

# 標本交代方式を採る統計調査の標本バイアス

山口幸三\*

## 要旨

公的統計調査のうち月次または四半期ごとの経常調査では、推定値の精度を高めるために、標本抽出において、標本を順次交代する方式を採っている。そのような標本交代方式を採る統計調査では、交代するそれぞれのグループが同質でないことや複数回調査される場合の世帯・世帯員の回答行動が変化することによって、偏りが生じていると考えられる。

本稿では、労働力調査のデータを用いて、そうした偏りの有無やその特徴を検証することとし、8組の副標本を組み合わせた組別標本を比較し分析する方法で行った。その結果、世帯・世帯員の回答行動によって、偏りが生じているものの、結果の推定値に与える影響は限定的であると確認できた。

## キーワード

標本交代方式、標本バイアス、副標本、集計用乗率、労働力調査

## 1. はじめに

### 1.1 問題意識

公的統計を作成するための統計調査のうち、月次または四半期ごとに調査する経常調査では、標本の選択について、抽出した標本を一定期間固定して調査する方法、月次または四半期ごとに新たな標本を抽出する方法、月次または四半期ごとに一部の標本を順次交代する方法が考えられる。経常調査では、調査結果の時系列データの精度を高め、記入者負担を考慮し、かつ標本を長期に固定化することにより母集団の代表性が損なわれないようにするために、一般的に標本を順次交代する方法を採っている<sup>1)</sup>。標本を交代する方法については、それぞれの経常調査で異なっている。

標本交代する方法を採っている調査の例として、内閣府の消費動向調査、総務省の労働

力調査、家計調査、家計消費状況調査、厚生労働省の毎月勤労統計調査などが挙げられる。これらの調査のうち、労働力調査以外では、一度交代すると、再び標本になることはない。

標本交代する方法を採る世帯・個人を対象とする世帯調査の標本では、交代するそれぞれの標本グループが同質でない、または同質性が保たれずに偏りが生じている、複数回調査される場合の世帯・世帯員の回答行動によって偏りが生じている可能性が考えられる。なお、各標本グループの推定値が特定の傾向や特徴がみられる場合に、偏りがあると考えている。したがって、この偏りは、非標本誤差のうちの標本抽出段階および実地調査段階で生じる系統的な誤差とみなせる。

そこで、その偏りの有無、その特徴について調べ、標本交代方式がもたらす標本構造の解明を行う。具体的には、わが国の就業・不就業の状態を毎月調べる労働力調査を用いる。分析データ<sup>2)</sup>の対象期間は、労働力調査が改

\* 総務省統計研修所

e-mail : k3.yamaguchi@soumu.go.jp

正された2002年前後の1995年～2008年とし、主たる調査項目である就業状態を分析指標として用いる。

## 1.2 標本バイアスに関する研究

先行研究としては、佐井・加納（2004）が1989年から2001年までの労働力調査を用いて、調査回数による影響について分析しており、失業率<sup>3)</sup>は、調査回数が多いと低くなることや減少—増加—減少という変動傾向を示すことを指摘している。さらに、2001年以前の調査票の「求職理由」という項目の有無<sup>4)</sup>が回答に影響している可能性を示唆している。本稿の後述する4組の標本の分析が、この調査回数の影響の分析に相当している。また、芳賀（1984）は労働力調査のデータを用いて、データ構造のモデルを仮定し、月間、調査区間、調査区内・世帯群内の分散構成の推定を行い、交代標本の分析を行っている。加納（2003）は労働力調査を例に取り、ローテーション・サンプリングによって得られる調査結果を分析するための時系列モデルを提示している。実際のデータではなく、擬似データによってローテーション構造の持つサンプリングの長所・短所について述べている。

わが国では、調査方法や調査結果の利用に関する研究は、個票データ利用の制約<sup>5)</sup>等もあり、海外に比べて遅れている。海外では、例えば、Bailar（1975）は様々な調査におけるローテーション・バイアスの存在を紹介し、併せて、米国のCPS（Current Population Survey）におけるバイアスが推定に与える影響について分析している。その他にも、Williams and Malloes（1970）、Ghangurde（1982）、Solon（1986）、McLaren and Steel（1997）などがある。各国のローテーションの方式は、日本や米国が1年後に再び調査されるのに対して、カナダやオーストラリアは継続して調査されるものの、1年後には調査されない、などさまざまである。また、ローテーション

構造によるバイアスや相関がもたらす推定値への影響やそれらを考慮した推定方法なども研究されている。例えば、Lents, Miller and Cantwell（1996）、Bell and Carolan（1998）などである。

なお、2002年の労働力調査の改正において、特定調査票を2年目2か月目の住戸で調査するに当たり、土生・高橋（2003）は調査結果への影響を検討しており、2年目2か月目の標本の結果と全標本の結果との乖離は小さいと結論付けている。

## 2. 労働力調査の標本設計と推定方法

ここで、本稿の分析の方法と結果を理解するのに必要であると思われる、分析対象期間における労働力調査の標本設計と推定方法について、簡潔に説明しておく。

### 2.1 標本設計

労働力調査の標本抽出は、第一次抽出単位を国勢調査の調査区<sup>6)</sup>、第二次抽出単位を住戸<sup>7)</sup>とする二段抽出法を採用し、第一次抽出単位の調査区は、いくつかの特性に分類（層化）され、各層ごとに抽出（層別抽出）されている。第一次抽出単位の標本の大きさは、約2900の調査区<sup>8)</sup>、第二次抽出単位の標本の大きさは、調査区内の住戸に居住する世帯約15世帯、全体で調査世帯約4万世帯、15歳以上の世帯員約10万人である。

#### (1) 層化

調査区の層化は、国勢調査の結果等に基づく特性によるものであり、基準としては、産業・従業上の地位別就業者数の構成、寮・寄宿舎、病院・療養所、社会施設、給与住宅の有無を用いている。産業別や従業上の地位別の就業者は利用上重要であり、寮・寄宿舎等に居住している人は就業状態が均質であるので、そうした施設が含まれるかどうか調査区の特性を左右するためである。世帯数の少ない調査区は層化の効果が小さいとして、こ

これらの地域をまとめて1つの層としている。層化基準の詳細は、総務省（総務庁）統計局（1996, 1999, 2005, 2008）を参照のこと。

このように調査区の層化は、就業者を把握するには適した層化基準となっているが、失業者や非労働力人口を把握するのに、層化の効果は限定的であると考えられる（近藤・山口（1990））。

**(2) 新設集団住宅地域**

労働力調査の第一次抽出単位である調査区内に国勢調査以後住宅が集団的に新設された地域の一部又は全部が含まれる場合、その地域が調査区の特徴を大きく変化させたと考えられる。層別抽出における層化の効果を維持するため、国勢調査以後住宅が集団的に新設された地域を新設集団住宅地域とし、その地域を調査区から分離している。この結果、労働力調査の標本調査区では新設集団住宅地域に係る部分を除いた範囲を調査することになる。一方、分離した新設集団住宅地域については、上述の層とは別に新設集団住宅地域全体を1つの層として、標本となる地域を抽出

し、さらに地域内を調査区と同規模の単位区に再分割し、その中から標本となる単位区を抽出し、単位区内を調査している。このような新設集団住宅地域は、1962年7月調査から設定され、2002年7月調査まで調査地域に充てられていた。

**(3) 標本交代**

標本交代は1961年から次のように行われている。標本調査区については、4か月継続して調査され、8か月離れて、翌年の同期に再び調査される。毎月、標本調査区全体の1/4は他の調査区に交代する。

標本調査区は、1/4ずつ交代に対応した標本調査区の調査開始月による4区分、今年新たに調査する標本調査区（1年目）と前年標本調査区となり今年再び調査する標本調査区（2年目）の2区分によって、8組に分けられる。それぞれの組の標本を副標本という。この8組の標本調査区は、同質性を持つように抽出されている。つまり、同一層からそれぞれ8組の調査区が抽出され、8組とも同じ層別構成になっている。ただし、02層、03層、

