

〔調査〕 日本の社会経済研究と統計データの利用形態*

はじめに

社会経済の研究のなかで、数量的分析手法に依存するものの比重が高まるにつれて、研究のあり方・組織化の仕方自体が、従来の伝統的なものとは異ってくるようになってきた。本調査は、その様な状況を、数量的分析手法の不可欠な素材の1つである統計データの利用状況を解析することによって、明らかにする試みの1つである。特に汎用計算機システムの急速な発展は、統計データのデータベースシステムによる提供と相まって、研究のあり方・組織化を変える大きな要因となってきている。従って本調査は、同時にこれからの統計データベース開発のあり方の方向付に対する資料であると位置付けすることが出来る¹⁾。社会経済の研究自体、伝統的な人文的研

究のあり方である個人の書斎のなかで何冊かの書物を読みながらの思索と冥想による研究から、組織としての研究へと急速に転化してきており、その限りではいわゆる研究・開発(Research and Development, 以下 R & D と略記することがある)の一環として営まれているものとみなすことが出来る。従ってそれは、R & D の社会会計の手法で解析することが、知識社会学または科学の社会学(sociology of science)の立場での分析と共になされる必要がある。諸科学²⁾のうちで、自然科学や工学の分野では、この様な分析は急速に進展してきている。この研究動向のなかに、社会科学の研究を位置付けることが、第Ⅰ部の展望の目的であり、さらにそこで得られたモデルから、実態調査の結果を再検討することが第Ⅱ部の目的である。

Ⅰ. 研究・開発の社会会計と生産関数

1.1 研究・開発の生産物とその評価

研究・開発(以下 R & D と略す)が社会経済の発展に不可欠の知的生産活動の1つであるとしても、そのあり方自体を組織的に管理するものとして検討する様になったのは、比較的最近のことである。2つの世界大戦の戦時動員体制が引き金になったことは、確かであるが、第1に R & D 自体が、自然科学・工学を中心に大規模かつ高価な実験装置と多くの人々の共同作業を必要とする様に成ったことが、かかる組織的管理という発想を引き出したし、第2に、R & D の急速な発展に伴う主題領域の細分化と多分野に互っての相互依存関係とが、R & D 自身を個々の研究者の自生的展開にゆだねるのではなく、

成のあり方を策定するために行われた調査である。この問題全体については、昭和57年度に実施したシンポジウムの記録(尾高・松田編[1983])を参照されたい。この記録に関しては中村隆英東京大学教授の書評を得ることが出来た(『経済研究』35-3)。

2) 科学(science)または学(Wissenschaft)という表現は、研究者(researcher)・学者(scholar)・科学者(scientist)といった表現と共に日本語では独特の響を持たせて使われているが、本稿ではその様な意味を付さずに使用することとした。

* 本調査の第Ⅱ部は、一橋大学経済研究所附属日本経済統計文献センターが、昭和57-58両年度に互って文部省国際学術局情報図書館課の御高配による特別事業費によって行われた「日本経済統計データベース編成の課題と方法に関する調査研究」の一環として行った「日本の社会経済統計データベース需要動向調査」(郵送調査)の調査結果を中心としている。この調査の実査に当っては、調査票の発送・回収・点検は、本稿の筆者の他に、秋山凉子文献センター助手を始め、文献センター事務職員の共同の作業による。また調査票の発送者名簿の計算機処理に関しては、一橋大学経済研究所計算機室野島教之助手の協力を得た。調査の設計段階から、終始、藤野正三郎文献センター長(当時)、溝口敏行・尾高煌之助両教授から貴重な助言と指導を頂いた。調査の過程で、第51回日本統計学会大会で報告(周防・松田・野島[1983])を行い、調査終了後、経済研究所定例研究会、第52回日本統計学会大会で報告(周防・松田[1984]、松田・周防・森[1984])を行い多くの方々に教示頂く機会を得た。記して謝意に替えたい。

第Ⅰ部は、先に発表した研究・開発の社会会計に関する調査(倉林・松田[1981])の延長であり、主として松田が担当し、第Ⅱ部は、松田と周防の協同により、上記の郵送調査結果ファイルの再集計・解析によって第Ⅰ部で検討したモデルに従って取りまとめたものである。調査方法の細部と調査結果自体については、別稿(松田編[1984])を参照されたい。

1) 第Ⅱ部の基礎となった需要動向調査は、本来今後の学術研究のため日本の社会経済統計データベースの編

その組織化を必要としたといえる。この様に組織的に管理された R & D という発想は、伝統的な大学という組織による研究者の自発的展開から、大学の外的研究組織の発達という方向をもたらしてきた(Shinn[1982], Keylor [1975])。

R & D の組織化を図るためには、まず社会会計的に R & D に投入される諸要素の計測と、その成果の計測とが必要となって来る。特に大学の外的多様な研究組織の役割の増大は、この様な計測の重要性を増すと同時に、その計測を困難にさせている(倉林・松田[1981])。計測上の実際問題としては、投入された費用の計測は調査がされていないという意味で社会会計を組み立てることが難しいが、理論的な問題としては、R & D の成果=サービスの生産をどう測定するかがある。

サービス生産の量化の方式としては、ゲルシャニイとマイルズが、サービス生産の生産性を測定するためにヨーロッパ諸国の統計作成の実態を要約して、4 類型にまとめている(Gershunny & Miles [1983])。(1)何等かの物量、例えば論文・書物の出版点数とか特許の数、(2)サービスの市場価格表示、(3)賃金支払額、(4)雇用量。)ゲルシャニイを始め、多くの経済学者・社会学者は、雇用量に着目して分析している。

一方情報学(information science, informatics)や図書館学(library science)からは、物量表示、特に出版物の量による分析がなされている(Price[1961], [1963], Menard [1971])。しかし、これでは質的变化を考慮することが出来ないで、引用される量による重みづけであるとか、出版された総量(発行部数)であるとか読者数による重みづけといった評価方法が導入されている。

この様な評価は、特定の分野に関しては、投下された費用の生産量の測定といった生産性の測定が可能であるけれども、分野を越えた社会会計のための生産量の代理変数としては機能しない。従って何等かの形での市場価格による表示が必要となってくる。

他方通常とられている市場価格表示の場合には、非市場のサービスに関しては、帰属価格として推計するといった方式をとる以外には推計し得ない。今1つの方法は、サービスの生産に投入された要素費用のすべてを生産額と看做す方法である。ただこの方法では、技術進歩等による費用の減少は、即生産額の減少となり、生産性の向上を測定出来ないという難点がある。

人的要素費用は、その質の差を教育期間等の代理変数としての賃金水準に反映しているとすれば、社会全体として簡便に測定可能でもある。多くの R & D の分析が

人的資源に集中する理由の1つでもある。しかし冒頭に述べた様に R & D への関心の1つが、物的資本設備への投下費用の巨大化から発していることを考慮に入れると、R & D の生産関数の総体として、資本と労働とを考慮に入れたモデルを設定する必要がある。

1.2 R & D の生産関数

R & D の分析に生産関数を明示的に導入するには、これ迄暗々裡に仮定されていた幾つかの基本的な前提を明らかにする必要がある。即ち、

(1) これ迄知的生産と呼ばれるものが、R & D の生産関数という形で定式化されずに来たのは、知識の集積が、何等かの積み重ねの可能なものとしては認識されていなかった事を意味する。科学史家の T. クーン³⁾は、「標準科学」(normal science)という概念を導入した。ひとたび標準科学が形成されたならば、それに準拠した活用範式(paradigm)に従った事実の観察と活用範式の活用例は、擬似的な正の生産物として蓄積可能になって来る。今仮りに、この様な知識を一定の社会的流通様式の下に表現したものが、論文であるとすれば、その生産物は量的に測定出来る。またその様な活用範式を訓練によって教えることが教育であるとすれば、それは教える必要のある知識の量に従って一定の教育活動として標準化された生産活動であると定義出来る。

(2) (1)によって、生産物が定義されるとすれば、それを生産するのに必要な費用のすべてで、生産関数が定義される。

(3) 知識の生産の社会的生産関数には、外部経済の利益が存在する。R & D は、一般的に知識の活用に依拠しており、既存の知識は、教育以外に書物等の印刷媒体などで蓄積されている(仮定(1))。これ等の印刷媒体等は、図書館等に蓄積されており、図書館には R & D のための特定の組織以外にも、一般的に利用可能な公共財として設立され知識の蓄積が、社会全体としては、個別組織に対して外部経済をもたらしている。

(4) 生産関数の形状は、資本と労働に関して、規模の優位が存在すると想定される。資本と労働の間にはある程度は代替関係が存在する。例えば、図書が資本の

3) Kuhn [1962] は、多くの議論を呼んでおり、クーン自身の定式化や(Kuhn [1977])、科学史の方法論としての定式化もかならずしも単一ではない(例えば、Barnes [1982])。筆者の行った解釈(松田 [1971])もその1つであり、またパラダイムは、クーンの用語法と異った意味で流布されているので、あえて「活用範式」という語義の明らかな訳語を用いた。訳語等も、訳本に従わなかったものもある。

