

平成 24 年度内閣府大臣官房統計委員会担当室請負調査

## 統計データの補完推計に関する調査

### 報告書

平成 25 年 3 月

株式会社 野村総合研究所



# 目次

|   |           |
|---|-----------|
| <b>第1章 調査研究と結果の概要</b>                             | <b>1</b>  |
| 1 調査研究の目的   | 1         |
| 2 調査研究の方法   | 2         |
| 1) 調査研究の全体フロー                                     |           |
| 2) 有識者による研究会の設置                                   |           |
| 3) 国内調査・海外調査の視点                                   |           |
| 3 調査結果の概要   | 6         |
| 1) 国内における補完推計に関する研究                               |           |
| 2) 海外における補完推計に関する取組                               |           |
| 3) 災害時における補完推計などの対応                               |           |
| <b>第2章 調査における欠測の分類と対応</b>                         | <b>14</b> |
| 1 欠測が生じる状況  | 14        |
| 2 欠測パターン  | 15        |
| 1) ユニット非回答(Unit Non Response)                     |           |
| 2) 項目非回答(Item Non Response)                       |           |
| 3 欠測メカニズム   | 16        |
| 1) MCAR(Missing Completely At Random: 完全にランダムな欠測) |           |
| 2) MAR(Missing At Random: ランダムな欠測)                |           |
| 3) NMAR(Not Missing At Random: ランダムでない欠測)         |           |
| 4 欠測を含むデータの取扱                                     | 17        |
| 1) 完全ケースに基づく分析(Complete Case Analysis)            |           |
| 2) 利用可能なケースに基づく分析(Available Case Analysis)        |           |
| 5 欠測値の補完  | 18        |
| 1) 単一代入法  |           |
| 2) 多重代入法  |           |
| (参考) 傾向スコア  |           |
| (参考) ユニット非回答に対応する手法 キャリブレーション(Calibration)        |           |
| 6 調査における欠測への対応                                    | 23        |
| <b>第3章 国内における補完推計に関する研究</b>                       | <b>25</b> |
| I 国内における補完推計に関する研究の概要(一覧)                         | 26        |
| II 国内における補完推計に関する研究事例                             | 28        |
| 1 経済センサスー活動調査                                     | 28        |
| 1) 調査の概要  |           |
| 2) 研究の経緯・目的                                       |           |
| 3) 研究の体制・時期                                       |           |
| 4) 研究手法の概要  |           |
| 5) 今後の課題  |           |
| 2 国民生活基礎調査  | 33        |
| 3 特定サービス産業実態調査                                    | 39        |

|  |           |
|--|-----------|
| 4 全国企業短期経済観測調査 .....   | 45        |
| 5 日本人の国民性調査 .....  | 49        |
| <b>第4章 海外における補完推計に関する取組</b> .....                            | <b>53</b> |
| I 海外における補完推計に関する取組の概要(一覧) .....                              | 54        |
| II 海外における補完推計に関する調査結果 .....                                  | 56        |
| 1 アメリカセンサス局 (United States Census Bureau) .....              | 56        |
| 1) ACS (American Community Survey: アメリカ地域社会調査) における取組        |           |
| (1) 調査概要   |           |
| (2) 欠測を防ぐ取組  |           |
| (3) 補完手法概要   |           |
| 2) Population Census (US Census:国勢調査) における取組                 |           |
| 2 アメリカ労働統計局 (United States Bureau of Labor Statistics) ..... | 64        |
| 1) OES(Occupational Employment Statistics: 職業雇用統計)における取組     |           |
| 3 カナダ統計局 (Statistics Canada) .....                           | 66        |
| 1) LFS(Labour Force Survey: 労働力調査)における取組                     |           |
| (コラム)補完を行うプログラム Banff .....                                  | 71        |
| 4 westat 社 (調査会社).....                                       | 73        |
| 5 ミシガン大学 .....   | 74        |
| <b>第5章 災害時における補完推計などの対応</b> .....                            | <b>75</b> |
| I 国内・海外における大規模災害への対応(一覧) .....                               | 76        |
| II 国内における東日本大震災への対応事例—労働力調査 .....                            | 78        |
| III 海外における大規模災害への対応・取組 .....                                 | 81        |
| 1 アメリカセンサス局 (United States Census Bureau) .....              | 81        |
| 1) ACS (American Community Survey: アメリカ地域社会調査) における対応・取組     |           |
| (1) ACS における対応の概要  |           |
| (2) 実査上の影響と対応  |           |
| 2 アメリカ労働統計局 (United States Bureau of Labor Statistics) ..... | 82        |
| 1) CES(Current Employment Survey: 雇用状況調査)における対応・取組           |           |
| 2) CPS(Current Population Survey: 就業状態調査)における対応・取組           |           |
| 3 カナダ統計局 (Statistics Canada) .....                           | 86        |
| 1) 統計調査ごとの BCP(Business Continuity Plan: 事業継続計画)の策定          |           |
| 2) 異常時の状況把握を迅速に行うための特殊モジュールの設定                               |           |
| <b>第6章 統計調査に欠測値補完を適用する際の示唆・課題</b> .....                      | <b>88</b> |
| 1 平常時における補完に対する示唆・課題 .....                                   | 89        |
| 2 災害時における補完に対する示唆・課題 .....                                   | 94        |
| <b>附属資料目次</b> .....  | <b>97</b> |
| I 国内における補完推計に関する研究   |           |
| II 海外における補完推計などに関する取組  |           |
| III 参考ウェブサイト・文献一覧  |           |

## **第1章 調査研究と結果の概要**

### **1. 調査研究の目的**

企業の情報管理意識や個人の情報保護意識の高まり等により、近年の統計調査を取り巻く環境が厳しさを増している中で、統計データの補完推計は、統計精度の維持・向上を図る上で重要となっている。

また、東日本大震災などの大規模災害の際に、特定の地域で調査の実施が困難になるという事態も発生しており、「平成 22 年度統計法施行状況に関する審議結果報告書」(※)においても、全国を対象とする基幹統計調査で調査対象地域の一部除外等の取扱をした場合、被災地の状況を踏まえて可能な限り補完的、補足的な調査、推計等の措置を講ずる必要があるとされている。

そこで本調査研究では、今後の統計委員会における審議等に資することを目的として、欠測値の補完方法の種類、国内における補完推計の研究事例、海外における欠測値補完の適用事例等について、文献調査及びヒアリング調査を通じて整理するとともに、各種の統計調査における欠測値補完方法の適用の方向性について考察を行った。

(※)「平成 22 年度統計法施行状況に関する審議結果報告書」(平成 23 年 9 月 22 日 統計委員会)  
<http://www5.cao.go.jp/statistics/report/report.html#11>

# 第1章 調査研究と結果の概要

## 2. 調査研究の方法

### 1) 調査研究の全体フロー

本調査研究では、まずプレ調査(文献調査)として、文献・書籍、各種学術機関(大学、学会等)、及び統計機関のウェブサイト等から得られる情報を収集・整理した。プレ調査では欠測値補完に関する国内の研究事例や海外の適用事例を収集・整理し、本調査における論点整理、及び有識者からの事前の論点収集を行った。

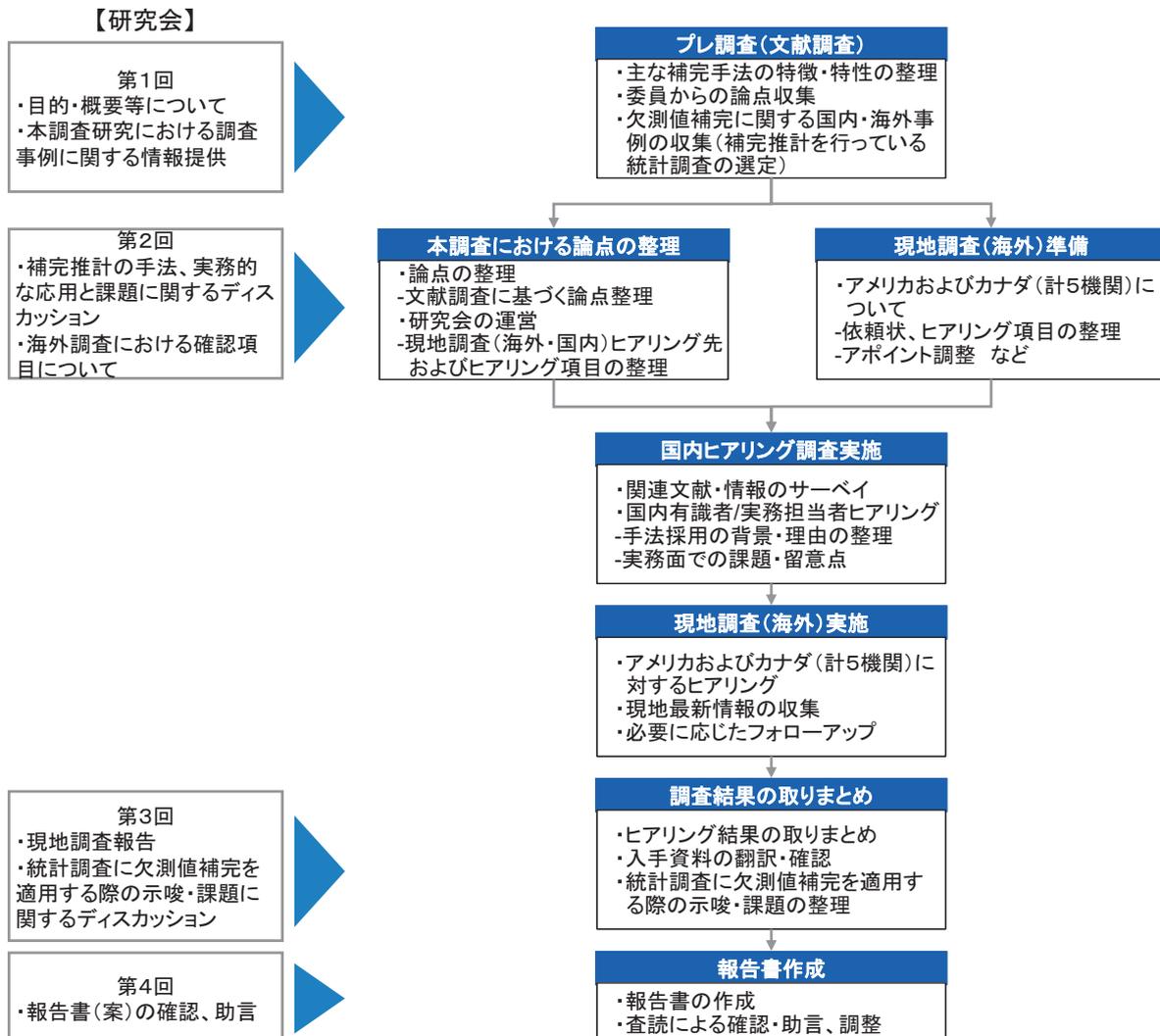
その上で、有識者研究会を実施し、本調査においてフォーカスすべき論点の整理を行った。

有識者研究会での検討を通じて、現地調査(海外)でなければ把握が難しい事項を中心に、現地調査による情報の深堀りを行った。海外事例については、アメリカ統計学会(American Statistical Association)の研究者が多く在籍するアメリカ、カナダを対象とすることで、学術的な視点を含めた検討を行うこととした。

また、国内における補完推計の研究については、実務担当者(総務省統計局を含め各府省担当者等)に対するヒアリングを実施することで、実務的な課題・留意点を意識した検討を行うことができるよう配慮した。

以上を踏まえ、研究会において「統計調査に欠測値補完を適用する際の示唆・課題」についてディスカッションを行い、その結果を最終報告書として取りまとめた。

本調査研究の全体フロー



本調査研究において調査対象とした調査(国内)及び調査機関(海外)

<国内>

|         |              |
|---------|--------------|
| 総務省     | 経済センサスー活動調査  |
| 厚生労働省   | 国民生活基礎調査     |
| 経済産業省   | 特定サービス産業実態調査 |
| 日本銀行    | 全国企業短期経済観測調査 |
| 統計数理研究所 | 日本人の国民性調査    |

<海外>

|      |  |
|------|--|
| アメリカ | 米国統計局(US Census Bureau)<br>米国労働統計局(BLS: US Bureau of Labor Statistics)<br>ミシガン大学(University of Michigan)<br>westat 社 |
| カナダ  | カナダ統計局(Statistics Canada)  |

アメリカにおいては、統計調査を受託する調査会社である westat 社に加えて、アカデミック分野における取組や見解を整理するためにミシガン大学を調査対象とした。

2) 有識者による研究会の設置

本調査研究では、文献調査、国内における研究事例、海外事例調査等の情報を基に、有識者の知識・視点等を反映するため、学識経験者を中心とする4名から構成される研究会を設置し、適宜、助言を受けながら研究を進めた。

「統計データの補完推計に関する調査」  
研究会委員名簿

|      |        |                             |
|------|--------|-----------------------------|
| 座長   | 岩崎 学   | 成蹊大学理工学部教授                  |
|      | 川崎 茂   | 日本大学経済学部教授                  |
|      | 保田 時男  | 関西大学社会学部准教授                 |
| 講師参加 | 鳩貝 淳一郎 | 日本銀行調査統計局経済統計課企業統計グループ企画役補佐 |

なお、海外現地調査には、吉森雅代 大阪大学大学院基礎工学研究科博士課程3年/学術振興会特別研究員(DC2)が同行することで、学術面からのアドバイス・支援を受けた。

## 第1章 調査研究と結果の概要

### 研究会における検討経過

| 回   | 実施時期               | アジェンダ  |
|-----|--------------------|--|
| 第1回 | 2012年<br>10月下旬     | <ul style="list-style-type: none"> <li>• 本調査研究の目的・概要等について</li> <li>• 補完推計に関する情報提供のお願い<br/>(対象調査・訪問先機関等のご推薦・ご紹介のお願い)</li> </ul>   |
| 第2回 | 2012年<br>11月21日(水) | <ul style="list-style-type: none"> <li>• 本調査研究の目的及び進め方について</li> <li>• プレ調査(文献調査)及び国内ヒアリングに基づく論点の整理</li> <li>• 補完推計に関するディスカッション               <ul style="list-style-type: none"> <li>✓ 補完推計の方法</li> <li>✓ 実務的な応用と課題</li> </ul> </li> <li>• 海外現地調査について               <ul style="list-style-type: none"> <li>✓ 訪問先機関について(委員からのご推薦・ご紹介の確認)</li> <li>✓ ヒアリング項目について(ディスカッション)</li> </ul> </li> </ul> |
| 第3回 | 2013年<br>2月19日(火)  | <ul style="list-style-type: none"> <li>• 各種調査結果及び報告書案について               <ul style="list-style-type: none"> <li>✓ 国内における補完推計の研究事例</li> <li>✓ 海外における補完推計の適用事例(海外現地調結果報告)</li> </ul> </li> <li>• 統計調査に欠測値補完を適用する際の示唆・課題について(ディスカッション)</li> </ul>  |
| 第4回 | 2013年<br>3月上旬      | <ul style="list-style-type: none"> <li>• 報告書案について(助言・最終確認)</li> </ul>  |

注)会議形式での委員会開催に加え、郵送・メール及び電話による意見収集により実施した。

### 3) 国内調査・海外調査の視点

国内においては、補完推計に関する取組に研究段階のものが多ことを踏まえ、アメリカ・カナダにおける統計調査において、どのような補完推計が行われているかを中心に整理を行った。

なお、海外調査の訪問先及び主なヒアリング項目は次ページのとおりである。

## 海外現地調査訪問先リスト

| 国名   | 都市・日付                | 訪問先  | ヒアリング項目   |   |
|------|----------------------|--|---|---|
| アメリカ | ワシントン DC<br>12/11(火) | 米国統計局<br>US Census Buerau                            | Dr. Roderick Little<br>Dr. Yves Thibaudeau  | <ul style="list-style-type: none"> <li>各種統計調査における補完状況について</li> <li>大災害時の補完推計</li> </ul>   |
|      | ワシントン DC<br>12/12(水) | 米国労働統計局<br>US Bureau of Labor<br>Statistics<br>(BLS) | Mr. Kenneth Robertson<br>Dr. Michael Horrigan<br>Mr. David Hiles<br>Mr. Thomas Nardone<br>Dr. John Eltinge<br>Mr. Larry Huff<br>Mr. Steven Paben<br>Ms. Deborah Sebastian Olsen | <ul style="list-style-type: none"> <li>各種統計調査における補完状況について</li> <li>大災害時(ハリケーン・カトリーナ)の補完推計について</li> </ul>                              |
| カナダ  | オタワ<br>12/13(木)      | カナダ統計局<br>Statistics Canada                          | Mr. Mike Sirois<br>Ms. Laurie Reedman<br>Mr. Yves DeGruire  | <ul style="list-style-type: none"> <li>補完プログラム BANFF について</li> </ul>  |
|      |                      |  | Mr. Edward Chen<br>Mr. Joseph Duggan<br>Ms. Cynthia Campeau   | <ul style="list-style-type: none"> <li>LFS(Labor Force Survey)</li> </ul>   |
|      |                      |  | Ms. Linda Ramsey  | <ul style="list-style-type: none"> <li>BCP(Business Continuity Plan)</li> </ul>   |
|      |                      |  | Dr. Jean-Francois Beaumont<br>Dr. Joel Bissonnette  | <ul style="list-style-type: none"> <li>SEVANI(System for Estimation of the Variance Due to Nonresponse and Imputation)について</li> </ul> |
| アメリカ | アナーバー<br>12/14(金)    | ミシガン大学<br>University of Michigan                     | Dr. Trivellore Raghunathan  | <ul style="list-style-type: none"> <li>補完プログラム IVEware について</li> </ul>  |
|      | ワシントン<br>12/17(月)    | westat 社   | Mr. Roger Tourangeau<br>Ms. Jane Li<br>Ms. Andrea Piesse  | <ul style="list-style-type: none"> <li>統計調査の実施機関から見た補完について</li> </ul>   |

## 第1章 調査研究と結果の概要

### 3. 調査結果の概要

#### 1) 国内における補完推計に関する研究

日本国内において、補完推計を実際の統計調査に適用している例はそれほど多くない。しかし将来的には回収率の低下等、調査環境が悪化していく可能性もあるため、それに備えるための手段の一つとして、補完推計の実務への適用可能性に注目が集まっている。こうした背景の下、公的調査や社会調査、ビジネスサーベイ等の各分野において、補完推計を実務に適用する場合、どのような手法が相応しいか、また、どのような課題があるのか等について、いくつかの研究が行われている。

本調査研究では、こうした国内における補完推計の研究事例について、文献調査及び聞き取り調査を実施した。その対象を選定する際に、統計調査の性質を捉えるために以下の4つの観点を用いた。

- ・ 実施機関 : 行政機関による調査か、それ以外による調査か
- ・ 調査対象 : 世帯調査か、事業所調査か
- ・ 実施頻度 : 高頻度調査か
- ・ 調査規模 : 大規模調査(全数調査)か

上記の観点に基づき、統計法で規定されている国内の56の基幹統計(平成24年11月現在)や、行政機関以外が実施している社会調査やビジネスサーベイの中から、調査の性質が偏ることのないように以下の5つの統計調査を選定した。

1. 経済センサスー活動調査  
行政機関が実施する、事業所・企業を対象とした大規模調査(全数調査)の例として選定。
2. 国民生活基礎調査  
行政機関が実施する、世帯を対象とした高頻度調査の例として選定。
3. 特定サービス産業実態調査  
行政機関が実施する、事業所・企業を対象とした高頻度調査の例として選定。
4. 全国企業短期経済観測調査  
行政機関以外が実施する、事業所・企業を対象とした調査の例として選定。
5. 日本人の国民性調査  
行政機関以外が実施する、個人を対象とした調査の例として選定。

### 2) 海外における補完推計に関する取組

本調査研究では、平常時における補完推計の実施・検討事例として、海外における統計調査においてどのような取組がなされているかの整理を行った。そのために、アメリカ及びカナダにおいて補完推計を実務に適用している調査機関を対象として、そこで用いられている手法の概要や、手法選定に至った経緯、欠測を防ぐ取組等について聞き取り調査を行った。

具体的には、米国統計局及び労働統計局、カナダ統計局の3つの統計機関における取組の整理に加えて、統計調査を受託する調査会社(westat 社)及びアカデミック分野(ミシガン大学)における取組や見解も整理することとした。

### 3) 災害時における補完推計などの対応

本調査研究では、災害などによって通常の調査実施が困難となる場合に、どのような対応がなされているのかについて整理を行った。国内の事例として、2011年3月に発生した東日本大震災に対する労働力調査における対応を整理した。海外の事例として、2005年8月に発生したハリケーン・カトリーナに対するACS(American Community Survey: アメリカ地域社会調査)、CES(Current Employment Survey: 雇用状況調査)及び、CPS(Current Population Survey: 就業状態調査)における対応を整理した。

なお、海外の事例を整理する中で、災害などに対する事前の備えの重要性が指摘されたため、カナダ統計局におけるBCP(Business Continuity Plan: 事業継続計画)や異常時の状況把握を迅速に行うための特殊モジュールの設定についても整理を行うこととした。

# 第1章 調査研究と結果の概要

(国内における補完推計に関する研究 概要一覧 1/2)

| 統計調査名 |          | 経済センサス活動調査  | 国民生活基礎調査  |
|-------|----------|---|---|
| 調査概要  | 実施機関     | 総務省   | 厚生労働省   |
|       | 調査対象     | 事業所・企業  | 世帯  |
|       | 実施頻度     | 平成24年に第1回調査を実施  | 3年毎(大規模調査)<br>中間年(簡易調査)   |
|       | 標本抽出方法   | 全数調査  | 層化2段集落抽出法   |
|       | サンプル数    | 約5,800,000事業所   | 世帯票・健康票:約277,000世帯<br>介護票:約6,000人<br>所得票・貯蓄票:約50,000世帯  |
|       | 調査方法     | 留置調査<br>郵送調査<br>インターネット調査   | 留置調査  |
|       | 調査項目例    | 開設時期、事業の内容、売上(収入)金額、従業員数等   | 家計支出、医療保険の加入有無、年金の受給状況、健康状態等  |
| 研究概要  | 研究主体     | 統計センター情報技術部統計技術研究課<br>(伊藤孝之)<br>(高橋将宜)  | 「国民生活基礎調査の標本設計・推定手法等に関する研究会」<br>・岩崎学(成蹊大学)<br>・石井太(社会保障・人口問題研究所)<br>・稲葉由之(慶應大学)<br>・西郷浩(早稲田大学)<br>・樋田勉(群馬大学)                          |
|       | 研究時期     | 平成24年9月11日(2012年度統計関連学会連合大会報告)等で発表  | 平成22年4月～平成23年3月(研究会)  |
|       | 研究の経緯    | ・平成24年2月の経済センサスー活動調査におけるデータエディティング方法の研究<br>・理論研究が進みつつある多重代入法の精度について、EDINETの実データを用いて単一代入法と比較、検討                        | ・統計委員会答申において「近年、非標本誤差を解析し、集計値を補正する理論の研究が進んできていることから、それらの利用可能性に関する検討」を行う必要があると指摘<br>・所得票の非回答により発生する非標本誤差を解析し、所得に関する集計値を補正する理論の利用可能性を検討 |
|       | 研究手法の概要  | ・売上高が欠測したデータセットを、複数の欠測パターンと欠測率で生成<br>・事業従事者数から売上高を推計する単回帰モデルを作成し、EMBアルゴリズムによる多重代入を実施<br>・単一代入法と多重代入法とで、補完値と真値の乖離具合を比較 | ・傾向スコアを所得票回答確率とみなし、調査対象世帯の傾向スコアを推定<br>・大きく分けて下記の3手法により総所得金額を推計<br>－IPW推定<br>－マッチングによる直接代入<br>－マッチングによる回帰代入                            |
|       | 今後の課題・方針 | ・概して単一代入法よりも多重代入法の方が精度が高いが、いつでも多重代入法が優位というわけではない<br>・様々なプログラムとの比較などが今後の課題   | ・真の所得分布が未知であるため補正結果の妥当性は検証できず、いずれの補正手法も集計に適用可能であると結論づけられないため、公的統計の結果として扱うには引き続き検討が必要  |

## (国内における補完推計に関する研究 概要一覧 2/2)

| 特定サービス産業実態調査  | 全国企業短期経済観測調査  | 日本人の国民性調査   |
|---|---|---|
| 経済産業省   | 日本銀行  | 統計数理研究所   |
| 事業所・企業  | 企業  | 個人  |
| 毎年  | 四半期毎  | 5年毎   |
| 層化抽出法(一部業種は全数調査)  | 層化抽出法   | 層化2段抽出法   |
| 約55,000事業所  | 約11,000社  | 約6,500人<br>(面接調査:3,000人、<br>留置調査:3,500人)  |
| 留置調査<br>郵送調査  | 郵送調査<br>インターネット調査   | 面接調査<br>留置調査  |
| 事業形態、売上高、従業者数等  | 判断項目、売上高、営業利益、減価償却費、人件費、負債、等  | 個人的態度、信条、人間関係、ジェンダー観、社会問題への態度、政治的<br>態度、人種観、等   |
| 「特定サービス産業実態調査等における推計手法の確立に関する調査研究 研究会」<br>・美添泰人(青山学院大学)<br>・西郷浩(早稲田大学)<br>・土屋隆裕(統計数理研究所)<br>・山本涉(電気通信大学)                  | 日本銀行調査統計局<br>・平川貴大<br>・鳩貝淳一郎  | 統計数理研究所<br>・土屋隆裕<br>・中村隆  |
| 平成21年12月～平成22年3月<br>(研究会)   | 平成23年11月(欧州委員会・OECD共済 第5回サーベイ調査に関するワークショップ報告)、平成24年8月(日本銀行ワーキングペーパー)等   | 平成21年9月30日(論文)等   |
| ・工業統計、特定サービス産業実態調査、企業活動基本調査について、回収率が100%に至らないなかで産業の実態を示すための検討が課題化<br>・各統計手法の特性に応じた欠測値の推計手法を検討                             | ・リーマンショックのように経済環境が大きく変動する局面においても精度を維持できる補完方法の検討が課題化<br>・短観における現行の補完方法とその代替方法について、統計精度を比較                                  | ・回収率の低下(平成20年調査で51.6%)を背景に、調査不能によるバイアスの補正が課題化<br>・バイアスを補正する際のウェイト調整において、調査不能の二大理由(不在・拒否)を区別する手法を検討  |
| ・実査における適用を想定して、一般的な下記手法を候補として検討<br>－横置き補完<br>－伸び率補完<br>－平均値補完<br>－平均値補完＋比補完<br>・売上高の欠測を対象にシミュレーションを実施                     | ・代表的な欠測値補完方法として、下記方法を比較・検討<br>－横置き補完(現行)<br>－伸び率補完<br>－項目間比率補完<br>・主要3項目(設備投資額、売上高、経常利益)を対象に、2004～2010年のデータを基にシミュレーションを実施 | ・ウェイトのキャリブレーションを各不能段階に応じて二段階で実施<br>・その際の補助変数として下記を使用<br>－人口統計学的な属性変数(国勢調査項目)<br>－調査への指向性変数(協力度や調査員の訪問回数等)   |
| ・平成21年調査において調査対象を拡大するため、前年データが利用できない業種が存在することを鑑み、当面の間は平均値補完を採用する<br>・全ての業種について前年データが利用できるようになった段階で、横置き補完、伸び率補完の採用を改めて検討する | ・主要3項目において、現行の補完方法より精度が高いかほぼ同等であるような代替方法が存在する<br>・ただし、短観への適用を検討する場合には、具体的な適用範囲について慎重に考慮する必要がある                            | ・真値が未知であるため手法の優劣は決定できないが、真値がどのような傾向にあるのかはある程度推測できる<br>・国勢調査が実施されない年は、他調査で代替しなければならない<br>・調査費や調査員への負担が大きい<br>・手法の専門性が高い<br>・1,000～2,000程度のサンプル数が必要 |

# 第1章 調査研究と結果の概要

## (海外における補完推計に関する取組 概要一覧 1/2)

| 国         | アメリカ   | アメリカ  |
|-----------|--|---|
| 統計調査名(英語) | Population Census  | American Community Survey (ACS)   |
| (日本語)     | 国勢調査   | アメリカ地域社会調査  |
| 調査概要      | 実施主体   | US Census Bureau  |
|           | 調査対象   | 世帯  |
|           | 実施頻度   | 10年毎  |
|           | 実施時期   | ・2010年調査では、2010年4月1日を基準日とし、2010年3月15日～17日にかけて調査票を発送   |
|           | 標本抽出方法   | 全数調査<br>・アメリカの全ての世帯、及びその構成員が対象  |
|           | サンプル数・回収数  | 1億3,170万4,730住戸への郵送   |
|           | 調査手法   | ・紙調査票による郵送自記調査<br>・回答は、郵送による返送に加えて、調査員の訪問によるフォローが行われる   |
| 取組概要      | 概要・目的  | ・紙調査票による郵送自記調査<br>・回答は、郵送による返送に加えて、電話、調査員の訪問、ウェブでの回答が可能となっている   |
|           | 欠測の状況  | ・10年ごとに実施される国勢調査の間を埋める形で、各自治体における政策・予算配分に関する意思決定を行うために利用される<br>・2000年国勢調査で用いられた、回答項目数の多い“Long Form”調査票を、2010年国勢調査では廃止し、ACSに移行 |
|           | 適用している手法   | ・項目非回答については、以下の3つを組み合わせた対応を行っている<br>－アサインメント (Assignment)<br>－アロケーション (Allocation)<br>－代替 (Substitution)                      |
| 経緯        | ・項目非回答については、以下の2つを組み合わせた対応を行っている<br>－アサインメント(assignment)<br>－アロケーション(allocation)   |   |
|           | ・10年ごとに実施される国勢調査<br>・補完前データ(CUF: Census Unedited File)と補完後データ(CEF: Census Edited File)を作成<br>・CEFでは、各質問について欠測・代入の状況を記録するフラグを作成している |   |

第1章 調査研究と結果の概要

(海外における補完推計に関する取組 概要一覧 2/2)

| 国         | アメリカ                                     | カナダ   |  |
|-----------|--|---|--|
| 統計調査名(英語) | Occupational Employment Statistics (OES) | Labour Force Survey (LFS)   |  |
| (日本語)     | 職業雇用統計                                   | 労働力調査   |  |
| 調査概要      | 実施主体                                     | US Bureau of Labor Statistics   | Statistics Canada  |
|           | 調査対象                                     | 事業所・企業  | 世帯・個人  |
|           | 実施頻度                                     | 半年毎(5月・11月)   | 毎月   |
|           | 実施時期                                     | ・毎年5月と11月に実施  | ・毎月15日を含む週を基準週として、翌週に実査を行う<br>・データ収集後、13日後には結果が公表されており、月次の統計として迅速な情報提供が行われている  |
|           | 標本抽出方法                                   | 層化抽出<br>・アメリカに所在する約670万事業所から、州・産業及び所有形態に応じて抽出された約20万事業所が対象<br>・年に2回、約20万事業所を抽出し、3年間のパネルを構成  | 層化抽出<br>・AR(Address Register: 住所登録)データと呼ばれる住所データに基づいて住戸を抽出<br>・抽出された住戸に居住する15歳以上の個人が対象<br>・全国(10州に加えて3準州を含む)<br>・6か月間にわたって回答を行うパネル調査(毎月新しいパネルと入れ替えるため、1/6の回答者は新しい対象者となる)<br>・約5万4千世帯、約10万個人からの回答を得ている  |
|           | サンプル数・回収数                                | 回答率77.3%<br>(1,110,296事業所のうち、858,474事業所が回答)<br>(2011年5月調査)  | 回答率95.1%<br>(2005年平均)  |
|           | 調査手法                                     | ・郵送による自記調査が主<br>・電話、電子メール、調査員による訪問での回答も行われている   | ・CATI及び調査員による訪問面接(ノートPCを用いたCATI)<br>・電話番号が不明の場合には、初回は訪問した上で、残る5か月間についてはCATIを用いる  |
| 概要・目的     | ・12の賃金区分に基づく各事業所における従業員数を回答              | ・雇用・非雇用状況の把握を行う<br>・生産年齢にあたる人口を「雇用状態」「非雇用状態」「非労働力人口」の3つに区分することで、労働力市場の状態を把握<br>・主な質問項目は以下のとおり<br>－雇用・非雇用状況<br>－労働時間<br>－産業<br>－職業<br>－賃金 など |  |
| 取組概要      | 欠測の状況                                    | ・対象事業所は3年間のパネルを構成するため、初回のコンタクト時の協力依頼を慎重に行うことで協力率を高める  | ・CATI及びCAPIによる調査票を利用することで、データ品質の向上を図る<br>・ただし、調査効率の観点から、そうした論理矛盾等をスキップすることも可能となっているため、補完が必要な欠測が生じる<br>・修正が発生した箇所については、調査員に対する周知が行われることで同様の修正が発生しないように注意を喚起する   |
|           | 適用している手法                                 | ・ホット・デック法による代入を行う<br>・パネル調査の特性上、過去の回答結果に基づいた代入も行われる<br>・ホット・デック法では、地域、産業、従業員数を基準にし、一部産業(病院、教育機関)では所有形態も基準にされている                             | ・項目非回答については、以下の3つの手法を用いている<br>－ホット・デック法による代入<br>－Carry-Forwardによる代入(前月からの据置き(横置き))<br>－推定による代入<br>・ユニット非回答については、回答履歴の状況に応じて、ホット・デック法あるいはウェイト付けによる補完を行っている<br>・回答者が6か月にわたって回答するパネル調査であるため、前月の回答を含め、ホット・デックのドナーとする、“Longitudinal Hot-deck”と呼ばれる手法を導入している |
|           | 経緯                                       | -   | ・調査項目の中でも把握の優先度を設定し、調査における重要指標に注目した補完を行っている<br>・最優先されるものは「就労状態」、次いで「収入」とされている<br>・2005年1月以前では、Carry-Forwardも行われていたものの、月次変化を過小に評価すること、また、クロスセクションでのホット・デック法についても過大に評価することから、現在のLongitudinal Hot-Deckを用いることとした   |

# 第1章 調査研究と結果の概要

(国内・海外における大規模災害への対応 概要一覧 1/2)

| 国        | 日本            | アメリカ  | アメリカ   |  |
|----------|---------------|---|--|--|
| 取組主体     | 総務省統計局        | US Bureau of Labor Statistics   | US Bureau of Labor Statistics  |  |
| 調査概要     | 統計調査名<br>(英語) | Current Population Survey<br>(CPS)  | Current Employment Survey<br>(CES)   |  |
|          | (日本語)         | 労働力調査   | 雇用状況調査   |  |
|          | 調査対象          | 世帯・個人   | 事業所・企業   |  |
|          | 実施頻度          | 毎月  | 毎月   |  |
| 災害への対応   | 災害名           | ハリケーン・カトリナ  | ハリケーン・カトリナ   |  |
|          | 発生時期          | 2011年3月   | 2005年9月  |  |
|          | 実査時の対応        | <ul style="list-style-type: none"> <li>・被災3県(岩手県、宮城県及び福島県)を除く全国結果を集計・公表した</li> <li>・岩手県及び宮城県は5月から、福島県は8月から調査を再開し、各県において調査地域を順次拡大した</li> <li>・2011年9月分結果から被災3県を含む全国結果の集計・公表を再開した</li> </ul>  | <ul style="list-style-type: none"> <li>・避難所やホテル、モーテルなどでの調査は行わなかったが、対象世帯に対しては、可能な限りコンタクトを取ることとした</li> <li>・避難地域に指定された地域についてはコンタクトは行わなかった</li> <li>・就業状態の判定基準に変更は行わなかった</li> </ul> | <ul style="list-style-type: none"> <li>・対象事業所に対しては、可能な限りコンタクトを取ることとした</li> <li>・通常よりも電話によるフォローアップに注力した</li> </ul>  |
|          | 適用した補完手法      | <ul style="list-style-type: none"> <li>・2011年3月～8月の各月を被災3県分について推計し、既公表の被災3県を除く全国の結果数値に加算することで、被災3県を含む全国結果を算出</li> <li>・就業者数は、他の複数のデータから求めた対前年同月減少率と同程度の就業者数の減少があったものと仮定して、就業者数の対前年同月減少数を推計</li> <li>・完全失業者は、概念的に近い職業安定所(ハローワーク)の「有効求職者数」をもとに作成した回帰式により推計</li> </ul> | <ul style="list-style-type: none"> <li>・回答を得ることができた世帯に対して、より重いウエイト付けを行うことで対応を行った</li> <li>・公式な見解として、アメリカ全体での推計に対しては、無視可能なレベルであるとされている</li> </ul>                                 | <ul style="list-style-type: none"> <li>・雇用の喪失を過小に推定しないよう、被害の影響が特に大きい地域に所在する事業所については、原則として事業は行われておらず、雇用数もゼロと見なした</li> <li>・2005年11月からは通常と同様の補完手法を行った</li> </ul> |
| 災害に対する備え | -             | -   | -  |  |

注) 労働力調査との比較のため本報告書の掲載順序と異なる

第1章 調査研究と結果の概要

(国内・海外における大規模災害への対応 概要一覧 2/2)

| 国        |               | アメリカ  | カナダ  | カナダ               |
|----------|---------------|---|--|-------------------|
| 取組主体     |               | US Census Bureau  | Statistics Canada  | Statistics Canada |
| 調査概要     | 統計調査名<br>(英語) | American Community Survey<br>(ACS)  | Labour Force Survey<br>(LFS)   | -                 |
|          | (日本語)         | アメリカ地域社会調査  | 労働力調査  | -                 |
|          | 調査対象          | 世帯・個人   | 世帯・個人  | -                 |
|          | 実施頻度          | 毎年 ※実査は毎月実施   | 毎月   | -                 |
| 災害への対応   | 災害名           | ハリケーン・カトリーナ   | -  | -                 |
|          | 発生時期          | 2005年9月   | -  | -                 |
|          | 実査時の対応        | <ul style="list-style-type: none"> <li>被害の影響が特に大きい地域では、調査票の郵送を取りやめた</li> <li>通常行われる、電話によるフォローアップも一部地域では行わなかった</li> <li>2005年11月以降は通常と同様の調査を行った</li> </ul>   | -  | -                 |
|          | 適用した補完手法      | <ul style="list-style-type: none"> <li>ユニット非回答、及び項目非回答の発生率を把握した上で、通常時に行っているユニット非回答に対するウエイト付けによる対応方法を変更</li> <li>避難地域に指定された地域に所在する住所を全てウエイト付けの対象から除外<br/>(回答を得ることができたケースに対して、より重いウエイトを付ける)</li> </ul>  | -  | -                 |
| 災害に対する備え | -             | <ul style="list-style-type: none"> <li>異常時の状況把握を迅速に行うために、2つの特殊モジュールを調査票に組み込む</li> <li>1. 追加項目 総選挙での投票行動やその理由など、通常は起こらない特殊なイベントの影響を把握するもの</li> <li>2. 災害による影響項目 事故や災害などの異常を想定したモジュール</li> <li>CAPIあるいはCATIによる調査票として、画面に表示されるようにしている</li> </ul> | BCP(Business Continuity Plan)について <ul style="list-style-type: none"> <li>統計調査ごとのBCPを策定中</li> <li>災害が発生した場合に備えて、どのような影響が生じるかを事前に分析する</li> <li>カレンダー形式とすることで、複数の統計調査を横断的に管理できるように工夫を行っている</li> </ul> |                   |

### 第2章 調査における欠測の分類と対応

本調査研究においては、統計データの補完推計について調査を行うため、調査において発生する欠測の分類と対応について、まず整理を行うこととした。ここでは、欠測を含むデータを特徴づける要素として「欠測が生じる状況」、「欠測パターン」及び「欠測メカニズム」について整理を行い、次に欠測を含むデータに対する対応方法として、欠測を欠測として扱う「欠測を含むデータの取扱」、及び代入を伴う「欠測値の補完」について整理を行った。

#### 1. 欠測が生じる状況

統計調査においては何らかの理由によって、データを得ることができない場合がある。欠測を含むデータが生じる状況には、主に以下の4つが挙げられる。

##### 1) 脱落(Drop out)

脱落とは、同一の対象者に対して長期にわたる調査を行う場合に、ある時点のデータが欠測する状況のことである。パネル調査等、一定の期間を置いて調査を繰り返す際に、特定の調査回以降の回答が得られない場合が該当する。

##### 2) 無回答

無回答とは、調査全体に対する拒否(ユニット非回答)や、一部質問に対する回答忘れ・拒否(項目非回答)等によって欠測が生じる状況のことである。

##### 3) 打ち切り(Censoring)

打ち切りとは、ある変数が決められた範囲外の値を取ることに伴って、関連するデータが欠測する状況のことである。例えば、一定の調査期間が経過したところで調査を打ち切った場合、その時点以降のデータは欠測となる。また、ある一定の所得以下の世帯については一部の質問項目を調査していない、というような場合も打ち切りに該当する。

##### 4) 切断(Truncation)

切断とは、上記の打ち切りに類似しているが、欠測となったデータの件数も欠測している状況のことである。例えば、特定の従業員規模以下の事業所については調査対象から除外したことで、そうした事業所の数も質問項目の回答もわからない、といった場合が該当する。

上記4つの状況のうち、本調査研究においては、1)脱落及び2)無回答について考察を行っている。さらに2)無回答については、統計調査の品質に大きく影響することからも、様々な対応がなされており、以降の第3章・第4章でもそうした対応について整理を行っている。

## 2. 欠測パターン

欠測を含むデータについて、欠測の生じ方を示すものが欠測パターンである。大きく分けると、全ての項目についてデータが欠測する「ユニット非回答」と、一部項目についてデータが欠測する「項目非回答」の2つがある。欠測パターンに応じた分析を行う必要がある。

### 1) ユニット非回答 (Unit Non Response)

ユニット非回答とは、調査対象者の不在や拒否、転居、対象把握誤り等の理由で、全ての項目のデータが欠測するものである。全ての項目が欠測するため、代わりとなる回答者を追加サンプリングすること、事前に多めにサンプリングすることや、ウエイトの調整等の対応が取られることもある。

### 2) 項目非回答 (Item Non Response)

項目非回答とは、記入漏れや誤記入等の理由で、一部の変数(質問)に対するデータが欠測するものである。項目非回答は、さらに欠測パターンが単調か、単調でないか、によって分類される。

#### (1) 単調な欠測パターン

単調な欠測パターンとは、欠測を含むデータの変数と回答者(ケース)を以下のように入れ替えられるものである。

図表 単調な欠測パターンのイメージ(○:回答、×:欠測)

| 変数<br>回答者 | ① | ② | ③ | ④ | ⑤ | ⑥ | ... |
|-----------|---|---|---|---|---|---|-----|
| A         | ○ | ○ | ○ | × | ○ | ○ |     |
| B         | × | ○ | × | × | ○ | × |     |
| C         | ○ | ○ | ○ | ○ | ○ | ○ |     |
| D         | × | ○ | × | × | ○ | × |     |
| E         | ○ | ○ | ○ | ○ | ○ | ○ |     |
| F         | × | × | × | × | ○ | × |     |
| ⋮         |   |   |   |   |   |   |     |

変数の並び順を入れ替え  
回答者の並び順を入れ替え

| 変数<br>回答者 | ⑤ | ② | ① | ⑥ | ③ | ④ | ... |
|-----------|---|---|---|---|---|---|-----|
| C         | ○ | ○ | ○ | ○ | ○ | ○ |     |
| E         | ○ | ○ | ○ | ○ | ○ | ○ |     |
| A         | ○ | ○ | ○ | ○ | ○ | × |     |
| B         | ○ | ○ | × | × | × | × |     |
| D         | ○ | ○ | × | × | × | × |     |
| F         | ○ | × | × | × | × | × |     |
| ⋮         |   |   |   |   |   |   |     |

上記のようにデータを表現したものは「データ行列」と呼ばれている。単調な欠測パターンとは、データ行列において、「ある回答者の変数が回答されていれば、上図のデータ行列の左あるいは上に位置する変数も回答されている状態」と捉えることが可能である。このパターンの場合、欠測値補完の計算が容易になる。

#### (2) 単調でない欠測パターン

単調でない欠測パターンとは、上記の単調な欠測パターンが成立しない状態である。このパターンの場合、欠測値補完にあたって反復計算が必要となり計算が煩雑になる。

## 第2章 調査における欠測の分類と対応

### 3. 欠測メカニズム

欠測メカニズムとは、どのように欠測が発生しているのか、を示すものである。欠測値を含むデータを分析する際にバイアスが生じるか、また、どのようなバイアスが生じるかについて、欠測メカニズムを理解することが重要である。ここでは、欠測について検討する際に頻繁に用いられている3つの欠測メカニズム<sup>1</sup>を取り上げる。

#### 1) MCAR (Missing Completely At Random: 完全にランダムな欠測)

MCARとは、ある値が欠測する確率が、その変数を含め他の変数に依存していない状態である。ある質問に対する回答を忘れる、データ集計の際の手違いによってデータの一部がランダムに欠測する、といった場合にMCARとなる。

項目非回答者と項目回答者とが類似している(欠測の有無によらずに両者が類似している)、という意味において、MCARは最も単純な欠測メカニズムである。しかしながら、現実の調査において、MCARが成立することは少ないと考えられる。

#### 2) MAR (Missing At Random: ランダムな欠測)

MARとは、ある値が欠測する確率が、その変数以外の他の変数に依存している状態である。例えば、以下のような場合にMARとなる。

(例1)

- ・それぞれの年代層(例えば高年代層と若年層)の中ではある値が欠測する確率が、その変数を含め他の変数に依存していないMCARの状態にある
- ・しかし、欠測メカニズムが高年代層と若年層の回答者の間では異なっている

(例2)

- ・大規模企業及び小規模企業のグループの中では、MCARの状態にある
- ・しかし、大規模企業と小規模企業のグループ間では異なる欠測メカニズムとなっている

MARはMCARに比べてより複雑な状態であるが、MCARとなる適切なグループ分けを行うことができれば、そのグループ内ではMCARを前提として欠測値を補完することも可能である。実務上では、欠測メカニズムはMARと仮定されることが多い。

#### 3) NMAR (Not Missing At Random: ランダムでない欠測)

NMARとは、ある値が欠測する確率が、その変数に依存している状態である。収入が高い人ほど収入を回答しない、売上が欠測するか否かが売上の多少に依存している、といった場合にNMARとなる。

NMARは最も複雑な欠測メカニズムであり、NMARの状態にあるデータにおいては、観測された値のみでは補完を行うことができないため、欠測している変数についての欠測メカニズムに何らかの仮定を置かなければならない。

<sup>1</sup> Rubin, D.B. (1987), Multiple Imputations for Non-Response in Surveys, John Wiley & Sons, New York.

**(参考) Ignorable (無視可能)と Nonignorable (無視不可能)**

MCAR、MAR、NMAR と類似した、欠測メカニズムの分類に「Ignorable (無視可能)」と「Nonignorable (無視不可能)」がある。この分類は、ある値が欠測する確率はその変数に依存するか否かに基づいたものである。

## ① Ignorable (無視可能)

Ignorable (無視可能)とは、欠測メカニズムが MCAR あるいは MAR の状態を指す。この分類では、欠測する確率はその変数には依存しないため、適切な手法を用いることで対応が可能であることから、「無視可能」と呼称されている。

## ② Nonignorable (無視不可能)

Nonignorable (無視不可能)とは、欠測メカニズムが NMAR の状態を指す。この分類では、欠測が生じていることによる影響を分析した上で欠測への対応を行うことが求められる。

**4. 欠測を含むデータの取扱**

欠測値を含むデータを分析する際に当たっての考え方について整理を行う。大きく分けると、代入等の補完を行わずに、欠測を欠測のままとして扱う考え方と、欠測値に値を代入することで完全なデータとして扱う考え方の2つに分かれる。ここでは、まず前者の考え方について整理を行う。

**1) 完全ケースに基づく分析 (Complete Case Analysis)**

完全ケースに基づく分析とは、一部の変数でも欠測を含むケース(回答者)は全て取り除いて分析を行うことである。最も簡単な方法ではあるものの、欠測を含むケースが多い場合には適さない、欠測メカニズムが MCAR 以外であると結果に影響が生じる、といった欠点がある。

**2) 利用可能なケースに基づく分析 (Available Case Analysis)**

利用可能なケースに基づく分析とは、ある変数について分析する場合に、その変数が得られているケース全てを対象として分析を行うことである。欠測を欠測のままとして扱う考え方としては、利用できる情報は可能な限り用いることができるものの、完全ケースに基づく分析と同様に、欠測メカニズムが MCAR 以外であると結果に影響が生じる。

## 第2章 調査における欠測の分類と対応

### 5. 欠測値の補完

次に欠測値を含むデータについて、その欠測値を補完、特に代入を行うことで対応する手法について概観する。

データには「量的データ」(間隔尺度、比率尺度)、「質的データ」(名義尺度、順序尺度)といった種類や、複数年データ、パネルデータ等の時系列データといった調査形態による差異も存在する。通常、欠測値の補完では、上記のようなデータ種類や調査形態に応じた欠測値補完がなされるため、複数の補完手法を用いることも多いが、補完手法における概念は下記の2つに大きく分けることが可能である。

1つ目は、欠測値を1つの値で代入する”Single Imputation”(単一代入法)である。これは、データ内の欠測値を1つの値で代入し、完全データセットとすることを指すものである。

2つ目は、複数の値を代入する”Multiple Imputation”(多重代入法)である。こちらは、1つの欠測値に対して複数の値をシミュレーションするものである。

一般的に、単一代入法では推定値の分散を過小評価する傾向が強く、全体分布が歪められてしまう可能性があることが指摘されるが、多重代入法では複数の完全データセットを作成することでその課題への対応が図られている。

まず、単一代入法において用いられる具体的な補完手法について概観する。

#### 1) 単一代入法

##### (1) 平均値による補完

回答を得られているケースの平均値を、欠測値の部分に代入するもの。平均値には通常、算術平均が用いられるが、中央値や最頻値を用いることも可能である。年収や売上高、消費金額等の数量値に対して行われることが多い。

|     |  |
|-----|--|
| 長所: | - 補完作業が容易である<br>- 集計値(平均値)に影響を及ぼさない      |
| 短所: | - 欠測値を含むケースが多い場合には不適切<br>- 回答者の属性が反映されない |

平均値を用いた補完における長所として、補完作業が容易という点が挙げられる。算術平均値の算出は、作業自体も容易で、再現性も備えている。加えて、平均値算出のベースが全体なのか、特定のセグメント(20代や30代といった年齢セグメント、関東地方や近畿地方といった地域セグメント等)なのか、といった条件に影響されるものの、集計値として広く用いられる平均値の集計には影響を及ぼさない、という点も長所として挙げられる。

一方で、短所としては、欠測値を含むケースが多く、平均値の算出が少数の回答ベースとなることで、平均値自体の信頼性が劣る可能性が存在する点が挙げられる。加えて、補完しようとする欠測値が回答者の属性に依存する場合(例えば、年収は回答者の就業形態や業種に影響されることが一般的である)には、そうした属性の影響を加味しない形で補完が行われるということがある。

なお、極端に小さい又は大きい値(外れ値)の影響を除くために、これらを除外した平均を使用する場合もある(刈り込み平均)。

**(2) Hot Deck**

回答を得られているケースから、背景データが類似したケース(ドナー)を探し出し、ドナーの値を欠測値の代わりとして代入するもの。数量値、カテゴリカルデータのいずれにおいても適用することが可能である。

- |     |  |
|-----|--|
| 長所: | - 回答者の属性等を反映することができる<br>- 欠測値の発生要因に依存しない |
| 短所: | - 補完作業が煩雑となる<br>- 欠測値の生じるケースが多い場合には不適切   |

Hot Deck 法の長所としては、属性等が類似したケースを探し出すため、就業形態に応じた年収が代入値として得られる等、属性を反映させた結果となる(納得感のある結果となる)ことが挙げられる。加えて、欠測値の発生要因が MCAR (Missing Completely At Random: 欠測は完全にランダムに発生)、MAR (Missing At Random: 欠測はランダムに発生)、NMAR (Not Missing At Random: 欠測はランダムでなく発生)のいずれの場合であっても適用可能である。

一方で、類似したケースを探し出すための作業が必要となるため、補完作業は煩雑となってしまう。特に、どのような定義をもって「類似している」と判断するかを慎重に決定する必要があるため、試行錯誤の繰り返しが発生することも考えられる。加えて、類似ケースを探し出すためのベースとなる「欠測値を含まないケース」が少ない場合には、結果に対する信頼性も劣る可能性がある。

なお、回答者の属性等を反映することができる Hot Deck 法の概念を基本として、類似したケースの平均値を代入する等、前述の手法との組合せも可能である。

**(3) Cold Deck**

背景データが類似したケース(ドナー)を他のデータセットから探し出し、そのケースの値を欠測値の部分に代入するもの。前述の Hot Deck では同一データセット内から類似ケースを探し出したが、Cold Deck では他のデータセットから探し出す点が異なる。例えば、企業売上高を調査ではなく、財務諸表等の情報から補完することも Cold Deck の 1 つと見なせるであろう。

- |     |  |
|-----|--|
| 長所: | - 回答者の属性等を反映することができる                   |
| 短所: | - 補完作業が煩雑となる<br>- 他のデータセットの利用可能性に制約される |

Cold Deck 法では、Hot Deck 法同様に、属性等が類似したケースを探し出すため、属性を反映させた結果を得られることが多いことが長所として挙げられる。

一方で、類似したケースを特定するために、補完作業は煩雑となってしまう。また、他のデータセットが利用できない場合には Cold Deck 自体を行うことができない。

## 第2章 調査における欠測の分類と対応

### (4) 回答データの据置き(横置き)による補完

据置き補完(横置き補完)(Carry Forward)とは、特定のケースについて、過去の回答結果をそのまま代入するもの。同一客体を継続的に追跡する調査(企業統計調査、パネル調査等の時系列データ)における適用が想定されるものである。

|     |   |
|-----|---|
| 長所: | - 補完作業が比較的容易である   |
| 短所: | - 外部環境変化(景気動向等)やケース属性の変化(従業員数増減等)が反映されない<br>- 長期にわたって回答していないケースには適用できない |

同一客体を継続的に追跡する調査においては、据置き補完(横置き補完)の作業は比較的容易である点が長所として挙げられる。

短所としては、景気動向による影響や従業員数増減等、ケース属性の変化が反映されない点が挙げられる。特に、季節変動が存在する場合やリーマンショックのような大きな変化が生じた場合には、その影響を加味することができない。また、少なくとも前回調査における回答は必須であり、前回からさらに遡って回答が欠測する場合には適用できない(代入するもとのデータが古いデータとなる)点も短所として挙げられる。

なお、据置き補完(横置き補完)を行う場合には、今回調査と前回調査において同一の質問項目であることが前提となり、新たに追加した質問に対しては、据置き補完(横置き補完)を適用することはできない。

### (5) 伸び幅又は伸び率を用いた補完

回答を得られているケースを基に、前期からの変動幅又は変動率(伸び率)を算出し、特定のケースの過去の回答結果にその値を乗じた値を代入するもの。

|     |   |
|-----|---|
| 長所: | - 外部環境変化(景気動向等)を反映することができる<br>- 補完作業が比較的容易である             |
| 短所: | - ケース属性の変化(従業員数増減等)が反映されない<br>- 長期にわたって回答していないケースには適用できない |

伸び幅又は伸び率を用いた補完の長所としては、回答を得られているケースからの情報に制約されるものの、景気動向等の外部環境変化を一定程度反映することが可能な点である。

一方で、短所としては代入する値が必ずしもケース属性の変化を反映していない点が挙げられる。具体的には、年数が経過するに従って従業員数が増加したにも関わらず、その影響を加味しない形で代入が行われる例が考えられる。また、長期にわたって回答がなされていないケースでは、元となるデータが存在しないあるいは古いものであるために、信頼性が劣る危険性も短所として挙げられる。さらに、伸び率を用いる場合には、前期が0であると計算できないということがある。

**(6) 重回帰式モデルによる補完**

回答を得られているケースを基に、複数の要因を加味した重回帰式モデルを作成し、欠測値を予測するもの。

|     |   |
|-----|---|
| 長所: | - ケース属性や季節変動等複数の要因を加味することができる<br>- 決定係数、調整済み決定係数等によるモデルの当てはまりを評価できる |
| 短所: | - 重回帰式に投入する変数の選定が困難<br>- 再現性が低い<br>- 分散を過小評価する                      |

長所としては、重回帰式には複数の変数を投入することになるため、複数の要因を加味することができる点が挙げられる。また、作成された重回帰式モデルについても、決定係数や調整済み決定係数を基に、当てはまりを評価することができる点も長所として挙げられる。

一方で短所としては、重回帰式に投入する変数の選定が困難である点が挙げられる。また、多重共線性の問題や係数の符号(正負)の整合性を考慮しなければならない点も短所として挙げられる。こうした点は、いったん重回帰式モデルを作成しても、そのモデルが次回調査には当てはまらない可能性がある、という再現性の低下につながる。同時に、都度こうしたモデルを作成する必要が生じる等、補完作業が極めて煩雑となる。また、重回帰式モデルの特性上、分散を過小評価する危険性も含んでいる。

**2) 多重代入法**

多重代入法では、1つの欠測値に対して複数個の値を代入するが、一般的にそのステップは以下の3ステップである。

- ステップ①: 欠測値に対して複数個の値を代入し、複数個の完全データセットを作成する
- ステップ②: 複数個の完全データセットのそれぞれについて、集計・分析を行う
- ステップ③: ステップ②で得られた複数個の結果を統合する

多重代入法の基本的な考え方は、Rubin によって 1980 年代に提唱されたものである。複数の完全データセットを作成する際には、例えば重回帰式モデルによる補完結果に、ランダムに発生させた誤差を加えた結果を代入する等、単一代入法によるデータセットを複数個作成するものが例として挙げられる。

以下では、多重代入法の概念に沿った主な手法について概観する。多重代入法の長所としては、複数のデータセットを作成することで、欠測値が生じることの不確実性を考慮し、分散の過少推計を避けることができる点が挙げられる。短所としては、変数の種類(尺度変数、名義変数等)や結果分析の目的に応じて適切な多重代入を行わないと結果にバイアス(標準誤差の過少評価)が生じることがある点が挙げられる。

## 第2章 調査における欠測の分類と対応

### (1) EM アルゴリズムによる補完

EM アルゴリズムにより、最尤推定値を予測するもの。欠測値を含むデータの尤度を最大化することで、“もっともらしい”結果を得ようとするものである。E ステップ (estimation: 期待値ステップ) と M ステップ (maximization: 最大化ステップ) から構成される。

E ステップ : 尤度関数の期待値を求める

M ステップ : E ステップにおける尤度の期待値の最大化を図る

### (2) マルコフ連鎖モンテカルロ法

マルコフチェーン・モンテカルロ法 (MCMC) と呼ばれ、I ステップ (Imputation: 代入ステップ) と P ステップ (posterior: 事後ステップ) の 2 ステップから構成される。それぞれのステップを繰り返すことで、代入値を算出するもの。

I ステップ : 欠測値を含むデータセットに対する条件付き分布に基づき、完全データセットを作成

P ステップ : 得られた完全データセットから母集団情報を推定

#### (参考) 傾向スコア

類似したケース (ドナー) を探し出す際の基準のひとつとして用いることができるものに、各ケースの 傾向スコア (propensity score) がある。傾向スコアとは各ケースが回答する確率として表現されるものである。傾向スコアが同一のケースをドナーとして使用する等、ドナー選定の基準に用いられることがある。

基本的な考え方は、事後確率を求めるベイズ統計に沿ったものといえる。

#### (参考) ユニット非回答に対応する手法 キャリブレーション (Calibration)

キャリブレーション (Calibration) とは、回収標本から母集団を推定する際に、母集団の値を補助変量として用いることでウェイトを調整する手法である。ウェイトを調整するものであるため、項目非回答ではなくユニット非回答に対して用いられる手法である。「日本人の国民性調査」(実施: 統計数理研究所) における調査不能バイアスの補正に関する研究において用いられた (本報告書第3章)。

回収標本から母集団を推定する際のウェイト調整には、以下の 2 段階のプロセスがある。

- 回収標本から抽出標本の推定
- 抽出標本から母集団の推定

前者には回答確率 (回収率、傾向スコア、等) の逆数を用いた推定 (IPW 推定) が用いられ、後者にはキャリブレーションが使用される場合が多い。回答確率の逆数を用いた推定においては、回答確率の低い回答者のウェイトが大きくなるため、推定量は不偏であっても分散が過大評価される、という欠点がある。

キャリブレーション先 (母集団) の情報があらかじめ分かっている (「ある趣味をもった人に対する調査」のように、母集団に関する情報が未知の場合には使用できない) といった限界もあるものの、キャリブレーションは定まった母集団の値を使用することで、その分散を抑えることが可能である。

## 6. 調査における欠測への対応

調査における欠測については、海外(アメリカ・カナダ)においては主にセンサス(国勢調査)における非回答に関して、どのように扱うのか、という議論が行われてきた。カナダ統計局においては、1950年のセンサスまでは、欠測値に対してデミング法(Deming's method)と呼ばれる方法で補完を行っていた。これは、前回実施のセンサスにおける頻度分布を用いて、手作業で欠測値に代入を行うものである。さらに1953年には、アメリカセンサス局の統計家Morris H. Hansenらによって、1948年に実施された小売シェア調査(Survey of Retail Shares)の欠測の扱いについての議論が出版<sup>2</sup>された。先月の売上高・賃金等の情報に基づいて「補完(代入)」を行う、と統計調査における「補完」に初めて言及したものとされている。

その後、アメリカにおけるセンサスでは1960年に実施されたものが、カナダにおけるセンサスでは1961年に実施されたものが、コンピューターを利用した補完が行われた初めての調査とされている。補完に関する技術の進歩は、コンピューターの進歩とともにあったが、初期においてはホット・デック法が主な補完手法であった。

その後、様々なプログラムがアメリカセンサス局やカナダ統計局によって開発されてきた。

アメリカセンサス局

- ・ SPEER (Structured Programs for Economic Editing and Referrals)  
1984年に開発。経済データにおける比率の補完に用いられる。

カナダ統計局

- ・ CANEDIT  
1976年に開発。Fell センサスに利用される。
- ・ GEIS (Generalized Edit and Imputation System)  
※本報告書4章にて GEIS の後継プログラム Banff について記載
- ・ CANCEIS/NIM  
(Canadian Census Edit & Imputation System / Nearest-neighbor Imputation Methodology)  
最近隣法によるホット・デック法を行う。

本調査研究において調査を行った海外事例(アメリカ・カナダ)では、上記のような経緯から、ホット・デック法が多く用いられている。多くの補完手法が利用可能となった現在においても、ホット・デック法が主要な補完手法として用いられ続けている。

<sup>2</sup> Hansen, M.H., Hurwitz, W.N. and Madow, W.G. (1953). *Sample Survey Methods and Theory*, Vol. I and II. New York: John Wiley and Sons.

