

公的統計の現代的意義
並びに作成技法及び利用の高度化
に関する研究

清水 誠

目次

はじめに	1
第 I 部 公的統計の現代的意義	
第 1 章 日本における公的統計の現代的意義の獲得と促進	3
1 公的統計の種類	4
2 公的統計の現代的意義	7
3 公的統計の歴史	11
4 公的統計の特徴	14
5 利用上の留意点	18
5.1 収集データの範囲 —知りたい事項は収集データの範囲内か—	19
5.2 概念の区別 —同じ用語でも統計によって意味が異なる—	21
5.3 作成方法の特徴 —作り方の性格を理解する—	23
5.4 作成時期の区別 —同一年月のデータは変化する—	25
5.5 誤差の取扱 —誤差の情報も利用する—	27
5.6 高度な利用に向けて	29
6 公的統計の課題	30
第 2 章 公的統計の国際基準の意義と変化	31
1 背景	32
2 統計の作成・提供方法に関する基準	33
2.1 公的統計の基本原則	33
2.2 品質	34
2.3 データ収集方法	35
2.4 行政記録	36
2.5 統計の基準年等の改定	37
2.6 分類	37
2.7 季節調整	40
2.8 2次利用	40
2.9 データエディティング	41
2.10 統計提供	41
3 統計結果に関する基準	43
3.1 持続可能な開発目標	43

3.2	主要経済・金融指標	44
3.3	その他の指標	45
4	今後の見通し	45

第Ⅱ部 統計作成技法の高度化研究

第1章	季節調整法の研究	46
1	X11-ARIMAの適用に向けた検討	47
1.1	推定値と改定値の確定値からの差による分析	48
1.2	SSAPによる適用結果の比較	49
1.3	CPI総合の季節指数の変化の比較	50
2	総務省統計局における時系列データの季節性を踏まえた検討	50
3	個人企業経済調査への季節調整法の適用に向けた検討	51
3.1	個人企業経済調査の結果の季節性	52
3.1.1	業況	52
3.1.2	売上高	54
3.1.3	営業利益	56
3.1.4	季節調整により把握される総括的特徴	58
3.2	個人企業経済調査の結果による季節調整法の比較	59
3.2.1	比較の指標	59
3.2.2	比較の結果	61
3.2.3	個人企業経済調査の季節調整値の導入に向けて	67
3.3	他の調査結果との比較	67
3.3.1	総務省統計局の月次調査との比較	67
3.3.2	他府省等の景況・経理調査との比較	68
4	X13-ARIMA-SEATSの適用に向けた検討	69
4.1	X13-ARIMA-SEATSの概要	70
4.2	X12-ARIMAとX13-ARIMA-SEATSの比較	71
4.3	全体評価	72
第2章	国勢調査の結果によるJGSSの結果の改善	72
1	海外の状況	74
2	比推定乗率の意義	75
3	比推定乗率の作成方法	75
4	比推定の方法	76

4.1	推定式	76
4.2	推定結果の比較	77
4.3	標準誤差	78
5	その後の対応	80
第Ⅲ部 統計利用の高度化研究		
第1章	人口が経済に及ぼす影響	82
1	時系列分析	83
1.1	人口と労働時間に着目した供給面の評価	83
1.2	人口要因と資産要因の区分	88
2	地域間比較	94
2.1	先進国間比較	94
2.2	都道府県間比較	97
3	定量的評価の留意点及び課題	101
第2章	所得が出生に及ぼす影響	102
1	時系列データの適用	103
1.1	1971年から2000年における Butz and Ward Model の適合状況	103
1.2	1975年から2004年における Butz and Ward Model の適合状況	103
2	JGSS-2002における夫婦データを用いた検証	104
2.1	検証の方法	105
2.2	使用するデータセット	105
2.3	所得・出生関数の作成方法	106
2.4	生涯レベルの推定方法	108
2.5	結果の評価	109
第3章	科学技術研究調査による企業の生産性の要因分析	110
1	企業の研究投資効果	111
1.1	研究費と売上高の関係	111
1.2	対前年度変化率による階級区分	112
1.2.1	全産業の研究費と売上高の関係	112
1.2.2	医薬品工業のケース	113
1.2.3	鉄鋼業のケース	113
2	生産性の要因分析	114

2.1	生産関数の構築	117
2.2	データの特徴	118
2.3	分析に用いた生産関数	119
2.4	生産性の要因分析	120
2.5	研究投資の効果とタイムラグ	127
2.6	生産性の要因分析に関する結論と課題	128

第IV部 消費者物価指数の意義、作成技法と利用の高度化

第1章	ボスキンレポートをめぐる議論とCPIへの影響	131
1	ボスキンレポートの背景	131
2	ボスキンレポートの概要	133
2.1	米国CPIにおける上方バイアス	133
2.2	CPIバイアスの財政への影響	134
2.3	勧告	134
2.4	BLSの対応	135
2.4.1	指数の目的と算式の選択	135
2.4.2	その他のバイアス	136
3	上位代替バイアス	137
3.1	ボスキンレポートの主張	137
3.1.1	生計費指数の作成	137
3.1.2	最良指数の採用	138
3.1.3	上位代替バイアスの評価	138
3.2	BLSの主張	139
3.2.1	生計費指数の作成について	139
3.2.2	最良指数の採用について	139
3.2.3	上位代替バイアスの評価について	140
3.3	指数算式に関する理論と見解	140
3.3.1	生計費指数の背景	140
3.3.2	生計費指数の問題点	141
3.3.3	各種指数の算式の比較	142
3.4	生計費指数と指数算式の留意点	144
4	下位代替バイアス	145
4.1	ボスキンレポートの主張	145
4.2	BLSの主張	146

4.3	幾何平均指数の作成状況	146
4.4	下位集計レベルの指数算式の採用条件	147
5	新店舗バイアス	149
5.1	ボスキンレポートの主張	149
5.2	BLSの主張	150
5.3	新店舗の価格反映の留意点	151
6	品質・新製品バイアスをめぐる議論	152
6.1	ボスキンレポートの主張	153
6.1.1	品質・新製品バイアスの評価	153
6.1.2	品質・新製品バイアスの縮小方策の提言	155
6.2	BLSの主張	156
6.2.1	品質・新製品バイアスの評価の妥当性	156
6.2.2	品質・新製品バイアスの縮小方策	158
6.3	品質バイアスをめぐる議論	159
6.3.1	ヘドニック回帰法について	159
6.3.2	コスト推定法について	161
6.3.3	リンク法及び直接比較法について	162
6.4	品質調整法の比較	163
7	新製品バイアスをめぐる議論	164
7.1	消費者余剰について	164
7.2	新製品のCPIへの採用の遅れについて	165
7.3	新製品バイアスに関する留意点	166
8	ボスキンレポートの妥当性	167
8.1	CPIの概念的枠組み	167
8.2	上位集計レベルの指数算式	168
8.3	下位集計レベルの指数算式	169
8.4	新店舗の取扱	170
8.5	品質調整法	170
8.6	新製品の取込み	171
9	21世紀におけるボスキンレポートの影響	171
第2章 日本のCPIの作成技法の高度化		173
1	総世帯ベース月次指数の作成	174
1.1	変更による影響	174

1.2	月次指数の計算方法	175
1.3	指数の接続	175
2	上位指数算式の検討	175
2.1	インフレ時及びデフレ時の実データの比較	176
2.2	適用すべき指数算式	178
3	月次連鎖指数の作成	179
3.1	月次連鎖指数の作成方法	180
3.2	月次連鎖指数の評価方法	181
3.3	2005年基準改定以降の状況	182
4	コア指数の作成	182
4.1	コア指数の作成状況	183
4.1.1	特定品目を除く方法	184
4.1.2	統計的処理を施す方法	185
4.2	コア指数の条件	186
4.3	日本の状況	187
4.4	「食料（酒類を除く）及びエネルギーを除く総合」の動向	189
5	新旧指数の差の要因分解	189
5.1	要因分解の考え方	190
5.2	2005年基準改定への適用	191
5.3	2000年基準改定への適用	195
5.4	その他の指数の差について	196
5.5	2010年以降の基準改定	197
6	家賃指数の回帰式による推定方法の検討	198
6.1	家賃指数の問題点	199
6.2	回帰式による品質調整の方法	199
6.2.1	ヘドニック法の適用の条件と方法	199
6.2.2	諸外国の状況	201
6.3	家賃に関する回帰式による推定手法の意味	202
6.4	家賃の回帰分析	203
6.4.1	回帰式の概要	203
6.4.2	タイムダミー法の適用可能性	205
6.5	地域別家賃指数の推定手法の検討	206
6.5.1	転出入の割合	206
6.5.2	地域別家賃指数の推定のための回帰式の利用	206

6.5.3	比較の条件及び手法	207
6.5.4	各手法の家賃推定値	207
6.5.5	各手法の推定家賃の前月比	208
6.6	全体評価と今後の課題	213
7	IMF/ROSC 対応	213
7.1	ROSC とは	214
7.2	ROSC における評価基準	215
7.3	CPI に対する評価の概要	216
7.4	今後の課題	216
第3章	物価の実感に関する定量的評価	218
1	EU の状況	219
2	物価安定下における日本の状況	221
2.1	CPI の動き	222
2.2	物価の実感と CPI	222
2.3	品質調整による影響	223
2.4	購入頻度による影響	225
2.5	直近の物価変動による影響	227
2.6	実感と指数の差の要因	227
2.7	2008 年の状況	228
3	EU と日本の共通点	228
おわりに		229
参考文献		231

はじめに

公的統計は国民が意思決定を行うための情報基盤である。これを第一条に掲げて統計法が全面改正されてから 10 年が経過しようとしている。その間に情報通信技術の進展によりビッグデータの時代を迎え、データサイエンスが芽吹いた。もはや統計は統計家だけのものでなく、誰もが意思決定を行うために不可欠なものになっている。

その間に、日本は、本格的な少子高齢社会を迎え、人口が減少に転じる中で、震災とデフレに対面し、より良い社会、より良い経済を実現するために処方箋を練り続けており、統計はその羅針盤となっている。グローバル化の進展は著しく、世界では、金融危機を乗り越え、安全確保を模索しながら、貧困の縮減と持続可能な開発に向けて、統計指標を目標に掲げて追求している。

日本においても、世界においても、様々な危機や出来事に直面し、必ずしも歓迎できない局面に入ってしまったとしても、客観的な証拠として、常に統計の重要性が増大してきた点は見逃せない。

このような状況を踏まえ、本論文は、公的統計の意義、作成技法の改善について研究し、統計データを利用して社会経済の構造を分析するとともに、それらの成果を包括的に体系化したものである。全体を 4 つの部で構成し、それぞれ第 I 部が意義、第 II 部が作成技法、第 III 部が利用について公的統計全般に貢献し得る研究内容を論じる。公的統計の中には加工度の高い統計など、一般論だけで意義、作成技法、利用を論じることができないものも存在することから、第 IV 部では、そのうち経済の体温計と呼ばれ経済・金融政策にも重要な CPI（消費者物価指数）に焦点を当て、具体的な研究を示す。

まず、第 I 部では、公的統計の現代的意義について、日本における歴史的変遷を踏まえた検討を行うとともに、作成者と利用者の双方の側から公的統計が本来の役割を發揮するために必要な観点を述べる。また、公的統計の意義の国際的な動向について、国際基準の観点から具体的な事例を含めて論じる。

次に、第 II 部では、公的統計の作成技法の高度化に向けた研究を論じる。作成技法は利用方法と不可分であり、大局的には、時系列で分析するための統計の作成技法と一時点の構造を分析するための作成技法に区別される。時系列分析をするための統計は、傾向、季節性、異常値を区別することが基本であることから、各種の時系列統計の季節調整法の高度化研究を論じる。他方、一時点の構造を分析するためには、センサや大規模標本調査の結果を利用することが一般的である。それらに含まれない項目について分析をする場合には、その他のデータ源に頼らざるを得ないが、精度確保についてはリスクが伴うことになる。しかし、そのような場合にも、大規模統計調査

の結果を利用すると、よりの確な分析が可能になる。その典型的な事例として、ここでは、国勢調査の結果による社会調査の結果の改善に係る研究を示す。このような研究はビッグデータの作成技法の高度化にも資する。

第Ⅲ部では、公的統計の利用の高度化に関する研究を論じる。公的統計の利用は、大括りで見ると、時系列分析と一時点の構造分析に区別される。そこで、経済成長と出生率の低下が問題となる状況にかんがみ、経済における人口要因、所得と出生の関係、企業の生産性の要因分析について、時系列と地域別の両面から、実データを各種のモデルに適用することにより、経済や社会に関して得られた知見を示す。

最後の第Ⅳ部では、CPIの意義、作成技法の高度化、利用の高度化全体についての研究を少し詳細に論じる。CPIについては、物価の考え方、指数算式、指数の接続の仕方、価格の把握方法、品質の調整方法など、配慮すべき特有の事項が多数存在し、それぞれ相互に関連していることから、意義、作成技法、利用のいずれについても深い検討が必要だからである。

第 I 部 公的統計の現代的意義

第 I 部では、公的統計の現代的意義について、日本における歴史的変遷を踏まえた研究成果を示すとともに、作成者と利用者の双方の側から公的統計が本来の役割を發揮するために必要な観点を述べる。また、公的統計の意義の国際的な動向について、国際基準の観点から具体的な事例を含めて論じる。

まず、第 1 章では、日本における公的統計の意義の変遷と利用の高度化について述べている。公的統計は、古くから行政を的確に運営する上で基本となる情報として重要な役割を果たしてきたが、近年、国民が社会・経済を正しく把握し、適切に意思決定するための情報基盤を提供するという、公的統計が有する本来の役割の方向に見直されるようになってきた。公的統計が情報基盤として有効に機能するには、公的統計の利用が質、量ともに向上することが必要であるとして、利用における留意点を整理して解説している。

次に、第 2 章では、公的統計に関する国際基準の意義が変化している点を考察している。国際基準は、以前は統計の作成・提供方法の原則を設けることを主としていたが、近年は目標に照らしての統計そのもの、あるいはアウトカムの評価にルールを設ける方向に変わる傾向にある。

第 1 章 日本における公的統計の現代的意義の獲得と促進

公的統計の意義は公共の利益のために必要な情報を提供することである。公的統計は、古くから行政を的確に運営する上で基本となる情報として重要な役割を果たしてきたが、近年、国民が社会・経済を正しく把握し、適切に意思決定するための情報基盤を提供するという公的統計が有する本来の役割の方向に見直されるようになってきた。公的統計は行政だけでなく、社会全体で利用される情報基盤と位置付けられている点が重要である。そうした観点から、公的統計の意義が認識されるようになったのは最近のことであり、その意味で現代的意義と呼ぶにふさわしいものである。本章は、清水誠[1]をもとに、統計法（平成 19 年法律第 53 号）の全面改正等を踏まえて記述している。なお、統計法の全面改正を踏まえた統計組織の体系に関わる記述については、Makoto Shimizu[2]に基づいている。

清水誠[1]が執筆される以前にも、公的統計の現代的意義を主張する論文は存在した。永山貞則[3]は、政府統計の役割について、国の行政目的に資するためだけでなく、国民の必要とする統計情報をサービスすることと明確に位置付けることが必要である

と指摘している。大屋祐雪[4]も、近代社会における政府は国家、国民のための組織であるから、行財政のためだけでなく、企業やその他の社会構成員のためにも、政治、経済、社会の現状と動向について、定期的に、あるいは臨時的に、各種の統計を作成し、社会に提供しなければならないと指摘している。竹内啓[5]も、民主主義国家において統計の果たすべき最も重要な役割は、国の社会・経済の実情を正しく反映している情報を提供し、主権者としての国民が正しい判断を下すための基礎データを提供することであると指摘している。

公的統計が現代的意義を満たし、本来の役割を果たすためには、社会・経済の構造や動向を的確に捉えることに資するよう、公的統計の利用が質、量ともに向上することが必要であると考え。このため、清水誠[1]では、公的統計の現代的意義を述べるにとどまらず、公的統計の利用のメリットと適切な利用の方法を体系的かつ具体的に述べている。すなわち、公的統計は、その性格をよく理解して正しく利用すれば非常に有益な情報が得られるとし、利用に際しては、データを客観的に扱うという基本的な視点に立って、①統計データの体系の中での位置付けを認識し、②概念を明確にし、③作成方法の特徴を理解し、④作成時期を確認し、⑤誤差の情報を利用することが必要であると指摘している。

1 公的統計の種類

統計は政府、企業、個人が意思決定し、行動する上で不可欠なものである。とりわけ公的統計の果たす役割は大きく、公的統計は社会・経済の様々な局面で活用されている。公的統計は、統計法において、行政機関、地方公共団体又は独立行政法人等が作成する統計と規定している。公的統計は、統計作成のための情報源と手法によって大きく3つに区分される。1つは、統計の作成を単一又は主目的として、調査票に回答の記入を求める実地調査、すなわち、統計調査と呼ばれる方法を経て、その結果を集計するもので調査統計と称される。もう1つは、行政機関が指揮命令系統や法律に基づく権限を通して、又は政策上の要請により、統計以外の目的のために収集・保有する各種データを集計するもので業務統計と称される。もう1つは、調査統計や業務統計の各種のデータに推計・加工を行って作成するもので、加工統計と称される。

ここでいう「行政機関」とは、法律の規定に基づき内閣に置かれる機関若しくは内閣の所轄の下に置かれる機関、宮内庁、内閣府設置法（平成11年法律第89号）第49条第1項若しくは第2項に規定する機関又は国家行政組織法（昭和23年法律第120号）第3条第2項に規定する機関をいう。すなわち、府、省、委員会及び庁のほか、地方支分部局、審議会等が含まれる。他方、国の機関であっても、国会、裁判所、会計検査院などは除かれる。また、「独立行政法人等」とは、独立行政法人、政府関係機

関 1、法律により直接設立された法人等（国立大学法人等）である。一方、「地方公共団体」とは、地方自治法（昭和 22 年法律第 67 号）第 1 条の 2 に基づく都道府県、市町村、特別区、地方公共団体の組合、財産区及び地方開発事業団である。また、「作成する」とは「企画する」ことであり、調査の実施や集計が民間に委託される場合であっても、作成主体が行政機関であれば公的統計である。

国の行政機関が作成する統計のうち、重要度が特に高い統計を基幹統計と呼ぶ。

具体的には、統計法で規定されている国勢調査による国勢統計及び国民経済計算の他、次のいずれかに該当するものとして総務大臣が指定するものである。

- ① 全国的な政策を企画立案し、又はこれを実施する上において特に重要な統計
- ② 民間における意思決定又は研究活動のために広く利用されると見込まれる統計
- ③ 国際条約又は国際機関が作成する計画において作成が求められている統計その他国際比較を行う上において特に重要な統計

上記に該当する基幹統計以外の公的統計は一般統計と称される。

公的統計の統計調査について、基幹統計を作成する統計調査を基幹統計調査、一般統計を作成する統計調査を一般統計調査という。

基幹統計調査には、対象者に報告を求めることができ、報告を求められた者は、これを拒み、又は虚偽の報告をしてはならないとされている。その場合、報告を拒み又は虚偽の報告をした者には 50 万円以下の罰金に処するとされている。また、基幹統計調査であると誤認させる表示、説明をすること、いわゆるかたり調査により、情報を取得する行為を未遂も含め罰則付きで禁止している。これに違反すると 2 年以下の懲役又は 100 万円以下の罰金に処するとされている。

2016 年 11 月現在、基幹統計に指定された統計は 56²あるが、この中の大半は国勢調査、経済センサス、農林業センサスなど、統計調査により作成される調査統計である。

また、基幹統計調査の中で、国勢調査、住宅・土地統計調査など世帯又は住戸のみを対象とするものは 8³あり、このうち国民生活基礎調査以外は、国勢の基本となる統計を所管する総務省統計局が実施している。諸外国においても、この種の世帯調査は、国の中立を保つ機関が、居住者を対象に報告義務を課する統計調査として実施している場合が多い。

¹ 日本銀行、（株）日本政策金融公庫、国家公務員共済組合連合会等

² 基幹統計に指定された統計の中で、人口推計については、2017 年 4 月以降に公表される結果から基幹統計として扱われる。

³ 国勢調査、住宅・土地統計調査、労働力調査、家計調査、就業構造基本調査、全国消費実態調査、社会生活基本調査、国民生活基礎調査

一方、経済センサス、法人企業統計調査など事業所（港湾及び船舶を含む。）又は企業のみを対象とするものは 33 もあり、このうち、すべての産業を対象として、事業所・企業活動を概括的に把握するものが 5⁴、事業所ベースで雇用・賃金を把握するものが 3⁵、土地の実態を把握するものが 1（法人・土地基本調査）あるが、残り 24 は、特定産業の活動状況を把握するもので、製造業を対象とするものが多い。

調査対象をすべて調べる調査を全数調査又はセンサス、母集団の一部の構成単位について調査し、それに基づいて母集団全体について推計する方法を標本調査と呼ぶ。両者を組み合わせて統計体系を整備しているのが通例であり、製造業を例にとれば、まず、工業統計調査で製造業の基本的構造を捉え、生産動態統計調査で産業別に毎月の生産動向を調査している。

業務統計は統計作成が副次的であるので、データの入手の方法と根拠は区々である。例えば、基幹統計の人口動態統計は、戸籍法（昭和 22 年法律第 224 号）に基づいて行われた届出を転記して作成されている。建築着工統計は統計法と建築基準法（昭和 25 年法律第 201 号）に基づいて、建築着工に際して提出される情報から作成されている。

基幹統計ではないが、税関を通過する輸出入品についての書類を集計した貿易統計（通関統計）や、気象統計、選挙統計、災害統計、犯罪統計などは代表的な業務統計である。企業・事業所に関する統計に関していえば、業務統計が充実している産業については、概して、統計調査の対象となることが少ない。建設業、電気・ガス・熱供給・水道業、金融業、保険業、公務などの産業は、法令等の規定するところにより、データが収集される仕組みになっており、その結果、統計作成のための調査を実施する必要性が少なかった。逆に言えば、業務統計の作成が困難な産業で多くの統計調査が実施されており、産業に係る公的統計は調査統計と業務統計を合わせて体系化されている。

加工統計が基幹統計として指定されるようになったのは統計法の全面改正の結果である。加工統計の代表である国民経済計算は、統計法の条文で基幹統計に指定されている。統計法の改正以来、産業連関表、生命表、鉱工業指数、社会保障費用統計、人口推計の 5 つの加工統計が新たに基幹統計に指定された。

なお、調査の中には意識調査と呼ばれるものがあるが、意識調査の結果は統計法においては公的統計の範疇から除外されている。意識調査はその名のとおり意識を調査するものであり、思想や感情といった主観に左右される内容や移ろいやすい判断の調

⁴ 経済センサス、個人企業経済調査、科学技術研究調査、法人企業統計調査、経済産業省企業活動基本調査

⁵ 民間給与実態統計調査、毎月勤労統計調査、賃金構造基本統計調査

査結果は客観性を欠くものであるとされることによる。もっとも、法人企業統計調査や個人企業経済調査の業況判断や、労働力調査（特定調査票）の「転職などの希望」のような意識を問う項目が含まれる調査においても、客観的な経営状態を踏まえた状況判断・先行きの見通しや調査時点の状況に関する判断のような内容は、実態を反映する客観性を有しており、他の事実に関する調査項目に相当するとして扱われている。

2 公的統計の現代的意義

公的統計の意義は公共の利益のために必要な情報を提供することである。このために、公的統計の多くは、国民から情報を得ることにより、特に、統計調査の多くは国民から直接回答してもらうことにより作成されている。公的統計は、これまで行政を的確に運営する上で基本となる情報として重要な役割を果たしてきたが、国民が社会・経済を正しく把握し、適切に意思決定するための情報基盤を提供するという、公的統計が有する本来の役割の方向に見直されるようになってきた。

しかしながら、政府が有効な施策を企画立案し、行政を効率的に運営するための公的統計の役割は、現代においても極めて重要であるので、公的統計の現代的意義を論じるに当たって、このことについて解説することから始めることとする。

