

CIRJE-J-291

経済史研究におけるマイクロ・データの利用

東京大学大学院経済学研究科
岡崎哲二

2017年9月

CIRJE ディスカッションペーパーの多くは
以下のサイトから無料で入手可能です。
http://www.cirje.e.u-tokyo.ac.jp/research/03research02dp_j.html

このディスカッション・ペーパーは、内部での討論に資するための未定稿の段階にある論文草稿である。著者の承諾なしに引用・複写することは差し控えられたい。

経済史研究におけるマイクロ・データの利用

岡崎哲二（東京大学）*

Abstract

This paper is written for a chapter of a primer of the study in economic and business history. The paper explains advantages of micro-data in economic history research and how to exploit the advantages, referring to three seminal articles in economic history, Sokoloff(1984), Kim(2005) and Bresnahan and Raff(1991).

Key words: マイクロ・データ、経済史、工場、都市化、大恐慌

* okazaki@e.u-tokyo.ac.jp

1.はじめに

今日、日本を含む各国の政府が公表している統計のほとんどは、国・県・市町村等の地域別、あるいは繊維・金属・機械等の産業別に表示されている。しかし、これらの統計を作成するにあたっては、企業・工場・個人等のマイクロな主体ごとにデータが収集されており、公表される統計はそれらのデータを集計した結果である。こうした集計プロセスを経ない企業・工場・個人等の主体別のデータをマイクロ・データという。

例えば、ある産業の生産額が 100 万円であるという場合、それは 100 社の企業が均等に 1 万円ずつ生産した結果かもしれないし、1 社が 80 万円、他の 1 社が 15 万円、他の 1 社が 5 万円生産した結果かもしれない。仮に 100 万円の生産額が実際に後者のように分布していたとして、その分布を示すデータは、単に産業の生産額合計が 100 万円という情報より正確で豊かである。

この簡単な例からわかるように、マイクロ・データは集計されたデータよりはるかに多くの情報を含んでいる。したがって経済現象を理解するためにマイクロ・データは集計されたデータより高い有用性を持っている。近年、各国で政府統計のもとになっているマイクロ・データ（個票データ）が一定の条件の下で研究者に広く利用可能となり、そのことが経済研究の広さと深さを広げている。

経済史研究においては、政府が古い統計の個票データを電子化された形で保有し、それを提供するケースは限られているが、紙媒体の資料からデータを電子化する等、研究者の努力によってマイクロ・データを用いた研究が進められてきた。この章ではそのいくつかを紹介することを通じて、経済史研究におけるマイクロ・データの利用方法と有用性を例証することにしたい。

2.マイクロな歴史事象を理解する：工場の生産性効果

経済史あるいはより広く歴史研究者は、古くからマイクロな現象に関心を持ち続けてきた。特定の個人や特定の企業の属性や活動に関する史料を詳細に調査し、そこから得られた情報を独自に構成することによって、さまざまな知見・洞察を得ることは経済史研究の本質的かつ普遍的な営みである。そして今日の経済史研究はこうして蓄積された膨大な知見・洞察のうえに成り立っている。

このように、歴史上の特定の個人や企業に焦点を当てて深くそれを研究することは経済史研究にとって本質的な重要性を持っている。一方で、こうした研究に対してしばしば提起される懸念に、得られた結果をどの程度一般化できるかというものがある。経済史研究者の多くは、学会発表でのコメントや学術誌に論文を投稿した際の査読者のコメントの中で、こうした懸念を提起されたことがあるだろう。そしてこれはもっともな懸念といえる。こうした懸念を解消するためにはどのようにしたらよいか。その有力な手段となるのが、広いサンプルをカバーしたマイクロ・データに基づいて、個別的な対象の深い観察から得られた結論を統計的に検証することである。そのような検証を行った古典的な研究

の一つとして、アメリカの経済史研究者 Kenneth Sokoloff 論文をとりあげよう (Sokoloff 1984)。

19 世紀に欧米で生じた「産業革命」は経済史上のもっともよく知られた出来事の一つであり、長く経済史研究の主要なテーマとされてきた。産業革命の過程で、多数の独立した職人たちによって担われていた工業生産は、機械化された大規模な工場の中で、多数の労働者が経営者の管理の下で行う生産に代替された。

この出来事に関する論点の一つに、機械化と切り離して、多数の労働者が経営者の管理の下で働くという生産の組織のしかたそのものが生産性を引き上げる効果があったのか、それとも機械化こそが工場の生産性にとって本質的な条件であったのかという問題がある。経済史研究者の多くは後者の見方をとってきたが、前者の見方も無視できない影響力を持っている。

