

7 21 世紀出生児縦断調査における脱落要因の分析

福田 節也
金子 隆一

はじめに

21 世紀出生児縦断調査は、全国の 2001 年 1 月 10-17 日および 7 月 10-17 日に出生したすべての子を調査客体としたパネル調査である。パネル調査では、同一個人あるいは世帯を対象として複数年にわたり継続的にサーベイを行う。パネル調査における利点として、同一対象の時間的変化を特定できること、また横断調査に比べてより直接的に標本のもつ異質性を統制し、説明要因と非説明変数との因果関係を考察できることなどが挙げられる（樋口他，2004）。しかし、一方でパネル調査特有の問題として、調査拒否や転居等の理由によりサンプルが脱落してしまうパネル減少（panel attrition）がある。サンプルの脱落が非無作為に生じる場合、データの代表性が損なわれ、データから得られる知見に偏りが生じることとなる。また、サンプルの脱落は調査回数とともに累積する為、脱落が著しい場合には調査の長期的な実施が困難となる。したがってパネル調査においては、いかに脱落を減らすかが調査実施上の要諦となる。

本稿では、21 世紀出生児縦断調査における脱落の発生状況について考察し、同パネル調査における脱落要因の分析を行う。また、脱落への対処法についても述べる。本稿における考察を通じて、21 世紀出生児縦断調査における脱落の防止に資する提言を行うことが目的である。

1. 21 世紀出生児縦断調査における脱落の発生概況

本節では、21 世紀出生児縦断調査の第 1 回調査（以後、第 1 波）および第 2 回調査（以後、第 2 波）の確定データを用いて脱落の発生概況およびその留意点について報告する。

表 1-1. 21 世紀出生児縦断調査第 2 波における脱落の発生割合

| 年度 | 回収数 | 回収率 | 脱落数 | 脱落率 |
|----|-------|---------|------|-----|
| 1 | 47015 | (87.8%) | - | - |
| 2 | 43926 | 93.4 | 3089 | 6.6 |

表 1-1 において明らかなように、同調査第 2 波における脱落サンプルは 3089 ケース、割合にして初年度サンプルの 6.6%となっている。一般にパネル調査では、毎年継続的に調査することについて調査対象者より了解をとる必要がある。そのため、初年度の調査では調査協力を得にくいのが、第 2 年度以降の調査では協力が得やすいことが指摘されている（坂本，2004）。厚生労働省による統計調査であることを反映し、当調査では初年度の回収率が

87.8%と高いことが指摘できる。わが国における先駆的なパネル調査として知られる『消費生活に関するパネル調査』（財団法人家計経済研究所により 1993 年より実施）では初年度（1993 年）の回収率が 41.4%，以後 1994 年から 2003 年までの各年度調査の平均回収率が 95.5%（いずれも初年度調査よりサンプルとなっているコーホート A の値）となっている（家計経済研究所，2004）。したがって，21 世紀出生児縦断調査においては初年度の回答率が高い代わりに，脱落率に関しては同様の調査と比較して同程度か若干高い傾向にあるといえる。

次に，パネル調査における脱落の留意点について述べたい。パネル調査における脱落は累積するという点に注意が必要である。表 1-2 は，21 世紀出生児縦断調査における脱落の状況ならびに 4 回調査以降の回収率を 95%と仮定した場合の推計値を表したものである。

表 1-2 をみると，各年度の回収率を 95%に仮定した場合，5 年度調査における累積脱落率の推計値は 17.8%，10 年度調査においては 36.4%にも達している。このことは，初年度調査の対象サンプルのうち 5 年度調査まで継続して回答している者は約 82%，10 年度調査では約 64%にまで減少することを意味している。さらに初年度の回収率が 87.8%であったことを考慮すると 5 年度調査では標本母集団の 72.1%（ $=87.8 \times 0.822$ ），10 年度調査では 55.8%（ $=87.8 \times 0.636$ ）しか捕捉されていないことを意味している。

パネル調査における脱落がすべての回答項目に対して無作為に生じる場合（missing completely at random）には，各変数の分布は不変であるため統計量の推定において脱落によるバイアスは生じない。ただし，脱落によりサンプル数が減少するため，推定される統計量の標準誤差が大きくなるが，このことは統計的検定に影響を及ぼさない（Allison, 2002）。しかし，ほとんどすべての場合において，脱落は調査対象者のなんらかの属性に依存して生じる。例えば，共働きや高所得などの特定の傾向をもつサンプルほど調査から脱落する確率が高い。したがって，脱落の発生したデータを用いる場合，実際には脱落と関連のある変数の統計量の推定にはバイアスが生じることとなる。

表 1-2 21 世紀出生児縦断調査における脱落の状況ならびに推計値

| 年度 | 回収数 | 回収率* ¹ | 脱落数 | 累積脱落率* ² |
|-----------------|---------|-------------------|--------|---------------------|
| 1 | 47015 | (87.8%) | - | 0.0 |
| 2 | 43925 | 93.4 | 3090 | 6.6 |
| 3* ³ | 42803 | 97.4 | 1122 | 9.0 |
| 4 | 40662.9 | 95.0 | 2140.2 | 13.5 |
| 5 | 38629.7 | 95.0 | 2033.1 | 17.8 |
| 6 | 36698.2 | 95.0 | 1931.5 | 21.9 |
| 7 | 34863.3 | 95.0 | 1834.9 | 25.8 |
| 8 | 33120.1 | 95.0 | 1743.2 | 29.6 |
| 9 | 31464.1 | 95.0 | 1656.0 | 33.1 |
| 10 | 29890.9 | 95.0 | 1573.2 | 36.4 |

* 4 年度以降は回収率を 95%と仮定した場合の推計値

*¹ 当年 (t 年) の回収数 / 前年 (t-1 年) の回収数 × 100

*² 当年 (t 年) までの脱落数累計 / 47015 人 (初年度回収数) × 100

*³ 回収数には第 2 回調査に無回答であったサンプルも含む

一方で、多変量解析を行って回帰係数を推定するような場合には、パネル調査における脱落は係数の推定にそれほど大きな影響を与えないとの結果もある（Lillard & Panis, 1998; Alderman, et. Al., 2001; Falaris, 2003）。しかし、21世紀パネル調査が主として記述統計を旨とする官庁統計であることを考慮すると、脱落による標本の歪みは当調査に基づき公表する統計データの信頼性や妥当性に大きな影響を与えるものと思われる。また、脱落は調査の長期的な継続を困難にするという点からも望ましいものではない。そこで本稿では21世紀出生児縦断調査の第2回までのデータを用いて、脱落発生のメカニズムに関する考察を行う。

2. 21世紀パネル調査における脱落要因の2変量解析

本節では、21世紀出生児縦断調査の第2回調査における脱落が調査対象者の属性とどのように関わっているのかを明らかにする。具体的には、第1波における対象者の属性や回答状況と第2波における脱落との相関関係について考察する。

2-1. 量的変数と脱落の関係

はじめに、第1波の量的変数と第2波における脱落との2変数の関係を考察する。表2-1は、第1波における量的変数について、第2波まで回答を継続したサンプルと脱落したサンプルの平均値を比較したものである。

夫妻の属性と脱落との関係をみると、脱落サンプルは継続回答サンプルに比べて父母の年齢が若いもしくは父母の同居期間（おそらくは結婚期間）が短いといった特徴がある。また、夫婦の年齢差が大きいことも脱落サンプルの特徴といえる。さらに脱落サンプルは継続回答サンプルに比べて父母ともに年収が低いという傾向がみられる。前述の家計研パネルにおいては年収の高いサンプルほど脱落する傾向があるが（坂本, 2004）、これとは逆の結果となっている。また、出生児の父母が喫煙者である場合、脱落サンプルほど1日の喫煙本数が多い傾向にある。特に脱落者と継続回答者の間で母親の喫煙本数の差が大きい。

