



RIETI Policy Discussion Paper Series 05-P-001

工業統計パネルデータの作成 —産業構造データベースの一環として—

新保 一成
経済産業研究所

高橋 睦春
経済産業研究所

大森 民
経済産業研究所



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/>

工業統計パネルデータの作成 —産業構造データベースの一環として—[†]

新保 一成*
高橋 睦春**
大森 民***

要 旨

20 世紀における我が国の産業構造の変化を追跡し、経済・産業政策を評価するためには、観察されたデータの発生機構を記述するモデルの構築が要請される。われわれは産業連関モデルを基礎として、20 世紀における我が国の産業構造の変化を記述するモデルを構築することを念頭においている。この目的のためには、20 世紀に我が国が作成・公表した諸データを、産業連関表を軸にデータ間及び時系列の整合性を保持して再構成し、観察事実を整理することが必要である。具体的には作成・公表された産業連関表の時系列に接続し、接続された産業連関表と工業統計の詳細な商品分類に対応する産出デフレーターを整備する。さらに工業統計及び石油等消費構造統計を事業所ベースの経年追跡データ（パネルデータ）として再構成し、接続された産業連関表との接合を行うことを目指している。われわれは、このように産業連関表を軸に各データを整合的に整理したものを産業構造データベースと呼ぶ。本稿は、マイクロデータとしても活用意義の大きい工業統計パネルデータ作成作業の実際について詳細に述べるものである。

[†] 本稿は独立行政法人産業経済研究所の産業構造データベースプロジェクトの成果の一部を取りまとめたものである。本稿の内容や意見は、筆者らに属し、経済産業研究所の公式見解を示すものではない。

* 経済産業研究所平成 15 年度ファカルティフェロー、慶應義塾大学商学部助教授

** 経済産業研究所 研究部 計量分析・データ室 計量分析・データ担当マネージャー

*** 経済産業研究所 研究部 計量分析・データ室 計量分析・データ担当

1. はじめに

われわれの研究グループは、これまで産業構造データベースの設計および開発を進めてきた。本稿は、その中でも工業統計のパネルデータ化に関する作業マニュアル的な意味合いを持つ。

最初に、われわれの産業構造データベースの構想について簡単にまとめておこう。われわれの究極的な研究目的は、20世紀における我が国の産業構造の変化を追跡し、経済・産業政策を評価することである。経済・産業政策の評価を行うためには、観察されたデータの発生機構を記述するモデルの構築が要請される。われわれは産業連関モデルを基礎として、20世紀における我が国の産業構造の変化を記述するモデルを構築することを念頭においている。この目的のためには、20世紀に我が国が作成・公表した諸データを、産業連関表を軸にデータ間及び時系列の整合性を保持して再構成し、観察事実を整理することが必要である。具体的には作成・公表された産業連関表の金額表示取引表と物量表を時系列に接続し、接続された産業連関表に対応する産出デフレーターを整備する。さらに工業統計及び石油等消費構造統計を事業所ベースの追跡データ（パネルデータ）として再構成し、接続された産業連関表との接合を目指す。

われわれは、これまでに、産業連関表に関しては1980年から1995年の産業連関表を1995年基準に組み替え、さらに1995年基準の2000年表の延長推計を行った。また、産業連関表の部門分類および工業統計の商品分類6桁に対応可能な長期卸売物価指数の推計作業を行ってきた。政府が公表する接続産業連関表では3時点の産業連関表が接続される。したがって、1995年基準に関しては、1985年までが政府によって1995年基準で作成・公表されている。われわれは、工業統計を1980年まで遡ってパネルデータ化しようと試みているので、産業連関表に関しても1980年まで遡って1995年基準の表を推計することとした。

本稿では、工業統計のパネルデータ化に関する作業の実際を、研究所における今後の作業継続のためにまとめておくことにする。最初に工業統計をパネルデータとして再構成することの意味について、一般的なパネルデータのメリットという視点から整理する。続いて工業統計の仕組みについて概説し、工業統計の個票データを使ったパネルデータ化の作業の詳細について解説する。最後に工業統計のパネルデータ整備の課題について述べる。

2. パネルデータを使うメリット

2.1. 個票と集計資料：ユール・シンプソンのパラドックス

2つの変数が因果関係にあったとしても、観察された両者の相関関係の大きさが因果関係の大きさを表すとは限らない。そのことは、歴史的には質的変数の分割表(contingency Table)において Yule(1903)¹や Simpson(1951)²によって指摘され、ユール・シンプソンのパ

¹ Yule, G.U., "Notes on the Theory of Association of Attributes in Statistics", *Biometrika*,

ラドックスと呼ばれている³。これは、パネルデータに限らず、個体レベルのデータを使用することのメリットあるいは集計されたデータしか使用できない場合のデメリットを示す格好の例でもある。以下、シンプソンの1951年の論文から例をひいて説明してみよう。

表 1: ユール・シンプソンのパラドックス (個票)

	男性		女性	
	処置なし	処置あり	処置なし	処置あり
生存	4/52	8/52	2/52	12/52
死亡	3/52	5/52	3/52	15/52

表 2: ユール・シンプソンのパラドックス (集計資料)

	処置なし	処置あり
生存	6/52	20/52
死亡	6/52	20/52

表 1 は、処置の有無を要因として生死が観測された医学データで、個体属性としての性別ごとに記録された 3 元分割表である。表によれば、男性に関して、処置をした場合の生存割合は 8/13(61.5%)で、処置をしなかった場合の生存割合 4/7(57.1%)よりも高い。女性の場合にも同様に、処置をした場合の生存割合は 12/27(44.4%)で、処置をしなかった場合の生存割合 2/5(40%)よりも高い。このように表 1 は、男女の別にかかわらず処置の有効性を示唆している。

表 2 は、表 1 を性別に関して集計して 2 元分割表したものである。表 2 では処置の有無にかかわらず生存割合は 50%であり、処置の有効性が消滅している。これは、女性の生存割合が 14/32(43.75%)と男性の 12/20(60%)よりかなり低いにもかかわらず、処置を受けた個体の割合が 27/32(84.375%)と男性の 13/20(65%)より多いために、処置の有効性がうちけされてしまったためであると考えられる。

このように個体情報が使用できることのメリットは明らかであり、逆に集計情報しか使えない場合には誤った結論を導き出してしまう可能性も否定できないのである。

2.2. 横断面資料, 時系列資料, そしてパネルデータ

工業統計を使って産業別の労働需要を測定することを考えてみよう。生産者の主体均衡

2, pp. 121-34.

² Simpson, E. H., "The Interpretation of Interaction in Contingency Tables", *Journal of the Royal Statistical Society, Series B(Methodological)*, 13, pp. 238-41.

³ 詳細は、宮川正巳、『統計的因果推論—回帰分析の新しい枠組み—』(2004年, 朝倉書店)

図式が教えるところによれば、労働需要関数は、可変生産要素価格、生産量、準固定生産要素の賦存量、技術水準の関数であり、特に可変生産要素価格の 0 次同次関数でなければならない。ここでは、可変生産要素として、労働、燃料、原材料の 3 種類を考えよう。生産要素需要関数はシステムとして推定するほうが推定の精度が高いというのが計量経済学の教えるところであるが、ここでは労働需要関数を単独で推定することを考えよう。次のように労働需要関数を特定化しよう。

$$L_{it} = b_{LL}(y_{it}, z_{it}) + b_{LE}(y_{it}, z_{it}) \left(\frac{w_{Et}}{w_{Lit}} \right)^{1/2} + b_{LM}(y_{it}, z_{it}) \left(\frac{w_{Mt}}{w_{Lit}} \right)^{1/2}$$

L_{it} は事業所 i の t 期の労働需要量で、工業統計の総従業者数がそれにあたる。 w_{Lit} は、賃金で、工業統計の現金給与総額を総従業者数で割った値とする。 w_{Et}, w_{Mt} は燃料および原材料の価格で、卸売物価指数統計の適当な価格指数が対応するものとし、同じ産業に属する事業所は同じ価格に直面しているとしよう。 b_{LL}, b_{LE}, b_{LM} は労働需要関数のパラメーターであるが、生産量 y_t および機械、装置、土地などの準固定要素 z_t の水準によって変動することを認めている。たとえば、

$$b_{LL} = \beta_{LL} y_t^{\beta_{Ly}} z_t^{\beta_{Lz}}$$

という具合に定式化される。 $\beta_{LL}, \beta_{Ly}, \beta_{Lz}$ が推定されるべきパラメーターである。

従来の分析では、生産要素間の代替を分析するためには 2 桁レベルの時系列データ、規模の効果を分析するためにはある時点の横断面データが利用されることが多かった。図 1 は、2 つの事業所の労働投入量と生産量の関係を示している。それぞれの事業所は準固定生産要素の量のみが異なるものとする（事業所 1 の準固定生産要素の賦存量を z_1 、事業所 2 のそれを z_2 とする）。すなわち同じ産業に属する 2 つの事業所は、きわめて同質的な財を異なる生産技術で生産していることを仮定している。点 A, B は事業所 1 の異なる 2 時点に関する労働投入と生産量に関する観察値、同様に点 C, D は事業所 2 に関する観察値を示している。A と C、および B と D はそれぞれ同じ時点に関する観察値であるとする。点 E は、点 A と C の集計値、点 F は点 B と D の集計値で、E および F は産業レベルに集計された時系列データを示している。また、G および H は、各時点の平均値からなるデータを示している。

を参照されたい。

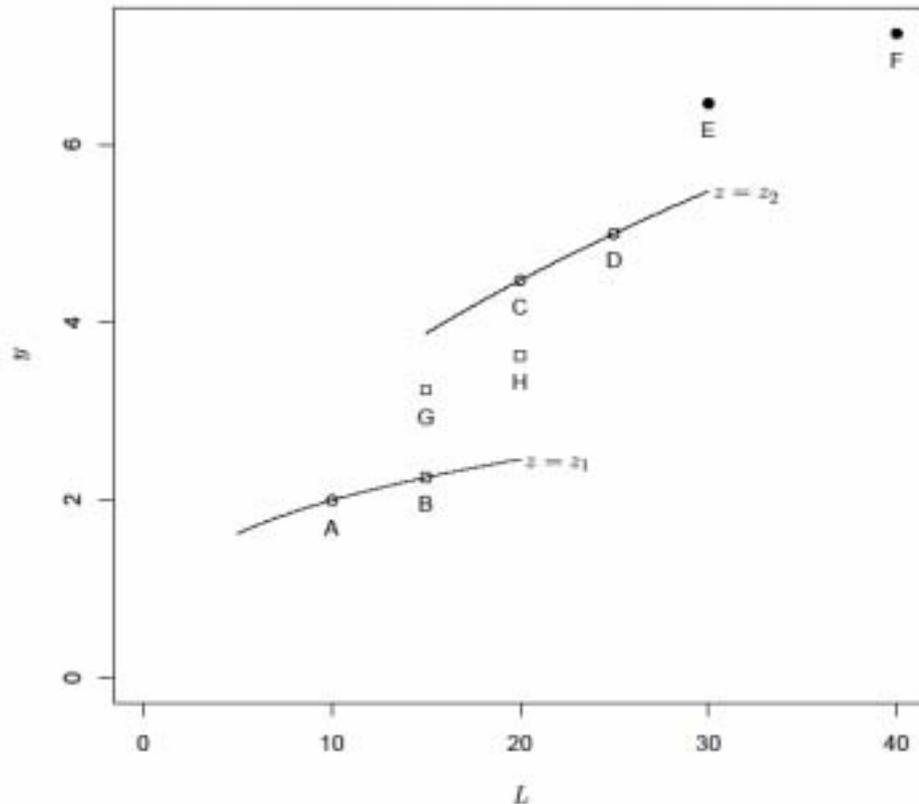


図 1: 横断面資料, 時系列資料, そしてパネルデータ

いま, ある時点の横断面資料しか利用できないものと考えよう (点 A および C (ウェーブ 1 と呼ぶ), または点 B および D (ウェーブ 2 と呼ぶ))。実際には 2 つの事業者は異なる技術を用いて操業しており, 産出量と労働投入量の関係は, 事業所 1 については点 A および B を通る曲線, 事業所 2 については点 C および D を通る曲線で示されるものとしよう。さらに技術を表す変数 z が観測できない変数である場合を考えてみよう。知りたい関係は技術を一定にしたときの, 産出量と労働投入量も関係であるから図に示されている 2 つの曲線である。横断面資料しか利用できないときには, 資料から観察される点 A と C の組み合わせ, または点 B と D の組み合わせによって, 産出量と労働投入の関係を捉えてしまうという誤りを犯してしまう可能性がある。事業所 1 と同じ技術を用いる事業所, 事業所 2 と同じ技術を用いる事業所が多数観察されたとしても, 変数 z をコントロールできないのであれば状況は同じである。また, 変数 z をコントロールできる場合でも, たとえばウェーブ 1 において点 A および点 C の周辺に観測値が集中して観察されるような場合には, 2 つの曲線とはかなり異なる傾きの曲線を推定してしまう可能性がある。

次に, ある事業所だけの時系列だけが利用可能であるとしよう。図をみても明らかなように, いまの例の場合, 事業所間で曲線の形は相当に異なる。したがって, 一つあるいは

少数の事業所の時系列データのみが利用可能であるという状況は、それらからの推定結果を産業全体のものと看做せるかどうかの代表性の問題が残る。

それでは、各事業所の集計値（点 E および F）または平均値（点 G および H）のみが利用可能である場合はどうであろう。集計値または平均値の場合には、先に述べたユール・シンプソンのパラドックスが生じている可能性を否定できない。F または H の観測値が、E または G の観測値に対してどの位置に観測されるかは、事業所 1 と 2 のタイプの事業所のウェイトの変化にも依存する。F へ移動したのは、技術 z_1 を使用する事業所のウェイトが高まった結果なのか、 z_2 の技術のウェイトが高まった結果なのか、見かけ上同じように観測されたとしても、その含意は全く異なる。

以上概観してきた横断面資料、単独個体の時系列資料、集計時系列資料における分析上の短所は、パネルデータが利用可能な場合に克服できる点が多い。それは、パネルデータが上記 3 種類のタイプの資料の特徴を全て併せ持っているからである。パネルデータによれば、個々の事業所の動的な変化と分布の変化を考慮した上での産業全体の平均的な姿を捉えることも可能であるし、産業全体をいくつかのグループに分けてそれぞれの動的な変化の分析を行うことも可能である。

3. 工業統計の仕組み

3.1. 集計データ利用の限界

我が国の全製造事業所は、表 3 に示されているように 1983 年の 78.0 万事業所をピークに減少に転じたが、大・中規模事業所を中心に従業者数や出荷額が増加したことを反映して製造業全体の生産活動は拡大を続けた。その後、従業者数や出荷額は 7 年遅れの 90 年をピークに減少に転じているが、特に事業所数の減少は著しく、2000 年では 59.0 万事業所、ピーク時に比べ約 19 万事業所、25.7%の減少となっている。

このような製造業の生産活動に係わる観測データを体系的にとりまとめたものが『工業統計表』である。一般に「統計表(Statistical Table)」は、個別事象の観測値(マイクロデータ)を収集し、それを何らかの方法で集計・公表したものである。『工業統計表』は、おおむね製造業に属する事業所⁴の生産活動を特定の指標によって観測し、それらを事業所の属性に応じて分類・集計したものである。その産業別の集計値(集計結果、あるいは公表値)からは