図1 労働力調査の標本交代

組	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月
1	a <sub>13-1</sub>	a <sub>13-2</sub>	a <sub>14-1</sub>	a <sub>14-2</sub>	a <sub>15-1</sub>	a <sub>15-2</sub>	a <sub>16-1</sub>	a <sub>16-2</sub>	a <sub>17-1</sub>	a <sub>17-2</sub>	a <sub>18-1</sub>	a <sub>18-2</sub>	a <sub>19-1</sub>	a <sub>19-2</sub>
2	A <sub>7-1</sub>	A <sub>7-1</sub>	A <sub>8-1</sub>	A <sub>8-2</sub>	A <sub>9-1</sub>	A <sub>9-2</sub>	A <sub>10-1</sub>	A <sub>10-2</sub>	A <sub>11-1</sub>	A <sub>11-2</sub>	A <sub>12-1</sub>	A <sub>12-2</sub>	A <sub>13-1</sub>	A <sub>13-2</sub>
3	b <sub>12-2</sub>	b <sub>13-1</sub>	b <sub>13-2</sub>	b <sub>14-1</sub>	b <sub>14-2</sub>	b <sub>15-1</sub>	b <sub>15-2</sub>	b <sub>16-1</sub>	b <sub>16-2</sub>	b <sub>17-1</sub>	b <sub>17-2</sub>	b <sub>18-1</sub>	b <sub>18-2</sub>	b <sub>19-1</sub>
4	B <sub>6-2</sub>	B <sub>7-1</sub>	B <sub>7-1</sub>	B <sub>8-1</sub>	B <sub>8-2</sub>	B <sub>9-1</sub>	B <sub>9-2</sub>	B <sub>10-1</sub>	B <sub>10-2</sub>	A <sub>11-1</sub>	A <sub>11-2</sub>	A <sub>12-1</sub>	A <sub>12-2</sub>	A <sub>13-1</sub>
5	c <sub>12-1</sub>	c <sub>12-2</sub>	c <sub>13-1</sub>	c <sub>13-2</sub>	c <sub>14-1</sub>	c <sub>14-2</sub>	c <sub>15-1</sub>	c <sub>15-2</sub>	c <sub>16-1</sub>	c <sub>16-2</sub>	c <sub>17-1</sub>	c <sub>17-2</sub>	c <sub>18-1</sub>	c <sub>18-2</sub>
6	C <sub>6-1</sub>	C <sub>6-2</sub>	C <sub>7-1</sub>	C <sub>7-1</sub>	C <sub>8-1</sub>	C <sub>8-2</sub>	C <sub>9-1</sub>	C <sub>9-2</sub>	C <sub>10-1</sub>	C <sub>10-2</sub>	C <sub>11-1</sub>	C <sub>11-2</sub>	C <sub>12-1</sub>	C <sub>12-2</sub>
7	d <sub>11-2</sub>	d <sub>12-1</sub>	d <sub>12-2</sub>	d <sub>13-1</sub>	d <sub>13-2</sub>	d <sub>14-1</sub>	d <sub>14-2</sub>	d <sub>15-1</sub>	d <sub>15-2</sub>	d <sub>16-1</sub>	d <sub>16-2</sub>	d <sub>17-1</sub>	d <sub>17-2</sub>	d <sub>18-1</sub>
8	D <sub>5-2</sub>	D <sub>6-1</sub>	D <sub>6-2</sub>	D <sub>7-1</sub>	D <sub>7-1</sub>	D <sub>8-1</sub>	D <sub>8-2</sub>	D <sub>9-1</sub>	D <sub>9-2</sub>	D <sub>10-1</sub>	D <sub>10-2</sub>	D <sub>11-1</sub>	D <sub>11-2</sub>	D <sub>12-1</sub>

(注) 英字は調査開始月の区分を表し、英小文字は1年目、英大文字は2年目を表す。組符号の1～8は調査開始月の区分と1年目・2年目の区分によるものである。添え字の最初の数字は、調査区内の住戸（調査世帯）グループを表し、「-」の後の数字「1」、「2」は、住戸（調査世帯）がそれぞれ1か月目、2か月目であることを表す。

例えば、a<sub>13-1</sub>は1月開始の調査区の住戸（調査世帯）で1か月目の調査に当たることを示し、a<sub>13-1</sub>、a<sub>13-2</sub>、A<sub>13-1</sub>、A<sub>13-2</sub>は、継続して4回調査される住戸（調査世帯）の1か月目、2か月目、3か月目、4か月目を表している。

0401層, 0402層, 0403層, 0404層（以下「02層～0404層」と呼称<sup>9)</sup>）については、層ごとに8組揃わないため、併せて1つの層とみなして標本抽出している。

標本調査区内の住戸は、2か月継続して調査され、2か月後に同一標本調査区内で他の住戸と交代するが、翌年の同期に再び調査される。つまり、2か月調査され、10か月離れ、2か月調査され、計4回調査される。

労働力調査では、標本交代のほかに、5年に1度、国勢調査の標本抽出関係資料を用いて、最新の国勢調査の調査区に段階的に切り替える。まずは1年目の調査区を4か月かけて切り替え、その8か月後に2年目の調査区を4か月かけて切り替えるので、16か月目ですべての調査区が切り替わることになる。

#### (4) 調査の改正

労働力調査は、2002年に見直しが行われ、改正されている。その改正は、年1回（1999年～2001年は年2回）実施していた労働力調査特別調査を労働力調査に統合して、2種類の調査票を導入するものであった。従来の労働力調査の調査票を基礎調査票、労働力調査特別調査の調査票を特定調査票とし、基礎調査票については全調査世帯に、特定調査票については2年目2か月目の世帯（調査世帯全体の1/4の世帯）を対象に調査するようになった。

## 2.2 推定方法

労働力調査の毎月の基本集計の全国結果は、大都市部とそれ以外の非大都市部、男女、年齢5歳階級（14区分）別に、国勢調査に基づく推計人口を基準人口とする比推定によって算出している。

算出の基本式は、就業者数を例にとれば、次のとおりである。ここで、線型推定とは、調査で得られた人口に抽出率の逆数（＝線型推定乗率）を掛け、全体の人口を推定することである。

就業者数＝

$$\text{線型推定による就業者数} \times \frac{\text{基準人口}}{\text{線型推定による人口}}$$

なお、 $\frac{\text{基準人口}}{\text{線型推定による人口}}$  は比推定乗率

という。

この調査では、調査員が住戸に居住する世帯を調査することになっているが、自衛官の営舎内居住者および刑務所・拘置所等の矯正施設収容者のデータについては、それぞれ関係の省から資料を得て、直接推定されている。

集計するために、次のように各調査客体に集計用乗率を算出している。

客体の集計用乗率＝

客体の線型推定乗率×客体の比推定乗率

## 3. 組別調査区数による分析

前述した問題意識から、実際の労働力調査の標本において、①8組の副標本の同質性が保たれずに偏りが生じている、②4回調査される住戸に居住する世帯の回答行動によって偏りが生じている可能性が考えられるので、その偏りの有無、特徴について検証する。また、労働力調査は2002年に改正されているために、2002年以前と以後のデータを用いることによって、改正による影響、つまり改正前後での偏りの有無、特徴に変化が生じているのかについても併せて検証する。

### 3.1 組別調査区数による分析方法

労働力調査の8組の副標本は、それぞれ独立して、同質の標本になるように抽出されている。ここでの同質というのは、層別の調査区数の構成が同じであることを想定している。

8組の副標本ごとに標本調査区の層符号（調査区の層化基準）別調査区数および就業者数等を集計する。これらの結果を比較することによって、8組の標本のそれぞれの調査区の構成や特徴が同質であるかどうかを調べる。実際の調査においては、実地調査上の問

