

【論文】

古典派経済学の普及過程に関するテキストマイニング分析 —リカード、ミル、マーティノーを中心に—

下平 裕之・福田 進治

1 課題設定

1.1 研究の背景

リカードの『経済学原理』（1817年初版刊行、以下『原理』と略記）は古典派経済学の理論的基礎を確立し、その後の経済学の歴史を決定づけた著作である。しかし、リカード『原理』は必ずしもそのまま社会に受け入れられたわけではない。リカードの経済学が普及する過程で、ジェームズ・ミルやハリエット・マーティノーといった、第一級の理論家とは言えない人々が重要な役割を果たしたことが知られている。すなわち、古典派経済学は、リカード『原理』のような専門書だけでなく、ミルの『経済学要綱』（1821年初版刊行、以下『要綱』）やマーティノーの『経済学例解』（1832-34年初版刊行、以下『例解』）のような通俗的な教科書や啓蒙書が一般大衆の目に触れることを通して普及したのである。従って、そこには経済学の専門的な内容に関わる問題だけでなく、専門的な内容が通俗的な内容に変容していく過程という問題が存在する。こうした問題は古典派経済学の普及過程に限らず、広く経済学の歴史全般に存在すると思量される¹⁾。

こうした事情を考慮するなら、一般に、経済学の普及過程を研究するためには、経済学の専門的な文献だけでなく、より通俗的な文献を検討しなければならない。そのような文献は恐らく膨大な分量に上るだろう。歴史上、経済学の主要な専門書だけでも決して少数ではなく、それらの内容を解明することだけでも決して容易ではない。従って、それらに加えてより通俗的な膨大な文献を検討するためには、従来の経済学史研究＝文献資料の質的研究の方法に加えて何か別の工夫が必要になるだろう。そのような工夫の一つとして、本稿ではテキストマイニング分析を提案する。テキストマイニング分析とは、諸々の文献のテキストデータを統計的に処理し、定量的な分析を行うものである（次節で詳述する）。本稿では、古典派経済学の普及過程に注目しながら、こうした量的分析の方法が経済学史研究のために有効であるか否かを探っていきたい。

なお、本研究は「経済思想の受容・浸透過程に関する実証研究：人々は経済学をどのように受け入れたか」²⁾というプロジェクトの一翼を担うものである。このプロジェクトは、経済学の理論や思想が経済学の専門家から一般大衆である非専門家に普及していく過程を、(a)古典派経済学以降の200年という射程で、(b)イギリスに焦点を絞り、(c)質的および量的に特定化・類型化する試み

である。そして、その一次的接近の方法として、その時代を代表する経済学の専門的な文献（アダム・スミス『国富論』、カール・マルクス『資本論』、J. M. ケインズ『一般理論』等）に込められた理論や思想が、より通俗的な評論・入門書・教科書・解説書・通俗書・詩・寓話・小説・パンフレット・ビラ等を通じて、どのように一般大衆に受け入れられていったのかを明らかにするものである。

1.2 本稿の課題

本稿の課題は、上のプロジェクトの一環として、テキストマイニング分析を用いて古典派経済学の普及過程を検討することである。本稿では、とくにリカードの経済学が通俗化していく過程の一端を明らかにする。そのために、テキストマイニング分析を用いて、リカード『原理』（第3版）、ミル『要綱』（第3版）、マーティノー『例解』（第3版）第25話を比較検討する。マーティノーは古典派経済学の普及のためにとくに大きな役割を果たした人物として知られているが、リカード『原理』の議論がマーティノー『例解』によって広められる際にどのような変容を被ったのか、こうした過程をミル『要綱』を間に挟んで明らかにしたい。これらと同時に、こうした経済学の普及過程の研究のためにテキストマイニング分析を用いることの妥当性を探っていききたい。

本稿では、次のような仮説をあらかじめ設定する。すなわち、マーティノー『例解』は一般大衆を讀者として想定していたため、リカード『原理』やミル『要綱』と比較して、

- (1) 専門的な用語を排除し、一般大衆が理解しやすい通俗的な言葉を多く用いている。
- (2) 基礎的な理論分析よりも、その背景にある世界観や政策論に注目している。

以下では、こうした仮説が成立するか否か、そしてそれらがテキストマイニング分析によって正当化されるか否かを探っていききたい。

2 テキストマイニング分析の概要と文書データ

2.1 経済学研究とテキストマイニング分析

テキストマイニング分析とは「テキストデータを計算機で定量的に解析して有用な方法を抽出するための様々な手法の総称」（松村・三浦 2009:1）であり、大量の文字情報を一括処理して、テキストの中に隠された法則や知見を発見する分析手法である。「マイニング」とは元々「採鉱する」「坑道を掘る」という意味であり、テキストマイニング分析には2つの基本的な特性がある。第1に、大量の情報の中に隠された法則や知見を抽出することである。第2に、こうして抽出された法則の中から、意義のある、有意味な、有用な法則を厳選することである。第1の段階では主に統計処理ソフトやアルゴリズムの理解等、数理的な処理（量的分析）が助けとなる。しかし、第2の段