⁴ 工業統計調査は、『日本標準産業分類』の F・製造業に属する事業所が対象で、地方公共団体経営の事業所で製造加工を行っているものは対象に含まれるが、国に属する事業（郵政事業、国有林野事業、印刷業、造幣業の製造加工事業所）所は対象外である。これら対象事業所の「製造 Activity」についてみの調査している。また、集計対象事業所数は、総務省（旧：総務庁）が実施する「企業・事業所統計調査」の実施の後では増加し、それ以降漸減するという、いわゆるノコギリ型の現象が見られる。これは、製造事業所か否かを外観で判断することが難しいことなどから小規模事業所を中心に他産業への逃げ込みがあることや、新設事業所等の捕捉が困難であること、等の要因による。

個別事業所のミクロの変動を読み取ることはできない。したがって、『工業統計表』を使っている調査期間について産業別の生産活動の変化とその要因を詳細に分析するには、基礎的な統計資料である個別事業所のミクロデータに遡って検討する必要があるが、研究者が個票データを利用するには、統計制度的にも、また統計情報の入手にも様々な制約がある。

表3：製造業の事業所数，従業者数，出荷額等，付加価値額の推移

	全数調査年	全事業所数	従業者4人以上の事業所						(参考) 事業所番号の変更年
			事業所数			従業者数	出荷額等	付加価値額	
			計	従業者30人以上	従業者29人以下	万人	億円	億円	
1970	○	652,931	405,515	58,460	347,055	1,116	683,763	257,152	昭和45年
1971	○	642,952	397,124	58,259	338,865	1,095	721,716	274,875	
1972	○	702,586	427,172	59,734	367,438	1,122	800,142	312,813	
1973	○	708,447	432,450	59,753	372,697	1,138	1,021,559	389,718	
1974	○	696,795	417,876	57,431	360,445	1,089	1,259,474	440,144	
1975	○	735,970	430,491	56,340	374,151	1,066	1,258,409	449,689	昭和50年
1976	○	727,427	423,904	56,211	367,693	1,054	1,435,254	511,812	
1977	○	714,177	415,014	54,392	360,622	1,025	1,549,428	551,846	
1978	○	744,337	431,102	54,153	376,949	1,023	1,625,936	611,162	
1979	○	739,304	430,439	54,203	376,236	1,021	1,818,409	670,187	昭和55年
1980	○	734,623	429,336	53,868	375,468	1,029	2,121,243	698,381	
1981			436,463	55,430	381,033	1,057	2,247,119	731,673	
1982			427,998	55,207	372,792	1,048	2,299,340	762,228	
1983	○	780,280	446,942	55,703	391,239	1,065	2,355,269	791,455	
1984			428,998	56,584	372,414	1,073	2,530,298	859,379	昭和61年
1985	○	749,366	438,518	57,626	380,892	1,089	2,653,206	905,227	
1986			436,009	58,349	377,660	1,089	2,546,886	892,054	
1987			420,804	58,154	362,650	1,074	2,535,153	928,245	
1988	○	745,108	437,574	59,003	378,571	1,091	2,744,007	1,027,288	
1989			421,757	59,596	362,161	1,096	2,988,931	1,109,927	平成3年
1990	○	728,853	435,997	60,386	375,611	1,117	3,233,726	1,190,283	
1991			430,414	61,669	368,745	1,135	3,408,346	1,257,476	
1992			415,112	60,974	354,138	1,116	3,295,206	1,211,358	
1993	○	696,090	413,670	59,209	354,461	1,089	3,111,995	1,149,739	
1994			382,825	57,164	325,661	1,042	2,990,274	1,123,492	平成9年
1995	○	654,436	387,726	56,749	330,977	1,032	3,060,296	1,172,043	
1996			369,612	56,106	313,506	1,010	3,130,684	1,193,040	
1997			358,246	55,386	302,860	994	3,230,718	1,198,728	
1998	○	643,468	373,713	54,518	319,195	984	3,058,400	1,131,931	
1999			345,457	52,469	292,988	938	2,914,496	1,078,596	平成14年
2000	○	589,713	341,421	51,434	289,987	918	3,004,776	1,102,426	
2001			316,267	49,364	266,903	887	2,866,674	1,033,051	
2002			290,848	47,046	243,802	833	2,693,618	974,587	

資料：各年工業統計表より作成

3.2. 実施機関

経済産業省経済産業政策局調査統計部（旧：通商産業大臣官房調査統計部）

3.3. 根拠法規

工業統計調査は、統計法(昭和22年法律第18号)及びこれに基づく工業統計調査規則(昭和26年通商産業省令第81号、最終改正平成12年通商産業省令第278号)により、指定

統計第 10 号として実施している。

3.4. 工業統計の沿革

工業統計の歴史は古く、1968 年(明治初年)に民部省が「府県物産表」調査を開始したことから始まると言われており、日本を代表する経済センサス⁵のひとつである。1883 年(明治 16 年)からは農商務省統計の「工場調査」(従業員 10 人以上)として毎年実施されたが、1910 年(明治 43 年)からは、5 年毎の「工場統計調査」(職工 5 人以上)として実施されるようになった。更に 1920 年(大正 9 年)からは再び毎年調査となるなど、調査周期は度々変更されてきた。その後、1939 年(昭和 14 年)からは調査の名称を「工場調査」に変え、対象範囲も全ての工場、作業場にまで拡大された。

戦後に入って、1947 年(昭和 22 年)からは統計法に基づく指定統計に指定(第 10 号として指定)され、調査の名称も「工業調査」に改められた。1950 年(昭和 25 年)には世界センサスへの協力から調査の名称も「昭和 25 年工業センサス」と一時的に変えられたが、翌 1951 年(昭和 26 年)以降は「工業統計調査」に改称し、1981 年(昭和 56 年)調査からは予算的な制約から西暦末尾 0, 3, 5, 8 年は全数調査、それ以外の年は従業者 4 人以上の事業所について調査する、いわゆる裾切調査(Cut off 調査)の導入を図り、現在に至っている。なお、1991 年(平成 3 年)には、長崎県の島原市及び深江町については雲仙普賢岳の噴火災害により調査の範囲から除外され、2000 年(平成 12 年)には、東京都三宅村については、三宅島(雄山)の噴火災害により調査の範囲から除外されている。

3.5. 調査の仕組み

工業統計調査は、日本標準産業分類(平成 5 年総務庁告示第 60 号)に掲げる大分類 F—製造業に属する事業所(国に属する事業所を除く)業種別、従業者規模別、地域別等に従業者数、製造品出荷額等を把握し、わが国工業の実態を明らかにし、工業に関する施策の基礎資料を得ることを目的として実施されている。なお、第 11 回の日本標準産業分類改訂は 2002 年(平成 14 年)10 月から施行。

調査は、毎年、12 月 31 日現在で実施され、約 60 万事業所を対象に、経済産業省→都道府県→市区町村経由の調査員調査として実施されている。なお、1981 年(昭和 56 年)調査からは、予算的な制約から西暦末尾 0, 3, 5, 8 年は全数調査、それ以外の年は従業者 1～3 人規模以下の事業所については特定業種についてのみ調査する、いわゆる裾切調査(Cut off 調査)の導入を図っている。

この調査は、甲調査(従業者 30 人以上)と乙調査(従業者 29 人以下)の 2 種類⁶から構成され

⁵ 経済センサスとは、総務省実施の「企業・事業所統計調査(5 年毎)」と経済産業省実施の「工業統計調査(毎年)」「商業統計調査(5 年毎)」の 3 つの大規模な事業所ベースの統計の俗称として使われている。

⁶ 工業統計調査の歴史の中で、個々の事業所(工場)を統括する本社・本店のみを対象(一定規模以上)とした丙調査(「本社本店調査」)が昭和 31 年から開始されたが、この丙調査

ており、調査年の翌年 10 月頃に速報、翌々年 3 月頃に確報が公表されている。最終的な確報は、毎年 3 月～5 月頃に財務省印刷局から、『工業統計表』として各編⁷(産業編、品目編、用地・用水編、工業地区編、企業統計編⁸)が順次刊行されている他、(財)経済産業調査会(旧:通商産業調査会) 経済統計情報センターから磁気媒体でデータ提供(有料)されている。

①工業調査票甲

調査対象 地域： 全 国

単位： 事業所

属性： 従業者 30 人以上の製造業に属する事業所（製造、加工又は修理を行っていない本社又は支店を除く）

調査方法 選定： 全 数

対象数： 約 50,000

配布・収集 調査員

記 入 自計方式

周期・期日 周 期： 年

実施期日： 毎年 12 月 31 現在

調査系統 経済産業省→都道府県→市（区）町村→調査員→報告者

調査事項

- 1.事業所の名称及び所在地, 2.本社又は本店の名称及び所在地, 3.他事業所の有無,
- 4.経営組織, 5.資本金額又は出資金額（会社に限る）, 6.従業者数, 7.常用労働者毎月末現在数の合計, 8.現金給与総額, 9.原材料, 燃料, 電力の使用額及び委託生産費, 10.有形固定資産, 11.リース契約による契約額及び支払額, 12.製造品在庫額, 半製品, 仕掛品の価格及び原材料, 燃料の在庫額, 13.製造品の出荷額, 在庫額等,
- 14.品目別製造品出荷額, 加工賃収入額及び修理料収入額の合計金額, 15.内国消費税額, 16.製造品出荷額に占める直接輸出額の割合, 17.主要原材料名, 18.作業工程,
- 19.工業用地及び工業用水

は昭和 59 年調査をもって廃止された。代わって、昭和 62 年と平成元年には、製造企業を対象とした「多角化等調査」が丙調査として新たに実施されたが、この丙調査も平成 4 年に「通商産業省企業活動基本調査」が新たに発足したのに伴って廃止された。

⁷ 品目編、工業地区編、企業統計編は従業者 4 人以上の事業所、用地・用水編は従業者 4 人以上の事業所について集計したものである。

⁸ 平成 10 年調査からは従業者 4 人以上の事業所に集計対象範囲を拡大しているが、それ以前では、従業者 20 人以上の事業所について、企業単位に組み替えて再集計したものである。なお、丙調査（本社本店調査）の廃止に伴い、従来加えられていた本社本店の従業者数は、この『企業統計編』のデータには加算されていない。また、この企業統計編は、過去『企業編』として刊行されていたが、企業活動基本調査の創設に伴い、『企業統計編』に名称が変更されているほか、丙調査（本社本店調査）結果を使った再集計から甲・乙調査結果のみによる再集計方式に変更している。

②工業調査票乙

調査対象 地域： 全 国
単位： 事業所
属性： 従業者 29 人以下の製造業に属する事業所（製造，加工又は修理を行っていない本社又は支店を除く）
調査方法 選定： 全 数
対象数：約 550,000
配布・収集 調査員
記入 自計方式
周期・期日 周期： 年
実施期日： 毎年 12 月 31 現在
調査系統 経済産業省→都道府県→市（区）町村→調査員→報告者
調査事項

1.事業所の名称及び所在地，2.本社又は本店の名称及び所在地，3.他事業所の有無，4.経営組織，5.資本金額又は出資金額（会社に限る），6.従業者数，7.現金給与総額，8.原材料，燃料，電力の使用額及び委託生産費の合計金額，9.製造品出荷額等，10.品目別製造品出荷額，加工賃収入額及び修理料収入額の合計金額，11.内国消費税額，12.製造品出荷額に占める直接輸出額の割合，13.主要原材料名及び簡単な作業工程，14.有形固定資産（従業者 10 人以上の事業所）

③準備調査名簿

調査の実施に先立って調査員が準備調査名簿を作成。

3.6. 個票データ等の電算入力状況

この調査において個票の電算入力は 1968 年(昭和 43 年)調査から始まり，同年調査以降の個票データは磁気ファイルで保存されているものの，統計の信頼性の面から現在，個票の統計目的外利用申請ができるのは，1974 年(昭和 49 年)調査からである（目的外利用申請での遡及年次）。実際，磁気ファイル化されているのは，従業者 4 人以上の事業所のみで，1～3 人規模事業所については，サマリー（都道府県別産業別）のみの入力にとどまっている。なお，最近地方分散入力との関係で，1998 年調査と 2000 調査の全数調査年については 1～3 人規模の事業所についても個票入力されている。他方，工業準備調査名簿⁹の電算入

⁹工業準備調査名簿は，従来その利用が厳しく制限されおり，他の統計調査の母集団名簿として利用する以外には利用できなかった。また，紙ベースでの保存には物理的な制約もあり，新しい名簿が作成され次第，破棄され，過去の準備調査名簿は電子化されるまでは保存されていない。

力（漢字入力）は遅れて始まり、現在のように磁気ファイルにて利用できるのは、1990年調査分からである。

3.7. 工業統計の事業所コード等の設定の仕方

工業統計では、企業名あるいは事業所名は財政的な制約等もあって入力されず、何桁かの番号(コード)でもって識別されてきた。

3.7.1. 事業所コードの設定方法

個々の事業所データを追跡するには、各年の個票データ(磁気テープ)を使って、個々の事業所を過去に遡及して追跡する必要がある。そのとき重要になるのが個々の事業所を識別するキー項目(「属性指標」といもいう)である。個々の事業所を識別するキー項目とは、個々の事業所が、どの地域(都道府県・市区町村・基本調査区からなる地域コード¹⁰)の、どの産業(4桁の分類からなる産業コード)の、どの従業者数規模階級(従業者規模コード)に属しているか、を識別する項目である。工業統計調査は、行政区域いわゆる市区町村単位で調査が行われるため、事業所番号は市区町村一連番号で設定されている。

経済産業省(経済産業政策局調査統計部)では、原則として西暦末尾の0、5年をベースに翌年の調査で「工業調査事業所番号」の設定替えを行い、その後5年間は原則として変更せず使用している。しかし、事業所番号設定後、事業所の廃業や移転(他の都道府県や市区町村への移動)、他産業への転業により、対象外となる事業所が発生するが、それら事業所の番号は欠番処理を行い、新設事業所(新規事業所の他、他地域からの転入、他産業からの転業の事業所も含む)は最終事業所の次の番号から追加される。このほか、市区町村の合併¹¹や政令都市の出現¹²などによって、市区町村の行政区域が変更になり、それに伴って事業所番号の設定替え等が行われる。

3.7.2. 品目コードの設定方法

もう一つ、産業分類の決定に必要な製造品出荷額等は、「合計値」と「品目別製造品出荷額」とを識別するための「品目コード」(6桁の数字からなるコード)からなっている。これら品目コードについても日本標準産業分類の改訂(第11回改訂の概要については参考資料を参照)や新商品の出現や既存商品の生産縮小などによって変更される。

¹⁰ 工業統計調査では、基本調査区として、当初独自に「工業調査区」を設けて使っていたが、1980年調査から「事業所調査区」に統一化を図っている。

¹¹ 市区町村の合併により、ある市区町村に加わった事業所の番号は、当該市区町村における最終事業所の次の番号から、順次付け替えが行われている。また、当該市区町村の一部地域が他の市区町村に合併され管轄区域が変更した場合は、当該地域の事業所番号は欠番処理されている。

¹² 大阪市、仙台市等の政令都市が出現し、それに地域が幾つかの区に分けられたような場合は、それぞれの区毎に市区町村番号が設定され、その市区町村番号の基で、既存の事業所番号が付け替えられる。

日本標準産業分類の改訂事例

第 10 回改訂（旧分類）

告示日 1993 年 10 月 4 日 適用日 1994 年 4 月 1 日

第 11 回改訂（新分類）

告示日 2002 年 3 月 7 日 適用日 2002 年 10 月 1 日

2001 年調査と 2002 年調査を事例に示すと、次のようなコード設定替えを行っている。

旧分類コード	……………	2001 年調査（第 10 回改訂分類適用）
旧－新分類対応コード	………	2001 年調査の新分類（第 11 回改訂分類）
新分類コード	……………	2002 年調査（第 11 回改訂分類適用分類）