一種であるとするれば、その内容を目録化し、整理することによって、図書に関して通曉している労働力を節約することが出来る。計算用機器の発達、計算のための労働力と代替することが出来るようになった。しかし、労働と資本とが完全に代替しうるわけではなく、また代替の仕方も、従来知られていなかった資本設備の導入などは、生産関数の形態変化であって、技術進歩とみなすべきである。

(5) また規模の優位には限界が存在する。例えば1つの論文を生産するのに可能な分業には限界があるからである。しかし、R & Dの組織化に従って、自然科学および工学と呼ばれる分野(science and technology)では、共著論文の数が多くなり、また共著者の数も2名以上に拡大する傾向が指摘されているが、現在の人間の知識の交換のシステムでは、論文生産工程の分業には限度があるからである。クーンが、かつて、科学者(scientist)と実務家(practitioner)という形で科学者を2分割した図式を作ったのは、すべての研究者が創造的活動を行うわけではないという仮定と、アメリカ社会でのR & Dの組織的実行状況を暗々裡に仮定していたからに他ならない。プライスが、論文を書く研究者の数と論文数の間には、バレート分布が仮定出来るとして数多くの論文を書く研究者は多くはないことを統計的に示したのは(Price [1961])同じことを別な表現で行ったと解釈することが出来る。

このような生産物としての論文数と、その生産要素の人的資源としての著者という形で科学者が曲りなりにも計測可能になったのは、書誌情報の標準化、総合目録の編成、索引誌抄録誌の刊行といったドキュメンテーション活動の蓄積を背景に、その計算機処理が可能になったからである。即ち計算機可読型書誌情報データ・ファイルの蓄積が、引用文献分析(citation analysis)を始め各種の統計データの作成を可能にし、いわゆる計量書誌学(bibliometrics)と呼ばれる分析手法を発達させたからである⁴⁾。

然し人的資源以外の生産要素に関する統計データは、あまり調査公表されてはいない。書誌情報の場合には、

4) 原田 [1974] は、この分野の発展の当初の状況の簡潔な総説論である。日本の経済学分野での研究は、桜田 [1983] に引き届いた展望が付されている。その後の研究例に高多 [1984] がある。人的資源については、大学の研究者については、新堀 [1965] [1984], Cummings [1972] がある。経済の分野に限定したものとしては藤野・畑中 [1982] がある。

生産物である論文と、生産者である著者との対応関係がつきやすいからである。その場合でも共著者以外の人的資源との対応関係は求めにくいのは、他の諸費用と同様である。生産要素のなかの資本設備の1つである図書・資料は、図書館に集積された限りでは比較的統計データとして整備されている。R & D 活動が、大学を含めた組織体の活動であると定義すれば、各組織体の附属図書館施設は、その資本設備の1つの尺度である。1950-69年のアメリカ合衆国の学術図書館のデータを集め分析したW. ボーモルと M. マルクスの作業はその一例である(Baumol & Marcus [1973])。

ボーモルとマルクスの分析によると、議会図書館を始め大学等の学術研究の図書館の蔵書規模と収書による蔵書の増加率はかならずしも相関しておらず、逆に運営のための経費は、規模に対して逓増的であると同時に時系列的に増加の傾向にある。これをこれ迄述べたモデルに置き換えると、図書館の存在は個別研究者の研究にとっては外部経済の利益を保障するが、それぞれの R & D の組織にとっては、蓄積資本の規模拡大は検索等に時間を要しサービスが低下するおそれがある。これを防ぐためには、費用は規模の拡大に伴って増進する(規模の不経済)ことになる。

このコスト軽減のための手段として、ボーモルとマルクスが注目したのは、計算機の導入である。この後の1970年代に入ってから事態の推移は、(1) 同一の文献については1つの図書館で目録作成した情報を他の図書館では利用するという分担目録作業(shared cataloguing)の計算機によるオンライン・システムとして実現したこと(例えば、OCLC)であり、さらにその延長としての(2) 資料の分担蒐集である。この種のシステムが効率的に運用されたのは、計算機による書誌情報データベースの編成以前からの、LC印刷カードに見られる様なカード形式による書誌情報の統一と資源共有が行われていたからである⁵⁾。

図書以外の資本設備、計算施設や巨大実験装置に関する組織間の相互比較を可能にさせる様なストック統計はかならずしも十分に整備されていない。

日本における R & D の経費については、先に事例紹介的に検討した(倉林・松田 [1981])けれども、そこでは、

5) アメリカ合衆国の書誌情報データベースの編成と分担目録作業については、多くの紹介と研究がなされているが、ここでは日本で社会科学分野で総合目録が存在しないための代替的作業例との関係で、松井 [1979], 古瀬・松田 [1983] をあげておく。

ストック統計との結びつき、特に何等かの形で生産量統計との結びつきを示すことの出来るデータは示されていない。資本ストックの利用状況は、図書館の利用者調査以外には十分なデータはない。既存の調査で注目すべきものは、国立大学の研究者を対象に、1971-2年に文部省大学学術局情報図書館課の実施した「人文・社会科学関係学術情報の流通・利用の実態調査」「自然科学関係学術情報の流通・利用の実態調査」がある。われわれの当面の視点からは、この調査から(1)研究成果の発表形態と発表回数という生産量の統計と、(2)利用している図書館(室)等の種類(所属大学・所属大学以外の大学・国立国会図書館・大学以外の機関)といった資本ストック量、(3)学術図書・雑誌等の年間購入経費中私費負担額、(4)学会等出席旅費私費負担額などの調査項目がある。日本における図書館施設の未発達と、その資源の共有化の度合の低さから自己蔵書に依存する度合が高いだけに、(3)のデータは極めて重要である。ただ、この後1970年代に入って、共同利用施設としての大型計算機センターの利用の研究面での定着といった、図書館以外の資本ストックの重要性が増加してきているだけに、そのような調査項目を含めての追跡調査がなされなければ、現在の研究状況を知るデータとはなり難い面がある。

1.3 社会科学の R & D の生産関数

社会事象の研究をなす社会科学で果して「標準科学」が形成されているのか、形成されたとすれば、どの時点であるかについては議論が分れるところである。一部に1870年代の「限界革命」をクーンの「コペルニクス革命」と同様な、「科学革命」の一種と考えて、「標準科学」の成立とする学説史家もある様であるが、これはクーンの理論モデルの誤読と「革命(revolution)」という用語法にまどわされたからに過ぎない。本稿は、クーンの理論モデルを検討する場ではないが、1930年代に、国際連盟委嘱のティンバーゲンのマクロ・モデル推定の試みが、コールズ財団によるケインズ・モデルの推定の試みと合体して、計量経済学的モデル分析が「活用範式」として定着した時点に求めることにする。これは、クーンのモデルが、明示的ではないが科学研究を組織体(enterprise)として実施される様になった側面を包括的に説明出来るものである点を重視したからである。即ち、R & D の生産関数で言及した様に、大学という研究組織が、唯一の研究機関ではなく、数多くの R & D 関連の組織の一種に過ぎなくなった現状を説明しうるモデルだからである。日本の場合には、1947年に公表された山田勇モデル(Yomada Model I)から出発し、50年代の後半

には、通商産業省の MITI-I モデルを皮切りに、マクロ・モデル作成作業は、ほぼ官庁エコノミストのなかに定着したと考えられる(松田[1972]、森[1974])。

このマクロ・モデルの推定作業が、大学の研究室での実験から、官庁エコノミストの手に移る過程が、まさに経済学のなかでの有効な活用範式が形成され、科学者(scientist, researcher)から、実務家(practitioner)の手に移っていく過程の実例である。

官庁エコノミストの急速な台頭を可能にさせたのが、計量経済学的実証分析であり、大量の統計データを処理するために計算機という資本設備と、そのためのデータ入力用労働力と計算機によるデータ処理のためのプログラミング要員とが必要であったからである。かかる研究作業の分業処理と、図書館に蓄積された膨大な過去の文献ではなく現在政府の刊行している統計データの蒐集は、官庁エコノミストによってより効率的に実現されていた。

官庁エコノミストに続いて、民間企業の調査・企画部門に研究者が蓄積され始めた。これ等の研究者がどの程度の規模でありうるかを示したのが、第1表である。官庁における研究者数を推定する資料と歴史的に比較する資料はないが、少なくとも社会科学特に経済学の研究者のなかで、大学における研究者の占める比重が低下していることを裏付けている。

第1表 研究本務者の分野別数(1983)

所属機関	人文・社会科学系	内訳の一部 経済学	社会学	自然科学系
会 社	1,803	—	—	199,334
研究機関	2,001	1,139	402	31,170
大 学	32,877	10,590	—	109,930
計	36,681	—	—	340,434

(資料) 総理府統計局『科学技術研究調査報告、昭和58年』1984。

1.4 社会科学における統計データの利用

社会科学における「標準科学」の形成を、計量経済学的実証分析が、組織的に実施されることと結びつけて定義したとき、その組織的研究の主体と、その経済の方式がどの様であるかが問題である。計量経済学的分析の工程を、(1)モデルの設計、(2)モデルに適合した統計データの蒐集、(3)統計的推定方式の設計、(4)計算機処理のためのプログラムの設計、(5)データの計算機可読型への変換(初期入力)、(6)計算機操作、に分割したとしてそれぞれの工程で、資本(計算機、図書館)と労働(データ入力、プログラミング)とが組織されなければならない。研究が、(1)大学等での個人の営みとしてなされる場合でも、(2)官庁や企業における研究事業体の活動として