前者の立場の代表例として、経済学の古典であるアダム・スミスの『国富論』を挙げることができる。スミスは分業論の中で、作業プロセスを分割して個々の労働者を特定の作業に専門化させるという、機械化と直接に関わらない作業組織の編成の仕方が高い生産性をもたらすと論じた。また前者の見方に立ちつつスミスに批判的な見解として、工場の中で経営者が労働者を規律・監督することが労働者の労働強度を上げ、1 人一定時間あたりの生産量という意味での生産性を引き上げたという、ラディカル派経済学の旗手、Stephen Marglin の議論もある (Marglin 1974)。

それぞれの立場とも相手を説得する体系的な論拠を提示することができず、研究が膠着状態にあった時に、Sokoloff(1984)は 19 世紀アメリカの工業統計の個票、すなわちマイクロ・データを用いてこの問題にアプローチした。すなわち、1820 年と 1850 年に関するアメリカの工業統計 (センサス) の個票からサンプリングした事業所レベルのデータを使用して次のような分析を行った。

19 世紀前半のアメリカではた綿工業、羊毛工業、鉄工業はすでに機械化が進んでいたが、酒造、金属製品、製粉等のその他の産業は依然として「非機械化産業」であった。そこで、上のマイクロ・データを用いてそれら非機械化産業の生産関数を推定し、それにも基づいて規模の経済性の程度を検証した。コブ・ダグラス型の生産関数

$$Y=AK^{\alpha}L^{\beta}$$

を想定する。Y は産出量 (付加価値)、A は技術水準、K は資本ストック、L は労働力である。ここで、2 つの生産要素 K と L を同時に λ 倍すると、産出量 (付加価値) は $A(\lambda K)^{\alpha}(\lambda L)^{\beta} = AK^{\alpha}L^{\beta} \lambda^{\alpha+\beta}$ 、すなわち $\lambda^{\alpha+\beta}$ 倍になる。したがって $\alpha+\beta>1$ なら、算出は λ 倍以上になるため規模に対して収穫逓増である。そして $\alpha+\beta=1$ なら規模に対して収穫一定、 $\alpha+\beta<1$ なら規模に対して収穫逓減となる。Sokoloff は上記の生産関数の両辺を L で除して労働者 1 人当たりの形、

$$Y/L=A(K/L)^\alpha L^{\alpha+\beta-1}$$

としようえで両辺の対数を取り、

$$\log(Y/L) = \log(A) + \alpha \log(K/L) + (\alpha + \beta - 1)\log(L)$$

と変形し、この式を推定した。

表 1 は推定結果の一部を紹介している。(1)式で $\log(L)$ の係数は正で統計的に有意であり、これは機械化されていない工場にも規模の経済性が存在したことを示している。式(2)では、規模の経済性の程度に規模による不連続性があるかどうかを、従業員が 5 人を超える工場については 1、そうでない工場には 0 となるようなダミー変数を加えてテストしている。従業員 5 人超を示すダミー変数の係数は正で有意となっているから、従業員が 5 人を超えると不連続に生産性が上昇したことになる。一方で、式(2)にはもう一つの変数が加えられている。それは従業員 5 人超を示すダミー変数と $\log(L)$ を掛けた項（交差項）である。この変数はそれを構成する 2 つの変数の相互作用を捉えている。交差項の係数は負で統計的にも有意であり、興味深いことにその絶対値は $\log(L)$ の係数の絶対値とほぼ同じ大きさとなっている。これは、従業員数が 5 人を超えて生産性がジャンプした後は、それ以上従業員数が多くなっても、それにともなって生産性が上昇しないことを意味している。

表 1

これらの推定結果に基づいて Sokoloff は、機械化を伴わなくても労働者を工場に集めて管理すること自体が生産性を引き上げた、ただしこうした規模効果は比較的狭い範囲に限定されたという結論を引き出した。この節のはじめに述べた論争との関係では、Sokoloff は、前者の見方をマイクロ・データの分析によってサポートし、あわせてそれには限定が必要であることも明らかにしたことになる。

3. マクロの歴史事象を理解する

3-1 工業化と都市化

マイクロな事象を理解するためにマイクロ・データが有用なことはある意味で当然といえるが、マイクロ・データはマクロ的な事象の理解にも寄与し得る。その例として、アメリカにおける工業化と都市化の関係に関する Sukkoo Kim の論文 (Kim 2005) を取り上げよう。

