



日本銀行ワーキングペーパーシリーズ

「労働力調査」の標本誤差と非標本誤差 —家族従業者の短期の変動（振れ）や バイアスに関する考察を中心に—

塩谷匡介*

kyousuke.shiotani@boj.or.jp

肥後雅博**

masahiro.higo@boj.or.jp

No.09-J-3
2009年5月

日本銀行
〒103-8660 郵便事業（株）日本橋支店私書箱第30号

* 調査統計局企業調査担当

** 調査統計局物価統計担当

日本銀行ワーキングペーパーシリーズは、日本銀行員および外部研究者の研究成果をとりまとめたもので、内外の研究機関、研究者等の有識者から幅広くコメントを頂戴することを意図しています。ただし、論文の中で示された内容や意見は、日本銀行の公式見解を示すものではありません。

なお、ワーキングペーパーシリーズに対するご意見・ご質問や、掲載ファイルに関するお問い合わせは、執筆者までお寄せ下さい。

商用目的で転載・複製を行う場合は、予め日本銀行情報サービス局までご相談下さい。転載・複製を行う場合は、出所を明記して下さい。

「労働力調査」の標本誤差と非標本誤差[◆]

— 家族従業者の短期の変動（振れ）やバイアスに関する考察を中心に —

塩谷 匡介^{*}・肥後 雅博[†]

2009年5月

【要 旨】

「労働力調査」は雇用動向を把握する上で不可欠の統計であり、統計に含まれる短期の変動（振れ）の性格を理解することによって、その有用性は一層高まると考えられる。各計数のうち、雇用者や非労働力人口については変動の振幅は比較的小さいが、自営業主、家族従業者、完全失業者の変動の振幅は大きい。

標本誤差ならびに非標本誤差が、こうした短期の変動に対して、どの程度の影響を持っているのか、一定の前提の下で試算を行った。まず、標本誤差がもたらす影響は、標本理論で一般に示されているとおり、母数の値が大きい雇用者や非労働力人口では比較的小さいが、母数の値が小さい自営業主、家族従業者、完全失業者では相対的に大きい。標本誤差による短期変動は、雇用者、非労働力人口、自営業主では実データの短期変動のほとんどを、家族従業者と完全失業者ではその6割程度を占め、短期の変動は主として標本誤差から生じていることが分かる。

次に非標本誤差をみると、家族従業者では、回答者が調査票の記入要領を十分に参照しないため、本来雇用者に計上されるべき「有給の自家営業の手伝い」が計数に混在しており、その一部は調査票の記入要領の変更で解消されたが、まだ一定のバイアス（計数の恒常的な偏り）が生じている。もっとも他の計数では、不在によって回答が得られない世帯、調査には回答するが就業状態や従業上の地位を回答しない世帯、各々によるバイアスがみられるが、その値は大きくない。非標本誤差の短期変動に対する影響も、家族従業者を除き、多くの計数で小さな値にとどまる。

Keywords : 労働力調査、標本誤差、非標本誤差、短期の変動（振れ）、バイアス

JEL classifications : C82, J21, J82

[◆] 本稿の作成に際しては、西郷浩教授（早稲田大学）、黒田祥子准教授（東京大学）ならびに西村清彦副総裁、一上響氏、北村富行氏、福永一郎氏、前田栄治氏、門間一夫氏ら日本銀行スタッフから有益なコメントを頂いた。総務省統計局労働力人口統計室からは労働力調査に関する資料等を提供頂いた。記して感謝の意を表したい。ただしあり得べき誤りは筆者らに属する。本稿の内容・意見は筆者個人に属するものであり、日本銀行の公式見解を示すものではない。

^{*} 日本銀行調査統計局企業調査担当（E-mail:kyousuke.shiotani@boj.or.jp）

[†] 日本銀行調査統計局物価統計担当（E-mail:masahiro.higo@boj.or.jp）

1 はじめに

「労働力調査」は、雇用動向を把握するための基礎的な統計調査である。家計に対して調査を行い、月々の就業・失業の状態を把握し、就業者数、雇用者数、完全失業者数、完全失業率など、景気判断や雇用対策等の基礎資料となる結果を提供している。「労働力調査」の結果は、景気判断における雇用面の指標として広く用いられているほか、「国民経済計算」を作成する際にも、雇用者報酬の推計等の基礎統計として利用されるなど、経済統計としての重要性は極めて高い。

一方で「労働力調査」から就業状態の動向を把握する際には、計数の短期の変動（振れ）によって、変化が一時的なものか、趨勢的な変動なのか、識別するのが難しい場合がある。例えば、「労働力調査」の雇用者数の前年同月比と、事業所を調査対象とした「毎月勤労統計調査」の常用雇用指数（5人以上）の前年同月比を比較すると、「労働力調査」の雇用者数は、数ヶ月単位の不規則な短期変動がみられる（図表 1(1)）。「労働力調査」の雇用者の趨勢変動は信頼性が高いと考えられるが、上記のような短期変動の存在により、数ヶ月程度のデータの蓄積では、それが趨勢変動かどうかを判断するのが難しい。ちなみに、「毎月勤労統計調査」の常用雇用指数については、「労働力調査」の雇用者数と比べると、その短期の変動は小さく、動きは滑らかである。しかしながら、「毎月勤労統計調査」の調査対象企業は、「事業所・企業統計調査」の名簿から抽出されて、一定期間固定され、新規に開設された事業所がサンプルに含まれないなどの理由から、趨勢的な変動がバイアスを持つ場合があると考えられる。

「労働力調査」のうち自営業主や家族従業者については、雇用者数と比べ、短期の変動がさらに大きく、前年同月比で基調的な動きを把握するのはより難しくなる（図表 1(2)(3)）。例えば、「労働力調査」の家族従業者数の前年同月比をみる（図表 1(3)）と、減少率は2005年末から2006年にかけて年▲15%程度まで急速に拡大し、その後、ゼロ近傍まで縮小するなど、やや周期が長めの振幅の大きい変動も伴っている。緩やかな景気拡大局面であったこの時期において、このような家族従業者の減少率の一時的な加速とその後の減速が生じた背景には、統計誤差の一時的な拡大と縮小が存在した可能性もあるのではないかと考えられる。

「労働力調査」のような標本調査の計数の変動には、真の変動だけではなく、サンプル抽出に伴う「標本誤差」や標本誤差以外の誤差である「非標本誤差」の変動も含まれている。そうした誤差の特性を理解しておくことは、その統計の動きを正しく評価するうえで重要である。そこで本稿では、標本誤差や非標本誤差が「労働力調査」の主要な計数の変動に与える影響について分析する。

予め、本稿の要旨を述べると以下のとおりである。「労働力調査」の各計数の変動には、周期 12 ヶ月未満の短周期の変動が多く含まれている。そのうち、計数の値が大きい雇用者、非労働力人口については、その変動の振幅（変動係数＝計数の標準偏差／計数の平均）はさほど大きくない。もっとも、自営業主、家族従業者、完全失業者については、各計数の値が小さいこともあって、変動の振幅はより大きくなっている。また、時系列的な変化（1988～1997 年平均→1998～2008 年平均）をみると、多くの計数では変動の振幅はほぼ横ばいであるが、完全失業者は減少傾向に、家族従業者については増加傾向にある。

こうした短期の変動をもたらす要因として、2 種類の標本誤差（調査区抽出に伴う標本誤差、住戸抽出に伴う標本誤差）の影響を分析した（図表 2）。その結果によると、標本誤差はバイアス（計数の恒常的な偏り）には影響しないが、調査対象となる住戸を 2 ヶ月ごとに、調査区を 4 ヶ月ごとに、各々交代させることを主因に、周期数ヶ月程度の短期変動には影響していることが分かった。すなわち、「労働力調査」では、雇用者など特定の就業状態（従業上の地位）の構成比が高い住戸が多く選ばれた場合、次の月も続けて調査されるため、2 ヶ月続けて雇用者数が多くなり、同じように雇用者の多い調査区が選ばれた場合には、4 ヶ月間に亘って雇用者数が多くなる傾向がみられる。さらに、雇用者、自営業主、非労働力人口では、短期の変動のほとんどを標本誤差で説明できること、家族従業者や完全失業者でも短期変動の 6 割を標本誤差で説明できること、が分かった。

次に非標本誤差について、観察可能な 4 タイプ、①調査対象世帯の不在による無回答によるもの、②調査には回答するが、就業状態や従業上の地位については回答しないことによるもの（就業状態ならびに従業上の地位不詳）、③完全失業者と非労働力人口の区分について、完全失業者に該当する「仕事を探していた」者の条件（ILO 基準）が調査対象者に十分に伝わっていなかったことに起因するもの、④本来は雇用者に含まれるべき「有給の自家営業の手伝い」が家族従業者に混在することによって生じるもの、各々の影響について分析を行った（前掲図表 2）。まず、これらの非標本誤差によるバイアスへの影響をみてみた。①調査対象世帯の不在による無回答や②就業状態ならびに従業上の地位不詳がもたらすバイアスは総じて小さいが、④有給の自家営業の手伝いが家族従業者に混在することによる誤差で、家族従業者には無視できないバイアスが生じている。一方で、③完全失業者と非労働力人口の区分について、完全失業者に該当する「仕事を探していた」者の条件が調査対象者に十分に伝わっていなかったことに起因する誤差は、かつては一定のバイアスをもたらしていたが、2002 年に実施された調査票の質問内容の改善によって、現在ではバイアスは解消している。

さらに、これらの非標本誤差による短期の変動への影響をみると、④有給の自家営業の手伝いが家族従業者に混在することによる非標本誤差が、家族従業者の短期の変動に大きく寄与している。これは、主として 2006 年に生じたバイアスのレベルシフトに伴う一時的なものである。つまり、2006 年 1 月に「調査票の記入要領」の記述が変更されたことにより、バイアスが 2006 年以降恒常的に縮小したが、この際に生じたバイアスのレベルシフトに伴い、2006 年のみ変動が一時的に大きくなったことによるものである。2006 年前後の前年比計数を用いる時には注意を要する。また、①調査対象世帯の不在による無回答、②就業状態ならびに従業上の地位不詳によって生じる短期の変動は比較的小さな値にとどまっている。このように、非標本誤差が短期の変動に与える影響は限定的であり、「労働力調査」の短期の変動は主として標本誤差から生じている。

本稿の構成は以下のとおりである。2 節では、「労働力調査」の計数の変動に関するファクトを整理する。3 節では「労働力調査」の計数の変動を、真の変動、標本誤差による変動、非標本誤差による変動、に分けて議論する。4 節では、「国勢調査」の町丁字別集計データを用いて、「労働力調査」の標本誤差による短期の変動（振れ）の大きさを試算する。5 節では、「労働力調査」のフローデータを用いて、非標本誤差のバイアス（計数の恒常的な偏り）を試算する。6 節では、非標本誤差がもたらす短期の変動の大きさを試算し、標本誤差と非標本誤差による短期の変動によって各計数の変動をどの程度説明できるか、を明らかにする。7 節では、本稿の結論をまとめるほか、標本誤差・非標本誤差を削減しうる方策について考察する。

2 ファクトの整理

本節では「労働力調査」の短期の変動に関するファクトを整理する。

(1) 「労働力調査」の変動特性

「労働力調査」の主要計数の動きをみると、短周期の小刻みな変動が目立っている（図表 3）。雇用者や非労働力人口の変動の振幅（変動係数）は比較的小さいが、自営業主、家族従業者、完全失業者の変動の振幅がかなり大きくなっている。各計数の対数前月差（季節調整済み）の標準偏差（1988 年 1 月～2008 年 3 月）をみると、雇用者や非労働力人口の標準偏差は小さいが、自営業主、家族従業者、完全失業者において標準偏差が大きく、これらの計数の変動が大きいことが分かる（図表 4）。

Band-Pass フィルターを用いて各計数の対数前月差を周期 12 ヶ月未満の短周期変動と、周期 12 ヶ月以上の長周期変動に分解し、計数の全変動に対する寄与をみると、いずれの計数についても、短周期変動が占める寄与が大きい（前掲図表 4）。スペクトル分解により、周期別の変動の大きさをみても、短周期変動の寄与が大きく、パワースペクトルのピークは 4 ヶ月程度の短い周期に生じている（図表 5）。以上のようにいずれの手法でも、「労働力調査」の各計数は、短周期の変動の寄与が大きくなっている。

（2）他の統計との比較

次に「労働力調査」の雇用者数（季節調整済み）を、「毎月勤労統計」の常用雇用指数（30 人以上）や米国の「雇用統計」の雇用者数と比べると、「労働力調査」の雇用者数は、他の 2 つの統計と比較して、短期の変動が大きくなっている（図表 6）。対数前月差の標準偏差やスペクトル分解でも同様の傾向がある。すなわち、「労働力調査」の雇用者数の変動のうち、景気変動に対応する周期 12 ヶ月以上の長周期変動の大きさは、他の統計とほぼ同程度である一方で、周期 12 ヶ月未満の短周期変動の寄与は、他の統計と比べて大きい（前掲図表 4、図表 7）。また、「労働力調査」にみられる周期 4 ヶ月でのパワースペクトルのピークは、他の統計では観察されない。このように「労働力調査」は短周期の変動が占める比率が高い。

もっとも、この点については、「毎月勤労統計」は事業所規模 30 人以上について標本事業所を 3 年程度固定していることで、みかけの短周期変動が小さくなっているほか¹、米国の「雇用統計」については、調査結果が滑らかな動きとなるように、標本の前月差を重視する推計方法によって集計していることが、短周期変動の寄与を小さくする要因となっている点は割り引いてみる必要がある。同時に、「毎月勤労統計」には、調査対象企業が一定期間固定され、新規に開設された事業所がサンプルに含まれないことによる一定のバイアス（計数の恒常的な偏り）が存在すること、米国の「雇用統計」にも、前述の推計方法の採用に伴う一定のバイアスが生じる可能性があること、にも留意が必要である²。

¹ このほか、「毎月勤労統計」には、短期の非正規雇用の多くが含まれていないことがみかけの短周期変動が小さくなっている原因とも考えられる。もっとも、「労働力調査」の雇用者から短期雇用者（臨時雇、日雇）を控除した計数である雇用者（常雇）でも、対数前月差の標準偏差は減少しない。この点からは、短期の非正規雇用の影響は小さいと推測される。

² 米国の「雇用統計」では、標本の前月比を重視する推計方法によって推計されており、みかけの短周期の変動が小さくなるメリットがある一方で、次のような問題が存在する。この推計方法では、全サンプルの一部である継続サンプルの前月差の集計ウエイトを大きく与えることで変動を小さくしているが、継続サンプルにバイアスが存在するために、基準時点から時間が経過するにつれて、計数の推計値が全サンプルから推計された値から乖離する、すなわち時間

(3) 時系列での比較

さらに「労働力調査」の短周期の変動が、時系列でどのように変化してきたのかを、前半（1988年1月～1997年12月）と後半（1998年1月～2008年3月）の2つの期間に分けて比較する。季節調整済み計数の対数前月差の標準偏差をみると（図表8）、前半と後半の時期において、多くの計数の標準偏差に変化はないが、家族従業者については後半になって標準偏差が増加し、一方で完全失業者の標準偏差が小幅減少している。パワースペクトルを比較すると（図表9）、多くの計数で周期4ヶ月にピークがあるとの傾向には、いずれの時期でも変化はなく、パワースペクトルの値にも目立った変化はない。もっとも、家族従業者については、後半の時期にパワースペクトルの値が増加している一方、完全失業者ではパワースペクトルの値が減少している。

3 「労働力調査」における標本誤差と非標本誤差の特徴

本節では、「労働力調査」各計数の変動を、就業状態の真の変動と、標本誤差および非標本誤差による変動に分けて考える。

(1) 「労働力調査」の変動をもたらす要因：標本誤差と非標本誤差

標本調査の計数には2つの誤差が存在する。その一つは「標本誤差」である。これは、調査対象サンプルが母集団の一部にとどまることから生じる誤差であり、標本調査では避けることができない。もっとも標本誤差は、短期の変動（振れ）をもたらすが、長期的には計数にバイアス（恒常的な偏り）をもたらさない。

もう一つの誤差は、標本抽出に起因するものを除いた誤差であり、「非標本誤差」という。非標本誤差には、様々な要因が考えられるが、記入ミスや設問に対する理解不足などで調査が正確に行われない場合や、調査対象世帯の不在による無回答から調査結果に偏りが生じる場合などが、代表的な事例である。非標本誤差は、単に短期の変動をもたらすだけでなく、例えば家族従業者と誤って申告する傾向が存在する場合には、継続的に家族従業者が過大推計されるなど、長期的に計数にバイアスをもたらす可能性がある。

$$(\text{公表値}) = (\text{真の値}) + (\text{標本誤差}) + (\text{非標本誤差})$$

の経過に伴いバイアスが増加する。また、集計ウェイトに関するパラメータを外生的に与える必要があるが、この適切な値が時間とともに変化するために、それによるバイアスが増加する可能性がある。

従って、「労働力調査」の公表値に大きな短周期の変動が存在するのは、真の変動、標本誤差、非標本誤差、いずれかの短期の変動が大きいことを意味する。また、計数に恒常的な変動をもたらすのは、真の変動あるいは非標本誤差によって生じるバイアスの変化であり、標本誤差は影響を及ぼさない。

(2) 「労働力調査」の標本設計

次に、「労働力調査」の標本設計の仕組みをみていこう。「労働力調査」は、約4万世帯を対象とする標本調査であり、以下のような「層化2段抽出」で、調査対象世帯を選び出している(図表10)³。

第1段階で、全国98万の国勢調査区(国勢調査の際に設定される調査区)から2,912調査区を選定する。その際には、標本の抽出に偏りが生じにくいようにするため、全調査区を予め幾つかの特性に分類(層化)し、分類(層)ごとに一定数の調査区を選定している(「層化抽出」)。第2段階では、選定された調査区ごとに、調査区に属する約50世帯(住戸)の中から無作為に30世帯(15世帯×2)を選定する(図表11)。選定された15世帯について2ヶ月間継続して調査を行い、その後、選定された残りの15世帯について、次の2ヶ月間調査を行う。このように第1段階で選定された調査区で、4ヶ月間継続して調査が行われる。さらに4ヶ月経過後、第1段階に戻り、層化抽出のルールに従って調査区が新たに選定される。

このように2ヶ月ごとに調査対象世帯の交代、4ヶ月間ごとに調査区の交代、が行われる。このため、各月の調査サンプルには、以下の4つの種類が存在する。

(調査区1ヶ月目)：新規に調査対象となった世帯(世帯調査1ヶ月目)

(調査区2ヶ月目)：先月も調査対象であった世帯(世帯調査2ヶ月目)

(調査区3ヶ月目)：新規に調査対象となった世帯(世帯調査1ヶ月目)

(調査区4ヶ月目)：先月も調査対象であった世帯(世帯調査2ヶ月目)

このうち、世帯調査2ヶ月目のサンプルでは、先月から継続して同一の世帯を調査している一方、世帯調査1ヶ月目のサンプルでは、調査対象世帯が交代している。①長期間、調査対象世帯を固定した場合には、標本誤差の短期の変動は小さくなるが、調査区に新規に転入してきた世帯が調査できず、得られる計数にバイアスが生じる一方で、②毎月全てのサンプルを交代させると、計数のバイアスは解消できるが、サンプル交代に伴う短期の変動が大きくなる。こうした双方のメリット・デメ

³ 標本設計について詳しくは、総務省統計局[2008]を参照のこと。

リットを考慮すると、その中間の方式——「毎月、一部分のサンプルを交代させる方式」——を採用するのが適切であるという考え方で「労働力調査」は設計されている。すなわち、「労働力調査」は、短期変動の振幅がある程度大きくなることを許容しつつ、調査対象が長期間固定されることによるバイアスを回避して、正確な趨勢が把握できるように工夫された調査であるといえる。

なお、同一調査区での調査は4ヶ月で終了するが、調査終了後8ヶ月間の空白を置いて、1年後に同じ時期に再び同一の調査区が調査対象となり、1年前に選定された調査対象世帯で各々2ヶ月間ずつ、調査が行われる。そのため、世帯ベースで見ると、2ヶ月調査して、10ヶ月間の空白ののち、1年後に再度2ヶ月間調査が行われる。このように2年間に亘って同一世帯で調査が行われるのは、「労働力調査」においては、各計数の前年同月比を用いて分析することが多いことから、前年同月比の標本誤差による変動の振幅を小さくするように配慮されているためである。

（3）継続サンプルと非継続サンプルから観察される短期の変動（振れ）の特性

「労働力調査」の標本誤差についてみると、偶然、雇用者の多いサンプル（住戸）が選ばれた場合、次の月も続けて調査されるため、2ヶ月続けて雇用者数が多くなる。同じように雇用者の多い調査区が選ばれた場合には、4ヶ月間に亘って雇用者数が多くなる。従って、標本誤差の変動は数ヶ月の短い周期を持つ。

こうした標本誤差の特性を確認するため、「労働力調査」の各計数の変動を、前月から継続して調査している世帯調査2ヶ月目のサンプル（継続サンプル）の前月（調査1ヶ月目）から今月までの変動と、今月新たに調査対象世帯が変更された世帯調査1ヶ月目のサンプル（非継続サンプル）の先月から今月までの変動（すなわち、先月で調査が終了した調査対象世帯の先月の値から新規に調査が始まった調査対象世帯の今月の値への変化）に分けて考える。継続サンプルの変動には標本誤差に伴う変動を含まないのに対し、非継続サンプルの変動には、サンプル交替による標本誤差の変動が含まれる。その結果、両サンプルの動きの違いから、短期の変動（振れ）に対する標本誤差の寄与の大きさを知ることができる。もし、非継続サンプルの変動が継続サンプルと比較してはっきり大きい場合には、「労働力調査」の短期の変動の主な要因は標本誤差にある可能性が高い。

「労働力調査」の各計数を継続サンプルと非継続サンプルに区分して、その計数に季節調整を行った。その対数前月差の標準偏差をみると（図表12(1)）、非継続サンプルの変動は継続サンプルの変動よりも大きくなっている。公表値の対数前月差の分散に対する寄与を継続サンプル、非継続サンプル、およびその共分散に分解す

ると（図表 12(2)）、非継続サンプルの寄与度が大きいことが確認できる。さらに、スペクトル分解の結果をみると（図表 13）、非継続サンプルにおける短周期変動が継続サンプルの短周期変動と比べて大きい。周期 4 ヶ月におけるパワースペクトルのピークは、継続サンプルにはみられないが、非継続サンプルにおいては顕著である。この結果からは、「労働力調査」の短期の変動の多くは標本誤差から生じている可能性が高いことが分かる。

4 標本誤差によって労働力調査の短期変動ほどの程度説明可能か？

本節では、各計数の標本誤差に伴う短期の変動（振れ）の大きさを試算し、実際のデータの短期の変動と比較する。「労働力調査」には、調査区の抽出に伴う標本誤差、住戸抽出に伴う標本誤差、の 2 つの標本誤差が存在する。以下、一定の前提の下に各々の標準誤差率を計算し、最後に合算して標本誤差に伴う短期の変動の試算値を求める。

（1）調査区の抽出に伴う標本誤差

（標本誤差の算出方法）

①「人口を基準に調査区を母集団から確率比例抽出した場合の標準誤差率」と②「層化の効果（層化抽出による標本誤差削減後の標準誤差残存割合）」を各々求め、それを乗ずることにより、「調査区抽出に伴う標本誤差率」を求める。

（調査区抽出に伴う標本誤差率）＝

（人口を基準に調査区を母集団から確率比例抽出した場合の標準誤差率）

×（層化の効果＜層化抽出による標本誤差削減後の標準誤差残存割合＞）

このように「調査区抽出に伴う標本誤差率」を 2 つに分けると、①「人口を基準に調査区を母集団から確率比例抽出した場合の標本誤差率」からは「母集団における計数のばらつき度合い」を、②「層化の効果（層化抽出による標本誤差削減後の標準誤差残存割合）」からは「層化抽出による標本誤差削減の効果」を、各々把握することができる。

ここでの試算を行う上で必要となる項目に関する国勢調査区毎のデータは非公表であるため、本分析では、最も細分化した公表データである「町丁字別集計デー

タ」を用いる⁴。データの対象範囲は、北海道、福島県、茨城県、東京都、新潟県、愛知県、兵庫県、広島県、愛媛県、福岡県の 10 都道県（人口カバー率約 40%）⁵、分析対象年次は 1995 年と 2005 年である。なお、抽出する調査区の数、1995 年では 2,880、2005 年では 2,912 とする。同一調査区での調査は 4 ヶ月継続することから、毎月、新規に抽出される調査区の数（ n ）は、1995 年は 720（=2,880/4）、2005 年は 728（=2,912/4）となる。

（第 1 段階：人口を基準に調査区を確率比例抽出した場合の標準誤差率）

①「人口を基準に調査区を母集団から確率比例抽出した場合の標準誤差率」を計算する⁶。人口を基準に母集団から 1 つの調査区を確率比例抽出した場合、計数の調査区の人口に占める比率（ X_i ）の分散（ σ^2 ）を

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N w_i (X_i - \mu)^2$$

によって計算する。ここで N は母集団に存在する全調査区の数、 μ は母集団平均

（ $= \frac{1}{N} \sum w_i X_i$ ）、 w_i は各調査区の母集団における人口ウェイト（=調査区の人口/

母集団の人口）である。母集団から調査区を n 個、人口を基準に確率比例抽出した場合の計数の人口に占める比率（ X_i ）の平均値の標準誤差率（ $\sigma_{\bar{x}}$ ）は、人口を基準に母集団から 1 つの調査区を確率比例抽出した場合の標準偏差を抽出する調査区数（ n ）の平方根で除して求められる⁷。

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

1 調査区は約 50 世帯から構成されるため、調査区あたりの平均人口は、1995 年では 140 人（50 戸×2.8 人<世帯平均人員>）、2005 年では 125 人（50 戸×2.5 人<同>）である。町丁字の人口サイズが大きいほど標準誤差率が小さくなるため（図表 14）、推計に用いる町丁字別データの人口サイズを調査区の人口と一致させる必

4 「町丁字別集計データ」とは、市区の町・丁目ごと、町村の字ごとに統計値を集計したデータのことである。例えば、同データでは、東京都中央区日本橋 1 丁目、同 2 丁目、同 3 丁目の区分で就業状態の統計値が公表されている。

5 これらの 10 都道県は、地域の偏りを小さくするため、労働力調査の地域区分ごとに人口の多い 1 都道県を抽出した。なお近畿について、人口の一番多い大阪府は農林業が少ないなど業種に大きな偏りがあるため、人口が 2 番目に多い兵庫県を選んだ。

