

補論 | アジア諸国のマイクロ・データ 公開の現状

1 はじめに

先進国別のマイクロ・データ公開の現状は第1章で詳しく説明された通りであるが、アジアの相当数の国でもマイクロ・データが収集され、それをを用いた実証研究がなされている。筆者の知る限りでも、日本、大韓民国、中国、台湾、フィリピン、ヴェトナム、タイ、マレーシア、インドネシア、ネパール、インド、パキスタン、バングラディシュ、スリランカなどに関するマイクロ経済分析は経済学の専門誌、世界銀行刊行物、研究書などの形で発表されている。

しかしながら、アジア諸国のマイクロ・データ公開の現状を正確に把握できるような網羅的なサーベイや研究があるわけではなく、断片的に集めた研究論文や資料からその状況を拾い上げて行くしかない。本補論の内容も極めて限定的なものであることをあらかじめお断りしておきたい¹⁰⁾。

先進国の例からも明らかのように、マイクロ・データは主として人口センサスや家計調査を中心に公開されている。この事情はアジア諸国についても同様である。いうまでもなく、人口統計、とりわけ、全数調査による人口セン

10) ミクロデータが政府レベルでどのように調査され、利用されているか、現時点で不明なアジア諸国には、(旧、現)社会主義国や戦乱で国家が混乱している、アフガニスタン、ラオス、カンボジア、ミャンマー、モンゴル、北朝鮮がある。また、フィジー、キリバス、モルジブ、パプアニューギニア、ソロモン諸島、トンガ、バヌアツ、西サモア等の太平洋諸島あるいはインド洋諸島の国々、および旧ソ連邦から独立した中央アジア5カ国、カザフスタン、キルギス、ウズベキスタン、タジキスタン、トゥルクメニスタンにおける政府マイクロ・データ統計の実態についても掌握していない。

サスは、社会経済状況の基礎指標として、最も重要な統計であり、かつ実質的に唯一の人口統計となっている場合が多い。しかし、この人口センサスでさえ、多くの国では10年毎にしか実施されておらず、経済的困難や内戦などで調査が延期されたり、一部地域の調査が出来ないということも珍しくない。最も基本的なマイクロ・データである人口センサスがこのような状態にあるとすれば、それを基に標本抽出される各種のマイクロ・データ調査の統計的有意性にも疑問が残ることは否めない¹¹⁾。

アジア諸国の中には、発展途上国の立場からそろそろ先進国の仲間入りしようかという国も含まれているが、全体的な経済発展の現状では、製造業などの第2次産業は全体の10-20%程度に止まり、第1次産業、とりわけ農業が40%以上を占めている。このような状況では、マイクロ統計調査が農業や農家に重点がおかれるのは当然であろう。しかし、アジア経済は急速な発展を見せており、農村部における家内工業、あるいは、中国における郷鎮企業の増加など経済活動として極めて面白い現象が起こっていることも事実である¹²⁾。これらの現象を適切に捉えるためには、企業統計を整備する必要があるが、会計制度、税制などの不備もあり、マイクロ・レベルでの企業活動の把握は遅れており、当然、マイクロ・データとしての公開も行われていないことも指摘しておきたい。

発展途上国、とりわけアジア諸国のマイクロ統計調査、開示および利用に関

11) 人口センサスと個票データの利用については早瀬 [1998] を参照。早瀬は5年または10年に不完全にしか行われぬ人口センサスに代わる人口標本調査としてUSAIDの援助で実施されている「人口保健調査」(Demographic and Health Survey=DHS)を紹介している。これは世界保健機構(WHO)や国連などの資金でアメリカの調査機関マクロ・インターナショナルが、途上国約60カ国(アフリカ33カ国、アジア15カ国、ラテンアメリカ13カ国)の統計局、厚生省との協力の下に実施し、現在も継続中の標本調査である。この調査結果は個票データとして公開されている(問い合わせ先は<http://www.macrint.com/dhs>である)。

12) 中国の農村在来企業である郷鎮企業の存在意義についてはMatsuda [1990] を参照。

して主要な問題点は以下の通りである。

- (1) 情報開示に関する統計法が不備である。
- (2) 統計調査当局の標本設計、調査設計、調査実施、調査結果の記入、集計、管理の体制が不十分である。
- (3) ミクロ統計を利用した政策立案のための分析能力、分析道具、分析に対する評価が不足している。

