

200902003A

厚生労働科学研究費補助金
統計情報総合研究事業

パネル調査(縦断調査)に関する統合的高度統計 分析システムの開発研究

(課題番号H20-統計-一般-003)

平成 21 年度 総括研究報告書

研究代表者 金子 隆一

平成 22(2010)年3月

目 次

I. 総合研究報告

研究代表者（金子隆一）	3
研究分担者（北村行伸）	13
研究分担者（釜野さおり）	17

II. 個別研究報告

方法論的研究

1 21世紀出生児縦断調査における脱落サンプルのもたらす影響の評価 （北村行伸）	27
2 第1～7回21世紀出生児縦断調査の脱落・移動の動向および子育て感の 都市規模別変化 （西野淑美）	57
3 21世紀成年者縦断調査（第1回～第5回）における男性票の脱落者・ 継続回答者の特性に関する分析 （釜野さおり）	69
4 成年者縦断調査における希望子ども数の無回答に関する分析 （守泉理恵）	81
5 パネル調査における脱落の初婚に対する因果効果についての検討 傾向スコアを用いた共変量の調整 （鎌田健司）	97
6 21世紀出生児縦断調査における対象児の特徴：出生動向基本調査 夫婦調査の出生歴に基づく出生児データとの比較 （岩澤美帆・金子隆一・三田房美）	109

分析研究：子育て環境をライフコースの観点から探る

- 7 子どもの肥満に関連した社会・生活環境因子に関する分析
(井出博生)143
- 8 第2子出生タイミングに関する要因分析
(福田節也)165
- 9 中高年者の生活実態の変化に関する分析手法の検討(4)
—健康状態変化とその要因に関する分析—
(石井 太)181
- 10 21世紀出生児縦断調査における婚外出生児の現状
(岩澤美帆・三田房美)197
- 11 若年出産と高齢出産：21世紀出生児縦断調査第1～6回から
(相馬直子)227
- 12 子ども観と育児方針2—第1回～第6回「出生児縦断調査」の分析より—
(元森絵里子)241

Ⅲ. 資料編

- 1 パネル調査（縦断調査）に関する統合的高度統計分析システムの開発研究
[研究過程の流れ]285

※ 本報告書は本プロジェクト研究に基づき、研究代表者の監修により編集されたものであり、
記述に関する一切の責任は研究代表者にある。

研究組織

○ 研究代表者

金子 隆 一 国立社会保障・人口問題研究所 人口動向研究部長

○ 研究協力者

(企画委員会委員 (50 音順)) (敬称略、所属等は平成 22 年 3 月時点による)

阿 藤 誠 早稲田大学人間科学学術院 特任教授

津 谷 典 子 慶應義塾大学経済学部 教授

中 田 正 (株) リソースネット 顧問

(データ管理・分析法研究担当)

北 村 行 伸 一橋大学経済研究所 教授

釜 野 さおり 国立社会保障・人口問題研究所 人口動向研究部第二室長

岩 澤 美 帆 国立社会保障・人口問題研究所 人口動向研究部第三室長

阿 部 彩 国立社会保障・人口問題研究所 国際関係部第二室長

石 井 太 国立社会保障・人口問題研究所 国際関係部第三室長

三 田 房 美 国立社会保障・人口問題研究所 企画部主任研究官

守 泉 理 恵 国立社会保障・人口問題研究所 人口動向研究部主任研究官

鎌 田 健 司 国立社会保障・人口問題研究所 客員研究員

井 出 博 生 東京大学医学部附属病院 助教

藤 原 武 男 国立保健医療科学院生涯保健部 行動科学室長

福 田 節 也 マックスプランク人口研究所 研究員

西 野 淑 美 首都大学東京都市教養学部 助教

相 馬 直 子 横浜国立大学大学院 准教授

元 森 絵里子 明治学院大学社会学部 専任講師

○ ご協力いただいた方々

(敬称略、所属等は平成 22 年 3 月時点による)

廣 松 毅 情報セキュリティ大学院大学 教授、東京大学 名誉教授

(厚生労働省大臣官房統計情報部)

小 野 暁 史 人口動態・保健統計課長

篠 原 一 正 社会統計課長

上 田 響 国民生活基礎調査室長

市 川 隆 之 社会統計課 縦断調査室長

(同部 21 世紀縦断調査担当)

吉 川 英 樹 社会統計課 縦断調査室 室長補佐

松 岡 芳 弘 社会統計課 縦断調査室 室長補佐

巽 恵美子 社会統計課 縦断調査室 縦断調査専門官

ならびに、社会統計課縦断調査室の皆さん

I. 総合研究報告（要旨）

研究代表者 金子 隆一
（国立社会保障・人口問題研究所）

パネル調査（縦断調査）に関する総合的高度統計分析システムに関する開発研究
（平成 20～21 年）

研究代表者 金子隆一 国立社会保障・人口問題研究所

研究概要

本研究は、厚生労働省が各種の施策策定に資する科学的基礎資料を得るために実施しているパネル調査(21 世紀出生児縦断調査、成年者縦断調査、中高年者縦断調査(以下、21 世紀縦断調査))に対し、この調査が対象とする国民生活上の諸事象に関する要因や発生メカニズムの特定、施策効果の測定、ならびに行政ニーズの把握に結びつく総合的な分析システムを構築し、年々蓄積されるデータの速やかで効率的な結果公表に資するとともに、手法開発ならびに分析研究による学術的貢献を目指すものである。

本来、パネル調査は、施策効果測定、行政ニーズ把握等に有効な調査形態であるが、横断型調査とは異なる独自のデータ管理方法と分析法を要する。21 世紀縦断調査は、政府統計初の本格的なパネル調査であるが、それだけに行政目的に適したデータ管理や分析システムの環境は十分ではない。これに対し申請者は平成 16・17 年度および平成 18・19 年度の本研究費補助事業による研究プロジェクトにおいて、当該調査のデータ特性分析、管理システム構築および分析システムの検討を中心とした研究開発を行い、縦断調査分析に必要な基礎的環境の構築を行った。本研究ではその成果を踏まえつつ、年々蓄積される当該調査データについて、(1)データ管理から高度な統計分析手法の適用までを統合化するシステムの開発、(2)標本脱落などの分析上不可欠となるデータ特性の把握、(3) 調査事項を体系化して分析を行う枠組みの提示など、縦断調査分析に必要な研究インフラとしての総合的統計分析システムの構築を目指した。