題からすべての標本調査区のデータが得られないこともありうるので、データ欠落の影響も含まれる。

### 3.2 組別調査区数による比較

8組の副標本は、ほぼ同数の標本調査区が充てられている。しかし、2004年以前の標本調査区数は、正確には毎年同数ではなく、組ごとにも異なっている。その理由は、新設集団住宅地域を追加抽出していたこと、その追加抽出を廃止したこと、2000年国勢調査の調査区に切り替える時に標本調査区数を2880から2912に拡大したことである。その上、これらの標本の追加、廃止、拡大において、時系列結果の安定性のために、ある一時点で一斉に切り替えるのではなく、一定期間かけて副標本ごとに順次切り替えていることも影響している。

その結果として、標本設計上の標本調査区数は、1998年、1999年、2001年、2002年には2年目の標本調査区数が多く、2003年は1年目の標本調査区数が多くなる。しかし、標本調査区数の差がそのまま世帯・世帯員数の差とは言えない。標本拡大については、世帯・世帯員数の増加に直接的に影響するが、

新設集団住宅地域の設定により調査区数は純増するものの、世帯・世帯員数は新設集団住宅地域分が増加したとしても、既存の標本調査区では減少することになるので、ある程度は相殺されていると考えられるからである。実際に調査され、回答された標本調査区数は標本設計上の標本調査区数を若干下回っている。回答されない標本調査区の中には、世帯が居住していない場合も含まれる(図2<sup>10)</sup>。

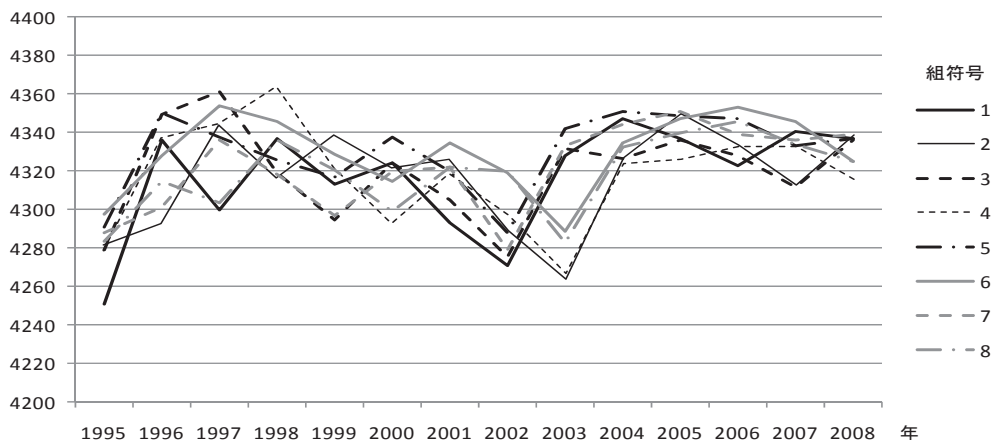
8組の標本調査区数の変動は、直接的には15歳以上人口に影響すると考えられるが、15歳以上人口には、標本調査区数と同じような動きを確認できなかった(図3)。

## 4. 組別標本による分析

### 4.1 組別標本による分析方法

8組の副標本ごとに最も主要な調査事項である就業状態(就業者、失業者、非労働力人口)別15歳以上人口のそれぞれの年平均を集計する。年平均にするのは、月次では標本規模が小さいこと、季節変動の影響を受けることを考慮したためである。この8組の副標本を組み替えて、組別標本に偏りが存在するかどうかを検証する。つまり、8組ごとに集計した値が、複数年にわたり特定の傾向や特

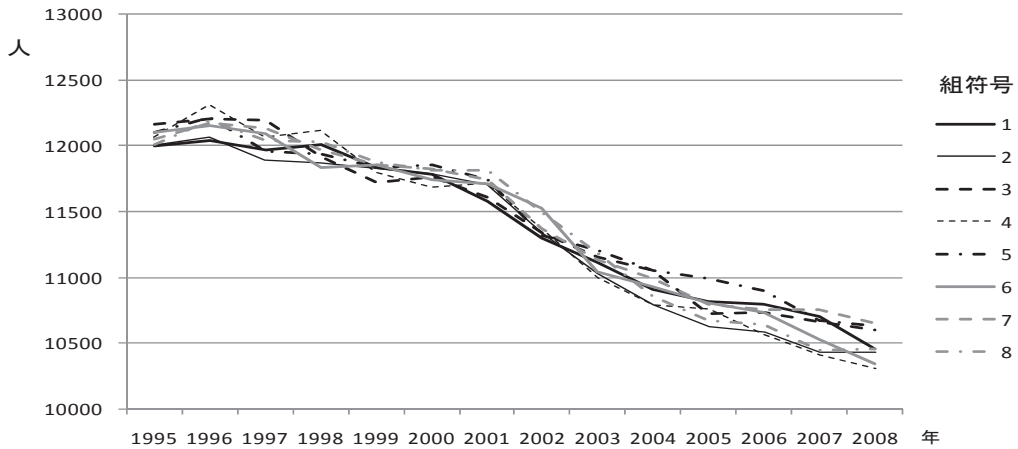
図2 組別調査区数の推移



資料) 総務省統計局「労働力調査」を用いて作成(以下の図も同じ)。



図3 組別15歳以上人口（原集計数）の推移



徴があるかどうかを調べる。

02層～0404層も含めた標本を用いた集計値の比較と、02層～0404層を除いた標本を用いた集計値での比較の2種類について行う。これらの層の標本を除くのは、前述のとおり組によって含まれる層が異なるからである。比較する集計値については、乗率を用いないで集計した値（以下「原集計数」と呼称）、線型推定乗率を用いて集計した値（以下「線型推定値」と呼称）、集計用乗率を用いて集計した値（以下「推定値」と呼称）の3種類で行う。なお、新設集団住宅地域については、推定値でも比較するために含めた方がよいと判断した。集計用乗率については、線型推定乗率と比推定乗率を掛けることで求められるが、8組を合わせた全組での比推定乗率を用いるのではなく、8組ごとに求めた比推定乗率を用いる。したがって、推定値<sup>11)</sup>において、15歳以上人口は、8組すべて同じ数値になる。比較するのは15歳以上人口および就業状態別15歳以上人口とし、実数または15歳以上人口に対する割合（構成比）によって比較する。

組別標本の組替えについては、月次ごとに8組の副標本を、4組（1年目1か月目、1年目2か月目、2年目1か月目、2年目2か月目）別に組み替えて、特定の傾向や特徴が

表れるかどうかを分析する。また、8組の副標本ごとに同一住戸における調査回数（1回目、2回目、3回目、4回目）別に組み替えて、調査回数別に特定の傾向や特徴があるかどうかを調べる。

#### 4.2 8組の標本による比較による比較

8組の副標本は、前述のとおり調査区における調査開始月による4区分と1年目、2年目の調査区の2区分によって分けられている。8組については、1組、3組、5組、7組の1年目と2組、4組、6組、8組の2年目で分けられるが、それぞれの4組は、標本調査区の1か月目、2か月目、3か月目、4か月目、住戸の1か月目、2か月目がすべて同じ構成になっている。したがって、8組については、1年目と2年目の2組での違いはあるが、1年目内の4組、2年目内の4組については、調査開始月の違いはあるものの、年単位で何か月目の調査区によって構成されているかということで見ると、それぞれ標本の構成は同じになっている<sup>12)</sup>。8組について、標本バイアスを生む構造的要因には、標本の調査からの脱落や調査への参入のような不規則なものでなく、調査区数の違い、02層～0404層の標本の違い、合併前の層の違いなどが考えら

れる。調査区数の違いは、前述の「3 組別調査区数による比較」では特になかったように、年次ごとに組別の15歳以上人口（原集計数）の違いに注目しても、1か月目と2か月目、1年目と2年目での違いは認められるものの、調査区数の違いによる特徴は確認できなかった（図3）。組別就業者（原集計数）でみると、組符号1組と2組は他の組よりも多数である傾向がみられたが、02層～0404層の標本の違いによると思われる、偏りは生じているが、推定値ではそうした特徴も表れないので、集計用乗率がその偏りを補正していることになる。合併前の層の違いについては、合併前の情報が得られないので、検証することは難しい。