## 第2章 調査における欠測の分類と対応

### 第3章 国内における補完推計に関する研究

日本国内において、補完推計を実際の統計調査に適用している例はそれほど多くない。しかし将来的には回収率の低下等、調査環境が悪化していく可能性もあるため、それに備えるための手段の一つとして、補完推計の実務への適用可能性に注目が集まっている。こうした背景の下、公的調査や社会調査、ビジネスサーベイ等の各分野において、補完推計を実務に適用する場合、どのような手法が相応しいか、また、どのような課題があるのか等について、いくつかの研究が行われている。

本調査研究では、こうした国内における補完推計の研究事例について、文献調査及び聞き取り調査を実施した。その対象を選定する際に、統計調査の性質を捉えるために下記の4つの観点を用いた。

- ・ 実施機関 : 行政機関による調査か、それ以外による調査か
- ・ 調査対象 : 世帯調査か、事業所調査か
- ・ 実施頻度 : 高頻度調査か
- ・ 調査規模 : 大規模調査(全数調査)か

上記の観点に基づき、統計法で規定されている国内の56の基幹統計(平成24年11月現在)や、行政機関以外が実施している社会調査やビジネスサーベイの中から、調査の性質が偏ることのないように以下の5つの統計調査を選定した。

1. 経済センサスー活動調査  
行政機関が実施する、事業所・企業を対象とした大規模調査(全数調査)の例として選定。
2. 国民生活基礎調査  
行政機関が実施する、世帯を対象とした高頻度調査の例として選定。
3. 特定サービス産業実態調査  
行政機関が実施する、事業所・企業を対象とした高頻度調査の例として選定。
4. 全国企業短期経済観測調査  
行政機関以外が実施する、事業所・企業を対象とした調査の例として選定。
5. 日本人の国民性調査  
行政機関以外が実施する、個人を対象とした調査の例として選定。

図表1 各統計調査の性質

| # | 統計調査名        | 発行機関 |     | 調査対象  |        | 実施頻度 |     | 調査規模 |      |
|---|--------------|------|-----|-------|--------|------|-----|------|------|
|   |              | 行政機関 | その他 | 世帯・個人 | 事業所・企業 | 高頻度  | 低頻度 | 全数調査 | 標本調査 |
| 1 | 経済センサスー活動調査  | ○    |     |       | ○      |      | ○   | ○    |      |
| 2 | 国民生活基礎調査     | ○    |     | ○     |        | ○    |     |      | ○    |
| 3 | 特定サービス産業実態調査 | ○    |     |       | ○      | ○    |     |      | ○    |
| 4 | 全国企業短期経済観測調査 |      | ○   |       | ○      | ○    |     |      | ○    |
| 5 | 日本人の国民性調査    |      | ○   | ○     |        |      | ○   |      | ○    |

以下では、これらの5の統計調査に関して、補完推計の研究事例を紹介する。そのためにまず、各統計調査の性質及び研究の概要について一覧表として整理し(第I節)、次いで各研究事例の順に詳細に述べていく(第II節)。

### 第3章 国内における補完推計に関する研究

#### I 国内における補完推計に関する研究の概要(一覧)

| 統計調査名 |          | 経済センサス活動調査  | 国民生活基礎調査  |
|-------|----------|---|---|
| 調査概要  | 実施機関     | 総務省   | 厚生労働省   |
|       | 調査対象     | 事業所・企業  | 世帯  |
|       | 実施頻度     | 平成24年に第1回調査を実施  | 3年毎(大規模調査)<br>中間年(簡易調査)   |
|       | 標本抽出方法   | 全数調査  | 層化2段集落抽出法   |
|       | サンプル数    | 約5,800,000事業所   | 世帯票・健康票:約277,000世帯<br>介護票:約6,000人<br>所得票・貯蓄票:約50,000世帯  |
|       | 調査方法     | 留置調査<br>郵送調査<br>インターネット調査   | 留置調査  |
|       | 調査項目例    | 開設時期、事業の内容、売上(収入)金額、<br>従業者数等   | 家計支出、医療保険の加入有無、年金の受給状況、健康状態等  |
| 研究概要  | 研究主体     | 統計センター情報技術部統計技術研究課<br>(伊藤孝之)<br>(高橋将宜)  | 「国民生活基礎調査の標本設計・推定手法等に関する研究会」<br>・岩崎学(成蹊大学)<br>・石井太(社会保障・人口問題研究所)<br>・稲葉由之(慶應大学)<br>・西郷浩(早稲田大学)<br>・樋田勉(群馬大学)                          |
|       | 研究時期     | 平成24年9月11日(2012年度統計関連学会連合大会報告)等で発表  | 平成22年4月～平成23年3月<br>(研究会)  |
|       | 研究の経緯    | ・平成24年2月の経済センサスー活動調査におけるデータエディティング方法の研究<br>・理論研究が進みつつある多重代入法の精度について、EDINETの実データを用いて単一代入法と比較、検討                        | ・統計委員会答申において「近年、非標本誤差を解析し、集計値を補正する理論の研究が進んできていることから、それらの利用可能性に関する検討」を行う必要があると指摘<br>・所得票の非回答により発生する非標本誤差を解析し、所得に関する集計値を補正する理論の利用可能性を検討 |
|       | 研究手法の概要  | ・売上高が欠測したデータセットを、複数の欠測パターンと欠測率で生成<br>・事業従事者数から売上高を推計する単回帰モデルを作成し、EMBアルゴリズムによる多重代入を実施<br>・単一代入法と多重代入法とで、補完値と真値の乖離具合を比較 | ・傾向スコアを所得票回答確率とみなし、調査対象世帯の傾向スコアを推定<br>・大きく分けて下記の3手法により総所得金額を推計<br>ーIPW推定<br>ーマッチングによる直接代入<br>ーマッチングによる回帰代入                            |
|       | 今後の課題・方針 | ・概して単一代入法よりも多重代入法の方が精度が高いが、いつでも多重代入法が優位というわけではない<br>・様々なプログラムとの比較などが今後の課題   | ・真の所得分布が未知であるため補正結果の妥当性は検証できず、いずれの補正手法も集計に適用可能であると結論づけられないため、公的統計の結果として扱うには引き続き検討が必要  |

第3章 国内における補完推計に関する研究

| 特定サービス産業実態調査  | 全国企業短期経済観測調査  | 日本人の国民性調査   |
|---|---|---|
| 経済産業省   | 日本銀行  | 統計数理研究所   |
| 事業所・企業  | 企業  | 個人  |
| 毎年  | 四半期毎  | 5年毎   |
| 層化抽出法(一部業種は全数調査)  | 層化抽出法   | 層化2段階抽出法  |
| 約55,000事業所  | 約11,000社  | 約6,500人<br>(面接調査:3,000人、<br>留置調査:3,500人)  |
| 留置調査<br>郵送調査  | 郵送調査<br>インターネット調査   | 面接調査<br>留置調査  |
| 事業形態、売上高、従業者数等  | 判断項目、売上高、営業利益、減価償却費、人件費、負債、等  | 個人的態度、信条、人間関係、ジェンダー観、社会問題への態度、政治的<br>態度、人種観、等   |
| 「特定サービス産業実態調査等における推計手法の確立に関する調査研究 研究会」<br>・美添泰人(青山学院大学)<br>・西郷浩(早稲田大学)<br>・土屋隆裕(統計数理研究所)<br>・山本渉(電気通信大学)                  | 日本銀行調査統計局<br>・平川貴大<br>・鳩貝淳一郎  | 統計数理研究所<br>・土屋隆裕<br>・中村隆  |
| 平成21年12月～平成22年3月<br>(研究会)   | 平成23年11月(欧州委員会・OECD共済 第5回サーベイ調査に関するワークショップ報告)、平成24年8月(日本銀行ワーキングペーパー)等   | 平成21年9月30日(論文)等   |
| ・工業統計、特定サービス産業実態調査、企業活動基本調査について、回収率が100%に至らないなかで産業の実態を示すための検討が課題化<br>・各統計手法の特性に応じた欠測値の推計手法を検討                             | ・リーマンショックのように経済環境が大きく変動する局面においても精度を維持できる補完方法の検討が課題化<br>・短観における現行の補完方法とその代替方法について、統計精度を比較                                  | ・回収率の低下(平成20年調査で51.6%)を背景に、調査不能によるバイアスの補正が課題化<br>・バイアスを補正する際のウェイト調整において、調査不能の二大理由(不在・拒否)を区別する手法を検討  |
| ・実査における適用を想定して、一般的な下記手法を候補として検討<br>－横置き補完<br>－伸び率補完<br>－平均値補完<br>－平均値補完＋比補完<br>・売上高の欠測を対象にシミュレーションを実施                     | ・代表的な欠測値補完方法として、下記方法を比較・検討<br>－横置き補完(現行)<br>－伸び率補完<br>－項目間比率補完<br>・主要3項目(設備投資額、売上高、経常利益)を対象に、2004～2010年のデータを基にシミュレーションを実施 | ・ウェイトのキャリブレーションを各不能段階に応じて二段階で実施<br>・その際の補助変数として下記を使用<br>－人口統計学的な属性変数(国勢調査項目)<br>－調査への指向性変数(協力度や調査員の訪問回数等)   |
| ・平成21年調査において調査対象を拡大するため、前年データが利用できない業種が存在することを鑑み、当面の間は平均値補完を採用する<br>・全ての業種について前年データが利用できるようになった段階で、横置き補完、伸び率補完の採用を改めて検討する | ・主要3項目において、現行の補完方法より精度が高いかほぼ同等であるような代替方法が存在する<br>・ただし、短観への適用を検討する場合には、具体的な適用範囲について慎重に考慮する必要がある                            | ・真値が未知であるため手法の優劣は決定できないが、真値がどのような傾向にあるのかはある程度推測できる<br>・国勢調査が実施されない年は、他調査で代替しなければならない<br>・調査費や調査員への負担が大きい<br>・手法の専門性が高い<br>・1,000～2,000程度のサンプル数が必要 |

## II 国内における補完推計に関する研究事例

### 1. 経済センサスー活動調査

#### 1) 調査の概要

「経済財政運営と構造改革に関する基本方針 2005」(いわゆる「骨太の方針」)(平成 17 年 6 月 21 日閣議決定)において、「経済活動を同一時点で網羅的に把握する経済センサス」の実施が提言された。これに基づいて関係府省等において検討が行われ、経済に関連した大規模統計調査の統廃合、簡素・合理化を行って、平成 21 年 7 月に「経済センサスー基礎調査」が、平成 24 年 2 月に「経済センサスー活動調査」が実施された。両者の目的は以下のとおりである。活動調査では、売上高や費用等の経理項目も把握することとされている。

図表 2 経済センサスの目的

| 調査種別 | 目的  |
|------|---|
| 基礎調査 | <ul style="list-style-type: none"> <li>・ 全産業分野における事業所及び企業の基本的構造を、全国及び地域別に明らかにする</li> <li>・ 各種統計調査実施のための基礎資料を得る</li> </ul>       |
| 活動調査 | <ul style="list-style-type: none"> <li>・ 全産業分野における事業所及び企業の経済活動の実態を、全国及び地域別に明らかにする</li> <li>・ 各種統計調査の精度向上に資する母集団情報を得る</li> </ul> |

図表 3 活動調査の概要

| 項目   | 内容  |
|------|---|
| 調査対象 | 全ての事業所及び企業(全数調査)<br>※農林漁家に属する個人経営の事業所、家事サービス業、外国公務、国及び地方公共団体に属する事業所を除く  |
| 調査方法 | 単独事業所<br>企業<br>留置調査<br>(調査員が事業所に赴き、調査票への記入依頼、調査票等の配布・回収を行う)   |
|      | 複数事業所<br>企業<br>郵送調査・インターネット調査<br>(行政機関が調査票を直接郵送により配布し、郵送(紙・電子媒体)又はインターネットで回収を行う)  |
| 調査事項 | <ul style="list-style-type: none"> <li>・ 基本情報(名称、電話番号、所在地、開設時期、事業内容、従業者数、等)</li> <li>・ 経理情報(売上(収入)金額、費用、設備投資の有無、土地・建物の所有有無、等)</li> <li>・ その他(売場面積、営業時間、施設・店舗の形態、等)</li> </ul> |

#### 2) 研究の経緯・目的

経済センサスー活動調査における経理項目の欠測に対する分析の一環として、統計センターにおいて、データエディティング及び補定<sup>1</sup>方法の研究(特に多重代入法に焦点を当てた研究)が行われた。

多重代入法に関する理論的研究は多くなされている一方で、実データを用いた研究は少ない。研究では、EDINET(Electronic Disclosure for Investors' Network: 金融商品取引法に基づく有価証券報告書等の開示書類に関する電子開示システム)における実データを用いて、補定の精度について多重代入法と単一代入法とを多角的に比較、検討が行われている。

<sup>1</sup> 「補定」と「補完」とは同義であるが、本節では統計センター情報技術部統計技術研究課における用語法に準ずる。

### 3) 研究の体制・時期

統計センター情報技術部統計技術研究課によって研究が進められ、平成 24 年 9 月 11 日の「2012 年度統計関連学会連合大会」において、結果が報告されている。

### 4) 研究手法の概要

#### (1) 欠測データの生成

本研究で使用された EDINET データの対象は、平成 23 年 3 月 31 日に決算を迎える上場企業 3,587 社であり、提出日が平成 23 年 6 月 30 日からさかのぼり 1 年以内の企業である。これらのうち、産業大分類が「製造業」、「卸売業、小売業」、「建設業」、「情報通信業」及び「学術研究、専門・技術サービス業」であるものを対象に、売上高が欠測したデータを生成する。その際、産業大分類ごとに欠測パターンを 6 種類、欠測率を 3 種類に設定し、合計 18 種類の欠測データが生成された。

図表 4 対象となった EDINET データ

| 産業大分類 |                 | 完全データ数 |
|-------|-----------------|--------|
| E     | 製造業             | 1,222  |
| I     | 卸売業、小売業         | 571    |
| D     | 建設業             | 158    |
| G     | 情報通信業           | 276    |
| L     | 学術研究、専門・技術サービス業 | 191    |

図表 5 欠測パターンの設定

| # | 欠測対象           | カテゴリ |
|---|----------------|------|
| 1 | 事業従事者数が小さい値の企業 | MAR  |
| 2 | 事業従事者数が中の値の企業  |      |
| 3 | 事業従事者数が大きい値の企業 |      |
| 4 | 事業従事者数が大小の値の企業 |      |
| 5 | 系統抽出された企業      | MCAR |
| 6 | ランダムな企業        |      |

図表 6 欠測率の設定

| # | 欠測率 |
|---|-----|
| 1 | 30% |
| 2 | 40% |
| 3 | 50% |

#### (2) モデルの作成

こうして生成された欠測データに対して、単一代入法と多重代入法とによってそれぞれ補定値を代入し、両者の精度を比較する。その際に使用されたモデルは、単一代入法も多重代入法も共通であり、事業従事者数から売上高を推計する以下の 2 種類の単回帰モデルである。

- ・ 1 次多項式  $y = ax + b$
- ・ 自然対数変換  $\ln y = a \ln x + b$   
( $y$ : 売上高、 $x$ : 事業従事者数、 $a$ : 傾き、 $b$ : 切片)

説明変数としては、産業の種類を問わず欠測が少ないことと、売上高と相関があると想定できることから、事業従事者数が採用されている。なお、産業別に事業従事者数を説明変数とし、売上高を被説明変数としたモデルを作成した。

### 第3章 国内における補完推計に関する研究

#### (3) 多重代入法の設定

多重代入法では EMB アルゴリズムが採用されている。EMB アルゴリズムでは第 2 章でも述べた EM アルゴリズムの E ステップの前にブートストラップによるデータセット生成のステップがあり、マルコフ連鎖・モンテカルロ法に比べて計算負荷が小さいというメリットがある。データセット数が増えるほど相対効率が上がるが、データセット数が 20 付近から相対効率は逡減するため、本研究ではデータセット数を 20 に設定した。

また、実際の多重代入には「R」のパッケージである「Amelia II」を使用する。EMB アルゴリズムを実装している点、及び扱えるデータセット数が多い点を考慮して採用されている。

なお、多重代入法における欠測補定値は、各データセットの補定値の算術平均が用いられている。

図表 7 R パッケージ「Amelia II」の概要

| 項目  | 内容   |
|-----|--|
| 開発者 | Gary King 教授(ハーバード大学)を中心とするチーム   |
| 開発年 | 2011 年<br>※Amelia I (汎用多重代入法プログラム)は 2001 年   |
| 特徴  | <ul style="list-style-type: none"><li>・ EMB アルゴリズムを実装</li><li>・ 巨大データセットの多重代入に対応</li></ul> ※EMB アルゴリズムの限界データサイズ <ul style="list-style-type: none"><li>・ 観測値: 32,000 個</li><li>・ 変数: 240 個</li><li>・ 総パラメータ数: 29,160 個</li><li>・ 個別要素数: 4 億 2516 万 7380 個</li><li>・ 反転可能な最大分散共分散行列: <math>29,160 \times 29,160</math></li></ul> |

#### (4) 評価

単一代入法と多重代入法とによって売上高の欠測補定を産業別に行った後、補定精度を以下の 3 種類の方法で評価される。

- A) 真値の売上高総額と、補定済みの売上高総額との乖離具合を比較する(小さい方をカウント)
- B) 散布図を用いて視覚的に比較する
- C) 欠測値補定データの標準偏差を算出し、真値の標準偏差と比較する

A による評価の結果、産業種別、欠測パターン別、モデル種別、欠測率別に比較したところ、概して多重代入法の方が単一代入法よりも精度が高かった。また、B 及び C による評価の結果、多重代入法は復元に近づける結果であることも確認された。

しかし、全てにおいて多重代入法が優位であるわけではないことや、元々のサンプル数が小さい産業に関しては結果の信頼性に留保をつけなくてはならないことに注意しなければならない。

### 第3章 国内における補完推計に関する研究

図表 8 評価方法 A 結果(産業別)

単位: 試行回数(カッコ内は構成比)

| 産業                 | 多重代入法   | 単一代入法   | 合計       |
|--------------------|---------|---------|----------|
| E(製造業)             | 4(67%)  | 2(33%)  | 6(100%)  |
| I(卸売業、小売業)         | 2(50%)  | 2(50%)  | 4(100%)  |
| D(建設業)             | 1(10%)  | 9(90%)  | 10(100%) |
| G(情報通信業)           | 6(86%)  | 1(14%)  | 7(100%)  |
| L(学術研究、専門・技術サービス業) | 12(92%) | 1(8%)   | 13(100%) |
| 合計                 | 25(63%) | 15(38%) | 40(100%) |

図表 9 評価方法 A 結果(欠測パターン別)

単位: 試行回数(カッコ内は構成比)

| 欠測パターン | 多重代入法   | 単一代入法   | 合計       |
|--------|---------|---------|----------|
| 従業者:小  | 3(75%)  | 1(25%)  | 4(100%)  |
| 従業者:中  | 9(90%)  | 1(10%)  | 10(100%) |
| 従業者:大  | 4(67%)  | 2(33%)  | 6(100%)  |
| 従業者:大小 | 1(50%)  | 1(50%)  | 2(100%)  |
| 系統抽出   | 2(20%)  | 8(80%)  | 10(100%) |
| ランダム   | 6(75%)  | 2(25%)  | 8(100%)  |
| 合計     | 25(63%) | 15(38%) | 40(100%) |

図表 10 評価方法 A 結果(モデル別)

単位: 試行回数(カッコ内は構成比)

| モデル    | 多重代入法   | 単一代入法   | 合計       |
|--------|---------|---------|----------|
| 1次多項式  | 11(65%) | 6(35%)  | 17(100%) |
| 自然対数変換 | 14(61%) | 9(39%)  | 23(100%) |
| 合計     | 25(63%) | 15(38%) | 40(100%) |

図表 11 評価方法 A 結果(欠測率別)

単位: 試行回数(カッコ内は構成比)

| 欠測率 | 多重代入法   | 単一代入法   | 合計       |
|-----|---------|---------|----------|
| 30% | 11(55%) | 9(45%)  | 20(100%) |
| 40% | 9(69%)  | 4(31%)  | 13(100%) |
| 50% | 5(71%)  | 2(29%)  | 7(100%)  |
| 合計  | 25(63%) | 15(38%) | 40(100%) |

### 第3章 国内における補完推計に関する研究

#### 5) 今後の課題

本研究を踏まえて、今後の課題として以下の事項が挙げられている。

図表 12 実用化に向けた今後の課題

| 目的          | 課題   |
|-------------|--|
| 手法の<br>精度向上 | <ul style="list-style-type: none"><li>・ 経済センサスー活動調査の(速報)データから、実態にあった欠測率を用いて検証する</li><li>・ 様々なプログラムと比較する</li><li>・ 外れ値が欠測値補定に与える影響を考慮する</li></ul> |
| 補定結果<br>の公表 | <ul style="list-style-type: none"><li>・ 個別データを公表する場合に、複数ある補定値の意味をわかりやすく説明する</li><li>・ 散布図を公表す場合に、回答者が特定されるリスクに対処する</li></ul>                     |

## 2. 国民生活基礎調査

### 1) 調査の概要

国民生活基礎調査は、保健、医療、福祉、年金、所得等国民生活の基礎的事項を調査し、厚生労働行政の企画及び運営に必要な基礎資料を得るとともに、各種調査の調査客体を抽出するための親標本を設定することを目的とした調査である。調査の種類としては、3年毎に「大規模調査」を実施し、中間年には世帯の基本的事項及び所得の状況について「簡易調査」を実施している。

大規模調査では、調査事項に応じて「世帯票」、「健康票」、「介護票」、「所得票」、「貯蓄票」の計5種類の調査票があり、調査票ごとに調査対象となる標本の抽出方法が異なる。特徴的なのは、「介護票」や「所得票」、「貯蓄票」における標本抽出は、「世帯票」・「健康票」の調査対象地区から無作為抽出するという二相抽出法が採用されていることである。以下に、平成25年の大規模調査における各調査票の調査事項、標本数及び標本抽出方法を示す。

図表 13 国民生活基礎調査(平成25年大規模調査)の調査事項

| 調査票 | 調査事項                                 |
|-----|--------------------------------------|
| 世帯票 | 世帯員数、住居の種類、家計支出総額、世帯員の出生年月・学歴・就業状況、等 |
| 健康票 | 入院・入所の状況、健康状態、健診等の受診状況、等             |
| 介護票 | 要介護度の状況、介護サービスの利用状況、介護サービスの費用、等      |
| 所得票 | 所得の種類別金額、課税等の状況別金額、企業年金・個人年金等の掛金、等   |
| 貯蓄票 | 貯蓄現在高、増減理由、借入金残高、等                   |

図表 14 国民生活基礎調査(平成25年大規模調査)の調査対象

| 調査票 | 標本数  | 標本抽出方法   |
|-----|--|--|
| 世帯票 | 5,530 地区(約 11,000 単位区)、<br>約 277,000 世帯(約 716,000 人) | 国勢調査による全地区のうち、後置番号「1」及び「8」の地区(約 96 万地区)から層化無作為抽出 |
| 健康票 |  |  |
| 介護票 | 2,500 地区、約 6,000 人                                   | 上記 5,530 地区から層化無作為抽出                             |
| 所得票 | 2,000 単位区、<br>約 50,000 世帯(約 130,000 人)               | 上記 5,530 地区に設定された単位区のうち、後置番号「1」の単位区から層化無作為抽出     |
| 貯蓄票 |  |  |

注1)「後置番号」とは、国勢調査区の種類を表す番号であり、「1」は一般調査区、「8」はおおむね50人以上の単身者が居住している寄宿舍・寮等のある区域を指す。

注2)「単位区」とは、推計精度の向上や調査員の負担平準化等を図るため、1つの国勢調査区(50～60世帯)をおおむね30世帯ごとに地理的に分割したものである。

### 第3章 国内における補完推計に関する研究

#### 2) 研究の経緯・目的

厚生労働省統計情報部によれば、近年の大規模調査における世帯票の回収率は約 80%であり、二相抽出を行う所得票の回収率はさらにその約 80%である。こうした事態を背景に、統計委員会の答申において、「今後の課題」として「近年、非標本誤差を解析し、集計値を補正する理論の研究が進んでいることから、それらの利用可能性に関する検討」を行う必要があるという指摘がなされた(平成 22 年 1 月 25 日)。そこで、欠測値補完に関する専門家である成蹊大学理工学部の岩崎学教授を座長として、「国民生活基礎調査の標本設計・推定手法に関する研究会」が開催された。

研究会においては、答申の「非標本誤差を解析し、集計値を補正する理論」として、「傾向スコアを用いた分析」(後述)が採用された。分析対象は所得票の非回答により発生する非標本誤差であり、所得に関する集計値を補正できる可能性について研究が行われた。これは、世帯票の非回答世帯(全体の約 20%)の情報を収集することは困難であるが、所得票の非回答世帯(世帯票回答世帯の約 20%)の情報については、その世帯票情報と所得票への回答状況とを用いて傾向スコアを推定することが可能であると考えられるためである。なお、傾向スコアを用いる手法は座長の岩崎学教授の提案によるものであり、他の手法との比較については検討の範囲外とされた。

#### 3) 研究の体制・時期

「国民生活基礎調査の標本設計・推定手法に関する研究会」は、成蹊大学の岩崎学教授を座長とし、他に 4 名の委員と 1 名のオブザーバーによって構成された。本研究会は平成 22 年 4 月 16 日(金)を第 1 回とし、平成 23 年 3 月 3 日(木)の最終回まで、計 5 回実施された。

図表 15 「国民生活基礎調査の標本設計・推定手法に関する研究会」委員名簿

| 役割     | 氏名    | 所属・役職                   |
|--------|-------|-------------------------|
| 座長     | 岩崎 学  | 成蹊大学理工学部教授              |
| 委員     | 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 |
|        | 稲葉 由之 | 慶應義塾大学経済学部教授            |
|        | 西郷 浩  | 早稲田大学政治経済学術院教授          |
|        | 樋田 勉  | 群馬大学社会情報学部准教授           |
| オブザーバー | 伊藤 伸介 | 明海大学経済学部講師              |

注)所属・役職は当時のものである。

## 4) 研究手法の概要

所得票対象世帯の総所得金額を推定する手法として、「傾向スコア (Propensity Score)」と呼ばれる指標を用いた分析が検討された。前述のとおり、所得票調査では、先に行う「世帯票」調査において抽出された調査区から再度調査区を抽出することで調査対象を選び出している。この二相抽出の性質を利用し、傾向スコアを所得票の回答確率として用いることで、所得票対象世帯の総所得金額を推定している。

## (1) モデルの作成

総所得金額の推定に際して、所得票非回答世帯の真の傾向スコアは分からないため、世帯票での回答結果を用いてロジスティック回帰モデルによって推定値を算出している。その際に説明変数として使用された世帯票項目は以下のとおりである。相関のある項目同士を変数として使用するリスクは考慮のうえで、総所得金額との相関が想定される項目は全て使用された。

- ・ 市郡
- ・ 地域ブロック
- ・ 世帯人員数
- ・ 有業人員数
- ・ 世帯構造
- ・ 世帯類型
- ・ 世帯業態
- ・ 家計支出額
- ・ 住居の種類
- ・ 世帯主年齢階級

## (2) 推定の方法

こうして算出された傾向スコア推定値を用いて、本研究では大きく分けて以下の3種類の方法により総所得金額の推定を行なっている。

- A) 傾向スコアを用いた IPW 推定
- B) 傾向スコアを用いたマッチングによる直接代入 (ノンパラメトリックな代入)
- C) 傾向スコアを用いたマッチングによる回帰代入 (パラメトリックな代入)

A) 傾向スコアを用いた IPW (Inverse Probability Weighting) 推定

傾向スコアが回答確率そのものを表していると考え、傾向スコアの逆数でウェイト付けして、非回答世帯も含めた母集団全体の総所得金額を推計する手法である。すなわち、回答確率が低いにもかかわらず実際には回答した世帯の結果が重くなるようにウェイト付けをしている。

$$(\text{修正後のウェイト}) = (\text{現行の拡大乗数}) \times \frac{1}{(\text{当該世帯の傾向スコア推定値})}$$

B) 傾向スコアを用いたマッチングによる直接代入 (ノンパラメトリックな代入)

傾向スコア推定値を昇順に並べ、順番に非回答世帯と回答世帯とを一对一にマッチングさせた後、マッチング先の回答世帯の総所得金額をそのまま非回答世帯の総所得金額として代入する手法である。その際のマッチングの手法は、以下の3パターンを使用した。

- ・ 試算①:最近傍マッチング (Nearest Neighbor matching) (復元)  
同一の回答世帯を何度マッチングさせてもよい手法。
- ・ 試算②:最近傍マッチング (Nearest Neighbor matching) (非復元)  
一度マッチングさせた回答世帯は二度目以降マッチングさせない手法。
- ・ 試算③:キャリパーマッチング (Caliper matching) (非復元)

傾向スコア推定値との差の絶対値が 0.001 より小さい回答世帯がない場合はマッチングを行わない手法。傾向スコアが小さい回答世帯数は少ないため、傾向スコア推定値が小さい世帯ほどマッチングできない世帯が多くなり、パラメトリックな代入ができない。

### 第3章 国内における補完推計に関する研究

#### C) 傾向スコアを用いたマッチングによる回帰代入(パラメトリックな代入)

マッチング先の回答世帯における世帯票の回答結果と総所得金額との間の重回帰式を作成し、それに基づいて非回答世帯の総所得金額を推定、代入する手法である。その際重回帰モデルは、マッチングできた $M$ 世帯の第 $k$ 回答世帯において、総所得金額の対数値 $\log(Y_k)$ を被説明変数、 $m$ 個の世帯票情報 $X_{1k}, X_{2k}, \dots, X_{mk}$ を説明変数として、

$$\log(Y_k) = \beta_0 + \beta_1 X_{1k} + \dots + \beta_m X_{mk} + \varepsilon_k$$

をあてはめ、パラメータ $\beta_l$ の推定値 $\hat{\beta}_l$ を最小二乗法により決定している。このとき、所得票非回答のある第 $t$ 世帯に対して、以下の3種類の回帰式によって補完値を算出している。

・ 不偏推定  $\hat{Y}_{1t} = \left\{1 + \frac{1}{2}s^2\right\} \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1t} + \dots + \hat{\beta}_m X_{mt})$

・ 最尤推定  $\hat{Y}_{2t} = \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1t} + \dots + \hat{\beta}_m X_{mt})$

・ 確率的回帰代入  $\hat{Y}_{3t} = \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1t} + \dots + \hat{\beta}_m X_{mt} + \hat{\varepsilon}_t)$

ただし、 $s^2 = \frac{1}{M-m-1} \sum_k \hat{u}_k^2$

$$\hat{u}_k = \log(Y_k) - (\hat{\beta}_0 + \hat{\beta}_1 X_{1k} + \dots + \hat{\beta}_m X_{mk})$$

$$\hat{\varepsilon}_t \sim N(0, s^2)$$

このうち、確率的回帰代入の重回帰式には残差項があり、これは正規分布に従って確率的に決定されるため、補完値に再現性がないという制約があることに注意が必要である。

なお、重回帰に用いた説明変数は、傾向スコアの推定の際に用いられた変数に加えて、以下の変数が使用された。

- ・ 有業者構成
- ・ 公的年金受給者数
- ・ 公的年金加入者数
- ・ 世帯主の性別

#### (3) 評価

以上の3種類の方法による推定結果を受けて、本研究では、「所得分布の補整方法としてはIPW推定やノンパラメトリック代入の結果が補整後のイメージに近い」としながらも、「母子世帯や父子世帯のような出現頻度の低い属性においては、属性の異なる世帯とマッチングされる可能性が高く、個別の所得分布を見ると実績値と著しく乖離した結果となることもあるため、ノンパラメトリック代入による補整には注意が必要である」としている。一方「パラメトリック代入は、結果がモデル選択の良否に左右される傾向にあるため、IPW推定及びノンパラメトリック代入と比較した推定の良し悪しは一概には言えない」としている。

また、本手法においては、「所得票の無回答が総所得の値に依存しないランダムな欠測(MAR)であり、なおかつ共変量の選択が適切であれば」という前提を置かなければならないことにも注意が必要である。

図表 16 A:IPW 推定結果

|                 | 実績値   | IPW 推定値 |
|-----------------|-------|---------|
| 1 世帯あたり平均所得(万円) |       |         |
| 全世帯             | 566.8 | 564.0   |
| 高齢者世帯           | 306.3 | 303.5   |
| 母子世帯            | 236.7 | 236.4   |
| 父子世帯            | 515.9 | 515.9   |
| その他の世帯          | 646.8 | 634.6   |
| 所得分位値(万円)       |       |         |
| 中央値             | 451   | 450     |
| 第1五分位           | 214   | 207     |
| 第2五分位           | 365   | 360     |
| 第3五分位           | 554   | 550     |
| 第4五分位           | 838   | 834     |

図表 17 B:ノンパラメトリック代入結果

|                 | 実績値   | 試算①   |       | 試算②   |       | 試算③(参考値) |       |
|-----------------|-------|-------|-------|-------|-------|----------|-------|
|                 |       | 補完値   | 補完後   | 補完値   | 補完後   | 補完値      | 補完後   |
| 1 世帯あたり平均所得(万円) |       |       |       |       |       |          |       |
| 全世帯             | 566.8 | 545.1 | 558.7 | 564.5 | 566.0 | 561.2    | 564.8 |
| 高齢者世帯           | 306.3 | 375.1 | 327.2 | 452.5 | 350.9 | 420.5    | 340.5 |
| 母子世帯            | 236.7 | 552.3 | 361.1 | 597.4 | 379.0 | 581.3    | 366.3 |
| 父子世帯            | 515.9 | 384.7 | 451.0 | 540.6 | 528.1 | 528.3    | 521.7 |
| その他の世帯          | 646.8 | 578.8 | 620.4 | 585.9 | 623.2 | 590.5    | 626.2 |
| 所得分位値(万円)       |       |       |       |       |       |          |       |
| 中央値             | 451   | 430   | 446   | 450   | 450   | 450      | 450   |
| 第1五分位           | 214   | 189   | 202   | 200   | 210   | 200      | 210   |
| 第2五分位           | 365   | 342   | 358   | 360   | 364   | 360      | 363   |
| 第3五分位           | 554   | 539   | 550   | 557   | 555   | 550      | 553   |
| 第4五分位           | 838   | 830   | 834   | 850   | 842   | 847      | 840   |

### 第3章 国内における補完推計に関する研究

図表 18 C:パラメトリック代入結果

|                | 実績値   | 不偏推定  |       |       | 最尤推定  |       |       | 確率的回帰代入 |       |       |
|----------------|-------|-------|-------|-------|-------|-------|-------|---------|-------|-------|
|                |       | 試算①   | 試算②   | 試算③   | 試算①   | 試算②   | 試算③   | 試算①     | 試算②   | 試算③   |
| 1世帯あたり平均所得(万円) |       |       |       |       |       |       |       |         |       |       |
| 全世帯            | 566.8 | 565.0 | 564.2 | 564.0 | 534.1 | 533.6 | 533.9 | 568.9   | 568.2 | 567.9 |
| 高齢者世帯          | 306.3 | 304.7 | 304.1 | 304.3 | 291.2 | 290.7 | 291.1 | 305.7   | 305.0 | 305.2 |
| 母子世帯           | 236.7 | 239.8 | 238.8 | 240.0 | 225.5 | 224.8 | 226.0 | 239.7   | 238.8 | 240.1 |
| 父子世帯           | 515.9 | 512.9 | 538.3 | 519.6 | 475.7 | 497.6 | 482.2 | 601.3   | 632.0 | 606.3 |
| その他の世帯         | 646.8 | 636.1 | 635.2 | 634.9 | 600.6 | 600.1 | 600.3 | 640.6   | 639.9 | 639.5 |
| 所得分位値(万円)      |       |       |       |       |       |       |       |         |       |       |
| 中央値            | 451   | 472   | 473   | 472   | 442   | 445   | 444   | 438     | 438   | 438   |
| 第1五分位          | 214   | 227   | 226   | 226   | 212   | 213   | 213   | 200     | 200   | 200   |
| 第2五分位          | 365   | 382   | 382   | 382   | 360   | 360   | 360   | 350     | 350   | 350   |
| 第3五分位          | 554   | 575   | 575   | 574   | 540   | 540   | 540   | 542     | 542   | 541   |
| 第4五分位          | 838   | 848   | 847   | 846   | 789   | 790   | 790   | 849     | 849   | 848   |

#### 5) 今後の課題

本研究では、「真の所得分布が未知であり、補整結果の妥当性の検証が行えない」ことや、確率的回帰代入においては補完値が一意に決定されないこと、詳細な項目についても結果表を全て作成するには相当な工数が必要であること等を踏まえて、「いずれの補整手法も集計に適用可能であると結論づけることまではできない」としている。

ただし、厚生労働省統計情報部によれば「補完値の多くは標準誤差率の内に収まっているため、元の集計値の確からしさを補強するための参考資料として使用することは可能だと考えている」と述べているとおり、補完推計の実務への適用可能性については、今後も継続的な研究が必要であろう。

### 3. 特定サービス産業実態調査

#### 1) 調査の概要

特定サービス産業実態調査は、サービス産業の実態を明らかにし、サービス産業に関する施策の基礎資料を得ることを目的とした統計調査であり、調査の概要は以下のとおりである。

本統計調査は、平成 19 年から平成 21 年にかけて調査対象の業種を大きく拡大している。平成 18 年調査では日本標準産業分類ベースで 7 業種であったが、平成 19 年調査では 4 業種を追加して 11 業種に、平成 20 年調査では 10 業種を追加して 21 業種に、さらに平成 21 年調査からは、7 業種(「冠婚葬祭業」、「映画館」、「興行場・興行団」、「スポーツ施設提供業」、「公園・遊園地・テーマパーク」、「学習塾」、「教養・技能教授業」)を追加して 28 業種となっている。

図表 19 特定サービス産業実態調査の概要

| 項目     | 内容  |
|--------|---|
| 実施頻度   | 毎年<br>※経済センサスー活動調査実施年の前年を除く   |
| 調査対象   | 全国の事業所(一部業種は企業)   |
| 標本抽出方法 | 層化抽出法(業種別、事業従事者規模別、都道府県別)<br>※一部業種については母集団数が小規模なため全数調査<br>※平成 20 年調査までは全数調査 |
| 標本数    | 約 55,000 事業所  |
| 調査方法   | 留置調査、郵送調査<br>(調査員又は郵送により調査票を配布し、調査員又は郵送により回収する)                             |
| 回収率    | 約 82%(平成 22 年調査結果)<br>※調査票回収数÷調査対象事業所数により算出                                 |
| 調査項目   | ・ 共通項目(フェース項目、従業者数、年間売上高、等)<br>・ 固有事項(部門別従業者数、契約高・取扱高、業務種別売上高、会員数、作品数、等)    |

#### 2) 研究の経緯・目的

『平成 21 年度サービス産業統計等経済産業統計の基盤整備事業「特定サービス産業実態調査等における推計手法の確立に関する調査研究」調査報告書』(平成 22 年 3 月)によれば、経済産業省が実施する主要な大規模全数統計である商業統計、特定サービス産業実態調査及び企業活動基本調査(当時)について、「統計調査の実施に当たり、調査客体への働きかけ等により回収率向上の努力は行なっているものの、100%の回収率には至らないことから、審議会等において当該産業の実態を示すために検討することが課題として指摘」されていた。そこで、「各統計調査の特性に応じた欠測値の推計手法を検討」することを目的として、「特定サービス産業実態調査等における推計手法の確立に関する調査研究 研究会」が設置され、本研究が実施された。

特定サービス産業実態調査については、「平成 20 年度の検討により、当面の間「平均値補完」によって行う方針としている」とした上で、「悉皆調査を行う大規模事業所層のうち、規模の上限のない最大の層(事業従事者数 500 人以上)については、「平均値補完」ではなく、母集団名簿の事業従事者数を補助変量とした比補完を行うことで、より精度の高い推計が実現可能かどうか」検討を行なっている。

### 第3章 国内における補完推計に関する研究

#### 3) 研究の体制・時期

「特定サービス産業実態調査等における推計手法の確立に関する調査研究 研究会」は、青山学院大学経済学部的美添泰人教授を座長とし、他3名の委員から構成された。本研究会は平成21年12月24日(木)を第1回とし、平成22年3月8日(金)の最終回まで、計4回実施された。

図表 20 「特定サービス産業実態調査等における推計手法の確立に関する調査研究 研究会」委員名簿

| 役割 | 氏名    | 所属・役職                           |
|----|-------|---------------------------------|
| 座長 | 美添 泰人 | 青山学院大学経済学部教授                    |
| 委員 | 西郷 浩  | 早稲田大学政治経済学術院教授                  |
|    | 土屋 隆裕 | 大学共同利用機関法人情報・システム研究機構統計数理研究所准教授 |
|    | 山本 渉  | 国立大学法人電気通信大学電気通信学部講師            |

注)所属・役職は当時のものである。

#### 4) 研究手法の概要

##### (1) 欠測値補完の方法

本研究会では、「実査における適用を想定した場合、時間的な制約から複雑な形で行うことが難しい」ことを前提として、一般的な手法のうち以下の手法を候補として検討している。

図表 21 検討対象の補完手法

| # | 手法名                | 具体的方法                              |                                   |
|---|--------------------|------------------------------------|-----------------------------------|
| 1 | 横置き補完 <sup>2</sup> | 過去の調査結果を代入する                       |                                   |
| 2 | 伸び率補完              | 過去の調査結果に層内の総和の前回調査比(伸び率)を乗じた値を代入する |                                   |
| 3 | 平均値補完              | 層内の平均値を代入する                        |                                   |
| 4 | 平均値補完<br>+比補完      | 500人未満の層                           | 層内の平均値を代入する                       |
|   |                    | 500人以上の層                           | 母集団名簿の事業従事者数を補助変量とした比補完により総和を推定する |

注)いずれの手法においても、最大規模の層(500人以上)においては、層内平均+3σを超える売上高の事業所(企業)は欠測させない

<sup>2</sup> 「横置き補完」と「据置き補完」とは同義であるが、本節では当該研究会における用語法に準ずる。

## (2) シミュレーションの方法

本研究では、「特定サービス産業実態調査において最も注目している数値」である売上高を欠測させ、以上の各手法により補完するシミュレーションを実施している。その際のシミュレーション条件は以下のとおりである。

図表 22 シミュレーションの条件

| 項目     | 内容  |
|--------|---|
| 母集団データ | 平成 19 年調査・平成 20 年調査からパネル化できた事業所データ<br>※対象業種は以下<br>「ソフトウェア業」「情報処理・提供サービス業」「各種物品賃貸業」「産業用機械器具賃貸業」「事務用機器器具賃貸業」「広告業-旧広告代理業」「広告業-旧その他の広告業」「映像情報制作・配給業」「クレジットカード業、割賦金融業」「デザイン業、機械設計業」「計量証明業」 |
| 欠測パターン | 層別に無作為に発生<br>※層は「業種 × 都道府県(47 区分) × 事業従事者規模(8 区分)」を設定   |
| 欠測率    | 平成 20 年調査実績より層別に設定  |
| 試行回数   | 各手法とも 1,000 回   |

## (3) 評価

シミュレーションの結果、以下が判明した。

- ・ 伸び率補完が最もばらつきが小さく、かつ、分布の歪みも少ない
- ・ 大規模事業所が欠測しないとしても、平均値補完は、伸び率補完、横置き補完と比較すると、ばらつきが大きい点は変わらない(若干は改善している)
- ・ 平均値補完を基本として、事業従事者規模 500 人以上の層については比補完によって総和を推定した場合、業種によっては平均値補完のみを用いる場合と比較して真値に近い推定が可能となり、より精度の高い推定が期待できるが、その差はわずかである
- ・ 「500 人以上」の層に該当する事業所(企業)が少ない等の業種では、比補完を行う場合の方が、ばらつきが大きくなる場合がある。

第3章 国内における補完推計に関する研究

図表 23 シミュレーション結果 1/2

|                | 補完なし               | 補完あり          | 平均値補完         | 他比率補完         | 比推定           |
|----------------|--------------------|---------------|---------------|---------------|---------------|
| 391 ソフトウェア業    | 真値                 | 1,165,864,867 | 1,165,864,867 | 1,165,864,867 | 1,165,864,867 |
|                | 平均値                | 1,060,285,251 | 1,169,042,932 | 1,189,655,299 | 1,165,384,548 |
|                | 最小値                | 991,501,719   | 1,125,709,371 | 1,104,161,850 | 1,119,859,814 |
|                | 第1四分位              | 1,048,259,810 | 1,165,960,531 | 1,179,165,459 | 1,161,629,679 |
|                | 中央値                | 1,064,436,897 | 1,169,436,036 | 1,191,738,194 | 1,165,634,963 |
|                | 第3四分位              | 1,075,141,228 | 1,172,955,519 | 1,202,711,692 | 1,170,953,177 |
|                | 最大値                | 1,102,344,859 | 1,188,205,873 | 1,243,118,583 | 1,183,986,332 |
|                | 標準偏差               | 29,194,016    | 7,486,156     | 20,847,316    | 7,888,713     |
|                | 乖度                 | -0.76         | -1.93         | -0.63         | -1.46         |
|                |                    |               |               |               | -0.80         |
|                | 392 情報処理・提供サービス業   | 真値            | 371,384,474   | 371,384,474   | 371,384,474   |
| 平均値            |                    | 324,436,230   | 376,116,164   | 379,865,186   | 370,658,707   |
| 最小値            |                    | 283,306,847   | 351,862,131   | 346,284,798   | 343,840,117   |
| 第1四分位          |                    | 317,940,383   | 373,529,279   | 374,104,583   | 367,720,659   |
| 中央値            |                    | 327,026,432   | 376,147,181   | 381,224,855   | 370,995,863   |
| 第3四分位          |                    | 332,423,546   | 376,973,569   | 397,150,675   | 374,126,036   |
| 最大値            |                    | 345,751,325   | 395,302,140   | 413,103,804   | 387,937,503   |
| 標準偏差           |                    | 11,021,380    | 4,677,245     | 11,013,281    | 5,508,850     |
| 乖度             |                    | -0.81         | 0.78          | -0.45         | -0.76         |
|                |                    |               |               |               | -0.42         |
| 411 映像情報制作・配給業 |                    | 真値            | 77,329,974    | 77,329,974    | 77,329,974    |
|                | 平均値                | 45,540,239    | 78,367,694    | 77,080,867    | 74,253,529    |
|                | 最小値                | 19,226,169    | 55,812,005    | 38,731,912    | 46,358,758    |
|                | 第1四分位              | 40,396,245    | 74,921,853    | 70,376,049    | 68,464,542    |
|                | 中央値                | 46,088,072    | 76,325,699    | 77,198,050    | 74,808,765    |
|                | 第3四分位              | 50,864,202    | 81,973,618    | 84,039,712    | 80,274,488    |
|                | 最大値                | 65,980,426    | 99,071,705    | 109,498,842   | 100,988,845   |
|                | 標準偏差               | 7,755,270     | 5,654,280     | 9,898,855     | 9,062,686     |
|                | 乖度                 | -0.26         | 0.01          | -0.16         | -0.12         |
|                |                    |               |               |               | -0.11         |
|                | 643 クレジットカード業 割賦金融 | 真値            | 5,472,428,292 | 5,472,428,292 | 5,472,428,292 |
| 平均値            |                    | 5,140,484,779 | 5,446,032,690 | 5,450,392,919 | 5,420,860,226 |
| 最小値            |                    | 3,599,149,351 | 4,882,498,828 | 4,282,903,392 | 4,593,768,217 |
| 第1四分位          |                    | 4,895,935,275 | 5,409,795,590 | 5,401,342,178 | 5,400,589,562 |
| 中央値            |                    | 5,346,860,226 | 5,444,815,598 | 5,470,651,121 | 5,430,994,818 |
| 第3四分位          |                    | 5,398,773,102 | 5,472,416,262 | 5,631,578,343 | 5,474,637,619 |
| 最大値            |                    | 5,471,076,847 | 5,905,106,019 | 6,880,219,216 | 5,979,167,162 |
| 標準偏差           |                    | 375,366,016   | 80,435,375    | 286,723,605   | 125,357,583   |
| 乖度             |                    | 1.34          | 0.63          | -0.38         | -1.57         |
|                |                    |               |               |               | 0.44          |
| 701 各種物品買貨業    |                    | 真値            | 633,868,814   | 633,868,814   | 633,868,814   |
|                | 平均値                | 555,683,953   | 642,321,073   | 634,243,142   | 623,799,559   |
|                | 最小値                | 484,736,196   | 604,764,439   | 568,465,396   | 571,824,180   |
|                | 第1四分位              | 543,077,398   | 635,450,254   | 623,345,146   | 616,537,014   |
|                | 中央値                | 556,935,944   | 641,361,792   | 634,521,532   | 624,803,153   |
|                | 第3四分位              | 569,418,460   | 649,611,662   | 645,178,119   | 632,280,130   |
|                | 最大値                | 596,128,768   | 673,625,145   | 700,178,028   | 660,334,721   |
|                | 標準偏差               | 18,886,363    | 10,992,271    | 17,376,807    | 13,509,177    |
|                | 乖度                 | -0.40         | 0.13          | -0.09         | -0.48         |
|                |                    |               |               |               | -0.09         |
|                | 702 産業用機械器具買貨業     | 真値            | 157,103,151   | 157,103,151   | 157,103,151   |
| 平均値            |                    | 130,219,149   | 157,931,350   | 157,169,403   | 155,356,467   |
| 最小値            |                    | 123,090,013   | 152,920,524   | 149,061,911   | 148,197,642   |
| 第1四分位          |                    | 128,871,202   | 157,234,423   | 155,586,237   | 154,007,632   |
| 中央値            |                    | 130,423,962   | 158,061,783   | 157,295,828   | 155,441,612   |
| 第3四分位          |                    | 131,696,205   | 158,867,979   | 158,851,539   | 156,739,592   |
| 最大値            |                    | 136,171,062   | 161,556,468   | 163,755,130   | 160,542,991   |
| 標準偏差           |                    | 2,093,675     | 1,329,001     | 2,371,258     | 1,881,237     |
| 乖度             |                    | -0.41         | -0.76         | -0.26         | -0.35         |
|                |                    |               |               |               |               |

