標・中期計画（実施期間 2016 年～2022 年）を策定するにあたり国立大学の組織の見直しの方針を示したものだ。文部科学省の方針は一般に「人文社会科学不要論」と受け取られ、研究者コミュニティのみならず社会にも衝撃を与えた。日本学術会議は「総合的な学術の一翼を成す人文・社会科学には、独自の役割に加えて、自然科学との連携によってわが国と世界が抱える今日的課題解決に向かうという役割が託されている。このような観点からみると、人文・社会科学のみをことさらに取り出して「組織の廃止や社会的要請の高い分野への転換」を求めること」に強く疑問を呈する幹事会声明（平成 27 年 7 月 23 日）を發した。

「2018 年科学技術研究調査」(総務省統計局)によると、大学等を本務とする研究者 294、257 人の中で人文社会科学を専門とする研究者は 62、376 人（全体の 21%）である。文部科学省科学技術・学術審議会学術分科会（第 69 回、平成 30 年 8 月 22 日）に配布された資料¹によると、平成 29 年度の科学研究費の人文社会科学分野の採択件数は全体の 21.1%と研究者の比率と同じであった。配分額は 13.3%にとどまったが、大規模設備や研究材料を必要とする自然科学・生命科学と研究のスタイルが異なるため配分額の差は不自然ではない。今のところ、人文社会科学は日本の学術において相応な扱いを受けていると言ってい

いだろう。

文部科学省²は人文社会科学不要の姿勢を示したことを否定したが、「社会的要請の高い分野への転換」の余地が人文社会科学においてより大きいという観点から、むしろ人文社会科学が「社会的要請をリードするような積極的な提案」をすることを求めた。実際に文部科学省は国立大学の各部局が営々に行っている基礎的な教育研究に安定的に配分する運営費交付金の基盤的経費の配分を毎年定率で削減し、その分を機能強化経費として「積極的な提案」に対してコンテスト形式で再配分している。さらに、国立大学を様々な観点で格付けし、基盤的経費の削減と機能強化経費の配分に差をつけている。

研究費および運営費交付金の競争的配分と事後評価の基準として、研究業績の評価指標

¹ http://www.mext.go.jp/kaigisiryoy/2018/08/_icsFiles/afieldfile/2018/08/21/2-2.pdf

² 文部科学省高等教育局「新時代を見据えた国立大学改革」

が重視されている。評価指数の中で特に重視されるのが、Web of Science (WoS) あるいは SCOUPS といった国際的な学術誌のデータベースにおいて、インパクト・ファクターが高い学術誌に掲載された論文がどれだけあるか、あるいは、短期間に引用件数が同一分野の学術誌論文の中でトップ 1% となった高頻度引用論文がどれだけあるか、といったビブリオメトリック指数である。

ところが、ビブリオメトリック指数に基づく評価は理系向きで、人文社会科学には、その分野の特性に適していないと考えられている。ここで言う分野の特性とは、研究に要する時間の長さ、論文よりも学術書を通じた成果発表や日本語を含む英語以外の言語による成果発表の多さなどである。理系と同じ評価指標を用いられることにより、国の学術国際競争順位あるいは大学の評価につながりにくい人文社会科学に配分される資源が少なくなる懸念がある³。これは、日本に限った事情ではなく、諸外国も同じ状況に直面している (Ochsner et al. 2017)。人文社会科学不要論を払しょくするためにも日本学術会議は人文・社会科学に求められる自己改革の一つとして、「分野の特性に応じた評価指標の確立」を提言したが⁴、具体的な指標の策定にまで至っていない。

本稿は、文献レビューを通じて諸外国で進められている人文社会科学研究の評価の議論を理解し、日本において検討すべき課題を明らかにすることを目的とする。

2. インパクト・ファクターの問題点

WoS や SCOUPS がデータベースに掲載する研究業績は学術誌掲載論文、なかでも英語で発行されていて他の研究者により引用されやすい学術誌に絞られている。もっともよく参照される WoS に収録されている学術誌から計算されるインパクト・ファクターは、自然科学をカバーする SCIE (Science Citation Index Expanded) と社会科学をカバーする SSCI (Social Science Citation Index) 等のシリーズがある。SCIE および SSCI はがアメリカの Institute of Scientific Information (ISI) によって提唱され、1992 年に ISI がトムソン社に買収され、現在はトムソン・ロイターから事業を継承した Clarivate Analytics 社により運営されている。