なされる場合でも、これらの諸工程が必要であり、それが、1人の人間によってなされるか、複数の人間の分業と協業の下になされるかによって、研究の規模(例えば、方程式と変数の数と対象期間の長さ等のモデルの規模)が決定される。

社会科学のなかでも、社会学の場合には、統計データの蒐集工程自体が、実態調査の形で1つの共同作業として営まれる必要があることが早くから大学の研究者の間でも自覚されていたが、経済学の場合には、統計データの作成自体は、官庁統計部局の編成するものとして、研究組織の外に出ていたために、計算処理・統計解析が大型化する迄は、共同作業としての側面が十分に認識されていたとは思われない。計量経済学的分析が、大型化することにつれて、大学の研究室内で、学生の実習・演習と結びつけて行う初期の方式は効率的でなくなってきた。このような研究の現状を詳細に調査した事例研究は、管見に入った限りでは見当たらない。ただ、実証分析の不可欠な素材である統計データの利用を中心にした調査、統計データの利用者調査がその研究の実態を知る手掛になると考えられる。統計データが不可欠の情報であることは、1971年の国立大学の研究者を対象とした「人文・社会科学学術情報の流通・利用の実態調査」で、社会科学系研究者中、経済学分野では、統計データの常用度比率は50%を越え(経営・会計学のみは27.4%)、他の社会科学分野では社会学が同じく50%台、政治学・教育学は20%台、法学が14%であることから明らかである(松田編[1984])。

統計データが、どの様に使われているかは1970年に総理府統計局が野村総合研究所に委託した「統計利用状況と統計に対する要望調査」、1978年の統計研究会「統計利用と要望に関するアンケート調査」、「通産統計の利用・要望・新規統計に関するアンケート調査」などが、政府・官庁、民間企業、大学・研究所の研究者を対象に行っている(これらの諸調査については松田編[1984]を参照)。

統計データの利用方式を、資本設備である図書館施設、計算処理機器、労働投入量との関連でプログラミング、データ入力といった側面と結びつけた調査に、1982-3年の一橋大学経済研究所日本経済統計文献センターの「日本の社会経済統計データベース需要動向調査」と1983年の全国統計協会連合会(委員長:竹内啓)の「民間企業における統計利用の実態に関する調査」がある。前者は、大学に勤務する社会科学系研究者中主として経済学・社会学・統計学の専攻者中日本経済の実証分析を行っているものと目される2,213人に対して行い、1,063人の回答

(回収率48.0%)を得ている。後者は、東京証券取引所一部・二部上場企業500社に調査を行い、213社(366部課)の回答(回答率47.2%)を得ている。また地方自治情報センターが同年に地方自治体、全国企業・地方企業、情報取扱業者を対象に調査を行っている。第II部に於ては、上記の大学の研究者に対する実態調査を基にその統計データ分析の特質を検討する。

II 社会科学系研究者の統計データの利用の実態

2.1 大学における社会科学系研究者

大学の研究者は、社会科学系研究者のなかで、かつては最も多い研究者集団であった。現在では、特に実証研究という側面では、その様な比重は占めていない。ただその果す役割は、政府部門や民間部門の研究者とは異っていると思われる。それが具体的にどの様な形で現われるかの分析は、研究の成果から分析する方法と、研究に使用しているデータから分析する方法とがありうる。ここでは、前述の「社会経済統計データベース需要動向調査」で得られた結果から、後者の方法で検討して見る。

調査の概要 この調査では、日本の四年制大学・短期大学・高等専門学校の教員でありかつ経済学・社会学・統計学の研究者である者の中から、日本の社会経済統計データを利用していると目される(少なくとも過去5年間に日本経済に関する論考を書いた、または実証分析を行っている可能性のある者)2,213名を調査対象とし、「今後の日本の社会経済統計データを利用する研究者の間で作成することが望しい「日本の社会経済統計データベース」がどのようなものであるか」を知るための調査である。

調査事項と調査票の設計 調査票は、フェース・シートとAからF票の全11葉から構成されており、A:日本の社会経済統計の利用経験の有無、B:日本の社会経済統計の用途、C:日本の社会経済統計の入手方法と使用・処理方法、D:「自分で設計・調査した統計」について、E:「自分で設計・調査した以外の統計」の利用について、F:データベースの購入・維持管理について、全体で14問と、副次質問20問からなっている。選択肢方式で回答は示されていると同時に、各人が回答すべき質問項目は、前問の回答の選択肢の選択いかんで次に誘導

6) 藤野正三郎前センター長からは、全研究者を対象とした標本調査にすべきであるという指摘がなされたが、本調査票が極めて詳細記入を必要とするので、回収率の低下による非標本誤差の増大を恐れ、日本経済の実証分析を行っていると思われる人の重点調査として実施した。費用が許せば調査票Aのみを切り離して調査し、B以下を該当者について別途調査する方式が望ましかった。

される様になっており、インタビュー調査に近い形に定式化されている。D票とE票の前半(E₁票)は、記入者が、各統計調査毎に任意の点数だけ記入してもよい旨の指示がなされているので、この2票から得られるデータは統計調査単位であり、その他の票については回答者単位である。

調査対象者と回収率 母集団は、経済学(統計学・日本経済史および日本史を含む)と社会学の研究者であるが、このような研究分野別の完全な調査対象者リストは今のところ編成されていない⁷⁾。文部省の1980年調査(文部省学術国際局[1980])等を利用した推定では、経済学研究者6,393人、社会学研究者1,477人合計約7,800人と想定される。このなかから日本社会・経済実証分析者と想定して調査対象者としたのは、経済学1,828人、社会学385人である。母集団に占める比重は、それぞれ28.59%、26.06%であり、全体で約28%前後と思われる。回答者は1,063人であり、回収率は48.03%である(年令別、所属機関別、専攻分野別の集計結果は、それ自体多くの情報を与えている。松田編[1984])。

分析対象の限定 調査に対する回答者総数は1,063人であるが、このなかには、回答項目によっては部分的な回答もれ、非整合的な回答をした人々も含まれており、回答の修正を依頼した人も居る。このなかで、1983年12月迄に修正回答をされた861人について回答データ・ファイルを編成した。このなかで、「日本の社会経済関係の統計データを利用しなかった人」120人(約14%)を除くと741人(約86%)が、何等かの形で統計データを利用していることになる。これ等の人々(以下「統計データ利用者」と略記する)について、その利用の実態を分析してみる。

ただ、利用の程度は区々であり、733人が再加工の有無について回答しており、そのまま数字の引用(単純な%計算や複数の表の数値の合体等はこのなかを含めている)であるものが605人(82.54%)、再加工(推計・統計計算(回帰分析・シミュレーション等)の複雑な計算を行っているものが468人(63.85%)である。重複回答を許しているもので、単純計算のみは、265人(36.15%)である。

利用している統計種類 利用している統計の種類を「自分で設計・調査した統計」(以下「自分の統計データ」と略記する)と「自分で設計・調査した以外の統計」(以下「他人の統計データ」と略記する)に分けると、「自分の

統計データ」のみの人は、13人(1.75%)、「他人の統計データ」のみの人は445人(60.05%)であり、双方を使用している人は、278人(37.51%)である。これらの利用データの種類の完全に未記入のものは、5人(0.67%)である。

「自分の統計データ」のみの利用者は、極めて僅かであり、社会学者の場合でも自分の統計データである特定地域の実態調査が中心の研究者であっても、その大部分は、官庁統計等の全国的統計データも参照していることが明らかとなった。先の13人のなかには「村明細帳」データの再集計等で「自分の統計データ」を編成している者も居るところから、歴史家で、近代統計調査による

第2表 集計用合成変数と変数値コード表

〈研究主題〉

SUBJECT=PERIOD×10+DTAREA

PERIOD(対象期間)	DTAREA(対象地域)
1. 無記入	1. 無記入
2. 戦後のみ	2. 日本本土のみ
3. 戦前のみ	3. 日本本土と特定地域
4. 戦前戦後(含明治以前)	4. 特定地域のみ

〈分析方法〉

ANALTECH=DTANAL×10+HOWUSE

DTANAL(分析手法)	HOWUSE(再計算の有無)
1. 無記入	1. 無記入
2. 時系列分析のみ	2. 引用又は単純計算 (含む%計算等)
3. 横断面分析のみ	3. 複雑計算 (推計・回帰分析等)
4. 2と3	

〈データ入手形態〉

DTSEARCH=HOWGET×10+WHYTRBLK

HOWGET	WHYTRBLK
(図書館利用形態)	(統計書の欠号探しの経験)
0. 無記入	0. 無し
1. 手持の資料のみ	1. 有り
2. 所属機関の図書館も利用	
3. 他機関の時は自分で借用する	
4. 他機関の時は相互利用制度で借用する	
5. 3と4	
6. その他の手段のみ	