第3章 国内における補完推計に関する研究

図表 23 シミュレーション結果 2/2

|                 | 補完なし               | 模擬き補完       | 平均値補完       | 伸び率補完       | 比推定         |
|-----------------|--------------------|-------------|-------------|-------------|-------------|
| 703 事務用機械器具賃貸業  | 真値                 | 64,259,756  | 64,259,756  | 64,259,756  | 64,259,756  |
|                 | 平均値                | 37,657,092  | 55,641,852  | 47,781,824  | 53,306,254  |
|                 | 最小値                | 12,352,591  | 21,819,733  | 18,086,797  | 18,217,071  |
|                 | 第1四分位              | 23,765,917  | 50,331,797  | 33,873,856  | 47,537,691  |
|                 | 中央値                | 42,602,558  | 59,179,230  | 48,982,313  | 56,352,684  |
|                 | 第3四分位              | 50,657,460  | 64,826,757  | 61,403,848  | 62,800,833  |
|                 | 最大値                | 62,464,071  | 80,155,785  | 91,638,515  | 83,721,999  |
|                 | 標準偏差               | 14,755,206  | 12,754,164  | 16,082,570  | 13,904,330  |
|                 | 歪度                 | -0.04       | -0.85       | -0.01       | -0.69       |
|                 | 726743 デザイン業 機械設計業 | 真値          | 52,613,146  | 52,613,146  | 52,613,146  |
| 平均値             |                    | 43,054,270  | 52,957,940  | 52,712,169  | 52,291,513  |
| 最小値             |                    | 40,968,494  | 51,725,080  | 50,969,222  | 50,919,168  |
| 第1四分位           |                    | 42,588,230  | 52,595,988  | 52,340,363  | 51,884,291  |
| 中央値             |                    | 43,073,658  | 52,921,456  | 52,698,982  | 52,280,535  |
| 第3四分位           |                    | 43,562,813  | 53,298,893  | 53,089,506  | 52,662,630  |
| 最大値             |                    | 45,018,007  | 55,127,428  | 54,754,784  | 54,265,679  |
| 標準偏差            |                    | 727,756     | 520,984     | 557,754     | 568,478     |
| 歪度              |                    | -0.12       | 0.41        | 0.19        | 0.17        |
| 7311 広告業-旧広告代理業 |                    | 真値          | 616,911,098 | 616,911,098 | 616,911,098 |
|                 | 平均値                | 520,116,882 | 618,927,594 | 617,456,854 | 603,583,767 |
|                 | 最小値                | 310,305,318 | 474,372,749 | 390,719,903 | 397,263,241 |
|                 | 第1四分位              | 495,269,617 | 617,120,569 | 601,832,385 | 603,173,798 |
|                 | 中央値                | 545,317,778 | 619,270,532 | 626,264,349 | 613,410,197 |
|                 | 第3四分位              | 559,438,150 | 621,846,281 | 649,818,843 | 618,423,919 |
|                 | 最大値                | 579,305,548 | 699,883,708 | 801,456,394 | 735,597,842 |
|                 | 標準偏差               | 55,797,747  | 21,171,673  | 60,225,045  | 39,134,197  |
|                 | 歪度                 | -1.39       | -2.98       | -0.80       | -1.82       |
|                 | 7312 広告業-旧その他の広告業  | 真値          | 110,974,314 | 110,974,314 | 110,974,314 |
| 平均値             |                    | 97,632,949  | 111,694,420 | 111,320,706 | 108,684,948 |
| 最小値             |                    | 89,619,409  | 104,566,876 | 100,935,966 | 99,561,077  |
| 第1四分位           |                    | 95,550,424  | 110,573,068 | 108,658,886 | 106,426,693 |
| 中央値             |                    | 97,867,222  | 111,777,078 | 111,357,916 | 108,911,023 |
| 第3四分位           |                    | 99,850,655  | 112,869,336 | 113,807,011 | 110,794,512 |
| 最大値             |                    | 104,403,402 | 119,882,550 | 125,553,057 | 118,874,779 |
| 標準偏差            |                    | 2,868,250   | 1,932,719   | 3,919,780   | 3,149,242   |
| 歪度              |                    | -0.28       | -0.13       | 0.11        | -0.00       |
| 745 計量証明業       |                    | 真値          | 11,875,564  | 11,875,564  | 11,875,564  |
|                 | 平均値                | 11,236,815  | 11,897,320  | 11,904,224  | 11,815,372  |
|                 | 最小値                | 10,789,057  | 11,678,757  | 11,639,414  | 11,511,463  |
|                 | 第1四分位              | 11,133,328  | 11,849,705  | 11,849,845  | 11,756,826  |
|                 | 中央値                | 11,243,162  | 11,888,342  | 11,901,875  | 11,814,571  |
|                 | 第3四分位              | 11,347,879  | 11,941,431  | 11,961,207  | 11,872,754  |
|                 | 最大値                | 11,639,944  | 12,122,936  | 12,308,761  | 12,103,037  |
|                 | 標準偏差               | 150,164     | 72,962      | 83,895      | 95,926      |
|                 | 歪度                 | -0.20       | 0.43        | 0.09        | -0.01       |

### 第3章 国内における補完推計に関する研究

#### 5) 今後の課題

以上のような結果を踏まえ、本研究会は「当面の方針」として以下のように結論づけている。

- ・ 平成 21 年調査において調査対象を 28 業種に拡大することで、未経験の業種を標本調査することになり、前年データを利用できない業種が存在する制約を鑑み、平均値補完を用いる
- ・ 「平均値補完」と「平均値補完＋比補完」の結果に大きな差が見られないこと、「500 人以上」の層に該当する事業所(企業)が少ない等の業種では、比補完を行う場合の方が、ばらつきが大きくなる場合があること等を勘案し、「平均値補完＋比補完」は採用しない

また、「将来的に向けた考え方」として以下のように結論づけている。

- ・ シミュレーションの結果において、いずれの業種でも「平均値補完＋比補完」と比較して「横置き補完」「伸び率補完」の方が望ましい結果となっていることを鑑み、全ての業種について前年データを利用できるようになった段階で、「横置き補完」、「伸び率補完」を採用することを改めて検討する
- ・ また、その際には、「モデルを用いた補完」についても検討を行うものとし、モデル構築の手法(層別にモデルを作成する、層を統合した上でモデルを作成する等)やモデルを見直す周期等について併せて検討を行うこととする

#### 4. 全国企業短期経済観測調査

##### 1) 調査の概要

全国企業短期経済観測調査(以下、短観)は、統計法(平成19年法律第53号)に基づいて日本銀行が実施する代表的なビジネスサーベイの一つであり、「全国の企業動向を的確に把握し、金融政策の適切な運営に資すること」を目的としている。調査の概要は以下のとおりである。

図表 24 短観の概要

| 項目     | 内容  |
|--------|---|
| 実施頻度   | 四半期毎  |
| 調査対象   | 全国の資本金2千万円以上の民間企業(金融機関を除く)(約210,000社)<br>※総務省の「事業所・企業統計調査」(2006年10月実施分)に基づく<br>※金融機関に対しては補完的な調査を実施している  |
| 標本抽出方法 | 層化抽出法(業種別、資本金規模別)   |
| 標本数    | 約11,000社  |
| 調査方法   | 郵送調査、インターネット調査  |
| 回収率    | 約99%  |
| 調査項目   | <ul style="list-style-type: none"> <li>・ 判断項目(業況、国内外での製商品・サービス需給、雇用人数、等)</li> <li>・ 年度計画(売上高、材料費、人件費、減価償却費、営業利益、経常利益、設備投資額、等)</li> <li>・ 四半期項目(金融機関借入金、社債、現金・預金、有価証券、等)</li> <li>・ 新卒者採用状況(新卒採用者数)</li> </ul> |

##### 2) 研究の経緯・目的

日本銀行ワーキングペーパーシリーズ「ビジネスサーベイにおける欠測値補完の検討—全国企業短期経済観測調査(短観)のケース—」によれば、短観は「調査先企業のご協力を得て高い回答率を達成しているものの、調査回によって欠測値が相応に存在」しており、このうち、「売上高や設備投資計画等の計数項目において欠測値が発生した場合は、『欠測した企業の前年度の回答計数』を補完した上で、集計結果を取りまとめている」という。

こうした補完方法(横置き補完)は、本報告書第2章でも述べたとおり、「経済環境が安定している時期においては、前期から今期にかけての計数の変動が比較的小さいことから、妥当と考えられる」が、その一方で、「経済環境が大きく変動するような局面においては、この間の動きを適切に反映していない前期の回答計数がそのまま代入されるため、景気の実感からやや離れた集計結果につながる恐れがある」という特徴がある。特に、近年の「リーマンショックのような大きな経済変動」が発生した場合、後者の側面が強く出ることになる。

そこで日本銀行調査統計局の平川氏と鳩貝氏は、短観において「より精度の高い補完方法の有無を検討」するための研究を実施した。

##### 3) 研究の体制・時期

本研究の内容は、日本銀行調査統計局の平川貴大氏と鳩貝淳一郎氏によって、2011年秋に開催された、欧州委員会・OECD共催の「Fifth Joint EC-OECD Workshop on Recent Developments in Business and Consumer Surveys(2011年11月17～18日、於ブリュッセル)」において報告された。

### 第3章 国内における補完推計に関する研究

#### 4) 研究手法の概要

##### (1) 欠測値補完の方法

本研究では、現行の補完方法である横置き補完<sup>3</sup>も含め、以下の3手法について比較、検討している。具体的には、設備投資額と売上高の補完には伸び率補完を、経常利益の補完には項目間比率補完を適用し、それぞれ3種類ずつの方法によって補完値を算出している。

図表 25 検討対象の補完手法

| # | 手法名           | 適用する欠測項目           | 具体的方法   |
|---|---------------|--------------------|---|
| 1 | 横置き補完<br>(現行) | 設備投資額、売上高、<br>経常利益 | 前期の回答値をそのまま代入   |
| 2 | 伸び率補完         | 設備投資額、売上高          | <p>前期の回答値を基に、回答企業から得られた「前期から今期の変化(伸び率)」を加味した値を代入</p> <p>①業種×規模<br/>欠測企業が属する「業種×規模」グループの回答企業の平均伸び率を加味</p> <p>②雇用者分割<br/>欠測企業が属する「雇用者分割」グループの回答企業の平均伸び率を加味</p> <p>③近接10社<br/>欠測企業が属する「業種×規模」グループのうち、さらに設備投資額又は売上高が近い「近接10社」グループの回答企業の平均伸び率を加味</p> |
| 3 | 項目間比率補完       | 経常利益               | <p>「欠測項目と他項目の関係(項目間比率)」を利用して算出した値を代入</p> <p>①欠測前年<br/>欠測企業の前期の「経常利益/売上高」率を利用</p> <p>②回答平均<br/>欠測企業が属する「業種×規模」グループの回答企業の、今期の平均「経常利益/売上高」率を利用</p> <p>③伸び幅<br/>欠測企業の前期の「経常利益/売上高」率と、欠測企業が属する「業種×規模」グループの回答企業の、「経常利益/売上高」率の平均伸び幅を利用</p>             |

<sup>3</sup> 「横置き補完」と「据置き補完」とは同義であるが、本節では当該ワーキングペーパーにおける用語法に準ずる。

## (2) シミュレーションの方法

本研究では、以上の手法を比較、検討するために、短観の実データを用いたモンテカルロ・シミュレーションを実施している。その際のシミュレーション条件は以下のとおりである。

図表 26 シミュレーションの条件

| 項目     | 内容   |
|--------|--|
| 母集団データ | 2004～2010 年度の実績値を基に作成                              |
| 欠測パターン | グループごとに無作為に発生<br>※グループは「業種×資本金規模×雇用者数」により 400 弱に設定 |
| 欠測率    | 各年度の 3 月調査における新年度計数の実績をグループごとに設定                   |
| 試行回数   | 各手法とも 500 回  |

## (3) 評価

各補完方法の評価は以下のとおりに行なっている。

- ・ 欠測値を補完したうえで、「大／中堅／中小企業」×「製造／非製造業」の 6 区分における、当該項目の総額(母集団推計値)を算出
- ・ 評価当該項目の総額について、欠測値を補完した値と真値(もとのデータセットの値)との誤差を、PRMSE (Relative Root Mean Squared Error)として算出

PRMSE は、推定値と真値の誤差の指標の 1 つであり、「推定値と真値の差の 2 乗」の期待値につき平方根をとった値を、真値で除した値である。したがって、この値が小さいほど真値に近く、補完精度が高いと解釈できる。本研究では、補完方法の統計精度を PRMSE の水準とし、「代替方法が現行方法よりもどれだけ PRMSE を低下させたか」を検証している。

この結果、以下の 2 点が判明した。したがって、主要 3 項目(設備投資額、売上高、経常利益)においては、「現行の補完方法より精度が高いかほぼ同等であるような代替方法が存在するといえる」と結論づけている。

- ・ 前年比変化率が大きい時期については、代替方法の採用による精度改善効果が大きい。一部で悪化しているケースもあるが、現行方法とほぼ同等の精度との評価が可能
- ・ 前年比変化率が大きい時期以外でも、代替方法は、現行方法より精度が高いかほぼ同等との評価が可能

### 第3章 国内における補完推計に関する研究

図表 27 シミュレーション結果:PRMSE の改善度合い

| 欠測項目      | 補完手法                      | 対象業種 | リーマンショックを背景に前年<br>比変化率が大きくなった時期         | 左記以外の時期                         |   |
|-----------|---------------------------|------|---|---------------------------------|---|
| 設備<br>投資額 | 伸び率補完<br>①業種×規模<br>③近接10社 | 製造業  | ・ 全ての規模で改善                              | ◎<br>・ 大・中堅で若干の改善<br>・ 中小で若干の悪化 | ○ |
|           |                           | 非製造業 | ・ 中堅で改善<br>・ 大で若干の悪化<br>・ 中小で若干の悪化または改善 | ○<br>・ 全ての規模で若干の悪化              | ○ |
| 売上高       | 伸び率補完<br>①業種×規模<br>③近接10社 | 製造業  | ・ 全ての規模で改善                              | ◎<br>・ 全ての規模で改善                 | ◎ |
|           |                           | 非製造業 | ・ 全ての規模で改善                              | ◎<br>・ 全ての規模で改善                 | ◎ |
| 経常利益      | 項目間<br>比率補完<br>③伸び幅       | 製造業  | ・ 全ての規模で改善                              | ◎<br>・ 中堅・中小で改善<br>・ 大で若干の悪化    | ○ |
|           |                           | 非製造業 | ・ 全ての規模で改善                              | ◎<br>・ 大・中堅で改善<br>・ 中小で若干の悪化    | ○ |

注) ◎:全ての企業規模で精度が改善しているケース

○:若干の悪化がみられるものの、ほぼ同等の精度といえるケース

#### 5) 今後の課題

本研究では以上のような結論に至っているが、以下については注意を呼びかけている。

- ・ あくまで現状での実務上の実現可能性を踏まえて、代替方法となり得る考え方を選り出し検討を加えたものであり、今後の経済環境や欠測値の発生状況によって異なる結論が得られる可能性もある
- ・ 「どの項目、時期、業種、企業規模においても、代替方法がより適切である」ことを意味しているわけではない。したがって、代替方法の短観への適用を検討する場合には、具体的な適用範囲について慎重に考慮する必要がある

平川氏・鳩貝氏は、「経済環境が大きく変動する局面においては、欠測値の補完方法により推計値が左右され得るため、ビジネスサーベイにおける欠測値補完方法の改善は、統計作成上の重要テーマである」としたうえで、今後の課題として、「より良い欠測値補完方法について、検討を続けることが重要である」と述べている。

## 5. 日本人の国民性調査

### 1) 調査の概要

日本人の国民性調査(以下、国民性調査)は、統計数理研究所が行っている統計調査の1つで、日本人のものの見方や考え方とその変化を、社会調査によって捉えようとするものである。統計数理研究所によれば、本調査には以下の3つの目的がある。

- ・ 国民性の解明  
調査結果をとおして、日本人のものの見方や考え方とその変化を明らかにする
- ・ 調査手法の研究開発  
実際の調査を行いながら、今後の社会変化にも対応できる新たな統計調査手法を研究する
- ・ 統計手法の研究開発  
現実の調査データを用いることで、実際に役に立つ統計的方法を開発する

なお、調査の概要は以下のとおりである。

図表 28 国民性調査の概要

| 項目     | 内容  |
|--------|---|
| 実施頻度   | 5年毎   |
| 調査対象   | 全国の20歳以上80歳未満の男女個人  |
| 標本抽出方法 | 層化2段抽出法   |
| 標本数    | 約6,500人<br>(面接調査:3,000人、留置調査:3,500人)                        |
| 調査方法   | 面接調査、留置調査<br>※回答者の居住地により異なる                                 |
| 回収率    | 約52%(第12次全国調査(2008年秋)実績)<br>※回収標本数(3,302)÷抽出標本数(6,398)により算出 |
| 調査項目   | 基本項目、個人的態度、宗教、子供・家、身近な社会、男女の差異、一般の社会問題、政治的態度、日本人・人種         |

### 2) 研究の経緯・目的

2008年秋に実施された「日本人の国民性 第12次全国調査」では、抽出標本数6,398に対し回収標本数3,302で、回収率は約51.6%となり過去最低を記録した。「調査への指向性変数を用いた調査不能バイアスの二段補正—『日本人の国民性 第12次全国調査』への適用—」によると、国民性調査では、原則として全ての個人が等しい抽出ウエイトをもつ自己加重標本が抽出されるように標本設計を行なっている。それゆえ、「抽出標本の半数近くが調査不能であれば、標本割合は母集団割合の推定値としてバイアスをもつ場合が少なくない」と考えられる。こうした問題意識の下で、統計数理研究所の土屋隆裕氏・中村隆氏は、「推定のためのウエイト調整を行い、調査不能によるバイアスの軽減を試みる」研究を開始した。

その際、調査不能の理由には「拒否」だけでなく「不在」も少なくなく、ウエイト調整に際しては、「それぞれに対応した指向性変数を用いる必要がある」という。そこで、本研究では、「複数の指向性変数を用いたウエイト調整を試み、標本割合や、属性変数のみを用いたウエイト調整の結果と比較することで、その妥当性を検討する」ことが目的とされている。

### 第3章 国内における補完推計に関する研究

#### 3) 研究の体制・時期

本研究は統計数理研究所の土屋隆裕氏・中村隆氏によって行われ、「調査への指向性変数を用いた調査不能バイアスの二段補正—『日本人の国民性 第12次全国調査』への適用—」として平成21年9月30日に報告された(これは受付日であり、採択は同年12月24日)。

#### 4) 研究手法の概要

##### (1) 二段階のキャリブレーション

本研究で用いられるウエイト調整の方法は、「事後層化やレイキング、一般化回帰推定量をその特別な場合として含む非常に一般的な方法」である。むしろ本研究の独自性は、そのキャリブレーションを二段階で行う点にある。一般に近年の調査不能の二大理由は拒否と不在であり、第12次調査ではこの2つの理由で調査不能の8割を占めるという。したがって本研究における不能バイアスの補正に際しては、拒否と不在の2つの軸を考慮している。すなわち、まず拒否による調査不能を補正し、その後不在による調査不能を補正している。

図表 29 第12次調査における各標本の関係と数

| 標本名称 | 標本数   |
|------|-------|
| 母集団  | -     |
| 非標本  | -     |
| 抽出標本 | 6,398 |
| 不在標本 | 1,436 |
| 在宅標本 | 4,962 |
| 拒否標本 | 1,660 |
| 回収標本 | 3,302 |

##### (2) 補助変数

キャリブレーションにおける補助変数については、ウエイト調整に属性変数だけを用いると、推定値のバイアスはかえって大きくなるおそれがあるため、人口統計学的な属性変数に加えて、「調査への指向性変数」を用いている。例えば、若年層の男性は一般に不在がちであるが、家にいることが多いゆえに回収できた若年男性に大きなウエイトを与えると、項目によってはかえって誤差が大きくなってしまう。そこで、調査をしたときにどの程度不在になりにくい、あるいは積極的な協力の得られやすさを示す変数である「調査への指向性変数」を加味することで、補正の精度を高めているのである。具体的には以下のような変数を補助変数として用いている。

図表 30 補正に用いた補助変数

| 種別        | 補正対象  | 変数                     | 取得元   |
|-----------|-------|------------------------|-------|
| 属性変数      | 拒否・不在 | 性別、年齢層、市郡、地方、住居形態      | 国勢調査  |
|           | 不在    | 婚姻、学歴、勤務形態、住居の種類、世帯人員数 | 国勢調査  |
| 調査への指向性変数 | 拒否    | 調査協力理由、再協力の意向          | 調査票内  |
|           |       | 対象者の反応                 | 調査員記入 |
|           | 不在    | 休日は在宅か                 | 調査票内  |
|           |       | 訪問回数                   | 調査員記入 |

#### 5) 今後の課題

本研究を実務へ適用することを念頭においた場合に、今後の課題として、中村氏は以下を挙げている。

- ・ 母集団における真値が未知である以上、複数の手法の優劣を明確に決定することはできない
- ・ 国勢調査のデータを使用しているため、国勢調査が実施されない年は他の調査で代替しなければならない
- ・ 調査への指向性変数を取得するためには、調査費や調査員への負担が増加してしまう
- ・ サンプル数は1,000～2,000程度は必要である

また、「回収率が90%になる可能性は低い以上補完・補正は必要である」としながらも、「補正は欠測を100%リカバーできるものではないため、回収率向上の工夫は必要である」とも述べている。

### 第3章 国内における補完推計に関する研究

### **第4章 海外における補完推計に関する取組**

本研究調査では、平常時における補完推計の実施・検討事例として、海外における統計調査においてどのような取組がなされているかの整理を行った。そのために、アメリカ及びカナダにおいて補完推計を実務に適用している調査機関を対象として、用いられている手法の概要や、手法選定に至った経緯、欠測を防ぐ取組等について聞き取り調査を行った。

具体的には、アメリカセンサス局及び労働統計局、カナダ統計局の3つの統計機関における取組の整理に加えて、統計調査を受託する調査会社(westat 社)及びアカデミック分野(ミシガン大学)における取組や見解も整理することとした。

## 第4章 海外における補完推計に関する取組

### I 海外における補完推計に関する取組の概要(一覧)

| 国         | アメリカ   | アメリカ   |
|-----------|--|--|
| 統計調査名(英語) | Population Census  | American Community Survey (ACS)  |
| (日本語)     | 国勢調査   | アメリカ地域社会調査   |
| 調査概要      | 実施主体   | US Census Bureau   |
|           | 調査対象   | 世帯   |
|           | 実施頻度   | 10年毎   |
|           | 実施時期   | ・2010年調査では、2010年4月1日を基準日とし、2010年3月15日～17日にかけて調査票を発送  |
|           | 標本抽出方法   | 全数調査<br>・アメリカの全ての世帯、及びその構成員が対象   |
|           | サンプル数・回収数  | 1億3,170万4,730住戸への郵送  |
|           | 調査手法   | ・紙調査票による郵送自記調査<br>・回答は、郵送による返送に加えて、調査員の訪問によるフォローが行われる  |
| 取組概要      | 概要・目的  | ・紙調査票による郵送自記調査<br>・回答は、郵送による返送に加えて、電話、調査員の訪問、ウェブでの回答が可能となっている  |
|           | 欠測の状況  | ・10年ごとに実施される国勢調査の間を埋める形で、各自治体における政策・予算配分に関する意思決定を行うために利用される<br>・2000年国勢調査で用いられた、回答項目数の多い“Long Form”調査票を、2010年国勢調査では廃止し、ACSIに移行 |
|           | 適用している手法   | ・項目非回答については、以下の3つを組み合わせた対応を行っている<br>－アサインメント (Assignment)<br>－アロケーション (Allocation)<br>－代替 (Substitution)                       |
| 経緯        | ・項目非回答については、以下の2つを組み合わせた対応を行っている<br>－アサインメント(assignment)<br>－アロケーション(allocation)   |  |
|           | ・10年ごとに実施される国勢調査<br>・補完前データ(CUF: Census Unedited File)と補完後データ(CEF: Census Edited File)を作成<br>・CEFでは、各質問について欠測・代入の状況を記録するフラグを作成している | ・ユニット非回答の発生状況を把握すべく、非回答の理由を分類・集計している<br>・拒否は全体の1%程度にとどまる   |

## 第4章 海外における補完推計に関する取組

| 国         | アメリカ                                     | カナダ   |  |
|-----------|--|---|--|
| 統計調査名(英語) | Occupational Employment Statistics (OES) | Labour Force Survey (LFS)   |  |
| (日本語)     | 職業雇用統計                                   | 労働力調査   |  |
| 調査概要      | 実施主体                                     | US Bureau of Labor Statistics   | Statistics Canada  |
|           | 調査対象                                     | 事業所・企業  | 世帯・個人  |
|           | 実施頻度                                     | 半年毎(5月・11月)   | 毎月   |
|           | 実施時期                                     | ・毎年5月と11月に実施  | ・毎月15日を含む週を基準週として、翌週に実査を行う<br>・データ収集後、13日後には結果が公表されており、月次の統計として迅速な情報提供が行われている  |
|           | 標本抽出方法                                   | 層化抽出<br>・アメリカに所在する約670万事業所から、州・産業及び所有形態に応じて抽出された約20万事業所が対象<br>・年に2回、約20万事業所を抽出し、3年間のパネルを構成  | 層化抽出<br>・AR(Address Register: 住所登録) データと呼ばれる住所データに基づいて住戸を抽出<br>・抽出された住戸に居住する15歳以上の個人が対象<br>・全国(10州に加えて3準州を含む)<br>・6か月間にわたって回答を行うパネル調査(毎月新しいパネルと入れ替えるため、1/6の回答者は新しい対象者となる)<br>・約5万4千世帯、約10万個人からの回答を得ている   |
|           | サンプル数・回収数                                | 回答率77.3%<br>(1,110,296事業所のうち、858,474事業所が回答)<br>(2011年5月調査)  | 回答率95.1%<br>(2005年平均)  |
|           | 調査手法                                     | ・郵送による自記調査が主<br>・電話、電子メール、調査員による訪問での回答も行われている   | ・CATI及び調査員による訪問面接(ノートPCを用いたCATI)<br>・電話番号が不明の場合には、初回は訪問した上で、残る5か月間についてはCATIを用いる  |
| 概要・目的     | ・12の賃金区分に基づく各事業所における従業員数を回答              | ・雇用・非雇用状況の把握を行う<br>・生産年齢にあたる人口を「雇用状態」「非雇用状態」「非労働力人口」の3つに区分することで、労働力市場の状態を把握<br>・主な質問項目は以下のとおり<br>－雇用・非雇用状況<br>－労働時間<br>－産業<br>－職業<br>－賃金 など |  |
| 取組概要      | 欠測の状況                                    | ・対象事業所は3年間のパネルを構成するため、初回のコンタクト時の協力依頼を慎重に行うことで協力率を高める  | ・CATI及びCAPIによる調査票を利用することで、データ品質の向上を図る<br>・ただし、調査効率の観点から、そうした論理矛盾等をスキップすることも可能となっているため、補完が必要な欠測が生じる<br>・修正が発生した箇所については、調査員に対する周知が行われることで同様の修正が発生しないように注意を喚起する   |
|           | 適用している手法                                 | ・ホット・デック法による代入を行う<br>・パネル調査の特性上、過去の回答結果に基づいた代入も行われる<br>・ホット・デック法では、地域、産業、従業員数を基準にし、一部産業(病院、教育機関)では所有形態も基準にされている                             | ・項目非回答については、以下の3つの手法を用いている<br>－ホット・デック法による代入<br>－Carry-Forwardによる代入(前月からの据置き(横置き))<br>－推定による代入<br>・ユニット非回答については、回答履歴の状況に応じて、ホット・デック法あるいはウェイト付けによる補完を行っている<br>・回答者が6か月にわたって回答するパネル調査であるため、前月の回答を含め、ホット・デックのドナーとする、“Longitudinal Hot-deck”と呼ばれる手法を導入している |
|           | 経緯                                       | －   | ・調査項目の中でも把握の優先度を設定し、調査における重要指標に注目した補完を行っている<br>・最優先されるものは「就労状態」、次いで「収入」とされている<br>・2005年1月以前では、Carry-Forwardも行われていたものの、月次変化を過小に評価すること、また、クロスセクションでのホット・デック法についても過大に評価することから、現在のLongitudinal Hot-Deckを用いることとした   |

## 第4章 海外における補完推計に関する取組

### II 海外における補完推計に関する調査結果

#### 1. アメリカセンサス局 (United States Census Bureau)

アメリカセンサス局は、商務省の下で人口・経済に関する統計情報を提供する機関である。ここでは、センサス局が管轄する統計調査の中から、ACS(American Community Survey)及びPopulation Census 国勢調査の2つの調査について、どのような補完推計がなされているかを整理する。

##### 1) ACS(American Community Survey: アメリカ地域社会調査)における取組み

###### (1) 調査概要

|      |  |
|------|--|
| 調査名  | : American Community Survey  |
| 概要   | : 所得や居住形態等を把握する調査  |
| 対象   | : 世帯・個人  |
| 実施時期 | : 毎年(実査は毎月実施)  |
| 調査手法 | : 郵送調査・電話調査・訪問調査<br>郵送調査に未回答の場合には電話によるフォローアップを、さらに調査員の訪問によるフォローアップを行う。             |
| 抽出方法 | : 層化抽出<br>Census Bureau が保有する住所データ(Master Address File)より、3,142 の郡を代表するようサンプリングする。 |
| 回収数  | : 世帯票では年間約 327 万世帯を抽出 集合住宅票では約 20 万人を抽出<br>最終的な実回収数は世帯票が約 213 万件、集合住宅票が約 15 万件     |

ACS は、世帯・個人を対象として毎年実施される調査で、4,000 億ドル以上の予算が連邦政府及び州政府より毎年支出されている。10年に1回実施される国勢調査における調査項目の一部について、よりタイムリーな情報を提供するために1995年から検討が始められた。国勢調査における調査項目の一部とは、所得や居住形態等を質問する”long-form”と呼ばれる、詳細な調査項目を指す。”long-form”は、サンプリングされた一部の世帯を対象とするもので、全ての世帯が回答する”short-form”に追加して質問されるものであった。Housing Units(HU)及びGroup Quarters(GQ)の2種類の調査票が利用される。

ACSでは、2000年を開始年としてデータ収集・公開を行っている。自治体や州政府、連邦政府の各種事業を策定するための情報を提供することを目的として、主に以下のような質問項目が設定されている。

図表1 ACSにおける主要質問項目

|           |           |
|-----------|-----------|
| -年齢       | -教育水準     |
| -性別       | -退役状況     |
| -人種       | -障害の有無    |
| -家族構成     | -勤務地と通勤手段 |
| -収入/給付    | -居住地と支払状況 |
| -健康保険加入状況 | 等         |

なお、調査票は通常の住宅(戸建て、マンション等)を想定した世帯票と、老人ホームや寮等を想定した集合住宅票の2種類が用いられている。

(2) 欠測を防ぐ取組

センサス局では、欠測を完全に防ぐことは不可能であるとしながらも、実査及び集計作業を通じて欠測への対応を図っている。

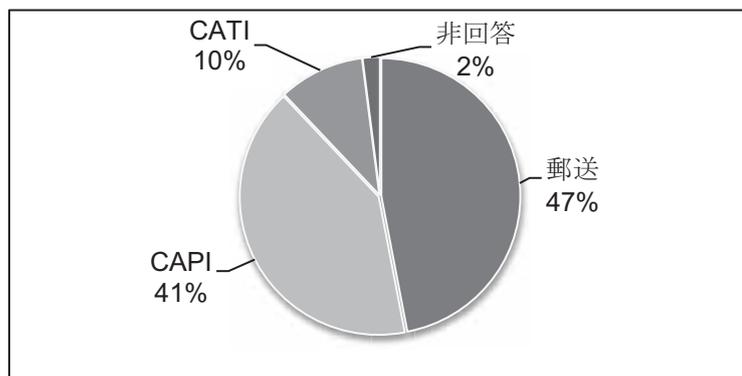
ここではユニット非回答及び項目非回答の2つに対する対応について整理する。

① ユニット非回答を防ぐ取組

ACS では、回答を全く得ることができない「ユニット非回答」を防ぐことにまず注力している。

第一に、ACS では、複数のコンタクト方法及び回答方法(郵送による回答に加えて、CATI (Computer-Aided Telephone Interview)<sup>1</sup>、CAPI (Computer-Aided Personal Interview)<sup>2</sup>)を用意することで、回答率の向上を図っている。実査においては、まず調査票を郵送し、郵送による返送を依頼し、調査票の返送を促すために3~4回の依頼状を郵送する。郵送による返送がない対象者に対しては、電話によるフォローアップが行われるが、その際にはCATIによる回答も可能とすることで、コンタクトと同時に回答できるように工夫されている。さらに、郵送、電話による依頼に対しても回答しなかった対象者に対しては、調査員が訪問することで回答を依頼している。訪問に際しては、CAPIによる回答も可能となっている。なお、2007年における回答手法の構成比は以下のとおりであった。

図表2 ACS 調査における回答手法の構成比(2007年)



第二に、ユニット非回答の発生状況とその理由を把握するために、以下のような非回答の理由を分類・集計している。その結果に応じて、コンタクト方法に工夫を図っている。センサス局では、ACS への回答を義務付けているが、一定程度のユニット非回答が発生する。下記の分類は、調査員によって判断・記入され、整理される。

<sup>1</sup> PC を利用した電話調査のこと

<sup>2</sup> 質問応答に PC などを利用する調査のこと

## 第4章 海外における補完推計に関する取組

図表3 ACS 調査におけるユニット非回答の理由分類

|                              |   |
|------------------------------|---|
| 1.拒否 (Refusal)               | 調査対象となった世帯・対象者から、調査への参加を拒否されたもの                           |
| 2.所在地不明 (Unable to locate)   | 指定された住所を調査員が見つけられなかった                                     |
| 3.不在 (No one home)           | 調査実施期間において、指定された世帯に誰も居住していなかった                            |
| 4.一時不在 (Temporarily absent)  | 調査実施期間において、休暇、出張、看病等により、不在であることが確認された                     |
| 5.言語問題 (Language problem)    | 言語障壁により調査を行うことができない、通訳を手配できない、スーパーバイザー等によっても調査を行うことができない等 |
| 6.データ不十分 (Insufficient data) | 最低限の回答が行われていない  |
| 7.その他                        | 稀ではあるが、疾病等による隔離や道路の通行止め等、上記に該当しないもの                       |

過去調査におけるユニット非回答理由(アメリカ全体)の割合は以下のとおり<sup>3</sup>であった。その多くは、拒否によるものであった。

図表4 ACS 調査におけるユニット非回答理由の分布

| 年    | 世帯票  |     |       |     |      |      |        |     |
|------|------|-----|-------|-----|------|------|--------|-----|
|      | 回収率  | 拒否  | 所在地不明 | 不在  | 一時不在 | 言語問題 | データ不十分 | その他 |
|      | %    | %   | %     | %   | %    | %    | %      | %   |
| 2011 | 97.6 | 1.1 | 0.0   | 0.6 | 0.1  | 0.0  | 0.2    | 0.4 |
| 2010 | 97.5 | 1.1 | 0.0   | 0.6 | 0.1  | 0.0  | 0.2    | 0.4 |
| 2009 | 98.0 | 0.8 | 0.0   | 0.6 | 0.1  | 0.0  | 0.2    | 0.3 |
| 2008 | 97.9 | 0.8 | 0.0   | 0.6 | 0.1  | 0.0  | 0.2    | 0.3 |
| 2007 | 97.7 | 0.9 | 0.2   | 0.5 | 0.1  | 0.0  | 0.4    | 0.2 |
| 2006 | 97.5 | 1.0 | 0.3   | 0.5 | 0.1  | 0.0  | 0.4    | 0.3 |
| 2005 | 97.3 | 1.0 | 0.4   | 0.5 | 0.1  | 0.0  | 0.3    | 0.3 |
| 2004 | 93.1 | 1.0 | 0.3   | 0.5 | 0.1  | 0.0  | 0.4    | 4.7 |
| 2003 | 96.7 | 1.7 | 0.3   | 0.6 | 0.1  | 0.0  | 0.3    | 0.2 |
| 2002 | 97.7 | 1.0 | 0.0   | 0.5 | 0.1  | 0.0  | 0.4    | 0.2 |
| 2001 | 96.7 | 1.3 | 0.0   | 0.7 | 0.1  | 0.0  | 0.7    | 0.4 |
| 2000 | 95.1 | 1.7 | 0.0   | 1.1 | 0.2  | 0.1  | 1.0    | 0.8 |

| 年    | 集合住宅票 |     |       |     |      |        |     |     |     |
|------|-------|-----|-------|-----|------|--------|-----|-----|-----|
|      | 回収率   | 個人  |       |     |      |        |     | 全体  |     |
|      |       | 拒否  | 所在地不明 | 不在  | 言語問題 | データ不十分 | その他 | 拒否  | その他 |
| %    | %     | %   | %     | %   | %    | %      | %   | %   |     |
| 2011 | 96.9  | 0.8 | 0.2   | 0.1 | 0.0  | 0.1    | 0.7 | 0.4 | 0.8 |
| 2010 | 97.6  | 0.9 | 0.2   | 0.1 | 0.0  | 0.2    | 0.4 | 0.1 | 0.5 |
| 2009 | 98.0  | 0.9 | 0.1   | 0.1 | 0.0  | 0.1    | 0.4 | 0.1 | 0.3 |
| 2008 | 98.0  | 0.5 | 0.1   | 0.1 | 0.0  | 0.1    | 1.0 | 0.2 | 0.0 |
| 2007 | 97.8  | 0.4 | 0.2   | 0.1 | 0.0  | 0.0    | 1.2 | 0.3 | 0.0 |
| 2006 | 97.4  | 0.8 | 0.2   | 0.1 | 0.0  | 0.1    | 0.6 | 0.5 | 0.2 |

<sup>3</sup> 回答率を含め、ユニット非回答の理由割合については、州別にも集計されており、下記 URL より参照が可能。  
[http://www.census.gov/acs/www/methodology/response\\_rates\\_data/](http://www.census.gov/acs/www/methodology/response_rates_data/)

② 項目非回答を防ぐ取組

ユニット非回答同様に、ACSではCATI、CAPIによる実査を通じて、項目非回答の防止を図っている。郵送によって返送された調査票については、論理矛盾等を確認し、必要に応じて電話による再確認を行うことで、項目非回答の発生を防いでいる。例としては、幼児が収入を回答する、15歳以下の回答者が結婚している、といった論理矛盾の修正が挙げられる。論理矛盾の修正に当たっては、矛盾する回答のどちらを修正するか、どのように修正するか、といった詳細なルール(マニュアルとして別途整備されている)を設定し、既に得られている情報(回答結果)を最大限に活用することを重視している。

(3) 補完手法概要

上記の取組を行ったにも関わらず、項目非回答が生じた場合には、以下の2つの方法による補完が行われている。

① アサインメント(Assignment)

アサインメントとは、欠測した値に対して、他の質問に対する回答に基づいて推測される回答を代入する方法。論理的に見て代入結果の確からしさが高いものと考えられている。

- 例； 回答者の名前に基づいて、欠測した“性別”を代入する
- 回答者の妊娠履歴に基づいて、欠測した“性別”を代入する

② アロケーション(Allocation)

アロケーションとは、世帯内あるいは類似のドナー(回答を代入するもとなる対象者)に基づいて、最近隣法等の統計学的な処理を含む手法の総称で、具体的にはホット・デック法が用いられている。

ACSでは、主に居住地に関する変数(州、郡、国勢調査ブロック等)をもとに、ホット・デック法のドナーを決定する。これは、特定の地域内において、ACSの質問項目(住居形態や人口属性等)が、より類似する傾向が見られたためである。

例えば、勤務地についてアロケーションを行う場合には、ドナーを以下の条件に沿って探し出す。

- 産業分類
- 勤務地への通勤手段
- 通勤時間
- 居住州
- 居住郡

なお、世帯票と集合住宅票ではプロセスが異なる。世帯票では、勤務地質問を除き、州ごとにホット・デック法を行っている。一方、集合住宅票では、同一の集合住宅タイプ(老人ホーム等)のデータについて、全国のデータを対象にドナーを探し出している。

センサス局では上記の補完について、個別の質問項目ごとにアロケーション比率(代入処理を行った件数比率)を公開することに加え、センサス局内の担当者が補完前後のデータを比較することで、データ品質を測定できるようにしている。またその結果は補完方法の見直しにも活用されている。アロケーション比率は「代入処理(アロケーション)されたケース数/各質問に回答すべきケース数」として定義されている。アロケーションが多く行われているのは、住宅の建築年や収入等、数値での回答が求められるものが多い傾向が見られている。なお、アサインメントやアロケーションの記録については、データ上でフラグ管理しており、どのような代入が行われたのかも把握できるようにされている。

第4章 海外における補完推計に関する取組

図表 5 ACS 調査におけるアロケーション比率<sup>4</sup>

|                 | 2011 年 |                     | 2011 年 |
|-----------------|--------|---------------------|--------|
|                 | %      |                     | %      |
| <b>住居全体</b>     | 5.2    | <b>教育</b>           |        |
| <b>個人全体</b>     | 5.8    | 過去 3 か月の就学有無        | 3.6    |
| <b>住宅特性</b>     |        | 学歴                  | 5.8    |
| 不在状況            | 2.9    | 就学先                 | 5.4    |
| 保有状況            | 1.0    | 学位分野                | 10.0   |
| 部屋構成            | 1.4    | <b>モビリティ</b>        |        |
| 引越し年            | 3.2    | モビリティ状況             | 4.0    |
| 引越し月            | 0.7    | 移民状況                | 8.0    |
| 建築年             | 16.7   | 移民郡                 | 9.2    |
| 面積              | 4.1    | 移民小区                | 8.6    |
| 農作物収入           | 4.3    | 移民場所                | 9.6    |
| 事務所有無           | 2.9    | <b>健康保険</b>         |        |
| 部屋数             | 5.2    | 勤務先からの健康保険          | 6.2    |
| 寝室数             | 4.3    | 個人での健康保険            | 7.1    |
| 水道              | 1.9    | Medicare での健康保険     | 5.1    |
| 水洗トイレ有無         | 1.9    | Medicaid での健康保険     | 7.9    |
| 浴室              | 2.0    | Tricare での健康保険      | 8.1    |
| 流し台             | 1.9    | VA での健康保険           | 8.0    |
| ストーブ            | 2.5    | 先住民医療サービスでの健康保険     | 8.5    |
| 冷蔵庫             | 2.7    | <b>障害</b>           |        |
| 電話              | 1.0    | 視覚障害                | 3.3    |
| 自動車             | 1.2    | 聴覚障害                | 3.0    |
| <b>住宅機器</b>     |        | 身体障害                | 3.4    |
| 暖房用燃料           | 3.2    | 記憶障害                | 3.4    |
| 電気代(月)          | 7.1    | 着衣時の不自由             | 3.4    |
| ガス代(月)          | 9.8    | 外出時の不自由             | 3.4    |
| 水道代(年間)         | 8.0    | <b>祖父母と子育て</b>      |        |
| その他燃料費(年)       | 10.4   | 孫の同居                | 0.9    |
| <b>住宅(給付)</b>   |        | 孫の世話                | 11.9   |
| フードスタンプ支給       | 1.2    | 孫の世話期間              | 12.8   |
| <b>不動産項目</b>    |        | 出産状況                | 3.7    |
| 固定資産税(年)        | 16.4   | <b>従軍関連</b>         |        |
| 不動産保険額(年)       | 23.7   | 従軍状況                | 3.8    |
| 抵当状況            | 1.9    | 従軍期間                | 6.3    |
| 住居費用(月)         | 10.5   | 障害等級                | 3.7    |
| 第二抵当            | 3.1    | 障害評価                | 0.6    |
| 住宅ローン利用有無       | 3.9    | <b>労働力</b>          |        |
| その他住居費用(月)      | 18.3   | 雇用状況                | 5.2    |
| <b>その他の経済状況</b> |        | 最後の就労時期             | 5.8    |
| 不動産価値           | 12.3   | 過去 12 か月の就労週数       | 6.9    |
| 移動式住宅費用(年)      | 19.5   | 週あたり就労時間            | 7.8    |
| コンドミニアム費用(月)    | 0.7    | 勤務州・国               | 6.5    |
| 賃貸料(月)          | 9.3    | 勤務郡                 | 7.2    |
| 食事有無            | 1.9    | 勤務郡区                | 2.4    |
| <b>個人関連</b>     |        | 勤務場所                | 7.8    |
| 人種              | 1.5    | <b>通勤</b>           |        |
| ヒスパニック系出自       | 1.8    | 通勤手段                | 5.6    |
| 性別              | 0.1    | 駐車場の規模              | 6.7    |
| 年齢              | 1.3    | 家を出る時刻              | 13.0   |
| 家族構成            | 1.1    | 通勤時間                | 9.6    |
| 婚姻状況            | 3.0    | <b>産業と職業</b>        |        |
| 過去 12 か月以内の婚姻有無 | 4.6    | 地位                  | 7.3    |
| 過去 12 か月以内の死別有無 | 4.5    | 産業                  | 7.8    |
| 過去 12 か月以内の離婚有無 | 4.5    | 職業                  | 8.2    |
| 結婚回数            | 4.9    | <b>収入</b>           |        |
| 最後の結婚年          | 11.4   | 賃金                  | 16.3   |
| <b>出自と言語</b>    |        | 自営業収入               | 5.9    |
| 出生地             | 6.2    | 金利収入                | 8.8    |
| 国籍              | 2.7    | 社会保障                | 8.8    |
| 帰化年             | 17.0   | その他保障               | 6.7    |
| 米国居住開始年         | 10.4   | 公的扶助                | 6.7    |
| 自宅内での他言語使用      | 3.3    | 退職手当                | 7.5    |
| 使用言語            | 5.6    | その他収入               | 7.3    |
| 英語能力            | 3.9    | 収入の一部あるいは全てのアロケーション | 22.6   |

<sup>4</sup> ACS 調査では全ての質問のそれぞれについて実施年ごとのアロケーション比率を公開している。過去の結果については、[http://www.census.gov/acs/www/methodology/item\\_allocation\\_rates\\_data/](http://www.census.gov/acs/www/methodology/item_allocation_rates_data/) を参照。

2) Population Census (US Census:国勢調査)における取組

(1) 調査概要

|      |  |
|------|--|
| 調査名  | :US Population Census  |
| 概要   | :アメリカにおける国勢調査として性別・生年月日・人種等を把握する調査   |
| 対象   | :世帯・個人   |
| 実施時期 | :10 年ごと<br>2010 年調査では 2010 年 4 月 1 日を基準日として 2010 年 3 月 15 日～17 日にかけて調査票を発送 |
| 調査手法 | :郵送調査・電話調査・訪問調査<br>郵送調査に未回答の場合には電話によるフォローアップを、さらに調査員の訪問によるフォローアップを行う       |
| 抽出方法 | :全数調査  |
| 配布数  | :1 億 3,170 万 4,730 住戸への郵送  |

Population Census はアメリカにおける下院議会の定数を決定するとともに、連邦政府から州政府への予算配分を決定するために用いられる等、最も重要な公的統計調査の 1 つである。

2000 年までは”Long-Form”と呼ばれる詳細な調査票を、約 300 万世帯を対象に配布・回収していたが、2010 年からは”Short-Form”と呼ばれる簡易な調査票のみで実施されている。なお、”Long-Form”の質問事項は ACS で質問されている。

質問は下記の 10 問に限定されており、基本的な属性を質問するにとどめている。

図表 6 Census における調査項目 (2010 年)

|                          |                     |
|--------------------------|---------------------|
| 1.2010 年 4 月 1 日時点での世帯人数 | 7.世帯主の年齢/生年月日       |
| 2.質問 1 以外の人数             | 8.世帯主のヒスパニック系出自分類   |
| 3.不動産所有状況                | 9.世帯主の人種            |
| 4.電話番号                   | 10.世帯主の一時転居有無       |
| 5.世帯主の氏名                 | 以降、世帯構成員について質問      |
| 6.世帯主の性別                 | 世帯主との続柄を含め、同様の質問を行う |

(2) 欠測を防ぐ取組

2000 年以前では、代入率のみを補足していたが、データ品質を高めるために、2000 年の Census から項目非回答に関する分析を行っている。その結果を利用しながら、欠測を防ぐ取組を行っている。

① ユニット非回答を防ぐ取組

Census においては、ユニット非回答を可能な限り発生させないためのフォローアップについて、NRFU (Nonresponse Followup)と呼ばれる取組を行っている。NRFU には、初期の郵送による調査票の回収ができなかった住戸から、可能な限り回答を得るための取組が含まれている。NRFU に対して 2000 年の Census では約 10 億ドル以上(1 回答あたり約 27ドル)の、2010 年の Census では約 15 億ドル以上(1 回答あたり約 34ドル)の実査費用が費やされており、その費用削減も図ることも目的とし、コストとのバランスを意識している。

2010 年の Census では、約 1 億 3 千万住戸のうち、約 3 分の 1 (33.5%)が郵送による回答を行わなかったため、約 4,700 万住戸が NRFU の対象として、電話あるいは調査員の訪問によるフォローアップを行った。ユニット非回答を防ぐため、調査員は 3 回の訪問と 3 回の電話によるコンタクトを行い、最低 6 回のコンタクトを取ることが指示されている。また、調査員の訪問時には、”Language Identification

## 第4章 海外における補完推計に関する取組

Flashcard”と呼ばれる回答者がどの言語を用いているかを判断するカード(指差し)を携行することで、言語障壁によるユニット非回答を防いでいる。カードには 51 の言語が記載されており、日本語では以下のように記されている。

図表 7 Census 2010 における Language Flash Card(日本語)

「こんにちは。私は米国勢調査局の係員です。こちらには英語を理解できこの調査にご協力いただける方がいらっしゃいますか？もしない場合は、あなたのお電話番号をお書きいただければ、日本語を話す係員が連絡をいたします。

調査員による訪問以前においても、事前の依頼状や再度の依頼状を送付する、対象者からよく聞かれる質問に対する回答集を準備する、といった取組を通じてユニット非回答の発生を防ぐ努力がなされている。

### ② 項目非回答を防ぐ取組

2010 年 Census では、それまで行われていた”Long-Form”調査票を ACS に移管し、質問数を絞りこんで行うことでも、項目非回答を防ごうとしている。また、調査項目ごとによく聞かれる質問に対する回答を準備することで、誤回答等を含めて項目非回答を防ごうとしている。

### (3) 補完手法概要

Census では、以下の 3 つの手法を組み合わせて、項目非回答に対する補完を行っている。

#### ① アサインメント(Assignment)

回答が欠測あるいは他の回答と整合していない、かつ、同一の回答者内で補完することができる場合に行われる。例えば、「人種」質問の欠測に対しては、「ヒスパニック系出自」に関する質問の回答に基づいて代入を行っている。なお、2010 年の Census では、「人種」質問に対して、2000 年の Census あるいは ACS に対する回答結果を基にした代入も行われており、過去調査あるいは類似調査の結果を活用している。

#### ② アロケーション(Allocation)

回答が欠測あるいは他の回答と整合していない、かつ、同一の回答者内で補完できない場合に行われる。世帯/集合住宅内の他の回答者の回答、あるいは、近隣の世帯/集合住宅内の他の回答者の回答に基づいて代入を行う。

#### ③ 代替(Substitution)

世帯内の全ての対象者に関する回答が欠測している場合、近隣の世帯(世帯人数は 6 人を上限とする)の回答に基づいて補完する。世帯の選択では、まず回答方法(郵送、訪問等)が同一の世帯を優先する等、回答方法を考慮した補完を行っている。

上記 3 つの手法の比率は以下のとおり。

## 第4章 海外における補完推計に関する取組

図表 8 Census 2010 における項目非回答に対する補完手法比率

|          | 個人          |     |             |               |     |     | 世帯<br>不動産<br>所有状況 |
|----------|-------------|-----|-------------|---------------|-----|-----|-------------------|
|          | 世帯主<br>との関係 | 性別  | 年齢<br>/生年月日 | ヒスパニック系<br>出自 | 人種  | %   |                   |
|          | %           | %   | %           | %             | %   | %   |                   |
| 項目非回答発生率 | 1.5         | 1.5 | 3.5         | 3.9           | 3.3 | 4.5 |                   |
| 補完あり     |             | 2.1 | 1.6         | 5.1           | 4.5 | 4.1 |                   |
|          | アサインメント     | 0.5 | 1.3         | 1.5           | 1.7 | 1.2 |                   |
|          | アロケーション     | 1.7 | 0.3         | 3.6           | 2.8 | 2.9 |                   |
| 代替       |             |     |             | 1.9           |     |     |                   |

センサス局では、個別の質問についてどのような補完を行ったかについても記録しており、各質問における記録は、以下のような分類に基づいている。

図表 9 Census2010 における質問別の補完内容の分類

| 質問項目(英語)          | (日本語)     | 変数名   | 分類            | 値  | 内容                                  |
|-------------------|-----------|-------|---------------|----|-------------------------------------|
| Tenure            | 不動産保有状況   | FTEN  | 補完なし          | 0  | 回答通り                                |
|                   |           |       | 補完あり(アサインメント) | 1  | 整合性チェックによる修正                        |
|                   |           |       |               | 4  | 非回答のため代入                            |
|                   |           |       | 補完あり(アロケーション) | 5  | 「自己保有/非保有」の両方に回答したため代入              |
|                   |           |       |               | 6  | 「賃貸/賃貸料なし」の両方に回答したため代入              |
|                   |           |       |               | 7  | 「自己保有/賃貸」の両方に回答したため代入               |
|                   |           |       |               | 0  | 回答通り                                |
| Age/Date of Birth | 年齢/生年月日   | FAGE  | 補完なし          | 1  | 年齢のみ回答通り                            |
|                   |           |       |               | 2  | 生年月日のみ回答通り                          |
|                   |           |       | 補完あり(アサインメント) | 3  | 「年齢/生年月日」が1年以上不整合のため修正              |
|                   |           |       |               | 8  | 「世帯主あるいは配偶者年齢」と「子供の数」の整合性チェックによる修正  |
|                   |           |       | 補完あり(アロケーション) | 4  | ホット・デック法による代入                       |
|                   |           |       |               | 7  | 代替値                                 |
|                   |           |       |               | 0  | 回答通り                                |
| Hispanic Origin   | ヒスパニック系出自 | FSPAN | 補完なし          | 00 | 回答通り                                |
|                   |           |       |               | 01 | 複数のヒスパニック系出自に回答: 回答の中から1つを選択        |
|                   |           |       |               | 02 | 複数のヒスパニック系出自に回答: 回答の中からランダムに選択      |
|                   |           |       |               | 03 | 複数のヒスパニック系出自に回答: 回答の中からランダムに選択      |
|                   |           |       |               | 04 | 複数の非ヒスパニック系出自に回答: 非ヒスパニックを選択        |
|                   |           |       | 補完あり(アサインメント) | 05 | 「人種」質問の回答に基づく修正                     |
|                   |           |       |               | 07 | 前回センサスの回答に基づく修正                     |
|                   |           |       | 補完あり(アロケーション) | 06 | 世帯員からの代入                            |
|                   |           |       |               | 08 | ホット・デック法による代入(スペイン系姓を利用)            |
|                   |           |       |               | 09 | ホット・デック法による代入(非スペイン系姓を利用)           |
|                   |           |       |               | 10 | ホット・デック法による代入(姓を利用しない)              |
|                   | 11        | 代替値   |               |    |                                     |
| Race              | 人種        | FRACE | 補完なし          | 0  | 回答通り                                |
|                   |           |       | 補完あり(アサインメント) | 1  | 整合性チェックによる修正                        |
|                   |           |       |               | 3  | 「ヒスパニック系出自」質問の回答に基づく修正              |
|                   |           |       |               | 9  | 前回センサスの回答に基づく修正                     |
|                   |           |       | 補完あり(アロケーション) | 4  | 世帯員からの代入                            |
|                   |           |       |               | 5  | ホット・デック法による代入                       |
|                   |           |       |               | 7  | 代替値                                 |
| Relationship      | 世帯主との関係   | FREL  | 補完なし          | 0  | 回答通り                                |
|                   |           |       | 補完あり(アサインメント) | 1  | 「世帯主と配偶者が同一の性別」: 配偶者の関係を未婚のパートナーに修正 |
|                   |           |       |               | 2  | 世帯における整合性チェックによる修正                  |
|                   |           |       |               | 3  | 集合住宅(GQ: Group Quarters)            |
|                   |           |       | 補完あり(アロケーション) | 4  | ホット・デック法による代入                       |
|                   |           |       |               | 5  | 整合性チェックによる代入                        |
|                   |           |       |               | 8  | 名簿に基づく代入                            |
|                   | 7         | 代替値   |               |    |                                     |
| Sex               | 性別        | FSEX  | 補完なし          | 0  | 回答通り                                |
|                   |           |       | 補完あり(アサインメント) | 1  | 名前による修正                             |
|                   |           |       |               | 2  | 世帯における整合性チェックによる修正                  |
|                   |           |       | 補完あり(アロケーション) | 4  | ホット・デック法による代入                       |
|                   |           |       |               | 5  | 整合性チェックによる代入                        |
|                   | 7         | 代替値   |               |    |                                     |

## 第4章 海外における補完推計に関する取組

### 2. アメリカ労働統計局(United States Bureau of Labor Statistics)

アメリカ労働統計局は、労働省の下で労働に関する統計情報を提供する機関である。ここでは、労働力に関する代表的な統計である OES(Occupational Employment Statistics)における補完に関する取組を整理する。

#### 1) OES(Occupational Employment Statistics: 職業雇用統計)における取組

##### (1) 調査概要

- 調査名 : Occupational Employment Survey
- 概要 : 各事業所について 12 の賃金区分ごとの従業員数を把握する調査
- 対象 : 事業所・企業
- 実施時期 : 毎年 5 月と 11 月に実施
- 調査手法 : 郵送調査・電話調査・訪問調査  
郵送による回答が中心となっている。
- 抽出方法 : 層化抽出  
アメリカに所在する約 670 万事業所から、州・産業及び所有形態に応じて抽出された約 20 万事業所が対象(従業員数では約 1.2 億人の労働者のうち約 7,800 万人をカバーしている。)
- 回収数 : 約 85 万事業所  
(2011 年 5 月調査では、回答率 77.3%(1,110,296 事業所のうち、858,474 事業所が回答)

OES は、半年ごとに約 20 万事業所を抽出し、3 年間のパネルを構築する調査である。例えば、2011 年 5 月調査は、2010 年 11 月、5 月、2009 年 11 月、5 月、2008 年 11 月に抽出されたパネルからの回答となる。いったんパネルとなった事業所は以降の 5 回分(2 年半)のパネルからは除外される。産業分類については、NAICS コード(North American Industry Classification System)に準拠している。

OES は、職業訓練に関する計画策定や、労働省による FLC(Foreign Labor Certification)が就労ビザ保有者に対して支払うべき賃金計算にも用いられており、下記 12 の賃金区分ごとの従業員数が質問される。

図表 10 OES における賃金区分

| 範囲 | 賃金区分        |                 |
|----|-------------|-----------------|
|    | 月次(ドル)      | 年次(ドル)          |
| A  | ～9.24       | ～19,240         |
| B  | 9.25～11.49  | 19,240～23,919   |
| C  | 11.50～14.49 | 23,920～30,159   |
| D  | 14.50～18.24 | 30,160～37,959   |
| E  | 18.25～22.74 | 37,960～47,319   |
| F  | 22.75～28.74 | 47,320～59,799   |
| G  | 28.75～35.99 | 59,800～74,879   |
| H  | 36.00～45.24 | 74,880～94,119   |
| I  | 45.25～56.99 | 94,120～118,559  |
| J  | 57.00～71.49 | 118,560～148,719 |
| K  | 71.50～89.99 | 148,720～187,199 |
| L  | 90.00～      | 187,200～        |