本年度の研究としては、これまで開発を進めてきたデータ管理・分析統合支援システム（PDA21）ならびにパネル調査情報ベースシステム、さらには分析手法の検討・体系化を進めることに加えて、とくに調査回を重ねるごとに重要性を増す脱落、不詳発生分析結果への影響などデータ特性を正しく把握する研究、またはその結果を事例分析に反映させる研究に重点を置いた。また、昨年度から取り組んできた事例分析の組み合わせによる日本人のライフコース分析枠組みの体系化に向けて基礎事例分析を進めた。本報告書では、分担研究報告書とともに、それらの個々の研究について要約する。これらを含め、現時点における本事業の成果を総覧すると、本事業における個々の研究領域はそれぞれ一定の到達点に接近しつつあると認められ、今後はそれらの成果を短期集中的に統合を図り、一般における利用を前提とした実用化に向けての作業に進む段階に至っていると考えられる。

A. 研究目的

本研究は、厚生労働省が各種の施策策定

に資する科学的基礎資料を得るために実施

しているパネル調査(21 世紀出生児縦断調

査、成年者縦断調査、中高年者縦断調査(以下、21世紀縦断調査)に対し、この調査が対象とする国民生活上の諸事象に関する要因や発生メカニズムの特定、施策効果の測定、ならびに行政ニーズの把握に結びつく総合的な分析システムを構築し、年々蓄積されるデータの速やかで効率的な結果公表に資するとともに、手法開発ならびに分析研究による学術的貢献を目指すものである。

本事業は同様の目標を目指して実施された先行事業(平成16-17年度および平成18-19年度)の後継研究であり、それらの成果として開発されたデータ管理システム、ならびに有効性が検証された統計手法等を踏襲して、当該調査データに関する(1)データ管理から高度な統計分析手法の適用までを統合的に支援するシステムの開発、(2)標本脱落などの分析上不可欠となるデータ特性に関する分析と把握、(3)調査事項の事例分析を実施し、これらを縦断的視点から体系化することによって得られるライフコース分析枠組みの提示など、縦断調査、とりわけ21世紀縦断調査データの分析に有効であり、また必要となる研究インフラとしての総合的統計分析システムの構築を目指している。

そうした目標への過程における本年度の研究目的としては、システム開発関係においては実用化への整備の促進、またデータ特性の把握については、新たに得られたウェイブ(調査回)における脱落、不詳発生状況の継続的分析による変化の捕捉、さらには調査事項のライフコース的体系化において、基礎的事例研究の積み重ねによる日本人のライフコース分析枠組みの構成強化などを目指して行われた。

B. 研究方法

本研究では、平成16-17年度および平成18-19年度の本研究費補助事業による先行

研究において、当該調査のデータ特性分析、管理システム構築および分析システムの検討を中心とした研究開発を行い、縦断調査分析に必要な基礎的環境の構築を行った。本研究ではその成果を踏まえつつ、主として以下の領域についてそれぞれ継続して研究を進めている。

- (1) データ管理・統計分析システムの開発
- (2) パネル調査に関する情報ベースの開発
- (3) 分析手法の確立・体系化
- (4) データ特性の分析・把握
- (5) 事例研究とその体系化

これらの領域の研究は並行して行われているが、本年度の研究においては、上述の年次目標にしたがって、(4)ならびに(5)に関する研究に重点を置いて進めた。

すなわち、「(4) データ特性の分析・把握」として、

- ・ 第1～7回21世紀出生児縦断調査の脱落・移動の動向および子育て感の都市規模別変化〔北村〕
 - ・ 21世紀成年者縦断調査(第1回～第5回)における男性票の脱落者・継続回答者の特性に関する分析〔釜野〕
 - ・ 21世紀出生児縦断調査における脱落サンプルのもたらす影響の評価〔北村〕
 - ・ 成年者縦断調査における希望子ども数の無回答に関する分析〔釜野〕
 - ・ 21世紀出生児縦断調査における対象児の特徴：出生動向基本調査夫婦調査の出生歴に基づく出生児データとの比較〔金子〕
 - ・ パネル調査における脱落の初婚に対する因果効果についての検討〔釜野〕
- の各研究が行われ、それぞれ堅実な成果が得られた(〔 〕内は研究の集約担当)。

また、「(5) 事例研究とその体系化」の領域では、

- ・子どもの肥満に関連した社会・生活環境因子に関する分析 [金子]
- ・第2子出生タイミングに関する要因分析 [金子]
- ・若年出産と高齢出産 [釜野]
- ・子ども観と育児方針2 [釜野]
- ・21世紀出生児調査における婚外出生児の現状 [釜野]
- ・中高年者の生活実態の変化に関する分析手法の検討(4) 一健康状態変化とその要因に関する分析 [金子]

の各研究が行われ、これらもまた分析テーマに関する新たな知見の獲得をとまなう成果を得ている。

[北村][釜野]の分担となる各テーマについては、それぞれの分担研究報告が参照されるものとして、ここでは研究領域(1)～(3)に関する研究進捗の概略と、上記(4)(5)の領域の[金子]が集約を担当した研究について、以下に報告するものとする。

(1) データ管理・統計分析システムの開発

21世紀縦断調査は3つの大規模な調査で構成され、それぞれ毎年新たな調査が行われる。しかも、成年者調査のように4種類もの調査票を用いるものもあり、その調査データの扱いは、通常の調査のノウハウでは追いつかないため独自の仕組みを考える必要が有る。

すなわち、複数年の時系列変化を追うためには、複数回の回答結果を統合して分析を行う必要があり、さらにたとえば夫婦の状況については、個人の調査票をも複数合わせて分析する必要がある。

こうした多数の調査票にまたがる分析を行うためには、これまでの単独の横断調査を分析する場合とは異なる独自の支援システムが必要だから、本事業では21世紀縦断調査に特化したこの種の支援システムの開発を行っている。

当該支援システム(PDA21)は、調査の

一次データの管理から、集計や基礎分析(基礎統計指標や分布指標の算出など)を経て、高度な統計分析(多変量解析、マイクロシミュレーション分析等)までをカバーするものとし、複数の調査票から必要な分析項目を指定することによって自動的に抽出・統合した分析用一次データを生成し、分析ソフトに受け渡すなどの仕組みによって、一連の分析作業を支援するものである。

