

## 日本の研究パフォーマンスと研究実施構造の変遷

Transition of Research Performance and Research Structure in Japan

林 隆之, 富澤 宏之

HAYASHI Takayuki and TOMIZAWA Hiroyuki

1. はじめに	57
2. 国レベルの研究パフォーマンスの変化	58
2.1 先行研究と分析方法	58
2.2 分析結果：日本の論文生産の特徴と他国との差異	60
2.3 データのバイアスの検討	62
3. 日本の研究実施セクター・実施機関の変化	63
3.1 セクターごとのシェアの遷移	63
3.2 大学セクター内部での変化	64
3.2 集中と分散	66
4. 論文生産の特徴の背景	67
5. おわりに	70
ABSTRACT	73

## 日本の研究パフォーマンスと研究実施構造の変遷

林 隆之\*, 富澤 宏之\*\*

### 要 旨

本稿では過去20年間の日本の論文データを定量的に分析することにより、日本の研究活動のパフォーマンスおよび研究実施者の構造がいかに変化したかを明らかにする。被引用数の高さに基づく論文分布を分析した結果、日本は20年間に被引用数の高い論文数を増してきた一方で、被引用数の低い論文のシェアが継続的に高く、他の先進諸国とは異なる論文分布を有していることが示された。論文を生産している機関をみると1990年代半ばまでは多くの大学が論文を生産するようになる分散化傾向を示していたが、被引用数が上位10%に入る論文に限ればその半数が8つの大学により産出されており、1990年代半ばからはわずかに集中化に転じている。博士課程学生数や研究費の増加は多くの大学が研究活動を推進する基盤を形成してきたが、それらの資源は総量の増加にもかかわらず少数の大学への集中度合いを変えておらず、これらの大学の競争力を増してきた一因となってきたと考えられる。

### キーワード

ビブリオメトリクス, 研究評価, 研究パフォーマンス, 引用数, 集中化と分散化

### 1. はじめに<sup>1</sup>

知識や情報が経済・社会を駆動する主要な基盤となる知識基盤社会が進展する中で、大学などによって担われる研究活動は、その成長を左右する重要な要因となり、その卓越性は国際的な競争の中で捉えられるようになってきている。そのため、国や大学などの研究成果の量や質を共通的な方法で測定し比較することへの要望は増しており、その際にしばしば定量的な指標として用いられるのが、研究成果の代表的存在である学術論文のデータである。

日本においても、論文のデータを用いて研究活動の成果を定量的に測定している例はこれまでも見られる。文部科学省科学技術政策研究所の『科学技術指標 平成16年度版』では、トムソン・サイエンティフィック社(旧・ISI)の *National*

*Science Indicators 1981-2002 (Deluxe version)* を用いて分析を行っており、人文・社会科学を除く自然科学・工学分野において、世界の中での日本の論文発表数のシェアは1981~1985年には7.0%であったが、1998~2002年には10.1%へと増加しており、米国に次いで世界第二位の論文産出国になっていることを示している(文部科学省科学技術政策研究所 2004)。

一方で、研究成果の質的側面の測定については、論文が他の論文から引用された回数(被引用数)がしばしば用いられる。被引用数は論文がその後の研究に与えた影響の大きさを示すものであり、論文の質を測定するものではないが、被引用数とピアレビュー結果との相関がある程度は認められることから(Anderson et al. 1978, Irvin 1989, Rinia et al. 1998, 2001, 林2003など)、研究の質を一定程度には反映する指標として頻繁に用いられ

\* 大学評価・学位授与機構 評価研究部 助教授

\*\* 文部科学省科学技術政策研究所 科学技術基盤調査研究室 室長

<sup>1</sup> 本稿は Hayashi and Tomizawa (2006) を基に、新たなデータを用いて4節の相関分析を追加し、論点を主に大学セクターに絞るなどして、全体的に加筆修正したものである。

ている。日本の論文の被引用数については、これまで『科学技術指標』や『科学技術白書』などで、他の先進諸国の論文の被引用数と比べて低いことが指摘されてきている。被引用数のシェアを論文数のシェアで除した値を「相対被引用比率 (RCI)」と呼ぶが、日本の RCI は 1981~1985 年には 0.86 であり、1998~2002 年は 0.88 であり、20 年間に渡って 1 以下のままで大きく変わっていない。すなわち、日本の論文は平均してみれば、数は多いがあまり引用されていないということになる。

しかし、他方で *Nature* や *Science* などの国際的に著名な学術誌をみると、そこに掲載される日本の論文数はこの 20 年間で着実に数を増しており、科学の知識生産への日本の貢献は増しているという見方もできる。

では、日本の研究パフォーマンスは影響力が平均より低いまま変化していないのであろうか、それとも改善しているのであろうか。また、そのような傾向はいかなる学術的要因や社会的要因によって生じてきたのであろうか。

本稿では、論文データを詳細に分析することにより、日本の研究パフォーマンスや研究活動を担う実施者がいかに変化してきたかを明らかにする。分析では、国レベル、大学や企業といったセクターレベル、ならびに個別大学や大学の設置形態別レベルの 3 つの階層的レベルに渡って、過去 20 年間の論文生産の分布の変遷を追う。この 20 年間には、1980 年代の科学技術政策の基礎研究シフト、1990 年代からの経済不況、1995 年の科学技術基本法と翌年の科学技術基本計画の策定と、研究を取り巻く社会的状況に大きな変化を見ている。このような中で、日本の研究実施者の構造がいかに変化し、それにより日本の研究パフォーマンスがどのように変化してきたかを本分析では明らかにする。これにより、現在の日本の科学技術政策における課題を把握し、大学評価をはじめとして、研究活動に外部から影響を与えようとする方策が、どのような方向へと研究活動を誘引していくべきであるかを検討する。

## 2. 国レベルの研究パフォーマンスの変化

### 2.1 先行研究と分析方法

上述のように国全体の被引用総数だけを見ても、あるいは逆にトップレベルの学術誌における論文数だけを見ても、日本の論文生産の全体的特徴は明らかとはならない。そのため、被引用数の高さごとに論文がどのように分布しているかを明らかにすることが求められる。

このような分析は、少なくとも日本の全論文を対象に被引用数を計測することを必要とする。さらに、研究分野によって被引用数の平均値は大きく異なるため、研究分野間での比較可能性を担保するためには被引用数の標準化を行わなければならない。そのためには日本以外も含めた全論文の被引用数の分布を研究分野ごとに把握することが必要となる。これには大規模なデータ処理が必要とされるため、このような分析は国内外でほとんど行われてこなかった。

先行研究として、Butler (2003a) は、論文ではなくジャーナルをその平均被引用数によって 4 つのグループに分類し、オーストラリアの大学のシェアの分布を分析している。オーストラリアでは大学の研究活動への一括補助金を、論文数などの複数の指標によって算出するという方法をとっている。Butler の分析では、このような制度の導入移行に、研究者が平均被引用数が低く査読が通りやすいジャーナルを選択して、論文数を増すようになったという興味深い結果を示している。この分析では、研究者によるジャーナルの選択行為を分析するために、ジャーナルの査読の厳しさを代替する指標として、ジャーナルごとの平均被引用数を用いている。一方で、本分析では査読という事前の評価ではなく、論文が発表後にどれほどの影響を与えたかを把握することが求められる。雑誌の平均被引用数と個々の論文の実際の被引用数にはずれがあることも指摘されており<sup>2</sup>、本分析では論文ごとの被引用数を実際に計測する。

また、van Raan らライデン大学の研究グループは、大規模なデータ処理により研究分野ごとの平

<sup>2</sup> 通常、一雑誌の中でも被引用数の分布は、極めて少数の論文が多くの引用を受けるという歪んだ形になる。そのため、論文の実際の被引用数と掲載雑誌の平均被引用数（たとえばインパクトファクター）にはずれが生じることになる。詳細は、調 (2004) を参照。

均被引用数を算出し、大学評価などにおいて、評価対象研究グループの論文の被引用数を平均被引用数と比較している（その手法については van Raan 1996を参照）。一方で、本分析は、国全体の論文を対象に分析を行い、そこから国レベルやその内部のセクターや機関レベルの論文生産の構造的特徴を明らかとするものであり、より大規模な分析が必要とされる。

本分析では、トムソン・サイエンティフィック社が発行している *Science Citation Index*（以下 SCI と略す）CD-ROM 版の1982年から2003年までに収録された論文を分析対象とする。SCI は自然科学系の全分野を対象とし、論文に記された参考文献（reference）リストのデータを含んでいることから、引用分析に標準的に用いられているデータベースである。SCI には Web of Science で利用可能なオンライン版の SCI-expanded もあるが、本分析ではデータ収集が行いやすい CD-ROM 版を用いている。CD-ROM 版はオンライン版より収録論文数が若干少ないという問題があるが、それでも各年の収録論文数は1982年は550,096本、2003年は886,804本と巨大であり、年々増加している。収録雑誌数は2003年には3,733誌である。

