



RIETI Discussion Paper Series 10-J-001

## 発明者による先行特許認識と特許後方引用

和田 哲夫  
学習院大学



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<http://www.rieti.go.jp/jp/>

## 発明者による先行特許認識と特許後方引用

和田哲夫（学習院大学経済学部）

### 要 旨

特許引用により、知識フローないし技術累積関係を計量することができる、という前提に基づいて、多数の実証研究が行われてきている。一方、ある発明にとって複数の先行技術が基礎にあった場合には、それら先行技術が等しく影響を与えたというよりは、寄与度に差があると考えの方が自然と考えられる。今まで明示的に扱われていないこの寄与度の差を検討するため、2008年に実施されたRIETI発明者サーベイの追加調査を活用し、発明者が指摘した重要先行特許の選択要因を分析した。その結果、先行引用特許の審査官による被引用（前方引用）特許数が多いとき、発明者も自発明にとって重要な先行技術であったとサーベイ回答にて指摘する傾向があることがわかった。さらに、2007年に先行実施された発明者サーベイの結果を再検討した結果、サーベイ対象特許が引用する先行特許それぞれが得た審査官被引用数のうち、最大値を取り出してサーベイ対象特許を代表させると、発明者の認識する先行技術の有無と正の有意な相関がみられることがわかった。従来の知識フロー計測ではノイズと考えられてきた審査官による特許引用は、企業間の技術的累積関係を計量する補助手段として有用性を持つことが確認された<sup>1</sup>。

キーワード：特許引用、累積的技術開発、技術経営、イノベーション政策

JEL classification: O33, O34

RIETI ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、(独)経済産業研究所としての見解を示すものではありません。

本稿は、独立行政法人経済産業研究所におけるプロジェクト「日本企業の研究開発の構造的特徴と今後の課題」の一環として執筆されたものである。有益なコメントと支援を与えて頂いた(独)経済産業研究所の及川耕造理事長、藤田昌久所長、長岡貞男教授と研究会のメンバー、ほか多数の方々に深く感謝申し上げます。

## 1. 研究目標

特許が私的技術情報を政策的に公開させようとしているのは、新たな技術的知識が個人や企業の中にとどまらず、地理的距離や企業組織の枠を超えて社会的に交換されることによって、イノベーションが促進されるという考えに基づく。近年、「オープン・イノベーション」という標語のもと、研究開発委託・技術ライセンス・人材異動などを通じ、技術的知識を外部から積極的に得ようとする動きが活発化している（Chesbrough, 2003）のは、個別企業の戦略としても、外部からの技術獲得が以前より意識され重視されている表れといえる。ただ、技術知識の社会的交換が進展すると、技術的知識の専有度を下げる可能性もあるので、社会的にはイノベーション促進の目的に正にも負にも作用しうる。累積的な技術開発過程と個別発明との関係を統計的に把握することと、そのための計量手段は、したがってイノベーション政策を考える基礎として重要である。

これまで、知識フローないし技術累積関係の統計的な追跡手段としてもっとも用いられてきているのは特許引用である。しかし、特許引用を用いた多数の実証研究において、一つ一つの引用の重み付けを考慮しないものが多い。つまり、ある特許が多数の後続特許から引用されている場合に、先行特許はどの後続特許にも同じように影響を与えたかのように計量上は取り扱われている。また、ある特許が複数の先行特許を引用している場合に、どれか特定の先行特許が飛び抜けて後続発明にとって重要な先行技術であったかもしれないが、その可能性は明示的に取り扱われていない。実際は、ある発明にとって複数の先行技術があった場合には、それら先行技術が等しく影響を与えたというよりは、寄与度に差があると考えの方が自然と考えられる。そうであれば、引用数を単純カウントすることによって技術依存関係を量的に計測する試みには修正する余地ないし必要があることになる。

本稿では、この問題意識に基づき、個々の特許となった発明の先行特許のうちには重要性の違いが存在するかもしれない、という前提を導入する。そして、2008年に実施された RIETI 発明者サーベイの追加調査を活用し、発明者が自身の発明にとって重要であったと回答した先行引用特許（後方引用特許、**backward patent citation**）の属性を、企業内外との関係や先行引用特許の重要度指標に照らして検討することとする。具体的には、以下の3つの目標を立てる。

1) 先行引用特許のうちどれかが発明者が重要であったか、サーベイ回答にて指摘する要因として、引用している特許の被引用特許（または前方引用特許、**forward patent citation**）の数が説明力を持つのではないか、という仮説について、審査官引用と発明者引用の区別に留意しつつ検証する。

2) 先行引用特許のうち、発明者の属する企業内の特許について、サーベイ上で発明者が重要であったと指摘する可能性が企業外特許よりも高いのではないか、という仮説について検証する。

3) 2007年に実施された発明者サーベイ（長岡・塚田 2007）において明らかになった特許引用数と先行特許認識の関係における偏り（和田, 2008）が、一般的な特定の特許への

引用集中とも関係があるか、検討する。これにより、特許引用を企業間の技術的累積関係を計量する手段として用いる上で歪みが存在するか、考察する。

以下では、まず次節で先行研究を概観し、現在まであまり取り組まれることのなかった問題の所在を述べる。また、2007年に実施された発明者（一次）サーベイの解析結果（和田，2008）が、別の側面からの研究動機となることを説明する。次に第3節では、特許引用の分布に関する先行研究を引きつつ、発明者の先行技術認識に関する考え方を述べ、先行特許引用の累積の影響と、組織内引用バイアス、さらに引用累積と組織内バイアスに関する仮説を導き出す。続く第4節では、検定方法と、分析に用いるデータとして、今回の発明者サーベイ（二次調査）の概要、特徴を述べ、5節で分析結果を解説する。最後に意味合いを議論し、残された課題について触れる。

## 2. 先行研究と研究動機

### 2-1. 特許引用の均質仮定

特許間の引用は、先行特許に対する後続特許の関連を示すものであり、同時に先行特許に比した後続特許の新しさを明示し権利範囲を画定する働きをもつ、と米国特許に基づく多くの実証研究では理解されている。発明時点で発明者が知らなかった先行特許も引用されることがあるなど、特許引用と、学術研究論文における引用関係とは大きな差異がある。しかし、ある特許を別の特許が引用しているとき、後続特許が実際に先行特許に関する知識を利用した蓋然性がある、つまり知識フローを反映しているのではないか、という一般論も受け入れられた。この根拠として、例えば、特許引用が実際に先行知識からのスピルオーバーを示すかどうか、について発明者サーベイによって調べた研究（Jaffe, Trajtenberg, and Fogarty, 2000）があり、特許引用のうち少なからぬ割合は実際の発明者の得た知識源を示している、と示されている。多くの引用特許を生み出した特許は、学術論文の影響と似た意味で技術的な影響が多かった蓋然性が高い、と理解できることになる。

特許引用が知識フローであることを前提とすると、研究ツールとして多方面への展開が可能となる。たとえば地理的な距離が近い方が、研究開発間の正の外部性を享受しやすい、ということの実証に使われた例がある（Jaffe, Trajtenberg, and Henderson 1993）。また、大学や公的研究機関の研究成果による外部インパクトに関する研究にも用いられた（Trajtenber, Jaffe, and Henderson, 1997; Jaffe, Fogarty, and Banks 1998）。さらに、企業間契約関係が知識フローと関係がある、という発見（Mowery, Oxley and Silverman, 1996, Oxley and Wada, 2009）もあった。

しかしながら、上述の研究の多くでは、特許引用のうち一定割合が知識フローを代理する蓋然性が高いと考えてはいるが、それら特許引用ごとに重み付けは与えていない。暗黙のうちに、それぞれの引用関係は同等、ないし確率的にみれば代表値を取り扱えば十分であるという前提に立っているものと考えられる。しかし、特許引用ごとの重みが均質に分布しているという保証はないし、平均など代表値で取り扱えるかどうかは不明である。

特許引用を知識フロー以外の考え方に基づいて利用した研究の多くにおいても同じ問題は存在する。たとえば、特許引用が経済学的な実証研究に使われるようになった端緒となる論文(Trajtenberg, 1990) のタイトルが“A penny for your quotes”ということからわかるように、特許の経済価値と特許引用の関係において、もともと特許引用ごとの違いはあまり考えられていなかった。特許引用を用いて特許ごとの技術的一般性(Generality)や獨創性(Originality)を計ろうとした論文(Henderson, Jaffe and Trajtenberg 1998)においても、計測しようとしている特許ごとの特徴の差異には着目しているが、それら差異を計測する手段としての引用ごとの違いには触れられていない。引用特許の権利者の集中度を「fragmentation index(権利分散インデクス)」と名付け、補完的技術の所有分散の程度を表す指標として用いた研究(Ziedonis, 2004)においても同じであり、特許引用ごとに重み付けが異なるのであれば、集中度の計測手段として問題があることになる。

## 2-2. 特許引用ごとの違いについて

引用されている特許の間に違いがあることに着目した研究は今までに全くないわけではない。代表的なものとして、発明者により付加された引用と審査官によって付加された引用の違いに関する研究がある(Alcacer and Gittelman, 2006, Alcacer, Gittelman, and Sampat 2009, Hedge and Sampat 2009)。これら研究では、地理的距離や技術的距離と、引用が与えられているかどうかの関係が審査官引用か発明者引用かによって異なるか、などについて解明されてきた。特許引用は、発明者以外に弁理士や特許審査官によっても付加されるため、技術知識フローの追跡・計量手段として歪みを持つ可能性がある、また審査官はイノベーション間の知識フローには関与しないから審査官引用は引用情報におけるノイズを生み出している、という可能性がこれら研究で指摘された。