このデータベースに基づいて計算されるインパクト・ファクターは、各学術誌に毎年掲載された論文が引用された件数を掲載された論文の数で割って計算されたもので、各学術誌の 1 論文当たり引用回数の単純平均である。

Nederhof (2006) が指摘しているように、人文社会科学と理系の研究成果報告スタイルの違いは以下のような理由から、インパクト・ファクターの数値レベルの差に表れる。新しい技術、実験結果を誰が最初に発表するのか、研究室等のグループ単位で国際的に競って

³ 実際に人文社会科学の研究者数は 2000 年以降頭打ちから減少に転じている。

⁴ 日本学術会議『学術の総合的発展をめざして —人文・社会科学からの提言—』平成 29 年 6 月 1 日。

いる理系分野では、理論・実証結果の更新サイクルが短い。このため理系の論文では、一般に論文の分量が短い回数が多く、共著者の数が多い。新しい研究は前に行った研究を引用するため、同じ研究グループ内の自己引用（self-citation）が多く、研究グループが大きく共著者が多いほど、自己引用も多くなる。

また、普遍性を追求する自然科学・生命科学では、最先端の研究であれば引用すべき文献に国籍は関係ない。したがって、優れた研究は世界中から引用される可能性がある。しかし、人文社会科学が分析対象とする文化・社会は多様性が大きく、ある国に関する研究から得られたエビデンスが必ずしも他の国にも通用するとは言えず、分析対象とする同じ国・地域に関する文献を多く引用することになる。例えばアメリカ経済に関する研究は日本経済を分析した論文を引用することは少ない。SSCIでカバーされている学術誌はアメリカで発行されているものが中心になっている偏りもみられる。

さらに人文社会科学の研究者は研究成果を個別に学術誌論文にするよりも、集成して学術書として刊行することを優先する傾向がある。学術書からの引用も多い。個人研究あるいは少数の共同研究が中心である人文社会科学は自己引用の頻度も少ない。

このような研究報告と引用の習慣の違いから、理系学術誌の引用件数の多さは掲載論文の数の多さで部分的に相殺されるが、インパクト・ファクターが人文社会科学より大きくなる傾向がある。

図1はWoSに掲載されている学術誌のインパクト・ファクターの分布のヒストグラムをSCIEとSSCIの間で比較したものである。収録されている学術誌の数はSCIEが9153、SSCIが3380である⁵。インパクト・ファクターの中央値はSCIEが1.907、SSCIが1.3965で、図1では縦の直線で示されている。なおここで表示しているのはインパクト・ファクターが10までに限定している。外れ値として除外したインパクト・ファクター10を超える学術誌は、SCIEで256、SSCIでは21存在する。極端に引用件数が多い理系の学術誌の影響により、インパクト・ファクターの単純平均はSCIEが2.7376、SSCIが1.7899と中央値の比較よりもさらに大きい。

また、図1により分布を見ると、SSCIは中央値がSCIEよりも低いだけでなく、中央値よりも低い範囲により多くの雑誌が集まっており、右側のテール（裾）が薄い。ただし、この指標から社会科学の研究活動の水準が自然科学・生命科学のそれを下回っていると単純に判断することはできない。第1節で紹介した日本で人文社会科学の研究者が研究者全体に占めるシェア21%が世界的にも当てはまると仮定すると、SSCIの収録誌数がSCIEとSSCI収録誌合計に占めるシェア27%はこれを上回り、人文社会科学の学問的多様性が大きいことを示している。上で述べたように人文社会科学がもともと学術誌を研究成果報告として理系ほど重視していないことに加えて、研究者人口と比較して相対的に多い学術誌に引用が分散する影響もあるため、平均値であるインパクト・ファクターを領域間で比較することには問題がある。