本分析では、SCI データベースに収録された全論文の被引用数を計測し、それらを研究分野ごとに標準化した後に、集計を行う。具体的には、まず SCI に収録された全論文の参考文献（reference）リストから各参考文献の出現回数を計測し、SCI に収録されている論文と照合して、被引用数を求める。なお、参考文献リストには筆者によって表記揺れがあるため、その修正を行った。すなわち、ミドルネームの表記の有無の統合、ジャーナル名の表記揺れの修正、および、年間5回以上引用されているものについては、参考文献の著者名が研究グループ名の場合に第一著者名に修正した。参考文献リストと論文との照合は、第一著者の氏名、ジャーナルのイニシャル、巻数、ページ数の4つを用いた。また、被引用数の測定は、2003年までに引用された回数の合計値とした。そのため、出

版年が古い論文ほど被引用期間は長くなる。一方で、2003年に出版された論文は被引用期間が最大でも1年間しかなく、ほとんど引用されていないため、妥当な分析を行うことができない。そのため、以後に示す被引用分析では1982年から2002年までに出版された論文を対象とした分析結果を示す<sup>3</sup>。

次に論文の被引用数の標準化を行う。標準化の方法はいくつか提案されているが（Shubert and Braun 1996）、本分析では、分析対象の論文と同じ年に出版され、同じ論文形態（article, review, letter, note のいずれか）であり、同じ研究分野のジャーナルに掲載された全論文を比較対象群とし、その中で被引用数によって論文をランキングした場合に上位何%に位置するかという値により標準化する<sup>4</sup>。

分野分類については、トムソン・サイエンティフィック社がジャーナルごとに付与している168の分野分類（2002年の場合。分野の新設・統廃合があるために年により変化する）を用いる。たとえば、*Journal of Biochemistry* は「biochemistry & molecular biology」という分野分類に区分されている。そのため、1996年の *Journal of Biochemistry* 誌に掲載された article は、同年の「biochemistry & molecular biology」の分野分類に区分された雑誌に掲載された全 article が比較対象群になる。一雑誌に分野分類が複数付与されている場合には、論文はそれぞれの分野分類に分数で計上する。また、*Nature* や *Science* などの「multidisciplinary」（学際分野）という分類が付与された学術雑誌については、雑誌ごとではなく、論文ごとに再分類した。すなわち、各論文が引用している論文（参考文献）の掲載雑誌の分野分類の出現回数を集計し、最も多く出現した分野分類に区分した。

以上の方法によって、各論文の比較対象群を設定し、その群の中での被引用数によるランキングにより標準化する。たとえば10,000本の比較対象群の中で、被引用数で上位100位に位置する論文は、「被引用数上位1%」とされる。この標準ランクによって、各論文を、出版年・論文形態・分野

<sup>3</sup> ただし、出版年が2002年でも被引用期間は1～2年でしかなく、論文が引用されるまでの平均期間が長い研究分野にとっては十分な期間ではない。そのため、以降の節で最近の状況について詳細な分析を行う場合には、2001年以前の過去3年程度を分析対象年として用いている。

<sup>4</sup> 標準化の方法の詳細については、林（2003）を参照。

ごとに上位25%、25-50%、50-75%、75-100%の4つのグループに区分する。また、上位25%の中でも、さらに被引用数の高いグループとして上位10%というグループを重複する形で作成した。

## 2.2 分析結果:日本の論文生産の特徴と他国との差異

図1は、25%ごとの4つのグループと上位10%論文の合計5つのグループについて、日本の論文のシェアの変遷を出版年による時系列で示したものである<sup>5</sup>。これは、前述のように、まず168分野および4論文形態に分けて5つの被引用数グループを作成して、その中の論文総数および日本の論文数を計測した上で、次に、全分野・全論文形態の合計値を5つの被引用数グループごとに集計して日本の論文のシェアを計測したものである<sup>6</sup>。なお、ここでの「日本の論文」の計測は、著者の所属機関に一つでも日本の住所の機関が入っている論文は一本とカウントしている（このような計測の仕方は「全数カウント」と呼ばれる。他方、後の分析で用いるように、複数の機関からの著者

の共著の場合に、一論文当たりの所属機関数の逆数を用いてカウントする計測方法は「分数カウント」と呼ばれる)。

図1からは、この20年間にどのグループにおいても日本のシェアは増したことがわかる。被引用数上位10%論文の日本のシェアは1982年には5.5%であったが、2002年には8.7%にまで増加している。前述のように、日本は他国よりも被引用数の平均値が低いことが批判されてきたが、実際には日本からの影響力のある研究成果の産出は年々増している。

しかしながら、日本のシェアが最も高いグループは20年間に渡って、被引用数で上位50-75%という平均以下のグループである。また、被引用数の順位が最下位のグループは1990年代から急速に増加し、1999年にピークに達している。この結果は、日本は影響力の高い論文の産出を増してきたが、被引用数の低い論文をそれよりも高い割合で産出しており、とくに1990年代にこの傾向が強まったことを示している。

だが、このような特徴は必ずしも全ての研究分

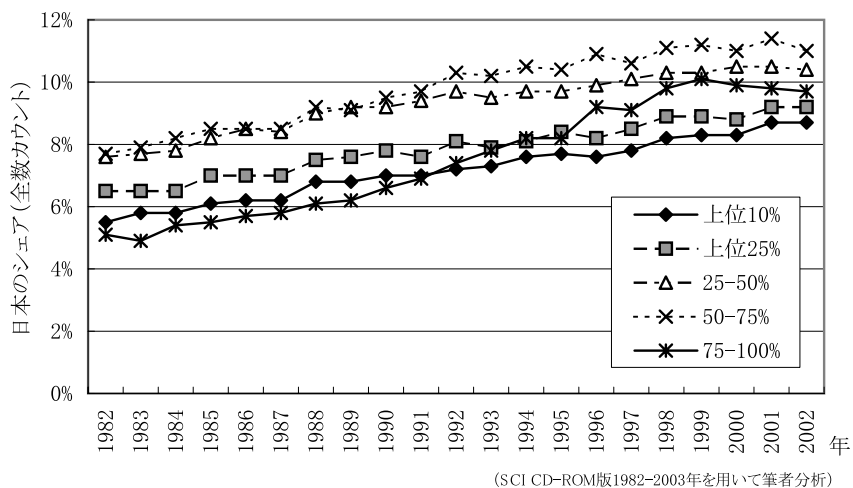


図1 被引用数ごとの日本の論文数シェアの推移

<sup>5</sup> 本稿では、全ての分析において、論文数そのものではなく、各年のSCIに収録された全論文の中でのシェアを示している。これは、SCIの収録論文数自体が毎年増加しているため、論文数そのものを用いると、いずれも増加傾向を示すことになり、論文グループ間での差異が明確にならないためである。一方で、中国や南米などの新たな論文生産国の出現によって、日本の論文数が増したとしてもシェアは相対的に低下する。SCIにおけるシェアは、国や機関の国際的な存在感の高さを示す指標として解釈すべきである。

<sup>6</sup> なお、このような合計をとることによって、SCIに比較的に良くカバーされている研究分野の特徴が反映されやすいという問題はある。ただし、2001年の場合には、SCIに収録されている論文の中で生命科学領域（臨床医学、薬学、生物・生化学、農学など）の論文は53.7%であり、たとえば科研費における生物系の配分額割合が51.4%、配分件数割合が46.8%（ともに2001年）であることを考えると、生命科学系と非生命科学系のバランスはほぼ妥当と言える。しかし、より詳細に見た場合には、分野によって影響力に差異が生じている可能性は十分にある。

野に共通しているものではない。図2は1999年から2001年の3年間に出版された論文について、見やすいように168分野を20分野に集計して、それらを二次元上に配置したものである。ただし、20分野のうち3分野は日本からの論文数が、極めて少ないか0であるために図には現れていない。横軸は被引用数上位10%論文における日本のシェアであり、影響力の高い論文を日本がどの程度産んでいるかという、日本の存在感の高さを示している。縦軸は、全論文のシェアに対する被引用数上位10%論文のシェアの比であり、日本による論文産出が被引用数の高い論文が多いのか、低い論文が多いのかという論文産出の分布を示す。円の大きさは日本の論文数に比例しており、生命科学領域とそれ以外とで色分けしてある。