このほか、上述した知識フローとしての特許引用の利用に際して、特許引用が均質ではなく、時間、組織属性、技術分野など特許階層によって引用頻度が変わることが分析に取り込まれている(Jaffe and Trajtenberg, 1996)。しかし、階層属性によって引用頻度が異なることは認識されているものの、個別の特許が引用する際に、どれか一つの先行特許が他と比べて大きな貢献度を持つ、というようなマイクロな引用の違いは捨象されている。

## 2-3. 発明者サーベイ(一次サーベイ)結果における疑問について

以上のように、先行研究では、特許引用どうしの重要度の違いに対して個別特許のレベルで関心を払っていないので、これ自体が未解明の分野として残されているといえる。それ以外にも、発明者からみて自企業内からの先行特許引用と企業外からの先行特許引用には意味に違いがあるかもしれない、という可能性を示唆する事実がRIETI発明者サーベイの一次サーベイ(2007年実施、長岡・塚田 2007)から以下の通り明らかになった(和田, 2008)。これは、特許引用を計量手段として用いる上で疑問を投げかける。

発明者サーベイの一次サーベイにおいては、そもそも先行技術に依拠しているかどうか、

先行特許が存在するかどうか、について、特許となった発明の発明者に調査し回答を求めている。これにもとづいて、企業内・外からの特許引用数（審査官引用および発明者引用）それぞれと、「先行特許に基礎をおいていた」と回答する確率の関係を探った結果、先行特許がある、と回答した発明者について、企業組織の内外からの特許引用数と、発明者が認識する先行特許の社内外の区別認識は良く一致した。また、発明者にとっての企業内からの特許引用数(企業単位でみた自己引用、**firm-level self-citation**)が多いとき、発明者の先行技術が存在した、という認識ともよく一致した。しかし企業外からの特許後方引用数は、発明者の認識とは統計的な関係がみられなかった。結果として、企業内外をあわせた後方引用特許の総数は、「先行特許に基礎をおいていた」という回答確率との関係も見いだせなかった。ここから、発明者の属する同一企業組織の中での先行技術利用と、企業組織外から得た先行技術利用が、特許引用数に同等に反映していない可能性が示唆された（和田，2008）。

得られた結果について、社内技術を他社ソースより重要視する認識バイアスがある、または知っていても他社ソース依拠を認めると訴訟リスクがある、ということが原因となっている可能性もある。しかし、発明者引用での結果ならばともかく、審査官引用で測定しても、企業内引用数のみが先行特許への依拠肯定の確率と相関を持つ、という結果が得られることに対して説得的な理由とは考えがたい。したがって、①企業内からと、企業外からの特許引用は、発明者認識に対してそれぞれ量的には異なる影響を与える、という可能性のほか、他に統計的に観察される表面上の認識バイアスを説明する原因があるか、検討する意味がある。次節では、その一つとして、②それぞれの引用ごとの重み、ないし分布が異なるために（単純な集計量として用いた場合には）発明者認識との関係が異なってみえる、という可能性を考える。

### 3. 仮説

直前で述べた②のアイディアは、ネットワーク一般に関する物理学研究から手がかりを得たものである。学術引用数の頻度分布がどのようなものであるか、物理学では古くから研究がされている。近年の展開として、ネットワークの接続次元の分布がベキ法則(**Power law**)に従うという考えが有力になっている。その代表例であるバラバシ・アルバートモデル（**Barabashi and Albert, 1999**）では、学術文献の引用数の分布やインターネット **WWW** の接続状況を例にとり、すでにリンクが多いノードには、より多く新たなリンクが付加される「優先的付加（**preferential attachment**）」によってベキ法則が説明できる、という。その優先的付加の単純なモデルでは、ある時点でのリンク数（特許引用では被引用数）に比例して新たな特許が引用を付加する、と仮定され、この結果スケール・フリー性が表れることが示されている。「スケール・フリー」とは、ネットワークの接続次元の分布がベキ分布に従うとき、ネットワークを特徴付ける特定のスケールを持たないことに由来する。

実際に、特許引用（前方引用）数の出現頻度を図1～図4に示す。図1は、日本の審査

官引用<sup>1</sup>を用いており、横軸は被引用回数（ $k$ ）の対数表示である。被引用回数 $k$ 回となるような特許が全体のうちどれくらいの割合で存在するか、を引用確率  $P(k)$ と解釈して縦軸にやはり対数で表示している。バラバシ・アルバートモデルに従えば対数軸では直線上に並ぶことになるが、実際に一定以上の引用回数を得ている特許の分布は直線的ともみることができ。

(図 1 を挿入)

日本の審査官引用データではなく、欧州・米国の特許データについては先行研究がある。図 2 は EPO の 1989 年引用データをプロットしたもの (Silverberg and Verspagen, 2007) であるが、やはり一定以上の引用回数を得ている部分に限っていえば直線的である。ただし、この Silverberg らによる研究では、図中の中央より左 (つまり低い引用回数) に関心を持つ場合には、log-normal に近いことを認めている。この考え方に近いものとして、被引用回数とともに発明の経済価値に関して log-normal の説明力が高い、という研究は他にも存在する (Harhoff et al. 1999)。

(図 2 を挿入)

日本の審査官引用の代わりに日本の公開・公告上の発明者引用<sup>2</sup>を引用ラグを考慮に入れずプロットしたものが図 3 であり、やはりバラバシ・アルバートモデルと似た分布が観察される。図 3 は引用ラグに関わらず用いているところ、引用・被引用の関係が長いか短い、をコントロールするため、引用・被引用の間隔を 5 年間に限定してプロットしても、やはり非常に似た分布が観察される (図 4)。バラバシ・アルバートモデルによれば、スケール・フリー性はネットワークの時間的成長によって生成される、とされるが、ネットワークの成長期間は、特許引用の場合にはそれほど長く要しないと推測できる。

(図 3, 図 4 を挿入)

このような全体のプロットから分布がべき法則に従っているかどうかを厳密に検定することは困難である。しかし、バラバシ・アルバートモデルのいうように既に引用を集めている特許が比例的に引用される確率が高いような「優先的付加」のメカニズムがもし存在するとすれば、それは発明者サーベイの時点で発明者が行った過去の関連特許評価でも表れていることが期待できる。優先的付加の示唆するように、特許付与時点までに各先行特許が集めた被引用数が引用確率に影響しているとすれば、サーベイ時点において発明者が

---

<sup>1</sup> ここで用いた日本の審査官引用のデータ源は 4-1-2 で後述。

<sup>2</sup> 同様に、ここで用いた日本の発明者引用のデータ源は 4-1-2 で後述。

重要と回答する先行特許は、さらに被引用を集めているであろう、と考えられるからである<sup>3</sup>。既に累積された被引用数の種類を審査官引用と発明者引用の2つに区別することによって次の2つの仮説を導く。

H1 A) 先行引用特許のうちどれが発明にとって重要であったか、発明者が指摘する要因として、引用している特許の審査官による被引用特許数が正の説明力を持つ。

H1 B) 先行引用特許のうちどれが発明にとって重要であったか、発明者が指摘する要因として、引用している特許の発明者による被引用特許数が正の説明力を持つ。

一方、バラバシらが当初提唱したスケール・フリーなネットワークでは、純粋に優先的付加のみによって重要度評価がなされ、引用が付加される、と仮定されており、ネットワークが空間的・組織的条件によって局所的に密なリンクを持つ可能性は捨象されている。しかし実際の発明者の多くは特定の企業に属し、また一定の人的関係の中で活動しているから、ネットワーク全体を等しく引用可能対象とするよりは、一定の人的範囲に優先的に重要度が認知され、引用が付加されていくのではないかと考えられる。具体的には、発明者の属する企業内の先行特許は、より自身の発明にとって重要であったと評価する確率が高まると考えられる。従って次の仮説を立てる。

H2) 先行引用特許のうち、発明者の属する企業内の特許については、発明者が重要であったとサーベイ上で指摘する可能性が企業外特許よりも高い。

先に述べた H1 は、発明者がサーベイ時点において先行引用特許の重要度評価を行った結果、被引用数と相関を持つであろう、という仮説であった。さかのぼって特許の出願から審査時点においても、すでに引用を集めている特許については比例的に重要度が高いと評価され、新たな特許引用の付加に際して優先的に引用される確率が高いような優先的付加メカニズムが存在すると仮定する。さらに、H2 について、発明者は、発明者の属する企業内の先行特許を優先的に引用する、と仮定する。これらの条件により、引用特許の構成するネットワークはスケール・フリー性とクラスタ性（同じノードにつながる2つのノードが、互いにも直接リンクを他と比べて多く持つ傾向があること）を備えるであろう。

クラスタ性を企業組織と関連して持つ場合、自社内の先行特許2つが、それぞれの間でも引用関係があり互いに技術的に近いものである確率が、社外の任意の先行引用特許2つの場合よりも高くなる。そして、社内からの重要技術を引用している場合、周辺特許

---

<sup>3</sup> 発明者のサーベイ時点での評価は、発明や特許出願審査までの間から相当の年数をおいており、特許取得時点で先行特許を選んだ基準をさかのぼって検定することは困難であるため、サーベイ時点での被引用数と重要性評価の関係から類推しようとしている、ともいえる。