⁵ 2019年10月24日閲覧。

インパクト・ファクターは理系と人文社会科学の間だけでなく人文社会科学の分野別にも多様である。図2は社会科学の中で親和性があると思われる、経済学、経営学、地域研究・国際開発学、都市・地域・交通・環境・計画の4つの領域のインパクト・ファクターの分布を比較している。経営学および都市・地域・交通・環境・計画でインパクト・ファクターの中央値が高くなっているのはオペレーションズ・リサーチや工学の研究者が投稿する学術誌の数値が全体を押し上げているためである。経済学や開発学は数学や統計学で理系の手法を頻繁に用いるが、数学と統計学は理系であっても個人研究の多さや研究サイクルの長さという点でインパクト・ファクターの傾向は人文社会科学系と似たところがある。

Mongeon and Paul-Hus (2016)は、世界で発行されている学術誌のもっとも網羅的なリストとして Ulrich Knowledgebase を基準とすると、SCOUPS と WoS は自然科学・工学でそれぞれは Ulrich の 38%、33%をカバーしていて両社の差が小さいのに対して、社会科学は 25%、15%とカバー範囲が小さいとともに両社の差が大きいことを指摘している。Ulrich では社会科学で発行されている学術誌の数が自然科学・工学を上回っているにもかかわらず、WoS に収録されている社会科学の学術誌の数は自然科学・工学の半分しかないことも指摘している。

インパクト・ファクターに基づく研究評価では人文社会科学の研究者は理系の研究者よりも不利な状況に置かれていることがわかるが、理系の研究者もインパクト・ファクターを研究評価に用いることを必ずしも肯定しているわけではない。アメリカ細胞生物学会が2012年にサンフランシスコで開催した年次大会で採択した「研究評価に関するサンフランシスコ宣言」(San Francisco Declaration on Research Assessment, DORA)⁶では、インパクト・ファクターはもともと図書館員が購読すべき学術誌を選別するために作られた指標であって、研究の学術的評価をするためのものではないことを確認したうえで、研究の評価は研究自体の重要性に基づいて行われるべきで、どの学術誌に掲載されたかではないと主張している。この宣言は、インパクト・ファクターは、学術誌の平均的な引用回数に過ぎず、同じ学術誌の中で論文によって引用の頻度は非常にばらばらであるため各論文の評価と必ずしも一致しないこと、インパクト・ファクターの数値は分野により異なること、インパクト・ファクターは学術誌の編集方針によって操作可能であること、インパクト・ファクターを計算するための原データは明白でもなければ誰でも検証できるようになっていないこと、などの問題を指摘し、インパクト・ファクターを研究資金の審査、研究者の採用と昇任に用いるべきではないという見解を示した。

このほかにも、ジョージア工科大学公共政策学部のダイアナ・ヒックス教授とライデン大学科学技術研究センターのポール・ウォウタース教授による「研究指標に関するライデン・マニフェスト」(Hicks et al. 2015)も、インパクト・ファクターに基づいて簡便に研究業績評価が行われる傾向に警鐘を鳴らしたものとして知られている。このマニフェスト

⁶ <https://sfdora.org/read/> (2012年12月16日)

は以下の 10 項目からなる。

1. 定量的評価は、質的・専門的評価をサポートするものとして用いるべきである
2. 実績の測定は所属機関、グループまたは研究者個人のミッションに照らして行え
3. 特定地域に意味のある研究を評価せよ
4. 評価のデータ収集と分析のプロセスは常にオープン、透明、わかりやすくあるべきだ
5. 評価を受けるものがデータと分析過程を確認できるようにすべきである
6. 分野により公刊と引用の慣行が異なることに留意せよ
7. 個々の研究者の評定は、生涯の研究実績の定性的判断に基づくべきである
8. 評価指標を無理に具体化したり確証もなく詳細にしたりするべきではない
9. 評定と指標によってシステム全体が影響を受けることを認識せよ
10. 指標を定期的に吟味し、改善せよ