### (2) 欠測を防ぐ取組

OESでは、対象となった事業所に対して、初回は郵送によるコンタクトが行われる。コストの制約もあるため、一部ではあるものの、調査結果に大きな影響を及ぼすと考えられる大規模事業所を優先して、調査員による訪問も行われる。また、2回目以降の調査においても、約4週間ごとに督促依頼の手紙を最大3通送付することで、ユニット非回答の防止を図っている。さらに回答のなかった対象事業所については、電話あるいは電子メールによるフォローアップが行われる。

### (3) 補完手法概要

OESにおいては、類似した特性を持つ回答者の回答を代入するホット・デック法が用いられている。調査に回答しなかった事業所を”unit nonrespondents”、一部項目に回答しなかった事業所を”partial non respondents”と呼んでいる。本報告書では、他の調査との整合性を取るため、前者をユニット非回答、後者を項目非回答と名付ける。ユニット非回答に対してはステップ1とステップ2、項目非回答に対してはステップ2による代入が行われている。

#### ステップ1:

ユニット非回答の事業所の従業員構造について、最近隣法に基づくホット・デック法による代入を行う。事業所が所在する地域、産業及び従業員数が、ホット・デック法の基準に用いられている。一部の産業(病院及び教育機関)では、事業所の所有形態も基準に含めることで属性が類似するよう配慮している。

上記の基準に近いドナーが1つ以上存在する場合には、より新しい調査時点のデータが優先される。なお、上記の基準でドナーが見つからなければ、地域、産業の順に基準を緩和しながら、ドナーを探し出す。ドナーとして用いる回数には上限が設けられており、ドナーが特定された後、非回答の事業所の従業員構造に、ドナーの回答を代入する。

#### ステップ2:

ユニット非回答、あるいは項目非回答の事業所について、賃金ごとの従業員数分布を代入する。代入を行うに当たっては、MSA/BOS<sup>5</sup>による地域区分、産業コードであるNAICS5/6、企業規模が基準とされている。一部の産業(病院の及び教育機関)では、事業所の所有形態に基づいて代入が行われる。

代入を行う際には、十分なサンプル数が確保できることに留意し、ドナーの回答分布に基づいて欠測した従業員数分布を代入する。

<sup>5</sup> MSA(Metropolitan Statistical Area)とは、アメリカ合衆国大都市統計地域の意味で、都市部を指す BOS(Balance Of State)とは、郊外部を指す

## 第4章 海外における補完推計に関する取組

### 3. カナダ統計局(Statistics Canada)

カナダ統計局は、同国における統計調査の企画・実施を行う機関である。

原則として、カナダ統計局では、ユニット非回答に対してはウエイト付けが行われるが、事業所対象の調査においては、ユニット非回答に対しても代入が行われる場合がある。項目非回答に対しては、得られている回答を最大限に活用するために、欠測値にのみ代入を行うことを原則としている。なお、変数の1つ1つにウエイトをかける必要があるため、項目非回答に対してウエイト付けによる対応は行われていない。

ここでは、代表的な労働力に関する統計調査であるLFS(Labour Force Survey)における補完推計に関する取組を整理する。

#### 1) LFS(Labour Force Survey: 労働力調査)における取組

##### (1) 調査概要

|       |   |
|-------|---|
| 調査名   | : Labour Force Survey   |
| 概要・目的 | : 個人の就労状況を把握する調査  |
| 対象    | : 15歳以上の個人を対象   |
| 実施時期  | : 毎月15日を含む週を基準週として翌週に実査を行う  |
| 調査手法  | : CATI及び調査員による訪問面接(ノートPCを用いたCAPI)<br>電話番号が不明の場合には、初回は訪問した上で、残る5か月間についてはCATIを用いる。                |
| 抽出方法  | : 全国(10州に加えて3準州を含む)<br>層化抽出された世帯対象者が6か月間にわたって回答を行うパネル調査<br>(毎月新しいパネルと入れ替えるため、1/6の回答者は新しい対象者となる) |
| 回収数   | : 約5万4千世帯 約10万個人  |

LFSは、雇用・失業に関する統計調査である。データ収集後、13日後には結果が公表されており、月次の統計として迅速な情報提供が行われている。LFSの主な目的は、生産年齢にあたる人口を「雇用状態」、「失業状態」及び「非労働力人口」の3つに区分することで、労働力市場の状態を把握することにある。

一般的に知られている失業率を算出する基礎となるほか、就業者人口や雇用率といった指標を提供する調査である。加えて、産業別、職業別、公共部門、民間部門といった区分での就業者人口、労働時間等も提供している。

主な質問項目は、雇用・失業状況、労働時間、産業、職業、賃金等である。

##### (2) 欠測を防ぐ取組

LFSでは、欠測を防ぐために以下のような取組を行っている。

###### ① ユニット非回答を防ぐ取組

カナダでは、統計調査に対する回答義務が必須のものとして任意のものに分かれている。LFSに対する回答は必須に区分されているものの、調査に対する非協力(Person Non Response)が一定数存在する。

ユニット非回答に対しては、発生状況を理解できるよう、非回答の理由を以下のように区分して捕捉している。

図表 11 LFS におけるユニット非回答の理由分類

- |                             |
|-----------------------------|
| 1.拒否 (Refusal)              |
| 2.コンタクト不能 (No contact)      |
| 3.一時不在 (Temporary absence)  |
| 4.技術的問題 (Technical problem) |
| 5.その他 (Other reason)        |

LFSでは、回答率の向上を図るため、CAPI及びCATIの2種類の回答方法を提供している。対象者は6か月にわたって回答を行うこととなるため、初回のコンタクト時の協力依頼に注力し、対象者の状況(協力的、あるいは非協力的か)に応じて依頼状の内容を変えることでもユニット非回答を防ごうとしている。また、特定の回答者に固執することで過大な労力が発生しないよう、代理回答も認めることでユニット非回答を防いでいる。代理回答率は全体の65%程度である。

LFSに対する回答率は、2005年時点では約95%であった。直近のヒアリングでは90%程度で推移しているとのことであった。

図表 12 LFS に対する回答率 2005年

|                    | 平均(年) | 最低   | 最高   |
|--------------------|-------|------|------|
|                    | %     | %    | %    |
| カナダ全体              | 95.1  | 94.5 | 96.2 |
| ニューファンドランド・ラブラドール州 | 95.8  | 94.6 | 97.0 |
| プリンスエドワードアイランド州    | 96.5  | 95.2 | 97.6 |
| ノバスコシア州            | 93.7  | 92.7 | 95.4 |
| ニューブランズウィック州       | 95.4  | 94.6 | 96.9 |
| ケベック州              | 94.6  | 93.4 | 96.3 |
| オンタリオ州             | 95.2  | 94.3 | 96.3 |
| マニトバ州              | 96.4  | 94.6 | 97.9 |
| サスカチュワン州           | 96.4  | 95.4 | 97.6 |
| アルバータ州             | 95.1  | 93.7 | 96.9 |
| ブリティッシュコロンビア州      | 94.3  | 93.3 | 95.5 |

## ② 項目非回答を防ぐ取組

データ品質の向上を図るため、CATI及びCAPIによる調査票を利用し、論理矛盾や極端な外れ値を、実査の時点で防ぐように工夫が行われている。ただし、調査効率の観点から、そうした論理矛盾等をスキップすることも可能にしており、一定程度の項目非回答については、補完を行うことで対応を図っている。

## (3) 補完手法概要

上記の取組を経ても残った論理矛盾や欠測値に対しては代入が行われる。LFSでは3つの手法を組み合わせた補完が行われている。

### ① ホット・デック法

類似の属性の回答者の回答を、欠測値に対して代入する。なお、修正が発生した箇所については、調査員に対する周知が行われることで同様の修正が発生しないように注意が喚起される。

### ② キャリー・フォワード(Carry Forward)

欠測した月に対して、過去の回答を基にして代入する。

### ③ 推定による代入(Imputation by deduction)

欠測値に対して、他の回答結果から論理的に推定を行い、代入する。

## 第4章 海外における補完推計に関する取組

LFS では補完に関する 6 つのステップを通じて、上記 3 つの補完手法が組み合わされている。

### ステップ 1. 地域オフィスからのデータ受領と初期の補完

このステップでは主に「推定による代入」と「キャリア・フォワード」が行われる。前者は、欠測値を他の回答結果から推定できる場合に、後者は過去の回答結果が利用できる場合に用いられる。このステップでは、どのような補完がなされたかを記録するフラグ作成に加えて、後のステップで行われるホット・デック法が必要であると判断された箇所についても記録される。

### ステップ 2. 産業及び職業のコーディング

回答結果をコード化する。

### ステップ 3. 論理矛盾のチェック

回答結果の矛盾を除去するため、論理矛盾のチェックを行う。社会人口統計学的な変数については、ホット・デック法を行う前に代入される。

### ステップ 4. ホット・デック法による代入

実際にホット・デック法による代入が行われる前に、以下の事前作業が行われる。まず、回答者は 3 つのグループに分けられる(グループ A、B、C と名付けられる)。

グループ A: ホット・デック法のドナーになり得る回答者(全ての回答が記録されている回答者)

グループ B: グループ A に属さない回答者(ステップ 3 までに補完が行われた回答者)

グループ C: ホット・デック法を適用する回答者(グループ A・B 以外)

次に、グループ A に属する回答者の賃金に関するデータを全て時給に換算する。このプロセスを経ることで、ホット・デック法を行った際に労働時間と賃金総額が矛盾しないように留意している。

最後に、極端に高い、あるいは低い外れ値を特定し、欠測値と見なした上で、補完を行う対象とする。外れ値を回答した回答者は、ホット・デック法におけるドナーから除外し、グループ B に含める。外れ値の特定に当たっては、四分位法(quartile method)を用いている。基本的には、中央値からの距離を基準としており、ある一定の距離を超えた回答を外れ値と見なす。

上記の作業が行われた後、欠測値を代入する。代入の際には、ランダム・ホット・デック法を行う。ドナーを選定する際に用いる変数は以下の 18 変数である。この変数は、LFS の調査結果として最も重要な「就業状態」に対して、影響を及ぼす程度を分析した結果に応じて決められた。下記の優先度に従って、ホット・デック法のドナーを選定することで、「就業状態」をより良く補完できるように工夫されている。

図表 13 LFS におけるホット・デック法の適用基準

| 項目非回答の補完における優先順序 |                                     |
|------------------|-------------------------------------|
| 変数名              | 項目                                  |
| 1) TPATH         | 就業状態                                |
| 2) LMLFS         | 前月の就業状態(3 区分)                       |
| 3) COW           | Class of worker (給与所得者/自営業者/家族内従業員) |
| 4) OCC4          | 職業分類(4 区分)                          |
| 5) PROV          | 地域                                  |
| 6) AGE3P3        | 年齢区分(3 区分)                          |
| 7) ABQ1          | 原住民出生                               |
| 8) IMM           | 出生国                                 |
| 9) LMLFS7        | 前月の就業状態(7 区分)                       |
| 10) LMINDG       | 前月の産業分類                             |
| 11) MULTJOB      | 兼業有無                                |
| 12) AGE5P1       | 年齢区分(5 区分)                          |
| 13) SEX          | 性別                                  |
| 14) OCC10        | 職業分類(10 区分)                         |
| 15) AGE8P2       | 年齢区分(8 区分)                          |
| 16) STUD         | 学生区分(フルタイム学生/非フルタイム学生)              |
| 17) EDUC         | 高校卒業資格                              |
| 18) DWELRENT     | 不動産保有状況(持ち家/賃貸)                     |

| ユニット非回答の補完における優先順序 |                 |
|--------------------|-----------------|
| 変数名                | 項目              |
| 1) PROV            | 地域              |
| 2) AGE5P1          | 年齢区分(5 区分)      |
| 3) LMLFS7          | 前月の就業状態(7 区分)   |
| 4) SEX             | 性別              |
| 5) ABQ1            | 原住民出生           |
| 6) IMM             | 出生国             |
| 7) EIER            | 雇用保険地域区分(55 区分) |
| 8) EDUC            | 高校卒業資格          |

### ステップ 5.変数の作成

#### a)項目非回答に対する変数の作成

ホット・デック法を行うに当たっては、十分な数のドナーが存在することが重要であるが、LFS では下記 2 つの条件を設定している。

- ①ドナーの数は、代入される回答者数よりも多いこと
- ②それぞれの変数(上記図表)について、ドナーが最低 3 ケース存在すること

上記のいずれかが満たされない場合には、優先度の低い項目を除外して、ドナーを探すこととなる。最初は変数 1)から 18)まで全てを用い、次いで変数 1)から 17)まで、といったようにドナーを選定する条件を緩和していく。選定されたドナーから 1 つのドナーをランダムに決定し、論理矛盾等が生じない場合は代入を行い、矛盾等が生じる場合には他のドナーの結果を代入する、というプロセスを繰り返していく。

変数 1)から 5)まで緩和してもなお、ドナーが決定しない場合には、ユニット非回答として扱い、図表下の条件に従って改めてドナーを探すこととしている。

なお、LFS ではどのような補完を行ったかを記録することで、補完プロセスの管理や品質向上を図っている。

#### b)ユニット非回答に対する変数の作成

LFS は、6 か月間にわたるパネル調査であるため、ユニット非回答が生じた場合には、その回答者の前月の回答を代入することとなる。前月も回答されていない場合には、欠測値として代入は行わずに、ウエイト付けによる対応が行われている。

2005 年 1 月までは、ユニット非回答に対してはクロスセクションでのホット・デック法とキャリア・フォワードを組み合わせて行っていた。前月の回答が利用可能な場合に、社会人口統計学的な変数を含め全ての変数に対してキャリア・フォワードを行っていた。前月は非回答でも過去に回答している場合には、クロスセクションでのホット・デック法を行っていた。しかしながら、キャリア・フォワードでは月次変化を過小に推測しやすいこと、クロスセクションでのホット・デック法では過大に推測しやすいため、”longitudinal hot deck” (時系列ホット・デック法)を採用することとなった。なお、社会人口統計学的な変数については、時系列での変化は大きくないことが予想されるため、ホット・デック法による代入を行う前に、キャリア・フォワードによる代入を行うこととした。

### ステップ 6.ウエイト付けと最終クリーンマイクロデータ

上記のプロセスを経て最終クリーンマイクロデータが作成される。

補完を行った後には、LFS によるデータの一部を用いて、他の関連する統計調査との整合性を確認し、補完の精度についても見直しを行っている。例えば、SEPH (Survey of Employment, Payroll and Hours) や SLID (Survey of Labour Income and Dynamics)、雇用保険データやセンサス等の結果との比較がなされている。

**(コラム) 補完を行うプログラム Banff**

カナダ統計局では、補完作業を行うBanffと呼ばれるプログラムを2002年より独自に開発、運用している。

**Banffの概要**

Banffは、SAS上で動作する9つのプロセスを組み合わせた補完プログラムである。

1990年代にカナダ統計局で利用されていたGEIS (Generalized Edit and Imputation System)と呼ばれるデータ・エディティングと補完を行うプログラムを発展させ、より使いやすく、かつ、柔軟性のある仕組みにしたものである。主な違いは、GEISがOracle上で動作していたのに対して、BanffはSAS上で動作する点と、GEISが全てのモジュール(プログラム)を一連の流れとして利用しなければならないのに対して、Banffは個別のモジュールを独立して利用することができる点である。

一部のモジュールでの例外はあるものの、Banffが扱うことのできるデータは、数値データ、かつ連続変数で正の値を取るものに限られる。

補完作業において必要とされる機能について、Banffでは以下の9つのモジュールを備えている。

**1. エディットの定義**

回答内容の論理的な整合性をチェックするために、それぞれの変数がどのような関係にあるべきかを定義する。

**2. エディットの要約表の作成**

エディットの定義に従い、データの全体像(どの程度のエラーが発生しているのか)を把握する。Banffでは、それぞれの回答者について、以下の3つのエラー発生フラグを作成し、要約表を5つ作成する。

- エラーなし (Pass) : 整合性がとれている
- 欠測あり (Miss) : 欠測値があるために論理的な矛盾が生じている
- エラーあり (Fail) : 論理的な矛盾が生じている(欠測値はない)

- 1) エディット定義ごとのエラー発生件数
- 2) エディット定義ごとのエラー発生分布
- 3) エディット定義全体を通してのエラー発生件数
- 4) エラー発生箇所の検出回数
- 5) エディット定義全体を通してのエラー発生箇所の検出回数

上記の要約を行うことで、エディット定義が厳しすぎる(エラー発生件数が多すぎる)ために、定義を緩和すべきである、といった判断を行うことを可能としている。同時に、この後の補完作業(代入)がどの程度の負荷となるかの見通しを立てることも可能となる。

**3. 外れ値の検出**

このモジュールでは、外れ値の検出に用いられる理論の1つであるHidiroglou and Berthelot 法<sup>6</sup>に基づいて、ある変数における外れ値を回答者全体の中から検討する。

**4. 補完が必要な箇所の特定**

論理矛盾のチェック及び外れ値の検出を経て、補完が必要な箇所を特定する。このモジュールでは実際の修正や補完(代入)は行われない。特定に当たってのBanffの基本的な考え方は、変更の大きさを小さくするよりも、変更を加える箇所の数を可能な限り減らすことにある。背景には、既に得られているデータを可能な限り利用しようとする考えが存在する。

<sup>6</sup> 外れ値検出については、以下の文献に詳しい。

日本銀行ワーキングペーパーシリーズ「ビジネスサーベイにおける外れ値対応—全国企業短期経済観測調査(短観)のケース—」

### 5.一意に決定可能な箇所への代入

論理的に判断して、一意の値が代入できるものについては、このモジュールにおいて代入を行う。

### 6.ドナーによる代入

このモジュールでは、ホット・デック法に基づいた代入が行われる。Banffのユーザーは、ドナーを探し出す基準(変数)を指定することで、類似したドナーの回答を欠測値に代入することが可能となっている。

### 7.推定量による代入

平均値や中央値等による代入、あるいは線形回帰による代入が行われる。

### 8.変数の分解

個別の値の合計が、全体の合計と一致するよう、個別の値を分解する。全体の合計は正しいものと仮定して、その内訳を補完する。

### 9.マス代入

一部の回答者のみが回答する調査項目がある場合に用いる特殊なモジュールで、回答結果のないグループについての補完を行うものである。基本的にはホット・デック法に基づいて回答結果を代入するものである。

現状の Banff は主に事業所調査に用いられることが多い。今後の課題として、扱うことのできる変数を増やすことと、社会調査等への適用が検討されている。カナダ統計局では、Banff の利用者からのフィードバックを今後の機能向上に活用することを目指している。

#### 4. westat 社(調査会社)

westat 社はアメリカ連邦政府が管轄する統計調査の一部を受託する調査会社である<sup>7</sup>。メリーランド州の本拠を含め、アメリカ全土に6拠点を有する。アメリカには同社の他にも、ノースカロライナ州に所在する RTI 社<sup>8</sup>のような、公的統計調査の受託先として有力な企業が存在している。

westat 社の場合、連邦政府からの受託業務が98%程度を占めており、主な受託調査として NCES (National Center for Education Statistics: 国立教育統計センター)が行う NAEP (National Assessment of Educational Progress) や NCHS (National Center for Health Statistics: 国立健康統計センター)等が挙げられる。現在、westat 社内には60~70人程度の統計家(Statistician)が在籍しており、補完に関する取組も彼らが担い手となっている。

##### 1) ユニット非回答に対する対応・見解

統計調査における回答率を高く維持するためには、何らかのインセンティブが必要と westat 社では考えている。特に近年では、電話調査における回答率の低下傾向が見られており、他の調査手法(郵送等)も検討されている。ユニット非回答を防ぐためには、事前のコンタクト(挨拶状の送付等)が重要と考えており、謝礼の同封と合わせて同意を得るようにする、といった取組を行っている。また、目標以上の回収を行った場合には懸賞を提供する等、調査員に対するインセンティブを設定することもある。

アメリカにおける統計調査では、ユニット非回答に対しては、ウエイト付けによる対応が一般的である。westat 社では線形回帰モデルに基づく傾向スコア補正を行うことが多い。回答確率に応じて5つのグループ分けを行い、その平均回答傾向を計算するものである。なお、補完の結果については、他の統計データ(ACS等)と比較しながら、その精度の検証を行っている。

統計調査によっては、未回答や回答拒否の割合を捕捉し、重要な回答傾向として扱うこともある。

##### 2) 項目非回答に対する対応・見解

westat 社では、項目非回答を補完するよりも、その防止を重視している。CAPI や CATI による調査票によって、実査時のエラーを防いでいる。通常、項目非回答が発生する割合は、全体の1%程度に抑えることが可能となっている。

一般論ではあるが、社会調査に比べると、ビジネスサーベイの方が項目非回答を補完することが多いと考えられている。なお、必ずしも項目非回答に対して補完を行うとは限らず、仕様によって補完の有無、あるいは補完手法が定められている統計調査も、数は少ないが存在している。

また、調査票の回答しやすさを事前に確認することでも、項目非回答の防止を図っている。

##### 3) westat 社における補完システムについて

独自に開発を行った補完システム”Auto Impute”を社内で利用しており、調査票のロジックに基づいたチェックを行うとともに、回帰モデルによる欠測値の予測・補完や、ホット・デック法による補完を行うことが可能となっている。このシステムは後述するミシガン大学が開発した IVEware とは異なり、非正規分布の仮定においても対応が可能となっている。

補完を行うに当たっては、ユニット非回答、項目非回答のいずれにおいても MAR (Missing At Random: ランダムな欠測)を仮定することとしている。

<sup>7</sup> 企業ウェブサイト <http://www.westat.com/index.cfm>

<sup>8</sup> 企業ウェブサイト <http://www.rti.org/>

## 第4章 海外における補完推計に関する取組

### 5. ミシガン大学

ミシガン大学は ISR (Institute for Social Research: 社会調査研究所) を有し、PSID (Panel Study of Income Dynamics: 収入動態に関するパネル調査) や HRS (Health and Retirement Study: 健康と退職に関する調査)<sup>9</sup>等の社会調査を行っている。ミシガン大学では、IVEware (Imputation and Variance Estimation Software) と呼ばれる補完プログラムを開発し、無償で配布<sup>10</sup>している。補完推計に関する取組や見解についてアカデミックの視点からも整理を行った。

#### 1) IVEware について

IVEware は、社会調査に用いるために ISR からの要求によって 1995 年から開発が始められ、1997 年には初期のバージョンが完成した。当初は ISR が資金を提供していたが、NIH (National Institute of Health: 国立衛生研究所) からの資金援助も受けた。有料で販売することも検討されたが、ミシガン大学の研究成果として無料で公開を行っている。SAS 機能を利用するもののほか、SAS のライセンスが比較的高額であるために、単独動作が可能なパッケージも用意している。

補完の手法としては、多重代入法の考え方に基づいて、回帰モデル (sequential regression: 逐次回帰) が用いられている。個人の収入等の数量値に加えて、カテゴリカルデータにも対応している。主な機能は以下の 6 つのモジュールである。

図表 14 IVEware の主要モジュール

|                       |                                      |
|-----------------------|--------------------------------------|
| IMPUTE<br>(代入)        | 逐次回帰による代入を行う                         |
| DESCRIBE<br>(記述)      | 母数の平均、分布等を推定する                       |
| REGRESS<br>(回帰)       | 直線回帰や対数回帰等のモデルに基づいて代入を行う             |
| SASMOD<br>(SAS 機能)    | SAS 機能を利用する(主に共分散構造分析や対数線形モデル分析等の機能) |
| SYNTHESIZE<br>(データ合成) | 逐次回帰による合成データを作成する                    |
| COMBINE<br>(データ結合)    | 複数のデータソースからデータを結合し、多重代入法を適用する        |

#### 2) 補完推計に関する見解

補完推計の手法には様々なものがあるが、調査の内容やデータ量、統計家の能力によってもふさわしい補完手法は異なる、との意見が聞かれた。他のデータ (登録データ等) から情報を補完することも重要であり、一概にどの手法がふさわしいのかを論ずることは難しいと考えられている。センサス局ではホット・デック法が 40 年以上にわたって用いられているが、ドナーの数が不足する場合もあること、ホット・デック法ではパラメトリックではない、といった課題も指摘された。

しかしながら、MAR (Missing At Random) を仮定することになるものの、補完推計を行うことでバイアスを減少させ、調査効率の向上 (可能な限り多くのサンプルと変数を活用し、欠測値を放置しない) を図ることが可能であると考えられている。補完を行うことの必要性を明らかにすることに加えて、手法の妥当性の検討が重要な事項として指摘されていた。

<sup>9</sup> 2 つの調査の概要については平成 23 年度内閣府大臣官房統計委員会担当室請負調査「日本におけるパネルデータの整備に関する調査」報告書を参照 [http://www5.cao.go.jp/statistics/nenpou/nenpou\\_1206.html](http://www5.cao.go.jp/statistics/nenpou/nenpou_1206.html)

<sup>10</sup> IVEware の入手が下記 URL より可能 <http://www.isr.umich.edu/src/smp/ive/>

## **第5章 災害時における補完推計などの対応**

本研究調査では、災害等によって通常の調査実施が困難となる場合に、どのような対応がなされているのかについて整理を行った。国内の事例として、2011年3月に発生した東日本大震災に対する労働力調査における対応を整理した。海外の事例として、2005年8月に発生したハリケーン・カトリーナに対するACS (American Community Survey: アメリカ地域社会調査)、CES (Current Employment Survey: 雇用状況調査) 及び、CPS (Current Population Survey: 就業状態調査)における対応を整理した。

なお、海外の事例を整理する中で、災害等に対する事前の備えの重要性が指摘されたため、カナダ統計局におけるBCP (Business Continuity Plan: 事業継続計画)や異常時の状況把握を迅速に行うための特殊モジュールの設定についても整理を行うこととした。

第5章 災害時における補完推計などの対応

**I 国内・海外における大規模災害への対応(一覧)**

| 国        | 日本   | アメリカ  | アメリカ  |
|----------|--|---|---|
| 取組主体     | 総務省統計局   | US Bureau of Labor Statistics   | US Bureau of Labor Statistics   |
| 調査概要     | 統計調査名(英語)  | Current Population Survey (CPS)   | Current Employment Survey (CES)   |
|          | (日本語)  | 労働力調査   | 就業状態調査  |
|          | 調査対象   | 世帯・個人   | 世帯・個人   |
|          | 実施頻度   | 毎月  | 毎月  |
| 災害への対応   | 災害名  | 東日本大震災  | ハリケーン・カトリーナ   |
|          | 発生時期   | 2011年3月   | 2005年9月   |
|          | 実査時の対応   | <ul style="list-style-type: none"> <li>被災3県(岩手県、宮城県及び福島県)を除く全国結果を集計・公表した</li> <li>岩手県及び宮城県は5月から、福島県は8月から調査を再開し、各県において調査地域を順次拡大した</li> <li>2011年9月分結果から被災3県を含む全国結果の集計・公表を再開した</li> </ul> | <ul style="list-style-type: none"> <li>避難所やホテル、モーテルなどでの調査は行わなかったが、対象世帯に対しては、可能な限りコンタクトを取ることとした</li> <li>避難地域に指定された地域についてはコンタクトは行わなかった</li> <li>就業状態の判定基準に変更は行わなかった</li> </ul> |
| 適用した補完手法 | <ul style="list-style-type: none"> <li>2011年3月～8月の各月を被災3県分について推計し、既公表の被災3県を除く全国の結果数値に加算することで、被災3県を含む全国結果を算出</li> <li>就業者数は、他の複数のデータから求めた対前年同月減少率と同程度の就業者数の減少があったものと仮定して、就業者数の対前年同月減少数を推計</li> <li>完全失業者は、概念的に近い職業安定所(ハローワーク)の「有効求職者数」をもとに作成した回帰式により推計</li> </ul> | <ul style="list-style-type: none"> <li>回答を得ることができた世帯に対して、より重いウエイト付けを行うことで対応を行った</li> <li>公式な見解として、アメリカ全体での推計に対しては、無視可能なレベルであるとされている</li> </ul>  | <ul style="list-style-type: none"> <li>雇用の喪失を過小に推定しないよう、被害の影響が特に大きい地域に所在する事業所については、原則として事業は行われておらず、雇用数もゼロと見なした</li> <li>2005年11月からは通常と同様の補完手法を行った</li> </ul>                  |
| 災害に対する備え | -  | -   | -   |

注) 労働力調査との比較のため本報告書の掲載順序と異なる

第5章 災害時における補完推計などの対応

| 国        |               | アメリカ  | カナダ  | カナダ               |
|----------|---------------|---|--|-------------------|
| 取組主体     |               | US Census Bureau  | Statistics Canada  | Statistics Canada |
| 調査概要     | 統計調査名<br>(英語) | American Community Survey<br>(ACS)  | Labour Force Survey<br>(LFS)   | -                 |
|          | (日本語)         | アメリカ地域社会調査  | 労働力調査  | -                 |
|          | 調査対象          | 世帯・個人   | 世帯・個人  | -                 |
|          | 実施頻度          | 毎年 ※実査は毎月実施   | 毎月   | -                 |
| 災害への対応   | 災害名           | ハリケーン・カトリーナ   | -  | -                 |
|          | 発生時期          | 2005年9月   | -  | -                 |
|          | 実査時の対応        | <ul style="list-style-type: none"> <li>被害の影響が特に大きい地域では、調査票の郵送を取りやめた</li> <li>通常行われる、電話によるフォローアップも一部地域では行わなかった</li> <li>2005年11月以降は通常と同様の調査を行った</li> </ul>   | -  | -                 |
|          | 適用した補完手法      | <ul style="list-style-type: none"> <li>ユニット非回答、及び項目非回答の発生率を把握した上で、通常時に行っているユニット非回答に対するウエイト付けによる対応方法を変更</li> <li>避難地域に指定された地域に所在する住所を全てウエイト付けの対象から除外<br/>(回答を得ることができたケースに対して、より重いウエイトを付ける)</li> </ul>  | -  | -                 |
| 災害に対する備え | -             | <ul style="list-style-type: none"> <li>異常時の状況把握を迅速に行うために、2つの特殊モジュールを調査票に組み込む               <ol style="list-style-type: none"> <li>追加項目 総選挙での投票行動やその理由など、通常は起こらない特殊なイベントの影響を把握するもの</li> <li>災害による影響項目 事故や災害などの異常を想定したモジュール</li> </ol> </li> <li>CAPIあるいはCATIによる調査票として、画面に表示されるようにしている</li> </ul> | BCP(Business Continuity Plan)について <ul style="list-style-type: none"> <li>統計調査ごとのBCPを策定中</li> <li>災害が発生した場合に備えて、どのような影響が生じるかを事前に分析する</li> <li>カレンダー形式とすることで、複数の統計調査を横断的に管理できるように工夫を行っている</li> </ul> |                   |

## II 国内における東日本大震災への対応事例—労働力調査

### 1) 調査の概要

労働力調査は、我が国における就業及び不就業の状態を毎月明らかにすることを目的として実施される調査である。概要は以下のとおり。

図表 1 労働力調査の概要

| 項目     | 内容   |
|--------|--|
| 実施頻度   | 毎月   |
| 調査対象   | 全国の世帯・個人<br>※外国政府の外交使節団、領事機関の構成員(随員を含む)及びその家族、外国軍隊の軍人・軍属(その家族を含む。)は除外する      |
| 標本抽出方法 | 層化2段抽出法<br>国勢調査の約100万調査区から約2,900調査区を選定し、その調査区内から選定                           |
| 標本数    | 約40,000世帯に居住する15歳以上約100,000人   |
| 調査方法   | 留置調査<br>・基礎調査票:2年にわたり同一の2ヶ月を調査<br>・特定調査票:2年目2ヶ月目のみ調査                         |
| 調査項目   | ・基礎調査票:就業状態、所属事業所の名称、仕事の種類、求職の理由、等<br>・特定調査票:雇用形態、求職活動の方法・期間、就業希望の有無、前職の種類、等 |

### 2) 推計の経緯・目的

労働力調査では、東日本大震災の影響により、岩手県、宮城県及び福島県(以下「被災3県」)において調査の実施が一時困難となったため、2011年3月以降において被災3県を除く全国結果を集計・公表する等の対応が実施された。この後、岩手県及び宮城県は5月から、福島県は8月から調査が再開され、各県において調査地域が順次拡大し、9月分結果から、被災3県を含む全国結果の集計・公表を再開した。

これについて、「平成22年度統計法施行状況に関する審議結果報告書」(2011年9月22日統計委員会)においては、「今後の取り組むべき統計整備等の方向性」として、「全国を対象とする基幹統計調査等については、調査対象地域の一部を除外する等の特別の取扱いをした場合、被災地の状況を踏まえて可能な限り補完的、補足的な調査や推計を行う等の措置を講ずる必要がある」とされた。

こうしたことなどから、長期時系列比較や2012年における前月比較等に資する参考値を示すことを目的として、労働力調査の補完推計が実施された。

### 3) 推計の体制・時期

補完推計の検討に当たっては、外部の有識者にも意見を伺いながら進められた。

補完推計の結果については、2012年4月24日に公表されたほか、2012年9月9日から12日に行われた「2012年度統計関連学会連合大会」の一般講演報告においても報告が行われた。

4) 手法の概要

本補完推計では、まず、2011年3月～8月の各月を被災3県分について推計し、その結果を既公表の被災3県を除く全国の結果数値に加算することで、被災3県を含む全国結果を算出している。

その際の被災3県分の推計は、一定の仮定の下、利用可能な他の関連するデータを基に、可能な限り実態を反映するように推計している。また、被災3県の各県ごとに、『就業者数』と『完全失業者数』の総数をそれぞれ別に推計し、非労働力人口は15歳以上人口からこれらを差し引いて計算している。

(1) 就業者数の推計

労働力調査では、就業者について、従業者(調査期間(月末1週間)中に1時間以上仕事をした者)と休業者(仕事を持ちながら一時的に仕事に従事しなかった者)を合わせて人数を把握している。本推計においては、就業に関して利用が可能な他の複数のデータを用い、これらから求めた対前年同月減少率と同程度の就業者数の減少があったものと仮定して、就業者数の対前年同月減少数を推計している。具体的な推計式は以下のとおりである。

図表2 就業者数の推計式

| 業種    | 推計式  |
|-------|--|
| 非農漁業  | $\begin{aligned} & \text{(非農漁業就業者の対前年同月減少数)} \\ & = \text{(前年同月の非農漁業就業者数)} \\ & \quad \times \text{(雇用保険被保険者数の対前年同月減少率)} \times \text{(減少率幅の調整率)} \\ & \text{(減少率幅の調整率)} = \frac{\text{(毎月勤労統計調査等に基づく同期間の対前年減少率)}}{\text{(雇用保険被保険者数の同期間の対前年減少率)}} \end{aligned}$ |
| 農業・漁業 | $\begin{aligned} & \text{(農業・漁業就業者の対前年同月減少数)} \\ & = \text{(前年同月の農業・漁業就業者数)} \times \text{(農業・漁業を再開していない割合)} \end{aligned}$  |

なお、使用された主な外部データは、以下のとおりである。

- ・ 雇用保険事業月報(厚生労働省)に基づく雇用保険被保険者数
- ・ 毎月勤労統計調査(厚生労働省)に基づく常用労働者数(パート含む)
- ・ 農林水産省が実施した農業・漁業の被害・再開状況の確認結果

## 第5章 災害時における補完推計などの対応

### (2) 完全失業者数の推計

労働力調査では、完全失業者について、調査期間(月末1週間)中、①仕事に就いていない、②仕事があればすぐ就くことができる、③求職活動をしていた、の3つを要件としてその人数を把握している。本補完推計においては、完全失業者と概念的に近いデータとして、公共職業安定所(ハローワーク)の『有効求職者数』が用いられており、このデータを基に作成した回帰式により、完全失業者数が推計されている。具体的な推計式は以下のとおりである。

$$\Delta \text{完全失業者数}_t = \alpha + \beta \times \Delta \text{有効求職者数}_t + \sum \gamma_i d_{it}$$

- ・  $t$ は四半期。回帰式の作成には2008年(平成20年)1-3月期～2010年(平成22年)10-12月期のデータを使用
- ・ 完全失業者数は労働力調査都道府県別結果(モデル推計値)、有効求職者数は一般職業紹介状況(職業安定業務統計)(厚生労働省)を使用
- ・  $\Delta \text{完全失業者数}_t = (\text{t期の完全失業者数}) - (\text{t-1期の完全失業者数})$
- ・  $\Delta \text{有効求職者数}_t = (\text{t期の有効求職者数}) - (\text{t-1期の有効求職者数})$
- ・  $d_i$ は当該四半期を1とするダミー変数( $i:1=1-3$ 月期、 $2=4-6$ 月期、 $3=7-9$ 月期)
- ・ 推計値は2010年(平成22年)10-12月期を起点として逐次的に四半期値を算出、月次の値は有効求職者数の月次の動きで調整

なお、本補完推計では、個票レベルでの代入というよりは、マクロ的な推計が行われている。

### 5) 留意点

補完推計値は2012年3月～8月の前年同月比較や長期時系列比較の参考値として公表されており、補完推計値の利用上の注意として、以下の3点が挙げられている。

- ・ 補完推計値は、利用が可能な他の関連するデータを用いて一定の仮定の下に推計したものである
- ・ 補完推計は就業者数及び完全失業者数の総数のみ推計しており、非労働力人口は15歳以上人口からこれらを差し引いて計算している
- ・ 内訳となる各項目についても、被災3県の総数の対前年同月増減を内訳に按分した後、被災3県を除く全国結果に加算する方法により算出している

### Ⅲ 海外における大規模災害への対応・取組

本調査研究では、大規模災害によって通常の調査の実施が困難となった場合に、どのような対策が行われているかについて、海外における対応・取組を整理することとした。

#### 1. アメリカセンサス局(United States Census Bureau)

ここではアメリカセンサス局が管轄する統計調査のうち、ACS(American Community Survey)を取り上げる。

##### 1) ACS(American Community Survey: アメリカ地域社会調査)における対応・取組

2005年8月24日に発生したハリケーン・カトリーナでは、メキシコ湾沿岸に位置するルイジアナ州、ミシシッピ州、アラバマ州における被害が深刻であった。避難命令が出された地域では調査を行うことが不可能となった。そのため、通常の補完とは異なる方法を行うこととなった。

###### (1) ACSにおける対応の概要

ハリケーン・カトリーナによる被害が大きかった4州(アラバマ州、ルイジアナ州、ミシシッピ州、テキサス州)における一部の郡では、アメリカ全体での回答率は通常と同程度であったものの、ユニット非回答の発生率が高まった。ただし、項目非回答の発生率については通常時と大きな変動は見られなかった。

ACSではユニット非回答に対応するために、通常のウエイト付けの方法に変更を加えることとした。項目非回答に対する補完では通常と同じ方法を取ることとした。通常は、調査員による訪問の結果、住所が見つからない、空き家であると判断された住所を除外した上で、ウエイトを付けるが、ハリケーン・カトリーナによる影響で避難地域に指定された住所についても、Administrative Noninterview(行政上の理由による調査不能)と分類した上で、ウエイト付けの対象から除外することとした。

###### (2) 実査上の影響と対応

###### 2005年9月(被災翌月)

被災地の中でも特に影響が大きかった地域については、センサス局では調査票の郵送を取りやめることとした。また、通常の電話によるフォローアップについても一部地域では行わなかった。調査員による訪問についても一部地域では行わなかった。

なお、センサス局では、地域の特定にあたって、FEMA(Federal Emergency Management Agency: 連邦緊急事態管理庁)が指定した、深刻な罹災地域の定義を利用することとした。

###### 2005年10月(被災2か月後)

被害の大きい地域の一部に対しては、調査票の郵送を通常タイミングよりも遅らせた。電話によるフォローアップは10月下旬以降から通常どおり行うこととした。調査員による訪問は可能な限り行うこととし、11月以降は通常時と同様に調査を行った。

## 第5章 災害時における補完推計などの対応

### 2. アメリカ労働統計局(United States Bureau of Labor Statistics)

ここでは、ハリケーン・カトリーナによる被害に対する、CES(Current Employment Survey)及びCPS(Current Population Survey)の2つの統計調査の対応を取り上げる。

#### 1) CES(Current Employment Survey: 雇用状況調査)における対応・取組

ハリケーン・カトリーナによる被害が発生する中で、災害規模の把握や政策策定に資する情報を提供するために、CES 調査における雇用数、労働時間と賃金に関する正確な情報を提供することが求められた。

ハリケーンによる避難や建造物への被害等により、BLS が実施する各種調査においては、世帯あるいは事業所へのコンタクトができなくなった。特に、ハリケーンによる被害範囲が大きく、かつ、被害が長期にわたるものであったため、欠測値を補完する通常のプロセスでは補正しきれないと判断し、2005年9月と10月のCESでは個別の補完を行うこととなった。これはCESの歴史上、初めて補完プロセスを変更した例となっている。

##### (1) CES における対応の概要

2005年9月のCESにおいて行われた補完は、9月時点の雇用状況を可能な限り正確に反映するために、通常の補完から以下のような変更が行われた。

- ・被災地域においては、事業が一時的な休業あるいは廃業しているかの状況判断を踏まえた上で、ユニット非回答に対する代入プロセスを変更した(ユニット非回答に対しては前月の回答に基づいた代入が行われるが、被災地域ではそうした通常の代入を行わなかった)。
- ・通常よりも回収率が低かったため、ウェイト付けを変更した。
- ・”Birth/Death Model”(新規事業の開業時には雇用が創出されたとし、廃業時には雇用が喪失されたと仮定するモデル)については、被害を受けた事業所から報告されたゼロ(雇用なしの状態・以降、“ゼロ回答”という)を反映させたが、アメリカ全体の推計では変更せずに、州あるいは大都市での変更を行った。州あるいは大都市における被害をより精緻に反映するためである。

2005年10月も同様の変更を行ったものの、2005年11月からは通常の補完方法を行った。

Birth/Death Model については、2つの修正が加えられた。

1点目は、通常の推計においては、ゼロ回答は推計算出から除外されるが、代入された回答のうち、ゼロと代入されたもの(以降、“ゼロ代入”という)も含めることとした。これは、通常時では調査時点でリアルタイムでは反映できない新規雇用創出と喪失が互いに打ち消されることが多いものの、9月における被災地では新規雇用創出が少ないと想定されたためである。9月の推計には111事業所が、10月には57事業所がゼロ代入として含まれることとなった。

2点目は、Birth/Death Model において用いられる時系列モデルを使用しないこととした。なお、BLSでは上記の修正を行った場合においても、アメリカ全体における推計に対する影響は無視可能なレベルと判断している。

上記の修正は、ゼロ回答が通常の水準に戻っていることを確認した上で、11月以降は行われていない。

(2) 実査上の影響と対応

2005年9月（被災翌月）

CES調査の対象となった事業所に対しては、可能な限りコンタクトをとることとした。しかしながら、被災3州のうち、ルイジアナ州とミシシッピ州では通常よりも低い回収率となった。アラバマ州ではほぼ通常と同等の回収率であった。

図表3 2005年9月のCESの回収率

|        | 州全体 | 被害の深刻な地域(郡) |
|--------|-----|-------------|
| ルイジアナ州 | 57% | 50%         |
| ミシシッピ州 | 62% | 53%         |

事業所を複数有する大企業が、被災3州以外の本拠地から報告を行うケースも全体の約3分の1程度存在している。回収率を可能な限り高めるため、通常よりも電話によるフォローアップに注力することとした。

被災地域の中でも、特に被害が甚大な箇所については、9月12日時点で給与支払いが行われていることが判明した場合を除いて、事業は行われておらず、雇用数もゼロであると扱うこととした。その結果、通常時には回答が行われるものの、9月の調査に回答しなかった被害が甚大な地域に所在する82の事業所がゼロと見なされることとなった。

仮に、通常の補完手法(ユニット非回答に対するもの)を適用すると、回答が得られている他の事業所の結果に、月次変化を加味した補完が行われるが、この手法では、被害が大きな地域での正確性に欠け、雇用の喪失を過小に推定することが危惧された。BLSでは、地域の特定に当たって、FEMA (Federal Emergency Management Agency: 連邦緊急事態管理庁)が指定した、深刻な罹災地域の定義を利用することとした。

ウエイト付けによる対応も行われ、1,260の事業所に対して通常よりも30%程度高いウエイトを付加することとした。

2005年10月（被災2か月後）

依然として回収率は、通常よりも低かったものの、徐々にその差が小さくなった。11月には被災3州と全国での差も小さくなった。電話によるフォローアップも続けられた。

9月の時点で雇用数がゼロと扱われた事業所については、10月の回答がない限り引き続きゼロとして扱うこととした。結果として71事業者が10月もゼロとして扱われた。

## 第5章 災害時における補完推計などの対応

### 2) CPS (Current Population Survey: 就業状態調査) における対応・取組

CPSは、世帯内の16歳以上の個人を対象に、就業状態を把握するための統計である。毎月、全国の約6万世帯を対象とする調査を行っており、主な指標として、失業率や失業理由等が含まれる。毎月の12日が含まれる週が基準週とされている。したがって2005年8月の調査についてはハリケーン・カトリーナによる影響はなかったものの、9月の調査において以下のような対応が取られることとなった。

#### (1) CPS における対応の概要

- ・避難命令が出された地域を除いて、対象となった世帯には可能な限りコンタクトを取ることとした。
- ・避難先から帰宅した世帯が少なく、被災地域における非回答率が高かったが、基準週に滞在していた避難所やホテル、モーテル等での調査は行われなかった。
- ・ただし基準週(2005年9月11日～17日)における活動に応じて、就業状態を判定することとした。判定に当たっては、通常の判定基準を適用することとした。

図表4 CPS における就業状態の判定基準

|      |  |
|------|--|
| 雇用者  | 基準週において給与を受ける仕事に就いている、あるいは天候による理由で勤務できていない場合を含め仕事に就いている者。例えば、勤務先の指示で一時的に休業状態の場合も雇用者としてカウントされる。基準週における給与支払いが行われているかは問わない。 |
| 失業者  | 基準週において給与を受ける仕事に就いていない、かつ、調査の前4週間に求職活動をした/仕事に就ける状態にあった者。この基準を踏襲し、被災地域において後片付け等のために仕事に就ける状態になかった者は失業者として扱われた。             |
| 非労働者 | 雇用者でも失業者でもない者。被災後は、求職活動も困難となったため、被災地域では多くの対象者が非労働者として区分された。  |

#### (2) 実査上の影響と対応

##### 2005年9月 (被災翌月)

ハリケーンによる被害が発生した地域についても、対象となった世帯について電話あるいは訪問によるコンタクトを試みることとした。ただし、ルイジアナ州のオーリンズ郡・ジェファーソン郡では、調査時期に避難命令が出されていたためにコンタクトは行われなかった。

2005年9月の調査では、被災地域においては8月に比べると回答数が600世帯程度、減少した。2005年8月と比較した減少割合は以下のとおり。なお、同時期に発生したハリケーン・リタによる避難が行われたテキサス州においても10%程度の減少が生じた。

##### 2005年10月 (被災2か月後)

2005年10月からは全ての地域について対象世帯へのコンタクトを試みた。8月に比べると被災州では250世帯ほどの減少が見られたが、9月からは回収率の改善が見られた。なお、テキサス州では2%程度の減少にとどまった。

なお10月から、世帯構成員の中で、ハリケーン・カトリーナによる被害で避難をしている構成員がいるかどうかを質問することとした。この質問を行うことで、世帯構成員が調査時点で避難先から自宅に戻っているかを確認できるようにした。

2005年11月以降（被災3か月後以降）

被災州についても、徐々に回収率が改善し始めた。11月にはテキサス州、フロリダ州での回収数が通常の水準に戻った。

図表5 2005年9月以降のCPSの回収率(8月と比較した減少割合)

|        | 州全体(対2005年8月) |      |      |      |         |      |      |
|--------|---------------|------|------|------|---------|------|------|
|        | 9月            | 10月  | 11月  | 12月  | 2006年1月 | 2月   | 3月   |
| ルイジアナ州 | -36%          | -26% | -21% | -20% | -17%    | -17% | -16% |
| ミシシッピ州 | -13%          | -10% | -9%  | -14% | -10%    | -11% | -16% |

CPSでは、回答が得られない状況に対して、回答を得ることができた世帯に、重いウエイトをかけることで全体の推計を行うこととした。2005年11月以降、州を越えた避難者数を反映させた推計を行うこととしたが、アメリカ全体での推計に対する影響は無視可能であるとの見解を示している。

## 第5章 災害時における補完推計などの対応

### 3. カナダ統計局(Statistics Canada)

カナダにおいては、広大な国土を有する一方で、人口密度が極めて低いため、人々の生活に影響を及ぼすような大規模な自然災害は少なく、人口密度の低い地域で山火事の発生が挙げられる程度である。しかしながら、1989年、2006年に大規模な停電が発生し、生活に影響を及ぼした例があることから、カナダ統計局では、そうした異常事態に備えた様々な工夫を行っている。特に、CAPIを用いる調査においては、停電によってノートPCが利用できなくなるため、実査を行うことが不可能になること、また、データを保存するサーバーも利用できなくなる、といったインフラ面を意識した対策を図っている。

#### 1) 統計調査ごとのBCP(Business Continuity Plan: 事業継続計画)の策定

2012年からカナダ統計局では、災害等の異常時に備えるためにBCPの策定を進めている。BCPには、以下のような要素が含まれている。

図表6 カナダ統計局におけるBCPに関する取組の構成要素

|          |   |
|----------|---|
| ガバナンス    | ・非常時においては、誰が何をやるべきかを明確にし、責任を与えることが重要である、という考え方に基づいて、個々の職員の責任を定義する。  |
| 影響の分析    | ・各統計調査について、中断が生じることによる影響を想定する。<br>・品質維持の許容レベルを決める。  |
| 運営の継続    | ・非常時における前提条件(例: 主な責任を担う職員と連絡を取れる、等)を設定し、最低許容レベルを維持するための要件を、以下のような視点から整理する。<br>-IT インフラ(サーバー等)<br>-コミュニケーション<br>-オフィス環境<br>-紙の資料 等 |
| 対応計画     | ・BCPを実行するための基準を設定する。<br>・BCPの進行状況の把握方法、報告の流れ等を、フローチャートを用いて図示する。   |
| 再開のための計画 | ・統計調査の中断後に、正常な状態に戻すために必要な行動計画を策定する。   |
| 準備状況     | ・BCPについての訓練状況、テスト方法等を整理する。<br>・BCP自体の保存方法(媒体や保存場所等)やアクセス権限等を設定する。<br>・訓練の結果からの改善点や更新頻度を整理する。                                      |
| 懸案事項     | ・カナダ統計局におけるSCC (Security Coordination Committee:安全管理委員会)からのフィードバックや懸念事項への対応。  |

上記以外にも、主要な職員の連絡先情報や重要な資料・資材のリストといった、業務遂行に必要となるものをBCPに含めている。SCCからの評価を向上させるべく、現在は改善を重ねている段階にあるが、早期の完成を目指している。

さらに、カナダにおける多数の統計調査の状況を同時に確認できるようにするため、BCC(Business Continuity Calendar: 事業継続カレンダー)と呼ばれる、カレンダー形式のツールを作成しようとしている。災害や事故は予測不可能であり、発生した場合には、様々な箇所障害が生じることを前提としている。BCCでは、以下の4つの視点から、相互の依存性(影響範囲、程度、内容)を整理することで非常時への対応を図るものである。

図表7 カナダ統計局におけるBCCの視点

- |   |
|---|
| <ul style="list-style-type: none"> <li>①施設(庁舎、倉庫等)</li> <li>②IT インフラ(CATI、CAPIのデータサーバー等)</li> <li>③職員(調査員を含めた人員)</li> <li>④外部データ(住民登録データ等)</li> </ul> |
|---|

上記の視点に基づいて、各日ごとの作業内容及び4つの視点の相互依存性を整理していく。カレンダー形式とすることで、複数の統計調査について、最終的な結果の公表に向けての残日数や影響の範囲を横断的に把握できるよう工夫を行っている。

図表8 カナダ統計局におけるBCCのイメージ

|     | 統計調査A |         |             |     |       | 統計調査B |                 |             |     |     |
|-----|-------|---------|-------------|-----|-------|-------|-----------------|-------------|-----|-----|
|     | フェーズ  | 施設      | IT インフラ     | 職員  | 外部データ | フェーズ  | 施設              | IT インフラ     | 職員  | ... |
| 1日目 | 実査中   | 統計局 x 階 | サーバー<br>XYZ | 〇〇課 | -     | 集計中   | データセ<br>ンター y 階 | サーバー<br>PQR | ××課 | -   |
| 2日目 | 実査中   | ↓       | ↓           | ↓   |       | 集計中   |                 |             |     |     |
| 3日目 | 実査中   | ↓       | ↓           | ↓   |       | 集計中   |                 |             |     |     |
| ... |       | ↓       | ↓           | ↓   |       |       |                 |             |     |     |

ヒアリング時点(2012年12月)では、主に統計調査における実査オペレーションを中心とする検討がなされていたものの、今後の課題としてより細かい作業レベル(例として、実査の進捗状況管理や集計作業)でのBCC策定も検討されているとのことであった。

## 2) 異常時の状況把握を迅速に行うための特殊モジュールの設定

ここでは具体的な統計調査の例としてLFSを取り上げる。

LFSの調査票には2つの特殊なモジュールが含まれている。これらのモジュールは通常は使用されないもので、何らかのきっかけが生じた場合、CAPIあるいはCATIによる調査票として表示されるものである。 CAPI、CATIを調査に利用していることで、モジュールの追加を容易に行うことが可能となっている。

### 1.追加項目 (Additional Content Component)

このモジュールでは、通常は起こらない特殊なイベントの影響を測定するもので、総選挙での投票行動やその理由等が含まれる。CAPIあるいはCATIを利用しているため、ある特定の州にのみ表示する、といった管理も可能となっている。

### 2.災害による影響項目 (Diastrophic Catastrophe Effect Module)

主に事故や災害等の異常時を想定したモジュールである。例えば、新型インフルエンザ(H1N1型)発生時にも利用された。

上記のモジュールをあらかじめ作成し、準備することで、異常時の状況把握、それによる影響等を迅速に捉えられるような仕組みを有している。

第6章 統計調査に欠測値補完を適用する際の示唆・課題

**第6章 統計調査に欠測値補完を適用する際の示唆・課題**

前章までの国内外における欠測値補完に関する取組の整理及び研究会における有識者のディスカッションを通じて、今後の統計調査に欠測値補完を適用する際の示唆・課題について、以下のように「1. 平常時における補完」と「2. 災害時における補完などの対応」に分けた整理を行った。

～統計調査に欠測値補完を適用する際の示唆・課題 概要一覧～

| 1. 平常時における補完          |  |
|-----------------------|--|
| 1) 補完推計をなぜ行うのか        | <ul style="list-style-type: none"> <li>✓ 補完推計は、統計調査の品質を高めるための様々な取組の1つとして位置づけられる。</li> <li>(留意すべき点)               <ul style="list-style-type: none"> <li>－ 補完推計自体にはその限界があることを認識した上で、<u>予期可能なバイアスを減らす</u>とする試みとして捉える必要がある。</li> <li>－ 統計調査の結果に含まれる欠測値に対して、<u>周辺情報の利用やベンチマークとなる数値との比較等</u>を通じて可能な限り真の値に近いと考えられる値を推計し、代入することによって、利用者が調査結果を分析する際にできるだけ不便が生じないように留意する必要がある。</li> </ul> </li> </ul>   |
| 2) 補完推計を誰が行うのか        | <ul style="list-style-type: none"> <li>✓ 補完推計を行う際には、統計学に精通した「<u>統計家(Statistician)</u>」による研究・分析・支援が必要とされる。</li> <li>✓ 一方で、学術的な視点に過度に偏らないよう、<u>個別統計調査ごとに統計家を配置</u>する等、統計調査の事情も考慮する必要がある。</li> <li>(留意すべき点)               <ul style="list-style-type: none"> <li>－ 調査実施の際に得られる周辺情報をできるだけ活用するため、また、<u>統一的な補完推計結果を全ての利用者が共通して用いられる</u>ようにするため、統計作成者が行うことが必要とされる。</li> <li>－ なお、統計利用者において補完推計を行うことも可能ではあるが、利用できる補助情報の制約や、利用者ごとに補完の結果が異なる等の問題が生じやすい。</li> </ul> </li> </ul> |
| 3) 補完推計をどのように行うのか     | <ul style="list-style-type: none"> <li>✓ 補完手法に関する研究はアカデミック分野においても行われているが、統計調査においては比較的わかりやすい概念であり、以前から行われてきた<u>ホット・デック法</u>が用いられることが多い。</li> <li>✓ また、ソフトウェアの普及、進歩に伴い、統計調査に求められる効率性・即時性に対応するためにソフトウェアを活用した補完推計が行われている。</li> <li>✓ どの補完手法を用いるかは、<u>個別統計調査の特性を踏まえた検討</u>が必要となる。</li> <li>(留意すべき点)               <ul style="list-style-type: none"> <li>－ ソフトウェアの活用にあたっては、その背後にある理論を理解した統計家との協働も必要とされる。</li> </ul> </li> </ul>   |
| 4) 補完推計の内容をどのように伝えるのか | <ul style="list-style-type: none"> <li>✓ 補完推計の限界の理解と、<u>よりよい補完推計を行うための議論に必要な内容</u>が公開されている。</li> <li>✓ 統計家の支援のもと、補完推計を行うためにどのような情報が必要となるかを検討することが重要である。</li> <li>(留意すべき点)               <ul style="list-style-type: none"> <li>－ 真の値が不明であること、また多くの手法が <u>MAR(Missing at random)</u>を前提としているが、補完手法の選択も含めて補完推計における正解はないことを統計利用者が認識できるよう工夫が必要である。</li> </ul> </li> </ul>  |
| 2. 災害時における補完などの対応     |  |
| 1) どのように対応するのか        | <ul style="list-style-type: none"> <li>✓ 災害時においては、<u>可能な限り調査を行うことが前提</u>とされている。</li> <li>✓ やむを得ず調査を実施することができない場合、<u>ウエイト付けによる対応</u>が行われている。</li> </ul>   |
| 2) どのように備えるのか         | <ul style="list-style-type: none"> <li>✓ 個別統計調査ごとに、災害時の対応について平常時から<u>の検討・研究</u>が行われている。</li> <li>✓ <u>迅速に災害の状況や影響に関する情報を収集</u>するために、調査票上に予め、被災状況等を把握するための調査項目を特殊モジュールとして用意しておくといった工夫も行われている。</li> <li>✓ 平常時において行われる、調査の存在感を高める取組も、災害時における調査協力意欲を高める効果が期待される。</li> </ul>  |