をあわせて引用する確率は、社外からの引用の場合に比べて高くなるであろう。したがって、同程度の被引用数を持つ特許を引用している場合に、発明者の属する同一企業内からの場合に比べ、社外からの引用の場合には関連特許の引用数が少なくなるであろう。同様に、被引用数で計った技術的重要度を等しく保った比較において、社内での後続発明に対しては社外に比べて前方引用が多く発生するであろう、と予想できる。

このようなクラスタ性があれば、2-3節において述べた RIETI（一次）発明者サーベイ結果の疑問（和田、2008）への説明となりうる。昨年度の結果として、同一企業組織内からの特許引用数が多いほど、「先行特許に基礎をおいていた」と回答する確率も高いことが統計的に有意に確かめられた。しかし企業組織外からの特許引用数が多いことと、「先行特許に基礎をおいていた」と回答する確率とは関係が見いだせなかった。この事実が、審査官引用、発明者引用ともに、観察された。審査官という客観的な情報を把握している立場から付与された引用においても、組織内からの引用が多いときだけ、発明者の先行特許依拠の意識と符合することは簡単には説明がつけられなかった。しかし、引用特許群が構成するネットワークが企業組織と関係してクラスタ性を持つ場合には、企業内からと企業外からの引用数を比較したとき、発明者にとって先行特許が存在したかどうかという認識との相関の度合いが異なることと矛盾しない。

ここでは、クラスタ性が実際の組織などとの関係でどのように分布し存在するか、詳しく検討することは将来の課題としておくが、スケール・フリー性の結果としての引用数の極端な偏りとクラスタ性が同時に作用することによって、（2007年発明者サーベイでみられるような）発明者の評価する先行特許の存在認識と先行特許の被引用数との関係にどのような追加的な特徴が表れるか、を考える。

ある特許が多くの先行特許を引用しているが、スケール・フリー性にみられるように、そのうちのごく一部の先行特許だけが重要な役割を果たし、発明者にとって「たしかに先行特許が存在した」と感じさせる場合を検討する。その重要先行特許は、非常に多くの被引用数を持つであろうし、その存在は H1 の予想するように発明者の認識にも反映しやすいだろう。しかし、そのような多くの被引用数を持つ先行特許が社内の先行特許であれば、引用する側からみたとき、関連して（後方）引用特許も多くなることが期待できる。ところが、社外の先行重要特許が多くの被引用数をもっている場合でも、社内ほどには後方引用特許数が多くないであろう<sup>4</sup>。もしこのような差異がある場合でも、後方引用群は総体として発明者に先行技術認識を与えているわけではなく、個々の非常に重要な個別先行特許が問題であろう。そして、そのような非常に重要な個々の先行特許は非常に多くの被引用数を持っているはずだから、引用特許それぞれの持つ被引用数のうち、最大値だけを取り出して

---

<sup>4</sup> したがって、社内からの重要特許引用であれば、発明者が「たしかに先行特許に依拠した」と答える場合とは、社内の後方引用特許数が多い場合と正の相関関係を持つと予想できるが、社外からの重要特許引用であれば、その重要特許の被引用数と関係が表面上みられなくなる、という予想をここでは立てている。

サーベイ対象特許の先行引用特許群を代表させたとき、社内・社外どちらの場合でも発明者の先行特許の存在認識に表れるのではないかと予想できる。仮説として述べると、次のようになる。

H3A) 発明者が「先行特許に基礎をおいていた」と回答する確率に対して、先行引用特許の被引用の最大値が説明力を持つ。

このようなスケール・フリー性とクラスタ性があまり重要でないとすれば、後方引用特許群全体が及ぼす重要性は、それら引用特許群の総体として重要性を表すインデックスを用いればよいはずである。たとえば、後方引用となっている先行特許それぞれの被引用数が H1 によって確かめられたのであれば、引用されている後方引用全体について被引用数総和をとれば説明力を持つかもしれない。この推論に基づくと、次の予想を立てられる。

H3B) 発明者が「先行特許に基礎をおいていた」と回答する確率に対して、先行引用特許の被引用の総和が説明力を持つ。

#### 4. 検証方法とデータ

##### 4-1. データソース

##### 4-1-1. サーベイデータと特許データ

H1 および H2 の検証に用いたのは、特許発明者に対する RIETI 発明者サーベイ（2007年実施）の回答者に対して、さらに追加質問を行った2008年の追加サーベイのデータである。いくつかの追加質問に対して発明者の貴重な協力が得られた。このうち、本研究に用いたのは、特許の審査官引用および発明者引用（Tamada et al. 2006）によって提示した引用特許に関する発明者の回答である。この追加サーベイでは、一次サーベイの対象となった特許が引用している特許を提示した。その引用特許リストのうち「当該発明にとって相当程度に基礎となった先行特許があれば、その選択肢番号をお答え下さい。なお相当程度に基礎となった先行特許がない場合、それに最も近いものを選んでください。」という質問に対する回答を用いた。

##### 4-1-2. 特許データおよび特許引用データ

日本特許に関する審査官引用と、日本特許に関する発明者引用の2種類を用いる。日本の審査官引用は、2002年から2005年にかけて、後藤晃教授を中心として特許庁整理標準化データから作成された特許データベースの2005年12月時点版に基づく（和田・後藤 2003）。当該データベース中の特許引用データから計数されているが、企業内・外は出願人コードを基準として定義されている。日本特許に関する発明者引用は、玉田俊平太教授の考案に

よる特許公報明細書に含まれた日本の特許を用いており (Tamada et al. 2006)<sup>5</sup>引用する特許を基準として1976年から2005年までを含んでいる。企業内外は、やはり上記の整理標準化データにおける出願人コードを同じく基準として定義されている。

#### 4-2. 推定したモデル

##### 4-2-1. 仮説1および2の検定方法と使用データ

まず、H1A、H1BとH2に関する検証を行うため、発明者が後方引用特許それぞれについて相当程度に基礎としたかどうか、に関するバイナリ変数を被説明変数にとり、次のようなLogit質的検定モデルを用いる。

$$\begin{aligned} \text{DEPENDENCE}_{ij} = f(\text{WITHIN\_FIRM\_B\_CITE\_DUMMY}_{ij}, \text{INVENTOR\_BCITE\_DUM} \\ \text{MMY}_{ij}, \text{E\_SIBLINGS}_{ij}, \text{I\_SIBLINGS}_{ij}, \text{RECENTLY\_ADDED\_RATIO}_{ij}, \text{WITHIN\_FI} \\ \text{RM\_E\_SIBLINGS\_RATIO}_{ij}, \text{WITHIN\_FIRM\_I\_SIBLINGS\_RATIO}_{ij}, \text{TOP\_OF} \\ \text{\_LIST}_{ij}, \text{B\_CITE\_LAG}_{ij}, \text{USCAT}_{nj}) \end{aligned} \quad \dots (i)$$

ここで従属変数 *DEPENDENCE<sub>ij</sub>* は、発明者に引用特許リストを示し、その「特許リストのうち、当該発明にとって相当程度に基礎となった先行特許」または「相当程度に基礎となった先行特許がない場合、それに最も近いもの」に対する回答である。選択された特許についてのバイナリ変数であり、発明者からみて、その特許 *i* が引用する後方引用特許それぞれのうち特定の特許 *j* が先行特許として重要であった、と発明者が回答した場合に1をとる。なお、「相当程度に基礎とした」という回答に限らず、それに近いものを選択してもらったのは、複数の引用特許から相対的にみてもっとも重要な先行技術であった、と発明者が判断したものを選ぶ基準を検証する目的がある。このため、サーベイで発明者に複数の引用特許を提示できた場合にサンプルを限定している。このような特許と引用特許の組み合わせとしては3393ペアが計量検定に利用可能であり、そのうち665個の引用特許が実際に「相当程度に基礎とした」またはそれにもっとも近い特許であった、という回答が得られた。

説明変数のうち最初の *WITHIN\_FIRM\_B\_CITE\_DUMMY<sub>ij</sub>* は、引用している後方引用特許も発明者の属する企業と同じ所属発明者であったときに1をとるバイナリ変数である。H2が正しければ、この係数は正となる。

続く説明変数 *INVENTOR\_BCITE\_DUMMY<sub>ij</sub>* は、やはりバイナリ変数で後方引用特許が発明者引用であるときに1をとる。発明者がみずからの発明にとって重要だったと指摘するのは、審査官による引用特許よりは発明者による公報上の引用である可能性が高いと予想できるので、それが事実であれば係数は正となる。

<sup>5</sup> 鈴木潤教授および内藤祐介氏を通じて利用を許可された。研究の基礎を与えた玉田教授、長岡教授に対してとともに、深く謝意を表す。

H1A, H1B の2つの仮説を検証するための変数が  $E\_SIBLINGS_{ij}$  と  $I\_SIBLINGS_{ij}$  の2つの説明変数であり、それぞれ後方引用特許  $j$  に付された被引用特許の数を審査官引用 (examiner citation) と発明者引用 (inventor citation) の2つの方法でカウントしたものである。サーベイ対象である発明者が生み出した特許が引用している先行特許 (後方引用) を事後的に引用する特許であり、前方引用カウントである。サーベイ対象の特許からみれば同一の先行特許から平行引用している「兄弟」の位置にあるので、”sibling”とここでは名付けている。従来の特許引用数の研究 (Trajtenberg, 1990) から意味づければ、後方引用特許の技術的重要性ないし経済的重要性を表していることになる。

