

SUT 作成シミュレーションのための擬似マイクロデータの作成
Creation of Pseudo-micro Data for Simulating SUT Estimation

菅 幹雄

統計研究研修所客員教授

法政大学経済学部教授

SUGA Mikio

SRTI Guest Professor

Professor, Faculty of Economics, HOSEI University

宮内 環

統計研究研修所客員教授

慶應義塾大学産業研究所兼任所員

MIYAUCHI Tamaki

SRTI Guest Professor

Associate Research Fellow, Keio Economic Observatory, Keio University

宮川 幸三

統計研究研修所客員教授

立正大学経済学部教授

MIYAGAWA Kozo

SRTI Guest Professor

Professor, Faculty of Economics, RISSHO University

令和 5 年 10 月

October 2023

総務省統計研究研修所

Statistical Research and Training Institute (SRTI)

Ministry of Internal Affairs and Communications

受理日：令和5年9月29日

本ペーパーは、総務省統計研究研修所の客員教授である執筆者が、その責任において行った統計研究の成果を取りまとめたものであり、その内容については、総務省統計局又は統計研究研修所の見解を表したものではありません。本ペーパーの内容については、執筆者に問い合わせ願いたい。

SUT 作成シミュレーションのための擬似マイクロデータの作成

菅 幹雄
宮内 環
宮川 幸三

概要

統計改革推進会議「最終取りまとめ」（平成 29 年 5 月 19 日）により、産業連関表から供給・使用表（Supply Use Tables, SUT）への移行が決定した。だが、「供給表・使用表を直接推計し、それを用いて産業連関表を推計する」予定である令和 7 年（2025 年）の SUT 及び産業連関表（Input Output Table, IOT）の作成方法は未確定である。SUT 及び IOT の作成方法は各種ある。各種の作成方法について、それぞれ SUT 及び IOT を試算し、評価・比較することが望ましい。そのためには企業・事業所の産出、投入に関するすべての情報があり、そのまま集計すれば IOT、SUT になるようなデータがあることが理想的である。また、推計が偶然成功する可能性もあるため、推計作業を複数回実行できるように、複数のセットのデータがあることも理想的である。だが、実際にはそのような理想的なデータは存在しない。そこで経済センサス-活動調査の集計結果に基づいて SUT 及び IOT を推計するための擬似マイクロデータを 100 セット作成した。

キーワード：供給・使用表、経済センサス-活動調査、擬似マイクロデータ

Creation of Pseudo-micro Data for Simulating SUT Estimation

SUGA Mikio

MIYAUCHI Tamaki

MIYAGAWA Kozo

Abstract

The “Final Report of the Statistics Reform Promotion Council” (Decision by the Statistics Reform Promotion Council on May 19, 2017) decided to shift from input-output tables to Supply Use Tables (SUT). However, the method of creating the SUT and the Input Output Table (IOT) in 2025, which "directly estimates the supply and use tables and uses them to estimate the input-output table," has not yet been established. There are various methods of simulating SUTs and IOTs. It is desirable to estimate the SUT and IOT for each method, and evaluate and compare them. For that purpose, it is ideal to have all information related to the output and input of companies and establishments, and to have data that can be aggregated as IOT and SUT. It is also ideal to have multiple sets of data so that the estimation task can be run multiple times. However, actually, such ideal data do not exist. Therefore, 100 sets of pseudo-micro data for estimating SUT and IOT were created based on the aggregated results of the Economic Census- Activity Survey.

Keywords: supply/use table, Economic Census-Activity Survey, pseudo-micro data

1. はじめに¹

統計改革推進会議「最終取りまとめ」（平成29年5月19日）により、産業連関表（Input Output Table, IOT）から供給・使用表（Supply Use Tables, SUT）への移行が提言され、第三期基本計画（変更後／令和2年6月2日閣議決定）に基づいて移行に向けた検討と準備を関係府省が一体となって推進することになった。そこでは令和2年（2020年）表について「供給表とサービス分野の産業（『サービス産業・非営利団体等調査』の対象産業）に関する使用表を推計」、令和7年（2025年表）については「供給表・使用表を直接推計し、それを用いて産業連関表を推計する」方針が定められた。だが、令和7年（2025年表）については、SUTおよびIOTを推計する具体的な方法が未確定である。

一般的に言えば、SUT及びIOTの作成方法は各種ある。例えば、我が国のようにIOTと供給表から使用表を推計する方法、国際的によく行われる供給表と使用表からIOTを推計する方法などさまざまである。各種の作成方法について、それぞれSUT及びIOTを試算し、比較・評価することが理想的である。また、推計が偶然成功する可能性もあるため、推計作業を複数回実行できるように、複数のセットのデータがあることが理想的である。

本稿ではSUT及びIOTを推計し、評価・比較するための企業・事業所のデータの存在を想定する。そこでは企業・事業所の産出、投入に関するすべての情報があり、そのまま集計すればIOT、SUTになる。しかも、複数のセットのデータが存在する。だが、実際にはそのような理想的なデータは存在しない。そこで経済センサス-活動調査の集計結果に基づいて擬似的に作成した。

また擬似マイクロデータは、一般にも提供できるため、第三者による検証が可能というメリットがある。科学的方法においては、第三者の追試を経て、同じ結果が出た後に、実証分析結果が事実であると認められる。そのためには同じデータを第三者が利用できることが不可欠である。この場合で言えば、SUT及びIOTの推計の基礎となるマイクロデータを第三者が利用できることが不可欠である。公的統計のマイクロデータについては、各国の法制度や利用者のニーズによってマイクロデータの提供形態が異なっている（伊藤(2020)）。様々な提供形態の中でもアクセスの敷居が低いものに匿名データがあるが、我が国では、事業所・企業系の統計調査に関する匿名データの作成は現状では実現していない（横溝・伊藤(2022)）。

なお横溝・伊藤(2023)は「経済センサス - 活動調査」の個票データを用いて、事業所・企業系の統計調査を対象にした合成データの作成可能性を追究し、それだけでなく有用性と秘匿性に関する各種の評価方法に基づいて、合成データの作成手法の有効性に関する定量的な評価を行っている。

¹ 本研究は総務省統計局、統計研究研修所との共同研究の一環としてなされたものである。また本研究の一部は、文部科学省科学研究費補助金（基盤研究（C））、研究課題：SUT体系における商業サービスの表章および推計手法に関する調査研究、研究代表者：宮川幸三、研究課題番号：(20K01.596)、令和2年度～令和5年度）の助成を受けて行っている。

2. 先行研究

2.1 擬似マイクロデータ作成

我が国の統計センターでは、総務省統計局から委託を受け、統計演習など教育用に利用可能な「一般用マイクロデータ」の無償提供を行っている²。「一般用マイクロデータ」とは、集計表から作成するなど、調査票情報を直接的に用いない方法により作成した擬似的なマイクロデータのことである。これはもともと「教育用擬似マイクロデータ」として開発されたものである。河野・和田（2018）によれば、統計センターにおいて開発された「教育用擬似マイクロデータ」は、「平成 23 年 8 月に試行提供開始、平成 24 年 8 月に利用範囲が教育用以外にも拡大され、名称も『擬似マイクロデータ』となった。...中略...ところが、その後平成 26 年を始期とする『公的統計の整備に関する基本的な計画』（第Ⅱ期基本計画）において、『一般用マイクロデータ』の提供推進が謳われ、その作成方法に関する検討が行われた後、平成 28 年 3 月に『一般用マイクロデータ』が提供されるようになり、『擬似マイクロデータ』の試行提供は終了した」。

山口・伊藤・秋山（2013）によれば「教育用擬似マイクロデータ」の作成方法の基本的な考えとは、①個票データから高次元の集計表を作成し、②個票データの量的属性値が多変量（対数）正規分布にしたがうことを仮定し、高次元の集計表を用いて、多変量（対数）正規乱数を作成することである。「教育用擬似マイクロデータ」の後継の「一般用マイクロデータ」も、集計表から作成する方法を採用している。これを参考として、本稿では、事業活動別売上(収入)金額および商業マージン額を除き、擬似マイクロデータを集計表から作成することにした³。

統計センターは 2023 年 8 月現在、「全国消費実態調査」（平成 21 年）および「就業構造基本調査」（平成 4 年～24 年）の一般用マイクロデータを提供している。いずれも世帯統計であり、産業統計ではない。なお、海外では産業統計についても擬似マイクロデータを作成し、提供している事例がある。Vilhubera, Abowda and Reite (2016)によれば、縦断的ビジネスデータベース（Longitudinal Business Database, LBD）を主に用いて、Synthetic Longitudinal Business Database（SynLBD）と呼ばれる擬似マイクロデータセットが作成され、Web アクセスが容易なコンピューティング環境で提供された。SynLBD と同時期にドイツのビジネスデータセットを用いた擬似マイクロデータも提供された。したがって、産業統計の擬似マイクロデータは我が国ではまだないが、海外では既に提供されている。

