

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 9月 12日現在

機関番号：82640

研究種目：挑戦的萌芽研究

研究期間：2010～2011

課題番号：22653032

研究課題名（和文） イノベーションファクターによる新たなサービス産業分類の構築

研究課題名（英文） Analysis of the Characteristics of the Innovation System for the JSIC Service Sector

研究代表者

尾崎 雅彦 (OZAKI MASAHIKO)

独立行政法人経済産業研究所・研究グループ・上席研究員

研究者番号：50470068

研究成果の概要（和文）：潜在成長力の低下が懸念される日本経済において、GDPの7割を超えるサービス産業の活性化は重要課題である。しかし、サービス産業の分析においてはデータ制約問題（製造業に比して業種別データが乏しい）と業種分類問題（同一業種にイノベーション生成上異なる性質を持つ企業が混在）が存在する。本研究では、業種分類問題緩和のため、テキストマイニング手法を用いて日本標準産業分類／細分類業種のイノベーション・システム特性を明らかにし、その上で各業種を同特性に基づきセクター分類（カテゴリー化）した。

研究成果の概要（英文）：The vitalization of the service sector is one of the major economic issues in Japan, where the service sector accounts for more than 70% of GDP. However, empirical studies on the innovation factors and productivity in the service sector have faced significant difficulties due to problems of the sectoral classification (e. g. existence of establishments with varying characteristics of the innovation system in each sector). This paper challenges the problem by making categories based on the characteristics of the innovation system for the Japan Standard Industrial Classification (JSIC) service sectors (4-digit). Applying a text mining method, we quantitatively measure sensitivities of three innovation factors, and reveal the characteristics of the innovation system which each service sector has. We also classify the JSIC service sectors (4-digit) into 8 eight categories that have specific innovation potential in according to the proper innovation system obtained from the revealed characteristics.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	2,000,000	0	2,000,000
2011年度	800,000	240,000	1,040,000
総計	2,800,000	240,000	3,040,000

研究分野：社会科学

科研費の分科・細目：経済学・応用経済学

キーワード：サービス産業、生産性、イノベーション、テキストマイニング、JIP、JSIC

## 1. 研究開始当初の背景

(1) 1998年以降労働力人口の減少局面にある日本経済において、GDPの7割を占めるサービス産業の生産性向上は重要な課題である。しかし、米国において1990年代

後半にサービス産業における生産性成長の決定要因や制度などの環境要因に関する分析が急速に進展し、サービス産業に対する認識を「非イノベティブな停滞産業」から「新たな知識集約型成長産業」に転換させるに至

ったのに対し、わが国における研究蓄積は現状では豊富とは言えず体系的で明確な研究成果が求められている。

(2) サービス産業分析を困難にしている背景には、統計未整備などによるデータ制約と分析対象である産業分類が生産性分析の観点で適切でないことがあげられる。後者に関して言えば、日本標準産業分類（以下、**JSIC**）を用いた産業分析では、同分類が品目中心の分類であり近年の需要の多様化や情報技術の進歩等によるビジネスモデルの変化に十分には対応していないため、異なるビジネスモデルを持つ業種が混在する集合を一つの産業として分析することにならざるをえず、規制緩和などの外部要因による変化やIT投資による経営改革などの内部要因の変化が生産性に与える影響を産業レベルで明らかにすることが困難となっている。そのため、業種毎の特性を分析し、特性に応じたセクター分類（カテゴリー化）を実施する必要がある。

## 2. 研究の目的

(1) 業種分類問題に対応するため、**F. Malerva** が提唱する **SSI**（セクトラル・システムズ・オブ・イノベーション）の概念に示されたイノベーション生成要因（以下、イノベーション・ファクター）、具体的には i 「知識と技術」、ii 「アクターとネットワーク：アクター間の相互作用」および iii 「制度：規制等」に対する、各業種におけるアクターの感応度をテキストマイニングにより定量的に計測し、その高低に基づき **JSIC** 細分類（4桁分類）業種を8種のセクターに分類（カテゴリー化）することを試みる。