これら引用している後方引用特許の被引用については、元のサーベイ対象特許の発明者と同じ企業にどれくらいの割合が属しているか、検討する価値がある。なぜなら H2 が予想するように「発明者の属する企業内の先行特許について、サーベイ上で発明者が重要であったと指摘する可能性が企業外特許よりも高い」が支持されたとして、その理由が同一社内にある特許の重要性がより高く評価されるからだとすれば、その引用特許も企業内にある場合に高く評価されるかもしれないからである。審査官引用の場合と発明者引用の場合にわけ、後方引用特許の被引用数全体に対する社内被引用の割合として定義した。より具体的には、特定の後方引用特許  $j$  について事後に付された審査官引用特許数のうち同一企業内のものを  $WITHIN\_FIRM\_E\_SIBLINGS_{ij}$  とすると

$$WITHIN\_FIRM\_E\_SIBLINGS\_RATIO_{ij} = \frac{WITHIN\_FIRM\_E\_SIBLINGS_{ij}}{E\_SIBLINGS_{ij}} \quad /$$

とした。同様に、発明者による引用数の割合は、特定の後方引用特許  $j$  について事後に付された発明者引用特許数のうち同一企業内のものを  $WITHIN\_FIRM\_I\_SIBLINGS_{ij}$  とすると

$$WITHIN\_FIRM\_I\_SIBLINGS\_RATIO_{ij} = \frac{WITHIN\_FIRM\_I\_SIBLINGS_{ij}}{I\_SIBLINGS_{ij}} \quad /$$

で与えた。

以上の説明変数は、引用特許について、およびその被引用数と企業組織の内外区別について定義されたものであるが、被引用数を発明者が評価する時点の影響も検討した。バラバシらの提唱するように、ある時点でのリンク数 (特許引用では被引用数) に比例して新たな特許が引用を付加する、というような優先的付加メカニズムが、発明者による過去の関連特許評価にも反映するだろう、という予想が H1 の根拠となっている。つまり、発明者が特許を出願し審査された時点よりもずっとあとのサーベイ時点で再評価を求めているので、サーベイ時点に近い最近に多くの被引用を集めている後方引用特許がより重要だ、と回答される可能性がある。この効果が存在するかどうかをみるため  $RECENTLY\_ADDED\_RATIO_{ij}$  を発明者引用について定義した。発明者引用データが2005年までのサンプルであり、引用数を観察する期間が特許によって異なるので、発明者被引用数から、出願か

ら5年間の被引用を差し引き、さらに「2006 - (出願年) - 5」の年数で割る<sup>6</sup>ことによって最近の「引用特許の被引用数を近似し、さらにそれを当初出願から5年間の被引用で除することによりごく最近に被引用が増えたかどうかをみている。

このほか、サーベイ対象の特許と、それぞれの先行後方引用特許との間にどれくらい時間間隔が開いているか、をコントロールするため、*B\_CITE\_LAG<sub>ij</sub>* も追加した。あまり古い特許は発明にとって重要性が下がると予想されるので、この予想が正しければ係数は負となる。

サーベイでは、後方引用特許リストの中から発明者が選択して回答を寄せたが、後方引用特許のうちどれがもっとも重要かの判断は難しいケースがあろう、と予想された。そこで、提示した特許リストのうちもっとも上にある後方引用にはダミー変数 *TOP\_OF\_LIST<sub>ij</sub>* を付した<sup>7</sup>。リストの上にある特許が選ばれやすい、という選択バイアスがある場合には、このダミー変数の係数は正が予想される。

このほか、技術分野ごとの特殊性をコントロールするため、*USCAT<sub>n</sub>* として米国特許分類にもとづく最も広い6分類の分野ダミーを付した。この区別はNBER特許データベース(Hall, Jaffe and Trajtenberg, 2001)から採用され多くの研究で利用されている。

#### 4-2-2. 仮説3の検定方法

H3 に関する検証方法は、昨年のディスカッションペーパー (和田, 2008)において扱った方法に沿っており、発明者の先行特許への依拠意識に関する Logit 質的検定モデルを用いる。

$$\begin{aligned}
 \text{DEPENDENCE}_{i,t} = f(\text{MAX\_WITHIN\_FIRM\_E\_SIBLINGS}_{i,t}, \\
 \text{MAX\_OUTSIDE\_E\_SIBLINGS}_{i,t}, \\
 \text{MAX\_WITHIN\_FIRM\_I\_SIBLINGS}_{i,t}, \text{MAX\_OUTSIDE\_FIRM\_I\_SIBLINGS}_{i,t}, \\
 \text{TOTAL\_WITHIN\_FIRM\_E\_SIBLINGS}_{i,t}, \text{TOTAL\_OUTSIDE\_FIRM\_E\_SIBLINGS}_{i,t}, \\
 \text{TOTAL\_WITHIN\_FIRM\_I\_SIBLINGS}_{i,t}, \text{TOTAL\_OUTSIDE\_FIRM\_I\_SIBLINGS}_{i,t}, \\
 \text{SELF\_CITE}_{i,t}, \text{NON\_SELF\_CITE}_{i,t}, \text{NON\_SERENDIPITY}_{i,t}, \text{PATENT\_DOC\_USEFUL}_{i,t}, \\
 \text{INTERNAL\_INFO\_USEFUL}_{i,t}, \text{NON\_PATENT\_NON\_INTERNAL}_{i,t}, \\
 \text{COMMERCIALIZED}_{i,t}, \text{LICENSED}_{i,t}, \text{USCAT}_{n,t})
 \end{aligned}$$

. . . (ii)

これは 2007 年に実施された第一次発明者サーベイのデータを用いて、「先行特許に基礎をおいていた」と回答する確率が後方引用特許数によって有意に影響をうけるか、を、同一

<sup>6</sup> この変数を用いる推定では2001年以降の後方引用特許はサンプルから除いている。

<sup>7</sup> サーベイでは、出願番号の最終桁が偶数の時、引用番号を古いものから時系列で提示し、最終桁が奇数の時、新しいものから逆順に提示することによって提示順の影響が引用特許の古さにのみ影響されないようにした。

企業組織内の後方引用特許数と企業組織外後方引用特許数それぞれに分けて検討するものである。従属変数は、組織内・組織外を問わず「先行特許に相当程度依拠して発明がなされた」と回答するかどうかのバイナリ変数である。従属変数 *DEPENDENCE<sub>i</sub>* は特許の後方引用ごとに定義されているのではなく、一次サーベイで調査対象となった特許 *i* ごとに定義されたバイナリ変数であり、発明者サーベイ問 4-7 の主設問の回答に相当する（長岡・塚田, 2007）。

今回の H3 の焦点である被引用数の最大値や総和は次の 8 つの変数により検討する。最初に、*MAX\_WITHIN\_FIRM\_E\_SIBLINGS<sub>i</sub>*, *MAX\_OUTSIDE\_E\_SIBLINGS<sub>i</sub>*, *MAX\_WITHIN\_FIRM\_I\_SIBLINGS<sub>i</sub>*, *MAX\_OUTSIDE\_FIRM\_I\_SIBLINGS<sub>i</sub>* の 4 つの変数は、第二次発明者サーベイの対象特許それぞれの先行引用特許（後方引用）が受けた被引用のうち、サーベイ特許ごとの最大値をとったものであり、企業内外と審査官引用・発明者引用に分けて定義してある。具体的には、発明者の属する企業内部での審査官被引用数の最大値（*MAX\_WITHIN\_FIRM\_E\_SIBLINGS<sub>i</sub>*）、その企業外での審査官被引用の最大値（*MAX\_OUTSIDE\_E\_SIBLINGS<sub>i</sub>*）、企業内部での発明者被引用数の最大値（*MAX\_WITHIN\_FIRM\_I\_SIBLINGS<sub>i</sub>*）、企業外での発明者被引用数の最大値（*MAX\_OUTSIDE\_FIRM\_I\_SIBLINGS<sub>i</sub>*）、をそれぞれ意味する。昨年の検討範囲では、企業内における後方引用数は発明者の先行特許存在意識と一致していたが、企業外への後方引用特許数は統計的関連がみられなかった。これは、個々の後方引用特許の重要性が極端に偏っていることに原因があるとすると、後方引用数のカウントではなく、どれか一つのもっとも重要な特許について後方引用がどれくらい被引用を持つか、を検討すべきという予想に基づいており、H3A を検証する目的がある。

上記と平行して、後方引用特許の被引用を個別 1 つだけ取り出さずに後方引用全体について総和をとり、同時に説明変数としても加える（仮説 H3B に対応する）。

*TOTAL\_WITHIN\_FIRM\_E\_SIBLINGS<sub>i</sub>*, *TOTAL\_OUTSIDE\_FIRM\_E\_SIBLINGS<sub>i</sub>*, *TOTAL\_WITHIN\_FIRM\_I\_SIBLINGS<sub>i</sub>*, *TOTAL\_OUTSIDE\_FIRM\_I\_SIBLINGS<sub>i</sub>* の 4 つの変数は、それぞれの先行後方引用特許が受けた被引用カウントを、第一次サーベイの対象特許ごとに和をとったものであり、企業内外と審査官引用・発明者引用に分けて定義してある。具体的には、発明者の属する企業内部での審査官被引用数の和（*TOTAL\_WITHIN\_FIRM\_E\_SIBLINGS<sub>i</sub>*）、その企業外での審査官被引用の和（*TOTAL\_OUTSIDE\_E\_SIBLINGS<sub>i</sub>*）、企業内部での発明者被引用数の和（*TOTAL\_WITHIN\_FIRM\_I\_SIBLINGS<sub>i</sub>*）、企業外での発明者被引用数の和（*TOTAL\_OUTSIDE\_FIRM\_I\_SIBLINGS<sub>i</sub>*）、をそれぞれ意味する。