結果は、材料科学、物理学、化学などの自然科学分野では上位10%論文における日本のシェアは高く、また論文分布においても被引用数が高いグ

ループのシェアが比較的高い。一方で、論文数の最も大きい臨床医学や、農学、薬理学などの生命科学領域では、米国などの他国の論文産出の多さも影響し、双方の指標とも低い。このように、分野ごとの違いは確かに存在する。しかしながら、材料科学を除いた全ての分野において、全論文中の日本のシェアと比べて、上位10%論文中の日本のシェアは低く、被引用数の高い論文の産出率が相対的に低いことが示されている。

しかし、この日本の分析結果を、直ちに否定的に捉えることはできない。それは、被引用数が高く影響力のある論文を生むためには、長期間の継続的な研究活動が必要なことは多く、その過程で被引用数の低い論文が多く産出されることは十分に考えられるからである。もし日本に見られた特徴が他の先進諸国でも共通に見られるのであれば、このような解釈が成り立つ可能性がある。そのため、図3には、2000年に出版された論文を対象に、

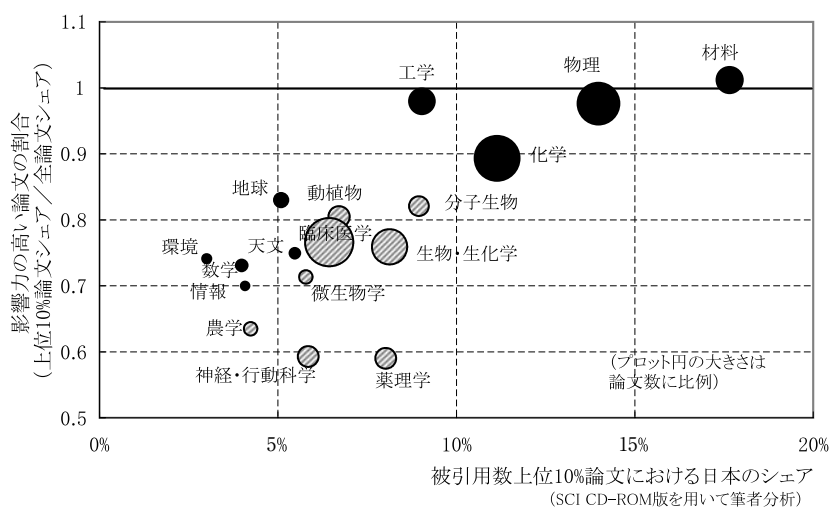


図2 高被引用論文における日本のシェアと分布状況 (1999—2001年)

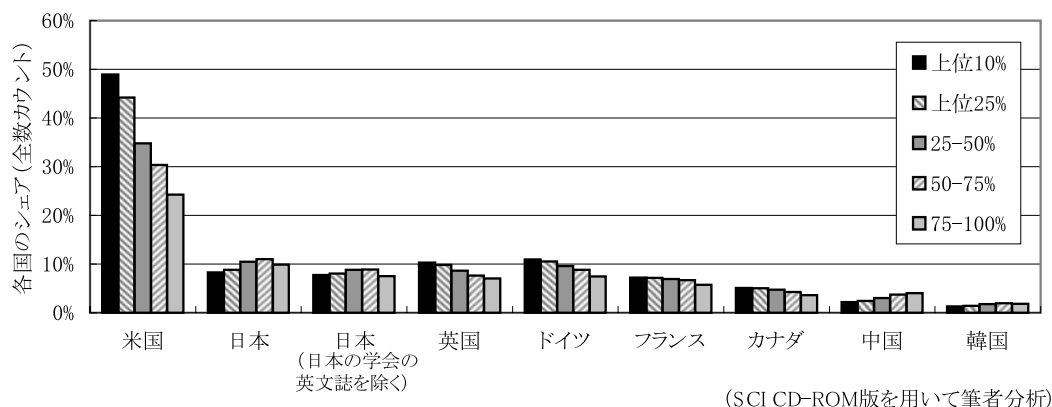


図3 各国の被引用数ごとの論文シェアの状況 (2000年)

日本を含めた幾つかの国について、同様の方法で被引用数ランキングに基づく論文グループごとの各国シェアを示している。

米国についてみると、被引用数上位10%の論文の4割以上に米国の機関が著者として入っており、被引用数が少ないグループになるにつれてシェアを下げている。英国、独国、仏国、カナダは、米国よりもシェアはだいぶ低い、被引用数が高いほうがシェアが高いという傾向を同様に示している。日本はこれらの国とは逆であり、被引用数75-100%を除けば、被引用数が高いグループほどシェアが低い。この傾向は中国や韓国に近いものであり、日本は論文数では第2位であっても、分布の特徴は依然としてキャッチングアップ国に近いものになっている。

### 2.3 データのバイアスの検討

この結果を詳細に分析する前に、まずはデータベースやデータ処理のバイアスの可能性を検討しておく必要がある。文部科学省(2000年以前は科学技術庁)が毎年、大学や企業などの研究者を対象に行っている調査「我が国の研究活動の実態に関する調査報告」の1998年版では、「日本人論文の被引用度が先進諸国に比べて低い理由」について質問している。その中で最も回答が多かったのは、「アブストラクトは英語であっても、本文が和文の論文が多い」(48.3%)であった。確かに、和文の論文は外国人からはほとんど引用されないために、全体の被引用数を下げることになる。しかしながら、実際には2000年のSCIに収録された和文論文は1,226本であり、これは日本人が著者の論文のわずか2%でしかなく、被引用数の分布に大きく影響するものではない。

次に多かった回答は、「英語表現上の問題で、同じレベルの内容でも引用されにくい」(43.5%)であった。このような言語上の問題が認識されている一方で、多くの研究分野では英文で論文を書くことは通常の営みとなり、さらに、日本の学会も多くの英文雑誌を出版するようになっている。実際、2000年のSCIには日本の学会が出版する88の英文雑誌の11,342編の論文が収録されている。こ

れは日本の論文の18.2%に相当する。

しかし、これらの雑誌の中には、全ての掲載論文の著者が日本人であるものも多くあり、その場合には国際的認知度も低く、引用されることも限られる。前出の図3には、日本の学会の英文雑誌の論文を除いて、日本のシェアを算出した結果も示してある。被引用数75-100%のグループでは36.7%が日本の学会の英文雑誌の論文であり、実際に日本の被引用数を引き下げる効果を持っていた。これらを除けば、分布状況は望ましい状態にある程度は近づく。しかし、日本のシェアが50-75%のグループで最も高いという傾向までは変わらなかった。

また、言語の問題は別の問題も引き起こしうる。日本人によって書かれた論文は、他国の論文と比べて参考文献の数が少ない。この原因には、英文文献を網羅的に調査することの労力が大きいことが挙げられよう。SCIの2001年のデータを対象とすると、日本人が著者に入っている論文は平均して24.8本の論文を参考文献として引用しているのに対して、米国は34.1本、英国は30.5本、ドイツは32.0本、フランスは30.9本である。通常、著者は自国の他の著者の論文群から引用を行う確率のほうが、他国の著者の論文群から引用を行う確率よりも高い。そのため、日本人の論文に参考文献数が少ないことは、日本の論文への被引用数が減少することになる。もし日本が自国および他国の論文を、米・英・独・仏の4カ国の平均値と同様の率で引用したという仮想的条件のもとで計算してみると、2001年の日本のRCIは0.93から0.97へと上昇する<sup>8</sup>。

以上のように、言語にとまなういくつかの問題が日本の被引用数を下げている可能性は考えられる。しかし、それらがたとえ解消した状態を仮想しても、他の先進諸国のように、被引用数が高い論文グループのシェアが最も高い状態にはならない。そのため、日本の研究活動の何らかの特徴が被引用分析の結果には反映されていると考えるべきであろう。先述の科学技術庁によるアンケートにおいて3番目に多く挙げられた回答は「研究内容に新規性・独創性のあるものが少な