6 実際の調査区の抽出は系統抽出によって行われている。系統抽出の際に、母集団名簿の順番には何の意味もなく並んでいると考えれば人口を基準に確率比例抽出と同様であると考えられる（松田・伴・美添[2000]・p25）。本稿では、こうした考えに基づき、調査区の抽出は人口を基準に確率比例抽出であるとみなして標準誤差率を計算している。

7 人口を基準に確率比例抽出ならびに層化抽出を行った場合の標準誤差率の計算方法については、例えば松井[2005]を参照。

要がある。本分析では、1995年について15歳以上人口100～200人の町丁字データ、2005年について同100～150人の町丁字データ、を用いる。

（第2段階：層化の効果：層化抽出による標準誤差削減後の標準誤差残存割合）

次に、②「層化の効果（層化抽出による標準誤差削減後の標準誤差残存割合）」を計算する。層化抽出を行うことで、農家の多い調査区を多く選んでしまうといった確率的な「偏り」が生じる可能性が減少することから、層化抽出を行った場合の標準誤差率は、人口を基準に確率比例抽出した場合の標準誤差率よりも小さくなる。「労働力調査」では、調査区における業種別・従業上の地位別の就業者比率を層化基準としたルールを用いて層化抽出を行っている（図表15）。本分析では、町丁字別データを用いて可能な範囲で現実の層化抽出ルールを再現し、「層化の効果（層化抽出による標準誤差削減後の標準誤差残存割合）」を「層化抽出を行った場合の標準誤差率／母集団から人口を基準に確率比例抽出した場合の標準誤差率」から求める。

調査区を母集団から確率比例抽出した場合の標準誤差率（ σ_r ）は①の標準誤差率（ $\sigma_{\bar{x}}$ ）と同様に求める。ただし、①とは異なり10都道府県におけるすべての町丁字別データを用いる⁸。層化抽出を行った場合の分散（ $\sigma_{stratified}^2$ ）は各層内の分散 σ_j^2 の加重平均となる。

$$\sigma_{stratified}^2 = \frac{1}{L} \sum_{j=1}^L W_j \sigma_j^2$$

ここで L は層の数、 W_j は各層の人口ウェイト（＝各層の人口／母集団の人口）である。よって層化抽出を行った場合の各副標本の標準誤差率（ σ_s ）は、

$$\sigma_s = \frac{\sigma_{stratified}}{\sqrt{n}}$$

従って、層化の効果（層化抽出による標準誤差削減後の標準誤差残存割合）は、

$$\text{層化抽出による標準誤差削減後の標準誤差残存割合} = \sigma_s / \sigma_r$$

となる。

⁸ なお、②において10都道府県におけるすべての町丁字別データを用いるのは、以下の理由による。②では、調査区における業種別就業者比率を層化基準とした層化抽出ルールを用いているため、層化抽出を行った場合の標準誤差率をより高い精度で算出するには、多数のデータを利用することで業種別就業者の偏りをできるだけ少なくする必要があるためである。

(調査区の抽出による標本誤差の計算結果)

以上の方法で、①「人口を基準に調査区を母集団から確率比例抽出した場合の標準誤差率」、②「層化の効果(層化抽出による標本誤差削減後の標準誤差残存割合)」、および、①と②の積である「調査区の抽出による標本誤差率」を計算した(図表 16)。

①「人口を基準に調査区を母集団から確率比例抽出した場合の標準誤差率」は、雇用者や非労働力人口では小さい一方で、家族従業者、完全失業者、自営業主では大きく、最大5倍の格差が生じている(図表 16(1))。このように母集団におけるばらつき度合いは、標本理論に示されているとおり、母数の計数の大小で大きく異なる。母数が大きい(計数の人口に占める比率が高い)計数については標準誤差率が小さく、母数が小さい(計数の人口に占める比率の低い)計数については標準誤差率が大きくなっている。なお、1995年から2005年までの変化をみると、家族従業者の標準誤差率が増加している一方、完全失業者の標準誤差率は減少している。

②「層化抽出による標本誤差削減後の標準誤差残存割合」をみると、自営業主が就業者に占める比率を層化基準に採用して層化抽出を行っていることから、自営業主や家族従業者では標準誤差が大きく減少し、標準誤差の残存割合が6割程度と小さくなっており、層化抽出による標本誤差の削減効果が大きい。一方で、完全失業者については、層化基準には採用されていないことから、標準誤差残存割合が9割以上と大きく、層化抽出による標準誤差の削減効果は小さくなっている(図表 16(2))。1995年から2005年までの変化をみると、自営業主および家族従業者では、標準誤差残存割合が幾分増加しており、従来の層化基準による層化抽出の効果が若干だが低下している。

①と②の積である「調査区の抽出による標本誤差率」をみると、完全失業者の標本誤差が最も大きく、次いで、家族従業者、自営業主の標本誤差が大きくなっている(図表 16(3))。一方で、雇用者や非労働力人口の標本誤差はかなり小さい。トータルの標本誤差率では、家族従業者よりも完全失業者の方が大きくなっているのは、家族従業者に関する「層化抽出による標準誤差削減後の標準誤差残存割合」が完全失業者よりも小さく、層化抽出による誤差削減効果が大きくなっているためである。

なお、1995年から2005年までの変化をみると、計数の減少に伴い母集団から確率比例抽出した場合の標準誤差率が拡大したこと、層化抽出の効果が低下していること、の2つの理由から、自営業主や家族従業者の標本誤差率が拡大している。一方で、完全失業者については、人口を基準に確率比例抽出した場合の標準誤差率が縮小したことから、その標本誤差率が低下している。

(2) 住戸の抽出に伴う標本誤差

(標本誤差の計算方法)

次に、住戸（世帯）の抽出に伴う標本誤差を、超幾何分布を用いて試算する。超幾何分布とは、箱の中に玉が M 個入っていて、そのうち K 個が赤い玉、残りが黒い玉の場合において、 m 個の玉を箱へ戻さずに取り出した（非復元抽出を行う）場合に含まれる赤い玉の個数 Y の分布である。その確率関数は以下のとおりである⁹。

$$f_Y(y) = \begin{cases} \frac{\binom{K}{y} \binom{M-K}{m-y}}{\binom{M}{m}} & y = 0, 1, 2, 3, 4, \dots, m \\ 0 & \text{otherwise} \end{cases}$$

「労働力調査」において、調査区内の人々の就業状態が（箱の中の玉と同様に）均一に分布していると考えると、超幾何分布を用いて住戸抽出に伴う標本誤差を計算することができる。 M が調査区内に住んでいる人の数、 K が調査区内で対象とする就業状態（例えば雇用者）である人数、 m が調査区から抽出される人の数とすると、超幾何分布の期待値と分散は

$$E[Y] = m \frac{K}{M}, \quad \text{Var}(Y) = m \frac{K}{M} \cdot \frac{(M-K)}{M} \cdot \frac{(M-m)}{(M-1)}$$

から得られる。超幾何分布に従う試行を抽出する調査区の数、すなわち n 回、繰り返すことになるため、各副標本の標準誤差率（ $\sigma_{dwelling}$ ）は以下で求めることができる。

$$\sigma_{dwelling} = \frac{\sqrt{\text{Var}(Y)}}{\sqrt{nE[Y]}}$$

「労働力調査」では調査区内の約 50 世帯から 15 世帯を 2 回、抽出する。そのため、1 調査区当たりの人口サイズを前提に、1995 年には 140 人中 42 人を 2 回抽出するものとし（ $M=140$ 、 $m=42$ ）、2005 年には 125 人中 37.5 人を 2 回抽出する（ $M=125$ 、 $m=37.5$ ）と考える¹⁰。また、調査区内で特定の就業状態である人数 K は、調査区に

⁹ なお、この確率関数において、 $\binom{n}{r} = {}_n C_r = \frac{n!}{r!(n-r)!}$ は、互いに異なる n 個の要素から、 r 個

を取り出す組み合わせ（combination）が何通りあるかを表す。

¹⁰ 「労働力調査」では、世帯（住戸）単位で抽出が行われるため、人ではなく世帯の抽出で標準誤差率を計算すべきである。しかしながら、町丁字別データからは、世帯ごとの就業状態のデータが得られないため、直接、人を抽出することを前提に計算を行う。

おける当就業状態の生起確率（K/M）がすべて同一であり、公表計数から算出される全国の生起確率（当該就業状態となる人数/15歳以上人口）と一致すると仮定して求める¹¹。

（標本誤差の計算結果）

住戸抽出による標準誤差率を試算してみると（図表 17）、家族従業者や完全失業者など母数が小さい計数では標準誤差率が大きい一方で、母数が大きい雇用者や非労働力人口では小さくなっている。1995年から2005年への変化をみると、自営業主、家族従業者、雇用者では2005年の標準誤差率が増加している。これは1世帯あたり人員の減少に伴い調査対象人口が減ったことや各計数の人口に占める比率が低下したこと、が原因である。その一方で、完全失業者については、人口に占める比率が上昇した効果から、標準誤差率は2005年の方が低くなっている。

（3）標本誤差に伴う短期の変動（振れ）

（標本誤差の算出方法）

以上で求めた調査区、住戸、各々の抽出に伴う標本誤差を用いて、標本誤差がもたらす「労働力調査」各計数の対数前月差の標準偏差を試算する。

t 期の調査の結果（ X_t ）は、調査区1ヶ月目から4ヶ月目までの4つの副標本の結果（ X_{it} ）の平均として計算される。

$$X_t = \frac{1}{4} \sum_{i=1}^4 X_{it}$$

調査結果の前月差は

$$X_t - X_{t-1} = \frac{1}{4} (X_{1t} - X_{4t-1} + X_{2t} - X_{1t-1} + X_{3t} - X_{2t-1} + X_{4t} - X_{3t-1})$$

母集団の真の値および非標本誤差が前月と不変だと仮定すると、今月の世帯調査2ヶ月目の副標本と先月の世帯調査1ヶ月目の副標本との結果は等しくなるため（ $X_{1t-1} = X_{2t}$ および $X_{3t-1} = X_{4t}$ ）、調査結果の前月差は以下のようになる。

¹¹ なお、調査区における当該就業状態の生起確率として層別に異なる値を採用しても、住戸抽出の標準誤差率にはほとんど影響しない。また、本稿では、近所に住んでいる人々の就業状態が相互に無相関であることを前提に住戸抽出の標準誤差率を算出している。実際には、近所に住んでいる人々の就業状態には一定の相関があることを考えると、ここでの試算値は上限値であると考えられる。

$$X_t - X_{t-1} = \frac{1}{4}(X_{1t} - X_{4t-1} + X_{3t} - X_{2t-1})$$

ここで前月の調査区4ヶ月目から今月の調査区1ヶ月目への変化($X_{1t} - X_{4t-1}$)が調査区抽出に伴う変動および1回目の住戸抽出に伴う変動を含むのに対し、前月の調査区2ヶ月目から今月の調査区3ヶ月目への変化($X_{3t} - X_{2t-1}$)は、2回目の住戸抽出に伴う変動のみを含むことから、

$$Var(X_t - X_{t-1}) = \frac{1}{4^2}(2\sigma_{town}^2 + 2\sigma_{dwelling,1}^2 + 2\sigma_{dwelling,2}^2)$$

ただし σ_{town}^2 は各副標本における調査区抽出に伴う標本誤差の分散、 $\sigma_{dwelling,1}^2$ は各副標本における1回目の住戸抽出に伴う標本誤差の分散、 $\sigma_{dwelling,2}^2$ は各副標本における2回目の住戸抽出に伴う標本誤差の分散である¹²。なお、以下の試算結果では、前月差の標準誤差を各計数の値で基準化した「標準誤差率ベース」で議論する。

(標本誤差の計算結果)

本稿での標本誤差の試算結果は、「国勢調査区データ」の代わりに「町丁字別集計データ」を用いるなど、一定の前提に基づくものであり、結果については幅を持つてみる必要がある。この点に留意しつつ、標本誤差率の試算結果をみると(図表18(1)(2))、2005年では、雇用者や非労働力人口では0.4%とかなり小さい一方で、家族従業者(1.8%)、完全失業者(1.6%)、自営業主(1.1%)では大きな値となっている。次に、同時期(2003~07年)における実データの対数前月差の標準偏差と比べると、雇用者、非労働力人口、自営業主では、標本誤差率の試算値が実データにおける標準偏差(雇用者0.3%、非労働力人口0.4%、自営業主1.1%)とほぼ等しくなる。標本誤差による変動が実データの変動のほとんどを説明している。一方で、家族従業者と完全失業者については、標本誤差による寄与は、実データ(家族従業者2.8%、完全失業者2.5%)の変動の6割程度にとどまっている。このように家族従業者と完全失業者については、標本誤差ではその変動を十分には説明できておらず、非標本誤差の寄与が大きい可能性がある。

次に、1995年から2005年までの変化をみる(図表18(3))と、自営業主、家族従業者では標本誤差率が増加している(自営業主+0.1%、家族従業者+0.3%)。これは、母集団での計数のばらつきが増加したこと、層化抽出による標本誤差の削

¹² 1回目に抽出される住戸と、2回目に抽出される住戸は、調査区が選定された時点で、同時に独立して選ばれる。従って、両者の分布は等しくなり、 $\sigma_{dwelling,1} = \sigma_{dwelling,2}$ となる。

減効果が低下したこと、住戸抽出における標準誤差率が増加したこと、各々が寄与している。一方で、完全失業者の標本誤差率は低下している（▲0.2%）。これには、母集団における計数のばらつきの低下が寄与している。なお、雇用者、非労働力人口についてはこの10年間でほとんど変化はない。

実データにおける対数前月差の標準偏差の変化（1993～97年→2003～07年）をみると、家族従業者の変動が増加（+1.1%）しているのが目立つ。しかし、この変動の増加分のうち、標本誤差の増大によって説明できる部分は4分の1程度にすぎない。家族従業者の非標本誤差がこの10年間で増加していることが示唆される。

5 「労働力調査」における非標本誤差がもたらすバイアス

（1）フローデータを用いた非標本誤差の把握

本節では、非標本誤差が「労働力調査」の計数にどのようなバイアス（計数の恒常的な偏り）を与えているかについてみてみる。「非標本誤差」は標本誤差以外の誤差の総称であり、その要因は幅広い。記入ミスなどで調査が正確に行われない場合や、調査対象者の不在による無回答から計数にバイアスが生じるなどが代表的な事例である¹³。

標本誤差はサンプル数と直接的な関係があり、量的把握が可能なのに対して、非標本誤差は原因の特定も容易ではなく、その量的な把握も困難である。ところが「労働力調査」では、「フローデータ」という形で、前月（世帯調査1ヶ月目）と今月（同2ヶ月目）、2ヶ月間継続調査されるサンプルの調査結果を、「前月の就業状態と今月の就業状態とのマトリックス」形式で利用可能である（図表19）¹⁴。このフローデータから、非標本誤差に関する以下の点を明らかにできる。

① 調査対象世帯の不在による無回答サンプルは計数に歪みをもたらすか？

不在によって回答が得られなかったサンプルにおける実際の就業状態・従業上の地位の構成比が、調査できたサンプルの就業状態・従業上の地位の構成比と異なる場合には、就業状態の公表計数にバイアスが生じる。このバイアスは通常は測定不可能であるが、前月（世帯調査1ヶ月目）と今月（同2ヶ月目）、いずれか1ヶ月

¹³ 非標本誤差の問題については、松田・伴・美添[2000]のP55～77に詳しい解説がある。

¹⁴ 具体的には、「労働力調査」の基本集計（全国）の第7表「今月及び前月の就業状態、農林業・非農林業、従業上の地位別15歳以上人口」に掲載されたデータである。フローデータについて、詳しくは黒田[2002]および桜[2006]を参照。

が無回答であり、もう1ヶ月については調査できた場合には、2ヶ月間とも調査できたサンプルとの間で就業状態にバイアスが存在するか否かを検証することが可能である。本稿では、(2) 調査対象世帯の不在による無回答がもたらす非標本誤差を取り上げるほか、類似例として(3)「労働力調査」自体には回答するが、就業状態や従業上の地位については無回答(就業状態ならびに従業上の地位不詳)のサンプルがもたらす非標本誤差、についても分析対象とする。

② 調査票・記入要領の形式が変化すると、調査対象者の回答に変化が生じるか？

調査対象者が就労していない場合、当該回答者の就業状態が完全失業者となるか、それとも非労働力人口となるかは曖昧である。というのは、完全失業者と非労働力人口との区分については、完全失業者がILOの基準を満たす者となるように、その条件が「調査票の記入要領」に記載されているが、調査対象者(回答者)がそれを十分に参照しないままに回答するケースがあるため、非労働力人口に区分されるべき者が完全失業者に区分されるケースが存在するためである。その場合、調査票における求職活動に関する質問内容が変化すると、その変化が回答者に影響を与え、就業状態に関する回答が変動する可能性がある。

また、「労働力調査」では、家族従業者は「自営業主の家族で、その自営業主の営む事業に無給で従事している者」と定義されている。このため、家族による自家営業の手伝いのうち、「無給の自家営業の手伝い」を家族従業者として計上し、「有給の自家営業の手伝い」については雇用者として計上するルールとなっている。しかし、このルールが「労働力調査」の調査対象者には十分に浸透しておらず、「有給の自家営業の手伝い」を家族従業者として誤って申告している可能性が少なくない。その場合、調査票の記入要領の説明内容を変更すると、これまで「有給の自家営業の手伝い」を家族従業者として申告していた人の一部が申告の誤りに気づき、雇用者へ申告を変更するなど、回答結果に変化をもたらす。

以上のような調査票の質問内容や調査票の記入要領の変化が、完全失業者(非労働力人口)や家族従業者(雇用者)の非標本誤差に与える影響について取り上げ、一定の前提の下で非標本誤差がもたらすバイアスの大きさを試算する。

(2) 調査対象世帯の不在による無回答がもたらす非標本誤差

調査対象世帯の不在による無回答によって生じる非標本誤差について検討する。

(無回答によるバイアスが生じるメカニズム)

フローデータから、世帯調査1ヶ月目と2ヶ月目、どちらかの月で無回答となっ

たサンプルの数や比率を知ることができる。フローデータにおける「前月いて今月いなかった者」は「前月（同1ヶ月目）は調査できたが今月（同2ヶ月目）は無回答となったサンプル」に相当し、「今月いて前月いなかった者」は「前月は無回答であったが、今月調査することができたサンプル」に相当する（前掲図表19）。

「労働力調査」の無回答は近年増加傾向にある。そのうち、「前月は調査できたが、今月は無回答となった」サンプル数はほぼ横ばいなのに対し、「前月は無回答であったが、今月調査することができた」サンプル数が一貫して増加しており、この15年間でほぼ2倍となっている（図表20）。前者の「前月は調査できたが、今月は無回答となった」サンプルは、調査対象世帯の転居に伴い、2ヶ月目に無回答となった世帯が大半を占めると推測される。これに対し、後者の「前月は無回答であったが、今月調査することができた」サンプルは、調査対象世帯の不在から1ヶ月目には調査できなかつたが、その後1ヶ月間、統計調査員が繰り返し訪問するなどの努力により、2ヶ月目によりやく調査することができた世帯が多く含まれていると考えられる。単身世帯や共働き世帯の増加によって、後者の調査対象世帯の不在による無回答の事例は増加していると考えられる¹⁵。

無回答サンプルの就業状態からその点が推測できる。前月の調査サンプル（世帯調査1ヶ月目）を、①前月も今月も調査できたサンプル、②前月は調査できたが今月は無回答となったサンプル、の2つのグループに、今月の調査サンプル（同2ヶ月目）を、①前月も今月も調査できたサンプル、③先月は無回答となったが今月は調査できたサンプル、の2つのグループに区分して就業状態を比較する。

②前月は調査ができたが今月は無回答となったサンプルは、①前月も今月も調査ができたサンプルと比べ、雇用者や自営業主、家族従業者が少なく、非労働力人口が多い（図表21(1)）。これは、転居をする世帯の特徴を反映している。すなわち、自営業を営んでいる世帯が少なく、また転居前には世帯員のうち主婦がパート労働を辞めるケースが多いのではないかと推測される。一方、③先月は無回答となったが今月は調査できたサンプルは、①前月も今月も調査できたサンプルと比べて、雇用者が多く、自営業主、家族従業者が少ない（図表21(2)）。これは、対象世帯の不在による無回答の特徴を反映している。すなわち、雇用者のいる世帯では昼間不在となる場合が多く、無回答が多くなりやすい一方で、自営業主・家族従業者の多くは在宅であるため、不在による無回答が生じにくいためである。

¹⁵ ちなみに、1ヶ月目、2ヶ月目双方とも調査できなかつた世帯については、フローデータのどこの欄にも計上されないことから、その数は不明である。本分析では、両月とも無回答となったサンプルによるバイアスは考慮していない。

（無回答によるバイアスの試算方法）

①～③のサンプルで就業状態が異なるため、前月の調査サンプル（世帯調査1ヶ月目）と今月の調査サンプル（同2ヶ月目）の間で計数に乖離が生じる。前月と比べ、今月の調査サンプルにおける無回答の比率が小さいため、今月のサンプルによる結果の精度は前月よりも高い。従って、前月のサンプルには、少なくとも、前月と今月のサンプルとの乖離分のバイアスが存在すると考えられる。その結果、世帯調査1ヶ月目と2ヶ月目双方のデータから算出される公表計数には、少なくとも、先月のサンプルがもつ2分の1のバイアスが含まれている。ここでは、以上の前提に基づいてバイアスを算出する。

（無回答によるバイアスの試算結果）

バイアスの試算結果によると（図表22）、2007年の公表計数においては、雇用者は▲0.1%程度、完全失業者は▲0.2%程度、真の値よりも過小となっている。一方で、自営業主は+0.1%程度、家族従業者は+0.4%程度、非労働力人口は+0.1%程度、過大である。これは「労働力調査」では雇用者や完全失業者の無回答の比率が高く、調査されたサンプルに占める自営業主、家族従業者の比率が実勢と比べて過大となるためである。もっとも、そのバイアスの大きさは比較的小さい。

（3）就業状態ならびに従業上の地位不詳がもたらす非標本誤差

（就業状態ならびに従業上の地位不詳によるバイアスが生じるメカニズム）

無回答と同じ状況にあるのが「就業状態ならびに従業上の地位不詳」によるバイアスである。「労働力調査」自体の調査には応じたが、就業状態や従業上の地位について回答しなかったサンプルは、プライバシー意識の高まりもあって、近年緩やかに増加する傾向にある（図表23）。世帯調査1ヶ月目、2ヶ月目どちらかの月において、「就業状態」や「従業上の地位」の回答をしなかった者のうち、他の月で完全失業者と回答した者の比率は全体と比べて目立って高い。これは、調査対象者が完全失業者である場合には、就業状態や従業上の地位に関する回答をしないケースが多いことを示している。この乖離がバイアスをもたらす（図表24）。

（就業状態ならびに従業上の地位不詳によるバイアスの試算方法）

ここでは、就業状態や従業上の地位について回答しなかった者のうち、どちらかの月で就業状態や従業上の地位を回答した者については、回答した月の就業状態・従業上の地位と回答しなかった月の就業状態・従業上の地位が一致すると考える。その方法で判明した分の就業状態・従業上の地位を各公表計数に加算した値を真の

値とみなし、バイアスを計算する。すなわち、

(世帯調査 1 ヶ月目の真の公表値) = (世帯調査 1 ヶ月目の結果) + (世帯調査 1 ヶ月目の「就業状態・従業上の地位不詳」のうち、調査 2 ヶ月目で判明した分)

(世帯調査 2 ヶ月目の真の公表値) = (世帯調査 2 ヶ月目の結果) + (世帯調査 1 ヶ月目では就業状態・従業上の地位が判明していたが、2 ヶ月目で「就業状態・従業上の地位不詳」となった分)

から求める。真の公表計数は、世帯調査 1 ヶ月目、同 2 ヶ月目の平均で求める。

(就業状態ならびに従業上の地位不詳によるバイアスの試算結果)

以上の方法でバイアスを試算すると、2002 年以前は、就業状態ならびに従業上の地位不詳により、完全失業者では▲1.0%程度、真の値よりも過小となっていたと推測される。しかし、2002 年以降は、1 ヶ月目に就業状態について回答せず、2 ヶ月目には完全失業者と回答したサンプルが減ったため、バイアスは縮小しており、2007 年の完全失業者では▲0.6%程度となっている。一方、雇用者など他の計数については▲0.2%程度、真の値よりも過小となっている(図表 25)。

(4) 調査票の質問の変化に伴う非標本誤差の変動：完全失業者と非労働力人口

(完全失業者と非労働力人口との曖昧な境界：調査票における質問内容の影響)

「労働力調査」において、調査対象者が就労していない場合、当該回答者の就業状態が完全失業者となるか、それとも非労働力人口となるかは曖昧である。というのは、完全失業者と非労働力人口との区分については、完全失業者が ILO の基準を満たす者となるように、その条件が「調査票の記入要領」に記載されているが、調査対象者(回答者)がそれを十分に参照しないままに回答するケースがあるため、非労働力人口に区分されるべき者が完全失業者に区分されるケースが存在するからである。その場合、調査票における求職活動に関する質問内容が変化すると、その変化が回答者に影響を与え、就業状態に関する回答が変動する可能性がある。

基礎調査票において、就業状態として「完全失業者」を選択した場合には、「仕事を探し始めた理由」についての質問を答える必要がある。「完全失業者」を選択した回答者は、「仕事を探し始めた理由」の設問をみたところで、完全失業者となるには求職活動をしていることが条件となることに気づき、回答者のうち一部が一旦回答した就業状態を「完全失業者」から「非労働力人口」へ変更すると推測され

る。このように「仕事を探し始めた理由」についての追加質問は、完全失業者と非労働力人口との区分を正確なものにする役割を果たしている。

ところが、2001年までは、世帯調査2ヶ月目の調査のみにおいて「仕事を探し始めた理由」の質問が行われており、世帯調査1ヶ月目の調査にはこの質問が存在しなかった。1ヶ月目においては、追加質問によって完全失業者と非労働力人口との区分をより正確なものにすることができていなかった。このため、「仕事を探していない者」の一部が完全失業者として申告され、完全失業者を過大に、非労働力人口を過少に推計するバイアスが生じていたと考えられる。

フローデータから、2001年までの期間における世帯調査1ヶ月の完全失業者・非労働力人口と同2ヶ月目の完全失業者・非労働力人口の値とを比較すると、1ヶ月目の完全失業者は2ヶ月目の完全失業者よりも+3~4%程度大きく、逆に非労働力人口は▲1%程度小さくなっていた（図表26(1)）。しかし、世帯調査1ヶ月目にも「仕事を探し始めた理由」の質問が実施されるようになった2002年以降では、完全失業者ならびに非労働力人口における1ヶ月目と2ヶ月目との計数の乖離が解消した。これは、2002年に実施された質問の追加によって、世帯調査1ヶ月目における完全失業者の過大推計、非労働力人口の過少推計が是正されたことを意味しており、完全失業者と非労働力人口の精度が向上したと考えられる。