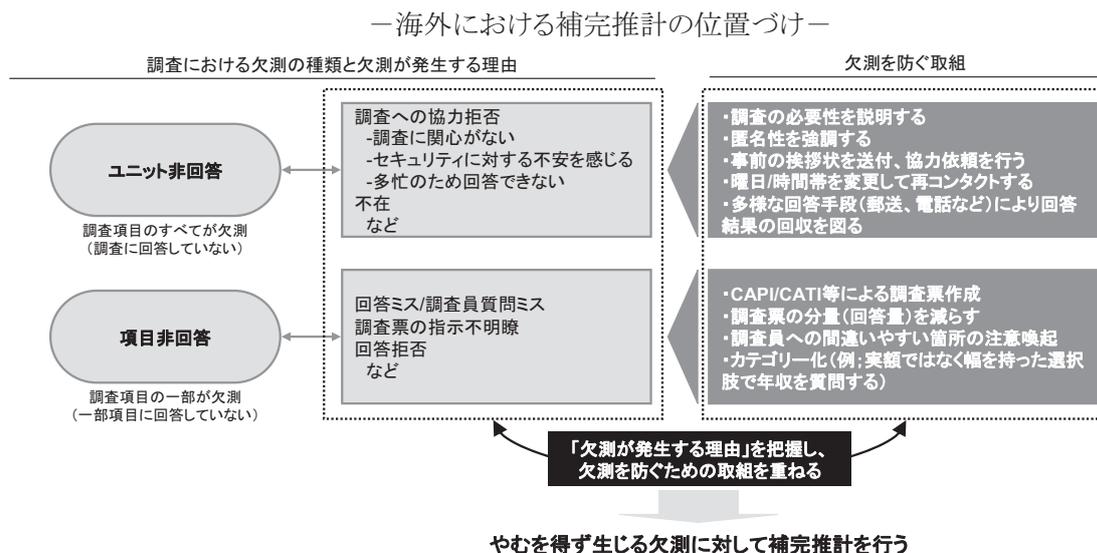
次ページ以降で、上記の示唆・課題のそれぞれについて、海外事例を踏まえた整理を行っている。本調査研究で対象とした海外事例は、アメリカ及びカナダにおいて補完推計を行っている統計調査の一部である。

## 1. 平常時における補完に対する示唆・課題

平常時においては、「補完推計をなぜ行うのか」、「補完推計を誰が担うのか」、「補完推計をどのように行うのか」及び「補完推計の内容をどのように伝えるのか」を検討することの重要性が挙げられる。

### 1) 補完推計を行う目的と背景の明確化（補完推計をなぜ行うのか）

補完推計を行っている海外事例においても、欠測値補完を行うよりも前に、ユニット非回答の防止（回収率の向上等）や項目非回答の防止（わかりやすい調査票の作成等）に注力しており、補完推計は統計調査の品質を高めるための取組の1つとして位置づけられている。



海外事例における項目非回答に対する取組では、CAPI(Computer-Aided Personal Interview)やCATI(Computer-Aided Telephone Interview)の活用により工夫が図られており、調査票のロジックに反する回答については一定の防止効果が得られていると考えられる。これらの手法は、面接調査や電話調査の場合に利用されるものであり、日本では留置法による調査が主流であることから、同様の方法をとることは必ずしも容易ではない。

一方で、ユニット非回答を減らすために多様な回答手法を用いて結果の回収を図る、といった回収率向上のための取組が行われている。統計調査の品質を向上させるため、回収率向上のための各種の取組を行いながらも、その両輪となる取組として補完推計が位置づけられている。

そのため、補完推計自体にはその限界があることを認識した上で、予期可能なバイアスについて、そのバイアスを減らそうとする試みである、と捉える必要がある。また、欠測値に対して周辺情報の利用やベンチマークとなる数値との比較を通じて、可能な限り真の値に近いと考えられる値を推計し、統計利用者に不便が生じないように留意する必要がある。

### 2) 補完推計の担い手となる統計家との協働（補完推計を誰が担うのか）

海外事例を通じて、補完推計を行うに当たって、統計学に精通した専門家である「統計家」(Statistician)による研究、分析、支援が行われていることが確認された。調査の設計・企画時に加えて、収集後のデータ検証においても統計家との協働が行われている。例えば、統計調査を受託する調査会社であるwestat社においても、統計学科を修士レベルで卒業したStatisticianの採用が行われており、60～70人程度（総スタッフ数は2000人以上）が雇用されている。

## 第6章 統計調査に欠測値補完を適用する際の示唆・課題

－westat 社における Statistician の募集内容の例－

| 職種名                 | 応募要件  |
|---------------------|---|
| Statistician        | <ul style="list-style-type: none"> <li>・サンプルサイズの計算や空間解析、ベイジアンモデル等に精通していることが望ましい。</li> <li>・統計学(Statistics)あるいは生物統計学(Biostatistics)において修士号を取得しており、公的機関の調査における3年以上の実務経験を有すること。</li> </ul>                                   |
| Senior Statistician | <ul style="list-style-type: none"> <li>・サンプルサイズの計算や空間解析、ベイジアンモデル等に精通していることが望ましい。</li> <li>・統計学(Statistics)あるいは生物統計学(Biostatistics)において修士号を取得しており、公的機関の調査における10年以上の実務経験を有すること。</li> <li>・PhD(博士号保持者)であることが望ましい。</li> </ul> |

統計家との協働を行う中では、理論的な概念に過度に偏らないよう、個別統計調査ごとに統計家を配置することで、各調査の事情を反映した支援が行われるような配慮も行われている。統計調査の品質を維持・向上させるべく、こうした統計家との協働が図られている。

海外事例の1つとして取り上げたBLSでは、毎年1~2つの統計調査を選定し、統計学の専門家によるレビューが行われている。加えて、外部の専門家に対しても定期的にレビューを依頼し、補完手法に関するアドバイスを受けている。アメリカにおいては主にCNSTAT(Committee on National Statistics: 国家統計委員会)と呼ばれる組織が、統計関連部局に対するレビューを提供している。

－統計局にレビューを提供する団体CNSTAT(Committee on National Statistics)の概要－

### CNSTATの概要

NAS(National Academy of Sciences: 米国科学アカデミー)の一部として1972年に設立された組織である。CNSTATでは、統計調査における手法を研究し向上させることを目的として、様々な研究活動が行われている。また、連邦政府が実施する統計調査の評価を通じて、統計学と政策決定の橋渡しを行っている。活動予算はNSF(National Science Foundation: アメリカ国立科学財団)によって支援されている。

構成委員には、ASA(American Statistical Association: アメリカ統計学会)の理事・委員等が含まれており、統計学の専門家で組織されている。

(URL) <http://sites.nationalacademies.org/DBASSE/CNSTAT/index.htm>

本調査研究において調査対象とした海外事例では、調査実施の際に得られる周辺情報を活用しながら、統計作成者である統計機関が補完推計を行っている。例えば、ACSにおいては以下のようにユニット非回答の理由を分類して把握している。

－ACS調査におけるユニット非回答の理由分類－

1. 拒否(Refusal)  
調査対象となった世帯・対象者から、調査への参加を拒否されたもの
2. 所在地不明(Unable to locate)  
指定された住所を調査員が見つけられなかった
3. 不在(No One Home)  
調査実施期間において、指定された世帯に誰も居住していなかった
4. 一時不在(Temporarily absent)  
調査実施期間において、休暇、出張、看病等により、不在であることが確認された
5. 言語問題(Language problem)  
言語障壁により調査を行うことができない、通訳を手配できない、スーパーバイザー等によっても調査を行うことができない等
6. データ不十分(Insufficient data)  
最低限の回答が行われていない
7. その他  
稀ではあるが、疾病等による隔離や道路の通行止め等、上記に該当しないもの

3) 統計調査に求められる効率性・即時性に対応するための工夫（補完推計をどのように行うのか）

欠測値補完を適用するには調査の特性を踏まえた上で、どの手法を用いるのか、という点についても検討が必要とされる。海外事例からは補完推計を行っている調査の特性として以下の点が指摘される。

－海外において補完推計を行っている調査の特性－

| 調査における6つの要素        | 海外事例調査からの結果  |
|--------------------|--|
| 対象者<br>(事業所・世帯・個人) | <ul style="list-style-type: none"> <li>✓今行行った海外事例調査からは、回答者が<b>事業所あるいは世帯・個人のいずれにおいても補完</b>が行われている。</li> </ul>   |
| 調査項目               | <ul style="list-style-type: none"> <li>✓態度質問ではない属性質問(企業売上高や就業状態など)では、<b>他属性との相関が高く、補完による推計も合理的</b>と見なされていた。(世論調査のような態度を聴取するものについては補完はふさわしくないとする意見も聞かれた。)</li> <li>✓調査事項ではないが、回答状況(拒否、不在、移転などの別)などを補完の参考情報として収集している。</li> </ul> |
| 規模・抽出方法            | <ul style="list-style-type: none"> <li>✓調査規模の大小については、補完の確からしさが高まる<b>大規模調査での補完</b>が行われている。</li> <li>✓同一回答者の前回調査時点での回答を確認できることから、<b>補完が行いやすいパネル調査</b>において補完が行われていることも確認された。</li> </ul>   |
| 実施頻度               | <ul style="list-style-type: none"> <li>✓頻度が高い月次統計から数年に1回の統計まで、<b>頻度の高低に関わらず補完</b>が行われている事例が存在する。</li> </ul>  |
| 調査手法               | <ul style="list-style-type: none"> <li>✓CAPI、CATIなどを用いた調査では、回答が得られないために実査効率を低めることがないように、調査員に対してある程度の自由度(回答しなくても次の質問に進める、など)を持たせている。</li> <li>✓郵送・電話・訪問調査といった、<b>調査手法を問わずに補完</b>が行われている。</li> </ul>                             |
| 回収率                | <ul style="list-style-type: none"> <li>✓<b>実査上の工夫によっても、なお回収率の向上が見込み難い</b>調査において補完を行っている。</li> </ul>  |

本調査研究において調査対象とした海外事例では、結果的に**補完手法として比較的わかりやすい概念であるホット・デック法が用いられていた。**

補完手法を検討する際には、なぜその補完手法を用いるのか、という理論的な背景を理解することも重要である。海外事例からは、他の補完手法に比べてシンプルなホット・デック法を適用する場合であっても、その手法の限界や課題を認識した上で補完推計を行うことで、理論的な裏付けをもって統計調査を実施することが重視されている。カナダ統計局におけるLFSの事例では、LFSの調査目的である就業状態に対する各変数の影響度を分析することでホット・デック法の基準が決定される等、理論的な裏付けを意識した運用が行われている。

## 第6章 統計調査に欠測値補完を適用する際の示唆・課題

### －カナダ統計局 LFS におけるホット・デック法の適用基準－

| 項目非回答の補完における優先順序 |                           |
|------------------|---------------------------|
| 変数名              | 項目                        |
| 1) TPATH         | 就業状態                      |
| 2) LMLFS         | 前月の就業状態(3 区分)             |
| 3) COW           | 従業上の地位(給与所得者/自営業者/家族内従業員) |
| 4) OCC4          | 職業分類(4 区分)                |
| 5) PROV          | 地域                        |
| 6) AGE3P3        | 年齢区分(3 区分)                |
| 7) ABQ1          | 原住民出生                     |
| 8) IMM           | 出生国                       |
| 9) LMLFS7        | 前月の就業状態(7 区分)             |
| 10) LMINDG       | 前月の産業分類                   |
| 11) MULTJOB      | 兼業有無                      |
| 12) AGE3P1       | 年齢区分(5 区分)                |
| 13) SEX          | 性別                        |
| 14) OCC10        | 職業分類(10 区分)               |
| 15) AGE3P2       | 年齢区分(8 区分)                |
| 16) STUD         | 学生区分(フルタイム学生/非フルタイム学生)    |
| 17) EDUC         | 高校卒業資格                    |
| 18) DWELRENT     | 不動産保有状況(持ち家/賃貸)           |

| ユニット非回答の補完における優先順序 |                 |
|--------------------|-----------------|
| 変数名                | 項目              |
| 1) PROV            | 地域              |
| 2) AGE3P1          | 年齢区分(5 区分)      |
| 3) LMLFS7          | 前月の就業状態(7 区分)   |
| 4) SEX             | 性別              |
| 5) ABQ1            | 原住民出生           |
| 6) IMM             | 出生国             |
| 7) EIER            | 雇用保険地域区分(55 区分) |
| 8) EDUC            | 高校卒業資格          |

補完推計を行っている海外事例においては、補完推計に対して多くのリソースをかけることは難しいという認識のもと、プログラム等を活用しながら、統計調査に求められる効率性・即時性への対応が図られている。プログラムの活用にあたっては、その背後にある理論を理解している統計家との協働も必要とされている。統計ソフトウェアの発達も進んでおり、補完推計を行うソフトウェアとして以下のようなものが利用可能である。カナダ統計局における Banff、westat 社における社内利用を前提とした補完推計プログラムも開発されている。

### －補完推計に用いられるソフトウェア(パッケージ)の例－

|           |  |
|-----------|--|
| Amelia II | <ul style="list-style-type: none"> <li>多重代入法(EMB アルゴリズム)を実行することができる。</li> <li>左記 URL より入手できる。 <a href="http://gking.harvard.edu/amelia">http://gking.harvard.edu/amelia</a></li> </ul>                       |
| R         | <ul style="list-style-type: none"> <li>MICE (Multivariate Imputation by Chained Equation) によって多重代入法を実行することができる。</li> </ul>   |
| SPSS      | <ul style="list-style-type: none"> <li>Missing Value パッケージによってマルコフ連鎖モンテカルロ法、回帰モデル等の多重代入法が利用できる。</li> </ul>   |
| SAS       | <ul style="list-style-type: none"> <li>Proc MI によって下記 3 つの手法が利用できる。 <ul style="list-style-type: none"> <li>傾向スコア(Propensity score method) による代入</li> <li>回帰モデル</li> <li>マルコフ連鎖モンテカルロ法</li> </ul> </li> </ul> |
| Banff     | <ul style="list-style-type: none"> <li>カナダ統計局によって開発・提供されている(4 章に記載)。</li> <li>SAS 上で動作する。</li> </ul>   |
| IVEware   | <ul style="list-style-type: none"> <li>ミシガン大学による開発・提供されている(4 章に記載)。</li> <li>SAS 上での動作に限らず、単独でも動作する。</li> </ul>  |

4) よりよい補完を検討するために必要な情報内容の検討(補完推計の内容をどのように伝えるのか)

本調査研究で調査対象とした海外事例では、例えば、アメリカにおける ACS 調査や Census 等で、アロケーション比率、ホット・デック法におけるドナーの選定基準・方法等の情報が整理され、どの程度補完が行われているのか、また、どの程度の曖昧さが残されているのか、といった補完推計の限界や、補完による効果(バイアス傾向の理解等、補完を行うことへの価値)等について検討されており、一部については公開されている。

ー補完推計を行っている統計調査(海外事例)において示されている内容の例ー

|                         |
|-------------------------|
| (1)行ったこと                |
| ・補完プロセス                 |
| ・補完手法                   |
| ・ホット・デック法によるドナーの選定基準    |
| (2)補完を行う前の状態            |
| ・ユニット非回答を防ぐ取組内容         |
| ・項目非回答を防ぐ取組内容           |
| ・回収率                    |
| ・ユニット非回答の発生理由           |
| (3)補完を行った結果             |
| ・補完方法分類(フラグ条件)          |
| ・代入による補完の発生率(アロケーション比率) |

補完推計を行う場合には、どのような欠測が生じているのか、なぜ欠測が生じているのか、といった欠測に関する情報を分析することが必要となる。統計家との協働のもと、どのような情報が必要となるかを検討した上で、補完推計の適用に備えることが求められる。なお、これらの情報の公開については、実査への影響等も考慮した上で、慎重に検討する必要がある。

欠測値に対する真の値は不明であることに加え、多くの手法が欠測の発生メカニズムとして MAR を前提としているが、補完手法の選択も含めて補完推計における正解はないことが海外事例においても指摘されている。上記の点を統計利用者が認識できるよう工夫を行うことも必要である。

## 第6章 統計調査に欠測値補完を適用する際の示唆・課題

### 2. 災害時における補完に対する示唆・課題

災害時については、災害が起こった際に「どのように対応するのか」、について調査を行ったが、災害が起こる前に「どのように備えるのか」という点が重要であるとの示唆が得られた。災害時における統計調査には、災害による被害と復興の進捗を把握する(政策効果を確認する)ことが求められるため、事前に備えることの重要性が指摘された。

#### 1) 迅速な調査の再開とウエイト付けを用いた対応 (どのように対応するのか)

アメリカにおけるハリケーン・カトリーナによる統計調査に対する影響では、災害が発生して3か月後には通常の調査体制に戻る等、迅速な調査の再開が図られている。被災地の事情(避難命令地域の指定等)や回収率等を把握することで、災害が調査に及ぼす影響の分析を行いながら調査の再開が目指されている。例えば、アメリカにおけるCESの事例では、廃業によるものなのか、災害対応によって回答できないのか、といった非回答理由の特定の困難さが指摘されている。推計を行う際にもBLS担当者と州担当者との間で調整が行われる等、各州の事情を踏まえながら推計が行われている。

また、災害発生後には、BLS内部や州政府機関の代表者との会合も行いながら、補完方法を含めた対応について検討がなされた。分散型統治機構を有するアメリカにおいては、各部局が自立しているが、ホワイトハウス内に各部局間のコーディネーションを行うグループが置かれており、各統計局からの代表による会合も行われる等、部局間の連携にも留意されていた。

調査を行うことができない場合には、ウエイト付けによる対応が行われている。その過程には、「正しいと思われる数値を出す方法を求める」、あるいは「正しいと考えられる方法で算出される数値を求める」という考え方について議論が存在しており、通常時と異なるウエイト付けの方法を用いることの是非が検討された。

災害時の統計調査については、その調査結果に基づいて、どの程度の被害が生じたか、どのような復興・支援が行われたかを示すことも重要と考えられる。災害の発生は、雇用喪失につながることから、海外事例では労働分野における統計調査における対応が見られたが、雇用以外にも医療分野等、災害後の復興・支援を見越した対応も重要である。

2) 個別調査の実施プロセスを踏まえた事前の備え (どのように備えるのか)

本調査研究で行った海外事例調査からは、大規模な災害が起こりうることを踏まえ、平常時から災害時の対応について検討・研究が行われていることが明らかになった。

カナダ統計局におけるBCPに関する取組では、個別調査の実施プロセスを踏まえることで、災害時にあっても複数の統計調査を横断的に把握できるように工夫を行うことが検討されている。また、カナダ統計局が実施するLFSでは、労働時間に特化して質問数も盛り込んだ特殊モジュールDCE(Disaster / Catastrophe Effects:災害/大災害による影響)を事前に作成しておくことで災害時の影響把握を迅速に行う工夫も行われている。

ーカナダ LFS における特殊モジュールの質問イメージー

※斜体下線部分は、調査実施時期及び災害内容に応じたものが表示される

|   |   |
|---|---|
| 今月は数問の追加質問をお伺いしたいと思います。<br>1. 回答(→次の質問へ)<br>2. 拒否(→理由を記入して終了) |   |
| DC_Q01  | 1月の最初の2週間において、●●さん(世帯内の回答対象者)は、 <u>寒波</u> によって仕事を離れていましたか？<br>1. はい(→DC_Q02 へ)<br>2. いいえ(→DC_Q03 へ)<br>3. 仕事には就いていなかった(→終了) |
| DC_Q02  | では、 <u>寒波</u> によって仕事を離れていたのは何時間でしたか？<br>( )時間(→DC_Q03 へ)  |
| DC_Q03  | 1月の最初の2週間において、●●さん(世帯内の回答対象者)には、 <u>寒波</u> によって残業等の追加的な就業時間がありましたか？<br>1. はい(→DC_Q04 へ)<br>2. いいえ(→終了)                      |
| DC_Q04  | では、その残業等の追加的な就業時間は何時間でしたか？<br>( )時間(→終了)  |

(参考)

災害内容としては、寒波(Ice Storm)や停電(Power Outage)等が想定されているが、それ以外にも対応可能。

上記の質問内容は、災害によって事業が行われずに就業時間が減少する労働者が存在する一方、消防士等災害対応のために追加的な就業時間が発生する労働者も存在することから、災害が労働市場に対してどの程度の影響を及ぼしたかを就業時間の視点から把握するために設定されている。災害発生後の一定期間にわたって質問を行うことで、災害の影響が消失するまでの期間についても把握することが可能となっている。加えて、モジュールに質問項目の柔軟性(災害内容や時期の設定)を持たせることで、複数の災害が同時発生した場合にも対応が図られている。

また、平常時における統計調査の存在感・レピュテーション(評判)を高める取組も、災害に対する備えとして重要である。統計調査の存在感・レピュテーションを平常時から高め、被災した対象者にとって回答が負担となり得る場合でも、可能な限り協力しようとする意識が持たれるような状況につなげていくことが求められる。

## 第6章 統計調査に欠測値補完を適用する際の示唆・課題

## 附属資料 目次

|   |              |
|---|--------------|
| <b>I.国内における補完推計に関する研究</b>                               | <b>附属 1</b>  |
| 1 経済センサスー活動調査 .....                                     | 附属 1         |
| 1) 調査票例   |              |
| 2) 研究概要   |              |
| 2 国民生活基礎調査 .....  | 附属 31        |
| 1) 調査の概要・体系   |              |
| 2) 調査票例   |              |
| 3) 研究概要   |              |
| 3 特定サービス産業実態調査 .....                                    | 附属 59        |
| 1) 調査業種   |              |
| 2) 調査票例   |              |
| 3) 研究概要   |              |
| 4 全国企業短期経済観測調査 .....                                    | 附属 65        |
| 1) 調査票例   |              |
| 2) 研究概要   |              |
| 5 日本人の国民性調査 .....                                       | 附属 68        |
| 6 労働力調査 .....   | 附属 69        |
| 1) 調査票例   |              |
| 2) 取組概要   |              |
| <b>II.海外における補完推計などに関する取組</b>                            | <b>附属 73</b> |
| 1 ACS (American Community Survey: アメリカ地域社会調査) .....     | 附属 73        |
| 2 Population Census (US Census:国勢調査) .....              | 附属 87        |
| 3 カナダ統計局におけるBCP(Business Continuity Plan: 事業継続計画) ..... | 附属 93        |
| <b>III.参考ウェブサイト・文献一覧</b>                                | <b>附属 97</b> |

# 附属資料 I 国内における補完推計に関する研究

## 1. 経済センサスー活動調査

### 1) 調査票例

経済センサスー活動調査(平成 24 年) 単独事業所調査票(農業、林業、漁業)

The form is titled '経済センサスー活動調査' (Economic Census Activity Survey) and is for '単独事業所調査票(農業、林業、漁業)' (Individual Business Survey Form for Agriculture, Forestry, and Fishing). It contains various tables and text boxes for data entry, including sections for business identification, financial statements, and production volumes. A large blue 'SAMPLE' watermark is overlaid across the center of the form.

出所) 総務省統計局 経済センサス ウェブサイト

経済センサスー活動調査(平成 24 年) 単独事業所調査票(農業、林業、漁業)

This is a detailed data table from the survey form, showing various categories and their corresponding values. The table is organized into columns and rows, with a large blue 'SAMPLE' watermark overlaid diagonally across it. The categories include various types of agricultural, forestry, and fishing activities.

出所) 総務省統計局 経済センサス ウェブサイト

経済センサス-活動調査(平成 24 年) 単独事業所調査票(鉱業、採石業、砂利採取業)

出所) 総務省統計局 経済センサス ウェブサイト

経済センサス-活動調査(平成 24 年) 単独事業所調査票(鉱業、採石業、砂利採取業)

出所) 総務省統計局 経済センサス ウェブサイト

経済センサスー活動調査(平成 24 年) 単独事業所調査票(製造業)

出所)総務省統計局 経済センサス ウェブサイト

経済センサスー活動調査(平成 24 年) 単独事業所調査票(製造業)

出所)総務省統計局 経済センサス ウェブサイト





## 2) 研究概要

### 経済センサスー活動調査 研究概要

2012/11/13

#### 目次

- 研究の目的
- 単一代入法 (Single Imputation) の欠点
- 多重代入法 (Multiple Imputation) のメカニズム
- 単一代入法と多重代入法の検証
- 多重代入法の精度評価
- 結論と今後の課題



6

#### 研究の目的1

- 目標  
今後の経済統計における経理項目の欠測値補定を確立
- 現時点 (平成24年度)  
個別データの欠測値補定方法として多重代入法を研究



7

4

出所) 統計センター情報技術部統計技術研究課 提供資料

## 研究の目的2

### □本報告の目的

売上高を欠測値補定した場合の当てはまりの精度を、単一代入法と多重代入法で多角的に比較



4

## 単一代入法 (Single Imputation) の欠点

### □従来から使用されている欠測値対処法

### □不確実性の無視→分散の過小推定

単一代入法の具体例

| ID | 身長  | 年齢 | 性別 | 体重 | 補定    |
|----|-----|----|----|----|-------|
| 1  | 欠測  | 40 | 男  | 63 | 170.3 |
| 2  | 174 | 31 | 男  | 62 | 174   |
| 3  | 161 | 45 | 女  | 48 | 161   |
| 4  | 158 | 24 | 女  | 42 | 158   |
| 5  | 欠測  | 40 | 男  | 63 | 170.3 |
| 6  | 163 | 52 | 女  | 58 | 163   |
| 7  | 172 | 29 | 男  | 70 | 172   |
| 8  | 153 | 38 | 女  | 46 | 153   |
| 9  | 178 | 28 | 男  | 70 | 178   |



5

### 多重代入法 (Multiple Imputation) のメカニズム

Rubin (1978) により理論化

多重代入法の具体例

| ID | 身長  | 年齢 | 性別 | 体重 | 補定間の不確実性 |       |       |
|----|-----|----|----|----|----------|-------|-------|
|    |     |    |    |    | 補定1      | 補定2   | 補定3   |
| 1  | 欠測  | 40 | 男  | 63 | 171.5    | 165.3 | 173.2 |
| 2  | 174 | 31 | 男  | 62 | 174      | 174   | 174   |
| 3  | 161 | 45 | 女  | 48 | 161      | 161   | 161   |
| 4  | 158 | 24 | 女  | 42 | 158      | 158   | 158   |
| 5  | 欠測  | 40 | 男  | 63 | 168.6    | 172.2 | 168.1 |
| 6  | 163 | 52 | 女  | 58 | 163      | 163   | 163   |
| 7  | 172 | 29 | 男  | 70 | 172      | 172   | 172   |
| 8  | 153 | 38 | 女  | 46 | 153      | 153   | 153   |
| 9  | 178 | 28 | 男  | 70 | 178      | 178   | 178   |

補定内の不確実性

統計センター 10

### 多重代入法1 ~モデル~

$$D \sim N_p(\mu, \Sigma)$$

$D = n \times p$  データセット  
 $n$  は観測数、 $p$  は変数の数

統計センター 11

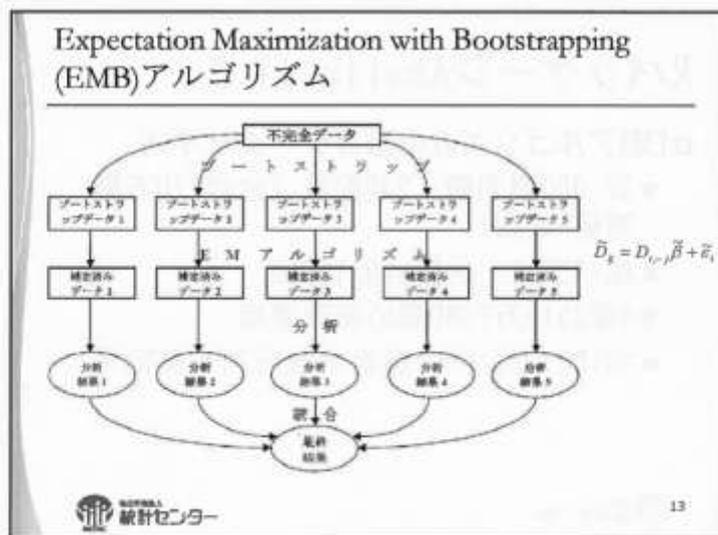
### 多重代入法2 ～モデル～

$$\tilde{D}_{ij} = D_{i,-j} \tilde{\beta} + \tilde{\varepsilon}_i$$

$D_{ij}$  : 観測値*i*及び変数*j*の欠測値  
 $\tilde{D}_{ij}$  : 観測値*i*及び変数*j*のシミュレーション値  
 $D_{i,-j}$  : 変数*j*を除く*i*行のすべての観測値  
 $\sim$  : 適切な事後分布からの無作為抽出



12



## RパッケージAmelia II 1

- 開発者
  - ハーバード大学のGary King教授を中心とするチーム
- Amelia I (2001年)
  - 汎用多重代入法プログラム
- Amelia II (2011年)
  - EMBアルゴリズムを実装
  - 巨大データセットの多重代入に対応



14

## RパッケージAmelia II 2

- EMBアルゴリズムの限界データサイズ
  - 32,000観測値、240変数（合計720万観測値-変数）
  - 総パラメータ数=29,160個
  - 4億2516万7380個の個別要素
  - $29160 \times 29160$ 分散共分散行列を反転可



15

### 産業情報～EDINETデータ～

| 産業大分類 |                 | 完全データ数 |
|-------|-----------------|--------|
| E     | 製造業             | 1222   |
| I     | 卸売業、小売業         | 571    |
| D     | 建設業             | 158    |
| G     | 情報通信業           | 276    |
| L     | 学術研究、専門・技術サービス業 | 191    |

**備考**

EDINETデータは、Electronic Disclosure for Investors' Networkの略であり、「金融商品取引法に基づく有価証券報告書等の開示書類に関する電子開示システム」である。今回使用したデータの対象となっているのは、2011年3月31日に決算を迎える上場企業3,587社である（提出日：2011年6月30日からさかのぼり1年以内の企業）。

16

### 欠測パターンと欠測率1

- MCAR : *Missing Completely At Random* (欠測は完全にランダム)

$$P(K|D) = P(K)$$

- MAR : *Missing At Random* (欠測はランダム)

$$P(K|D) = P(K|D_0)$$

- NI : *Nonignorable* (欠測は無視できない)

$P(K|D)$ は単純化することができない



17

## 欠測パターンと欠測率2

産業大分類ごとに事業従事者を昇順にして、

事業従事者が

- 小さい値
  - 中の値
  - 大きい値
  - 大小の値
- の売上高を欠測 (MARのパターン)

□ ランダム欠測 (MCARのパターン)

□ 系統抽出による欠測発生 (MARのパターン)



⇒欠測率は30%、40%、50% 18

## モデル紹介 ~単一代入法と多重代入法に共通~

使用したモデル

### ■ 1次多項式

$$\text{式: } y = a \cdot x + b$$

y : 売上高

x : 事業従事者

a : 傾き、b : 切片

### ■ 自然対数変換

$$\text{式: } \ln y = a \ln x + b$$

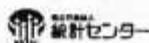


19

### 評価方法～単一代入法と多重代入法の比較～

#### 単一代入法と多重代入法で

1. 完全データの売上高総額（真値）と欠測値補定後の売上高総額の差（差が小さかった数をカウント）
2. 散布図による視覚的アプローチ
3. 欠測値補定データの標準偏差



20

### 多重代入法の設定1 ～欠測値補定～

- 補定済データセット数Mを20に設定
- 欠測補定値については、各々のデータセットの補定値の平均値を設定

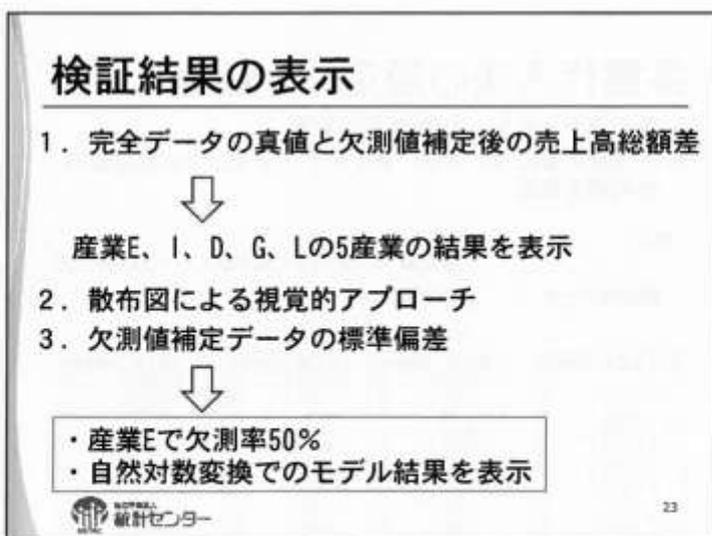
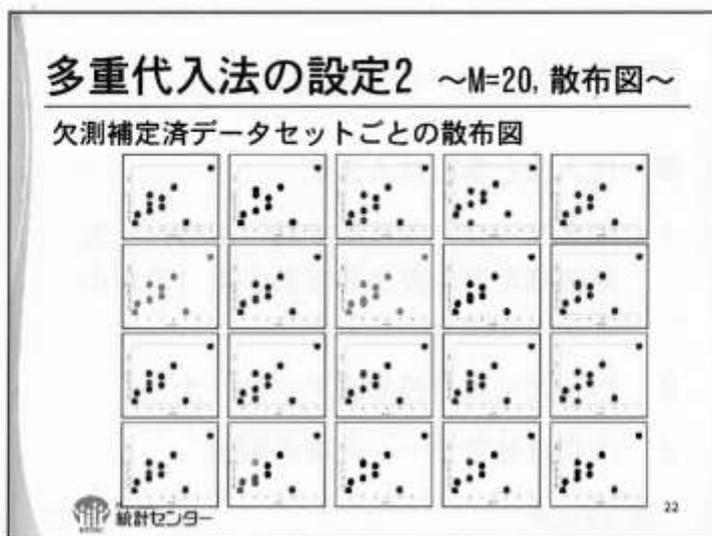
例：

$$\text{補定値} = (57128 + 92188 + \dots + 55784) / 20$$

補定前データ

欠測補定済データ

| No. | 補定前データ |      | n=1    |      | n=2    |      | n=20   |      |
|-----|--------|------|--------|------|--------|------|--------|------|
|     | 売上高    | 標準誤差 | 売上高    | 標準誤差 | 売上高    | 標準誤差 | 売上高    | 標準誤差 |
| 1   |        | 20   | 57128  | 20   | 92188  | 20   | 55784  | 20   |
| 2   | 72000  | 30   | 72000  | 30   | 72000  | 30   | 72000  | 30   |
| 3   | 80000  | 20   | 80000  | 20   | 80000  | 20   | 80000  | 20   |
| 4   | 30000  | 10   | 30000  | 10   | 30000  | 10   | 30000  | 10   |
| 5   | 100000 | 40   | 100000 | 40   | 100000 | 40   | 100000 | 40   |
| 6   | 10100  | 50   | 10100  | 50   | 10100  | 50   | 10100  | 50   |
| 7   | 150000 | 70   | 150000 | 70   | 150000 | 70   | 150000 | 70   |
| 8   | 8000   | 8    | 8000   | 8    | 8000   | 8    | 8000   | 8    |
| 9   | 50000  | 30   | 50000  | 30   | 50000  | 30   | 50000  | 30   |
| 10  | 40000  | 20   | 40000  | 20   | 40000  | 20   | 40000  | 20   |



### 検証結果1 ～売上高総額差、検証結果の一部～

完全データ情報

| 産業     | データ数 | 完全データの売上高総額(真値) |
|--------|------|-----------------|
| E(製造業) | 1222 | 145,785,642     |

欠測値補定情報 単位：百万円

| 欠測パターン    | 欠測率 | モデル    | 欠測補定済売上高総額 | 真値との差(大きさ)  |           |
|-----------|-----|--------|------------|-------------|-----------|
| 従事者小を欠測   | 40% | 1次多項式  | 単一代入       | 147,438,713 | 1,653,071 |
|           |     |        | 多重代入       | 145,651,014 | 134,628   |
| 系統抽出による欠測 | 30% | 自然対数変換 | 単一代入       | 142,479,531 | 3,306,111 |
|           |     |        | 多重代入       | 138,683,476 | 7,102,166 |

備考  
完全データの総額と比較して、1%ほどの乖離があるペアのみ対象  
ペア同士の差が約2倍ほどある場合に優劣差をチェック

24

### 検証結果1.1 ～売上高総額差、産業別～

真値との差が小さい方をカウントする

産業別

| 産業                 | 実数値 [回] |       |    | 構成比 [%] |       |     |
|--------------------|---------|-------|----|---------|-------|-----|
|                    | 多重代入法   | 単一代入法 | 合計 | 多重代入法   | 単一代入法 | 合計  |
| E(製造業)             | 4       | 2     | 6  | 66.7    | 33.3  | 100 |
| I(卸売業、小売業)         | 2       | 2     | 4  | 50.0    | 50.0  | 100 |
| D(建設業)             | 1       | 9     | 10 | 10.0    | 90.0  | 100 |
| G(情報通信業)           | 6       | 1     | 7  | 85.7    | 14.3  | 100 |
| L(学術研究、専門・技術サービス業) | 12      | 1     | 13 | 92.3    | 7.7   | 100 |
| 合計                 | 25      | 15    | 40 | 62.5    | 37.5  | 100 |

統計センター

25

**検証結果1.2 ～売上高総額差、欠測パターン別～**

欠測パターン別

| 欠測パターン | 実数値[回] |       |    | 構成比[%] |       |     |
|--------|--------|-------|----|--------|-------|-----|
|        | 多重代入法  | 単一代入法 | 合計 | 多重代入法  | 単一代入法 | 合計  |
| 従事者:小  | 3      | 1     | 4  | 75.0   | 25.0  | 100 |
| 従事者:中  | 9      | 1     | 10 | 90.0   | 10.0  | 100 |
| 従事者:大  | 4      | 2     | 6  | 66.7   | 33.3  | 100 |
| 従事者大小  | 1      | 1     | 2  | 50.0   | 50.0  | 100 |
| ランダム   | 6      | 2     | 8  | 75.0   | 25.0  | 100 |
| 系統抽出   | 2      | 8     | 10 | 20.0   | 80.0  | 100 |
| 合計     | 25     | 15    | 40 | 62.5   | 37.5  | 100 |

**検証結果1.3 ～売上高総額差、モデル別～**

モデル別

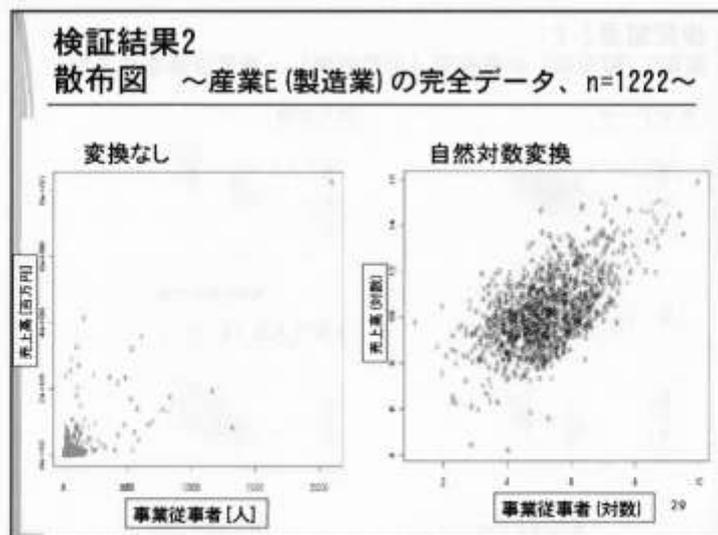
| モデル    | 実数値[回] |       |    | 構成比[%] |       |     |
|--------|--------|-------|----|--------|-------|-----|
|        | 多重代入法  | 単一代入法 | 合計 | 多重代入法  | 単一代入法 | 合計  |
| 1次多項式  | 11     | 6     | 17 | 64.7   | 35.3  | 100 |
| 自然対数変換 | 14     | 9     | 23 | 60.9   | 39.1  | 100 |
| 合計     | 25     | 15    | 40 | 62.5   | 37.5  | 100 |

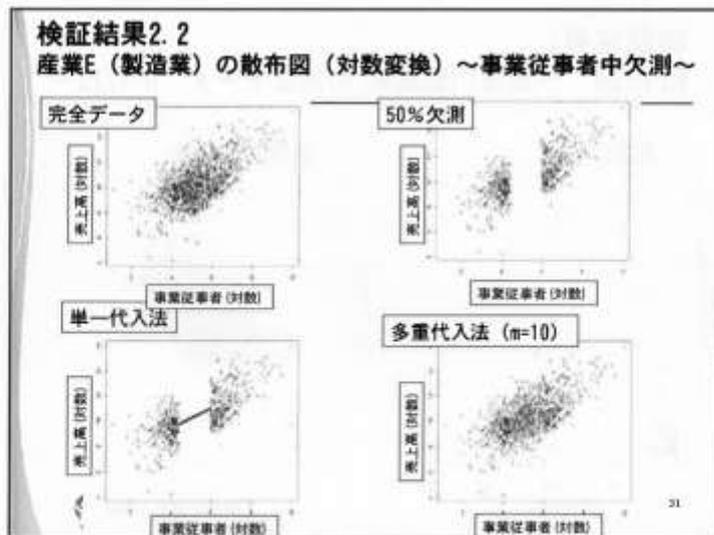
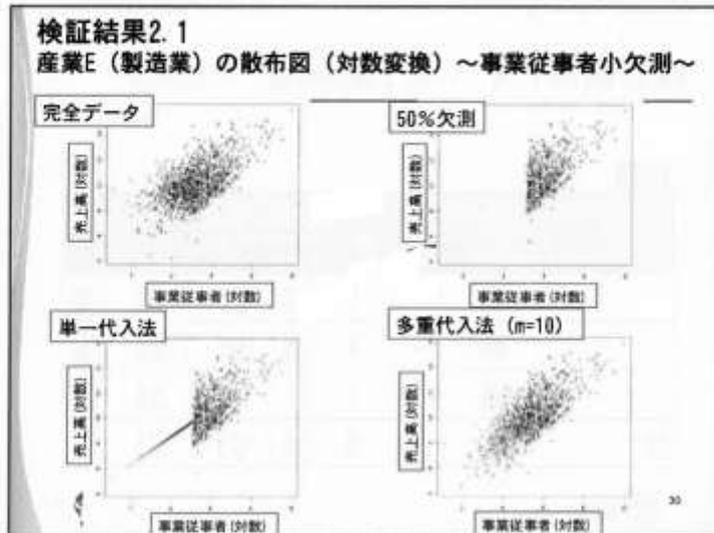
**検証結果1.4 ～売上高総額差、欠測率別～**

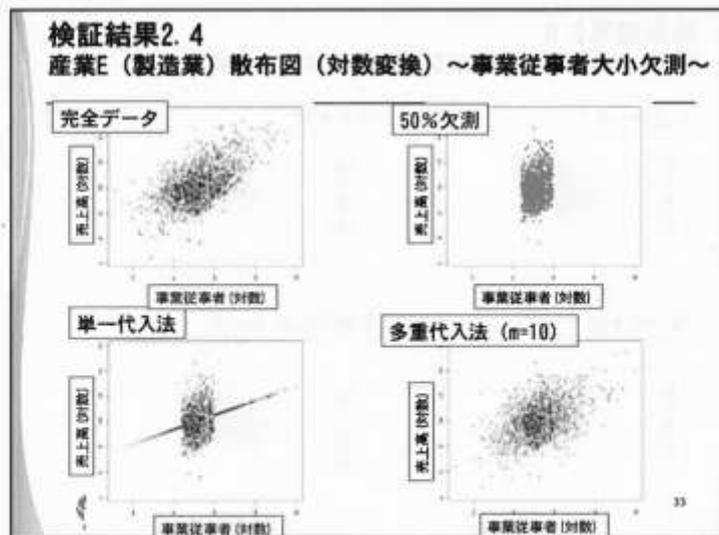
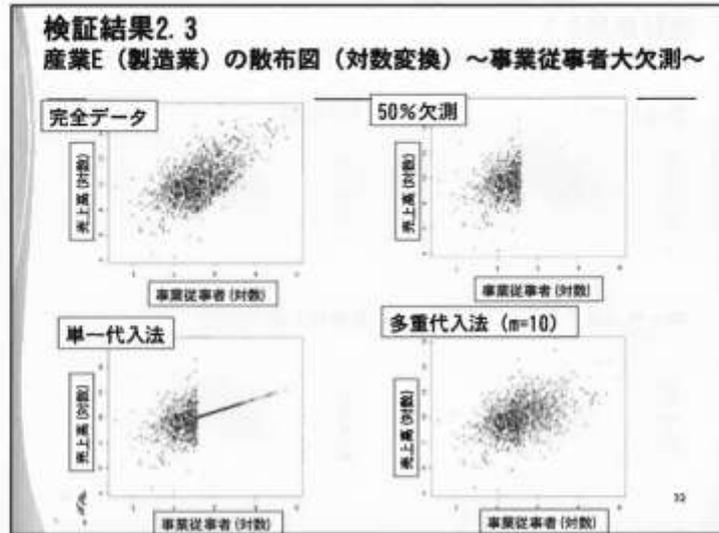
欠測率別

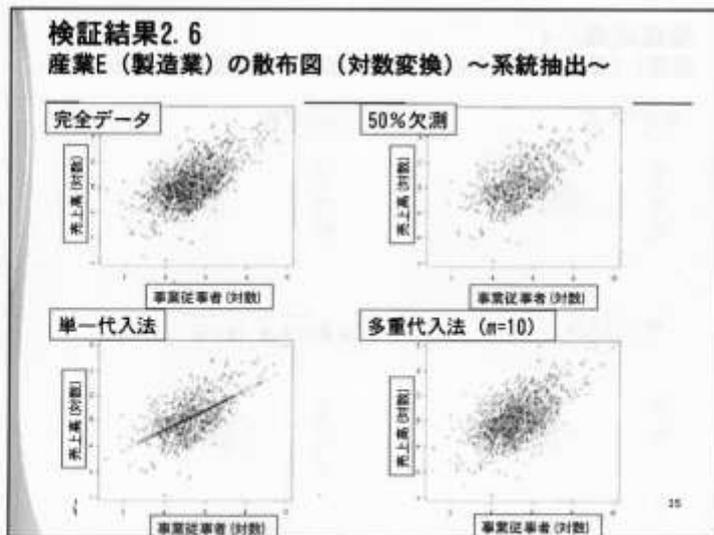
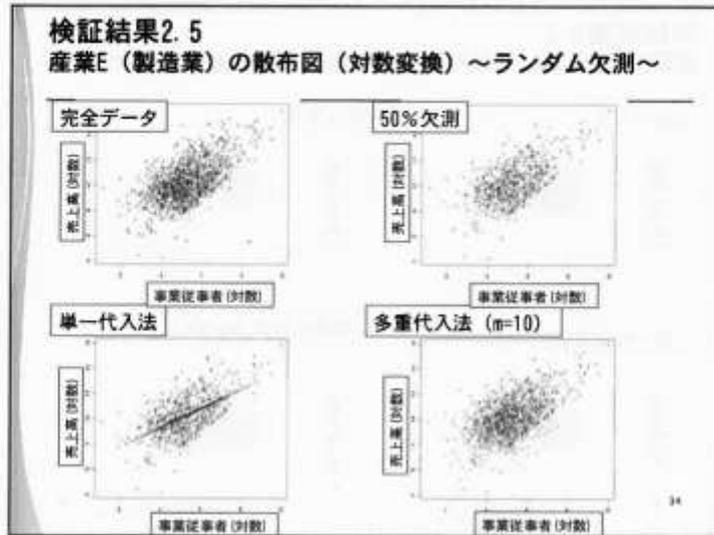
| 欠測率 | 実数値[回] |       |    | 構成比[%] |       |     |
|-----|--------|-------|----|--------|-------|-----|
|     | 多重代入法  | 単一代入法 | 合計 | 多重代入法  | 単一代入法 | 合計  |
| 30% | 11     | 9     | 20 | 55.0   | 45.0  | 100 |
| 40% | 9      | 4     | 13 | 69.2   | 30.8  | 100 |
| 50% | 5      | 2     | 7  | 71.4   | 28.6  | 100 |
| 合計  | 25     | 15    | 40 | 62.5   | 37.5  | 100 |

統計センター 20







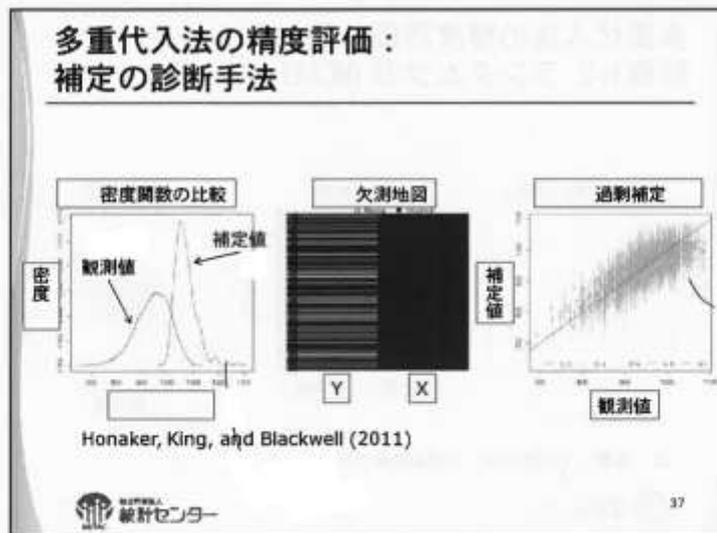


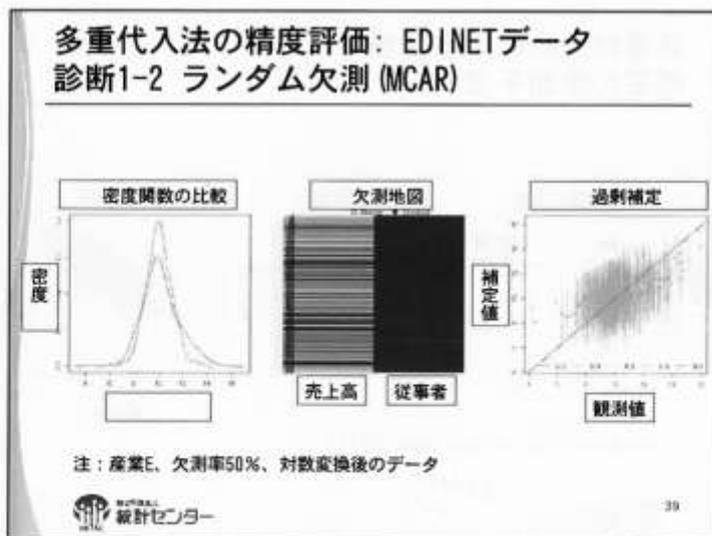
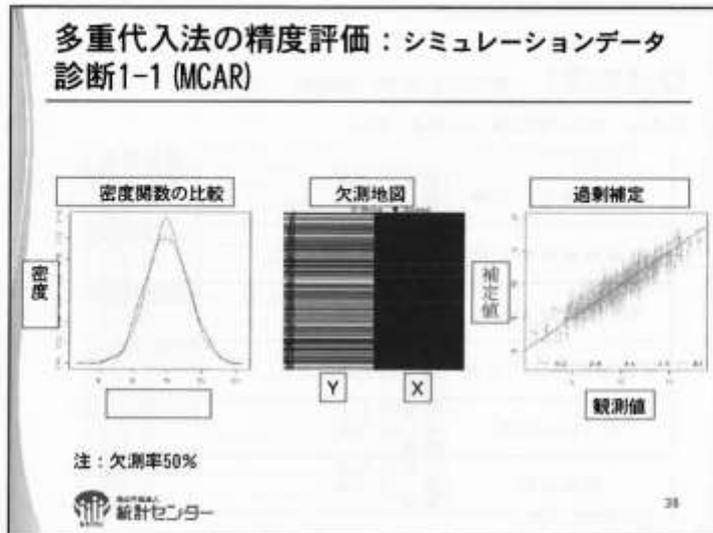
**検証結果3** ~標準偏差(産業E(製造業)、欠測値補定データ)~

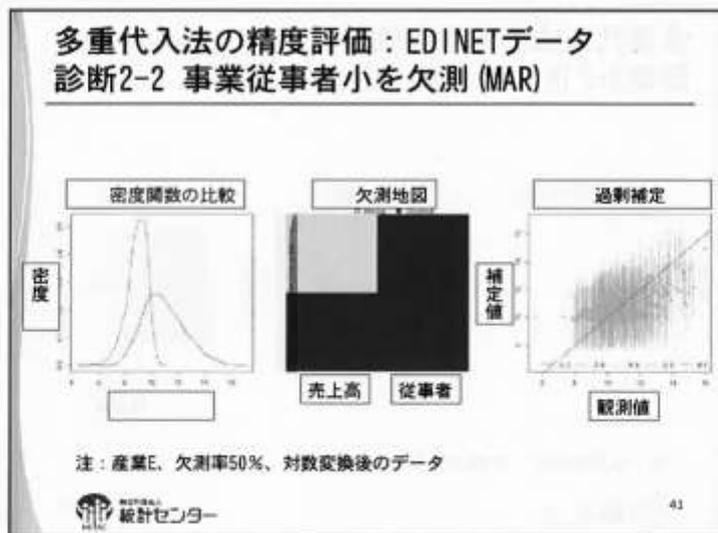
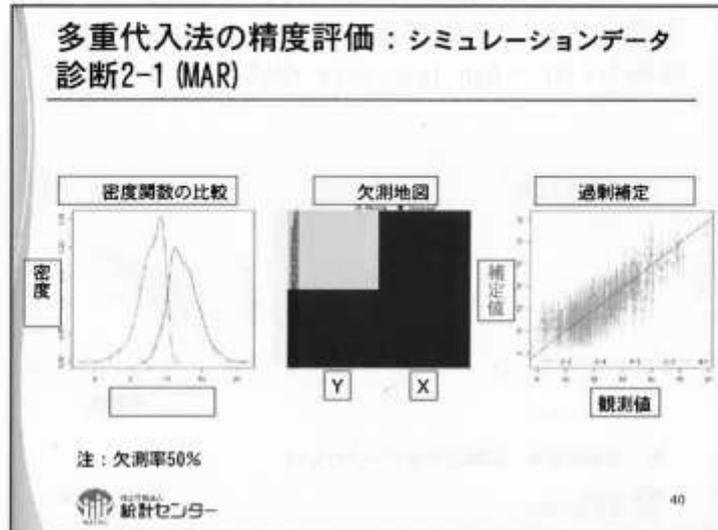
モデル：自然対数変換 欠測率：50%

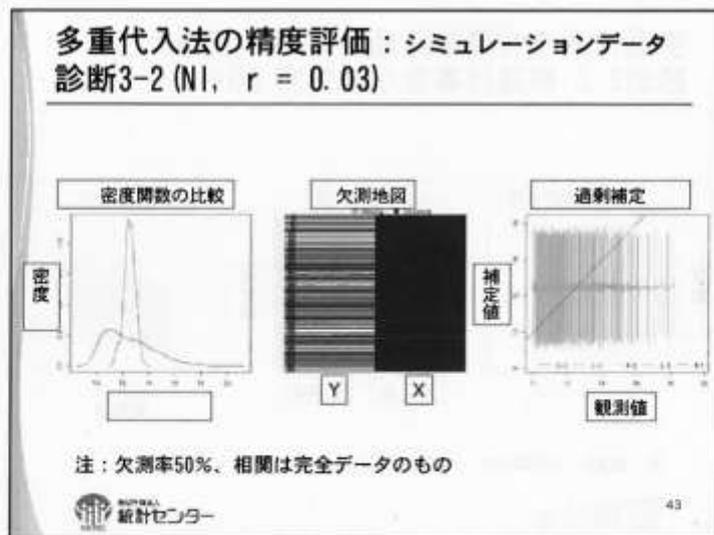
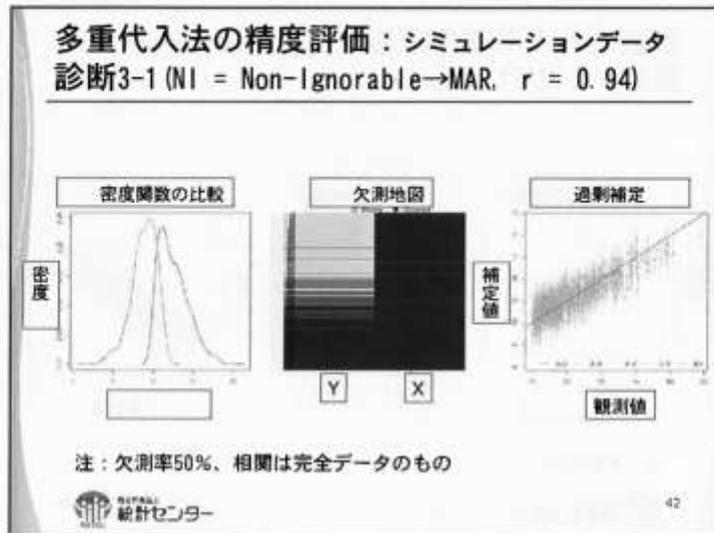
| 欠測パターン    |       | 標準偏差  |
|-----------|-------|-------|
| 事業従事者小欠測  | 多重代入法 | 1.426 |
|           | 単一代入法 | 0.736 |
|           | 真値    | 1.240 |
| 事業従事者中欠測  | 多重代入法 | 1.245 |
|           | 単一代入法 | 0.331 |
|           | 真値    | 1.282 |
| 事業従事者大欠測  | 多重代入法 | 1.257 |
|           | 単一代入法 | 0.418 |
|           | 真値    | 1.554 |
| 事業従事者大小欠測 | 多重代入法 | 1.599 |
|           | 単一代入法 | 0.937 |
|           | 真値    | 1.760 |
| ランダム欠測    | 多重代入法 | 1.525 |
|           | 単一代入法 | 0.889 |
|           | 真値    | 1.532 |
| 系統抽出      | 多重代入法 | 1.586 |
|           | 単一代入法 | 0.956 |
|           | 真値    | 1.489 |

36









### 結論1

#### □ 欠測値補定の総額差の結果（検証結果1）

- 産業別
- 欠測パターン別
- モデル別
- 欠測率別



多重代入法の優位が確認



44

### 結論2

#### □ 散布図による視覚的な観点（検証結果2）

#### □ 欠測値補定データの標準偏差（検証結果3）



多重代入法は復元に近づける結果であった



45

### 結論3

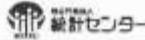
- Ameliaにおける補定の結果診断方法は実務的に有用
- 多重代入法を行うプログラムとして、RパッケージのAmelia IIは有用

### 今後の課題

- 具体的な調査データから実態にあった欠測率にて検証
- 様々なプログラムとの比較
- 外れ値の欠測値補定に与える影響を考慮

## 参考文献1

- Abayomi, Kobi, Andrew Gelman, and Marc Levy. (2008). "Diagnostics for Multivariate Imputations," *Applied Statistics* vol.57, no.3: 273-291.
- Congdon, Peter. (2006). *Bayesian Statistical Modelling*, Second Edition. West Sussex: John Wiley & Sons Ltd.
- DeGroot, Morris H. and Mark J. Schervish. (2002). *Probability and Statistics*. Boston: Addison-Wesley.
- Drechsler, Jörg. (2009). "Far From Normal - Multiple Imputation of Missing Values in a German Establishment Survey," *Work Session on Statistical Data Editing, Conference of European Statisticians*, Neuchâtel, Switzerland, 5-7 October 2009.
- Gelman, Andrew, and Jennifer Hill. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.
- Gill, Jeff. (2008). *Bayesian Methods—A Social Sciences Approach*, Second Edition. London: Chapman & Hall/CRC.
- Honaker, James and Gary King. (2010). "What to do About Missing Values in Time Series Cross-Section Data," *American Journal of Political Science* vol.54, no.2: 561-581.
- Honaker, James, Gary King, and Matthew Blackwell. (2011). "Amelia II: A Program for Missing Data," *Journal of Statistical Software* vol.45, no.7.
- Honaker, James, Gary King, and Matthew Blackwell. (2012a). *Amelia II: A Program for Missing Data* version 1.6.1. <http://cran.r-project.org/web/packages/amelia/vignettes/amelia.pdf>



48

## 参考文献2

- Honaker, James, Gary King, and Matthew Blackwell. (2012b). Package 'Amelia' Version 1.6.1 <http://cran.r-project.org/web/packages/amelia/amelia.pdf>
- 星野康忠. (2009). 『調査観察データの統計科学—因果推論・選択バイアス・データ融合』。東京: 岩波書店
- 岩崎学. (2002). 『不完全データの統計解析』。東京: エコノミスト社
- King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. (2001). "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation," *American Political Science Review* vol.95, no.1: 49-69.
- Little, Roderick J. A. and Donald B. Rubin. (2002). *Statistical Analysis with Missing Data*, Second Edition. New Jersey: John Wiley & Sons.
- Rubin, Donald B. (1978). "Multiple Imputations in Sample Surveys—A Phenomenological Bayesian Approach to Nonresponse," *Proceedings of the Survey Research Methods Section, American Statistical Association*: 20-34.
- Rubin, Donald B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.
- Schafer, Joseph L. (1999). "Multiple Imputation: A Primer," *Statistical Methods in Medical Research* vol.8: 3-15.