これらの問題点は個別に独立したものではなく、相互に依存して起こっており、その解決策も包括的なものでなければならない。

まずミクロ統計に関しても、国民所得体系（SNA）のように国際比較が可能になるような標準化を進める必要がある。後述するように、世界銀行が行っている生活水準指標調査（LSMS: Living Standard Measurement Survey）は途上国の家計調査の標準化に大きな影響を与えている。しかし、LSMS 調査がカバーしている国はまだ少なく、とりわけアジア諸国では5カ国を数えるにすぎず、この調査がさらに広範に実施されることが望ましい。とはいえ、世界銀行としては、限られた予算の中でプライオリティーの高い国に対してLSMS 調査を実施する方針であり、相対的に発展したアジア諸国が対象に選ばれる可能性は低いかもしれない。その場合、世界銀行に限らず、先進国政府や独立研究所、大学との間で、同様のプロジェクトを立ち上げ、国際基準に順じたミクロ統計調査のノウハウを習得する努力をすべきである。このようなプロジェクトを通して、標本設計、調査設計、調査実施、調査結果の記入、集計、管理の体制が標準化されていくものと期待される。いうまでもなく、このような標準化のためには、各種の国際会議、研修、勧告書などを通じて国際比較が可能となるような標準的な調査項目の設定やその検討が統計関係者の間で継続的に行われる必要がある。

また、世界銀行のLSMS 調査や同様の独立研究機関が実施したミクロ統計調査に関しては、研究者の利用が前提になっており、公開も比較的自由である。このような国際共同プロジェクトを通して、様々な研究が行われ、そ

れが政策立案に用いられたり、調査自体の不備を発見することによって、次回からの改善につながることもあるだろう。

このような経験を通して、途上国政府の中にも、マイクロ・データ公開の意義に気づき、自国政府が調査したマイクロ統計も公開する方針を打ち出すようになった国もある¹³⁾。

2 公開の現状

公開されている発展途上国のマイクロ・データの中で、先ず挙げなければならぬのは、世界銀行が実施している生活水準指標調査 (LSMS) である¹⁴⁾。世界銀行は1980年頃から活動を開始し、1985年以後、20以上の発展途上国で調査を行ってきた¹⁵⁾。サンプルサイズはおおよそ2,000-5,000家計であるが、ルーマニアでは36,000家計の大調査が行われている¹⁶⁾。調査内容は家計質問票 (household questionnaire)、コミュニティ質問票 (community questionnaire)¹⁷⁾、価格質問票 (price questionnaire) に分かれており、かなり詳細な情報を収集している。

この調査の目的は、発展途上国の消費、所得、健康状態、教育、住環境など生活水準に関する広範なデータを集めることの他に、これまで大規模標本

13) 例えば、フィリピン政府はアメリカ合衆国政府の公開形式にならって、1990年人工・住宅センサス、1985-1994年家計調査等いくつかの政府統計について公共利用用にリサンプルしたファイル (Public Use Files) を提供している。また今後は農林水産センサスや農業統計も公共利用ファイルとして公開する方針である。このファイルは海外の研究者も利用可能である。

14) 世界銀行の生活水準指標調査については、Grosh and Glewwe [1998]、黒崎 [1998]、Blank and Grosh [1999] 等を参照されたい。またこのLSMS調査のウェブサイトには利用可能な発展途上国のマイクロデータの内容 (調査票、調査方法、データフォーマット) や入手手続きについての情報が載っている。そのうち、幾つかのデータは簡単な手続きを取れば、オンラインで即座にダウンロード出来る (<http://www.worldbank.org/lsm/lsmshome.html>)。

家計調査をあまり実施できなかった国に、標本の設定、調査票の設定、データの収集、データの編集、データの分析、政策提言への利用にまで及ぶ調査方法を伝えるという教育的配慮もある。さらに、これらのデータは原則的に個票データの形で公開されることになっており、多くの研究者、政策関係者に利用されている。

アジア諸国に関する生活水準指標調査 (LSMS) はこれまでパキスタン (1991)、ヴェトナム (1992/93)、キルギスタン (1993)、カザフスタン (1996)、ネパール (1996) で行われてきた。例えば、ヴェトナムにおける調査は以下のような内容になっている¹⁵⁾。