階では、自明ではないが有意味な法則・解釈・知見は何かという判断が不可欠であり、ここに質的分析も同時に求められていることに留意したい。言い換えるなら、テキストマイニング分析は質的分析を補完するための量的分析なのである。

このような観点から見ると、テキストマイニング分析は従来の経済学史研究と接続する可能性を持つ。テキストマイニング分析では、デジタル化された文献データの収集から始め、統計的手法を導入するための処理を行い、処理されたデータを分析して結果を得る。しかし、分析はこうした技術的な処理だけで完結するわけではない。実際には、さまざまな段階で「解釈」を施さなければならない。すなわち、仮説、収集された文献データの意義、カテゴリーの名辞付け、結果の意義と限界、そして解析全体を総合的に検討する考察が限りなく重要な位置を占める。このためにテキストマイニング分析には、従来から蓄積されてきた経済学史研究の知見が絶対的に必要となる。こうした知見を前提として、大量の文献データを統一的に処理することによって、従来見逃されていた新しい知見——例えば、匿名文書の筆者の特定、当該書籍の版別異同における思想的変化の明確化、通俗化・単純化された支配的言説の変形等——が見出される可能性が存在する。

つまり、従来、研究者の直観や長年の研究の蓄積によって支えられてきた個別的な文献の質的分析の間に、大量の文献データを統一的に処理する量的な分析を挟み込むことによって、それまで個別的だった事例報告を越えて、より一般的な、より深みのある、より説得力のある知見を見出すことが期待できる。この意味で、テキストマイニング分析は従来の経済学史研究と完全に適合的であり、従来の経済学史研究をさらに発展させ、より「発見的」heuristicな知見を生み出す新しい研究方法となる可能性を秘めている³⁾。

2.2 テキストマイニング分析の概要

本稿では、リカード『原理』、ミル『要綱』、マーティノー『例解』第25話について、テキストマイニング分析の代表的な手法である頻度分析と共起ネットワーク分析を適用し、それぞれの文献の特徴を明らかにしながら、これらを比較検討する。そこで、頻度分析と共起ネットワーク分析の概要についてあらかじめ説明しておきたい。なお、これらの分析のためのテキストマイニングソフトとして、本研究ではKH Coder⁴⁾を用いた。

(1) 頻度分析

頻度分析とは、対象文献における単語の出現頻度を検討するものである。今回は、KH Coderを用いて出現頻度の高い名詞を抽出する。名詞に注目するのは、認知心理学のカテゴリー化研究における「ある理念・概念・思想がもっとも象徴的に現れるのが名詞に他ならない」（喜田 2008:151）という命題に依拠したためである。頻度分析の基本的考え方は「出現頻度が高い単語ほど重要度が高い」、または「出現頻度が低い単語ほど重要度が低い」というものであり、対象文献の特徴を知るための最も基本的な分析である。

(2) 共起ネットワーク分析

共起ネットワーク分析とは、対象文献における諸々の単語と単語の間の関連性を検討するものである。頻度分析によって単語の出現頻度を検討するだけでは文献の特徴が十分に明らかにならない場合には、単語間の関連性を分析することによって、それぞれの文献の特徴を明らかにすることが可能となる。KH Coderを用いると、このような単語間の関連性を捉えて、出現パターンが似通った単語、すなわち共起の程度が強い単語を線で結んだネットワーク図を描くことができる。

共起関係の強弱については、分析対象となった単語のすべての組み合わせについて、Jaccard係数の値が計算される。Jaccard係数は0から1までの値をとり、2つの単語について同じ文章中に同時に出現する（共起する）と関連が強いと見なし、その値は1に近づく。または、2つの単語について同じ文章中に同時に出現しない（共起しない）と関連が低いと見なし、その値は0に近づく。任意の単語Xと単語Yに関するJaccard係数(J)は次式で求められる。

$$J=A/(A+B+C)$$

ここで、Aは単語Xと単語Y両方が同時に出現している文章数、Bは単語Xのみが出現している文章数、Cは単語Yのみが出現している文章数を表す。

KH CoderはこうしたJaccard係数の値を自動的に計算し、単語間のネットワーク図を自動的に描画する。ネットワーク図の描画に際しては、単語の「中心性」を規準としたネットワークと「サブグラフ検出」によるネットワークのいずれかを選択することが可能であり、文献の特徴を考察する際の一助とすることができる⁵⁾。