〈利用計算機器種類〉

HOWCALPT

$$= \sum_{i=1}^5 i \times 10^{(i-1)}$$

1. 手計算(電卓・ソロバン)
2. PCS等
3. パーソナル・コンピュータ
4. 汎用コンピュータ
5. その他

〈プログラム作成者〉

WHOPROPT

$$= \sum_{i=1}^4 i \times 10^{(i-1)}$$

1. 自分又は共同作者
2. 既存のプログラム・パッケージ利用
3. プログラム外部委託
4. その他

7) 経済学に関しては藤野正三郎教授の好意により、藤野・畑中の編成した(藤野・畑中[1982])四年制大学の研究者に関するファイルを使用することが出来た。

(注) 1) 最後の2変数は各コード*i*に対し $i \times 10^{(i-1)}$ を施し合計したもので、各パターンを示している。
2) パターン化されたコード中の?はその桁に該当のコードを無視した事を示す。

第3表 研究主題・分析方法別研究者数

(イタリックは総計に対する百分率)

ANALTECH 分析方法		(分析手法無記入)			時 系 列 分 析			横 断 面 分 析		時系列・横断面分析			計
研究主題 SUBJECT	コード	(無記入) 11	単純計 12	複雑計 13	(無記入) 21	単純計 22	複雑計 23	単純計算 32	複雑計算 33	(無記入) 41	単純計 算 42	複雑計 算 43	
(対象期間無記入)	(無 記 入) 11	1 <i>0.13</i>	9 <i>1.21</i>	10 <i>1.35</i>	0	0	1 <i>0.13</i>	0	0	0	0	0	21 <i>2.83</i>
	日 本 本 土 12	0	0	0	0	0	0	1 <i>0.13</i>	0	0	1 <i>0.13</i>	1 <i>0.13</i>	3 <i>0.40</i>
	日本・特定 13	0	0	1 <i>0.13</i>	0	0	0	0	0	0	1 <i>0.13</i>	1 <i>0.13</i>	3 <i>0.40</i>
	特定地域 14	0	0	0	0	1 <i>0.13</i>	0	0	0	0	0	0	1 <i>0.13</i>
現 状 関 心 派	(無 記 入) 21	0 <i>0.27</i>	2 <i>0.13</i>	1 <i>0.13</i>	2 <i>0.27</i>	3 <i>0.40</i>	4 <i>0.54</i>	0	0	0	3 <i>0.40</i>	3 <i>0.40</i>	18 <i>2.43</i>
	日 本 本 土 22	0 <i>0.27</i>	2 <i>0.54</i>	4 <i>0.54</i>	1 <i>0.13</i>	49 <i>6.61</i>	53 <i>7.15</i>	17 <i>2.29</i>	13 <i>1.75</i>	0	36 <i>4.86</i>	126 <i>17.00</i>	301 <i>40.62</i>
	日本・特定 23	0 <i>0.27</i>	2 <i>0.13</i>	1 <i>0.13</i>	0	12 <i>1.62</i>	14 <i>1.89</i>	3 <i>0.40</i>	3 <i>0.40</i>	1 <i>0.13</i>	25 <i>3.37</i>	58 <i>7.83</i>	119 <i>16.06</i>
	特定地域 24	0	0	0	0	8 <i>1.08</i>	7 <i>0.94</i>	0	5 <i>0.67</i>	0	1 <i>0.13</i>	12 <i>1.62</i>	33 <i>4.45</i>
歴 史 派	(無 記 入) 31	0 <i>0.27</i>	2 <i>0.13</i>	0	0	0	0	0	0	0	0	0	2 <i>0.27</i>
	日 本 本 土 32	0 <i>0.13</i>	1 <i>0.13</i>	0	0	2 <i>0.27</i>	1 <i>0.13</i>	0	0	0	3 <i>0.40</i>	2 <i>0.27</i>	9 <i>1.21</i>
	日本・特定 33	0	0	0	0	2 <i>0.27</i>	5 <i>0.67</i>	0	0	0	6 <i>0.81</i>	5 <i>0.67</i>	18 <i>2.43</i>
	特定地域 34	0	0	0	0	2 <i>0.27</i>	2 <i>0.27</i>	1 <i>0.13</i>	0	0	1 <i>0.13</i>	2 <i>0.27</i>	8 <i>1.08</i>
長 期 動 向 派	(無 記 入) 41	0 <i>0.13</i>	1 <i>0.27</i>	2 <i>0.27</i>	0	0	0	0	0	0	2 <i>0.27</i>	0	5 <i>0.67</i>
	日 本 本 土 42	0 <i>0.13</i>	1 <i>0.13</i>	0	0	11 <i>1.48</i>	9 <i>1.21</i>	5 <i>0.67</i>	3 <i>0.40</i>	1 <i>0.13</i>	14 <i>1.89</i>	44 <i>5.94</i>	88 <i>11.88</i>
	日本・特定 43	0 <i>0.27</i>	2 <i>0.27</i>	0	0	13 <i>1.75</i>	10 <i>1.35</i>	0	0	2 <i>0.27</i>	15 <i>2.02</i>	63 <i>8.50</i>	105 <i>14.17</i>
	特定地域 44	0	0	0	0	3 <i>0.40</i>	1 <i>0.13</i>	0	0	0	2 <i>0.27</i>	1 <i>0.13</i>	7 <i>0.94</i>
計		1 <i>0.13</i>	22 <i>2.97</i>	19 <i>2.56</i>	3 <i>0.40</i>	106 <i>14.30</i>	107 <i>14.44</i>	27 <i>3.64</i>	24 <i>3.24</i>	4 <i>0.54</i>	110 <i>14.84</i>	318 <i>42.91</i>	741 <i>100.00</i>

統計データは存在しない様な明治以前の時期の研究者も含まれている。

従って、少なくとも「他人の統計データ」を利用している723人(97.57%)について、「他人の統計データ」の利用状況(E₁票の記入事項)で、その人が、どの様な研究主題をどの様な分析方法で分析しているかを分類することが出来る。即ち、E₁票には複数の使用調査名が記入できて、その各々の調査名毎に、その利用状況を記す様になっているので、複数の調査名を記入した者については、あたかも重複回答を許すかの様にして、統計調査単位の情報を回答者単位の情報に変換することが可能である。「他人の統計データ」を利用しているという項目に該当していると記した712人の他に、11人がE₁票に記

入しているので、少なくとも上記の723人については、E₁票の記入事項によって、その人の統計利用形態の属性が定義出来る。E₁票の未記入者18人については、その点は欠損値となる。

2.2 利用統計データの種類と利用方法

研究者がどの様な研究を行っているかを、研究者の主観による位置付けではなく、利用している統計のうちの「他人の統計データ」の種類とその利用方法で格付する方法は、E₁票で記入された個別の統計調査単位の利用状況の回答を類型化して、個人に関する情報の次元に写像する事に他ならない。この統計調査単位の情報は、人によって異った数だけ記入されており、最大点数は、40点が1人であり、10点の133人と2点の96人の双峰型

分布をしている。723 人の E₁ 票記入者に対して、延 3,811 点の記載であり、1 人平均 5.27 点となる。

統計調査単位の情報と個人単位の情報に縮約するさいに、調査数によって重み付けをする方法が考えられる。ただ個人別の総統計調査の記入点数は、かならずしも統計の利用頻度・重要性を示しているわけではないので、ここではその種の重み付けをせずに、個人情報に変換する。これ等の個人情報は、他の質問項目に関する回答で得られる個人情報と組み合わせて、個人の属性を定義する諸変数とすることが出来る。第 2 表は、以下の分析で使用するこの実態調査の回答から誘導された諸変数の一覧表である。

研究主題とデータの分析方法 研究主題(SUBJECT 以下変数名は太字で示す)は、E₁ 票の使用統計調査の「データ対象期間」(PERIOD)と「調査対象地域」(DTAREA)の 2 変数で定義される。ここで第二次世界大戦後のデータのみを使用している人を現状関心派、戦前のデータのみを使用している人を歴史派、戦前・戦後のデータを使用している人を長期動向派と名付けると、現状関心派は 6 割強の比重を占めるに過ぎない(第 3 表)。このことは大学の研究者の研究関心が長期的視点から観察しているという点で、他の政府官庁・民間企業の研究者とは際立って異なる特徴を示している。全体の 5 割強の人が、日本本土全体を対象としているのに対して、日本本土全体と特定地域のデータの双方を使用している人が 3 割強である。歴史派であっても特定地域にのみ関心をもっていたのでは分析が出来ない状況を示している。ただ現状関心派では、他の時期にも関心のある者と異って、日本本土全体のデータのみを使用しているものが圧倒的比重を占めている(40.62%)。

この研究主題の差は、分析手法の差に反映している。いま分析手法(DTANAL)を時系列、横断面、時系列と横断面の双方に分類すると、全体では時系列のみは 29.15%であるのに対し、現状関心派では、他のグループと比べて、時系列のみの占める比重が高く、32.48%に達している。これに対し、長期動向派は、22.93%である。これは一見矛盾している様に見えるが、現状関心派の分析技法が、計算機によるマクロ・データ中心の場合には、時系列解析に依拠することが多いことを考慮に入れるとかならずしも不思議ではない。従って、分析手法とどの程度の計算処理を行っているかを考慮に入れなければ分析方法の全体を明らかにすることは出来ない。C₁ 票から得られる統計データの利用の際の「再計算の有無」(HOWUSE)を基準に、そのまま統計数字を引用するか%