<sup>7</sup> ただし、20分野のうち3分野は日本からの論文数が、極めて少ないか0であるために図には現れていない。

<sup>8</sup> 詳細については Hayashi and Tomizawa (2006) を参照。



い」(34.4%)であった。

### 3. 日本の研究実施セクター・実施機関の変化

#### 3.1 セクターごとのシェアの遷移

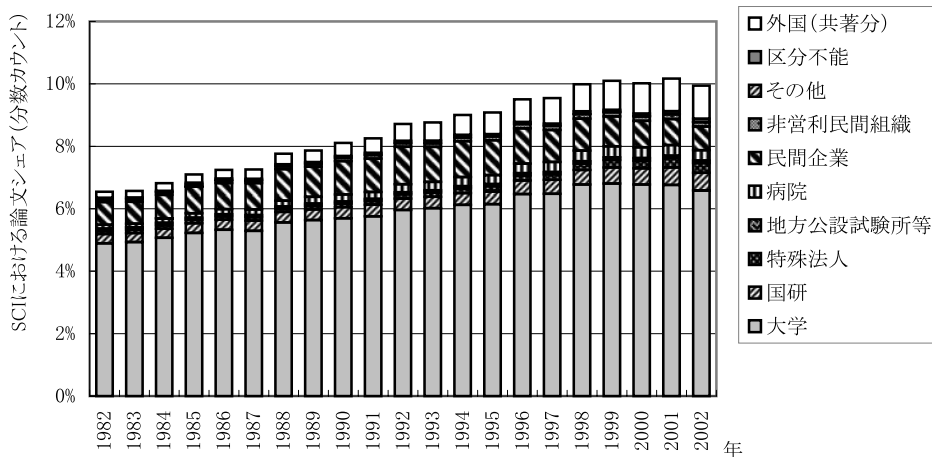
もし被引用分析の結果が日本の研究パフォーマンスを反映しているとすれば、なぜこのような特徴が生じているのか。その社会的な要因を把握するためには、まず、どのような機関やセクターが論文を産出しているのかを特定する必要がある。各論文を機関やセクターごとに分類するためには、論文に記された著者の所属機関名を用いる(たとえばKatz 1995, Godin and Gingras 2000, Hayashi 2003)。本分析では日本の研究実施機関を次のようなセクターに分類した。大学(短大や大学校などの高等教育機関、および大学共同利用機関も含む)、国立研究所、準公的研究所(特殊法人など)、民間企業、非営利民間機関(財団法人など)、病院(大学病院を除く)、その他である。なお、分析対象期間が1982年~2002年であるため、2001年以降に実施された独立行政法人化よりも前の分類を全年に渡って用いる。

分類では、まず「UNIV」「COLL」「HOSP」「CORP」「LTD」などの、セクターを代表するキーワードを用いて暫定的に分類を行った。次に、日本の研究機関のリストである『全国試験研究機関名鑑』および、JSTの「研究開発支援総合ディレクトリ(ReaD)」を用いて、英語名の機関リストを作成し、論文著者の所属機関名と照合した。なお、分類が不可能であったもののうち、1年に10回以上

出現したものについては手作業で可能な限り分類を行った。これにより、日本の所属機関名のうちでいずれかのセクターに分類不可能であったものは0.1%であった。

図4は、SCIに収録されている全論文の中での、日本の各セクターのシェアを示している。複数機関の共著の場合は機関数に応じて各セクターに分数で計上している。そのため日本と共著を行っている外国機関のシェアが、棒グラフの一番上に示されている。

まず、大学セクターをみると、データベース内での日本の大学セクターの論文シェアは、1982年の4.9%から上昇し続け、1999年の6.8%で最大値となり、それ以降は2002年の6.6%まで微減している。すなわち、世界全体の科学研究の中での日本の大学セクターの貢献度合いは1990年代の終わりまで増大し続けてきたが、近年は新興国の台頭等もあり、相対的に頭打ちとなっている。一方、日本の論文の中だけでの大学シェアをみると、大学セクターは1982年の77.1%から1992年には72.9%まで減少したが、それでも7割以上という高い値を維持し続けており、1992年以降はおおむね上昇に転じ2002年には74.1%となっている。また、全数カウントでみれば、1982-1995年では、日本の論文のほぼ80%に大学が少なくとも一機関は著者として入っており、それ以降は2003年の85%まで値がわずかに上昇している。これらの結果からは、知識生産における大学セクターの重要性は、他セクターの台頭のために分数カウントでは相対的な貢献度合いは減少しているように見えるが、全数



(SCI CD-ROM版1982-2003年を用いて筆者分析)

図4 セクターごとの論文数シェア

カウントでは8割の論文生産に関与し続けており、さらに近年は国内での大学セクターの貢献度合いが若干増加する傾向にあることが示されている。

一方で大学以外のセクターをみると、その多くが論文数を増加させている。特に顕著な伸びを見せたのは、理研、JSTなどに代表される準公的機関セクターであり、1996年の第一期科学技術基本計画の開始以降に論文数で1.9倍になっている。また、国立研究所は1980年代の基礎研究シフトにより論文数を継続して増加させており、同様に第一期基本計画以降には1.5倍になっている。逆に、民間企業は1992年から1996年をピークとして減少に転じている。この間、民間企業は自社内の基礎研究所を廃止や改組するなどして外部へ依存するよう転換し、バブル崩壊以降の経済不況がこの傾向を促進させてきた。論文分析の結果では、特に物理や化学分野がそれまで民間企業からの論文数が多い分野であったが、1992年から2002年の10年間でそれぞれ0.62倍、0.71倍へと減少している。

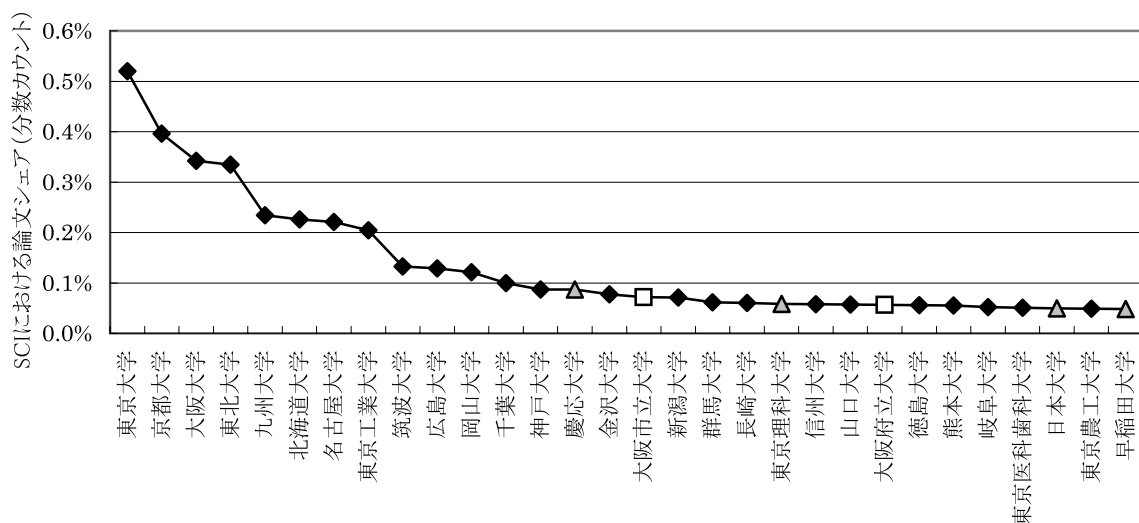
大学以外のセクターについて、国レベルの分析と同様に、被引用数によって論文をグループ化すると、準公的機関および非営利民間組織を除くセクターでは、国レベルと同様に50-75%の論文グループにおいて最もシェアが高い。一方で、準公的機関は上位10%論文において最も高いシェアを示しており、米国や英国などの傾向と類似している。これらの機関では、研究業績の高い大学研究者等をプロジェクトリーダーに据え、大学や企業、

公的研究機関からの期限付きの研究者の参加によるフレキシビリティの高い研究組織を構成するなどしている例も多く、その結果として実際に影響力の高い研究成果が産出されてきたことがわかる。

### 3.2 大学セクター内部での変化

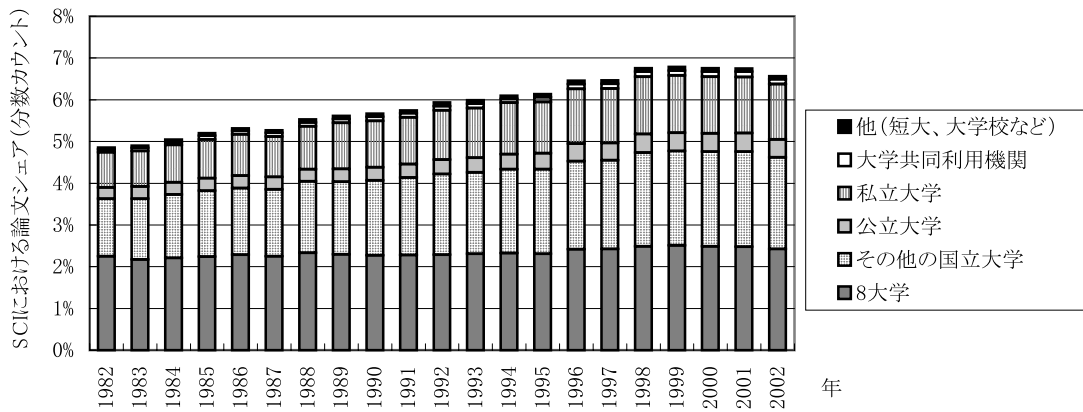
大学セクターは日本の論文の80%に関与しているため、図1で示した日本の国レベルの特徴は大学セクターの特徴を強く反映している。その内部構造を調べるため、国立大学(2002年で99校)、公立大学(75校)、私立大学(512校)、および、短大(541校)、高等専門学校(62校)、大学共同利用機関(15校)にさらに区分して分析する。ただし、大学を単純に論文数の多さで並べてみると(図5)、上位の8つの大学(いずれも国立大学)はその他の大学と比して論文数が多いため、国立大学99校から区分して図に示す。