ライデン・マニフェストは研究評価を数値化して行うことを否定するものではないが、分野の多様性の尊重、指標の透明性と検証可能な妥当性、短期的成果への偏りの排除を要求するとともに、第 9 項目においてインパクト・ファクターに基づく評価が研究活動に歪みを与えていることを真剣に認識すべきだと提言している。評価は学術の進歩に貢献するために行うべきであって、それを歪めることがあってはならないが、一部では引用を増やすことを目的とした引用ドーピングと呼ばれるような組織的な行動が顕在化している。例えば、次節で述べるようにイタリアは近年研究評価システムを導入した。その後、イタリア人研究者の引用件数が増加したが、これは自己引用とイタリア人研究者同士が相互に引用しあう「引用クラブ」の存在によるところが大きく、「引用ドーピング」の疑いが強いと批判されている⁷。自己引用と国内研究者間の引用によって引用数を急速に増やしていることは、研究開発投資を増加し研究者数を増やしている中国において顕著であるとも報告されている⁸。

3. 社会科学における研究評価の取り組み事例

第 2 節で論じたようにインパクト・ファクターに代表されるビブリオメトリック指標に基づく研究評価は、理系でも問題にされているが、特に人文社会科学では適切でないことを批判する研究者は多い。とはいえ、Ochsner et al. (2017)が言っているように、評価そのものをボイコットすることは人文社会科学研究にとって適切ではなく、研究の生産性を高めるとともに、社会からの信頼と社会への貢献を促進することに寄与するような評価方法の在り方について研究者自身が積極的に関心を持つ必要がある。

この節では諸外国における取り組みを紹介する。最初に取り上げる欧州科学技術協力

⁷ “Italy’s rise in research impact pinned on ‘citation doping’” *Nature News*, September 13, 2019

⁸ “Citations strength begins at home,” *Nature Index*, December 12, 2018

(European Cooperation in Science and Technology, COST) は、欧州連合 (EU) の資金援助のもとで活動している欧州研究者の国際組織である。COST では、加盟する 36 か国から 125 人の専門家が参加して 2016 年 4 月から 4 年間の計画の下に、人文社会科学における研究評価のための欧州ネットワーク (European Network for Research Evaluation in the Social Sciences and the Humanities、<https://enressh.eu/>) を運営している。この取り組みでは、人文社会科学の研究評価方法の改善と社会への貢献の促進にむけて、人文社会科学の研究者が研究課題を明確にし、細分化された専門分野に閉じこもる蝸壺化を克服することを目的とし、①研究評価の枠組み、②社会的インパクトと人文社会科学の重要性、③人文社会科学のデータベース、④研究成果公開の 4 つの分科会を設けている。

分科会③でフォローアップされているように、画一的な WoS や SCOPUS は多様性が強い人文社会科学の評価に適さないという認識のもと、欧州では、各国で独自の研究データベースの開発が進められている。例えばノルウェーの CRISTIN、フランドル地方の VABB-SHW、イタリアの CINECA、英国の Research Outcomes System、スロヴェニアの SICRIS である。

分科会①で扱われているテーマは、研究評価における数値化と査読制それぞれのメリットとデメリット、国際通用性と地域固有性のバランス、学際性と専門性のバランスなどが挙げられている。ここで議論されているバランスの問題が人文社会科学に適した研究評価方法を確認する際の重要なポイントであろう。例えば、論文が掲載された学術誌や学術書が出版された出版社の格付けで研究の水準を評価するのをやめて、ひとつひとつの論文や学術書を査読して評価しなければならないという意見がある。しかし、そうするために膨大な労力を必要とする。研究評価においてインパクト・ファクターのような単純な数値に頼るのは、読まなくてもすむ省力化のメリットがあるからだと言うこともできる。数値化と査読性にはそれぞれメリットとデメリットがあり、最適な評価方法を決めることは難しい。また、評価の基準として、地域社会の固有性に基づく研究や英語以外の言語で書かれた研究、あるいは学問を統合していくような学際的研究といった、単純化のために従来のインパクト・ファクターからこぼれ落ちた研究を拾い上げていくことは評価基準を複雑にもする。

ノルウェーの事例では、研究費の配分と大学・研究機関の機関評価の両方の機能を持つノルウェー学術委員会 (The Research Council of Norway、2018) が主導して改革が進められた。Sivertsen (2016)によると、ノルウェーの CRISTIN は WoS 等ではカバーされていないものも含めて、査読を経たすべての研究業績を網羅したデータベースである。