これらのことを踏まえて、はっきりとした特定の傾向や特徴はみられず、標本バイアスがあったとしても、どの程度のものかは捉えられない。そこで、4回調査される住戸に居住する世帯・世帯員の回答行動によって生じるとされる偏りに焦点を合わせて、詳細にみていくことにする。ここでの回答行動とは、複数回調査される間の回答の揺らぎや非回答などを想定している。

#### 4.3 4組の標本による比較

8組の副標本を調査月ごとに住戸が1年目1か月目、1年目2か月目、2年目1か月目、2年目2か月目の4組になるように再編する。4組に再編するのに、例えば、奇数の調査月は、それぞれ組符号を1と5、3と7、2と6、4と8、図1の記号を使えば、調査月1月は $a_{13-1}$ と $c_{12-1}$ 、 $b_{12-2}$ と $d_{11-2}$ 、 $A_{7-1}$ と $C_{6-1}$ 、 $B_{6-2}$ と $D_{5-2}$ のように組み合わせる。再編した4組別に年平均を求め、就業状態別15歳以上人口について比較する。

原集計数でみると、図4によれば、就業者は、2001年以前、2002年、2003年以後では動きが異なる。2001年以前では、ほとんど増減はなく、横ばい傾向である。2002年は直線的な増加と少し特異な動きをしている。2003年以後では、増加—減少—増加の動きをしながら、2年目に減少している。図5によれば、失業者は、2001年以前と2002年以後では動きが逆になっている。2001年以前は減少—増加—減少、2002年以後では増加—減少—増加の動きをしている。図6によれば、非労働力人口は、増加—減少—増加の動きをしながら、2002年以前は2年目が増加、2003年以後は2年目が減少している。