図6は、SCIの全論文の中でのシェアを示したものである。SCI全体の中での国立大学のシェアは2001年には4.8%であり、日本の大学セクター論文のおよそ70%を産出している。また、論文数の多い8大学だけのシェアはこの20年間で2.2-2.5%の間で推移しており、8大学だけで日本の大学セクターの論文の36.7%を産出している(共著は分数カウントによる)。一方で、8大学以外の国立大学および、公立大学、私立大学のSCIにおける論文シェアは上昇しており、論文数で見れば、20年間で8大学の論文数は1.6倍に増加したのに対



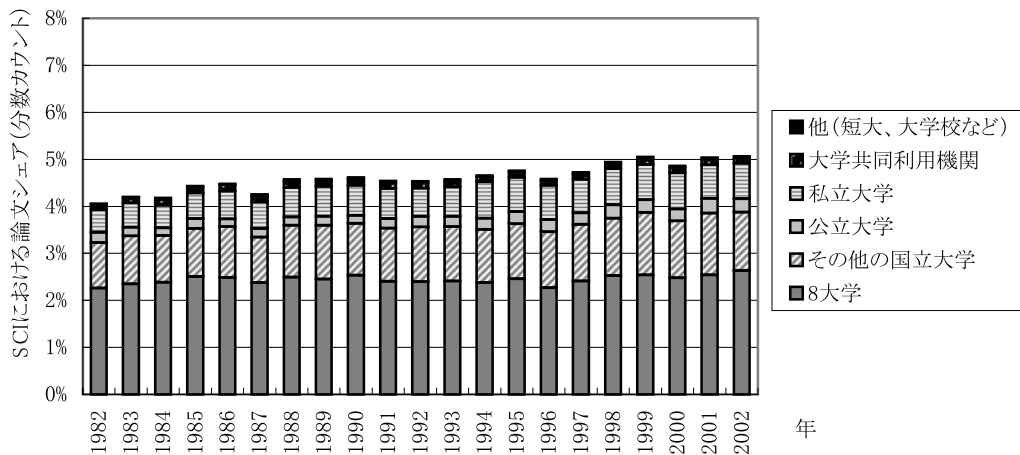
(SCI CD-ROM版を用いて筆者分析)

図5 SCIに論文数の多い大学(2001年)



(SCI CD-ROM版1982-2003年を用いて筆者分析)

図6 大学セクターの論文シェア (全論文を対象)



(SCI CD-ROM版1982-2003年を用いて筆者分析)

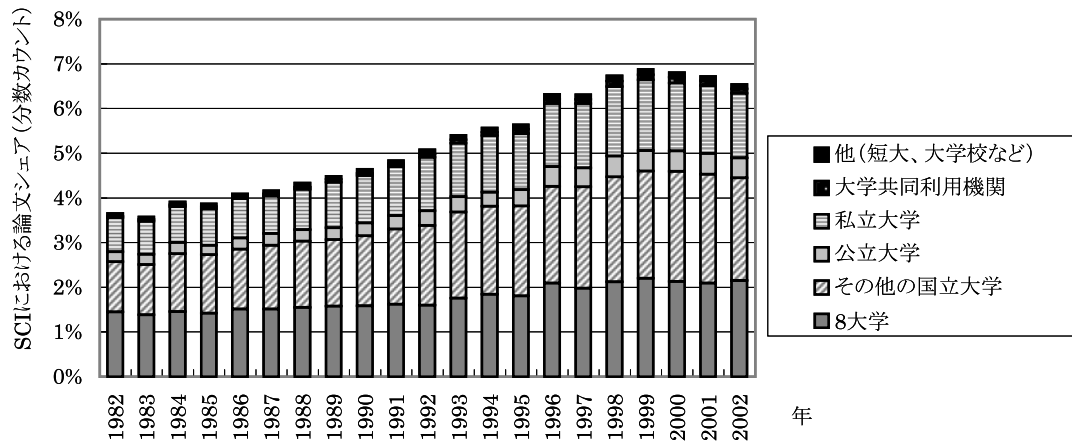
図7 大学セクターの論文シェア (引用数上位10%を対象)

し、それ以外の国大、公大、私大はそれぞれ2.3倍、2.5倍、2.3倍に増加している。つまり、この20年間の日本の論文シェアの伸びは、論文数の多い少数の大学よりも、それ以外の多くの大学から論文生産が増加したことがより貢献している<sup>9</sup>。

次に、被引用数によって論文をグループ化した場合をみる。被引用数上位10%の論文に限った場合においても同じ8大学が論文数の多い上位大学であるが、この8大学だけで日本の大学セクターの被引用数上位10%論文の50.4%を産出している(図7)。この値は、前述の全論文の場合(36.7%)

よりも高い。また、上位10%論文における8大学のSCIにおけるシェアは2.3-2.6%と、全論文の場合のシェアと変わらず、被引用数の高い論文もほぼ同じ割合で産出している(ただし、8大学の中でも、1999-2001年に、上位10%論文のシェアが全論文のシェアよりも高くなっているのは4大学のみである)。一方で、それ以外の大学を見ると、全論文における世界シェアと比して、上位10%論文でのシェアがだいぶ低いことが図からもわかる(ただし、これらの大学の中にも、論文数は少ないが、上位10%論文のシェアが全論文のシェアよ

<sup>9</sup> 本多, 慶伊 (2005) は、化学論文データベースであるケミカル・アブストラクトを分析し、1970年以降の日本全体の論文数増加率は東京大学の論文数増加率よりも高いことを明らかにした。他国ではトップ大学の論文数増加率のほうが国全体の論文数増加率よりも高いため、日本はトップ大学に限らずに国全体で論文産出を増しているという特異的傾向があることを明らかにしている。



(SCI CD-ROM版1982-2003年を用いて筆者分析)

図8 大学セクターの論文シェア (被引用数上位75-100%を対象)

り高い大学はある。岡崎国立共同研究機構の3研究所、核融合研究所などの大学共同利用機関4機関や、奈良先端科学技術大学院、北陸先端科学技術大学院、豊橋技術科学大学といった科学技術系の国立大学、および5つの私立大学である<sup>10)</sup>。

逆に、あまり引用されていない被引用数上位75-100%のグループを見ると(図8)、8大学は全論文のシェアよりも低い値で推移しており、1990年代前半に上昇し、2%程度という他の論文グループにおけるシェアと近い値になっている。一方、その他の大学のシェアは上昇しつづけており、特に1990年代に大きく増加し1999年にピークを迎えている。

### 3.3 集中と分散

これまでの分析結果からは、日本の2つの特徴がわかる。一つは日本の研究活動が多数の大学によって担われるようになってきたという、アクターの拡大(分散化)の傾向である。もう一つは、この拡大傾向がありながら、被引用数の上位の論文に限れば、少数の大学に変わらずに集中しているという傾向(卓越した研究の集中)である。この特徴を集中度の指標を用いて定量的に示してみよう。集中度の指標はいくつかのものが提案されているが、アクターの絶対数の増減とその中での相対的な集中度の両方に感度を有する指標としてハーフィンダール指標を用いる(芳鐘2003)。アクター*i*のシェアを $S_i$ とすると、ハーフィンダール指

