

Title	米国特許における非特許引用の分析とその政策評価に対する含意
Author(s)	調, 麻佐志
Citation	年次学術大会講演要旨集, 25: 833-838
Issue Date	2010-10-09
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/9422
Rights	本著作物は研究・技術計画学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Science Policy and Research Management.
Description	一般講演要旨

○調 麻佐志（東工大）

1. はじめに

科学の技術に与える影響や科学と技術の連関を分析するためのデータとして注目を集めている非特許引用（NPR: Non Patent Reference）には、学術論文に加えて、会議録や書籍、特許抄録、さらに工業標準、製品パンフレットなど、あらゆる文献への引用が含まれている。その中で、学術論文の引用は科学研究と技術の連関を分析する上で要となるデータを提供するとされている。そこで、本論文では、米国特許に引用された非特許引用（以下、NPR と記す）、特に学術論文への引用に着目して科学研究と技術の連関の網羅的な分析を試行し、その政策評価に対する含意を検討する。

特許の包括的分析を実施する過程には様々な課題があるが、特に NPR の数は難題を突き付ける。なぜなら、論文引用の分析に先立って NPR とされた無数の文献から論文だけを抽出する time consuming な作業が求められるからである。もちろん、少数の NPR を抽出して人手によりこれら作業を実施することはさほど困難ではない。しかし、2005 年登録の特許で NPR は 55 万件を超えており、包括的な分析に利用可能な手法は限られてしまう。

そこで、本研究では、包括的な分析を試行するに先立って、各 NPR の論文レベルの同定に代えて学術誌レベルの同定（各 NPR を特定の学術論文データベースに収録された学術誌に掲載されているかを基準に学術論文であるかの判定を行い、あわせて当該掲載誌を同定すること）を行った。なお、基準となる収録誌データベースとしては Thomson 社の JCR2006 年版を用いている。

論文レベルの同定を行うには、「雑誌名」「巻号」「発刊年」「掲載頁」「(筆頭) 著者名」などの照合が必要であるのに対して、学術誌レベルの同定では最低限「雑誌名」の照合ができればよい。米国特許における NPR の書誌情報は自由書式であるため、各々の項目は表現形式が多様であるのはもちろんのこと、欠落もあるため、複数項目の照合では精度が急激に落ち、結果として自動化を試みても人手に頼る部分が残る。一方、「雑誌名」の照合だけであれば、記載形式にブレや誤記は当然あるものの、そのブレなどのパターンそのものを抽出して人手による照合を経てデータベース化することにより、バリエーションにも半自動的に対処できるという利点がある。しかも、NPR 内に含まれる学術誌名の抽出に関しては、専門業者が存在するという利点がある。実際、本研究においても、その種の業者の一つである ipIQ 社（旧 CHI 社）に抽出作業の下処理を委託している。

なお、分析に用いたデータは、米国特許による JCR 収録誌論文の引用、当該特許が属す技術分野（政策研究大学院大学鈴木潤教授提供の IPC コードをベースとした 27 分野および IPC コードの上位 4 桁に基づく技術分野区分を利用した）、特許の申請年月日、被引用論文の発刊年および所属する学術分野（2006 年版 JCR による 172 分野）などである。

2. 米国特許に引用された学術論文の学術誌レベルの同定

本研究が、対象とした米国特許（utility patents のみを対象とした）は 1995 年から 2005 年に登録されたものであり、引用はフロントページに掲載されたもののみを対象とした。表 1 は、対象となった特許の総数および総 NPR 数、推定論文数を示す。推定論文数は、対象となった NPR の中で ipIQ 社が論文として抽出したものの数を示しており、該当 NPR については雑誌名ならびに発刊年（記載がある場合）が自由書式のテキストから抽出され、付与されている。ただし、この推定作業の精度は十分に高いものの、誤りが含まれていることには注意が必要である。

これら論文と推定された NPR に対して付与された雑誌名は 12396 件であり、この雑誌名に対して JCR の 2006 年版（以下、JCR2006 とする）に収録されている雑誌名との対応をとることが、米国特許に引用された学術論文の学術誌レベルでの同定に必要な作業である（作業の具体的な内容については調（2008）を参照）。