システム全体はオープンな仕組みを取り入れ、分析ソフトには、たとえば市販の統計分析ソフト(現状ではSAS、SPSS)を用いることが出来る。これは一部のインターフェースを作成することにより、その他の任意のソフトウェアに拡張することが出来るように設計されている。

本年においては、実用化に向けてのインターフェースの再検討、改善などを行った。

(2) パネル調査に関する情報ベースの開発

21世紀縦断調査はわが国における最初の政府による本格的パネル調査であるが、諸外国においてはパネル調査は戦後の早い時期から盛んに行われており、データとともに実施や分析の手法、事例などの蓄積が著しい。しかし、それらの情報を個別に検索することは、その蓄積の量からして不効率であり現実的ではない。

このため、本事業では世界中で行われているパネル調査について、キーワード検索等によって効率的に必要な情報が参照できるような独自の情報ベースを開発している。

これは、すでに諸外国において存在しているミシガン大学のデータバンクと、社会科学系の縦断調査のデータをまとめて所有しているハーバード大学のデータバンク(Murray Center)などを中心として、これまで行われたパネル調査に関する情報を収集し、日本語による検索や情報の提供が可能となるようにしたものである。

本システムは、インターネットを利用して公開することを目指しており、本年の事業では、キーワード検索部分の実用化に向けて整備等を行った。また、効率的なデータ更新の方式について検討を行った。

(3) 分析手法の確立・体系化

縦断調査には、従来から行われている横断調査に対する分析手法に加えて、複数回の時系列データを扱うことから、独自の分析手法が必要となる。パネル型調査が盛んな諸外国においては、こうした手法の発達が著しく、分析事例も豊富であり、さらに現在においても手法の開発が進められていて、有用なものは市販の統計ソフトなどへ反映されることも多い。

本事業では、パネル型データに関する分析手法等についてサーベイを行いレビューとしてまとめており、縦断調査における特定のテーマや目的に応じて適切な分析法が選べるように、それらの手法を体系的に位置づけることを目指している。

たとえば、21世紀縦断調査のテーマの分析にとって重要なものとしては、事象歴分析法(event history analysis) 共分散構造分析＝構造方程式モデル(structural equation model)、マルチレベルモデル(multi-level model)などを挙げることができる。

本年の事業においては、これら分析手法を支援システム(PDA21)に組み込むための仕組みについて検討し、設計を行った。

また、もう一つのパネルデータに対する強力な分析手法として、マイクロシミュレーションを取り上げ、21世紀縦断調査への応用するためのシステム構築に取り組んでいるが、本年は昨年開発した日付・時間管理を行うクラスモジュールを用いた標本モジュールの作成に取り組んだ。

(4) データ特性の分析・把握

本年度の事業においても、引き続き脱落のデータ特性に関する分析を中心として、21世紀縦断調査データの特性の分析、把握に関する研究を行った。ただし、以下に要約する研究以外の研究については、それぞれ後続の分担研究報告を参照されたい。

ここでは「21世紀出生児縦断調査における対象児の特徴：出生動向基本調査夫婦調査の出生歴に基づく出生児データとの比較」に関する研究について報告する。

21世紀出生児縦断調査は、2001年に生まれた子どもの成長を追うことによって、子育て環境の実態把握や子どもに影響を与える要因の特定に活用されている。しかしながら、分析の対象が特定の時期に産まれた集団であるため、結果を解釈する際には、縦断調査の対象児が、過去に生まれた子どもや今後生まれる子どもに比べ、どのような特徴をもっているのかを十分把握しておく必要がある。そこで、国立社会保障・人口問題研究所が実施している出生動向基本調査の結果との比較によって、縦断調査による標本の特徴を分析した。出生動向基本調査は、横断調査であり、第6回調査までの縦断調査と比較可能な6歳未満の子どもの情報を時系列で把握することができるため、縦断調査データとの比較には最適である。

すなわち、出生動向基本調査夫婦票には、妻が現在の結婚で産んだ子どもの出生歴が含まれているが、これにより各出生児ごとに出生年月や両親に関する情報、子育て環境等に関する情報を含んだ出生児データを作成した。そして縦断調査と比較可能な項目、あるいは子育て環境の変化をとらえるために重要な項目について、子どもの生まれ年別の変化を分析し記述した。出生動向基本調査の夫婦票に基づく出生児データは、

調査時点で両親が結婚生活を送っているケースに限られる。従って、厳密な比較を行うため、縦断調査の対象児も、調査時点で両親と同居している日本在住の子どもに限定して比較を行った。

(5) 事例研究とその体系化

本事業では、縦断調査における調査テーマとその分析手法の体系化に貢献するための事例分析を行っている。本年度は、出生児調査による児童の生育、出産・育児に関する状況、ならびに中高年者の生活実態に関する研究を行った。以下に示す3つの研究以外については、それぞれ分担研究報告を参照されたい。

ここでは「子どもの肥満に関連した社会・生活環境因子に関する分析」、「第2子出生タイミングに関する要因分析」、さらに「中高年者の生活実態の変化に関する分析手法の検討(4) 一健康状態変化とその要因に関する分析」について要約する。

まず、「子どもの肥満に関連した社会・生活環境因子に関する分析」について報告する。肥満が生活習慣病の危険因子であることは、数多くの研究が示しているなか、近年、成人期における肥満が、幼児期のAdiposity Rebound (AR)の時期によって予測できることがわかるようになった。ARとは幼児期にいったんBody-Mass Index (BMI、体重/身長²)が最小値を示した後に上昇に転じる現象のことで、ARの時期が早い子どもほど将来のBMI値が高くなると考えられている。現在、ARの生物学的、社会環境的な背景因子を明らかにすることが求められており、本研究は1)日本人の子どもにARが観察されるのかを分析し、2)肥満状態にある子どもの社会・生活環境因子を明らかにすることを目的として実施された。

すなわち、両親ともに日本人であり、

出生時体重、妊娠週数が明らかである子どもを分析の対象とし(分析対象となった子どもの数は男児23,608人、女児21,784人、合計45,392人で、第6回調査での該当者は38,537人であった)、各調査回の調査対象者の身長と体重からBMIを計算した。記述統計、behaviorの軌跡を明らかにするためにNaginが開発したセミパラメトリックな手法を用いた解析、子どもの社会・生活環境因子を被説明変数としてGeneralized Estimating Equation (GEE)を用いた多変量解析を行った。