3.8. 事業所の開業・廃業等の定義

調査は、工業統計調査員(都道府県知事が任命)が『工業調査準備調査名簿』に従って、調査票(「個票」)を配布・回収する調査員調査方式をとっている。工業準備調査名簿は、「ある地域(市区町村)」の「ある調査区(基本調査区)」にある製造事業所を全て網羅した名簿である。最近では、当年の調査票に記載された「事業所の名称及び所在地」等の情報を、電算機に入力し、それを基に翌年の準備調査名簿を作成する方式をとっている。

準備調査名簿で対象事業所となっている前回調査の事業所の他に、新規事業所(新設、転入、他産業からの転業)があった場合は、その事業所が対象事業所(製造事業所か否か)であることを確認した上で、製造業を営む事業所¹³であれば、名簿に追加し、調査票を配布することになっている。調査票(「個票」)は、調査員が配布・回収するが、事業所番号の変更等は市区町村段階で行っている。したがって、この調査では、同一事業所か否かは「同一市区町村内に事業所が存続しているかどうか」で判断している。前節で詳述したように、対象事業所を把握する最小単位は市区町村で、事業所番号も市区町村単位で設定されている。したがって、比較年次に同一市区町村に引き続き事業所がある場合、基準年は(残存)で、比較年は(存続)となる。

①開業及び廃業

開業事業所及び廃業事業所の定義¹⁴は、一般に言われているより広く、基準年と比較して

¹³ 事業所とは、「一区画を占めて経済活動を行っている場所」をいう。製造業とは、①主として新製品(必ずしも完成品だけではなく部品等も含む)の製造加工を行う事業所と②製造加工した新製品を主として卸販売する事業所の2つの条件を満たしている事業所をいう。工業統計調査では、このほか自家発電と自家倉庫に限って付随事業所の活動として、製造事業所分に合算して調べている。

¹⁴ 工業統計調査では、操業準備中の事業所や操業開始後未出荷の事業所も調査の対象となる。

比較年に、〈開業〉及び〈廃業〉の事由により追加された事業所、あるいは削除された事業所をいう。

②転出及び転入

転出事業所及び転入事業所の定義は、いずれの時点において同一の事業所が存在しているが、その事業所の分類属性が異なっている事業所をいう。例えば、ある産業(2桁, 3桁, 4桁)に属する事業所が比較年に他の産業に移動した場合は、転出事業所、逆に、比較年に当該産業(2桁, 3桁, 4桁)に入って来た場合は、転入事業所という。これら転出、転入事業所の数は、各産業別に見ると当然違うが、製造業全体で見れば、同じ現象を表と裏の両面から見たもので同数となる。

表 4: 開業・廃業の理由

事業所番号	同一市区町村の事業所番号
〈開業〉 ・製造業に新規参入 ・他の産業から製造業に転入 ・他の市区町村からの転入 ・1～3人規模からの規模上がり	・比較年に追加されたもの
〈廃業〉 ・製造業を廃業 ・製造業以外の産業への転出 ・他の市区町村への転出 ・1～3人規模への規模下がり	・比較年に削除されたもの

3.9. 商品と産業の対応と産業の決定方法

工業統計における「商品(品目)」と「産業」の関係は、商品分類の体系は『工業統計調査用商品分類』、産業分類の体系は『日本標準産業分類』によっている。このふたつの分類体系は基本分類で対応しており、「事業所」を通じて結び付けられている。この調査の対象事業所は、自ら生産しているあらゆる「製造品名」「賃加工品目名」「番号」「数量単位名」などを、調査票と同時に配られた『商品分類表』に従って個票に記入する自計方式によっている。このとき、「製造品名」「加工賃名」は6桁品目分類コードで記入する。調査票には製造品目の入欄(甲調査:製造品7品目,加工賃4品目,乙調査:製造品5品目,加工賃2品目)が刷り込まれているが、これ以上の品目を生産している事業所については、調査関係書類に添付されている『補助用紙』を使って記入してもらう。

工業統計の産業格付けは、調査結果の出荷額等をもとに『日本標準産業分類』に基づい

また、廃業及び休業事業所の定義は、次のとおり。

- ・廃業事業所とは、現在操業を中止しており、将来再開する意志がないものをいう。
- ・休業事業所とは、現在操業を中止しているが、将来再開する意志があるものをいい、調査対象とする。

て決定している¹⁵。具体的には、製造品の6桁分類番号をもとに、単品を生産している事業所の場合は、品目番号の上4桁番号でもって産業細分類を決定し、3桁番号で小分類、2桁番号で産業中分類を決定している。これに対して、複数の品目を生産している事業所の場合は、品目6桁番号の上2桁番号の最も大きなもので2桁の産業中分類を決定し、以下同様の方法で、3桁の産業小分類、4桁の産業細分類を決め、最終的な産業格付けとしている。

3.10. 工業統計表の項目の定義等

工業統計の主な項目の定義等は、以下のとおり。

3.10.1. 事業所数

事業所数は、12月31日現在の数値である。事業所とは、一般工場、製作所、加工所などと呼ばれている一区画を占めて主として製造又は加工を行っているものをいう。

3.10.2. 従業者数

従業者数は、12月31日現在の常用従業者数と個人事業主及び無給家族従業者との合計である。

①常用従業者とは、次のいずれかの者をいう。

(a) 期間を決めず、又は1か月を超える期間を決めて雇われている者。

(b) 日々又は1か月以内の期限を限って雇われていた者のうち、その月とその前月にそれぞれ18日以上雇われた者。

(c) 重役、理事などの役員のうち、常勤して毎月給与の支払いを受けている者。

(d) 事業主の家族で、その事業所に働いている者のうち、常時勤務して毎月給与の支払いを受けている者。

②個人事業主及び無給家族従業者とは、業務に従事している個人事業主とその家族で、無報酬で常時就業している者。したがって、実務にたずさわっていない事業主とその家族で手伝い程度のものは含まれていない。

③現金給与総額とは、1年間に常用労働者に対し決まって支給された給与および特別に支払われた給与の額とその他の給与の合計である。

3.10.3. 製造品出荷額等

製造品出荷額等とは、製造品出荷額、加工賃収入額、修理料収入額、製造工程から出た

¹⁵ 工業統計調査の産業格付けは、日本標準産業分類の準じ、原則として、製造品出荷額等の大小でもって決定しているのに対し、同じ製造事業所を対象に調査している「企業・事業所統計調査」では、原則として、従業者数あるいは就業者数の大小でもって決定している。前者では、生産活動の結果であるアウトプットで測っているのに対し、後者では、労働投入のインプットでもって測っている。これは、1人当たりの労働生産性が産業部門間で等しいと仮定したことになる。

「くず及び廃物の出荷額」及び「その他の収入額」の一年間の合計である。

①製造品の出荷

- (a) その事業所の所有に属する原材料によって製造されたもの（原材料を他に支給して製造されてものを含む）を当該年中にその事業所から出荷した場合。
- (b) 同一企業に属する他の事業所へ引き渡したものの。
- (c) 自家使用されたもの（その事業所において最終製品として使用されたもの）
- (d) 委託販売に出したもの（販売済みでないものを含み、各当該年中に返品されたものを除く）

②製造品出荷額は、工場出荷額によっている。特に、

- (a) 内国消費税を課せられたものは、その税額を含めた工場出荷額。
- (b) 割引、値引きされたものは、その分を差し引いた販売価額によっている。

③加工賃収入額

該当年中に他の企業の事業所から支給された主要原材料によって製造し、あるいは他の所有する製品、半製品に加工処理を加え、これによって加工賃を受け取った、又は受け取るべき加工賃。

④修理料収入

他の物を修理して受け取った、又は受け取るべき修理料。ただし、事業所全体の修理料収入額が出荷額等の2分の1を超える場合はサービス業で対象外となる。

⑤製造工程からでたくず、廃物

酒かす、精米かす、精麦かす、製材くず、鉄くず、非鉄金属くずなど。

⑥その他の収入額

冷蔵保管料、新聞・雑誌広告料、自家発電の余剰電力の販売電力料など。

なお、新聞・雑誌広告料については、2002年（平成14年）調査から日本標準産業分類の変更に伴って報告がなくなっている。

3.10.4. 生産額（算式）

生産額 = 製造品出荷額等 + (製造品年末在庫額 - 製造品年初在庫額) + (半製品および仕掛品年末価額 - 半製品および仕掛品年初価額)

3.10.5. 付加価値額（粗付加価値額）（算式）

①従業者30人以上

付加価値額 = 生産額 - 内国消費税 - 原材料使用額等 - 減価償却額

②従業者10人～29人

付加価値額 = 製造品出荷額等 - 内国消費税 - 原材料使用額等 - 減価償却額

③従業者9人以下

付加価値額 = 製造品出荷額等 - 内国消費税 - 原材料使用額等

4. 工業統計パネルデータの作成

以下に説明する方法によって、1988年から1999年までの工業統計の個票データとして磁気媒体に記録されているすべて事業所に関して、経年的に接続に成功した。1987年以前に関しては、名簿情報が完備していないので、データに記録されている情報（地域、産業を固定したときの従業者数や敷地面積など）から接続する必要があり、まだ接続が完了した事業所は限られている。

4.1. TASK1: 工業統計表のデータベースへの登録

統計の目的外利用申請¹⁶によって利用可能になった工業統計の個票データには、1年間で約35万から40万以上の事業所のデータが記録されている。これらのデータを事業所に関する経年追跡データ(パネルデータ)にするためには、必然的に複数年のデータを同時に処理することが必要である。この膨大なデータをプログラム上に一気に読み込みデータ処理を施すのは、コンピューター資源が豊富になった昨今といえども限界があり、かつ無駄も多い。なぜならば、あるデータ処理についてすべての変数が同時に必要なわけではないし、事業所を経年接続するという目的からすれば、コンピューターのメモリー上に存在すべき事業所データは都道府県および市区町村でコントロールされた範囲である。つまり、事業所および変数の範囲を条件付けてプログラム上にデータをロードしながら処理を施すことができれば無駄がない。そこで、すべてのデータをリレーショナルデータベース上に格納し、データ処理プログラムからSQL文を発行して必要なデータのみを処理するという方針で作業を進めることとした。データベースプログラムとして、オープンソースで豊富なユーザーの存在するPostgreSQLを選択した。

TASK1では、PostgreSQLに各年の工業統計の個票データ(産業編甲乙、商品編甲乙、甲は従業者30人以上の事業所、乙は従業者4人~29人の事業所に関する調査を指す)を登録する。

4.1.1. 作業ディレクトリ

以下、工業統計の個票データのパネルデータ化を行うユーザーをrietiとし、作業ディレクトリをCMFWKとする(具体的には、/home/rieti/cmfである)。また、CMFWK/task1をTASK1ディレクトリと呼ぶ。

4.1.2 TASK1の作業フロー

¹⁶ この研究プロジェクトでは、工業統計の個票データを利用しているが、このために統計の目的外利用申請手続きを行った上で使用している。したがって、本システム利用に際しても同様の手続きが必要になる。

- **[TASK1-1]** 工業統計の個票データの一時的コピー。TASK1 ディレクトリの下にサブディレクトリ `CMFWK/task1/data` を作成し、そこに目的外申請によって利用可能になった工業統計の個票データを一時的にコピーする。提供されたデータは、各年 4 つのファイルから構成され、産業編甲は `sankouYYYY.txt`、産業編乙は `sanotuYYYY.txt`、商品編甲は `hinkouYYYY.txt`、商品編乙は `hinotuYYYY.txt` というファイル名である。YYYY は、西暦で表記された調査年を示す。`CMFWK/task1/data` の下に `YYYY(1980~1999)` というディレクトリを作り、各年の 4 つのファイルをコピーする。
- **[TASK1-2]** 登録定義ファイルの作成。コピーされた工業統計の個票データは、`cmf2db` という Perl スクリプトによって、PostgreSQL が読み込めることができる形式に変換され、PostgreSQL への問い合わせツール `Psql` によって登録される。`cmf2db` は、前もって用意された定義ファイルの指示にしたがって工業統計の個票データを PostgreSQL が読み込める形式に変換する。
- **[TASK1-3]** 工業統計の個票データのデータベースサーバーへの登録。`cmf2db` と `Psql` によって各年の工業統計の個票データを PostgreSQL データベースサーバーに登録する。各年の産業編と商品編それぞれに 1 つのテーブルを作成する。つまり、各テーブルには甲調査と乙調査の両調査のデータが格納される。甲調査と乙調査は、`type` という変数で識別される(`type=1` が甲、`type=2` が乙)。テーブル名は、産業編の場合 `indYYYY`、商品編の場合 `comYYYY` である(ここでも YYYY は調査年を示す)。`cmf2db` は、提供されたファイルにおいて空欄になっている調査項目を NULL 値として扱う(¥N に変換)。また、乙調査では調査されていない項目についても NULL 値扱いとする。1980 年から 1999 年のデータの登録に 6 時間程度必要である。`cmf2db` は、`Psql` の実行に必要な SQL 文(CREATE TABLE 文および¥copy 文)を同時に作成する。ファイル名は、`indYYYY.sql` および `comYYYY.sql` である。`cmf2db` によって変換されたファイルは、`Psql` の `COPY table FROM file` によって PostgreSQL に読み込む。この作業には大した時間はかからない(80 分程度)¹⁷。
- **[TASK1-4]** 常用労働者数と現金給与総額(常用労働者)の調整。1980 年～1984 年の甲調査では、常用労働者数が男女別に生産労働者と管理事務労働者に分けて調査されていたが、1985 年以降は生産労働者と管理事務労働者の区別が廃止され、男女別の常用労働者数のみが調査されている。同様に、現金給与総額に関しても 1980 年～1984 年では、生産労働者と管理事務労働者別に調査されていたが(男女の区別はない)、1985 年以降はこの区別が廃止されている。1980 年～1984 年のファイルには、常用労働者男子計、常用労働者女子計、現金給与総額常用労働者計が計上されていないので、これらの合計値を計算する。この作業は `Psql` で行う。
- **[TASK1-5]** バックアップの作成。TASK1-3 に膨大な時間を要するので、以降の作業に

¹⁷ CPU Pentium III 500Mhz, メモリー250MByte, OS Vine Linux 2.5 で実行した結果である。

おけるエラーの発生に備えてバックアップを作成しておく。これは Psql の COPY table TO file コマンドによって行う。エラーが発生したときのリカバリーは、COPY table FROM file で対応できるので、大して時間はかからない。すべての作業が終了した時点でバックアップファイルも消去されなければならない。

- **[TASK1-6]** 工業統計の個票データの消去。TASK1-1 で一時的に TASK1 ディレクトリにコピーした工業統計の個票データを消去する。

以下、TASK1-2, 3, 4, 5, 6 について説明する。

4.1.3. TASK1-2: 登録定義ファイルの作成

経済産業省から提供されたファイルはそのままでは PostgreSQL で読み込める形式になっていないので、前処理が必要である。また、1985 年の前後で調査項目に相違があり、その調整も前もって行いたい。さらに、データの値がゼロであるものと未記入を区別するために、未記入に関しては NULL 値を設定したい。これらの前処理を施すプログラムが Perl で記述された cmf2db である。cmf2db は、工業統計の個票データの種類ごとに用意された定義ファイルにしたがって工業統計の個票データを PostgreSQL が読み込める形式に変換するプログラムである。以下、定義ファイルの仕様について説明する。