² 独立行政法人統計センター「一般用マイクロデータの利用」、アクセス：

<https://www.nstac.go.jp/use/archives/ippan-microdata/>、最終アクセス日：2022 年 5 月 6 日

³ 第 4 節で示すように、事業活動別売上(収入)金額および商業マージン額については、経済センサス-活動調査の個票データに基づいて擬似マイクロデータを作成している。

2.2 オランダの DPI キューブ

本稿では、DPI キューブ (Domestic-Product-and-Imports Cube) の考え方に基づき、SUT 及び IOT を推計するための擬似マイクロデータを作成した。DPI キューブとは、オランダ統計局が約 30 年前に IOT から SUT へ移行した際に考案したものである (図 1)。

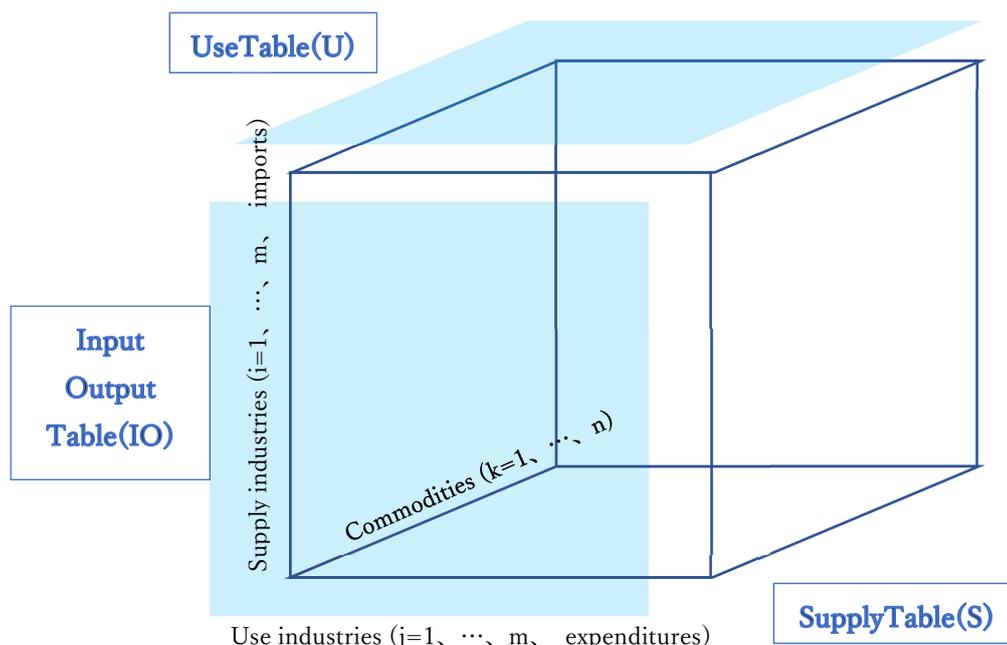


図 1 オランダ統計局の DPI キューブ

(出所) Kazemier, Driessen & Hoogbruin(1991)

DPI キューブは三次元の表であり、その周辺分布が供給表、使用表、IOT になっている。SUT へ移行するまでオランダの IOT は産業×産業表であった。オランダでは SUT は 1987 年分から推計されていたが、国民経済計算の基準改定のため 1970 年以降の時系列が必要となった。そこで IOT から SUT を導出する方法が開発された。その中で DPI キューブが考案された。

DPI キューブのアイデアは供給表、使用表、IOT の関係を明確に示している点で優れている。ただし、このようなデータを構築するためには膨大なデータを必要とすると考えられ、行政記録が充実し、かつ規模が小さい (人口は 2021 年で 1755 万人、日本の約 7 分の 1) 国であるからこそ作成できたと考えられる。

DPI キューブはその周辺分布が供給表、使用表、IOT になっており、三つの表の関係を明確に示すことができるという意味で優れた概念である。そして DPI キューブを構築できる理想的なデータがあるのであれば、供給表、使用表、IOT を整合的に作成することができる。本稿の擬似マイクロデータのフォーマットは DPI キューブが作成できるように設定した。

なお、本稿における DPI の形式のイメージは、オランダ統計局の DPI キューブを日本向けに改変したものであり、オランダ統計局の DPI キューブの産業×産業表が、生産物×生産物表に入れ替わっている（図 2）。

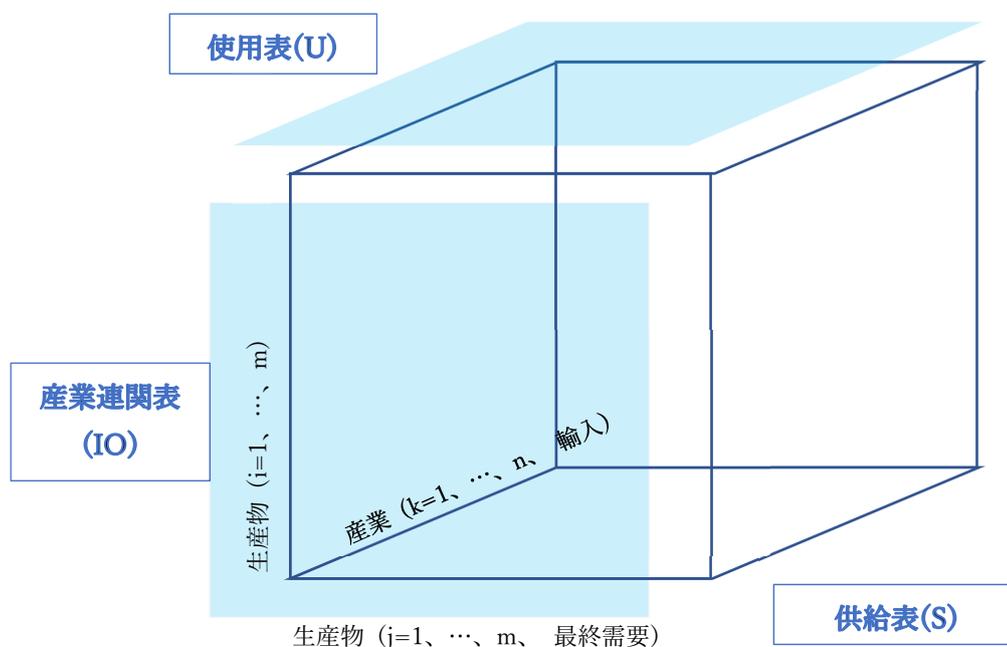


図 2 オランダ統計局の DPI キューブを日本向けに改変
(出所) 筆者作成。

2.3 米国センサス局の欠測値補完シミュレーション

実際の経済センサス-活動調査のデータを用いて IOT 及び SUT を推計した場合、計算が偶然に成功することがありうる。本当は調査によってデータを複数回収集し、それに基づいて IOT 及び SUT を推計しないと、推計の妥当性の検証はできない。だが、経済センサス-活動調査は 5 年毎の調査であり、それに基づいて IOT 及び SUT を推計することができるのは、10 年間で 2 回、15 年間で 3 回となる。あまりにも時間がかかりすぎる。そこで本稿では 100 個の擬似マイクロデータを作成し、それぞれについて IOT 及び SUT を推計するというシミュレーションを実施することを考えた。100 回であるから 5 年毎の経済センサス-活動調査なら 500 年分のデータである。

このアイデアは、米国センサス局が 2017 年米国経済センサスにおける生産物に関する調査事項の補定に関して実施した調査研究から来ている。その考え方は、次の通りである。まず 4 種類の補定手法 (①比率補定、②ランダム・ホットデック(Hot Deck Random, HDR)、③最近隣ホットデック (Hot Deck Nearest Neighbor, HDN)、④逐次回帰多重代入法 (Sequential

Regression Multivariate Imputation、RMI) のそれぞれを用いて、欠測値 (missing data) があるオリジナル・ファイル (Original File) から、(欠測値がない) 母集団データ (“POPULATION” no missing data) を 4 種類の補定手法をそれぞれについて作成する。次に、その母集団データに欠測値 (Induce Nonresponse) を 50 回発生させる。さらにそれぞれの (欠測値がある) 母集団 (“POPULATION” missing data) について 4 種類の補定手法を適用し、それぞれ 100 回ずつ補定 (Impute and Multiple impute) する。これによって、それぞれの補定手法の精度の評価を行う (図 3)。考えられる可能性をすべて試みて、その中から最良のものを選択するという考え方である。欠測値補完のシミュレーションは、欠測値のない、あるいは少ない観測データを用いて実施されることがあるが、その結果から欠測値が多いデータに適用しても、その補完方法が依然として有効であるかはわからない。欠測値のない母集団データを作成し、それを用いてシミュレーションを実施する、それも考えられる有力な手法を網羅して実施するというアイデアはその弱点を補っている。