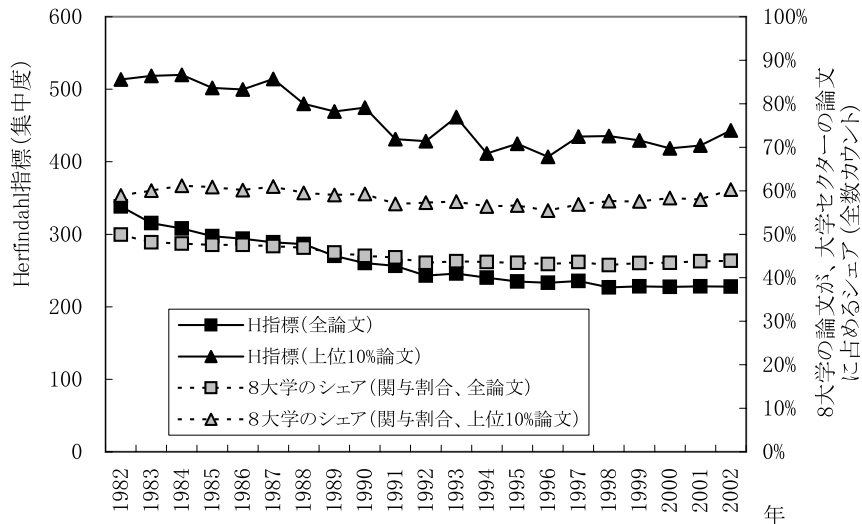
標は次の式で表される。

$$H = \sum S_i^2$$

ハーフィンダール指標の最大値は1であるが、値を見やすくするために指標の値を10,000倍して用いる(あるいはパーセンテージの値の二乗をとる)ことが多く、ここでもそのようにして示す。図9には実線で、大学セクターから生み出された全論文と上位10%論文のそれぞれにおける集中度*H*を示している。全論文の場合は、集中度は1990年代半ばまで減少し、1996年以降、 $H = 230$ 周辺のほぼ一定値をとっている。上位10%論文においても同様の傾向であり、集中度は1995年まで減少し、それ以降は $H = 430$ 周辺の値をとっている。この結果は、1980年代および90年代前半には多様な大学が研究活動に加わることによる分散化が進行したが、1990年代半ばにこの傾向が終了したことを示している。

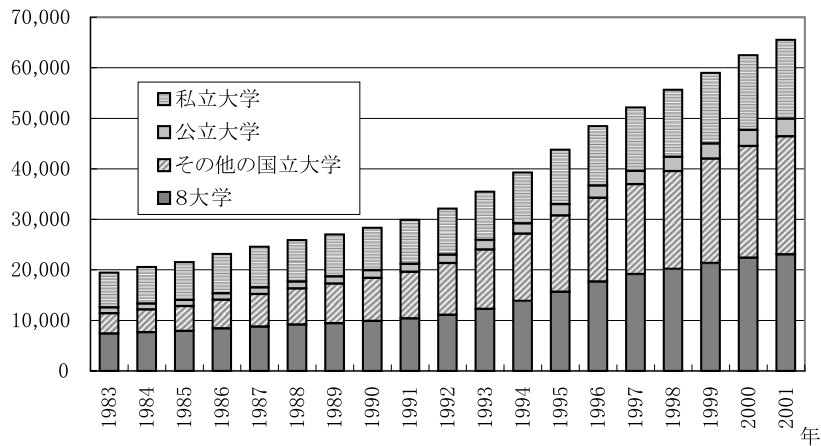
また、図9には、日本の大学セクターの論文の中で、論文数の多い上位8大学が著者に入っている論文のシェア(全数カウント)の推移も点線で同時に示している。上位10%論文においては、8大学のうち1大学でも関与した論文シェアは1990年代の半ばからわずかに増している。

<sup>10)</sup> 1999-2001年の3年間の論文数(分数カウント)が100本以上の大学のみを対象とした場合。



(SCI CD-ROM版1982-2003年を用いて筆者分析)

図9 集中度と8大学のシェアの推移



(データ:学校基本調査(文部科学省)および大学一覧(大学基準協会))

図10 博士課程学生数の推移

#### 4. 論文生産の特徴の背景

では、このような論文シェアの増加傾向や、論文生産者の分散と被引用度の高い論文の集中の傾向はなぜ生じたのであろうか。

第一に考えるべき背景は、大学の数および大学教員の数の増加に起因する分散化である。大学数はこの20年間で1.5倍となり、教員の数も1.4倍となっており、分散化傾向の要因となっていることが考えられる。しかし、教員数で見ると、国立大学に属している教員は39%であり、論文数の多い8大学に属している教員はわずか13%である。そのため、論文数の7割が国立大学によって産出されている現象を説明するためには十分ではない。

また、国立大学の教員数の伸びは20年間で1.2倍であり、論文数の伸びよりもだいぶ小さい。

一方で、教員以外の研究実施者としては大学院生、特に、博士課程学生を考えることができる。図10に示すように、博士課程学生数は、大学審議会が1991年に大学院生の10年間での倍増を答申して以降に増加し、1990年の28,354人から2001年の65,525人へと増加した。これにより、多くの大学で研究を実施する人的基盤が増強され、研究実施者の分散化を促進したと考えられる。しかし、このように博士課程学生の絶対数は増大したにもかかわらず、博士課程学生のうちで国立大学に在籍する者の割合は1980年代半ばには60%であったものが1995年以降は70%になり、さらに、論文数の

多い8大学へ在籍している割合はおよそ35%で一定である。すなわち、少数の大学に在籍する博士課程学生数が増大しているのである。この背景には、旧帝大を中心に大学院重点化が実施され大学院生の定員が増加したことがある。このような少数の大学における博士課程学生数の増加は、研究・教育環境が恵まれた大学に優れた学生が集中することをも可能とし、大学間での研究パフォーマンスに差を生じさせようの要因となっている。

また、人材と並び重要な資源である研究費についても、総額は増大してきた一方で少数の大学への集中は増している。Asonuma (2002)によれば、国立大学の研究活動のための基盤的な経費である教官当積算校費は、1980年代には実質値で総額は停滞、単価は減少していたが、1992年より増加し研究活動の基盤を強化した。その一方で、一般大学経費においても1980年代には「特別教育研究経費」などの選択的に配分される費目の額が増えるとともに、1990年代には大学院重点化を行った大学への配分額が増加し、差別化が進んだ。

また、競争的研究費の代表的存在である科学研究費補助金(科研費)は、その予算額の推移を見ると、1980年代は実質値でみれば毎年数%程度の増加であるが、1992年以降は毎年、前年度比10%近い伸びを示し、10年間で2.6倍に増加している。このような科研費の増額による採択者の増加は多様な研究者の研究活動を促進していった面がある一方、2001年には大学(短大等を除く)へ配られた科研費のうちで、80%が国立大学、49%が論文数の多い8大学へ配分されている。

また、1996年の科学技術基本計画実施以降には科研費以外の競争的資金制度がいくつか創設され、大学はよりいっそう複数の資金源から競争的資金を獲得できるようになった。しかし、科研費が少額の研究費を多数の研究者に分配する傾向が強いに対し、新設された競争的資金制度である「未来開拓学術研究推進事業」や「戦略的基礎研究推進事業」は、高額な研究費を少数のプロジェクト

へ配分するものであり、結果的に研究能力の高い大学へ重点的に配分されるという特徴を有する。そのため、2001年の競争的研究費の全体額(科研費はその中のほぼ半分の額を占める)では、その52%が8大学へ配分されている(内閣府2004)。

これらの科研費総額と博士課程学生数の2つの変数について、上位10%論文数との関係を大学を分析単位として図示すると、図11、12のようになる。これは、自然科学系学部を有し、SCIに論文が1本以上ある国立75校(総研大を除く)、公立39校、私立157校を対象としたものであり、規模による見かけ上の相関を防ぐためにいずれも教員数で除した値を用いている。また、SCIが自然科学系の論文のみを対象としていることから、科研費総額と博士課程学生数のデータも自然科学系のみを集計したものを用いており<sup>11</sup>、論文数と資源のデータには1年のタイムラグをおいた。図上のプロット円の大きさは教員数に比例する。

教員あたりの被引用数上位10%論文数と科研費総額との相関係数は0.76、博士課程学生数との相関係数は0.63であり(ともに $p < 0.01$ )、特に図11において科研費との間にはきれいな相関関係を確認できる。一方、図13、図14は、同様な方法で教員当たりの被引用数上位75-100%論文数と2変数の関係を示したものであるが、相関係数はそれぞれ0.61と0.55であり( $p < 0.01$ )、上位10%論文の相関係数と比べれば有意に低い(それぞれ $p < 0.01$ ,  $p < 0.05$ )。特に科研費額との関係において、被引用数上位10%論文では科研費額の増大に比例して論文数が増加するという線形の見られるが(図11)、被引用数上位75-100%論文では教員当たり科研費が百万円程度までは、科研費の増加に伴い論文数が増加するが、それ以上になると論文数の増加率が減速するという収穫逓減傾向が認められる(図13)。4つの各図に対して近似式を当てはめると、図13以外の3つの図では、線形式が対数式や累乗式よりも当てはまりが良いが、図13では収穫逓減傾向を示す対数式が最も $R^2$ 値が高