（調査票の質問の変化に伴うバイアスの変動の試算方法）

世帯調査1ヶ月目に「仕事を探し始めた理由」についての質問がなかったことによって生じていたバイアス（2001年まで）をフローデータから推計する。これは、調査1ヶ月目の完全失業者・非労働力人口の値と2ヶ月目の値との乖離には、無回答や従業上の地位不詳など、他の非標本誤差の影響を含むため、こうした影響を受けない「完全失業者から非労働力人口へのフロー」を用いるのが適当であるためである。2001年までのバイアスは以下のように算出できる。世帯調査1ヶ月目に完全失業者の過大な申告が行われていたが、調査2ヶ月目には「仕事を探し始めた理由」に関する質問の影響から、一部の完全失業者が非労働力人口へ回答を変更し、完全失業者から非労働力人口へのフローが増加していたと考えられる。一方で、非労働力人口から完全失業者へのフローには、こうした影響が含まれないことから、2つのフローの差分が世帯調査1ヶ月目におけるバイアスとなる。2001年以前の完全失業者、非労働力人口の公表計数のバイアスは、この両フローの差分の2分の1として求めることができる（図表26(2)）。

なお、2002年以降については、「仕事を探し始めた理由」の質問によって非標本誤差が解消し、完全失業者と非労働力人口のバイアスがゼロとなったと考える。

(調査票の質問の変化に伴うバイアスの試算結果)

以上の仮定を前提に 2001 年までのバイアスを求めると、完全失業者が+2%程度過大である一方、非労働力人口が▲0.1%程度過小となっていたと推測される(図表 27)。このように完全失業者に上方バイアスが生じていた。しかし、2002 年には世帯調査 1 ヶ月目にも「仕事を探し始めた理由」の質問が実施されたことから、2002 年以降については、完全失業者のバイアスは解消して、ゼロとなった。

(5) 調査票の記入要領の変更に伴う非標本誤差：有給の自家営業の手伝いの扱い

(家族従業者に関する定義と非標本誤差)

「労働力調査」では、家族従業者については、「自営業主の家族で、その自営業主の営む事業を無給で手伝っている人をいい、給料・賃金をもらっている場合は、家族であっても、雇われている人(雇用者)に含める」と定義している。しかし、「労働力調査」の調査対象者に「家族従業者は無給の従業者のみに限定されている」という正確な定義を理解してもらうことが必ずしも容易ではないため、従業上の地位に関する回答が正しいものとはなっていない可能性がある。特に、日本の場合、自営業主の配偶者に給料・賃金を支払っている場合でも、事業専従者控除の税制メリットを享受することを目的に、年に 1 度の税務申告時に賃金を纏めて計上するケースが多いだけに、調査対象世帯サイドでも、有給の自家営業の手伝いと無給の自家営業の手伝いとの違いが十分には認識されていないと考えられる。こうした状況の下で、家族従業者に関する定義を厳格に適用しようとしても、その回答結果は曖昧になりやすく、非標本誤差が生じてしまう余地がある。

(2006~07 年の家族従業者数の 2 年目・2 ヶ月目調査における大幅減少)

図表 1 でもみたように、家族従業者数は 2006 年から 2007 年にかけて大きく減少している。この減少の原因を知るために、世帯調査の何ヶ月目のサンプルでこうした減少が生じていたのかをみしてみる。既に述べたように「労働力調査」では、同一世帯に対して、1 年目に 2 ヶ月間、2 年目に 2 ヶ月間、合計 4 ヶ月間調査を行う。その際に利用される調査票が「基礎調査票」であり、その集計結果が「基本集計」として公表される。さらに同一世帯への調査の最後の 4 ヶ月目(2 年目の 2 ヶ月目)に、「特定調査票」による追加調査を行っている。この集計結果が「詳細結果」として公表される。

まず、「基本集計」から、世帯調査 1 ヶ月目と 2 ヶ月目との家族従業者数を比較すると、2 ヶ月目の値が 1 ヶ月目の値と比べて小さくなっている。特に、2006 年以

降に両者の乖離が急に拡大しており、2007年にはその乖離は▲9%にも達している（図表28(1)）。フローデータをみると、「雇用者から家族従業者へのフロー」がほぼ横ばいである一方で、「家族従業者から雇用者へのフロー」が2006年以降大幅に増加している。これは、1ヶ月目に家族従業者と答えていた回答者のなかで、2ヶ月目に雇用者へ回答を変更したケースが、2006年以降に大幅に増加したことを意味している（図表28(2)）。

次に、調査2年目・2ヶ月目のサンプルから集計される「詳細結果」を用い、各調査月間の乖離を詳しくみると、2年目・2ヶ月目のサンプルでは、他の調査月と比較して家族従業者数が顕著に少なくなっている。例えば、2007年では、2年目・2ヶ月目のサンプルは、他のサンプル（1年目・1ヶ月目、同2ヶ月目、2年目・1ヶ月目）と比べて、家族従業者数が▲20%（▲30万人）も少なくなっている（図表29(1)）¹⁶。一方、同じ2ヶ月目でも1年目・2ヶ月目のサンプルでは、1年目および2年目の1ヶ月目とほとんど同一の値となっている。この事実からは、過去の3回の調査では、家族従業者と回答していた回答者のかなりの部分が、2年目・2ヶ月目になって雇用者へ回答を変更したことが分かる。また、こうした変化は2006年から顕著となっており、2006年から2007年にかけての家族従業者数の大幅減少は、2年目・2ヶ月目における回答の変更によって多くが生じていたことが分かる。

（調査票の記入要領の変更による影響：非標本誤差の縮小）

この時期、「基礎調査票」「特定調査票」いずれについても、その様式や質問内容には変化はなかったが、「基礎調査票の記入のしかた（記入要領）」については、2006年1月より、以下のような変更が行われている。

<基礎調査票の記入のしかた：「従業上の地位」>

（旧）「自家営業の手伝い」とは、自営業主の家族で、その自営業主の営む事業を手伝っている人をいいます。

↓

（新）「自家営業の手伝い」とは、自営業主の家族で、その自営業主の営む事業を無給で手伝っている人をいいます。給料・賃金をもらっている場合は、家族であっても、「雇われている人」とします。

以上のように、新しい記入要領では、有給の自家営業の手伝いは家族従業者には含まれず、雇用者として計上されることを明示しており、従来、家族従業者として申告されていた可能性が高い有給の自家営業の手伝いが、2006年以降、雇用者へ

¹⁶ 公表データの制約から、ここでの家族従業者は「非農林家族従業者」である。

回答がシフトしたのではないかと推測される。

しかしながら、全ての調査月で調査が行われる「基礎調査票」の記入要領が変更されたにもかかわらず、この家族従業者から雇用者へのシフトは「特定調査票」の調査月である2年目・2ヶ月目でのみ生じており、他の3ヶ月（1年目・1ヶ月目、同・2ヶ月目、2年目・1ヶ月目）では生じていないことが不可解である。ここからは推測の域を出ないが、以下の考察が一つの答えとなりうると考えられる。

「特定調査票」では、「今の仕事について、雇われている人は勤め先での呼称を記入してください」との質問（A4）がある。しかしながら、この質問では、雇用者の呼称として「正規の職員・従業員」「パート」「アルバイト」「派遣労働者」「契約社員・嘱託」の各項目が、さらに自営業主として「会社などの役員・自営業主など」の項目が用意されているが、「自家営業の手伝い（家族従業者）」に直接該当する項目は設定されておらず、家族従業者がどの項目を選択してよいか、迷うことになりやすいと推測される¹⁷。その際に、記入対象者は、「特定調査票の記入のしかた」のほか「基礎調査票の記入のしかた」を読み返すことで、初めて有給の自家営業の手伝い（家族従業者）の正しい扱いに気づき、有給の自家営業の手伝いを雇用者として申告するようになったと考えられる。

すなわち、「特定調査票」が存在しない1年目・1ヶ月目、同・2ヶ月目、2年目・1ヶ月目では、調査対象者の相当数が、「基礎調査票の記入のしかた」を読まずに有給の自家営業の手伝いを家族従業者として申告していた。しかし、そうした調査対象者も、2年目・2ヶ月目については「特定調査票」の質問（A4）が迷いやすいものであるがゆえに、初めて「基礎調査票の記入のしかた」を読んだ結果、有給の自家営業の手伝いの正しい扱いに気づき、有給の自家営業の手伝いを雇用者として申告するようになったのではないかと推測される。

このように「基礎調査票」「特定調査票」の記入要領の改訂により、従来、家族従業者扱いされていた有給の自家営業の手伝いの一部が雇用者に変更され、非標本誤差が2006年以降、縮小している。しかし、現在のところ、この変化は「特定調査票」による調査が行われる2年目・2ヶ月目に限定されている。残された他の調査月については、相当数の有給の自家営業の手伝いが家族従業者に含まれたままになっており、家族従業者数はその分過大な値となっていると推測される。

（家族従業者数の減少の加速／家族従業者から雇用者へのフローの増大）

2006年入り後、家族従業者から「有給の自家営業の手伝い」の一部が除外されて、

¹⁷ 本来は、自営業主と同様に「会社などの役員・自営業主など」の項目を選択するのが正しい。

家族従業者数が大幅に減少したことから、家族従業者の近年の趨勢的な減少傾向が加速している（図表 29(2)）。また、フローデータでは、2年目・2ヶ月目において家族従業者から雇用者へ従業上の地位を変更する人が増加することから、「家族従業者から雇用者へのフロー」が大幅に増加するとの変化が生じている。家族従業者から雇用者へ大きなフローが生じた点を捉えると、「自営業の衰退→家族従業者の労働市場への進出→労働供給増を受けた雇用者の増加」との労働市場の構造変化が2006年以降、一時的に加速したとの解釈も可能である。

しかし、実際には、家族従業者から雇用者へのフローには、「基礎調査票の記入要領」の改訂の影響を受けて生じた回答の変化分が含まれており、それを控除した実勢ベースでみると、家族従業者から雇用者へのフローは逆に減少している可能性が高く、そうした解釈は妥当ではない（図表 29(3)）。このように、基礎調査票の記入要領の改訂を受けた回答の変更による影響は小さくない。

（「有給の自家営業の手伝い」が家族従業者に混在することによるバイアスの試算方法）

2年目・2ヶ月目の家族従業者数が正しいと仮定して、他の調査月の回答の誤りによる家族従業者ならびに雇用者のバイアスを試算する。記入要領の改訂による効果が一巡したと考えられる2007年の2年目・2ヶ月目の家族従業者は、他の調査月と比べ30.5万人少なくなっていることから、2006年以降の他の調査月については、家族従業者が30.5万人（+14%）過大であると考えられる。また2005年以前については、回答を誤った家族従業者が全家族従業者に占める比率は、2006年以降と同一であると仮定し、全調査月で家族従業者数が+14%過大となっていると考える。

（「有給の自家営業の手伝い」が家族従業者に混在することによるバイアスの試算結果）

上記の前提に基づいて、有給の自家営業の手伝いが家族従業者に含まれることによるバイアスを試算する。2005年以前の家族従業者数は約+14%過大となっており、一方で、雇用者数は▲0.7～1.1%程度過小となっていたことが分かる。さらに2006年以降については、基礎調査票の記入要領の改訂によって、バイアスは幾分縮小しており、家族従業者数は約+11%過大に、雇用者数は▲0.5%程度過小となっている（図表 30）。このように有給の自家営業の手伝いが家族従業者に含まれることによって、家族従業者数や雇用者数にバイアスが生じている。また、2006年1月以降、記入要領が改訂されたことによって、家族従業者数のバイアスが縮小する過程では、よく用いられる前年比計数に逆方向のバイアスが発生したと考えられる。すなわち、2006～07年には、家族従業者数の減少率と雇用者数の増加率が、ともに過大な数値

となっている可能性が高いことから、この時期の労働市場に関する分析においては、注意する必要がある。

(6) 非標本誤差がもたらす公表計数への影響

本稿で試算した非標本誤差によるバイアスは、無回答サンプルや「就業状態ならびに従業上の地位不詳」サンプルにおける真の就業状態に関して一定の仮定を設けるなど、様々な前提に基づいて試算されたものであり、結果は幅を持つてみる必要がある。この点に留意しつつ、(2)～(5)で求めた4つの「非標本誤差によるバイアス」の合計値をみてみよう(図表31)。非標本誤差によるバイアスが最も大きいのは家族従業者であり、2005年までよりは縮小したとはいえ、2007年でも有給の家族従業者の過大計上を主因に、計数が+11%過大であると試算される。しかしながら、家族従業者以外の計数については、非標本誤差によるバイアスは小さな値に止まっている。雇用者は、本来雇用者に計上されるべき有給の家族従業者が正しく計上されていないことを主因に▲0.7%程度の過小な値となっているが、そのバイアスは近年縮小傾向にある。完全失業者については、2001年以前には+1%程度の過大推計となっていたが、2002年に実施された「仕事を探し始めた理由」の質問を世帯調査1ヶ月目にも追加したことによってバイアスが大幅に縮小し、現在では「就業状態ならびに従業上の地位不詳」の増加を主因に▲0.6%程度の過小推計¹⁸となっている。自営業主や非労働力人口については、各種のバイアス要因が相殺しており、非標本誤差によるバイアスは、ほとんど無視できる水準である。

さらに、この「非標本誤差によるバイアス」が各公表計数の前年比に与える影響をみてみる(図表32)。家族従業者のバイアスが縮小した2006年においては、家族従業者の前年比が実勢と比べて▲3.5%過小な推計となるなど大きな乖離が生じている。同様に、2002年においても、完全失業者のバイアスが小さくなったことを主因に、完全失業者の前年比が▲1.3%過小な推計となっている。しかしながら、この2年を除くと、非標本誤差によるバイアスが公表計数の前年比に与える影響は、いずれの計数についても小幅にとどまっている。

¹⁸ 直近の完全失業者数(300万人)および労働力人口(6,500万人)のレベルでみると、完全失業者数の▲0.6%程度の過小推計は、計数にすると約2万人に相当し、完全失業率に対しては▲0.03%程度の過小推計となる。

6 非標本誤差がもたらす「労働力調査」の短期の変動（振れ）

（1）非標本誤差がもたらす短期の変動（振れ）：標準誤差率の試算方法

5節では、非標本誤差がもたらす「労働力調査」の各計数のバイアス（計数の恒常的な偏り）について検討してきたが、本節では、非標本誤差がもたらす「労働力調査」の短期の変動への影響をみてみよう。公表計数（ X_t ）は、標本誤差（ s_t ）と各要因による非標本誤差（調査対象世帯の不在による無回答によるもの： n_{1t} 、就業状態・従業上の地位不詳によるもの： n_{2t} 、完全失業者と非労働力人口との区分の曖昧さによるもの： n_{3t} 、有給の自家営業の手伝いが家族従業者に混在することによるもの： n_{4t} ）によるバイアス分、真の値（ X^* ）から乖離している。

$$X_t = X^* + s_t + n_{1t} + n_{2t} + n_{3t} + n_{4t}$$

真の値に変化がないと仮定すると、公表計数の前月差は

$$X_t - X_{t-1} = s_t - s_{t-1} + n_{1t} + n_{2t} + n_{3t} + n_{4t} - n_{1t-1} - n_{2t-1} - n_{3t-1} - n_{4t-1}$$

となる。本稿で取り上げた各非標本誤差、標本誤差はそれぞれ独立であると考えられるため、公表計数の短期の変動（前月差）の分散は、以下のように書ける。

$$\begin{aligned} \text{Var}(X_t - X_{t-1}) &= \text{Var}(s_t - s_{t-1}) \\ &+ \text{Var}(n_{1t} - n_{1t-1}) + \text{Var}(n_{2t} - n_{2t-1}) + \text{Var}(n_{3t} - n_{3t-1}) + \text{Var}(n_{4t} - n_{4t-1}) \end{aligned}$$

すなわち、公表計数の短期の変動は、標本誤差の変動と各要因によって生じる非標本誤差の変動の合計となる。以下では、まず4種類の非標本誤差の変動（標準誤差率）を計算し、それらを合算することで非標本誤差に伴う各計数の短期の変動を計算する。

まず、「調査対象世帯の不在による無回答」および「就業状態ならびに従業上の地位不詳」による非標本誤差（ n_{1t} 、 n_{2t} ）の変動については、図表31で計算した公表計数のバイアスを月次で算出し、当該バイアスの前月差の標準誤差率を求める。

次に、「完全失業者と非労働力人口の区分の曖昧さに伴って生じる非標本誤差」（ n_{3t} ）については、公表計数のバイアスから直接、計算できないため、フローデータを用いて計算する。1993～97年の「完全失業者から非労働力人口へのフロー」は非標本誤差と真の変動を含み、1993～97年の「非労働力人口から完全失業者へのフロー」、2003～07年の「完全失業者から非労働力人口へのフロー」、同「非労働力人口から完全失業者へのフロー」、各々には上記の非標本誤差が存在せず、真の変動のみを含むことから、1993～97年における「完全失業者と非労働力人口の区分の曖昧さに伴って生じる非標本誤差の分散」は以下のように計算できる。

(1993～97年における「完全失業者と非労働力人口の区別の曖昧さに伴って生じる非標本誤差」の分散) = (1993～97年の両フローの差の分散) - (2003～07年の両フローの差の分散)

また2003～7年における「完全失業者と非労働力人口の区別の曖昧さに伴って生じる非標本誤差」による標準誤差率はゼロと考える。

さらに、「有給の自家営業の手伝いが家族従業者に混在することによる非標本誤差」(n_{4i})については、「詳細結果」が4半期ごとの公表であり、2年目・2ヶ月目のみの計数が月次で入手できないことから、標準誤差率を直接算出できない。そこで「家族従業者から雇用者へのフロー」の標準誤差率を用いて、非標本誤差の標準誤差率を計算する。1993～97年のフローデータは非標本誤差の変動を含まない一方、2003～7年のデータは非標本誤差の変動を含んでいる。

(2003～7年における「有給の自家営業の手伝いが家族従業者に混在することによる非標本誤差」の分散) = (2003～7年のフローの分散) - (1993～97年のフローの分散)

から2003～7年における標準誤差率を計算する。なお、1993～97年の標準誤差率はゼロと仮定する。

(2) 非標本誤差がもたらす短期の変動(振れ): 標準誤差率の試算結果

まず、「調査対象世帯の不在による無回答」による非標本誤差の変動をみると(図表33(1))、近年の無回答の増加を反映して、完全失業者を除き1993～97年から2003～7年にかけて標準誤差率は増加しており、2003～7年の家族従業者では0.3%程度に達している。一方、完全失業者では、2003～7年の標準誤差率は0.4%と値は大きめであるが、この10年間で標準誤差率は減少している。

この傾向は「就業状態ならびに従業上の地位不詳」による非標本誤差の短期の変動についても同様である。もっとも、「就業状態ならびに従業上の地位不詳」による標準誤差率の水準は小さい。比較的值が大きい家族従業者、完全失業者においても、標準誤差率は0.1～0.2%程度にとどまっている(図表33(2))。

次に「完全失業者と非労働力人口の区別の曖昧さ」に伴う非標本誤差の変動は、1993～97年では、完全失業者に標準誤差率で1.0%程度の短期の変動をもたらしている。しかしながら、調査1ヶ月目にも「仕事を探し始めた理由」の質問を追加した2003～7年においては、非標本誤差が解消したことから、標準誤差率はゼロとな

っている（図表 33(3)）¹⁹。

さらに、「有給の自家営業の手伝いが家族従業者に混在すること」による非標本誤差の変動についてみる。2006 年以降、2 年目・2 ヶ月目に限定して家族従業者に関する過大計上が修正され、2 年目・2 ヶ月目において「家族従業者から雇用者へのフロー」が増加していると推測される。そのため、家族従業者の短期の変動が顕著に増加しており、2003～7 年における標準誤差率は 1.7%に達している（図表 33(4)）。しかしながら、これは、有給の自家営業の手伝いの家族従業者への混在が一部修正されたことによる短期の変動（標準誤差率）の増大であり、その短期の変動の増大は、バイアスのレベルシフトに伴う一時的なものと考えられる。

（3）標本誤差による短期の変動と非標本誤差による短期の変動：合計の試算値

本稿での標本誤差や非標本誤差の試算結果は、以上述べたように一定の前提に基づくものであり、結果については幅を持ってみる必要がある。この点に留意しつつ、非標本誤差による短期の変動を図表 18 で算出した標本誤差による短期の変動に加算し、その合計値を実データの対数前月差の標準誤差と比較する（1995 年：図表 34(1)、2005 年：図表 34(2)）。家族従業者や完全失業者については、標本誤差のみでは最近の実データの大きな短期の変動を十分に説明することができなかった。しかし、そのうち家族従業者については、非標本誤差がもたらす短期の変動を追加した「標本誤差＋非標本誤差」合計による短期の変動で標準偏差が 2.5%に達しており、実際のデータの短期の変動（2.8%）の大部分を説明できるようになっている。また、1995 年から 2005 年までの変化についてみても、実データの家族従業者の短期の変動の増大（+1.1%）は、標本誤差に非標本誤差を合算した場合の短期の変動（振れ）の増加（+1.0%）でほぼ説明可能となっている（図表 34(3)）。一方で、完全失業者については、本稿で試算された非標本誤差では、その水準も変化も十分には説明できていない。

以上の結果から、「労働力調査」の各計数について短期の変動をもたらしているのは、主として標本誤差であるが、家族従業者については非標本誤差が短期の変動に大きな役割を果たしていることが分かる。もっとも、家族従業者における近年の短期の変動の増大には、有給の自家営業の手伝いが家族従業者に混在していた分が是正されたことによる非標本誤差の短期の変動の影響がほとんどであり、その寄与

¹⁹ より細かくみると、2002 年に世帯調査 1 ヶ月目にも「仕事を探し始めた理由」に関する質問が追加されたことにより、バイアスのレベルシフトが生じ、一時的に振れが増加している。ただし、本稿の推計期間（1993～97 年、2003～07 年）からはずれているため、分析では影響が生じていない。

は一時的なものである。このように非標本誤差が各計数の短期の変動に与える影響はかなり限定的である。

7 まとめ：「労働力調査」の精度改善に向けて

本稿では、「労働力調査」の公表計数の短期の変動（振れ）の性質について整理を行った。さらに、一定の前提の下で標本誤差や非標本誤差の変動が短期の変動に与える影響を試算したほか、非標本誤差がもたらすバイアス（計数の恒常的な偏り）についても試算を行った。本稿の分析の結果、分かったことは以下のとおりである。

「労働力調査」のいずれの計数についても、全変動に対する周期 12 ヶ月未満の短周期変動による寄与が大きい。計数別にみると、雇用者や非労働力人口については、変動の振幅は比較的小さいが、短周期の変動の寄与により、自営業主、家族従業者、完全失業者の変動の振幅はかなり大きい。時系列的な変化をみると、近年、家族従業者の短期の変動が増加している。このように短期の変動の増加は、2006 年の調査票の記入要領の改善がもたらした一時的なものと考えられるが、2006 年前後の前年比計数を用いる時には注意を要する。

一定の前提の下で、標本誤差ならびに非標本誤差の値を試算した結果、こうした短期の変動（振れ）をもたらしているのは、主として標本誤差である。非標本誤差によるバイアスについては、調査対象世帯の不在による無回答、就業状態ならびに従業上の地位不詳（就業状態・従業上の地位に関する無回答）の影響があるほか、有給の自家営業の手伝いが家族従業者に混在することによって生じる非標本誤差が比較的大きくなっている。

最後に、以上の分析結果を受けて、「労働力調査」の精度を一段と向上させるにはどのような方策を採りうるのかについて、考察してみたい。

標本誤差を縮小するに有効な改善策としては、①調査サンプル（調査対象世帯）数を増加させる、②層化抽出ルールを変更する、③同一世帯への継続調査期間（現行 2 ヶ月）をより長期に延長する、の 3 つが考えられる。

そのうち、①調査世帯数（現行 4 万世帯）を増加させるのは、あらゆる計数の誤差を縮小させることから最も抜本的な対策である。しかし、調査サンプル数を n 倍にしても標本誤差は $1/\sqrt{n}$ への減少に止まる。例えば、調査サンプル数を 2 倍にしても、標本誤差は▲ 3 割減少するに止まる。一方で、調査コスト面での負担が極めて大きくなる。

②層化抽出ルールの変更は、比較的低いコストで対処可能であることから、一見、魅力的な対策である。例えば、標本誤差が大きい家族従業者や完全失業者の精度を高めるためには、これらを層化基準に採用して層化を行えば、当該計数については標本誤差の一定程度の改善が見込める。しかし、層化基準に採用した計数の標本誤差は縮小する一方で、層化基準からはずされた計数の標本誤差は増加するというトレードオフがある²⁰。

③同一世帯への継続調査期間の延長も、調査コストの増加は小さく、標本誤差を改善する有力な対策となりうる。米国の雇用統計においては、同一世帯に対して4ヶ月間連続して調査を行っており、日本でも現行の調査期間を連続2ヶ月から4ヶ月へ延長することが検討課題となりうる。仮に同一世帯に対する連続調査期間を4ヶ月に延長すれば、標本誤差は▲2割程度減少させることができると推測される。もっとも、同一世帯に対する調査期間が延長され、調査対象世帯の報告者負担が重くなることから、調査の協力が得られない世帯が増加することに伴う非標本誤差が増加し、かえって統計精度が低下する可能性もある。このほか、継続サンプルの比率が高まることから、継続サンプルがもつ非標本誤差の影響が増加する可能性も存在する。

本稿で指摘した非標本誤差のうち、「有給の自家営業の手伝いが家族従業者に混在している」問題は、調査対象者が「調査票の記入のしかた」を十分に参照しないことに起因するものであり、根本的な是正には困難が伴う。むしろ、統計ユーザーが、こうした問題に統計の精度が左右されている点を認識しておく必要がある。ただし、調査票の記入要領の変更などが行われると、その前後の前年比計数等は実勢と乖離するため、それが大きいと考えられる場合には、統計作成部署によるユーザーに対する説明が開示されることが望ましい。

また、「調査対象世帯の不在による無回答」については、計数の遡及訂正による対応が一案ではないかと考えられる。すなわち、現在のところ、統計調査員の努力により、調査2ヶ月目にはかなりの数が調査可能となっている。2ヶ月目に調査可能となった対象者に対して、無回答であった1ヶ月目についても就業状態ならびに従業上の地位を遡及調査し、無回答で処理されていた1ヶ月目の公表計数を遡及訂正することができれば、本稿で指摘した「調査対象世帯の不在による無回答」から生じている非標本誤差は相当程度解消できると考えられる。

²⁰ 例えば、層化基準として就業状態（自営業主、家族従業者、雇用者、完全失業者、非労働力人口）を用いれば、これらの計数の標本誤差を最大▲1割程度削減することが可能であると試算される。しかし、その場合、層化基準からはずれる産業別就業者数の標本誤差は増加する。