以下に、1999 年の産業編を登録するための定義ファイル(ind1999.def)を示す。

```
# Database Definition
# Census of Manufacture, Industry Table 1999

# Table Name
ind1999

# Input File Information
/home/rieti/cmf/task1/data/sankou1999.txt, /home/rieti/cmf/task1/data/sanotu1999.txt
86, 33 # no. of items in a record

# Field Definition
1, 0, 0, Year, INT, 1999 # 調査年
2, 2, 2, Type, INT # 甲・乙区分
3, 3, 3, State, CHAR(2) # 都道府県
4, 4, 4, City, CHAR(3) # 市町村番号
5, 5, 5, Code_Est, CHAR(5) # 事業所番号
6, 6, 6, Code_Ind, CHAR(4) # 産業コード
7, 7, 7, Afl, INT # 他事業所有無
8, 8, 8, Org, INT # 経営組織
```

9,	9,	9,	SuvArea,	INT	# 基本調査区
10,	10,	10,	Scale_Lab,	INT	# 従業規模
11,	11,	11,	Scale_Cap,	INT	# 資本規模
12,	12,	12,	Scale_Ship,	INT	# 出荷規模
13,	13,	13,	Scale_Area,	INT	# 敷地規模
14,	14,	14,	Capital,	INT	# 資本金または出資金
15,	15,	15,	L_O_M,	INT	# 従業者数:常用労働者:男
16,	16,	16,	L_O_F,	INT	# 従業者数:常用労働者:女
17,	17,	17,	L_S_M,	INT	# 従業者数:個人事業主:男
18,	18,	18,	L_S_F,	INT	# 従業者数:個人事業主:女
19,	19,	19,	L_Total,	INT	# 従業者数:合計
20,	20,	NA,	L_O_Cumsum,	INT	# 延常用従業者数
21,	21,	20,	Shipment,	INT	# 出荷額:合計
22,	22,	NA,	C_L_O,	INT	# 現金給与総額:常用労働者
23,	23,	NA,	C_L_Other,	INT	# 現金給与総額:その他
24,	24,	21,	C_L_Total,	INT	# 現金給与総額:合計
25,	25,	NA,	C_Mat,	INT	# 原材料使用額
26,	26,	NA,	C_Fuel,	INT	# 燃料使用額
27,	27,	NA,	C_Elec,	INT	# 電力使用額
28,	28,	NA,	C_Consig,	INT	# 委託生産費
29,	29,	22,	Cost,	INT	# 原材料・燃料・電力・委託生産費合計
30,	30,	23,	Tax,	INT	# 国内消費税
31,	31,	24,	K_Land,	INT	# 有形固定資産:年初現在高:土地
32,	32,	25,	K_DepAsset,	INT	# 有形固定資産:年初現在高:有形固定資産計
33,	33,	NA,	K_Bld,	INT	# 有形固定資産:年初現在高:建物・構築物
34,	34,	NA,	K_Mac,	INT	# 有形固定資産:年初現在高:機械・装置
35,	35,	NA,	K_Other,	INT	# 有形固定資産:年初現在高:その他
36,	36,	26,	R_Land,	INT	# 有形固定資産:除却額:土地
37,	37,	27,	R_DepAsset,	INT	# 有形固定資産:除却額:有形固定資産計
38,	38,	28,	Dep,	INT	# 有形固定資産:減価償却額
39,	39,	NA,	Bld_Inc,	INT	# 有形固定資産:建設仮勘定増
40,	40,	NA,	Bld_Dec,	INT	# 有形固定資産:建設仮勘定減
41,	41,	29,	A_Land,	INT	# 有形固定取得額:土地
42,	42,	30,	A_N_DepAsset,	INT	# 有形固定取得額:新規:有形固定資産計
43,	43,	NA,	A_N_Bld,	INT	# 有形固定取得額:新規:建物・構築物
44,	44,	NA,	A_N_Mac,	INT	# 有形固定取得額:新規:機械・装置

45,	45,	NA,	A_N_Other,	INT	# 有形固定取得額:新規:その他
46,	46,	31,	A_U_DepAsset,	INT	# 有形固定取得額:中古:有形固定資産計
47,	47,	NA,	A_U_Bld,	INT	# 有形固定取得額:中古:建物・構築物
48,	48,	NA,	A_U_Mac,	INT	# 有形固定取得額:中古:機械・装置
49,	49,	NA,	A_U_Other,	INT	# 有形固定取得額:中古:その他
50,	50,	NA,	S_B_Finish,	INT	# 在庫額:年初:製造品
51,	51,	NA,	S_B_Prg,	INT	# 在庫額:年初:半製品・仕掛品
52,	52,	NA,	S_B_Mat,	INT	# 在庫額:年初:原材料・燃料
53,	53,	32,	S_B_Total,	INT	# 在庫額:年初:合計
54,	54,	NA,	S_E_Finish,	INT	# 在庫額:年末:製造品
55,	55,	NA,	S_E_Prg,	INT	# 在庫額:年末:半製品・仕掛品
56,	56,	NA,	S_E_Mat,	INT	# 在庫額:年末:原材料・燃料
57,	57,	33,	S_E_Total,	INT	# 在庫額:年末:合計
58,	58,	NA,	Area_Lot,	INT	# 敷地:建築面積:敷地面積
59,	59,	NA,	Area_Bld,	INT	# 敷地:建築面積:建築面積
60,	60,	NA,	Area_Total,	INT	# 敷地:建築面積:延べ建築面積
61,	61,	NA,	Land_Farm,	INT	# 地面別面積:田畑
62,	62,	NA,	Land_Resid,	INT	# 地面別面積:宅地
63,	63,	NA,	Land_Forest,	INT	# 地面別面積:山林・原野
64,	64,	NA,	Land_Landfill,	INT	# 地面別面積:埋立地
65,	65,	NA,	Land_Other,	INT	# 地面別面積:その他
66,	66,	NA,	Land_Total,	INT	# 地面別面積:合計
67,	67,	NA,	WF_S_Pub_Ind,	INT	# 淡水水源別:公共(工業用)
68,	68,	NA,	WF_S_Pub_Sup,	INT	# 淡水水源別:公共(上水道)
69,	69,	NA,	WF_S_Surf,	INT	# 淡水水源別:地表水・伏流水
70,	70,	NA,	WF_S_Well,	INT	# 淡水水源別:井戸水
71,	71,	NA,	WF_S_Other,	INT	# 淡水水源別:その他
72,	72,	NA,	WF_S_Recycle,	INT	# 淡水水源別:回収水
73,	73,	NA,	WF_S_Total,	INT	# 淡水水源別:合計
74,	74,	NA,	WF_U_Boiler,	INT	# 淡水用途別:ボイラー用水
75,	75,	NA,	WF_U_Mat,	INT	# 淡水用途別:原料用水
76,	76,	NA,	WF_U_Treat,	INT	# 淡水用途別:処理・洗浄用水
77,	77,	NA,	WF_U_Cool,	INT	# 淡水用途別:冷却用水
78,	78,	NA,	WF_U_Heat,	INT	# 淡水用途別:温調用水
79,	79,	NA,	WF_U_Other,	INT	# 淡水用途別:その他
80,	80,	NA,	WF_U_Total,	INT	# 淡水用途別:合計

```

81, 81, NA, WS_U_Mat,      INT      # 海水用途別:原料用水
82, 82, NA, WS_U_Treat,   INT      # 海水用途別:処理・洗浄用水
83, 83, NA, WS_U_Cool,    INT      # 海水用途別:冷却用水
84, 84, NA, WS_U_Heat,    INT      # 海水用途別:温調用水
85, 85, NA, WS_U_Other,   INT      # 海水用途別:その他
86, 86, NA, WS_U_Total,   INT      # 海水用途別:合計

```

End of Field Definition

なお、# は、それ以降がコメントであることを意味する。cmf2db は、# 以下および空行を無視する。

定義ファイルは、作成するテーブル名を指定することから始める(# 以下および空行は無視されるので、テーブル名 **ind1999** が実質的な 1 行目である)。

次に入力ファイルの情報を記入する。例では、**ind1999** テーブルに **sankou1999.txt**(1999 年産業編甲)と **sanotu1999.txt**(1999 年産業編乙)の 2 つのファイルのデータを格納することを示している。入力ファイルが 2 つ以上ある場合には、それぞれをカンマで区切って記入する。ファイル名は、フルパスで指定する。

続いて、入力ファイルの各レコードに記録されている項目数を記入する。例では、甲調査 86 項目、乙調査に 33 項目記録されていることを示している。これは、入力データを読み込みながら指定された項目数が実際にあるかどうかをチェックするために使われる。もし、指定された項目数と異なるレコードが発見されたときには、何レコード目で検出されたかを出力してプログラムは停止する。

コメント # **Field Definition** 以下が入力するデータに関する情報である。入力ファイルに含まれるデータのうちデータベースに格納したいデータのそれぞれにつき一行(途中で改行しないで)で記述する。フォーマットは、下に示す順序に従ってカンマ区切りで記述していく。

番号	入力ファイル 1 におけるデータの位置	入力ファイル 2 におけるデータの位置	フィールド名	フィールドの型	代入値
----	---------------------	---------------------	--------	---------	-----

最初に記述する番号は、定義ファイルにおける連番で、定義ファイルを記述する際のメタ的なもので、データベースへの登録に関しては何の役割もない。以下、例を見ながら各項目の説明をする。番号 2 の行は、

```

2, 2, 2, Type, INT      # 甲乙区分

```

となっている。これは、入力ファイル 1(sankou1999.txt の 2 番目、および入力ファイル 2(sanotu1999.txt の 2 番目に記録されている甲乙区分をフィールド名 **Type**、フィールドの型 **INT** (整数)で登録するように指示している。同様に、次の行の

```
3, 3, 3, State, CHAR(2) # 都道府県番号
```

は、入力ファイル 1 の 3 番目および入力ファイル 2 の 3 番目に記録されている都道府県番号をフィールド名 **State**、フィールドの型を **CHAR(2)**、すなわち 2 文字の文字列として登録するように指示している。

同じ種類のデータが異なるファイルにおいて同じ位置に記録されているとは限らない。例えば、定義ファイルの 24 番目のエントリーである現金給与総額・合計は甲調査(入力ファイル 1)では 24 番目、乙調査(入力ファイル 2)では 21 番目に記録されている。それをフィールド名 **C_L_Total** の **INT** 型で登録するために次のように指示している。

```
24, 24, 21, C_L_Total, INT # 現金給与総額・合計
```

また、ある種のデータは、全ての入力ファイルに存在するとは限らない。例えば、22 番目のエントリーである現金給与総額・常用労働者は、甲調査では調査されているが、乙調査では調査されていない。このような場合には、記録がない入力ファイルのデータ位置に **NA** を記入する。したがって、22 番目の指示は次のようになる。

```
22, 22, NA, C_L_0, INT # 現金給与総額・常用労働者
```

最後に、一番目のエントリーである、

```
1, 0, 0, Year, INT, 1999 # 調査年
```

について説明しておく必要がある。ここでは、入力ファイルにおけるデータ位置の全てに **0** が置かれ、さらにフィールドの型の後に **1999** と書かれている。cmf2db では、入力ファイルのデータの位置に **0** が書かれた場合には、フィールドの型に続く値を入力ファイルに記録

されているデータの代わりに代入するようになっている。すなわち、INT 型の Year というフィールドには、入力ファイルのデータを用いずに、強制的に 1999 を代入せよということを示している。これは、工業統計の個票データの調査年が、平成 4 年までは和暦で、平成 5 年から西暦で記録されているため、統一して西暦 4 桁で登録するために施した処置である。

1980 年から 1999 年の工業統計個票データに関する定義ファイルは、TASK1 ディレクトリ以下の def ディレクトリにテーブル名に拡張子 def を付して保存してある。例に示した 1999 年産業編の定義ファイルは ind1999.def であり、対応する商品編の定義ファイルは com1999.def となっている。

4.1.4. TASK1-3: 工業統計個票データのデータベースサーバーへの登録

定義ファイルの指示に従って、cmf2db でデータを変換し、Psq1 によって PostgreSQL にデータを登録する。この作業は、TASK1 ディレクトリ以下の CMFWK/task1/entry ディレクトリで実行する。

4.1.4.1. データベースの作成

最初に、PostgreSQL にデータベースを作成する。データベース名は cmf とする。データベースを作成するためには、その権限を持った PostgreSQL ユーザーとして登録しておかなければならない。ユーザー rieti をデータベースの作成できるユーザーとして登録するためには次のようにする。

```
$ su -
Password: root のパスワードを入力する
# su - postgres
$ createuser rieti
Shall the new user be allowed to create database ? (y/n) y
Shall the new user be allowed to create new user ? (y/n) n
CREATE USER
$ exit
# exit
```

そして、データベースを次のように作成する。

```
$ dbcreate cmf
CREATE DATABASE
```

そして、作業ディレクトリに移動する。

```
$ cd CMFWK/task1/entry
```

4.1.4.2. cmf2db の実行

定義ファイル ind1999.def を用いて 1999 年工業統計の個票データ産業編甲乙を登録するには次のようにする。

```
$ ../cmf2db ../def/ind1999.def
```

初期作業では、1980 年から 1999 年までの 20 年分の工業統計個票データを登録しなければならない。したがって、産業編と商品編を合わせて **cmf2db** を 40 回実行しなければならない。そこで、これを一括処理するためのシェル・スクリプト **do_all** を **TASK1** ディレクトリに用意してある。**do_all** を利用して一括登録を行うには次のようにする。

```
$ ../do_all
```

cmf2db は実行ログを、それを実行したディレクトリに残す。ログファイルには、実行開始時間、終了時間、登録レコード数が記録される。ログファイルのファイル名は、テーブル名に拡張子 **.log** をつけたものである。例えば、1999 年の産業編のログファイルは **ind1999.log** で、商品編のログファイルは **com1999.log** である。

cmf2db にはテストモードが用意されている。

```
$ ../cmf2db ../def/ind1999.def test
```

とすればテストモードで実行され、ログファイルに最初のレコードが定義ファイルとの対応で出力されるので、予定したとおりにデータの格納が行われるかどうかの確認をすることができる。

4.1.4.3. Psql による PostgreSQL への格納

cmf2db は、**Psql** によって **PostgreSQL** にデータを格納するための SQL 文を記述したファイルを、その実行ディレクトリに生成する。ファイル名は、テーブル名に拡張子 **.sql** をつ

けたものである。例えば、1999年の産業編は `ind1999.sql` で、商品編は `com1999.sql` である。`ind1999.sql` の内容は次のようになっている。

```
CREATE TABLE ind1999 (Seq_no INTEGER PRIMARY KEY, Year INT, Type INT,
  State CHAR(2), City CHAR(3), Code_Est CHAR(5), Code_Ind CHAR(4),
  Afl INT, Org INT, SuvArea INT, Scale_Lab INT, Scale_Cap INT,
  Scale_Ship INT, Scale_Area INT, Capital INT, L_O_M INT, L_O_F INT,
  L_S_M INT, L_S_F INT, L_Total INT, L_O_Cumsum INT, Shipment INT,
  C_L_O INT, C_L_Other INT, C_L_Total INT, C_Mat INT, C_Fuel INT,
  C_Elec INT, C_Consig INT, Cost INT, Tax INT, K_Land INT,
  K_DepAsset INT, K_Bld INT, K_Mac INT, K_Other INT, R_Land INT,
  R_DepAsset INT, Dep INT, Bld_Inc INT, Bld_Dec INT, A_Land INT,
  A_N_DepAsset INT, A_N_Bld INT, A_N_Mac INT, A_N_Other INT,
  A_U_DepAsset INT, A_U_Bld INT, A_U_Mac INT, A_U_Other INT,
  S_B_Finish INT, S_B_Prg INT, S_B_Mat INT, S_B_Total INT,
  S_E_Finish INT, S_E_Prg INT, S_E_Mat INT, S_E_Total INT,
  Area_Lot INT, Area_Bld INT, Area_Total INT, Land_Farm INT,
  Land_Resid INT, Land_Forest INT, Land_Landfill INT, Land_Other INT,
  Land_Total INT, WF_S_Pub_Ind INT, WF_S_Pub_Sup INT, WF_S_Surf INT,
  WF_S_Well INT, WF_S_Other INT, WF_S_Recycle INT, WF_S_Total INT,
  WF_U_Boiler INT, WF_U_Mat INT, WF_U_Treat INT, WF_U_Cool INT,
  WF_U_Heat INT, WF_U_Other INT, WF_U_Total INT, WS_U_Mat INT,
  WS_U_Treat INT, WS_U_Cool INT, WS_U_Heat INT, WS_U_Other INT,
  WS_U_Total INT);
\copy ind1999 FROM 'ind1999' USING DELIMITERS ','
```