<sup>11</sup> 科研費に関するデータは、国立情報学研究所 NACSIS-IR の「科学研究費補助金採択課題データベース」を用いて、理学、工学、農学、医学、および複合領域において自然科学系と考えられる研究分野コードの課題のみを集計した。教員数と博士課程学生数については総務省の科学技術研究調査の個票を入手し自然科学系の学部のみを集計した。なお、これらデータの入手と相関分析は文部科学省科学技術政策研究所「基本計画の達成効果の評価のための調査」(H15-16年度科学技術振興調整費)の一環で実施したものである。詳細については科学技術政策研究所レポート No.79および No.88を参照。

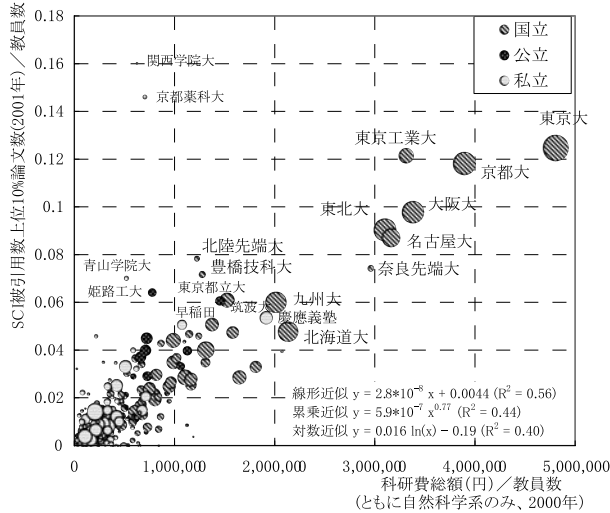


図11 科研費総額と引用数上位10%論文との関係

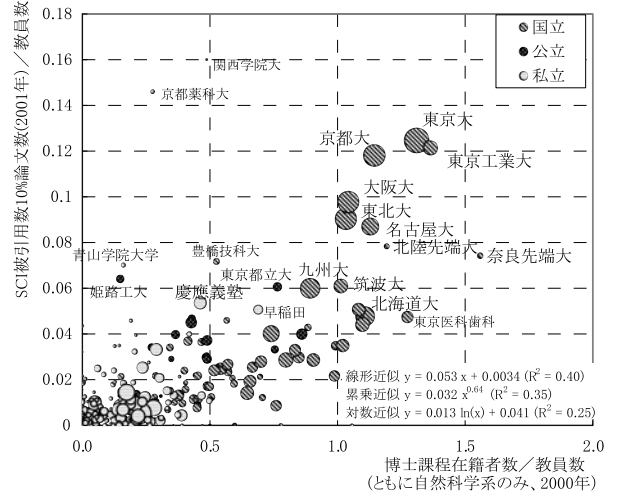


図12 博士課程学生数と引用数上位10%論文との関係

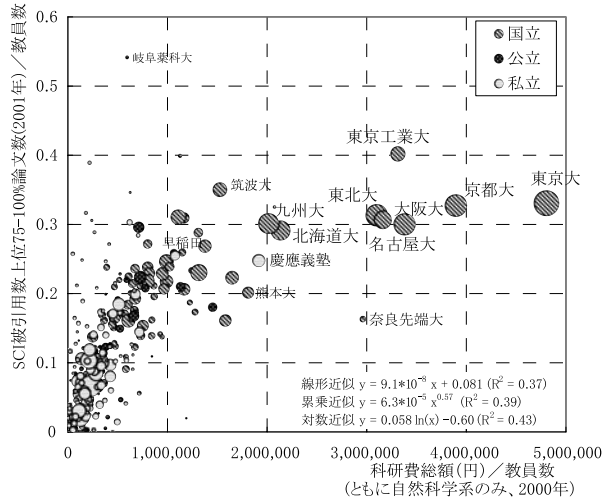


図13 科研費総額と引用数上位75-100%論文との関係

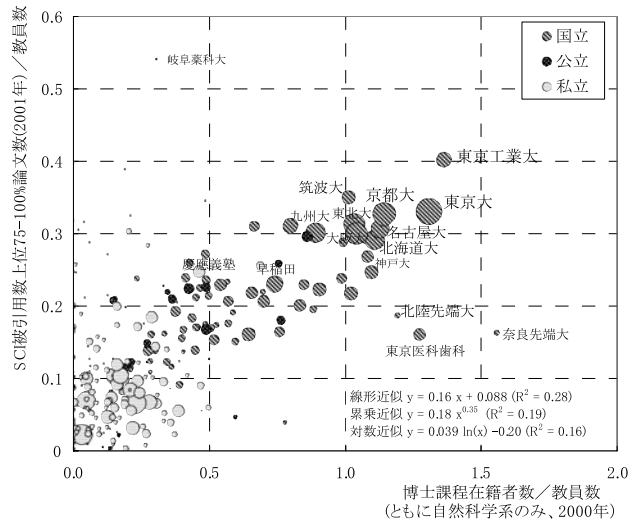


図14 博士課程学生数と引用数上位75-100%論文との関係

く当てはまりが良い<sup>12</sup>。

以上の結果をまとめると、教員当たり科研費額と博士課程学生数の多い大学ほど論文を多く生産するという傾向が全体的にあり、特に、教員当たり科研費額が一定程度以上に達した場合には、単なる論文総数の増加よりも、被引用数の高い論文を多く産出するという傾向があると言える。この結果は、研究費や博士学生の獲得が論文生産に好影響を与えるという可能性を示唆するが、逆に被引用数の高い論文を多く産出した大学がその後に研究費や博士課程学生を多く獲得できるという逆方向の因果関係も存在することは容易に予想され

る。このような両方向の因果関係によって、「富むもの(大学)がさらに富む」という傾向はマタイ効果と称されており (Merton 1967), 本分析結果は、そのような資源と論文の集中化が同時展開している様相の一時点を切り出したものとも考えられよう。

また、このような資源以外にも様々な要因がSCIにおける論文数の増加へ影響してきたことは想定される。一つには英文のジャーナルへ研究成果を公表するという、研究発表の文化自体の変化が挙げられる。たとえば、英語以外の論文も数多く収録している化学分野のデータベースであるケ

<sup>12</sup> 近似式の当てはめでは、被説明変数(教員当たり論文数)あるいは説明変数(教員当たり科研費総額あるいは博士課程数)の少なくとも一方が0の大学は除いている。

ミカル・アブストラクト (CAplus) を用いて検索すると、著者所属国に日本を含むジャーナル論文 (大学以外も含む) のうちで、英文で書かれたものの割合は、1982年には57.8% (20,832本) であったが、2002年には68.5% (53,935本) へと上昇している<sup>13</sup>。すなわち、少なくとも化学領域においては、日本語よりも英語で論文を発表する比率が上昇している。

また論文生産に影響を与える別の要因として評価制度の導入が挙げられる。前述のようにオーストラリアでは、論文数を指標の一つとする資金配分方式の導入によって平均被引用数の低い論文誌への投稿論文数が増加した (Butler 2003)。日本の場合には、そのような資金配分との結びつきはないが、1991年には自己点検・評価が導入され、1999年に行われた調査 (米澤2000) では90.2%が研究活動の自己点検・評価を行ったと回答し、82.2%が「著者や論文・報告書のリストアップ」を「大変重視した」あるいは「ある程度重視した」と回答している。2000年からは大学評価・学位授与機構が大学の研究評価を開始し、教員はこれまで以上に研究成果の産出を求められる環境におかれるようになった。このような評価の制度化が論文生産数の増加を後押ししてきた可能性はある。

これらの各種要因を総合的に踏まえて考察すると、1980年代までは校費などの基盤的資金の一定程度の配分や、科研費の増加、ならびに研究成果の国際雑誌への発表の一般化によって多くの大学が英文論文の産出数を増し、研究実施者の分散化が緩やかに展開した。1990年代に入ると大学院生数、基盤的資金、科研費が急増し、さらに評価制度の導入が教員の論文生産を誘引することにより、大学セクターの論文数の増大をもたらした。しかし、一方で影響力の低い論文数が増加するという結果も生んだ。同時に、資源はその総体の増加にもかかわらず少数の大学へ重点化することで、それら大学での被引用数の高い論文産出数の増加を支え、さらに1990年代後半になると科学技術基本計画に伴う科研費以外の競争的資金がいっそう少数の大学への集中を増し、論文生産においても分

散化傾向から集中傾向へと緩やかに転じたと考えることができる。しかし各種の資源や制度が実際にどの程度影響を及ぼしたかは、論文数を増やした研究グループや学科などを対象としたアンケート調査などのミクロレベルでの調査を実施して実態を把握することが必要であり、それは論文データを分析対象とする本稿の目的を超えるものである。