次に「第2子出生タイミングに関する要因分析」では、第2子出生の生起およびタイミングがどのような要因によって規定されているのかを明らかにすること目的として行われた。

第2子出生はコーホートの出生力水準を決定する重要な人口学的事象であり、近年の夫婦出生力の低下に際しては、有配偶女性における第2子出生確率の低下が著しく進行していくことが予測されている。

そこで「21世紀出生児縦断調査」(以後、出生児調査)の第1回から第6回(パネル1からパネル6)までの個票データを使用して、第2子出生ハザード率についての Kaplan-Meier 生存分析ならびにイベントヒストリー分析を行い、第2子の出生生起を左右する要因について分析を行った。

次に「中高年者の生活実態の変化に関する分析手法の検討(4) 一健康状態変化とその要因に関する分析」について述べる。

中高年世代は、子どもの父母・祖父母になりうるという観点から子育て環境の重要な要素を構成しているが、かれらはまた離職・再就職による就労環境の変化や、親世代の介護負担の発生など、その生活実態の

変化が顕著な世代である。このような生活実態の変化のうち、特に健康状態の変化の分析は、これら中高年世代が今後とも活力ある高齢者として社会参画を続けるために何が必要かを考える上で重要な検討素材となるものと考えられる。そこで、本研究は、中高年世代の健康状態の変化がどのような要因によって引き起こされているかについて、生活実態の変化に関する要因の分析手法の検討を行う観点から、健康状態の変化に焦点を当て、健康状態の変化に関する要因分析を行うことを目的として実施した。

すなわち、第1～3回中高年者縦断調査のデータに基づき、健康状態の変化を被説明変数としたロジスティック回帰分析を行った。これは、前年の健康状態別に2つのモデルとなるが、これを統合的な一つのモデルとして分析するため、離散時間2方向ハザードモデルによるモデリングを試みたものである。

C. 研究成果、ならびにD. 考察

(1) データ管理・統計分析システムの開発

現在開発中の縦断データ管理・分析支援システム(PDA21)は、複数回、あるいは複数調査票に対応するオリジナルのデータセット群から、分析に必要な変数群を指定することによってそれら当該データが抽出し、必要な変数群のみから群から成る分析用の一時ファイルを自動的に生成し、分析用のソフトウェア(通常市販のパッケージ)へと受け渡される。

通常、市販パッケージ等で分析プログラムを作成する際には、データセットからプログラム上の変数(たとえば学歴)に対応するデータを読み込み、カテゴリ(高卒)を表す数値(2)を分析に対して適切な値(各調査に共通したコード)に変換するなどの作業をプログラムとして記述する必要がある。この工程は多大な時間と労力を要

することが普通である。したがって、通常的环境下では、一時ファイル等、多種のデータセットファイルを生成して取り扱うことは、そのたびに読み込みプログラムを作成することになるため、現実的ではない。

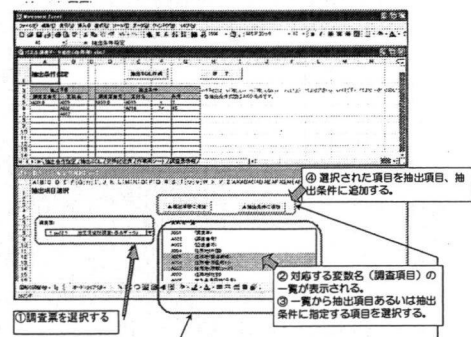
しかし、この工程を自動化することができれば、個々の分析に特化したデータセットを生成することは、計算効率のみならず、分析者にとっても操作が楽になることが期待できる。

とりわけ、調査回が重なり、10回を超えるようになると、不要な計算不可が雪だるま式に膨らむだけでなく、分析者にとっても操作力の限度を超え、誤り等を引き起こす原因となりかねないため、分析用一時ファイル作成は必須となる。

本システムは、この工程の自動化を実現することによって、縦断調査データ全体を一種のリレーショナルデータベースとし、データ管理、分析に必要な操作を支援するものである。

本年においては、昨年に引き続き実装への調整に関するシステムの開発が進められ、インターフェースの改良が行われた。下図は分析に必要な変数を指定する際の画面イメージである。

パネルデータ自動集計・分析支援システム

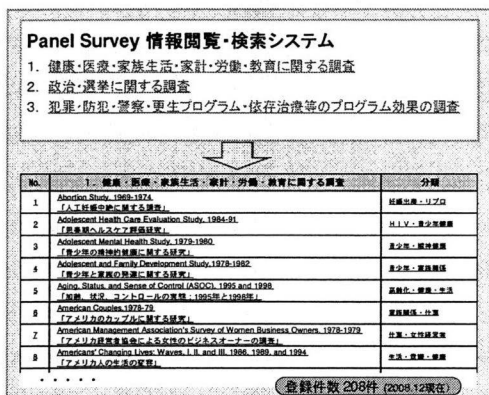


(2) パネル調査に関する情報ベースの開発

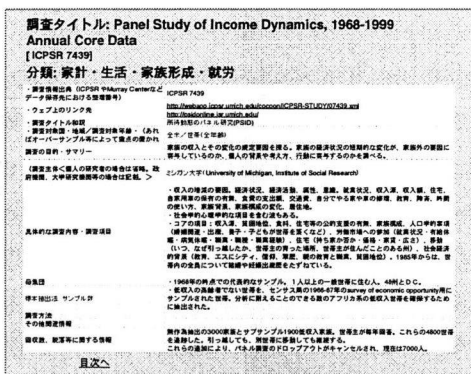
本研究では、先行事業において収集した諸外国のパネル調査に関する情報を追加す

るとともに、ウェブサイトで検索、閲覧できる形式のシステムを開発した。

これはその画面イメージを下図に示す。まず分野別にこれまでに実施されたパネル調査の一覧が表示される。



個別の調査を指定すると、下図のように当該調査に関する詳細な情報が提供される。



また、これらの全情報はキーワードによって検索することが可能であり、該当する調査のリストやここの情報を得ることができる。本年においては、この検索部分の開発、改良が行われた。

(3) 分析手法の確立・体系化

本年の事業においては、パネル型データに関する各種の統計手法については、先行研究における各種手法の特徴に関する検討に基づいて、各事例研究への応用として用いられ、これらを通して個々に有効性の検証がなされている。とりわけ、縦断調査の