19 世紀後半にアメリカでは工業化と都市化が並行して進んだ。すなわち、1850 年から 1880 年の間に労働力に占める工業労働力の比率は 10% から 20% に倍増し、この間に都市

人口の比率も 15%から 30%に上昇した。これら二つのマクロ的な変化の間にどのような関係があったのだろうか。Kim(2005)は、この問題に Sokoloff(1984)が用いた 19 世紀の工業センサスのマイクロ・データを用いて取り組んだ。

興味深いことに、その際、Kim は労働者が集まって働く工場という作業組織を機械化と独立の概念として捉えるという Sokoloff(1984)の視点を継承している。すなわち彼は、蒸気機関の使用と工場制度を別々に捉えたと同様、それぞれが工業立地の都市化に与えたインパクトを定量的に評価することを試みた。

Kim は 1850、1860、1870、1880 年の 5 年分の工業統計のマイクロ・データを使用している。これによると、都市で雇用される工業労働者の比率は 1850 年の 40.4%から 1880 年の 71.0%に上昇した。この間に変動はあったが、傾向としては工業立地の都市化が進行したことになる。

上記のデータを用いて、Kim は次のような事業所の立地選択に関するモデルを推定した。

$$\ln[P_i/(1-P_i)] = \beta_0 + \beta_1 \text{Factory}_i + \beta_2 \text{Steam}_i + \beta_3 \text{Water}_i + \beta_4 \text{Steam}_i \times \text{Factory}_i \\ + \beta_5 \text{Water}_i \times \text{Factory}_i + \beta_6 \text{Women}_i + \beta_7 \text{Children}_i + u_i$$

P_i は事業所 i が都市に立地する確率で、それを都市以外（地方）に立地する確率 $(1 - P_i)$ で除した値はオッズ比、その対数値はロジット(logit)と呼ばれる。このロジットが右辺のいくつかの変数の線形関数になっていることが想定されている。このモデルはロジット・モデルと呼ばれ、2 つの選択肢の間の選択の決定要因を分析する場合にしばしば用いられる。右辺の変数 Factory_i は事業所 i が 15 人を超える労働者を雇用している場合、すなわち Kim の用語法ではその事業所が工場である場合に 1 となるダミー変数、 Steam_i は事業所 i が蒸気機関を動力としている場合に 1 となるダミー変数、 Water_i は事業所 i が水力を動力としている場合に 1 をとるダミー変数である。動力に関しては蒸気機関、水力以外の馬力、人力等の動力を用いる工場が基準となっている。 Women_i と Children_i は、それぞれ労働者の中の女性比率と児童比率、 u_i は誤差項である。推定にあたっては他に各郡を示すダミー変数、各産業を示すダミー変数が加えられている。

推定結果は表 2 の通りである。表の中の数字は、自然対数の底 e を推定された係数でべき乗してオッズ比に対する効果に変換したものである。例えば Factory の係数 β_1 の場合、 e^{β_1} である。したがって説明変数がダミー変数の場合、表の中の数字が 1 を超えていれば、その変数が、基準となっているケース、 Factory の場合であれば工場以外の小規模な事業所のケースに比べて、事業所の都市への立地確率を高めたことになる。 Factory に関する値は一貫して 2~3 という大きい値となっている。工場、すなわち規模の大きな事業所は、小規模な事業所に比べて都市への立地確率が格段に高かったことになる。例えば 1850 年の値 2.132 は、工場が都市に立地するオッズ比が小規模事業所のその 2.132 倍であったことを示している。一方で、 Steam に関する値は 0.664 と 1 より小さい。すなわち蒸気機関を使

用する事業所は、動力を使用しない事業所に比べて都市に立地する確率がむしろ低かった。

表 2

この結果は、蒸気機関の普及による工業の動力化が都市化を進めたとは言えないことを示唆している。他方で、**Steam** と **Factory** の交差項に関するオッズ比はほぼ一貫して 1 より大きな値となっている。これは蒸気機関が工場で使用される場合には、工場の都市への立地確率を高めたことを示している。しかしこの効果を考慮しても、蒸気機関が事業所の都市立地を促進したという効果は認められない。

工業化と都市化はいずれも経済のマクロ的な変化であるが、それらは個々の事業所の選択の結果でもある。**Kim(2005)**はこの点に着目して、マイクロ・データを用いて両者の関係にアプローチし、蒸気機関の普及は工業立地の都市化を促進しなかったという直感とは異なる結論を導いたのである。