本稿で作成する擬似マイクロデータは、欠測値がないデータである。そこで欠測値がないデータの一部の情報を隠して欠測値としたうえで、様々な欠測値補完の方法を適用して完全データを再推計し、元の欠測値がないデータと比較するシミュレーションを行えば、欠測値補完方法の精度比較・評価を行うことも可能となる。このように、本稿で作成した擬似マイクロデータについては、SUT 作成シミュレーション以外にも様々な活用方法を考えることができる。

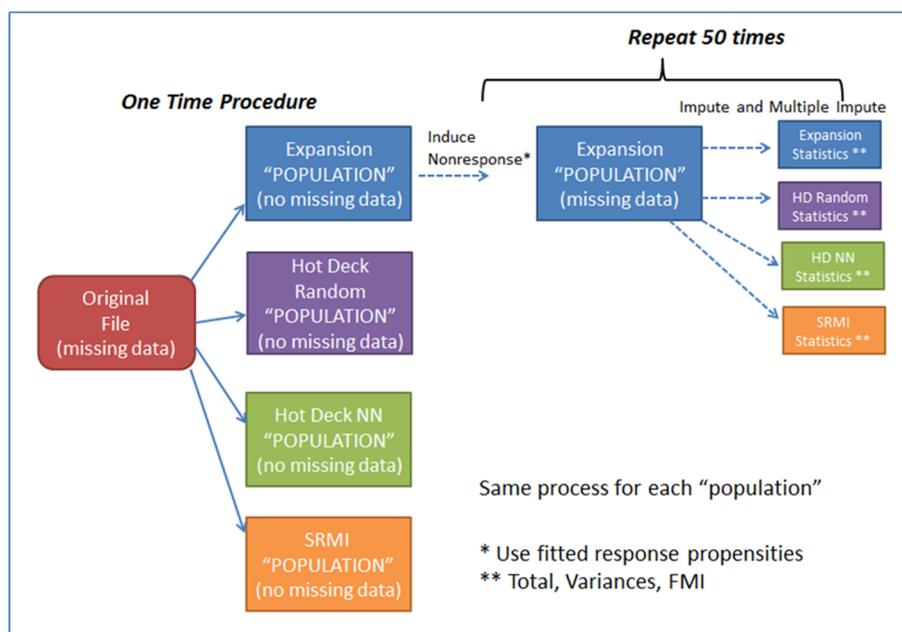


図 3 米国センサス局の欠測値補定シミュレーションのイメージ
 (出所) U.S. Census Bureau (2014)

3. 擬似マイクロデータの作成方針および SUT・IOT の推計手順

先行研究のサーベイに基づき、以下のような(1)~(5)の作成手順によって擬似マイクロデータを作成した。

(1) 経済センサス-活動調査の集計表および個票データより得られる分布の特性値に基づき、上額に関する企業単位の擬似マイクロデータを 100 セット作成する。

(2) 尾崎型生産関数の推定結果を用いて、(1)で作成した売上(収入)金額から従業者数を企業毎に推定する。

(3) 企業単位の擬似的な売上(収入)金額を、さらに事業活動別に細分化する。

(4) 平成 27 年産業連関表の部門別中間投入比率を企業毎の事業活動別売上(収入)金額にあてはめて中間投入額を推計する。

(5) 商業企業における商品販売額とマージン額の関係を表す関数の推定結果を用いて商品販売額擬似マイクロデータからマージン額擬似マイクロデータの作成を試みる。

さらに次の(6)に基づいて分析を実施する。

(6) 標本データの一部の情報を隠した上で 4 種類の手法で SUT から IOT を推計し、(完全データから作成した) 真の IOT と比較する。

4. 集計表から擬似マイクロデータ (母集団データ) を作成

4.1 売上(収入)金額

産業別売上(収入)金額の規模別分布を当初は対数正規分布で近似することを試みた。だが、試行錯誤の結果、対数正規分布以外の分布で近似するべきであるという結論に至った。最終的に、Fleishman(1978)による変換 (以下「Fleishman 変換」と呼ぶ) を応用して非正規分布に従う乱数を作成した。Fleishman 変換とは、平均、標準偏差、尖度、歪度が既知の場合に、標準正規分布に従う乱数 X から非正規分布に従う乱数 Y を以下の数式で作成する手法である。

$$Y = a + bX + cX^2 + dX^3$$

Fleishman 変換に用いた係数は表 1 の通りである。これらは経済センサス-活動調査による公表集計表の売上(収入)金額の自然対数値の平均、標準偏差、尖度、歪度から推定した。

表1 Fleishman 変換に用いた係数

1.5桁産業分類	a	b	c	d
AB 農業, 林業, 漁業	8.51785	1.49968	-0.00893	0.03004
C 鉱業, 採石業, 砂利採取業	9.25092	1.62653	0.05502	0.07904
D 建設業	8.23403	1.61304	0.07739	0.01370
E 製造業	8.09867	2.06860	0.23571	0.00413
F 電気・ガス・熱供給・水道業	10.34440	2.95016	0.16851	0.00780
G1 情報通信業 (通信業, 放送業, 映像・音声・文字情報制作業)	8.39385	2.06813	0.26027	0.03578
G2 情報通信業 (情報サービス業, インターネット附随サービス業)	8.59234	2.12849	0.13794	0.00957
H 運輸業, 郵便業	8.88107	2.24231	0.02347	-0.02552
I1 卸売業	9.12440	1.99946	0.14805	0.02100
I2 小売業	7.41865	1.75413	0.15230	0.01181
J 金融業, 保険業	7.33667	1.94708	0.53318	0.05176
K1 不動産業	6.75182	1.46392	0.19712	0.02790
K2 物品賃貸業	8.20747	2.19680	0.22767	-0.00186
L 学術研究, 専門・技術サービス業	7.42134	1.36329	0.13579	0.06268
M1 宿泊業	7.35843	1.78763	0.18945	-0.00562
M2 飲食店, 持ち帰り・配達飲食サービス業	6.65440	1.09614	0.14387	0.04666
N 生活関連サービス業, 娯楽業	5.81520	1.31302	0.43039	0.02463
O1 教育, 学習支援業 (学校教育)	9.66752	1.38168	0.11660	0.07505
O2 教育, 学習支援業 (その他の教育, 学習支援業)	5.39092	1.51424	0.43295	-0.04045
P 医療, 福祉	8.11370	1.61505	0.02486	0.03998
R2 サービス業 (政治・経済・文化団体, 宗教を除く)	7.91895	1.75379	0.13593	0.00850

注)「Q 複合サービス事業」と「R1 サービス業 (政治・経済・文化団体, 宗教)」は収束しないため係数を求められなかった。

(出所) 筆者作成。

Fleishman 変換により売上(収入)金額の擬似マイクロデータを 100 セット作成した。その企業 1 社あたり売上(収入)金額の 1.5 桁産業分類別平均値は表 2 の通りである。ちなみに、経済センサス-活動調査の売上(収入)金額は消費税に関して「税込み」であることから、生産者価格表示である。参考までに企業 1 社あたり売上(収入)金額の 1.5 桁産業分類別実績値、そして擬似マイクロデータの平均値と実績値の乖離及び比率を示した。擬似マイクロデータの平均値と実績値の比率は、21 産業中、15 産業が 0.5 から 2 の範囲にある。なお、ここでの擬似マイクロデータは、経済センサス-活動調査における企業の売上規模別の集計表に基づいて分布の特性値を求め、それらの特性値をもとに Fleishman 変換により擬似マイクロデータを作成したものである。従って、データ作成に用いた分布の特性値が実際の特性値(個票データに基づいて計算される特性値)に完全には一致しないうえに、規模別分布の形状が特異な産業については、作成された擬似マイクロデータによって実際の規模別分布(公表されている集計表)を正確に再現できていないケースもある。このような産業については、表 2 における比率が 1 から大きく乖離することになる。産業別の詳細な結果については補

論を参照のこと。ただし乖離していてもそもそも擬似マイクロデータであるので実績値と一致する必要はない。

表2 企業1社あたり売上(収入)金額の1.5桁産業分類別平均値
(単位：生産者価格表示百万円、擬似マイクロデータ100セットの平均値)