計算程度の単純計算と、推計、回帰分析、シミュレーション等の再加工を行う複雑計算を行うグループとに分類し、分析方法(ANALTECH)を定義するとこの間の関係を明らかにすることが出来る。

長期動向派のなかで時系列分析のみを行う者では、単純計算は、57.44%であるのに対し、現状関心派のそれは 47.05%である。全体として、引用・単純計算の者は、35.76%であるけれども、時系列・横断面分析の双方は、引用・単純計算が 25.46%であり、他方、時系列のみは、引用・単純計算が 49.07%であることから、この特徴を知ることが出来る。概して、長期動向派は、時系列・横断面分析を組み合わせた複雑計算を行うものの比重が高いといえる。これに対して歴史派の場合には、引用・単純計算が 54.05%を占めており、既存のデータの存在量が、分析手法を制約している状況をうかがわせる。

計算用機器の使用状況と研究主題 C₁ 票で個人単位に計算に使用した機器のすべてが列挙されている。何等かの計算用機器を使用しているのは、482 人で全体の 65.04%であるが、そのなかからソロバン・電卓の類のみを利用しているものを除くと、338 人で全体の 45.61%であって、そのなかでも汎用電子計算機の利用者は 269 人で、汎用機の使用はほぼ定着したと見ることが出来る。この使用計算機の組み合わせ類型変数(HOWCALPT)と研究主題の組み合わせは、第 4 表に示してある。

使用機器の中で、主な計算作業がここ 10 年以前に属すると想定される PCS 等のみ依存したものとソロバン・電卓等の手計算のみとその他(外部委託等)のみを非コンピュータ派とし、パソコンと汎用機のいずれか又は双方を使用している者をコンピュータ派と定義すると、現状関心派のなかではコンピュータ派が 44.58%、長期動向派のなかでは 36.09%である。他方、非コンピュータ派で複雑計算を行っているものは、現状関心派では 21.86%、長期動向派は 30.24%であり、長期動向派の利用するデータの方が現状関心派の利用するデータに比べて、断片的かつ少量データであることが多いことを推測させる。

この点を分析手法との組み合わせで見ると第 5 表となり、単純計算処理が主であるもののうちコンピュータ派は、僅かに 1 名でありほぼ無視しうる数である。非コンピュータ派で単純計算のものも 13 名であり、HOWCALPT での単純計算と ANALTECH から見た単純計算・複雑計算はほぼ完全に整合的な結果を示している。

対象地域との関係で見ると、現状関心派で日本本土を対象とするものに、コンピュータ派が多く、しかもそ

の大多数が少なくとも汎用機は使用しているということは、先のマクロ・データによるマクロ・モデル依存が多いという推定を裏付けている。パソコンを使用する者は現状関心派から増していると推定される。

計算機器を使用する以上は、当該研究者が資本設備として利用可能なものがあるか、外部委託をするかのいずれかである。完全に外部委託にのみ依存するものは、僅かに全体で5人(0.67%)であるけれども、自分又は共同作業で計算処理もするが外部委託もするものもそれ程多くはなく43人(5.80%)である。自分又は共同作業で計算処理するものは423人(57.09%)である。残りは270人(36.44%)であり、これ等の人々は、計算機利用方法に詳しくないという287人にほぼ対応すると考えられるので、共同利用施設の大規模計算機センターの発達状況からして、計算機器の不足が制約条件であるよりも、利用方法に関する知識と新しい機

器を使おうとする積極性が重要な要素になって来る。

計算用機器に関する知識のなかでも重要なものはプログラム作成能力であり、コンピュータの利用もプログラム・パッケージの開発によって急速に拡大する様になった。それ以前には、自分自身でプログラムを組むかプログラミング能力のあるものとの共同作業が不可欠であった。誰れがプログラム作成者(WHOPROPT)であるかという視点から類型化すると、コンピュータ使用の300人中、自分または共同作業者がプログラムを組むものは、252人(84%)であり、既存のプログラム・パッケージしか利用しないかプログラムは外部委託又はそれ以外の手段に頼る者は、48人(16%)に過ぎない。いづれにしてもプロ

第4表 研究主題別・利用計算機器別研究者数 (イタリックは総計に対する百分率)

HOWCALPT 利用計算機 種類		単純計算 の み	複 雑 計 算						計	
研究主題 SUBJECT	コード		0	コンピュータ (C)非使用者		コンピュータ使用者				C非利用者
				手計算 ?0001	PCS等 ?002?	パソコン ?03??	汎用機 ?40??	併 用 ?43??		そ の 他 50000
(対象期間無記入)	(無 記 入)	9	5	0	0	6	0	1	21	
	11	1.21	0.67			0.81		0.13	2.83	
	日本本土	2	0	0	0	1	0	0	3	
	12	0.27				0.13			0.40	
	日本・特定	1	1	1	0	0	0	0	3	
13	0.13	0.13	0.13					0.40		
特定地域	1	0	0	0	0	0	0	0	1	
14	0.13								0.13	
現 状 関 心 派	(無 記 入)	10	2	0	0	4	2	0	18	
	21	1.35	0.27			0.54	0.27		2.43	
	日本本土	100	49	13	12	99	27	1	301	
	22	13.50	6.61	1.75	1.62	13.36	3.64	0.13	40.62	
	日本・特定	39	26	5	4	30	15	0	119	
23	5.26	3.51	0.67	0.54	4.05	2.02		16.06		
特定地域	9	5	1	2	12	3	1	33		
24	1.21	0.67	0.13	0.27	1.62	0.40	0.13	4.45		
歴 史 派	(無 記 入)	2	0	0	0	0	0	0	2	
	31	0.27							0.27	
	日本本土	6	1	0	0	1	1	0	9	
	32	0.81	0.13			0.13	0.13		1.21	
	日本・特定	7	8	0	3	0	0	0	18	
33	0.94	1.08		0.40				2.43		
特定地域	4	0	0	1	3	0	0	8		
34	0.54			0.13	0.40			1.08		
長 期 動 向 派	(無 記 入)	2	3	0	0	0	0	0	5	
	41	0.27	0.40						0.67	
	日本本土	32	17	4	4	26	5	0	88	
	42	4.32	2.29	0.54	0.54	3.51	0.67		11.88	
	日本・特定	30	27	10	4	25	8	1	105	
43	4.05	3.64	1.35	0.54	3.37	1.08	0.13	14.17		
特定地域	5	0	0	1	1	0	0	7		
44	0.67			0.13	0.13			0.94		
計		259	144	34	31	208	61	4	741	
		34.95	19.43	4.59	4.18	28.07	8.23	0.54	100.00	

グラム・パッケージを使用しているものは189人(63%)であり、そのうち186人迄が自分または共同作業者がプログラムを組んでいるなかに入っている。プログラム・パッケージを利用しない111人中、プログラムは一切外部委託または「その他」の8人を除いた103人がプログラム・パッケージで処理しきれないプログラムを開発しているか、何等かの理由でパッケージ利用の機会がないことになる(第6表)。

2.3 図書館と統計データ

計算機施設(ハードウェア)およびその利用技術(ソフトウェア)に匹敵する重要な分析用の資本設備は、統計資料である。この統計資料は、他の資料と異り、図書館

施設に依存する度合いが高い。

通常、雑誌や事典等の参考資料は単行書と異り、長期間欠なく累積する事が有効であり、また一点当りの金額が高価なこともあって、通常の研究者は図書館施設に依存して利用することが多い。統計資料の場合には、この性質がさらに強まる。

雑誌の場合には、欠号があってもその号の特定の論文が必要なことが多く、記事索引等で必要な書誌事項が判っていると他で求めることが可能である。これに対して、一回限りの特殊な統計調査でない限り、統計調査資料は時系列的に揃える必要があり、その作業は通常分析開始の時点で行うべき予備作業である。また「国勢調査報告書」の様に、膨大な数の報告書が一調査毎に刊行されるものを、事前に必要な巻だけ取揃えるべく選別することは難しい。

従って、雑誌や単行書以上に図書館に依存する度合いが高い。問題は、それぞれの大学の図書館が十分に欠号なく揃えていないだけでなく、欠号のある時にそれがどこに所蔵されているかを知るための情報源である総合目録(union catalogue)が日本では未整備なことである⁸⁾。統計データの利用にあたって困った事のなかで、資料の欠号探しの問題がある(C₃票)。欠号探しの経験のあるものは(WHYTRBLK), 40.08%に達していることから明らかである(第7表)。

印刷媒体の統計データの入手形態の類型変数(HOWGET)では、手持ち資料や自分の所属機関の図書館にのみ依存している人は、46.56%に過ぎず、48.45%の人は、

8) 明治末期から、第二次世界大戦直後迄を主な対象とした統計資料の総合目録(経済資料協議会編『日本経済統計資料総合目録, 明治42-昭和22年』(農林業編, 鉱工業・エネルギー産業編, 財政・金融・経営・商業・貿易・運輸編)全6冊, 1979-1980, 同朋舎)の所蔵編の統計を見ると、日本の現状の問題点を明らかにすることが出来る。この解析は紙幅の関係で割愛した。