図4 4組別就業者数（原集計数）

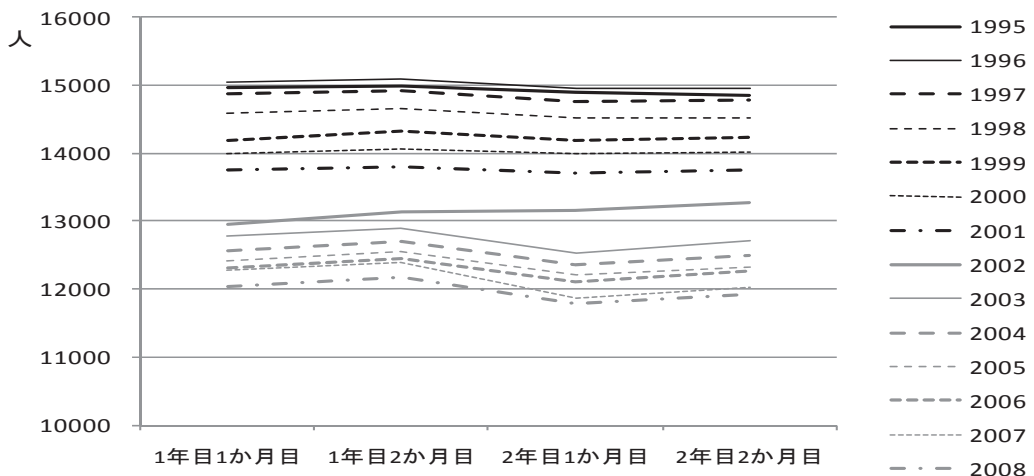


図5 4組別失業者数（原集計数）

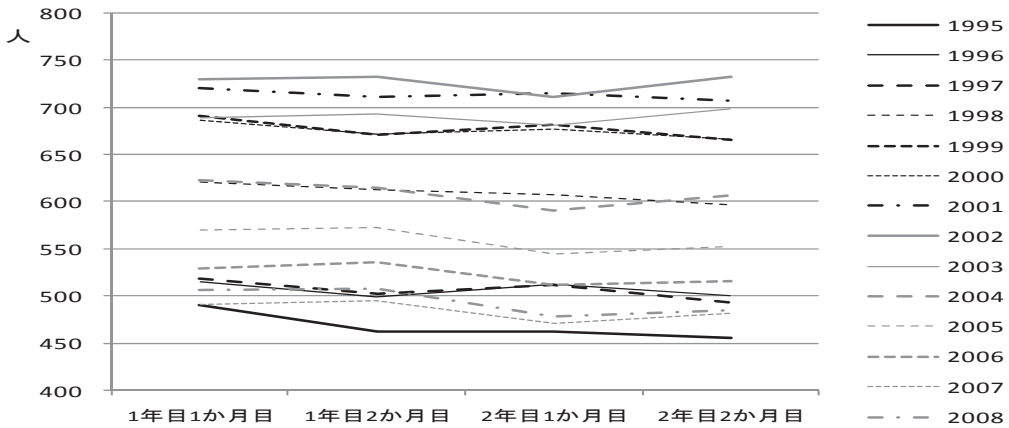


図6 4組別非労働力人口（原集計数）

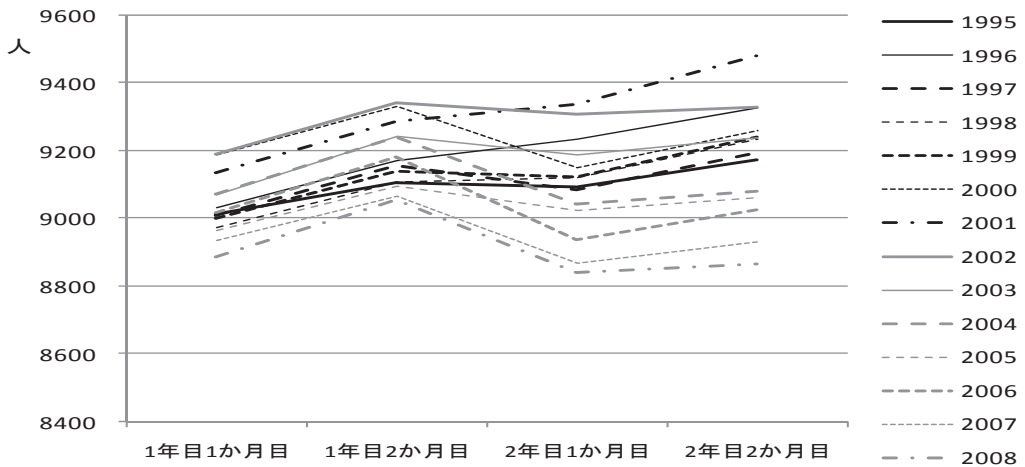


図7 4組別就業者数（推定値）

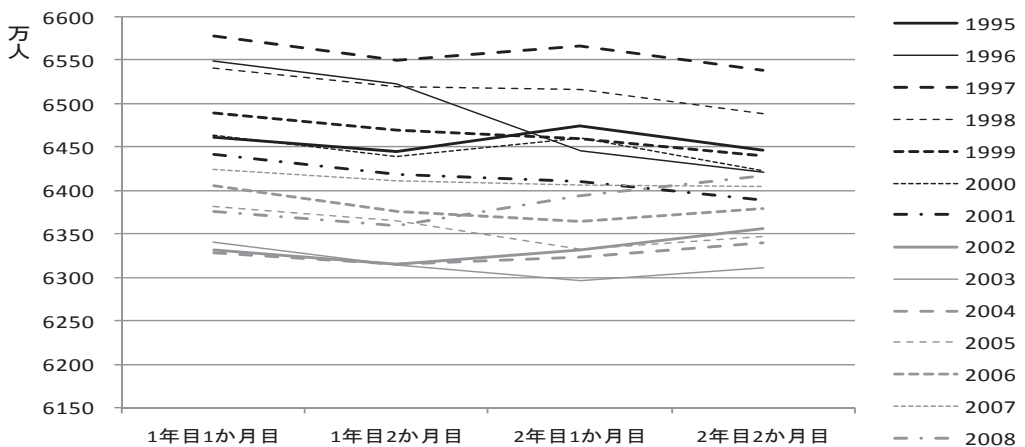




図8 4組別失業者数（推定値）

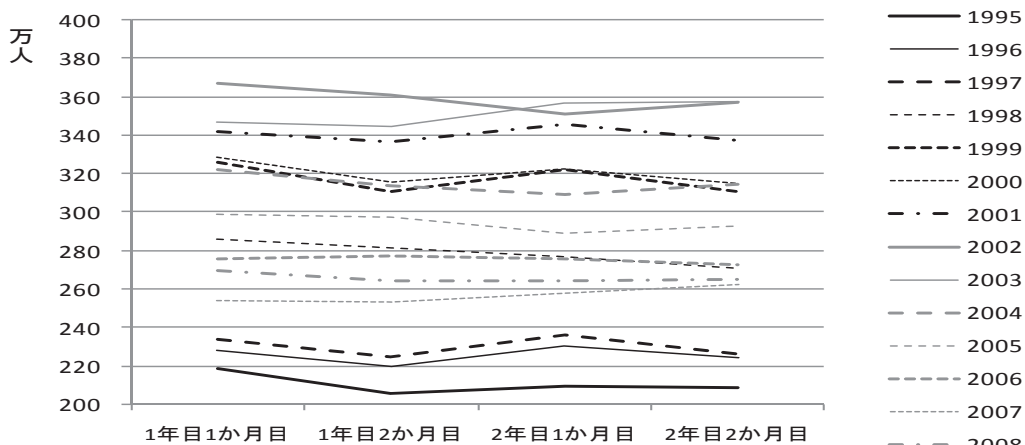
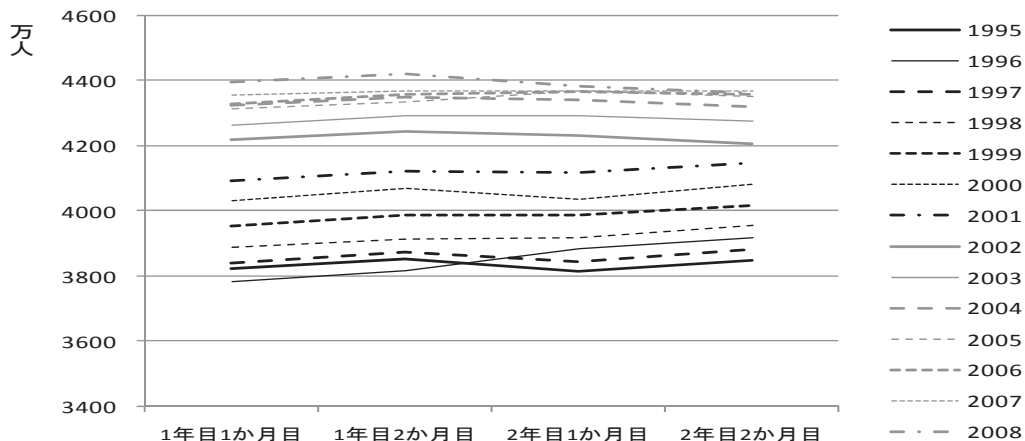


図9 4組別非労働力人口（推定値）



02層～0404層を除いた原集計数でみると、上述の02層～0404層を含めた原集計数と比べて、数値の水準は低くなるものの、動きは変わらない。

推定値でみると、世帯・世帯員の移動や世帯・世帯員の回答行動による人口構成の偏りを補正するとともに、15歳以上人口が4組とも同じ水準になる。そのために、図7によれば、就業者は、2003年以後の原集計数の動きのように、一律に増加と減少を繰り返す動きはない。図8によれば、失業者は、原集計数の場合と同じ動きをしている。図9によれば、非労働力人口は、増加と減少の動きは

原集計数に比べ小さい。

なお、次の「4.4 調査回数別の標本による比較」も併せて、線型推定値、構成比でみた場合については、上述した原集計数、推定値の実数でみた場合と比べて、特に違った特徴はみられなかったため、本稿では取り上げないこととした。

4.4 調査回数別の標本による比較

住戸は1年目の2か月、2年目の2か月の合計4か月、つまり4回調査される。住戸に居住する調査世帯・世帯員も移動がなければ、4回調査されることになる。8組の副標本を

調査回数1回目、調査回数2回目、調査回数3回目、調査回数4回目の4組になるように再編する。調査回数別に再編するのに、例えば、調査月1月の1回目は組符号1と5が該当し、1回目が組符号1の場合は、2回目以降は2月の組符号1、翌年の1月の組符号2、2月の組符号2、図1の記号を使えば、 $a_{13-1}$ 、 $a_{13-2}$ 、 $A_{13-1}$ 、 $A_{13-2}$ のように組み合わせる。再編した調査回数別に年平均を求め、就業状態別15歳以上人口について比較する。

調査回数別は、時間とともに変化する15歳以上人口を把握することになり、1回目か

ら4回目までの期間は14か月ということになる。前述の4組はある一時点の標本であり、1か月目、2か月目、1年目、2年目となっても、求めている推定値は、時間的な違いはなく、同じ調査時点のものである。

原集計数でみると、図10によれば、就業者は、増加—減少—増加の動きを示しながら、2年目には減少している。図11によれば、失業者は、2001年以前と2002年以後とで異なる動きをしている。2001年以前は減少—増加—減少、2002年以後は横ばい—減少—増加になっている。図12によれば、非労働力