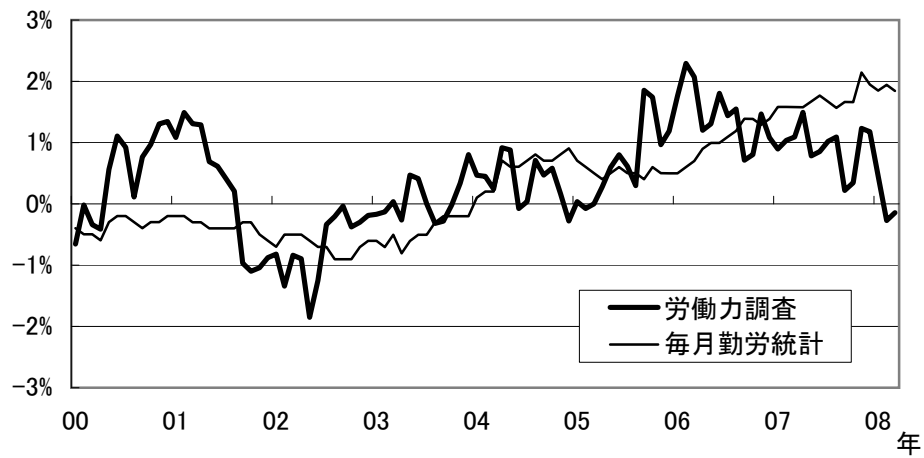
【参考文献】

- 黒田祥子、「わが国失業率の変動について—フロー統計からのアプローチ」、『金融研究』第21巻第4号、日本銀行金融研究所、2002年
- 桜健一、「フローデータによるわが国労働市場の分析」、日本銀行ワーキングペーパーシリーズ、06-J-20、2006年
- 総務省統計局、「平成17年国勢調査の聞き取り調査等の状況及び「国勢調査の実施に関する有識者懇談会」における検討状況」、総務省統計局ホームページ、2006年5月30日
- 、『労働力調査 標本設計の解説』、総務省統計局、2008年
- 総務庁統計局、『労働力調査の解説』、総務庁統計局、1993年
- 松井博、『標本調査法入門—基礎から学ぶ、標本調査の理論と実際—』、日本統計協会、2005年
- 松田芳郎・伴金美・美添泰人編、『講座「ミクロ統計分析」第2巻 ミクロ統計の集計解析と技法』、日本評論社、2000年

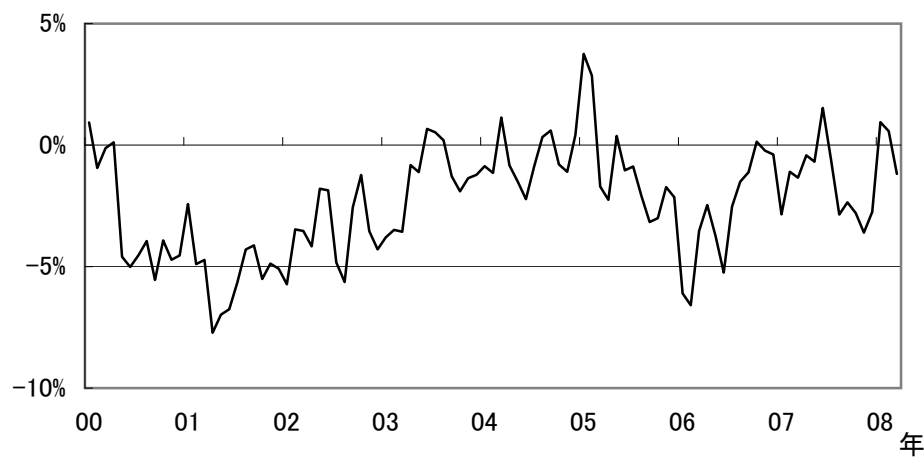
「労働力調査」各計数の動き

—2000年1月～2008年3月、男女計、前年同月比—

(1) 労働力調査（雇用者）と毎月勤労統計常用雇用指数（5人以上）の比較



(2) 自営業主



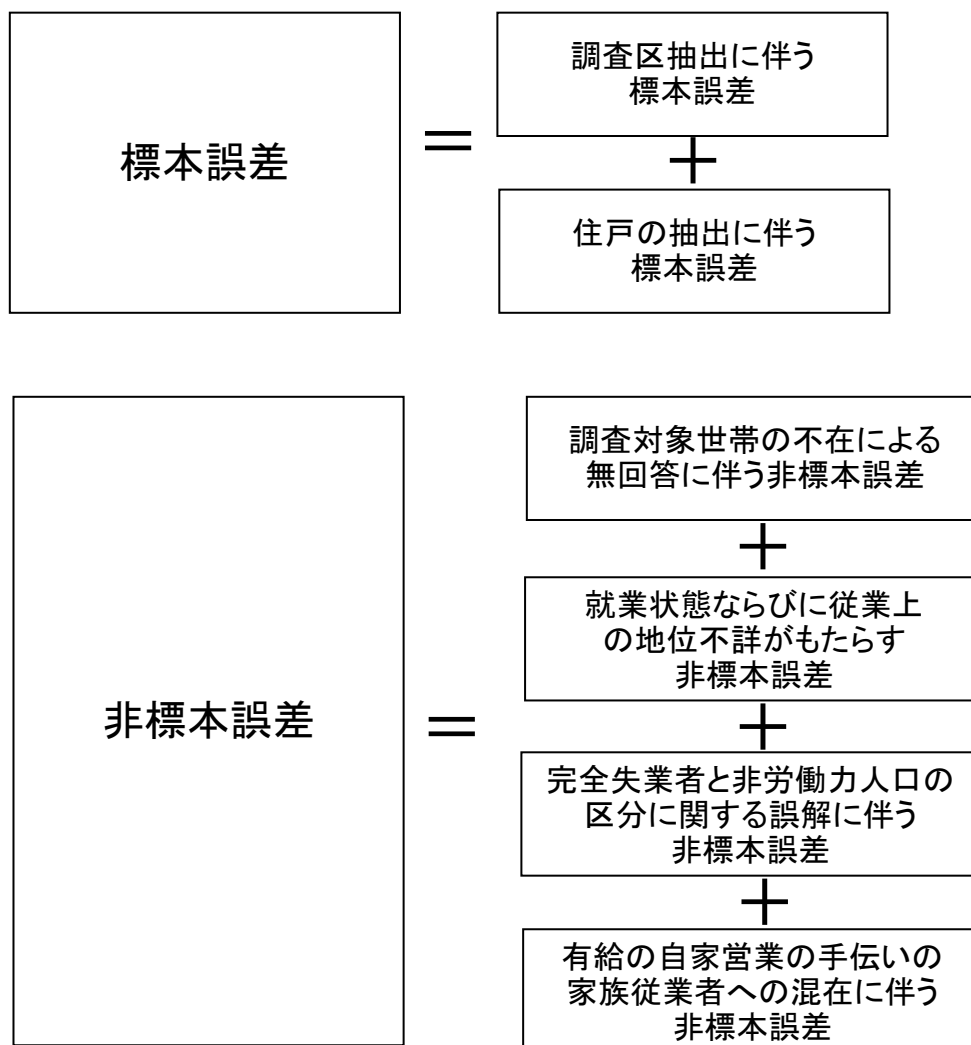
(3) 家族従業者



* 労働力調査と毎月勤労統計、米国雇用統計を比較する際の留意点については、4ページの本文ならびに注2を参照。

労働力調査の標本誤差と非標本誤差

(1) 労働力調査における標本誤差と非標本誤差の要因



(2) 標本誤差と非標本誤差の特徴

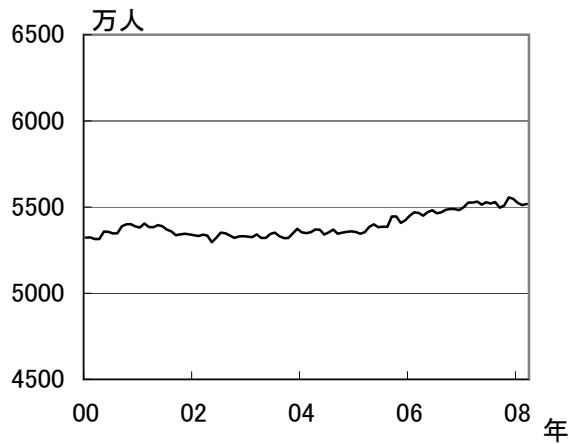
	バイアス	振れ
標本誤差	× 長期的な偏りには影響なし	○ 短期の変動をもたらす
非標本誤差	○ 長期的な偏りに影響する	△ 短期の変動をもたらす

○: 影響大、△: 影響小、×: 影響なし

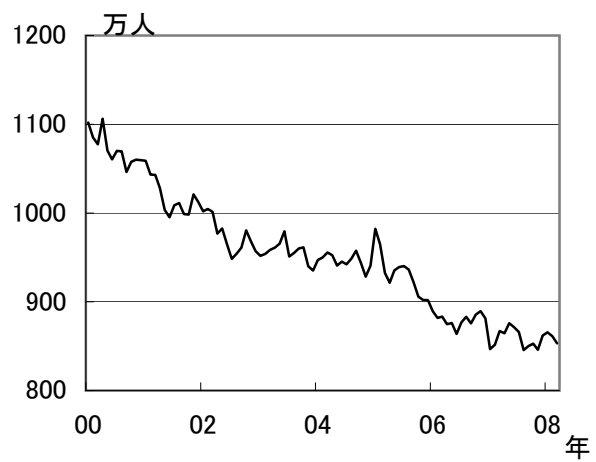
「労働力調査」各計数の動き

—2000年1月～2008年3月、男女計、季節調整済み—

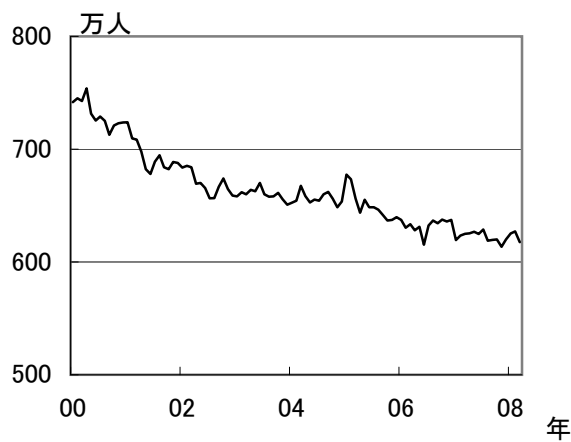
(1) 雇用者



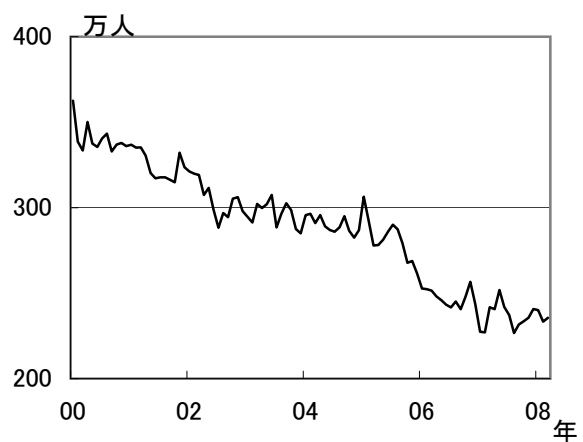
(2) 自営業主+家族従業者



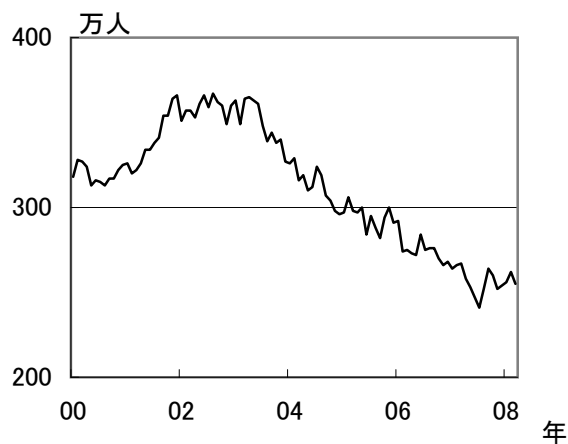
(2) - (a) 自営業主



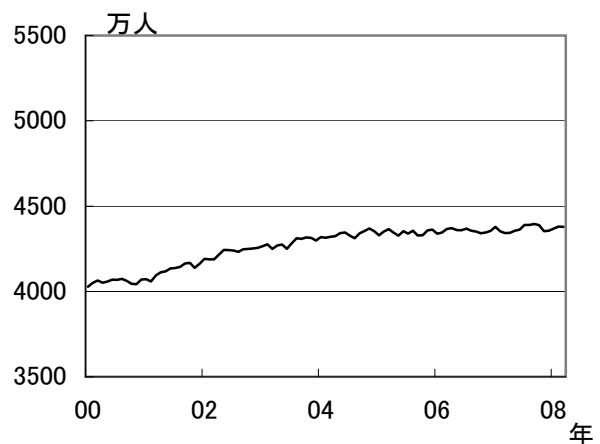
(2) - (b) 家族従業者



(3) 完全失業者



(4) 非労働力人口



各計数（対数前月差）の標準偏差
—1988年1月～2008年3月、季節調整済み—

		原計数	標準偏差	
			2-12ヶ月周期	12-60ヶ月周期
労働力調査 (男女計)	雇用者	0.33%	0.31%	0.06%
	自営・家族計	1.20%	1.17%	0.26%
	自営業主	1.14%	1.12%	0.25%
	家族従業者	2.13%	1.97%	0.38%
	完全失業者	2.60%	2.42%	0.55%
	非労働力人口	0.36%	0.35%	0.08%
雇用米 統計	雇用者	0.26%	0.25%	0.09%
	完全失業者	2.40%	2.17%	0.75%
	非労働力人口	0.38%	0.38%	0.10%
毎勤常用雇用指数(30人以上)		0.18%	0.13%	0.05%

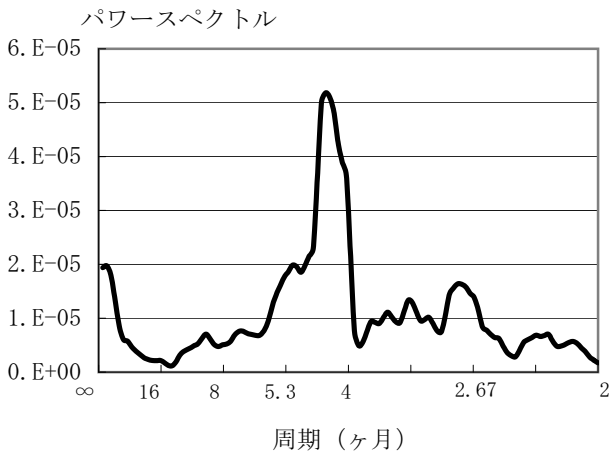
*ただし2-12ヶ月周期の分散については、Band-Passフィルターで周期変動を抽出しているため、データ期間が1988年1月～2005年3月。

* 労働力調査と毎月勤労統計、米国雇用統計を比較する際の留意点については、4ページの本文ならびに注2を参照。

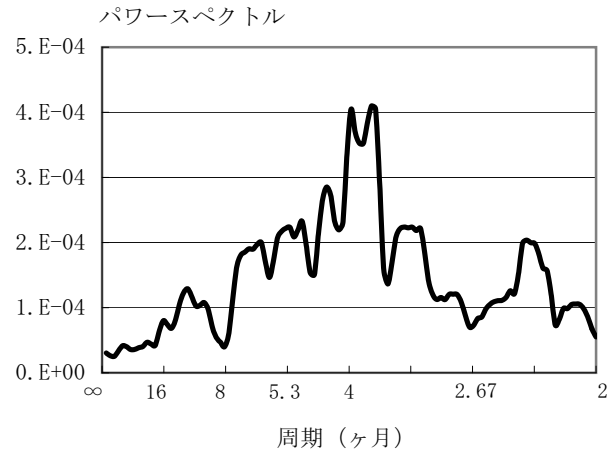
「労働力調査」各計数のスペクトル分解

—1988年1月～2008年3月、季節調整済み—

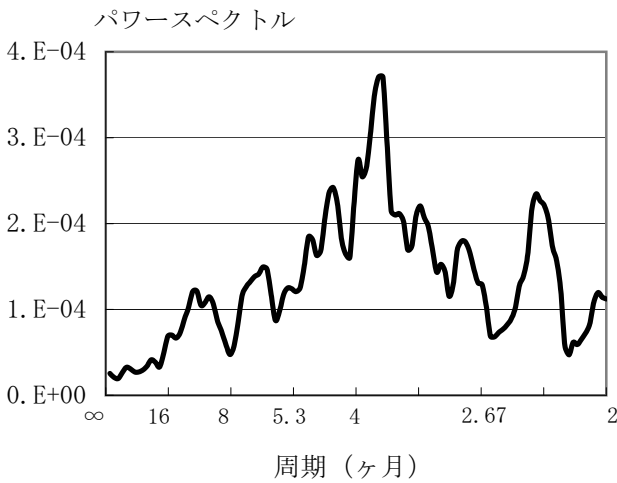
(1) 雇用者



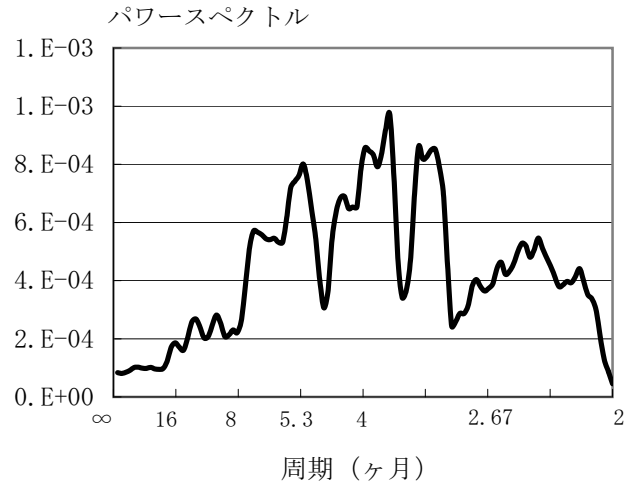
(2) 自営業主+家族従業者



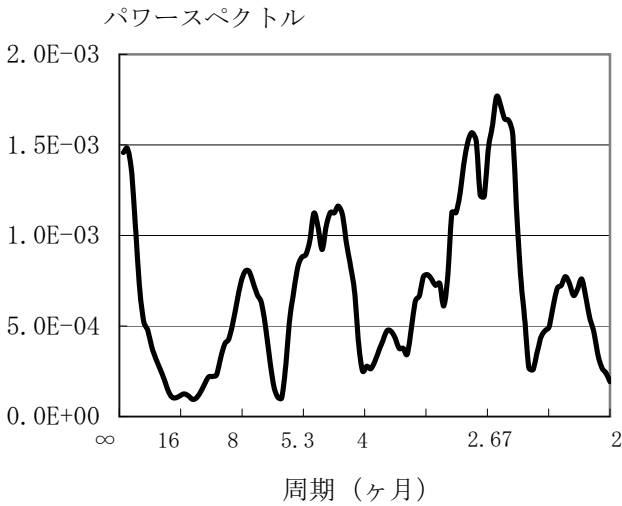
(2) - (a) 自営業主



(2) - (b) 家族従業者



(3) 完全失業者



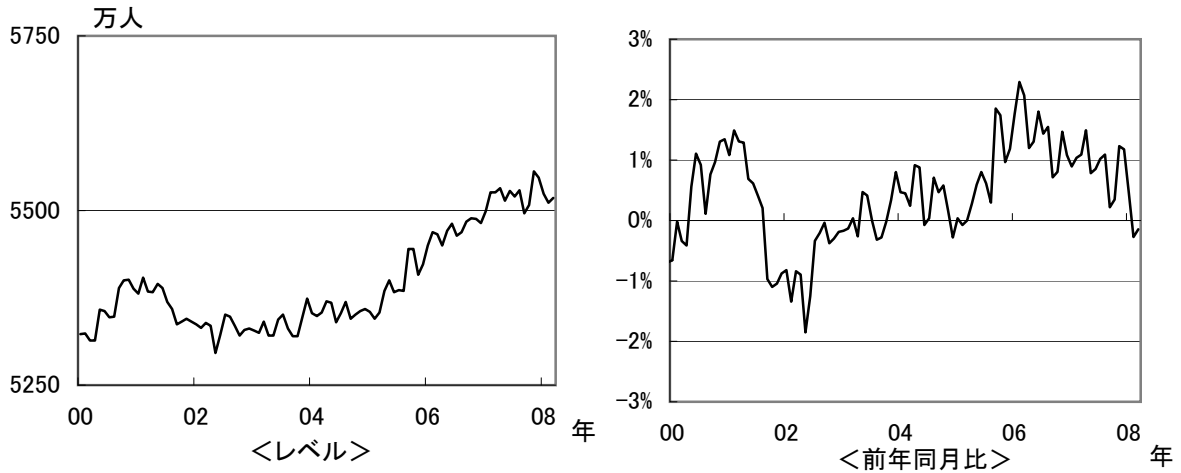
(4) 非労働力人口



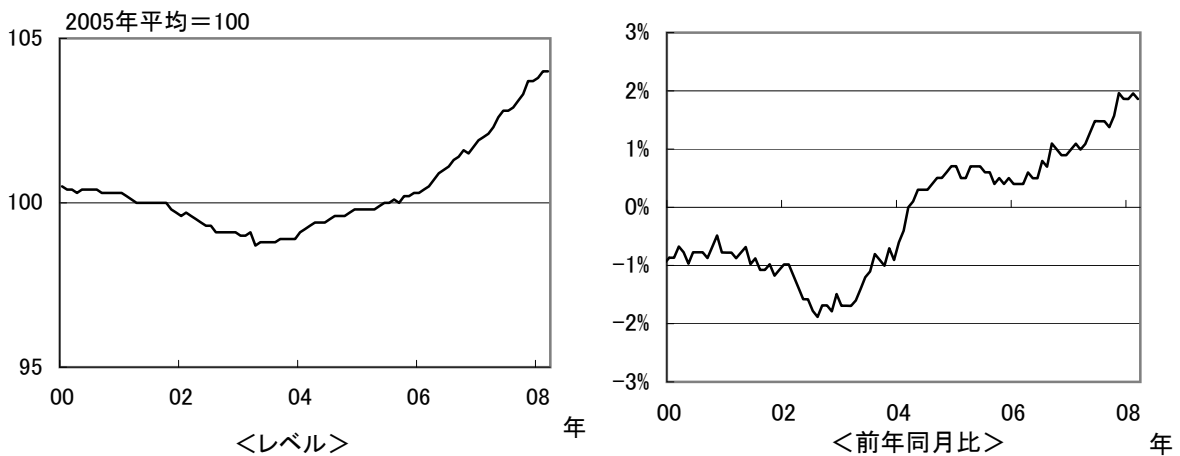
雇用者数の推移

—2000年1月～2008年3月、男女計、季節調整済み—

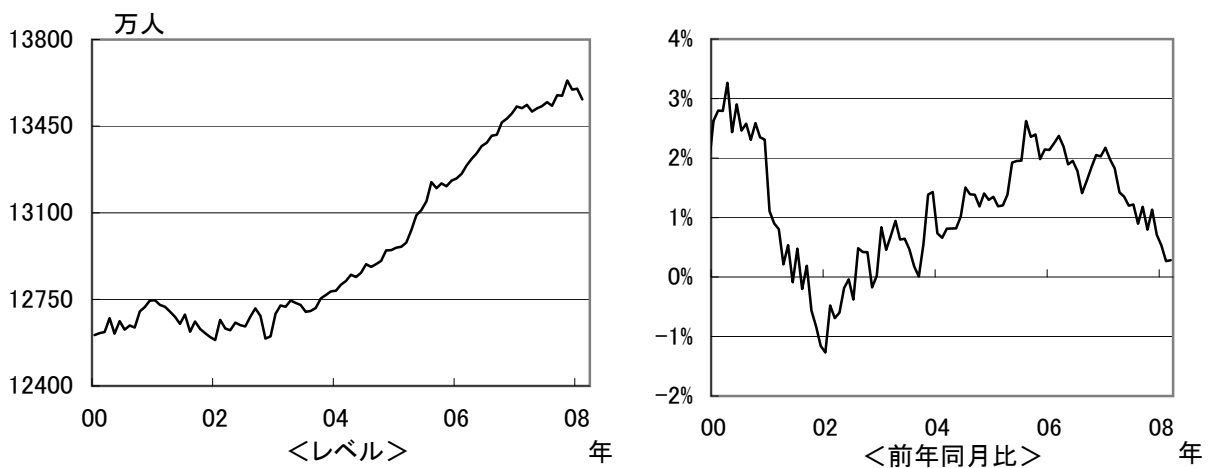
(1) 労働力調査 (雇用者数)



(2) 毎月勤労統計常用雇用指数 (30人以上)