表 1 米国特許の登録数と NPR 数

登録年	1995	1996	1997	1998	1999	2000
特許数	101417	109642	111982	147510	153482	157489
NPR 数	183470	237708	319917	441284	449298	465203
推定論文数	97638	133956	195392	273398	270973	270623
登録年	2001	2002	2003	2004	2005	
特許数	166033	167326	169020	164291	143804	
NPR 数	519514	548650	584916	556703	556886	
推定論文数	300750	303636	323920	290206	276682	

表 2 に同定作業を行った結果の一部を示す。本表では類似の先行研究と比較を行っており、表から明らかなように、全数を対象とした本研究の同定作業の結果は、限られたサンプルを対象とした先行研究のそれと同様の値（SCI 雑誌収録論文の占める割合において）を示しており、本研究の分析に用いるデータの信頼性が十分高く、かつカバレッジにおいて優れていることが示唆された。

表 2 同定された学術論文の比率

	期間	学術論文比率（うち SCI 収録誌論文）
Narin & Noma (1985)	1978-80	48% (37%)
Van Vianen et al (1990) *	1982-85	55.7% (46%)
Callaert et al. (2003)**	1996-2001	55% (NA)
本研究	2001-2005	NA (49.1%***)

*出願国がオランダに限られている

**標本（n=5000）に対する同定結果である

***SCI 収録誌ではなく、JCR 収録誌が対象となっている

3. 1 審査官による引用とそれ以外による引用

米国特許のフロントページに掲載された引用には審査官によるものとそれ以外（発明者、出願者、弁理士、等々）によるものの二種類が混在している。「審査官による引用と発明者による引用は、各々異なる論理に由来して（調・富澤・山下・玉田（2007）」おり、それらの違いについては注意を払う必要がある。幸いにも、近年の米国特許ではこの二種類の引用（審査官とそれ以外による引用）が峻別されており、これらの比率について比較を行った結果が以下の図 1 である。

図に示されるように、米国特許フロントページ中に収録された特許の引用に関しては審査官引用とそれ以外の引用が 1:2 の比率で混在しているのに対して、NPR や JCR 収録雑誌論文の引用についてはほぼすべてが審査官以外に由来する引用である。残念ながら、このことが直ちに学術論文の引用のほぼすべてを発明者が行ったことを意味するわけではなく、学術論文引用には「発明が発明者の頭の中に閃いたときに発明者の頭の中に存在していた論文等の知識以外のノイズが混入している可能性がある（ibid.）」ものの、審査官由来の「ノイズ」は相対的に小さいといえる。

3. 2 特許の出願年および登録年

本論文は特許に引用された学術論文の動向を通じた科学研究と技術の連関について分析と試行することを目的としている。したがって、ある時点で生み出された特許とそれが関連する学術論文とのリンクは最も重要な情報であり、原則として出願年ベースで特許の分析を実施することが望ましい。ところが、特定時点（たとえば、特定年）に出願された特許の全貌が明らかになるのは原理的には無限時間後、現実的にもそれなりの時間が経過してからであり、出願年ベースで分析を行う際には、その点の見極め

が必要である。調（2008）において 1995 年から 2005 年に登録された特許が各々の年に出願されたかをクロスで集計したところ、登録された特許の多くは出願年の翌年から 3 年後までに登録されており、登録数が最大となるのが出願後 2 年目であった。さらに、登録数が最大となる出願後 2 年目の登録件数を 100 として、登録特許数の出願後経過年による推移を示したのが図 2 である。グラフの形状に関する出願年による違いはわずかであるが、出願後 4 年・5 年後の登録件数が年を追って上昇する傾向がみられる。それでも、出願後 7 年以降は登録件数が激減しており（ピーク時の 2% 未満）、この部分のデータが欠けても全体の傾向に大きな影響を及ぼすことはないと推定できる。実際、2005 年までに登録された 1995 年出願特許の 98.8% が 6 年目の 2001 年までに登録されており、出願後 6 年までをカバーすれば最低限必要な精度は十分確保できると推測できる。そこで、次項の集計では、1995 年から 99 年出願分の登録特許を対象とした。