公的統計の利用価値の高さを他の統計と比較して客観的に表現することは困難であるが、まず、法令（法律又は政令）の規定における公的統計の引用例を挙げよう。

衆議院議員選挙区画定審議会の事務は、衆議院議員選挙区画定審議会設置法（平成6年法律第3号）の第2条で、衆議院小選挙区選出議員の選挙区の改定に関し、調査審議し、必要があると認めるときは、その改正案を作成して内閣総理大臣に勧告するものとなっている。この改正案の作成に当たっては、第3条で、各選挙区の人口（官報で公示された最近の国勢調査又はこれに準ずる全国的な人口調査の結果による人口）の均衡を図り、各選挙区の人口のうち、その最も多いものを最も少ないもので除して得た数が2以上とならないようにすることを基本とし、また、第4条では、勧告の期限を、国勢調査⁶の結果による人口が最初に官報で公示された日から1年内に行うものとするとしている。

国民年金法（昭和34年法律第141号）第27条の2には、老齢基礎年金の改定率について、毎年度、物価変動率を反映させた名目手取り賃金変動率を基準として改定し、当該年度の4月以降の年金たる給付に適用するとした上で、物価変動率を当該年度の初日の属する年の前々年の物価指数（総務省において作成する年平均の全国消費者物価指数）に対する当該年度の初日の属する年の前年の物価指数の比率としている。

⁶ 統計法第5条第2項本文の規定により10年ごとに行われる国勢調査に限るとしている。

法令以外にも、公的統計は客観的事実として白書や各種行政計画に数多く取り込まれ、社会・経済政策の立案や評価の基礎資料として役立っている。例えば、家計調査の結果は、経済政策の基礎資料、需要予測、給与ベース算定、国民経済計算、公共料金改定、CPIのウエイトの算定など幅広く利用されている。さらに、最近では政策評価において、統計に表章される結果を目標に設定し、政策の具体的な評価をすることが多くなっている。日本銀行は、2013年1月に「物価安定の目標」を消費者物価の前年比上昇率2%と定め、これをできるだけ早期に実現するとしている。また、2016年6月に閣議決定された「ニッポン一億総活躍プラン」では、誰もが活躍できる一億総活躍社会を創っていくため、名目GDP600兆円、希望出生率1.8、介護離職ゼロという目標を掲げている。

次に、国民が社会・経済を正しく把握し、適切に意思決定するための情報基盤を提供するという本来の役割に焦点を当ててみることにする。国民が誰でも国の状況を正しく把握し、それを踏まえて意思決定し、行動することは、民主主義国家の基本である。公的統計の果たすべき最も重要な役割は、主権者としての国民が適切な判断を下せるように、国の社会・経済の実情を正しく反映している情報を提供することである。公的統計はこのような役割を果たすために整備されている。全面改正された統計法においても、第1条に記されているように、公的統計が国民にとって合理的な意思決定を行うための基盤となる重要な情報であることにかんがみ、公的統計の作成及び提供に関し基本となる事項を定めることにより、公的統計の体系的かつ効率的な整備及びその有用性の確保を図り、もって国民経済の健全な発展及び国民生活の向上に寄与することを目的としている。

実際、情報化の進展は、個人が日常的に公的統計に目が触れる機会を多くし、利用しないと不利な状態に追いやられる状況を作り出している。企業も、経営者の勘に頼るのではなく、客観的なデータに基づいて合理的に判断することが一般的になっている。公的統計は、利用のされ方も人によって様々であるので、その対象はすべての地域、属性をカバーするものとなっている。もちろん、特定の地域、属性の利用に多くの利用上の需要があれば、それに対応し得るように調査の設計段階で工夫されていることが肝要である。しかし、公的統計はあくまで公共の利益のために行政機関等が多数の国民や関係者の協力を得ながら作成しているものであり、個人の利益や関心のために作成されるものではない。特定の対象に関する分析に対する興味や関心が高いからといって、その需要に必ず応えるのが妥当とも言えない。例えば、町丁字別犯罪発生率、学区別子供の平均知能指数の集計結果は、個人的関心からすれば需要の多い社会・経済の情報であるかもしれない。しかし、このような情報は、行政や研究においては有意義ではあっても、それが広く知れ渡る結果、無益な地域間差別や偏見を生み

出す怖れが十分にあると考えられ、統計として公表することは避けられるべきであろう。

近年、国民主導の動きが進む中で、政治・行政のあるべき姿が問いただされ、官庁自体がその役割を問われている。官庁の役割は行政であり、行政は国民のためにあるとの認識が改めて強調されている。同様に、公的統計も直接国民に資する役割を有しているか否かが問われている。最近、行政や政策の外部評価、専門家による計画策定など、行政機関が直接関与しない形で公共の利益の増進を目指す機会が重視されているが、その際にも、中立で客観的な指針が必要であり、公的統計はその基礎資料を提供する役割を担っている。公的統計を利用した研究についても、社会・経済を対象とした有用な研究成果は何らかの意味で行政や政策の方向に活用できるものである。したがって、公共の利益を目的とする利用の範囲は行政機関に限定する必要はなく、国民による幅広い利用に設定されるべきであり、公的統計の役割においては、公共の利益の意味をかなり幅広く解釈すべきであると考えられる。当面の行政や政策に直接関連がなくとも、社会科学の研究であれば、将来、又は間接的に貢献する場合もあり得るからである。

昨今、公的統計の民間利用は増加しており、その背景を考察すると、次の5つの点を指摘できる。

第1に、右肩上がりの時代が終焉し、不確定な時代になったことである。かつてはほぼ一律に発展し、そのバロメータとされた人口や経済指標の増減が0（ゼロ）近傍で推移している。そこから世の中の変化の動きを把握するためには、変化をもたらしている根幹についての詳細、例えば、年齢別、産業別、事業所規模別等から捉えなければ実態が理解できなくなってきた。しかし、行政が個々の構成単位ごとの的確な結果表章を行うことは困難であり、また、個々の構成単位のどこに関心があるかは企業、団体、研究機関ごとに相違するので、それぞれの関心に沿って公的統計を自らが解剖しなければならないこととなる。つまり、不確定な時代においては、国民が公的統計を入手し、正確に、迅速に、かつ詳細に読解する能力が問われるのである。

第2に、急速な情報化の進展である。インターネットの普及、データの収容力の増大、統計解析ソフトウェアの進歩と価格低減などが統計データの入手と読解を個人ベースでも可能とした。公的統計は、かつては、基本的な集計事項を除けば、事実上、資金と時間を有する者しか使いこなせないものであったが、今では多少複雑なクロス集計の結果についても誰でも容易に利用できるようになった。これにより、公的統計を国民が社会・経済を正しく把握し、意思決定し判断するための情報とする基盤が整った。

第3に、多様化の進展である。社会・経済が成熟化し、国民の関心や活動、生活様式、価値観や意識が多様化する中で、平等や均質を求めるよりも個性や個人の多様な価値観を重視する傾向が強まった。それとともに、国民が判断のよりどころとする統計も多様化し、求められる表章形式も総数や平均値を中心とする単純クロス表にとどまらなくなってきた。しかし、行政機関が事前にそれらのすべてを把握し対応することは困難である。そこで、国民が自らその情報を得ることができるようなシステムが求められるようになった。オーダーメイド集計や個別データの利用に対する需要はこのような状況から発生している。

第4に、国際化である。グローバル化が進展する中で、日本の社会・経済に関する基本的情報は国際的にも世界全体について知る上で不可欠な構成要素である。しかも近年、市場経済化が国際的により一層進んだ結果、必要な統計が増加している。各国の政府も企業も諸外国の統計が必要になるからである。公的統計の国際的需要の増大は、利用者層の飛躍的拡大につながった。

第5に、民主化の進展である。民主主義社会においては、主権は国民にあり、行政は国民からの負託に基づいて市場の失敗を補完するために活動を代行する存在である。このため、行政活動を遂行するに当たっては国民のコントロールが必要である。近年、このような国民本位が再認識される中で、国民から透明で公正な行政への要請が強まり、行政は従来にも増して政策決定の根拠やプロセスの説明を求められるようになった。また、行政は政策の実施についてもその結果責任を一層厳しく問われることとなり、国民の政策決定過程への参加の要請も強まっている。行政に対する国民の理解と協力は、行政からの一方的な説明だけでなく、国民による評価を踏まえて制度・運営を改善するなど、国民との双方向の意思の伝達を図ることにより達成される。このような評価には、公正な根拠が必要なことから、そのよりどころとして公的統計への需要が高まってきた。

ここまでは、公的統計それ自体の役割に焦点を当ててきたが、公的統計のいくつかには、他統計の作成のための基盤としての役割もある。国勢調査、経済センサス、工業統計調査、商業統計調査、医療施設調査などのセンサス調査の結果は他の統計の調査設計、推計に活用されている。例えば、国勢調査を基に推計された人口は、住宅・土地統計調査、就業構造基本調査、労働力調査などで比推定を行う際のベンチマークとして使用される。経済センサスの従業者数も、毎月勤労統計調査などで比推定のベンチマークとして使用される。さらに、大規模センサスの調査対象や調査区に関する情報は、他の統計調査の調査対象や調査区の抽出にも利用されている。例えば、国勢調査の調査区は多くの標本調査の調査区抽出に利用されている。経済センサスの結果は毎月勤労統計調査の調査対象抽出の名簿であり、商業統計調査の結果は小売物価統

計調査などの調査対象抽出の名簿である。また、国民生活基礎調査のように、大規模標本調査の中にもこのような役割を果たしているものがある。大規模調査の結果から標本を抽出することは、これにより複数の標本調査で同一対象が選ばれないようにすることができるので、調査対象の重複を避ける意味でも効果的である。このように、大規模な公的統計は他統計の作成のための基盤として様々な役割を果たしており、特に、国勢調査は、その基本単位区が世帯調査の多くで直接又は間接に用いられ、経済センサスの調査区の枠組みにもなっていることから、大半の統計調査に結果が活用されている。

なお、公的統計が充実したとしても、民間統計の役割が減ずることはない。民間統計調査の結果の中には、行政機関が加工統計の基礎データとしているもの、白書や官公庁の報告にそのまま又は加工して掲載されているものなどがある。これらは、かつて業界動向の把握、マーケティング、需要予測などの必要から開発されたが、社会・経済の変化に伴い次第に公共の利益にとっても有用性が増したことによって、行政での利用が拡大している。平均株価、マンション販売額、倒産件数などは、社会・経済の基本的知識として公的統計と同等の役割を果たしている。民間の加工統計にも、市街地価格指数のように、様々な場面で公的統計と同等の扱いを受けるもの、労働生産性指数のように公的統計を民間機関が加工したもの、月次 GDP のように四半期ごとに作成される公的統計を月次単位で推計するもの、人口や経済の各種将来推計の中には公的統計との比較において、時には推計精度が高かったものなどがある。また、公的統計の調査対象が事業所・企業の場合は、調査票への記入事項や業務報告内容は当該事業所・企業の業務記録や業務統計であることも少なくない。

3 公的統計の歴史

ここで世界に目を広げ、公的統計の歴史について国家の誕生以来を長期で捉えれば、為政者が利用するための統計から、行政が活用するための統計を経て、国民が幅広く利用しうる公共財としての統計に至る歴史と言える。公的統計の意義も、官庁そのものの性格と軌を一にしている。つまり、為政者は、いにしえは国家を統治するために、自らや自らの取り巻きのみが兵役・徴税に資する情報を得るための統計調査を行っていた。近代国家においては、行政を有効に実施するために、国家の情勢を広範囲に把握するための統計調査を行っていた。現代においては、国民が適切に意思決定し行動するための情報基盤を提供することに立って統計が作成されている。

日本では、公的統計の立ち上がりが欧米よりも遅く、第2次世界大戦時に民主化に逆行する動きがあったとしても、長期的には類似の軌道をたどっている。

統計調査は、古くは紀元前 3050 年頃のエジプトの人口調査、紀元前 3000 年頃の中国での人口調査、土地調査、農工及び商業に関する調査、日本では紀元前 86 年の崇神天皇の時代に行われた人口調査などに遡るが、これらは、そこから実態を映す事実を引き出すとか、当時の課題を明らかにするという目的からではなく、専ら納税、徴兵、強制労働を達成するための情報収集として行われた。欧州では、ギリシア時代やローマ時代にも大規模な戸口調査が実施されたが、同様の目的で行われた。ただし、これらは、為政者が最も必要としている人口、耕地面積、軍隊及び納税に対する負担能力などの情報入手が統計調査の目的であった。また、これらの調査には、原始的とはいえ、現在の統計調査における実地調査や業務統計の作成過程が導入されており、こうした大作業を組織できる統計主体は国家に限られていた。

加工統計については生命表の誕生が古代に遡る。生命表は、ローマ帝国の Ulpianus が作成した平均余命表が最古であるが、その後の著名な実績は 17 世紀にオランダで保険数学を開祖した de Witt の死亡表に下る。しかし、彼等はいずれも時代の権力者であり、この間、統計学の材料になり得る統計調査の結果は、深く政府の倉庫に秘蔵され、公開されることもなく、学者も利用することができなかった。古典統計学の始まりは、17 世紀に英国の Graunt がロンドンの発表した約 80 年間の死亡表に統計的方法を適用したことである。これを端緒に、英国では数量的な観察の中から政治体、いわば社会現象の構造を知るための政治算術が発達した。同派の先駆者は統計資料を収集したことと、そこから統計法則を発見したことの 2 重の意味で偉大であった。また、英国では、1675 年に Vaughan が「硬貨及び鑄貨制度」において 1352 年を基準年次として 1650 年の物価を比較し、物価指数の端緒となった。

一方、17 世紀にドイツでは Conring が、国家の行動は国家目的を追求することにあるとし、国情論を著した。ただし、国情論では、数量的な関連の把握を重視せず、平板な国家状況の記述に固執し、各国家を別々に扱った。この意味で、国情論は現代的な意味での統計的認識を欠くが、統計学の名の下で統計調査を含めた社会の実証的観察を広める契機となり、その意味で近代統計学のもう 1 つの柱であった。

17 世紀の欧州では、国家主義的な経済思想や経済政策を目指す重商主義のもとで、人口増加や輸出額の拡大が国家繁栄の象徴とみなされ、統計調査は重要政策の 1 つであった。絶対主義的な国家であったフランスでは、公的統計が異常な発展を見せ、1665 年に貿易統計が作成された。

18 世紀には欧州各国で統計を所管する官庁や統計協会が設立された。また、18 世紀は人口統計が各地で発展した時代でもあり、1790 年には米国で近代的人口センサス

が初めて開始された⁷。「近代的」とは、徴兵や課税のような特定の行政目的との結びつきを離れ、社会・経済の情勢を客観的に把握することを第1の目的としているという意味である。

19世紀に入ってから、各国政府が政府の仕事として人口センサスを本格的に始めた。また、ベルギーの Quetelet が、英国政治算術派、ドイツ国情論派、これに古典確率論派を加えた3つの源流を統合して1つの統計学を作り上げるとともに、人口に関する研究を主体とする従来の研究範囲を拡張し、すべての社会現象に大量観察法を媒介として実態を把握した。さらに、19世紀には、多様性に対応する必要が顕在化し、近代的な統計制度が始まった。こうして、ベルギー、ドイツなどの欧米各国で公的統計の整備が進展し、1853年には Quetelet を議長として第1回国際統計会議が開催された。

日本においても明治期になって近代的な統計調査の方法によって統計が作成されるようになった。

明治期で最も大きな流れは、経済の発展段階で、1870年に作成された物産表を始め、生産又は産業統計が発達したことであった。この過程で農林業センサス、工業統計調査、商業統計調査につながる調査が近代的調査方法のもとに実施された。一方、人口統計が発達したことも大きな出来事の1つであるが、これには、2つの流れがあった。1つは、維新の新制度を導入・創設していく上で不可欠な戸籍編成であり、制度的発展に伴う業務統計の整備過程を経て今日の人口動態調査につながるものである。もう1つは人口センサスの準備、試験調査であり、1920年以降5年ごとに実施される国勢調査につながるものである。

20世紀になると、戦争や不況により国家の社会・経済に対する関与が拡大し、それとともに統計の実際的な必要性も強まった。日本においても、大正期には、米騒動や労働争議が多発する中で、家計調査が熱狂的に流行し、労働統計が体系的に確立された。しかし、昭和初期は政治的にも経済的にも危機の時代と呼ぶにふさわしく、統計も権力の都合で隠されたり歪められたりすることもあった。戦後の公的統計の大きな転換は、真実を伝える統計への転換である。このために基本的な制度の枠組みが整備された。

終戦までの昭和前期の流れには、不況対策に関連していたか、あるいは戦時動員に役立てることを狙いとした統計が非常に多かった。しかし、戦時統制を運用するためには、統計そのものが全体の姿、又はその正確な部分を表すものでなければ使い物にならず、そのためには大規模センサスが必要だが、それは経費的にも時間的にも戦時

⁷ 1665年にニューフランスの植民地（カナダ）において実施された人口調査を最古の近代的人口センサスとする見解もある。

経営には適さなかった。終戦前までのこうした問題は、戦後の標本理論の導入によって初めて打開された。

20世紀後半における経済統計上の最大の発展は国民経済計算体系の確立である。国民経済計算は、統計に対する需要を高め、それにつれて調査の内容も詳細になり、統計利用の総合化と精緻化を促した点も見逃せない。また、公的統計の利用の仕方として、研究者が社会・経済の研究に、企業がマーケティングに利用する機会も増えた。このような流れの中で公的統計の本来の役割が見直されるようになった。1994年の国際連合統計委員会の決議で「公的統計の基本原則」が採択され⁸、公的統計は、民主的社会の情報システムに不可欠なものであり、政府、経済、国民のために経済、人口、社会、環境に関するデータを提供するものとしている。

同原則について、国際連合統計委員会は、公的統計は、経済学、人口学、社会学、環境学の発展基盤であるとともに国家と人類が相互に理解し取引するための基盤であるとの認識をしている。将来、国際化が進展し、公的統計の最も重要な役割が国家の枠を越え、このように地球規模的になる可能性も十分あり得ることである。

4 公的統計の特徴

公的統計はその意義を達成するために、実験調査、市場調査、世論調査、民間統計などと比較していくつかの特徴がある。これらは統計により区々であり、一言では表せないものであるが、公的統計の有する性質として次の4つを提示する。

第1に、公的統計は網羅性が高い。つまり、公的統計は広範囲に及ぶので、行政機関の運営や国民が生活を進める上で最低限必要な統計は公的統計の中に見出すことができる。

一方、公的統計調査は、統計法に基づいて重複しないように総合調整されている。また、他の統計調査の結果を利用することにより補完できる調査事項は盛り込まれない。これも国民の負担を軽減し、統計調査の質の向上を図ることに寄与している。

第2に、公的統計は民間統計に比べて信頼性が高い。統計の信頼性とは、統計の対象範囲が、本来の把握対象である社会・経済とどの程度合致しているかになるが、公的統計では、例えば、失業者、生産額、事業所という重要な対象範囲を規定するに当たって、作成者の主観に頼るのではなく専門家も含めた多数の関係者の意見が聴取され、国際的な議論も反映されている。もちろん、統計であるからには、現実には作成又は推計可能な範囲に収めることによる対象範囲の歪みは避けられない。それでも、公

⁸ 2014年の国際連合総会でも採択

的統計は、社会・経済の実像を得るべく理想の対象範囲を歪めぬようできるだけ広範囲のデータを収集して作成されている。

個人が統計を作成しようとする場合、予算、人手、時間などの制約から、特定の地域、特殊な属性、手軽な方法などで、入手しやすいデータだけを収集することが多い。もちろん、自然科学、医学、心理学などの分野では、ある特定の現象、対象を実験的に捉えるために個人やグループで統計を作成することがあり、これらは非常に重要な役割を果たしている。しかし、社会科学の分野で利用価値の高い統計を得るには、広範囲の対象を収集したデータが必要になる。行政機関では行政に反映させるべく社会・経済を的確に把握する必要があることから、データ入手が困難ともいえる対象をも含むことが多い。民間機関は、利潤を追求する以上、基本的にはマーケット戦略の対象となる範囲からデータを収集する機会が多く、その意味では個人と同じ状況下にある。もっとも、大学、団体、研究所の中には、社会・経済全般の把握を目的としているところが少なからず存在し、これらの統計の対象は広範囲に分布している。また、独立行政法人や特殊法人のように、国の機能に近い役割を果たしている機関についても同様である。

公的統計の対象の広さは次の点からうかがえる。現在作成されている基幹統計調査の対象地域は、特殊な地域を除き、すべて原則全国であり、国勢調査や経済センサスなど全数を対象とするものが 27 ある。国が実施する一般統計調査の対象地域も大半は全国である。地方公共団体が実施するそれらの対象も所掌地域全域に及ぶ場合がほとんどである。業務統計にも人口動態調査、建築着工統計、金融統計、貿易統計など、捉えられている対象が全国網羅的で全数のものが多数存在している。

一般に、対象範囲を広くすると対象数も多くなる。ただし、対象数が多いことをもって利用価値の高さを評価すべきではない。対象数が数十万程度であれば民間統計調査の中にも多数存在するが、その中には、例えば、調査対象の協力を得るために景品を付随するなど、調査結果に偏りを含みやすいものもあるからである。公的統計には大規模統計が多いが、港湾調査、鉄道車両等生産動態統計調査など、対象数が数百程度のものもある。

また、公的統計の対象範囲の広さは時間軸で見ても当てはまる。家計調査を実施している国は数多いが、1946年以降毎月実施している国は日本だけである。公的統計の中には、その作成開始時期もさることながら、社会・経済の体温計として、四半期ごと、月ごとに作成されているものもある。一般社団法人などでは乗用車生産台数やエアコン出荷台数などのいわゆる業界統計などの統計を月単位で作成しているところもあるが、大学、団体、研究所の中で、このように高頻度で継続的に社会・経済の統計を作成しているところは少ない。もっとも、株価、視聴率、販売額などに見られるよ

うに、企業には情報通信技術を利用して秒単位で作成する統計があるが、公的統計でこれほど作成頻度の高い統計は気象統計などに限られる。しかし、企業の統計作成の目的は、同一業界における自企業の株価、視聴率、販売額などの位置付けであるのに対し、公的統計の目的は株を含めた総資産、テレビの視聴を含めた生活時間、小売業全体の販売額など社会・経済の実態である。

第3に、公的統計は正確性が高い。すなわち、公的統計の中には比較的正確な統計が比較的多い。正確な統計を作成するには真実を伝えようとする努力とこれを可能にする制度が不可欠だが、これは戦時中の苦い経験が薬に転じて獲得されたものである。

これは、公的統計はその重要性故に、まず、多数の国民や関係者の協力を得て作成されていることによる。つまり、公的統計は国民生活に必要なものなので、国民もこれを作成するために自分に関する情報を提供することが半ば義務となっているのである。特に、基幹統計調査については報告義務が課されるなどにより、未回収や無回答が非常に少なくなっている。しかし、実際の統計調査においては、情報の入手は強制力ではなく協力により行われる。それでも、一般に、民間統計調査と比較すると、公的統計調査では、個々の入手情報に対する秘密の保護と使用の限定が徹底されていることから、また、実施主体が行政機関であることから、調査対象者の信用を得やすく、このために情報の入手が可能となっている。

公的統計の正確性の高さは作成方法の工夫、統計作成に携わる人々の努力にも依存している。統計調査では、調査員が巡回して調査対象から調査票に記入してもらう、あるいは調査員が自分で記入する調査員調査という方法を採用することが多いが、一般に、調査員調査はオンライン調査、郵送調査及び電話調査よりも調査票の回収率が高い。これは、公的統計を国際比較したときに正確性で日本が優位に立つ大きな要因でもある。現在のようにプライバシー意識が強い中での調査員調査の実現は、調査員を始めとする関係者の努力の賜物である。特に、基幹統計調査では、調査員は回収を前提に調査対象との接触を図らなければならないので体力的にも精神的にも苦勞が多い。もし未回収もあり得るのであればその負担が軽減される代わりに正確性が犠牲になる。

また、調査員調査のうち個人や世帯を対象とする調査について、公的統計では、実地調査の多くで、国や地方公共団体の指導のもとに調査の前に調査員などが巡回して対象者名簿が作成されるのに対し、民間統計調査では、経費の節減を重視するために、このような作業を捨象し、住民基本台帳や選挙人名簿に頼ることが多い。このため、調査時において名簿の内容が古くなり、調査対象の実態との乖離が起りやすい。