回が進むにつれて次第に深刻となると考えられる脱落のデータ特性ならびに統計分析結果への影響を考慮した手法を多く取り入れている。それらについては、事例研究に関する記述を参照されたい。

また、パネル型データの分析に不案内な利用者に対して、分析の支援やガイド資料の提供が行えるような仕組みを作成する基礎として、引き続きパネル型データに関する分析手法の体系化に取り組んでいる。

(4) データ特性の分析・把握

ここでは、「21世紀出生児縦断調査における対象児の特徴に関する出生動向基本調査夫婦調査の出生歴に基づく出生児データとの比較研究」の結果について述べる。

縦断調査の対象児は、横断調査である出生動向基本調査と比較すると、第1子の割合が高い、婚前妊娠による出生が多い、親が再婚である割合が高い、望んだ妊娠による割合が高い、多胎児が多い、親の年齢が高齢化し分散が大きくなっている、親が高学歴化している、父母の就業にパートや派遣などが増えている、育児休業利用者が増えている、祖父母との近居が増えている、親のきょうだい数が減少し、親が「跡継ぎ」役割である割合が高い、といった特徴を持っていることがわかった。

親の年齢の分散化や、婚前妊娠による出生の増加、多胎児の増加など、変化が観察される現象については、こうした条件の違いによって子育て環境にどのような違いがあるのかを把握し、今後の子育て環境の変化の見通しに役立てることができると思われる。例えば、親の年齢の分散化は、親の収入の分散化にもつながる可能性があり、子どもの経済的環境について格差を生む可能性がある。また親への負担が大きい多胎児の増加も、子育て支援のあり方として考慮する必要があるかもしれない。

(5) 事例研究とその体系化

まず、「子どもの肥満に関連した社会・生活環境因子に関する分析」について、その結果についてみると、調査回毎のBMIの平均値は順に12.61、16.33、16.09、15.75、15.48、15.38であり、第6回調査における男児17,656人のBMIの平均値は15.43(95%信頼区間15.41-15.45)、女児16267人の平均値は15.34(15.31-15.36)であり、男女間に有意差があった。

($p<0.001$)第2回調査を最高にBMI値は低下していたが、BMI値の99 percentileは第4回、95 percentile、90 percentileでは第5回が最低値で、以降の調査ではこれらの値は上昇した。同様にして、BMI95 percentile以上および85 percentile以上のグループでは、第6回調査までの間にBMIが上昇、下降、再度上昇するという発育の軌跡が発見された。

つまり、このグループでは、5歳半頃までにARを経験するという発育パターンであることがわかった。GEEによる分析の結果は、出生時のBMI、時間、性、妊娠週数の他に、社会・生活環境因子として子どもの世話をする人、同居の家族状況、起床時刻、テレビゲームで遊ぶ時間、おやつの摂り方が有意であることを示していた。

この研究からは、1)日本人の子どもにもARと呼ばれるBMI値の変化があり、BMIが高いグループではARが早期に起きていることがわかった、2)出生時のBMI、性の他にも、いくつかの社会・生活環境因子が肥満状態を説明することがわかった。本年度までの解析結果から体系的に社会・生活環境因子の影響を説明することは容易ではなく、更に解析方法を工夫するなどして検討を行う必要がある。

次に「第2子出生タイミングに関する要

因分析」では、従来の分析において用いられてきた人口学的要因や夫妻の社会経済的要因に加えて、出産後6ヵ月の時点における第1子の保育状況、夫の家事・育児参加ならびに養育者(主として母親)の育児体験などの要因が第2子の出生確率に重要な影響を与えていることを明らかにした。

第2子の出生ハザード率には、第1子出生時の妻の年齢、居住地方、都市規模などの人口学的要因が大きく作用している。また、夫の職業が大企業雇用や公務員である場合や、妻の学歴が高く、育児休業制度を利用している場合に、第2子の出生が促されるなど、夫妻の経済的な安定も重要な要因である。

これらに加えて、本研究では次の点について明らかにした。夫の育児への協力は、第2子出生を促す傾向にあるが、夫の家事参加は第2子の出生に対して負の効果をもつ。また、第1子3歳未満で保育施設を利用している場合、第2子出生ハザード率が低い傾向がある。第1子6ヵ月時において、養育者(主として母親)が育児に強い不安や悩みを抱えている場合、子育ての負担感が高い場合に第2子の出生が生じにくい傾向が認められた。

次に「中高年者の生活実態の変化に関する分析手法の検討(4)一健康状態変化とその要因に関する分析」において、まずロジスティック回帰分析によるモデリングからは、前年の健康状態が「よい」であったグループについて、健康状態が悪化するオッズを増大させることに関連性が高いと考えられる説明変数は、介護、診断の有無等であり、健康状態が悪化するオッズを減少させることに関連性が高いと考えられる説明変数は、就労、学歴、過去1か月間感じたこと等であった。

一方、前年の健康状態が「わるい」であ

ったグループについて、健康状態が好転するオッズを増大させることに関連性が高いと考えられる説明変数は、配偶者の有無就労等であり、健康状態が好転するオッズを減少させることに関連性が高いと考えられる説明変数は、診断の有無等であった。

また、離散時間2方向ハザードモデルによる分析からは、前年における健康状態がよい者の場合、就業をやめた者の健康悪化のリスクは、前年・当年を通じて就業をしていない者に比べて高いのに対し、前年における健康状態が悪い者の場合には、前年・当年を通じて就業をしていない者よりも、前年において就業している方が健康改善の可能性が高いという結論を得た。なお、他の説明変数を投入した分析を行い、ロジスティック回帰分析結果との比較を行い、両者が整合的であるとの成果を得た。

すなわち、ロジスティック回帰分析の結果からは、この年代に特有ともいえる介護に関する負担の発生が、本人の健康状態の悪化と関連している可能性が示唆され、また、前年の診断状況の健康悪化への影響は病状によっても異なり、特に糖尿病で大きい値となっていた。就労についての結果からは、就労継続者に対して就労をやめた者の健康状態悪化のオッズの増加は相当大きいものとなっていることが示唆された。また、健康状態が悪化するオッズは高学歴の方が低いことが観察された。一方、健康状態好転の要因について、配偶者ありが健康状態が好転するオッズを増大させることに関連性が高いとの結果が得られ、健康状態が「よい」であったグループとは異なる特徴が観察された。