- ① 中心性の分析：それぞれの単語が当該文献の中でどの程度中心的な役割を果たしているかを、ネットワーク図によって示すものである。それぞれの単語は色分けされ、赤色～桃色～白色～水色の順に、中心性の高さが示される。
- ② サブグラフ検出：相対的に強く互いに結びついている単語同士を自動的に検出してグループ分けを行い、ネットワーク図によって示すものである。それぞれの単語グループは色分けされ、緑色、黄色、青色、赤色他によって示される⁶⁾。色分けにおいて、囲み枠の中が白色であれば、その単語が他の単語とグループを形成していない単独の単語であることを意味している。サブグラフ検出を行った場合、同じサブグラフに含まれる単語は実線で結ばれるのに対して、互いに異なるサブグラフに含まれる単語は破線で結ばれる。

なお、いずれのネットワーク図においても、単語間を結ぶ線の太さは、単語間の共起の程度が強いほど太く描画される。

2.3 文書データ

本研究では、リカード『原理』、ミル『要綱』、マーティノー『例解』第25話についてテキストマ

イニング分析を行うために、以下の文献データをデジタル化して用いた⁷⁾。

Ricardo, D. 1821, *The Principles of Political Economy, and Taxation*, 3rd. ed., London: John Murray.

Mill, J. 1844, *Elements of Political Economy*, 3rd. ed., London: Henry G. Bohn

Martineau, H. 1834, *Illustrations of Political Economy*, 3rd. ed., vol.9, London: Charles Fox.

これらの文献データの各々について、次のような手順で処理した。すなわち、(1)テキストをスキャンし、OCR (optical character reader) 処理をする。(2)OCR処理されたテキストをMicrosoft Wordなどのソフト上で整形する(手作業で文字化けを直したり、無意味な記号を排除したりする)。(3)文献ごとのタグ付けを行った後、テキストファイルとして保存する。その上で、KH Coderを用いて分析を行った。

3 分析結果

3.1 頻度分析

リカード『原理』、ミル『要綱』、マーティノー『例解』第25話について頻度分析を行い、各文献の出現頻度1位から30位までの名詞を抽出した結果は、表1のとおりである⁸⁾。各文献の上位の単語を比較することによって、次の点が明らかになった。

- (1) リカード『原理』の1-10位の名詞(赤字)は、ミル『要綱』の1-30位にすべて含まれているが、マーティノー『例解』第25話の1-30位には5語しか含まれていない。
- (2) リカード『原理』の11-20位の名詞は、ミル『要綱』の1-30位に9語含まれているが、マーティノー『例解』第25話の1-30位には6語しか含まれていない。
- (3) リカード『原理』の21-30位の名詞は、ミル『要綱』の1-30位にも、マーティノー『例解』第25話の1-30位にも、2語しか含まれていない。
- (4) ミル『要綱』の1-30位には、リカード『原理』の1-30位に含まれていなかった名詞が9語含まれている(マーティノー『例解』第25話の1-30位には、このうち3語しか含まれていない)。
- (5) マーティノー『例解』第25話の1-30位には、リカード『原理』の1-30位にも、ミル『要綱』の1-30位にも含まれていなかった名詞(青字)が14語含まれている。

3.2 共起ネットワーク分析(その1): 中心性の分析

リカード『原理』、ミル『要綱』、マーティノー『例解』第25話に出現する名詞について共起ネットワーク分析を行い、中心性を規準としたネットワーク図で表すと、図1.1～図1.3のとおりである⁹⁾。これらの図を比較することによって、次の点が明らかになった。

- (1) リカード『原理』において、price、profitが中心的な役割を果たしている。また、produce、laborがそれらに準ずる役割を果たしている。
- (2) ミル『要綱』において、labor、commodityが中心的な役割を果たしている。また、wages、capital、moneyがそれらに準ずる役割を果たしている。
- (3) マーティノー『例解』第25話において、capital、societyが中心的な役割を果たしている。
- (4) 3つの文献について、中心的な役割を果たす名詞は、互いにまったく異なっている。

3.2 共起ネットワーク分析（その2）：サブグラフ検出

リカード『原理』、ミル『要綱』、マーティノー『例解』第25話に出現する名詞について共起ネットワーク分析を行い、サブグラフ検出を適用したネットワーク図で表すと、図2.1～図2.3のとおりである。これらの図を比較することによって、次の点が明らかになった。