さらに、公的統計について、個人や世帯を対象とする統計調査では、まず、世帯を捉え、その中の一定年齢以上のすべての個人を調査することが多いが、民間統計調査では、直接個人を捉えることが多い。というのは、住民基本台帳や選挙人名簿では世

帯を区分することが非効率又は不可能だからである。しかし、世帯調査では世帯員の誰かが捕捉できれば実地調査が成立するのに対し、個人調査では当人との接触が必要になるので、報告義務が課されない統計調査においては回収率が低くなる。また、未回収者には、若者、単身者、未婚者、就業者が多いなど、基本的な属性区分に影響が出やすい。同じことは意識調査にも当てはまる。意識調査では意識を調査する性格から、同一世帯員であっても第3者が介在しないほうが好ましく、このためにやはり個人が調査対象になるからである。また、意識調査には客観化しにくい意識を尋ねることから生ずる困難もある。人の意識は漠然としたものであって、自らそれをはっきり表現できるとは限らない。そのはっきりしないものを明らかにしようとする困難である。

第4に、公的統計は統計間の比較可能性が高い。これは、複数の公的統計で共通に用いられる用語や分類について、相互に比較できるような基準が定められているからである。しかもこのような基準は、国際的枠組みに沿って定められる場合が多く、この場合には国際比較が可能になる。

例えば、産業は日本標準産業分類、職業は日本標準職業分類、商品は日本標準商品分類など共通の分類に従って区分することとなっている。このため、産業、職業、商品などについて複数の統計調査の結果を比較することが容易である。しかもこのような分類基準は国際機関で定められているので、世界各国の統計を比較することができる。これらの他にも、統計については国際連合等で国際的な基準を定めることが多い。失業の定義、物価指数の作り方などはILOの基準で定められており、日本でもこれに沿って作成されているので国際比較が可能である。

国民経済計算は、所得水準や経済成長率などの国際比較を行い、各国の経済の実態を明らかにする目的にも用いられるので、世界の各国が共通の基準に基づいて作成する必要があり、国際連合が世界の専門家を集めて作成している。国際連合が2008～2009年に2008SNAとしてSNAを改定し、これに準拠したSNAの作成を各国に勧告したことを受け、日本も2016年12月に2008SNAに移行する予定である。国際収支表についても、IMFが2008年に改定したマニュアル第6版に合わせて2014年から作成方法を変更した。このように、加工統計の改定は国際的な要請を踏まえて行われることが多い。遡れば、国勢調査が始まった最初のきっかけは世界人口センサスへの参加であり、農林業センサスも経済統計に関する国際条約（1952年条約第19号）に基づきFAOが提唱する世界農業センサス計画に参加することが目的の1つとなっている。

統計調査の結果についても、国際連合、OECD、IMFなどの国際機関が各国の基本的な統計を国際比較が可能な様式で報告書に取りまとめて発行している。統計調査の

調査項目には国外での利用を見込んで設定されるものもある。各国の人口センサスは、国外に住む自国籍の人を調査することができないので、国際連合は、各国が人口センサスでその国に住む外国人の国籍別資料を得るよう勧告している。

このように、公的統計は世界全体の社会・経済を把握する上で不可欠な基本的情報なので、世界各国から貴重なデータ源としての位置付けを得ている。そもそも統計にはグローバルスタンダードに馴染みやすい面があるが、公的統計は国際比較に耐え得るように設計されている。

比較可能性という点では、地域間のみならず時間間、すなわち、時系列比較の可能性も忘れてはならない。公的統計は時系列比較を確実にするための工夫がなされている。例えば、労働力調査では、前月及び前年同月との比較を可能にするために、標本を4組⁹に区分して各組2か月継続の調査を同一月で2年間調査し、組ごとに開始年月を変えることにより、毎月、前月と比較しても前年と比較しても標本全体の2分の1だけが入れ替わる。しかも継続調査における標本の脱漏を防ぐために、標本として世帯ではなく住戸を捉えている。家計調査でも調査世帯は6か月間調査され、その交替は毎月6分の1ずつ行われる。また、小売物価統計調査では、同一品目内で銘柄の売れ筋が変わる前後で新旧銘柄両方の価格の動きが調べられて物価指数に反映される。一般に、報告書においては、時系列で比較できる調査対象のみを集計した増加幅、増加率、指数などが掲載される。時系列での利用が想定される統計においては、各時点の正確性と時系列としての安定性とのバランスが図られるが、公的統計は、他の統計と比較して後者にシフトした設計になっている。

5 利用上の留意点

公的統計が本来の役割を発揮するには、統計作成者が統計に対する国民の需要を捉える努力が不可欠であるが、同時に、社会・経済の構造・動向を捉えるための公的統計の利用が質、量ともに向上することも必要であると考ええる。

統計作成者以外の方が当該統計を利用する場合には、利用の仕方を誤ることによる誤解を避ける必要があるが、このことは公的統計を利用する場合にあっては特に重要なことと言える。数字は一人歩きしやすいので、誤った使い方をすれば公共の利益が損なわれるおそれがあるからである。逆に、公的統計の性格をよく理解して正しく利用すれば、非常に有益な情報が得られる。

統計利用に当たっては、まず、データを客観的に公平に取扱う姿勢が重要である。先入観にとらわれて数字の示す真実を曲げてはいけない。統計を説得の材料として利

⁹ 実際には調査区を2つに分け、それぞれ前半2か月、後半2か月調査するので、標本は8組に分けられている。

用したいという気持ちが過ぎると、部分的な誇張をしがちであるが、誤解を与えないようにしなければならない。逆に言えば、統計を利用した主張を追う場合にこのような誤解をしないよう、客観的な利用の仕方をしているかどうかを意識する訓練が必要である。

また、単独の統計だけで結果を出そうとするのではなく、複数の統計を比較したり、組合せたり、切り分けたりして、多角的に統計を眺める姿勢が重要である。複数の統計から同一の現象を捉えようとする努力の過程で、統計ごとの性格を見出すことができるからである。

技術的にも、例えば、平均値を解釈する際には分布についても考慮する、グラフを見る場合には目盛りを確かめる、相関係数の強さをもって因果関係の強さと決め付けないなど、様々な注意が必要である。これらのことは公的統計に限ったことではなく、統計であればどれを利用する場合でも言えることである。そこで、これらの詳細は他の文献に譲ることとし、ここでは、公的統計の特徴を踏まえつつ公的統計を利用する場合に注意しなければいけない事項に絞って論じる。

しかし、このような留意点は統計ごとに区々であり、個々の統計を理解するには、その統計に関する解説を読み、その統計の作成者の話を聞くことが最良かつ最短の方法である。このような解説を統計ごとに比較すれば、分析目的にかんがみてどの統計を利用すべきか見当がつく。しかし、すべての公的統計についてこのような解説を束ねると、膨大な紙数を要することとなる。

そこで、本説では、このような方法で個々の公的統計に対する理解を深めようとしたときに、多くの統計に共通に現れる基本的な視点だけをまとめ、それらの中から重要な5点を提示する。

5.1 収集データの範囲 —知りたい事項は収集データの範囲内か—

公的統計を利用するには、まず、その目的を見る必要がある。目的により収集データの範囲、すなわち、カバレッジが異なり、それが調査結果の性格として現れる場合があるからだ。

この観点からすれば、同じ公的統計でも統計調査と業務統計とで性格が異なる。前者が社会集団の表現が目的であるのに対し、後者は行政組織の組織活動の表現が目的である。したがって、後者における業務対象の範囲は利用者の観察対象である社会・経済の範囲と一致しないこともある。例えば、職業安定業務統計は、職業安定所を通じて処理された求人・求職状況を取りまとめたものであり、情報誌やインターネットなどを通じた求人・求職状況を含んでいない。

最近は行政機関内に電子化された行政記録データが大量に存在するので、これらを使用した業務統計の活用が検討されている。このようなデータには、収集した目的以外には使用できないという問題が生じるが、それ以前に、収集目的が異なるためにそのままでは活用や接続が困難という問題もある。

他方、公的統計調査においては、収集データの範囲は比較的広い。しかし、個々の統計調査を見ると、作成目的に対応する収集データの範囲を100%把握しているものはない。収集データの範囲は作成目的に沿っていることが原則であるが、実地調査やデータ収集上の制約、費用などにより限定せざるを得ないからである。

収集データの範囲が限定されているということは、利用の際にはその範囲でものを言う姿勢が重要である。例えば、経済センサスでは、農業、林業及び漁業に属する個人経営の事業所は調査対象から除外されている。個人企業経済調査の対象産業は製造業、卸売業、小売業、宿泊業、飲食サービス業及びサービス業のみであり、他の産業は含まれていない。また、特定サービス産業実態調査において学習塾を調べているが、予備校などの各種学校、通信教育及び家庭教師は対象外である。

もちろん、収集データの範囲が限定されても、統計作成の目的を逸脱するほどのことではなく、主要な情報を把握することにより概括的構造と動向を把握するには支障がない場合がほとんどである。例えば、事業所を調査することにより日本の経済情勢を把握するのであれば、すべての事業所を調べる必要はなく、比較的大きな事業所だけを網羅的に把握しておけば良いからである。実際、工業統計調査では従業者数3人以下の事業所が除外されており、「裾切り調査」と呼ばれている。経済産業省生産動態統計調査も全品目を調査しているのではなく、品目により従業者数でみた裾切りの水準が異なっている。標本調査では、事業所・企業を対象にする場合、科学技術研究調査や法人企業統計調査のように、小規模単位については比較的少数の対象しか抽出しない調査が多いが、結果の推定は抽出率の逆数を乗じることにより行われている。

特定産業の事業所を対象とする統計調査ではその産業の選び方も問題である。このような統計調査では、各事業所の産業を主な事業所で区分することが多く、主な事業所を見分けるために従業者数や販売額などが使用される。しかし、工業統計調査と商業統計調査は、それぞれ製造業、卸売業又は小売業を少しでも営むところを対象にしているので、双方で同一事業所が選ばれる可能性もある。農林業センサスや各種の生産動態統計調査のように、行政目的に沿った府省独自の基準に基づき調査対象を選定している統計調査も多い。

地域的広がりという視点からも同様である。地域の限定を念頭に置き、除外されている地域にまで結果を延長しないようにしなければならない。いくつか例を挙げよう。
①同じ地域でも統計により区分の仕方が異なる。新潟県は東北地方に含める統計と北

陸地方に含める統計がある。②同じ統計でも対象地域が異なることがある。学校基本調査の報告書では、大学に関する都道府県別結果の中で、学校数は大学本部の所在地で、学生数は学部の所在地で集計されている。③時点を遡れば地域区分が異なることもあり得る。1972年以前のデータは沖縄県の分を含むかどうかを確認する必要がある。④都道府県の行政事務の一部が政令指定都市に移管されている場合、都道府県別結果は必ずしも当該県の結果に相当するとは限らない。⑤災害が発生したなどの理由で調査そのもののできない地域もある。

対象範囲への配慮は加工統計にも必要である。企業向けサービス価格指数においては、企業間サービス取引の需給動向を正確に捉えるために実際の市場取引価格として高精度で収集できる価格だけが収集されている。このため、金融機関の帰属利子、商業マージン、教育、研究、医療、保険などは対象外である。また、食料自給率は食料自給表の国内生産量を国内消費仕向量¹⁰で除した比率であるが、畜産物については、国産であっても輸入した飼料を使って生産された分は国内生産量に算入されていない。

国民経済計算の各種指標は、国内を対象にするものと国民を対象にするものの2通りあり、前者は領土、後者は居住者に着目する指標である。県内、県民の区分も同様である。したがって、複数の統計データを組合せて分析をする際には、対象範囲ができるだけ合致するようにデータを選ぶようにすべきである。例えば、1人当たり県民所得を計算する際には、分母はその県の夜間人口（常住人口）が適当であるが、1人当たり県内総生産を計算する際には、分母としては、むしろ従業地・通学地で人口を集計した昼間人口のほうが適当である。

また、把握しようとする現象は同じであっても、統計の作成経緯により、収集データの範囲が複数の統計に区別されるときもある。このような場合には、これらの統計を結合することにより全体像を描き出すことも可能であり、このような結果に対する需要が強い場合には、作成者がこれを公表することもある。経済センサスのデータを基に事業所・企業に関する各種のデータを収集した事業所母集団データベースから作成した統計は、典型的な事例である。

5.2 概念の区別 —同じ用語でも統計によって意味が異なる—

公的統計は、その目的により調査項目の定義や意味が異なり、それが調査結果の性格として現れる場合がある。特に、公的統計調査については、調査対象、調査事項、時、場所が集団観察の4要素と呼ばれているので、最低限これらの情報は正確に把握する必要がある。

¹⁰ 国内消費仕向量＝国内生産量＋輸入量－輸出量－在庫の増加量

例えば、国勢調査は国内の人口と基本的な生活事情を把握するために行われるが、農林業センサスは農家¹¹及び林家¹²の経済的性格区分をするために実施される。このために両調査で「世帯員」の意味が異なる。国勢調査では不在期間が3か月以上にわたるような出稼ぎをしている人などは、その家の世帯員とせず、出稼ぎ先で調査することになっているのに対し、農林業センサスでは、不在期間が1年未満の場合は世帯員としている。

また、社会・経済が時代とともに変化し、それに伴って概念の重要性、ひいては統計の価値が変化することもある。例えば、企業物価指数について、明治時代の作成開始期には、国内品、輸出品、輸入品の区別をすることが困難又は無意味であったために総合指数が作成されたが、現在ではこれら3種をすべて手掛ける企業は一般的ではないなどの理由から、総合指数の意味が薄れており、国内指数、輸出指数、輸入指数ごとに分析をすることが望ましいとされている。これとは逆の流れが、サービス業における包括的統計調査の誕生や総世帯による家計収支結果の公表である。

さて、収集目的が概念に影響を及ぼすとはいうものの、目的の記述は概括的であり、統計表に現れる個々の用語の定義は、統計により定められたある種の約束事である。用語の定義は報告書とホームページに掲載されている。また、利用者の誤解を避けるために詳細な注釈は統計表ごとに掲載されている場合が多い。しかし、これらは利用者にとって必ずしも理解しやすい言葉で説明されているとは限らず、行政特有の専門用語が多用されることもあり、少なくとも複数の統計間で整合性がある書き方ではない。これは無論統計作成者側の問題であるが、専門用語を駆使して説明された統計の根本の作成目的は個別行政の指針であるという事情を考えると、少なくとも現状では、統計を利用する前のある程度関連する行政の基礎知識を身に付けておく以外に方法がない。特に、経済統計を使用するに際しては、経済の流れの中での位置付けや前提を理解しておくことも重要である。例えば、産業連関表は、各種経済活動の生産波及などの効果の推計・分析にも用いられるが、一般に、このような推計・分析には、生産能力に限界がなくあらゆる需要に応えられることと過剰在庫が存在せず効果の中断がないことが前提とされている。

また、各種統計を頻繁に利用する場合には、分野ごとに統計体系の概略を理解しておくことで統計の選択が容易になる。統計体系の概略とは、例えば、①人口については、静態を国勢調査で5年に1回、その間の動態を人口動態統計や住民基本台帳移動報告などで毎月把握することにより、静態も毎月推計、②事業所については、経済センサ

¹¹ 経営耕地面積が10a以上の農業を営む世帯又は経営耕地面積が10a未満であっても調査日前1年間の農産物販売額が15万円以上の世帯

¹² 保有山林面積が1ha以上の世帯

スで基本的事項を、詳細を産業ごとに全数調査や標本調査で把握、③土地については、個人であれば住宅・土地統計調査で、法人であれば法人・土地基本調査で把握、④国民経済計算は、国民所得勘定を中心として、産業連関表、資金循環表、国際収支表、国民貸借対照表から構成されるというような、分野ごとの統計の位置付けである。

さらに、用語の定義や意味の誤用を払拭するには、統計の原典に当たること以外に解決策はない。統計調査の場合には調査票様式、業務統計の場合には報告様式を入手し、個々の用語の定義がどの部分に記入したどの内容に該当するものなのか、記入説明書を見ながらチェックしてみると良い。

用語の定義を確認した後は、これが利用者の個々の分析目的と一致するかという検証が必要である。分析目的に応じて概念の選び方は異なるからである。地域内の病院や消防署の数を評価する場合、社会基盤の充実度を評価するのであれば分母は人口であるのに対し、利便性を評価するのであれば分母はむしろ面積とするほうが良い。

同じ働いている人を調べる調査でも、労働力調査の「就業者」は月末1週間に働いた人であるのに対し、就業構造基本調査の「有業者」はふだん働いている人である。就業者は「雇用者」、「自営業主」、「家族従業者」に区分されるが、このうちの自営業主は個人経営の事業主であり、法人組織の社長や経営者はすべて雇用者に分類される。このため、資本家や経営者について分析をする場合には、例えば、職業分類を用いるなどにより、雇用者を切出し、自営業主に加える必要がある。また、雇用者はいわゆる賃金労働者に等しいと考えてはならない。「失業」とは、労働力調査ではILO基準に沿って、「仕事がなく、就業可能で、仕事を探している」人と定義されているので、適当な仕事がありそうもないので仕事を探していない人も合わせて分析しようとするならば、非労働力人口の就業希望者で非求職の理由を「適当な仕事がない」としている人に加える必要がある。

5.3 作成方法の特徴 —作り方の性格を理解する—

概念の相違は統計作成方法に起因する場合もある。また、作成方法が異なれば同じ概念であっても結果は異なる。とりわけ、世帯から調査した結果と事業所・企業から調査した結果は、様々な事情により一致しないことが多い。いくつか例を挙げよう。

国勢調査では就業者を世帯から捉えるのに対し、経済センサスでは事業所から捉えるので複数事業所で働く同一人は、それぞれの職場でカウントされる。一方、組織に属さず家庭で仕事をしている人は、経済センサスでは就業者として捉えられない。

物やサービスの販売又は購入額を調査する場合、基本的には世帯から調べれば家計の購入額であるのに対し、小売店から調べれば企業が購入した分も含む。資金循環統計の家計部門の金融資産残高による1人当たり貯蓄額は、家計調査によるそれを上回

る金額になるが、資金循環統計の家計部門には、①純粋な個人だけでなく、個人事業主も含まれるため事業性資金が含まれる、②金融資産に年金準備金、預け金、未収・未払い金など、自らの金融資産とは認識し難い金融商品が含まれる、などの事情がある。

公的統計から家計の消費支出を推計するには2つの方法がある。1つは世帯を対象にした統計調査の結果、例えば、家計総世帯集計結果をそのまま用いる方法であり、これによると2015年度において家計の消費支出額は1人当たり1か月10万4千円である。他方、国民経済計算において2015年度の国内家計最終消費支出額から持ち家の帰属家賃を除くと237兆9千億円であるが、2015年国勢調査の人口が1億2711万人であることを利用して1人当たりへ換算すると、1か月15万6千円となる。この値は先に計算した家計総世帯集計結果に基づくものと大きく離れている。両者の差は、国民経済計算における家計最終消費支出額の推計に金融機関の帰属サービスが含まれる、集計や推計の基となる原データの調査又は収集時点が両者で異なるなど、ここでは定量化していない他の概念の違いに依存している。しかし、家計総世帯集計結果では、すべて家計簿に記入されたものを集計、国民経済計算では、単身者の動向を人口5万人以上の都市における勤労者世帯の消費支出を代理変数として推計するなど、推計方法の違いによる影響も大きい。

業務統計の原データは統計を主たる目的として収集されていない。このため、統計目的を掲げて分析をする際には、作成方法に依存して発生する問題点に配慮する必要がある。例えば、金融統計では、業務上の必要により勘定経理が変更されたり、勘定項目の取扱が金融機関の業態により異なったりする場合がある。職業安定業務統計の求人倍率は労働力需給の実態を把握する指標の1つとして公表されているが、労働力需給の実態を把握するために収集したデータを基に作成されているわけではないので、求職者が求職活動をあきらめれば求職申込み件数は減少する場合もあることを前提に使用する必要がある。そもそも、業務統計では制度の適用が強制か任意かによって性格が大きく異なる。その意味で、業務報告が法的に義務付けられている金融統計と任意の求職者集団を対象とする職業安定業務統計では見方も変える必要がある。

加工統計でも、国民経済計算のGDPは、支出系列ではcommodity flow法、生産系列では付加価値法という別個の方法で推計されるために推計結果に乖離が発生し、統計上の不突合¹³の要因となっている。このように、作成方法に起因する概念との問題は数多い。

次に、概念とは異なるが、利用に当たり留意することが必要な作成方法を例示する。

¹³ 2014年度においてGDPの0.5%

戦後、公的統計調査の標本抽出方法はほとんど無作為抽出である。しかし、機械受注統計調査は、迅速に調査を実施するために調査対象は有意抽出されているので、その結果から機械受注の総額を把握することはできない。

業務統計の基となる行政記録は、上意下達の組織系統で作成され、個人や組織の利害に直接関連する行政上の決定に用いられることもあるので、報告者が行政に対し特定の利害関心を抱き、虚偽の報告・届出が行われるおそれがある。このため、業務統計は構成や増減を概観するにとどめ、水準としては業務遂行組織の自己観察としての利用に限るべきである。例えば、税金のデータは、税の金額としては正確であるものの、税制を踏まえて逆算した収入や資産は過少になりやすい。統計調査の結果を利用する場合も、府省が所管行政に関して実施する統計調査についてはこの問題を視野に入れる必要がある。回収率は高くても何らかの利益を期待するような回答に傾いているおそれがあるからである。

加工統計は利用の仕方にふさわしい指標がある。景気動向指数の DI (Diffusion Index) は、景気の変化の方向だけしか表すことができず、変化の大きさや量感を把握するには、CI (Composite Index) を用いる必要がある。これは、DI が単に系列総数に対する拡張系列の割合であるのに対し、CI は DI の変化率を合成していることによる。

加工統計にも作成上の制約がある。地域メッシュ統計については、統計を正方形の枠目に区分することのメリットが強調されるが、約 1 km 四方の基準地域メッシュは、正確には緯度・経度を基に作成されているため、沖縄県那覇市では縦の長さが 0.923km、横の長さが 1.249km と、横長の長方形となる。地域メッシュ統計はフィンランドで生まれたが、これを国土が細長い日本に導入したことによる留意点である。なお、地域メッシュ統計は、調査個票から直接集計したものではなく、国勢調査の基本単位区別に集計結果を再編成したものなので、基本単位区が複数の地域メッシュにまたがる場合、当該基本単位区の結果は便宜的に配分される。

5.4 作成時期の区別 —同一年月のデータは変化する—

同じ統計でも、時間を経るにつれて作成方法が異なる場合がある。当節では作成方法に起因する留意点の中で、特に、時系列で結果を見る際に気をつけるべき点をまとめる。

公的統計で特記すべき点は、同じ統計、同じ時期・区間のデータでも、利用者がアクセスする時点に依存して異なる場合があることである。公的統計では、過去に遡って結果を訂正することがあるからである。したがって、統計を見るに当たっては、対象時点だけでなく、作成時点も確認しておく必要がある。

国民経済計算の四半期別結果は、1次速報値が当該四半期終了後約1月と2週間程度後に、2次速報値がその約1月後に、確報値が当該年度終了後の12月末に公表される。例えば、2015年1～3月の季節調整済み実質GDP1次速報値は前期比0.6%であったが、2次速報値は1.0%、確報値は1.3%に改訂されている。しかも確報値は次期公表時に遡及して改訂される。国民経済計算は、定期的に早い公表が求められる反面、多数の統計調査の結果や業務統計を加工したものであり、これらがすぐにはそろわず、代替指標で置き換えた上で暫定値として公表され、後で改める方法が採られるからである。

材料の有無の他にも、四半期や月ごとの統計には、季節、曜日、日数などの変動により現れる特徴があるので、前期又は前月の数値と比較する際にはこのような変動を取り除く必要がある。このために、多くの場合、米国センサス局法による季節調整値が用いられるが、これらは計算の元となる原データのとり方により結果が異なるため、新しいデータが原データに追加される度に同年同期（月）の数値が異なってくる。また、米国センサス局法にはARIMA（Auto Regressive Integrated Moving Average）モデル、管理限界など選択枝の種類が多く、それらは統計ごとに、その特性を踏まえて選択されている。