これによって、WoS 収録誌でなくとも、査読を経た研究成果であれば必ず評価の対象になった。とくに単著・共著の学術書を明確に評価の対象とした。ノルウェーでは表 1 のような定められたポイント加算に従って研究活動が評価されている。重要性の高い学術誌や出版社から公刊された研究成果はポイントを高くした。レベル 1 とレベル 2 のポイントの差は特に学術誌論文で大きい。どの学術誌や学術書にレベル 2 の評価を与えるかを定める

のは、インパクト・ファクターではなく、各領域を代表する専門家の合議によるものとした。

表1 ノルウェー・モデルにおける研究成果別のポイント

	レベル1	レベル2
学術誌論文	1	3
学術書分担執筆論文	0.7	1
学術書	5	8

(出所) Sivertsen (2016)

このポイント加算方法は研究分野による成果発表形態の違いを考慮することによって横断的比較がしやすくした。これに基づいて機関ごとのパブリケーション・ポイントが計算され、研究費の一部がパブリケーション・ポイントの大きさに従って分配されるようになった。現在、ベルギー、デンマークもこの方式を採用している。ノルウェー方式は WoS に収録されていない雑誌の査読付き論文や学術書の研究業績を評価の対象に加え、地域固有性、特に自国に関する研究をより積極的に評価することになった。Sivertsen (2016)はこの方式を取り入れた結果、WoS 論文が増加したと述べている。

この評価方法において加算されるポイントは成果発表媒体に応じて決まっている点は従来のインパクト・ファクターと同じであり、個別の研究の内容を評価しているとは言えない。研究の社会的インパクトをどのように評価するのかについてもまだ論争がある (Hammarfelt, 2016)。

学術の評価制度の確立が遅れていたイタリアでは、2011 年に大学研究機関評価庁 (ANVUR) が設置され、最初の 2004 年～2010 年の評価結果を 2013 年に公表した。理系はビブリオメトリックに基づく評価がなされた。経済学は理系と同様にビブリオメトリックに基づく評価が選択されたが、それ以外の人文社会科学は査読に基づく評価を行った、1 万 4000 人の研究者を査読に動員し、18 万点の業績を評価し、非公開で各研究者に点数を配布した。このプロセスは Bonaccorsi (2018)に詳しく報告されている。

人文社会科学における重要な成果発表媒体である学術書を評価するために通読することは労力が大きいですが、代替的に簡便な数値指標として、図書館の蔵書カタログに掲載された冊数、出版社の評価、出版前の査読の有無を評価に用いることの是非は定まらず、査読をすることになった。査読において評価者により基準がバラバラになってしまうことが問題点として指摘されているが、Bonaccorsi (2018)は意見交換を繰り返すことによって評価基準を共通化していくことは可能だと述べている。

一方、Liu et al. (2005)によると、中国はインパクト・ファクターによる評価の文化を人文社会科学においても完全に受け入れ、WoS 収録論文の生産で急速にアメリカに迫っている。かつて大半は香港の大学の研究者による論文であったが、世界が中国に向けた関心

の大きさを背景に、北京や上海を拠点とする研究者による中国を題材にした研究が急速に増加している。アメリカの研究者との国際共著も積極的に行っている。

このように、諸外国ではヨーロッパで人文社会科学についてインパクト・ファクターではない独自の評価方法を確立しようとする動きがあり、そのなかでも多様な出版形態を考慮した数値化を行うノルウェーと、人文社会科学では実際に成果物を読んで評価する査読方式を取り入れたイタリアの間で対応が分かれている。一方、中国ではインパクト・ファクターによる評価を人文社会科学においても受け入れ、WoS 収録論文を公刊した場合に報奨金を支払う等のインセンティブをつけて数値の改善を奨励している。