- (1) リカード『原理』には、money、value、corn他からなる第1グループ（緑色）、land、produce、rent他からなる第2グループ（黄色）他が存在する。
- (2) ミル『要綱』には、proportion、case、labourer他からなる第1グループ（緑色）、price、commodity、quantity他からなる第2グループ（黄色）他が存在する。
- (3) マーティノー『例解』第25話には、labourer、capitalist、wages他からなる第1グループ（緑色）、country、commodity、money他からなる第2グループ（赤色）、food、supply、demand他からなる第3グループ（青色）他が存在する。また、前2者と比較して、グループは分散化・細分化されている。
- (4) リカード『原理』の第1グループ、ミル『要綱』の第2グループ、マーティノー『例解』第25話の第2グループには、money、commodity等、2者間または3者間で共通する構成要素が見出せるが、全体的な構成は互いに大きく異なっている。
- (5) リカード『原理』の第2グループ、ミル『要綱』の第1グループ、マーティノー『例解』第25話の第1グループには、produce、labourer等、2者間で共通する構成要素があるが、3者間で共通する構成要素はない。

4 考察

4.1 リカード『原理』の特徴

リカード『原理』は、労働価値論を理論的基礎として、価値、分配、成長、貿易、租税、貨幣等に関わる理論的・政策的問題を検討するものである。

頻度分析より、リカード『原理』の1-10位の名詞は、price、value、commodity、labor、capital、quantity、profit、tax、corn、moneyであることが分かった。この結果は、リカード『原理』の議

論を十分に彷彿させるものである。とくに、1位のprice(1088回)と2位のvalue(925回)が抜きんでていることは、リカードが価格決定または価値決定の問題を重視していたことを反映していると思われる。また、11-20位の名詞にも、rent、wagesを始め、リカードにとって重要な概念が含まれている。21-30位の名詞にも、やや焦点がぼやけてくるようにも思われるが、gold、tradeを始め、重要な概念が含まれている。

共起ネットワーク分析による中心性を規準としたネットワーク図(図1.1)を見ると、リカード『原理』の議論の中で、price、profitが中心的な役割を果たしていること、produce、laborがそれらに準ずる役割を果たしていることが分かった。この結果も、価格決定と利潤率低下の関係を中心とするリカード『原理』の議論を彷彿させるものである。

共起ネットワーク分析によるサブグラフのネットワーク図(図2.1)を見ると、リカード『原理』の関連名詞の第1グループ(緑色)は、money、value、corn、commodity、quantity、production、labor、priceからなることが分かった。このグループは、主として、リカード『原理』の労働価値論と穀物価格の変化をめぐる議論を反映していると思われる。また、第2グループ(黄色)は、land、produce、rent、profit、tax、landlordからなる。このグループは、主として、リカード『原理』の、所得分配と租税をめぐる議論を反映していると思われる。

以上のとおり、テキストマイニング分析の結果は、リカード『原理』の特徴を大まかにではあるが、適切に示している。

4.2 ミル『要綱』の特徴

ミルはリカードの友人で、リカードに『原理』執筆を促した人物である。ミルは息子のジョン・ステュアート・ミルに『原理』を教えながら作成したノートをまとめ、リカードの経済学の入門書として『要綱』を執筆した。ここでミルはリカードの経済学の諸原理を「生産」、「分配」、「交換」、「消費」の4部門に整理して論じている。

頻度分析より、リカード『原理』の1-10位の名詞は、ミル『要綱』の1-30位にすべて含まれていることが分かった。また、リカード『原理』の11-20位の名詞のうち、ミル『要綱』の1-30位に9語含まれていることが分かった。これらの結果はミル『要綱』がリカード『原理』の議論を比較的忠実に踏襲していることを反映している。ただし、リカード『原理』の1位のpriceはミル『要綱』では16位に、同じく9位のcornは23位に順位を落としている。また、リカード『原理』の21-30位の名詞のうち、ミル『要綱』の1-30位に2語しか含まれておらず、gold、trade等のリカードにとって重要だった概念は含まれていない。これらの結果はリカード『原理』の中心的な議論がミル『要綱』においてやや希薄化されていることを示唆している。

共起ネットワーク分析による中心性を規準としたネットワーク図(図1.2)を見ると、ミル『要綱』において、labor、commodityが中心的な役割を果たしていること、wages、capital、moneyがそれらに準ずる役割を果たしていることが分かった。この結果も、リカード『原理』では中心的

だった議論がミル『要綱』では必ずしも中心的な位置を占めていないということを示唆している。

共起ネットワーク分析によるサブグラフのネットワーク図(図2.2)を見ると、ミル『要綱』の関連名詞の第1グループ(緑色)は、proportion、case、labourer、wages、profit、labor、capital、portion、rent、produce、landからなることが分かった。このグループは、リカード『原理』の第2グループと共通する要素(下線)を複数含んでいることから、ある程度、これに対応していると言える。また、ミル『要綱』の第2グループ(黄色)は、price、commodity、country、quantity、money、value、metalからなる。このグループは、リカード『原理』の第1グループと共通する要素(下線)を複数含んでいることから、ある程度、これに対応していると言える。これらの結果は、ミル『要綱』の議論がリカード『原理』の議論を比較的忠実に踏襲していることを反映していると思われる。ただし、ミル『要綱』には、リカード『原理』に見られなかった小グループ(橙色、赤色、紫色、青色)がいくつも見られる。とくに、supply、demandの2語からなる小グループが見られることは興味深い。これらの結果はミル『要綱』の議論の構成がリカード『原理』よりも分散していることを示唆している。