*SELF\_CITE* および *NON\_SELF\_CITE* は企業内と企業外の後方引用数であり、それぞれについて、日本特許に関する審査官引用と発明者引用を定義するので 4 種類の変数となる。日本の審査官引用で計測した同一企業組織内引用は、*SELF\_JP\_EXAMINER\_BCITE* であり、組織外引用は *NON\_SELF\_JP\_EXAMINER\_BCITE* である。発明者引用でカウ

トした企業内引用は、*SELF\_JP\_INVNTOR\_BCITE* であり、企業外引用は *NON\_SELF\_JP\_INVENTOR\_BCITE* である。

#### 4-2-3. その他のデータ

H3 に関する推定式のコントロール変数はすべて発明者サーベイの第一次結果に基づき長岡教授が作成されたデータベースによっており（長岡・塚田, 2007） 昨年のディスカッションペーパーにおいて利用したのと同じ方法で定義されている（和田, 2008）。*NON\_SERENDIPITY<sub>i</sub>*は、当該発明の創造プロセスがほぼ予定されたとおり進行した、ということを示すダミー変数である。成果が目標のとおりである、または直接ではないが予想できた副産物である、という回答は、セレンディピティーではない場合を示している。予想どおりの成果として生まれた特許には、先行技術がすでにあつた可能性が高いと考えられる。したがって、従属変数が1をとる可能性を高める変数と考えられ、(ii)式を推定した結果には正の係数が予測される。

発明者サーベイでは、発明着想にさまざまな知識源がどの程度有用であつたか、詳細にたずねているが、*PATENT\_DOC\_USEFUL<sub>i</sub>*は、発明着想に関する特許情報の重要性指標を示す。特許情報が有用であつた場合とは、すでに公になっていた技術がもとになった発明である可能性が高くなるであろう。一方、すでに社内に先行技術があり、その改良を目的としていた場合を表す *INTERNAL\_INFO\_USEFUL<sub>i</sub>*は、組織内の知識源の重要性指標を示す。特許が重要な情報源ではないにせよ、企業内に存在した既存技術情報がもとになっていることを意味する。*NON\_PATENT\_NON\_INTERNAL<sub>i</sub>*は、特許でも組織内でもない情報源が発明着想にどの程度有用であつたか、を示しており、特許情報や同一企業組織内情報の重要性指標を相対化するために加えてある。

*COMMERCIALIZED<sub>i</sub>*は、当該発明が自社内で商業化されたかどうかを示すダミー変数である。逆に、*LICENSED<sub>i</sub>*は、社外に技術がライセンスされたことを示すダミー変数である。社内での補完的資産や、関連技術知識がむしろ少ないことを示唆する。

技術分野ごとの特殊性をコントロールするための *USCAT<sub>h</sub>*は、前述したように米国特許分類にもとづく6分類の分野ダミーである。

## 5. 推定結果

### 5-1. 先行特許の中での重要度に関する発明者認識

H1 および H2 に関する式(i)の推定結果を表1に示す。まず H2 に関する *WITHIN\_FIRM\_B\_CITE\_DUMMY* をみると、係数は正かつ有意であり、発明者からみると社内特許が先行技術として相対的に重要であつたと評価する傾向にあると判断される。また *INVENTOR\_BCITE\_DUMMY* 変数の係数が正であることから、発明者がみずからの発明にとって重要だつたと指摘するのは、審査官による引用特許よりは発明者による引用の方が確率的に多いということも読み取れる。なおモデル6および7は、それぞれ分析対象である後方

引用が発明者の企業内であるケースと企業外にあったケースにサンプルを限っているが、サンプル数が大幅に少ないためにモデル6では有意な係数が得られていない。それ以外は、審査官引用数や発明者引用数を説明変数として追加した場合でも省略した場合でも安定した結果がみられている。

H1A, H1Bの2つの仮説を検証するための *E\_SIBLINGS* と *L\_SIBLINGS* については、上述のサンプル不足となったモデル6以外、一貫して審査官引用 (examiner citation) による後方引用特許に付された被引用特許数の係数だけが有意に正となった。一方、発明者引用によるカウントは有意な係数を持たなかった。したがって H1A は支持されるが、H1B については不明、という結果となる。特許被引用数は、引用される特許の技術的重要性ないし経済的重要性を表していることが共通理解となっているが、ここでは、審査官による評価 (前方引用数) と、発明者による重要度評価 (サーベイ回答) が一致している。しかし、発明者前方引用が蓄積している、つまり発明者が後々多く引用しているということと、その特許に関する発明者による重要度評価 (サーベイ回答) には関連がみられない、ということになる。

米国特許における審査官引用と出願者引用を比較した結果から、審査官被引用数は経済価値に対する説明力を持つが、出願者被引用数は説明力がない、という結果が最近報告されている (Hedge and Sampat, 2009)。本稿での結果は、この報告内容と矛盾なく、一貫性を持つものと思われる。

以上のほか、サーベイ時点に近い最近に多くの被引用を集めている後方引用特許がより重要だ、と回答される可能性を考えて *RECENTLY\_ADDED\_RATIO* を検討したが、この処理の範囲では有意な結果は得られていない。このほか、後方引用特許の被引用が元のサーベイ対象特許の発明者と同じ企業にどれくらいの割合が属しているか、を審査官引用・発明者引用ともに評価した *WITHIN\_FIRM\_E\_SIBLINGS\_RATIO*, *WITHIN\_FIRM\_L\_SIBLINGS\_RATIO* ともに有意な係数は観察されていない。

コントロール変数の中では、サーベイ提示の中で最初に提示された特許を選ぶ傾向を示す *TOP\_OF\_LIST* は有意に正であり、選択肢の順序に影響されていることがわかる。また、*B\_CITE\_LAG* が有意に負であることから、相対的に新しい特許を重要視する傾向があることも確認された。

## 5-2. 先行特許の存在

発明者への一次サーベイにおいて発明者に回答を求めた「先行特許が存在するかどうかの意識」について、発明者の属する同一企業組織の中での後方引用数とは相関を持つが、企業組織外から得た後方引用数とは表面上相関をもたない、という問題が昨年度のディスカッションペーパー (和田, 2008) において確認された。このことにつき、基本的に同じ推定式に基づいて、別の仮説 H3A, H3B の変数を加えて検証した結果を表3に示す。モデル9は審査官引用のみに基づく説明変数で、H3A に関する審査官被引用数最大値のみを加え

たもの、モデル10はH3Bに関する審査官被引用数総和を追加したもの、モデル11は発明者引用のみにもとづき、H3Aに関する被引用数最大値を用いたもの、モデル12は発明者引用でH3Bに関する被引用数総和を追加したものである。モデル13、14では審査官引用と発明者引用の双方を用いており、H3Aに関する被引用数最大値だけを用いたものと、H3Bに関する被引用数総和を追加したものである。

発明者引用を用いて後方引用に付加された被引用数の総数や最大値を追加したとき、その中にはとくに有意な係数はみられない。一方、審査官引用を用いて、後方引用に付された被引用数の最大値を求め説明変数とした場合には、企業外でのカウントを用いると一貫して正で有意な係数を持つことがわかった。つまり、H3Aは企業外での審査官被引用数による場合には、支持されている。一方、H3Bは支持される結果となっていない。

ところで、後方引用に付加された被引用数の総数や最大値を説明変数として追加する以前と同様に、企業内後方引用数は審査官引用でも発明者引用でも(`self_jp_inventor_bcite`, `self_jp_examiner_bcite`)有意に正の係数を持つので、企業内の後方引用数単純カウントは同じように発明者による先行特許存在意識と一致している。しかし、後方引用特許が得た被引用の最大値を追加した今回の推定では、企業外からの審査官後方引用数そのもの(`non_self_jp_examiner_bcite`)は有意に負な係数を持つようにみえる結果となった。企業外から少数の非常に影響力のある特許を引用している場合には、その一部の特許のもつ大量の審査官特許被引用数が極端な偏りをもって分布していることが想像できる。

これらの結果は、5-1で述べたように「企業外における重要度評価では、審査官による評価と、発明者による重要度評価が一致している」という理解と整合的な解釈が可能である。つまり発明者の属する企業外で重要な発明が先行特許として少数存在し、それに相当程度もとづいて発明が行われたとき、その特定少数の特許の重要度だけに限定して審査官被引用数という(ある程度)客観的な方法で測定すれば、発明者の「先行特許があった」という意識と一致する、という意味がよみとれる。この最大値でなく、企業外からの後方引用数をそのままとったとき、あるいは最大値だけでなく被引用数の総和を追加したとき(すなわち被引用数で加重した後方引用数全体を説明変数として追加したとき)、どちらにしても焦点となる先行特許とは関わりない先行特許が影響し、重要度情報にノイズが多量に乗せられるので、発明者の先行特許の存在に関する意識と関わりが見られなくなるのであろう。いいかえれば、ある特許が多く先行特許を引用しており、そのうち一部の特許だけが重要な役割を果たした、という場合に、ごく少数の特許に被引用が集中する傾向があるので、発明者が「たしかに先行特許に依拠した」と答える場合と、後方引用特許数や後方引用特許数を被引用で加重したものと表面上の関係がみられなくなるのかもしれない。このような非常に偏った重要度分布に関し、審査官引用が重要度ウエイトとして有用であるが、極端な被引用数の偏りを考慮することも必要になるといえそうである。