図10 調査回数別就業者数（原集計数）

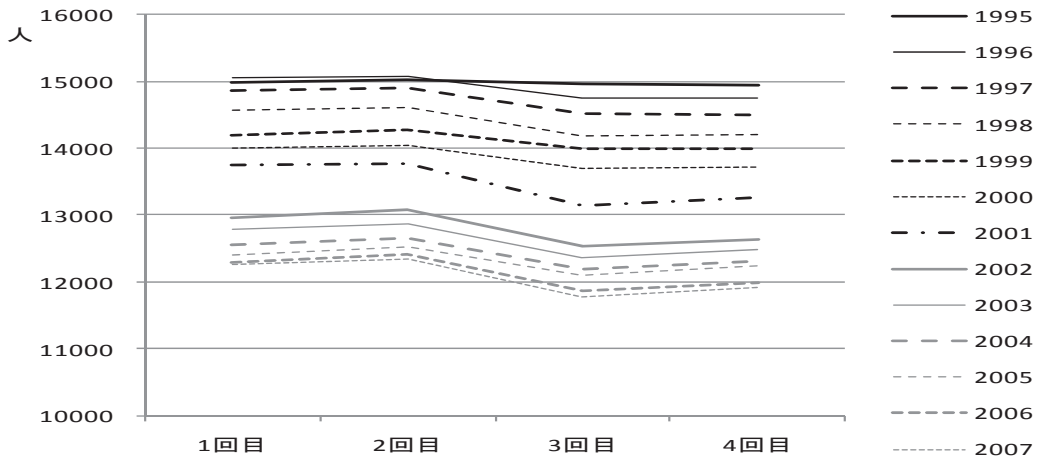


図11 調査回数別失業者数（原集計数）

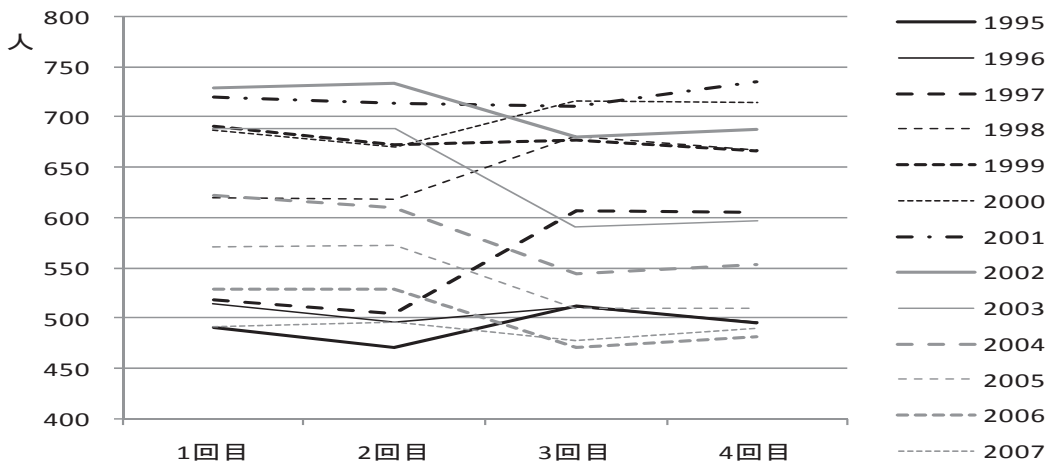


図12 調査回数別非労働力人口（原集計数）

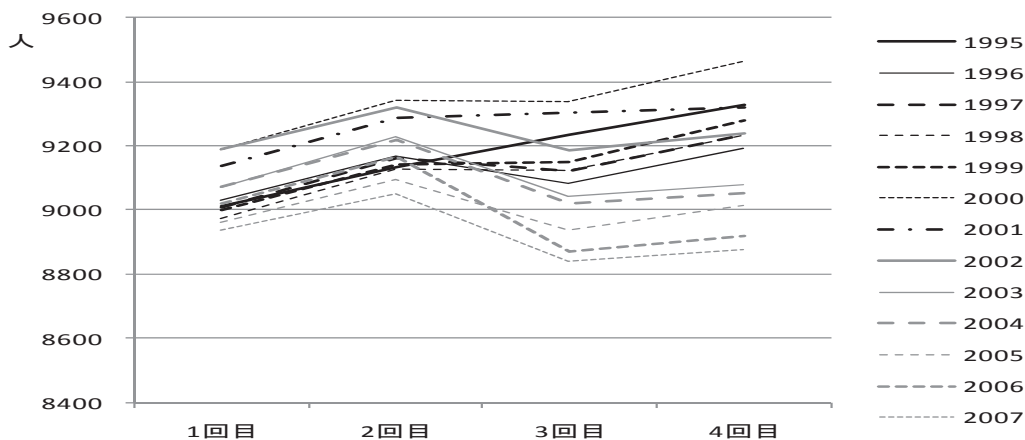
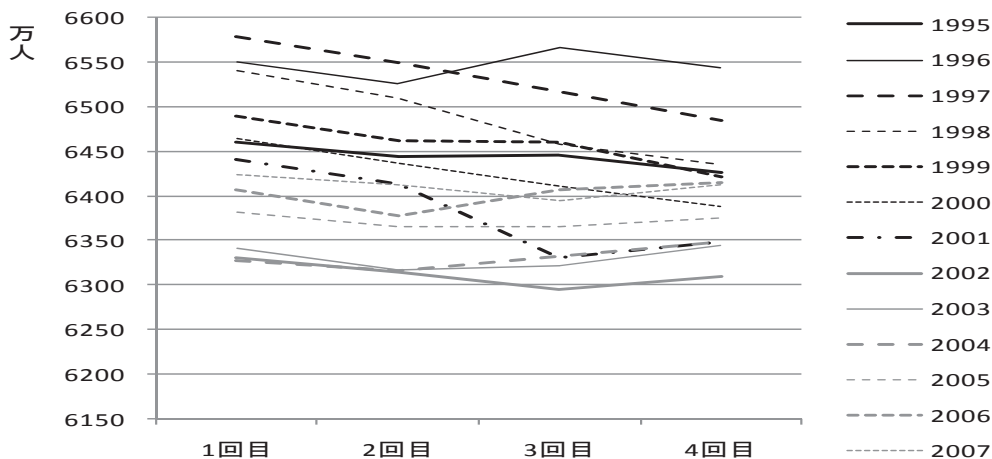


図13 調査回数別就業者数（推定値）



人口は、増加—減少—増加の動きを示しながら、2001年以前は増加傾向、2002年以後では減少傾向になっている。

時系列でみて、2年目が翌年の1年目と同じ水準ならば、増減傾向と合致した実態を表していることになるが、実際には就業者は2年目が2002年以前は高い水準、2003年以後は低い水準となっている。失業者は2年目が低い水準、非労働力人口は2年目が2003年以前は高い水準、2004年以後は低い水準となっている。

推定値でみると、図13によれば、就業者は、

2000年以前では減少—減少—減少の動きを示しながら、1年目から2年目には減少しているが、2001年・2002年では減少—減少—増加、2003年以後では減少—増加—増加と変化している。図14によれば、失業者は、2001年以前と2002年以後とで異なる動きしている。2001年以前は減少—増加—減少、2002年以後は減少—減少—増加になっている。図15によれば、非労働力人口は、2000年以前では増加—増加—増加、2001年以後では増加—増加—横ばいとなっている。

図14 調査回数別失業者数（推定値）

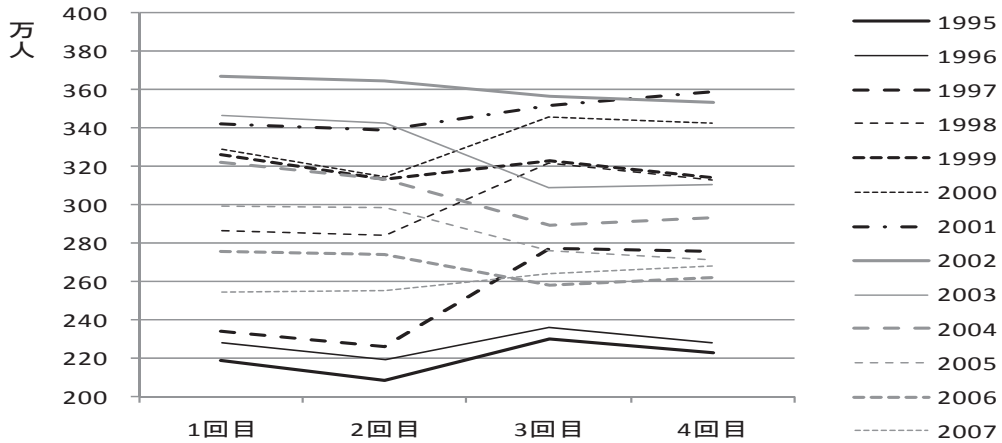
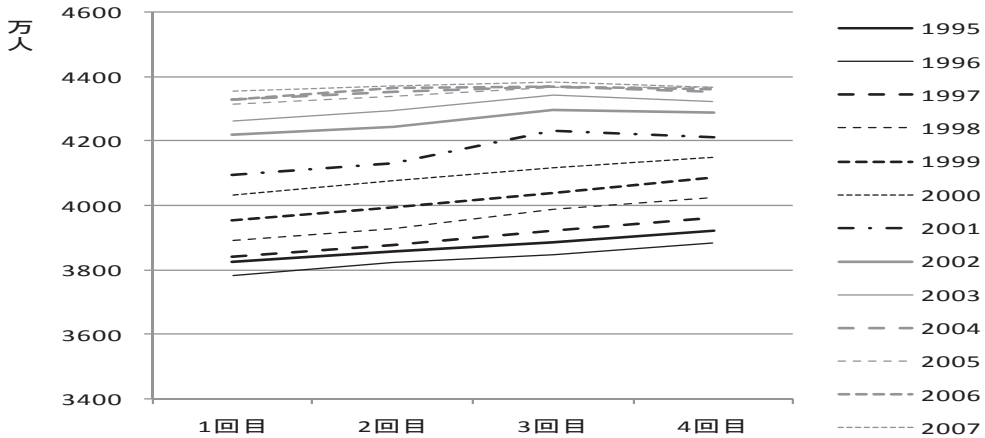


図15 調査回数別非労働力人口（推定値）



5. おわりに一組別標本の比較による考察

組別の標本における原集計数は、それぞれの世帯員数を集計しているのので、比較においては、世帯員数の増減が直接的に表れる。4組の標本の場合は、同じ調査時点の標本であるので、複数の年次で同じ動きをする場合には、世帯・世帯員の回答行動による偏りがあることを示していると考えられる。4組の標本の原集計数でみると、1か月目と2か月目、1年目と2年目の動きに違いがあり、2つの動きが合成されていると考えられる。年次別には、2002年前後で異なる動きをしている。就業者と非労働力人口は2か月目が1か月目