3-2 大恐慌と工場のダイナミクス

Kim(2005)が使用した 10 年毎の工業統計はそれぞれ異なる事業所のグループを捉えている。すなわちその間に多数の事業所の退出と参入があったはずである。しかし、**Kim(2005)**はこうした事業所の入れ替わりについては考慮していない。経済史研究の文脈で、マクロ的事象の背後にある事業所や企業の退出・参入に着目した、先駆的な研究として **Bresnahan and Raff(1991)**がある。アメリカの大恐慌が実体経済を大きく縮小させたことはよく知られている。1929 年から 1933 年にかけて実質 GDP は 27%¹、実質工業生産は 41%減少した (**Bresnahan and Raff 1989, p.320**)。一方で、この時期のアメリカでは、大量生産の発達等によって個々の産業の内部構造が大きく変化したことが知られている。すなわち個々の産業の中で新しい大量生産技術を用いる大規模な企業・工場の構成比が上昇した。すなわち、**Bresnahan** と **Raff** は個々の産業内部の企業・工場の異質性に着目し、こうした異質性とマクロの生産縮小との関係に着目した。

より具体的には、彼らは、自動車工業の工場別データを用いて、大恐慌が異質な工場群の間で閉鎖と生産の変化にどのように異なる影響を与えたか、そしてそうした影響が自動車工業の構造変化にどのような意味を持ったかを分析した。この目的のためには、単に工場別のデータというだけでなく、個々の工場のデータが異時点間で接続されていること、いいかえればパネル・データであることが必要とされる。彼らは工業統計の個票を用いて、1929、1931、1933、1935 年の 4 時点の工場別パネル・データを作成した。

表 2 は彼らの分析結果の一部を示している。大恐慌の出発点である 1929 年にはアメリカに 211 の自動車工場が存在した。そのうち約半数にあたる 105 の工場が 1933 年までに閉鎖された。これは 1929 年と 1933 年の間で操業している工場の顔ぶれが大きく異なってい

¹ Historical Statistics of the United States online.

たこと、大恐慌期の自動車生産の縮小が大きな工場の構成変化を伴っていたことを示している。そして、閉鎖工場の労働者数 51 万 7,755 人は存続工場の労働者数減少 106 万 6,784 人の 48.5%に登った。この数字は工場閉鎖のインパクトの大きさを示している。

表 3 は同時に存続工場と閉鎖工場の間には明確な属性の差があったことを示している。明確なのは規模の差である。1929 年の労働者数で見ると、存続プラントは 2 万 599 人、閉鎖プラントは 4,931 人で存続プラントの方が 4 倍以上大規模であった。自動車生産台数で見ても、存続プラント 27.24 が 100 万ドル、閉鎖工場が 7.83 百万ドルと、存続工場の方が格段に大きい。このことは、大恐慌期の多数の工場閉鎖が平均的な工場規模を大きく引き上げたことを示している。

表 3

次に Bresnahan と Raff は、存続した工場の中で構成変化が生じたか否かを、簡単な回帰分析によって調べている。すなわち彼らは、工場退出によるサンプル・セレクションの影響を補正したうえで、1933 年の各存続工場の規模（対数）を 1929 年の同じ工場の規模（対数）に回帰して、後者の変数の係数がおおよそ 1 になるという結果を得た。労働者、販売収入、生産台数のいずれを規模の指標として用いた場合も結果は同様であった。この結果は、1929 年から 1933 年にかけての存続工場の規模変化率は期首の規模に依存しなかったことを示している。彼らはこの結果から、存続工場の中での構成変化の影響はほとんどなかったと結論している。この他、存続工場と閉鎖工場間に生産性の差があった点も重要である。すなわち、1929 年における労働者 1 人当たり実質生産額は存続工場が 1,806 ドルであったのに対して閉鎖工場は 1,244 であった。このことは、大恐慌期の大規模な工場閉鎖では、相対的に生産性の低い工場が淘汰され、その結果、自動車工業全体の平均的な生産性が上昇したことを示している。

マクロ的な事象の背後にある企業や工場の異質性は、1990 年代以降、経済学のさまざまな分野で理論的にも実証的にも注目され、非常に多くの研究が行われてきた (Dunne, Jensen and Roberts eds., 2001; Melitz 2003 等を参照)。Bresnahan と Raff の歴史研究はその先駆けとなった論文の 1 つである。