-
- 15) これまで LSMS 調査やそれに先立つ調査に基づいて多数の優れた研究が生まれてきた。例えば、Kakwani [1993 a, b], Anand [1983], Glewwe and Hall [1994], Grootaert [1995], Lanjouw and Ravallion [1995], Newman, Jorgensen and Pradhan [1991], Sahn and Sarris [1991] 等は所得分布と貧困に関する研究で大きな成果を挙げた。Alessie et al [1992], Angrist and Lavy [1997], Glewwe [1996], Hoddinott [1996], Jacoby [1993], Newman and Gertler [1994], Stelcner et al. [1989], van der Gaag and Vijverberg [1988], Vijverberg [1993] は労働市場と教育投資に関する研究を行っている。Gertler and Glewwe [1990], Gertler and van der Gaag [1990] は医療・教育費用の分析を行っている。Glewwe et al [1995], Glewwe and Jacoby [1994 a, 1994 b], Handa [1996], Jacoby [1994], Levy [1996] は教育投資の効果に関する研究を行っている。Cox and Jimenez [1992], Deaton [1992] は貯蓄と消費に関する研究を行い、Deaton [1989], Higgins and Alderman [1997], Jacoby [1995], Thomas [1994] は性差別の問題に取り組んだ。これらの成果は LSMS 調査を用いた研究のほんの一部であり、アジア諸国に関する研究は遅れているが、中には、マイクロ・データが公開されるまでは、到底考えられなかった新しい研究が含まれており、いかにマイクロ・データの公開が学問研究に多くの成果をもたらすかを物語っている。とりわけ、Kakwani [1980] は所得分布に関する研究で、Deaton [1997] は発展途上国のマイクロ・データ分析における計量経済学手法に関する研究で、金字塔的業績を挙げたと言える。
- 16) 調査対象となった家計が、どのような標本抽出方法に基づいて抽出されているかは、それぞれの国によって事情が違ふようであり、利用に際しては、個別に確認する必要がある。
- 17) 時として、コミュニティー質問票に代わって施設質問票 (facility questionnaire) が入れられることもある。

ヴェトナムでは国連開発計画（UNDP: the United Nations Development Programme）の資金により1992年と1993年に標本数4,800家計を対象に行われた。標本抽出はヴェトナム政府統計局（The General Statistical Office）の人口センサスに基づいて三段階の無作為抽出が行われた〔詳細はThe World Bank（1998）を参照〕。全国約10,000村から選ばれた150村（都市部30村，地方120村），4,800家計の内，わずか3%に当たる155家計を入れ替え，4家計のみが調査を拒否したと言われている。コミュニティー質問票は地方120村全てで集められ，価格質問票は全国150村中148村で集められた。各家計には最大5桁の識別番号が付けられている。

調査は国家計画委員会（The State Planning Committee）と統計局の共同プロジェクトとして行われ，実際には1チーム6人の15チームが全国に分散して，1家計当たり2回の訪問で調査票に記入するという体制をとった。

この調査はヴェトナム政府にとっては色々な意味で大きな変化をもたらしたと言われている。すなわち，(1) 調査票入力を全てコンピュータで行った。(2) 政府職員に対して調査トレーニングを行った。(3) 初めて調査用統計ソフトが用いられた。(4) 調査終了後6ヶ月後には統計分析がヴェトナム人研究者によって行われた。その結果，ヴェトナム政府はこの調査の意義を十分認め，1997-98年には国連開発計画とスウェーデン国際開発庁（SIDA: the Swedish International Development Agency）の資金援助により2回目の調査を行った。

ヴェトナム政府はこのヴェトナム生活水準指標調査（VNLSS）は外部の研究者に対して公開しているが，その他の政府マイクロ・データ統計は一切公表していない。

それに対して，パキスタン政府，インド政府は，マイクロ統計調査を世界銀行のLSMS調査と連動して実施したり，調査に対する援助を受けるなどし

18) この調査については，The World Bank [1998]，Blank and Grosh [1999] 等を参照。

て、1990年代以降の家計調査等に関しては、研究者が利用目的を明示化した目的外申請を出し、それが認められれば公開するという方針を打ち出している。

世界銀行のLSMS調査ほど大々的なものではなく、小規模であるが、調査目的、項目などに特色をもたせた調査が、主として農業経済に関する研究機関によって行われている。

インドのデカン高原にある国際半乾燥熱帯作物研究所(ICRISAT: International Crops Research Institute for the Semi-Arid Tropics)が集めた村落調査データは広範に利用されている。このデータは3ヶ村104家計について9年間(1975-1984)に亘り調査された詳細なパネルデータを含んでおり、発展途上国の農村経済活動のダイナミックな側面を捉えるのにうってつけのデータである。標本としては地域的に限定され、しかも小規模調査であるにもかかわらず、多くの研究者がこのデータを利用し、農村経済におけるリスクと保険機能に関するマイクロ経済分析の新しいフロンティアを開拓してきた¹⁹⁾。