この配偶者の影響の違いについては、離散時間2方向ハザードモデルによる分析でも整合的であることが観察され、統合的な一つのモデルで表現できる離散時間2方向ハザードモデルの利点が明らかとなった。

E. 結論

事業の研究は、データ管理だけでなく高度な統計分析を支援し、有用な結果を導くためのインフラとしての総合的システムの開発を行い、年々蓄積される縦断調査データを効率的に分析し、有用な政策的、学術的知見を得るための支援を行なうことを目指している。

現時点においては、出生児調査で7回分、成年者調査で6回分、さらには中高年調査で3回分の調査データが蓄積したことから、これまでに比べ分析の幅が大きく広がり、とくに今後は縦断調査本来のテーマに即した本格的な分析が可能となってくると考えられる。

とくに今期においては事例研究をライフコース的視点に沿って体系化を行うことで、個々の分析事例の知見がより有効なものとして捉えられることが確認できた。今後さらに、この体系化の試みを強めることで、21世紀縦断調査の3調査をより有機的に連携させることができると考えている。こうした試みは国際的にも見られないものであり、学術的にも有意義な試みであると言える。

システム開発においては、複数のデータセットを有機的に結びつけ、統計パッケージ・ソフトウェアに、必要なデータとある程度のプログラムを自動生成して供給するシステム PDA21 を完成し、実装に成功し、インターフェースの改良に取り組んだ。今後は本格的な実用に付し、その中で改良を続けて行くことになる。

また、事例研究による統計的分析の実践の経験はシステムに要求される機能や使い勝手等について検討によって、より有効なシステムの開発に役立てて行くものとして、これらシステム開発、事例研究の全体を通して、縦断調査に必要な分析支援体制の構築はさらに進んだものと考えられる。

「(4) データ特性の分析・把握」の分野における「21世紀出生児縦断調査における対象児の特徴に関する出生動向基本調査夫婦調査の出生歴に基づく出生児データとの比較分析」の結果からは、縦断調査から得られた結果を一般化し、他の年に生まれた出生児集団に当てはめる場合に、本研究で明らかになった構造的な違いに十分配慮する必要があるといえる。

「(5) 事例研究とその体系化」において、まず、「子どもの肥満に関連した社会・生活環境因子に関する分析」からえられる結論としては、発育の測定精度等の問題はあるものの、21世紀出生児縦断調査のデータは規模と継続性という点で、他国における本研究と同種の研究で用いられたデータと比較しても有意義であり、わが国で明らかではなかった子どもの発育に関して新しい事実を提示しうるものである。

次に「第2子出生タイミングに関する要因分析」から得られる結論は、第2子の出生には、経済的な安定と夫による可能な限りの育児参加が、重要な要件である。3歳未満保育の活用が、なぜ第2子の出生に対して負の効果をもつのか、今後検証していく必要がある。母親の育児ストレスは出生に関する意思決定にネガティブな影響を与えている可能性が示唆される。したがって、育児ストレスの軽減は、広義の少子化対策としても重要である。今後は、より詳細な分析によって、第2子出生に関する意思決定とそれを取り巻く諸要因との関係が明らかにされていくことを期待したい。

次に「中高年者の生活実態の変化に関する分析手法の検討(4) 一健康状態変化とその要因に関する分析」では、本研究では、健康状態変化のモデリングにあたり、ロジスティック回帰分析と離散時間2方向ハザ

ードモデルの有効性を検討したが、ロジスティック回帰分析からは、前年の健康状態別の2つのモデルを用いて健康状態変化の要因についての分析結果を得た。また、一方の離散時間2方向ハザードモデルの検討からは、健康と就業の関係を一つのモデルとして表すこと、また、ロジスティック回帰分析では前年の健康状態に対応した2つのモデルが必要であったのに対し、統合的に一つのモデルで分析を行うことができ、かつ、ロジスティック回帰分析による分析結果と極めて整合的な結果が得られるということがわかった。

しかしながら、これらのモデルについても、説明変数の選択や、就労以外の変数との同時分析など、さらなる検討点が考えられる。また、これらのモデルの発展や、他の分析手法の検討なども必要であろう。このような観点から、健康状態の変化とその要因に関するモデリング技術のさらなる向上を行うことが今後の検討課題である。

以上を総括すると、分担研究報告を含めた本事業内のすべての研究課題について、一定の到達点に接近しつつあると認められ、本事業は全成果をまとめる時期にさしかかっていると考えられる。したがって、本事業を踏襲し、短期間内において精力的に成果を統合、実用化するための事業に進むことが望ましいと考えられる。

F. 研究発表

システムの検討・開発、分析の準備を事業の目的としているため、個別テーマの研究発表は行わないものとしている。

厚生労働科学研究費補助金（統計情報高度利用総合研究事業）
分担研究報告書

パネル調査（縦断調査）に関する統合的高度統計分析システムに関する開発研究
（平成 21 年度）

研究分担者 北村行伸 一橋大学経済研究所

研究要旨

本分担研究は、21 世紀縦断調査のうち、出生児縦断調査について脱落サンプルのもたらす影響を分析するものである。まず、(1)「21 世紀出生児縦断調査における脱落サンプルのもたらす影響の評価」では、21 世紀出生児縦断調査における脱落サンプルのもたらす統計上のバイアスについて、調査した。パネルデータとクロスセクション・データを比較したところ、子どもの成長に関するパラメータは、ほぼ安定しており、調査回数が増えても大きな変化はないが、子育て費用のパラメータは正から負へと変化しており、その有意性も高まっていることがわかった。総体として見た場合には、経済主体が脱落によって減少するデメリットと時系列方向のデータが増えるメリットを比較すると、明らかに時系列方向のデータの増加のメリットが大きいいといえる。次に家計に関する結果において脱落サンプルがパネルデータ推定にバイアスをもたらすかを分析したところ、脱落サンプル家計と継続サンプル家計のサンプルサイズの違いにもかかわらず、パラメータの大きさも、符号条件も、有意性も全てほとんど同じであり、パネルデータ推定という統計的にかなり集約したレベルでは、脱落サンプル・バイアスはほとんど認められないと結論づけることができる。