1.5桁産業分類	①擬似マイクロ	②実績値	乖離 (=①-②)	比率 (=①/②)
AB 農業、林業、漁業	5203263	4991597	211666	1.04
C 鉱業、採石業、砂利採取業	4067564	2044079	2023485	1.99
D 建設業	98947183	108450918	-9503735	0.91
E 製造業	736646033	396275421	340370612	1.86
F 電気・ガス・熱供給・水道業	54428879	26242446	28186433	2.07
G1 情報通信業（通信業、放送業、映像・音声・文字情報制作業）	90724170	31207946	59516224	2.91
G2 情報通信業（情報サービス業、インターネット附随サービス業）	44138915	28737690	15401225	1.54
H 運輸業、郵便業	44561557	64790606	-20229049	0.69
I1 卸売業	637172227	345578658	291593569	1.84
I2 小売業	152933505	155215598	-2282093	0.99
J 金融業、保険業	420332860	125130273	295202587	3.36
K1 不動産業	43667307	31809378	11857929	1.37
K2 物品賃貸業	27780189	14245933	13534256	1.95
L 学術研究、専門・技術サービス業	83800370	41501702	42298668	2.02
M1 宿泊業	7039412	6132951	906461	1.15
M2 飲食店、持ち帰り・配達飲食サービス業	28803888	19348540	9455348	1.49
N 生活関連サービス業、娯楽業	179433647	45661141	133772506	3.93
O1 教育、学習支援業（学校教育）	30695468	11784701	18910767	2.60
O2 教育、学習支援業（その他の教育、学習支援業）	3989613	3625355	364258	1.10
P 医療、福祉	77836197	111487956	-33651759	0.70
R2 サービス業（政治・経済・文化団体、宗教を除く）	41339347	35905017	5434330	1.15
合計	2813541594	1610167906	1203373688	1.75

(出所) 筆者作成。

4.2 従業者数

統計調査では一般に調査事項の数と回収率はトレードオフの関係にある。SUT作成に関連しては、詳細な投入及び産出の情報があることが望ましいが、一方で記入負担から未回収が増える懸念がある。経済センサス-活動調査では売上(収入)金額に欠測値があった場合、もっぱら従業者数から補完している。したがって、本研究における擬似マイクロデータには従業者数が必要である。本研究では宮内(2023)による制約付き尾崎型生産関数の推定結果(表3)を用いて、売上(収入)金額から従業者数を企業毎に推定した。

表3 制約付き尾崎型生産関数の推定結果

1.5桁産業分類	α_{ik} 平均値	β_{ik} 平均値	σ_i
AB農業, 林業, 漁業	2.9947	0.362	0.398
C鉱業, 採石業, 砂利採取業	0.8605	0.4972	0.2403
D建設業	0.6408	0.5752	0.2624
E製造業	0.7292	0.6177	0.2288
F電気・ガス・熱供給・水道業	0.5766	0.6366	0.3262
G1情報通信業 (通信業, 放送業, 映像・音声・文字情報制作業)	0.6305	0.6216	0.2326
G2情報通信業 (情報サービス業, インターネット附随サービス業)	0.5048	0.7207	0.2633
H運輸業, 郵便業	0.5205	0.7548	0.4266
I1卸売業	0.432	0.6047	0.2142
I2小売業	0.5777	0.699	0.1638
J金融業, 保険業	0.6337	0.6451	0.2045
K1不動産業	0.8252	0.5211	0.201
K2物品賃貸業	0.6991	0.6107	0.3204
L学術研究, 専門・技術サービス業	0.5441	0.6831	0.1416
M1宿泊業	0.8575	0.7129	0.2413
M2飲食店, 持ち帰り・配達飲食サービス業	0.6075	0.843	0.1175
N生活関連サービス業, 娯楽業	0.8777	0.6264	0.1216
O1教育, 学習支援業 (学校教育)	0.8864	0.7158	0.2919
O2教育, 学習支援業 (その他の教育, 学習支援業)	1.189	0.7228	0.1339
P医療, 福祉	0.6449	0.7596	0.2419
R2サービス業 (政治・経済・文化団体, 宗教を除く)	0.4755	0.8184	0.1894

(出所) 筆者作成。

4.3 事業活動別売上(収入)金額

供給表を作成するためには事業活動別売上(収入)金額が必要である。そこで4.1で作成した企業単位の擬似的な売上(収入)金額を、さらに事業活動別に細分化する。集計表から擬似マイクロデータを作成するという原則に従えば、統計局から公表された企業産業別事業活動(22区分)別売上(収入)金額の集計表を用いて、企業単位の擬似的な売上(収入)金額を事業活動別に細分化することがまず考えられる。だがこの集計表を用いた場合、同じ産業に格付けされた企業は、どれも同じ構成比で事業活動を行うことになってしまう。実際には同じ産業に格付けされた企業でも、事業活動の構成比は異なっている。リアリティのあるシミュレーションを実施するためには、企業産業別事業活動パターン別売上(収入)金額の集計表が必要である。本研究では、菅(2022)による「平成28年経済センサス-活動調査」による企業の副業パターンの集計表を用いた。この集計表は、統計センターのオンサイト施設において平成28年経済センサス-活動調査の個票データから集計を行ったものである。特別に集計されたデータではあるものの、集計表から擬似マイクロデータを作成するという原則は守っている。表4は1.5桁産業分類別事業活動別売上(収入)金額(擬似マイクロデータ100セットの平均値)を示している。表側が1.5桁産業分類、表頭が事業活動(22区分)になっている。

表4 1.5 桁産業分類別事業活動別売上(収入)金額
(単位：生産者価格表示百万円、擬似マイクロデータ 100 セットの平均値)

事業活動 (22 区分)	農業, 林業, 漁業	鉱物, 採石, 砂利採取事業	建設事業	製造業	電気, ガス, 熱供給, 水道事業
1.5 桁産業分類					
AB 農業, 林業, 漁業	5009744	668	182	55293	1237
C 鉱業, 採石業, 砂利採取業	1280	3908197	65191	9039	522
D 建設業	105180	13833	97592378	120432	114573
E 製造業	578361	59353	2023332	715648179	109607
F 電気・ガス・熱供給・水道業	0	0	400124	0	53232679
G1 情報通信業 (通信業, 放送業, 映像・音声・文字情報制作業)	4488	0	23626	84809	1180
G2 情報通信業 (情報サービス業, インターネット附随サービス業)	98	0	35492	107032	1607
H 運輸業, 郵便業	26010	12271	68268	66261	5465
I1 卸売業	248863	25694	2164566	4391059	180531
I2 小売業	23328	536	341396	131689	83572
J 金融業, 保険業	13704	0	318323	32488	20525
K1 不動産業	38357	518	118098	13934	18206
K2 物品賃貸業	26219	0	440437	52473	2838
L 学術研究, 専門・技術サービス業	9585	170	240178	54619	12859
M1 宿泊業	21641	0	6177	1795	574
M2 飲食店, 持ち帰り・配達飲食サービス業	7689	0	8894	3326	380
N 生活関連サービス業, 娯楽業	41583	0	19257	10457	646
O1 教育, 学習支援業 (学校教育)	0	0	0	0	66
O2 教育, 学習支援業 (その他の教育, 学習支援業)	435	0	514	1071	97
P 医療, 福祉	4327	0	13314	18217	1737
R2 サービス業 (政治・経済・文化団体, 宗教を除く)	13264	1189	238361	115495	12064
合計	6174156	4022429	104118108	720917668	53800965

(出所) 筆者作成。

表4 1.5 桁産業分類別事業活動別売上(収入)金額 (つづき)
(単位：生産者価格表示百万円、擬似マイクロデータ 100 セットの平均値)

	情報通信事業 通信, 放送, 映像・音声・文字情報制作事業	情報通信事業 情報サービス, インターネット付随サービス事業	運輸, 郵便事業	商業 卸売業	商業 小売業	金融, 保険事業	不動産・物品賃貸事業 不動産事業	不動産・物品賃貸事業 物品賃貸事業
AB	25143	135	1289	36623	26785	0	7577	2229
C	0	0	29496	6047	2458	0	1518	853
D	13011	13728	56102	133812	231077	1294	271930	17853
E	60531	124037	314806	11245103	3609303	4510	242752	58321
F	0	0	0	195082	317925	0	140886	673
G1	89025759	146160	936	143970	372560	2843	56641	16472
G2	41565	42763696	714	280190	268120	30476	56076	14549
H	1220	2889	43923441	100751	49590	2977	107481	13669
I1	58716	268740	257057	601421926	20625422	93574	389959	463992
I2	19468	34296	53266	1703354	147606083	45172	127408	61138
J	297	340749	59468	2100003	3411994	410159720	752797	1108338
K1	1117	8048	8095	38044	84504	25883	43067712	23611
K2	3931	5037	39310	443840	809974	52367	153671	25150611
L	28299	146589	1442	48004	214825	35722	194267	14291
M1	0	360	1400	3426	30156	728	58731	1331
M2	1513	1703	898	7999	61908	356	89738	1222
N	3954	3459	13666	54986	732930	342	96697	20558
O1	0	0	0	924	4693	387	16861	0
O2	174	3046	97	659	15762	120	4366	254
P	360	2908	15255	9762	92399	559	38933	1308
R2	5810	59984	49812	131108	428392	27538	99955	48790
合計	89290868	43925564	44826550	618105613	178996860	410484568	45975956	27020063