第5表 分析手法別・利用計算機別研究者数 (イタリックは総計に対する百分率)

HOWCALPT 利用計算機 種類		単純計算 の み 0	複 雑 計 算					計	
分析手法 ANALTECH	コード		コンピュータ (C)非使用者		コンピュータ使用者		C非使用者		
			手計算 ?0001	PCS等 ?002?	パソコン ?03??	汎用機 ?40??	併用 ?43??		そ の 他 50000
(分析手法無記入)	(無 記 入) 11	1 0.13	0	0	0	0	0	0	1 0.13
	単 純 計 算 12	19 2.56	3	0	0	0	0	0	22 2.97
	複 雑 計 算 13	0	8	1	1	8	0	1	19 2.56
			1.08	0.13	0.13	1.08		0.13	
時 系 列 分 析	(無 記 入) 21	3 0.40	0	0	0	0	0	0	3 0.40
	単 純 計 算 22	104 14.04	2	0	0	0	0	0	106 14.30
	複 雑 計 算 23	0	41	4	6	45	11	0	107 14.44
			5.53	0.54	0.81	6.07	1.48		
横 断 面 分 析	単 純 計 算 32	26 3.51	1	0	0	0	0	0	27 3.64
	複 雑 計 算 33	0	7	0	2	11	3	1	24 3.24
			0.94		0.27	1.48	0.40	0.13	
横 時 断 系 面 列 分 析	(無 記 入) 41	4 0.54	0	0	0	0	0	0	4 0.54
	単 純 計 算 42	102 13.77	7	0	0	1	0	0	110 14.84
	複 雑 計 算 43	0	75	29	22	143	47	2	318 42.91
			10.12	3.91	2.97	19.30	6.34	0.27	
計		259 34.95	144 19.43	34 4.59	31 4.18	208 28.07	61 8.23	4 0.54	741 100.00

他機関の資料をも利用している(残りの4.99%は欠損値である)。前者のうちで欠号探しの経験のあるものは24.63%であるのに対し、後者のそれは58.49%であり、欠号探し、他機関の図書館を利用する大きな誘因になっていると推定される。しかも、この他機関の図書館の利

第6表 計算機利用者のプログラム作成者による分類

「利用計算機種類」によるコンピュータ 使用者の内訳		(1) パソコン 使用者	(2) 汎用機 使用者	(3) 併 用 者	計算機利用者計
「プログラム作成者」による再分類					
再分類対象者数		31	208	61	300(100%)
プログラムの 自給による分類	有	26	166	60	252 (84%)
	無	5	42	1	48 (16%)
プログラムの・パッ ケージの使用の有 無による分類	有	17	125	47	189 (63%)
	無	14	83	14	111 (37%)
プログラムの外注の有 無による分類	有	0	29	8	37 (12%)
	無	31	179	53	263 (88%)

第7表 研究主題別・データ入手形態別研究者数

(イタリックは総計に対する百分率)

DTSEARCH データ 入手形態			(図書館利用) 形態無記入	手持ちの資料又は所属機関の図書館を利用する				他機関の図書館も利用する						計
				手持ちの資料のみ		所 属 機 関 の 館		他機関図書館 (自分で借用)		他機関図 書 館 (相互利用制度)		他 機 関 図 書 館 (自己借用・相互)		
研究主題 SUBJECT	コード	欠号無 0	欠号有 1	欠号無 10	欠号有 11	欠号無 20	欠号有 21	欠号無 30	欠号有 31	欠号無 40	欠号有 41	欠号無 50	欠号有 51	
(対象期間無記入)	(無 記 入)	14	0	1	0	4	0	1	0	1	0	0	0	2
	11	1.89		0.13		0.54		0.13		0.13				2.83
	日本本土	1	0	2	0	0	0	0	0	0	0	0	0	3
	12	0.13		0.27										0.40
	日本・特定	0	0	0	0	0	0	2	1	0	0	0	0	3
現 状 関 心 派	13							0.27	0.13					0.40
	特定地域	0	0	0	0	1	0	0	0	0	0	0	0	1
	14					0.13								0.13
	(無 記 入)	3	0	0	0	6	2	5	2	0	0	0	0	18
	21	0.40				0.81	0.27	0.67	0.27					2.43
歴 史 派	日本本土	10	1	19	1	117	43	39	33	12	11	5	10	301
	22	1.35	0.13	2.56	0.13	15.79	5.80	5.26	4.45	1.62	1.48	0.67	1.35	40.62
	日本・特定	2	0	4	1	36	12	22	25	2	5	3	7	119
	23	0.27		0.54	0.13	4.86	1.62	2.97	3.37	0.27	0.67	0.40	0.94	16.06
	特定地域	2	0	2	0	10	2	7	4	3	1	1	1	33
長 期 動 向 派	24	0.27		0.27		1.35	0.27	0.94	0.54	0.40	0.13	0.13	0.13	4.45
	(無 記 入)	0	0	0	0	0	0	0	1	0	0	0	1	2
	31								0.13				0.13	0.27
	日本本土	0	0	0	0	1	1	1	4	0	1	0	1	9
	32					0.13	0.13	0.13	0.54		0.13		0.13	1.21
計	日本・特定	0	1	0	0	1	1	1	6	2	2	0	4	18
	33		0.13			0.13	0.13	0.13	0.81	0.27	0.27		0.54	2.43
	特定地域	0	0	0	0	3	0	1	4	0	0	0	0	8
	34					0.40		0.13	0.54					1.08
	(無 記 入)	0	0	0	0	3	0	1	1	0	0	0	0	5
長 期 動 向 派	41					0.40		0.13	0.13					0.67
	日本本土	0	0	4	0	23	10	18	17	1	3	2	10	88
	42			0.54		3.10	1.35	2.43	2.29	0.13	0.40	0.27	1.35	11.88
	日本・特定	3	0	1	0	22	12	11	34	1	4	4	13	105
	43	0.40		0.13		2.97	1.62	1.48	4.59	0.13	0.54	0.54	1.75	14.17
計	特定地域	0	0	0	0	0	0	2	3	1	0	0	1	7
	44							0.27	0.40	0.13			0.13	0.94
	計	35	2	33	2	227	83	111	135	23	27	15	48	741
		4.72	0.27	4.45	0.27	30.63	11.20	14.98	18.22	3.10	3.64	2.02	6.48	100.00

用者のうち、図書館の相互貸借制度を通じて事務的に行われる手続にのみ依存している者は、僅かに13.92%に過ぎない。残りは自分で借用手続に歩かなければならず、極めて時間使用的作業である。また欠号探しの経験がないという人のなかでも、自分で借用手続に足を運んでいる者が、そのうちの28.37%を占めている所から、広い意味で統計書をあさる作業の必要のない人は極めて少ないと推定される。

これ等の統計書の探索作業の有無は、研究主題とも密接に関連している。現状関心派の54.14%が、手持ち資料と自機関の図書館の資料で分析を行っているのに対し

て、長期動向派は61.95%が、他機関の図書館をも利用しているし、歴史派に至っては、78.37%に達している。さらに欠号探しの経験の有無を加えると、この関係はさらに明瞭になり、現状関心派は、62.20%が欠号探しの経験が無くその比重は、手持ち資料と所属機関の図書館にのみ依存するもの場合にはさらに高く76.07%となる。

これに対して、歴史派はいうまでもなく長期動向派でも欠号探しの経験の無いのは、45.85%であり、手持ち資料と所属機関の図書館のみに依存するもの場合でも70.66%に留っている。

第8表 データ集計処理方法

データの種類		自 分 の 統計データ	他 人 の 統計データ
機械処理			
記 入 者 総 数(人)		291	723
総 記 入 点 数		403	3811
集計処理 方法 ¹⁾ (複数回答)	〈有効記入点数〉	394	3294
	手 集 計	187 (47.46%)	2584 (78.45%)
	機 械	211 ²⁾ (53.55%)	1115 (33.85%)
機械処理 の場合の 使用機種 (複数回 答)	〈有効記入点数〉	234	
	集計機、会計機、パン チ・カードシステム	69 (29.49%)	
	パーソナルコンピュ ータ	19 (8.12%)	
	汎用コンピュータ	160 (68.38%)	
機械処理 の場合の 初期入力 方式 (複数回 答)	〈有効記入点数〉		1351
	自分・共同作業で入 力		1168 (86.45%)
	外部の作成した媒体で 入力 ³⁾		244 (18.06%)
	オンライン検索による ファイル編成		15 (1.11%)

(注) 1) ()内の%は、有効記入点数に対する比率。

2) この他社計算機サービス会社等に外部委託したものが65点ある。

3) カード・磁気テープ・ディスク。

2.4 計算機可読型媒体の資本・労働節約効果

統計データを印刷媒体で利用している限り、そのデータを計算機処理するに当っては、計算機可読型データ(以下データ・ファイルと呼ぶ)に変換・再入力する必要がある。第8表は、利用された統計データが、それぞれの様な形で処理されているかを示したものである。「他人の統計データ」の約35%が、機械処理されていて、再入力されているが、そのうちの18%しか、計算機可読型媒体で入手されていない。オンライン検索によるデータ・ファイル編成というデータベースの利用は1%強であって、まだ結についたばかりである。