以上のとおり、テキストマイニング分析の結果は、ミル『要綱』がリカード『原理』の議論を比較的忠実に踏襲しようとしていること、同時に、ミル『要綱』では、リカード『原理』に固有の労働価値論や穀物価格の決定をめぐる議論の構成がやや希薄化していることを示している。

4.3 マーティノー『例解』第25話の特徴

マーティノーは古典派経済学の普及のために、とくに重要な役割を果たした人物である。マーティノーは古典派経済学から導き出される諸々の教訓を、一般大衆が平易に読めるように物語形式にまとめ、『例解』全9巻25話を執筆した。そして『例解』最終話である第25話では、第1話から第24話までの物語の理論的基礎となった経済学の諸原理を提示した。ここでマーティノーはミル『要綱』を参考にし、その4部門構成を継承している¹⁰⁾。

頻度分析より、リカード『原理』の1-10位の名詞のうち、マーティノー『例解』第25話の1-30位には、value、commodity、labor、capital、moneyの5語しか含まれていないこと、リカード『原理』の1位のpriceを始め、quantity、profit、tax、cornの5語は含まれていないことが分かった。それらの順位も、リカード『原理』の2位のvalueがマーティノー『例解』第25話では21位に下がるなど、大きく分散している。これらの単語に代わって、マーティノー『例解』の1位はlabor、2位はcapitalとなっている。また、リカード『原理』の11-20位の名詞のうち、マーティノー『例解』第25話の1-30位には、6語しか含まれていないこと、リカード『原理』の21-30位のうち、マーティノー『例解』第25話の1-30位には2語しか含まれていないことが分かった。その代わりに、マーティノー『例解』第25話の1-30位には、リカード『原理』の1-30位にもミル『要綱』の1-30位にも含まれていなかった名詞が14語も入っている。とくに1-10位に、society、principle、classの3語が入っていることが目に付く。これらの結果は、リカード『原理』の労働価値論を始めとす

る中心的な議論が、マーティノー『例解』第25話では、大きく希薄化されていることを示唆している。

共起ネットワーク分析による中心性を規準としたネットワーク図（図1.3）を見ると、マーティノー『例解』第25話において、capital、societyが中心的な役割を果たしていることが分かった。この結果も、リカード『原理』では中心的だった議論がマーティノー『例解』第25話では中心的な位置を占めていないことを示唆している。

共起ネットワーク分析によるサブグラフのネットワーク図（図2.3）を見ると、マーティノー『例解』第25話の関連名詞の第1グループ（緑色）は、labourer、capitalist、population、wages、capital、labor、proportionからなることが分かった。このグループは、ミル『要綱』の第1グループと共通する要素（下線）を複数含んでいることから、ある程度、これに対応していると言える。また、マーティノー『例解』第25話の第2グループ（赤色）は、commodity、country、money、placeからなることが分かった。このグループは、ミル『要綱』の第2グループと共通する要素（下線）を複数含んでいることから、ある程度、これに対応していると言える。さらに、マーティノー『例解』第25話の第3グループ（青色）は、food、supply、demand、productionからなることが分かった。このグループは、ミル『要綱』の、supply、demandからなる小グループに対応していると言える。これらの結果は、マーティノー『例解』第25話の議論がミル『要綱』の議論に比較的類似した構成を持っていることを反映していると思われる。ただし、マーティノー『例解』には、ミル『例解』よりも、さらに多くの小グループ（黄緑色、黄色、桃色、紫色、灰色、橙色）が見られる。これらの結果はマーティノー『例解』第25話の議論がミル『要綱』の議論よりも、さらに分散していることを示唆している。

以上のとおり、テキストマイニング分析の結果は、マーティノー『例解』第25話の議論がミル『要綱』の議論とある程度類似しているものの、中心的な概念や問題構成について大きく異なっていることを示している。マーティノー『例解』第25話では、リカード『原理』に固有の労働価値論や穀物価格の決定をめぐる問題構成はさらに希薄化しており、これらに代わって、資本と労働の関係、あるいは階級社会の問題が中心的な位置を占めるようになったのである。