(3) 米国雇用統計

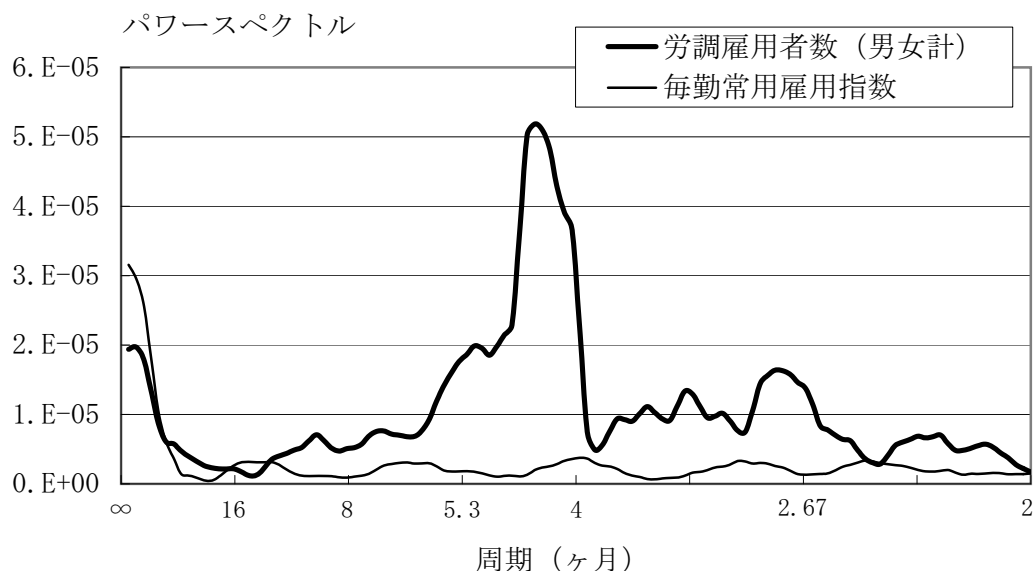


* 労働力調査と毎月勤労統計、米国雇用統計を比較する際の留意点については、4ページの本文ならびに注2を参照。

スペクトル分解：他の統計との比較

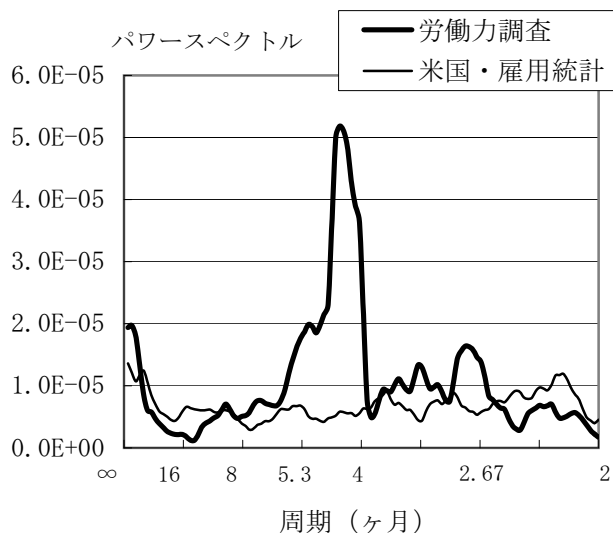
—1988年1月～2008年3月、季節調整済み—

(1) 毎月勤労統計（常用雇用指数（30人以上））との比較

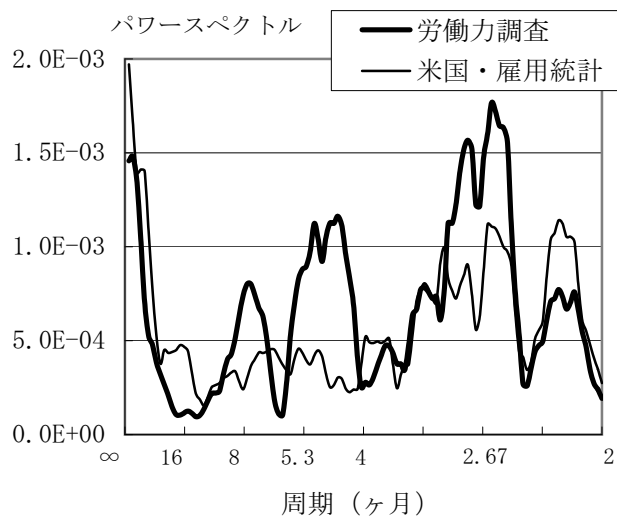


(2) 米国雇用統計との比較

①雇用者



②完全失業者



* 労働力調査と毎月勤労統計、米国雇用統計を比較する際の留意点については、4ページの本文ならびに注2を参照。

各計数の変動の時系列変化
—対数前月差(前月比)の標準偏差、季節調整済み—

		1988～1997年		1998年～直近	
			2-12ヶ月周期		2-12ヶ月周期
労働力調査 (男女計)	雇用者	0.33%	0.32%	0.32%	0.29%
	自営・家族計	1.09%	1.06%	1.30%	1.32%
	自営業主	1.13%	1.11%	1.16%	1.14%
	家族従業者	1.71%	1.67%	2.47%	2.34%
	完全失業者	2.78%	2.66%	2.41%	2.05%
	非労働力人口	0.37%	0.36%	0.36%	0.35%
雇用米 統計	雇用者	0.27%	0.25%	0.25%	0.25%
	完全失業者	2.27%	2.02%	2.52%	2.38%
	非労働力人口	0.40%	0.39%	0.34%	0.35%
毎勤常用雇用指数(30人以上)		0.20%	0.16%	0.14%	0.10%

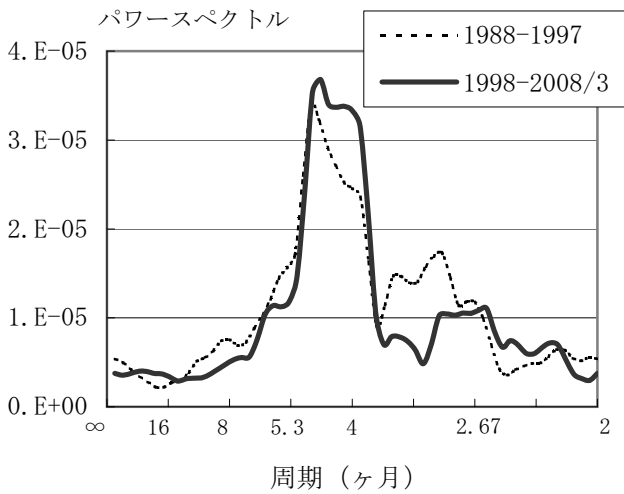
*ただし2-12ヶ月周期の分散については、Band-Passフィルターで周期変動を抽出しているため、データ期間がそれぞれ1988-1997年、1998年-2005年3月。

* 労働力調査と毎月勤労統計、米国雇用統計を比較する際の留意点については、4ページの本文ならびに注2を参照。

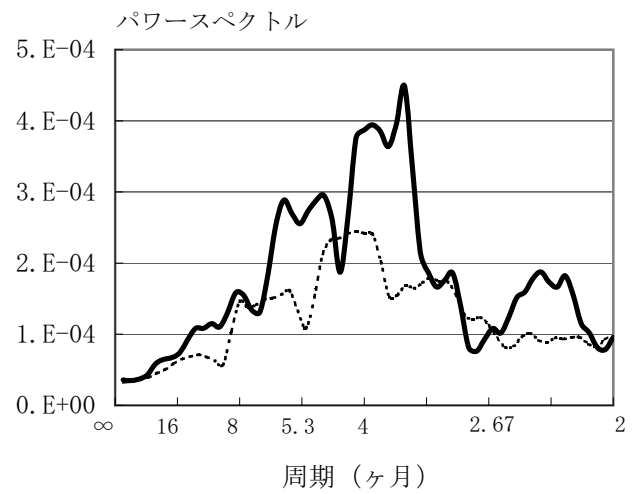
スペクトル分解の時系列変化

—男女計、季節調整済み—

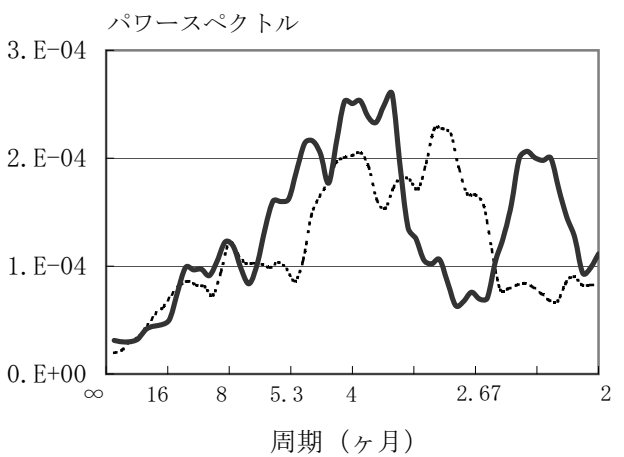
(1) 雇用者



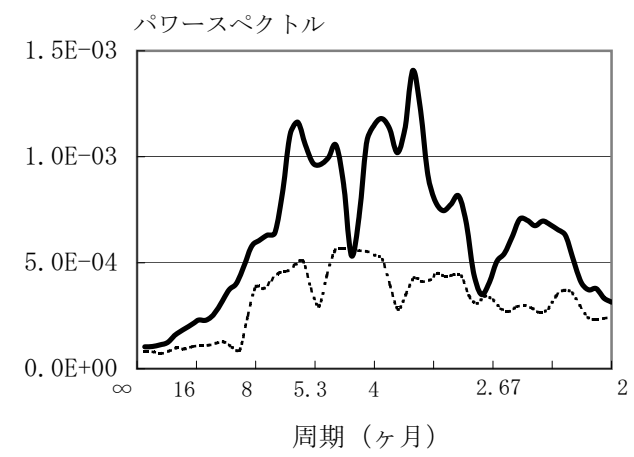
(2) 自営業主+家族従業者



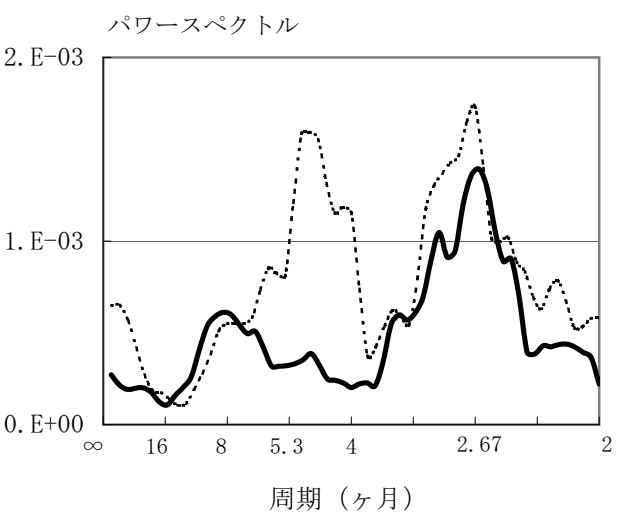
(2) - (a) 自営業主



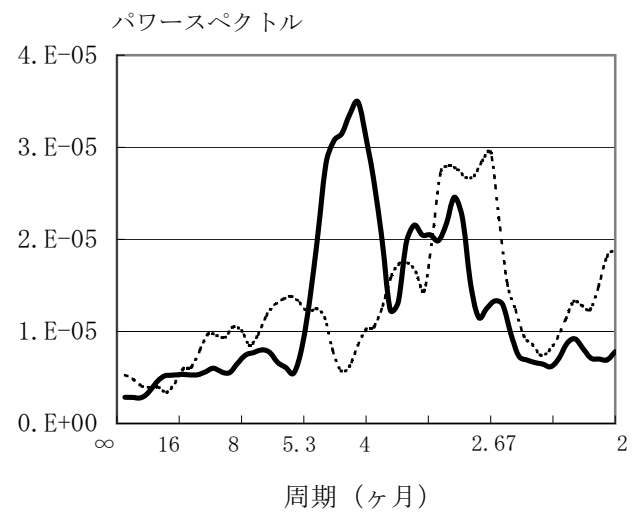
(2) - (b) 家族従業者



(3) 完全失業者



(4) 非労働力人口



労働力調査の概要

調査戸数	約 4 万世帯 (日本の世帯数：4,960 万世帯)
サンプリング方法	層化 2 段抽出 第 1 次：全国 98 万の国勢調査区から 2,912 の調査区を層化抽出 第 2 次：住戸を無作為抽出（1 調査区約 50 世帯から 15 世帯を選ぶ）。
標本の交代方法 (ローテーション構造)	調査区 ：4 ヶ月連続して調査（よって毎月、1/4 が交代）。ただし 8 ヶ月休んだ後、また 4 ヶ月連続して調査する。 住戸 ：調査区の中では 2 ヶ月ずつに分けて前期と後期で異なる住戸を調査（よって毎月、1/2 が交代）。ただし 2 ヶ月連続して調査し、10 ヶ月休んだ後、また 2 ヶ月連続して調査する。
特定調査票	2 年目の世帯調査 2 ヶ月目（最終月）に、通常の基本調査票に加えて実施される。より詳しい従業上の地位や、求職活動の状況、前職について調査する。詳細集計として公表される。
推定値の計算法	以下の各月の調査サンプル内における 4 つの種類のサンプル（副標本）の平均

<ローテーション構造の例>

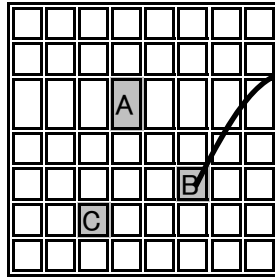
	前年			当年				
	10月	11月	12月	1月	2月	3月	4月	5月
調査区 1 ヶ月目、世帯調査 1 ヶ月目 (新規に調査対象となった世帯)				1	2	①	②	
調査区 2 ヶ月目、世帯調査 2 ヶ月目 (先月も調査対象であった世帯)			1	2	①	②		
調査区 3 ヶ月目、世帯調査 1 ヶ月目 (新規に調査対象となった世帯)		1	2	①	②			
調査区 4 ヶ月目、世帯調査 2 ヶ月目 (先月も調査対象であった世帯)	1	2	①	②				

一年後に同じサンプルを調査

労働力調査のサンプル抽出方法

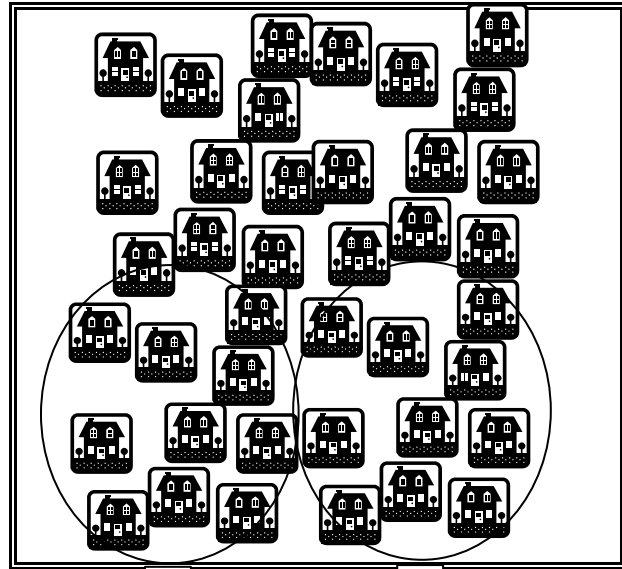
調査区変更時

調査区の決定



約98万の調査区から
2,912調査区を抽出

<調査区B>



抽出

抽出

約50世帯から、1ヶ月目、2ヶ月目の調査のための15世帯を抽出。

同時に、3ヶ月目、4ヶ月目の調査のための15世帯を抽出。

継続・非継続サンプルの短期変動

—1988年1月～2008年3月、季節調整済み—

(1) 標準偏差 (対数前月差)

		原計数	2-12ヶ月周期	12-60ヶ月周期
継続 サンプル	雇用者数	0.40%	0.28%	0.09%
	自営・家族計	0.98%	0.71%	0.35%
	自営業主	0.98%	0.77%	0.50%
	家族従業者	2.45%	1.43%	0.48%
	完全失業者数	2.66%	2.13%	0.95%
	非労働力人口	0.65%	0.41%	0.11%
非継続 サンプル	雇用者数	0.70%	0.58%	0.15%
	自営・家族計	2.46%	2.29%	0.63%
	自営業主	2.25%	2.15%	0.71%
	家族従業者	4.21%	3.51%	0.77%
	完全失業者数	5.10%	4.27%	0.99%
	非労働力人口	0.87%	0.71%	0.18%

*ただし2-12ヶ月周期の分散については、Band-Passフィルターで周期変動を抽出しているため、データ期間が1988年1月～2005年3月。

(2) 継続・非継続サンプルの公表値の分散に対する寄与

		公表値 (分散×10,000)	継続サンプル	非継続サンプル	その他 (共分散)
寄与 度	雇用者数	0.112	0.041	0.123	-0.051
	自営・家族計	1.440	0.241	1.518	-0.318
	自営業主	1.302	0.238	1.262	-0.198
	家族従業者	4.521	1.500	4.433	-1.412
	完全失業者数	6.753	1.775	6.501	-1.524
	非労働力人口	0.133	0.106	0.189	-0.161

* 公表値の変動 = $\frac{\text{継続サンプルの変動} + \text{非継続サンプルの変動}}{2}$ なので、

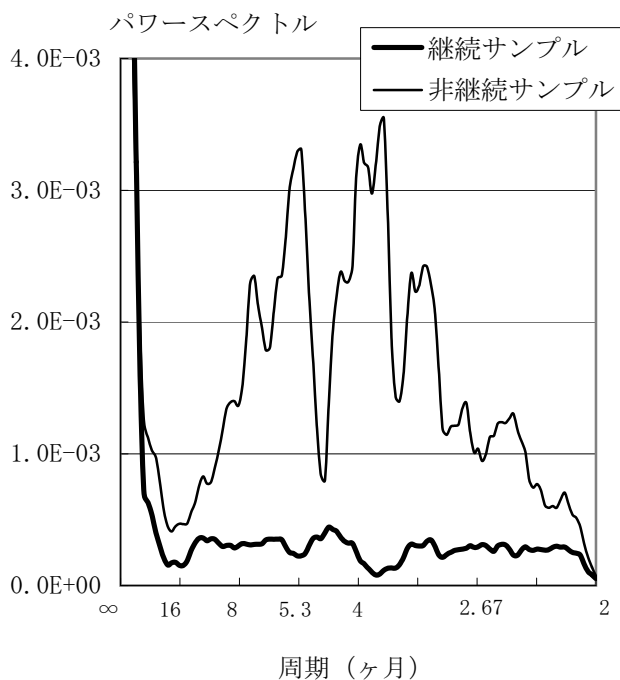
$$\text{Var}(\text{公表値の変動}) = \frac{1}{4}\text{Var}(\text{継続サンプルの変動}) + \frac{1}{4}\text{Var}(\text{非継続サンプルの変動})$$

$$+ \frac{1}{2}\text{Cov}(\text{継続サンプルの変動}, \text{非継続サンプルの変動})$$

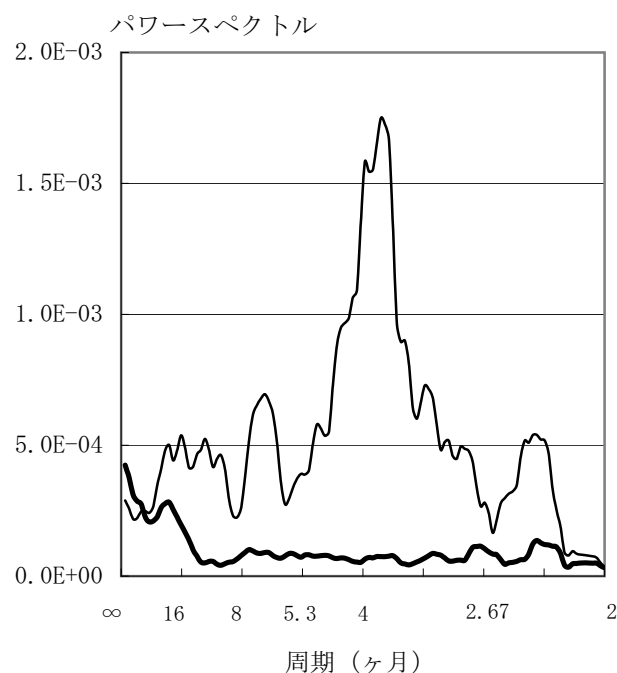
継続サンプルと非継続サンプルのスペクトル分解

—1988年1月～2008年3月、季節調整済み—

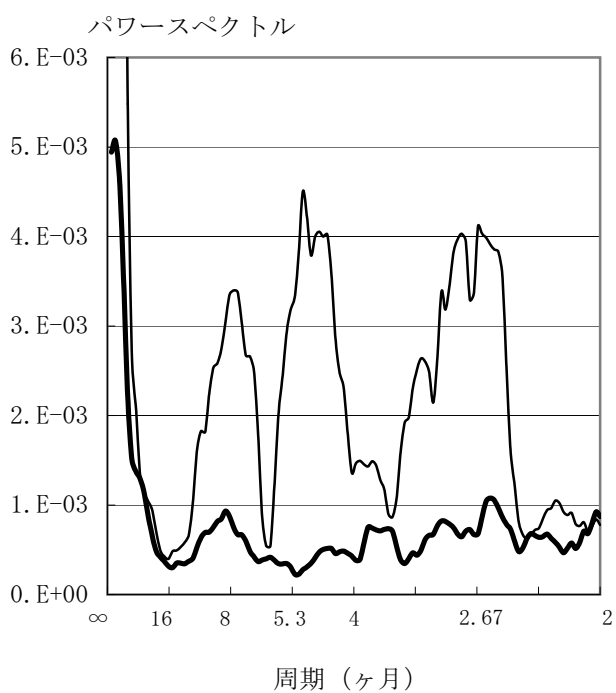
(1) 雇用者



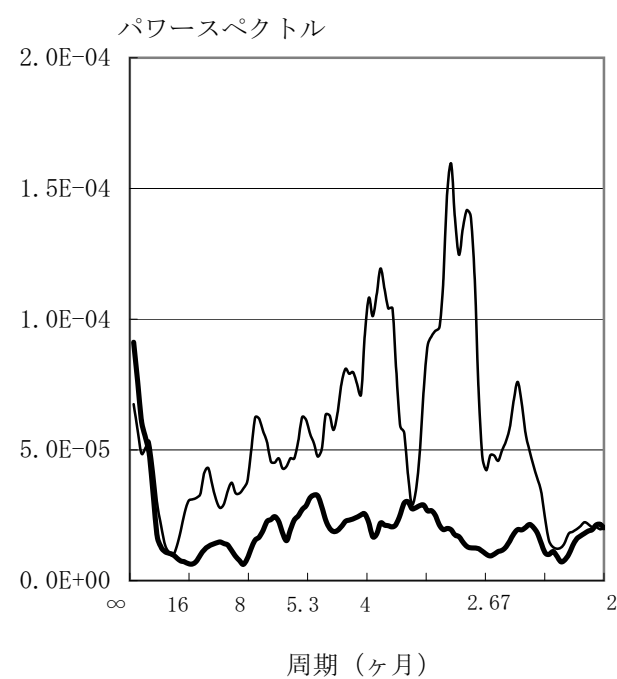
(2) 自営業主+家族従業者



(3) 完全失業者



(4) 非労働力人口



町丁字の規模別、各計数の標準誤差率

—母集団から調査区を確率比例抽出した場合—

1995年 国勢調査

町丁字の 人口サイズ	サンプル数 (平均)	雇用者	自営・家族		完全失業者	非労働力 人口	
			自営業主	家族従業者			
100～200	147	0.90%	2.83%	2.20%	4.13%	3.28%	0.92%
201～300	247	0.78%	2.71%	2.10%	4.06%	2.59%	0.83%
301～400	348	0.74%	2.71%	2.08%	4.12%	2.38%	0.82%
401～500	448	0.71%	2.58%	2.01%	3.84%	2.18%	0.79%
501～600	549	0.66%	2.53%	1.98%	3.87%	2.23%	0.76%
601～700	649	0.61%	2.39%	1.86%	3.68%	1.84%	0.71%
701～800	748	0.60%	2.42%	1.87%	3.73%	1.82%	0.71%
801～1000	892	0.57%	2.24%	1.76%	3.48%	1.78%	0.70%
1001～2000	1382	0.52%	2.12%	1.64%	3.40%	1.56%	0.68%
2001～	3274	0.37%	1.51%	1.19%	2.55%	1.35%	0.54%
全町丁字	678	0.57%	2.38%	1.79%	3.87%	1.82%	0.69%

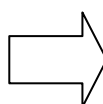
2005年 国勢調査

町丁字の 人口サイズ	サンプル数 (平均)	雇用者	自営・家族		完全失業者	非労働力 人口	
			自営業主	家族従業者			
100～150	123	0.98%	2.96%	2.29%	4.52%	2.79%	1.01%
100～200	146	0.93%	2.95%	2.27%	4.53%	2.62%	0.95%
201～300	247	0.80%	2.77%	2.10%	4.42%	2.12%	0.86%
301～400	347	0.76%	2.66%	2.00%	4.37%	1.87%	0.82%
401～500	448	0.71%	2.45%	1.87%	4.05%	1.76%	0.78%
501～600	549	0.67%	2.46%	1.82%	4.25%	1.63%	0.74%
601～700	648	0.66%	2.32%	1.74%	3.97%	1.57%	0.75%
701～800	749	0.61%	2.15%	1.65%	3.66%	1.46%	0.71%
801～1000	893	0.57%	1.94%	1.53%	3.28%	1.43%	0.67%
1001～2000	1385	0.52%	1.86%	1.43%	3.37%	1.34%	0.65%
2001～	3195	0.42%	1.41%	1.08%	2.78%	1.20%	0.55%
全町丁字	654	0.60%	2.30%	1.68%	4.17%	1.54%	0.70%

労働力調査における国勢調査区の層化基準

＜労働力調査の層化基準＞

分類番号	分類基準
01	刑務所、拘留所などの地域、自衛隊区域、駐留軍区域、水面
02	人口が0
03	換算世帯数が15以下
04	社会施設、大きな病院のある区域、寄宿舎、寮などのある地域
05	鉱業の就業者の比が0.1以上
06	漁業の就業者の比が0.2以上
07	漁業の就業者の比が0.1以上0.2未満
08	建設業、製造業の業主の比が0.1以上
09	卸売・小売業、飲食店の業主の比が0.1以上
10	電気・ガス・熱供給・水道業、運輸・通信業、金融・保険業、不動産業、サービス業の業主の比が0.1以上
11	農林業の就業者の比が0.3以上
12	農林業の就業者の比が0.1以上0.3未満
13	公務の就業者の比が0.1以上
14	金融・保険業、不動産業の雇用者の比が0.1以上
15	製造業の雇用者の比が0.3以上
16	建設業の雇用者の比が0.1以上
17	卸売・小売業、飲食店の雇用者の比が0.2以上
18	サービス業の雇用者の比が0.2以上
19	電気・ガス・熱供給・水道業、運輸・通信業の雇用者の比が0.1以上
20	製造業の雇用者の比が0.2以上0.3未満
21	製造業の雇用者の比が0.1以上0.2未満
22	卸売・小売業、飲食店の雇用者の比が0.1以上0.2未満
23	サービス業の雇用者の比が0.1以上0.2未満
24	上記のいずれにも属さない



＜本分析の層化基準＞

分類番号	分類基準
01	鉱業の就業者の比が0.1以上
02	漁業の就業者の比が0.2以上
03	漁業の就業者の比が0.1以上0.2未満
04	農業の就業者の比が0.3以上
05	農業の就業者の比が0.1以上0.3未満
06	自営業主の比が0.1以上
07	公務の就業者の比が0.1以上
08	金融・保険業、不動産業の就業者の比が0.1以上
09	製造業の就業者の比が0.3以上
10	建設業の就業者の比が0.1以上
11	卸売・小売業、飲食店の就業者の比が0.2以上
12	サービス業の就業者の比が0.2以上
13	電気・ガス・熱供給・水道業、運輸・通信業の就業者の比が0.1以上
14	製造業の就業者の比が0.2以上0.3未満
15	製造業の就業者の比が0.1以上0.2未満
16	卸売・小売業、飲食店の就業者の比が0.1以上0.2未満
17	サービス業の就業者の比が0.1以上0.2未満
18	上記のいずれにも属さない