毎日公表されている公的統計は金融統計、気象統計などに限られるが、これらにも定期的な変動が起こることがある。例えば、資金需給統計は、毎月法人税等の租税や保険料の受入のある2日前後と民間給与払いのある25日前後に資金不足となる。このような高頻度の統計を月間比較する際には、日数が大きく異ならないという確認が必要である。また、天候や事件などの影響を受けやすいことから、移動平均をとるなどにより異常値を取り除く必要性も高い。

他方、長期的には、物価指数、国民経済計算などには時代の変化に対応させて計算に盛り込むデータの範囲を見直すために、5年ごとに基準改定がある。これらのデータを利用するに当たっては、基準年の確認が必要である。また、毎月の現在推計人口は次の国勢調査の結果が判明した時点で過去5年間遡って補間される。このような見直しは5年ごととは限らない。商業動態統計調査により作成される業種別商業販売額は、商業統計調査の結果が公表された時点で過去に遡って修正される。景気動向指数も景気への感応度や精度を向上させるため、適宜採用系列が改定される。

さらに、国民経済計算は、1968年、1993年、2008年のように、国際的な基準改訂が提唱される都度、合わせて改定される。これは、作成対象や計算方法をも含めた概念そのものに切り込むものなので、同じ年月のデータを見る場合でも先に述べた概念の確認は必要である。1993SNAから2008SNAへの改定に伴う概念の変更により、名目GDPは3%以上も増加する見込みである。国民経済計算に見られるようにこの

ような改定は即実施されるとは限らない。業務統計についても、新しい制度が創設されても、経過措置などにより、すぐに統計に反映されるとは限らない。

さて、次に作成方法に起因する時系列上のバイアスについて概観することとする。

標本を固定することにより上方の偏りが発生することがある。毎月勤労統計調査では、標本は経済センサスを母集団情報として標本事業所が抽出された後、当分の間、標本の変更はないため、業績が良くない事業所が廃業や規模縮小で脱落する結果、継続標本から算出される平均賃金には上方バイアスが生じることとなる。

計算式に起因する問題としては、指数について、Laspeyres 指数であれば上方バイアス、Paasche 式であれば下方バイアスが発生する。日本では5年ごとにウエイトが更新されているので一般的な分析には支障はないが、生産量や販売量が急速に変化する品目を含めた時系列分析をする際には配慮すべきである。

このように、ある時点を基準に時系列で追跡したときに結果にバイアスが発生する場合はいくつか想定されるが、一方で、バイアスが発生しても各期（月）同一方向とみなして支障がないことも多いので、総量ではなく前月又は前年同期（月）比で分析をするなどの工夫をすると効果的である。

5.5 誤差の取扱 —誤差の情報も利用する—

一般に、標本調査については、①標本の抽出方法、②結果の推定方法、③誤差の計算方法を理解した上で、標本誤差を考慮し、過度に正確さにこだわるような分析は慎むべきである。例えば、GDP 変化率の僅か 0.1 ポイントの動きが社会問題化することがあるが、これは 0.1 ポイントよりも大きな標準誤差率を有する各種統計調査の結果を加工したものなので、この程度の幅の動きを評価することは無意味である。このような統計はもっと大幅に長期的に眺める必要がある。

2012 年 9 月には家計調査の 2 人以上世帯の消費支出が前年同月比で実質 0.9% 減少した。これは、数字の上では 8 か月ぶりの減少であるが、消費支出の標準誤差率は 1.4% 程度（2000 年 10 月）であるから、これだけをもって直ちに消費が下向いたとは言えない。その後も、変化率の幅が小さな状況が 12 月まで続き、2013 年になって再び増加傾向となる。

一方、統計調査には調査対象者の把握漏れ、不在、報告拒否、解読不可、分類不可、虚偽報告、集計段階の符号付けの誤り、演算ミスなどにより、非標本誤差が発生する。これは標本調査のみならず全数調査についても言えることである。一般に、調査対象数が多くなれば誤差における非標本誤差の重みは標本誤差のそれよりも大きくなるので、大規模な公的統計における非標本誤差は大きな問題である。しかし、非標本誤差は標本誤差のように計測することは困難である。

非標本誤差には統計表でわかるものとわからないものがある。統計調査の調査項目には世帯数、住宅数、世帯人員数など、対象者本人でなくても外観や近所の聞きとりによりおおむね把握できる基本的なものがある。例えば、国勢調査では、調査票への記入を基本としつつも、やむをえず協力が得られなかった場合には、聞き取り調査で補うことにより人口と世帯数を極めて正確に把握している。このような基本項目の総量が把握されていれば、年齢、配偶関係、就業状況などの他の項目について、不在、報告拒否、解読不能及び分類不能のために把握できなかった数量は、統計表上で不詳として逆算により把握できる。しかし、不詳は、統計表に明示されているとは限らないので、明示されていなければ属性ごとに区分ごとの結果を加えて総数から引く作業が必要である。また、不詳数は項目ごとに異なる。就業状況には回答する人が、年齢や配偶関係については回答しないなど、調査対象により回答する項目が異なるからである。

公的統計には比較的正確なものが多いという特徴があるが、このために、却って不詳に気をつける必要がある。というのは、回収率の低い統計調査の多くで、結果は構成比や平均値という形でしか表示されないのに対し、回収率が高い統計調査では、人数、組織数、金額など、実数により表示されることが多いからである。このため、例えば、国勢調査の年齢別構成比を時系列で比較する場合には、年齢不詳を按分などにより補正する必要がある。国勢調査の年齢不詳は2010年に98万人と、総人口の0.8%であるが、これをどう処理するかによって将来推計人口の試算は大きく異なる。僅かな量でも数十年先の推計値には蓄積されて現れるからである。

また、統計表上の分類不能は非標本誤差ばかりではない。産業別構成については、時代を経るにつれてすき間産業が増えるために分類を変えなければ分類不能は増加傾向で推移する。すなわち、分類不能はその他の産業に該当するものも含んだものであり、記入の不備だけを集めたものではない。同じことが家計の収支項目など、分類が困難な項目についても言える。このため、分類を変えれば、過去との接続に注意する必要がある、先述の時系列の問題が生じる。そもそも、時系列のデータの振れをすべて誤差が要因とみなしてしまうのは問題である。世帯の支出はセールや制度改正など全国規模の特殊要因の影響を受けることは明らかであるし、輸出入額は国際的な政治、経済、環境などの微妙な変化で大きく変動することがある。時系列分析をするに当たっては、これらの要因を定量化するか、移動平均を行うなどにより長期的な趨勢だけを取り出す工夫が必要である。

統計表には現れない非標本誤差もある。世帯調査では、事業所の中に住居がある場合などの調査対象者の把握漏れや、虚偽の報告は聞き取りが通用しない誤差である。また、基本項目を数量として正しく捉えようとする統計調査は国勢調査だけで、他の

統計調査では、回収された調査票の枚数が報告書の総数に反映され、各項目の分類不能や無回答を不詳としている。ただし、標本調査では、基本項目について、全数調査や大規模標本調査などの結果をベンチマークとして利用することが多い。

なお、業務統計にも非標本誤差は存在する。非標本誤差には、制度の未適用、報告誤りなど収集段階で含まれるものと、転写誤り、計算誤りなど統計を作成する時点で含まれるものがある。前者の程度を把握するには、制度の適用・運用状況に関する知識が必要になる。社会保険については、国民年金の未加入者数と未納者数が統計報告により把握されている。雇用保険についても、労働力調査の結果と比較することにより、適用漏れを把握できる。しかし、犯罪統計、金融統計、貿易統計など比較的正確であると言われている業務統計についても、社会・経済の変化に伴う新しい型の現象や、いわゆる *underground* の実態を数量的に把握できないために、非標本誤差の程度は定かでない。

加工統計にも誤差はある。国民経済計算における統計上の不突合の他にも、国際収支表では、膨大な取引について様々な種類の報告書や資料を基に集計するため、必ずしも1つの取引に係る2つの計上資料が同一時期に入手できるとは限らず、また、評価方法のずれ等から同じ取引であっても資料によって金額が異なる場合もあり得る。このため、標準構成項目の中にこうした統計作成上の誤差を調整する項目として誤差脱漏が設けられている。

5.6 高度な利用に向けて

以上のように、公的統計を利用者が見て分析する際には気をつけるべき点が多々あり、上記の視点単独で整理できない場合、これらを組合せなければならない場合もある。また、これらに留意するだけでは済まない場合もある。しかし、公的統計を利用することに慣れていない人は、まず上記の視点でチェックをかけることが、誤用の防止に寄与するものと考えられる。詳しくは第2章で取り上げるが、上記は、統計を国際比較する際にも必要な視点である。というのは、類似の統計であっても、国により、目的、カバレッジ、概念、作成方法などは大きく異なることが普通だからである。公的統計の特徴のところ、国際比較が可能である点を述べたが、詳細は各国・地域の判断に任されているため、国間の比較に当たっては誤差を覚悟する必要がある。そもそも、歴史、文化、社会的背景事情により対象となる人や物が違い過ぎることもある。あまり違いが大き過ぎるものは比較の仕方を変えるべきもので、同一のものさしで測ったら誤解を生むこととなる。

これらの視点をまとめたのは公的統計の利用を制約しようという意図からではない。むしろ公的統計は多数の人手と多額の費用をかけて作成しているものなので、もっと

多方面で縦横に利活用すべきものとする。上記に掲げた利用上の注意は、簡潔にまとめると、原資料をしっかりと読んで理解するという学問に共通の基本的態度を守ることには過ぎないので、細かな誤りを恐れることなく、まずは公的統計の利活用を試みることで公的統計を理解し、有用な成果を得る第一歩であるとする。

さらに、公的統計の個々の対象者にとって統計の重要性を実感として認識することは必ずしも容易ではないので、対象者の協力を得るためにも印象的かつ有用な利用成果が存在することが望ましい。ある統計の利活用が増大、高度化すれば、対象者、協力者、作成者及び関係者の改善意欲も増し、その統計の質が向上することにより、更に利活用が増大、高度化する。統計の作成側と利用側の間のこのような相乗効果は、公共の利益の増進に大いに寄与する。

6 公的統計の課題

これまで、公的統計は網羅性、信頼性、正確性、比較可能性で優れていると、良い点ばかりを羅列したが、個々の統計の個々の項目にはこれらの特徴で包括するのに疑問が残る点もある。また、公的統計には長所ばかりではなく、改善すべき点もあるとする。例えば、迅速性、柔軟性、利便性という特徴も、統計として備えているに越したことはないが、公的統計を他の統計、他の国と比較してみると、これらの点については一般的に課題が残るのではないかと考える。

まず、迅速性について、公的統計の膨大なデータ量にかんがみればその作成期間は決して長いとは言えない。また、各種統計の公表時期は他の先進国と比較して遅いとは言えない。しかし、21世紀に入って、情報通信技術の発達により市場調査や民間統計の作成が急速化している中、特に、重要な判断指標だけを先行して作成すべきであるとの要求が強くなってきている。特に、公的統計には経済統計が多いが、経済対策には即断・即決が要求されることもあるので、このための判断材料である経済統計も即座に必要となるからである。また、民間の利用者は、高度の正確性の確保に比べ結果の早期入手に関心が向きやすい。

次に、柔軟性について、統計体系の変化は社会・経済の変化に比べてかなり遅い。例えば、産業の7割を占める第3次産業に関する統計は、少しずつ充実しているものの、決して十分とは言えない。個々の統計を見ても、何らかの事情で必要な内容が突然発生しても即座に統計に盛り込めるような柔軟性は低い。これは、公的統計が国民の負担に依存する国民の共有財産なので変更には多数の合意形成が必要なこと、時系列としての利用が重視されるので変更そのものが馴染まないことなどが理由である。

最後の利便性の欠如は、社会・経済の実像を公的統計から把握しようとする場合に起こる問題点である。公的統計は体系としては網羅性に優れているものの、個々の統

計の多くは個々の行政需要によって誕生したことから、個々の行政需要を満たすことを第1の目的として作成されている。つまり、公的統計は、純粹に統計という視点から全体を眺めると、合理的、整合的、総合的に作成されていない。例えば、国民経済計算や産業連関表などの加工統計は、個々の統計調査の結果を組合せて作成されているが、統計調査により対象範囲、作成方法などが異なるために推計部分が広がる。また、事業所・企業については、基本的な事項については経済センサスなどで把握できるが、少し詳細な事項を広く分析しようとするならば、産業ごとの様々な統計調査の結果を合わせる必要が生じ、この場合にも推計域を拡大しなければ全体像を描きにくい。このことは、そもそも行政が縦割りであることから付随して起こる問題である。類似の問題は国と地方公共団体との間でも起きている。例えば、県民経済計算は、都道府県が作成しているが、国が作成したマニュアルを参考にしているために、細部に踏み込まなければ分析に大きな支障はない。しかし、県民経済計算は国民経済計算の内訳として存在するほうが合理的である。いずれも、統計としての役割を重視するならば、本来最初に全体の姿をどう描くかという視点から個々の統計を設計すべきであり、極論すれば枝葉のある1つの統計として企画すべきであるが、そうっていない。このために、社会・経済の実像を描こうとするならば、公的統計は必ずしも使い勝手が良いとは言えないし、幾分正確性を犠牲にしなければならない。また、統計調査の個別データを学術目的で使用できる場合が諸外国と比較して限定されている、報告書や統計表の様式、配列、表記法などが統計ごとに異なるなど、行政目的以外の利用にとって便宜が十分図られているとは言えない。

このように、公的統計には改善すべき問題点が多々ある。これらについては様々な観点から多数の指摘があり、また、公的統計の作成機関や調整機関が社会・経済の変化を踏まえつつ計画を立て、これらの解決・改善に努めているところである。このような見直しに当たっては、利用者の視点が大きな手掛かりになる。

もっとも、公的統計の問題は、利用者から見えるところにあるとは限らない。現在の公的統計は、調査環境の悪化、予算人員の制約、グローバル化に伴う対象把握の困難等多くの問題に直面しており、それは日本だけでなく世界共通の問題となっている。それにどのように対処するかは、現代の公的統計の重要な課題である。

第2章 公的統計の国際基準の意義と変化

第1章では、日本における公的統計の現代的意義について、過去から変遷を踏まえ、解説し、利用における留意点を述べたが、海外においても、公的統計の重要性は強

まっている。各国とも、公的統計は社会・経済の状況を把握するための重要な根拠になっていることから、この作成に相当なリソースをかけている。グローバリゼーションが進展する中で、特定国の統計の欠落は全体像の把握を困難にすることから、国際機関が中心となり、重要な統計について、各国ともに国際比較が可能な統計の作成・提供が可能となるように勧告、基準、ガイドライン等の公的統計の枠組みを企画・調整・推進している。世紀をはさんでこのような国際統計基準に変化が生じている。20世紀は、統計の作成・提供方法の原則を設けることが主であったが、21世紀になり、目標に照らして統計そのものあるいはアウトカムの評価にルールを設けることが主になった。国際統計基準についてこのような傾向があることを指摘する論文は、筆者の知り得る限り海外を含めて存在しない。第2章では、このような国際統計基準の意義と変化について具体例を含めて論じる。

1 背景

統計の国際的動向を把握するには、各国それぞれの統計事情を知ることが必要である。また、地理的、歴史的、文化的な事情は各国とも様々であることから、必ずしも国際的な事情が各国にそのまま当てはまるとは限らない。しかし、気象、経済、病気、安全などの地球規模の課題を解決するには、各国が共通の目標のもとに統一された方法で国際比較可能な統計を作成・提供することが必要である。国際統計基準はこのための指針となるものである。国際統計基準は、強制力を有するものではないが、対象国の範囲内で多くの専門家が議論を重ねて合意を得たものであることから、実測可能性という制約の中で域内の統計の理想像を示すものである。また、柱の立て方や分類の仕方などは公的統計の体系を考えるための重要な手掛かりになる。

国際的に見て、公的統計に関する最も普遍的な基準を決定している機関は国際連合であり、国際連合統計委員会が毎年開催されている。しかし、国際連合には開発途上国が数多く含まれていることから、そこでは統計に関する先端的な取組はあまり反映されない。そのような取組を反映した基準はむしろ EU、OECD、IMF など特定の国や分野を対象とする国際機関や、国際連合でも地域機関の欧州経済委員会 UNECE

(United Nations Economic Commission for Europe) などで検討されており、域内の同意が得られたところで基準、マニュアル、ガイドラインなどにまとめられる。それらが普及したところで、今度は開発途上国も含めた適用を目指して国際連合統計委員会で議論されるという段取りが一般的である。そのような段取りを踏まえると、国際統計基準はとくに欧州の事情に傾倒しがちである。したがって、日本としては、日本の事情も国際統計基準に適切に反映されるよう必要な働きかけを早期かつ強力に進めることが重要となる。

21世紀に入り、このような国際的な決め事に変化が生じている。20世紀は、統計の作成・提供方法の原則を設けることが主であったが、21世紀になり目標に照らしての統計そのものあるいはアウトカムの評価にルールを設けることが主になった。

この背景には以下の事情が考えられる。

まず、グローバリゼーションの進展により、統計の欠落に伴う全体像や比較可能性の喪失が以前よりも深刻になってきたことがある。

また、データ量の拡大、情報通信技術の発展、統計能力の向上などにより、開発途上国でも統計の作成が可能になったことが挙げられる。

さらに、ビッグデータの時代において、政策目標を達成することとそれを評価するための統計が、より密接になっていることが挙げられる。

なお、本章で示す勧告、基準、ガイドライン等は、データの収集から分析に至る実務を対象にしたもので、自然科学に見られるような対象の測定・分析手法を理論的に解説した手引書及び研修用に編集したマニュアル等を除外している。さらに、作成機関は国際機関とし、特定国、大学、民間機関等によるものも除外している。

本章では多数の勧告、基準、ガイドライン等を取り上げるが、それらは一般への普及を目指す書類であることから、その名称で検索するだけでインターネットから入手が可能なものである。したがって、本章に関して参考文献に掲げる文書は、本章の中で位置付けを示すだけでなく、当該文書の中から何らかの方向性を示唆する内容を引用した場合に限定する。

2 統計の作成・提供方法に関する基準

本節では、公的統計の作成・提供方法に関する国際基準の概要を示す。個々の統計にはそれぞれ作成・提供方法に関する基準が存在するが、ここでは、統計を横断的に方向付ける基準のみを取り上げることとする。

2.1 公的統計の基本原則

公的統計の意義と役割を10の原則により概括的に定めたものが United Nations[6] Fundamental Principles of Official Statistics (公的統計の基本原則) である。同原則は、まず、1992年に UNECE で採択された後、1994年に国際連合統計委員会で採択された。その中で公的統計の役割が政府の政策決定を超えた社会・経済の情報提供にあるとされたことが、2007年の日本の統計法の全面改正に影響を与えた。同原則に関連して、2006年に Principles Governing International Statistical Activities として条文ごとにグッドプラクティスが示された。その後、同原則については、前文のみが改定され、2013年に国際連合統計委員会で採択されるとともに、

2014年に国際連合総会で決議された。なお、同原則の第9条には、国際的な概念、分類及び方法を各国統計機関が用いることは、すべての公的レベルの統計システムの整合性及び効率性を向上させると書かれている。

同原則を確保するための実践的なツールとして、UNECEはGSBPM (Generic Statistical Business Process Model)、GAMSO (Generic Activity Model for Statistical Organizations) 及びGSIM (Generic Statistical Information Model) を定め、各国での適用を呼びかけている。GSBPMは2009年に作成され、当初はメタデータの発展を支援するものであったが、徐々に適用範囲が広まり、版を重ねるごとに、データ源にかかわらず、集計の枠組みの調整、ソフトウェアの共有、後述する統計品質の確保など、統計の作成・提供全般を支援する1つのモデルとなっている。GAMSOは、2015年に作成され、統計組織の活動全体を支援するモデルで、GSBPMの他、戦略・リーダーシップ、能力管理、組織的支援から構成される。GSIMは2012年に作成、2013年に改定され、他のモデルと連動しながら、作業工程におけるGSBPMの両側の部分、すなわち、データ収集とデータ提供について、情報を整理・区分する概念モデルとなっている。

公的統計を含めた統計全般に係る規格として、国際商取引に関する統計的方法の適用に関して、国際標準化機構第69技術委員会 (ISO TC 69 “Application of Statistical Methods”) が1960年頃から国際標準化活動を行い、多くのISO規格を発行している。

また、分野ごとに統計についてこのような基本原則を定めることもある。臨床試験はその典型であり、1996年におけるGood Clinical Practice: Consolidated Guidelineを踏まえて1998年にICH (日・米・EU三極医薬品規制調和国際会議) における合意に基づきStatistical Principles for Clinical Trialsが定められた。同基本原則は、臨床試験から得られる結果の偏りを最小にし、精度を最大にすることを目標として臨床試験における統計的原則について記載したものである。

2.2 品質

統計のQuality (品質あるいは質) については、1970年代まではデータの収集から集計までの統計作成過程における統計精度管理、すなわち、製造業工程の狭義の品質管理 (Quality Control) 分野の精度管理と同様の意味で用いられてきた。しかし、1970年以降に顧客要求への合致を第一目的としたQM (Quality Management) が発展し、1980年代にISO9000規格で国際合意されたQMP (Quality Management Principle) 8原則の第1原則である顧客のためのQMが合意され、Qualityの概念は、公的統計の世界でも利用者満足度という視点を含め、対象を制度に広げた総合的な観点で用い

られるようになってきた。これをいち早く目標として取り入れたのはカナダであり、1985年にはガイドラインを作成している。

統計の品質に関する framework（枠組み）は、1990年代から EU 各国を中心に検討が進められ、2000年前後にいくつかの国で作成された。また、EU、OECD、IMF といった国際機関も 2003年に同様の枠組みを策定している。これらは、relevance（妥当性）、accuracy（正確性）、timeliness（適時性）といった統計について目指すべき規範をいくつか掲げ、それらのチェック項目を具体化したものであり、統計のチェックのポイントを示すだけのもの、評価の指針にもなるもの、評価の過程も書き込んだものなどがある。

EU の例を挙げると、2005年に欧州統計計画委員会が採択し、欧州委員会が欧州連合理事会と欧州議会に勧告した CoP（European Statistics Code of Practice）が、品質評価の広範な概念的枠組みを提示し、欧州統計体系の組織環境、統計作成過程及び統計結果の基準をまとめたものとなっている。CoP の制定を受けて、2009年に欧州連合理事会と欧州議会は欧州統計に関する規則を改定し、加盟国は Eurostat（欧州委員会統計局）に品質報告書を提出することが義務化された。品質報告書については、2003年に How to Make a Quality Report が作成されていたが、規則の改定を踏まえて 2009年に ESS Standard for Quality Reports として統計作成過程及び統計結果について包括的な品質報告をするよう勧告し、2014年に改定した。また、Eurostat が 2009年に作成した ESS Handbook for Quality Reports では、品質報告の詳細なガイドラインと事例を紹介し、2014年に作成した ESS Guidelines for the Implementation of the ESS QPI（Quality and Performance Indices）では、標準誤差や回答率など品質を評価する指標を解説している。

国際連合も、NQAF（National Quality Assurance Framework）について、generic NQAF template（骨子）を 2010年国際連合統計委員会において議論し、2012年同委員会で採択した。

2.3 データ収集方法

公的統計のデータ源としては、統計調査、行政記録、地理情報データ、民間データなど様々なものが想定される。その収集方法については、総括的には先述の GSIM が存在し、統計調査については古くから分野や調査ごとにマニュアル等が整備されている。

分野横断的な指針としては、データ収集において男女別把握が重要だとして 2010年に UNECE が Developing Gender Statistics: A Practical Tool を作成し、男女別把握・提供の意義、方法等を解説している。

ISO 規格の中には、ISO3534-4 のように標本調査の用語を規定した国際規格も各国投票により成立している。また、ISO TC 225 は、市場調査、社会調査、世論調査などの調査プロセスに関する第3者認証規格を策定しており、ISO20252 による調査機関が実施する調査プロセスの妥当性・有効性を保証する第3者認証は、日本でも実施されている。