(2) また、各セクターのイノベーション・システム特性から予測されるイノベーション生成の潜在的可能性（以下、イノベーション・ポテンシャル）の有無を確認するため、イノベーション促進的イノベーション・ファクターである上述 i および ii に対する感応度と長期パフォーマンス指標（**JIP** データベース産出額および **TFP** の変化率）との相関関係を分析するとともに、イノベーション抑制的イノベーション・ファクターである iii の影響を把握するため、iii の感応度が高い業種グループと低い業種グループに二分し同様の分析を行う。

(3) 以上の分析により、**JSIC** 細分類業種をイノベーション・システム特性に基づきセクター分類することができれば、**JSIC** 大～小分類に拘わらず、比較的類似の同特性を持つ業種集団での比較分析が可能となる。

## 3. 研究の方法

(1) 本研究では、イノベーション・ファクターの感応度を、分析対象業種について書かれたテキストデータの中にイノベーション・ファクターに関連する言葉（以下、**SSI** キーワード）が出現する頻度によって計測する。ある業種について書かれた文章は、執筆者が（当該業種の利害関係者である）読者の問題意識を念頭に作成したものであるから、その文章中において特定の **SSI** キーワードが多用されるということは、その **SSI** キーワードに対応するイノベーション・ファクターの変化に対して当該業種の利害関係者が強い関心を持っていると言える。よって、感応度の代理変数として **SSI** キーワードの出現率を用いる。これにより感応度を客観的・定量的に計測することが可能となる。

出現率算出のためには、テキストマイニング技術を用いる。テキストマイニングは、定型化されていない文章（自然文）を自然言語解析し、特定の単語やフレーズの出現頻度を計測することで、有用な情報を抽出する技術・手法であり、応用例の一つとして、所定のテキストデータにおける特定の単語の出現率を算出することができる。

具体的には、以下の工程により感応度の計測（出現率の算出）を行った。

- ① テキストデータベース構築
- ② 業種分類キーワード作成
- ③ 業種別テキストデータベース構築
- ④ **SSI** キーワード作成
- ⑤ 業種別テキストデータにおける **SSI** キーワード出現率の算出および業種別のイノベーション・ファクター感応度の算出

①については、企業ヒアリング、業種別貸出審査辞典（金融財政事情研究会）、各種市場調査レポート、業界団体発行物および論文等を収集し比較検討したが、最終的には執筆者のバイアスが最も少なくかつ情報量が最も多い新聞記事データを採用した。分析対象となる新聞記事データは、日経新聞朝夕刊、日経産業新聞および日経地方経済面の全文記事データ（1991～2010年）であり、データサイズは、記事数516万件、文字数23億字である。このうち、文化欄、スポーツ欄、人事または広告欄など産業、企業関連情報が含まれていない記事を除外し、また、解析ソフトで対応できない記号等を含む記事は修正または修正できない場合は除外した。その結果、分析に用いるテキストデータベースのサイズは記事数194万件、文字数9億字となった。

②に関しては、まずどの業種を分析対象とするかを決定する必要がある。サービス産業の範囲を、**JSIC** 大項目 G～T（細分類 702 業種）と想定し、このうち、公的機関など、経

営スタンスにおいて事業性のウェイトが低い業種は、特異な（比較困難な）イノベーション・システム特性を持つ可能性が高いと判断し分析非対象とした。また、主として管理事務を行う本社等やその他の管理、補助的経済活動を行う事業所等の事業実体のない細分類業種も分析非対象とし、残る435細分類業種を分析対象業種とした。

業種分類キーワードは、第12回改訂（2007年改訂）JSIC業種説明テキストデータおよび社団法人金融財政事情研究会の第11次業種別審査辞典（2008年刊行）本文テキストデータから抽出を行っている。

③業種別テキストデータベースは、①のテキストデータベースに収納される記事を②から得られた業種分類キーワードが文章中に含まれているか否かによってどの業種に関して書かれた記事であるかを判定し集約して構築する。