5 結 語

5.1 要 旨

テキストマイニング分析とその結果に基づく考察より、リカード『原理』、ミル『要綱』、マーティノー『例解』第25話の間の議論の異同が鮮明になった。頻度分析より、リカード『原理』で出現頻度の高い名詞が、ミル『要綱』ではその順位が分散化し、マーティノー『例解』第25話では大きく順位を下げ、代わって多くの名詞が新たに上位に入っていることが分かった。次第に順位を下げた名詞の多くは、価格決定や所得分配の問題に関わる理論的概念に相当するものであり、順位を

上げた名詞の多くは、比較的大雑把に社会観を表す現実的概念に相当するものだった。共起ネットワーク分析より、リカード『原理』で中心的な役割を果たす名詞が、ミル『要綱』では別の名詞に入れ替わり、マーティノー『例解』第25話ではさらに別の名詞に入れ替わっていることが分かった。ここでも価格決定と所得分配に関わる理論的概念に相当する名詞が、次第に、比較的大雑把な現実的概念にその役割を奪われていったと言える。また、リカード『原理』の関連名詞のグループ分けが、ミル『要綱』ではある程度引き継がれながらも変化・分散し、マーティノー『例解』第25話では決定的に変化・分散していることが分かった。ここでも厳密な理論的概念によって構成された問題群が、次第に分散化され、通俗化されていったと言えるだろう。こうした考察の一定の結論として、リカード『原理』の専門的な議論は、ミル『要綱』において簡略化・平易化され、マーティノー『例解』第25話において決定的に通俗化されることを通して、広く一般大衆に受け入れられていったのである。

従って、本稿の冒頭で提示した仮説は正当化される。すなわち、マーティノー『例解』は一般大衆を読者として想定していたため、リカード『原理』やミル『例解』と比較して、(1)専門的な用語を排除し、一般大衆がより理解しやすい通俗的な用語を用いており、また、(2)理論的な分析よりも、その背景にある世界観や政策論に着目している。本稿では、こうした仮説がテキストマイニング分析によって正当化されることが明らかになった。

5.2 残された課題

上の結論または仮説は、テキストマイニング分析によって正当化されるが、テキストマイニング分析だけによって十分に論証されるというわけではない。言うまでもなく、こうした結論または仮説を十分に論証するためには、テキストの内容を検討し、質的な分析を行うことが不可欠である¹¹⁾。すでに述べたように、テキストマイニング分析はあくまで質的分析を補完するための量的分析であるにすぎない。本稿では、古典派経済学の普及過程を十分に解明したわけではないし、本格的なテキストマイニング分析を行ったわけでもない。こうした課題を完遂することは今後の課題とせざるを得ない。

しかしながら、本稿では、古典派経済学の普及過程における通俗化の一端を明らかにするとともに、経済学史研究におけるテキストマイニング分析の有効性の一端を示した。本稿は古典派経済学の普及過程に関するテキストマイニング分析の試みであり、本格的なテキストマイニング分析を行うための布石であった。

[謝辞] 本稿は文部科学省科学研究費補助金基盤研究(B)「経済思想の受容・浸透過程に関する実証研究：人々は経済学をどのように受け入れたか」(課題番号22330064)の助成を受けた研究成果である。

注

- 1) こうした問題は以前より「経済学の制度化」または「経済政策思想史」として取り組まれてきた問題と密接に関連している。代表的な先行研究として、Coats (1981)、Coats (1993)、Furner & Supple (1990)、池尾 (2006)、西沢・服部・栗田 (1999) を挙げることができる。
- 2) 文部科学省科学研究費補助金基盤研究 (B)、課題番号22330064、研究期間2010-2014年度、研究代表者は下平裕之 (山形大学)、研究分担者は金井辰郎 (東北工業大学)、小峯敦 (龍谷大学)、福田進治 (弘前大学)、船木恵子 (武蔵大学)、古谷豊 (東北大学)、本郷亮 (関西学院大学)、松山直樹 (兵庫県立大学)。
- 3) 経済学史研究にテキストマイニング分析を導入する際の諸問題については、下平・小峯・松山 (2012) 第1節を参照。
- 4) KH Coderは計量テキスト分析のためのフリーソフトウェアであり、テキストデータの中から単語を自動的に取り出し、その結果を集計・解析することにより、データの特徴を探ったりデータを要約したりすることができる。インターネット上のサイト (<http://khc.sourceforge.net/>) より入手可能。使用法については、樋口 (2012) を参照。
- 5) 本研究では、「中心性」を規準としたネットワーク図は「媒介中心性」に基づいて作成される。また「サブグラフ検出」によるネットワーク図は「共起関係の媒介性に基づく方法」を用いて作成される (樋口 2012: 63)。
- 6) 単語グループの色分けはKH Coderの自動処理によって行われる。ただし、これらの色分けにつねに重要な意味があるわけではない。これらの色分けは過度に深読みせず、グラフを解釈する際の補助として利用することが望ましい (樋口 2012: 62)。
- 7) これらの文献のデジタル化されたテキストデータは、インターネット上のサイト「The Online Library of Liberty」 (<http://oll.libertyfund.org/>) より入手することもできる。ただし、このデータには原本のページ数が記載されていない等の問題がある。
- 8) 出現頻度の集計の際、複数形はKH Coderの自動処理によって単数形に変換され、単数形とともにカウントされる。また、英綴り (例えば、labour) は自動的に米綴り (labor) に変換される。こうした諸々の調整が自動的に行われるので、結果の数値に過度に厳密な解釈を与えることは控えなければならない。
- 9) それぞれの単語に関する Jaccard 係数の具体的数値は、KH Coderの「関連語検索」というコマンドを用いて確認することができる。例えば、リカード『原理』の出現頻度1位の名詞 price と他の名詞に関する Jaccard 係数の数値 (1-10位) は、下の付表のとおりである。