2.4 行政記録

統計を作成するために、かつては統計調査でデータを得る手法が一般的であったが、近年各国とも行政記録の活用を視野に入れている。Eurostat は、2003 年に行政記録に基づく統計の品質評価の枠組みを改善するために **Quality Assessment of Administrative Data for Statistical Purposes** を統計の品質評価 Working Group に提出した。2011 年には UNECE[7]が **Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices** を作成し、**Administrative Sources** の定義を統計機関による1次データと非統計機関による2次データに区別し、2次データについてはこれまで狭く公的機関によるものだけとされていたが、今後は広く民間機関によるものを含めることとし、それらを利用することによる利点、利用のための枠組み、問題点・解決策等を解説している。

一方で、行政記録については、統計調査によるデータとは異なり、統計作成機関が一括して管理できないことによる課題がある。2009 年には UNECE[8]**Principles and Guidelines on Confidentiality Aspects of Data Integration Undertaken for Statistical or Related Research Purposes** で、データ接合は統計及び関連目的のためだけに国家統計機関によって行われるべきこと、公開された透明な方法によって行われるべきこと等としている。

行政記録は企業や事業所を把握する場面で先行している。これらのすべてを把握しようとする場合、日本ではまだセンサスが主流であるが、EU では、まず、行政記録からデータを手し、それで把握できない場合に統計調査や民間データ等を用いる。これらのデータはビジネスレジスタに蓄積される。EU は 2003 年に **Business Registers: Recommendations Manual** を策定し、2010 年に改定している。同マニュアルは、統計目的のビジネスレジスタのための共通基盤を確立するとした規則（**BR Regulation**）を踏まえ、同規則を解釈し、ビジネスレジスタの発展を目指すために策定されている。他方、企業統計においてビジネスレジスタの利用を促進しているマニュアルもある。ビジネス動態統計に関する 2007 年改定版マニュアル **EU and OECD[9]Eurostat-OECD Manual on Business Demography Statistics** では、データ源としてビジネスレジスタとその他を挙げている。後者には統計調査の他に税と事

務的チェック（行政記録、直接連絡、インターネット等）がある。同マニュアルでは、ビジネス動態データの収集のための統計単位として企業を推奨している。

2015年のUNECE[10]Guidelines on Statistical Business Registersでは、経済センサスはコストや対象把握の困難性からビジネスレジスタを構築する手段として推奨できないが、信頼できる行政記録が存在しない場合には周期的な経済センサスの実施は適当であるとしている。

北欧にはレジスタにより人口を把握している国があるが、UNECEは2007年にベストプラクティスをRegister-based Statistics in the Nordic Countries: Review of Best Practices with Focus on Population and Social Statisticsにまとめている。

2.5 統計の基準年等の改定

日本では、国勢調査が西暦0及び5の年に行われ、数多くの周期調査が5年ごとに実施されていることから、多くの指数の基準が西暦0又は5の年となっている。多くの国でこのような統計の改定方針が定められているが、国際的なガイドラインも存在している。

OECDとEurostatは2006年にOECD短期経済統計Working Groupを設立し、ウェブ上にGuidelines on Revisions Policy and Analysisを立ち上げ、政策分析結果の作成・利用、利用者のニーズを効果的に支援する改定方針の策定に向けたベストプラクティス等を掲載している。

また、2013年にEurostat[11]ESS (European Statistical System) Guidelines on Revision Policy for PEEIs (Principal European Economic Indicators)の中で、EUの共通改定方針の原則として、ESSを所管する統計機関は統計の改定に関し各統計分野の共通戦略を定義し、質的・量的情報を提供すること、作成過程で誤りや不測の事態が生じた場合にはあらかじめ決められていない改定を実施すべきことなど書いている。

基準改定については、個々の統計指標ごとに指標の性格を踏まえてマニュアル等に必要事項が記載されている。

2.6 分類

統計を比較するには、国際的にある程度統一された分類に基づいていることが前提となる。これらの分類は、単に種類や区分を示すだけでなく、各統計のデータ収集、加工、提供の指針となるとともに、当該統計と他の統計との統合を推進する役割を果たすものとなっている。

まず、国際連合は、1969年に Standard Country or Area Codes for Statistical Use として国や地域の名称を定め、最近では1998年に改定している。

また、国際連合は、産業分類に相当する ISIC (International Standard Industrial Classification of All Economic Activities) について1948年に経済社会理事会で初版を採択している。その後、統計委員会で1958年に改定第1版を採択し、最近では2008年に改定第4版を採択している。

さらに、1989年に商品分類に相当する CPC (Central Product Classification) を国際連合統計委員会で採択し、2015年に改定している。このうちエネルギーについては2011年に SIEC (Standard International Energy Product Classification) として改定している。貿易分類については、標準的な分類である SITC (Standard International Trade Classification) を1950年の国際連合統計委員会で採択し、2006年に改定している。また、SITC と対比して産業供給、資本設備などの広範な経済分類による分析が可能になるように1971年に BEC (Classification by Broad Economic Categories) を採択し、2003年に SNA (System of National Accounts、国民経済計算体系)、SUT (Supply and Use Table) 等での利用が可能な大分類として改定している¹⁴。

支出については1999年に Classification of Expenditure According to Purpose を定め、個人については COICOP (Classification of Individual Consumption According to Purpose)、法人については COPP (Classification of the Outlays of Producers According to Purpose)、政府については COFOG (Classification of Functions of Government)、非営利団体については COPNI (Classification of the Purposes of Non-Profit Institutions Serving Households) で具体的な分類を定めている。これらは1993SNA の概念と定義に沿ったものであり、例えば、COICOP は世帯調査に適用することが実用的でない部分もあるとしている。

さらに、生活時間について2004年に ICATUS (International Classification of Activities for Time-Use Statistics)、犯罪について2015年に ICCS (International Classification of Crime for Statistical Purposes) を定めている。

分類は統計の共通語であることから、国際連合が定めている場合が多いが、必ずしも統計部が主導しているとは限らない。

職業については産業も含めて ILO が古くから分類を作成している。ILO は、1923年に第1回国際労働統計家会議に報告された Systems of Classification of Industries and Occupations を踏まえて産業・職業分類を決議した。1952年に International

¹⁴ その後も、サービスを含める等の改定版が2016年国際連合統計委員会で採択され、ISIC、CPC 等との対応表を完成させることとなっている。

Classification of Occupations for Migration and Employment Placement を作成し、1958 年には、ILO 決議に基づいて職業分類に相当する ISCO (International Standard Classification of Occupations) を作成し、1968 年、1988 年、2008 年に改定している。雇用者、自営業主などの従業上の地位についても、ILO は ICSE (International Classification of Status Employment) を 1958 年に設定しているが、1993 年に改定して以来、現在まで改定していない。

貿易統計については、WCO (World Customs Organization) が 1988 年に HS (Harmonized Commodity Description and Coding System) を定め、4～6 年に 1 度改定している。HS は 6 桁コードで認識される約 5 千商品群から成る。

非関税貿易については、UNCTAD が 2012 年に International Classification of Non-Tariff Measures をまとめ、技術的手法 (衛生又は環境保護手法) と商業政策の手段として伝統的に使われるその他の手法 (分配、価格制御、輸出規制等) を分類している。

農畜産物については FAO が Commodity Classification を定め、ホームページに掲載している。

医療、健康については WHO がいわゆる健康統計分類 family を定めている。参照分類として、疾病に関する ICD (International Statistical Classification of Diseases)¹⁵、障害に関する ICF (International Classification of Functioning, Disability, and Health) を定めるとともに、健康の支障について ICHI (International Classification of Health Intervention) を作成しつつある。また、派生分類として、腫瘍学について ICD-O (International Statistical Classification of Diseases for Oncology)、精神・行動疾患について The ICD-10 Classification of Mental and Behavioural Disorders、歯科・口腔科学について ICD-DA (Application of the ICD to Dentistry and Stomatology)、神経学について ICD-10-NA (Application of the ICD to Neurology) 及び ICF の子供・若者版がある。さらに、関連分類として、主な治療について ICPC (International Classification of Primary Care)、外傷要因について ICECI (International Classification of External Causes of Injury)、解剖学、治療学及び化学について The ATC (Anatomical Therapeutic Chemical) Classification System with Defined Daily Doses 等がある。

教育に関する ISCED (International Standard Classification of Education) は UNESCO が 1970 年に作成し、1997 年、2011 年に改定した。

¹⁵ ICD は、1948 年に初版が定められ、2016 年現在では 1990 年に定められた第 10 版が最新で ICD-10 と呼ばれる。2018 年までに ICD-11 が策定される見込みである。

なお、UNECE は 2011 年に *Classification of Statistical Activities* という国際機関が統計活動を進めるための分類を取りまとめた。

この他にも、分類については統計ごとに概念、定義などとともに記述された文書が数多く存在する。

2.7 季節調整

統計の季節調整に関する基準は EU で検討されており、2009 年に季節調整に関するガイドラインの初版である Eurostat[12]ESS Guidelines on Seasonal Adjustment で、季節調整の利点、留意点、コスト及びリスクを紹介している。手法については EU 内で利用されている X12-ARIMA 及び TRAMO (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers) -SEATS (Signal Extraction in ARIMA Time Series) に焦点を当て、留意点については、事前調整の目的、モデル選択、改定方針などの項目ごとに必要な確認すべき点を一覧表にまとめている。2015 年に改定された Eurostat[13]ESS Guidelines on Seasonal Adjustment では、初版と異なり、手法については、本文で限定せずに、付録の中で、外れ値やカレンダー要因の調整に regARIMA モデルを使うことを推奨するという概括的な説明をした上で、推奨すべき手法としては SEATS のような Signal Extraction に基づくパラメトリック法、センサス局法 II X11 family 及び X13-ARIMA-SEATS のようなあらかじめ決められた移動平均に基づくセミパラメトリック法であると述べている程度である。

2.8 2次利用

マイクロデータの提供については、公的統計の基本原則第 6 条に基づき、統計機関が統計作成のために収集した個別データは、自然人に関するものであれ、法人に関するものであれ、厳重に秘匿されなければならない、統計目的以外に用いてはならないことが原則となっている。2007 年に、UNECE はそれを踏まえて CES (Conference of European Statisticians) [14]Managing Statistical Confidentiality & Microdata Access: Principles and Guidelines of Good Practice をまとめている。CES[14]は、各国の手法を統一化し、研究者集団によるマイクロデータの入手を容易にするとともに、各国のマイクロデータの提供手法を改善するために、主要原則のほか、法整備、研究機関の支援方法、国家統計機関と研究者との間の調整、マイクロデータの提供に関する管理的問題等を説明している。また、各国のマイクロデータの提供手法のグッドプラクティスについて、長所と短所を紹介し、併せてそれらがなぜグッドプラクティスなのかの理由を説明している。さらに、ESS Net SDC (Statistical Disclosure Control) は、2009 年に、マイクロデータによる研究に基づく成果物のチェックについて

Guidelines for Checking of Output Based on Microdata Research を作成している。加えて、2010 年に利用者のニーズと回答者のプライバシーのバランスへの対処技術について Handbook in Statistical Disclosure を作成している。

OECD はマイクロデータの国際的利用を可能にするための取組を進めている。また、マイクロデータに限らないが、公的統計を含む公的資金による研究データについて、2007 年に OECD Principles and Guidelines for Access to Research Data from Public Funding をまとめている。

2.9 データエディティング

データエディティングについては欧米を中心に研究が進められており、UNECE が 1994 年に Statistical Data Editing と題して方法と技術に関する解説書を作成し、1997 年に改定している。また、2000 年には用語集として Glossary of Terms on Statistical Data Editing をまとめるとともに、UNSC (United Nations Statistical Committee、国際連合統計委員会) and CES[15]Evaluating Efficiency of Statistical Data Editing: General Framework Evaluating Efficiency of Statistical Data Editing を作成し、その中で、典型的な統計調査においてデータエディティングに要する費用は 40%にもなるとした上で、データエディティングの成果の測定方法を説明している。2006 年には、その視点を踏まえてデータの品質に焦点を当て、Statistical Data Editing を改定している。

2.10 統計提供

統計については、正確かつタイムリーに作成するだけでなく、結果を的確かつわかりやすく表現することも重要である。例えば、エネルギー統計については、IEA (International Energy Agency) が OECD などと共同して 2005 年に電気、ガス、石油などの統計上の概念などの基本的事項を解説した Energy Statistics Manual を作成した。統計の政策利用の実例を示したものもある。ESCAP は 2013 年に Making Data Count を作成し、政策及び意思決定における公的統計の高度利用のベストプラクティスを提示した。

結果の表現方法について解説したガイドラインも存在する。UNECEは2000年に Guidelines for Statistical Metadata on the Internetを作成し、インターネット上の統計メタデータについて最低限の基準を定めた。また、2009～2012年に4刊から成る Making Data Meaningfulを作成している。この中のUNECE[16]Making Data Meaningful, Part 1, A Guide to Writing Stories about Numbersは、統計についてストーリーを語る理由として、①国家統計機関の使命は国民に国家の人口、社会・経済、

文化を知らせることであること、②政府及び国民にデータの妥当性を証明する手段になること、③統計の利用可能性は、調査回答者がデータは自分達が居住する世界を映す鏡という重要な目的に叶うことを理解する程度にも依存することを挙げている。ストーリーに書き込む内容としては、現在の興味、日常生活との関連、特別な集団との関連、個人的経験、休暇、現行イベント、カレンダーのテーマ、新発見、通常のシリーズとしている。同書は、文体、表の作り方、ビジュアル化等について、良い例、悪い例を比較しながら、統計作成者がデータを明確かつ意味のある方法で表現する実践的方法をわかりやすく解説している。

さらに、高度な統計提供・説明手法について、OECDは2007年に *Data and Metadata Reporting and Presentation Handbook* を発行した。

UNECEは1995年にデータ及びメタデータのモデル化を解説するため、*Guidelines for the Modeling of Statistical Data and Metadata* を作成した。UNECEは *The Common Metadata Framework* として、A. 組織的管理及び統治に関する統計的メタデータ（2009年）、B. メタデータの種類、概念、標準、モデル及び登録（2011年）、C. メタデータ及び統計業務の過程（2013年）、D. 実行（各国及び国際機関の経験）ごとに解説をまとめた。Cに関しては先述の GSBPM が存在している。

メタデータについては、1990年代から社会学者を中心にデータの入手から分析結果にまとめるまでの必要事項を文書化するための DDI (*Data Documentation Initiative*) が存在しており、2003年には連合組織である *DDI Alliance* が設立された。統計数値及びメタデータを提供できるようにするだけでなく、それらを組織間で交換する際の効率性及び比較可能性を高める取組も進んでいる。2002年に国際機関の間で取り組むことが協定された *SDMX (Statistical Data and Metadata eXchange)* は、組織間で統計数値及びメタデータを交換する際の共通様式であり、これにより統計数値及びメタデータの交流が促進され、そのための技術及び基準の利用が効率化し、組織の負担が減り、さらには、統計数値及びメタデータの利用可能性が高まる。国際機関から構成される *SDMX* 共同体は *User Guide*、*Content-oriented Guidelines* などを作成し、適宜更新している。

地図上にデータを表示することについては、EUが2007年に空間情報基盤を設立することについて規則を設け、用語の定義、メタデータの内容、ネットワークサービスの機能、データ共有の範囲及び条件、構成国の報告事項等を定めている。他方、国際連合は、2002年に6か国語で比較可能な地理用語集 *Glossary of Terms for the Standardization of Geographical Names* を作成した。また、統計の国際比較にとって地理用語（市、郡、山、湖、空港、漁獲域等）の基準を設けることは重要だとして、2006年に *Manual for the National Standardization of Geographical Names* を、

2007年にアルファベット等の略式記号の比較を示した **Technological Reference Manual for the Standardization of Geographical Names** を作成した。

3 統計結果に関する基準

国際統計基準をもとに国際協力が進められ、各国での取組も進み、多くの国で基本的な統計を作成するための基盤が整ってきた。また、情報通信技術が発達し、データ量が増え、それを解析するソフトウェアが進歩した。

このような背景の中で、21世紀に入り、国際機関は、国際統計基準を定めるだけでなく、貧困縮減や経済成長などの世界共通の目標の実現のために広範な分野に及ぶ統計そのものを各国から提供させる動きが表れている。提供は義務でないとは言え、それに同意することにより逐次国際機関から統計結果そのものを求められることになるため、指標を定める際には国際機関、各国ともに十分な議論を重ねる。また、これらの統計は政策目標と密接に関わりを持っていることから、統計部門が政策部門と協働し、国際機関においても高いレベルでオーソライズされる。

3.1 持続可能な開発目標

人口・社会統計分野の国際基準としては、社会人口統計体系が存在していたが、全体の枠組みを示すのみで、個々の指標については各国が個別に整備を進めていた。しかし、最近この分野で大きな変革が現れた。

1980年代には、多くの開発途上国で市場経済メカニズムに依拠する構造調整政策を通じた開発手法が採用された。しかし、この手法は順調に進むとは限らず、貧困の悪化をも引き起こすことがあった。このため、1990年代に貧困に対する関心が高まり、1995年の世界社会開発サミットでは、人間中心の社会開発を目指し、世界の絶対的貧困にある人口を半減させるという目標が提示された。これが発展して、2000年に開催された国際連合ミレニアムサミットにおいて2015年を達成期限として8つの目標であるMDG (Millennium Development Goals) と21のターゲットと60の指標が設定された。MDGについては、各国が指標の作成を可能にするような指針が必要であった。このため、国際連合経済社会局持続可能開発部が **Indicators of Sustainable Development: Guidelines and Methodologies** を作成・改定するとともに、国際連合人間居住計画が2009年に **Urban Indicators Guidelines** としてスラム住民の課題に対処するためのMDG指標を整理するなど、課題ごとに横断的なガイドラインも作成された。

MDGの達成年である2015年には、国際連合持続可能な開発サミットにおいて、持続可能な開発のための2030アジェンダが定められ、2030年を達成期限として17の

持続可能な開発目標である SDGs (Sustainable Development Goals) と 169 のターゲットと 229 の指標が設定された。ターゲットの中には、「現在 1 日 1.25 ドル未満で生活する人々と定義されている極度の貧困をあらゆる場所で終わらせる。」「世界の妊産婦の死亡率を出生 10 万人当たり 70 人未満に削減する。」などの統計数値で評価すべきものが数多く盛り込まれている。しかし、実際に指標を整備するとなると様々な困難に直面する。2016 年に開催された国際連合統計委員会では、技術的な精緻化を条件に実務的開始点として SDG の指標一覧が同意されたが、指標の発展と改善に向けて更に検討を重ねることとなった。

SDGs の策定に先立ち、これに関連する地域の取組として、UNECE は 2009 年に Measuring Sustainable Development を作成し、2013 年に CES Recommendations on Measuring Sustainable Development を定めた。しかし、これらには、幸福や資本（自然資本、人間資本、社会資本）の測定など SDGs には含まれていない方向も含まれていた。ESCAP 地域でも、2013 年 ESCAP 統計委員会で採択された人口社会統計の指標に関する Core Set に沿って、構成国に対して Core 指標の整備を求めている。

3.2 主要経済・金融指標

IMF は経済・金融指標に関する 2 種類の公表基準を定めた。1 つ目は 1996 年に定めた SDDS (Special Data Dissemination Standard) であり、国際資本市場にアクセスを持っているか、いずれ持つかもしれない国々に指針を与えるために設立された。2 つ目は、すべての国々がより信頼性の高いデータを提供することを支援するために 1997 年に定めた GDDS (General Data Dissemination System) である。GDDS が統計システムの改善を目的としているのに対し、SDDS はすでにデータの質において高い基準を満たしている国々がデータ公表基準を自主的に守ることを目的としている。両方とも参加は各国の自由意志によるが、SDDS では一旦加入すると遵守の義務を負う。具体的には、GDP、就業者数、為替レート、平均株価などの基本的な指標を逐次 IMF に報告する義務がある。SDDS については、2008 年に生じた世界金融危機を踏まえ、より迅速かつ正確に経済・金融情勢を把握することができるように、2012 年に SDDS プラスとして指標の見直しを行った。各国がこれに参加する場合、参加後 5 年以内に一般政府収支、証券データ等についてそれぞれ定められた公表頻度・公表ラグの基準を満たすことが必要である。日本は 2016 年にアジア太平洋地域の中で唯一 SDDS プラスに加入した。

OECD も、1985 年から MEI (Main Economic Indicators) を構築し、各国の協力を得て構成国といくつかの非構成国の経済関係指標を収集・公開している。ESCAP 地域においては、人口社会統計に関する Core 指標の整備と同様に、経済統計につい

でも Steering Group で検討し、2010 年 ESCAP 統計委員会で採択された経済統計に関する Core Set に沿って、構成国に対して Core 指標の整備を求めている。

3.3 その他の指標

国際機関が統計そのものを各国から提供させる動きは分野ごとにも存在する。

例えば、情報通信技術については、以前から ITU が Handbook を作成し、指標の作成を促進してきたが、Core ICT Indicators が定められ、2007 年に国際連合統計委員会で採択され、2012 年に改定されている。

4 今後の見通し

情報通信技術が進歩し、データ量が増大し、その処理速度が迅速化する中で、各国にとっても、国際機関にとっても、統計をはじめとする数値データを根拠として政策を立案、実施、評価することが普通になってきた。このため、従来のような統計のために統計の品質を向上するための基準、マニュアル、ハンドブック等を改善、整備するだけでなく、政策評価に使用すべき統計そのものを定め、それを追跡するという動きが一般的となってきた。しかも、その対象となる統計の種類は増大かつ細分化し、それを提供する各国の負担も重くなっている。また、提供元は必ずしも統計部門とは限らず、政策部門、ひいては民間部門にも広がっている。今後も情報通信技術が進歩することにかんがみると、このような流れが強まることが予想される。

しかしながら、これをもって従来型の国際統計基準の必要性が弱まるということではない。国際機関が各国に統計を求めれば、それをどう作るかという課題が発生し、具体的な側面では各統計とその基盤の品質を高めることの重要性も強まっていくので、従来型の統計の共通基盤や個別統計に関する国際統計基準に対する必要性も強まるからである。

他方、国際統計基準の中で、公的統計の基本原則を確保するための支援ツールについては、GSBPM、GAMSO、GSIM、SDMX、DDI など類似の枠組みが、それぞれの経緯を踏まえて独自に発展しているので、各国が適用するツールとしての普及を目指すのであれば、わかりやすい形に整理・集約されることが望ましい。

なお、本章で紹介した基準、マニュアル、ハンドブック等は社会・経済情勢の変化、各国と国際機関の取組状況等により常に変化するものなので、適宜メンテナンスをすることが必要なものである。

第Ⅱ部 統計作成技法の高度化研究

第Ⅱ部においては、統計の作成技法の高度化に向けた研究を紹介する。

第1章の季節調整法の研究では、労働力調査、CPI、家計調査の結果について、X11と比較してX11-ARIMAを適用するほうが原系列の追加に伴う推定季節指数及び最新季節指数の改訂幅が小さく、季節調整値に安定性があることを実証している。また、個人企業経済調査の結果は季節調整されていないが、それらに各種の季節調整法を適用し、AIC (Akaike's Information Criterion)、安定性 (原系列の追加に伴う季節調整値の改定幅が小さいこと)、適切性 (同一季節調整を再適用したときの改定幅が小さいこと) により季節調整法の候補をX12-ARIMAの(0,1,1)(0,1,1)及びDECOMPの定常変動次数0であるとしている。さらに、完全失業率、CPI、家計調査消費支出の結果にX13-ARIMA-SEATSを適用し、X11を選択するよりもSEATSを適用するほうが、安定性と適切性が高いことを実証している。

他方、一時点の構造を分析するためには、センサスや大規模標本調査の結果を利用することが一般的である。それらに含まれない項目について分析をする場合には、その他のデータ源に頼らざるを得ないが、精度確保についてはリスクが伴うことになる。しかし、そのような場合にも、大規模統計調査の結果を利用すると、よりの確な分析が可能になる。その典型的な事例として、第2章では、国勢調査の結果によるJGSS (Japanese General Social Surveys) の結果の改善に係る研究について説明する。JGSSの個別データに対し、地域、男女、年齢階級ごとに合算して国勢調査の人口を割り、比推定の乗率として合算すると、属性別結果の多くで偏りが縮小する。このような研究はビッグデータの作成技法の高度化にも資すると考える。