業種別感応度の計測結果に正確を期するため、このプロセスにおいて、複数の業種が関連する或いは業種特性に関連しない等の記事を極力除外し、その上で記事数が30件未満の業種を分析非対象とするとの条件設定を行った結果、分析対象業種は42業種に絞られた。

④SSIキーワードは、F. Malerva 論文中に示されたイノベーション・ファクターに関連する単語を和訳し、独立行政法人情報通信機構が開発した「日本語 Wordnet」を用いて類義語を抽出しそれらを付加することで拡充し策定した。拡充後のキーワード群を実際に新聞記事データで形態素解析を行い、出現頻度が0であったもの或いは特定の業種に著しく偏って使用されているもの（偏差値が100以上のもの）は除外し、i「知識と技術」に関連するキーワード（以下、SSI-1キーワード）10語、ii「アクターとネットワーク：アクター間の相互作用」に関連するキーワード（以下、SSI-2キーワード）9語およびiii「制度：規制等」に関連するキーワード（以下、SSI-3キーワード）17語を特定した。

⑤SSIキーワード出現率は、③で得られた業種別テキストデータベースの総名詞数を分母に、SSIキーワードの出現頻度を分子にして算出する。この結果を、SSI-1キーワードグループ、SSI-2キーワードグループおよびSSI-3キーワードグループ毎に集計することにより、各業種固有のイノベーション・ファクター出現率（感応度）を算出する。

(2)次に、業種別の3種のイノベーション・ファクター感応度を分析対象業種間で相对比较することにより、セクター分類（カテ

ゴリー化）が可能となる。具体的には、3種のイノベーション・ファクター感応度を業種間で比較するため偏差値で表現し、各イノベーション・ファクター感応度が各々他の業種に比較して相対的に高いか低いかを判定した。この結果により、平均値以上をH、未満をLと表すと2の3乗の組み合わせが得られ8種のセクターが仮定される。

F. Malervaが提唱するSSIの概念では、「産業セクター内において、共通した知識ベースや技術を有する（しかし、特化された習熟過程、能力、組織構造、信念および目的を持つ）アクターが、コミュニケーション、交換、取引、協力、競争および命令等のネットワーク活動を通じて相互作用を生み出し、イノベーション・生産活動を行う。そして、ルールや規制などの制度によってその活動はshapeされる」と考えられており、これに基づけばSSI-1キーワードによって特徴付けられるイノベーション・ファクターをSSI-1、同様にその他の2つのイノベーション・ファクターをSSI-2、SSI-3とすると、SSI-1およびSSI-2はイノベーション促進的なイノベーション・ファクターであり、SSI-3は抑制的なイノベーション・ファクターである。

従って、分析対象のサービス産業JSIC細分類42業種は、イノベーション促進的なイノベーション・ファクター（SSI-1およびSSI-2）の高低によって4種、さらにSSI-3の高低によって2分割され計8種のイノベーション・システム特性の異なるセクターに分類される。

なお、研究当初において想定していた需要の共通性および産業関連性に係るイノベーション・システム特性については、分析対象となる企業や産業固有のイノベーション生成上の問題点把握やイノベーション政策等の政策評価分析を行うことを想定した場合には、イノベーション・システム特性を3種のイノベーション・ファクターに限定し分析することが望ましいと判断し、本研究での分析対象とはしなかった。

(3)3種のイノベーション・ファクターの感応度（SSIキーワードの出現率）により決定されたイノベーション・システム特性によって形成された各セクターは、同特性に基づき異なるイノベーション・ポテンシャルを持つ。たとえば、SSI-1およびSSI-2の感応度が高く、かつSSI-3の感応度が低いイノベーション・システム特性を持つ業種は、イノベーション・ポテンシャルが高いことが予測され、逆は逆であることが予測される。