Psql で PostgreSQL に格納するには次のようにする。

```
$ psql -d cmf -f ind1999.sql
```

これをすべてのデータについて一括して実行するためのシェルスクリプト `do_sql` が `TASK1` ディレクトリに用意されている。`do_sql` を実行するには次のようにする。

```
$ ../do_sql
```

さて、これ以降 `CMFWK/task1/data` に保存されている経済産業省から提供されたファイルを直接使うことはない。これらのファイルは大量のディスクスペースを占有し、今や同じ情報が PostgreSQL 上に存在するので、消去してしまってもかまわない。しかし、今後の作

業で何らかのエラーが発生して、TASK1 を繰り返す可能性も考えられるので、`data` ディレクトリを圧縮保存しておく。TASK1 ディレクトリで次を実行する。

```
$ tar cvzf data.tar.gz data/
$ cd data
$ rm *.*
$ cd ..
```

同じことは、`cmf2db` によって生成されたファイル群について当てはまる。これらのファイルも圧縮保存しておく。

```
$ tar cvzf sql.tar.gz entry/*.sql
$ tar cvzf log.tar.gz entry/*.log
$ rm entry/*.sql
$ rm entry/*.log
$ tar cvzf entry-data.tar.gz entry/
$ cd entry
$ rm *
$ cd ..
```

4.1.5. TASK1-4: 常用労働者数と現金給与総額(常用労働者)の調整

1980年～1984年の甲調査では、常用労働者数が男女別に生産労働者と管理事務労働者に分けて調査されていたが、1985年以降は生産労働者と管理事務労働者の区別が廃止され、男女別の常用労働者数のみが調査されている。同様に、現金給与総額についても1980年～1984年では、生産労働者と管理事務労働者別に調査されていたが(男女の区別はない)、1985年以降はこの区別が廃止されている。1980年～1984年のファイルには、常用労働者男子計、常用労働者女子計、現金給与総額常用労働者計が計上されていないので、これらの合計値を計算する。

すなわち、`type = 1` (甲調査)の `year = 1980, 1981, 1982, 1983, 1984` について、次を計算する。

```
l_o_m = l_o_b_m + l_o_w_m
l_o_f = l_o_b_f + l_o_w_f
c_l_o = c_l_o_b + c_l_o_w
```

<code>l_o_m</code>	従業者数・常用労働者・男
<code>l_o_b_m</code>	従業者数・常用生産労働者・男
<code>l_o_w_m</code>	従業者数・常用管理事務労働者・男
<code>l_o_f</code>	従業者数・常用労働者・女
<code>l_o_b_f</code>	従業者数・常用生産労働者・女
<code>l_o_w_f</code>	従業者数・常用管理事務労働者・女
<code>c_l_o</code>	現金給与総額・常用労働者
<code>c_l_o_b</code>	現金給与総額・常用生産労働者
<code>c_l_o_w</code>	現金給与総額・常用事務管理労働者

対応する 1980 年に関する SQL 文は次の通り。

```
UPDATE ind1980 SET l_o_m = l_o_b_m + l_o_w_m WHERE type = 1 ;
UPDATE ind1980 SET l_o_f = l_o_b_f + l_o_w_f WHERE type = 1 ;
UPDATE ind1980 SET c_l_o = c_l_o_b + c_l_o_w WHERE type = 1 ;
```

1980 年～1984 年の UPDATE 文のスク립トが TASK1 ディレクトリの `update-labor.sql` に用意してある。

この計算を行う前に、変更を施す 1980 年～1984 年の産業編のバックアップをとっておく。

```
\copy ind1980 TO /home/rieti/cmf/task1/backup/ind1981
\copy ind1981 TO /home/rieti/cmf/task1/backup/ind1982
\copy ind1982 TO /home/rieti/cmf/task1/backup/ind1983
\copy ind1983 TO /home/rieti/cmf/task1/backup/ind1984
\copy ind1984 TO /home/rieti/cmf/task1/backup/ind1985
```

このスク립トファイルが TASK1 ディレクトリに `bkup-1980-1984.sql` として用意してあるので、TASK1 ディレクトリにおいて、

```
$ psql -d cmf -f bkup-1980-1984.sql
```

とする。そして、次を実行してデータの更新を行う。

```
$ psql -d cmf -f update-labor.sql
```

PostgreSQL で行の更新を行うと、元の情報をテーブルファイルに残し、終了したことをマークたうえで、新しい行を書き加える。したがって、テーブルファイルは更新とともに肥大化する。終了マークを付けられた行を削除するのが **VACUUM** コマンドである。1980年～1984年において、従業者数 30 人以上の事業所は 5 万 5 千以上存在する。これらの事業所につき 3 回の更新を施したことになるので、16 万 5 千行以上の無駄なレコードが各テーブルファイルに残っていることになるので、ここで **VACUUM** を実行しておく。

```
VACUUM ind1980 ;  
VACUUM ind1981 ;  
VACUUM ind1982 ;  
VACUUM ind1983 ;  
VACUUM ind1984 ;
```

これを実行するスクリプトが **TASK1** ディレクトリの **vacuum.sql** にあるので、

```
$ psql -d cmf -f vacuum.sql
```

とする。

4.1.6. TASK1-5: バックアップの作成

TASK1-3 に膨大な時間を要するので、以降の作業におけるエラーの発生に備えてバックアップを作成しておく。この作業は、TASK1-4 のバックアップファイルの作成と全く同じである。バックアップを **CMFWK/TASK1/backup** ディレクトリに作るスクリプト **backup-all.sql** が **TASK1** ディレクトリに用意してあるので、**TASK1** ディレクトリでこれを実行する。また、ディスクスペースを節約するために圧縮保存することにする。

```
$ psql -d cmf -f backup-all.sql
$ tar cvzf backup.tar.gz backup/
$ cd backup
$ rm *
$ cd ..
```

4.1.7. TASK1-6: 工業統計個票データの消去

これ以降の作業は、PostgreSQL に格納したデータにアクセスして行うので、経済産業省から提供されたオリジナルファイルと、TASK1-3 で生成した中間ファイルをディスクから消去する。TASK1 ディレクトリで、以下を実行する。

```
$ rm data.tar.gz
$ cd entry
$ rm entry-data.tar.gz
```

4.2. TASK 2: 事業所コンバーターの作成

事業所コンバーターとは、工業統計個票データにおける事業所コード(都道府県番号 2 桁+市区町村番号 3 桁+事業所番号 5 桁)の調査間対応表である。例えば、1980 年工業統計個票データにおいて 0110100005 番という事業所が、1995 年工業統計個票データの 0110100001 番の事業所であることを示す一連の対応表をコンバーターと呼んでいる。

工業統計の事業所コードは、全体で 10 桁のコードで識別されているが、事業所コンバーターを整備するためには、下 8 桁の市区町村番号 (3 桁) と事業所番号 (5 桁) の同定化が必要になる。これらは特定の研究目的で行っており、基盤的な業務となっておらず、経済産業省調査統計部や、通商産業研究所 (現在の独立行政法人経済産業研究所の前身) 等で度々整備を行ってきたが、それらの成果は必ずしも満足のいくものではなかった。その要因として、①パネルデータ作成に必要な情報の制約、②作業体制、作業期間、③コンピューターの容量やソフト面の制約、等である。本研究プロジェクトでは、過去の成果の課題等を検討しつつ、その上にたって、人的な作業で直近の事業所コードをベースに、過去に遡及して整備する方法をとった。詳しくは、次節でその作成方法について概略を述べる。

現行システムでは事業所コンバーターは、原則として各都道府県について一つの Excel ファイルとして整備され提供されているが、東京、愛知、大阪など事業所数の大きい都道府県については複数の Excel ファイルに分割されて提供される。

事業所コンバーター作成の作業フローは以下のとおりである。

- **[TASK 2-1] テキストファイルへの変換。** Excel で提供される事業所コンバーターをテ

キストファイルに変換する。各 Excel ファイルは、`xls2txt.xls` に記録されている VBA マクロプログラムによって一括してテキストファイルに変換される。

- **[TASK 2-2] PostgreSQL への登録。** テキストファイルに変換された事業所コンバーターを PostgreSQL データベースサーバーに登録する。PostgreSQL コマンドを記録したコマンドファイル、`Jconv.sql` を `Psql` で実行することによって事業所コンバーターを PostgreSQL に登録する。
- **[TASK 2-3] 工業統計個票データと事業所コンバーターのマッチング。** 工業統計の個票データ個票データに存在する事業所が、事業所コンバーターに存在しているかを確認する。`match-cnvr.R` という R プログラムが事業所コンバーターと工業統計の個票データのマッチングを行ない、`sum-match-cnvr.R` がマッチング結果を Excel 等で読み込み可能なテキストファイルで出力する。
- **[TASK 2-4] 事業所コンバーターの修正。** TASK 2-3 において、工業統計の個票データに存在して事業所コンバーターに存在しないコードが見つかった場合には、該当する Excel ファイルを修正し、事業所コンバーターに漏れがなくなるまで TASK 2-1～TASK 2-3 を繰り返す。
- **[TASK 2-5] TASK 2-4 の結果による事業所コンバーターの修正。** 事業所コンバーターは、その 100%が過去の事業所番号付替記録に基づいて作成されているわけではないので、TASK 2-4 において従業者数、資本金、敷地面積、期末在庫残高一期首在庫残高等の情報を用いながら、同一事業所が接続されているかどうかのチェックを行なわなければならない。チェックの方法については TASK 2-4 で後述するが、このチェックによってコンバーターを修正する必要があるときには、該当する Excel ファイルを修正し、TASK 2-1～TASK 2-3 をもう一度実行し、Excel ファイル、テキストファイル、データベース(PostgreSQL)が全て同じ情報を持つようにする。

4.2.1. 作業 1: テキストファイルへの変換

Excel で提供される事業所コンバーターをテキストファイルに変換する。各 Excel ファイルは、`xls2txt.xls` に記録されている VBA マクロプログラムによって一括してテキストファイルに変換される。

`xls2txt` は以下のように実行する。

1. 最新の事業所コンバーターが保存されているフォルダーに `xls2txt.xls` をコピーする。以下このフォルダーを作業フォルダーと呼ぶ。
2. `xls2txt.xls` を開く。このとき、「マクロを無効にする」か「有効にする」か問われるので、「マクロを有効にする」を選択する(図 2)。

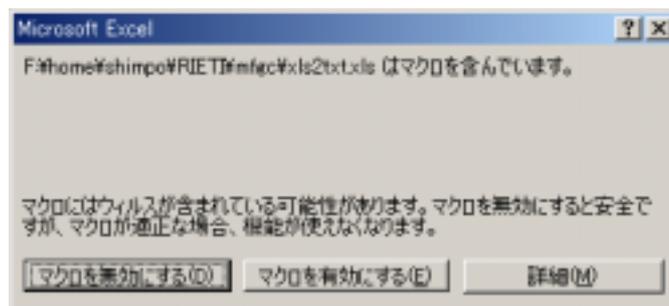


図 2: マクロを有効にする

もし、マクロの無効/有効を問われない場合は、セキュリティレベルがマクロを実効できないように「高」に設定されているので、次のようにセキュリティレベルを変更する。「ツール(T)」→「マクロ(M)」→「セキュリティ(S)」をクリックしセキュリティレベルの設定ウインドウを表示して、「セキュリティレベル(S)」で「中(M)」を選択し、「OK」をクリックする(図 3)。そして、一端 xls2txt.xls を閉じ、再び xls2txt.xls を開くとマクロの無効/有効を問われるので、「マクロを有効にする」を選択する。

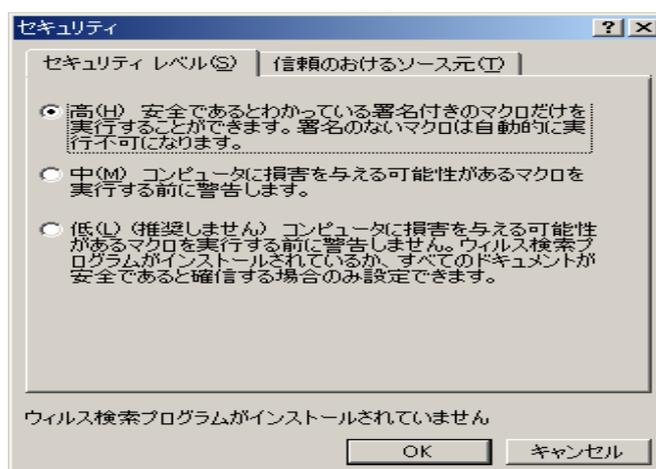


図 3: セキュリティレベルの変更

3. xls2txt.xls を開くと、「定義」というワークシートがあるので、必要に応じて所定の項目を記入する(図 4)。

開始年	最終年	ファイル名	シート名	出力ファイル
1976	1978	.\form01.xls	01 北海道	.\form01el
		.\form02.xls	02 青森	
		.\form03.xls	03 岩手	
		.\form04.xls	04 宮城	
		.\form05.xls	05 秋田	
		.\form06.xls	06 山形	
		.\form07.xls	07 福島	
		.\form08.xls	08 茨城	
		.\form09.xls	09 栃木	
		.\form10.xls	10 群馬	
		.\form11.xls	11 埼玉	
		.\form12.xls	12 千葉	
		.\form13-105.xls	13 東京都	
		.\form13-112.xls	13 東京都	
		.\form14.xls	14 神奈川	
		.\form15.xls	15 新潟	
		.\form16.xls	16 富山	
		.\form17.xls	17 石川	
		.\form18.xls	18 福井	
		.\form19.xls	19 山梨	
		.\form20.xls	20 長野	
		.\form21.xls	21 岐阜	
		.\form22.xls	22 静岡	
		.\form23-105.xls	23 愛知県	
		.\form23-206.xls	23 愛知県	
		.\form24.xls	24 三重	
		.\form25.xls	25 滋賀	
		.\form26.xls	26 京都	
		.\form27-105.xls	27 大阪府	
		.\form27-206.xls	27 大阪府	
		.\form28.xls	28 和歌山	
		.\form29.xls	29 奈良	

図 4: 「定義」ワークシート

入力項目は以下のとおりである。

- A) A 列「開始年」および B 列「最終年」の 2 行目に、何年から何年のコンバーターを抽出するかを記入する。
 - B) E 列「出力ファイル」の 2 行目に変化されるテキストファイルのファイル名を記入する。テキストファイルに変換されたコンバーターは、作業フォルダーに指定した名前で保存される。
 - C) C 列「ファイル名」、D 列「シート名」には、都道府県単位で提供される Excel の事業所コンバーター名とそれぞれの Excel ファイルにおけるコンバーターのワークシート名を記入する。ファイル名は拡張子(.xls)も含めて記入する。複数の Excel ファイルに分割されて提供される都道府県に関しては、その全てを記入しなければならない(図中 14, 15 行目の東京, 25, 26 行目の愛知, 30, 31 行目の大阪を参照)。
4. 「ツール(T)」から「マクロ(M)」を選択し、「マクロ(M)」で「マクロ(M)」を選択すると「マクロ」のサブウィンドウが現れるので、xls2txt を選択し、「実行(R)」をクリックすれば、コンバータの変換作業が開始される(図 5)。