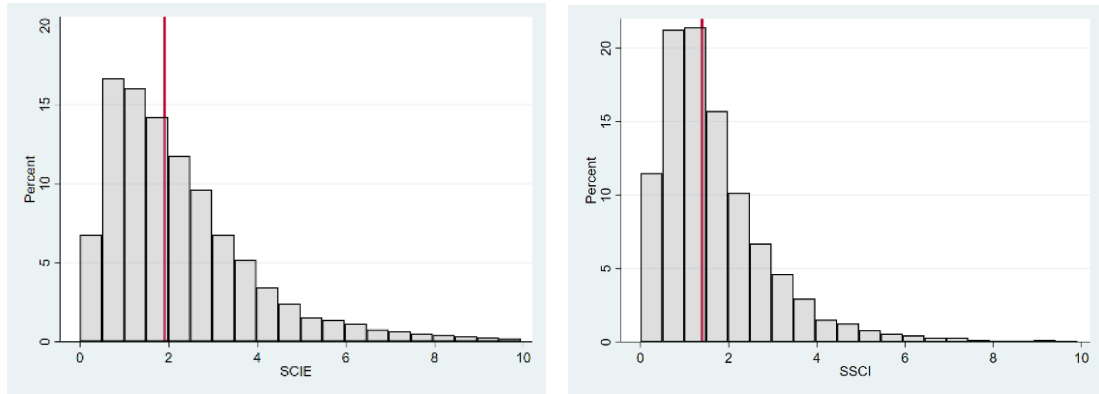
4. おわりに

筆者は神戸大学経済経営研究所長の任務を務める経験の中で、日本において国立大学が年を追って財政的に厳しい状況に追い込まれていることを強く認識したが、同時に機関評価の在り方においてもっと評価されて良いはずの人文社会科学の価値が矮小化されていることを感じた。その理由の一つが、研究評価で参照指標となっているインパクト・ファクターが、人文社会科学の特徴と合致していないことにあることを感じていた。このことが、本稿を書くに至った直接的な動機である。

第2節で述べたように、実際にインパクト・ファクターに基づく評価は人文社会科学の研究方法や成果発表方法と大きくずれていることが、これまでも研究者によって指摘されており、根本的な修正が行われないまま今日に至っている。とはいえ、評価そのものをボイコットすることは人文社会科学研究にとって適切ではなく、人文社会科学に適した評価方法の在り方について研究者自身が積極的に関心を持つ必要がある。