子の属性についてみると、わずかではあるが対象児の体重や母親の妊娠週数が短い場合に脱落する傾向があるようである。出生児が未熟児であるほど脱落する傾向が高いのであろうか。このことは後に質的変数と脱落の関係において考察する。また、脱落サンプルには母の出産した子の数が多いという傾向が認められる。しかし、出生児の兄弟の数は脱落に影響していない。このことは双子などの多胎児の出産が脱落を促すことを示唆している。これについても後に考察する。

さらに父母の育児意識や育児・家事頻度と脱落の関係について着目する。脱落サンプルには母親の育児・家事頻度が低い傾向が認められる。特に、育児頻度の差が大きい。一般にパネル調査からの脱落は、調査の主目的によって異なることが指摘されている。例えば、調査の目的が政治意識に関するものである場合、政治意識の高い対象者ほど継続回答する傾向があるだろう。本調査の主目的が子育てに関する意識や行動であることを考慮すると、

母親の子育てへの意識や参加度が低い場合に脱落が生じやすいのかもしれない。こうした傾向は、脱落サンプルほど子育てで意識して行っていることや子どもをもってよかったことの得点が低いことにも表れている。先の喫煙本数と脱落の関係についても、母親の喫煙本数が育児意識の低さを反映していると考えられる。しかし、父親の家事・育児頻度は脱落に影響していない。母親のみが調査の回答者である割合は約 9 割 (87.6%) であり、調査継続に関する意思決定は主に母親によって行われていることが示唆される。

表 2-1 回答継続者と脱落者の比較 (量的変数)

| 変数 | 継続回答サンプル (1) | 脱落サンプル (2) | 差 (2) - (1) |
|----------------------|------------------|-----------------|----------------|
| 子の体重 (g) | 3035.92 43915 | 3018.99 3086 | -16.93 * |
| 妊娠週数 (週) | 38.89 43895 | 38.83 3084 | -0.06 * |
| この母の出産した子の数 (出生子) | 1.68 43926 | 1.73 3089 | 0.05 ** |
| 妊娠満22週以後の死産子の数 | 0.01 43926 | 0.01 3089 | 0.00 |
| 父の年齢 (歳) | 31.66 43431 | 30.03 2971 | -1.63 *** |
| 母の年齢 (歳) | 29.50 43926 | 27.60 3089 | -1.90 *** |
| 同居期間月数 | 44.94 43246 | 40.22 2944 | -4.72 *** |
| きょうだい等の人数 (兄弟の人数) | 0.66 43926 | 0.69 3089 | 0.03 |
| 父母の喫煙状況 (お母さん) 1日の本数 | 11.32 7071 | 12.92 1053 | 1.59 *** |
| 父母の喫煙状況 (お父さん) 1日の本数 | 18.72 26324 | 19.48 2125 | 0.76 *** |
| 母の家事・育児の点数 (育児) | 16.69 42917 | 16.51 2969 | -0.18 *** |
| 母の家事・育児の点数 (家事) | 17.08 43002 | 16.98 2998 | -0.09 *** |
| 父の家事・育児の点数 (育児) | 10.51 41525 | 10.52 2733 | 0.01 |
| 父の家事・育児の点数 (家事) | 6.48 42035 | 6.34 2787 | -0.14 |
| 子育てで意識して行っていること(尺度) | 2.78 43835 | 2.70 3065 | -0.08 ** |
| 子どもをもってよかったこと(尺度) | 3.29 43809 | 3.19 3060 | -0.11 *** |
| 子どもをもって負担に思うこと(尺度) | 1.76 43681 | 1.75 3040 | -0.01 |
| 母親の年収 (万円) | 101.35 42352 | 75.12 2867 | -26.23 *** |
| 父親の年収 (万円) | 457.37 41214 | 378.35 2634 | -79.02 *** |
| その他の年収 (万円) | 11.22 43406 | 12.11 2984 | 0.89 |
| 夫婦の年齢差(絶対値) (歳) | 3.10 43431 | 3.35 2971 | 0.25 *** |
| 無回答項目総数 | 3.53 43926 | 5.35 3089 | 1.82 *** |

各行の上段は実数値、下段はサンプル数を表す。

***: 0.1%水準、**: 1%水準、*: 5%水準で統計的に有意(t検定)

調査票における無回答項目の数¹⁾もまた脱落に影響を与えている。脱落サンプルほど初年度調査における無回答項目数が多い傾向にある。無回答項目の多さは、調査に対する理解度や関心の低さを反映するものと思われる。したがって、育児意識とは別に調査に対する協力度が低い場合にも脱落が生じやすいものと思われる。しかし、同時に当調査では父母のどちらかが出生児と同居していない場合や出生児が非嫡出である場合、多くの質問項目において非該当として処理されている。そのため、無回答項目の数が調査への協力度の低さを反映しているのか、それとも父母の非同居や非嫡出による影響なのかについては、2変量解析では明らかにすることができない。この問題については、後に多変量解析によって明らかにする。

2-2. 質的変数と脱落の関係

次に、第1波における質的変数と第2波における脱落との2変量の関係について考察する。表2-2は脱落率が質的変数の各カテゴリーにおいてどのように異なるのかを比較している。なかにはカテゴリーのサンプル数が非常に少ない（全サンプルの1%以下）変数もあるが、脱落に重要な影響を与えているものもあるので敢えて表に含めた。

表中のほとんどの変数において、脱落率との間に統計的に有意な関係が認められる。しかしカイ2乗検定ではサンプル数が多いほど統計的な有意性が検出されやすいという特性があるのでその解釈には注意が必要である。表の解釈においては、むしろカテゴリーの脱落率とカテゴリーのサンプル数という2つの観点から脱落への影響を判断する必要がある。主な知見について要約すると以下である。

脱落率が10%以上のカテゴリーをもち、かつそのカテゴリーに占めるサンプル数が全体の10%程度である変数として、母親の喫煙状況、父母の喫煙状況、母親の就業状況、そして婚前妊娠の有無を挙げることができる。喫煙本数は脱落サンプルほど多い傾向がみられたが、今回の分析でも母親の喫煙が脱落と関係することが明らかとなった。母親が室内で喫煙する割合は全サンプルの12.2%にも上り、脱落率は13.4%となっている。母親の喫煙は父親の喫煙と強い相関をもっており、両親ともに喫煙する場合の脱落率は12.6%となっている。次に、第1波において母親が求職中である場合にも高い脱落率がみられる。母親が「仕事を探している」と回答した割合は9.5%であったが、そのうちの10.7%が第2波において脱落している。就業による多忙によるものであろうか。そして近年増加傾向にある婚前妊娠はサンプルの13.3%を占めており、うち10.7%が第2波で脱落している。なお、本分析における婚前妊娠とは、夫婦の同居期間が妊娠週数より短い場合と定義している。婚前妊娠は早婚とも関連しているので（厚生労働省大臣官房統計情報部、2002）、夫婦の年齢が若く、同居期間が短いほど脱落しやすいという先の分析結果は、婚前妊娠と関係しているのかもしれない。これについては後に多変量解析によって確認する。