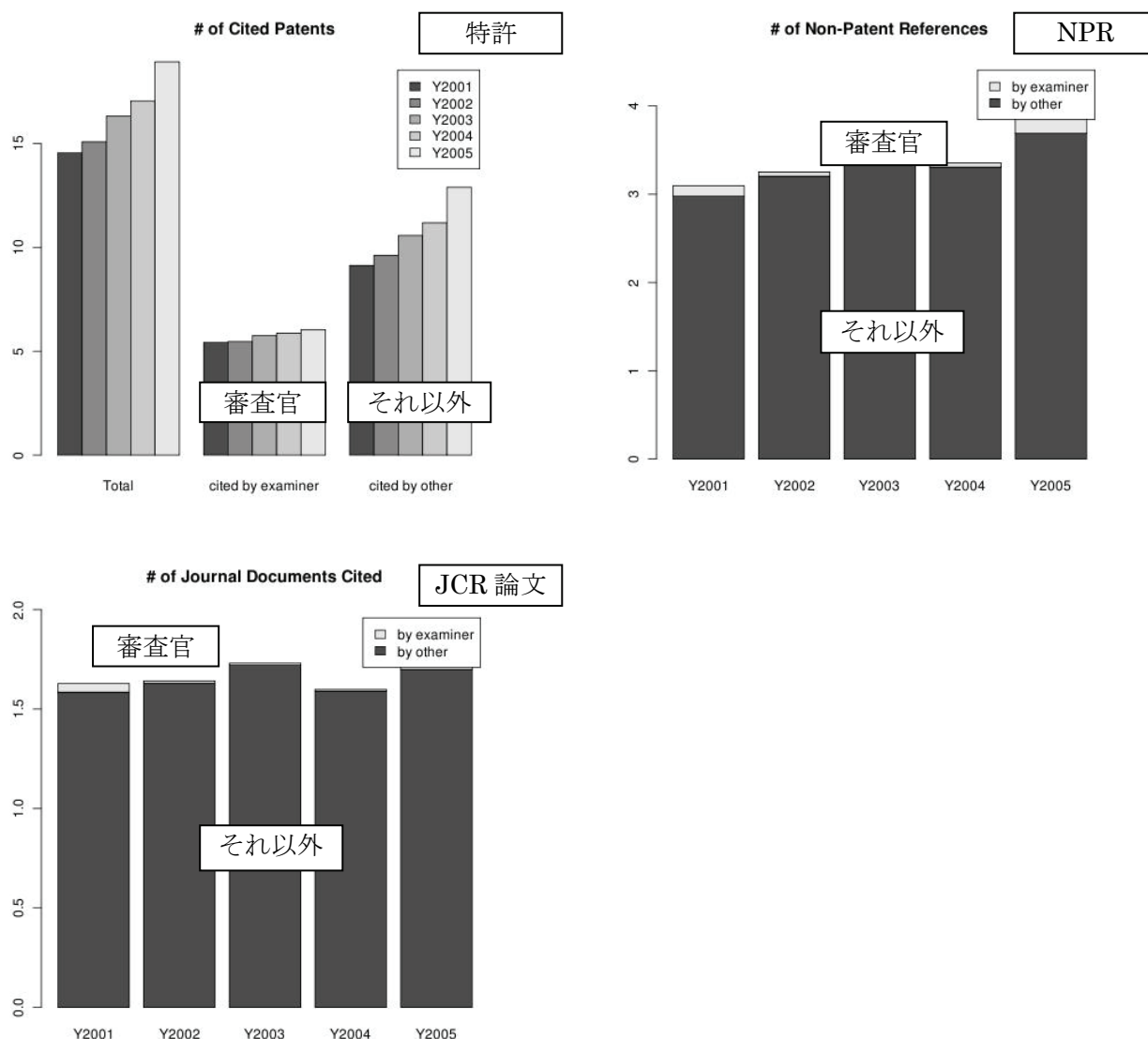


図 1 審査官による引用件数とそれ以外による引用件数の比較（平均値）

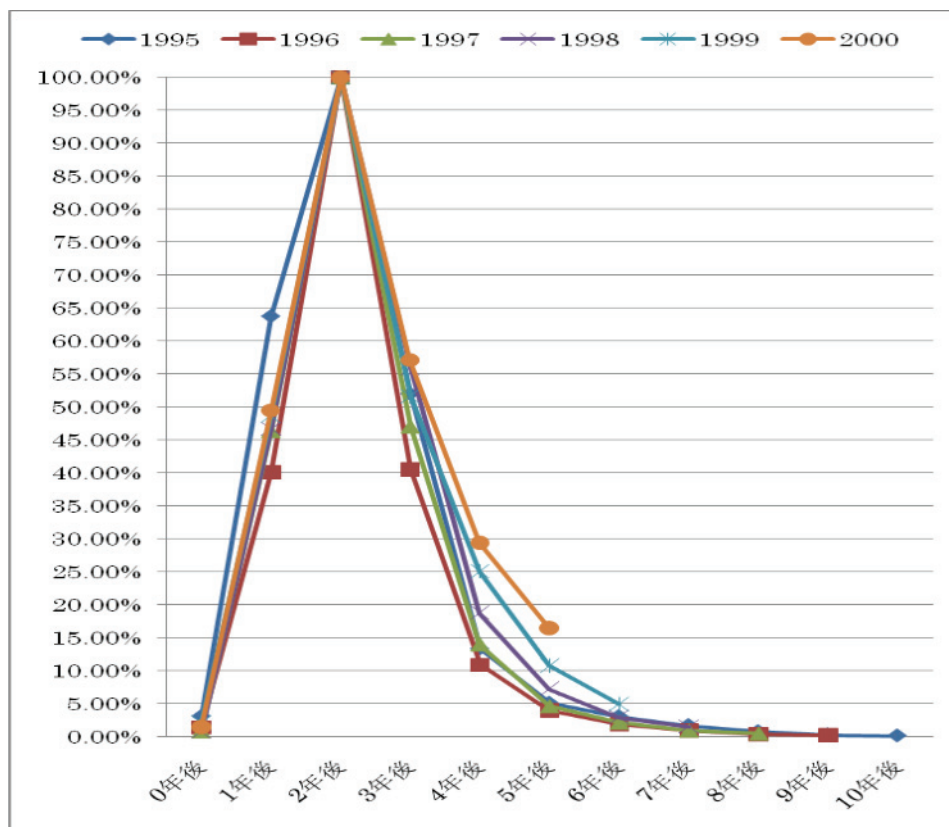


図2 出願後経過年による出願年別登録特許数の推移（出願後2年=100）

3. 3 各技術分野における技術と科学の連関

JCR 収録論文引用特許数の平均が年 500 件を超えている 7 つの技術分野（「医薬品」「化学肥料等」「家庭電器」「機械」「精密工業」「通信電気計測器」「窯業」）について、引用特許比率（当該分野の特許中論文を引用した特許の割合）、平均論文引用数、平均引用ラグ（特許出願年と引用対象となった学术论文の出版年とのラグ）の三軸でプロットしたのが図 3 である。引用特許比率と平均論文引用数の間には強い相関があり（1995 年に 0.92、99 年は 0.90 でいずれも 1%水準で有意）、一方で、平均引用ラグと他の 2 変数の間には有意な相関がない。また、各分野の位置関係については、「家庭電器」分野（図中、下赤線）を除けば大きな変動がないものの、「家庭電器」では平均引用ラグの著しい短縮がみられる。すなわち、比較的新しい科学の成果を取り込む技術が多く誕生したことが示唆される。