(出所) 筆者作成。

表4 1.5 桁産業分類別事業活動別売上(収入)金額 (つづき)
(単位：生産者価格表示百万円、擬似マイクロデータ 100 セットの平均値)

	学術研究、 専門・技術 サービス事 業	宿泊・飲食 サービス事 業 宿泊事 業	宿泊・飲食 サービス事 業 飲食 サービス事 業	生活関連サー ビス、娯楽事 業	学校教育・ 学習支援事 業 学校教 育事業	学校教育・ 学習支援事 業 社会教 育、学習支 援事業	医療、福祉 事業	サービス事 業（他に分 類されない もの）政 治・経済・ 文化団体	サービス事 業（他に分 類されない もの）そ の他のサー ビス事業	合計
AB	1870	934	5974	3326	45	324	6418	2480	14987	5203263
C	0	31016	0	0	0	0	0	0	11947	4067564
D	39603	2776	14508	9753	554	1688	9441	299	183359	98947183
E	149050	6552	177860	44827	19051	17893	35249	7342	2110014	736646033
F	67607	0	0	0	0	0	0	0	73904	54428879
G1	480279	0	59426	10687	1374	17954	4735	15714	254558	90724170
G2	131845	0	13032	15876	310	13147	9031	0	356058	44138915
H	3183	3781	20192	14920	271	2808	9473	535	126098	44561557
I1	211585	7336	330399	100035	19675	40749	81725	22364	5768260	637172227
I2	44840	3719	150144	233614	11753	67170	46928	2501	2142130	152933505
J	226526	254	61978	147501	2158	8927	422396	24252	1120461	420332860
K1	49013	8696	34108	23230	28	3576	10313	2191	90026	43667307
K2	175863	961	57250	71116	0	295	76034	586	217375	27780189
L	82303148	92	8010	60841	1084	10219	260529	58269	97329	83800370
M1	659	6498850	330772	54842	49	2707	2373	0	22842	7039412
M2	2814	4621	28536847	62129	134	1924	1404	40	8349	28803888
N	23986	10373	60889	178269988	277	30066	4152	276	35105	179433647
O1	9590	2249	257	774	30008534	11486	626004	7500	6141	30695468
O2	5676	329	1998	9564	1121	3938928	719	467	4216	3989613
P	9277	366	10743	9769	15033	8395	77527531	4353	51649	77836197
R2	67430	1493	23585	36711	253	12730	5776	8161	39951447	41339347
合計	84003844	6584398	29897972	179179503	30081704	4190986	79140231	157330	52646255	2813541594

(出所) 筆者作成。

(注) 単位未満で四捨五入しているため、合計と内訳の計は必ずしも一致しない。

4.4 中間投入額

中間投入額に関しては、経済センサス-活動調査に詳しい情報がなく、実際の産業連関表の推計においては、産業連関構造調査（所謂「投入調査」）の結果を用いている。ただし、同調査の結果をそのまま用いているケースもあれば、別途、行政情報やヒアリング等などから得た情報を用いて作成しているケースもある。ここでは平成 27 年産業連関表の部門別中間投入比率を企業の事業活動別売上(収入)金額にあてはめて作成した(表 5)。その際にノイズ(平均 1、標準偏差 0.5 の正規確率乱数)を乗じて中間投入係数にばらつきが発生するように設定した。なお正規確率乱数はボックス=ミュラー法(一様分布に従う確率変数から標準正規分布に従う確率変数を作成させる手法)により作成した。

表5 生産物別 1.5 桁産業分類別中間投入額

(単位：購入者価格表示百万円、擬似マイクロデータ 100 セットの平均値)

生産物	1.5 桁産業分類	AB 農業、林業、漁業	C 鉱業、採石業、砂利採取業	D 建設業	E 製造業	F 電気・ガス・熱供給・水道業
農業、林業、漁業		665410	1432	232274	23004248	1083
鉱物、採石、砂利採取事業		2989	11049	979925	32858195	13669157
建設事業		12112	12053	66490	1424579	797464
製造業		1563636	425807	34706381	354084212	4082925
電気、ガス、熱供給、水道事業		54758	177806	597615	16541473	5112655
情報通信事業 通信、放送、映像・音声・文字情報制作事業		9700	21743	488159	870487	119798
情報通信事業 情報サービス、インターネット附随サービス事業		20791	18015	507987	4430615	777534
運輸、郵便事業		153391	808968	2164764	7513489	754348
商業 卸売業		24	3	139	5505	209
商業 小売業		0	0	0	0	0
金融、保険事業		33305	145519	1248548	4747980	1066659
不動産・物品賃貸事業 不動産事業		10560	34867	457783	1523963	250675
不動産・物品賃貸事業 物品賃貸事業		17877	65753	2550225	3099031	237880
学術研究、専門・技術サービス事業		0	0	0	0	0
宿泊・飲食サービス事業 宿泊事業		0	0	0	0	0
宿泊・飲食サービス事業 飲食サービス事業		71	550	172	1440	0
生活関連サービス、娯楽事業		1578	1350	26606	94874	5883
学校教育・学習支援事業 学校教育事業		0	0	0	0	0
学校教育・学習支援事業 社会教育、学習支援事業		342	1105	15628	236749	22833
医療、福祉事業		2037	31	447	3320	3159
サービス事業（他に分類されないもの） 政治・経済・文化団体		5630	11686	108269	610900	140944
サービス事業（他に分類されないもの） その他のサービス事業		114163	189231	8088566	22507988	4149321
中間投入合計		2668374	1926968	52239978	473559048	31192527
(参考) 売上(収入)金額		5203263	4067564	98947183	736646033	54428879

(出所) 筆者作成。

表5 生産物別 1.5 桁産業分類別中間投入額（つづき）
 （単位：購入者価格表示百万円、擬似マイクロデータ 100 セットの平均値）

生産物	1.5 桁産業分類	G1 情報通信業（通信業、放送業、映像・音声・文字情報制作業）	G2 情報通信業（情報サービス業、インターネット附随サービス業）	H 運輸業、郵便業	I1 卸売業	I2 小売業
農業、林業、漁業		9627	4676	10248	236268	27704
鉱物、採石、砂利採取事業		4403	5376	5509	258450	32834
建設事業		490103	77744	293800	371343	89708
製造業		2036318	4049512	7159119	8602037	1698822
電気、ガス、熱供給、水道事業		1794039	203696	953112	3309162	804688
情報通信事業 通信、放送、映像・音声・文字情報制作事業		13954732	1490074	141347	1921293	494137
情報通信事業 情報サービス、インターネット附随サービス事業		8220588	3426057	371500	3443560	865583
運輸、郵便事業		1952916	1033791	5118123	6240943	1505690
商業 卸売業		224	200	57	234377	56418
商業 小売業		0	0	0	0	0
金融、保険事業		648226	186954	1040981	2171004	521290
不動産・物品賃貸事業 不動産事業		915019	1421368	858937	3405115	821602
不動産・物品賃貸事業 物品賃貸事業		1034639	623072	1458824	1178587	283927
学術研究、専門・技術サービス事業		0	0	0	0	0
宿泊・飲食サービス事業 宿泊事業		0	0	0	0	0
宿泊・飲食サービス事業 飲食サービス事業		304	140	227	3002	1444
生活関連サービス、娯楽事業		857624	445164	30788	118376	33869
学校教育・学習支援事業 学校教育事業		0	0	0	0	0
学校教育・学習支援事業 社会教育、学習支援事業		370842	162849	53885	31522	7595
医療、福祉事業		93581	7008	51826	6138	2028
サービス事業（他に分類されないもの） 政治・経済・文化団体		116643	55054	56256	90737	23168
サービス事業（他に分類されないもの） その他のサービス事業		10552219	7596371	4210396	11288655	2764274
中間投入合計		43052047	20789106	21814935	42910569	10034781
（参考）売上（収入）金額		90724170	44138915	44561557	637172227	152933505