1) 無回答項目数の定義については付表1を参照のこと。

次に、カテゴリーにおける脱落率が10%以上で、かつそのカテゴリーのサンプル数が1000以上（全サンプルの約2%）である変数として、平日日中の保育者、父親との非同居、家族形態、そして父親の就業状況および労働時間が挙げられる。平日の日中に子どもを施設やベビーシッターに預けている場合に脱落率が高くなる。また、対象児が父親と同居していない場合や母子世帯において脱落率が約17%と高い傾向がある。父親が無職やパートである場合や労働時間が20時間以下である場合にも脱落率が10%を超えている。先の分析では脱落者のほうが継続回答者よりも年収が低いという傾向がみられた。民間の調査機関が行う調査とは異なり、当調査では謝金などのインセンティブが作用しないことが低所得層や母子のみの世帯における脱落を促している可能性がある。

さらに、カテゴリーに占めるサンプル数は少ないが、脱落率に非常に大きな影響を与えている変数として、嫡出子・否の別、父母の国籍、ならびに母親との非同居が挙げられる。非嫡出児の脱落率は約20%にも上る。また、父母ともに外国籍であるサンプルは217サンプルと少ないが27.7%と高い脱落率を示している。これらの世帯はもともと今回のような標本調査において捕捉することが困難な層であり、パネル調査における脱落率も大きいものと思われる。さらに54サンプルと少数ではあるが、対象児が母親と同居していないケースも約30%と高い脱落率を示している。したがって、当調査のサンプルは比較的標準的かつ安定的な世帯属性をもつサンプルに収束していくことが予測される。

その他の知見として、低体重児、早産、もしくはこの両方である未熟児といった出生児の属性が脱落率と正の関係をもつことが明らかである。このことは脱落サンプルほど出生児の体重が軽く、妊娠週数が短いという先の分析結果と一致する。なぜ未熟児に脱落が多いのかについては、脱落理由を精査しない限り断定はできないが、乳幼児の死亡率に関連しているのかもしれない。つまり、未熟児で出生したサンプルほど死亡による脱落が多い可能性がある。データでは「子どもをもって負担に思うこと」として、「子どもが病気がちである」を挙げる割合が、未熟児で7.3%、未熟児以外では3.4%と未熟児の方が2倍以上高くなっている。また、統計的な有意性は5%と低いですが、多胎児であるほど脱落率が高いという傾向がみられる。これは出生児の兄弟の数は脱落に影響しないが、母親の出生児数が多いほど脱落がみられるという量的変数の分析結果を支持するものである。多胎児の回答者もまた「子どもが病気がちである」を挙げる割合が6.0%と高い傾向にある。また多胎児の場合、回答者は出生児と同じ数の調査票に記入しなければならないため負担が大きい。このことが脱落に繋がりがやすいということも考えられる。

表 2-2 回答継続者と脱落者の比較 (質的変数)

| 変数 | 値 | 脱落ケース/全体 | 脱落率(%) |
|--------------|-------------------------|------------|-----------|
| 居住地方 | 北海道 | 121/1804 | 6.71 |
| | 東北 | 262/3498 | 7.49 |
| | 関東 | 927/14893 | 6.22 |
| | 北陸 | 111/2132 | 5.21 |
| | 中部 | 417/7112 | 5.86 |
| | 近畿 | 504/7915 | 6.37 |
| | 中国 | 201/2824 | 7.12 |
| | 四国 | 104/1399 | 7.43 |
| | 九州・沖縄 | 442/5438 | 8.13 *** |
| 市郡規模 | 13大都市 | 635/10061 | 6.31 |
| | その他の市 | 1833/27943 | 6.56 |
| | 郡部 | 621/9011 | 6.89 |
| 嫡出子・否の別 | 嫡出子 | 2971/46402 | 6.40 |
| | 嫡出子でない子 | 118/613 | 19.25 *** |
| 性別 | 男 | 1600/24436 | 6.55 |
| | 女 | 1489/22579 | 6.59 |
| 出生年月日時 (月) | 1月 | 1500/23423 | 6.40 |
| | 7月 | 1589/23592 | 6.74 |
| 国内外(生まれたところ) | 日本 | 3084/46963 | 6.57 |
| | 日本外 | 5/52 | 9.62 |
| 父母の国籍 | 父母とも日本人 | 2795/45399 | 6.16 |
| | 父母のどちらかが日本人 | 234/1399 | 16.73 |
| | 父母の両方が外国籍 | 60/217 | 27.65 *** |
| 世帯主の主な仕事 | 農家 | 82/1455 | 5.64 |
| | 自営 | 306/3871 | 7.90 |
| | 勤労者 I | 1241/16160 | 7.68 |
| | 勤労者 II | 925/19473 | 4.75 |
| | その他 | 344/4436 | 7.75 |
| | 無職 | 113/745 | 15.17 *** |
| 出産場所 | 病院 | 1595/25329 | 6.30 |
| | 診療所 | 1454/21156 | 6.87 |
| | 助産所 | 30/449 | 6.68 |
| | 自宅・その他 | 10/81 | 12.35 * |
| 単胎・多胎の別 | 単胎 | 3006/46050 | 6.53 |
| | 多胎(双子以上) | 83/965 | 8.60 * |
| 低体重児 | 低体重児ではない(出生時2500g以上) | 2778/42996 | 6.46 |
| | 低体重児(出生時2500g未満) | 308/4005 | 7.69 ** |
| 早産 | 正期産(妊娠37-43週) | 2879/44596 | 6.46 |
| | 早産(妊娠23-36週) | 205/2383 | 8.60 *** |
| 未熟児(早産の低体重児) | 未熟児ではない | 2955/45439 | 6.50 |
| | 未熟児 | 129/1538 | 8.39 ** |
| 回答者 | 母のみ | 2661/41198 | 6.46 |
| | 父のみ | 266/3597 | 7.40 |
| | 父母両方 | 120/1973 | 6.08 |
| 平日日中の保育者 | 父母のどちらか | 2725/43199 | 6.31 |
| | 祖父母 | 150/1810 | 8.29 |
| | 保育士・保育ママ・ベビーシッター | 182/1793 | 10.15 *** |
| 母親と同居 | 非同居 | 16/54 | 29.63 |
| | 同居 | 3073/46961 | 6.54 *** |
| 父親と同居 | 非同居 | 186/1098 | 16.94 |
| | 同居 | 2903/45917 | 6.32 *** |
| 家族形態 | 核家族 | 2250/36030 | 6.24 |
| | 3世代同居 | 650/9873 | 6.58 |
| | 母子世帯 | 173/1058 | 16.35 |
| | その他 | 13/41 | 31.71 *** |
| 住まいの広さ | 手狭だと感じている | 1137/15865 | 7.17 |
| | 必要な広さはある | 1204/20310 | 5.93 |
| | 十分にゆとりがある | 522/8502 | 6.14 |
| | あまり考えたことがない | 198/2144 | 9.24 *** |
| 引越しや増築の状況 | 引越しや増築をしていない | 1932/32889 | 5.87 |
| | 今回の妊娠・出産をきっかけに引越しや増築をした | 523/5438 | 9.62 |
| | 他の理由で引越しや増築をした | 552/8077 | 6.83 *** |
| 母親の喫煙状況 | 吸っていない | 1990/38565 | 5.16 |
| | 室外で吸う | 266/2288 | 11.63 |
| | 室内で吸う | 768/5740 | 13.38 *** |

(次ページへ続く)