なお、発明者引用を用いた結果からは、単純な企業内からの後方引用数以外には有意な係数がみられず、つまり発明者の重要度意識と発明者被引用の関連がみつからない。発明

者が評価した関連先行研究の重要度評価ががまったく無意味と断ずる根拠はないが、審査官引用の場合と何らかの意味で前方引用の分布が大きく異なっていると考えられる。

## 6. 当面の結語と課題

引用特許のうち発明者にとって重要だったと指摘されたものは、審査官被引用（前方引用）数の多いものであること、また自社内にある傾向を持つこと、という2つの発見事実は、直感的にも予測できる内容である。先行研究では、発明者引用は発明者の認識を反映しているが、審査官引用が付加されることでノイズが混入する、と強調されてきたきらいがあった。従来議論では知識フローの代理変数としてノイズだといわれてきた審査官引用は、被引用を通じた重要度評価の意味では、むしろ発明者引用よりも有用である、といえそうである。

本稿の3番目の課題、すなわち2007年に実施された発明者（一次）サーベイにおいて示された特許後方引用数と、発明者による先行特許認識の関係における偏り（和田, 2008）については、十分な解明にはいたっていないが、興味深いネットワーク構造が示唆されたと考える。つまり、スケール・フリー性とも呼ばれる一般的な特定の特許への引用集中と、おそらくはクラスタ性の存在により、特許引用を企業間の技術的累積関係を計量する手段として用いる上で一定の歪みがある、と考えられる。

この結果の背後には、引用の分布という未解明分野が存在することも今回明らかになった。重要な特許にさらに引用が集中する傾向が、発明者認識と審査官被引用数が相関を持つことで示唆されたが、これによって優先的付加がある程度働いている可能性が示唆されている。ある特許にたいし、多数の前方引用がなされている場合とはごく限られた場合である、いいかえると、ごく一部の特許に引用が集中している、という事実は、スケール・フリーネットワークで提示されているようなネットワークの生成過程の結果に非常に似ている。

今回の検討には、いくつかの方法上の困難があり、今回の作業結果は端緒的なものにとどまっている。特許引用を付加していた時点とサーベイ時点の時間差により認識の差異が存在するかもしれないが、検出することは難しい。さらに、先行技術として重要であると答えた発明者認識が、優先選択によって影響されたかどうか、を検定するのも困難である。本質的な障害として、発明者からみて「引用されていないものも含めてもっとも重要な先行技術」や「現に引用している中で発明者認識バイアスの影響を受けない、真にもっとも重要な引用技術」を知るすべはなく、実際に引用されている特許の中から発明者が重要と答えたものをもとに分析せざるを得ないことがある。このような問題点は残されているが、引用のネットワークがダイナミックに形成されていく過程にネットワーク科学の知見が利用できることから、単純に引用数、被引用数をカウントすることで行われてきた先行研究に対して、新たな視角からの分析の必要性を提起できたものとする。

<References>

Alcácer, Juan J. and M. Gittelman, 2006, Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations, *Review of Economics and Statistics*, Vol. 88, No. 4, Pages 774-779.

Alcácer, Juan J. and M. Gittelman, Bhaven Sampat, 2009, Applicant and examiner citations in U.S. patents: An overview and analysis, *Research Policy*, vol. 38, no.2, pp.415-427.

Barabási, Albert-László, Réka Albert, 1999, Emergence of Scaling in Random Networks, *Science*, vol. 286. no. 5439, pp. 509 - 512.

Chesbrough, Henry William, 2003, *Open Innovation: The New Imperative for Creating and Profiting from Technology*, Harvard University Press.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg, 2001, The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools, NBER Working Paper 8498.

Harhoff, Dietmar, Frederic M. Scherer, Katrin Vopel, 2003, Citations, family size, opposition and the value of patent rights, *Research Policy*, Volume 32, Issue 8, September 2003, Pages 1343-1363.

Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson, 1993, Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics*, pp.577-98.

Jaffe, Adam, and Manuel Trajtenberg, 1996, Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries, NBER Working Paper No. W5712.

Jaffe, Adam, and Manuel Trajtenberg, 1999, "International knowledge flows: evidence from patent citations," *Economics of Innovation and New Technology*, vol.8 pp.105-136.

Jaffe, A., M. Trajtenberg, and M. Fogarty, 2000, Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors, *American Economic Review*, vol.90, no.2, pp.215-218.

Hedge, Deepak, and Bhaven Sampat, 2009, Examiner citations, applicant citations, and the private value of patents, *Economics Letters*, forthcoming.

Mowery, David C., Joanne E. Oxley and Brian S. Silverman, 1996, Strategic Alliances and Interfirm Knowledge Transfer, *Strategic Management Journal*, Vol. 17, Special

Issue: Knowledge and the Firm (Winter, 1996), pp. 77-91.

Nakamoto, 1988, Synchronous and diachronous citation distributions, in *Informetrics 87/88: Select Proceedings of the 1st International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval*, Diepenbeek, Belgium, 25-28 August 1987, L. Egghe, R. Rousseau, eds. (Elsevier, New York, 1988), p. 157.

Oxley, Joanne, and Tetsuo Wada, Alliance Structure and the Scope of Knowledge Transfer: Evidence From US-Japan Agreements, *Management Science*, Vol. 55, No. 4 April 2009, pp. 635-649.

Silverberg and Verspagen, 2007, The size distribution of innovations revisited: An application of extreme value statistics to citation and value measures of patent significance, *Journal of Econometrics*, Volume 139, Issue 2, August 2007, Pages 318-339.

Tamada, Schumpeter, Yusuke Naito, Fumio Kodama, Kiminori Gemba, Jun Suzuki, 2006, Significant difference of dependence upon scientific knowledge among different technologies, *Scientometrics*, Volume 68, Number 2.

Trajtenberg, Manuel, 1990, "A penny for your quotes: patent citations and the value of innovations," *Rand Journal of Economics*, vol.21 no.1 pp.173-187.

Trajtenberg, Manuel, Rebecca Henderson, and Adam Jaffe, 1997, "University versus Corporate Patents: A Window on the Basicness of Invention," *Economics of Innovation and New Technology*, vol.5 no.1 pp.19-50.

Ziedonis, Rosemarie Ham, 2004, "Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms," *MANAGEMENT SCIENCE*, Vol. 50, No. 6, June 2004, pp. 804-820.

長岡貞男・塚田尚稔, 2007, 『発明者から見た日本のイノベーション過程: RIETI 発明者サーベイの結果概要』 RIETI Discussion Paper Series 07-J-046.

和田哲夫, 2008, 『先行技術の量的指標としての特許引用数』 RIETI Discussion Paper Series 08-J-038.

和田哲夫・後藤晃, 2003, 「特許に関する経済学的な分析の必要性と特許データのありかた」, 『産業界におけるライセンス契約の動向及び経済学の問題に関する調査研究報告書 (平成14年度特許庁産業財産権制度問題調査研究報告書)』 第9章、財団法人知的財産研究所。

図 1 審査官引用の頻度分布 (縦軸: k 回引用される確率 P(k)対数表示、横軸: k 対数表示)