また、ワシントンDCにある国際食料政策研究所(IFPRI: International Food Policy Research Institute)でも家計に関する詳細なマイクロ・データを収集し、公開している。

シカゴ大学のタウンゼンド(Townsend)教授は生活環境が厳しく、従って、経済上のリスクの高いタイの農村において、生産変動や所得変動に対する家計の対応などについての聞き取り調査(field research)を自ら組織し、ユニークなマイクロ・データ・セットを蓄積、整備しようとしている²⁰⁾。このプロジェクトは現在進行中のものであり、全体像はまだ見えてこないが、概

19) 例えば、Rosenzweig [1988], Rosenzweig and Stark [1989], Walker and Ryan [1990], Rosenzweig and Wolpin [1993], Townsend [1994], Ligon [1998]等を参照。

20) これまでの成果については、Townsend [1995, 1999]を参照。

要は Townsend (1999) に要約されている。すなわち、標本抽出は過去のタイ社会経済調査 (SES: the Thai Socio-Economic Survey) に必ず含まれている4県を第一段階として選び、その4県から無作為に12郡を選ぶ(計48郡)。さらに、各郡から無作為に4村(計192村)を選ぶ。そして、最終的に各村から無作為に15家計を選ぶという方法をとった(計2,880家計)。調査は選ばれた家計に対してだけでなく、192村の村長、262の農業銀行支店 (BAAC: the Bank for Agriculture and Agricultural Cooperative)、389の地域金融機関に対して保険機能や貸出に関する質問を行なった。第一回調査は1997年4-5月に行なわれた。さらに、パネルデータを構築する目的で、各県から4村を選び、その村に関しては、さらに30家計を加え、合計45家計(計720家計)に対して毎月調査を実施している。

調査内容は、農地の保有状況、場所、質、作物、用水の有無など生産情報の他に、農村金融の不完全性と流動性制約、相互扶助的擬似保険の機能などについて細かく調べている。

また、このプロジェクトではタイ政府が行ったマイクロ・データ統計を入手し、自らの調査の補完的資料として利用している。具体的には、過去8回の全てのタイ社会経済調査 (SES)、過去5回の全ての地域振興局 (CDD: the Community Development Department) の村レベル全国調査、過去全ての県レベルの所得データを保有し、さらに人口センサスと労働力調査を加える計画がある。

このプロジェクトはタウンゼンド教授の個人的なイニシアチブの下に行なわれているもので、一般公開されているものではないが、共同研究、研究補助などの形で、研究者の間での利用が徐々に増えてきている。

社会主義国でのマイクロ・データの公開はほとんどないが、中国では国際機関や先進国の研究機関と共同研究という形で、資金援助、技術支援を受ける見返りにマイクロ・データを公開するという方法がとられることがある。

例えば、国家統計局 (SSB: the State Statistical Bureau) による農村住

戸調査 (RHS: the Rural Household Survey) を世界銀行のエコノミストと共同で利用し、中国南部の農家の生活水準や貧困ラインを推計するというプロジェクトがある²¹⁾。

また、中国社会科学院 (the Chinese Academy of Social Sciences) と欧米の政治社会研究に関する大学共同研究機関 (ICPSR: Inter-university Consortium for Political and Social Research) との共同研究の形で、1988年度家計標本調査 (the 1988 Household Sample Survey) が行われた²²⁾。この調査は国家統計局 (SSB) の1988年度農村住戸調査 (RHS) の67,186家計から10,258家計 (構成員51,352名)、同じく国家統計局の都市住戸調査の34,945家計から9,009家計 (構成員31,827名) を、対称等距抽様方式 (3段階ジグザグ系統抽出法) により選びだし、1989年3-6月ごろにかけて実施された。調査の目的は、中国の都市と農村の所得分布を測定することであり、調査内容も現金所得、実物所得、公的所得移転、自家消費農産物、帰属家賃など広義の所得の他に、世帯属性、生産状況、資産保有、住宅環境、食料消費などが含まれている。このデータはミシガン大学に設置されている大学共同研究機関 (ICPSR) に申請し、一定の手続きを踏めば利用できる。

ここに挙げたのは、中国のマイクロ・データを使った研究のほんの一部であり、欧米及び日本の多くの大学、研究機関が中国側のカウンター・パートと共同で調査、研究を行っている。