(前ページの続き)

| 変数 | 値 | 脱落ケース/全体 | 脱落率(%) |
|------------------|----------------------|------------|-----------|
| 父親の喫煙状況 | 吸っていない | 723/16772 | 4.31 |
| | 室外で吸う | 762/11932 | 6.39 |
| | 室内で吸う | 1365/16589 | 8.23 *** |
| 母親の喫煙状況 | 父母とも非喫煙 | 724/16601 | 4.36 |
| | 父のみ喫煙 | 1264/21841 | 5.79 |
| | 母のみ喫煙 | 137/904 | 15.15 |
| | 父母ともに喫煙 | 912/7244 | 12.59 *** |
| 子育てで意識して行っていること | 意識して行っていることがある | 2883/44419 | 6.49 |
| | 特に意識して行っていることはない | 182/2481 | 7.34 |
| 子どもをもってよかったこと | よかったと思うことがある | 3033/46669 | 6.50 |
| | 良かったと思うことはない | 27/200 | 13.50 *** |
| 子どもをもって負担に思うこと | 負担に思うことがある | 2398/37357 | 6.42 |
| | 負担に思うことがない | 642/9364 | 6.86 |
| 子育ての不安や悩み | すごくある | 229/3111 | 7.36 |
| | 少しある | 1601/26078 | 6.14 |
| | ほとんどない | 1230/17679 | 6.96 *** |
| 相談者の有無 | 相談する | 2869/45102 | 6.36 |
| | 誰にも相談しない | 58/435 | 13.33 *** |
| 授乳状況 | 母乳のみ | 425/9801 | 4.34 |
| | 人工乳のみ | 97/795 | 12.20 |
| | 人工乳(初乳のみ) | 191/2137 | 8.94 |
| | 混合乳 | 2321/33883 | 6.85 *** |
| 移動の有無(出生時→初年度調査) | なし | 2871/42900 | 6.69 |
| | 県内移動 | 74/1378 | 5.37 |
| | 県間移動 | 144/2737 | 5.26 ** |
| 母親の育児休業取得経験 | 育休対象外(無職・学生・常勤以外の雇用) | 2188/31866 | 6.86 |
| | 育休取得(済み・中・予定) | 336/5818 | 5.78 |
| | 出産を機に常勤退職 | 429/7870 | 5.45 |
| | 育休不取得で常勤継続 | 129/1396 | 9.24 *** |
| 母親の就業状況 | 仕事を探している | 477/4447 | 10.73 |
| | 仕事を探していない・学生 | 1700/30165 | 5.64 |
| | 育児休業中 | 273/4725 | 5.78 |
| | 常勤 | 197/2532 | 7.78 |
| | パート・アルバイト・内職・その他 | 243/2578 | 9.43 |
| | 自営業・家業 | 138/2020 | 6.83 *** |
| 父親の就業状況 | 無職・学生 | 91/764 | 11.91 |
| | 常勤 | 2273/38485 | 5.91 |
| | パート・アルバイト・内職・その他 | 109/817 | 13.34 |
| | 自営業・家業 | 397/5423 | 7.32 *** |
| 母親の労働時間 | なし | 2496/40095 | 6.23 |
| | 20時間未満 | 184/2333 | 7.89 |
| | 20時間以上40時間未満 | 187/2177 | 8.59 |
| | 40時間以上60時間未満 | 131/1721 | 7.61 |
| | 60時間以上 | 18/166 | 10.84 *** |
| | 労働時間なしのとき | 2496/40095 | 6.23 |
| 母親の通勤時間 | 通勤に時間はかからない | 129/1906 | 6.77 |
| | 30分未満 | 283/3173 | 8.92 |
| | 30分以上1時間未満 | 74/921 | 8.03 |
| | 1時間以上1時間30分未満 | 9/150 | 6.00 |
| | 1時間30分以上 | 4/33 | 12.12 *** |
| | 労働時間なしのとき | 2496/40095 | 6.23 |
| 父親の労働時間 | なし | 126/971 | 12.98 |
| | 20時間未満 | 262/2174 | 12.05 |
| | 20時間以上40時間未満 | 248/3709 | 6.69 |
| | 40時間以上60時間未満 | 1396/25878 | 5.39 |
| | 60時間以上 | 712/11737 | 6.07 *** |
| | 労働時間なしのとき | 126/971 | 12.98 |
| 父親の通勤時間 | 通勤に時間はかからない | 299/4226 | 7.08 |
| | 30分未満 | 1254/20265 | 6.19 |
| | 30分以上1時間未満 | 694/11846 | 5.86 |
| | 1時間以上1時間30分未満 | 228/4944 | 4.61 |
| | 1時間30分以上 | 73/1405 | 5.20 *** |
| | 労働時間なしのとき | 299/4226 | 7.08 |
| 夫婦の年齢差 | 妻年上 | 611/9140 | 6.68 |
| | 同年齢 | 504/7934 | 6.35 |
| | 夫年上 | 1856/29328 | 6.33 |
| 婚前妊娠 | 婚前妊娠なし | 2269/39902 | 5.69 |
| | 婚前妊娠あり | 670/6253 | 10.71 *** |

***: 0.1%水準、**: 1%水準、*: 5%水準で統計的に有意(カイ2乗検定)

3. 21世紀パネル調査における脱落要因の多変量解析

前節における2変量解析によって、諸変数と脱落の相関関係が明らかとなった。しかし、2変量解析では説明変数間の相関が統制されないため、どの変数が脱落との間に因果関係をもつのかを特定することができない。そこで本節では、脱落を2値の従属変数とする多変量解析を行い、諸変数と脱落との因果関係を考察する。分析には脱落の生起確率を推定するロジスティック回帰分析を行う²⁾。

前節までの分析において、脱落には1) 夫婦の人口学および社会経済的属性、2) 未熟児や多胎児といった対象児の属性、3) 育児の状況や4) 親の育児意識、そして5) 調査への関心や理解度といった要因が関係していることが示唆された。そこで本節では以下の表3-1に示した説明変数を用いて脱落確率を推定するモデルを構築する。表中の符号条件は変数が順序尺度以上である場合に限り、2変量解析の結果をもとに導いている。

表 3-1 脱落確率の推定に用いた説明変数

| 要因 | 変数 | 符号条件 |
|------------|---------------------|------|
| 夫婦の人口学的属性 | 父親の年齢 | - |
| | 母親の年齢 | - |
| | 夫婦の年齢差(絶対値) | + |
| | 夫婦の同居期間(年) | - |
| | 婚前妊娠による出産 | + |
| | 母親の既往出生児数(対象児を除く) | + |
| | 家族形態 | |
| | 父母の国籍 | |
| | 居住地方 | |
| | 市郡規模 | |
| 夫婦の社会経済的属性 | 世帯年収(父年収+母年収+その他年収) | - |
| | 母親の就業状況 | |
| | 父親の労働時間 | + |
| | 父親の企業規模 | + |
| | 住まいの広さ | - |
| 対象児の属性 | 未熟児 | + |
| | 単胎・多胎の別 | + |
| 育児の状況 | 平日日中の保育者 | |
| | 授乳状況 | |
| | 母親の育児頻度 | - |
| | 母親の家事頻度 | - |
| | 父親の育児頻度 | - |
| 親の育児意識 | 父親の家事頻度 | - |
| | 子育てで意識して行っていること(尺度) | - |
| | 子どもをもってよかったこと(尺度) | - |
| | 子育ての不安や悩み | + |
| | 母親の喫煙状況 | + |
| 調査への理解・関心 | 父親の喫煙状況 | + |
| | 無回答項目の数 | + |
| | 各変数の不詳・無回答ダミー変数 | |

²⁾ 正確には以下の式で表される対数変換された脱落の生起オッズを推定する。

$$\log(p/(1-p)) = a + b_i \sum x_i, \quad p: \text{脱落確率}, b_i: \text{係数}, \sum x_i: \text{説明変数}$$