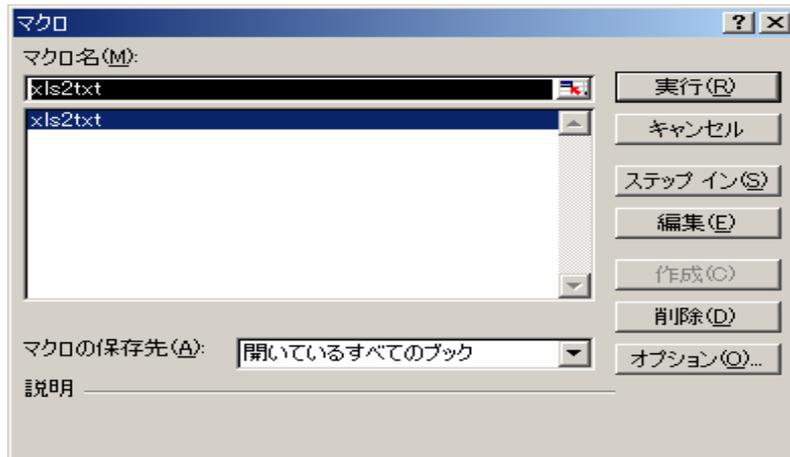


図 5:マクロの実行と編集

5. 「定義」シートの「ファイル名」列に記入した Excel ファイルが次々に開かれていくので、最後のファイルが開かれ(図 4 の「定義」シートの例では Jconv47.xls), 実行中を示す砂時計が消えたら Excel を終了する。
6. テキストファイルには、連番(8 桁), 都道府県番号(2 桁), 事業所コード(開始年), ..., 事業所コード(最終年)がカンマで区切られた 1 レコードとして保存される。ここで事業所コードは市区町村番号(3 桁)と事業所番号(5 桁)を結合した 8 桁のコードである。また、Excel ファイルにおいて事業所コードが空白, ?, w, ww となっているセルについては NULL 値として¥N が代入される。

4.2.2. プログラム使用上の注意

1. 「定義」シート第 1 行目で黄色く塗りつぶされている「開始年」, 「最終年」, 「ファイル名」, 「シート名」, 「出力ファイル」の各セルには名前が定義されており, それぞれの名前をマクロプログラムが参照しているので変更しないこと。
2. 「ファイル名」および「シート名」に記入できるファイルの数は, 最大で 100 個に指定されている。現状ではあまり考えられないケースではあるが, 100 を超過するような場合には, プログラムを修正する必要がある。直接修正する場合には, 「ツール(T)」→「マクロ(M)」→「マクロ(M)」でマクロサブウインドウを表示させ, 「編集(E)」をクリックするとプログラムが表示されるので(図 6), 10 行目および 11 行目の

```
Dim infile(100) As String
```

```
Dim state(100) As String
```

の 100 を適当な数値に置き換えればよい。

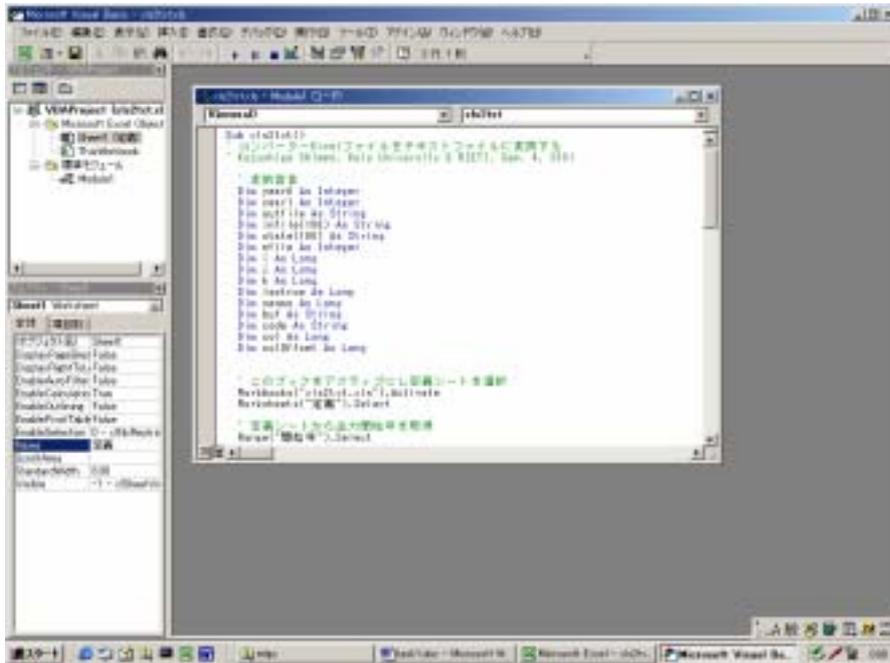


図 6: マクロプログラムの編集

3. 現状では連番を 8 桁(00000001~99999999)としているが、これを変更する必要がある場合には、プログラムを直接修正しなければならない。プログラムを直接修正するときには、上記手順に従ってプログラムを編集可能な状態にして(図 6)、75 行目の

```
buf = Format(seqno, "00000000") & "," & Cells(i,2)
```

の"00000000"を適当に修正すればよい。

4. 現状では、事業所コードが空白, ?, w, ww の場合には NULL 値を代入しているが、その他にも NULL 値に設定すべきものがあるときには、プログラムを直接修正しなければならない。プログラムを直接修正する場合には、上記手順に従ってプログラムを編集可能な状態にして(図 6)、85 行目の

```
If code = "" Or code = "w" Or code = "ww" Or code = "?" Then
```

を修正すればよい。

5. また、NULL 値には¥N を代入しているがこれを変更する場合には、86 行目の

```
code = "¥N"
```

を変更すればよい。

6. xls2txt は現行の事業所コンバーターの記録型式に基づいて書かれている。特に、各ワークシートにおいて

- A) 第 2 列目が都道府県番号である。
- B) 第 3 列目から事業所コード(上 3 桁市区町村番号, 下 5 桁事業所番号)が 1 列おきに記録されている。
- C) 第 3 列が 1976 年の事業所コードになっていることが、マクロプログラムには重要

である。これらの点に変更がある場合には、プログラムの多くの部分を書き換える必要がある。また、事業所コンバーターの作成作業環境が Excel から変更された場合には、このプログラムは当然のことながら使用できないので、新たなプログラムを開発する必要があることは言うまでもない。

4.2.3. TASK 2-2: PostgreSQL への登録

TASK 2-1 でテキストファイルに変換された事業所コンバーターを PostgreSQL データベースサーバーの cmf データベースに、テーブル名 jconv として登録する。

本作業は、TASK2 ディレクトリで行う。本作業に先立って、TASK 2-1 で作成した事業コンバーターのテキストファイル Jconv. dat を TASK2 ディレクトリに移しておかねばならない。PostgreSQL への登録は、コマンドファイル Jconv. sql を実行してすることによって行う。

```
> psql -d cmf -f Jconv.sql ↵
```

とすれば事業所コンバーターが PostgreSQL に登録される。

PostgreSQL の cmf データベースに登録された事業所コンバーター、jconv テーブルは、**エラー! 参照元が見つかりません。**のようになっている。各レコードは、同じ事業所に関する情報を含んでいる。第 1 フィールド(id)は、その事業所固有の 8 桁の識別番号で、これをプライマリー・キーに設定してある。第 2 フィールド(State)は、2 桁の都道府県番号(**エラー! 参照元が見つかりません。**参照)で、都道府県単位での検索が高速になるようにインデックスを設定している。第 3 フィールド以降が各年の事業所コードで、y に続けて年次からなるフィールド名が付けられている。事業所コードは、上 2 桁が市区町村番号、下 5 桁が事業所番号を結合した 8 桁のコードである。

表 5: 事業所コンバーター(jconv テーブル)

フィールド型	CHAR(8) 主キー	CHAR(2) Index	CHAR(8)	...	CHAR(8)
フィールド名	id	State	y1980	...	y1999
レコード 1	00000001	01	10100015	...	10100293
レコード 2	00000002	01	10100746	...	10100294
...

4.2.4. TASK 2-3: 工業統計個票データと事業所コンバーターのマッチング

事業所コンバーターは、その 100%が過去の事業所番号付替記録に基づいて作成されるのではなく、人的作業によって事業所番号を接続しなくてはならない部分が少なくない。そ

れが故に、工業統計の個票データ個票データに存在する事業所がコンバーターから漏れている可能性があり、その漏れをチェックする必要がある。ここでは、工業統計の個票データの産業編に記録されている事業所コードと事業所コンバーターの事業所コードのマッチングを行い、マッチング結果のサマリー表の作成とコンバーターに存在しない事業所コードの出力を行う。

事業所コンバーターに記録されている事業所が、工業統計の個票データに確かに存在しているか確認する。`match-cnvr.R` という R プログラムが事業所コンバーターと工業統計の個票データとのマッチングを行ない、`sum-match-cnvr.R` がマッチング結果を Excel 等で読み込み可能なテキストファイルで出力する。

4.3. 工業統計の事業所コンバーターの作成方法

この節では、工業統計の各年の事業所コードを使って、事業所コンバーターを作成する方法について説明する。ただし最近経済産業省で新しいコンピューター処理システムに移行したことにより、この方法は更されている模様。その実態把握は今後の研究課題である。

工業統計の事業所コードは、主として各年の調査対象事業所を識別するための機能を第1番に考え、作業が行われていることから、必ずしもパネルデータ作成の機能を持っているものの、その機能が発揮できる状況になっていない。幸い、従来コンピューターの機能の制約と作業の繁雑さから、事業所コードの切り替えは5年毎に行われてきた。国勢調査が西暦末日の0、5年に実施され、その実施に伴って国勢調査の調査区の設定替えが行われ、その設定替えをベースに事業所調査の基本調査区や工業調査の調査区¹⁸の設定替えが行われ、それに伴って工業統計の事業所コードの付け替えが行われ、その後原則として5年間固定されている。

既に第3章でも簡単に述べたように、工業統計の事業所コードは10桁から構成されており、その構成は都道府県番号2桁、市区町村番号3桁、事業所番号5桁からなる。そのうち、都道府県番号(2桁)はこれまで変更されたことが無いことから、実際は下8桁の市区町村番号(3桁)と事業所番号(5桁)の変更状況の把握になる。なお、最近では市区町村の越県合併が行われるようになってきていることから、今後都道府県コードの変更も発生する可能性もある。

4.3.1. 工業統計の事業所番号付け替え年

工業統計の事業所コンバーターの作成は、すでに何回か試みられている。その都度精度は向上しているものの、依然としてパネルデータ分析の都度コンバータ作成作業から始め

¹⁸ 当初、事業所統計調査の基本調査区と工業統計調査の調査区は必ずしも同一ではなく、工業統計調査では独自に調査区を設定していた。これが両統計調査データのリンケージ作業を進める上で大きなネックとなっていた。そこで、両統計調査の調査区設定の統一化が図られ、工業統計調査側においても、工業調査区から基本調査区をベースとした工業調査区の設定が行われることとなった。

る必要がある状態が続いている¹⁹。本来、パネルデータ作成には、個々の事業所の名所、住所等の属性情報が必要であるが、これが過去の調査では電子化されておらず、殆どが破棄処分されていることによる。

近年の工業統計の事業所コードは、1970年、1975年、1980年、1986年、1991年、1997年に付け替えが行われている（次回は2002年に行われる予定）が、1970年、1975年、1980年の新旧事業所番号の対照表が保存されていない。書き換え、あるいは誤って破棄された可能性が大である。このため、1980年代始めの2年間のコンバーター作成が進まない状況にある。

また、これ以外に工業統計は毎年調査を実施し、電算処理を行っていることから、各年の「工業統計調査事業所マイターファイル」が作成されているが、1984年以前及び1987年、1988年、1992年の同マイターファイルも不明となっている。

経済産業省経済産業政策局調査統計部(商工統計課、旧：工業統計課)提供のマスタヘファイルは、以下のとおりである。

「85conv.txt.」、「86conv.txt.」、「89conv.txt.」、「90conv.txt.」、「91conv.txt.」、「93conv.txt.」、「94conv.txt.」、「95conv.txt.」、「96conv.txt.」、「97conv.txt.」の10年分のマスターファイルがあり、この研究では、このうち「97conv.txt.」(設定)をベースに作成作業を行った。

4.3.2. 工業統計のコードコンバータ作成作業

これら現存するマスターファイルのうち、直近の「97conv.txt.」(設定)をベースに、過去に遡及してコードコンバータを作成するとともに、同時に将来の1998年、1999年に延長する作業を行った。

事業所のコードコンバータの作成は、二段階に分けて作業を実施した。最初は市区町村番号(3桁)の変更状況を把握し、そのコードコンバータを作成することである。続いて市区町村番号の変更状況を考慮した上で、事業所番号(5桁)の変更状況を追跡する作業である。

4.3.2.1. 市区町村番号の変更及びそのコードコンバータの作成

¹⁹ これは、統計調査を実施している経済産業省(経済産業政策局調査統計部)としてパネルデータの作成を基盤的業務として位置づけていないことによる。そのため、各研究プロジェクト毎に、研究に必要な範囲でパネルデータの作成作業を行っているためでもある。最大の問題は、パネルデータ作成には、個々の事業所の名所、住所等の属性情報が必要であるが、工業統計ではこの漢字情報の電子化は、財政的な問題もあり、最近である。それ以外にも、パネルデータ作成の意識の欠如から、貴重な情報が不明なことから、どこまで時間と労力とお金を掛けて作業を行うかで、そのコードコンバータの精度に依然として差が生じてしまう。しかし、これら情報や知識は、散逸しており、ますます整備が困難になってきている状況にある。筆者らは、現在作成できるコンバータはこれが限界で、後は本文中でも述べられているように、実際個票データを使った事業所のデータの連続性の検証作業を地道に積み上げていくしか方法がないものと判断している。

まず、最初は、1976年（昭和51年）以降の市区町村番号の変更状況の把握、整理である。市区町村番号は、総務省（統計局統計基準部）で作成している「統計に用いる標準地域コード」（全国、都道府県、市区町村、郡・支庁）に1971年（昭和46年）4月1日（官報告示）以降の改正経過がとりまとめられている。原則的には、この改正に対応しているが、工業統計の調査期日（12月31日現在）の関係で切り替え時期が一部地域でずれている年次があり、これのチェックは最終的には、個票データとの突き合わせを繰り返すことによってチェックした。

4.3.2.2. 事業所番号の同定作業

引き続き、上5桁（都道府県番号2桁+市区町村番号3桁）をキーコードに、下5桁の事業所番号について追跡し、同定作業を行う必要がある。

工業統計は、市区町村単位で調査が行われるため、事業所番号は市区町村一連番号で設定されており、その番号は原則として5年間固定されている。

経済産業省（産業政策局調査統計部産業統計室）では、原則として西暦末尾の0、5年をベースに翌年の調査で「工業調査事業所番号」の設定替えを行っているが、事業所番号設定後、事業所の廃業や移転（他の都道府県や市区町村への移動）、他産業への転業により、対象外となる事業所が発生するが、それら事業所の番号は欠番処理を行い、新設事業所は最終事業所の次の番号から追加される（新設事業所には、新設の事業所や他地域から移転してきた事業所、他産業から転業してきた事業所が含まれる）。このため、市区町村の行政区域が変更にならない限り、事業所番号の切り替え年のコードと切り替え年の前年のコードがあれば、その間の事業所コンバーターを整備することができる。問題は、事業所番号の切り替え年の前後の年である。