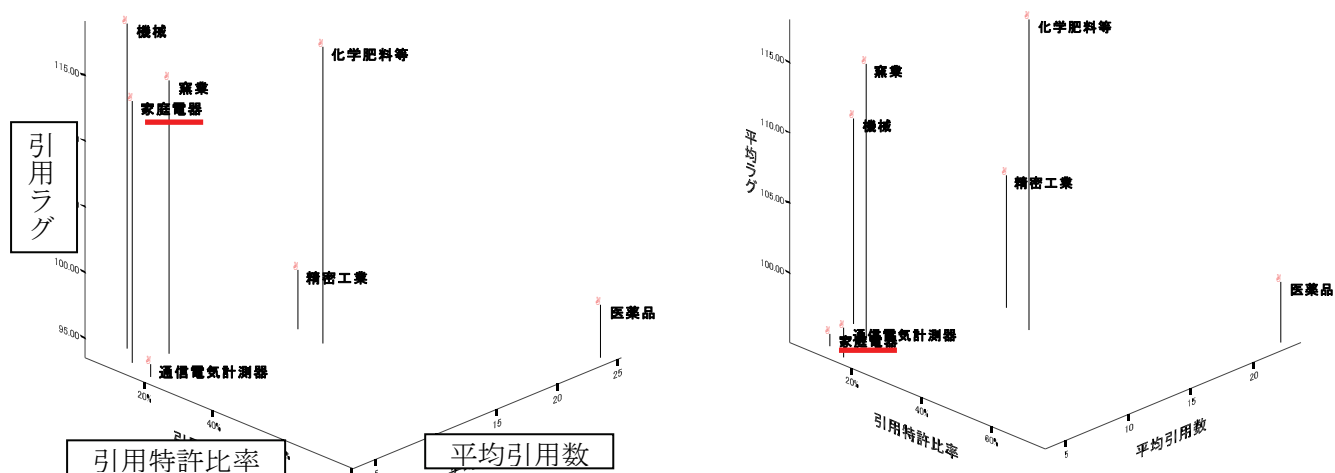


図3 各技術分野の学术论文引用の特徴（左 1995、右 1999）

この「家庭電器」分野における変化は、(1)技術と連関する科学知識の内容に変化が生じた、つまり、新しい科学知識を必要とする分野への参照が相対的に増加した可能性と、(2)これまで通りの科学分野と連関しながらも、より新しい知識を参照するようになった可能性の二つが主として考えられる。そこで、95年と99年の引用特許数および被引用JCR論文数に関する学術分野×技術分野のクロス表を使って、多次元尺度法により技術分野を配置したのが図4である。この図から判断する限り、各技術分野の配置に顕著な変動はなく、「家庭電器」分野における変化は、後者に起因するものと推測される。