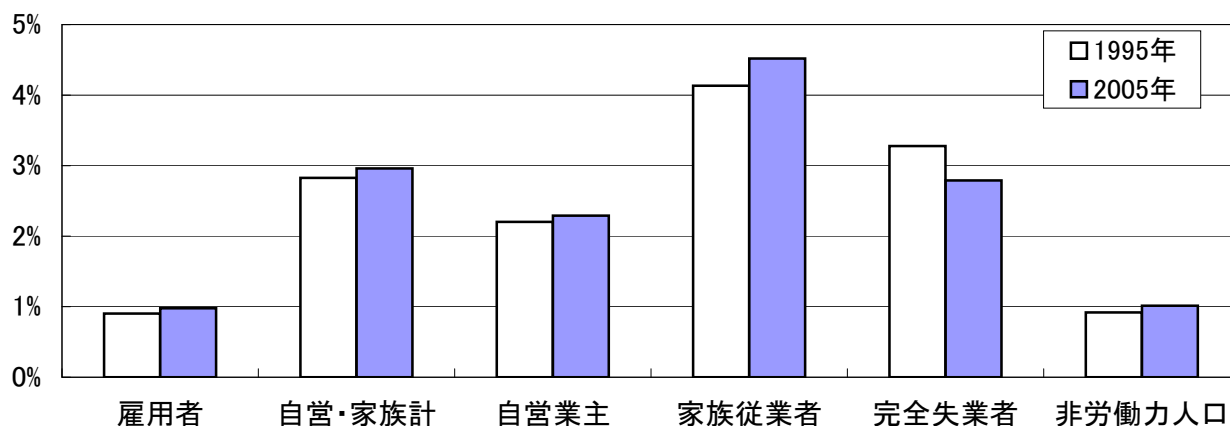
昭和60年（1985年）の国勢調査に基づき、1989年に変更して以降、変更無し。

(1) 産業、従業の地位別就業者数のデータが平成17年について入手できないため、就業者数を基準に層化。

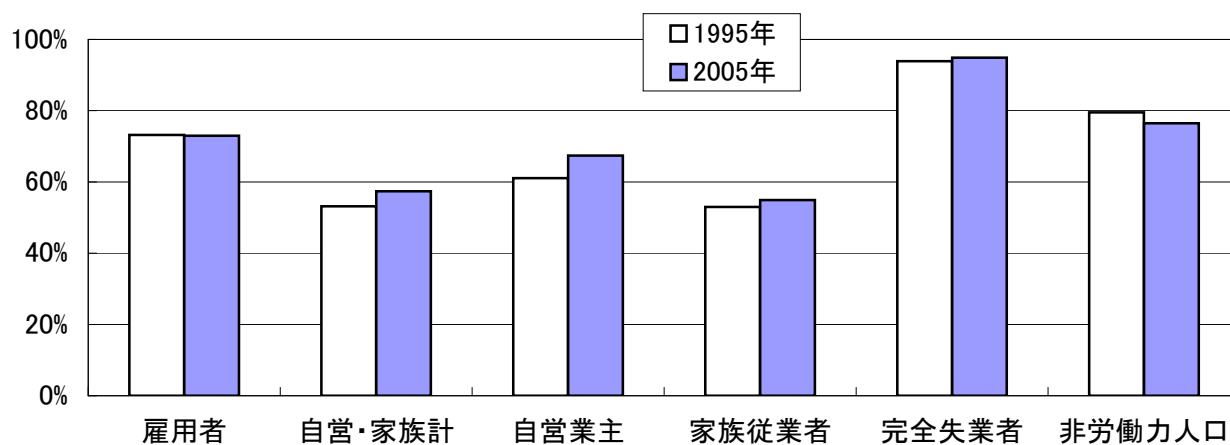
(2) 農業の就業者はほとんどが自営業主・家族従業者であるため、自営業主の比で層化されないよう若い分類番号を設定。

調査区の抽出による標本誤差（標準誤差率）

(1) 人口を基準に調査区を母集団から確率比例抽出した場合の標準誤差率

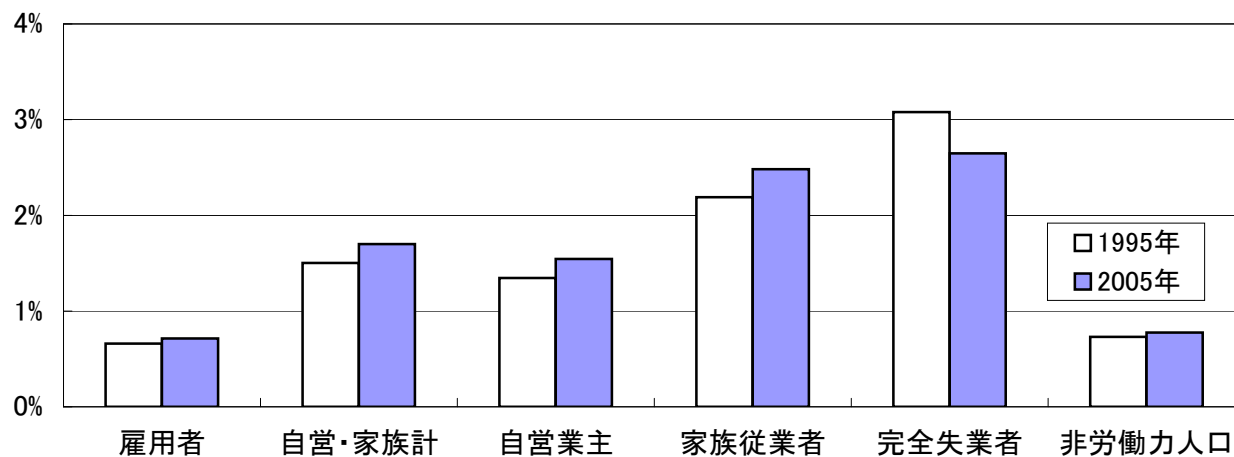


(2) 層化の効果：層化抽出後の標準誤差残存割合

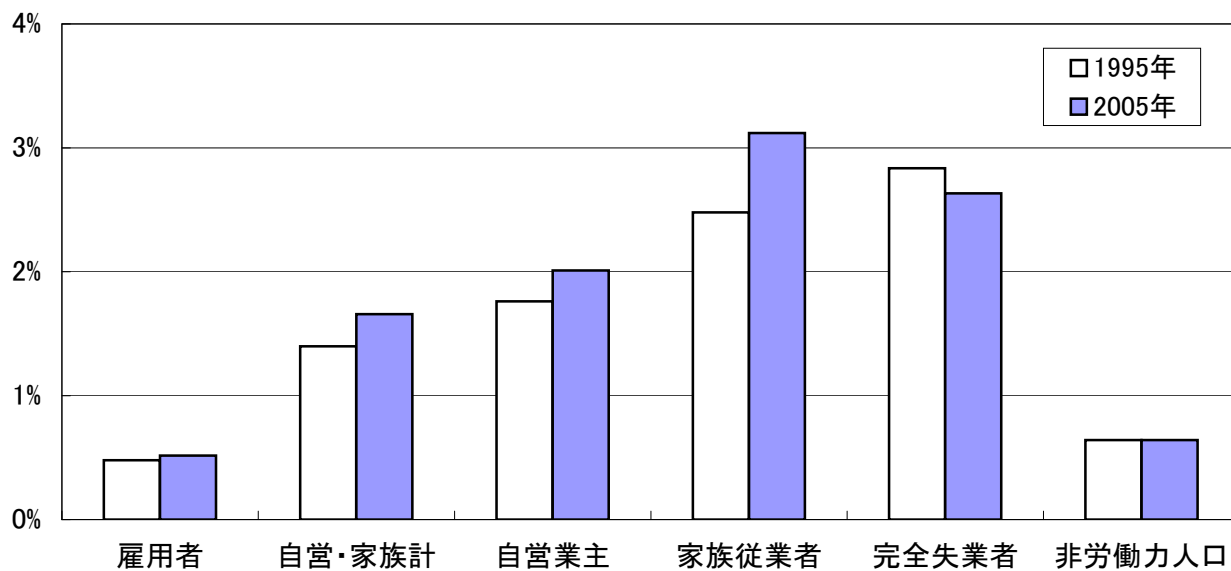


* 標準誤差残存割合 = 層化抽出の誤差率 / 確率比例抽出した場合の誤差率

(3) 調査区の抽出による標本誤差率 (= (1) × (2))

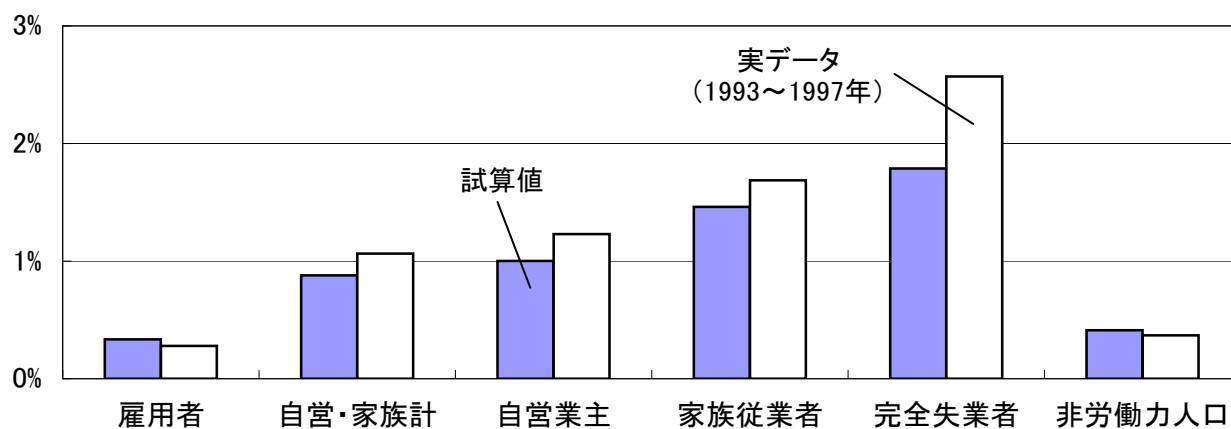


住戸の抽出に伴う標準誤差率

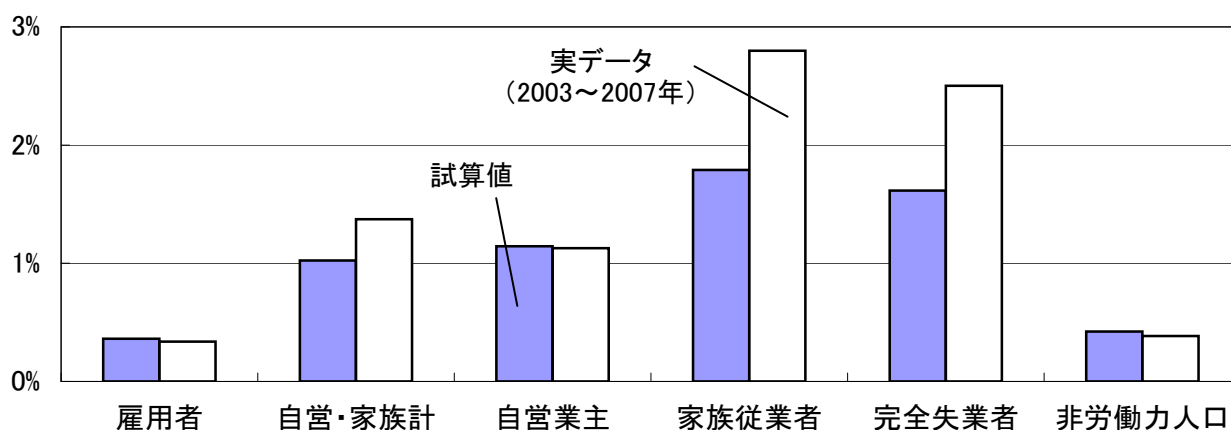


標本誤差に伴う変動と実データとの比較

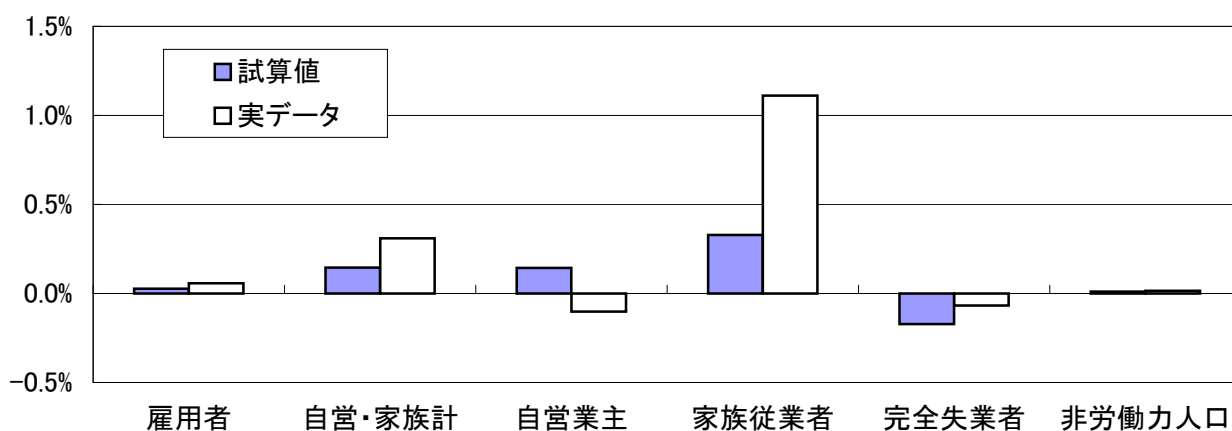
(1) 標本誤差に伴う変動 (1995年) と実データの対数前月差の標準偏差 (1993~1997年)



(2) 標本誤差に伴う変動 (2005年) と実データの対数前月差の標準偏差 (2003~2007年)



(3) 1995年から2005年への変化



フローデータの仕組み

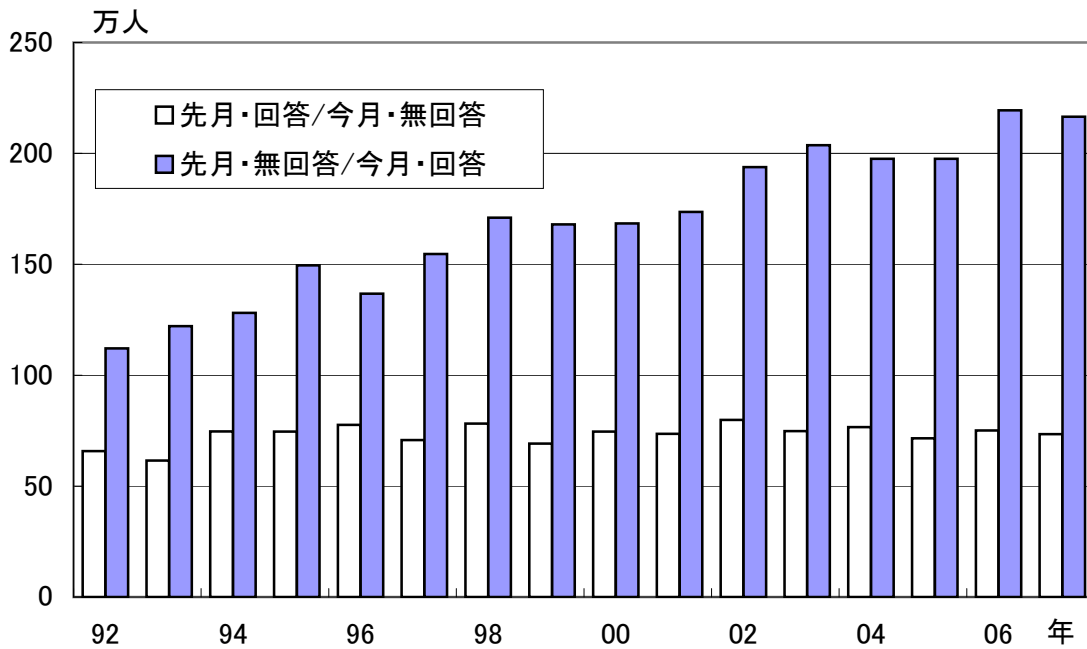
2007年平均

(万人)

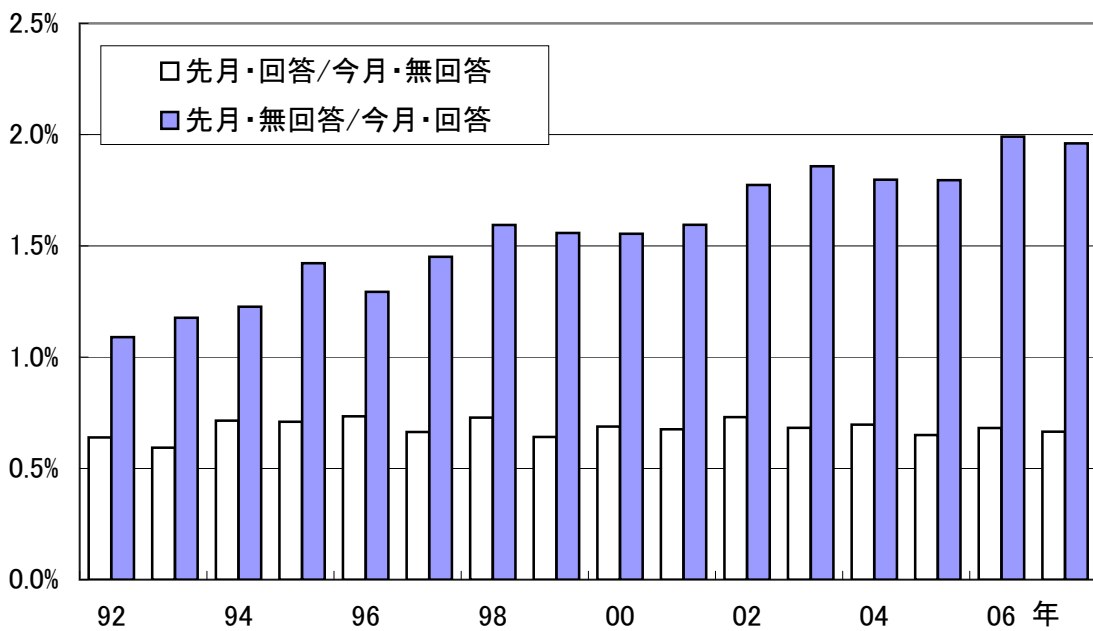
今月の状態	前月の状態									
	総数	雇用者	自営業主	家族従業者	従業上の地位不詳	完全失業者	非労働力人口	就業状態不詳	前月14歳で今月15歳の者	今月いて前月いなかった者
総数	11116	5428	618	243	34	254	4305	8	8	216
雇用者	5546	5266	33	30	13	28	57	1	0	117
自営業主	618	26	557	11	1	2	11	0	0	8
家族従業者	225	15	9	187	0	0	10	0	0	2
従業上の地位不詳	26	3	1	0	13	2	4	0	0	2
完全失業者	258	30	1	0	2	191	27	0	0	6
非労働力人口	4363	55	14	13	5	29	4159	2	8	78
就業状態不詳	6	0	0	0	0	0	0	5	0	1
前月いて今月いなかった者	73	33	1	0	0	2	35	0		

無回答・無回答率の推移

(1) 実数 (母集団推計値、年平均)



(2) 15歳以上人口に占める比率 (年平均)



無回答サンプルと回答サンプルとの乖離 (2007年平均)

(1) 前月 (世帯調査1ヶ月目)

前月の回答サンプル (世帯調査1ヶ月目)	=	前月も今月も 回答した サンプル	+	前月: 回答 今月: 無回答 サンプル
(98.6%)		(98.0%)		(0.6%)

雇用者	49.84%	49.87%	46.48%
自営業主	5.67%	5.70%	1.41%
家族従業者	2.23%	2.25%	0.00%
完全失業者	2.33%	2.33%	2.82%
非労働力人口	39.53%	39.47%	49.30%
その他	0.39%	0.39%	0.00%

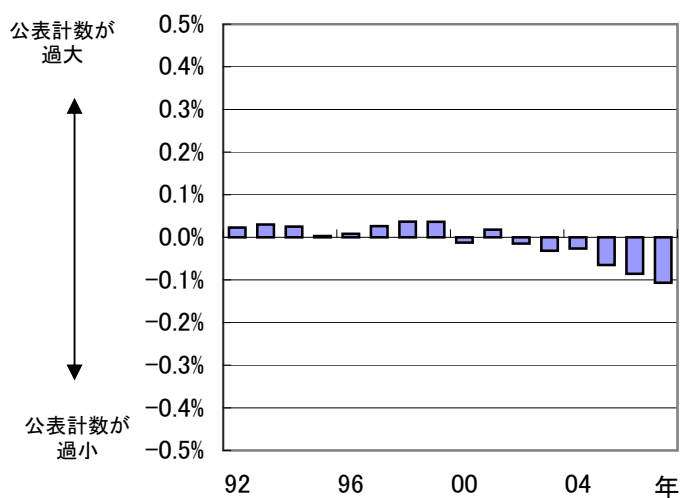
(2) 今月 (世帯調査2ヶ月目)

今月の回答サンプル (世帯調査2ヶ月目)	=	前月も今月も 回答した サンプル	+	前月: 無回答 今月: 回答 サンプル
(100.0%)		(98.0%)		(2.0%)

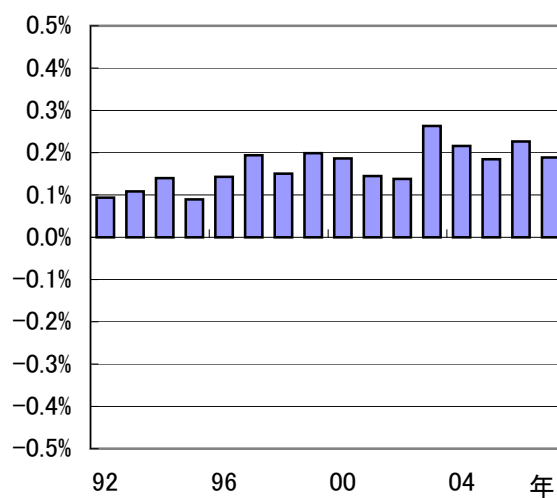
雇用者	49.91%	49.87%	52.23%
自営業主	5.66%	5.70%	3.57%
家族従業者	2.22%	2.25%	0.89%
完全失業者	2.34%	2.33%	2.68%
非労働力人口	39.46%	39.47%	39.29%
その他	0.41%	0.39%	1.34%