よりも増加し、就業者は2年目が1年目よりも減少し、非労働力人口は年次によって2年目が1年目よりも増加する。1か月目と2か月目、1年目と2年目の動きの要因を推測すると、1か月目と2か月目については、1か月目において調査漏れや調査への非協力などで脱落していた標本が、2か月目には調査に参入したことによって、増加の動きになっていると考えられる。1年目と2年目については、2年目において転入者の把握が十分でないことや、就業状態によって調査への協力度の差が出ているのではないかと考えられる。また、非労働力人口は、組間のばらつきが、

就業者や失業者よりも大きい。これは、標本を選択する際に層別抽出を用いているが、非労働力人口では、その層別効果が低いことを意味していると思われる。

4組の標本の推定値でみると、原集計数ほどに増減の動きはなく、むしろ2002年前後で動きが異なるのは、調査の改正の影響よりも時系列的な就業者、失業者、非労働力人口の動きを反映した結果になっていて、偏りについては、集計用乗率によって補正されていると考えられる。特に、失業者については、原集計数、推定値ともに2002年を境に動きが逆になるのは、世帯・世帯員の回答行動の違いによって生じていると考えられる。これは、①2002年の調査の改正で調査項目は1か月目と2か月目とで同じになっていること、②それにもかかわらず、2002年以後の動きにも特定の傾向があること、③世帯・世帯員の回答行動に違いが生じていることは就業者、非労働力人口で認められるので、失業者でもその違いが出ていると考えられることを踏まえて、④2002年の失業率が5.4%と最も高く(月次で言えば、季節調整値<sup>13)</sup>で2002年6月および8月の5.5%が最も高い)、雇用情勢が悪化する局面と改善する局面によって、世帯・世帯員の回答行動に影響しているとみられるからである。推測できる1つの可能性として、「失業状態」であることを回答することへの抵抗感の強弱が、雇用情勢の悪化局面と改善局面とで違ってくることが考えられる。

調査回数別の標本における原集計数では、調査回数ごとに増減の動きが、4組の標本に比べてはっきりしているものの、4組の標本

とあまり変わらないように思われる。調査回数別の標本は、調査回数を追うごとに一方向に変化していくと予測していたが、実際には、4組の標本と同じく、1か月目と2か月目、1年目と2年目の2つの動きを反映していると考えられる。したがって、調査回数を重ねることによって偏りが生じるのは、2回目と3回目の間に10か月空くこともあって、限定的であると思われる。また、調査回数別の標本の推定値でみると、4組の標本の推定値と動きは似ているとみられる。

今回の標本バイアスの検討を要約すると、労働力調査の標本においては、原集計数による比較から、8組の副標本において、世帯・世帯員の回答行動によって偏りが生じていると考えられる。推定値による比較からは、集計用乗率を用いることによりその偏りも軽減されていることも確認できた。失業者については、2002年を境に動きが異なることが、2002年の調査改正というよりも、雇用情勢の悪化局面と改善局面の違いによって生じていると考えられる。また、調査回数別の標本からは、調査回数が多くなるに従って、回答にある傾向を持つとは必ずしも言えず、あるとしても限定的であると考えられる。

このように本稿の分析で偏りが生じる原因を必ずしも特定できたわけではないが、労働力調査の標本交代方式がもたらす標本構造の解明の一端を示せたと考えている。今後は、分析方法を工夫するとともに詳細な分析を行って、偏りの原因の特定を含めた標本構造の実態の解明を進めたいと思っている。

## 謝辞

本稿の作成に当たって、匿名の査読者から貴重な指摘と有益な助言をいただいた。ここに記して、感謝の意を表したい。



## 注

- 1) 「交代」のほかに「交替」という言い方をする場合もあるが、本稿では「交代」を用いる。また、統計調査の標本交代については、浅井(1987)、土屋(2009)などに解説がある。
- 2) 分析に用いたデータは、2012年に総務省統計調査部国勢統計課労働力人口統計室の小池邦彦氏と共同で統計法第32条による利用を申し出て、承認された労働力調査の個票データである。
- 3) 労働力調査では、「完全失業者(率)」と表しているが、本稿では「失業者(率)」と「完全」を除いて表している。労働力調査で、1950年から「完全」を付けたのは、1949年以前の「失業者(率)」と区別するためであり、現在では「完全」を外しても混同することもなく、特に問題はない。
- 4) 基礎調査票について、従来の調査票の失業者の求職理由は、2か月目の調査票のみであったが、1か月目の調査票にも追加されて、1か月目と2か月目で調査項目は同じになっている。
- 5) 統計法が2007年に全部改正される以前については、個票データの利用は学術研究であっても厳しく、高い公益性が認められる研究等のみに限られ、研究者が利用できる機会は制約されていた。
- 6) 国勢調査の調査区とは、5年ごとに行われる国勢調査の時に、調査員の担当地域を漏れなく重複なく定めるため、およそ50世帯の集まりで日本全国を区分した地域区分である。
- 7) 住戸とは、住宅やその他の建物の各戸で、一つの世帯が居住できるようになっている建物または建物の一区画である。
- 8) 1983年1月から調査区数は約2000から約2900に拡大している。
- 9) 02層～0404層の層化基準を示すと、02層は人口が0の調査区、03層は換算世帯数が15以下の調査区、0401層は学生の寮・寄宿舎のある標本単位区、0402層は病院、療養所のある標本単位区、0403層は社会施設のある標本単位区、0404層は社会施設、大きな病院のある区域で、0401層～0403層に属さない標本単位区である。04層において、単身世帯に属する人員が50人以上からなる場合は、調査区内に単身世帯ごとに単位区を設定している。なお、換算世帯数とは、世帯人員2人以上の一般世帯以外の世帯人員1人の一般世帯数と施設等の世帯人員を1/3にして1世帯と換算した世帯数である。
- 10) 図2～図15は、動きの傾向や特徴を把握しやすいように、目盛を省略していることに注意されたい。
- 11) この8組の推定値の標準誤差は、労働力調査本体の標準誤差の $\sqrt{8}$ 倍になっている。なお、月次の失業者の標準誤差率は約1.5%である。
- 12) 独立した8組の標本を抽出するが、標本交代においては、そのまま1組、2組などに充てられているわけではない。例えば、前年1組であった標本は今年2年目として2組の標本になり、今年1組は前年の2組の標本の交代標本が充てられることになる。
- 13) 季節調整値は、季節指数を算出する対象データに、新たなデータが入り、古いデータが抜けることにより、数値が変わる可能性があることに留意する必要がある。

## 参考文献

- [1] 浅井晃(1987)『調査の技術』日科技連出版社
- [2] 加納悟(2003)「労働力調査とローテーション・サンプリング」『統計数理』第51巻第2号、199-222頁。
- [3] 近藤登雄・山口幸三(1990)「労働力調査における層別効果の測定」『統計局研究彙報』第48号、49-67頁。
- [4] 佐井至道・加納悟(2004)「労働力調査のローテーションサンプリングに関する検討」、『官庁統計におけるサンプリング法の改善と個票データとしての開示に関する諸問題の研究』(科学研究費補助金(基盤研究(B)(1))研究成果報告書)、115-130頁。
- [5] 総務省(総務庁)統計局(1996～2002)『労働力調査年報』(平成7～13年)
- [6] 総務省統計局(2003～2007)『労働力調査年報』(平成14～18年)
- [7] 総務省統計局(2003～2007)『労働力調査年報 詳細結果』(平成14～18年)