付表 リカード『原理』における名詞priceに関する Jaccard係数 (1-10位)

順位	抽出語	全体	共起	Jaccard係数
1	commodity	592	264	0.2435
2	corn	411	214	0.2246
3	profit	452	171	0.1649
4	money	369	157	0.1622
5	wages	323	136	0.1442
6	produce	366	141	0.1437
7	rise	206	119	0.1412
8	labor	514	156	0.1400
9	tax	410	133	0.1288
10	quantity	450	132	0.1229

ただし、表中の「全体」とは、分析対象ファイル内でその単語がいくつの文章に出現したのかという度数を、「共起」とは、その単語が price と同時に出現している文章の出現度数を示している。こうした数値に基づいて、「中心性」を規準としたネットワーク図と「サブグラフ検出」に基づくネットワーク図が描画されるのである。なお、このコマンドの詳細については、樋口 (2012: 49-52) を参照。

- 10) マーティノー『例解』の概要またはその影響については、船木2006、櫻井2003、上宮1996を参照。
- 11) 文献データの量的分析と質的分析の架け橋を行う機能として、KH Coderには「KWICコンコーダンス」というコマンドがある。このコマンドを用いることにより、抽出された語が分析対象ファイル内でどのように用いられていたのかという文脈を探ることができる。詳細については、樋口（2012:44-49）を参照。

参考文献

- Coats, A.W. (ed.) 1981, *Economists in Government: An International Comparative Study*, Durham: Duke University Press
- Coats, A.W. 1993, *The Sociology and Professionalization of Economics*, London & New York: Routledge
- Furner, M.O. & Supple, B. (eds.) 1990, *The State and Economic Knowledge: The American and British Experiences*, Cambridge: Cambridge University Press
- Martineau, H. 1932-34, *Illustrations of Political Economy*, 9 vols., London: Charles Fox.
- Mill, J. 1821, *Elements of Political Economy*, London: Printed for Baldwin, Cradock and Joy. 渡辺輝雄（訳）1948『経済学綱要』（古典経済学叢書）春秋社
- Ricardo, D., Sraffa, P. (ed.) 1951-73, *The Works and Correspondence of David Ricardo*, 11 vols., Cambridge: Cambridge University Press. 堀 経夫他（訳）1969-99『デイヴィッド・リカード全集』全11巻、雄松堂書店
- 船木恵子 2006「経済学史におけるハリエット・マーティノウ『経済学例解』」『武蔵大学総合研究所紀要』16: 151-169
- 樋口耕一 2012「KH Coder 2.x リファレンス・マニュアル」(<http://khc.sourceforge.net/>)
- 池尾愛子 2006『日本の経済学-20世紀における国際化の歴史』名古屋大学出版会
- 喜田昌樹 2008『テキストマイニング入門』白桃書房
- 金 明哲 2009『テキストデータの統計科学入門』岩波書店
- 松村真宏・三浦麻子 2009『人文・社会科学のためのテキストマイニング』誠信書房
- 西沢 保・服部正治・栗田啓子（編）1999『経済政策思想史』有斐閣
- 下平裕之・小峯 敦・松山直樹 2012「経済学史研究におけるテキストマイニング分析の導入：ケインズ『一般理論』と書評の関係」*Discussion Paper Series, Research Group of Economics and Management*, No.2012-E02, Yamagata University
- 櫻井 毅 2003「マーティノウ『経済学例解』（1832-1834）の刊行まで」『武蔵大学論集』50 (2): 201-63
- 上宮正一郎 1996「H・マルティノウと経済学」『国民経済学雑誌』173 (4): 31-51