なお、オッズとは、ある事象が生起する確率に対してその事象が生起しない確率の比 $(p/(1-p))$ である。

なお、分析からは非嫡出子、父母ともに外国籍であるサンプル、父母と同居していないサンプルは除外した。その理由は3つある。第1に、非嫡出子や父母ともに外国籍であるサンプル、および父母と非同居のサンプルについては、すでに2変量解析において脱落と深く関わるということが明らかにされている。しかし、各変数とも脱落に関わるカテゴリのサンプル数が少ないので多変量解析では安定した推定値が得られない。第2に、父母の就業状況や家事・育児頻度などのいくつかの重要な変数において父母の同居が前提となっている。そして第3に、脱落の分析においては、これらのマイノリティ層よりもむしろ標準的な属性をもつサンプルの中でどのような特性が脱落と因果関係をもつのが重要な関心となることが挙げられる。なお、多胎児は同じ父母の属性をもつことから、分析にバイアスを与えないよう出生順位が1番目のサンプルのみ分析に含めた。

脱落確率の推定に関するロジスティック回帰分析の結果を表3-2に示す。ロジスティック回帰分析では通常のコэффициент (b) の他にオッズ比 (exp(b)) を算出する。オッズ比では対象とするカテゴリの脱落オッズが、各変数の準拠カテゴリ (reference category) に対して何倍高いのか (あるいは低いのか) を知ることができる。オッズ比が1より高いほど脱落する可能性が高いことを意味しており、1より低いほど脱落する可能性が低いことを意味している。なお、表中では係数が0、オッズ比が1のカテゴリが準拠カテゴリである。また、量的変数のオッズ比はその変数が1単位増えることによるオッズの変化を表している。

はじめに夫婦の人口学的属性と脱落との関係から考察する。父母の年齢が低い場合、および夫婦の年齢差が大きい場合に脱落が生じやすいという傾向が多変量解析においても確認されている。また、婚前妊娠による対象児はそうでない対象児に比べて46%も脱落するオッズが高いことが明らかである。これらの要因は、結婚生活の不安定要因となりうることから、離婚が脱落確率を促していることが示唆される。しかし、これらの要因を統制すると、夫婦の同居期間や既往出生児数が多いほど脱落確率が上昇する傾向にある。特に、母親の既往出生児数は2変量解析では有意ではなかったが、ここでは3人以上である場合には既往出生児数がゼロの母親に比べて脱落のオッズが2倍以上と高い値を示している。家族形態に関しては、核家族世帯より3世代同居世帯において脱落が少ない。したがって、調査の対象となる子のはじめての出産である場合や祖父母と同居している世帯ほど調査への協力度が高いといえる。

一方、父母のどちらかが外国籍である場合に脱落確率が高い。国際結婚の夫婦は日本人同士の夫婦よりも脱落するオッズが約3倍となっている。外国への移動などによる理由もあろうが、母親が外国人である場合、調査票を理解して記入するのが困難であるということも影響しているのかもしれない。この点に関する改善がない限り、外国籍の母親をもつ出生児サンプルは年々減少していくことが予測される。

居住地方については、他の説明変数による影響を統制した結果、九州や沖縄地方のみに脱落率が高い傾向が認められた。また、統計的に有意ではないが中国・四国地方に

においても脱落率が高い。その理由については本分析の結果からは明らかではないが、中国地方以西では調査協力が得にくい状況にあるようだ。都市規模は単独では脱落に作用しないが、婚前妊娠との間に10%水準の交互作用効果が認められた。婚前妊娠による出産は、母親が13大都市の居住者である場合により強く脱落を促し、母親が郡部の居住者である場合には脱落を緩和している。このことは婚前妊娠による結婚生活の安定性が都市規模によって異なることを示唆しているのではないかと。

次に夫婦の社会経済的屬性についてみる。2変量解析によって確認された通り、世帯年収が高いほど脱落確率は低い。しかし、その影響は年収の自然対数値³⁾によって表される非線形である。したがって、世帯年収が一定以上になると脱落確率には影響を与えない。このことは世帯年収の低い層ほど脱落する傾向が強いことを意味している。また、世帯年収が不詳であるサンプルは、平均的な世帯年収をもつサンプルに比べて脱落するオッズが73%も高い傾向にある⁴⁾。年収は標本調査において最も回答を得にくい項目の1つである。年収への回答の有無が調査協力の度合いを反映していると考えると、調査協力が得にくいサンプルほど脱落しやすいことが明らかである。

母親の就業状況については、求職中であった母親において脱落確率が高い傾向がある。しかし、分析は第1波において育児休業中であった母親において最も脱落確率が高かったことを示している。これらの結果は母親の就業状態が無業から有業に変化することが脱落に大きな影響を与えることを示唆している。また、統計的には有意ではないが、常勤であった母親の脱落オッズも大きい。女性の育児と就業の両立は少子化政策の要である。働く母親の脱落は当調査の政策的意義を減ずるものであり、早急な対策が必要である。

一方、母親とは反対に、父親の就業については父親の週当たり労働時間が40時間未満、すなわち無職やパートにおいて脱落確率が高い。こうした傾向は世帯年収が高いほど脱落率が低いという知見と一致するものである。また、父親が安定した職をもたないことは家庭の不安定要因となっている可能性もある。さらに、父親が週に60時間以上働いている場合や父親の労働時間が不詳である場合にも脱落率が高い傾向にある。父親の企業規模については、従業員数5-99人の中小企業に比べて、従業員数が100人以上の大企業や官公庁勤務である場合に脱落確率が低い。特に、当調査が官庁統計調査であることを反映してか、父親が官公庁勤務である場合に脱落確率が低い。

住まいの広さについては、手狭であると感じていた世帯ほど脱落する傾向がみられた。このことは世帯年収との相関を統制した上で得られたものであるため、転居による追跡不能が脱落に繋がっている可能性を示唆している。

3) 世帯年収が0円である場合は1万円として対数値を算出した。その他は実数を対数化した。

4) 世帯年収が不詳であるケース(3093ケース)には、世帯年収の平均値を代入した。この場合、世帯年収不詳ダミーは、世帯年収の平均値に対するオッズ比を表す。父母の育児・家事得点の不詳ダミーについても同様の操作を行った。