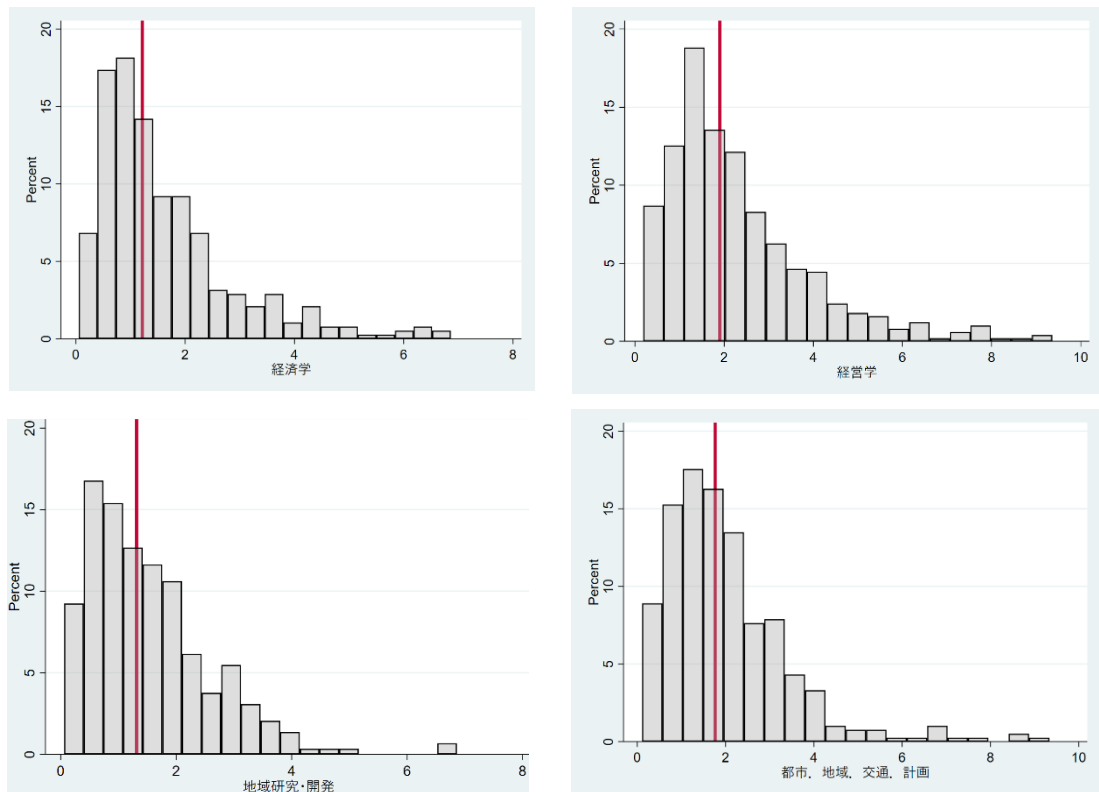
第3節で述べたように欧州において、研究成果に関するデータベースを再構築する動きは進んでいる。日本における同様の取組は遅れており、たとえば **researchmap** のような研究者共通のプラットフォームや、WoS・SCOUPS よりも広い範囲をカバーする **Google Scholar** あるいは **Microsoft Academic** を活用し、データベースを構築していくことの検討が求められるであろう。このようにして、業績の範囲を広げたうえで、どのような評価システムを構築すべきか、については、掲載媒体や出版元の格付けに基づくノルウェー方式か、査読を行うイタリア方式か、簡便さと多様性の尊重のバランスをとることは困難であり、どちらかに極端に寄った解決策に分化していくのではないか。労力が大きいイタリア方式を選択する国は少ないだろうが、どちらの方式がより大きく研究の質の向上をもたらすかは明らかでなく、今後も活発な議論が必要とされる。

図1 インパクト・ファクターの比較：SCIE（理系）とSSCI（人文社会科学）



(出所) Web of Science に基づいて筆者作成

図2 ジャーナル・インパクト・ファクターの比較：社会科学分野間



注：WoS に収録されている術誌の数は、経済学 381、経営学 496、地域研究・国際開発学 293、都市・地域・交通・環境・計画 395。インパクト・ファクター中央値は経済学 1.217、経営学 1.904、地域研究・国際開発学 1.314、都市・地域・交通・環境・計画 1.764

(出所) Web of Science に基づいて筆者作成

参考文献

- Bonaccorsi, A. ed. (2018) *The Evaluation of Research in Social Sciences and Humanities: Lessons from the Italian Experience*, Springer Cham/Switzerland.
- Hammarfelt, B. (2016) "Beyond Coverage: Toward a Bibliometrics for the Humanities," in in Ochsner, M. Hug, S.E., Daniel, H.-D. eds. *Research Assessment in the Humanities: Toward Criteria and Procedures*, Springer Open.
- Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). "Bibliometrics: the Leiden Manifesto for research metrics." *Nature News*, 520, 429-431.
- Liu, W., Hu, G., Tnag,, L., and Wang, Y. (2005) China's global growth in social science research: Uncovering evidence from bibliometric analyses of SSCI publications (1978–2013), *Journal of Informetrics* 9: 555-569.
- Mongeon, P. Paul-Hus, A. (2016) "The journal coverage of Web of Science and Scopus: a comparative analysis," *Scientometrics* 106(1): 213-228.
- Nederhof, A. (2006) "Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review," *Scientometrics* 66: 81-100.
- Ochsner, M., Hug, S. and Galleron, I. (2017) "The future of research assessment in the humanities: bottom-up assessment procedures," *Palgrave Communications Humanities and Social Science*, Article number 17020, 2017年3月21日.
- Sivertsen, G. (2016) "Publication-Based Funding: The Norwegian Model" in Ochsner, M. Hug, S.E., Daniel, H.-D. eds. *Research Assessment in the Humanities: Toward Criteria and Procedures*, Springer Open.
- The Research Council of Norway (2018) *Evaluation of the Social Sciences in Norway*.