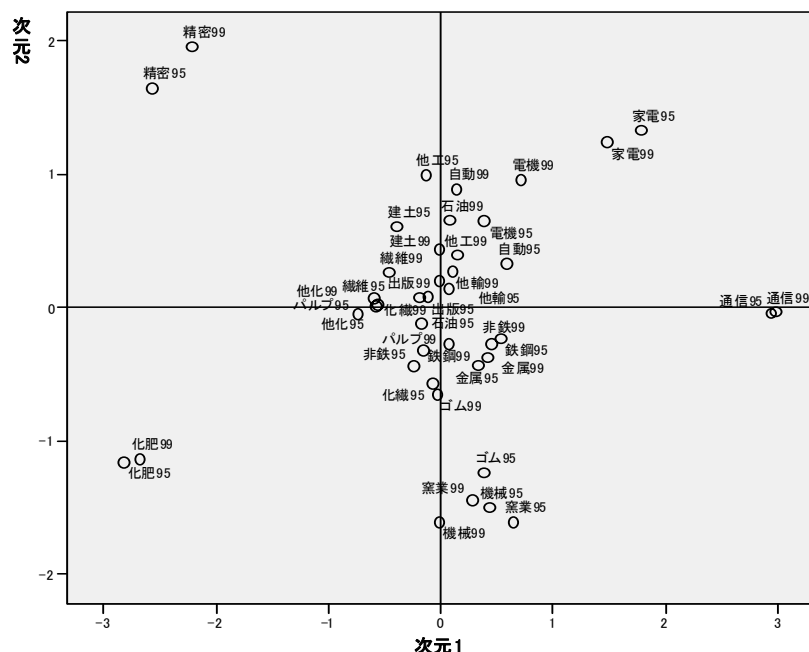


図4 被引用JCR論文数による（バイオ系分野を除く）技術分野の配置

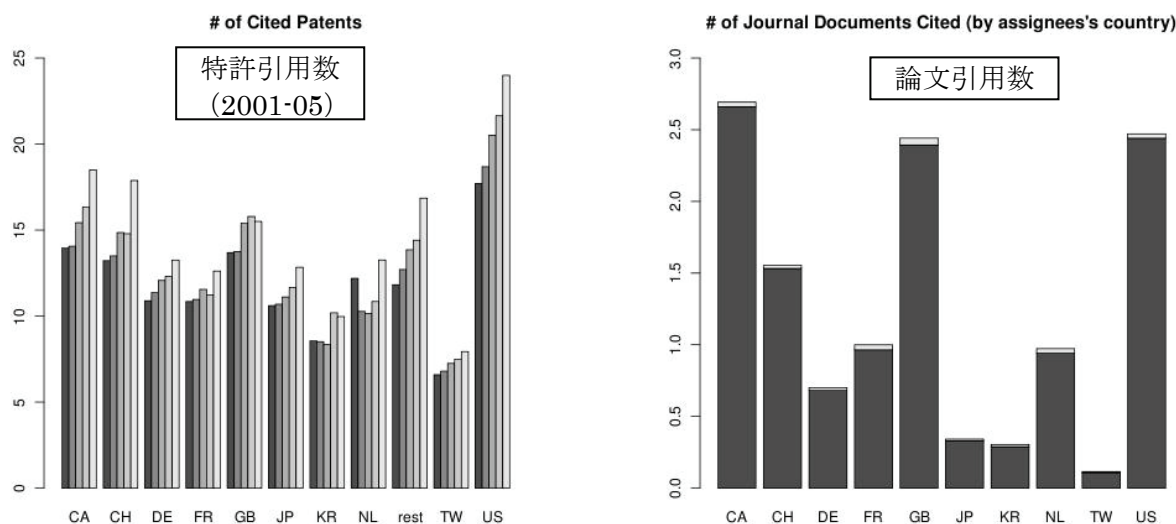


図5 各出願国の引用行動

3. 4 出願国による引用行動の違い

図5は、2001年から05年までの米国特許登録数上位10カ国（左より、カナダ、スイス、独、仏、英、日、韓、蘭、台、米、特許についてはその他を含む）に関して、平均の特許引用数およびJCR収録論文引用数をまとめたものである。特許の引用数と学術論文の引用数の傾向は国ごとで類似しており、引用数の多い（talkativeな）英および北米、中庸な欧州、引用数の少ない東アジア、という特徴がある。

この特徴は、各国の技術編成の違い、すなわち、学術論文引用の多い技術分野（医薬品）の登録特許が多い国と相対的に依存性の低い技術分野（半導体、電器等）の登録特許が多い国の違いによると考えられるかもしれない。しかし、それは一因であっても、十分な説明を提供するわけではない。なぜなら、技術分野を限定しても、各国の引用数の多寡に関する特徴がそのまま残っているからである（図6）。

さらに、特許の請求項数に示されるような一特許あたりのカバレッジの大小に説明を求めることも可能かもしれない。しかし、請求項数でも talkative な英および北米、中庸な欧州、静かな東アジア、という類似の特徴があるものの、請求項数でそろえて国別の比較を行うと同様の特徴が概ね残る（Shirabe, 2010）。すなわち、請求項、特許引用、論文引用に共通して情報の盛り込み度合いに特徴が存在する可能性が示唆される。