表1 リカード『原理』、ミル『要綱』、マーティノー『例解』第25話の名詞出現頻度1-30位

順位	リカード『原理』		ミル『要綱』		マーティノー『例解』第25話	
	名詞	頻度	名詞	頻度	名詞	頻度
1	price	1088	capital	446	labor	159
2	value	925	commodity	408	capital	148
3	commodity	750	labor	400	man	83
4	labor	684	value	341	time	80
5	capital	637	quantity	314	society	77
6	quantity	596	country	254	country	73
7	profit	573	case	242	labourer	72
8	tax	568	money	213	principle	70
9	corn	552	wages	208	production	64
10	money	500	production	194	class	60
11	country	496	produce	187	commodity	58
12	rent	472	profit	184	proportion	56
13	produce	452	tax	179	capitalist	55
14	land	432	man	170	power	54
15	wages	410	land	168	money	51
16	production	284	price	159	case	50
17	proportion	277	demand	142	demand	44
18	rise	243	exchange	130	population	44
19	demand	242	thing	127	exchange	43
20	labourer	216	portion	126	law	42
21	effect	212	labourer	122	value	42
22	landlord	202	cloth	114	food	41
23	rate	201	corn	114	supply	41
24	goods	196	metal	114	state	40
25	market	191	rent	114	increase	38
26	gold	182	goods	112	wages	38
27	trade	176	proportion	106	means	37
28	time	164	cost	98	wealth	37
29	man	160	whole	97	number	36
30	farmer	159	supply	93	people	36

注) 赤字はリカード『原理』の1-10位の名詞を示す。
 青字はマーティノー『例解』第25話のみで1-30位に入った名詞を示す。

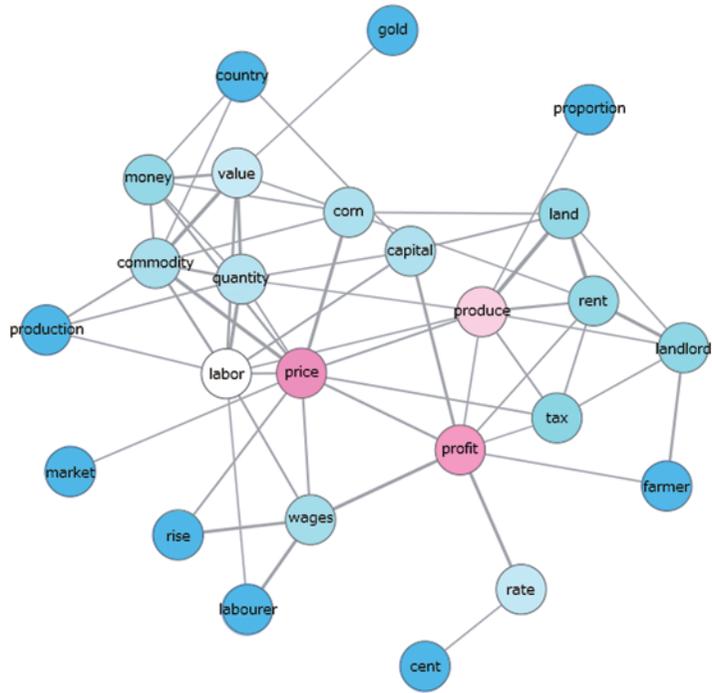


図 1.1 リカード『原理』出現名詞の共起ネットワーク（中心性の分析）

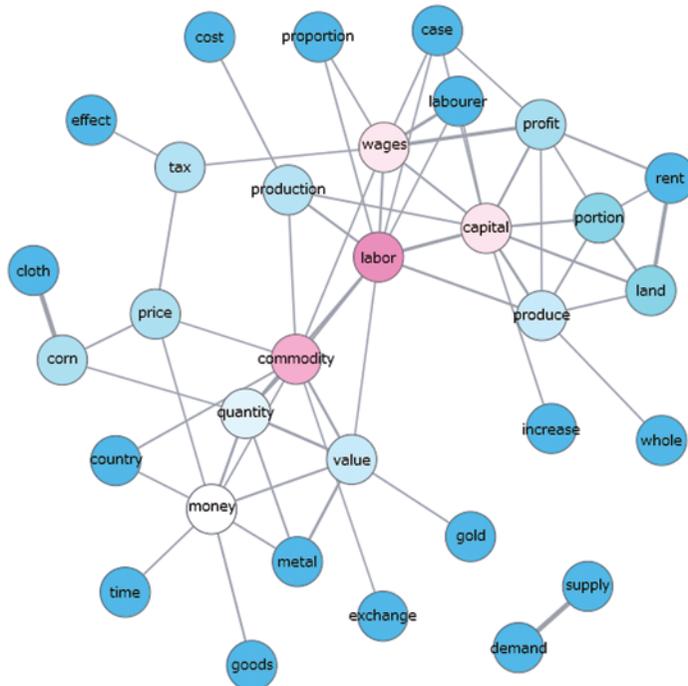


図 1.2 ミル『要綱』出現名詞の共起ネットワーク（中心性の分析）