3 利用例の紹介

今回の「文部省科学研究費補助金：特定領域研究(A)：マイクロ統計データ統計情報活用のフロンティアの拡大—マイクロデータによる社会構造解析—」

21) このプロジェクトと農村住戸調査の詳細については Chen and Ravallion [1996], Jyotsna and Ravallion [1999] を参照。

22) この調査については Griffin and Renwei [1993] を参照。

でも二つのアジア諸国のマイクロデータセットを扱い、報告書を作成した。すなわち、「スリランカ家計調査（1990/91年）のデータセット作成に関する解説書」（1998）と「中国農業部固定観察点調査データに関する検討」（1999）の二つである。

3.1 スリランカ家計調査（1990/91年）

本調査はスリランカ政府による統計調査である。今回の利用に関しては、スリランカ政府センサス統計局と一橋大学経済研究所附属日本経済統計情報センター教授（当時）松田芳郎との間で交わされた契約に基づき、スリランカ政府センサス統計局のご好意により利用可能になったものである²³⁾。

ここで調査の概要とそれに基づいて作成したデータセットについて説明しておきたい。

この調査は1990年6月から1991年5月までの期間に実施されたもので、スリランカ・センサス局が実施した第6番目の世帯調査である。調査方法は層化2段階抽出法に基づいている。1981年国勢調査における調査区を第1次抽出単位とし、家屋単位を第2次抽出単位としている。

調査の内容は3つのセクションに分かれており、収入および属性情報に関しては世帯員単位、消費状況等に関しては世帯単位に調査を実施している。おおまかには次のような区分になっている。

セクションⅠ 世帯員の属性情報

- I.1. 世帯員の年齢、性別、世帯主との関係、人種、宗教、学歴、婚姻関係等
- I.2. 世帯員の経済活動
- I.3. 土地所有/賃貸状況

セクションⅡ 家計の消費状況

23) スリランカ家計調査に関する先行研究には Deaton [1981] がある。

II.1. 飲食料品に関する支出状況

II.2. 耐久消費財の購入状況

II.3. 下宿人等による消費状況

II.4. 預金や負債に関する状況

セクションIII 世帯員の収入に関する情報

III.1. 賃金収入等の状況

III.2. 雇用主等としての経済活動

この他にコントロール・データ・セクションがあり、調査員の訪問日時や調査所要時間、ならびに調査結果（調査完了、回答拒否など）について記録している。

本調査は16種類のデータ・ファイルに分割整理されている。データのunit単位としては世帯単位と世帯員単位の2種類あり、単位として統一されていない。また、データは世帯や世帯員単位だけではなく、item単位となっているデータ・ファイルも存在する。従って、マイクロ経済分析を行うためには、このデータ・ファイルを世帯あるいは世帯員単位に整理統合する必要がある。特定領域研究では、マイクロ経済分析に必要な世帯単位、世帯員単位のデータ・ファイルの作成を行い、CD-ROM形式でデータ・ベース化した²⁴⁾。

データ作成上の問題点としては、現在、国際的な基準となっている世帯単位、世帯員単位の調査となっていないために、そのままのデータでは、家計の経済行動を計量経済学的に分析することが出来ないということが挙げられる。これは、調査を設計した時点では、現在のようなデータの利用方法を想定していなかった上に、その後も、基本的な変更を加えずに調査を続けてきた結果であると思われる。特定領域研究では、1990/91年度の家計調査を経済分析可能な形に変換したが、今後は1995/1996年度に行われた家計調査も

24) 詳しくは、松田 [1998] を参照。

同様の手法で変換し、さらには、調査設計自体の改善に向けた政策提言を行うことを視野に入れている。

3.2 中国農業部固定観察点調査

本調査は、本来、利用目的の限定された政府統計調査であるが、今回の利用に関しては、辻井博教授（京都大学）と松田芳郎教授（一橋大学名誉教授、東京国際大学）が、中国農業部農村経済研究中心の張文宝、曹力群両氏と協議を重ね、共同研究の一環として立ち上げたものである。

ここで調査の概要と特定領域研究で行った研究について説明しておきたい²⁵⁾。中国政府農業部が実施した固定観察点調査は中国農戸²⁶⁾を対象とした標本調査であり、調査対象は31自治区323県に及んでいる。抽出方法は2段階有意抽出法を用いている。すなわち、第1次抽出として、各県毎に「一人当たり収入」や「地形要素」を考慮して、「良」(30%)、「中」(40%)、「不良」(30%)に分け、「中」に属する村のうち、少し上のレベルの1村を抽出する。第2次抽出として、抽出した村から100戸を抽出する。農戸の抽出に当たっては、収入順に並べた名簿から、系統抽出法を用いて抽出する。選ばれた農戸は継続的に同じ調査に繰り返し答える。移転等により調査対象の農戸がなくなった場合には、類似の農戸を調査対象に入れ、調査農戸数が減少しないように配慮している。