第1章 季節調整法の研究

季節調整法には様々な方法が存在し、系列すべてに当てはまる最適な方法は存在しない。総務省統計局が作成する統計に適用する季節調整法についても、統計により、系列により個々の特性を踏まえて最適と判断される方法が適用されているが、どの時代においても技術的發展を踏まえた改善の余地が存在する。本章では、総務省統計局が作成する統計の季節調整法の改善に向けた研究を示す。

1 X11-ARIMA の適用に向けた検討

総務省統計局では労働力調査、CPI、家計調査に対して米国センサス局法によって季節調整を行っている。1980年代にはこのうちX11を適用していた。X11は移動平均と割り算の繰り返しによって季節指数、季節調整値及び向こう1年間の推定季節指数を算出するが、移動平均による欠項を最近年の同月値などで補項し、推定季節指数の算出には固定された計算式を用いている。そのために、原数値の期間を変えるごとに季節指数が変化し、季節調整値を過去に遡って改訂しなければならないという問題が生じていた。

カナダ統計局では、1975年に時系列分析における研究の成果を取り入れたARIMAモデルをX11に採用し、X11-ARIMAを開発した。この方法は原数値をARIMAモデルによって延長した後でX11を適用するものであり、これにより、原数値そのものは移動平均による影響を受けにくくなると同時に、予測季節指数も原数値の性格を反映したものとなる。実際、同局がこの方法をカナダ及び米国の主要経済及び労働力時系列に適用してテストを重ねたところ、X11より安定性がある季節指数が得られることが明らかとなり、X11-ARIMAによる季節調整値を公表している。

ここで、原数値を x_t 、定数項を δ 、 a_t をホワイトノイズ、 φ_i を係数、 p を i の最大値とすると、

$$x_t = \sum_{i=1}^p \varphi_i x_{t-i} + \delta + a_t$$

を、 p 次のARモデルと呼ぶ。また、 θ_j を係数、 q を j の最大値とすると、

$$x_t = a_t - \sum_{j=1}^q \theta_j a_{t-j}$$

を q 次のMAモデルと呼ぶ。

作用素 B を

$$Bx_t = x_{t-1}$$

で定めると、 s 期前の値は、

$$B^s x_t = x_{t-s}$$

となり、 d 次の階差は

$$(1-B)^d x_t = \sum_{k=0}^d (-1)^k \binom{d}{k} x_{t-k}$$
$$(1-B^s)^d x_t = \sum_{k=0}^d (-1)^k \binom{d}{k} x_{t-sk}$$

と表すことができる。

上記の仮定のもとで、X11-ARIMA が想定する ARIMA モデルは、以下の式で表現される。

$$\left(1 - \sum_{i=1}^p \varphi_i B^i\right) (1-B)^d (1-B^s)^D \left(1 - \sum_{l=1}^P \Phi_l B^{ls}\right) x_t = \left(1 - \sum_{j=1}^q \theta_j B^j\right) \left(1 - \sum_{J=1}^Q \Theta_J B^{Js}\right) a_t$$

これを $(p, d, q)(P, D, Q)_s$ で表す。ここで φ_i 、 Φ_l 、 θ_j 、 Θ_J は定数、 p と P は AR モデルの次数、 q と Q は MA モデルの次数、 d と D は階差の次数を意味する。 s は階数であり、月次であれば、 $s=12$ 、四半期であれば $s=4$ となる。

清水誠[17]では、総務庁統計局で季節調整している系列に対しても X11-ARIMA を適用してテストを行い、X11 を適用した結果と比較、検討した。評価の基準としては、X11-ARIMA が季節調整値の改訂の問題を少しでも解決する手法となり得るかどうか、すなわち、その適用によって季節指数及び季節調整値の安定性が増加するかどうかに着眼した。結果は、X11 と比較して X11-ARIMA を適用するほうが推定及び最新季節指数の改訂幅が小さく、季節調整値に安定性があるということになった。本節では、この検討結果の概要を示す。

1.1 推定値と改定値の確定値からの差による分析

まず、原数値を l 年 1 月から m 年 12 月までとし、X11 と X11-ARIMA を用いて季節調整して、 $(l+7)$ 年 1 月から $(m-1)$ 年 12 月までの季節指数を確定値とする操作を行った。次に、原数値を 8 年間ずつの期間に分け、それぞれ季節調整して向こう 1 年間の季節指数を推定値として確定値との差の絶対値を比べた。同様に、それぞれ最新年の季節指数を改定値として確定値との差を比べた。これらの結果は清水誠[17]に記載しているが、総平均のみ表 1 にまとめた。なお、X11-ARIMA のモデルは l 年 1 月から m 年 12 月までに適用したときの自動選択モデルとした。

(1) 労働力調査完全失業率への適用

原数値を 1959 年から 1987 年までの労働力調査完全失業率にした場合、推定値と改定値のどちらに対しても確定値との差の絶対値の総平均は X11 よりも X11-ARIMA のほうが小さい。また、月平均で見たときに、推定値と改定値のどちらに対しても X11 のほうが小さくなるのは 5 月と 6 月だけである。5 月と 6 月の季節指数は緩やかでほとんど単調に増加しているので X11 による予測が適合しやすい。

(2) CPI 総合への適用

原数値を 1970 年から 1987 年までの CPI 総合にした場合、推定値に関する結果を総平均すると、X11 よりも X11-ARIMA のほうが誤差が小さかった。年平均に関しては 1979 年と 1984 年以外はすべて X11-ARIMA のほうが誤差が小さく、月平均に関し

でも、7月と10月以外は同じであった。推定値と改定値の両方に対して X11-ARIMA のほうで誤差が大きくなる月は7月だけだが、7月は他の月と比べて季節指数が安定している。

改定値に関しても総平均で比較すると X11-ARIMA のほうで誤差が小さかった。

(3) 家計調査勤労者世帯消費支出に対する適用

原数値を1970年から1987年までの家計調査勤労者世帯消費支出にした場合には、総平均に関しては推定値、改定値のどちらも X11 よりも X11-ARIMA のほうで誤差が小さかった。12月の季節指数は他の月と比較して著しく大きく、激しい減少傾向を示しているが、X11 で大きな誤差が生じている。特に、推定値に対しては 2.0 を超える値が3つも存在し、平均値はかなり大きくなり、X11-ARIMA の2倍近くになっている。

表1 確定値との差の絶対値の平均

		労働力調査 完全失業率	CPI	家計調査 勤労者世帯 消費支出
推定値	X11	2.2235	0.1492	0.7288
	X11-ARIMA	1.9685	0.1266	0.5818
改定値	X11	1.5702	0.1029	0.4956
	X11-ARIMA	1.3793	0.0969	0.3896

1.2 SSAP による適用結果の比較

次に、SSAP (Sliding Span Analysis Program) における Maximum % Difference を利用した。Maximum % Difference は、ある期間の*i*年*j*月の季節調整値を $A_{i,j}(1)$ 、期間を*k*-1年ずらしたときの季節調整値を $A_{i,j}(k)$ 、 $N_{i,j} = \{k: i\text{年}j\text{月の}A_{i,j}(k)\text{が存在する}\}$ としたときに、以下の式で定義される。

$$\text{Maximum \% Difference} = \frac{\max_{k \in N_{i,j}} A_{i,j}(k) - \min_{k \in N_{i,j}} A_{i,j}(k)}{\min_{k \in N_{i,j}} A_{i,j}(k)} \times 100$$

1979年から1983年までのCPI総合の指数、前月比、前年同月比の季節調整値について X11、X11-ARIMA のうち自動選択モデルを適用、(0,2,2)(0,1,1)₁₂を適用、(0,1,1)(0,1,1)₁₂を適用、(2,1,2)(0,1,0)₁₂を適用の間で Maximum % Difference の和及び二乗和を比較すると、(0,1,1)(0,1,1)₁₂を適用、(2,1,2)(0,1,0)₁₂を適用の前月比以外は X11 を適用した場合より X11-ARIMA を適用した場合のほうが小さくなった。

1.3 CPI 総合の季節指数の変化の比較

さらに、1981年から1986年までのCPI総合指数の季節指数の変化を比較した。X11を適用してもX11-ARIMAを適用しても各月の年による季節指数の変化の方向は同じだが、その大きさは後者のほうが大きかった。したがって、X11-ARIMAを適用した場合のほうがX11を適用した場合より季節性の変化が大きく現れたことになる。これに1年間のデータを加え、1981年から1987年までのCPI総合の季節指数の変化を比較すると、X11-ARIMAを適用した場合にはどの月も転換点が高々1つであるのに対し、X11を適用した場合には、わずか8年の間に転換点が3つ以上ある月がほとんどである。センサス局法は、季節性は徐々にゆっくりと確率論的に変化するという仮定に基づいているので、X11を適用した場合のような激しい季節性の変化は結果として望ましくない。

2 総務省統計局における時系列データの季節性を踏まえた検討

総務省統計局では、労働力調査、CPI、家計調査について季節調整値を毎月公表していたが、木村武[18]、経済指標部会季節調整法検討小委員会[19]、上田聖[20]、奥本佳伸[21]が、ARIMAモデルを適用すると、ほとんどの場合に安定性が高まるとした清水誠[17]の結論を追認した。

他方、清水誠[17]以降も、様々な季節調整法が開発され、季節調整に利用するデータの期間の変更に伴う結果の安定性と与えられた情報を最大限利用する最適性を中心とする観点から実証分析が蓄積されていた。しかし、季節調整法を適用するに当たっては、適用されるデータの特徴を踏まえた当該データに基づく検討が必要である。清水誠[22]では、主要系列の季節性の特徴を把握し、それを踏まえて季節調整法の適合状況を検討した。検討に際しては、計算プログラムとしての季節調整法にとどまらず、各指標の原数値の期間、季節調整値の利用状況、公表のタイミング、他の指標との関係など実務的な事情にも配慮した。清水誠[22]の結論の概要は以下のとおりである。

総務省統計局が公表している時系列データのうち、完全失業率は1990年代に季節性は小さかったが、2000年代に入り僅かではあるが季節性が大きくなりつつある。他方、CPI及び家計調査の結果については、2000年代になっても季節性は小さくなる傾向にある。

また、2000年代の季節指数の変化幅が小さいことから、これらの最近の系列に季節調整法を適用する場合、米国センサス局法のうち、例えば、X11と、ARIMAモデルのようなあらかじめ原数値にモデルを導入する方法とで結果の安定性に大差がないことが想定される。仮に、ARIMAモデルを導入するにしても、完全失業率、CPI及び実収入については、12か月階差を2回とるか、又は1か月階差と組み合わせることに

より、原系列全体がほぼ定常になることから、可能な限り階差回数の少ない単純なモデルを選択すべきであると結論付けた。これに対して、消費支出については、原数値の期間の短縮、他のオプションの導入などの工夫が必要であることがわかった。

今後、これらの系列の季節性に変化が現れる場合には、季節調整法により結果が大きく異なることもあり得るので、引き続き各公表系列の季節性を注視する必要がある。また、季節性の検討に当たっては、安定性や最適性を包括的に評価するのみならず、まず、原数値における各月の推移や相対的位置関係に注意するとともに、これらに影響を及ぼす自然、社会・経済現象を踏まえる必要がある。さらに、それらの公表系列は社会・経済に大きな影響を及ぼすことから、データの蓄積に伴い過去の季節調整値が改定されるという実務的な問題に対する対応策も併せて講じる必要がある。

3 個人企業経済調査への季節調整法の適用に向けた検討

清水誠[22]以降、労働力調査、CPI及び家計調査の季節調整法には X11-ARIMA の改訂版である X12-ARIMA が適用された。

他方、個人企業経済調査の結果は四半期ごとに公表されているが、結果を前期と比較することがある。その際、前期差又は前期比の要因として、季節的なものとそうでないものとの区別が重要になる。同調査の結果の中には、系列により、又は産業によりほとんど季節性が見られないものも存在しているが、多くの場合、各期特有の季節性が含まれている。同調査は 2002 年度に大規模な見直しが行われたことから、調査結果の時系列的な比較が可能なのは 2002 年 4-6 月期からであるが、2010 年において 8 年近くが経過してデータの蓄積が進んだために、季節調整法の比較・検証ができるようになった。

そこで、個人企業経済調査については初めての試みとして、清水誠[23]で、同調査の代表的な系列の結果について標準的な季節調整法を適用し、調査結果の季節性と傾向を抽出するとともに、同調査の結果への適用という観点から各種季節調整法の評価を行い、今後同調査結果の参考系列を検討する際の基礎資料を提供することを目指した。

個人企業経済調査のような公的統計調査の場合には、公表値への影響という意味で安定性の喪失が最も致命的である。清水誠[23]は、多くの系列、産業である程度の安定性を保ち、しかも適切性や AIC で劣った結果を示すことが少ない X12-ARIMA の (0,1,1)(0,1,1) が季節調整法の候補となると結論付けた。しかし、2008 年 1-3 月期における卸売業、小売業の売上高の場合のように、データの動きによっては他の季節調整法と比べて安定性が低くなる場合があるので注意が必要であるとした。しかし、清水誠[23]の時点では、データとして利用可能な期間がわずか 7 年、四半期単位では 28

しかデータが存在しない状況であり、季節調整法を検討するには十分とは言えなかった。

それから6年が経過し、更にデータが蓄積したことを踏まえ、本節では、同様の手法により季節調整法の再評価をする。

3.1 個人企業経済調査の結果の季節性

個人企業経済調査の結果の季節性を概観するために、本節ではセンサス局法 X12-ARIMA のうち、清水誠[23]で良好な結果が得られた(0,1,1)(0,1,1)モデルを用いて産業大分類別に業況、売上高、営業利益を季節調整し、季節指数と季節調整値を時系列で観察する。原数値の期間は2002年4-6月期から2016年1-3月期までとする。また、異常値は季節調整値に含めることとする。なお、この間に業況は一貫してマイナスの値なので、符号を逆転させプラスの値に変換してから季節調整を行った。したがって、業況については季節指数が低いほど業況の悪さが緩和されたことになる。

3.1.1 業況

詳細な分析結果を提示する前に、まず分析に当たった期間の業況について産業別にもその特徴を把握することとする。

(1) 製造業

製造業については、図1-1-1のとおり、季節指数は10-12月期に低く（良く）、その他の期に高く（悪く）なるという季節性があったが、2011年度頃から4-6月期も低く（良く）なっている。

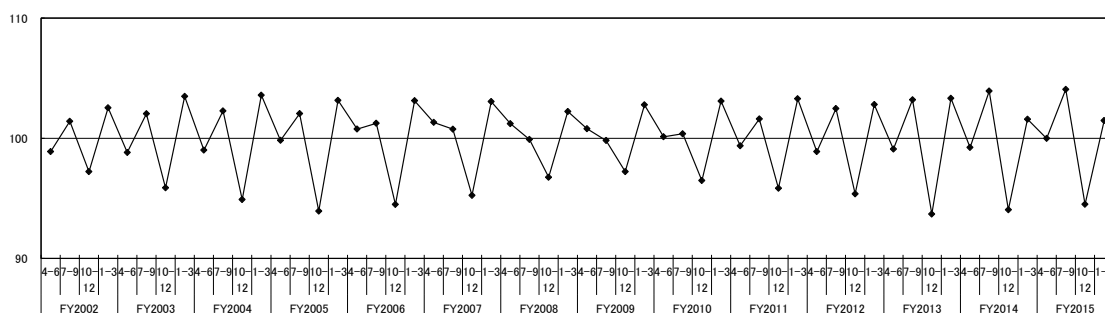


図1-1-1 製造業業況の季節指数

季節調整値を見ると、2005年10-12月期をピークに低下傾向で推移し、金融危機直後の2009年7-9月期をボトムとして上昇傾向で推移した。その後2013年1-3月期をピークに、消費税率上昇後に下降したものの、2015年1-3月期以降は上昇傾向で推移している。

(2) 卸売業、小売業

卸売業、小売業については、図1-1-2のとおり、2010年度頃までは1-3月期に高い（悪い）という特徴があったが、製造業ほど季節性は顕著でない。

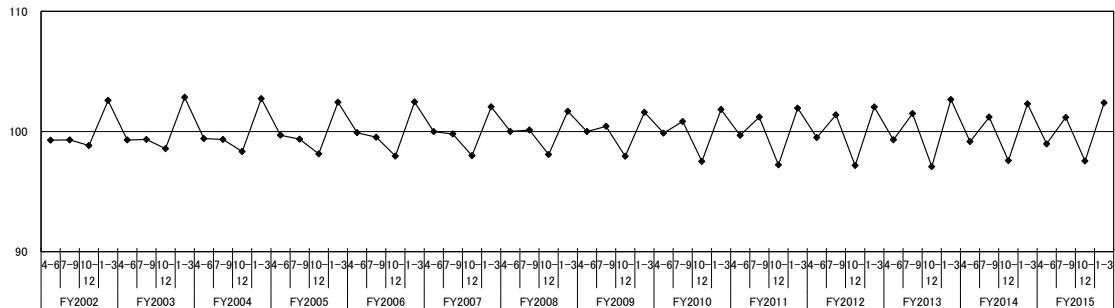


図1-1-2 卸売業、小売業業況の季節指数

季節調整値を見ると、製造業と比べてピークとボトムの時期や程度に違いがあるものの、傾向としては似た動きをしている。

(3) 宿泊業、飲食サービス業

宿泊業、飲食サービス業については、図1-1-3のとおり、2010年度頃までほとんど季節性が見られなかったが、その後1-3月期に高く（悪く）なっている。

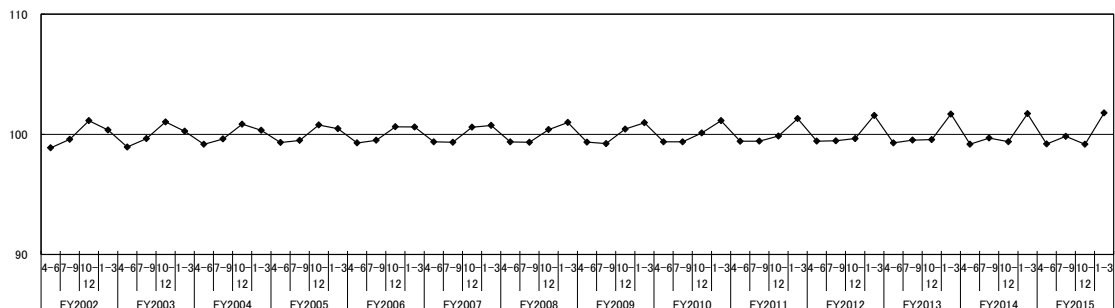


図1-1-3 宿泊業、飲食サービス業業況の季節指数

季節調整値を見ると、2009年10-12月期後は上昇傾向で推移している。

(4) サービス業

サービス業の季節性は著しく、図1-1-4のとおり、1-3月期に高く（悪く）、4-6月期に低い（良い）という特徴があるものの、年々その傾向は弱まっている。

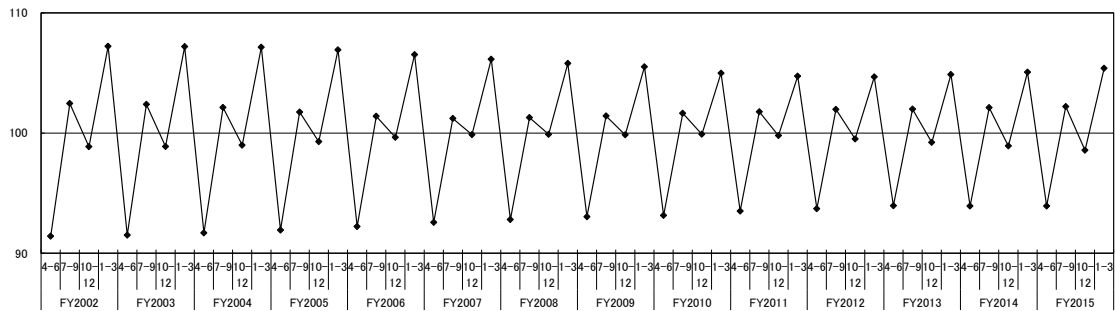


図 1-1-4 サービス業業況の季節指数

季節調整値の動向は宿泊業、飲食サービス業と類似している。

3.1.2 売上高

(1) 製造業

次に、四半期ごとの企業当たり売上高について産業別に比較をすることとする。売上高については、業況では季節性が弱い産業でも季節性が明確である。

製造業については、図 1-2-1 のとおり、10-12 月期に高く 1-3 月期に低いという特徴がある。

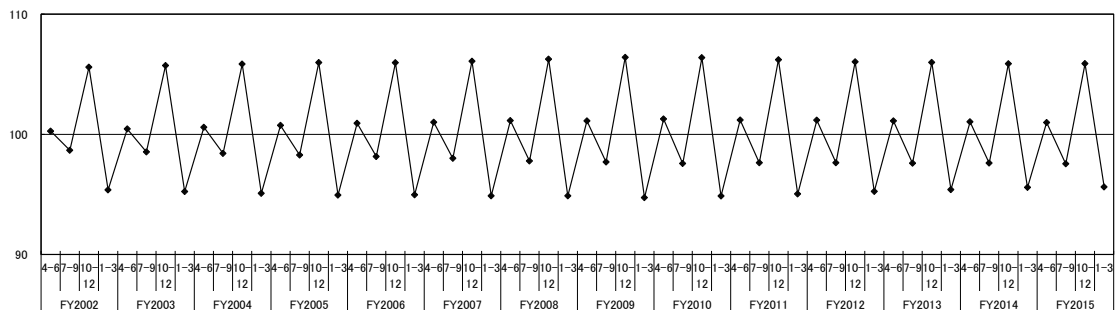


図 1-2-1 製造業売上高の季節指数

若干の凹凸が見られる原数値と比較して季節調整値の推移は滑らかである。

(2) 卸売業、小売業

卸売業、小売業についても、図 1-2-2 のとおり、製造業と同様に 10-12 月期に高く 1-3 月期に低いという特徴がある。

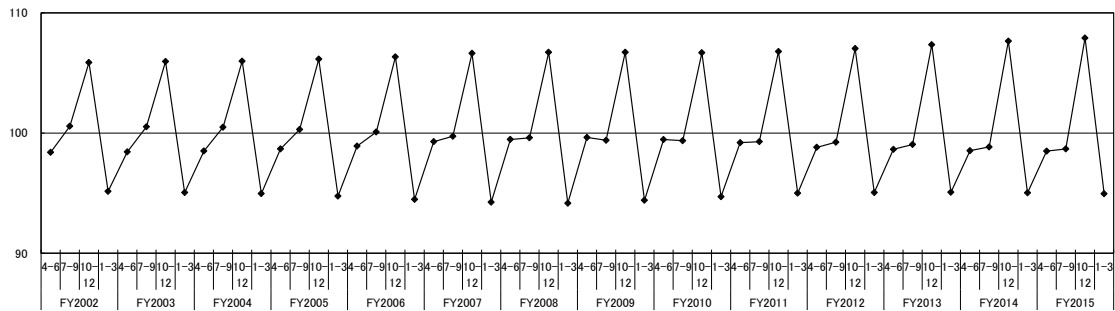


図 1-2-2 卸売業、小売業の売上高の季節指数

季節調整値の推移は原数値よりも滑らかであり、2007年7-9月期をピークに下降傾向である。

(3) 宿泊業、飲食サービス業

宿泊業、飲食サービス業の季節性は他の産業ほど大きくないが、図 1-2-3 のとおり、他の産業と同じく 10-12 月期に高く 1-3 月期に低いという特徴が見られる。しかし、2006 年度以降 7-9 月期でも高くなっている。