49

### 参考文献3

- Schmidt, Katrin. (2009). "Multiple Imputation with Standard Software: First Application Experiences," *Work Session on Statistical Data Editing, Conference of European Statisticians*, Neuchâtel, Switzerland, 5-7 October 2009.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.
- Shao, Jun. (2002). "Replication Methods for Variance Estimation in Complex Surveys with Imputed Data," in *Survey Nonresponse* edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, Roderick J. A. Little. New York: John Wiley & Sons, pp.303-314.
- Shao, Jun and Dongsheng Tu. (1995). *The Jackknife and Bootstrap*. New York: Springer.
- Waal, Ton de, Jeroen Pannekoek, and Sander Scholtus. (2011). *Handbook of Statistical Data Editing and Imputation*. Hoboken, NJ: John Wiley & Sons.
- 渡辺美智子, 山口和範 編著. 2010. 『EMアルゴリズムと不完全データの補完問題』, 東京. 多賀出版.
- Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.
- Yucel, Recai M. (2011). "State of the Multiple Imputation Software," *Journal of Statistical Software* vol.45, no.1.

ご清聴、ありがとうございました。

### 参考1

$$ARE = \left( \sqrt{1 + \frac{\delta}{M}} \right)^{-1} \times 100$$

- 有限のMと無限大のMとの相対効率(ARE: Asymptotic Relative Efficiency)
- ここで、 $\delta$ は欠測率( $0 \leq \delta \leq 1$ )
- Mが無限大の場合、極限值は100%となり、効率性が最大に到達
- AREは%であり、単位は標準偏差



52

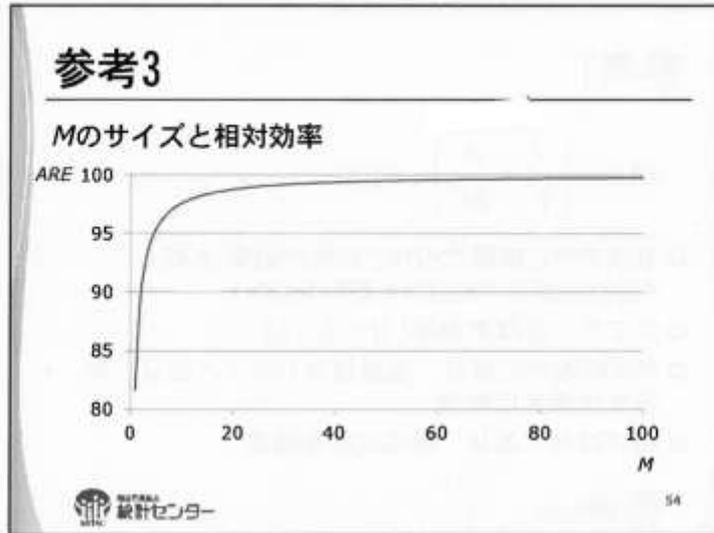
### 参考2

Mと相対効率

| M        | $\delta = 0.1$ | $\delta = 0.2$ | $\delta = 0.3$ | $\delta = 0.4$ | $\delta = 0.5$ |
|----------|----------------|----------------|----------------|----------------|----------------|
| 5        | 99.02          | 98.06          | 97.13          | 96.23          | 95.35          |
| 10       | 99.50          | 99.02          | 98.53          | 98.06          | 97.59          |
| 15       | 99.67          | 99.34          | 99.02          | 98.69          | 98.37          |
| 20       | 99.75          | 99.50          | 99.26          | 99.01          | 98.77          |
| $\infty$ | 100.00         | 100.00         | 100.00         | 100.00         | 100.00         |



53



### 参考4

$$\bar{\theta}_M = \frac{1}{M} \sum_{m=1}^M \hat{\theta}_m$$

$$T_M = \bar{v}_M + \left(1 + \frac{1}{M}\right) \tilde{v}_M$$

$$\bar{v}_M = \frac{1}{M} \sum_{m=1}^M v_m \quad \tilde{v}_M = \frac{1}{M-1} \sum_{m=1}^M (\hat{\theta}_m - \bar{\theta}_M)^2$$

統計センター

55



## 2) 調査票例

### 国民生活基礎調査(平成 25 年) 世帯票(案)

資料2-10

国民生活基礎調査【世帯票】(案)

世帯員3人

この調査は、統計法に基づき国が実施する基礎統計調査です。調査票情報の秘密の保護に万全を期していますので、ありのままを記入してください。

記入上の注意

- 『(世帯長・世帯員) 記入のしかた』をよくお読みになってから記入してください。
- もし記入方法がわからなかった場合は、調査員が受け取りにうかがったときにおたずねください。
- 選択肢は指示がない場合は、あてはまる番号1つに○をつけてください。
- 数字は右づめで記入してください。
- できるだけ黒のボールペンで記入してください。

あなたの世帯について、平成25年6月6日現在の状況をお答えください。

世帯とは、ふだん住居と生計を共にしている人々(世帯員)の集まりをいいます。世帯員には、旅行や出張などで一時的(3か月以内)に自宅を離れている人や勤員が就業場所を移動する人も含まれます。また、病院に入院している人も含まれますが、住民登録簿に記載している人は除きます。さらに、単身赴任や卒業で世帯を離れている人、老人福祉施設などの社会福祉施設に入院している人も除きます。

#### I 世帯の状況

質問1 ふだん一緒に住んでいる(世帯員)で、生計を共にしている(世帯長)は、あなたを指して何人ですか。  
(一時的に不在の方を含みます。)

1人 (世帯長) の場合は、補問1-Cをお読みください。

補問1-1 1人(世帯長)の方の場合は、その状況についてお答えください。

【世帯世帯の状況】

- 1 正み込み、世帯長等に相当する単身世帯
- 2 その他の世帯世帯

【単身赴任の有無】

- 1 単身赴任者である
- 2 単身赴任者でない

質問2 現在は、単身赴任などで世帯を離れているが、その期間、一緒に住んでいた方がいる場合は、あてはまるすべての番号に○をつけ、それぞれの人数を記入してください。(1つ1つ番号は、○に○をつけてください。)

- 1 単身赴任で世帯を離れている者がいる
- 2 学業のため世帯を離れている者がいる
- 3 老人福祉施設に入院している者がいる
- 4 社会福祉施設(老人福祉施設を除く)に入院している者がいる
- 5 病院に長期入院している(医療施設を病院に類している。)者がいる
- 6 1-5の者はいない

※ 1-5に該当する方は、この調査の世帯員とはなりませんので、質問1の人数には含まれません。

裏面に続きます。

出所)統計委員会 第34回人口・社会統計部会配布資料

### 国民生活基礎調査(平成 25 年) 世帯票(案)

質問3 お住まいの住居の種類と建て方についてお答えください。

【住居の種類】

- 1 持ち家
- 2 民間賃貸住宅
- 3 社宅・公務員住宅等の給与住宅
- 4 都市再生機構・公社等の公営賃貸住宅
- 5 借間・その他

【建て方】

- 1 一戸建て
- 2 共同住宅  
(マンション、アパート等)

質問4 お住まいの住居の室数について、居住用の部屋数(玄関や風呂等は含めないでください。)を記入してください。また、床面積は、玄関や廊下等も含めた住宅全体のおおよその床面積を記入してください。

【室数】

室

【床面積】

㎡

※ 床面積の換算  
1坪=3.3㎡  
2畳=3.3㎡

質問5 5月中の家計支出総額(世帯の全員の出支金部の合計額)を記入してください。

万円

※ 以下の費用は家計支出には含めないでください。  
税金、社会保険料、事業上の支払い(商家における租税や借料、商店における商品の仕入れに使った金等)、貯蓄、現金や住宅ローンなどの返済、掛け捨て以外の生命保険料・損害保険料

補問5-1 小学校入学前の方がいる場合は、5月中の家計支出総額のうち、育児にかかった費用を記入してください。

万円

※ 育児にかかった費用とは、ミルク代、母乳代、医療費、保険料、遊び場の費用、衣服費、おもちゃ代、その他育児にかかった費用のすべてが入ります。

補問5-2 5月中の家計支出総額のうち、別居している親又は子への仕送りの状況をお答えください。仕送りの目的については、あてはまるすべての番号に○をつけてください。

【仕送りの状況】

別居している親がいる

- 1 仕送りあり
- 2 仕送りなし

3 別居している親はいない

【仕送りの目的】

- 1 入所・入院費
- 2 その他

【仕送り額】

万円

【仕送りの状況】

別居している子がいる

- 1 仕送りあり
- 2 仕送りなし

3 別居している子はいない

【仕送りの目的】

- 1 学業
- 2 その他

【仕送り額】

万円

引き続き「II 世帯員の状況」についてもお答えください。

出所)統計委員会 第34回人口・社会統計部会配布資料



国民生活基礎調査(平成 25 年) 所得票(案)

資料 2-1

厚生労働省

## 国民生活基礎調査【所得票】(案)

(平成25年7月11日調査)

この調査は、統計法に基づき国が実施する基礎統計調査です。  
調査票情報の秘密の保護に万全を期していますので、ありのままを記入してください。

**<記入上の注意>**

- この調査票は、昨年1年間(平成24年1月1日～12月31日)に何らかの所得や税金、社会保険料、企業年金・個人年金等の掛金の支出があった方が1人1票ずつ、記入してください。  
(所得には、アルバイトによる所得や仕送り、年金も含まれます。)
- もし記入方法がわからなかった場合は、調査員が受け取りにうかがったときにおたずねください。
- ご自分で記入できない方については、ご家族の方が回答してください。
- できるだけ黒のボールペンで記入してください。
- 7月11日以前に調査員があらためておうかがいいたしますので、それまでに緑色の質問について記入してください。

所得や課税等の支出のあった方は、質問1から順に記入してください。

**質問1** あなたの性・出生年月を記入してください。  
性・元号はあてはまる番号1つに○をつけ、出生年月には数字を百の位で記入してください。

| 性   | 出生年月 |      |    |    |
|-----|------|------|----|----|
| 1 男 | 1 明治 | 3 昭和 | □年 | □月 |
| 2 女 | 2 大正 | 4 平成 | □年 | □月 |

※ 所得については2ページから、課税等の支出については6ページから記入してください。  
15ページは、世帯主又は世帯を代表する方が記入してください。

調査員が記入欄

|      |        |       |        |
|------|--------|-------|--------|
| 世帯番号 | □□□□□□ | 世帯員番号 | □□□□□□ |
|------|--------|-------|--------|

-1-

出所)統計委員会 第34回人口・社会統計部会配布資料

国民生活基礎調査(平成 25 年) 所得票(案)

昨年1年間に何らかの所得を受け取った場合は、2及び4ページの所開の種類ごとに、1年分の所得金額を万円単位で記入してください。  
所得のなかった方は、6ページへお進みください。

右ページの書類をお持ちの方は参考にしてください。

**【金額記入の注意】**

- 万円未満は四捨五入して、万円単位で必ず記入してください。  
(1～4,999円は「0万円」、5,000～14,999円は「1万円」)
- 生命保険の受取金、退職金、不動産や株の売却代金、宝くじの当せん金などの一時的なものは含みません。

**質問2** あなたは昨年1年間(平成24年1月～12月)に何らかの所得を受け取りましたか。

受け取った所得の種類ごとに金額を記入してください。

**雇用者所得** 01 □□□□□□ 万円  
□ 千 □ 百 □ 十

**事業所得** 02 □□□□□□ 万円  
□ 千 □ 百 □ 十

**農林・畜産所得** 03 □□□□□□ 万円  
□ 千 □ 百 □ 十

**家内労働所得** 04 □□□□□□ 万円  
□ 千 □ 百 □ 十

**財産所得** 05 □□□□□□ 万円  
□ 千 □ 百 □ 十

1年分の所得金額がわからずにはきは、1か月の収入の12倍にボーナス分を加えるなどして、1年分の金額を計算して記入してください。

**働いて得た所得**

勤め先から受け取った給料、賞与(ボーナス)を含めた精込み金額を記入してください。アルバイト等による所得も含みます。  
【参考書類】源泉徴収票(※本文は写し)給与明細書

事業(農林・畜産以外)による収入から、仕入額、従業員に対する給与などの必要経費を差し引いた所得金額を記入してください。漁業・林業による所得を含みます。  
【参考書類】確定申告書(控)

農業や畜産業による収入(自家消費分を除く)から、肥料代、農薬代、草畜・家畜の購入費、畜い人の賃金などの必要経費を差し引いた所得金額を記入してください。  
【参考書類】確定申告書(控)

従業主からの委託を受けて、品物の製造や加工等(校正業務やワープロ入力などを除く)を行って得た所得から必要な経費を差し引いた所得金額を記入してください。

**財産による所得**

借入金や土地を貸すことによって得た所得や、預貯金、公社債、株式などから得られた利息、配当金(源泉分離課税分を除く)の合計額を記入してください。宝くじの当せん金、宝くじの戻り金、株の売却代金、不動産の売却代金などは含みません。  
【参考書類】確定申告書(控)取引口座の通帳、配当金領収書など

-2-

4ページへ

出所)統計委員会 第34回人口・社会統計部会配布資料

国民生活基礎調査(平成 25 年) 所得票(案)

**給与所得者の方**

⇒ 平成24年分 給与所得の源泉徴収票 [原本又は写し] を参考にしてください。

**確定申告を行った方**

⇒ 平成24年分の所得税の確定申告書 [控] を参考にしてください。

【注意】 源泉徴収票に記された「給与」「配当」のあった年、事業が廃業処分をされた「給与」等も「給与所得」欄へ記入してください。

出所)統計委員会 第34回人口・社会統計部会配布資料

国民生活基礎調査(平成 25 年) 所得票(案)

**質問 2 (つづき)**

公的年金・恩給 万円

雇用保険 万円

児童手当等 万円

その他の社会保障給付金 万円

仕送り 万円

企業年金・個人年金等 万円

その他の所得 万円

**公的年金・恩給による所得**  
国民年金、基礎年金、厚生年金(厚生年金基金からの年金を含む。)、共済年金、福祉年金、恩給などからの受取額を記入してください。【参考書類】年金振込通知書など  
〔支払期(2か月)分しか受給額がわからないときは、その金額を8倍するなどして、1年分の金額を記入してください。〕

**公的年金・恩給以外の社会保障給付金による所得**  
雇用保険の失業等給付の受取額(育児休業給付、介護休業給付を含む。)を記入してください。【参考書類】雇用保険受給資格書証  
児童手当、子ども手当、児童扶養手当、特別児童扶養手当、児童障害手当などの児童に関する社会保障給付金の受取額を記入してください。  
生活保護法による扶助、医療保険による療養手当、出産手当金、労働者災害補償保険などその他法令に基づく支給金の受取額を記入してください。

**仕送りによる所得**  
定期的又は継続的に送られてきた金額の額を記入してください。品物は、価値に換算した額を記入してください。  
\*年金などの受取り口にある世帯で、家族を仕送りの目的で生活費等として定期的に仕送りしている場合は、その金額をこの欄に記入してください。

**企業年金・個人年金等による所得**  
企業年金、生命保険会社・かんぽ生活、銀行・証券会社などの個人年金及び年金型商品、国民年金基金、養老年金などからの受取額を記入してください。  
\*厚生年金基金からの受取額、「国民年金-副額」を記入してください。

**その他の所得**  
上記以外の所得調整の金、各種預金、せりふ、祝賀金などの受取額を記入してください。  
\*贈与金、返却しない貸付金等も記入してください。

6ページへ

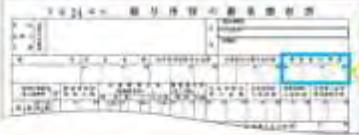
出所)統計委員会 第34回人口・社会統計部会配布資料



国民生活基礎調査(平成 25 年) 所得票(案)

**給与所得者の方**

⇒ 平成24年分 給与所得の源泉徴収票(原本又は写し)を参考にしてください。



源泉徴収税額を『所得税』欄へ記入

**確定申告を行った方**

⇒ 平成24年分の所得税の確定申告書(控)を参考にしてください。



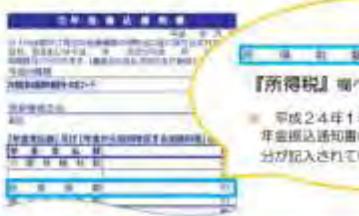
所得税額 = 差引所得金額① - 災害減免額②、  
外国税額控除③

『所得税』欄へ記入

※ 申告書Aでは、②と③の欄になります。

**公的年金を受給している方**

⇒ 公的年金から天引きされた方は、年金振込通知書を参考にしてください。



『所得税』欄へ記入

※ 平成24年1年分の金額を記入してください。  
年金振込通知書は、1支払期(2か月)  
分が記入されています。

**【金額記入の注意】**  
千円未満は四捨五入して、千円単位で右づめに記入してください。  
(1~499円は「0千円」、500~1,499円は「1千円」)

国民生活基礎調査(平成 25 年) 所得票(案)

**質問4** 平成25年度の住民税の課税はありましたか。

あてはまる番号に○をつけ、1に○をつけた方は金額も記入してください。

**住民税**

1 課税あり →  万  千円

千円未満は四捨五入して、千円単位で右づめに記入してください。  
(1~499円は「0千円」、500~1,499円は「1千円」)

2 課税なし

※ 住民税は、道府県(都)民税と市町村(区)民税の合計です。  
不動産課税にかかる住民税は、除いてください。

**【参考書類】**

- 平成25年度 給与所得等に係る市町村(区)民税・道府県(都)民税  
特別徴収税額の決定・変更通知書
- 平成25年度 市町村(区)民税・道府県(都)民税  
税額決定・納税通知書

10ページへ

国民生活基礎調査(平成 25 年) 所得票(案)

**給与所得者の方**

⇒ 給与天引き以外に住民税を納めている方は、次の「給与所得者以外の方」を参考にしてください。

平成25年度 給与所得等に係る市町村(区)民税・道庁県(都)民税特別徴収税額の決定・変更通知書を参考にしてください。

※ 様式は、各地の公共団体によって異なります。

**給与所得者以外の方**

⇒ 平成25年度 市町村(区)民税・道庁県(都)民税税額決定・納税通知書を参考にしてください。

※ 様式は、各地の公共団体によって異なります。

通知書の項目記載例:

- 「住民税精算」
- 「年税額」
- 「市町村(区)民税、道庁県(都)民税の合計税額」
- 「普通徴収と特別徴収の合計額」

記載されている額を『住民税』欄へ記入

**【金額記入の注意】**  
千円未満は四捨五入して、千円単位で右つめに記入してください。  
(1~499円は「0千円」、500~1,499円は「1千円」)

- 9 -

出所)統計委員会 第34回人口・社会統計部会配布資料

国民生活基礎調査(平成 25 年) 所得票(案)

**質問5** 平成24年分の社会保険料の支払いはありましたか。

医療保険(短期掛金)・年金保険(長期掛金)・介護保険・雇用保険のうち、ひとつでも支払いがあったときには、1に○をつけ、支払った金額も記入してください。ひとつも支払いがなかったときには、2に○をつけてください。

※ 保険料は、実際に支払った方が記入してください。

1 支払いあり      2 支払いなし

支払いのない方は、18ページへお進みください。

● 昨年1年間に支払った社会保険料の総額を記入してください。

記入のしかたは、11ページの『社会保険料の総額』の記入のしかたより、「給与所得者の方」又は「給与所得者以外の方」から選んで記入してください。

社会保険料の総額  万  千円

千円未満は四捨五入して、千円単位で右つめに記入してください。

千 百 + -

● 支払った保険料の内訳を記入してください。

記入のしかたは、12ページの『支払った保険料の内訳』の記入のしかたより「給与所得者の方」又は「給与所得者以外の方」から選んで記入してください。

**内 訳**

医療保険(短期掛金)  万  千円

年金保険(長期掛金)  万  千円

介護保険  万  千円

雇用保険  万  千円

千円未満は四捨五入して、千円単位で右つめに記入してください。  
(1~499円は「0千円」、500~1,499円は「1千円」)

千 百 + -

14ページへ

- 10 -

出所)統計委員会 第34回人口・社会統計部会配布資料

### 【『社会保険料の総額』の記入のしかた】

#### 給与所得者の方

下の書類をお持ちですか。複数お持ちの方は1つを選んでください。

- 源泉徴収票 (原本又は写し) ⇒ 参考資料①
- 確定申告書 (控) ⇒ 参考資料②
- 住民税の納税通知書 ⇒ 市町村課税課の『社会保険料』

持っていない方 ⇒ 12ページに読み、『支払った保険料の内訳』の記入のしかたを参考に、内訳を計算してから、内訳の合計を総額に記入してください。

#### 給与所得者以外の方

下の書類をお持ちですか。複数お持ちの方は1つを選んでください。

- 確定申告書 (控) ⇒ 参考資料②
- 住民税の納税通知書 ⇒ 市町村課税課の『社会保険料』

持っていない方 ⇒ 12ページに読み、『支払った保険料の内訳』の記入のしかたを参考に、内訳を計算してから、内訳の合計を総額に記入してください。

【注意】 昨年、年の途中で支払い方法を変更された方(例: 納付書から給与天引に変更)は、それぞれで支払った保険料の合計を記入してください。

**参考資料① 給与所得者の方**

⇒ 平成24年分 給与所得の源泉徴収票 (原本又は写し) を参考にしてください。

**参考資料② 給与所得者の方 給与所得者以外の方**

⇒ 平成24年分の所得税の確定申告書 (控) を参考にしてください。

【金額記入の注意】  
主計標準記号を記入して、主計単位で右之端に記入してください。  
(1~499円は「0千円」、500~1,499円は「1千円」)

### 【『支払った保険料の内訳』の記入のしかた】

#### 給与所得者の方

給与明細書をお持ちですか。  
(給与から天引されている方は、その「給与明細書」から選んでください。)

- 1年分を持っている方 ⇒ 月々(ボーナス分を含む)の給与明細書から、社会保険料の欄ごとに1年分の金額を合計して記入してください。
- 10ページの『社会保険料の総額』を記入した方 ⇒ 計算式A
- 10ページの『社会保険料の総額』を記入していない方 ⇒ 計算式B

1か月分を持っている方 ⇒ 1年分は持っていない方

#### 給与所得者以外の方

保険料は、どのような方法で納付されましたか。

- 口座振替又は納付書(普通徴収)で納付された方 ⇒ 納入(税)通知書又は各納付書から、国民健康保険料(税)、国民年金保険料、介護保険料の1年分の金額を記入してください。
- 公的年金から天引き(特別徴収)で納付された方 ⇒ 参考資料③  
年金振込通知書等から、国民健康保険料、介護保険料の1年分の金額を計算してください。

【注意】 昨年、年の途中で支払い方法を変更された方(例: 納付書から給与天引に変更)は、それぞれで支払った保険料の合計を記入してください。

- 40歳以上の方で、国民健康保険と介護保険の区分ができない場合は、国民健康保険にまとめて記入してください。
- 「厚生年金基金」の支払いは年金保険に含めますが、「国民年金基金」、「企業年金基金」の支払いは、14ページの欄で「企業年金・個人年金等」に含めてください。

**参考資料③ 給与所得者以外の方**

【社会保険料の内訳】  
以下のいずれか1つを  
「国民健康保険(税)額」  
「国民年金基金」  
「国民年金」  
「厚生年金基金」  
「企業年金」  
「個人年金等」

- 介護保険料額 ⇒ 『介護保険』欄へ記入
- 国民健康保険(税)額 又は 後期高齢者医療保険料 ⇒ 『医療保険』欄へ記入

⇒ 平成24年1年分の金額を記入してください。  
年金振込通知書は、1支出路「266円」が記入されています。

国民生活基礎調査(平成 25 年) 所得票(案)

**計算式A 『社会保険料の総額』を記入した方**

① 給与明細から転記

|                       |         |   |                 |   |
|-----------------------|---------|---|-----------------|---|
| (1) 1か月の「医療保険料」(短期掛金) | 円 × (6) | = | 医療保険 (短期掛金) (a) | 円 |
| (2) 1か月の「年金保険料」(長期掛金) | 円 × (6) | = | 年金保険 (長期掛金) (b) | 円 |
| (3) 1か月の「介護保険料」       | 円 × (6) | = | 介護保険 (c)        | 円 |
| (4) 1か月の「雇用保険料」       | 円 × (6) | = | 雇用保険 (d)        | 円 |

② 『社会保険料の総額』を記入する欄

社会保険料の総額 (e) 千円 = (a) + (b) + (c) + (d) 千円 = (6) 千円

計算式B 『社会保険料の総額』を記入していない方

① 給与の月数(回数) (f) 回 + ボーナス月数 (g) 回 = (1) 回

② 給与明細から転記

|                       |         |   |                 |   |
|-----------------------|---------|---|-----------------|---|
| (2) 1か月の「医療保険料」(短期掛金) | 円 × (1) | = | 医療保険 (短期掛金) (a) | 円 |
| (3) 1か月の「年金保険料」(長期掛金) | 円 × (1) | = | 年金保険 (長期掛金) (b) | 円 |
| (4) 1か月の「介護保険料」       | 円 × (1) | = | 介護保険 (c)        | 円 |
| (5) 1か月の「雇用保険料」       | 円 × (1) | = | 雇用保険 (d)        | 円 |

③ 計算した『02 医療保険(短期掛金)』から『05 雇用保険』の合計を千円単位で記入

社会保険料の総額 (e) 千円 ※ 10ペーシの『01 社会保険料の総額』欄へ必ず記入してください。

出所)統計委員会 第34回人口・社会統計部会配布資料

国民生活基礎調査(平成 25 年) 所得票(案)

**質問6 平成24年度の固定資産税の課税はありましたか。**

あてはまる番号に○をつけ、1に○をつけた方は金額も記入してください。

1 課税あり → 固定資産税 [ ] 万 [ ] 千円

2 課税なし

※ 個人所有の土地・家屋(事業関係分は除く)に対する税額です。土地・家屋の名称や面積がわからず、実際に支払った方が記入してください。

【参考書類】平成24年度 固定資産税・都市計画税通知書

---

**質問7 平成24年分の企業年金・個人年金等の掛金を支払いましたか。**

あてはまる番号に○をつけ、1に○をつけた方は金額も記入してください。

1 支払いあり → 企業年金・個人年金等 [ ] 万 [ ] 千円

2 支払いなし

※ 掛金は、実際に支払った方が記入してください。

【企業年金の例】 確定拠出企業年金・確定拠出年金(企業型)などの本人拠出分  
 【個人年金の例】 生命保険会社・かんぽ生命・銀行・証券会社などの個人年金や年金型商品  
 【その他】 国民年金基金、養老者年金、確定拠出年金(個人型)

出所)統計委員会 第34回人口・社会統計部会配布資料

国民生活基礎調査(平成 25 年) 所得票(案)

世帯主 又は 世帯を代表する方 は  
引きつづき下の 質問 8 にもお答えください。

質問 8 現在の暮らしの状況を総合的にみて、どう感じていますか。  
あてはまる番号1つに○をつけてください。

1 大変苦しい  
2 やや苦しい  
3 普通  
4 ややゆとりがある  
5 大変ゆとりがある

ご記入ありがとうございました。

### 3) 研究概要

#### 国民生活基礎調査 研究概要

**世帯票情報を用いた傾向スコアによる  
総所得金額の推定について**

国民生活基礎調査（大規模世帯）は、世帯票・健康票・介護票及び所得票・消費票のちつから成り、所得票は先に行う世帯票において抽出された国勢調査区から再度、地区を抽出することで、調査対象を選び出している（正確には、その後で単位区とよばれる国勢調査区を、3群に分割した単位を抽出する二段工程抽）。この二段抽出の性質を利用し、傾向スコアを国勢調査指標として用いることで、世帯票調査世帯における所得票の無回答データの補整を行った。なお、本補整においては、所得票の無回答は総所得の値に依存しないものと仮定している。

**1 本補整試算の概要**

本補整試算においては、まず、所得票が無回答の傾向スコア（所得票回答率）の推定を行い、その推定値を用いて、IPW (Inverse Probability Weighting) 推定及び傾向スコアマッチングによる傾向推定の処理を行った。前者においては、2段傾向スコア推定値による無回答率と回答世帯のマッチングを行い、マッチングした回答世帯の傾向値を用いて、回答世帯の傾向値にプロキシ変数として変換入及びプロキシ変数変換を行った。

傾向スコアマッチングとプロキシ変数変換入のイメージ

① 世帯票調査世帯（世帯票の回答）と所得票調査世帯（所得票の回答）のマッチングを行う。  
② マッチングされた所得票調査世帯に、回答世帯の傾向値をプロキシ変数として、傾向スコア変換モデルを用いて、傾向スコア変換を行う。

出所)厚生労働省統計情報部 提供資料

#### 国民生活基礎調査 研究概要

**2 傾向スコアの推定**

回答世帯の回答率のうち、第1世帯が回答したとき1、未回答のとき0をとる大規模インデックス変数を定義する。これを抽出したn個の世帯票回答率変数(無回答)を  $x_0, x_{10}, \dots, x_{n0}$  としたとき、第1世帯が回答する確率

$$p_1 = p(x_0 = 1 | x_1, x_2, \dots, x_n)$$

を第1世帯の傾向スコア (Propensity score) とする。 $x_n$  の負値は2以外の任意の値、その推定値をロジスティック関数モデルにより求める。すなわち、

$$p_1 = p(x_0 = 1 | x_1, x_2, \dots, x_n) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)]}$$

における  $\beta_j$  の最尤推定値  $\hat{\beta}_j$  を最尤法の手法により決定する<sup>1)</sup>。このとき

$$\hat{p}_1 = \frac{1}{1 + \exp[-(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_n x_n)]}$$

を第1世帯の傾向スコア推定値とする。

1) 最尤法としての実現  $\chi^2$  テスト。

回答・非回答に対して、“1”か“0”を割り当てる

|       | 回答世帯 (z=1) |       |     | 非回答世帯 (z=0) |     |                         |
|-------|------------|-------|-----|-------------|-----|-------------------------|
| 所属群 z | 1          | 1     | 1   | 1           | 0   | 0                       |
| 世帯番号  | 1          | 2     | ... | 23513       | ... | 36396 36397             |
| 総所得 y | $y_1$      | $y_2$ | ... | $y_{23513}$ | ... | $y_{36396}$ $y_{36397}$ |
| 共変量 x | $x_1$      | $x_2$ | ... | $x_{23513}$ | ... | $x_{36396}$ $x_{36397}$ |

抽出した世帯票項目とプロキシ変数(傾向スコア)の結果は次の表のとおり。

1) 統計解析ソフトウェアとして、SAS® Enterprise Guide®6.2 (SAS Institute Japan Ltd.) を用いた。本試算の本質も同様。

出所)厚生労働省統計情報部 提供資料

国民生活基礎調査 研究概要

表1: ロジスティック回帰による分析結果

| 説明変数(1/2)                    | 所得対象世帯         |                        | 説明変数(2/2)                        | 所得対象世帯         |                        |
|------------------------------|----------------|------------------------|----------------------------------|----------------|------------------------|
|                              | 係数<br>$\alpha$ | オッズ比<br>$\exp(\alpha)$ |                                  | 係数<br>$\alpha$ | オッズ比<br>$\exp(\alpha)$ |
| <b>市郡 (R=大都市)</b>            |                |                        | <b>世帯類型 (R=高齢者世帯)</b>            |                |                        |
| 人口15万人以上の市                   | 0.229          | 1.258 ***              | 母子世帯                             | -0.346         | 0.700 **               |
| 人口5万人以上15万人未満の市              | 0.539          | 1.715 ***              | 父子世帯                             | -0.498         | 0.608                  |
| 人口5万人未満の市                    | 0.932          | 2.540 ***              | その他の世帯                           | -0.310         | 0.734 ***              |
| 都府                           | 0.705          | 2.024 ***              | <b>世帯業態 (R=常雇者(会社・団体等の役員)世帯)</b> |                |                        |
| <b>地域ブロック (R=北海道)</b>        |                |                        | 常雇者(一般常雇者)世帯                     | 0.186          | 1.205 ***              |
| 東北                           | 0.341          | 1.406 ***              | 契約の雇用者世帯                         | 0.162          | 1.176 *                |
| 関東-I                         | 0.104          | 1.109                  | 自営業者                             | 0.087          | 1.091                  |
| 関東-II                        | 0.341          | 1.407 ***              | その他                              | 0.003          | 1.003                  |
| 北陸                           | 0.389          | 1.475 ***              | 不詳                               | -0.259         | 0.772 *                |
| 東海                           | 0.345          | 1.412 ***              | <b>家計支出額 (R=25万円未満)</b>          |                |                        |
| 近畿-I                         | 0.064          | 1.066                  | 25万円以上50万円未満                     | -0.063         | 0.939 *                |
| 近畿-II                        | 0.004          | 1.004                  | 50万円以上                           | -0.155         | 0.856 **               |
| 中国                           | 0.220          | 1.246 **               | 不詳                               | -0.638         | 0.528 ***              |
| 四国                           | 0.006          | 1.006                  | <b>住居の種類 (R=持ち家)</b>             |                |                        |
| 北九州                          | 0.388          | 1.474 ***              | 民間賃貸住宅                           | -0.153         | 0.858 ***              |
| 南九州                          | 0.373          | 1.453 ***              | 社宅・公務員住宅等の給与住宅                   | -0.049         | 0.952                  |
| <b>世帯人員数</b>                 | 0.045          | 1.046 *                | 公社・公団等の賃貸公営住宅                    | 0.009          | 1.009                  |
| <b>有業人員数 (R=0人)</b>          |                |                        | 借間・その他                           | -0.331         | 0.718 ***              |
| 1~2人                         | -0.097         | 0.908                  | <b>世帯主年齢階級 (R=25歳未満)</b>         |                |                        |
| 3人以上                         | -0.249         | 0.780 ***              | 25歳以上35歳未満                       | 0.303          | 1.353 ***              |
| 不詳                           | -0.484         | 0.616 ***              | 35歳以上45歳未満                       | 0.336          | 1.399 ***              |
| <b>世帯構造 (R=住み込み・寄宿舍等の単独)</b> |                |                        | 45歳以上55歳未満                       | 0.278          | 1.321 ***              |
| その他の単独世帯(男)                  | 1.051          | 2.859 ***              | 55歳以上65歳未満                       | 0.421          | 1.523 ***              |
| その他の単独世帯(女)                  | 1.120          | 3.088 ***              | 65歳以上75歳未満                       | 0.667          | 1.949 ***              |
| 夫婦のみの世帯                      | 1.346          | 3.840 ***              | 75歳以上                            | 0.628          | 1.874 ***              |
| 夫婦と未婚の子のみの世帯                 | 1.372          | 3.945 ***              | 不詳                               | -2.837         | 0.059 ***              |
| ひとり親と未婚の子のみの世帯               | 1.315          | 3.723 ***              |                                  |                |                        |
| 三世帯世帯                        | 1.402          | 4.064 ***              |                                  |                |                        |
| その他の世帯                       | 1.262          | 3.604 ***              |                                  |                |                        |
|                              |                |                        | <b>c統計量</b>                      | <b>0.657</b>   |                        |

※Rはダミー変数の基準カテゴリ(Reference category)を表す

\* = p < 0.05, \*\* = p < 0.01, \*\*\* = p < 0.001

出所)厚生労働省統計情報部 提供資料

□ 傾向スコア推定値の分布

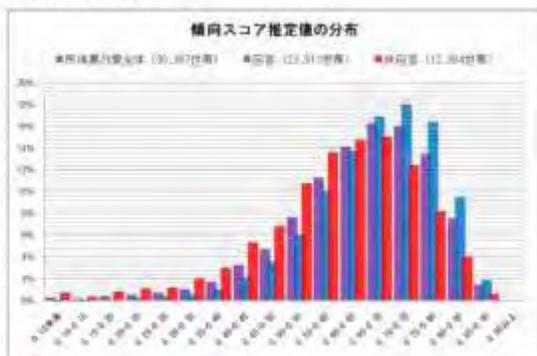


図1

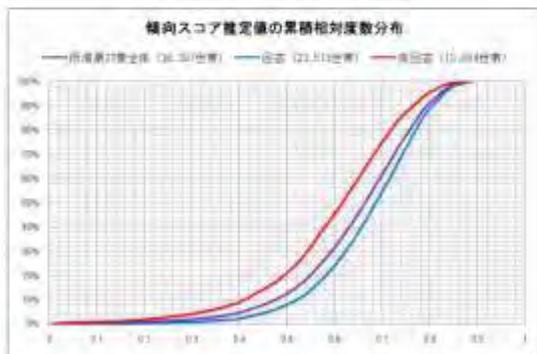


図2

出所)厚生労働省統計情報部 提供資料

□ IPW (Inverse Probability Weighting) 推定

各Y世帯の傾向スコア推定値を、傾向スコア $x_i$ に対して、実行の対大乗数 $w_i$ を $\frac{1}{K_i}w_i$ に置き換えて推定を行う。このとき、両の傾向スコアは異なるため、傾向スコア推定値 $x_i$ による $\frac{1}{K_i}w_i$ を対大乗数として用いた。結果については、図3及び表2のとおり。

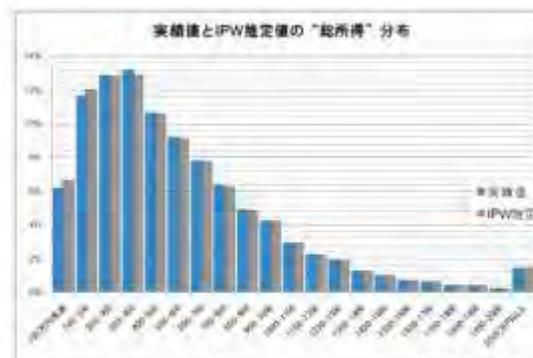


図3

表2: 実値及びIPW推定値における平均所得額と所得分位値

| 一世帯あたり平均所得(万円) | 所得分位値(万円) |        |       |        |     |
|----------------|-----------|--------|-------|--------|-----|
|                | 実値        | IPW推定値 | 実値    | IPW推定値 |     |
| 全世界            | 568.8     | 564.0  | 中央値   | 451    | 450 |
| 高齢者世帯          | 306.3     | 303.0  | 第1五分位 | 214    | 207 |
| 母子世帯           | 236.7     | 236.4  | 第2五分位 | 365    | 360 |
| 父子世帯           | 515.9     | 515.9  | 第3五分位 | 554    | 550 |
| その他の世帯         | 646.8     | 634.8  | 第4五分位 | 838    | 834 |

※所得額の対大乗数は、調査地区別をもとに総世帯数・世帯推定数割りに行なわれているため、所得層毎にも対大乗数と同じ対大乗数を算出したこととなる。

出所)厚生労働省統計情報部 提供資料

4 マッチングによる代入

4.1 マッチング

傾向スコア推定値の算出(※1)による、1対1マッチングを行い、推定値が一致(※2)した試算①、②

● 最近隣マッチング (Nearest Neighbor matching)

- ▶ 試算①、②
- ▶ 試算③、④

● キリシローマッチング (Caliper matching)

傾向スコア推定値の差の絶対値が 0.001 より小さくなる推定値がない場合、マッチングを行わないこととして、

- ▶ 試算③、④

マッチング対象における傾向スコア推定値、総所得及び各共業種の分布については、図4-15のとおり。凡例の試算①～④はすべて厚労省調査表の一部である。

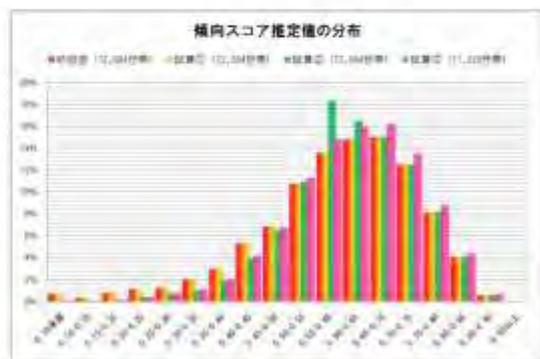


図4

※1傾向スコア推定値は、ここでは実行回数を1000回とした場合の平均傾向スコア推定値によるゲートを採用した。

出所)厚生労働省統計情報部 提供資料

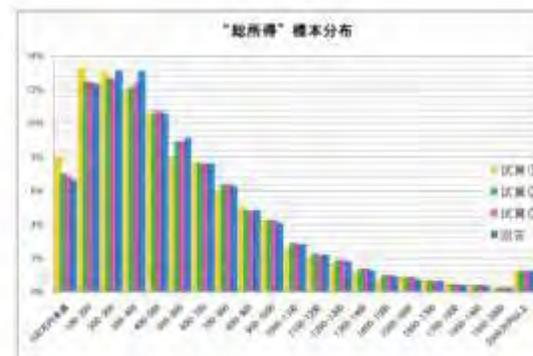


図5-1

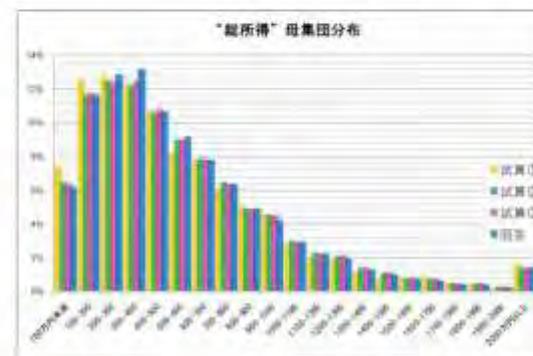


図5-2

出所)厚生労働省統計情報部 提供資料

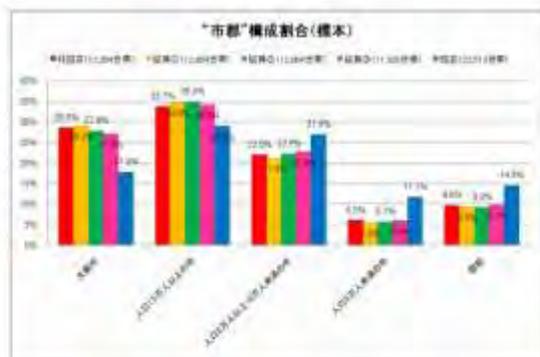


図 6-1

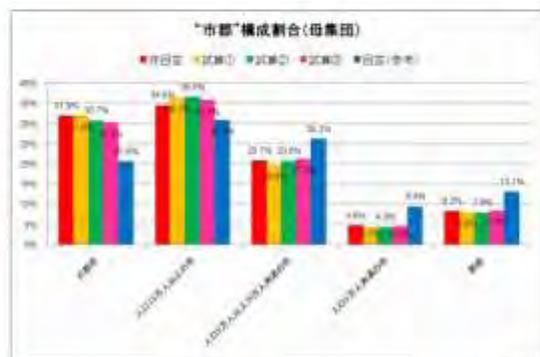


図 6-2

8

出所)厚生労働省統計情報部 提供資料

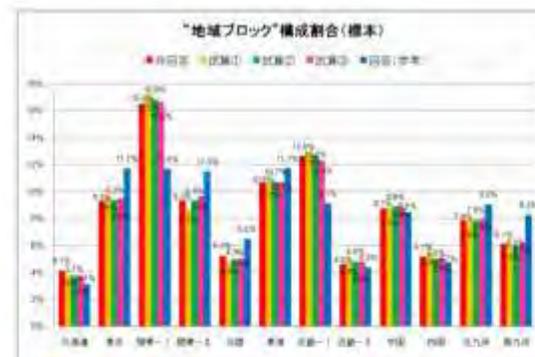


図 7-1

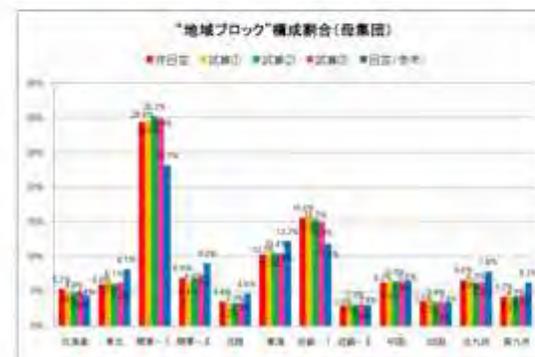


図 7-2

9

出所)厚生労働省統計情報部 提供資料

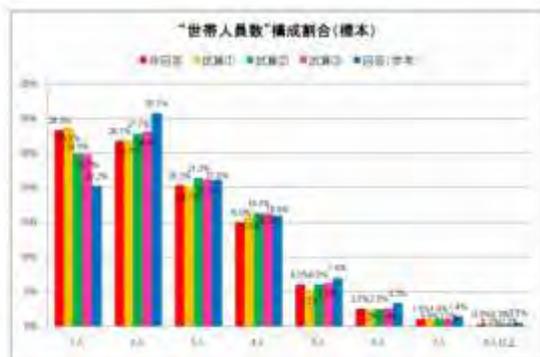


図 8-1



図 8-2

出所)厚生労働省統計情報部 提供資料

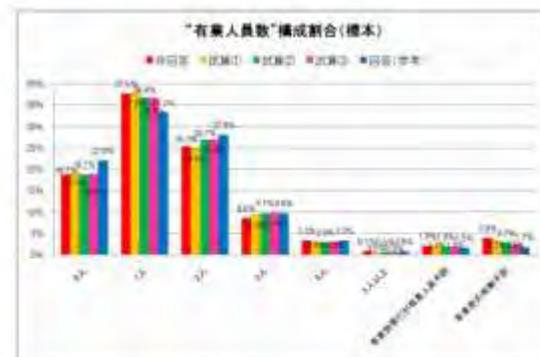


図 9-1

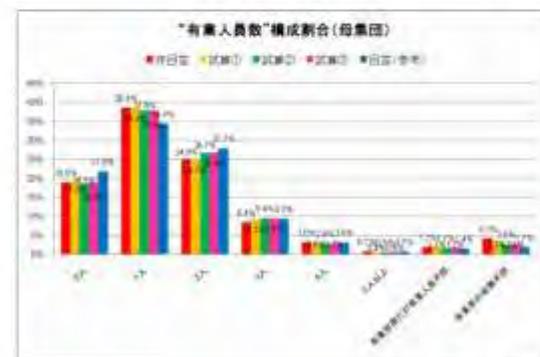


図 9-2

出所)厚生労働省統計情報部 提供資料

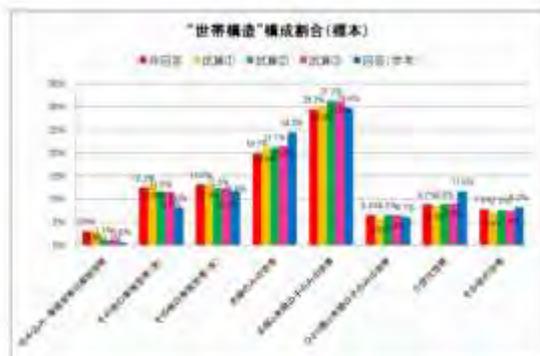


図 10-1

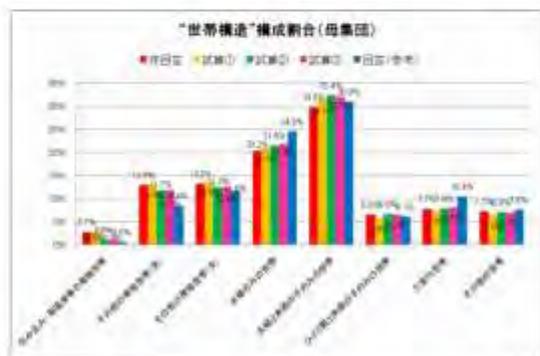


図 10-2



図 11-1

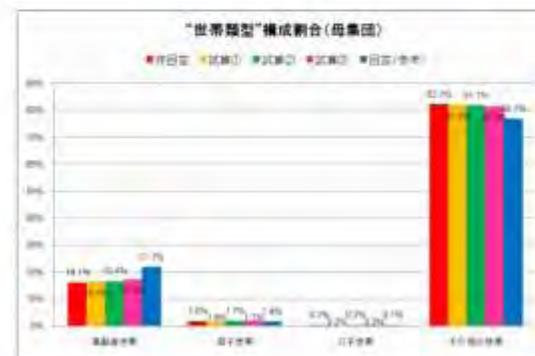


図 11-2

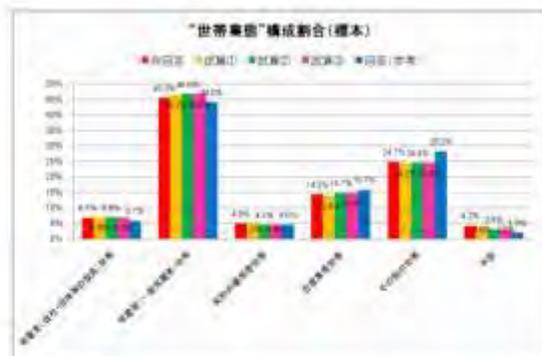


図 12-1

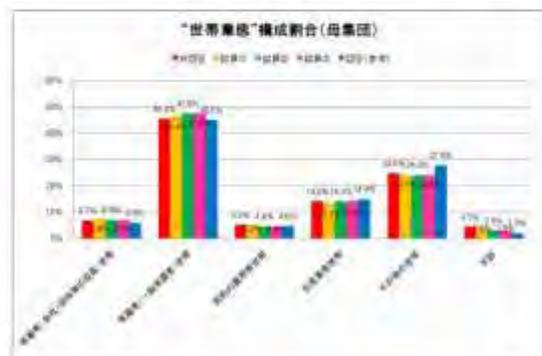


図 12-2

出所)厚生労働省統計情報部 提供資料

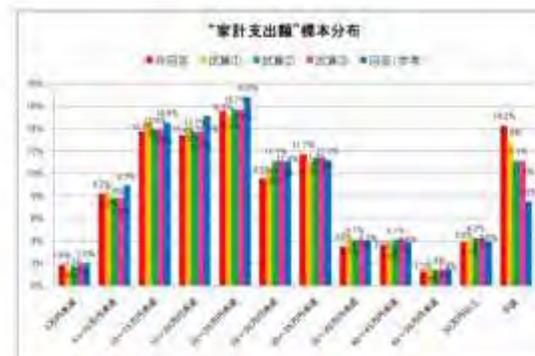


図 13-1

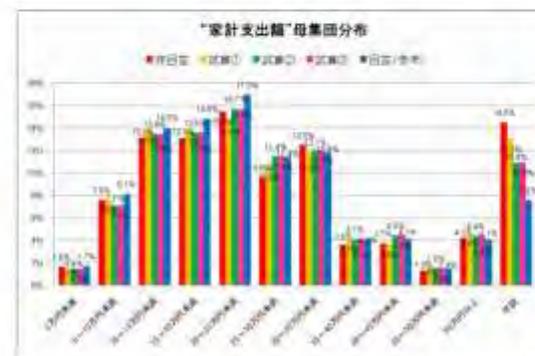


図 13-2

出所)厚生労働省統計情報部 提供資料

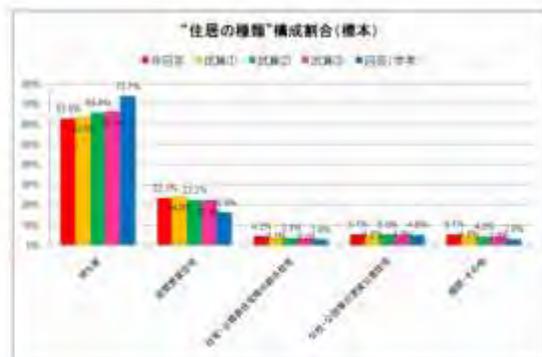


図 14-1

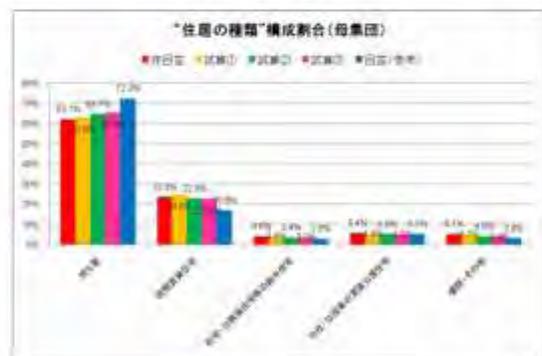


図 14-2

出所)厚生労働省統計情報部 提供資料



図 15-1



図 15-2

出所)厚生労働省統計情報部 提供資料

4.2 インパラメトリックな代入

一階一般化ランダムウォークにおいては、所得増減割合を基準に群して、相対となる所得増減割合が存在する。このランダムウォーク下の所得増減をそのまま無回答の補完値に代入することにより、無回答データの補完を行った。なお、補充ポイントでは(1)が参考として、(2)および(3)のみの場合も記載した。結果は表10-1及び図10-1のとおり。