例えば1995年と1996年調査を事例に示すと、次のようなコード設定替えを行っている。

旧事業所コード ……………1995年調査、1991年設定したコード体系、1991年から1995年まで使用（市区町村の合併、政令都市の出現、行政区域の変更）

旧→新事業所対応コード…1995年調査、1996年調査用のために新コードにコンバートしたもの（市区町村一連番号の設定処理）

新事業所コード ……………1996年調査結果に基づく新コード体系、1996年から2000年まで使用（新規事業所の追加等）パネルデータとして作成可能なのは、1980年代以降である。

現在、1982年まで事業所コンバータを遡及する作業は終了したが、1980年と1981年については、事業所番号の設定替えの時期に当たり、新旧コードマスターがなく、両年に存続する事業所のデータについてのみ遡及できる厳しい状況にある。

4.4. 工業統計の産業分類・品目分類のコードコンバータ作成作業

工業統計は、事業所コードコンバータの他に産業別の集計を行うために、産業分類（4桁）と商品（品目）分類（6桁）二つの分類体系があり、いずれも『日本標準産業分類』によっている。この二つの分類体系は基本分類で対応しており、商品（品目）分類（6桁）の上4桁が産業分類に対応している。

4.4.1. 産業分類改訂への対応

1980年～1999年の工業統計で、日本標準産業分類は2回改訂されている。さらに最近では、2002年に大幅な改正が行われている。

- ①1976年（昭和51年）5月改訂 → 1977年（昭和52年）1月適用
- ②1984年（昭和59年）1月改訂 → 1985年（昭和60年）4月適用
- ③1993年（平成5年）10月改訂 → 1994年（平成6年）4月適用
- ④2002年（平成14年）3月改訂 → 2002年（平成14年）10月適用

1980年調査～1984年調査	:	1977年（昭和52年）分類
1985年調査～1993年調査	:	1985年（昭和60年）分類
1994年調査～1999年調査	:	1994年（平成6年）分類 （今後）
2000年調査～2001年調査	:	同上
2002年調査～	:	2002年（平成14年）分類

このうちで、大規模な産業分類の改正は、1985年（昭和60年）分類の改正と、1994年（平成6年）分類の改正である。1985年（昭和60年）の分類改正は、中分類（2桁分類）レベルの改正で、「食料品製造業」を「食料品製造業」と「飲料・たばこ・飼料製造業」の分割したこと、「その他製造業」のうち、中分類（3桁）「プラスチック製造業（別掲を除く）」を中分類（2桁）「プラスチック製造業（別掲を除く）」として創設したこと、等が大きな改正で、コードコンバートの作成が比較的やりやすい。

しかし、1994年（平成6年）の分類改正は、「繊維工業」と「衣服、その他の繊維繊維製品製造業」の中分類間で分類の組み替えがあり、特に繊維、中でもニット関係の分類を中心に大幅な変更があり、3桁分類、4桁分類のコードコンバート作成は技術的にも困難な分類が多い。このため、統計データの利用面も考慮した上で、その利用目的に応じてコードコンバータを作成する必要がある。

4.4.2. 商品分類改訂への対応

工業統計の商品分類は、産業分類及び製造品分類並びに賃加工品分類と組み合わせて構

成されている。この他に「その他の分類」がある。

①製造品分類

製造品とは、事業所が所有する原材料によって製造するものをいい、原材料を他に支給して製造させた委託生産品を含む。ただし、仕入れてそのまま販売するものは含まれない。

これら製造品を区別するための分類番号は6桁で、原則として上4桁は、それぞれの製造品を主たる生産物とする産業細分類（4桁分類）の番号で、5桁目、6桁目の数字はそれぞれの製造品を区別するための番号である。また、製造品のうち「その他の○○○」とするものは、6桁目に「9」の数字を使っている。

②賃加工品分類

賃加工品とは、他の企業の所有する原材料又は製品に加工して引き渡したものをいい、そま加工品の分類番号は、製造品の番号と同様に6桁とし、5桁目に「9」の数字を用いて、製造品の分類番号と区別している。

③その他の分類

①、②以外に冷蔵保管料、新聞・雑誌広告料金、販売電力料金、「製造工程からでたくず、廃物」、修理料収入の分類番号がある。

・冷蔵保管料

(平成6年改正分類品目番号 634100, 平成14年改正分類品目番号 6044100)

冷蔵保管料が出荷額の50%を超える場合は、倉庫業の対象となり、対象外となる。

・新聞・雑誌広告料金

(平成6年改正分類品目番号 691100~692100)

ただし、広告収入のみの場合は、サービス業となり対象とならない。なお、2002年(平成14年)調査から「出版業」が大分類「F・製造業」から大分類「H・情報通信業」に分類替えになったことに伴い、廃止されている。

・販売電力料金

(平成6年改正分類品目番号 626400 など下4桁が6400のもの)

(平成14年改正分類品目番号 596400 など下4桁が6400のもの)

自家発電の余剰電力の販売電力料など。

・製造工程からでたくず、廃物

・品目が特掲れているもの (6桁番号があるもの)

清酒かす、精米かす、精麦かす、製材くず、鉄くず、非鉄金属くずなど

・品目が特掲れていないもの (平成6年改正分類品目番号 626600 など下4桁が6600のもの, 平成14年改正分類品目番号 596600 など下4桁が6600のもの)

・修理料収入（品目番号 880000）

他の物を修理して受け取った，又は受け取るべき修理料。ただし，事業所全体の修理料収入額が出荷額等の2分の1を超える場合はサービス業で対象外となる。

（参考）品目調査の計測単位

工業統計では，金額のほか，主要な製造品の数量を調べている。その計測単位は，下表のとおり。

工業統計の品目調査の計測単位（数量単位）

略称	使用単位名	略称	使用単位名
km	キロメートル	Kl	キロリットル
g	グラム	mm換算箱	ミリメートル換算箱
kg	キログラム	cmプライ	センチメートルプライ
t	トン	導体 t	導体トン
m ²	平方メートル	総 t	総トン
m ³	立方メートル	マッチ t	マッチトン

5. 最後に

われわれの研究グループでは，これまで産業構造データベースの設計および開発を進めてきたが，本稿は工業統計のパネルデータ化に関する作業の実際を，研究所における今後の作業継続のための防備録としてまとめたものである。

パネルデータによる分析手法そのものは確立しているが，パネルデータは分析目的に応じて作成する必要がある。工業統計のマイクロデータを使って計量経済的分析を行うためには，まず，データ整備が必要であるが，わが国においては工業統計等の指定統計の個票利用については，統計法上の規定があり，特定の目的のために統計目的以外の利用申請を行いそれが認められると利用することが出来るが²⁰，利用に関しては厳しい制約がある。

今回の研究では，データ整備の範囲を，工業統計甲調査(従業者 30 人以上の事業所)だけではなく，工業統計乙調査(従業者 29 人以下)にも広げて，パネルデータ作成を試みたものである。このため，過去の事業所コードコンバータの整備方法について検討し，入手することが出来なかった年次の事業所コードを工業統計における事業所コードの設定方法に準じて再現する方法をとって整備した。これは殆ど人手による作業になり，時間を要するこ

²⁰ 特定の使用目的のため，総務省(又は実施官庁)に統計目的外の利用申請を行い，承認されると特定の期間データを利用することが出来るが，承認するか否かの判断等の運用は官庁によって異なっていることがある。従前は，私立大学の研究者は利用出来なかったが，最近では利用することが出来るようになっている。

とになってしまった。1990年代の事業所コードコンバータを整備し、次いでそのコードをベースに過去の調査年次に順次遡及して事業所コードを補完し、1980年代の事業所コードコンバータを整備した。さらに、その事業所コードコンバータと個票データの事業所コードとを突合せることによって、事業所コードコンバータの欠落コードをさがしつつ、各調査年の事業所コードコンバータを整備していった。しかしながら、現時点では、1982年まで遡及して整備するのが限界である。1970年以降2000年の間に、1970(昭和45年)年、1975(昭和50年)年、1980(昭和55年)年、1991(平成3年)年、1997(平成9年)年5回の事業所番号の付け替えが行われている。更に、2002(平成14年)年にも事業所番号の付け替えが行われることになっている。事業所番号の付け替え時には、既に述べたように新旧事業所番号の対照表(工業統計で言うところの新旧事業所マスター)が必ず整備されるいるが、現在では、1970(昭和45年)年、1975(昭和50年)年、1980(昭和55年)年の新旧事業所番号の対照表は保存されていない。このため、この期間のパネルデータを作成するのは極めて困難であるといわざるを得ない。工業統計は歴史も古く、調査規模も大きいことから、パネルデータの整備には時間と労力と費用が必要になる。さらに、取り扱うデータ量の多さから、データベースの知識も必要になる。

工業統計のパネルデータの活用意義は大きく、特定の研究目的のためではなく、本来的には基盤的なデータベースとして、継続的に整備されるべきものといえるが、そのためにも研究者と調査実施主体との協力体制を整備する必要があるといえよう。

最後に、パネルデータを活用した実証分析を通じて、統計の整備・改善の一助となれば幸いである。

付録 データベースの変数定義

表 6: 工業統計テーブル変数表(産業編: ind1980~ind1999)

No.	調査項目	フィールド名	フィールド型	単位	データの有無			
					1980年 ～ 1984年		1985年 ～ 1999年	
					甲	乙	甲	乙
1	調査年度	Year	INT	西暦4桁	○	○	○	○
2	甲乙区分	Type	INT	表7	○	○	○	○
3	都道府県番号	State	CHAR(2)	表8	○	○	○	○
4	市区町村番号	City	CHAR(3)		○	○	○	○
5	事業所番号	Code_Est	CHAR(5)		○	○	○	○
6	産業分類	Code_Ind	INT		○	○	○	○
7	他事業所有無	Afl	INT	表9	○	○	○	○
8	経営組織	Org	INT	表10	○	○	○	○
9	基本調査区	SuvArea	INT		○	○	○	○
10	従業者規模	Scale_Lab	INT	表11	○	○	○	○
11	資本金規模	Scale_Cap	INT	表12	○	○	○	○
12	出荷規模	Scale_Ship	INT		○	○	○	○
13	敷地規模	Scale_Area	INT	表13	○	○	○	○
14	資本金又は出資金額	Capital	INT	万円	○	○	○	○
15	従業者数・常用労働者・男	L_O_M	INT	人		○	○	○
16	従業者数・常用労働者・女	L_O_F	INT	人		○	○	○
17	従業者数・常用生産・男	L_O_B_M	INT	人	○			
18	従業者数・常用生産・女	L_O_B_F	INT	人	○			
19	従業者数・常用管理事務・男	L_O_W_M	INT	人	○			
20	従業者数・常用管理事務・女	L_O_W_F	INT	人	○			
21	従業者数・個人事業主・男	L_S_M	INT	人	○	○	○	○
22	従業者数・個人事業主・女	L_S_F	INT	人	○	○	○	○
23	従業者数・合計	L_Total	INT	人	○	○	○	○
24	延常用従業者数	L_O_Cumsum	INT	人	○		○	
25	出荷額・合計	Shipment	INT	万円	○	○	○	○

表 6：工業統計テーブル変数表(産業編：ind1980～ind1999) (続き)

No.	調査項目	フィールド名	フィールド型	単位	データの有無			
					1980年 ～ 1984年		1985年 ～ 1999年	
					甲	乙	甲	乙
26	現金給与総額・常用労働者	C_L_0	INT	万円			○	
27	現金給与総額・常用生産労働者	C_L_0_B	INT	万円	○			
28	現金給与総額・常用管理事務労働者	C_L_0_W	INT	万円	○			
29	現金給与総額・その他	C_L_0other	INT	万円	○		○	
30	現金給与総額・合計	C_L_Total	INT	万円	○	○	○	○
31	原材料使用額	C_Mat	INT	万円	○		○	
32	燃料使用額	C_Fuel	INT	万円	○		○	
33	電力使用額	C_Elec	INT	万円	○		○	
34	委託生産費	C_Consig	INT	万円	○		○	
35	原材料・燃料・電力・委託生産費合計	Cost	INT	万円	○	○	○	○
36	内国消費税	Tax	INT	万円	○	○	○	○
37	有形固定資産・年初現在高・土地	K_Land	INT	万円	○	○	○	○
38	有形固定資産・年初現在高・有形固定資産計	K_DepAsset	INT	万円	○	○	○	○
39	有形固定資産・年初現在高・建物・構築物	K_Bld	INT	万円	○		○	
40	有形固定資産・年初現在高・機械・装置	K_Mac	INT	万円	○		○	
41	有形固定資産・年初現在高・その他	K_Other	INT	万円	○		○	
42	有形固定資産・除却額・土地	R_Land	INT	万円	○	○	○	○
43	有形固定資産・除却額・有形固定資産計	R_DepAsset	INT	万円	○	○	○	○
44	有形固定資産・減価償却額	Dep	INT	万円	○	○	○	○
45	有形固定資産・建設仮勘定増	Bld_Inc	INT	万円	○		○	
46	有形固定資産・建設仮勘定減	Bld_Dec	INT	万円	○		○	
47	有形固定資産取得額・土地	A_Land	INT	万円	○	○	○	
48	有形固定資産取得額・新規・有形固定資産計	A_N_DepAsset	INT	万円	○	○	○	○

表 6：工業統計テーブル変数表(産業編：ind1980～ind1999) (続き)

No.	調査項目	フィールド名	フィールド型	単位	データの有無			
					1980年 ～ 1984年		1985年 ～ 1999年	
					甲	乙	甲	乙
49	有形固定資産取得額・新規・建物・構築物	A_N_Bld	INT	万円	○		○	
50	有形固定資産取得額・新規・機械・装置	A_N_Mac	INT	万円	○		○	
51	有形固定資産取得額・新規・その他	A_N_Other	INT	万円	○		○	
52	有形固定資産取得額・中古・有形固定資産計	A_U_DepAsset	INT	万円	○	○	○	○
53	有形固定資産取得額・中古・建物・構築物	A_U_Bld	INT	万円	○		○	
54	有形固定資産取得額・中古・機械・装置	A_U_Mac	INT	万円	○		○	
55	有形固定資産取得額・中古・その他	A_U_Other	INT	万円	○		○	
56	在庫額・年初・製造品	S_B_Finish	INT	万円	○		○	
57	在庫額・年初・半製品・仕掛品	S_B_Prg	INT	万円	○		○	
58	在庫額・年初・原材料・燃料	S_B_Mat	INT	万円	○		○	
59	在庫額・年初・合計	S_B_Total	INT	万円	○	○	○	○
60	在庫額・年末・製造品	S_E_Finish	INT	万円	○		○	
61	在庫額・年末・半製品・仕掛品	S_E_Prg	INT	万円	○		○	
62	在庫額・年末・原材料・燃料	S_E_Mat	INT	万円	○		○	
63	在庫額・年末・合計	S_E_Total	INT	万円	○	○	○	○
64	敷地・建築面積・敷地面積	Area_Lot	INT	m ²	○		○	
65	敷地・建築面積・建築面積	Area_Bld	INT	m ²	○		○	
66	敷地・建築面積・延べ建築面積	Area_Total	INT	m ²	○		○	
67	地面別面積・田畑	Land_Farm	INT	m ²	○		○	
68	地面別面積・宅地	Land_Resid	INT	m ²	○		○	
69	地面別面積・山林・原野	Land_Forest	INT	m ²	○		○	
70	地面別面積・埋立地	Land_Landfill	INT	m ²	○		○	
71	地面別面積・その他	Land_Other	INT	m ²	○		○	
72	地面別面積・合計	Land_Total	INT	m ²	○		○	