この予測の正しさは、イノベーション・ファクターの感応度とイノベーション・ポテンシャルとの関係を実際のデータで観測することで確認できる。具体的には、SSIキーワー

ド感応度の高低を決定している出現率と中長期的なパフォーマンス指標との相関関係を計測することにより分析することができる。

本研究では、深尾一橋大教授および富川学習院大教授等により構築され、経済産業研究所で公開されている、全産業の長期時系列データを格納する JIP データベースにおける産出額成長率と TFP 伸び率をパフォーマンス指標として用いている。

#### 4. 研究成果

(1) 本研究では、テキストマイニング手法を用いて日本標準産業分類 (以下 JSIC) の細分類業種のイノベーション・システム特性を明らかにし、その特性の類似性によっての細分類業種をセクター分類 (カテゴリー化) することを試みた。SSI-1 と SSI-2 の感応度の高低の組み合わせを横に、SSI-3 の感応度の高低を縦に区分して示したマトリックスに、該当するイノベーション・システム特性を持った細分類業種を配したものが下図のセクター図である。

図 1

		SSI-1, SSI-2			
		SH	HL	LH	LL
SSI-3	L	372 移動電通業	4121 レコード制作業	5222 酒類販売業	4211 普通鉄道業
		3923 その他の情報処理・提供サービス業	4141 出版業	5301 コンビニストア/飲食料 品を中心とするものに限る	5611 百貨店・総合スーパー
		4012 アプリケーション・サービス・コンテンツソフト	5011 各種商品販売業(従業者が 100人以上のもの)	6020 ドラッグストア	6020 定期乗船業
		7241 公認会計士事務所	6090 中古品小売業(骨とう品を除く)	6091 ホームセンター	7311 旅館、ホテル
		7242 税理士事務所	6510 投資運用業	7011 広告業	7311 食堂、レストラン・専門料亭店 を除く
			6201 郵便業	7311 旅行業・旅行業付随業を除く	
				8011 映画館	
				8052 遊園地・テーマパークを除く	
				8053 テーマパーク	
				4711 倉庫業(冷蔵倉庫業を除く)	
H	3831 有線テレビジョン放送業	7011 総合リース業	3822 児童放送業	4711 倉庫業(冷蔵倉庫業を除く)	
	4013 インターネット利用サポート業	3211 一社病院	4201 一社専用旅客自動車運送業	5911 自動車(新車)小売業	
		4011 児童運送業	6020 郵便貯行	6221 普通貯行	
		6020 ガソリンスタンド	6222 協賛貯行	6711 生命保険業(郵便保険業、生 命保険再保険業を除く)	
		6411 消費者向け貸付業	6420 クレジットカード業		
		6721 損害保険業(損害保険再保険 業を除く)			

F. Malerva が提唱する SSI の概念を用いて、特定業種のケース・スタディ分析においてイノベーション・システム特性を明らかにした優れた研究はこれまでに多数見られるが、既存業種のセクター分類 (カテゴリー化) に応用された研究は現在のところ見当たらない。また、テキストデータを用いた業種分類の試みとしては、財務諸表内の製品説明 (product description) の記述から企業間の関係性をソーシャルネットワーク形成に類似する考え方で導出しカテゴリー化を行った Hoberg 他 (2010) の研究があるが、既存の理論的概念に基づきテキストデータによって定量的に業種カテゴリー化を行った研究はこれま

でのところ本研究のみである。

(2) 本研究結果であるセクター分類 (カテゴリー化) が現実経済に合致しているか否か、言い換えれば本研究によって明らかにされた各業種のイノベーション・システム特性が予測通りのイノベーション・ポテンシャルを有しているか否かを、SSI キーワード感応度の高低を決定している出現率と中長期的なパフォーマンス指標との相関関係を計測することにより明らかにした。

まず、イノベーション促進的なイノベーション・ファクターである SSI-1 および SSI-2 の感応度 (出現率) を合計した数値と、二つのパフォーマンス指標 (JIP 名目産出額成長率と JIP 全要素生産性伸び率) との相関係数は、以下の通りいずれも正の相関関係にあることが確認された。