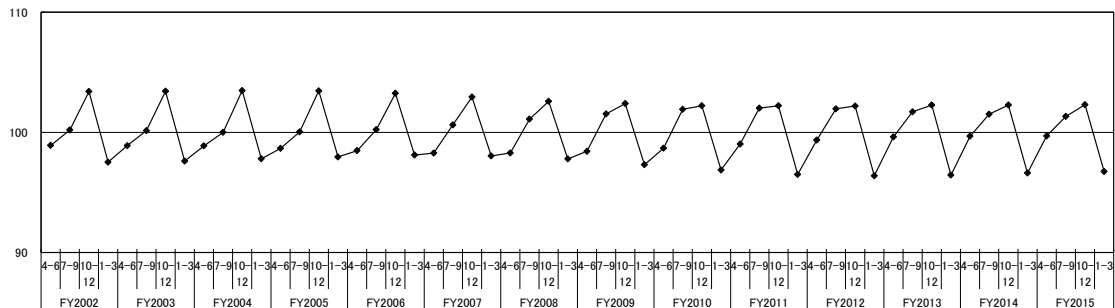


図 1-2-3 宿泊業、飲食サービス業の売上高の季節指数

季節調整値を見ると、最近では、2012 年 1-3 月期をボトムとして少しずつ回復する傾向にある。

(4) サービス業

サービス業についても、図 1-2-4 のとおり、10-12 月期に高く 1-3 月期に低いという特徴があったが、2007 年度以降 4-6 月期でも高くなっている。

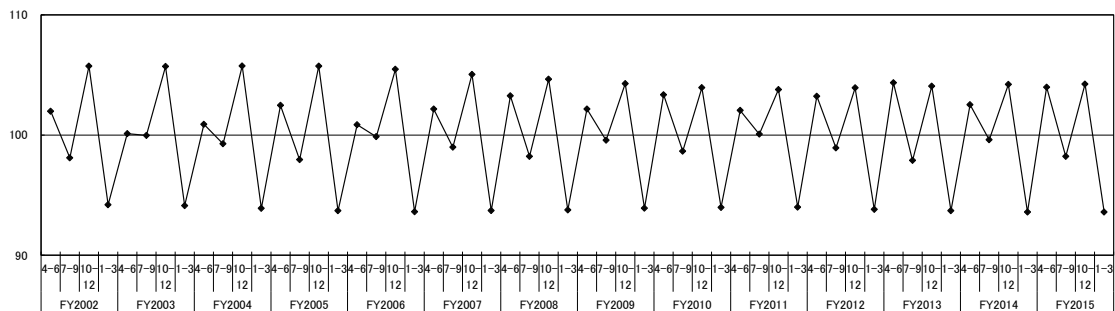


図 1-2-4 サービス業の売上高の季節指数

季節調整値を見ると、少しずつではあるが、一貫して下降傾向にある。

このように、個人企業の売上高についてはどの産業で見ても 10-12 月期に最も高く、1-3 月期に最も低いという特徴が見られる。個人企業と対比させるために法人企業について売上げ状況を見ると、年始期に低くはならない。実際、法人企業統計調査の金融保険業を除く全産業の売上高について公表されている原数値を季節調整値で割って季節指数を計算すると、1-3 月期が最も高い。個人企業経済調査の場合、調査対象のうち製造業、卸売業、小売業、宿泊業、飲食サービス業については食料品製造業、飲食料品小売業、飲食店を始め何らかの関係で食料に携わる企業がかなりの割合を占めている。食料についての需要側の消費動向に着目して、家計調査（総世帯）の 2002～2015 年度における消費支出のうち食料について、X12-ARIMA の (0,1,1)(0,1,1)モデルを適用して季節指数を計算すると、2015 年度においては 10-12 月期に 107.5 と最も高く、1-3 月期に 94.6 と最も低くなった。他方、個人企業経済調査のサービス業の調査対象の中には理容業及び美容業が数多く存在する。家計調査（総世帯）の 2002～2015 年度における消費支出のうち理美容サービスの季節性は強くないが、同様に季節指数を計算すると、2015 年度においては 10-12 月期に 103.9 と高く、他の期は 97.2～100.4 とわずかに低くなっている。

3.1.3 営業利益

さらに、企業当たり 3 か月間の営業利益について産業別に比較をすることとする。営業利益についてはどの産業も売上高よりも季節性の変化が顕著である。

(1) 製造業

製造業については、図 1-3-1 のとおり、売上高と同様に 10-12 月期に高く 1-3 月期に低いという特徴がある。

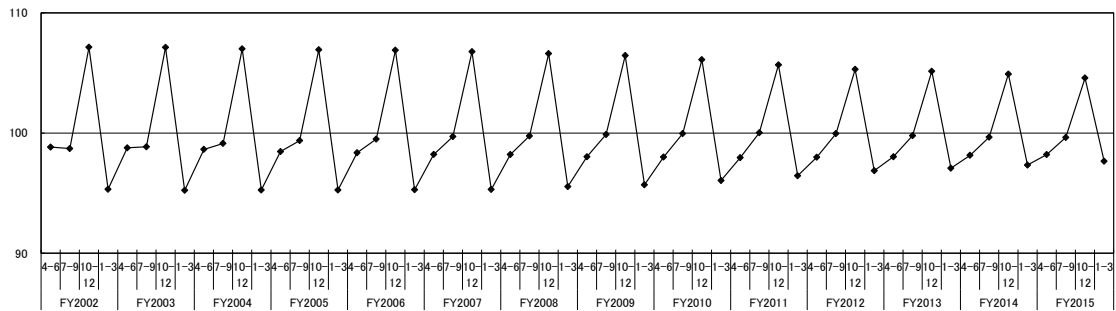


図 1-3-1 製造業営業利益の季節指数

季節調整値を見ると、2004年7-9月期をピークとして下降し、金融危機後の2009年10-12月をボトムとして少し回復したものの、2010年以降はほぼ一定の水準で推移している。

(2) 卸売業、小売業

卸売業、小売業についても、図1-3-2のとおり、10-12月期に高く1-3月期に低いという特徴がある。10-12月期の季節指数は、2007~2012年度を除き110を超えている。

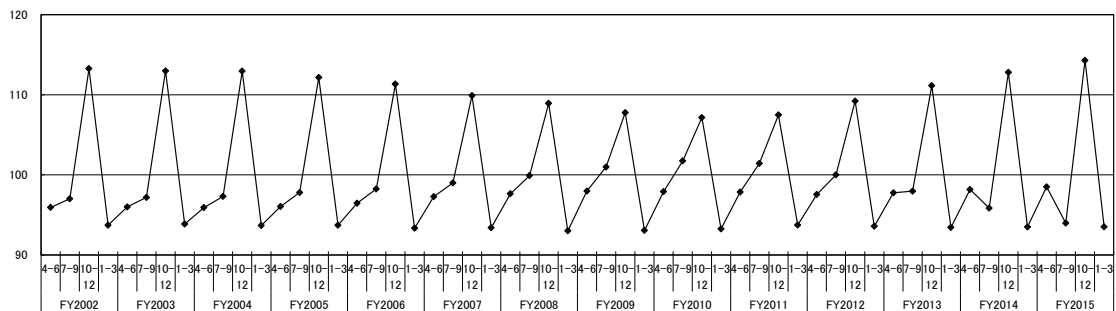


図 1-3-2 卸売業、小売業営業利益の季節指数

(3) 宿泊業、飲食サービス業

宿泊業、飲食サービス業の季節性は2006年度頃まではほとんど見られなかったが、少しずつ大きくなり、2007年度以降10-12月期に高く、1-3月期に低いという特徴が明確である。

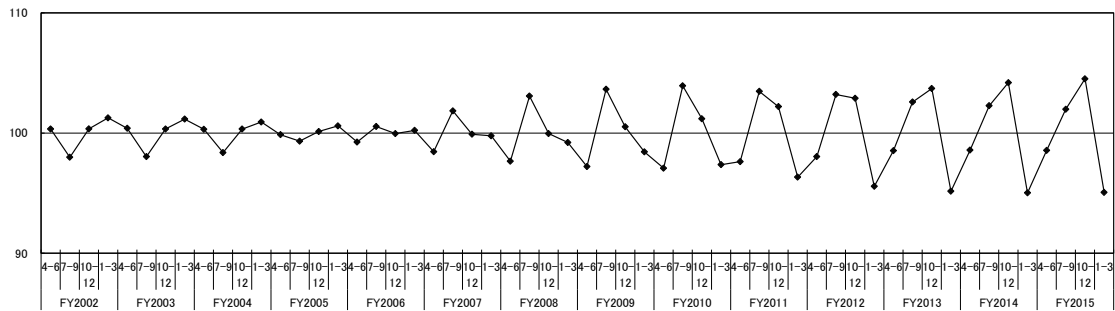


図 1-3-3 宿泊業、飲食サービス業営業利益の季節指数

(4) サービス業

サービス業は、図 1-3-4 のとおり、1-3 月期に低く、2007 年度以降 90 を下回っている。また、10-12 月期よりも 4-6 月期の方が高くなっている。

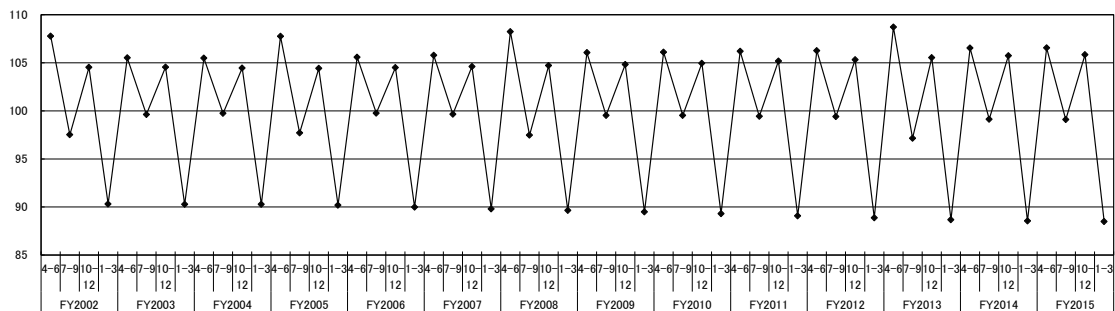


図 1-3-4 サービス業営業利益の季節指数

季節調整値を見ると、2004 年度 4-6 月期以降は売上高と同様に少しずつ下降する傾向にある。

営業利益は売上高から原材料費、賃金等を除いたものなので、季節性については売上高と類似の傾向が現れやすい。実際、営業利益については、売上高と同じように、10-12 月期に高く 1-3 月期に低いという特徴がある。

3.1.4 季節調整により把握される総括的特徴

最後に、季節性や季節調整値の特徴をより長期的、包括的に概観することとする。

どの産業においても 10-12 月期に売上高と営業利益が高く、したがって、業況の悪さも比較的緩和されるのに対し、1-3 月期にはその逆の現象が現れている。サービス業については 4-6 月にも売上高と営業利益が高く、業況の悪さが緩和される。

季節調整値を見ると、長期的な傾向を把握することができる。業況については、常にマイナスの値であるものの、産業にかかわらず 2002~2003 年度に上昇傾向で推移

し、2004～2006 年度をピークとして 2007～2008 年度に下降傾向で推移し、金融危機後の 2009 年度にボトムとなり、2010 年度以降は上昇傾向にあった。産業によっては、消費税率が上昇した 2014 年度に少し下降し、再び回復している。この傾向は景気動向指数の CI (Composite Index) の先行指数の動きと整合している。売上高については産業によって動きが異なり、製造業は 2004 年 7-9 月期をピークに上下動を繰り返してきたが、その振幅が小さくなり、230 万円程度に収束する動きを示している。卸売業、小売業は 2007 年 7-9 月期をピークに、サービス業は一貫して下降傾向で推移し、宿泊業、飲食サービス業については 2012 年 1-3 月期をボトムとして回復傾向にある。営業利益については、製造業は 2004 年 7-9 月期をピークに下降傾向で推移した後、2010 年以降はほぼ横ばいで推移している。他の産業については少しずつではあるが、低下傾向で推移している。

このように、季節調整を行うことにより、各系列、各産業の季節性と長期的傾向がより明確になった。

3.2 個人企業経済調査の結果による季節調整法の比較

上記で概観した各系列・各産業について、標準的な季節調整法の中から幾つかを選び、定量的指標により比較した。この指標としては、従来の実証分析で数多く用いられている AIC のほか、データを追加して同一手法を適用しても元の結果との変化が小さい安定性、同一データに同一手法を 2 度適用しても 1 度目の結果が再現される適切性とした。

3.2.1 比較の指標

個人企業経済調査について、3.1 で検討した 2002 年 4-6 月期から 2015 年 1-3 月期までの結果について以下の季節調整法を適用し、結果を比較した。なお、同調査は月次単位でなく四半期単位の調査であることから、曜日、祝日、日数等に関するきめ細かいオプションについては配慮していない。

(1) 季節調整法

最初に、適用・比較する 2 つの季節調整法についてモデルに関する部分を中心に簡単に解説する。

① X12-ARIMA

個人企業経済調査の原数値を x_t 、 a_t を平均 0、分散一定、互いに無相関、 x_t とも無相関とすると、ARIMA モデルは、個人企業経済調査が四半期単位であることを踏まえると、 $s=4$ となることから

$$(1 - \sum_{i=1}^p \varphi_i B^i)(1 - B)^d (1 - B^4)^D (1 - \sum_{l=1}^P \Phi_l B^{4l}) x_t = (1 - \sum_{j=1}^q \theta_j B^j)(1 - \sum_{J=1}^Q \Theta_J B^{4J}) a_t$$

と書くことができ、これを略して $(p, d, q)(P, D, Q)$ と書くこととする。

例えば、Airline モデルと呼ばれる $(0, 1, 1)(0, 1, 1)$ モデルは

$$(1 - B)(1 - B^4)Z_t = (1 - \theta B)(1 - \Theta B^4)\varepsilon_t$$

すなわち

$$Z_t - Z_{t-1} - Z_{t-4} + Z_{t-5} = \varepsilon_t - \theta\varepsilon_{t-1} - \Theta\varepsilon_{t-4} + \theta\Theta\varepsilon_{t-5}$$

を意味することとなる。

このモデルは、原数値に適合する状態空間モデルと近似的に同じという特徴もある。

なお、自動選択モデルとして、残差の無相関性をテストする Ljung-Box 統計量やその信頼係数、残差の平均値に対する t 値、残差の標準偏差の推定量、BIS (Basian Information Criterion) 等の指標などを用いてプログラムにより選択されるモデルが存在する。しかし、清水誠[23]によれば、この方法は、検討した個人企業経済調査の時系列データに関しては良好な結果を示さなかったため、本論の検討対象から除外した。また、X12-ARIMA が X11-ARIMA と異なる点は、regARIMA と呼ばれるレベルシフト、異常値、曜日変動等を表す変数を指定して事前調整を行う部分にあるが、ここではそれらを指定してない。したがって、今回選択した方法は X11-ARIMA と同じになる。

なお、適用したプログラムは米国センサス局からダウンロードした Version 0.3 である。

②DECOMP

情報・システム研究機構統計数理研究所の北川源四郎[24]によって開発された DECOMP については、個人企業経済調査の原数値 Y_t が傾向変動 T_t 、季節変動 S_t 、定常変動 A_t 、曜日変動 D_t 、偶然変動 E_t に区別され、 ε_{it} を平均 0、分散一定の正規分布、 d_{wt} を t 期に含まれる w 曜日の日数とすると、

$$Y_t = T_t + S_t + A_t + D_t + E_t$$

$$(1 - B)^d Y_t = \varepsilon_{1t}$$

$$(1 - B^4)^D S_t = \varepsilon_{2t}$$

$$\left(1 - \sum_{i=1}^p \varphi_i B^i\right) A_t = \varepsilon_{3t}$$

$$D_t = \sum_{w=1}^7 \gamma_w d_{wt}$$

$$E_t = \varepsilon_{4t}$$

と表される。

個人企業経済調査は原数値の期間が短いことから、DECOMP のモデルの次数はできるだけ小さなものから選ぶこととし、 $d = D = 1$ に固定した上で、 $p = 0$ 及び $p = 1$ の2通りについて比較した。プログラムは情報・システム研究機構統計数理研究所のホームページにある Web DECOMP を利用した。

(2) 比較の視点

季節調整法の比較の視点は、以下のとおり、AIC、安定性、適切性とした。

①AIC

2002年4-6月期から2016年1-3月期までについて、季節調整法ごとに各プログラムから算出されるAICを比較し、AICが小さいほど優れた季節調整法であるとした。ただし、X12-ARIMAで算出されるAICはモデルの当てはまりの良さを評価したものであり、移動平均部分の評価を含んでいない。したがって、AICを評価するに当たってはX12-ARIMAとDECOMPを区別することにする。

②安定性

2002年4-6月期から2015年1-3月期までの13年間について、季節調整法ごとに2015年1-3月期まで適用した場合の季節調整値に対する2016年1-3月期まで適用した場合の季節調整値の平均絶対変化率（変化率の絶対値の平均値）を計算した。平均絶対変化率が小さいほど安定性が強いとした。

公的統計の季節調整においては、季節調整値を公表した後に原数値を追加して季節調整値を改めて公表することが普通であることから、安定性は利用者が直面する重要な問題である。

③適切性

同一季節調整法を繰り返したときに、繰り返す前の数値と近いほど優れた季節調整法であるとの視点を踏まえ、2002年4-6月期から2006年1-3月期まで適用した場合の季節調整値に対してその値に繰り返し同一の季節調整法を適用した場合の季節調整値の平均絶対変化率を計算した。平均絶対変化率が小さいほど適切性が強いとした。

3.2.2 比較の結果

各季節調整法について上記の3つの指標を計算すると以下ようになる。

(1) AIC

X12-ARIMA について、 $d = D = 0$ という条件の下で、 p 、 q 、 P 、 Q について 0 又は 1 のすべての異なる 16 通りの組合せで AIC を計算した結果を表 2 に示す。これによれば、 $p = P = Q = 1$ のとき、すなわち、 $(1,0,0)(1,0,1)$ 又は $(1,0,1)(1,0,1)$ で AIC が最小になる系列・産業が多く、逆に p 又は P のいずれかのみが 1、すなわち、 $(1,0,0)(0,0,0)$ 又は $(0,0,0)(1,0,0)$ で AIC が最大になる系列・産業が多かった。系列・産業にもよるが、一般的には、 $(1,0,0)(0,0,0)$ 又は $(0,0,0)(1,0,0)$ というモデルを回避するほうが無難であると考えられる。

表 2 X12-ARIMA 各種モデルの AIC $d=D=0$

系列	産業	P=0								P=1							
		Q=0				Q=1				Q=0				Q=1			
		p=0		p=1		p=0		p=1		p=0		p=1		p=0		p=1	
		q=0	q=1	q=0	q=1	q=0	q=1	q=0	q=1	q=0	q=1	q=0	q=1	q=0	q=1	q=0	q=1
業況	製造業	659	636	616	613	655	635	615	613	686	652	615	613	605	644	608	609
	卸売業、小売業	600	579	563	561	589	570	537	548	625	578	547	547	626	601	537	561
	宿泊業、飲食サービス業	591	562	551	549	567	558	552	551	623	598	552	551	599	599	554	553
	サービス業	615	613	631	606	599	596	613	594	621	604	581	584	623	593	569	566
売上高	製造業	781	779	794	788	783	779	794	790	843	804	792	790	802	788	767	769
	卸売業、小売業	795	797	833	815	775	790	818	805	802	787	797	795	778	780	770	770
	宿泊業、飲食サービス業	692	683	700	702	693	662	701	703	759	724	701	702	755	722	694	698
	サービス業	655	656	708	680	656	669	688	675	696	666	660	645	693	652	629	631
営業利益	製造業	638	635	637	646	605	619	639	648	665	683	639	647	661	652	608	639
	卸売業、小売業	600	674	636	614	596	654	630	608	624	626	637	604	618	619	619	594
	宿泊業、飲食サービス業	559	569	596	592	562	559	598	594	636	607	598	594	631	616	599	595
	サービス業	596	596	644	609	576	577	618	598	597	564	563	581	592	573	549	549
		注	系列・産業ごとに最小													
			系列・産業ごとに最大													

清水誠[23]では、 $d = D = 0$ という条件の下では X11 の AIC が最も大きかったが、表 2 によれば、そのような結果は再現しなかった。

また、 $d = D = 1$ という条件の下で、 p 、 q 、 P 、 Q について 0 又は 1 のすべての異なる 16 通りの組合せで AIC を計算した結果を表 3 に示す。これによれば、 $Q = 1$ のときに AIC が最小になる系列・産業が多く、逆に $P = Q = 0$ のとき、すなわち、 $(0,1,0)(0,1,0)$ 、 $(0,1,1)(0,1,0)$ 、 $(1,1,0)(0,1,0)$ 、 $(1,1,1)(0,1,0)$ で AIC が最大になる系列・産業が多かった。

高についてどの産業でも X11 は最大であった。X12-ARIMA の(1,1,0)(0,1,1)、(0,1,0)(1,1,1)、(1,1,1)(1,1,1)については、平均絶対変化率が最も小さくなる場合と最も大きくなる場合があり、系列・産業によっては適用可能なモデルである。しかし、2002年4-6月期から2015年1-3月期までの安定性の観点から汎用的なモデルを選定するのであれば、X12-ARIMA の(0,1,1)(0,1,1)と DECOMP が適当である。

表5 2002年4-6月期から2015年1-3月期までの改定における平均絶対変化率

系列	産業	X12-ARIMA						DECOMP		
		X-11	(010)(011)	(011)(011)	(110)(011)	(010)(111)	(111)(111)	p=0	p=1	
業況	製造業	0.008	0.002	0.002	0.002	0.002	0.003	0.002	0.002	
	卸売業、小売業	0.002	0.001	0.001	0.001	0.001	0.005	0.001	0.001	
	宿泊業、飲食サービス業	0.004	0.008	0.003	0.003	0.006	0.002	0.002	0.002	
	サービス業	0.001	0.002	0.002	0.002	0.002	0.010	0.001	0.001	
売上高	製造業	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	
	卸売業、小売業	0.003	0.001	0.001	0.001	0.001	0.002	0.001	0.001	
	宿泊業、飲食サービス業	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	
	サービス業	0.001	0.001	0.001	0.001	0.000	0.000	0.001	0.001	
営業利益	製造業	0.003	0.010	0.003	0.003	0.009	0.003	0.004	0.005	
	卸売業、小売業	0.007	0.007	0.007	0.007	0.011	0.008	0.004	0.004	
	宿泊業、飲食サービス業	0.003	0.001	0.002	0.002	0.001	0.001	0.005	0.005	
	サービス業	0.002	0.005	0.006	0.006	0.001	0.001	0.001	0.001	
		注 系列・産業ごとに最小							
		 系列・産業ごとに最大							

表5の結果は2008年1-3月期までで検証した清水誠[23]の結果と異なるものであった。清水誠[23]では、X11が優れ、X12-ARIMAの(0,1,1)(0,1,1)とDECOMPが劣るという結果であった。これに対して、本研究の場合、利用できるデータ系列の期間が長期化したことにより、少し複雑なモデルを当てはめた効果が発揮されやすくなった。

営業利益については、どの産業も特異な動きをしており、季節調整をしても時期により季節性が大きく異なる結果となる。このような場合、X12-ARIMAの(0,1,1)(0,1,1)とDECOMPでも平均絶対変化率が大きくなるがあった。

(3) 適切性

適切性については、安定性と同じモデルを対象に比較した結果を表6に示す。

すべての系列、産業において平均絶対変化率はDECOMPで0.000と最も小さくなった。これは、DECOMPの計算方法から明らかなことである。逆に、多くの系列、産業において平均絶対変化率はX11で最も大きくなった。したがって、適切性の観点からはDECOMPが最も優れ、X11が最も劣っていることになる。

表 6 2002年4-6月期から2015年1-3月期までの再適用に伴う平均絶対変化率

系列	産業	X12-ARIMA						DECOMP	
		X-11	(010)(011)	(011)(011)	(110)(011)	(010)(111)	(111)(111)	p=0	p=1
業況	製造業	0.002	0.002	0.002	0.002	0.002	0.002	0.000	0.000
	卸売業、小売業	0.001	0.001	0.001	0.001	0.002	0.001	0.000	0.000
	宿泊業、飲食サービス業	0.002	0.002	0.001	0.002	0.001	0.001	0.000	0.000
	サービス業	0.002	0.002	0.001	0.001	0.001	0.001	0.000	0.000
売上高	製造業	0.003	0.001	0.001	0.001	0.002	0.002	0.000	0.000
	卸売業、小売業	0.002	0.001	0.001	0.001	0.001	0.001	0.000	0.000
	宿泊業、飲食サービス業	0.003	0.001	0.001	0.001	0.001	0.004	0.000	0.000
	サービス業	0.005	0.001	0.001	0.001	0.001	0.001	0.000	0.000
営業利益	製造業	0.002	0.004	0.002	0.002	0.004	0.002	0.000	0.000
	卸売業、小売業	0.005	0.003	0.004	0.004	0.003	0.004	0.000	0.000
	宿泊業、飲食サービス業	0.003	0.004	0.004	0.004	0.003	0.004	0.000	0.000
	サービス業	0.002	0.002	0.002	0.002	0.002	0.005	0.000	0.000
	注	
		