実際の調査では、村毎に指定された補助調査員の指導の下に、農家が日記形式で調査票に記入することになっている。それを補助調査員が月毎にチェ

25) 中国国家統計局も中国農戸調査を実施しているが、無作為抽出による「標本調査」であり、毎年標本が変化するのに対して、農業部固定観察点調査の標本は1985年に選択された同一農戸を10年間に亘って継続的に調査したものである。調査項目には重複するものが多いが、農業部固定観察点調査には家族類型や村幹部の有無など社会的な調査項目も含まれている。

26) 農戸とは法律上の定義であり、同じ家屋に居住している常住家族をまとめた単位である。

ックし、回収することになっている。

固定観察点調査で電子データとして整備されているのは1986年から1995年までのデータ(1992, 1994年を除く)である。具体的には、固定観察点調査は、調査票、調査内容から(1)1986-1991年、(2)1993年、(3)1995-1997年の3つに区分される。つまり調査内容が期間によって変更されているがために、それぞれの期間をつなぎ合わせたパネルデータの作成に当たっては、3期間に共通する項目、各期間に独自の項目、2期間のみに共通する項目に分けて、全体をカバーした共通調査票を作る必要がある。

主な調査内容は農戸における経営、労働等の投入状況、農戸の所得、資産保有状況、農戸の消費状況、農村の発展状況、土地・水面利用状況等である。調査内容は3期間で多少変更されているが全体の構成はおおよそ次のようになっている²⁷⁾。

- I. 世帯属性および経済活動：世帯員の職業、および職歴(軍属、村幹部、共産黨員)、労働者数、教育程度(識字、学歴)、雇用形態
- II. 経営土地、水面の保有状況：経営耕地面積、経営林地面積、経営水面面積
- III. 固定資産状況：家畜、農具、農漁業機械、工業機械、運輸機械、自動車、宅地面積、住宅家屋価格
- IV. 農林牧漁業生産状況：農地面積、農産物生産量(金額)、木材伐採量(金額)、家畜生産量(金額)、水産物生産量(金額)、労働投入量、その他投入量
- V. 経営会計収支：経営収入、各種費用支出、外部労働収入、残高
- VI. 家計支出：消費支出(細目、特に食料支出)、貯蓄、貸出、負債、耐久消費財保有

この調査は、同一農戸を多年度に亘って追跡することで、農戸の異時点間

27) 調査内容の違いなど詳しい報告は松田 [1999] を参照。

の最適化行動とクロスセクションの分布情報等を同時に検討することが出来るという統計上非常に魅力的な特性を備えたものになっている。

しかしながら、この調査にもいくつかの問題がある²⁸⁾。第一に、一般にあるデータ入力エラーの他に、調査員がパネルデータの意味を理解せずに農戸の識別番号を年毎に変更しているケースが見られる。農戸識別番号が一致していれば、パネル化はその識別番号を異年次データ間でマッチングさせれば簡単に出来るが、識別番号が一致しているケースと不一致のケースに分かれる場合は、マッチングの仕方に工夫が必要になる（適切なマッチング・キーを探す必要がある）。第二に、1992年と1994年については調査が行われておらず、パネル化を行うためにマッチングさせるためには、連続した年の場合とは違ったマッチング・キーを探す必要が出てくる。第三に、予算の関係で、1993年以後のサンプルが大幅に削減されているケースがあり、1986-1996年の完全パネルデータを作ろうとすると、サンプルが減少する可能性がある。

今回の特定領域研究では、上述したような問題点を克服しつつ、集められたデータを無駄にせず利用できるよう、完全パネルデータだけではなく、部分的なパネルデータの作成も意図した、マッチング実験を行った。その結果、調査の精度が村毎、年度毎にかなりばらついているためにパネル化が予想以上に難しいことが明らかになった。それらの限界を適切に考慮しながら、できるだけ多くの観察点を含んだパネルデータを作成することが日中共同研究の当面の課題となっている。

最終的には、このパネルデータ・セットを使った本格的な中国農村のパネルデータ分析を行い、経済研究上の新たな発見や政策立案に役に立つような提言を行うことを意図している。

28) 詳細は松田 [1999]、辻井・松田 [1999] を参照。