表 3-2 脱落確率に関するロジスティック回帰分析

| | B | Exp(B) |
|---------------|--------|----------|
| 父親の年齢 | -0.026 | 0.97 ** |
| 母親の年齢 | -0.056 | 0.95 *** |
| 夫婦の年齢差(絶対値) | 0.034 | 1.03 *** |
| 夫婦の同居期間(年) | 0.034 | 1.03 ** |
| 婚前妊娠の有無 | | |
| 婚前妊娠なし | 0.000 | 1.00 |
| 婚前妊娠 | 0.381 | 1.46 *** |
| 母親の既往出生児数 | | |
| 0人 | 0.000 | 1.00 |
| 1人 | 0.230 | 1.26 *** |
| 2人 | 0.278 | 1.32 ** |
| 3人以上 | 0.706 | 2.03 *** |
| 家族形態 | | |
| 核家族 | 0.000 | 1.00 |
| 3世代同居 | -0.173 | 0.84 ** |
| 父母の国籍 | | |
| 両親とも日本人 | 0.000 | 1.00 |
| どちらかの親が外国籍 | 1.051 | 2.86 *** |
| 居住地方 | | |
| 北海道 | -0.021 | 0.98 |
| 東北 | 0.055 | 1.06 |
| 関東 | 0.000 | 1.00 |
| 北陸 | -0.189 | 0.83 |
| 中部 | -0.032 | 0.97 |
| 近畿 | 0.006 | 1.01 |
| 中国 | 0.134 | 1.14 |
| 四国 | 0.108 | 1.11 |
| 九州・沖縄 | 0.266 | 1.30 *** |
| 市郡規模 | | |
| 13大都市 | -0.086 | 0.92 |
| その他の市 | 0.000 | 1.00 |
| 郡部 | -0.003 | 1.00 |
| 婚前妊娠×市郡規模 | | |
| 13大都市×婚前妊娠あり | 0.236 | 1.27 + |
| その他の市×婚前妊娠あり | 0.000 | 1.00 |
| 郡部×婚前妊娠あり | -0.096 | 0.91 |
| 世帯年収の対数値 | -0.179 | 0.84 *** |
| 世帯年収不詳ダミー | 0.548 | 1.73 *** |
| 母親の就業状況 | | |
| 仕事を探している | 0.386 | 1.47 *** |
| 仕事を探していない・学生 | 0.000 | 1.00 |
| 育児休業中 | 0.464 | 1.59 *** |
| 常勤 | 0.209 | 1.23 |
| パート・アルバイト・その他 | 0.166 | 1.18 + |
| 自営業・家業 | 0.040 | 1.04 |
| 父親の労働時間(週当たり) | | |
| 0時間 | 0.302 | 1.35 * |
| 1-19時間 | 0.351 | 1.42 *** |
| 20-39時間 | 0.166 | 1.18 * |
| 40-59時間 | 0.000 | 1.00 |
| 60時間以上 | 0.104 | 1.11 * |
| 労働時間不詳 | 0.341 | 1.41 ** |
| 父親の企業規模 | | |
| 常勤以外 | -0.016 | 0.98 |
| 1~4人 | 0.090 | 1.09 |
| 5~99人 | 0.000 | 1.00 |
| 100~499人 | -0.151 | 0.86 * |
| 500人以上 | -0.126 | 0.88 * |
| 官公庁 | -0.227 | 0.80 * |
| 企業規模不詳 | 0.121 | 1.13 |

(次ページに続く)

(前ページの続き)

| | B | Exp(B) |
|-----------------|----------|----------|
| 住まいの広さ | | |
| 手狭だと感じている | 0.134 | 1.14 ** |
| 必要な広さはある | 0.000 | 1.00 |
| 十分にゆとりがある | 0.029 | 1.03 |
| あまり考えたことがない | 0.037 | 1.04 |
| 未熟児の別 | | |
| 未熟児ではない | 0.000 | 1.00 |
| 未熟児 | -0.003 | 1.00 |
| 単・多胎の別 | | |
| 単胎児 | 0.000 | 1.00 |
| 多胎児(双子以上) | 0.461 | 1.58 * |
| 平日日中の保育者 | | |
| 父母のどちらか | 0.000 | 1.00 |
| 祖父母 | 0.159 | 1.17 |
| 保育士・ベビーシッター等 | 0.208 | 1.23 + |
| 授乳状況 | | |
| 母乳のみ | -0.412 | 0.66 *** |
| 人工乳のみ | 0.412 | 1.51 ** |
| 人工乳(初乳のみ) | 0.096 | 1.10 |
| 混合乳 | 0.000 | 1.00 |
| 母親の育児頻度得点 | -0.466 | 0.63 * |
| 母親の育児頻度得点の2乗 | 0.015 | 1.01 * |
| 母親の育児頻度不詳ダミー | -0.752 | 0.47 *** |
| 母親の家事頻度得点 | -0.218 | 0.80 ** |
| 母親の家事頻度得点の2乗 | 0.007 | 1.01 ** |
| 母親の家事頻度不詳ダミー | 0.215 | 1.24 |
| 父親の育児頻度得点 | -0.007 | 0.99 |
| 父親の育児頻度不詳ダミー | 0.342 | 1.41 ** |
| 父親の家事頻度得点 | 0.008 | 1.01 |
| 父親の家事頻度不詳ダミー | -0.127 | 0.88 |
| 子育てで意識して行っていること | -0.006 | 0.99 |
| 子どもをもってよかったこと | 0.016 | 1.02 |
| 子育ての不安や悩み | | |
| すごくある | -0.029 | 0.97 |
| 少しある | -0.061 | 0.94 |
| ほとんどない | 0.000 | 1.00 |
| 母親の喫煙状況 | | |
| 吸っていない | 0.000 | 1.00 |
| 室外で吸う | 0.424 | 1.53 *** |
| 室内で吸う | 0.567 | 1.76 *** |
| 父親の喫煙状況 | | |
| 吸っていない | 0.000 | 1.00 |
| 室外で吸う | 0.169 | 1.18 ** |
| 室内で吸う | 0.231 | 1.26 *** |
| 不詳 | 0.263 | 1.30 + |
| 無回答項目の数 | 0.045 | 1.05 * |
| 定数 | 6.768 | *** |
| サンプル数 | 43398 | |
| 脱落数 | 2586 | |
| カイ2乗値 | 1497.608 | *** |
| 自由度 | 70 | |

***: 0.1%水準、**: 1%水準、*: 5%水準、+: 10%水準で統計的に有意(Wald検定)

次に、対象児の属性について考察する。未熟児と脱落との因果関係は全く認められないことが明らかである。第一波の調査時点で対象児は生後6ヶ月となっている。したがって、未熟児として出生したことと死亡率との相関は弱くなっているのであろう。しかし、多胎児である場合は単胎児よりも57%も脱落するオッズが高い。未熟児の影響を統制していることを考慮すると、多胎児の母親は複数の同じ調査票に記入しなければならないことが負担となっている可能性が高い。

育児の状況についてみると、平日日中の保育者が家族以外である場合に脱落確率が高い。また、母乳のみで保育されている対象児は、母乳と人工乳との混合乳で保育されている対象児よりも脱落オッズが34%も低い。逆に人工乳のみで保育されている対象児は脱落オッズが51%高い。これらのことは乳幼児保育の利用者において脱落が多いことを示唆している。乳幼児保育の利用者には働く母親が多いことから、このことは母親の就業状況と脱落との関係を追認する結果である。一方、母親の家事・育児頻度の得点と脱落率の間には2次曲線の関係がみられた。実際の家事・育児頻度得点（各0-18点）をそれぞれのパラメーターに代入してみると、いずれもその影響は家事や育児の頻度が上がるほど脱落オッズが曲線的に低下するという単調減少を示していた。したがって、家事や育児への参加度が高い母親ほど脱落が少ないという傾向が確認できた。特に育児頻度得点において脱落確率を低下させる影響が強くみられる。しかし、これとは対照的に2変量解析と同様に父親の家事・育児頻度は脱落には影響していない。

育児意識については、脱落と有意な関係にある変数がほとんどない。おそらく育児意識は育児の頻度や喫煙状況などの行動変数と相関があるので、これらの変数を統制したことにより脱落との相関がなくなったものと思われる。これに対して、父母の喫煙状況は脱落の強い予測要因となっている。対象児の父母が喫煙している場合に脱落オッズが高まる。特に母親の喫煙は父親の喫煙よりも脱落に与える影響が大きい。また、父母の喫煙状況はそれが室内か室外かによって異なり、室内で喫煙する親ほど脱落が著しい傾向がみられる。たばこの副流煙が第三者の健康を害す恐れがあることは広く知られている。親が室内で喫煙するほど乳幼児はたばこによる健康被害を受けやすいものと思われる。保育者である父母の喫煙は育児に対する意識の低さを反映することによって脱落を促しているのではないか。