表10-1 実績値及びインパラメトリック代入における平均所得額と所得分位値

| 世帯あたり<br>平均所得(万円) | 実績値   |       | 試算①   |       | 試算②   |       | 試算③(参考値) |     |
|-------------------|-------|-------|-------|-------|-------|-------|----------|-----|
|                   | 補充前   | 補充後   | 補充前   | 補充後   | 補充前   | 補充後   | 補充前      | 補充後 |
| 全世界               | 588.8 | 549.1 | 530.7 | 564.5 | 586.8 | 581.2 | 564.8    |     |
| 高齢者世帯             | 306.8 | 376.1 | 327.2 | 452.5 | 390.8 | 470.5 | 348.5    |     |
| 母子世帯              | 238.7 | 352.3 | 281.1 | 591.4 | 379.0 | 581.3 | 368.3    |     |
| 父子世帯              | 515.9 | 334.7 | 481.0 | 540.8 | 526.1 | 528.3 | 521.7    |     |
| その他の世帯            | 848.8 | 578.8 | 620.4 | 585.9 | 823.2 | 580.5 | 428.2    |     |
| 所得分位値(万円)         |       |       |       |       |       |       |          |     |
| 中央値               | 451   | 439   | 448   | 450   | 450   | 450   | 450      |     |
| 第1五分位             | 214   | 189   | 203   | 206   | 210   | 200   | 210      |     |
| 第2五分位             | 365   | 342   | 358   | 380   | 364   | 360   | 385      |     |
| 第3五分位             | 554   | 539   | 550   | 557   | 550   | 550   | 553      |     |
| 第4五分位             | 838   | 850   | 834   | 850   | 842   | 847   | 840      |     |

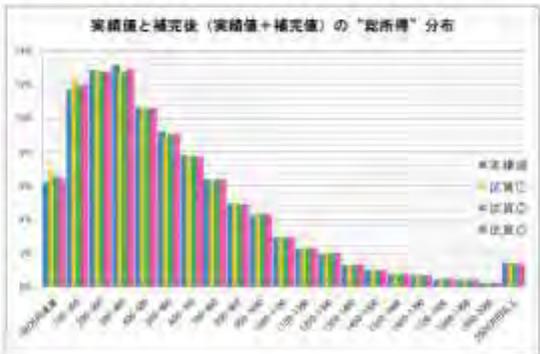


図 10-1

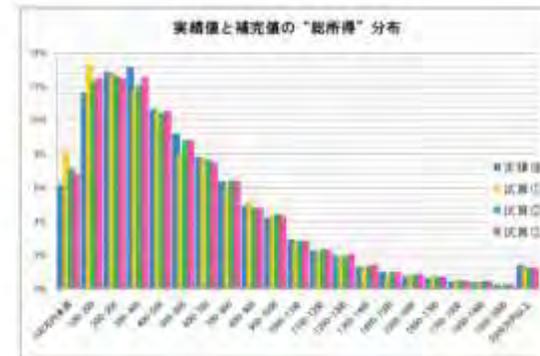


図 10-2 所得増減割合に対する個人属性で補正しているため、補完値と結果が異なる。

4.3 パラメトリックな代入

ランダムウォークモデルの推定結果に基づいて、所得増減割合  $\log(Y_t)$  を説明変数、 $X_t$  類の世帯属性変数  $X_{1t}, X_{2t}, \dots, X_{mt}$  を説明変数として、連続確率モデル  $\log(Y_t) = \beta_0 + \beta_1 X_{1t} + \dots + \beta_m X_{mt} + \epsilon_t$  を用いて、 $\epsilon_t$  の確率密度関数  $f(\epsilon)$  を見出し、

$$\begin{aligned} \hat{f}_t &= \left(1 + \frac{1}{2} \epsilon^2\right) \exp(\beta_0 + \beta_1 X_{1t} + \dots + \beta_m X_{mt}) \quad \text{--- (4)} \\ \hat{f}_t &= \exp(\beta_0 + \beta_1 X_{1t} + \dots + \beta_m X_{mt}) \\ \hat{f}_t &= \exp(\beta_0 + \beta_1 X_{1t} + \dots + \beta_m X_{mt} + \epsilon_t) \end{aligned}$$

をそれぞれ実際の総所得の補完値として代入する (4) の場合(1)が参照。

## 国民生活基礎調査 研究概要

ただし、 $\sigma^2 = \frac{1}{M-m-1} \sum \hat{\varepsilon}_i^2$  であり、 $\hat{\varepsilon}_i = \log(Y_i) - (\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \dots + \hat{\beta}_m X_{im})$  (第  
k世帯の回帰残差)、 $\hat{\varepsilon}_i \sim N(0, \sigma^2)$  とする。

国民生活基礎調査 研究概要

表4：世帯構による分析結果(1/2)

| 説明変数(1/4) | 世帯タイプA |        |        | 世帯タイプB |        |        | 説明変数(2/4) | 世帯タイプC |        |        | 世帯タイプD |        |  |
|-----------|--------|--------|--------|--------|--------|--------|-----------|--------|--------|--------|--------|--------|--|
|           | 係数     | 係数     | 係数     | 係数     | 係数     | 係数     |           | 係数     | 係数     | 係数     | 係数     | 係数     |  |
| 世帯タイプA    | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  | 世帯タイプB    | -0.200 | -0.200 | -0.200 | -0.200 | -0.200 |  |
| 世帯タイプB    | -0.200 | -0.200 | -0.200 | -0.200 | -0.200 | -0.200 | 世帯タイプC    | -0.400 | -0.400 | -0.400 | -0.400 | -0.400 |  |
| 世帯タイプC    | -0.400 | -0.400 | -0.400 | -0.400 | -0.400 | -0.400 | 世帯タイプD    | -0.600 | -0.600 | -0.600 | -0.600 | -0.600 |  |
| 世帯タイプD    | -0.600 | -0.600 | -0.600 | -0.600 | -0.600 | -0.600 |           |        |        |        |        |        |  |

11

出所)厚生労働省統計情報部 提供資料

国民生活基礎調査 研究概要

表4：世帯構による分析結果(2/2)

| 説明変数(1/4) | 世帯タイプA |        |        | 世帯タイプB |        |        | 説明変数(2/4) | 世帯タイプC |        |        | 世帯タイプD |        |  |
|-----------|--------|--------|--------|--------|--------|--------|-----------|--------|--------|--------|--------|--------|--|
|           | 係数     | 係数     | 係数     | 係数     | 係数     | 係数     |           | 係数     | 係数     | 係数     | 係数     | 係数     |  |
| 世帯タイプA    | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  | 世帯タイプB    | -0.200 | -0.200 | -0.200 | -0.200 | -0.200 |  |
| 世帯タイプB    | -0.200 | -0.200 | -0.200 | -0.200 | -0.200 | -0.200 | 世帯タイプC    | -0.400 | -0.400 | -0.400 | -0.400 | -0.400 |  |
| 世帯タイプC    | -0.400 | -0.400 | -0.400 | -0.400 | -0.400 | -0.400 | 世帯タイプD    | -0.600 | -0.600 | -0.600 | -0.600 | -0.600 |  |
| 世帯タイプD    | -0.600 | -0.600 | -0.600 | -0.600 | -0.600 | -0.600 |           |        |        |        |        |        |  |

12

出所)厚生労働省統計情報部 提供資料

(ア) 回答代入（不備推定）

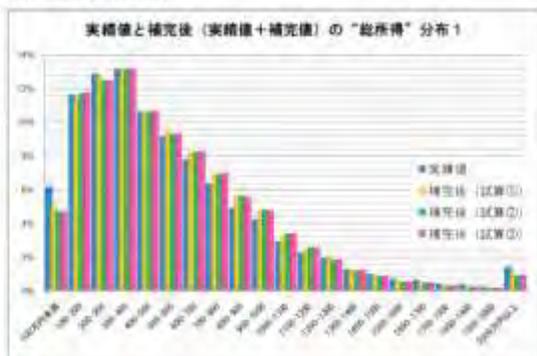


図 17-1

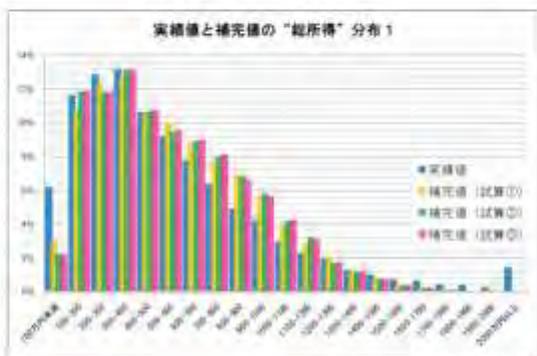


図 17-2

出所)厚生労働省統計情報部 提供資料

(イ) 回答代入（最尤推定）

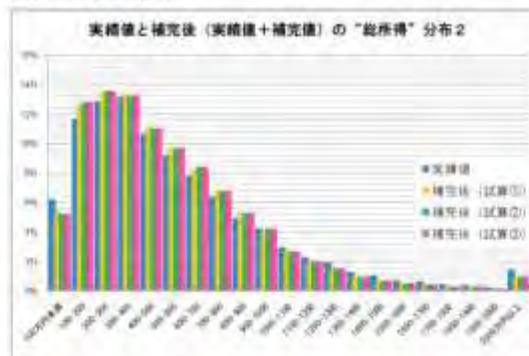


図 18-1

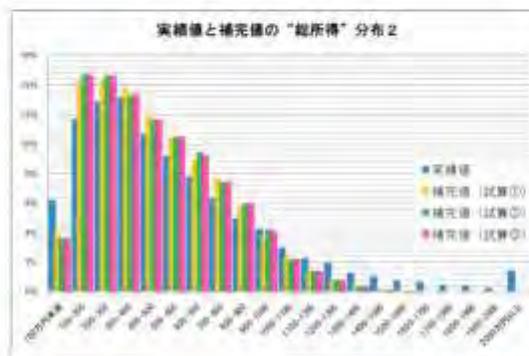


図 18-2

出所)厚生労働省統計情報部 提供資料

(ウ) 標準的所得代入（調査変動の比較については、付録A2を参照）

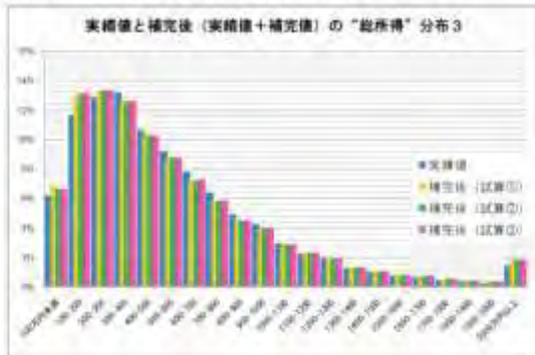


図 19-1

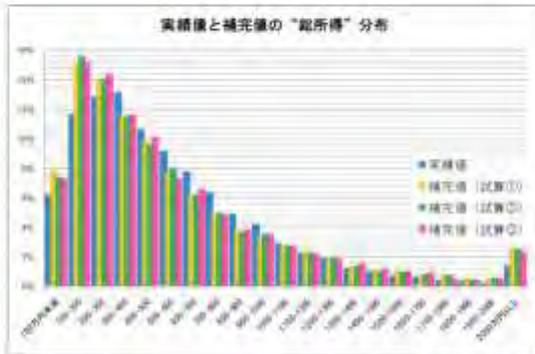


図 19-2

パラメトリック代入の結果をまとめると、表5のとおりとなる。

表5: 所得値の実総値及び補充後の平均所得額と所得分位値

|           | 実総値   |       |       | 所得代入(不偏推定) |       |       | 所得代入(最大推定) |       |       | 標準的所得代入 |     |     |
|-----------|-------|-------|-------|------------|-------|-------|------------|-------|-------|---------|-----|-----|
|           | 試算①   | 試算②   | 試算③   | 試算①        | 試算②   | 試算③   | 試算①        | 試算②   | 試算③   | 試算①     | 試算② | 試算③ |
| 一様値あたり    |       |       |       |            |       |       |            |       |       |         |     |     |
| 平均所得(万円)  |       |       |       |            |       |       |            |       |       |         |     |     |
| 全世帯       | 392.8 | 365.0 | 364.2 | 364.0      | 334.1 | 323.6 | 323.6      | 368.8 | 368.2 | 367.8   |     |     |
| 高齢者世帯     | 336.1 | 324.7 | 324.1 | 324.3      | 291.3 | 287.7 | 291.1      | 326.1 | 325.8 | 326.1   |     |     |
| 母子世帯      | 236.7 | 239.9 | 239.8 | 240.9      | 225.5 | 224.9 | 229.0      | 236.7 | 236.9 | 242.1   |     |     |
| 父子世帯      | 315.9 | 312.9 | 320.3 | 310.8      | 478.7 | 467.6 | 482.2      | 691.3 | 631.9 | 606.3   |     |     |
| その他の世帯    | 346.8 | 330.7 | 325.2 | 324.8      | 368.8 | 365.1 | 365.3      | 340.8 | 338.9 | 338.9   |     |     |
| 所得分位値(万円) |       |       |       |            |       |       |            |       |       |         |     |     |
| 中央値       | 451   | 472   | 473   | 472        | 442   | 445   | 444        | 438   | 438   | 438     |     |     |
| 第1五分位     | 214   | 227   | 226   | 226        | 212   | 213   | 213        | 200   | 200   | 200     |     |     |
| 第2五分位     | 265   | 282   | 282   | 282        | 260   | 260   | 260        | 250   | 250   | 250     |     |     |
| 第3五分位     | 354   | 375   | 375   | 374        | 340   | 340   | 340        | 342   | 342   | 341     |     |     |
| 第4五分位     | 430   | 540   | 547   | 540        | 368   | 368   | 368        | 340   | 340   | 340     |     |     |

### 5 まとめ

今回、経済的困難層における共定数調整の手法である、傾向スコアを所得層の回答確率の推定に用いることで、世帯別回答世帯の所得推測回答に対して、所得分布の補整を試みを行った。

傾向スコア推定の結果から、回答実績においては経済困難層の回答率が若干低いと考えられるため、所得分布の補整方法としてはIPW推定やノンパラメトリック代入の結果が補整後のイメージにより近いように思われる。しかしながら、表5のとおり母子世帯や父子世帯のような出現頻度の低い属性においては、属性の異なる世帯とマッチングされる可能性が高くなり、個別の所得分布を見ると高所得と著しく乖離した結果となることもあるため、ノンパラメトリック代入による補整には注意が必要である。一方で、パラメトリック代入は、結果がモデル選択の良否に左右される傾向にあるため、IPW推定及びノンパラメトリック代入と比較した推定の良し悪しは一断には言えない。

上述より推測のとおり、傾向スコア推定値でマッチングされた回答世帯の共定数分布は非回答世帯の共定数分布に近づくことが実際に検証できたため、所得層の無回答が所得層の偏りに起因しないランダムな欠測(Missing at random)であり、なおかつ共定数の推定が適切であれば、全体の所得分布に関する有効な補整法であると考えられる。しかしながら、真の所得分布が未知であり、補整結果の妥当性の検証が行えないため、いずれの補整手法も偏りに適用可能であると結論づけることまではできない。

付録 A 参考

A.1 不偏推定置換代人  $\hat{Y}_i$  の導出

$$\log(Y_i) = \beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im} + \varepsilon_i$$

から、

$$Y_i = \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im} + \varepsilon_i) \\ = \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im}) \exp(\varepsilon_i)$$

となり、 $Y_i$  の条件付き期待値を考えることで、

$$E(Y_i | X_{i1}, \dots, X_{im}) \\ = \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im}) E(\exp(\varepsilon_i) | X_{i1}, \dots, X_{im}) \\ = \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im}) E\left(\sum_{n=1}^{\infty} \frac{\varepsilon_i^n}{n!}\right)$$

と資料できる。ここで、Taylor 展開の二次近似を用いれば、

$$= \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im}) E\left(1 + \varepsilon_i + \frac{\varepsilon_i^2}{2}\right) \\ = \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im}) \left(1 + E(\varepsilon_i) + \frac{E(\varepsilon_i^2)}{2}\right)$$

さらに、 $\varepsilon_i \sim N(0, \sigma^2)$  なので、 $E(\varepsilon_i) = 0$  より、

$$= \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_m X_{im}) \left(1 + \frac{1}{2} E(\varepsilon_i^2)\right)$$

を推定すればよいことがわかる。

このとき、 $\beta_0$  を  $\hat{\beta}_0$ 、 $E(\varepsilon_i^2) = \sigma^2$  を  $\sigma^2 = \frac{1}{M-m-1} \sum_{i=1}^M \hat{\varepsilon}_i^2$  で推定することにより、 $\hat{Y}_i$  を求めることとした。

A.2 シード値を変えた場合の標準的置換代入の結果

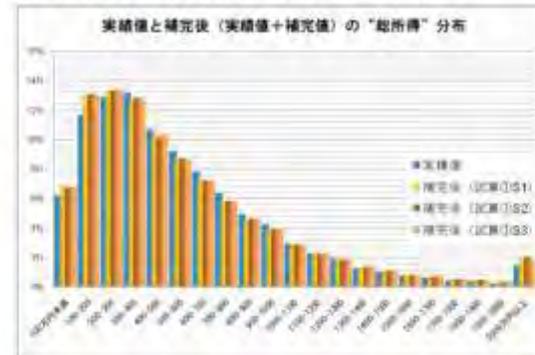


図 A-1

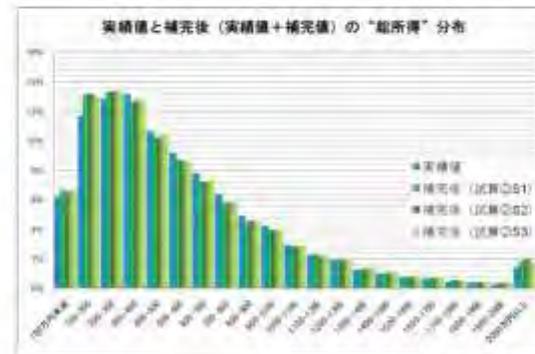
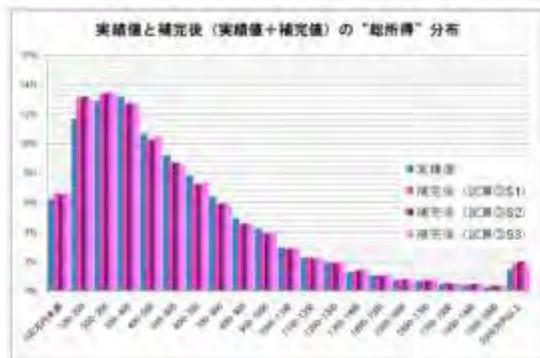


図 A-2

国民生活基礎調査 研究概要



図A-3





2) 調査票例

特定サービス産業実態調査(平成 22 年) 調査票(ソフトウェア業等)

This is a detailed survey form for the software industry. It includes sections for company information, business overview, and financial data. The form is filled out with handwritten text and includes several tables for data entry. A large watermark '見本' (Sample) is overlaid on the form.

出所)経済産業省 特定サービス産業実態調査 ウェブサイト

特定サービス産業実態調査(平成 22 年) 調査票(映像情報制作・配給業)

This is a detailed survey form for the video information production and distribution industry. It includes sections for company information, business overview, and financial data. The form is filled out with handwritten text and includes several tables for data entry. A large watermark '見本' (Sample) is overlaid on the form.

出所)経済産業省 特定サービス産業実態調査 ウェブサイト



特定サービス産業実態調査(平成 22 年) 調査票(各種物品賃貸業)

物品賃貸業調査票

事業者基本情報

経営概況及び事業内容

財務状況

従業員数

裏面に続く

出所)経済産業省 特定サービス産業実態調査 ウェブサイト

特定サービス産業実態調査(平成 22 年) 調査票(デザイン業等)

デザイン業、機械設計業調査票

事業者基本情報

経営概況及び事業内容

財務状況

従業員数

見本

出所)経済産業省 特定サービス産業実態調査 ウェブサイト

特定サービス産業実態調査(平成 22 年) 調査票(広告業)

This is a detailed survey form for the advertising industry. It includes sections for:
 

- 1. 調査票の取扱い (Handling of the survey form)
- 2. 調査票の記入方法 (Method of filling out the survey form)
- 3. 調査票の記入例 (Example of filling out the survey form)
- 4. 調査票の記入欄 (Survey form input fields)
- 5. 調査票の記入例 (Example of filling out the survey form)
- 6. 調査票の記入欄 (Survey form input fields)
- 7. 調査票の記入欄 (Survey form input fields)
- 8. 調査票の記入欄 (Survey form input fields)
- 9. 調査票の記入欄 (Survey form input fields)
- 10. 調査票の記入欄 (Survey form input fields)

 The form contains various tables and text boxes for data entry, with a large '見本' (Sample) watermark overlaid.

出所)経済産業省 特定サービス産業実態調査 ウェブサイト

特定サービス産業実態調査(平成 22 年) 調査票(計量証明業)

This is a detailed survey form for the metrology certification industry. It includes sections for:
 

- 1. 調査票の取扱い (Handling of the survey form)
- 2. 調査票の記入方法 (Method of filling out the survey form)
- 3. 調査票の記入例 (Example of filling out the survey form)
- 4. 調査票の記入欄 (Survey form input fields)
- 5. 調査票の記入欄 (Survey form input fields)
- 6. 調査票の記入欄 (Survey form input fields)
- 7. 調査票の記入欄 (Survey form input fields)
- 8. 調査票の記入欄 (Survey form input fields)
- 9. 調査票の記入欄 (Survey form input fields)
- 10. 調査票の記入欄 (Survey form input fields)

 The form contains various tables and text boxes for data entry, with a large '見本' (Sample) watermark overlaid.

出所)経済産業省 特定サービス産業実態調査 ウェブサイト

### 3) 研究概要

平成 21 年度サービス産業統計等経済産業統計の基盤整備事業

『特定サービス産業実態調査等における推計手法の確立に関する調査研究』調査報告書

[http://www.meti.go.jp/meti\\_lib/report/2010fy01/0022971.pdf](http://www.meti.go.jp/meti_lib/report/2010fy01/0022971.pdf)

4. 全国企業短期経済観測調査

1) 調査票例

全国企業短期経済観測調査(平成 25 年 3 月調査) 調査票

This is a detailed survey form for the March 2013 survey. It is divided into several main sections: 1. 会社情報 (Company Information) at the top right, including company name and industry. 2. 生産実績 (Production Results) in the middle left, with a table for monthly production and inventory. 3. 在庫計画 (Inventory Planning) in the bottom left, with a table for planned inventory. 4. 財務計画 (Financial Planning) in the middle right, with a table for planned financial data. 5. 参考情報 (Reference Information) at the bottom right, including a section for comments. The form includes various tables with columns for months and rows for different metrics. There are also checkboxes and text boxes for additional information.

出所) 日本銀行 全国企業短期経済観測調査 ウェブサイト

全国企業短期経済観測調査(平成 24 年 12 月調査) 調査票

This is a detailed survey form for the December 2012 survey. It is divided into several main sections: 1. 会社情報 (Company Information) at the top right, including company name and industry. 2. 生産実績 (Production Results) in the middle left, with a table for monthly production and inventory. 3. 在庫計画 (Inventory Planning) in the bottom left, with a table for planned inventory. 4. 財務計画 (Financial Planning) in the middle right, with a table for planned financial data. 5. 参考情報 (Reference Information) at the bottom right, including a section for comments. The form includes various tables with columns for months and rows for different metrics. There are also checkboxes and text boxes for additional information.

出所) 日本銀行 全国企業短期経済観測調査 ウェブサイト

全国企業短期経済観測調査(平成 24 年 9 月調査) 調査票

This is a detailed survey form for the 2012 national short-term economic survey, September survey. It is divided into several main sections:

- 1. 調査票の取扱い (Handling of the Survey Form):** Contains instructions for respondents regarding confidentiality and data usage.
- 2. 調査票の記入 (Filling out the Survey Form):** Includes a header for company name, address, and survey date (September 2012).
- 3. 生産項目 (Production Items):** A table for recording production and sales data for various categories, including a bar chart for visual representation.
- 4. 得意先別状況 (Customer Status):** A section for recording sales data by customer type, including a bar chart.
- 5. 参考情報 (Reference Information):** A section for providing additional information about the company's financial and operational status.

The form is marked with a 'Secret' (秘) and 'Confidential' (貴社控) label, indicating that the data collected is sensitive and should be handled accordingly.

出所) 日本銀行 全国企業短期経済観測調査 ウェブサイト

全国企業短期経済観測調査(平成 24 年 6 月調査) 調査票

This is a detailed survey form for the 2012 national short-term economic survey, June survey. It is divided into several main sections:

- 1. 調査票の取扱い (Handling of the Survey Form):** Contains instructions for respondents regarding confidentiality and data usage.
- 2. 調査票の記入 (Filling out the Survey Form):** Includes a header for company name, address, and survey date (June 2012).
- 3. 生産項目 (Production Items):** A table for recording production and sales data for various categories, including a bar chart for visual representation.
- 4. 得意先別状況 (Customer Status):** A section for recording sales data by customer type, including a bar chart.
- 5. 参考情報 (Reference Information):** A section for providing additional information about the company's financial and operational status.

The form is marked with a 'Secret' (秘) and 'Confidential' (貴社控) label, indicating that the data collected is sensitive and should be handled accordingly.

出所) 日本銀行 全国企業短期経済観測調査 ウェブサイト

## 2) 研究概要

日本銀行ワーキングペーパーシリーズ

ビジネスサーベイにおける欠測値補完の検討—全国企業短期経済観測調査(短観)のケース—

[http://www.boj.or.jp/research/wps\\_rev/wps\\_2012/data/wp12j08.pdf](http://www.boj.or.jp/research/wps_rev/wps_2012/data/wp12j08.pdf)

## 5. 日本人の国民性調査

統合数理研究所 日本人の国民性調査 ウェブサイト

<http://www.ism.ac.jp/kokuminsei/index.html>

6. 労働力調査

1) 調査票例

労働力調査(2013年1月調査) 基礎調査票

出所) 総務省統計局 労働力調査 ウェブサイト

労働力調査(2013年1月調査) 基礎調査票

出所) 総務省統計局 労働力調査 ウェブサイト







附属資料Ⅱ 海外における補完推計などに関する取組

1. ACS (American Community Survey: アメリカ地域社会調査) 調査票例

ACS 2012 年調査票(世帯票)

13192018

U.S. DEPARTMENT OF COMMERCE  
Economic and Statistics Administration  
U.S. CENSUS BUREAU



# THE American Community Survey

This booklet shows the content of the American Community Survey questionnaire.

**Please complete this form and return it as soon as possible after receiving it in the mail.**

This form asks for information about the people who are living or staying at the address on the mailing label and about the house, apartment, or mobile home located at the address on the mailing label.

**If you need help or have questions about completing this form, please call 1-800-354-7271. The telephone call is free.**

**Telephone Device for the Deaf (TDD):**  
Call 1-800-582-8330. The telephone call is free.

**¿NECESITA AYUDA?** Si usted habla español y necesita ayuda para completar su cuestionario, llame sin cargo alguno al **1-877-833-5625**. Usted también puede pedir un cuestionario en español o completar su entrevista por teléfono con un entrevistador que habla español.

For more information about the American Community Survey, visit our web site at: <http://www.census.gov/acs/www/>

### Start Here

➔ Please print today's date.  
Month Day Year

➔ Please print the name and telephone number of the person who is filling out this form. We may contact you if there is a question.  
Last Name   
First Name  MI   
Area Code + Number   -

➔ How many people are living or staying at this address?  

- **INCLUDE** everyone who is living or staying here for more than 2 months.
- **INCLUDE** yourself if you are living here for more than 2 months.
- **INCLUDE** anyone else staying here who does not have another place to stay, even if they are here for 2 months or less.
- **DO NOT INCLUDE** anyone who is living somewhere else for more than 2 months, such as a college student living away or someone in the Armed Forces on deployment.

 Number of people

➔ Fill out pages 2, 3, and 4 for everyone, including yourself, who is living or staying at this address for more than 2 months. Then complete the rest of the form.

FORM **ACS-1(INFO)(2012)KFI** OMB No. 0607-0810



出所)アメリカセンサス局 (United States Census Bureau) ACS ウェブサイト

| Person 1  | Person 2   |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
|---|--|--|--|---|---|--|--|--|---|--|--|---|---|---|--|--|---|--------------------------------|---------------------------------------|-----------------------------------|--|---|----------------------------------|---------------------------------|--|---|-----------------------------------|-------------------------------------|---------------------------------|---|---|--|--|
| <p>(Person 1 is the person living or staying here in whose name this house or apartment is owned, being bought, or rented. If there is no such person, start with the name of any adult living or staying here.)</p>  |  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <p><b>1 What is Person 1's name?</b><br/>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p>   | <p><b>1 What is Person 2's name?</b><br/>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p>  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <p><b>2 How is this person related to Person 1?</b><br/><input checked="" type="checkbox"/> Person 1</p>  | <p><b>2 How is this person related to Person 1? Mark (X) ONE box.</b></p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> Husband or wife</td> <td><input type="checkbox"/> Son-in-law or daughter-in-law</td> </tr> <tr> <td><input type="checkbox"/> Biological son or daughter</td> <td><input type="checkbox"/> Other relative</td> </tr> <tr> <td><input type="checkbox"/> Adopted son or daughter</td> <td><input type="checkbox"/> Roomer or boarder</td> </tr> <tr> <td><input type="checkbox"/> Stepson or stepdaughter</td> <td><input type="checkbox"/> Housemate or roommate</td> </tr> <tr> <td><input type="checkbox"/> Brother or sister</td> <td><input type="checkbox"/> Unmarried partner</td> </tr> <tr> <td><input type="checkbox"/> Father or mother</td> <td><input type="checkbox"/> Foster child</td> </tr> <tr> <td><input type="checkbox"/> Grandchild</td> <td><input type="checkbox"/> Other nonrelative</td> </tr> <tr> <td><input type="checkbox"/> Parent-in-law</td> <td></td> </tr> </table> | <input type="checkbox"/> Husband or wife | <input type="checkbox"/> Son-in-law or daughter-in-law | <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Other relative               | <input type="checkbox"/> Adopted son or daughter | <input type="checkbox"/> Roomer or boarder | <input type="checkbox"/> Stepson or stepdaughter | <input type="checkbox"/> Housemate or roommate  | <input type="checkbox"/> Brother or sister | <input type="checkbox"/> Unmarried partner | <input type="checkbox"/> Father or mother | <input type="checkbox"/> Foster child   | <input type="checkbox"/> Grandchild   | <input type="checkbox"/> Other nonrelative | <input type="checkbox"/> Parent-in-law |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Husband or wife  | <input type="checkbox"/> Son-in-law or daughter-in-law   |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Biological son or daughter   | <input type="checkbox"/> Other relative  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Adopted son or daughter  | <input type="checkbox"/> Roomer or boarder   |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Stepson or stepdaughter  | <input type="checkbox"/> Housemate or roommate   |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Brother or sister  | <input type="checkbox"/> Unmarried partner   |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Father or mother   | <input type="checkbox"/> Foster child  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Grandchild   | <input type="checkbox"/> Other nonrelative   |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Parent-in-law  |  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <p><b>3 What is Person 1's sex? Mark (X) ONE box.</b><br/><input type="checkbox"/> Male <input type="checkbox"/> Female</p>   | <p><b>3 What is Person 2's sex? Mark (X) ONE box.</b><br/><input type="checkbox"/> Male <input type="checkbox"/> Female</p>  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <p><b>4 What is Person 1's age and what is Person 1's date of birth?</b><br/>Please report babies as age 0 when the child is less than 1 year old.<br/>Print numbers in boxes.<br/>Age (in years) <input style="width: 30px;" type="text"/> Month <input style="width: 30px;" type="text"/> Day <input style="width: 30px;" type="text"/> Year of birth <input style="width: 30px;" type="text"/></p>   | <p><b>4 What is Person 2's age and what is Person 2's date of birth?</b><br/>Please report babies as age 0 when the child is less than 1 year old.<br/>Print numbers in boxes.<br/>Age (in years) <input style="width: 30px;" type="text"/> Month <input style="width: 30px;" type="text"/> Day <input style="width: 30px;" type="text"/> Year of birth <input style="width: 30px;" type="text"/></p>  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <p><b>5 Is Person 1 of Hispanic, Latino, or Spanish origin?</b><br/><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin<br/><input type="checkbox"/> Yes, Mexican, Mexican Am., Chicano<br/><input type="checkbox"/> Yes, Puerto Rican<br/><input type="checkbox"/> Yes, Cuban<br/><input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin - Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. <input style="width: 150px;" type="text"/></p>  | <p><b>5 Is Person 2 of Hispanic, Latino, or Spanish origin?</b><br/><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin<br/><input type="checkbox"/> Yes, Mexican, Mexican Am., Chicano<br/><input type="checkbox"/> Yes, Puerto Rican<br/><input type="checkbox"/> Yes, Cuban<br/><input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin - Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. <input style="width: 150px;" type="text"/></p>   |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <p><b>6 What is Person 1's race? Mark (X) one or more boxes.</b></p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> White</td> <td><input type="checkbox"/> Asian Indian</td> <td><input type="checkbox"/> Japanese</td> <td><input type="checkbox"/> Native Hawaiian</td> </tr> <tr> <td><input type="checkbox"/> Black, African Am., or Negro</td> <td><input type="checkbox"/> Chinese</td> <td><input type="checkbox"/> Korean</td> <td><input type="checkbox"/> Guamanian or Chamorro</td> </tr> <tr> <td><input type="checkbox"/> American Indian or Alaska Native - Print name of enrolled or principal tribe. <input style="width: 150px;" type="text"/></td> <td><input type="checkbox"/> Filipino</td> <td><input type="checkbox"/> Vietnamese</td> <td><input type="checkbox"/> Samoan</td> </tr> <tr> <td><input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. <input style="width: 150px;" type="text"/></td> <td><input type="checkbox"/> Other Pacific Islander - Print race, for example, Rapa Tongan, and so on. <input style="width: 150px;" type="text"/></td> <td></td> <td></td> </tr> </table> <p><input type="checkbox"/> Some other race - Print race. <input style="width: 150px;" type="text"/></p> | <input type="checkbox"/> White   | <input type="checkbox"/> Asian Indian    | <input type="checkbox"/> Japanese                      | <input type="checkbox"/> Native Hawaiian            | <input type="checkbox"/> Black, African Am., or Negro | <input type="checkbox"/> Chinese                 | <input type="checkbox"/> Korean            | <input type="checkbox"/> Guamanian or Chamorro   | <input type="checkbox"/> American Indian or Alaska Native - Print name of enrolled or principal tribe. <input style="width: 150px;" type="text"/> | <input type="checkbox"/> Filipino          | <input type="checkbox"/> Vietnamese        | <input type="checkbox"/> Samoan           | <input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. <input style="width: 150px;" type="text"/> | <input type="checkbox"/> Other Pacific Islander - Print race, for example, Rapa Tongan, and so on. <input style="width: 150px;" type="text"/> |  |  | <p><b>6 What is Person 2's race? Mark (X) one or more boxes.</b></p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> White</td> <td><input type="checkbox"/> Asian Indian</td> <td><input type="checkbox"/> Japanese</td> <td><input type="checkbox"/> Native Hawaiian</td> </tr> <tr> <td><input type="checkbox"/> Black, African Am., or Negro</td> <td><input type="checkbox"/> Chinese</td> <td><input type="checkbox"/> Korean</td> <td><input type="checkbox"/> Guamanian or Chamorro</td> </tr> <tr> <td><input type="checkbox"/> American Indian or Alaska Native - Print name of enrolled or principal tribe. <input style="width: 150px;" type="text"/></td> <td><input type="checkbox"/> Filipino</td> <td><input type="checkbox"/> Vietnamese</td> <td><input type="checkbox"/> Samoan</td> </tr> <tr> <td><input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. <input style="width: 150px;" type="text"/></td> <td><input type="checkbox"/> Other Pacific Islander - Print race, for example, Rapa Tongan, and so on. <input style="width: 150px;" type="text"/></td> <td></td> <td></td> </tr> </table> <p><input type="checkbox"/> Some other race - Print race. <input style="width: 150px;" type="text"/></p> | <input type="checkbox"/> White | <input type="checkbox"/> Asian Indian | <input type="checkbox"/> Japanese | <input type="checkbox"/> Native Hawaiian | <input type="checkbox"/> Black, African Am., or Negro | <input type="checkbox"/> Chinese | <input type="checkbox"/> Korean | <input type="checkbox"/> Guamanian or Chamorro | <input type="checkbox"/> American Indian or Alaska Native - Print name of enrolled or principal tribe. <input style="width: 150px;" type="text"/> | <input type="checkbox"/> Filipino | <input type="checkbox"/> Vietnamese | <input type="checkbox"/> Samoan | <input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. <input style="width: 150px;" type="text"/> | <input type="checkbox"/> Other Pacific Islander - Print race, for example, Rapa Tongan, and so on. <input style="width: 150px;" type="text"/> |  |  |
| <input type="checkbox"/> White  | <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese        | <input type="checkbox"/> Native Hawaiian               |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Black, African Am., or Negro   | <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean          | <input type="checkbox"/> Guamanian or Chamorro         |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> American Indian or Alaska Native - Print name of enrolled or principal tribe. <input style="width: 150px;" type="text"/>   | <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese      | <input type="checkbox"/> Samoan                        |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. <input style="width: 150px;" type="text"/>   | <input type="checkbox"/> Other Pacific Islander - Print race, for example, Rapa Tongan, and so on. <input style="width: 150px;" type="text"/>  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> White  | <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese        | <input type="checkbox"/> Native Hawaiian               |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Black, African Am., or Negro   | <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean          | <input type="checkbox"/> Guamanian or Chamorro         |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> American Indian or Alaska Native - Print name of enrolled or principal tribe. <input style="width: 150px;" type="text"/>   | <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese      | <input type="checkbox"/> Samoan                        |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |
| <input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. <input style="width: 150px;" type="text"/>   | <input type="checkbox"/> Other Pacific Islander - Print race, for example, Rapa Tongan, and so on. <input style="width: 150px;" type="text"/>  |  |  |   |   |  |  |  |   |  |  |   |   |   |  |  |   |                                |                                       |                                   |  |   |                                  |                                 |  |   |                                   |                                     |                                 |   |   |  |  |



| <b>Person 3</b>  | <b>Person 4</b>   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
|--|---|--|---|---|--|--|--|--|---|--|---|---------------------------------------|---|--|---|--|--|--|--|---|---|--|--|--|--|--|--|---|---|-------------------------------------|--|--|--|
| <p><b>1 What is Person 3's name?</b><br/>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 30px;" type="text"/></p>  | <p><b>1 What is Person 4's name?</b><br/>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 30px;" type="text"/></p>   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <p><b>2 How is this person related to Person 1? Mark (X) ONE box.</b></p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> Husband or wife</td> <td><input type="checkbox"/> Son-in-law or daughter-in-law</td> </tr> <tr> <td><input type="checkbox"/> Biological son or daughter</td> <td><input type="checkbox"/> Other relative</td> </tr> <tr> <td><input type="checkbox"/> Adopted son or daughter</td> <td><input type="checkbox"/> Roomer or boarder</td> </tr> <tr> <td><input type="checkbox"/> Stepson or stepdaughter</td> <td><input type="checkbox"/> Housemate or roommate</td> </tr> <tr> <td><input type="checkbox"/> Brother or sister</td> <td><input type="checkbox"/> Unmarried partner</td> </tr> <tr> <td><input type="checkbox"/> Father or mother</td> <td><input type="checkbox"/> Foster child</td> </tr> <tr> <td><input type="checkbox"/> Grandchild</td> <td><input type="checkbox"/> Other nonrelative</td> </tr> <tr> <td><input type="checkbox"/> Parent-in-law</td> <td></td> </tr> </table>   | <input type="checkbox"/> Husband or wife  | <input type="checkbox"/> Son-in-law or daughter-in-law | <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Other relative | <input type="checkbox"/> Adopted son or daughter | <input type="checkbox"/> Roomer or boarder     | <input type="checkbox"/> Stepson or stepdaughter | <input type="checkbox"/> Housemate or roommate | <input type="checkbox"/> Brother or sister  | <input type="checkbox"/> Unmarried partner   | <input type="checkbox"/> Father or mother | <input type="checkbox"/> Foster child | <input type="checkbox"/> Grandchild   | <input type="checkbox"/> Other nonrelative | <input type="checkbox"/> Parent-in-law    |  | <p><b>2 How is this person related to Person 1? Mark (X) ONE box.</b></p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> Husband or wife</td> <td><input type="checkbox"/> Son-in-law or daughter-in-law</td> </tr> <tr> <td><input type="checkbox"/> Biological son or daughter</td> <td><input type="checkbox"/> Other relative</td> </tr> <tr> <td><input type="checkbox"/> Adopted son or daughter</td> <td><input type="checkbox"/> Roomer or boarder</td> </tr> <tr> <td><input type="checkbox"/> Stepson or stepdaughter</td> <td><input type="checkbox"/> Housemate or roommate</td> </tr> <tr> <td><input type="checkbox"/> Brother or sister</td> <td><input type="checkbox"/> Unmarried partner</td> </tr> <tr> <td><input type="checkbox"/> Father or mother</td> <td><input type="checkbox"/> Foster child</td> </tr> <tr> <td><input type="checkbox"/> Grandchild</td> <td><input type="checkbox"/> Other nonrelative</td> </tr> <tr> <td><input type="checkbox"/> Parent-in-law</td> <td></td> </tr> </table> | <input type="checkbox"/> Husband or wife | <input type="checkbox"/> Son-in-law or daughter-in-law | <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Other relative | <input type="checkbox"/> Adopted son or daughter | <input type="checkbox"/> Roomer or boarder | <input type="checkbox"/> Stepson or stepdaughter | <input type="checkbox"/> Housemate or roommate | <input type="checkbox"/> Brother or sister   | <input type="checkbox"/> Unmarried partner | <input type="checkbox"/> Father or mother | <input type="checkbox"/> Foster child   | <input type="checkbox"/> Grandchild | <input type="checkbox"/> Other nonrelative | <input type="checkbox"/> Parent-in-law |  |
| <input type="checkbox"/> Husband or wife   | <input type="checkbox"/> Son-in-law or daughter-in-law  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Biological son or daughter  | <input type="checkbox"/> Other relative   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Adopted son or daughter   | <input type="checkbox"/> Roomer or boarder  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Stepson or stepdaughter   | <input type="checkbox"/> Housemate or roommate  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Brother or sister   | <input type="checkbox"/> Unmarried partner  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Father or mother  | <input type="checkbox"/> Foster child   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Grandchild  | <input type="checkbox"/> Other nonrelative  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Parent-in-law   |   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Husband or wife   | <input type="checkbox"/> Son-in-law or daughter-in-law  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Biological son or daughter  | <input type="checkbox"/> Other relative   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Adopted son or daughter   | <input type="checkbox"/> Roomer or boarder  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Stepson or stepdaughter   | <input type="checkbox"/> Housemate or roommate  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Brother or sister   | <input type="checkbox"/> Unmarried partner  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Father or mother  | <input type="checkbox"/> Foster child   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Grandchild  | <input type="checkbox"/> Other nonrelative  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Parent-in-law   |   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <p><b>3 What is Person 3's sex? Mark (X) ONE box.</b><br/><input type="checkbox"/> Male <input type="checkbox"/> Female</p>  | <p><b>3 What is Person 4's sex? Mark (X) ONE box.</b><br/><input type="checkbox"/> Male <input checked="" type="checkbox"/> Female</p>  |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <p><b>4 What is Person 3's age and what is Person 3's date of birth?</b><br/>Please report babies as age 0 when the child is less than 1 year old.<br/>Print numbers in boxes.</p> <table style="width: 100%; border: none;"> <tr> <td style="text-align: center;">Age (in years)</td> <td style="text-align: center;">Month</td> <td style="text-align: center;">Day</td> <td style="text-align: center;">Year of birth</td> </tr> <tr> <td style="text-align: center;"><input style="width: 30px;" type="text"/></td> <td style="text-align: center;"><input style="width: 30px;" type="text"/></td> <td style="text-align: center;"><input style="width: 30px;" type="text"/></td> <td style="text-align: center;"><input style="width: 60px;" type="text"/></td> </tr> </table>  | Age (in years)  | Month  | Day   | Year of birth                           | <input style="width: 30px;" type="text"/>        | <input style="width: 30px;" type="text"/>      | <input style="width: 30px;" type="text"/>        | <input style="width: 60px;" type="text"/>      | <p><b>4 What is Person 4's age and what is Person 4's date of birth?</b><br/>Please report babies as age 0 when the child is less than 1 year old.<br/>Print numbers in boxes.</p> <table style="width: 100%; border: none;"> <tr> <td style="text-align: center;">Age (in years)</td> <td style="text-align: center;">Month</td> <td style="text-align: center;">Day</td> <td style="text-align: center;">Year of birth</td> </tr> <tr> <td style="text-align: center;"><input style="width: 30px;" type="text"/></td> <td style="text-align: center;"><input style="width: 30px;" type="text"/></td> <td style="text-align: center;"><input style="width: 30px;" type="text"/></td> <td style="text-align: center;"><input style="width: 60px;" type="text"/></td> </tr> </table> | Age (in years)   | Month                                     | Day                                   | Year of birth   | <input style="width: 30px;" type="text"/>  | <input style="width: 30px;" type="text"/> | <input style="width: 30px;" type="text"/>  | <input style="width: 60px;" type="text"/>  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| Age (in years)   | Month   | Day  | Year of birth                                       |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input style="width: 30px;" type="text"/>  | <input style="width: 30px;" type="text"/>   | <input style="width: 30px;" type="text"/>              | <input style="width: 60px;" type="text"/>           |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| Age (in years)   | Month   | Day  | Year of birth                                       |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input style="width: 30px;" type="text"/>  | <input style="width: 30px;" type="text"/>   | <input style="width: 30px;" type="text"/>              | <input style="width: 60px;" type="text"/>           |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <p>→ <b>NOTE:</b> Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this survey, Hispanic origins are not races.</p>  |   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <p><b>5 Is Person 3 of Hispanic, Latino, or Spanish origin?</b></p> <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin</p> <p><input type="checkbox"/> Yes, Mexican, Mexican Am., Chicano</p> <p><input type="checkbox"/> Yes, Puerto Rican</p> <p><input type="checkbox"/> Yes, Cuban</p> <p><input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin – Print origin (for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on) →</p> <p><input style="width: 100%;" type="text"/></p>  | <p><b>5 Is Person 4 of Hispanic, Latino, or Spanish origin?</b></p> <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin</p> <p><input type="checkbox"/> Yes, Mexican, Mexican Am., Chicano</p> <p><input type="checkbox"/> Yes, Puerto Rican</p> <p><input type="checkbox"/> Yes, Cuban</p> <p><input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin – Print origin (for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on) →</p> <p><input style="width: 100%;" type="text"/></p> |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <p><b>6 What is Person 3's race? Mark (X) one or more boxes.</b></p> <p><input type="checkbox"/> White</p> <p><input type="checkbox"/> Black, African Am., or Negro</p> <p><input type="checkbox"/> American Indian or Alaska Native – Print name of enrolled or principal tribe →</p> <p><input style="width: 100%;" type="text"/></p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> Asian Indian</td> <td><input type="checkbox"/> Japanese</td> <td><input type="checkbox"/> Native Hawaiian</td> </tr> <tr> <td><input type="checkbox"/> Chinese</td> <td><input type="checkbox"/> Korean</td> <td><input type="checkbox"/> Guamanian or Chamorro</td> </tr> <tr> <td><input type="checkbox"/> Filipino</td> <td><input type="checkbox"/> Vietnamese</td> <td><input type="checkbox"/> Samoan</td> </tr> <tr> <td colspan="3"><input type="checkbox"/> Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. →</td> </tr> <tr> <td colspan="3"><input type="checkbox"/> Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on. →</td> </tr> </table> <p><input type="checkbox"/> Some other race – Print race. →</p> <p><input style="width: 100%;" type="text"/></p> | <input type="checkbox"/> Asian Indian   | <input type="checkbox"/> Japanese                      | <input type="checkbox"/> Native Hawaiian            | <input type="checkbox"/> Chinese        | <input type="checkbox"/> Korean                  | <input type="checkbox"/> Guamanian or Chamorro | <input type="checkbox"/> Filipino                | <input type="checkbox"/> Vietnamese            | <input type="checkbox"/> Samoan   | <input type="checkbox"/> Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. → |   |                                       | <input type="checkbox"/> Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on. → |  |   | <p><b>6 What is Person 4's race? Mark (X) one or more boxes.</b></p> <p><input type="checkbox"/> White</p> <p><input type="checkbox"/> Black, African Am., or Negro</p> <p><input type="checkbox"/> American Indian or Alaska Native – Print name of enrolled or principal tribe →</p> <p><input style="width: 100%;" type="text"/></p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> Asian Indian</td> <td><input type="checkbox"/> Japanese</td> <td><input type="checkbox"/> Native Hawaiian</td> </tr> <tr> <td><input type="checkbox"/> Chinese</td> <td><input type="checkbox"/> Korean</td> <td><input type="checkbox"/> Guamanian or Chamorro</td> </tr> <tr> <td><input type="checkbox"/> Filipino</td> <td><input type="checkbox"/> Vietnamese</td> <td><input type="checkbox"/> Samoan</td> </tr> <tr> <td colspan="3"><input type="checkbox"/> Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. →</td> </tr> <tr> <td colspan="3"><input type="checkbox"/> Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on. →</td> </tr> </table> <p><input type="checkbox"/> Some other race – Print race. →</p> <p><input style="width: 100%;" type="text"/></p> | <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese        | <input type="checkbox"/> Native Hawaiian               | <input type="checkbox"/> Chinese                    | <input type="checkbox"/> Korean         | <input type="checkbox"/> Guamanian or Chamorro   | <input type="checkbox"/> Filipino          | <input type="checkbox"/> Vietnamese              | <input type="checkbox"/> Samoan                | <input type="checkbox"/> Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. → |  |   | <input type="checkbox"/> Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on. → |                                     |  |  |  |
| <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese   | <input type="checkbox"/> Native Hawaiian               |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean   | <input type="checkbox"/> Guamanian or Chamorro         |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese   | <input type="checkbox"/> Samoan                        |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. →   |   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on. →  |   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese   | <input type="checkbox"/> Native Hawaiian               |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean   | <input type="checkbox"/> Guamanian or Chamorro         |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese   | <input type="checkbox"/> Samoan                        |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Other Asian – Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. →   |   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |
| <input type="checkbox"/> Other Pacific Islander – Print race, for example, Fijian, Tongan, and so on. →  |   |  |   |   |  |  |  |  |   |  |   |                                       |   |  |   |  |  |  |  |   |   |  |  |  |  |  |  |   |   |                                     |  |  |  |



| <b>Person 5</b>   |  | <b>Person 6</b>  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
|---|--|--|--|---|---|--|--|--|--|--|--|---|--|--|--|--|--|--|--|
| <p><b>1</b> What is Person 5's name? Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p>   |  | <p><b>2</b> If there are more than five people living or staying here, print their names in the spaces for Person 6 through Person 12. We may call you for more information about them. ↗</p>  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <p><b>2</b> How is this person related to Person 1? Mark (X) ONE box.</p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> Husband or wife</td> <td><input type="checkbox"/> Son-in-law or daughter-in-law</td> </tr> <tr> <td><input type="checkbox"/> Biological son or daughter</td> <td><input type="checkbox"/> Other relative</td> </tr> <tr> <td><input type="checkbox"/> Adopted son or daughter</td> <td><input type="checkbox"/> Roomer or boarder</td> </tr> <tr> <td><input type="checkbox"/> Stepson or stepdaughter</td> <td><input type="checkbox"/> Housemate or roommate</td> </tr> <tr> <td><input type="checkbox"/> Brother or sister</td> <td><input type="checkbox"/> Unmarried partner</td> </tr> <tr> <td><input type="checkbox"/> Father or mother</td> <td><input type="checkbox"/> Foster child</td> </tr> <tr> <td><input type="checkbox"/> Grandchild</td> <td><input type="checkbox"/> Other nonrelative</td> </tr> <tr> <td><input type="checkbox"/> Parent-in-law</td> <td></td> </tr> </table>  |  | <input type="checkbox"/> Husband or wife   | <input type="checkbox"/> Son-in-law or daughter-in-law | <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Other relative | <input type="checkbox"/> Adopted son or daughter | <input type="checkbox"/> Roomer or boarder     | <input type="checkbox"/> Stepson or stepdaughter | <input type="checkbox"/> Housemate or roommate | <input type="checkbox"/> Brother or sister | <input type="checkbox"/> Unmarried partner   | <input type="checkbox"/> Father or mother | <input type="checkbox"/> Foster child  | <input type="checkbox"/> Grandchild  | <input type="checkbox"/> Other nonrelative | <input type="checkbox"/> Parent-in-law |  | <p>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p> <p>Sex <input type="checkbox"/> Male <input type="checkbox"/> Female      Age (In years) <input style="width: 40px;" type="text"/></p> |  |
| <input type="checkbox"/> Husband or wife  | <input type="checkbox"/> Son-in-law or daughter-in-law |  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Biological son or daughter   | <input type="checkbox"/> Other relative                |  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Adopted son or daughter  | <input type="checkbox"/> Roomer or boarder             |  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Stepson or stepdaughter  | <input type="checkbox"/> Housemate or roommate         |  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Brother or sister  | <input type="checkbox"/> Unmarried partner             |  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Father or mother   | <input type="checkbox"/> Foster child                  |  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Grandchild   | <input type="checkbox"/> Other nonrelative             |  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Parent-in-law  |  |  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <p><b>3</b> What is Person 5's sex? Mark (X) ONE box.</p> <p><input type="checkbox"/> Male      <input type="checkbox"/> Female</p>   |  | <p><b>Person 7</b></p> <p>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p> <p>Sex <input type="checkbox"/> Male <input type="checkbox"/> Female      Age (In years) <input style="width: 40px;" type="text"/></p>  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <p><b>3</b> What is Person 5's age and what is Person 5's date of birth? Please report babies as age 0 when the child is less than 1 year old. Print numbers in boxes.</p> <p>Age (In years) <input style="width: 40px;" type="text"/>      Month <input style="width: 20px;" type="text"/> Day <input style="width: 20px;" type="text"/> Year of birth <input style="width: 40px;" type="text"/></p>   |  | <p><b>Person 8</b></p> <p>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p> <p>Sex <input type="checkbox"/> Male <input type="checkbox"/> Female      Age (In years) <input style="width: 40px;" type="text"/></p>  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <p><b>4</b> NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this survey, Hispanic origins are not races.</p>   |  | <p><b>Person 9</b></p> <p>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p> <p>Sex <input type="checkbox"/> Male <input type="checkbox"/> Female      Age (In years) <input style="width: 40px;" type="text"/></p>  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <p><b>5</b> Is Person 5 of Hispanic, Latino, or Spanish origin?</p> <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin</p> <p><input type="checkbox"/> Yes, Mexican, Mexican Am., Chicano</p> <p><input type="checkbox"/> Yes, Puerto Rican</p> <p><input type="checkbox"/> Yes, Cuban</p> <p><input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin - Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. ↗</p> <p><input style="width: 150px;" type="text"/></p>  |  | <p><b>Person 10</b></p> <p>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p> <p>Sex <input type="checkbox"/> Male <input type="checkbox"/> Female      Age (In years) <input style="width: 40px;" type="text"/></p> |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <p><b>6</b> What is Person 5's race? Mark (X) one or more boxes.</p> <p><input type="checkbox"/> White</p> <p><input type="checkbox"/> Black, African Am., or Negro</p> <p><input type="checkbox"/> American Indian or Alaska Native - Print name of enrolled or principal tribe. ↗</p> <p><input style="width: 150px;" type="text"/></p> <table style="width: 100%; border: none;"> <tr> <td><input type="checkbox"/> Asian Indian</td> <td><input type="checkbox"/> Japanese</td> <td><input type="checkbox"/> Native Hawaiian</td> </tr> <tr> <td><input type="checkbox"/> Chinese</td> <td><input type="checkbox"/> Korean</td> <td><input type="checkbox"/> Guamanian or Chamorro</td> </tr> <tr> <td><input type="checkbox"/> Filipino</td> <td><input type="checkbox"/> Vietnamese</td> <td><input type="checkbox"/> Samoan</td> </tr> <tr> <td colspan="2"><input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ↗</td> <td><input type="checkbox"/> Other Pacific Islander - Print race, for example, Bikin, Tongan, and so on. ↗</td> </tr> </table> <p><input type="checkbox"/> Some other race - Print race. ↗</p> <p><input style="width: 150px;" type="text"/></p> |  | <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese                      | <input type="checkbox"/> Native Hawaiian            | <input type="checkbox"/> Chinese        | <input type="checkbox"/> Korean                  | <input type="checkbox"/> Guamanian or Chamorro | <input type="checkbox"/> Filipino                | <input type="checkbox"/> Vietnamese            | <input type="checkbox"/> Samoan            | <input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ↗ |   | <input type="checkbox"/> Other Pacific Islander - Print race, for example, Bikin, Tongan, and so on. ↗ | <p><b>Person 11</b></p> <p>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p> <p>Sex <input type="checkbox"/> Male <input type="checkbox"/> Female      Age (In years) <input style="width: 40px;" type="text"/></p> |  |  |  |  |  |
| <input type="checkbox"/> Asian Indian   | <input type="checkbox"/> Japanese                      | <input type="checkbox"/> Native Hawaiian   |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Chinese  | <input type="checkbox"/> Korean                        | <input type="checkbox"/> Guamanian or Chamorro   |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Filipino   | <input type="checkbox"/> Vietnamese                    | <input type="checkbox"/> Samoan  |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <input type="checkbox"/> Other Asian - Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ↗  |  | <input type="checkbox"/> Other Pacific Islander - Print race, for example, Bikin, Tongan, and so on. ↗   |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |
| <p><b>4</b> </p>  |  | <p><b>Person 12</b></p> <p>Last Name (Please print) <input style="width: 150px;" type="text"/> First Name <input style="width: 100px;" type="text"/> MI <input style="width: 20px;" type="text"/></p> <p>Sex <input type="checkbox"/> Male <input type="checkbox"/> Female      Age (In years) <input style="width: 40px;" type="text"/></p> |  |   |   |  |  |  |  |  |  |   |  |  |  |  |  |  |  |

# Housing

**1** Please answer the following questions about the house, apartment, or mobile home at the address on the mailing label.