- [ 8 ] 総務省統計局 (2008, 2009) 『労働力調査年報 I 基本集計』 (平成19, 20年)
- [ 9 ] 総務省統計局 (2008, 2009) 『労働力調査年報 II 詳細集計』 (平成19, 20年)
- [10] 総務省 (総務庁) 統計局 (1996, 1999, 2005, 2008) 『労働力調査 標本設計の解説』 (各版)
- [11] 土屋隆裕 (2009) 『概説 標本調査法』 朝倉書店
- [12] 芳賀敏郎 (1984) 「労働力調査のデータを用いた解析—交替標本の分析—」 『昭和58年度統計調査におけるリンケージシステム開発に関する調査研究報告書』, 全国統計協会連合会, 138-153頁.
- [13] 土生勉・高橋裕之 (2003) 「労働力調査の改正について」 『統計研究彙報』 第60号, 85-122頁.
- [14] Bailar, B.A. (1975) “The Effect of Rotation Group Bias on Estimates from Panel Survey”, *Journal of the American Statistical Association*, 70(349), pp.23-30.
- [15] Bell, P.A. and M. Carolan (1998) “Using state space models and composite estimation to measure the effects of telephone interviewing on labour force estimates”, Australian Bureau of Statistics, Working Paper, No. 98/2.
- [16] Ghangurde, P.D. (1982) “Rotation Group Bias in the LFS Estimates”, *Proceedings of the Survey Research Methods Section*, American Statistical Association, pp.421-426.
- [17] Lent, J., S. Miller, and P. Cantwell (1996) “Effect of Composite Weights on Some Estimates from the Current Population Survey”, *Proceedings of the Survey Research Methods Section*, American Statistical Association, pp.130-139.
- [18] McLaren, C.H. and D.G. Steel (1997) “The effect of different rotation patterns on the sampling variance of seasonal and trend filters”, *Proceedings of the Survey Research Methods Section*, American Statistical Association, pp.790-795.
- [19] Solon, Gary (1986) “Effects of Rotation Group Bias on Estimation of Unemployment”, *Journal of Business & Economic Statistics*, 4(1), p.105-109.
- [20] Williams, W.H. and C.L. Mallows (1970) “Systematic Biases in Panel Surveys Due to Differential Non-Response”, *Journal of the American Statistical Association*, 65(331), pp.1338-1349.

# Estimation Bias in Statistical Survey applying the Sample Rotation System

Kozo YAMAGUCHI\*

## Summary

The sample rotation system is applied for the monthly and quarterly surveys in public statistics, and was adopted to improve the accuracy of estimation. Under the sample rotation system, a sub-sample group that acts as a replaces may not be homogenous, and this leads to bias over multiple rounds of surveys as the answer action changes.

In this paper, we confirm the presence of such a bias with the use of Labour Force Survey data, and identify its characteristics. We do so by conducting a comparative analysis, using the data set created by combining eight sub-samples. While we do find bias due to changes in the answer action, we confirm that the effect on the estimates is limited.

## Key Words

Sample Rotation System, Estimation Bias, Sub-Sample, Multiplier for Estimation, Labour Force Survey

---

\* Statistical Research and Training Institute

編集委員会からのお知らせ  
機関誌『統計学』の編集・発行について

編集委員会

1. 常時、投稿を受け付けます。
2. 次号以降の発行予定日は、  
第108号：2015年3月31日、第109号：2015年9月30日です。
3. 投稿に際しては、「投稿規程」、「執筆要綱」、「査読要領」などをご熟読願います。
4. 原稿は編集委員長（下記メールアドレス）宛にお送り願います。
5. 原稿はPDF形式のファイルとして提出して下さい。また、紙媒体での提出も旧規程に準拠して受け付けます。紙媒体の送付先は編集委員長宛をお願いいたします。
6. 原則としてすべての投稿原稿が査読の対象となります。
7. 通常、査読から発刊までに要する期間は、査読が順調に進んだ場合でも2ヶ月間程を要します。投稿にあたっては十分に留意して下さい。

編集委員会、投稿応募についての問い合わせは、  
下記メールアドレス宛に連絡下さい。  
また、編集委員長へのメールアドレスも下記になります。

[editorial@jsest.jp](mailto:editorial@jsest.jp)

編集委員長 岡部純一（横浜国立大学）

副委員長 長澤克重（立命館大学）

編集委員

栗原由紀子（弘前大学）

橋本貴彦（立命館大学）

山田 満（関東支部所属）

[注記] 2013年度より編集体制の見直しとして、第一次査読を従来のように支部選出委員が担当するのではなく、編集委員会全体で担当するように方針を変更しています。『統計学』の定期刊行にも力点をおく所存です。常時、投稿を受け付けていますので、できるかぎり早期のご投稿をお願いいたします。108号（2015年3月31日発行予定）への掲載を想定すると、A：「論文」・「研究ノート」の場合、2015年1月初旬、B：その他の場合、2015年1月末を目途に、それまでにご投稿いただく必要があります。

以上

編集後記

ご投稿いただいたすべての執筆者のみなさん、査読に関わってくださった会員のみなさんに心より御礼申し上げます。今回は書評や海外統計事情の執筆依頼にもご快諾いただきました。そうした掲載記事について、会員のみなさんから編集委員会にご提案ご推薦いただければ、紙面活性化にもつながりありがたいです。よろしく願います。

（岡部純一 記）

[訂正] 『統計学』第106号（2014年3月）p.40の「2013年度関西支部例会」5月19日(土)【報告者】  
(1) 桂政昭（誤）について、(1) 桂昭政（正）に訂正します。失礼いたしました。

## 執筆者紹介 (掲載順)

伊藤伸介	(中央大学経済学部)
星野なおみ	((独)統計センター)
山口幸三	(総務省統計研修所)
橋本貴彦	(立命館大学経済学部)
上藤一郎	(静岡大学人文社会科学部)
イリーナ・エリセーエワ	(ロシア統計学会会長)
山口秋義	(九州国際大学経済学部)

## 支部名

## 事務局

北海道	004-0042	札幌市厚別区大谷地西 2-3-1 北星学園大学経済学部 (011-891-2731)	古谷次郎
東北	986-8580	石巻市南境新水戸 1 石巻専修大学経営学部 (0225-22-7711)	深川通寛
関東	192-0393	八王子市東中野 742-1 中央大学経済学部 (042-674-3424)	芳賀寛
関西	525-8577	草津市野路東 1-1-1 立命館大学経営学部 (077-561-4631)	田中力
九州	870-1192	大分市大字且野原 700 大分大学経済学部 (097-554-7706)	西村善博

## 編集委員

岡部純一 (関東) [長]	長澤克重 (関西) [副]
山田満 (関東)	橋本貴彦 (関西)
栗原由紀子 (関東)	

## 統計学 No.107

---

2014年9月30日 発行	発行所	経済統計学会 〒194-0298 東京都町田市相原町4342 法政大学日本統計研究所内 TEL 042(783)2325 FAX 042(783)2332 <a href="http://www.jses.t.jp/">http://www.jses.t.jp/</a>
	発行人	代表者 菊地進
	発売所	音羽リスマチック株式会社 〒112-0013 東京都文京区音羽1-6-9 TEL/FAX 03(3945)3227 E-mail: <a href="mailto:otorisu@jupiter.ocn.ne.jp">otorisu@jupiter.ocn.ne.jp</a> 代表者 遠藤誠

---



# STATISTICS

---

No. 107

2014 September

---

## Articles

- Effectiveness of Data Swapping Based on the Microdata from Population Census  
..... Shinsuke ITO and Naomi HOSHINO (1)
- Estimation Bias in Statistical Survey applying the Sample Rotation System  
..... Kozo YAMAGUCHI (17)

## Book Reviews

- Tadashi YOSHIDA, *On the Progress of Probability Theory and Statistics in the Netherlands*,  
Hassakusha, 2014  
..... Ichiro UWAFUJI (33)
- Hiroshi IZUMI, *A Measurement of Embodied Labor and Basic Economic Indicators*,  
Ohtsuki Syoten, 2014  
..... Takahiko HASHIMOTO (38)

## Foreign Statistical Affairs

- Russian Association of Statisticians  
..... Irina ELISEEVA and Akiyoshi YAMAGUCHI (43)

## Activities of the Society

- The 58<sup>th</sup> Session of the Society of Economic Statistics ..... (46)

---

JAPAN SOCIETY OF ECONOMIC STATISTICS

---