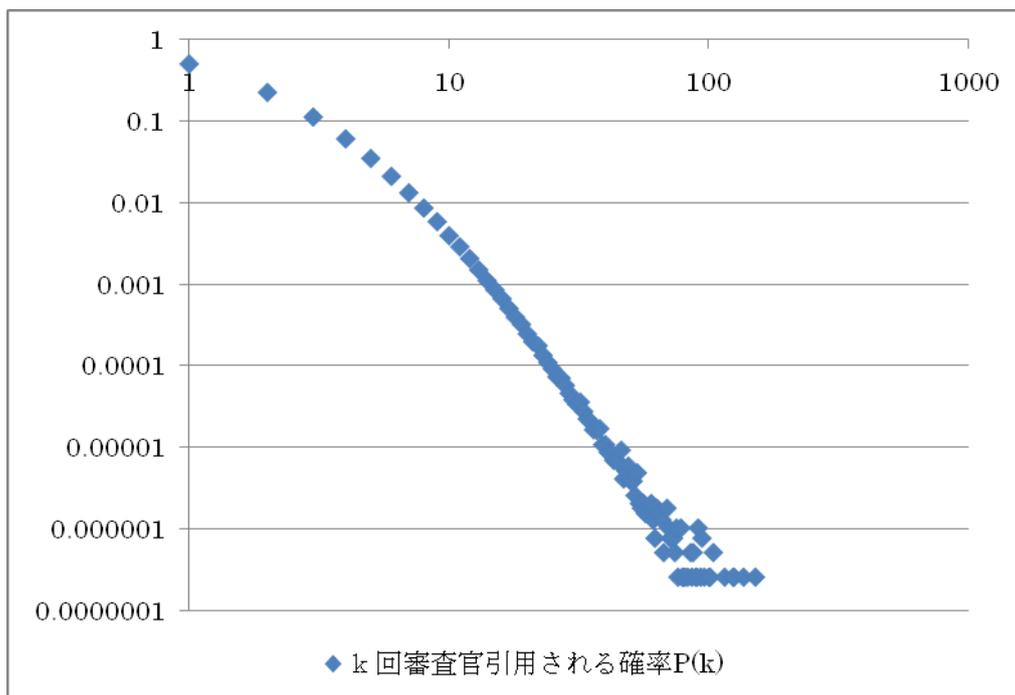
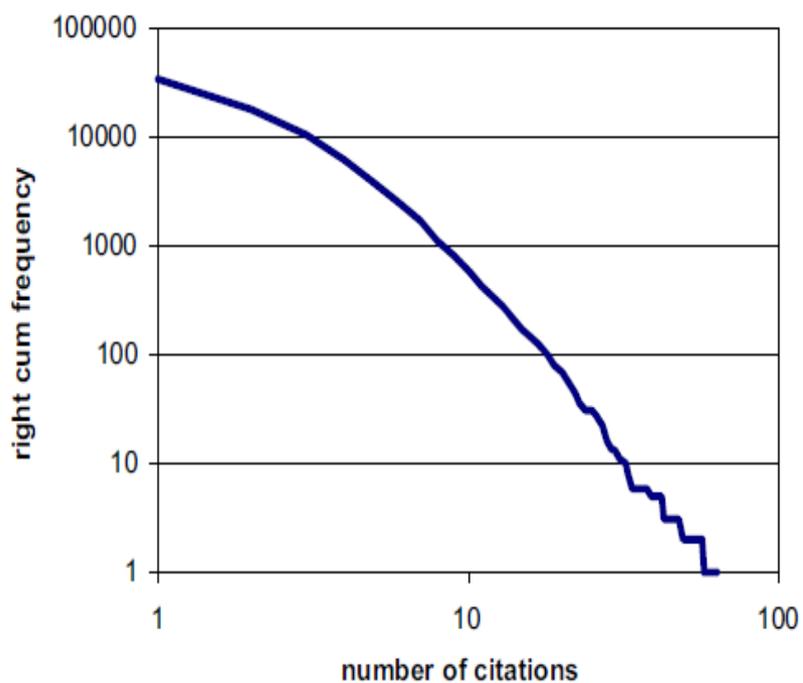
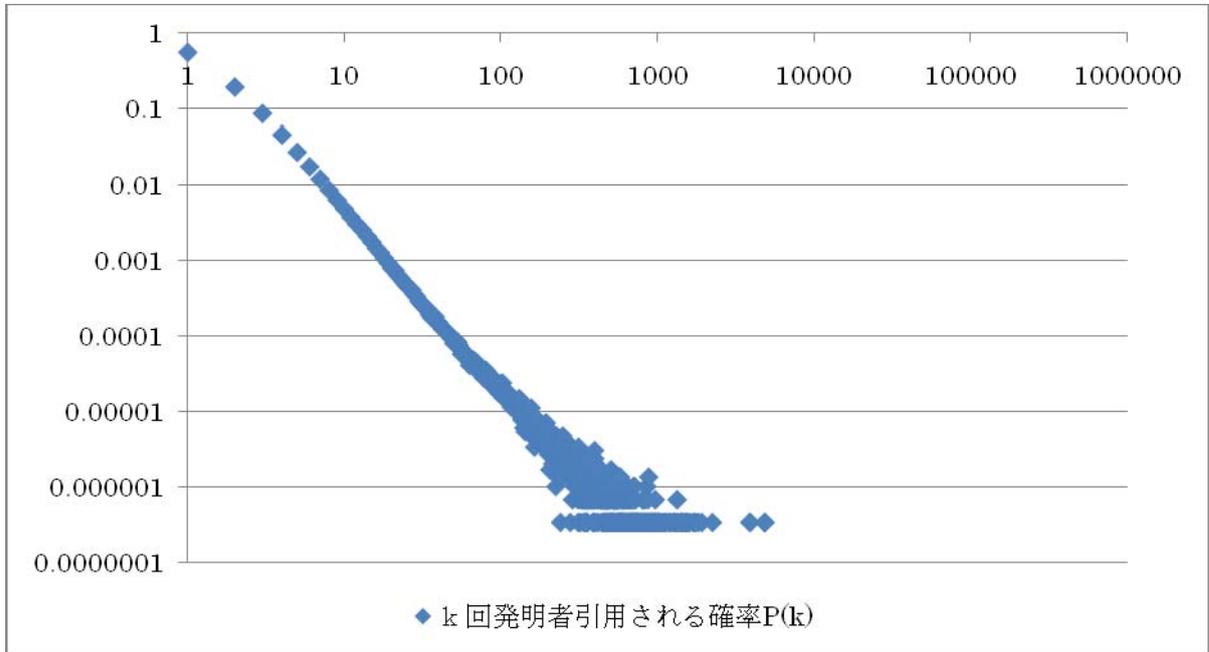


図 2 欧州での特許引用の頻度分布



Silverberg and Verspagen (2007) から引用。

図3 発明者引用の頻度分布 (縦軸:  $k$  回引用される確率  $P(k)$  対数表示、横軸:  $k$  対数表示)



(2008年までの公報上の引用データによる)

図4 発明者引用の頻度分布、引用・被引用が5年以内 (鈴木潤教授の作成したサンプル群)

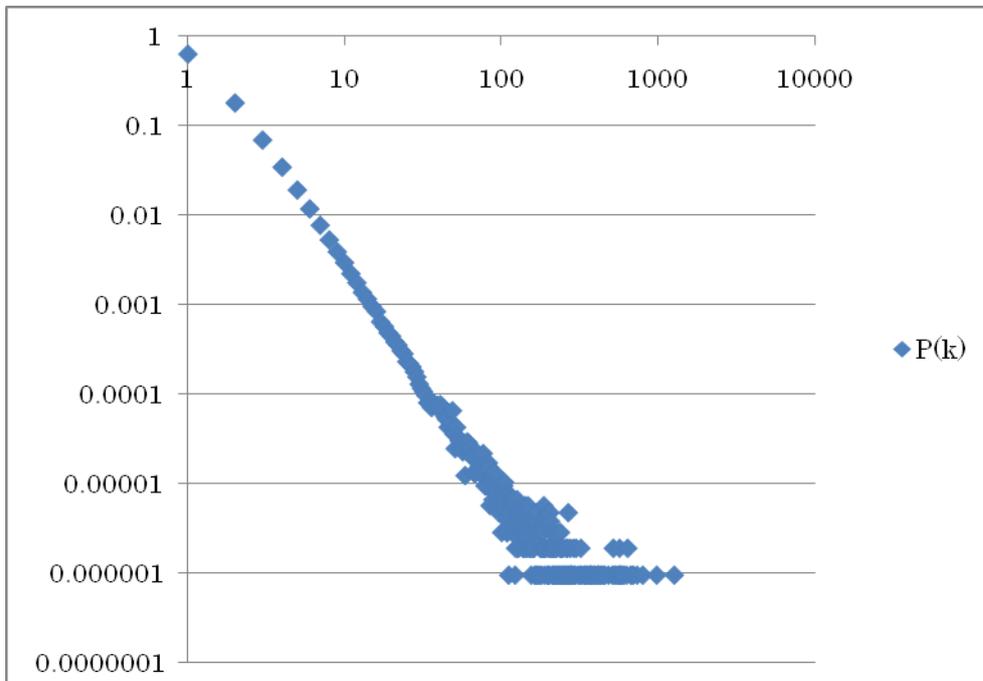


表 1. サーベイ対象特許が引用している後方引用特許群のうち発明者が先行特許として重要と回答する選択要因 (Logit)

Model	審査官後方引用				発明者後方引用		企業内後方引用		企業外後方引用	
	1	2	3	4	5	6	7	8	9	10
within_firm_b_cite_dummy	1.063**** (0.1371)	0.9797**** (0.1327)	1.0637**** (0.1634)	1.2131**** (0.2161)	0.7446**** (0.2787)					1.0808**** (0.19547)
inventor_bcite_dummy	0.2457** (0.1005)	0.2548*** (0.0977)	0.2363** (0.115)			0.0844 (0.264)	0.2730** (0.1358)	0.28149** (0.1392)		
e_siblings	0.0286**** (0.0097)		0.0364**** (0.0108)	0.0314** (0.0149)	0.0442**** (0.0137)	0.0239 (0.0333)	0.0375**** (0.01209)	0.0343*** (0.0124)		
i_siblings		-0.0002 (0.0008)	-0.0008 (0.0008)	0.0017 (0.0028)	-0.0011 (0.0009)	-0.00127 (0.00123)	-0.000523 (0.001208)	-0.00153 (0.001007)		
recently_added_ratio									0.04062 (0.07215)	
within_firm_e_siblings_ratio	0.109 (0.1693)		0.1869 (0.1973)	0.0077 (0.2866)	0.3212 (0.2593)	0.2441 (0.3147)	0.1518 (0.2641)	0.24605 (0.23657)		
within_firm_i_siblings_ratio		-0.0995 (0.1682)	-0.125 (0.2089)	-0.2545 (0.2436)	0.315 (0.4039)	-0.39459 (0.36002)	0.0251 (0.2791)	-0.11375 (0.2479)		
top_of_list	0.8905**** (0.1161)	0.9852**** (0.1132)	1.0057**** (0.1387)	0.8731**** (0.1751)	1.1364**** (0.1938)	0.9416**** (0.3059)	1.0311**** (0.156)	1.0169**** (0.17690)		
b_cite_lag	-0.0487**** (0.0091)	-0.0417**** (0.0084)	-0.0503**** (0.0105)	-0.0482**** (0.0154)	-0.0438**** (0.013)	-0.0302 (0.0234)	-0.0545**** (0.0118)	-0.0512*** (0.0199)		