**1** Which best describes this building? Include all apartments, flats, etc., even if vacant.

- A mobile home  
 A one-family house detached from any other house  
 A one-family house attached to one or more houses  
 A building with 2 apartments  
 A building with 3 or 4 apartments  
 A building with 5 to 9 apartments  
 A building with 10 to 19 apartments  
 A building with 20 to 49 apartments  
 A building with 50 or more apartments  
 Boat, RV, van, etc.

**2** About when was this building first built?

- 2000 or later - Specify year:
- 1990 to 1999  
 1980 to 1989  
 1970 to 1979  
 1960 to 1969  
 1950 to 1959  
 1940 to 1949  
 1939 or earlier

**3** When did PERSON 1 (listed on page 2) move into this house, apartment, or mobile home?

Month Year

**A** Answer questions 4 - 6 if this is a HOUSE OR A MOBILE HOME; otherwise, SKIP to question 7a.

**4** How many acres is this house or mobile home on?

- Less than 1 acre → SKIP to question 6  
 1 to 9.9 acres  
 10 or more acres

**5** IN THE PAST 12 MONTHS, what were the actual sales of all agricultural products from this property?

- None  
 \$1 to \$999  
 \$1,000 to \$2,499  
 \$2,500 to \$4,999  
 \$5,000 to \$9,999  
 \$10,000 or more

**6** Is there a business (such as a store or barber shop) or a medical office on this property?

- Yes  
 No

**7** a. How many separate rooms are in this house, apartment, or mobile home? Rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling.

- INCLUDE bedrooms, kitchens, etc.
- EXCLUDE bathrooms, porches, balconies, foyers, halls, or unfinished basements.

Number of rooms

b. How many of these rooms are bedrooms? Count as bedrooms those rooms you would list if this house, apartment, or mobile home were for sale or rent. If this is an efficiency/studio apartment, print '0'.

Number of bedrooms

**8** Does this house, apartment, or mobile home have -

- |   | Yes                      | No                       |
|---|--------------------------|--------------------------|
| a. hot and cold running water?  | <input type="checkbox"/> | <input type="checkbox"/> |
| b. a flush toilet?  | <input type="checkbox"/> | <input type="checkbox"/> |
| c. a bathtub or shower?   | <input type="checkbox"/> | <input type="checkbox"/> |
| d. a sink with a faucet?  | <input type="checkbox"/> | <input type="checkbox"/> |
| e. a stove or range?  | <input type="checkbox"/> | <input type="checkbox"/> |
| f. a refrigerator?  | <input type="checkbox"/> | <input type="checkbox"/> |
| g. telephone service from which you can both make and receive calls? include cell phones. | <input type="checkbox"/> | <input type="checkbox"/> |

**9** How many automobiles, vans, and trucks of one-ton capacity or less are kept at home for use by members of this household?

- None  
 1  
 2  
 3  
 4  
 5  
 6 or more

**10** Which FUEL is used MOST for heating this house, apartment, or mobile home?

- Gas: from underground pipes serving the neighborhood  
 Gas: bottled, tank, or LP  
 Electricity  
 Fuel oil, kerosene, etc.  
 Coal or coke  
 Wood  
 Solar energy  
 Other fuel  
 No fuel used



## Housing (continued)

- 11 a. LAST MONTH, what was the cost of electricity for this house, apartment, or mobile home?**

Last month's cost - Dollars

OR

- Included in rent or condominium fee  
 No charge or electricity not used

- b. LAST MONTH, what was the cost of gas for this house, apartment, or mobile home?**

Last month's cost - Dollars

OR

- Included in rent or condominium fee  
 Included in electricity payment entered above  
 No charge or gas not used

- c. IN THE PAST 12 MONTHS, what was the cost of water and sewer for this house, apartment, or mobile home? If you have lived here less than 12 months, estimate the cost.**

Past 12 months' cost - Dollars

OR

- Included in rent or condominium fee  
 No charge

- d. IN THE PAST 12 MONTHS, what was the cost of oil, coal, kerosene, wood, etc., for this house, apartment, or mobile home? If you have lived here less than 12 months, estimate the cost.**

Past 12 months' cost - Dollars

OR

- Included in rent or condominium fee  
 No charge or these fuels not used

- 12 IN THE PAST 12 MONTHS, did anyone in this household receive Food Stamps or a Food Stamp benefit card? (Include government benefits from the Supplemental Nutrition Assistance Program (SNAP). Do NOT include WIC or the National School Lunch Program.)**

- Yes  
 No

- 13 Is this house, apartment, or mobile home part of a condominium?**

- Yes → What is the monthly condominium fee? For renters, answer only if you pay the condominium fee in addition to your rent; otherwise, mark the "None" box.

Monthly amount - Dollars

OR

- None

- No

- 14 Is this house, apartment, or mobile home - Mark (X) ONE box.**

- Owned by you or someone in this household with a mortgage or loan? (Include home equity loans.)  
 Owned by you or someone in this household free and clear (without a mortgage or loan)?  
 Rented?  
 Occupied without payment of rent? → SKIP to **C**

- B Answer questions 15a and b if this house, apartment, or mobile home is RENTED. Otherwise, SKIP to question 16.**

- 15 a. What is the monthly rent for this house, apartment, or mobile home?**

Monthly amount - Dollars

- b. Does the monthly rent include any meals?**

- Yes  
 No

- C Answer questions 16 - 20 if you or someone else in this household OWNS or IS BUYING this house, apartment, or mobile home. Otherwise, SKIP to **E** on the next page.**

- 16 About how much do you think this house and lot, apartment, or mobile home (and lot, if owned) would sell for if it were for sale?**

Amount - Dollars

- 17 What are the annual real estate taxes on THIS property?**

Annual amount - Dollars

OR

- None

- 18 What is the annual payment for fire, hazard, and flood insurance on THIS property?**

Annual amount - Dollars

OR

- None





# Person 1

Please copy the name of Person 1 from page 2, then continue answering questions below.

Last Name

First Name MI

Where was this person born?

In the United States - Print name of state

Outside the United States - Print name of foreign country, or Puerto Rico, Guam, etc.

Is this person a citizen of the United States?

Yes, born in the United States → SKIP to 15a

Yes, born in Puerto Rico, Guam, the U.S. Virgin Islands, or Northern Mariana

Yes, born abroad of U.S. citizen parent or parents

Yes, U.S. citizen by naturalization - Print year of naturalization

No, not a U.S. citizen

When did this person come to live in the United States? Print numbers in boxes.

Year

a. At any time IN THE LAST 3 MONTHS, has this person attended school or college? Include only nursery or preschool, kindergarten, elementary school, home school, and schooling which leads to a high school diploma or a college degree.

No, has not attended in the last 3 months → SKIP to question 11

Yes, public school, public college

Yes, private school, private college, home school

b. What grade or level was this person attending? Mark (X) ONE box.

Nursery school, preschool

Kindergarten

Grade 1 through 12 - Specify grade 1 - 12

College undergraduate years (freshman to senior)

Graduate or professional school beyond a bachelor's degree (for example: M.A. or Ph.D. program, or medical or law school)

11 What is the highest degree or level of school this person has COMPLETED? Mark (X) ONE box. If currently enrolled, mark the previous grade or highest degree received.

**NO SCHOOLING COMPLETED**

No schooling completed

**NURSERY OR PRESCHOOL THROUGH GRADE 12**

Nursery school

Kindergarten

Grade 1 through 11 - Specify grade 1 - 11

12th grade - NO DIPLOMA

**HIGH SCHOOL GRADUATE**

Regular high school diploma

GED or alternative credential

**COLLEGE OR SOME COLLEGE**

Some college credit, but less than 1 year of college credit

1 or more years of college credit, no degree

Associate's degree (for example: AA, AS)

Bachelor's degree (for example: BA, BS)

**AFTER BACHELOR'S DEGREE**

Master's degree (for example: MA, MS, MEng, MEd, MSW, MBA)

Professional degree beyond a bachelor's degree (for example: MD, DDS, DVM, LLB, JD)

Doctorate degree (for example: PhD, EdD)

Answer question 12 if this person has a bachelor's degree or higher. Otherwise, SKIP to question 13.

12 This question focuses on this person's BACHELOR'S DEGREE. Please print below the specific major(s) of any BACHELOR'S DEGREES this person has received. (For example: chemical engineering, elementary teacher education, organizational psychology)

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

13 What is this person's ancestry or ethnic origin?

\_\_\_\_\_

(For example: Italian, Jamaican, African Am., Cambodian, Cape Verdean, Norwegian, Dominican, French Canadian, Haitian, Korean, Tagalog, Polish, Nigerian, Mexican, Taiwanese, Ukrainian, and so on.)

14 a. Does this person speak a language other than English at home?

Yes

No → SKIP to question 15a

b. What is this language?

\_\_\_\_\_

(For example: Korean, Italian, Spanish, Vietnamese)

c. How well does this person speak English?

Very well

Well

Not well

Not at all

15 a. Did this person live in this house or apartment 1 year ago?

Person is under 1 year old → SKIP to question 16

Yes, this house → SKIP to question 16

No, outside the United States and Puerto Rico - Print name of foreign country, or U.S. Virgin Islands, Guam, etc., below; then SKIP to question 16

\_\_\_\_\_

No, different house in the United States or Puerto Rico

b. Where did this person live 1 year ago?

Address (Number and street name)

\_\_\_\_\_

Name of city, town, or post office

\_\_\_\_\_

Name of U.S. county or municipio in Puerto Rico

\_\_\_\_\_

Name of U.S. state or Puerto Rico

ZIP Code

\_\_\_\_\_





**Person 1 (continued)**

**30** a. LAST WEEK, did this person work for pay at a job (or business)?

- Yes → SKIP to question 30
- No - Did not work (or retired)

b. LAST WEEK, did this person do ANY work for pay, even for as little as one hour?

- Yes
- No → SKIP to question 35a

**31** At what location did this person work LAST WEEK? If this person worked at more than one location, print where he or she worked most last week.

a. Address (Number and street name)

*If the exact address is not known, give a description of the location such as the building name or the nearest street or intersection.*

b. Name of city, town, or post office

c. Is the work location inside the limits of that city or town?

- Yes
- No, outside the city/town limits

d. Name of county

e. Name of U.S. state or foreign country

f. ZIP Code

**32** How did this person usually get to work LAST WEEK? If this person usually used more than one method of transportation during the trip, mark (X) the bar of the one used for most of the distance.

- |   |  |
|---|--|
| <input type="checkbox"/> Car, truck, or van       | <input type="checkbox"/> Motorcycle                            |
| <input type="checkbox"/> Bus or trolley bus       | <input type="checkbox"/> Bicycle                               |
| <input type="checkbox"/> Streetcar or trolley car | <input type="checkbox"/> Walked                                |
| <input type="checkbox"/> Subway or elevated       | <input type="checkbox"/> Worked at home → SKIP to question 39a |
| <input type="checkbox"/> Railroad                 | <input type="checkbox"/> Other method                          |
| <input type="checkbox"/> Ferryboat                |  |
| <input type="checkbox"/> Taxicab                  |  |

**J** Answer question 32 if you marked "Car, truck, or van" in question 31. Otherwise, SKIP to question 33.

**33** How many people, including this person, usually rode to work in the car, truck, or van LAST WEEK?

Person(s)

**34** What time did this person usually leave home to go to work LAST WEEK?

Hour  : Minute   a.m.  p.m.

**34** How many minutes did it usually take this person to get from home to work LAST WEEK?

Minutes

**K** Answer questions 35-38 if this person did NOT work last week. Otherwise, SKIP to question 39a.

**35** a. LAST WEEK, was this person on layoff from a job?

- Yes → SKIP to question 35c
- No

b. LAST WEEK, was this person TEMPORARILY absent from a job or business?

- Yes, on vacation, temporary illness, maternity leave, other family/personal reasons, bad weather, etc. → SKIP to question 38
- No → SKIP to question 36

c. Has this person been informed that he or she will be recalled to work within the next 6 months OR been given a date to return to work?

- Yes → SKIP to question 37
- No

**36** During the LAST 4 WEEKS, has this person been ACTIVELY looking for work?

- Yes
- No → SKIP to question 38

**37** LAST WEEK, could this person have started a job if offered one, or returned to work if recalled?

- Yes, could have gone to work
- No, because of own temporary illness
- No, because of all other reasons (in school, etc.)

**38** When did this person last work, even for a few days?

- Within the past 12 months
- 1 to 5 years ago → SKIP to L
- Over 5 years ago or never worked → SKIP to question 47

**39** a. During the PAST 12 MONTHS (52 weeks), did this person work 50 or more weeks? Count paid time off as work.

- Yes → SKIP to question 40
- No

b. How many weeks DID this person work, even for a few hours, including paid vacation, paid sick leave, and military service?

- 50 to 52 weeks
- 48 to 49 weeks
- 40 to 47 weeks
- 27 to 39 weeks
- 14 to 26 weeks
- 13 weeks or less

**40** During the PAST 12 MONTHS, in the WEEKS WORKED, how many hours did this person usually work each WEEK?

Usual hours worked each WEEK



**Person 1 (continued)**

**L** Answer questions 41 – 46 if this person worked in the past 5 years. Otherwise, SKIP to question 47.

**41 – 46 CURRENT OR MOST RECENT JOB ACTIVITY.** Describe clearly this person's chief job activity or business last week. If this person had more than one job, describe the one at which this person worked the most hours. If this person had no job or business last week, give information for his/her last job or business.

**41** Was this person – Mark (X) ONE box

- an employee of a PRIVATE FOR-PROFIT company or business, or of an individual, for wages, salary, or commissions?
- an employee of a PRIVATE NOT-FOR-PROFIT, tax-exempt, or charitable organization?
- a local GOVERNMENT employee (city, county, etc.)?
- a state GOVERNMENT employee?
- a Federal GOVERNMENT employee?
- SELF-EMPLOYED in own NOT INCORPORATED business, professional practice, or farm?
- SELF-EMPLOYED in own INCORPORATED business, professional practice, or farm?
- working WITHOUT PAY in family business or farm?

**42** For whom did this person work?

If now on active duty in the Armed Forces, mark (X) this box →

Name of company, business, or other employer

**43** What kind of business or industry was this?

Describe the activity at the location where employed. (For example, hospital, newspaper publishing, mail order house, auto engine manufacturing, bank)

**44** Is this mainly – Mark (X) ONE box

- manufacturing?
- wholesale trade?
- retail trade?
- other (agriculture, construction, service, government, etc.)?

**45** What kind of work was this person doing? (For example, registered nurse, personnel manager, supervisor of order department, secretary, accountant)

**46** What were this person's most important activities or duties? (For example, patient care, directing, setting policies, supervising order clerks, typing, and filing, reconciling financial records)

**47** INCOME IN THE PAST 12 MONTHS

Mark (X) the "Yes" box for each type of income this person received, and give your best estimate of the TOTAL AMOUNT during the PAST 12 MONTHS. (NOTE: The "past 12 months" is the period from today's date one year ago up through today.)

Mark (X) the "No" box to show type of income NOT received.

If net income was a loss, mark the "Loss" box to the right of the dollar amount.

For income received jointly, report the appropriate share for each person – or, if that's not possible, report the whole amount for (a) one person and mark the "No" box for the other person.

**a. Wages, salary, commissions, bonuses, or tips from all jobs.** Report amount before deductions for taxes, bonds, dues, or other items.

Yes →  TOTAL AMOUNT for past 12 months

No

**b. Self-employment income from own nonfarm businesses or farm businesses, including proprietorships and partnerships.** Report NET income after business expenses.

Yes →  TOTAL AMOUNT for past 12 months  Loss

No

**c. Interest, dividends, net rental income, royalty income, or income from estates and trusts.** Report even small amounts credited to an account.

Yes →  TOTAL AMOUNT for past 12 months  Loss

No

**d. Social Security or Railroad Retirement.**

Yes →  TOTAL AMOUNT for past 12 months

No

**e. Supplemental Security Income (SSI).**

Yes →  TOTAL AMOUNT for past 12 months

No

**f. Any public assistance or welfare payments from the state or local welfare office.**

Yes →  TOTAL AMOUNT for past 12 months

No

**g. Retirement, survivor, or disability pensions.** Do NOT include Social Security.

Yes →  TOTAL AMOUNT for past 12 months

No

**h. Any other sources of income received regularly such as Veterans' (VA) payments, unemployment compensation, child support or alimony.** Do NOT include lump sum payments such as money from an inheritance or the sale of a home.

Yes →  TOTAL AMOUNT for past 12 months

No

**48** What was this person's total income during the PAST 12 MONTHS? Add entries in questions 47a to 47h, subtract any losses. If net income was a loss, enter the amount and mark (X) the "Loss" box next to the dollar amount.

None OR  TOTAL AMOUNT for past 12 months  Loss

Continue with the questions for Person 2 on the next page. If no one is listed as person 2 on page 2, SKIP to page 28 for mailing instructions.



**Person 2**

The balance of the questionnaire has questions for Person 2, Person 3, Person 4, and Person 5. The questions are the same as the questions for Person 1.

INFORMATIONAL COPY



## Mailing Instructions

### Please make sure you have...

- listed all names and answered the questions on pages 2, 3, and 4
- answered all Housing questions
- answered all Person questions for each person.

### Then...

- put the completed questionnaire into the postage-paid return envelope. If the envelope has been misplaced, please mail the questionnaire to:

**U.S. Census Bureau**  
**P.O. Box 5240**  
**Jeffersonville, IN 47199-5240**

- make sure the barcode above your address shows in the window of the return envelope.

**Thank you for participating in  
 the American Community Survey.**

INFORMATIONAL COPY

### For Census Bureau Use

|  |   |                                   |                                  |                                  |
|--|---|-----------------------------------|----------------------------------|----------------------------------|
| POP<br><input type="checkbox"/>        | EDIT<br><input type="checkbox"/>            | PHONE<br><input type="checkbox"/> | JIC1<br><input type="checkbox"/> | JIC2<br><input type="checkbox"/> |
| EDIT CLERK<br><input type="checkbox"/> | TELEPHONE CLERK<br><input type="checkbox"/> |                                   | JIC3<br><input type="checkbox"/> | JIC4<br><input type="checkbox"/> |

The Census Bureau estimates that, for the average household, this form will take 38 minutes to complete, including the time for reviewing the instructions and answers. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to: Paperwork Project 0607-0810, U.S. Census Bureau, 4800 Silver Hill Road, AMSD - 3K138, Washington, D.C. 20233. You may e-mail comments to [Paperwork@census.gov](mailto:Paperwork@census.gov); use "Paperwork Project 0607-0810" as the subject. Please DO NOT RETURN your questionnaire to this address. Use the enclosed pre-addressed envelope to return your completed questionnaire.

Respondents are not required to respond to any information collection unless it displays a valid approval number from the Office of Management and Budget. This 8-digit number appears in the bottom right of the front cover of this form.

Form ACS-1 (INFO) (2012) (KFI) (07-14-2011)





2. Population Census (US Census:国勢調査)  
調査票

United States  
**Census  
2010**

This is the official form for all the people at this address.  
It is quick and easy, and your answers are protected by law.

U.S. DEPARTMENT OF COMMERCE  
Economics and Statistics Administration  
U.S. CENSUS BUREAU

---

Use a blue or black pen.

Start here

The Census must count every person living in the United States on April 1, 2010.

Before you answer Question 1, count the people living in this house, apartment, or mobile home using our guidelines.

- Count all people, including babies, who live and sleep here most of the time.

The Census Bureau also conducts counts in institutions and other places, so:

- Do not count anyone living away either at college or in the Armed Forces.
- Do not count anyone in a nursing home, jail, prison, detention facility, etc., on April 1, 2010.
- Leave these people off your form, even if they will return to live here after they leave college, the nursing home, the military, jail, etc. Otherwise, they may be counted twice.

The Census must also include people without a permanent place to stay, so:

- If someone who has no permanent place to stay is staying here on April 1, 2010, count that person. Otherwise, he or she may be missed in the census.

**1. How many people were living or staying in this house, apartment, or mobile home on April 1, 2010?**

Number of people =

**2. Were there any additional people staying here April 1, 2010 that you did not include in Question 1? Mark (X) all that apply.**

Children, such as newborn babies or foster children

Relatives, such as adult children, cousins, or in-laws

Nonrelatives, such as roommates or live-in baby sitters

People staying here temporarily

No additional people

**3. Is this house, apartment, or mobile home — Mark (X) ONE box.**

Owned by you or someone in this household with a mortgage or loan? *Include home equity loans.*

Owned by you or someone in this household free and clear (without a mortgage or loan)?

Rented?

Occupied without payment of rent?

**4. What is your telephone number? We may call if we don't understand an answer.**

Area Code + Number

-  -

OMB No. 0607-0919-C; Approval Expires 12/31/2011.

Form **D-61** (5-11-2008)

**5. Please provide information for each person living here. Start with a person living here who owns or rents this house, apartment, or mobile home. If the owner or renter lives somewhere else, start with any adult living here. This will be Person 1.**

What is Person 1's name? *Print name below.*

Last Name

First Name  MI

**6. What is Person 1's sex? Mark (X) ONE box.**

Male  Female

**7. What is Person 1's age and what is Person 1's date of birth? Please report babies as age 0 when the child is less than 1 year old. Print numbers in boxes.**

Age on April 1, 2010  Month  Day  Year of birth

→ **NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.**

**8. Is Person 1 of Hispanic, Latino, or Spanish origin?**

No, not of Hispanic, Latino, or Spanish origin

Yes, Mexican, Mexican Am., Chicano

Yes, Puerto Rican

Yes, Cuban

Yes, another Hispanic, Latino, or Spanish origin — *Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.*

**9. What is Person 1's race? Mark (X) one or more boxes.**

White

Black, African Am., or Negro

American Indian or Alaska Native — *Print name of enrolled or principal tribe.*

Asian Indian  Japanese  Native Hawaiian

Chinese  Korean  Guamanian or Chamorro

Filipino  Vietnamese  Samoan

Other Asian — *Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.*

Other Pacific Islander — *Print race, for example, Fijian, Tongan, and so on.*

Some other race — *Print race.*

**10. Does Person 1 sometimes live or stay somewhere else?**

No  Yes — *Mark (X) all that apply.*

In college housing  For child custody

In the military  In jail or prison

At a seasonal or second residence  In a nursing home

For another reason

→ If more people were counted in Question 1, continue with Person 2.

U.S. CENSUS BUREAU

1. Print name of **Person 2**

Last Name

First Name  MI

2. How is this person related to Person 1? Mark  ONE box.

- |   |  |
|---|--|
| <input type="checkbox"/> Husband or wife            | <input type="checkbox"/> Parent-in-law                 |
| <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Son-in-law or daughter-in-law |
| <input type="checkbox"/> Adopted son or daughter    | <input type="checkbox"/> Other relative                |
| <input type="checkbox"/> Stepson or stepdaughter    | <input type="checkbox"/> Roomer or boarder             |
| <input type="checkbox"/> Brother or sister          | <input type="checkbox"/> Housemate or roommate         |
| <input type="checkbox"/> Father or mother           | <input type="checkbox"/> Unmarried partner             |
| <input type="checkbox"/> Grandchild                 | <input type="checkbox"/> Other nonrelative             |

3. What is this person's sex? Mark  ONE box.

- Male  Female

4. What is this person's age and what is this person's date of birth?

Please report babies as age 0 when the child is less than 1 year old.

Print numbers in boxes.

Age on April 1, 2010      Month      Day      Year of birth

→ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this census, Hispanic origins are not races.

5. Is this person of Hispanic, Latino, or Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin — Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. ↗

6. What is this person's race? Mark  one or more boxes.

- White
- Black, African Am., or Negro
- American Indian or Alaska Native — Print name of enrolled or principal tribe. ↗

- |  |                                     |  |
|--|-------------------------------------|--|
| <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese   | <input type="checkbox"/> Native Hawaiian       |
| <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean     | <input type="checkbox"/> Guamanian or Chamorro |
| <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese | <input type="checkbox"/> Samoan                |
| <input type="checkbox"/> Other Asian — Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ↗ |                                     |  |
| <input type="checkbox"/> Other Pacific Islander — Print race, for example, Fijian, Tongan, and so on. ↗                  |                                     |  |

- Some other race — Print race. ↗

7. Does this person sometimes live or stay somewhere else?

- No  Yes — Mark  all that apply.
- |  |   |
|--|---|
| <input type="checkbox"/> In college housing                | <input type="checkbox"/> For child custody  |
| <input type="checkbox"/> In the military                   | <input type="checkbox"/> In jail or prison  |
| <input type="checkbox"/> At a seasonal or second residence | <input type="checkbox"/> In a nursing home  |
|  | <input type="checkbox"/> For another reason |

→ If more people were counted in Question 1 on the front page, continue with Person 3.

1. Print name of **Person 3**

Last Name

First Name  MI

2. How is this person related to Person 1? Mark  ONE box.

- |   |  |
|---|--|
| <input type="checkbox"/> Husband or wife            | <input type="checkbox"/> Parent-in-law                 |
| <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Son-in-law or daughter-in-law |
| <input type="checkbox"/> Adopted son or daughter    | <input type="checkbox"/> Other relative                |
| <input type="checkbox"/> Stepson or stepdaughter    | <input type="checkbox"/> Roomer or boarder             |
| <input type="checkbox"/> Brother or sister          | <input type="checkbox"/> Housemate or roommate         |
| <input type="checkbox"/> Father or mother           | <input type="checkbox"/> Unmarried partner             |
| <input type="checkbox"/> Grandchild                 | <input type="checkbox"/> Other nonrelative             |

3. What is this person's sex? Mark  ONE box.

- Male  Female

4. What is this person's age and what is this person's date of birth?

Please report babies as age 0 when the child is less than 1 year old.

Print numbers in boxes.

Age on April 1, 2010      Month      Day      Year of birth

→ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this census, Hispanic origins are not races.

5. Is this person of Hispanic, Latino, or Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin — Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. ↗

6. What is this person's race? Mark  one or more boxes.

- White
- Black, African Am., or Negro
- American Indian or Alaska Native — Print name of enrolled or principal tribe. ↗

- |  |                                     |  |
|--|-------------------------------------|--|
| <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese   | <input type="checkbox"/> Native Hawaiian       |
| <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean     | <input type="checkbox"/> Guamanian or Chamorro |
| <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese | <input type="checkbox"/> Samoan                |
| <input type="checkbox"/> Other Asian — Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ↗ |                                     |  |
| <input type="checkbox"/> Other Pacific Islander — Print race, for example, Fijian, Tongan, and so on. ↗                  |                                     |  |

- Some other race — Print race. ↗

7. Does this person sometimes live or stay somewhere else?

- No  Yes — Mark  all that apply.
- |  |   |
|--|---|
| <input type="checkbox"/> In college housing                | <input type="checkbox"/> For child custody  |
| <input type="checkbox"/> In the military                   | <input type="checkbox"/> In jail or prison  |
| <input type="checkbox"/> At a seasonal or second residence | <input type="checkbox"/> In a nursing home  |
|  | <input type="checkbox"/> For another reason |

→ If more people were counted in Question 1 on the front page, continue with Person 4.

1. Print name of **Person 4**

Last Name

First Name  MI

2. How is this person related to Person 1? Mark  ONE box.

- |   |  |
|---|--|
| <input type="checkbox"/> Husband or wife            | <input type="checkbox"/> Parent-in-law                 |
| <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Son-in-law or daughter-in-law |
| <input type="checkbox"/> Adopted son or daughter    | <input type="checkbox"/> Other relative                |
| <input type="checkbox"/> Stepson or stepdaughter    | <input type="checkbox"/> Roomer or boarder             |
| <input type="checkbox"/> Brother or sister          | <input type="checkbox"/> Housemate or roommate         |
| <input type="checkbox"/> Father or mother           | <input type="checkbox"/> Unmarried partner             |
| <input type="checkbox"/> Grandchild                 | <input type="checkbox"/> Other nonrelative             |

3. What is this person's sex? Mark  ONE box.

- Male  Female

4. What is this person's age and what is this person's date of birth?

Please report babies as age 0 when the child is less than 1 year old.

Print numbers in boxes.

Age on April 1, 2010  Month  Day  Year of birth

→ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this census, Hispanic origins are not races.

5. Is this person of Hispanic, Latino, or Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin — Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. ↴

6. What is this person's race? Mark  one or more boxes.

- White
- Black, African Am., or Negro
- American Indian or Alaska Native — Print name of enrolled or principal tribe. ↴

- |  |                                     |   |
|--|-------------------------------------|---|
| <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese   | <input type="checkbox"/> Native Hawaiian  |
| <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean     | <input type="checkbox"/> Guamanian or Chamorro  |
| <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese | <input type="checkbox"/> Samoan   |
| <input type="checkbox"/> Other Asian — Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ↴ |                                     | <input type="checkbox"/> Other Pacific Islander — Print race, for example, Fijian, Tongan, and so on. ↴ |

- Some other race — Print race. ↴

7. Does this person sometimes live or stay somewhere else?

- No  Yes — Mark  all that apply.
- |  |   |
|--|---|
| <input type="checkbox"/> In college housing                | <input type="checkbox"/> For child custody  |
| <input type="checkbox"/> In the military                   | <input type="checkbox"/> In jail or prison  |
| <input type="checkbox"/> At a seasonal or second residence | <input type="checkbox"/> In a nursing home  |
|  | <input type="checkbox"/> For another reason |

→ If more people were counted in Question 1 on the front page, continue with Person 5.

1. Print name of **Person 5**

Last Name

First Name  MI

2. How is this person related to Person 1? Mark  ONE box.

- |   |  |
|---|--|
| <input type="checkbox"/> Husband or wife            | <input type="checkbox"/> Parent-in-law                 |
| <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Son-in-law or daughter-in-law |
| <input type="checkbox"/> Adopted son or daughter    | <input type="checkbox"/> Other relative                |
| <input type="checkbox"/> Stepson or stepdaughter    | <input type="checkbox"/> Roomer or boarder             |
| <input type="checkbox"/> Brother or sister          | <input type="checkbox"/> Housemate or roommate         |
| <input type="checkbox"/> Father or mother           | <input type="checkbox"/> Unmarried partner             |
| <input type="checkbox"/> Grandchild                 | <input type="checkbox"/> Other nonrelative             |

3. What is this person's sex? Mark  ONE box.

- Male  Female

4. What is this person's age and what is this person's date of birth?

Please report babies as age 0 when the child is less than 1 year old.

Print numbers in boxes.

Age on April 1, 2010  Month  Day  Year of birth

→ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this census, Hispanic origins are not races.

5. Is this person of Hispanic, Latino, or Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin — Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. ↴

6. What is this person's race? Mark  one or more boxes.

- White
- Black, African Am., or Negro
- American Indian or Alaska Native — Print name of enrolled or principal tribe. ↴

- |  |                                     |   |
|--|-------------------------------------|---|
| <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese   | <input type="checkbox"/> Native Hawaiian  |
| <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean     | <input type="checkbox"/> Guamanian or Chamorro  |
| <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese | <input type="checkbox"/> Samoan   |
| <input type="checkbox"/> Other Asian — Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ↴ |                                     | <input type="checkbox"/> Other Pacific Islander — Print race, for example, Fijian, Tongan, and so on. ↴ |

- Some other race — Print race. ↴

7. Does this person sometimes live or stay somewhere else?

- No  Yes — Mark  all that apply.
- |  |   |
|--|---|
| <input type="checkbox"/> In college housing                | <input type="checkbox"/> For child custody  |
| <input type="checkbox"/> In the military                   | <input type="checkbox"/> In jail or prison  |
| <input type="checkbox"/> At a seasonal or second residence | <input type="checkbox"/> In a nursing home  |
|  | <input type="checkbox"/> For another reason |

→ If more people were counted in Question 1 on the front page, continue with Person 5.

1. Print name of **Person 6**

Last Name

First Name  MI

2. How is this person related to Person 1? Mark  ONE box.

- |   |  |
|---|--|
| <input type="checkbox"/> Husband or wife            | <input type="checkbox"/> Parent-in-law                 |
| <input type="checkbox"/> Biological son or daughter | <input type="checkbox"/> Son-in-law or daughter-in-law |
| <input type="checkbox"/> Adopted son or daughter    | <input type="checkbox"/> Other relative                |
| <input type="checkbox"/> Stepson or stepdaughter    | <input type="checkbox"/> Roomer or boarder             |
| <input type="checkbox"/> Brother or sister          | <input type="checkbox"/> Housemate or roommate         |
| <input type="checkbox"/> Father or mother           | <input type="checkbox"/> Unmarried partner             |
| <input type="checkbox"/> Grandchild                 | <input type="checkbox"/> Other nonrelative             |

3. What is this person's sex? Mark  ONE box.

- Male  Female

4. What is this person's age and what is this person's date of birth?

Please report babies as age 0 when the child is less than 1 year old.  
Print numbers in boxes.

Age on April 1, 2010    Month    Day    Year of birth  
           

→ NOTE: Please answer BOTH Question 5 about Hispanic origin and Question 6 about race. For this census, Hispanic origins are not races.

5. Is this person of Hispanic, Latino, or Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin  
 Yes, Mexican, Mexican Am., Chicano  
 Yes, Puerto Rican  
 Yes, Cuban  
 Yes, another Hispanic, Latino, or Spanish origin — Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. ↴

6. What is this person's race? Mark  one or more boxes.

- White  
 Black, African Am., or Negro  
 American Indian or Alaska Native — Print name of enrolled or principal tribe. ↴

- |  |   |  |
|--|---|--|
| <input type="checkbox"/> Asian Indian  | <input type="checkbox"/> Japanese   | <input type="checkbox"/> Native Hawaiian       |
| <input type="checkbox"/> Chinese   | <input type="checkbox"/> Korean   | <input type="checkbox"/> Guamanian or Chamorro |
| <input type="checkbox"/> Filipino  | <input type="checkbox"/> Vietnamese   | <input type="checkbox"/> Samoan                |
| <input type="checkbox"/> Other Asian — Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ↴ | <input type="checkbox"/> Other Pacific Islander — Print race, for example, Fijian, Tongan, and so on. ↴ |  |

- Some other race — Print race. ↴

7. Does this person sometimes live or stay somewhere else?

- No  Yes — Mark  all that apply.
- |  |   |
|--|---|
| <input type="checkbox"/> In college housing                | <input type="checkbox"/> For child custody  |
| <input type="checkbox"/> In the military                   | <input type="checkbox"/> In jail or prison  |
| <input type="checkbox"/> At a seasonal or second residence | <input type="checkbox"/> In a nursing home  |
|  | <input type="checkbox"/> For another reason |

→ If more than six people were counted in Question 1 on the front page, turn the page and continue.

INTERNATIONAL COPY

→ If more people live here, turn the page and continue.

Use this section to complete information for the rest of the people you counted in Question 1 on the front page. We may call for additional information about them.

|  |                      |                      |                      |   |
|--|----------------------|----------------------|----------------------|---|
| <b>Person 7</b>  |                      |                      |                      |   |
| Last Name  |                      | First Name           |                      | MI  |
| <input type="text"/>   |                      | <input type="text"/> |                      | <input type="text"/>  |
| Sex  | Age on April 1, 2010 | Date of Birth        |                      | Related to Person 1?  |
| <input type="checkbox"/> Male<br><input type="checkbox"/> Female | <input type="text"/> | Month                | Day                  | Year  |
|  |                      | <input type="text"/> | <input type="text"/> | <input type="text"/>  |
|  |                      |                      |                      | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <b>Person 8</b>  |                      |                      |                      |   |
| Last Name  |                      | First Name           |                      | MI  |
| <input type="text"/>   |                      | <input type="text"/> |                      | <input type="text"/>  |
| Sex  | Age on April 1, 2010 | Date of Birth        |                      | Related to Person 1?  |
| <input type="checkbox"/> Male<br><input type="checkbox"/> Female | <input type="text"/> | Month                | Day                  | Year  |
|  |                      | <input type="text"/> | <input type="text"/> | <input type="text"/>  |
|  |                      |                      |                      | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <b>Person 9</b>  |                      |                      |                      |   |
| Last Name  |                      | First Name           |                      | MI  |
| <input type="text"/>   |                      | <input type="text"/> |                      | <input type="text"/>  |
| Sex  | Age on April 1, 2010 | Date of Birth        |                      | Related to Person 1?  |
| <input type="checkbox"/> Male<br><input type="checkbox"/> Female | <input type="text"/> | Month                | Day                  | Year  |
|  |                      | <input type="text"/> | <input type="text"/> | <input type="text"/>  |
|  |                      |                      |                      | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <b>Person 10</b>   |                      |                      |                      |   |
| Last Name  |                      | First Name           |                      | MI  |
| <input type="text"/>   |                      | <input type="text"/> |                      | <input type="text"/>  |
| Sex  | Age on April 1, 2010 | Date of Birth        |                      | Related to Person 1?  |
| <input type="checkbox"/> Male<br><input type="checkbox"/> Female | <input type="text"/> | Month                | Day                  | Year  |
|  |                      | <input type="text"/> | <input type="text"/> | <input type="text"/>  |
|  |                      |                      |                      | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <b>Person 11</b>   |                      |                      |                      |   |
| Last Name  |                      | First Name           |                      | MI  |
| <input type="text"/>   |                      | <input type="text"/> |                      | <input type="text"/>  |
| Sex  | Age on April 1, 2010 | Date of Birth        |                      | Related to Person 1?  |
| <input type="checkbox"/> Male<br><input type="checkbox"/> Female | <input type="text"/> | Month                | Day                  | Year  |
|  |                      | <input type="text"/> | <input type="text"/> | <input type="text"/>  |
|  |                      |                      |                      | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <b>Person 12</b>   |                      |                      |                      |   |
| Last Name  |                      | First Name           |                      | MI  |
| <input type="text"/>   |                      | <input type="text"/> |                      | <input type="text"/>  |
| Sex  | Age on April 1, 2010 | Date of Birth        |                      | Related to Person 1?  |
| <input type="checkbox"/> Male<br><input type="checkbox"/> Female | <input type="text"/> | Month                | Day                  | Year  |
|  |                      | <input type="text"/> | <input type="text"/> | <input type="text"/>  |
|  |                      |                      |                      | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |

**Thank you for completing your official 2010 Census form.**

**FOR OFFICIAL USE ONLY**

|                      |                      |
|----------------------|----------------------|
| JIC1                 | JIC2                 |
| <input type="text"/> | <input type="text"/> |

*If your enclosed postage-paid envelope is missing, please mail your completed form to:*  
U.S. Census Bureau  
National Processing Center  
1201 East 10th Street  
Jeffersonville, IN 47132

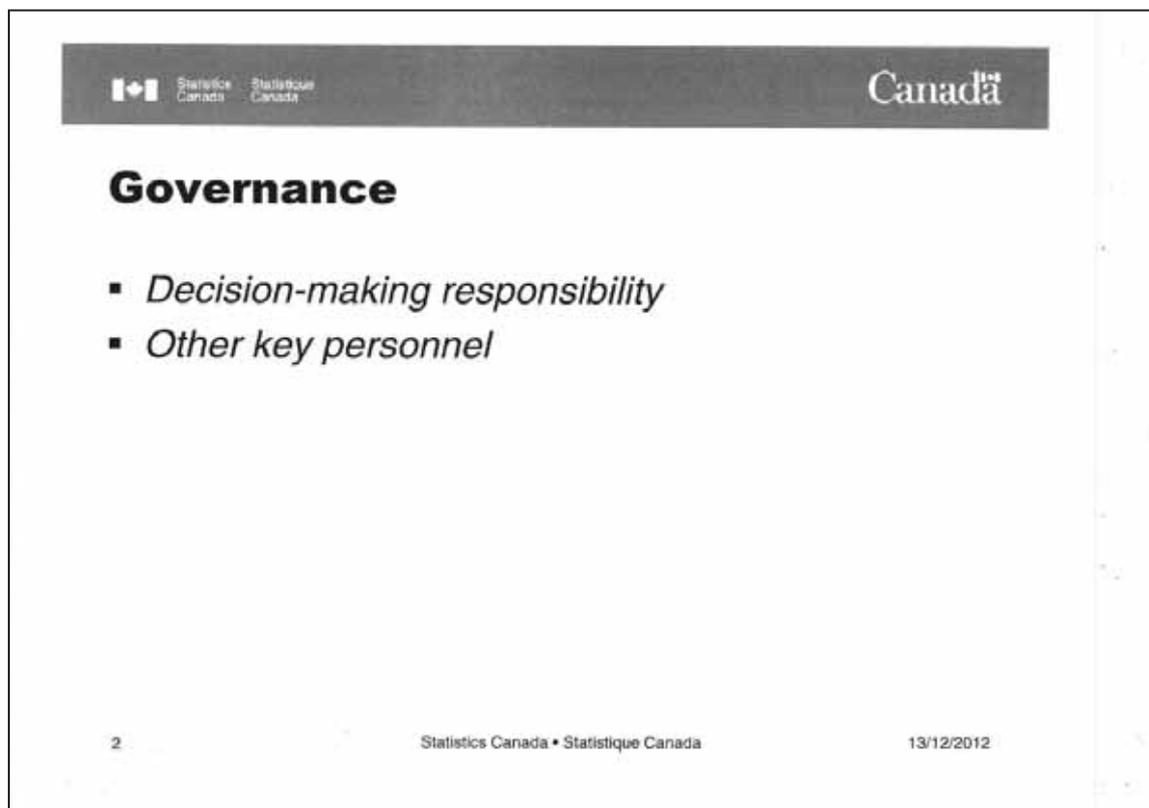
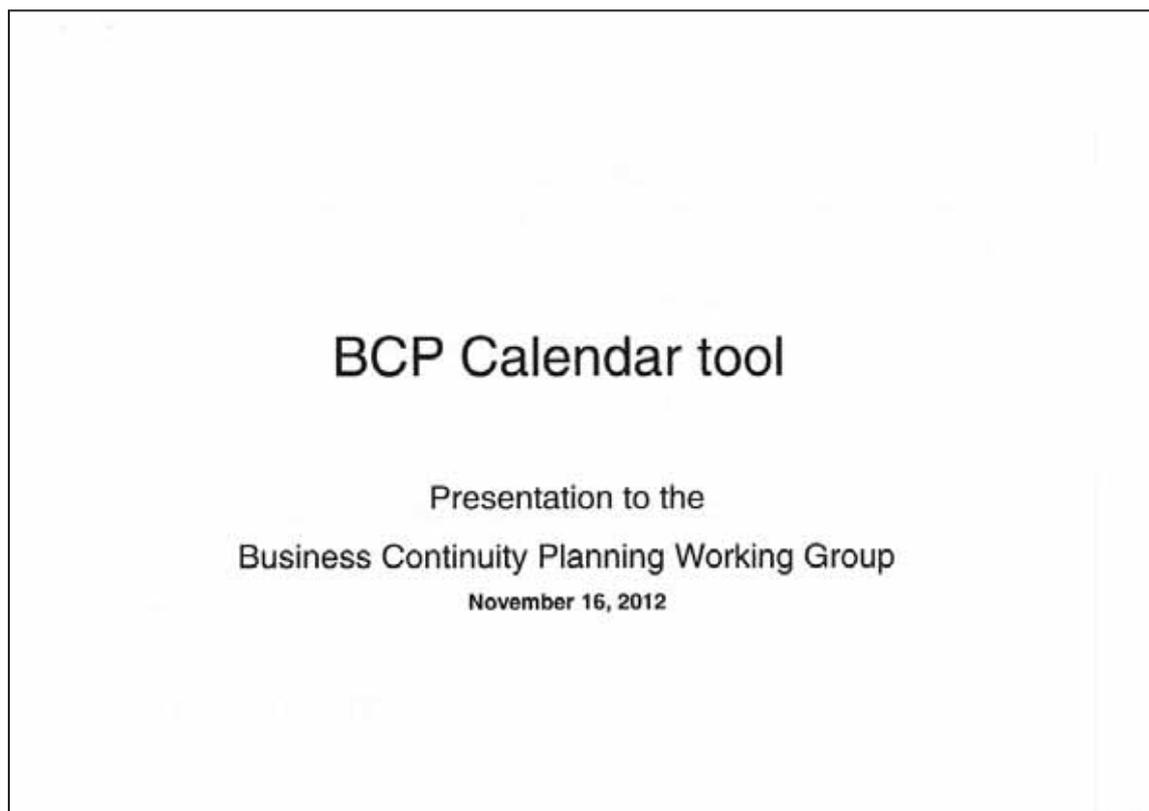
*If you need help completing this form, call 1-866-872-6868 between 8:00 a.m. and 9:00 p.m., 7 days a week. The telephone call is free.*

**TDD** — *Telephone display device for the hearing impaired. Call 1-866-783-2010 between 8:00 a.m. and 9:00 p.m., 7 days a week. The telephone call is free.*

**¿NECESITA AYUDA?** *Si usted necesita ayuda para completar este cuestionario, llame al 1-866-828-2010 entre las 8:00 a.m. y 9:00 p.m., 7 días a la semana. La llamada telefónica es gratis.*

The U.S. Census Bureau estimates that, for the average household, this form will take about 10 minutes to complete, including the time for reviewing the instructions and answers. Send comments regarding this burden estimate or any other aspect of this burden to: Paperwork Reduction Project 0607-0919-C, U.S. Census Bureau, AMSD-3K138, 4600 Silver Hill Road, Washington, DC 20233. You may e-mail comments to <Paperwork@census.gov>; use "Paperwork Project 0607-0919-C" as the subject.

Respondents are not required to respond to any information collection unless it displays a valid approval number from the Office of Management and Budget.



出所)カナダ統計局提供資料



## **Business Impact Analysis**

- *Mandate*
- *Priority for services or products*
- *Impact of disruption*
- *Minimum acceptable level of service*
- *Key dependencies*
  - *Internal*
  - *External*
- *Other*



## **Continuity Activities**

- *State any assumptions on which BCP is based*
- *List requirements to ensure acceptable level of continuity*

## Response Plan

- *Show activation decision flowchart or describe activation criteria*
- *Communications plan for response*
- *Who reports on progress/status of continuity effort?*
- *Describe how minimum acceptable service level will be maintained with increasing levels of severity of impact*

## Resumption Plan

- *What is required to return to full level of service following disruption?*

## **BCP Readiness – program level**

- *Status of training/awareness programs*
- *When will training be provided?*
- *What type of exercises are planned?*
- *Approach to lessons learned following incident/exercise*
- *Maintenance schedule (updates)*

## **Outstanding Issues**

- *Report on any issues/concerns outstanding for which input required by SCC*

## 附属資料Ⅲ 参考ウェブサイト・文献一覧

| 国内研究事例関連  |                   |
|---|-------------------|
| <a href="http://www.stat.go.jp/data/e-census/index.htm">http://www.stat.go.jp/data/e-census/index.htm</a>   | 経済センサス            |
| <a href="http://www.mhlw.go.jp/toukei/list/20-21.html">http://www.mhlw.go.jp/toukei/list/20-21.html</a>   | 国民生活基礎調査          |
| <a href="http://www.meti.go.jp/statistics/tyo/tokusabizi/index.html">http://www.meti.go.jp/statistics/tyo/tokusabizi/index.html</a>   | 特定サービス産業実態調査      |
| <a href="http://www.boj.or.jp/statistics/tk/">http://www.boj.or.jp/statistics/tk/</a>   | 全国企業短期経済観測調査      |
| <a href="http://www.ism.ac.jp/kokuminsei/index.html">http://www.ism.ac.jp/kokuminsei/index.html</a>   | 日本人の国民性調査         |
| 内閣府経済社会研究所 <i>ESRI Discussion Paper</i><br>統計からみた震災からの復興<br><a href="http://www.esri.go.jp/jp/archive/e_dis/e_dis286/e_dis286.html">http://www.esri.go.jp/jp/archive/e_dis/e_dis286/e_dis286.html</a> | 東日本大震災への対応        |
| 労働力調査における東日本大震災に伴う補完推計<br><a href="http://www.stat.go.jp/data/roudou/120424/index.htm">http://www.stat.go.jp/data/roudou/120424/index.htm</a>   | 東日本大震災への対応(労働力調査) |

## 海外事例関連

### ACS (American Community Survey: アメリカ地域社会調査)

|   |                 |
|---|-----------------|
| <a href="http://www.census.gov/acs/www/">http://www.census.gov/acs/www/</a>   | ACSに関するトップページ   |
| <a href="http://www.census.gov/acs/www/methodology/sample_size_and_data_quality/">http://www.census.gov/acs/www/methodology/sample_size_and_data_quality/</a>   | 回収率や代入率など       |
| <a href="http://www.census.gov/acs/www/methodology/methodology_main/Design_and_Methodology_Report">http://www.census.gov/acs/www/methodology/methodology_main/Design_and_Methodology_Report</a>   | 補完手法など          |
| <a href="http://www.census.gov/acs/www/Downloads/survey_methodology/acs_design_methodology.pdf">http://www.census.gov/acs/www/Downloads/survey_methodology/acs_design_methodology.pdf</a>   | 調査手法に関する解説      |
| <i>American Community Survey Accuracy of the Data (2011)</i><br><a href="http://www.census.gov/acs/www/Downloads/data_documentation/Accuracy/ACS_Accuracy_of_Data_2011.pdf">http://www.census.gov/acs/www/Downloads/data_documentation/Accuracy/ACS_Accuracy_of_Data_2011.pdf</a> | データ精度に関する報告     |
| <a href="http://www.census.gov/newsroom/emergencies/hurricane_katrina.html">http://www.census.gov/newsroom/emergencies/hurricane_katrina.html</a>   | ハリケーン・カトリーナへの対応 |

### Population Census (US Census: 国勢調査)

|   |  |
|---|--|
| <a href="http://www.census.gov/2010census/">http://www.census.gov/2010census/</a>   | Censusに関するトップページ   |
| <a href="http://www.census.gov/2010census/partners/materials/engguide-materials.php">http://www.census.gov/2010census/partners/materials/engguide-materials.php</a>   | 問合せ対応のためのマニュアル、Language Identification Flashcard (言語カード)など |
| <i>2010 Census Item Nonresponse and Imputation Assessment Report</i><br><a href="http://www.census.gov/2010census/pdf/2010_Census_INR_Imputation_Assessment.pdf">http://www.census.gov/2010census/pdf/2010_Census_INR_Imputation_Assessment.pdf</a>   | 欠測値補完に関する報告  |
| <i>2010 Census Nonresponse Followup Operations Assessment Report</i><br><a href="http://www.census.gov/2010census/pdf/2010_Census_NRFU_Operations_Assessment.pdf">http://www.census.gov/2010census/pdf/2010_Census_NRFU_Operations_Assessment.pdf</a> | ユニット非回答を防ぐ取組(NRFU)に関する報告                                   |
| <i>2010 Census Quality Survey</i><br><a href="http://www.census.gov/2010census/pdf/2010_Census_Quality_Survey.pdf">http://www.census.gov/2010census/pdf/2010_Census_Quality_Survey.pdf</a>  | 補完手法を含め Census 全体の品質に関する報告                                 |

| <b>OES(Occupational Employment Statistics: 職業雇用統計)</b>  |                                 |
|---|---------------------------------|
| <a href="http://www.bls.gov/oes/home.htm">http://www.bls.gov/oes/home.htm</a>   | OESに関するトップページ                   |
| <i>Survey Methods and Reliability Statement for May 2011</i><br><a href="http://www.bls.gov/oes/2011/may/methods_statement.pdf">http://www.bls.gov/oes/2011/may/methods_statement.pdf</a>   | 補完手法など調査手法に関する解説                |
| <b>LFS(Labour Force Survey: 労働力調査)</b>  |                                 |
| <a href="http://www23.statcan.gc.ca/imdb-bmdi/pub/3701-eng.htm">http://www23.statcan.gc.ca/imdb-bmdi/pub/3701-eng.htm</a>   | LFSに関するトップページ                   |
| <i>Methodology of the Canadian Labour Force Survey</i><br><a href="http://www.statcan.gc.ca/pub/71-526-x/71-526-x2007001-eng.pdf">http://www.statcan.gc.ca/pub/71-526-x/71-526-x2007001-eng.pdf</a>   | ホット・デック法の基準など調査手法に関する解説         |
| <i>Guide to the Labour Force Survey</i><br><a href="http://www.statcan.gc.ca/pub/71-543-g/71-543-g2012001-eng.pdf">http://www.statcan.gc.ca/pub/71-543-g/71-543-g2012001-eng.pdf</a>  | 調査内容などに関する解説                    |
| <b>CES(Current Employment Survey: 雇用状況調査)</b>   |                                 |
| <a href="http://www.bls.gov/ces/">http://www.bls.gov/ces/</a>   | CESに関するトップページ                   |
| <a href="http://www.bls.gov/katrina/cpscesquestions.htm">http://www.bls.gov/katrina/cpscesquestions.htm</a>   | ハリケーン・カトリーナへの対応                 |
| <a href="http://www.bls.gov/opub/mlr/2006/08/contents.htm">http://www.bls.gov/opub/mlr/2006/08/contents.htm</a>   |                                 |
| <i>Hurricane Katrina's effects on industry employment and wages</i><br><a href="http://www.bls.gov/opub/mlr/2006/08/art3full.pdf">http://www.bls.gov/opub/mlr/2006/08/art3full.pdf</a>  |                                 |
| <b>CPS(Current Population Survey: 就業状態調査)</b>   |                                 |
| <a href="http://www.bls.gov/cps/">http://www.bls.gov/cps/</a>   | CPSに関するトップページ                   |
| <a href="http://www.bls.gov/katrina/cpscesquestions.htm">http://www.bls.gov/katrina/cpscesquestions.htm</a>   | ハリケーン・カトリーナへの対応                 |
| <a href="http://www.bls.gov/opub/mlr/2006/08/contents.htm">http://www.bls.gov/opub/mlr/2006/08/contents.htm</a>   |                                 |
| <i>The Current Population Survey response to Hurricane Katrina</i><br><a href="http://www.bls.gov/opub/mlr/2006/08/art4full.pdf">http://www.bls.gov/opub/mlr/2006/08/art4full.pdf</a>   |                                 |
| <i>The effect of Hurricane Katrina on employment and unemployment</i><br><a href="http://www.bls.gov/opub/mlr/2006/08/art5full.pdf">http://www.bls.gov/opub/mlr/2006/08/art5full.pdf</a>  |                                 |
| <b>その他</b>  |                                 |
| <a href="http://www.whitehouse.gov/omb/inforeg_statpolicy">http://www.whitehouse.gov/omb/inforeg_statpolicy</a><br><i>Statistical Programs of the United States Government</i><br><a href="http://www.whitehouse.gov/omb/regulatory_affairs/reports_previous_yrs/">http://www.whitehouse.gov/omb/regulatory_affairs/reports_previous_yrs/</a> | アメリカにおける統計調査(プログラムの)の予算規模を含む全体像 |
| <a href="http://www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?lang=eng&amp;catno=10H0034">http://www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?lang=eng&amp;catno=10H0034</a>   | Banffに関するトップページ                 |
| <i>THE BANFF SYSTEM FOR AUTOMATED EDITING AND IMPUTATION</i><br><a href="http://www.ssc.ca/survey/documents/SSC2005_R_Kozak.pdf">http://www.ssc.ca/survey/documents/SSC2005_R_Kozak.pdf</a>   | Banffの機能について                    |
| <a href="http://www.isr.umich.edu/src/smp/ive/">http://www.isr.umich.edu/src/smp/ive/</a>   | IVEwareについて                     |
| <a href="http://www.westat.com/">http://www.westat.com/</a>   | westat社                         |
| <i>Edit and Imputation : From suspicious to scientific techniques</i><br><a href="http://isi.cbs.nl/iamamember/CD2/pdf/430.PDF">http://isi.cbs.nl/iamamember/CD2/pdf/430.PDF</a>  | 欠測値補完の歴史について                    |

## 補完手法に関連する資料

|   |                       |
|---|-----------------------|
| 不完全データの統計解析 岩崎学(エコノミスト社)  | 欠測値を含むデータの分析について      |
| 調査観察データの統計科学-因果推論・選択バイアス・データ融合- 星野崇宏(岩波書店)  | 欠測値を含むデータの分析について      |
| <i>Handbook of Statistical Data Editing and Imputation</i> , T.D.Waal, J. Pannekoek, S.Scholtus (Wiley Handbooks in Survey Methodology)   | 欠測値を含むデータの分析について      |
| <i>A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models</i><br><a href="http://www.statcan.gc.ca/ads-annonces/12-001-x/5857-eng.pdf">http://www.statcan.gc.ca/ads-annonces/12-001-x/5857-eng.pdf</a> | 多重代入法について             |
| 統計数理研究所「統計数理」第60巻<br>明治末期における小学生の理想人物調査 —キャリブレーション手法の比較—<br><a href="http://www.ism.ac.jp/editsec/toukei/pdf/60-1-219.pdf">http://www.ism.ac.jp/editsec/toukei/pdf/60-1-219.pdf</a>   | キャリブレーションについて         |
| <i>Variance estimation under composite imputation: The methodology behind SEVANI</i><br><a href="http://www.statcan.gc.ca/pub/12-001-x/2011002/article/11605-eng.pdf">http://www.statcan.gc.ca/pub/12-001-x/2011002/article/11605-eng.pdf</a>       | 分散分析(SEVANI)について      |
| 総務省統計研修所 統計研究彙報 第66号 No.4<br>ヨーロッパにおけるデータエディティング及び補定に関する調査報告 ～EDIMBUSプロジェクトを中心に～<br><a href="http://www.stat.go.jp/training/2kenkyu/2-2-664.htm">http://www.stat.go.jp/training/2kenkyu/2-2-664.htm</a>   | 欧州における統計調査の補完について     |
| 独立行政法人統計センター 製表技術参考資料<br>諸外国のデータエディティング及び混淆正規分布モデルによる多変量外れ値検出法についての研究<br><a href="http://www.nstac.go.jp/services/pdf/sankousiryu2408.pdf">http://www.nstac.go.jp/services/pdf/sankousiryu2408.pdf</a>  | データエディティング及び外れ値検出について |