表 1

JIP 名目産出額成長率との相関係数	+0.63
JIP 全要素生産性伸び率との相関係数	+0.31

また、42 業種をイノベーション抑制的なイノベーション・ファクターである SSI-3 感応度 (出現率) が低い SSI-3 (L) グループ (25 業種) と高い SSI-3 (H) グループ (17 業種) に二分し同様の分析を行った結果、SSI-3 (H) グループの業種では相関係数は下がり、逆に SSI-3 (L) グループの業種では相関係数は上がることが確認された。

表 2

JIP 名目産出額成長率 SSI-3 (L) グループ	+0.72
JIP 名目産出額成長率 SSI-3 (H) グループ	+0.50
JIP 全要素生産性伸び率 SSI-3 (L) グループ	+0.48
JIP 全要素生産性伸び率 SSI-3 (H) グループ	+0.16

このセクター分類 (カテゴリー化) により、イノベーション研究や生産性分析において、少なくとも 3 つの効果期待できる。まず、本稿の目的である業種分類問題の緩和が果たされる。JSIC 細分類 (4 桁分類) 業種のイノベーション・システム特性が明らかにされることで、上位の JSIC 業種分類 (1~3 桁分類) を取り扱う際にイノベーション・システム上、異質の JSIC 細分類 (4 桁分類) 業種の混在を排除することが可能となる。次に、イノベーション・システム特性がイノベーショ

ン・ファクターの感応度の高低によって表されているので、業種レベルのイノベーション生成における問題点抽出を容易にする。さらに、同一セクターに属する JSIC 細分類 (4 桁分類) 業種は JSIC 業種分類 (1~3 桁分類) において異なる業種であってもイノベーション・システム特性に類似性があるため、異業種間で分析結果を共用することが可能となり、データ制約問題の緩和に寄与することも期待される。

今回の研究で得られたイノベーション・ファクター情報は細分化可能であり、たとえば、知識・技術を科学と技能、相互作用を競争と協調、制度を内部と外部にブレークダウンして分析すれば多角的視点でより多様なセクター分類が可能となる。また、本研究プロセスで確立したイノベーション・ファクター感応度計測手法を用いて分析対象業種を全産業にまで拡張すれば、業種横断的な外部ショックの影響を考慮に入れたマクロレベルでの効率的なイノベーション生成プロセス解明を促すことが可能となる。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕 (計 1 件)

尾崎雅彦、「JSIC サービス産業業種のイノベーション・システム特性分析」～テキストマイニングによるイノベーション・ファクター感応度の計測～, RIETI Discussion Paper Series, 査読無, 12-J-032, 2012, pp. 1-33.  
[http://www.rieti.go.jp/jp/publications/act\\_dp.html](http://www.rieti.go.jp/jp/publications/act_dp.html)

〔学会発表〕 (計 0 件)

〔図書〕 (計 0 件)

〔産業財産権〕

○出願状況 (計 0 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年月日：  
国内外の別：

○取得状況 (計 0 件)

名称：  
発明者：  
権利者：

種類：  
番号：  
取得年月日：  
国内外の別：

〔その他〕

ホームページ等

[http://www.rieti.go.jp/jp/publications/act\\_dp.html](http://www.rieti.go.jp/jp/publications/act_dp.html)

#### 6. 研究組織

##### (1) 研究代表者

尾崎 雅彦 (OZAKI MASAHIKO)

独立行政法人経済産業研究所・研究グループ・上席研究員

研究者番号：50470068

##### (2) 研究分担者

加藤 篤行 (KATO ATSUYUKI)

独立行政法人経済産業研究所・研究グループ・リサーチアソシエイト

研究者番号：10470064

##### (3) 連携研究者

森川 正之 (MORIKAWA MASAYUKI)

経済産業研究所・その他部局・副所長

研究者番号：70272284

小滝 一彦 (ODAKI KAZUHIKO)

独立行政法人経済産業研究所・研究グループ・上席研究員

研究者番号：60314431