## 5. おわりに

本稿では、過去20年間の日本の論文を、被引用数の高さ、および論文を産出しているセクターや機関という2つの次元から内部構造を分析し、日本の研究活動の構造的な特徴を明らかにし、その背景を考察した。

ではこのような論文データの分析は、大学評価に対してどのような含意を持ち得るであろうか。方法論的含意としては、このような被引用分布を用いた分析を行うことによって、論文総数や被引用総数によって一律に大学をランキングするのは異なる形で、大学の研究活動の特徴や課題を定量的に示しうるということが明らかとなった。さらに、研究費や博士課程学生といった資源のデータは研究成果との相関を有しており、研究活動の状況を把握するために参照しうる指標であることも示された。

このような方法論的含意に加えて、分析結果からは、今後日本で実施する評価方法の方向性についても含意が得られる。研究者の研究活動が、評価方法によって少なからず影響を受けることを考慮すると、今後の評価で求められることは、論文の数をこれまで以上に増すことを強く誘因するような評価方法ではなく、研究成果の質を高めることを促す評価方法であると言える。日本は既に論文数では世界第2位を英国と争う立場にまでなったが、その過程で引用されない論文の数を急速に増してきた。被引用数が研究成果の質をある程度反映していると仮定すれば、このような現状において、単純に論文数を指標として評価することは有効ではなく、研究成果の質によりいっそう

<sup>13</sup> ここには企業による論文も入っており、それらは大学による論文よりも日本語の比率が高いことが想定される。本多、慶伊 (2005) がケミカル・アブストラクトを用いて、2003年の大学ごとの論文を調べた結果では、英文の割合が80%程度の値に集中していたという。



の重点をおいた評価方法の構築が期待される。

一方でこの分析結果が、評価に基づく資金配分に対して、どこまで含意を持ちうるかは不明である。本分析では、資源を多く獲得している大学において、特に被引用数の高い論文が多く生産されているという結果が得られた。しかし、この結果から、今後資源がより集中することで、さらに被引用数の高い論文が作成されうると単純に考えることはできない。なぜなら、これらの少数の大学の論文の多くは他の大学や機関との共著で書かれており、資源の過度な集中がそれらの大学や機関の研究基盤を弱体化させれば、共同研究関係が継続しづらくなる可能性もあるからである。

実際、1991年には、論文数の多い8大学が産出した論文15,432本のうちで、ほかの国内外の大学や機関との共著は46.2%であった。しかし、2001年には23,892本のうちの61.8%が共著となっている。被引用数上位10%論文に限ってみれば、2001年には68.1%が共著と、より高い<sup>14</sup>。このように半分以上の論文が他の大学や機関との共同で作成されていることを考えると、資源の過度な集中によってこれらの共同関係が壊れ、継続的な論文生産が困難になるリスクも存在していることを認識すべきであろう。

また、共著関係からは8大学以外の大学は、地域の公的研究所や病院との共同によって多くの論文を産出していることも明らかになる。これらには医療や環境分野の研究が多く、被引用数によって示される学術的影響ではない別の効果を生んでいることが予想される。このような研究を支援するための資金配分も別に十分考慮されなければならない。

本稿で分析を行った2002年以降には、21世紀COEプログラムをはじめとして競争的に配分される研究資金は増大し、また国立大学の法人化によって大学間の競争もいっそう顕在化するようになった。その中で国の研究実施構造がさらにどのように変化しているかを継続的に分析することは必要であり、特に研究資源の集中化の進展と多種多様な研究活動の確保とのバランスの関係を詳細に分析することによって、今後の政策形成や評価方法の形成に反映させていくことが求められよう。

## 参考文献

- Anderson, R.C., F. Narin and P. McAlister (1978), "Publication ratings versus peer ratings of universities", *Journal of the American Society for Information Science*, **29**, 91-103
- Asonuma, A. (2002), Financial reform in Japanese higher education, *Higher Education*, **43**, 109-126
- Butler, L. (2003a), Modifying publication practices in response to funding formulas, *Research Evaluation*, **12**, 39-46
- Butler, L. (2003b), Explaining Australia's increased share of ISI publications-the effect of a funding formula based on publication counts, *Research Policy*, **32**, 143-155
- 大学審議会 (1991) 「大学院の量的整備について」
- Godin, B. and Y. Gingras (2000), The place of university in the system of knowledge production. *Research Policy*, **29**, 273-278.
- 林隆之 (2003) 「ビブリオメトリクスによるピアレビューの支援可能性の検討—理学系研究評価の事例分析から」『大学評価』Vol.3, pp.167-187
- Hayashi, T. (2003), Effect of R&D programmes on the formation of university-industry-government networks: comparative analysis of Japanese R&D programmes, *Research Policy*, **32**, 1421-1442
- Hayashi, T. and H. Tomizawa (2006), "Restructuring the Japanese National Research System and its Effect on Performance" *Scientometrics* Vol.68 No.2 (印刷中)
- 本多卓也, 慶伊富長 (2005) 「自然科学系の研究評価」『大学論集』広島大学高等教育研究センター, 35, pp.403-418
- Irvin, J. (1989), "Evaluation of scientific institutions", *The Evaluation of Scientific Research*, John Wiley & Sons, 141-168
- 科学技術庁 (1999, 2000) 『我が国の研究活動の実態に関する調査報告』
- Katz, J.S., D. Hicks, M. Sharp, and B.R. Martin (1995), *The Changing Shape of British*

<sup>14</sup> 詳細については Hayashi and Tomizawa (2006) を参照。

- Science. STEEP special report no. 3*
- Kobayashi, S. and Y. Okubo (2004), *Demand articulation, a key factor in the reconfiguration of the present Japanese science and technology system*, *Science and Public Policy*, **31**, 55-67
- Merton R.K.(1968), "The Matthew Effect in Science: The reward and communication systems of science are considered", *Science*, **159**, 56-63
- 文部科学省科学技術・学術国際局監修 (2003)『全国試験研究機関名鑑』ラティス
- 文部科学省科学技術政研究所 (2004)『科学技術指標—日本の科学技術の体系的分析—』
- 文部科学省科学技術政研究所 (2005)「基本計画の達成効果の評価のための調査 科学技術研究のアウトプットの定量的及び定性的評価」NISTEP REPORT No.88
- 内閣府 (2004)『科学技術創造立国のための競争的研究資金の制度改革—政府研究開発データベースに基づく実態分析と改革設計 科学が変わる・大学を変える・日本は変わる』
- Rinia, E.J., Th.N. van Leeuwen, H.G. van Vuren and A.F.J. van Raan (1998), "Comparative analysis of a set of bibliometric indicators and central peer review criteria", *Research Policy*, **27**, 95-107
- Rinia, E.J., Th.N. van Leeuwen, H.G. van Vuren and A.F.J. van Raan (2003), "Influence of interdisciplinarity on peer-review and bibliometric evaluation in physics research", *Research Policy*, **30**, 357-361
- 調麻佐志 (2004)「学術論文データシステムを利用した研究評価」情報の科学と技術, Vol.54 No.6, 337-344
- Shubert, A. and T. Braun (1996), Cross-field Normalization of Scientometrics Indicators *Scientometrics*, **36**, 311-324
- 米澤彰純・編 (2000), 『大学評価の動向と課題』高等教育研究叢書62, 広島大学大学教育研究センター
- 芳鐘冬樹 (2000)「計量書誌学的分布における集中度：集中度の概念と指標の特徴」日本図書館情報学会誌, **46**, 18-32.
- Van Raan, A.F.J. (1996), "Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercise", *Scientometrics*, **36**, 397-420

[ABSTRACT]

Transition of Research Performance and Research Structure in Japan

HAYASHI Takayuki\*<sup>1</sup>, TOMIZAWA Hiroyuki\*<sup>2</sup>

This paper reveals changes in the performance and structure of research system in Japan by analyzing the past 20 years' bibliometric data. The analyses of the distribution of publications by citation frequency shows that although Japan gradually increased its production of highly cited publications, its share of low-cited publications is much higher than that of the former. Detailed analyses show that the top eight universities account for half of the highly cited publications in the university sector, while other hundreds of universities have radically increased their publications since 1990. The expansion of financial and human resources for research in the 1990s enabled new actors to be involved in scientific research. However, the resources were concentrated in a small number of universities, and the networks of collaboration around them were reinforced.

---

\*<sup>1</sup> Associate Professor, Faculty of University Evaluation and Research, National Institution for Academic Degrees and University Evaluation.

\*<sup>2</sup> Director, Research Unit for Science and Technology Analysis and Indicators, National Institute of Science and Technology Policy (NISTEP)