統計データをデータ・ファイルで入手することは、計算機で処理することが前提であり、その限りでは計算機処理に伴う資本投下は行われている。換言すれば、資本使用的な技術変化が実現した後の状況である。そこでデータの入力処理の作業が、データ・ファイルの入手で代替するのならば、極めて労働節約的技術変化が実現したことになる。周知の様に、現在の技術段階ではデータ・ファイルの編成は極めて労働集約的作業であるから、こ

種の技術進歩は、統計データの計算機処理による解析の普及に大きな役割を果たす。

さらに統計データ・ファイルの入手が、オンラインのデータベースの利用に代替された場合には、伝統的な図書館という資料の蓄積方式に対して大きな変化をもたらすといえる。

統計データのデータベース利用者 「計算機可読型共用データ・ファイル」と定義したときのデータベースを利用したことのあるものは143人で、741人中の19.29%を占めている。(ちなみにデータ・ファイルを入手利用したことのある者は132人、オンライン検索等でのデータベースの利用者は20人である。)

データ・ファイルで提供されるデータがあるならば使用したい人は499人で、前述のコンピューター使用派の300人を越し、再加工して利用している人の468人を越している。潜在的データベースの利用者は居るとして、問題は、果してこれ等の人々の欲している統計データが提供される可能性があるかになる(第9表)。

E₁票に記載の統計調査を調査名毎に再集計(名寄せ集計)すると、統計調査毎の利用者数が得られる。この利用者数当りの調査の種類数を両対数目盛で図示すると、第1図が得られる。利用者数の極めて少ない統計は多種類に互り、1人利用の統計が240種類であるのに対し、利用者数10人以上のものは、その種類が激減し、僅かに58種類になる。利用者の数の第1位は「国勢調査」で207人である。

本調査の回収率が約1/2であるから、非回答誤差が小さいとすれば、約2倍の利用者が想定される。さらに、前述の様にE₁票には利用した統計のすべてを記載しているわけではなく、多くは10点の統計調査で止めていると推定されるので、利用者数の多い統計の潜在的利用者はかなり多いと推測出来る。しかし、大部分の統計調査については、それ程多くの利用者が想定出来ない。これ等についてまでも、データベースの型で、維持更新管理をすることは難しいと考えられる。むしろ書誌情報データベースの領域であって、統計データベースに収録されるのは、利用者数の多い比較的少数のデータ・ファイルに限定されざるを得ない。問題は、その様に限定したとしても、なおかつデータベース供給者にとって採算のとれるものになりうるかどうかである。そのために研究者以外の潜在的需要を推定してみる。

民間企業の統計データ需要の実態 第I部で言及した民間企業213社の調査(全国統計協会連合会(竹内啓委員長)[1984])から⁹⁾統計を比較的良く使用している部課

第9表 データ入手形態別・利用計算機器別利用者数 (イタリックは総計に対する百分率)

HOWCALPT 利用計算機 種類	データ 入手形態 DTSEARCH	単純計算 の み 0	複 雑 計 算						計
			コンピュ-タ(C) 非 使 用 者		コンピュ-タ使用者			C非使用者	
			手 計 算 ?0001	PCS等 ?002?	パソコン ?03??	汎用機 ?40??	併 用 ?43??	そ の 他 50000	
(図書館利用) 形態無記入	欠無 0	10 1.35	4 0.54	0	1 0.13	18 2.43	0	2 0.27	35 4.72
	欠有 1	1 0.13	1 0.13	0	0	0	0	0	2 0.27
手 持 ち の 資 料 の み	欠無 10	20 2.70	5 0.67	1 0.13	0	5 0.67	2 0.27	0	33 4.45
	欠有 11	1 0.13	0	0	0	0	1 0.13	0	2 0.27
所 属 機 関 の 図 書 館	欠無 20	96 12.96	32 4.32	9 1.21	10 1.35	68 9.18	12 1.61	0	227 30.63
	欠有 21	24 3.24	16 2.16	7 0.94	3 0.40	24 3.24	8 1.08	1 0.13	83 11.20
他機関の図書館 (自分で借用)	欠無 30	38 5.13	23 3.10	5 0.67	3 0.40	34 4.59	7 0.94	1 0.13	111 14.98
	欠有 31	36 4.86	32 4.32	6 0.81	8 1.08	33 4.45	20 2.70	0	135 18.22
他機関の図書館 (相互利用制度)	欠無 40	5 0.67	7 0.94	3 0.40	1 0.13	5 0.67	2 0.27	0	23 3.10
	欠有 41	8 1.08	6 0.81	0	2 0.27	7 0.94	4 0.54	0	27 3.64
他機関の図書館 (自分借用・ 相互利用)	欠無 50	7 0.94	6 0.81	0	1 0.13	1 0.13	0	0	15 2.02
	欠有 51	13 1.75	12 1.62	3 0.40	2 0.27	13 1.75	5 0.67	0	48 6.48
計		259 34.95	144 19.43	34 4.59	31 4.18	208 28.07	61 8.23	4 0.54	741 100.00

366が良く使用する統計調査に関する情報が得られる。このなかから政府・地方公共団体の統計データ延2,219点(1部課当り6.06種)について、統計調査毎の名寄せ集計を行って、統計調査当りの部課数を求め図示したのが、第2図である。第1図と同様に両対数表示であり、利用部課数の少ない統計調査は多数あるにもかかわらず、利用部課数が増加するにつれて、統計調査の数は激減している。従って、大学の研究者集団の場合と同様に、利用部課数の多い統計データは少数のものに限定されてしまう。ちなみに利用部課数順で第1位(136部課)は大蔵省「有価証券報告書」であり、厳密な意味では再集計して統計化しない限り統計ではない、第2位(109部課)：総

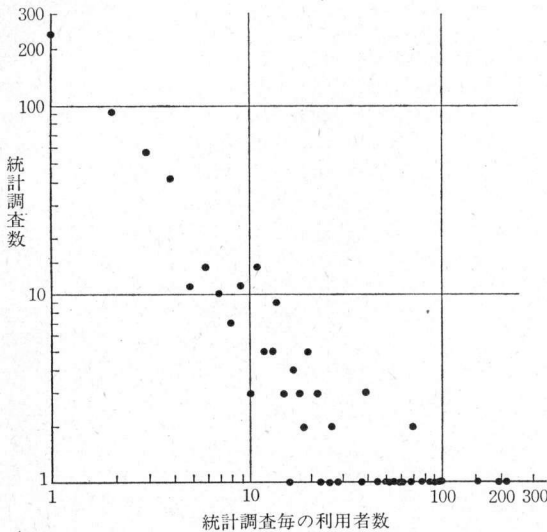
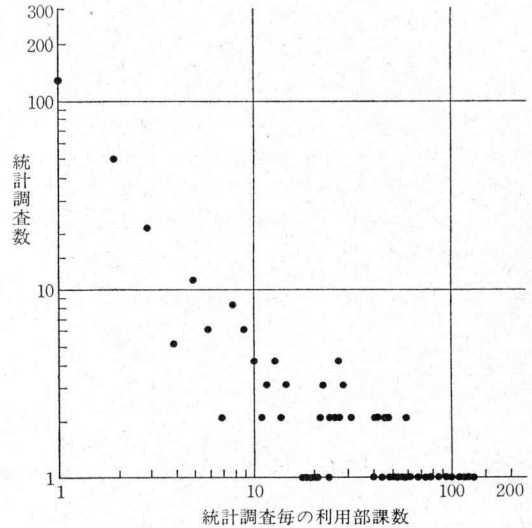
理府統計局「消費者物価指数」、第3位(103部課)：経済企画庁「国民経済計算年報」などである。

この第1図・第2図で示した統計調査が相互にどの様な対応関係になっているかを図示したものが、第3図である。利用者数の相関は、かならずしも高くなく、2つの利用者集団のなかで統計データの利用頻度の相対的地位の差を示している。注目すべきことには、図の欄外に表示した当該利用者集団のなかでのみ利用されている利用者数別統計数のなかに、当該利用者集団の利用統計の中では相対的に利用者数の多い統計調査が含まれている。

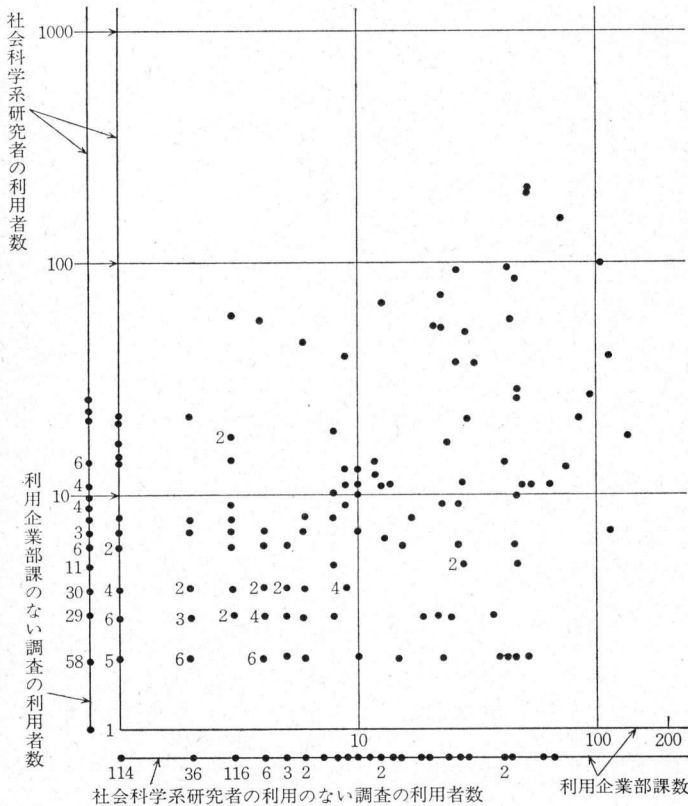
2つの利用者集団の間での利用されている統計の種類の違いは、両者の間の統計利用の目的の差を反映している¹⁰⁾。民間企業の場合特定プロジェクトと経常業務のいづれに利用しているかの点では、重複記入を許しており、前者が総点数に対し23.03%、後者が82.67%である。そ