(2)の「第 1 - 7 回 21 世紀出生児縦断調査の脱落・移動の動向および子育て感の都市規模別変化」では、①21 世紀出生児縦断調査について、脱落・復活回答と、自治体を超える転居の発生状況を確認し、脱落によるサンプルの歪みを検証した。対象児が学齢に達した第 7 回の時点で、第 1 回調査回答者総数 47015 人の 78.2%にあたるサンプル数を維持しており、第 1 回から全ての回で回答しているサンプルも 71.7%確保している。なお、第 1-7 回の間、自治体を越える転居を一度でも経験した人は、29.5%にのぼる。ひとり親、若い、外国籍、父親のプレゼンスが低い、相談相手がないなどのサンプルが抜け落ちる傾向がみられた。

A. 研究目的

本研究の目的は、厚生労働省による 21 世紀縦断調査のうち、21 世紀出生児調査における脱落に関する分析を行うことである。

大きく分けて、2つの研究を行った。

(1)「21 世紀出生児縦断調査における脱落サンプルのもたらす影響の評価」〔北村〕では厚生労働省（大臣官房統計情報部）に

よって始められた 21 世紀出生児縦断調査における脱落サンプルのもたらす統計上のバイアスについて調査した。

(2)「第 1 - 7 回 21 世紀出生児縦断調査の脱落・移動の動向および子育て感の都市規模別変化」〔西野〕では、①21 世紀出生児縦断調査について、脱落・復活回答と、自治体を超える転居の発生状況を確認し、脱落によるサンプルの歪みを検証した。また、転居者を捕捉しなかった場合に起こっていた影響を予測した。

B. 研究方法

(1) 厚生労働省雇用均等・児童家庭局による『乳幼児身体発育調査』と『21 世紀出生児縦断調査』の基本統計量を比較した。ここでは体重・身長ともに男女別・出生経過期間別の統計量はほぼ同じであることがわかった。また時系列変化を、やはり『乳幼児身体発育調査』から取り、直近の『21 世紀出生児縦断調査』と比べたが、大きくは外れていないことがわかった。

次に脱落サンプル家計と継続サンプル家計の個別変数の平均値を t 検定で比較した。全ての変数で継続サンプルと脱落サンプルの平均値は有意に違うことが明らかになったが、例えば、子育て費用の平均値を比べると、継続サンプル家計が 37320 円で、脱落サンプル家計が 40620 円であり、その差はわずか 3300 円である。また、父親就業と母親就業についての両者の差は 1% 程度であり、ほとんど無視できると判断できる。大きな差があったのは、親の国籍に基づく脱落と継続の違いである。

(2)では、以下のような分析を行った。脱落・復活回答と、自治体を超える転居の発生状況の集計、第 1 回調査に回答した全サンプル、すなわち「脱落なしで全員が第 7 回も回答したならば」という「理想」の仮想サンプルと、第 7 回調査に実際に回答し

たサンプルとで、第 1 回調査の諸変数の値を比較することで、脱落によりサンプルにどのような歪みが生じているかの検証、同様の方法で「理想」のサンプルと転居をしていない回答者のサンプルを比較し、もし転居した人を追っていないならば生じていたであろう偏りの検討である。

C. 研究成果

(1) まず、脱落サンプル問題を考える時に、パネルデータとクロスセクション・データを比較した場合のメリットとデメリットを見ることから始めた。ここで解ったことは、成長を経過日数と経過日数の 2 乗で回帰した自律的成長部分については、パラメータはほぼ安定しており、調査回数が増えても大きな変化はないが、子育て費用のパラメータは正から負へと変化しており、その有意性も高まっているということである。総体として見た場合には、経済主体が脱落によって減少するデメリットと時系列方向のデータが増えるメリットを比較すると、明らかに時系列方向 (T) のデータの増加のメリットが大きいと言える。

次に、脱落サンプルがパネルデータ推定にバイアスをもたらすかということ考えた。ここでは脱落サンプル家計と継続サンプル家計のサンプルサイズの違いにもかかわらず、パラメータの大きさも、符号条件も、有意性も全てほとんど同じである。パネルデータ推定という統計的にかなり集約したレベルでは、脱落サンプル・バイアスはほとんど認められないと結論できる。

(2) 対象児が学齢に達した第 7 回の時点で、第 1 回調査回答者総数 47015 人の 78.2%にあたるサンプル数を維持しており、第 1 回から全ての回で回答しているサンプルも 71.7%確保している。初めて脱落したサンプルは、その後の回で復活回答する確率が 4 割から 5 割近いが、その後再度脱落

した場合は、復活は4割を切っていた。

第1-7回間に、自治体を越える転居を一度でも経験した人は、29.5%にのぼる(市町村合併による自治体名変化と区別できない例は除いた比率)。そのうちおよそ2/3は県内(または二大都市圏内)移動だった。ひとり親、若い、外国籍、父親のプレゼンスが低い、相談相手がいないなどのサンプルが抜け落ちる傾向がある。移動者には、脱落者に近い傾向と、収入のある父親+専業主婦の母親の特徴に近い傾向とがある。

D. 考察

(1) 今後、日本社会は必ず移民受け入れ問題に直面するだろう。この調査に含まれている外国人の親とその子供の成育の記録は、移民の日本社会への同化の基礎資料にもなり得るものである。一つ懸念があるとすれば、もともとのサンプル数は小さいが親が外国人であるサンプルが急激に脱落していることである。外国人移民の受け入れなどの政策を検討する場合に重要になるであろう基礎資料が調査方法の煩雑さや説明不足で失われているとすれば、今からでも何らかの対応を立てることが必要だと思われる。

(2) 脱落者が支援を必要とする層に偏る傾向が、第6回までに行った同様の検証よりも一層進んでいる。また、第1-7回全ての回に回答した人のみに絞ると、上記の特徴が一層強く出ていた。比較的近距离の移動ではあるが、子供が小学校に入るまでは、家族の転居が比較的活発な時期であることは窺える。移動者を補足していなかった場合は、支援の必要な層が一層漏れたとともに、三世同居や祖父母の支援が比較的豊富なサンプルに偏っていたおそれがある。

E. 結論

(1) 21世紀出生児縦断調査の脱落率は6

年間で18%であり、通常のパネルデータ調査と比べて格段に低く、調査データにおける脱落バイアスは軽微である。さらに、この調査では、調査対象となっている子供の成長は調査記録者である親の属性とは比較的独立していることもあり、統計上の脱落サンプル・バイアスは現状では問題にならないことが解った。しかし、脱落サンプルを精査してみると、親が外国人である場合には、脱落率が高く、国籍にも偏りがあることが解った。調査方法の煩雑さや言語障壁によって脱落しているとすれば、何らかの対応を考える必要があるだろう。