最後に、調査への理解・関心を表す無回答項目について考察する。モデルでは世帯年収や父親の労働時間、企業規模、父母の家事・育児頻度、そして父親の喫煙状況などにおいて、回答が不詳・無回答であることを意味する不詳ダミーを投入した。各不詳ダミー変数は、主としてサンプル数の減少を防ぐことを目的として、不詳カテゴリーのサンプル数が1000ケース以上である時に使用した⁵⁾。しかし、無回答項目の発生が調査への協力態度に関わると仮定すれば、これらの変数と脱落との関係を考察することも可能である。これら

⁵⁾ ただし、父親の喫煙状況（917サンプル）および母親の育児得点（905サンプル）においては、1000サンプル以下であったが不詳ダミー変数を投入した。

の不詳ダミー変数のうち、世帯年収、父親の労働時間、父親の育児頻度、そして父親の喫煙状況において、無回答（不詳回答）が脱落を促していることが明らかとなった。分析に用いたサンプルはいずれも父母が同居しているケースに限られているが、特に父親に関する項目が不詳である場合に脱落する傾向がある。

また、2変量解析でみられたように、無回答項目の数が多いほど脱落確率が高まる。モデルで使用した変数についてはすべて回答していること（不詳ダミーがある項目には無回答も含める）という条件付ではあるが、モデルでは無回答項目がひとつ増える毎に脱落のオッズが5%高まることが示されている。分析対象サンプルにおける無回答項目数の平均は約3問となっている。したがって、無回答項目が3問ある場合の脱落オッズは無回答項目がないサンプルの1.16倍（ $=1.05^3$ ）となる。無回答項目が量的変数であることを考慮すると、無回答項目の数は脱落を予測する強い要因であるといえる。したがって、調査対象者の理解・関心を促し、自主的な調査協力を促すことが脱落を減らす最も重要な対策である。

4. 脱落への対処法

本節では、脱落への対処法について、その概略を述べる。パネル調査における脱落への対処法は3つに大別することができる。第1は調査票の設計、サンプリング、調査票の配布および回収の各段階において脱落を減らすための様々な施策を講じることである。前出の家計経済研究所による「消費生活に関するパネル調査」においては、転居先情報を捕捉するために調査月以外にも挨拶状等を送付し、さらには近親者や知人の情報を把握することに努めている（坂本，2004）。その他にも調査結果掲載新聞記事・雑誌資料冊子、および前年の調査結果をまとめた冊子を届ける、調査回収の際に謝金を支払う等の様々な工夫を行っている（坂本，2004）。また、家計経済研究所のパネル調査は留置き法であることから、同じ調査員が訪問することで、回答者との間に信頼関係が生まれ、調査に対して協力的になることが報告されているという（坂本，2004）。これらの施策が家計研パネル調査における各年度調査の回収率95%という高回収率に繋がっている。回答者の転居先情報を捕捉するのはもちろんのこと、回答者に調査結果を周知することで調査の意義に対する理解を促す、回答者の負担を減らすために項目数を減らす、そして謝金によるインセンティブを付与する等の施策を講じて対象者の調査協力を促していく必要があるだろう。

21世紀縦断調査においても、以下のような施策を講じて回答者の確保に努めている。1) 1月、7月にそれぞれ誕生日カードを送付。2) 調査票を送付する際及び誕生日カードを送付する際に、住所等が変わった場合の連絡用はがきを同封し、転居情報を得ている。3) 調査対象者に対しては、概況報告にあわせて調査結果の報告を送付。4) 記入者負担の軽減については、対象者へ直接の郵送調査であることから、第1回調査以降の質問項目を20問程度としている。5) 調査協力者に対する謝礼として、調査票送付の際に年齢に合わせた物品とボールペンを同封。6) 調査票の回収に際しては、3回の督促（3回目は調査票を同封）を行っている。7) 災害時の見舞い状の送付。こうした脱落を減らすための対処は、パネル

調査を行う上で最も重要かつ有効な手段であると考えられている⁶⁾。

もう 1 つの対処法としては、追加標本が挙げられる。追加標本とは、脱落によるサンプル減少を補うために計画標本に加えて新たな標本を追加することである。以下に、相田（刊行予定）をもとに追加標本の方法および問題点について述べる。標本の追加には 2 通りの方法が考えられる。ひとつは、脱落の頻度に逆比例した量の標本を標本抽出地点毎に追加する方法である。これにより標本抽出地点毎のサンプル数を常に一定数保つことが可能となる。しかし、この方法の問題点は標本の抽出確率が地点毎に異なることにより、無作為抽出（および層別多段抽出）を仮定できなくなり、平均および分散の推定にバイアスが生じることである。先の分析で明らかのように、中国地方以西において脱落の確率が高い。このことは四国地方以西に居住するサンプルが他の地方に居住するサンプルより多く抽出されることを意味し、計画された母集団から乖離した属性をサンプルに混入することとなる。もう 1 つの標本追加方法は、層化したサンプルに均一に標本を追加するというものである。つまり脱落の度合いに応じて抽出地点毎に標本を補充するのではなく、全抽出地点を対象とした層化抽出によって標本を追加する。この方法における利点は、無作為抽出の仮定が適用できるので、統計量の推定にバイアスが生じないことである。ただし、脱落が著しい地点の標本数の減少が防げないという欠点がある。

結論として、前者の方法ではバイアスを補正するために抽出確率の違いを考慮したウェイト（重み付け）を適用する必要があるが、またその結果として分散が増大するなどの望ましくない副作用が生じる。したがって、追加標本を使用する場合には後者の方法が望ましいものと思われる。しかし、後者の方法の欠点を補うための前提条件として、先に挙げた脱落を減らすための努力を十分に行うことが重要である（相田、刊行予定）。また、いずれの方法を用いたとしても、同一サンプルの変化を観察することによって得られるパネルデータの数々の利点（樋口他，2004）が失われてしまうことには変わりがない。そのため、追加標本はあくまで 2 次的な対処法であり、脱落があまりに大きく、調査の継続的な実施が困難となるなどの必要がない限り行わない方が望ましい。

最後に、収集されたデータに対して脱落サンプルの属性を加味したウェイトを適用する、または脱落によるバイアスを補正したパラメーター推定を行うなどの方法を挙げることができる。近年においては経済学や統計学などの分野において、パネル調査における脱落を補正するためのウェイトの使用（Lepkowski & Couper, 2002; Wooldridge, 2002）や欠損値の補完（Little & Rubin, 2002）、さらには脱落によるバイアスを補正する特殊な推定法（Heckman, 1974; Heckman, 1979; Wooldridge, 2001）などに関する様々な研究成果が報告されている。本稿における目的からは逸脱するため、これらの詳細については言及しない。しかし、一般に推定法によるバイアスの補正は、個々の分析者の力量に任されるため、分析の一意性を保つことが難しいという問題がある（相田、刊行予定）。そのため、欧米に

⁶⁾ 海外のパネル調査におけるサンプル追跡方法については、Farrington, et. al.(1990)に詳しい。

おける大規模調査では調査機関が標本設計や脱落過程を反映させた種々のウェイトを作成し、分析者にそれらのウェイトを使用することを推奨している。また、同様の理由から収入などの欠損値の多い変数に対して、欠損値の補完処置を施した変数を作成してユーザーに提供しているケースもある。これらの手法は第1、および第2の対処法に比べて、パネル脱落に対する事後的かつ間接的な対処方法であるといえることができる。しかし、パネル調査における脱落が避けられないものである以上、こうしたウェイトの使用やパラメータ推定における対処方法についても検討する必要があるだろう。特にウェイトの使用は記述統計の算出において有用であるものと思われる。