無回答によるバイアス

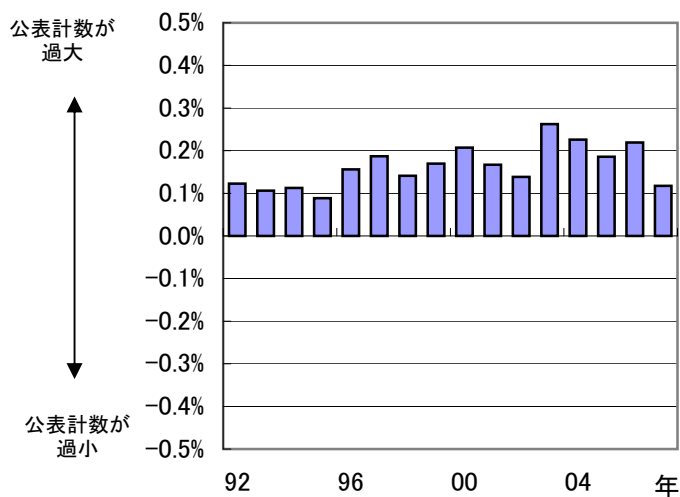
(1) 雇用者



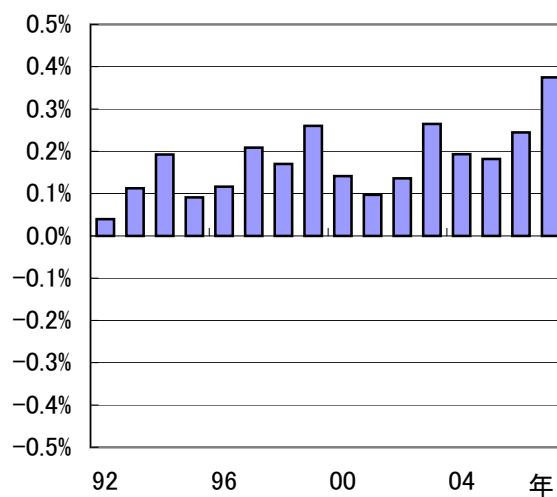
(2) 自営業・家族従業者



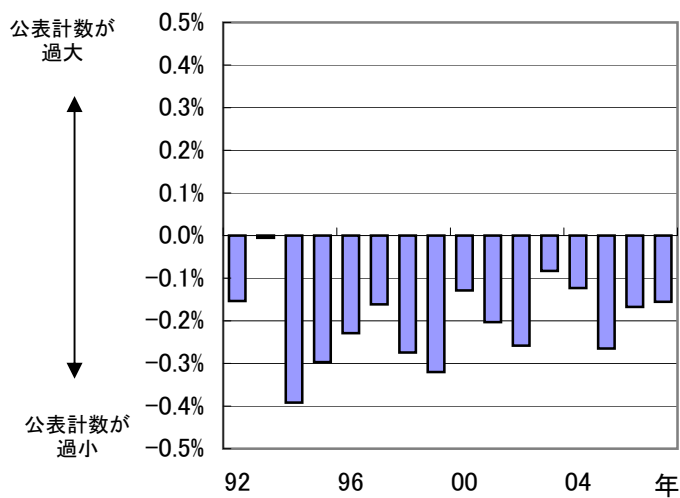
(2) - (a) 自営業主



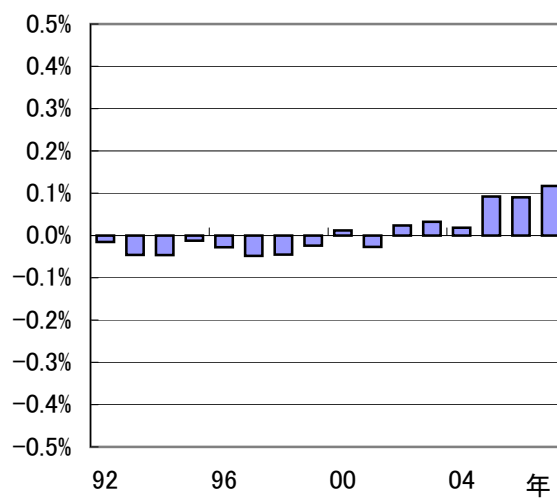
(2) - (b) 家族従業者



(3) 完全失業者

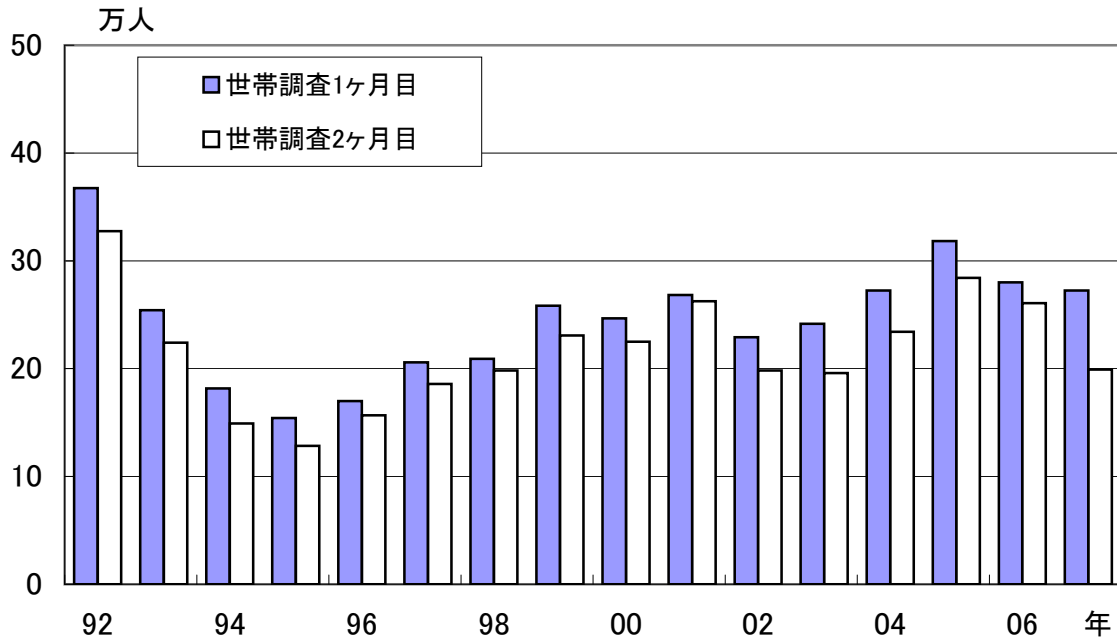


(4) 非労働力人口

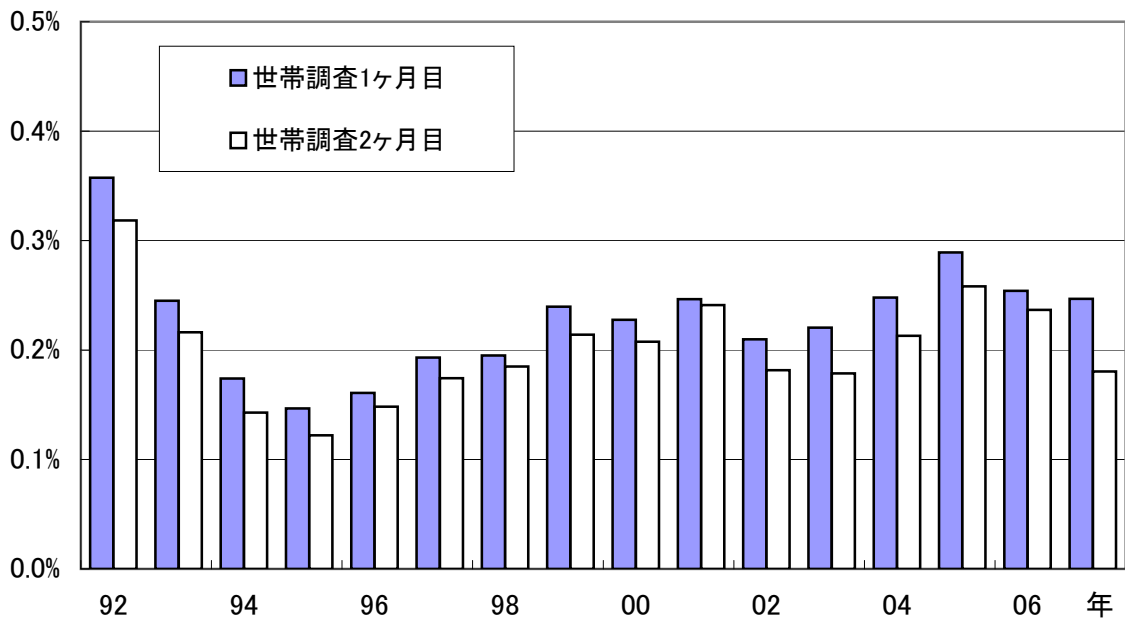


「就業状態ならびに従業上の地位不詳」の推移

(1) 実数 (母集団推計値、年平均)



(2) 15歳以上人口に占める比率 (年平均)



就業状態ならびに従業上の地位不詳サンプルと 判明サンプルとの乖離 (2007年平均)

(1) 前月 (世帯調査1ヶ月目)

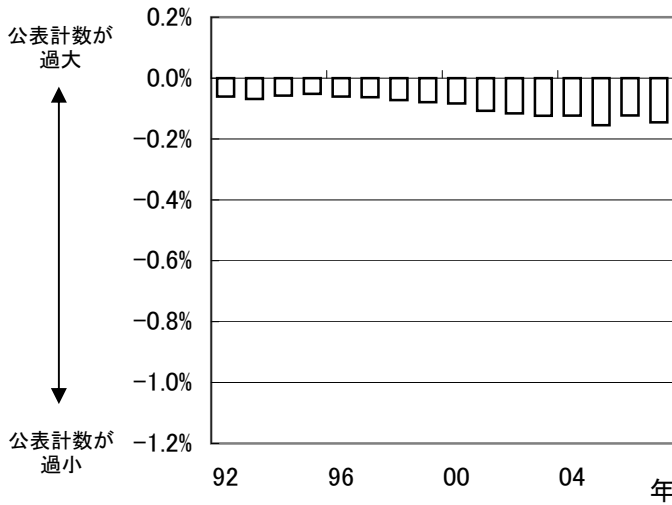
前月、就業状態、 従業上の地位が 判明している サンプル (99.9%)	=	前月も今月も 就業状態、 従業上の地位が 判明している サンプル (99.8%)	+	前月:判明 今月:不詳 サンプル (0.1%)
雇用者	50.05%			29.8%
自営業主	5.68%			10.7%
家族従業者	2.22%			0.8%
完全失業者	2.34%			14.9%
非労働力人口	39.57%			43.8%

(2) 今月 (世帯調査2ヶ月目)

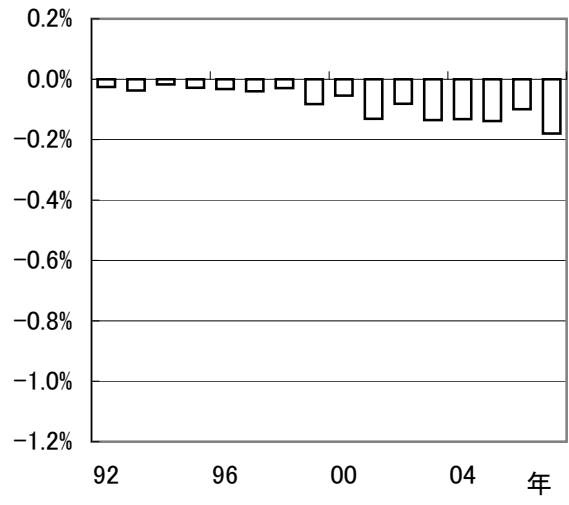
今月、就業状態、 従業上の地位が 判明している サンプル (100.0%)	=	前月も今月も 就業状態、 従業上の地位が 判明している サンプル (99.8%)	+	前月:不詳 今月:判明 サンプル (0.2%)
雇用者	50.15%			53.0%
自営業主	5.68%			5.7%
家族従業者	2.23%			2.0%
完全失業者	2.35%			7.4%
非労働力人口	39.60%			30.1%

就業状態ならびに従業上の地位不詳によるバイアス

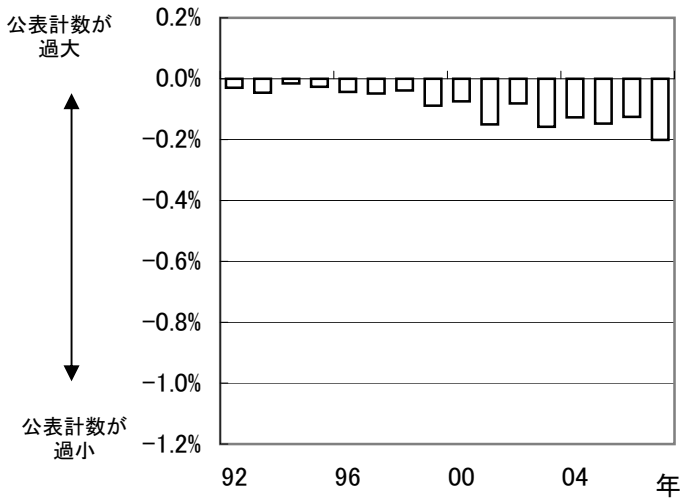
(1) 雇用者



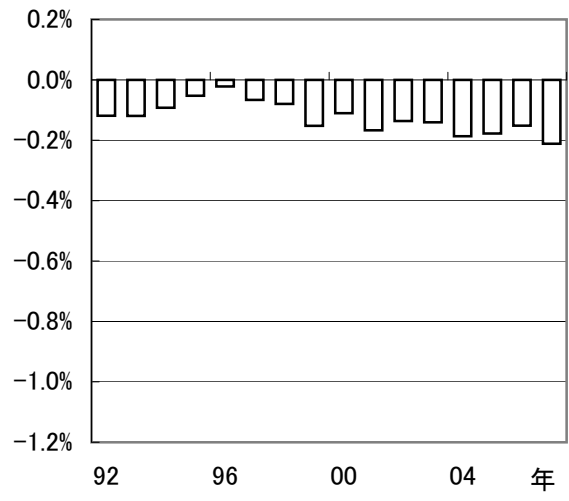
(2) 自営業・家族従業者



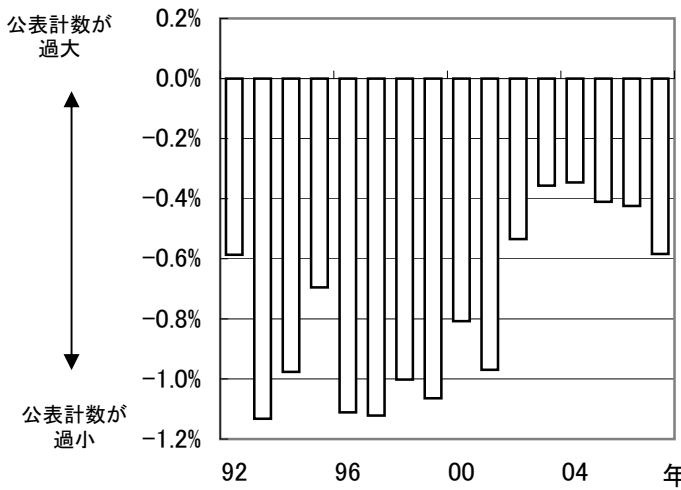
(2) - (a) 自営業主



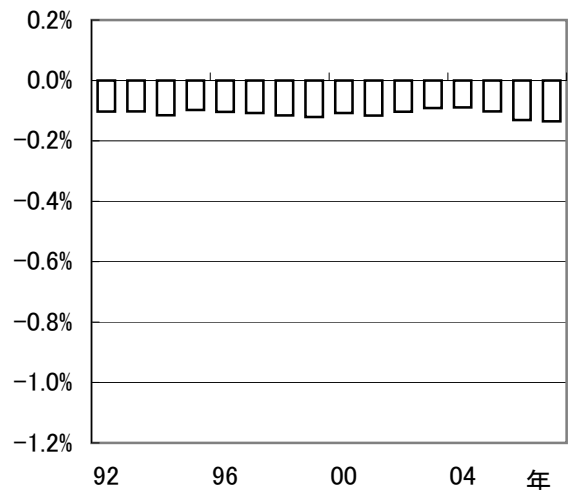
(2) - (b) 家族従業者



(3) 完全失業者

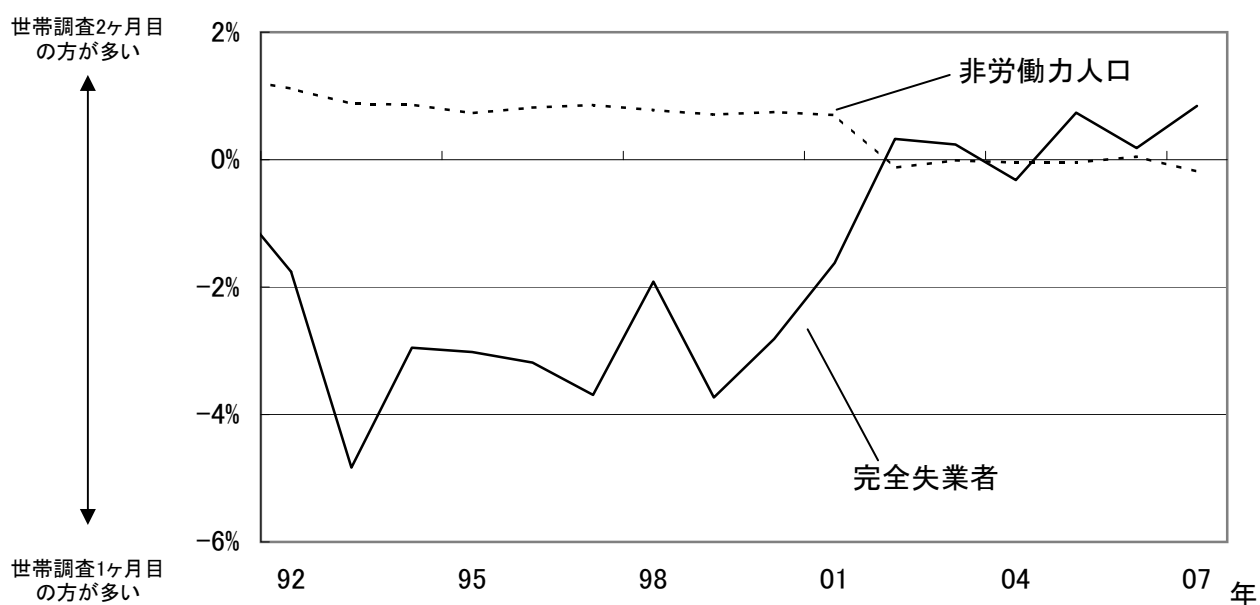


(4) 非労働力人口

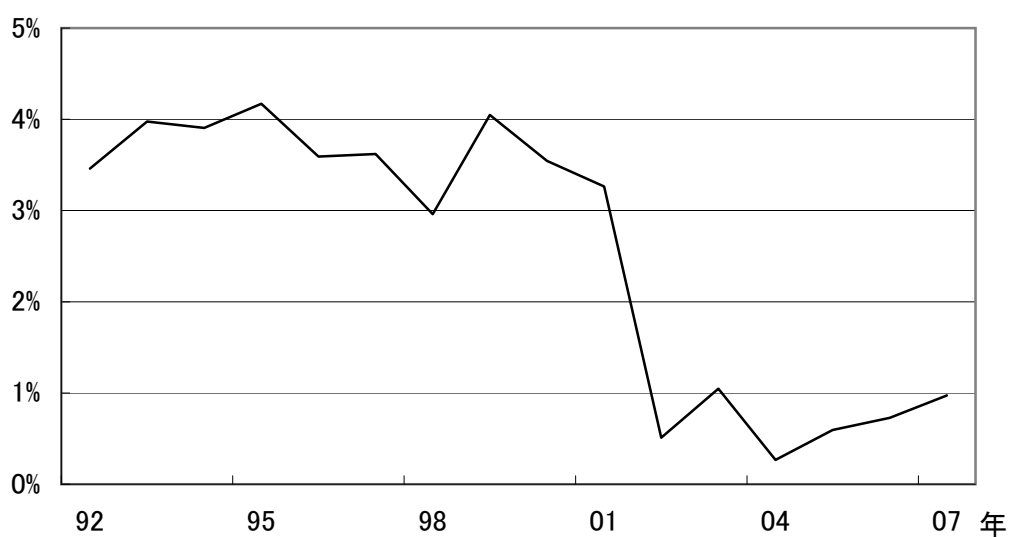


完全失業者と非労働力人口の区分に関する誤解に伴う乖離

(1) 完全失業者と非労働力人口における
継続サンプル（世帯調査2ヶ月目）と非継続サンプル（世帯調査1ヶ月目）の格差

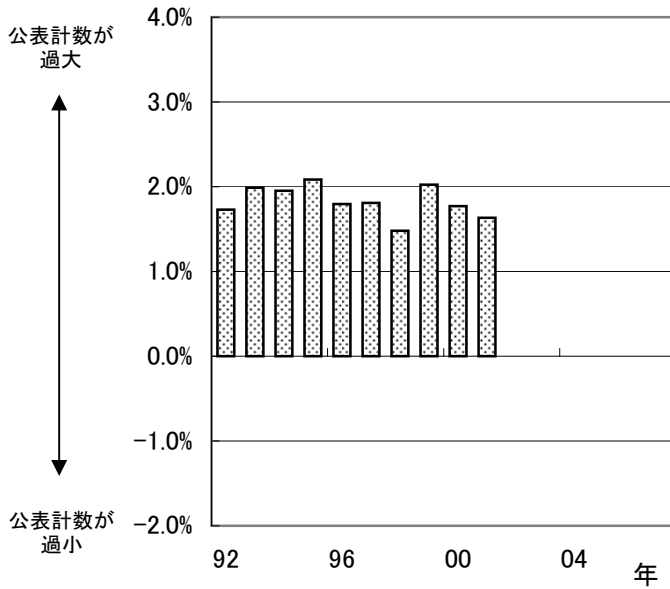


(2) 「完全失業者から非労働力人口へ」のフローと
「非労働力人口から完全失業者へ」のフローの差（完全失業者に対する比率）

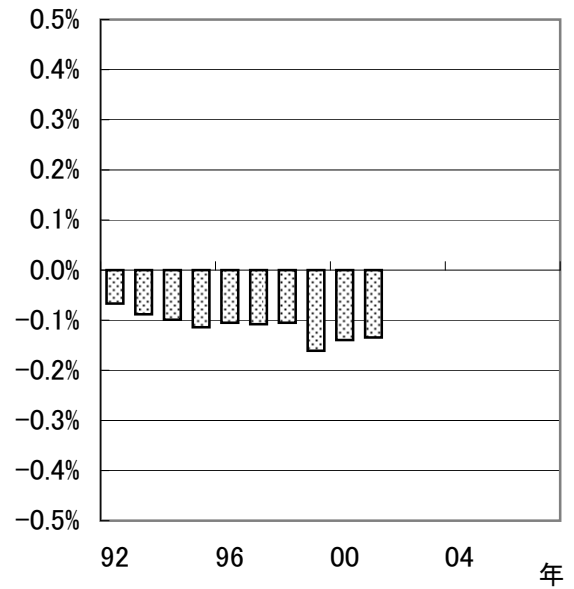


完全失業者と非労働力人口の区分に関する誤解に伴うバイアス

(3) 完全失業者

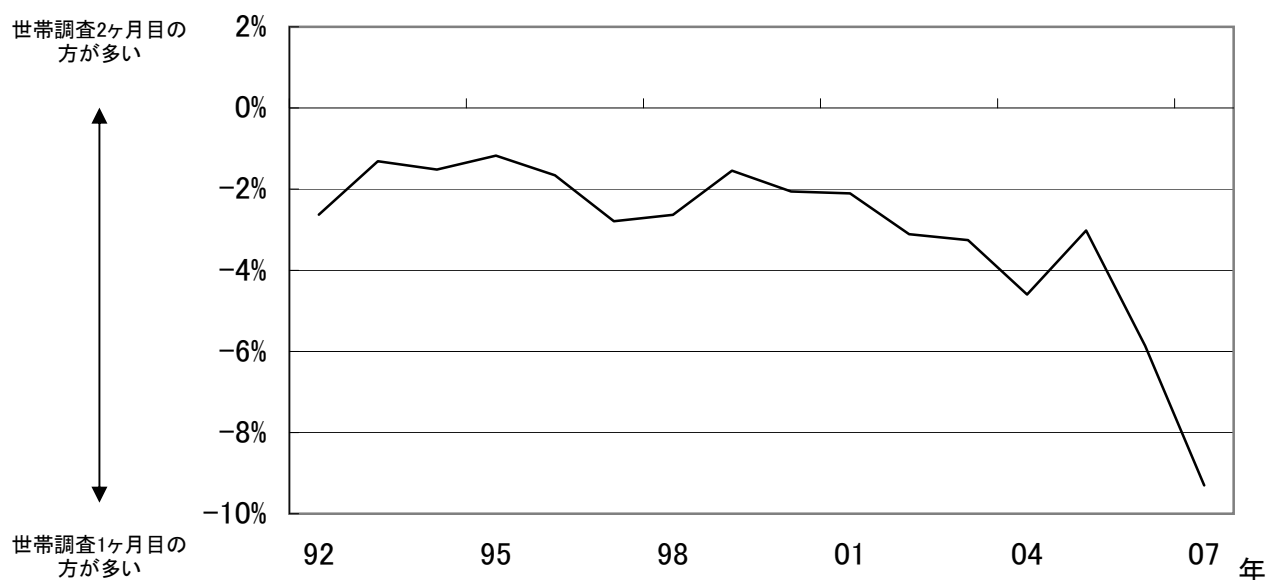


(4) 非労働力人口

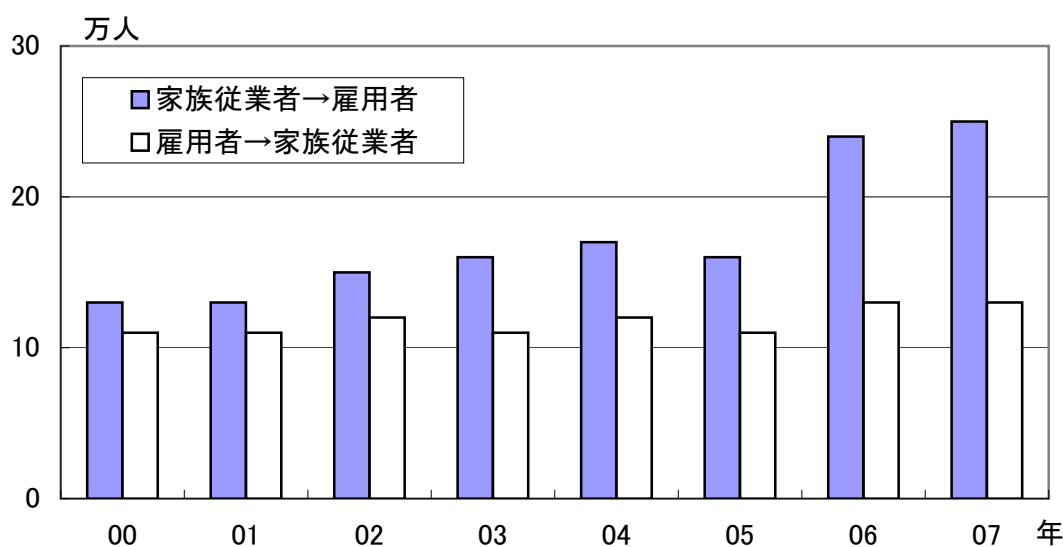


有給の自家営業の手伝いの 家族従業者への混在によるインパクト (1)

(1) 家族従業者における
継続サンプル（世帯調査2ヶ月目）と非継続サンプル（世帯調査1ヶ月目）の格差

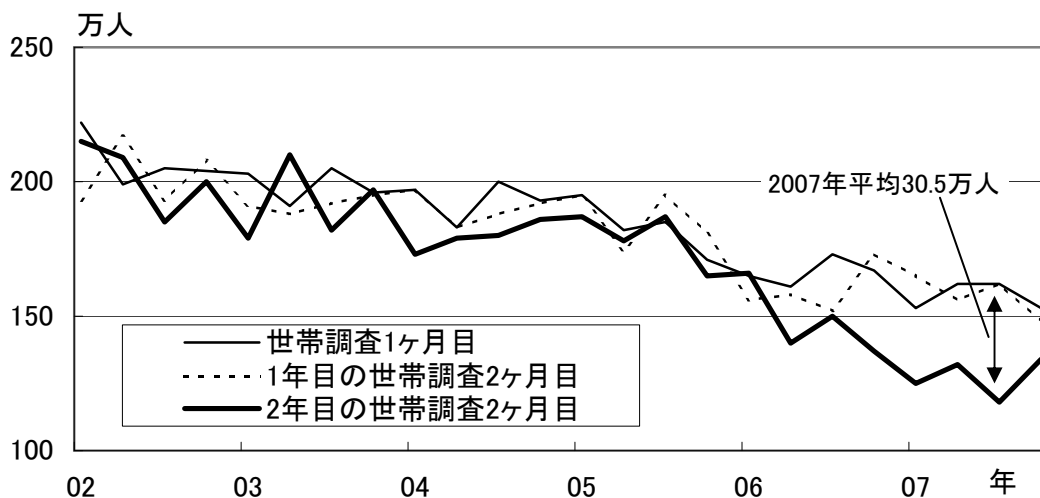


(2) フローデータの動き
（「家族従業者→雇用者」と「雇用者→家族従業者」、世帯調査1ヶ月目→2ヶ月目）

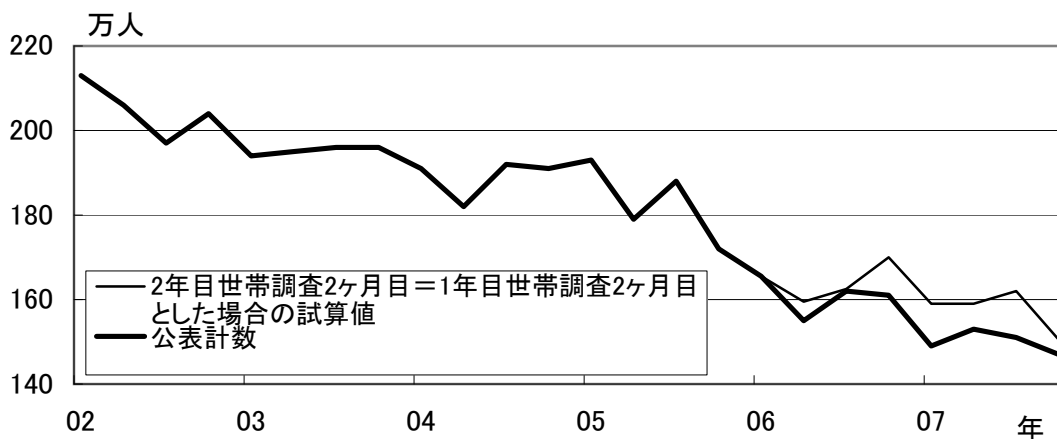


有給の自家営業の手伝いの 家族従業者への混在によるインパクト (2)