（出所）筆者作成。

表5 生産物別 1.5 桁産業分類別中間投入額（つづき）
（単位：購入者価格表示百万円、擬似マイクロデータ 100 セットの平均値）

生産物	J 金融業、保 険業	K1 不動産業	K2 物品賃貸 業	L 学術研究、 専門・技術 サービス業	M1 宿泊業
農業、林業、漁業	12097	8588	16697	171019	227824
鉱物、採石、砂利採取事業	8158	6256	9677	16790	626
建設事業	298628	384931	44583	313987	12752
製造業	4276863	239400	1255495	9634464	1091223
電気、ガス、熱供給、水道事業	1169100	201849	115527	1844317	777937
情報通信事業 通信、放送、映像・音声・文字情報制作事業	2318807	87852	59288	1569261	59595
情報通信事業 情報サービス、インターネット附随サービス事業	5562849	92055	414112	4144192	131728
運輸、郵便事業	3300807	108162	506170	1887644	655775
商業 卸売業	2163	49	501	78	14
商業 小売業	0	0	0	0	0
金融、保険事業	5499519	3365985	976848	216130	136077
不動産・物品賃貸事業 不動産事業	1904309	1281560	336295	1384074	75455
不動産・物品賃貸事業 物品賃貸事業	1085983	31227	92506	361702	56317
学術研究、専門・技術サービス事業	0	0	0	0	0
宿泊・飲食サービス事業 宿泊事業	0	0	0	0	0
宿泊・飲食サービス事業 飲食サービス事業	1324	425	535	961	97998
生活関連サービス、娯楽事業	39943	24153	30482	215929	107329
学校教育・学習支援事業 学校教育事業	0	0	0	0	0
学校教育・学習支援事業 社会教育、学習支援事業	28506	223	10633	840	1675
医療、福祉事業	25437	508	966	8677	117
サービス事業（他に分類されないもの） 政治・経済・文化団体	353823	14768	56775	356907	10540
サービス事業（他に分類されないもの） その他のサービス事業	14120300	1313802	4775614	10013537	157193
中間投入合計	40008616	7161793	8702704	32140509	3600175
（参考）売上（収入）金額	420332860	43667307	27780189	83800370	7039412

（出所）筆者作成。

表5 生産物別 1.5 桁産業分類別中間投入額（つづき）
（単位：購入者価格表示百万円、擬似マイクロデータ 100 セットの平均値）

生産物	1.5 桁産業分類	M2 飲食店、 持ち帰り・ 配達飲食 サービス業	N 生活関連 サービス業、 娯楽業	O1 教育、学 習支援業 (学校教 育)	O2 教育、学 習支援業 (その他の 教育、学習 支援業)	P 医療、福祉	R2 サービス業 (政治・経 済・文化団 体、宗教を除 く)
農業、林業、漁業		1933363	1159795	120849	15916	388227	7998
鉱物、採石、砂利採取事業		263	17677	190	97	2207	11322
建設事業		36881	578560	190619	25133	175646	43513
製造業		11060969	14352305	1123571	132144	16982853	5741233
電気、ガス、熱供給、水道事業		1739605	8911212	1217331	159397	1640905	320404
情報通信事業 通信、放送、映像・音声・文字情報制作事業		378933	1283707	50207	6837	311510	1527242
情報通信事業 情報サービス、インターネット付随サービス事業		140623	3501831	227976	29484	839635	1925164
運輸、郵便事業		239355	5718413	528671	72313	553784	667888
商業 卸売業		28	292	2	6	39	210
商業 小売業		0	0	0	0	0	0
金融、保険事業		171165	1103065	334352	44731	688206	182621
不動産・物品賃貸事業 不動産事業		283212	3016690	51489	5715	1239303	352754
不動産・物品賃貸事業 物品賃貸事業		47064	734309	64949	7911	687700	347890
学術研究、専門・技術サービス事業		0	0	0	0	0	0
宿泊・飲食サービス事業 宿泊事業		0	0	0	0	0	0
宿泊・飲食サービス事業 飲食サービス事業		198561	657	80934	10566	165395	233
生活関連サービス、娯楽事業		102184	3353234	25412	2524	791797	60295
学校教育・学習支援事業 学校教育事業		0	0	0	0	0	0
学校教育・学習支援事業 社会教育、学習支援事業		12653	88043	109	28	7108	20369
医療、福祉事業		2950	12868	11708	118	1225442	8363
サービス事業（他に分類されないもの） 政治・経済・文化団体		44819	816546	11147	1470	76619	86700
サービス事業（他に分類されないもの） その他のサービス事業		928706	8301817	1597410	205481	3510469	4789519
中間投入合計		17321334	52951021	5636926	719871	29286845	16093718
(参考) 売上(収入) 金額		28803888	179433647	30695468	3989613	77836197	41339347

(出所) 筆者作成。

4.5 商業マージン率・国内運賃率

供給表は生産者価格表示であり、使用表は購入者価格表示である。生産者価格に商業マージン、国内運賃を足したものが購入者価格である。ここでの目的は、4.3 で求めた事業活動別売上(収入)金額のうち、卸売および小売による商品販売額擬似マイクロデータ（企業別）より、それぞれに対応するマージン額の擬似マイクロデータを作成することである。ただし商業マージン額については、経済センサス-活動調査などの公表されている統計データから入手できる情報が限られているため、本研究では、「平成 28 年経済センサス-活動調査」の個票データを用いて商業企業における商品販売額とマージン額の間接関数を推定し、その推定結果を用いて商品販売額擬似マイクロデータからマージン額擬似マイクロデータの作成を試みている。

商品販売額とマージン額については、(1)式のようなモデルを想定している。

$$M_i = \alpha X_i^\beta \quad (1)$$

X_i は*i*番目の企業の商品販売額であり、 M_i はそれに対応するマージン額を表す。4.3 で求められた事業活動別売上(収入)金額のうち、卸売および小売の売上額が X_i に該当するものである。 $\beta = 1$ であれば、 α はマージン率（販売額に占めるマージン額の割合 M_i/X_i ）を表し、販売額規模が異なってもマージン率は一定であることを意味する。 $\beta < 1$ であれば、商品販売額が大きくなるほどマージン率は低下し、 $\beta > 1$ であれば、商品販売額が大きくなるほどマージン率は上昇することになる。

本研究では、(1)式の両辺の対数をとった(2)式を OLS で推定している。 ε_i は誤差項である。

$$\ln M_i = \ln \alpha + \beta \ln X_i + \varepsilon_i \quad (2)$$

平成 28 年経済センサス-活動調査では、企業調査票において年間商品販売額とともに年間商品仕入額および年初及び年末商品手持額を調査しているため、年間商品販売額から年間商品仕入額を差し引き、商品手持額の純増分を加えることにより、各企業のマージン額 M_i を計算している。この際に、手持商品の評価額調整などは行っていない。なお、個人企業については仕入額が調査されておらず、また仕入額に関して未回答企業も存在するため、それらの企業を除いて推定を行っている。個人企業や仕入額未回答企業には小規模企業が多く含まれているため、推定に用いた企業の規模分布は、実際の企業の規模分布よりも大規模に偏っている点に注意が必要である。

また推定に際しては、卸売業と小売業を区別し、さらに農林水産業、鉱業、製造業も区分している（ただし鉱業製品は卸売のみ）。実際にはより詳細な産業分類ごとに推定を行うことも可能であるが、本研究における擬似マイクロデータでは 1.5 桁産業分類レベルの区分が行われているため、それに準じている。表 6 は、推定結果をまとめたものである。

表 6 推定結果

	農林水産 卸売	鉱業 卸売	製造業 卸売	農林水産 小売	製造業 小売
$\hat{\alpha}$	1.062	1.129	0.822	0.418	0.620
$\hat{\beta}$	0.829 (0.000)	0.815 (0.000)	0.873 (0.000)	0.970 (0.000)	0.930 (0.000)
<i>N of obs.</i>	15,724	2,017	102,194	6,794	158,214
<i>Adj.R</i> ²	0.795	0.791	0.842	0.864	0.842

(出所) 筆者作成。

$\hat{\beta}$ の推定値の下段にある括弧内の数値は、 $H_0: \beta = 1$ 、 $H_1: \beta < 1$ として t 検定を行った場合の p 値を表している。 $\hat{\beta}$ はいずれも有意に 1 を下回っており、販売額規模の拡大に従ってマージン率が低下する傾向にあることがわかる。自由度修正済み決定係数は最も低い鉱業製品の卸売業でも 0.791 であり、全てのモデルで高い値を示している。

この推定結果をもとに、年間商品販売額が 1 億円、10 億円、100 億円の企業におけるマージン率を推定した結果が以下である。

表7 マージン率推定結果

商品 販売額	農林水産 卸売	鉱業 卸売	製造業 卸売	農林水産 小売	製造業 小売
1 億円	0.220	0.205	0.255	0.318	0.325
10 億円	0.149	0.134	0.190	0.297	0.277
100 億円	0.100	0.087	0.142	0.277	0.235

(出所) 筆者作成。

マージン率の水準は卸売業よりも小売業の方が高い一方、販売規模拡大に伴うマージン率の低下幅は小売業の方が小さいことがわかる。

ここで得られたパラメータ推定値 $\hat{\alpha}$ 、 $\hat{\beta}$ 、第3節で作成した商品販売額擬似マイクロデータ、および誤差項 ε_i をもとに、(2)式に基づいて企業別マージン額の擬似マイクロデータを作成することができる。この際に、 ε_i については、個票データを用いて推定を行った際の実際の残差の分布を考慮して決定している。具体的には、4.1 で述べた Fleishman(1978)の手法を用いて、残差分布の平均値・標準偏差・歪度・尖度と同じ統計量を持つ乱数を発生させ、 ε_i の推定値としている。