uscat1	-0.1102 (0.2321)	-0.5395**** (0.1468)	-0.3938** (0.1716)	0.2027 (0.2368)	-0.2092 (0.3108)	-0.3498 (0.3899)	-0.271 (0.289)	-0.0157 (0.2213)
uscat2	0.0756 (0.2367)	-0.3232** (0.1604)	-0.2725 (0.1825)	0.1806 (0.231)	0.0882 (0.3426)	0.2466 (0.4612)	-0.244 (0.3003)	0.09074 (0.23073)
uscat3	0.2236 (0.2356)	-0.1845 (0.2076)	-0.2261 (0.2616)	0.5485 (0.3805)	0.3934 (0.3391)	-0.4955 (0.555)		0.20172 (0.32422)
uscat4	-0.2019 (0.2412)	-0.4345**** (0.1637)	-0.3883** (0.189)	0.3324 (0.2367)	-0.0329 (0.3306)		0.14211 (0.2978)	0.40343* (0.2328)
uscat5	-0.0871 (0.2381)	-0.2923* (0.1577)	-0.3793** (0.1847)	-0.0458 (0.2494)	0.118 (0.3269)	-0.345 (0.3964)	-0.2712 (0.3086)	
uscat6	-1.0853**** (0.2301)	-0.6438**** (0.1367)	-0.8188**** (0.1622)	-1.2147**** (0.2297)	-1.0631**** (0.3172)	-0.8746** (0.4224)	-0.11473 (0.2984)	-0.01251 (0.2391)
N	2385	2440	1765	1065	982	347	1418	1136
Log likelihood	-1353.5	-1425.2	-1007.3	-607.9	-557.6	-227.3	-775.3	-664.68

\* significance level 0.1

\*\* significance level 0.05

\*\*\*significance level 0.01

\*\*\*\*significance level 0.005

Standard errors are in the parentheses.

表 2. 記述統計量

<i>Variable</i>	サンプル数	平均	標準偏差	最小	最大
<i>within_firm_b_cite_dummy</i>	3393	0.195992	0.39702	0	1
<i>inventor_bcite_dummy</i>	3393	0.437961	0.4962	0	1
<i>e_siblings</i>	3393	2.68936	4.255	0	74
<i>i_siblings</i>	3393	13.1957	51.452	0	894
<i>recently_added_ratio</i>	1666	.2929019	1.0064	0	21.52
<i>within_firm_e_siblings_ratio</i>	2385	0.18417	0.31873	0	1
<i>within_firm_i_siblings_ratio</i>	2440	0.20006	0.33141	0	1
<i>top_of_list</i>	3393	0.182434	0.38625	0	1
<i>b_cite_lag</i>	3393	7.342765	5.9133	0	35
<i>uscat1</i>	3393	0.254347	0.43555	0	1
<i>uscat2</i>	3393	0.17153	0.37702	0	1
<i>uscat3</i>	3393	0.055998	0.22995	0	1
<i>uscat4</i>	3393	0.167698	0.37365	0	1
<i>uscat5</i>	3393	0.16534	0.37154	0	1
<i>uscat6</i>	3393	0.185087	0.38842	0	1

表 3. RIETI 発明者サーベイ（原調査、2007 年度実施）において発明者が依拠した先行特許が存在すると回答する要因（Logit）

<i>Model</i>	<i>9</i>	<i>10</i>	<i>11</i>	<i>12</i>	<i>13</i>	<i>14</i>
max_within_firm_e_siblings	0.0416**** (0.0146)	0.0226 (0.0246)			0.0331** (0.0152)	0.0214 (0.0253)
max_outside_e_siblings	0.0175*** (0.0064)	0.0246**** (0.008)			0.0159** (0.007)	0.0214** (0.0094)
max_within_firm_i_siblings			0.0004 (0.0003)	0.0006 (0.0006)	0.0002 (0.0003)	0.0004 (0.0006)
max_outside_firm_i_siblings			0.001* (0.0006)	-0.0005 (0.0014)	0.0004 (0.0006)	-0.0007 (0.0012)
total_within_firm_e_siblings		0.0078 (0.0107)				0.0054 (0.0112)
total_outside_firm_e_siblings		-0.0074 (0.0045)				-0.0049 (0.0051)
total_within_firm_i_siblings				-0.0001 (0.0002)		-0.0001 (0.0002)
total_outside_firm_i_siblings				0.0009 (0.0008)		0.0005 (0.0007)
self_jp_examiner_bcite	0.229**** (0.068)	0.2294**** (0.0687)			0.212**** (0.0685)	0.2123**** (0.0692)
non_self_jp_examiner_bcite	-0.051** (0.0203)	-0.0481** (0.022)			-0.0442** (0.0205)	-0.0428* (0.022)

self_jp_inventor_bcite			0.1882****	0.1851****	0.1323**	0.1231**
			(0.0546)	(0.0549)	(0.0553)	(0.056)
non_self_jp_inventor_bcite			-0.0083	-0.0126	-0.0236	-0.0291
			(0.0306)	(0.031)	(0.0317)	(0.0324)
nonserendipity	0.3558****	0.3516****	0.3498****	0.3499****	0.3504****	0.3489****
	(0.0682)	(0.0682)	(0.068)	(0.068)	(0.0683)	(0.0683)
patent_doc_useful	0.3896****	0.3883****	0.39****	0.3897****	0.3902****	0.3895****
	(0.0238)	(0.0238)	(0.0238)	(0.0238)	(0.0239)	(0.0239)
internal_info_useful	0.0935****	0.0926****	0.0959****	0.0961****	0.0924****	0.0921****
	(0.0222)	(0.0222)	(0.0222)	(0.0222)	(0.0222)	(0.0222)
non_patent_non_internal	-0.0104****	-0.0103****	-0.0107****	-0.0107****	-0.0103****	-0.0103****
	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0032)
commercialized	-0.2467****	-0.2454****	-0.2412****	-0.2412****	-0.2466****	-0.2471****
	(0.0644)	(0.0644)	(0.0642)	(0.0642)	(0.0645)	(0.0645)
licensed	0.1892**	0.1951**	0.223***	0.2237***	0.1883**	0.1925**
	(0.0818)	(0.0819)	(0.0813)	(0.0813)	(0.0819)	(0.082)
uscat1	0.5248****	0.5412****	0.5587****	0.5755****	0.5586****	0.5442****
	(0.1409)	(0.1125)	(0.1115)	(0.1412)	(0.1122)	(0.1417)
uscat2	-0.0421	0.025	-0.0171	0.0178	0.0092	-0.0011
	(0.1447)	(0.1451)	(0.1446)	(0.1446)	(0.1456)	(0.1458)
uscat3	0.1742	0.2051*	0.2145**	0.2319*	0.2106*	0.2035
	(0.1396)	(0.1094)	(0.1088)	(0.1395)	(0.1093)	(0.1404)
uscat4	0.2458*	0.2723**	0.27**	0.2895**	0.2814***	0.274**

	(0.1386)	(0.1088)	(0.1079)	(0.139)	(0.1087)	(0.1393)
uscat5	0.1501	0.1769*	0.1886*	0.2056	0.1913*	0.1814
	(0.1337)	(0.1031)	(0.1017)	(0.1343)	(0.1028)	(0.1347)
uscat6	-1.8587****	-1.878****	-1.8475****	-1.863****	-1.9004****	-1.8852****
	(0.158)	(0.1307)	(0.1286)	(0.1588)	(0.1306)	(0.1594)
N	4821	4821	4821	4821	4821	4821
Log Likelihood	-3033.3	-3031.8	-3043.5	-3042.8	-3029.2	-3028.2

\* significance level 0.1

\*\* significance level 0.05

\*\*\*significance level 0.01

\*\*\*\*significance level 0.005

Standard errors are in the parentheses.

表 4. 記述統計量 (RIETI 発明者サーベイの一次サーベイ、2007年実施)

	サンプル数	平均	標準偏差	最小	最大
dummy_q4_7	5232	0.497133	0.5000396	0	1
max_within_firm_e_siblings	5278	1.200	2.838	0	76
max_outside_e_siblings	5278	3.076	5.424	0	80
max_within_firm_i_siblings	5278	10.636	112.284	0	3645
max_outside_firm_i_siblings	5278	11.420	75.972	0	2001
total_within_firm_e_siblings	5278	2.838	7.860	0	213
total_outside_firm_e_siblings	5278	3.023	10.566	0	267
total_within_firm_i_siblings	5278	49.936	931.749	0	36222
total_outside_firm_i_siblings	5278	25.958	229.373	0	6758
self_jp_examiner_bcite	5278	0.126942	0.580615	0	10
non_self_jp_examiner_bcite	5278	0.594543	1.665529	0	19
self_jp_inventor_bcite	5278	0.206707	0.678521	0	14
non_self_jp_inventor_bcite	5278	0.421751	1.034664	0	12
self_us_bcite	2002	0.772727	1.501888	0	16
non_self_us_bcite	2002	5.169331	3.915882	0	33
non_serendipity	5218	0.693561	0.461059	0	1
patent_doc_useful	5101	3.293668	1.6452	0	5
internal_info_useful	5049	3.258467	1.531736	0	5
non_patent_non_internal	4888	19.94947	11.51144	0	55
commercialized	5278	0.505684	0.500015	0	1
licensed	5278	0.195908	0.396935	0	1
uscat1	5278	0.177719	0.382312	0	1
uscat2	5278	0.140015	0.347036	0	1
uscat3	5278	0.06745	0.250823	0	1
uscat4	5278	0.17393	0.379085	0	1
uscat5	5278	0.187382	0.390255	0	1
uscat6	5278	0.253505	0.435059	0	1