おわりに

本稿では21世紀出生児縦断調査の第2回調査までの確定データを用いて、当調査の脱落要因に関する考察を行った。具体的には、第1回調査における対象サンプルの属性および回答内容と第2回調査における脱落との相関および因果関係について検証した。本稿で得られた主な知見は以下である。

1. 調査回答者の9割が母親である。脱落要因にも父親よりも母親の属性が強く反映されており、調査の継続回答に関する意思決定は主に母親によって行われている。
2. 夫婦の年齢が若い場合や夫婦の年齢差が大きい場合、そして婚前妊娠などの結婚の不安定要因は脱落を促す傾向がある。
3. 片親家庭や非嫡出生といった非標準的な世帯において脱落率が高い。
4. 対象児が母親のはじめての出産であった場合や祖父母と同居している世帯ほど調査への協力度が高い。
5. 外国籍の親をもつ出生児サンプルには脱落が多い。今後これらの層をフォローアップしていくには調査票の翻訳や記入指導などが必要である。
6. 中国地方以南、特に九州・沖縄地方で脱落率が高い傾向がある。
7. 働く母親の脱落が顕著である。有業女性の脱落は当調査の政策的意義を減ずるものであり、早急な対策が必要である。
8. 低所得世帯ほど脱落する傾向がある。近年における婚前妊娠の増加にみられるように、低所得世帯への子育て支援はニーズが大きく、家族政策の重要な課題となると思われるため、今後何らかの対処が必要である。
9. 多胎児の母親は複数の同じ調査票に記入しなければならないことが負担となっている可能性が高い。その場合、多胎児の回答者への負担を軽減する処置が必要である。
10. 家事や育児への参加度が高い母親ほど脱落が少ない。しかし、父親の家事・育児参加は脱落に影響しない。
11. 父母（特に母親）が喫煙している場合、脱落確率が高い。父母の喫煙は育児に対する意識の低さを反映することによって脱落を促しているものと思われる。

12. 無回答の発生は脱落を予測する強い要因であるといえる。調査対象者の理解・関心を促し、自主的な調査協力を促すことが脱落を減らす最も重要な対策である。

本分析において得られた結果は多変量解析に基づくものであり、説明変数間の相関を統制しているという点において、各変数と脱落との関係はより因果関係に近い数値であるといえる。しかしながら、モデルの当てはまりは決して良好とはいえない。例えば、分析サンプル中に脱落は2586ケースあったが、モデルで判別できたのはわずかに8ケースのみであった。これはひとつには脱落が全体に占める割合が6.6%（分析サンプルでは5.9%）と少なかったこと、そして恐らくは父母の離婚や対象児の死亡など脱落を予測するいくつかの重要な変数を欠いた事による。

しかし、他の先行研究をみてもパネル調査における脱落の規定要因に関するモデルの当てはまりは一様に低いことが報告されており（Fitzgerald, et. Al., 1998; 坂本, 2004）、モデル適合度の低さは今回の分析に特有の問題とはいえない。むしろ、このことは脱落がサンプルの属性によって選択的（selective）に発生しているものの、その大部分が攪乱要因（noise）であり、サンプルの代表性を損なうほどの大きな影響は与えていないことを示唆している（Fitzgerald, et. Al.）。ただし、脱落を補正するウェイトの作成には脱落確率を予測する、できるだけ精度の高いモデルを構築することが必要となる（Fitzgerald, et. Al., 1998; 相田, 刊行予定）。今後、そうした必要がある際には脱落の理由や父母の配偶関係の変化といった人口動態データとのリンケージなども視野に入れつつ、脱落発生のメカニズムを明らかにしていく必要があるだろう。

本稿では21世紀出生児縦断調査における脱落要因について論じてきた。しかし強調されるべきは、脱落によるバイアスが問題となるのは単変量の平均および分散の推定や2変量の相関を算出する場合においてであり、モデルが正しく設定されている限り、脱落は多変量解析による回帰係数の推定には有意な影響を与えないという事実である。このことは海外のパネルデータを用いた多くの脱落研究において実証されている（Fitzgerald, et. Al., 1998; Lillard & Panis, 1998; Alderman, et. Al., 2001; Falaris, 2003）。したがって、パネル調査には必ず脱落の問題が伴うが、そのことによって調査の意義が減じることはない。脱落という観点からみるならば、むしろパネル調査は記述統計の算出よりも多変量モデル等を用いた説明的研究に適したデータであるといえる。今後は記述統計の枠に留まらず、パネルデータに特有の高度な分析手法（固定効果・ランダム効果モデル等の計量経済分析や生存分析）を積極的に適用することで、21世紀縦断調査を少子化問題の解明に役立てていくことが肝要である。

参考文献

- Alderman, H., J. R. Behrman, H. Kohler, J. A. Maluccio, S. C. Watkins, 2001, "Attrition in Longitudinal Household Survey Data", *Demographic Research*, 5(4), 80-123.
- Allison, P. D., 2002, *Missing Data*, Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-136, Thousand Oaks, CA: Sage.
- Falaris, E. M., 2003, "The Effect of Survey Attrition in Longitudinal Surveys: Evidence from Peru, Cote d'Ivoire and Vietnam", *Journal of Development Economics*, 70, 133-157.
- Farrington, D. P., B. Gallagher, L. Morley, R. J. St. Ledger, and D. J. West, 1990, "Minimizing Attrition in Longitudinal Research: Methods of Tracing and Securing Cooperation in a 24-year Follow-up Study", in edited by D. Magnusson, and L. R. Bergman, *Data Quality in Longitudinal Research*, New York: Cambridge University Press, 122-147.
- Fitzgerald, J, P. Gottschalk, and R. Moffitt, 1998, "An Analysis of Sample Attrition in Panel Data", *The Journal of Human Resources*, 33(2), 251-299.
- Heckman, J.J., 1974, "Shadow Prices, Market Wages, and Labor Supply", *Econometrica*, 42, 679-694.
- Heckman, J.J., 1979, "Sample Selection Bias as a Specification Error", *Econometrica*, 47, 153-162.
- Lepkowski, J.M., and M.P. Couper, 2002, "Nonresponse in Second Wave of Longitudinal Household Surveys", Groves, Eltice, 2002, *Survey Non-response*, John Willey & Sons, 259-272.
- Lillard, L. A., and C. W. A. Panis, 1998, "Panel Attrition from the PSID", *The Journal of Human Resources*, 33(2), 437-457.
- Little, R.J.A. and Rubin, D.B. 2002, *Statistical Analysis with Missing Data*, 2nd.ed, John Wiley & Sons.
- Wooldridge, J., 2001, *Econometric Analysis of Cross Section and Panel Data*, MIT Press.
- Wooldridge, J., 2002, "Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification", *cemmap Working Paper CWP11/02*.
- 相田真彦, 刊行予定, 「縦断的調査における非等確率抽出と欠測の問題 - JESIII データを用いた検討 - 」, 『選挙学会紀要』。
- 家計経済研究所, 2004, 『共依存する家計 - 消費生活に関するパネル調査(第11年度)- 』, 東京, 国立印刷局。
- 厚生労働省大臣官房統計情報部, 2002, 『「出生に関する統計」の概況』, 『人口動態統計特殊報告』
- 坂本和靖, 2004, 「サンプル脱落に関する分析 - 脱落サンプルの規定要因と推計バイアスについて - 」, 2004年日本経済学会春季大会報告論文。
- 樋口美雄, 太田清, 新保一成, 2004, 「パネルデータによる経済分析①-⑤」, 『経済セミナー』, 6-10月号。