4.6 経済センサス-活動調査の対象外の活動等の取り扱い

平成 28 年経済センサス-活動調査は、農林漁家、公的部門などの活動をカバーしていない。また、輸入（普通貿易）についてもカバーしていない。これらについては擬似マイクロデータの範囲外とし、擬似マイクロデータから SUT の主要部分を作成したのち、別途、平成 27 年産業連関表を参考に補うことにした。

5. おわりに

以上のようにして作成された擬似マイクロデータは SUT 作成の各種シミュレーションに用いられる予定である。その目的は SUT 及び IOT の作成担当者が作成方法を検討する際に、作成担当者及び統計関係者の共通の理解のベースを提供することである。すなわち、各種の SUT から IOT を推計する手法を検討するシミュレーションをそれぞれ 100 回実施し、真の値との標準誤差率を求め、精度を評価することである。

また、このようにして作成された擬似マイクロデータは SUT を推計するための教材として活用することも考えている。全国表について SUT が作成された後は、都道府県においても IOT から SUT への移行作業が、近い将来ではないとしても、いずれ始まるであろう。その際の都道府県の推計担当者のトレーニングの教材として活用することである。

参考文献

- [1] 伊藤伸介 (2020) 「諸外国における公的統計と行政記録データの二次利用に関する展開方向」, 『経済学論纂』, 61(2), pp.1-16
- [2] 河野真理子, 和田かず美 (2018) 「マイクロデータ分析のための演習用教材の作成方法 : 一般用マイクロデータ詳細品目版及び擬似マイクロデータによる事例」, 『統計研究彙報』, 総務省統計研究研修所編, 75, pp.61-80
- [3] 坂下信之 (2012) 「教育用擬似マイクロデータの作成と試行提供 (特集 公的統計基盤サービスとその課題)」, 『統計』, 統計編集委員会編, 63(3), pp.17-20
- [4] 産業連関表作成府省庁, 内閣府 (2019) 「基準年 S U T ・産業連関表及び中間年 S U T に係る基本構成の大枠」 (総務省ホームページ, アドレス : https://www.soumu.go.jp/main_content/000676754.pdf?msclkid=9b44ada4cf9911ecbc3cea7038d6580f)
- [5] 菅幹雄 (2022) 「事業別売上(収入)金額構成比のパターンに関する考察」『統計研究研修所リサーチペーパー』, 52
- [6] 宮内環 (2023) 「制約付き尾崎型生産関数の推定:SUT 作成シミュレーションモデル開発のための試算」『統計研究研修所リサーチペーパー』, 59
- [7] 山口幸三, 伊藤伸介, 秋山裕美 (2013) 「教育用擬似マイクロデータの作成 : 平成 16 年全国消費実態調査を例として」, 『統計学』, 104, pp.1-15
- [8] 横溝秀始, 伊藤伸介 (2022) 「事業所・企業系のマイクロデータにおける匿名化措置の有効性の評価－経済センサス－活動調査を例として－」, 『統計研究彙報』, 79, pp.151-170
- [9] 横溝秀始, 伊藤伸介 (2023) 「合成データの作成手法の有効性に関する定量的な評価 : 事業所・企業系のマイクロデータを用いて」『統計研究彙報』, 80, pp.97-116
- [10] Fleishman, A.J. (1978) “A method for simulating non-normal distributions.” *Psychometrica* 43,

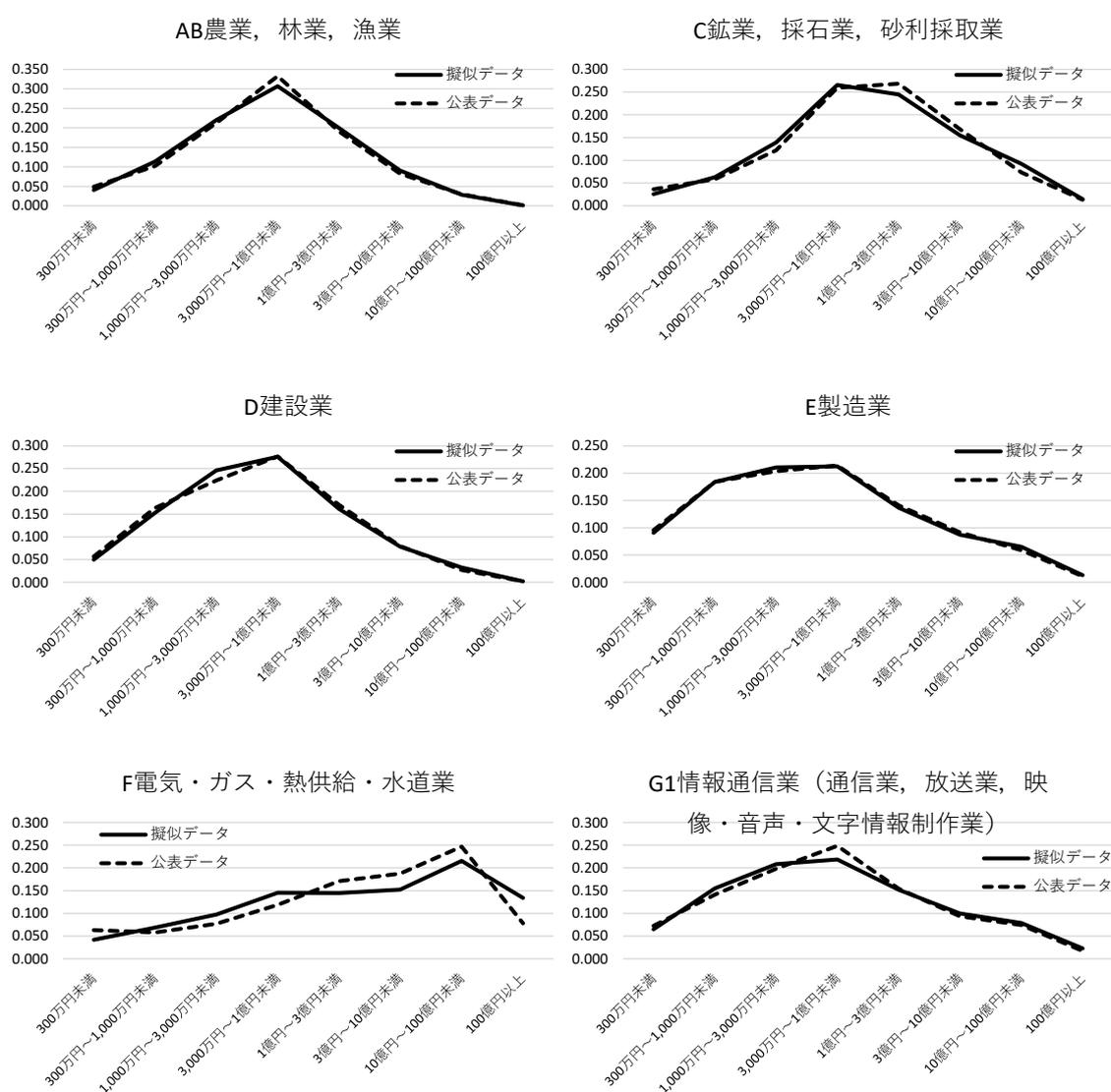
pp.521.-532

- [11] Kazemier, B. C. H. Driesen & E. Hoogbruin(1991)” From Input--Output Tables To Supply-And-Use Tables, ” *Economic Systems Research*, 2012, 24(3), pp. 319–327
- [12] Rueda-Cantuche, José M., Joerg Beutel , Frederik Neuwahl, Ignazio Mongelli & Andreas Loeschel (2009) “A Symmetric Input-output Table for EU27: Latest Progress, ” *Economic Systems Research*, . 21(1), March, pp. 59–79
- [13] U.S. Census Bureau (2014) , *Recommendation for Product Line Imputation for 2017 Economic Census*
- [14] Vilhubera, Lars, John M. Abowda and Jerome P. Reite (2016)”Synthetic establishment microdata around the world”, *Statistical Journal of the IAOS*, 32, pp.65–68

補論：擬似マイクロデータと公表データの規模分布比較

第4節で述べたように、本稿では、売上(収入)金額の擬似マイクロデータを作成するにあたって、「平成28年経済センサス-活動調査」の公表データより得られる規模別の企業数データに基づいて計算される平均、標準偏差、尖度、歪度を用いて、Fleishman変換によって擬似マイクロデータにおける売上(収入)金額を求めている。