4.おわりに

この章では、経済史研究におけるマイクロ・データの有用性とその利用の仕方を、3 つの先駆的研究の内容を紹介することを通じて説明した。上述のように、マイクロ・データは集計されたデータよりはるかに多くの情報を含んでおり、経済史に関する論点を分析し、検証するうえで有力な手段となる。いうまでもなく、マイクロ・データの利用可能性自体が経済史研究の質を保証するわけではない。個々のマイクロ・データが含む情報に合わせて有意義な問題を設定し、適切な分析方法を考えられるかが重要であり、この点に関して

も本章で紹介した3つの論文から学ぶべきことは多い。

References

- Bresnahan T. and D. Raff (1991) Intra-industry heterogeneity and the Great Depression: The American Motor Vehicle Industry, 1929-1935, *Journal of Economic History*, 51(2): 317-331, 1991
- Dunne, Jensen and Roberts eds. (2001) *Producer Dynamics: New Evidence from Micro Data*, Chicago: The University of Chicago Press
- Kim, Sukkoo (2005) “Industrialization and Urbanization: Did the Steam Engine Contribute to the Growth of Cities in the United States ?” *Explorations in Economic History* 42: 586-598
- Marglin, S. (1974) “What Do Bosses Do ?: The Origins and Functions of Hierarchy in Capitalist Production, *Review of Radical Political Economy* 6: 60-112
- Melitz, M. J. (2003) “The Impact of Intra-Industry Reallocation and Aggregate Industry Productivity,” *Econometrica*, 71(6): 1695-1725
- Sokoloff, K. L.(1984) “Was the Transition from the Artisan Shop to the Nonmechanized Factory Associated with Gains in Efficiency ?: Evidence from the U.S. Manufacturing Censuses of 1820 and 1850,” *Explorations in Economic History* 21: 351-382.

表1 工場制の生産性効果

被説明変数: $\log(Y/L)$	(1)	(2)
切片	3.309	3.181
$\log(K/L)$	0.362	0.359
$\log(L)$	0.097	0.236
$\log(\text{郡の農業就業者比率})$	-0.108	-0.102
ニューイングランド	-0.046	-0.063
従業員5人超		0.402
従業員5人超 $\times \log(L)$		-0.240
酒造	-0.420	-0.038
金属製品	-0.064	-0.050
製粉	-0.051	-0.048
皮革	-0.287	-0.290
雑工業	0.033	0.031
R^2	0.497	0.503

資料: Sokoloff (1984), p.364.

表2 事業所の都市立地の決定要因

	1850		1860		1870		1880	
Factory	2.132	(24.6)	2.02	(25.3)	3.296	(36.7)	1.732	(20.7)
Steam	0.664	(-7.1)	0.581	(-12.5)	0.659	(-8.4)	0.29	(-33.4)
Water	0.076	(-38.8)	0.167	(-35.4)	0.122	(-21.1)	0.074	(-39.5)
Steam × Factory	0.635	(-6.8)	1.662	(9.6)	1.189	(3.1)	3.533	(29.6)
Water × Factory	2.826	(14.1)	1.163	(2.5)	2.196	(7.5)	3.643	(16.3)
Women	0.993	(-0.2)	2.606	(21.7)	1.314	(5.7)	9.59	(47.9)
Children							0.931	(-0.8)
County dummies	Yes		Yes		Yes		Yes	
Industry dummies	Yes		Yes		Yes		Yes	
Pseudo R ²	0.13		0.11		0.08		0.15	
Obs.	4351		4679		3833		5535	

注: 本文参照。

()内はz値。

資料: Kim(2005), p.592.

表3 大恐慌期のアメリカ自動車工業における工場のダイナミクス

		1929	1933
工場数	存続工場	106	106
	閉鎖工場	105	0
	参入工場	0	16
労働者数(人)	存続工場	2,183,494	1,116,710
	閉鎖工場	517,755	0
	参入工場	0	57,376
工場当たり労働者数(人)	存続工場	20,599	10,535
	閉鎖工場	4,931	0
	参入工場	0	3,586
工場当たり生産台数(台)	存続工場	36,564	16,465
	閉鎖工場	13,173	0
	参入工場	0	6,128
工場当たり実質販売収入 (1929年価格、100万ドル)	存続工場	27.24	13.22
	閉鎖工場	7.83	0.00
	参入工場	0.00	4.32
労働者1人当たり実質販売収入 (1929年価格、ドル)	存続工場	1,806	1,424
	閉鎖工場	1,244	
	参入工場		1,055

資料: Bresnahan and Raff (1991), p.326より作成.