9) この調査の設計・集計処理には、松田も参加した。研究会の委員長竹内啓東京大学経済学部教授と全国統計協会連合会の好意により、報告書(全国統計協会連合会[1984])編成後さらに再集計し検討する機会に恵まれた(森・周防・松田[1984]、松田・周防・森[1984])。

10) 具体的な調査名の対応表は省略した。松田・周防・森[1984]を参照。

第1図 統計調査毎の利用者数と統計調査の数
(社会科学系研究者)第2図 統計調査毎の利用部課数と統計調査の数
(東証1・2部上場企業の部課)

第3図 統計調査当りの社会科学系研究者と企業の部課数



の利用の頻度は、毎日・週単位・月単位で38.37%を占めている。その結果として、毎月調査報告される小規模標本調査と景気動向に関連した予測データの利用頻度が高くなっている。これに対して、大学の社会科学系研究者の場合には、詳細集計されている3年内至5年に一度の大規模標本調査の利用頻度が高くなっている。さらにデータの週及の利用度は、民間企業の場合5年以内が57.95%、10年以内が17.46%、10年以上が14.25%である。センサスまたは大規模標本調査を2回以上比較して利用するといった利用法が比較的小さいことを含意している。

従ってデータベースの購入が、直接的経費節約効果を持つことが明らかな場合には、すぐ購入しうる企業にとって適切なデータベースと、大学の研究者の側からの需要度の高いデータベースとは、その内容が著しく異なっていると想定される。現行のNEEDSを始めとする各種市販のデータベースが企業を需要者と想定して設計されている以上、その内容がさして大きな購買力をもたない大学の社会科学系研究者を

念頭に置いて大幅に組替えられる可能性は低いといえる。学術情報としての統計データベースは、極めて現在の需要の高い現行社会経済統計が対象であったとしてもそれ自体独自の設計が必要であると想定される。

双方に共通な情報としては、統計データの書誌情報(内容検索も可能な広義なもの)データベースであると思われる。ただその場合であっても、過去100年に互る少なくとも明治以降の統計データに関心のある場合と、近々十年前後に関心のある私企業の場合とでは、遡及可能な期間の点で大きく異るといえる。

松田芳郎・周防節雄

(一橋大学経済研究所)

参 照 文 献

[1] Barnes, Barry [1982] *T. S. Kuhn and Social Science*. London (Theoretical Traditions in the Social Sciences. Anthony Giddens ed.).

[2] Baumol, W. J. & Matityahu Marcus [1973] *Economics of Academic Libraries*. Prepared for Council on Library Resources by Mathematica, Inc. Washington D. C.

[3] Brittain, J. M. [1970] *Information and its Users. A review with special reference to the social sciences*. Bath.

[4] Clark, Terry Nichols [1973] *Prophets and Patrons. The French University and the Emergence of the Social Sciences*. Cambridge (Mass.).

[5] Cummings, William K. [1972] "The Changing Academic Marketplace and University Reform in Japan" PhD dissertation, Harvard University (W. K. カミングス『日本の大学教授』(岩内亮一・友田泰正訳)1972, 至誠堂)。

[6] Gershuny, Jonathan [1983] *Social Innovation and the Division of Labour*. Oxford (The Library of Political Economy).

[7] ——— & I. D. Miles [1983] *The New Service Economy. The Transformation of Employment in Industrial Societies*. London.

[8] Helsey, A. H. & M. A. Trow [1971] *The British Academics*. Cambridge (Mass.).

[9] Keylor, William R. [1975] *Academy and Community. The Foundation of the French Historical Profession*. Cambridge (Mass.).

[10] Kuhn, Thomas [1962, 1970] *The Structure of Scientific Revolutions*, 2nd ed., enlarged. Chicago (International Encyclopedia of Unified Science 2-2).

[11] ——— [1977] *The Essential Tension. Selected studies in scientific tradition and change*. Chicago.

[12] Menard, Henry W. [1971] *Science: Growth and Change*. Cambridge (Mass.).

[13] Price, Dreek J. de Solla [1961] *Science since*

Babylon. New Haven.

[14] ——— [1963] *Little Science, Big Science*. New York.

[15] Schmookler, Jacob [1966] *Invention and Economic Growth*. Cambridge (Mass.).

[16] Shinn, Terry [1982] "Scientific Disciplines and Organizational Specificity: the social and cognitive configuration of laboratory activity," in *Scientific Establishments and Hierarchies. Sociology of the Sciences*.

[17] Stigler, George J. [1965] *Essays in the History of Economics*. Chicago.

[18] ——— & Claire Friedland [1975] The Citation Practices of Doctrates in Economics. in *Journal of Political Economy*, 83 in Stigler [1982].

[19] ——— & ——— [1979] The Pattern of Citation Practices in *Economics, History of Political Economy*. 11 in Stigler [1982].

[20] ——— [1982] *The Economist as Preacher and Other Essays*. Chicago.

[21] 藤野正三郎・畑中康一 [1982] 「エコノミストの労働市場——個票にもとづく分析」『経済研究』33-3。

[22] 原田勝 [1974] 「ビブリオメトリックスの方法とその応用」*Library and Information Science*. 12.

[23] 古瀬大六・松田芳郎(編) [1983] 『社会科学分野におけるデータのデータ構造の分析』。

[24] 倉林義正・松田芳郎 [1981] 「研究・開発の社会会計」『経済研究』32-2。

[25] 松田芳郎 [1971] 「日本経済の計量分析」山田勇編『計量経済学講義』所収。

[26] 松井幸子 [1979] 「所在源情報付き書誌情報データベースの管理——自然科学と対比した人文・社会科学分野でのファイルの分割の手法」『図書館短期大学紀要』16。

[27] 森敬 [1974] 「計量経済学とコンピュータ, I, II, III」『情報処理』15-1, 2, 3。

[28] ——— [1977] 「計量経済学」『bit』臨時増刊。

[29] 桜田忠衛 [1983] 「わが国の経済学分野におけるビブリオメトリックス——その概観と引用分析手法による一事例」『経済資料研究』17。

[30] 佐藤和夫 [1971] 「科学者生産性の分布——不比例効果法則の応用」『季刊理論経済学』22-2。

[31] 新堀通也 [1965] 『日本の大学教授市場』——(編) [1984] 『大学教授の総合的研究』。

[32] 高多亨 [1984] 「理論経済学分野における文献情報生産とその利用」『ドクメンテーション研究(1), (2)』34-5, 6。

調査結果報告書関係

[1] 文部省大学学術局情報図書館課 [1972] 「人文・社会科学学術情報の流通・利用の実態調査結果報告書」([3] および『学術月報』25-2 参照)。

[2] ——— [1974] 「自然科学関係学術情報の流通・利用の実態調査結果報告書」([4] および『学術月報』26-10)。

[3] 杉村優・佐藤隆司 [1971]「人文・社会科学学術情報の利用実態に関する比較研究——情報利用研究の展開のための予備的考察〔1〕」『図書館短期大学紀要』5。

[4] 桜井宜隆・—— [1973]「学術情報の利用実態に関する比較研究——情報利用研究のための予備的考察〔2〕」『図書館短期大学紀要』7。

[5] 全国統計協会連合会 [1984]『統計利用の促進に関する調査研究報告書』(委員長竹内啓)。

[6] 森博美・周防節雄・松田芳郎 [1984]「民間企業における統計の利用実態」『第52回日本統計学会講演報告集』。

[7] 周防節雄・松田芳郎・野島教之 [1983]「アンケート調査の集計処理のアルゴリズム——社会経済統計データベース需要動向調査を例として」『第51回日本統計

学会講演報告集』。

[8] 松田芳郎(編) [1984]『日本の社会経済統計データベース需要動向調査結果報告書(概要編)』(統計資料シリーズ: No. 28)。

[9] ——・周防節雄・森博美 [1984]「個別統計から見た統計利用の実態——企業と大学」『第52回日本統計学会講演報告書』。

[10] 周防節雄・松田芳郎 [1984]「大学の社会科学系研究者による統計利用の実態」『第52回日本統計学会講演報告集』。

[11] 船崎武男 [1984]「都道府県における統計利用の実態」『第52回日本統計学会講演報告集』。

[12] 地方自治情報センタ [1984]『情報資源の有効利用に関する調査研究報告書』。