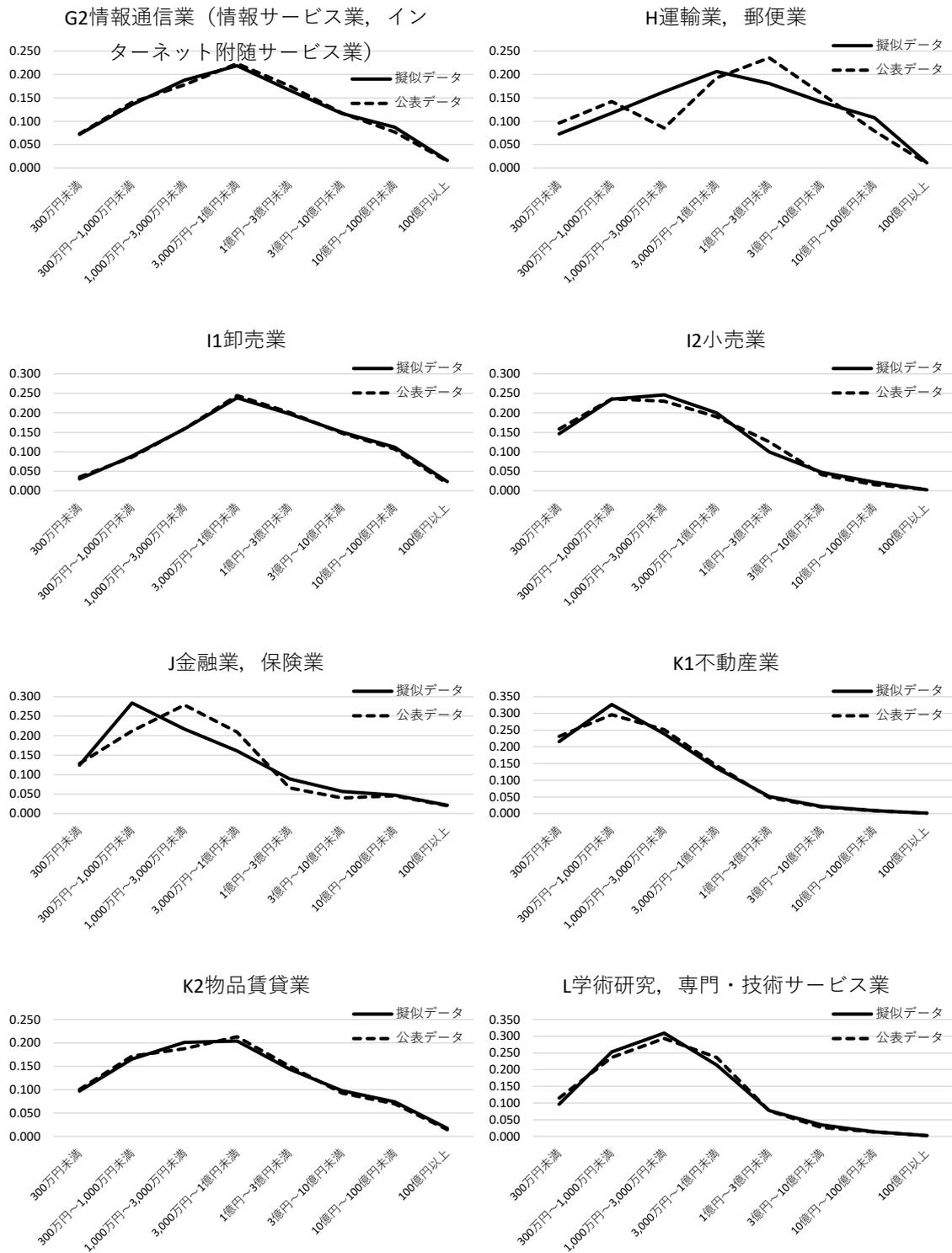
図4は、作成した擬似マイクロデータと公表データの売上(収入)金額規模別の企業数分布を1.5桁産業分類ごとに比較したものである。



※図中の「擬似データ」は「擬似マイクロデータ」を指す。

図4 擬似マイクロデータと公表データの規模分布比較

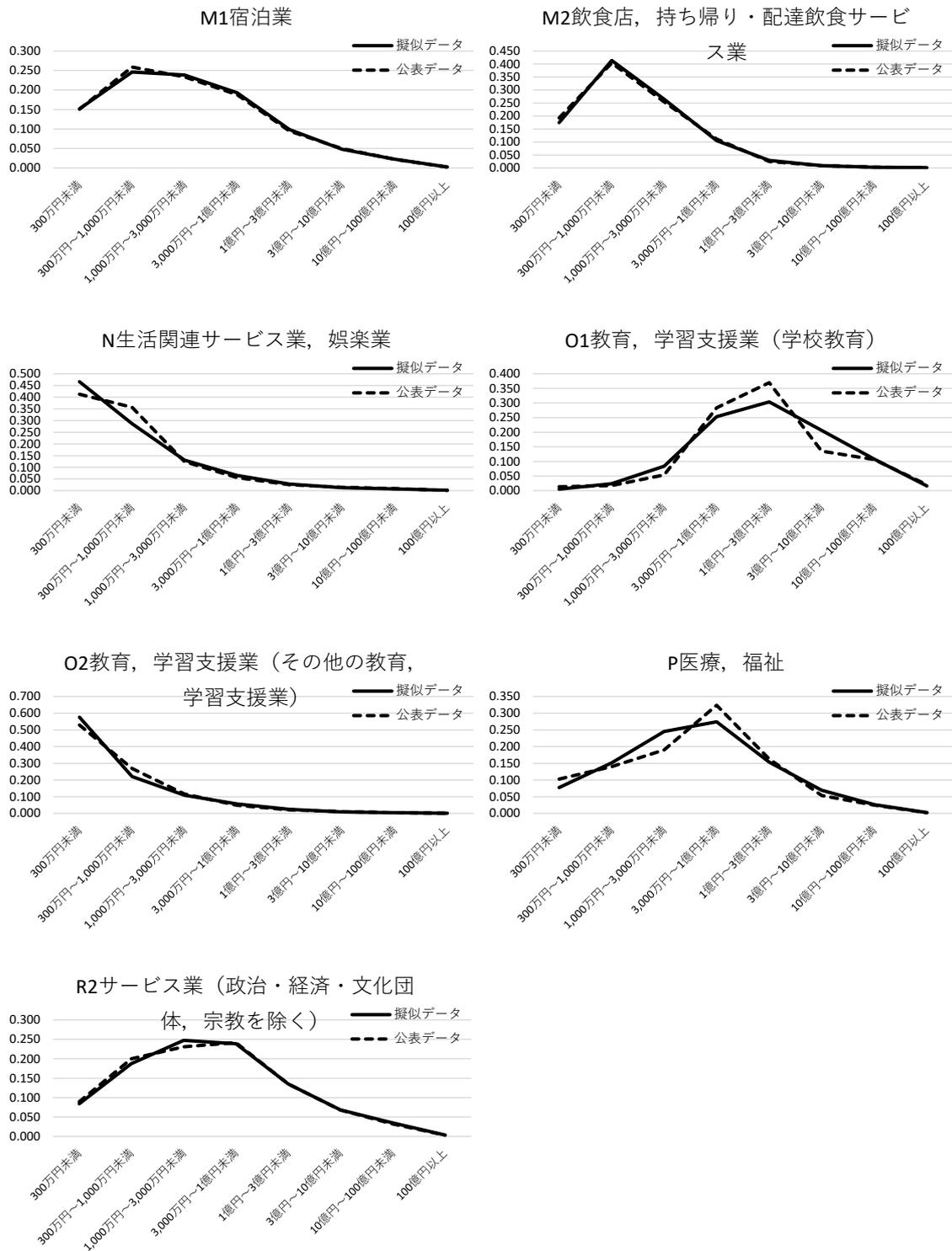
(出所) 筆者作成。



※図中の「擬似データ」は「擬似マイクロデータ」を指す。

図4 擬似マイクロデータと公表データの規模分布比較 (つづき)

(出所) 筆者作成。



※図中の「擬似データ」は「擬似マイクロデータ」を指す。

図4 擬似マイクロデータと公表データの規模分布比較 (つづき)

(出所) 筆者作成。

結果をみれば、多くの産業において擬似マイクロデータと公表データのグラフがほぼ重なっており、これらの産業では擬似マイクロデータが公表データの規模分布を再現できているといえる。ただし、「H 運輸業, 郵便業」、「J 金融業, 保険業」、「O1 教育, 学習支援業 (学校教育)」、「P 医療, 福祉」のように規模によっては両者が異なっている産業も存在する。また「Q 複合サービス事業」および「R2 サービス業 (政治・経済・文化団体, 宗教を除く)」については、Fleishman 変換のパラメータを推定することができなかった。これらの部門について、より再現性の高い擬似マイクロデータを作成することは、今後の課題の1つである。