表 6：工業統計テーブル変数表(産業編: ind1980~ind1999) (続き)

No.	調査項目	フィールド名	フィールド型	単位	データの有無			
					1980年 ～ 1984年		1985年 ～ 1999年	
					甲	乙	甲	乙
73	淡水水源別・公共(工業用)	WF_S_Pub_Ind	INT	m ³	○		○	
74	淡水水源別・公共(上水道)	WF_S_Pub_Sup	INT	m ³	○		○	
75	淡水水源別・地表水・伏流水	WF_S_Surf	INT	m ³	○		○	
76	淡水水源別・井戸水	WF_S_Well	INT	m ³	○		○	
77	淡水水源別・その他の淡水	WF_S_Other	INT	m ³	○		○	
78	淡水水源別・回収水	WF_S_Recycle	INT	m ³	○		○	
79	淡水水源別・合計	WF_S_Total	INT	m ³	○		○	
80	淡水用途別・ボイラー用水	WF_U_Boiler	INT	m ³	○		○	
81	淡水用途別・原料水	WF_U_Mat	INT	m ³	○		○	
82	淡水用途別・処理・洗浄用水	WF_U_Treat	INT	m ³	○		○	
83	淡水用途別・冷却水	WF_U_Cool	INT	m ³	○		○	
84	淡水用途別・温調用水	WF_U_Heat	INT	m ³	○		○	
85	淡水用途別・その他	WF_U_Other	INT	m ³	○		○	
86	淡水用途別・合計	WF_U_Total	INT	m ³	○		○	
87	海水用途別・ボイラー用水	WS_U_Boiler	INT	m ³	○			
88	海水用途別・原料水	WS_U_Mat	INT	m ³	○		○	
89	海水用途別・処理・洗浄用水	WS_U_Treat	INT	m ³	○		○	
90	海水用途別・冷却水	WS_U_Cool	INT	m ³	○		○	
91	海水用途別・温調用水	WS_U_Heat	INT	m ³	○		○	
92	海水用途別・その他	WS_U_Other	INT	m ³	○		○	
93	海水用途別・合計	WS_U_Total	INT	m ³	○		○	

表 6：工業統計テーブル変数表(商品編：com1980～com1999)

No.	調査項目	フィールド名	フィールド型	単位	データの有無			
					1980年 ～ 1984年		1985年 ～ 1999年	
					甲	乙	甲	乙
1	調査年度	Year	INT	西暦4桁	○	○	○	○
2	甲乙区分	Type	INT	表 7			○	○
3	都道府県番号	State	CHAR(2)	表 8	○	○	○	○
4	市区町村番号	City	CHAR(3)		○	○	○	○
5	事業所番号	Code_Est	CHAR(5)		○	○	○	○
6	産業分類	Code_Ind	INT		○	○	○	○
7	他事業所有無	Af1	INT	表 9	○	○	○	○
8	経営組織	Org	INT	表 10	○	○	○	○
9	基本調査区	SuvArea	INT		○	○	○	○
10	従業者規模	Scale_Lab	INT	表 11	○	○	○	○
11	資本金規模	Scale_Cap	INT	表 12	○	○	○	○
12	出荷規模	Scale_Ship	INT				○	○
13	敷地規模	Scale_Area	INT	表 13			○	○
14	品目番号	Code_Com	CHAR(6)		○	○	○	○
15	製造品出荷額・隻数	Ship_Q1	INT		○	○	○	○
16	製造品出荷額・数量	Ship_Q2	INT		○	○	○	○
17	製造品出荷額・金額	Ship_Val	INT	万円	○	○	○	○
18	製造品在庫額・隻数	Stock_Q1	INT		○	○	○	○
19	製造品在庫額・数量	Stock_Q2	INT		○	○	○	○
20	製造品在庫額・金額	Stock_Val	INT	万円	○	○	○	○

表 7: 甲乙区分

区分コード	甲乙
1	甲
2	乙

表 8: 都道府県番号

都道府県番号	都道府県名	都道府県番号	都道府県名	都道府県番号	都道府県名
01	北海道	17	石川県	33	岡山県
02	青森県	18	福井県	34	広島県
03	岩手県	19	山梨県	35	山口県
04	宮城県	20	長野県	36	徳島県
05	秋田県	21	岐阜県	37	愛媛県
06	山形県	22	静岡県	38	香川県
07	福島県	23	愛知県	39	高知県
08	茨城県	24	三重県	40	福岡県
09	栃木県	25	滋賀県	41	佐賀県
10	群馬県	26	京都府	42	長崎県
11	埼玉県	27	大阪府	43	熊本県
12	千葉県	28	兵庫県	44	大分県
13	東京都	29	奈良県	45	宮崎県
14	神奈川県	30	和歌山県	46	鹿児島県
15	新潟県	31	鳥取県	47	沖縄県
16	富山県	32	島根県		

表 9: 他事業所の有無

1	工場が一つで、本社・本店はこの工場と同じ場所にある。
2	工場が一つで、本社・本店はこの工場と異なった場所にある。
3	工場が二つ以上ある。(上記 1, 2 以外)

表 10: 経営組織

1	会社(株式, 有限, 合資, 合名)
2	組合・その他の法人
3	個人

表 11: 従業者規模

規模コード	従業者数合計
00	1 ~ 3 人
01	4 ~ 9 人
02	10 ~ 19 人
03	20 ~ 29 人
04	30 ~ 49 人
05	50 ~ 99 人
06	100 ~ 199 人
07	200 ~ 299 人
08	300 ~ 499 人
09	500 ~ 999 人
10	1000人以上

表 12: 資本金規模

規模コード	資本金
11	100 万円未満
12	100 万円以上 200 万円未満
13	200 万円以上 500 万円未満
14	500 万円以上 1,000 万円未満
15	1,000 万円以上 5,000 万円未満
16	5,000 万円以上 10,000 万円未満
17	10,000 万円以上 100,000 万円未満
18	100,000 万円以上 1,000,000 万円未満
19	1,000,000 万円以上

表 13: 敷地規模

規模コード	敷地面積
0	500 m ² 未満
1	500 m ² 以上 1,000 m ² 未満
2	1,000 m ² 以上 3,000 m ² 未満
3	3,000 m ² 以上 5,000 m ² 未満
4	5,000 m ² 以上 10,000 m ² 未満
5	10,000 m ² 以上 30,000 m ² 未満
6	30,000 m ² 以上 100,000 m ² 未満
7	100,000 m ² 以上

表14： 工業統計用産業中分類別品目数

(平成6年産業分類改正)

産業分類	産業分類		商品分類（6桁）					
	3桁分類	4桁分類	製造品	貸加工品	その他	うちくず	計	うちくず
計	155	565	1,810	570	63	12	2,443	12
食料品製造業	9	43	94	42	2		138	
飲料・たばこ・飼料製造業	6	13	28	11	3		42	
繊維工業	9	39	114	52	2		168	
衣服・その他の繊維製品製造業	7	31	71	31	2		104	
木材・木製品製造業	4	20	34	20	2		56	
家具・装備品製造業	4	10	21	12	2		35	
パルプ・紙・紙加工品製造業	6	21	59	22	2		83	
出版・印刷・同関連業	6	9	18	11	6		35	
化学工業	8	43	217	32	2		251	
石油製品・石炭製品製造業	6	7	24	7	2		33	
プラスチック製品製造業	6	23	43	23	2		68	
ゴム製品製造業	4	13	41	13	2		56	
なめし皮・同製品・毛皮製造業	9	10	34	11	2		47	
窯業・土石製品製造業	9	47	115	46	2		163	
鉄鋼業	7	26	65	25	2		92	
非鉄金属製造業	6	21	58	21	2		81	
金属製品製造業	9	31	104	35	4	2	143	2
一般機械器具製造業	9	46	231	46	4	2	281	2
電気機械器具製造業	9	37	199	37	4	2	240	2
輸送機械器具製造業	6	16	77	16	4	2	97	2
精密機械器具製造業	7	23	62	23	4	2	89	2
武器製造業	1	1	4	1	4	2	9	2
その他の製造業	8	35	97	33	2		132	

注：うちくずは「261168鉄くず」，「279931非鉄金属くず」で，その他，及び計の内数。

資料：通商産業省大臣官房調査統計部編「工業統計調査用産業分類」(平成11年11月)より引用。

(平成14年産業分類改正)

産業分類	産業分類		商品分類（6桁）					
	3桁分類	4桁分類	製造品	貸加工品	その他	うちくず	計	うちくず
計	150	562	1,822	573	69	16	2,464	16
食料品製造業	9	40	95	39	2		136	
飲料・たばこ・飼料製造業	6	13	28	11	3		42	
繊維工業	9	39	114	52	2		168	
衣服・その他の繊維製品製造業	6	32	72	32	2		106	
木材・木製品製造業	4	20	37	20	2		59	
家具・装備品製造業	4	10	21	12	2		35	
パルプ・紙・紙加工品製造業	6	19	59	20	2		81	
印刷・同関連業	4	5	9	11	6		26	
化学工業	8	40	216	31	2		249	
石油製品・石炭製品製造業	5	7	24	7	2		33	
プラスチック製品製造業	6	23	43	23	2		68	
ゴム製品製造業	4	13	41	13	2		56	
なめし皮・同製品・毛皮製造業	9	10	34	11	2		47	
窯業・土石製品製造業	9	49	116	48	2		166	
鉄鋼業	6	25	65	24	2		91	
非鉄金属製造業	6	20	57	20	2		79	
金属製品製造業	9	31	104	35	4	2	143	2
一般機械器具製造業	9	47	238	47	4	2	289	2
電気機械器具製造業	6	24	114	24	4	2	142	2
情報通信機械器具製造業	2	11	52	11	4	2	67	2
電子部品・デバイス製造業	1	9	41	9	4	2	54	2
輸送機械器具製造業	6	16	78	16	4	2	98	2
精密機械器具製造業	7	22	62	22	4	2	88	2
その他の製造業	9	37	102	35	4	2	141	2

注：うちくずは「231168鉄くず」，「249931非鉄金属くず」で，その他，及び計の内数。

資料：経済産業省経済産業政策局調査統計部構造統計課「工業統計調査用産業分類」(平成14年6月)より引用。

(参考)

第 11 回日本標準産業分類改訂について

今回の第 11 回日本標準産業分類改訂の概要は以下のとおりである。

(1)改訂の基本的視点

- (ア) 情報通信の高度化，サービス経済化の進展等に伴う産業構造の変化への適合
- (イ) 統計の継続性に配慮しつつ，的確な分類項目の設定と概念定義の明確化
- (ウ) 国際標準産業分類（ISIC）等国際的な産業分類との比較可能性の向上

(2)改訂に伴う分類項目数の増減

区 分	大分類	中分類	小分類	細分類
現行項目数(A)	14	99	463	1,322
改訂項目数(B)	19	97	420	1,269
増 減(B-A)	5	△2	△43	△53

(3)改訂の主な内容

第 4 回改訂（昭和 32 年 5 月）以来の大分類項目の新設。

中・小・細分類項目についても，産業構造の変化に適合させるため，全面的に見直し。

(ア) 大分類項目の新設

①「情報通信業」

電気通信分野と情報処理分野の技術の革新・進展等を踏まえ，「通信業」，「情報サービス業」，「インターネット附随サービス業」等 5 つの中分類で構成される大分類を新設

②「医療，福祉」及び「教育，学習支援業」

Lーサービス業は，全産業の事業所数，従業者数の約 4 分の 1 を占め，各種経済活動が混在していることから，産業の実態をより明確にするため，

- ・医療，福祉に関する分野は，介護福祉に係る新産業の出現・多様化等に伴い，産業規模が拡大していることから，Lーサービス業から分離して，大分類を新設
- ・教育・学習支援に関する分野は，教育機会の拡大，生涯学習の実践，余暇時間の増大等に伴い，産業規模が増大していることから，Lーサービス業から分離して新設

③「飲食店，宿泊業」

Iー卸売・小売業，飲食店のうち，飲食店はサービスの要素のウェイトが高まっていること，Lーサービス業の中分類「旅館，その他の宿泊所」は，飲食の提供に係る収入のウェイトが高くなっていることを踏まえ，それぞれが属する大分類から

分離して新設

④「複合サービス事業」

H－運輸・通信業の中分類「郵便業」とL－サービス業の中分類「協同組合」は、複数の大分類にわたる各種のサービスを行い、いずれが主たる事業であるか判別できないという実態を有していることから、それぞれが属する大分類から分離して新設

○ 大分類項目の新設により、国際的な産業分類との比較可能性が向上

日本標準産業分類(JSIC) 第11回改訂	国際標準産業分類 (ISIC) Rev.3 1989	北米産業分類システム (NAICS) 1997
H－情報通信業	(注)	51 情報産業
M－飲食店、宿泊業	H－ホテル及びレストラン	72 宿泊及び飲食業
N－医療、福祉	N－保健衛生及び社会事業	62 医療及び社会福祉業
O－教育、学習支援業	M－教育	61 教育サービス業

(注)国際標準産業分類の一部見直し(2002年)において、補助分類として情報関連項目(Information sector)が新設された。

(イ)中分類項目の見直し(新設 22 項目、廃止 24 項目)

①新設

i)「情報通信機械器具製造業」及び「電子部品・デバイス製造業」

情報技術の進展とこれに関連する産業の拡大等に伴い、F－製造業の中分類「電気機械器具製造業」から分離して、中分類を新設

ii)「インターネット附随サービス業」

「通信業」と「情報サービス業」の両産業のいずれにも分類し難い中間領域的な産業の受け皿として、H－情報通信業の中に、中分類を新設

②廃止

i)「金属鉱業」等；D－鉱業の中分類「金属鉱業」，「石炭・亜炭鉱業」，「原油・天然ガス鉱業」及び「非金属鉱業」を廃止し，中分類「鉱業」とする。

ii)「武器製造業」；F－製造業の中分類「武器製造業」を廃止して，中分類「その他の製造業」の小分類項目として位置付け

iii)「中小企業等金融業(政府関係金融機関を除く)」及び「農林水産金融業(政府関係金融機関を除く)」；J－金融・保険業の中分類「中小企業等金融業(政府関係金融機関を除く)」及び「農林水産金融業(政府関係金融機関を除く)」を廃止し，中分類「協同組織金融業」とする。

(ウ)小・細分類項目の見直し

①小分類（新設 76 項目，廃止 119 項目）

i)新設

「建築リフォーム工事業」，「床・内装工事業」，「産業用運搬車両・同部分品・附属品製造業」，「移動電気通信業」，「学習塾」，「教養・技能教授業」等

ii)廃止

「養蚕農業」，「貴金属鉱業」，「原油鉱業」，「屋根工事業」，「練炭・豆炭製造業」，「代理商，仲立業」，「在日外国銀行」，「洗張・染物業」等

②細分類（新設 162 項目，廃止 215 項目）

i)新設

「娯楽機械製造業」，「パーソナルコンピュータ製造業」，「衛星放送業」，「コンビニエンスストア（飲食料品を中心とするものに限る）」，「ペット・ペット用品小売業」，「訪問介護事業」，「特別養護老人ホーム」，「中等教育学校」，「外国語会話教授業」，「エステティック業」，「ゲームセンター」等

ii)廃止

「銅鉱業」，「井戸ポンプ工事業」，「寒天製造業」，「絹紡績業」，「経木・同製品製造業」，「木製履物製造業」，「馬具・むち製造業」，「旅客軽車両運送業」，「火薬類卸売業」，「石炭卸売業」，「古綿打直し業」等