(2) 脱落者が支援を必要とする層に明確に偏ってきており、結果の解釈は常にその点に注意を払うべきだろう。本調査のまず、本調査の場合、移動はあくまで、調査対象者からの自発的な住所変更連絡によって判明しているので、移動者の中でも調査に比較的熱心な人のみが残りと、他の移動は転居先不明によって脱落につながったと考えられる。今後の縦断調査では、移動者の追跡の手立てを検討すべきだろう。

F. 研究発表

1. 論文発表

なし

2. 学会発表

なし

G. 知的所有件の取得状況

なし

厚生労働科学研究費補助金（統計情報高度利用総合研究事業）
分担研究報告書

パネル調査（縦断調査）に関する統合的高度統計分析システムに関する開発研究
（平成 21 年度）

研究分担者 釜野さおり 国立社会保障・人口問題研究所

研究要旨

本分担研究は、2つの異なる目的をもって実施された。パート1では21世紀縦断調査のうち、成年者縦断調査について脱落サンプルのもたらす影響を分析した。パート2では、分析システム開発の一貫として昨年度実施したライフコースの観点から子育て環境を体系的に捉える試みを継続し、出産と子育てに関するテーマに焦点を当てたいくつかの分析テーマを見だし、分析システムをさらに強化した。

パート1の男性票の脱落状況および脱落によるデータのゆがみの検証においては、第1～5回の男性票データを用いて、継続者と脱落者の回答を比較し、さらに第1回目の調査を理想の調査とみなし、第3回および第5回までの継続者を比較した。年齢が高い、有配偶、高学歴、低学歴、子ども有、親と同居している、仕事を持っている、所得が高いといった特徴を持つ男性が継続者として残る傾向がみられた。次に女性票について希望子ども数の回答の事後的補正を試みた。脱落者と不詳者が次年度の調査でどのように回答しているか調べ、復活しているケースについてその回答分布を使うことで、補正希望子ども数を算出した。続いて、実験法における処置群と対照群を分けてその差を測定することによって求められる因果効果について、共変量を1変数に縮約した傾向スコアを用いることによって、マッチングという手法によって擬似的に実験法と同状況を作りだし、脱落の因果効果を測定した。

パート2のライフコースの観点から子育て環境を分析するシステムの開発では、出生児縦断調査という大規模な縦断調査の利点を生かし、婚外出生、若年出産および高齢出産など、少ないながらも増加傾向にある事象の実態の分析事例を提示した。また子育てに関しては、昨年度に引き続き保育者の子ども観による育児方針や育児行動にどのような差異があるかの分析事例を蓄積した。

A. 研究目的

本研究では、2つの異なる目的をもって実施された。パート1では21世紀縦断調査のひとつである成年者縦断調査について、その脱落に関する分析や方法論的な検討を

行った。パート2では、21世紀縦断調査の分析システム開発の一貫として、昨年度取り組んだ、ライフコースの観点から子育て環境を体系的に捉える試みを継続し、出産

と子育ての事項に関しての補強を行った。

各パートの構成を以下に示す。

●パート1

- (1): 21 世紀成年者縦断調査（第1回～第5回）における男性票の脱落者・継続回答者の特性に関する分析〔釜野〕
- (2): 成年者縦断調査における希望子ども数の無回答に関する分析〔守泉〕
- (3): パネル調査における脱落の初婚に対する因果効果についての検討〔鎌田〕

●パート2

- (4): 21 世紀出生児調査における婚外出生児の現状〔岩澤・三田〕
- (5): 若年出産と高齢出産〔相馬〕
- (6): 子ども観と育児方針2〔元森〕

それぞれの研究目的は、以下のとおりである。

(1)では、成年者縦断調査の男性票（配偶者票を除く）について、第1回～第5回の回収と脱落状況を調べ、それによるデータのゆがみを検証した。

(2)では、第1回(2002年)～第6回(2007年)の女性票データを用いて、希望子ども数の無回答（脱落・不詳）に関する分析を行った。希望子ども数の質問形式の違いに注意しながら、調査からの脱落者および希望子ども数の無回答者がどのような属性を持っているのか検証し、さらに縦断調査の特性を生かして、無回答の事後的補正について考察した。

(3)では、傾向スコアを用いた共変量の調整脱落の発生が初婚の発生に対する因果効果を、初婚を規定している共変量を調整したうえで測定した。

(4)は、「21 世紀出生児縦断調査」は、4万7千ほどの出生児の標本数があることから、従来の標本調査ではとらえることができなかった特殊な環境にある少数の出生児の実態の把握が可能であることを踏まえ、

「婚外出生児」に着目し、当該調査の活用事例を示したものである。婚外出生児の増加は、これまでの家族政策や福祉政策で対応できない問題を引き起こす可能性もあり、日本においては現時点では数としては少ないが、増加傾向にあるため、実態の把握を行った。

(5)では、地方自治体の子育て支援の現場で、若年出産・高齢出産と出産時期に着目した支援の必要性が提起されていること、および国際的にも *teenage pregnancy* の問題は、重要な社会政策の課題となっていることを踏まえ、若年出産・高齢出産の親の妊娠・出産・子育ての実態把握を目的とした。

(6)では平成19年度報告書の予備考察と、平成20年度報告書の考察の続きとして、「21 世紀出生児縦断調査」において、保育者の子ども観による育児方針や育児行動にどのような差異があるか明らかにした。

B. 研究方法

(1): 各回の男性票をマッチングしたファイルを用いて、第一に、継続回答者と脱落者の特性の比較として、第1回調査時の回答について、項目ごとにその後の継続者と脱落者の回答に有意な差がみられるかを検証した。第3回までの継続者と脱落者、第5回までの継続者と脱落者について、それぞれ、連続変数については母平均の差の t 検定を行い、離散変数については確率分布の差の χ^2 検定を行った。第二に、第1回調査の回答について、第1回のサンプル全体と、第3回・第5回の回答サンプルで平均値や分布に有意な差がみられるかどうか検証した。第1回調査の回答者（「回答者全員が継続回答する理想的な調査」を母集団として想定し、連続変数については t 検定、離散変数については χ^2 検定で、1サンプルによる検定を行った。