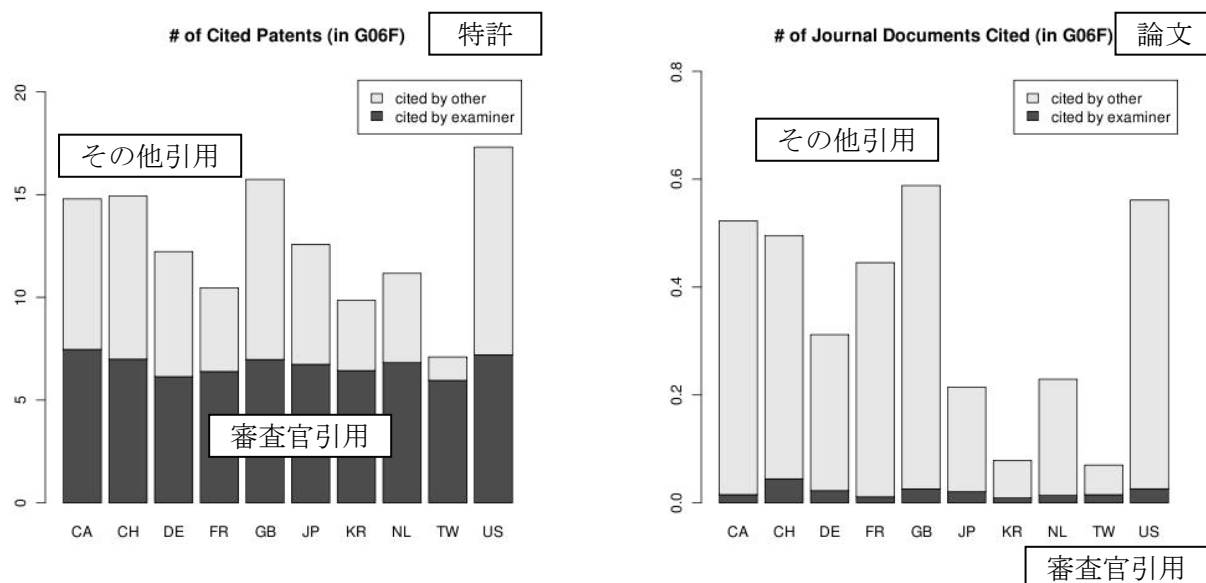


図6 半導体分野（G06F）における特許引用および JCR 収録論文引用の平均

4. まとめ

本論文では、米国特許に引用された学術論文を収録誌レベルで同定し、科学と技術の連関を包括的に分析する手法を検討するために、いくつかの分析を試みた。中でも、論文引用特許比率と論文平均引用数との間には強い正の相関がある一方で、平均引用ラグとこの2つの指数の間には相関関係はないことが示された。相関がある前2指数は特許と論文の結びつきの強度を示すものであり、特許（技術分野）側からみれば特許／分野の科学への依存度を、論文（学術分野）から見れば論文／分野の技術への貢献度を表すと考えられる。すなわち、この2つの相関は引用特許数や論文引用数をベースとした技術／学術分野の特徴を表す指標にある程度の妥当性があることを示唆する。他方、これら2つの指数と平均引用ラグとの間には相関がなく、たとえば、「医薬品」では科学と技術の連関は強く、しかも引用までのラグも短い、「機械」技術分野ではその逆であり、「通信電気計測器」では連関は強くないものの、引用までのラグは非常に短いといった特徴がある。つまり、連関の強さは科学が技術に転換するのにかかる期間の短さに直接はつながっていない。このことから、科学と技術の連関には少なくとも2種があり、この違いを考慮した政策評価・政策形成が必要といえる。

また、技術分野を超えて共有される特徴として、特許に盛り込む情報量に出願国による違いがあることも示唆された。これについてはさらに詳細な分析が求められるものの、この違いを前提とすれば、例えば、ある国の学術論文が当該国から出願された特許にはあまり引用されていないにもかかわらず、他国の特許に相対的にはよく引用されるといった現象が観察されたとしても、その政策的な含意は単純ではないと考えられる。すなわち、各国の技術編成や科学技術の連関の特徴などに加えて、ある種の文化的違いを考慮に入れた政策評価が求められることも考えられる。