(1) 各副標本の非農林業家族従業者数



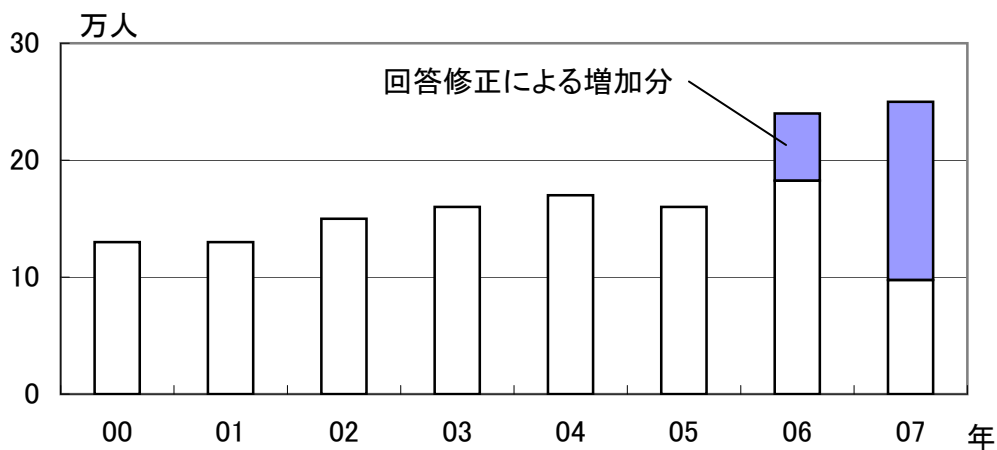
(2) 公表計数 (非農林業家族従業者) へのインパクト



*2006年1Q以降、回答修正が行われたものとして計算。

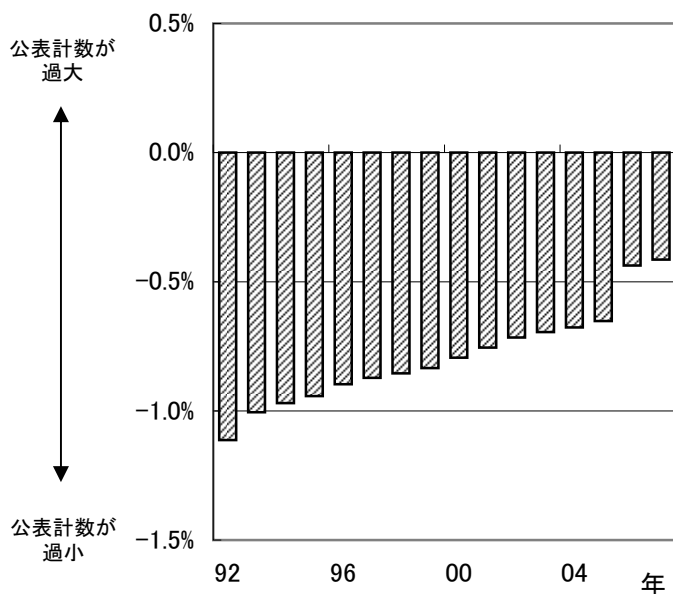
(3) フローデータ

(家族従業者→雇用者、世帯調査1ヶ月目→2ヶ月目) へのインパクト

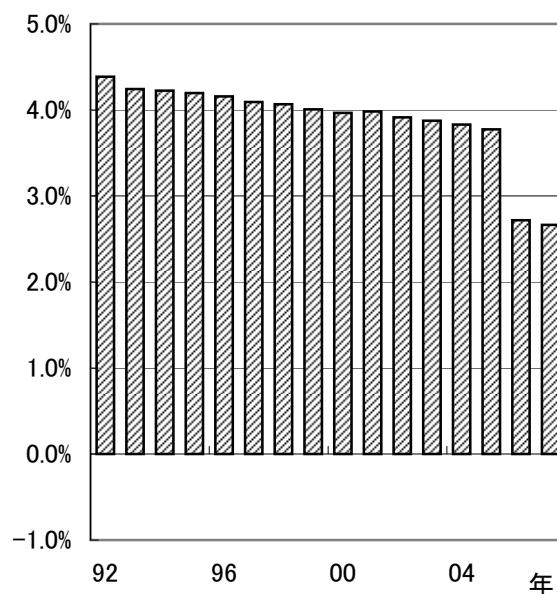


有給の自家営業の手伝いの 家族従業者への混在によるバイアス

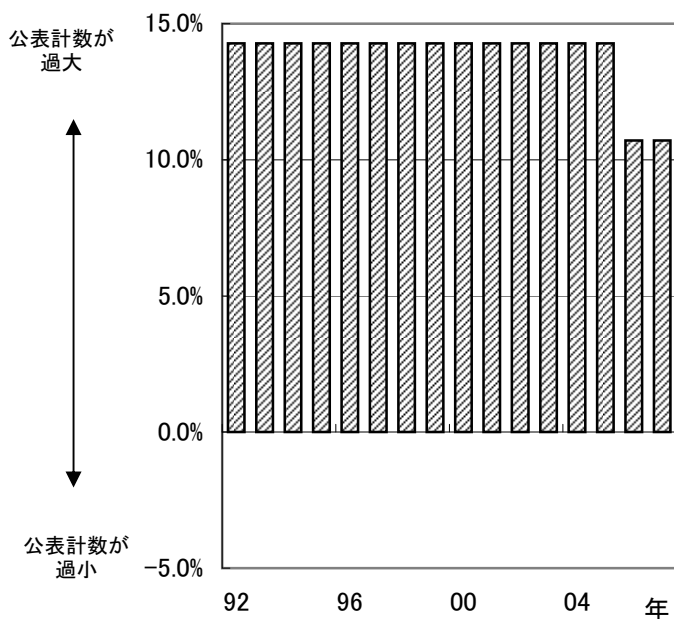
(1) 雇用者



(2) 自営業・家族従業者

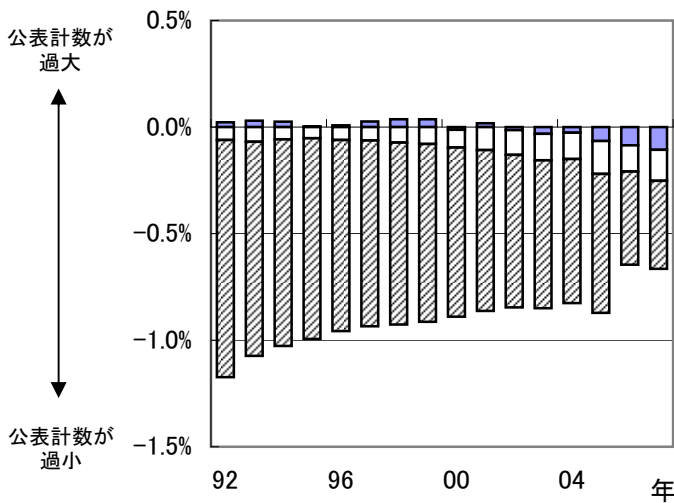


(2) - (b) 家族従業者

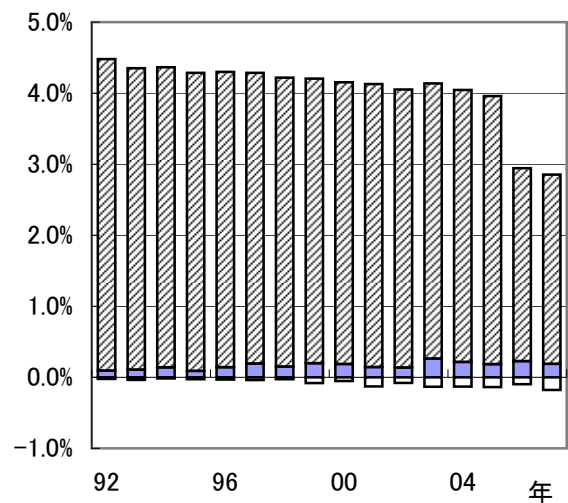


非標本誤差が各公表計数へ与えるバイアス：合計値

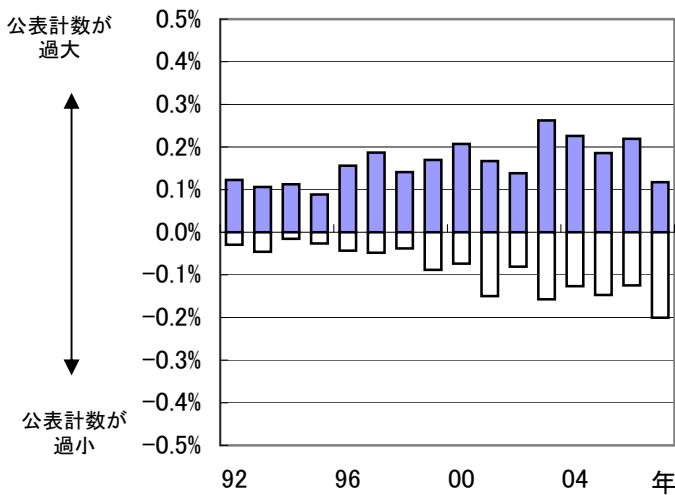
(1) 雇用者



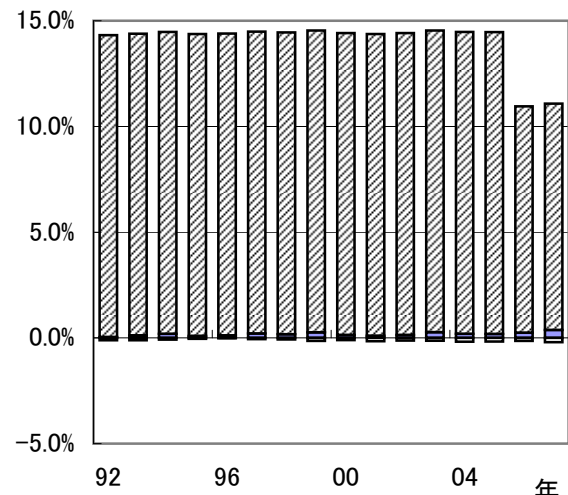
(2) 自営業・家族従業者



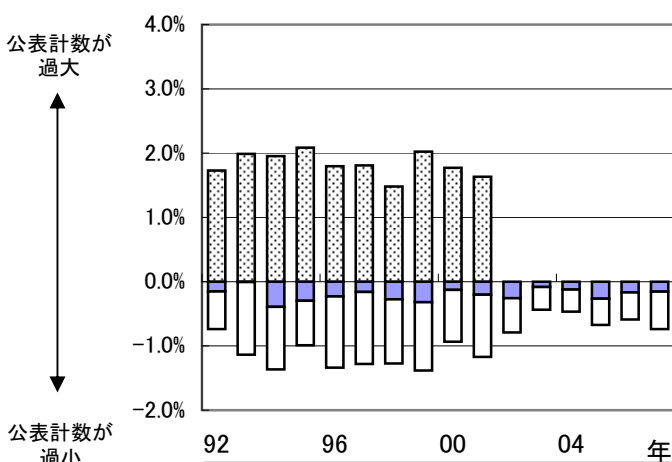
(2) - (a) 自営業主



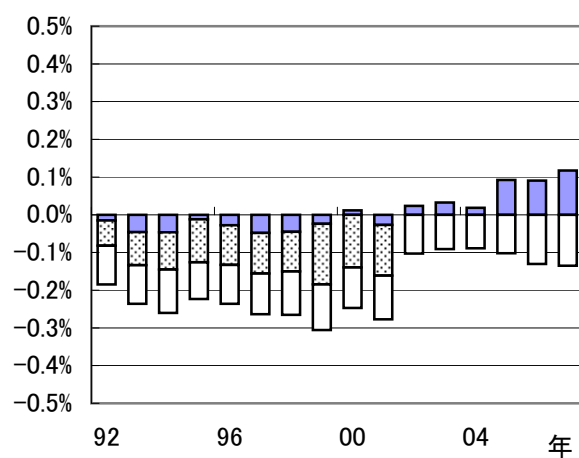
(2) - (b) 家族従業者



(3) 完全失業者



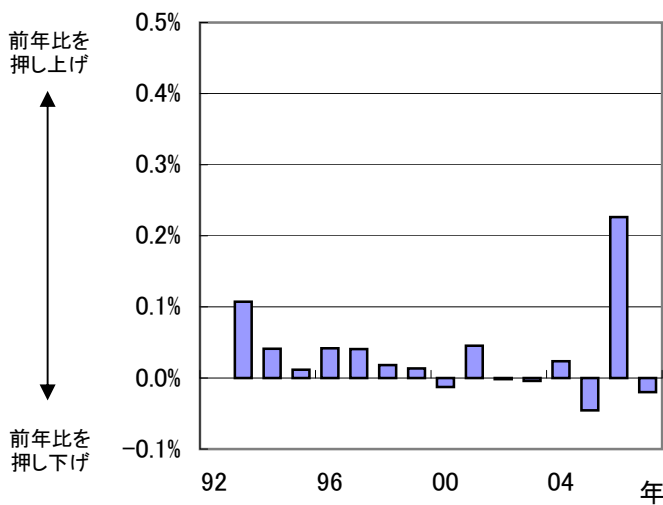
(4) 非労働力人口



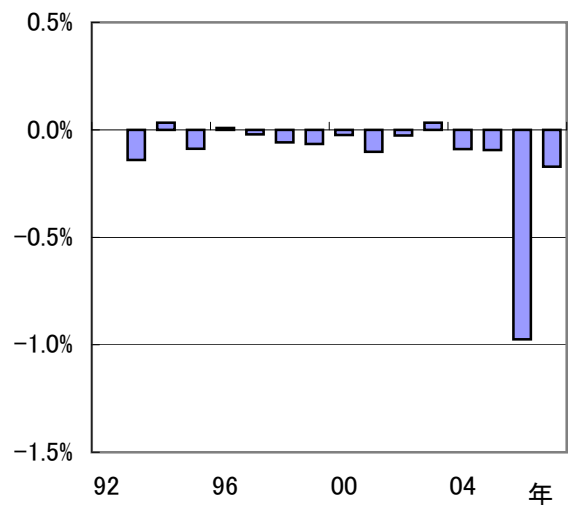
■ 有給の自家営業の手伝いの家族従業者への混在による非標本誤差
□ 就業状態・従業上の地位不詳による非標本誤差
▨ 完全失業者と非労働力人口の区分に関する誤解に伴う非標本誤差
■ 無回答による非標本誤差

非標本誤差が各公表計数の前年比に与える影響

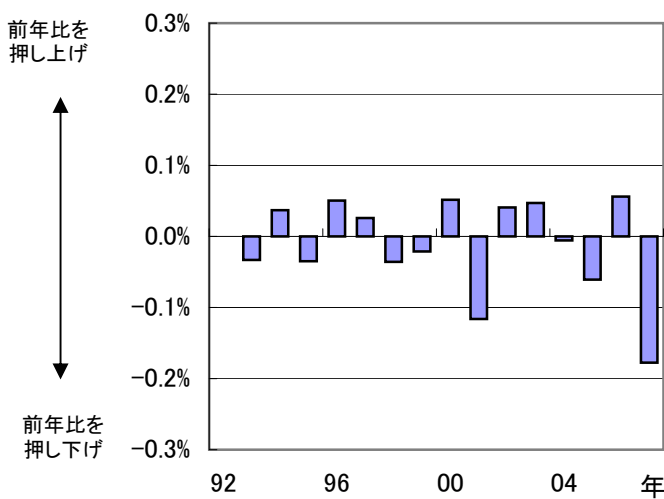
(1) 雇用者



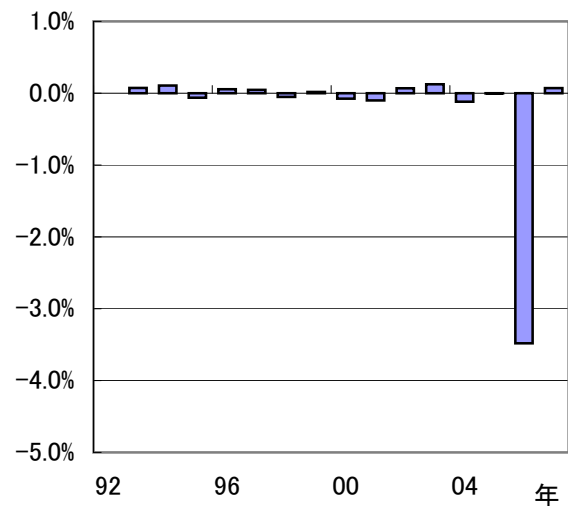
(2) 自営業・家族従業者



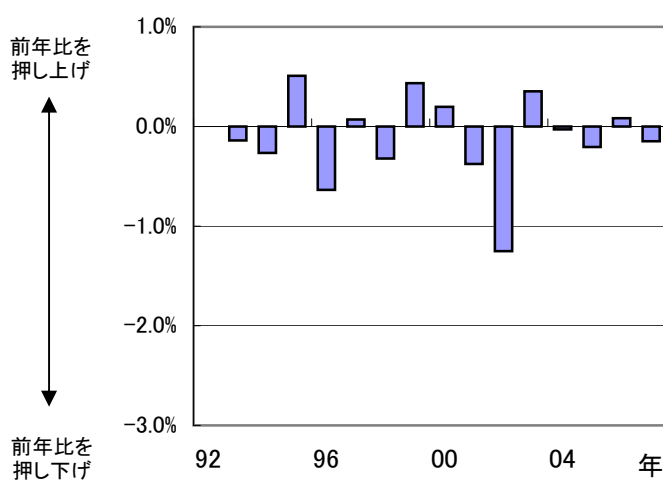
(2) - (a) 自営業主



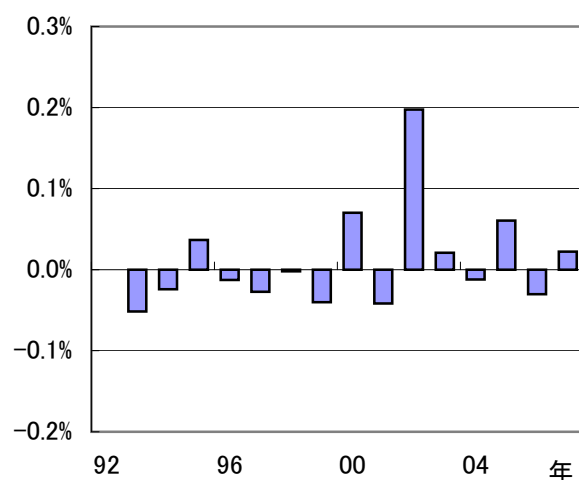
(2) - (b) 家族従業者



(3) 完全失業者

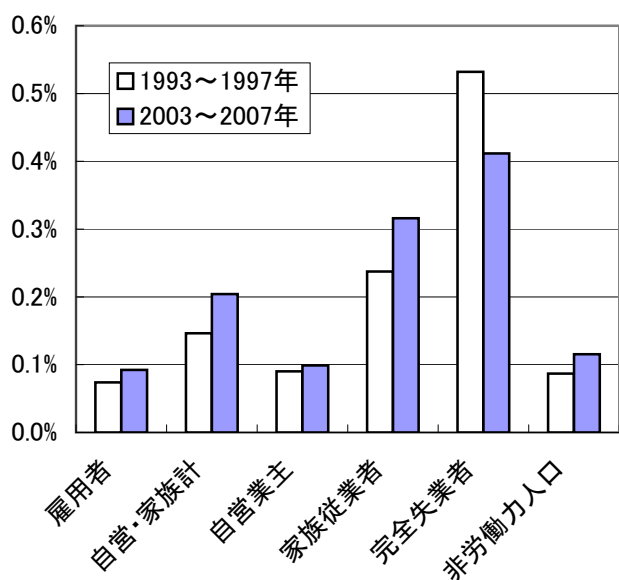


(4) 非労働力人口

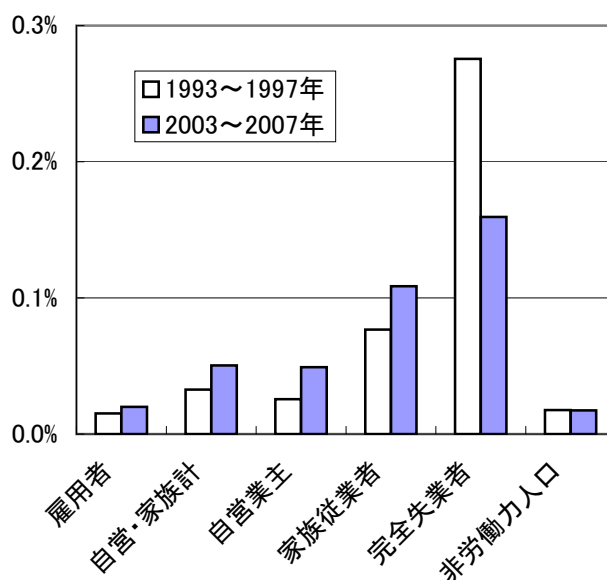


非標本誤差をもたらす変動 (標準誤差率)

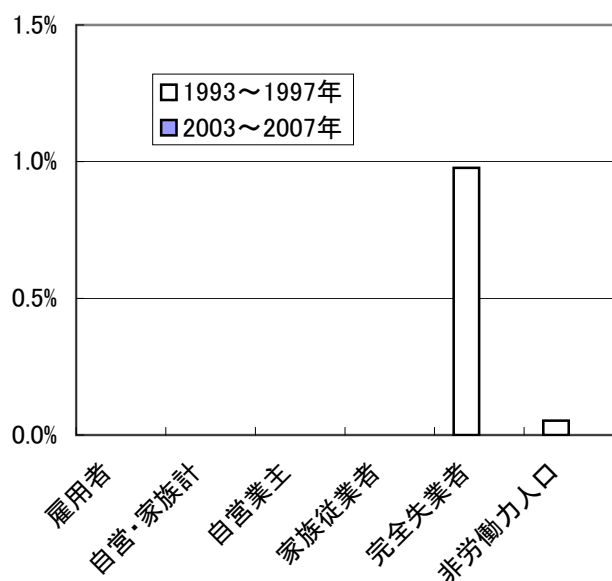
(1) 無回答による非標本誤差の変動



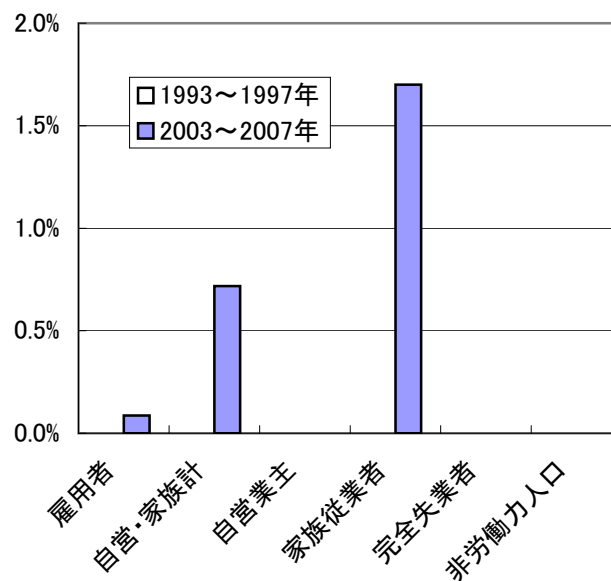
(2) 就業状態・従業上の地位不詳による非標本誤差の変動



(3) 完全失業者と非労働力人口の区分に関する誤解に伴う非標本誤差

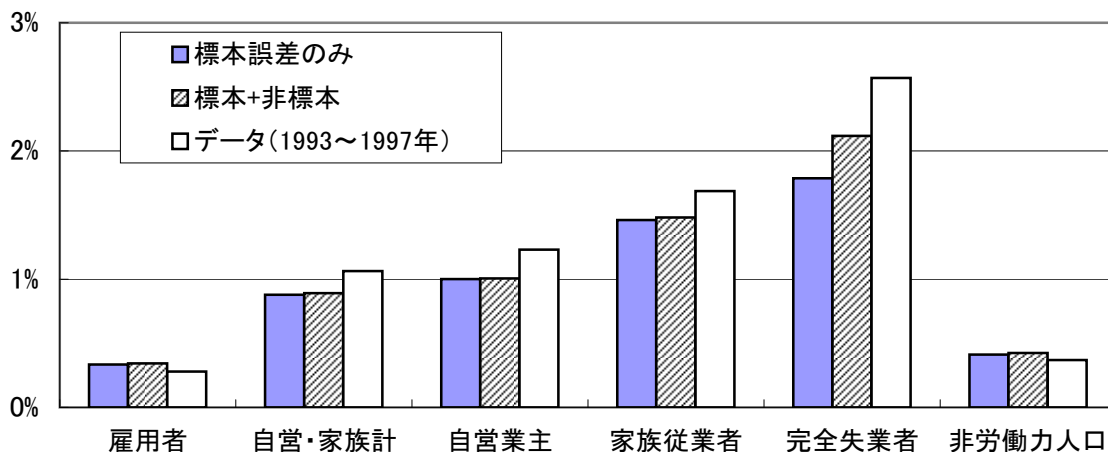


(4) 有給の家族従業者の過大計上による非標本誤差

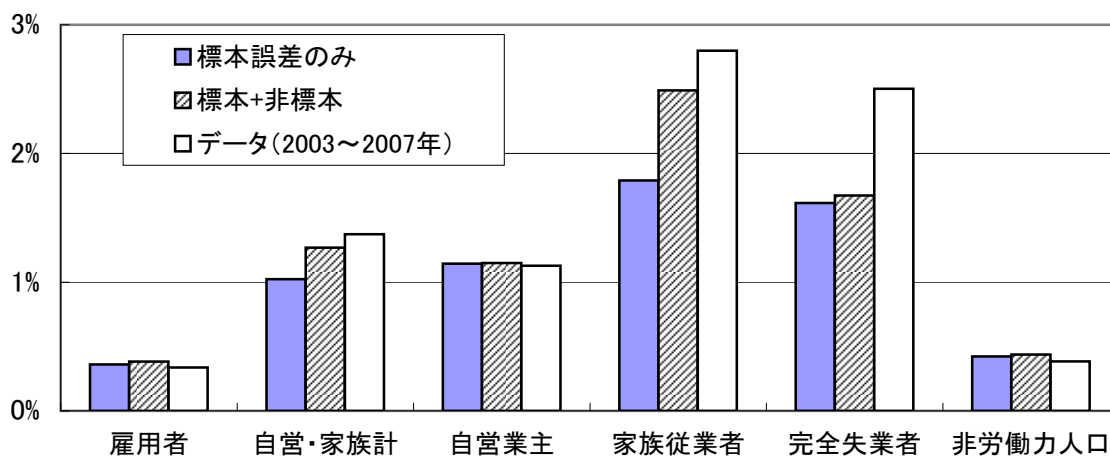


標本誤差・非標本誤差による変動と実データとの比較

(1) 標本誤差・非標本誤差に伴う変動(1995年)と
実データの対数前月差の標準偏差(1993~1997年)



(2) 標本誤差・非標本誤差に伴う変動(2005年)と
実データの対数前月差の標準偏差(2003~2007年)



(3) 1995年から2005年への変化