表 6 の結果については、清水誠[23]では X11 の優位性が比較的高かった点が異なる。

(4) 手法間の相違の程度

上記のように、適用する季節調整法、データの期間及びモデルの選択方法により季節調整値に相違が生じたが、その程度を見るために、2015 年度における業況、売上高、営業利益について産業ごとに原数値と各季節調整法による季節調整値を比較する。

全体として、X11 を除く他の季節調整値については、手法間で少し相違があるものの、結果の上下動については類似度が高い。このことは、どのような季節調整法、データ期間、モデルを選択しても、傾向を見い出すという点においては十分な役割を果たすことができることを示している。しかし、図 2-1、図 2-2、図 2-3 のとおり、X11 については、卸売業、小売業の業況、営業利益、サービス業の売上高などで原数値に引き寄せられて季節性が除外されていないと考えられる場合があった。

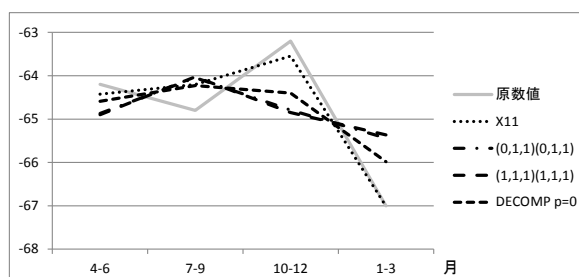


図 2-1 卸売業、小売業業況の原数値と季節調整値（2015 年度）

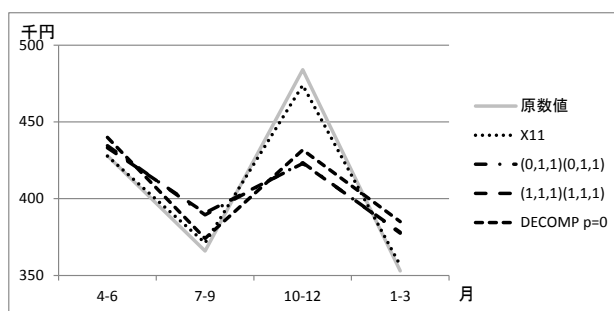


図 2-2 卸売業，小売業営業利益の原数値と季節調整値（2015 年度）

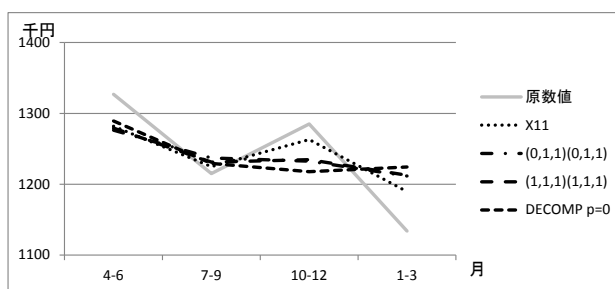


図 2-3 サービス業売上高の原数値と季節調整値（2015 年度）

(5) 総括

季節調整法については、評価の視点によって、また、系列、産業などのデータの特徴によって、優れた手法は異なる。個人企業経済調査のような公的統計調査の場合には、公表値への影響という意味で安定性の喪失が最も致命的であることから、汎用的なモデルを選定するのであれば、多くの系列、産業である程度の安定性を保ち、しかも適切性や AIC で劣った結果を示すことが少ない X12-ARIMA の $(0,1,1)(0,1,1)$ が季節調整法の候補となると考える。また、適切性だけでなく安定性や AIC においても優れた結果が得られる DECOMP の $p=0$ も候補として考えられる。

なお、季節調整法による季節調整値の相違は 3.1 で概観した季節調整値の傾向を左右するほどの大きなものではない。つまり、どの方法を用いても結果の解釈に大きな変更をもたらすほどのものではない。したがって、個人企業経済調査については、まず、季節調整値を計算することが重要であり、その方法のうちどれを選ぶかについては、他の調査結果に適用されているような標準的なものであれば大きな問題があるとは言えない。ただし、X11 については原数値に引き寄せられる場合があるので注意が必要である。

3.2.3 個人企業経済調査の季節調整値の導入に向けて

個人企業経済調査の結果には季節性が見られることから、前期と数字を比較して傾向を捉えようとするならば季節調整を行うことが必要である。個人企業経済調査の季節調整については、過去の実績が積まれていないという点で、本節に示した一連の比較研究は新しい試みである。

しかし、個人企業経済調査の結果は 2002 年以降蓄積されたとは言え、まだ系列ごとのデータ数は 14 年×4 四半期=56 に過ぎない。このため、今回は時系列モデルで選択可能なオプションを基本的なものに限定したが、最適なモデルを選択するためには、より多数のモデルを適用した結果について比較をすることが望ましい。比較に当たっては、四半期ごとに公表を実施、業況については速報値と確報値が存在、業況については業種合算値が存在、などの実務的側面にも配慮することが必要である。

3.3 他の調査結果との比較

経常的に作成される公的統計については、季節調整値を公表しているものが多い。季節調整値については、完全失業率のように主な結果として公表されるもの、CPI や家計調査の結果のように原数値（前年同月比）が主系列で、季節調整値は参考系列に近い扱いとなっているものなど統計によって様々である。

清水誠[23]では、個人企業経済調査の季節調整法を検討する際の基礎的な情報として、総務省統計局が実施する経常調査及び他府省等が実施する企業に対する景況・経理調査を中心に他の公的統計の季節調整について簡単に状況を述べている。本節の最後に、最近の状況を踏まえてその研究内容を改定した結果を示す。

3.3.1 総務省統計局の月次調査との比較

(1) 季節調整の状況

総務省統計局が実施する経常調査は月次のものが多く、それらの四半期単位の結果も公表されている。労働力調査及び CPI については、月次では完全失業率、CPI 総合など一部の系列について X12-ARIMA で季節調整された結果が公表されているが、四半期では季節調整された結果は公表されていない。

家計調査については、2 人以上の世帯に係る消費支出や実収入など主要系列の名目値及び実質値について、月次系列では X12-ARIMA で、四半期系列ではそれを 3 か月平均した結果が公表されている。家計調査において採用されている regARIMA モデルは、系列によって異なり、 $d = D = 1$ とした上で、 p 、 P 、 q 、 Q それぞれについて 0、1、2 及び 3 の中から、回帰係数の推定値の符号条件の妥当性を検証した上で、AIC が最小になるものが選択されている。2016 年現在総務省統計局が採用している

regARIMA モデルは、2000 年から 2015 年までの結果を原数値として選択されたものであるが、regARIMA モデルは、実質値を計算するための CPI の基準が改定される際に見直される予定である。なお、家計調査については、月次調査結果であっても 2 人以上の世帯に係る消費水準指数、総世帯及び単身世帯に係る主要系列（消費支出、住居等を除く消費支出）以外の結果は X11 で季節調整されている。また、家計消費指数についても 2002 年から 2015 年までのデータを原数値として季節調整値が試算されている。

（2）季節性の程度

家計調査の四半期単位の消費水準指数について、公表されている原数値を季節調整値で割って季節指数を計算すると、2015 年度においては 96.4～102.5 であり、個人企業経済調査の結果と比較して季節性が強いとは言えない。

一般に、月次の季節指数の範囲は四半期よりも大きくなるが、総務省統計局が実施する月次調査の結果を見ると、四半期単位の個人企業経済調査の結果よりも小さくなる場合が多く、例えば、2016 年における季節指数は、完全失業率については 92.8～106.0 と、個人企業経済調査のサービス業の営業利益より範囲が小さい。このように、個人企業経済調査の結果については、総務省統計局が従来から季節調整値を公表している指標と同等以上の季節性が含まれている。

3.3.2 他府省等の景況・経理調査との比較

他府省等には、主として法人企業を対象に景況、売上高、経常利益等を把握する調査が存在している。景況については、そもそも季節要因を除いた実勢についての回答が集計され、季節調整が行われないうちが多いが、売上高及び経常利益については、概して原数値と併せて季節調整した結果も公表されている。

（1）景況

法人企業の景況調査としては、内閣府及び財務省が四半期ごとに実施している法人企業景気予測調査が存在するが、調査票において判断は季節要因を除いた実勢であることとしており、調査結果の中に季節調整値は存在しない。しかし、財務省地方財務局又は地方公共団体が公表する同種の地域別の結果の中には季節調整されているものも存在する。

日本銀行が 3 か月に 1 回調査する短観（全国企業短期経済観測調査）でも業況を調査しているが、こちらも調査票（「調査表」と記載された書類）への回答は季節変動要因を除いた実勢ベースで行うこととされている。他方、中小企業庁及び中小企業基盤整備機構が産業別、地域別に四半期ごとに実施している中小企業景況調査の結果は X11 で全面的に季節調整されている。日本政策金融公庫総合研究所が毎月実施してい

る中小企業景況調査や全国小企業月次動向調査の結果も X11 で全面的に季節調整されている。

(2) 経理状況

法人を対象に経理状況を四半期ごとに把握する調査としては、財務省が実施する法人企業統計調査が存在し、売上高、経常利益及び設備投資について、製造業、非製造業（金融業及び保険業を除く。）ごとに X12-ARIMA を用いて季節調整した結果が原数値と併せて公表されている。regARIMA モデルは、 $d = D = 1$ とした上で、 p 、 P 、 q 、 Q それぞれについて0、1及び2の中から、AICが最小になるものが選択されており、毎年4-6月期に見直しが行われている。

(3) 各種景況・経理調査と比較した個人企業経済調査に関する留意点

景況についての調査結果を見る際には、調査結果に季節調整済という表示がなくても、季節調整された回答の集計結果である可能性があるので注意が必要である。しかし、調査方法として回答企業に季節調整済の報告を求めることには、問題がある。統計調査において季節要因を除いた実勢について回答が可能な企業の対象は限られるからである。特に、個人企業については、将来的にも調査回答者に一律にそのような回答を求めることは困難であることから、実勢を把握するためには調査結果について季節調整を行う方法が妥当である。

景況や営業（経常）利益については、プラス・マイナス両方の値をとることがあるが、このような場合、乗法型の季節調整法を適用することができず、法人企業統計調査では加法型を適用している。個人企業経済調査の場合、2002年度以降、業況については一貫してマイナス、売上高及び営業利益については一貫してプラスなので、このような問題は発生していないが、製品・商品・原材料の在庫状況の業況判断指数 DI (Diffusion Index) のようにプラス・マイナスが混在する系列については、加法型しか適用できないことに注意が必要である。

4 X13-ARIMA-SEATSの適用に向けた検討

日本では、公的統計については、2011年に内閣府統計委員会で設定された「季節調整法の適用に当たっての統計基準」の中で「季節調整法を適用する場合は、手法の適切性について国際的に一般的な評価を受けている手法を継続的に使用する」とされていることを踏まえ、X11を含めX12-ARIMAで季節調整が行われている。米国を始めとする太平洋沿岸国の多くの国でも、季節調整法は日本と同様X12-ARIMAである。しかし、欧州ではX12-ARIMAよりもTRAMO-SEATSを適用する国が多い。さらに、X12-ARIMAとSEATSを結合したX13-ARIMA-SEATSの実用化も進展している。

このように、国際的には、移動平均を主体とする方法から時系列統計モデルを主体

とする方法に少しずつシフトする傾向がうかがえる。

X13-ARIMA-SEATSについて、米国センサス局では2012年に実用を開始したが、2016年11月現在、総務省統計局の統計に適用した実証分析は見当たらない。そこで、本節では、X13-ARIMA-SEATSを労働力調査、CPI及び家計調査に適用し、第3節で検討した安定性と適切性の観点からX12-ARIMAと比較をしてみることにする。

なお、X13-ARIMA-SEATS の計算プログラムについては、米国センサス局のホームページよりダウンロードすることにより入手し、Time Series Research Staff, Centert for Statistical Research and Methodology, U.S. Census Bureau[25]に従い操作した。

4.1 X13-ARIMA-SEATSの概要

X13-ARIMA-SEATS は、事前調整と要因分解の部分に分かれている。事前調整の部分はX12-ARIMAと同じであり、regARIMAで数値の段差を取り除いてからARIMAモデルを適用する。要因分解の部分は2通りの中から選択が可能であり、1つはX11、他方はTRAMO-SEATSのSEATSである。したがって、前者を選べばX12-ARIMAと全く同じになる。

Victor Gómez and Agustín Maravall [26]によると、原数値を x_t 、ホワイトノイズを a_t としてSEATSの概要は次の式で示される。

$$\varphi(B) = (1 + \varphi_1 B + \dots + \varphi_p B^p)(1 + \Phi_1 B^s + \dots + \Phi_p B^{sp})$$

$$\theta(B) = (1 + \theta_1 B + \dots + \theta_q B^q)(1 + \Theta_1 B^s + \dots + \Theta_Q B^{sQ})$$

$$\varphi(B)(1 - B)^d(1 - B^s)^D x_t = \theta(B)a_t + c$$

$$x_t = \sum_{i=1, \dots, 4} x_{it}$$

ただし、各 x_{it} は、それぞれ x_{1t} 傾向変動、 x_{2t} 循環変動、 x_{3t} 季節変動、 x_{4t} 不規則変動である。また、これらの変動は、ARIMAモデルで表現する。

その後、正準相関分析を適用して、不規則変動成分と他の変動成分の分散比を最大化する。これにより変換された不規則変動成分の分散を大きくし、傾向変動、循環変動、季節変動は安定する。

これら正準変数に対して、Wiener-Kolmogorov型の予測、ないしはフィルタリングを行うことで、趨勢変動、傾向変動、季節変動の将来及び過去の系列を推定する。

4.2 X12-ARIMA と X13-ARIMA-SEATS の比較

ここでは、要因分解の部分で、X11 ではなく、SEATS を選んだときに、労働力調査、CPI、家計調査の月次系列の中で、それぞれの中で社会・経済の指標として最も注目されている指標である完全失業率、CPI 総合、2人以上世帯の消費支出について安定性と適切性の観点から X12-ARIMA、すなわち、要因分解の部分で X11 を選んだときと結果を比較してみることにする。

安定性については、以下の方法で比較する。

2015 年 12 月までの公表に適用している期間について季節調整を行い、2014 年 12 月までの公表に適用している期間について季節調整したものと比べて、同一年月の平均絶対変化率を比べ、小さいほうを安定性が高いとする。原数値の期間が追加されることにより大きく変化する部分は最新時点であることを踏まえ、平均をとる期間は①最も新しい共通の 1 年間である 2014 年の 1 年間、②2015 年の 1 年間の 2 通りとする。ただし、2014 年 12 月までを原数値とする場合については、2015 年の公表値と推定季節指数から季節調整値を推定する。

適切性については、2015 年 12 月までの公表に適用している期間について季節調整を行い、その結果を再度季節調整して、期間全体について同一年月の平均絶対変化率を比べ、小さいほうを適切性が高いとする。

なお、ARIMA モデルについては実際に公表に適用しているモデルを適用する。

それらの結果は以下のとおりである。

(1) 完全失業率

完全失業率については、季節調整の期間は 28 年間と長く。最も古い共通の 1 年間は 1987 年である。表 7 によれば、安定性、適切性のいずれも SEATS のほうが X11 よりも優れている。

表 7 完全失業率の平均絶対変化率 (%)

比較の視点		X11	SEATS
安定性	2014 年	0.43	0.30
	2015 年	0.51	0.36
適切性		0.15	0.01

(2) CPI

CPI 総合については、季節調整の期間が短く 2010 年を開始年としている。表 8 によれば、安定性、適切性のいずれも SEATS のほうが X11 よりも優れている。

表 8 CPI の平均絶対変化率 (%)

比較の視点		X11	SEATS
安定性	2014 年	0.07	0.04
	2015 年	0.09	0.07
適切性		0.01	0.00

(3) 消費支出

2人以上世帯の消費支出については、季節調整の開始年は2010年である。表9によれば、安定性のうち、①2014年の1年間についてはX11のほうが優れているものの、②2015年の1年間についてはSEATSのほうが優れている。

表 9 消費支出の平均絶対変化率 (%)

比較の視点		X11	SEATS
安定性	2014 年	0.23	0.33
	2015 年	0.54	0.26
適切性		0.11	0.01

4.3 全体評価

上記の結果をまとめると、安定性と適切性についてはX11よりもSEATSのほうが優れている場合が多く、季節調整法としてSEATSの選択が可能なX13-ARIMA-SEATSの有用性が認められた。今後、より多くの系列について同様の実証分析を重ねることが望まれるが、要因分解部分についてX11を超える優れた手法が存在していることは揺るぎないものとする。

センサス局法の発展経緯を振り返ると、以前は移動平均法を主体とするX11だけであったが、事前調整としてARIMAモデルが加わり、さらには、本体部分もモデル形式のSEATSの選択となった。EUではTRAMO-SEATSを採用することが多いという事情も加味すると、世界全体の傾向として、移動平均からモデルへの移行が認められる。今後も、事前調整と要因分解の両方でより洗練されたモデルが開発されることによって、いずれ世界標準から移動平均が消失するのではないかと予想する。このような傾向を踏まえると、日本においても、できるだけ早く事前調整部分だけでなく要因分解部分についてもモデルを採用するべきであるとする。

第2章 国勢調査の結果によるJGSSの結果の改善

民間統計調査には報告義務がないことから、公的統計のように高い品質を確保することが困難な場合が多い。しかし、その結果に公的統計の結果を活用することにより

結果の偏りを縮減することが可能となる。第2章では、国勢調査の結果により JGSS の結果の精度を向上する研究を紹介する。

JGSS は、大阪商業大学が 2000 年以降 1～3 年に 1 度、日本人の価値観、意識、行動を包括的に把握するために約 4,500 人¹⁶を対象に実施する国際比較が可能な社会調査である。JGSS-2000 では、あらかじめ設定された地域区分ごとに人口規模に比例するように 1995 年国勢調査（総務庁統計局）の調査区¹⁷を抽出し、そこからそれぞれ満 20～89 歳約 15 人の調査対象を無作為に抽出している。全国を 6 ブロック¹⁸に分け、各ブロック内で政令指定都市等（東京都特別区部を含む）、その他の市、町村別に層化している。各調査地点における対象者の抽出は、選挙人名簿¹⁹からの系統抽出により行っている。統計調査と異なり、調査結果については、集計よりも個別データとして研究用に利用することに重点を置いている。

JGSS では、確率比例抽出法を採用しているので、観測値に抽出率の逆数を掛けて加える線形推定の過程が単純化され、基本的には調査対象ごとの値を単に加算するだけで属性（回答の選択肢）別構成比を推定することができる。しかし、官公庁が標本調査により実施する統計調査では、結果の推定方法として、直近の全数調査などから得られた人口、従業者数等を基準とする比推定が良く用いられている。標本調査である以上、偶然に又は実務上の制約により偏った結果が現れることがあるので、このような比推定は、少しでも精度を高めるために既存の情報を活用しようという試みである。

一般に、単純無作為抽出を想定する線形推定に対して、層化比推定には以下の特徴がある。

- ①層化の基準指標について調査対象抽出に伴う偶然の偏りと未回収による情報の損失を一部回復することができる
- ②層化の基準指標について母集団との整合性を維持することが可能
- ③僅かではあるが偏りが発生

このような特徴を踏まえ、本章では、JGSS の個別データにおいて、国勢調査の結果をもとに層化比推定を行うことができる比推定乗率²⁰を導入するための研究について説明する。清水誠[27]は、第2回予備調査について、清水誠[28]は、JGSS-2000

¹⁶ 第2回予備調査の調査対象は約 1,200 人

¹⁷ 国勢調査の調査区は基本単位区を 2 つ程度結合して作成されている。基本単位区とは、住居表示実施地域においては街区を、住居表示実施地域以外の地域においては、街区に準じた区画を単位として設定されており、そこに含まれる世帯数はおおむね 25～30 世帯である。

¹⁸ 北海道・東北、関東、中部、近畿、中国・四国、九州別

¹⁹ 許可されない場合は住民基本台帳

²⁰ JGSS の個別データにおける比推定乗率の変数名は “weight” としている。

について個別データを用いて検証した研究であり、本章は、筆者のこの2つの研究に基づいている。

1 海外の状況

欧米には20世紀からJGSSに相当する一般社会調査GSS(General Social Surveys)が存在する。GSSや類似の調査を見ると、オーストリアのSSO(Sozialer Survey Osterreich)、イタリアのISI(Indagine Sociale Italiana)、スペインのASEP(Analisis Sociologicos Economicosy Politicos)、ブルガリアのGSS、ポーランドのPGSS(Polish General Social Survey)など、データセットに男女別×年齢別人口に一致させるための比推定乗率が付されているものがある。これらの比推定乗率は、更にその他の属性を調整する場合もあるが、その属性は、オーストリアSSOでは地域別×共同体規模別×就業形態別(又は職業別)に、イタリアISIでは職業別に、ブルガリアGSSでは教育の程度別に、ポーランドPGSSでは住宅の種類別にと、調査ごとに多様である。

他方、英国にはGSSに相当する調査としてBSA(British Social Attitudes Survey)がある。このデータセットには、分析をする前に“wtfactor”を利用して抽出率と世帯人員数に応じた換算をすべきであるとの指示があるが、比推定乗率は存在しない。また、ドイツのALLBUS(Allgemeine Bevölkerungsumfrage der Sozialwissenschaften)についても、旧東ドイツに偏った標本設計から全体像を適切に推計するための乗率付データが提供されているものの、比推定乗率は付されていない。

北米では、カナダのGSSが比推定²¹を系統的に実施している典型である。これは、月ごとに電話を通じたRDD(Random Digit Dialing)などにより世帯を対象に実施され、1996年に実施された第11期はSocial and Community Support、1998年に実施された第12期はTime Useというように、年ごとにテーマが設けられている。特に、1996年調査では、RDDによる通常の標本のほかに、労働力調査の終了者から選ばれた65歳以上の追加標本がある。このため、各調査の個別データは、これらの2種類の標本ごとに乗率が作成され、65歳以上については2種類の標本を結合するための調整係数もある。RDD標本の乗率は、電話番号に基づく抽出率の逆数、抽出世帯数に対する回答世帯数の割合の逆数、世帯内電話数の逆数、世帯人員数、労働力調査に基づく比推定に必要な地域別×男女別×年齢階級別乗率を乗じたものとなっている。

米国のGSSの変数には、世帯単位を個人単位に換算する“ADULTS”(1975年以降)、黒人の超過を調整する“OVERSAMP”(1982年)、複数の様式に基づく結果を

²¹ 正確には、Raking比推定法、すなわち、統計表において、行方向と列方向の合計(欄外)値を既知とした上で、原数値としてのセル値を行・列ごとに等倍して合計値に一致させる作業を、行・列ごとに繰り返すことによりセル値を推定する方法による。

調整するための“FORMWT”（1978、80、82 - 85年）、ISSP（International Social Survey Program）対象者の回答・無回答を区別するための“ISSP”が作成されているが、比推定は行われていない。

なお、比推定をするにはベンチマークとして対応する信頼される統計が必要であるが、その整備状況は国により異なるので、比推定の有無は GSS 側だけの事情ではない点に留意が必要である。また、ベンチマークの精度が低ければ、比推定は妥当な推計方法とは言えず、GSS との間で分布が似ていることを確認しても意味がないということも念頭に置く必要がある。

2 比推定乗率の意義

このように、諸外国のGSSや類似の調査の結果を見ても、カナダやイタリアなどのいくつかの国で国勢調査や労働力調査を基にした比推定が行われている。

また、米国のGSSについては、比推定は行われていないものの、結果の分布は全数調査などと大差がないことを確認しており、これらの調査と分布を比較・分析する視点は重視されている。実際、JGSSについても、このような視点に基づく検証を行ったが、調査対象抽出に伴う偶然と実地調査における調査対象者の不在や拒否などにより年齢構成に偏りがあることがわかった。

そこで、JGSSについても調査対象の抽出に用いた地域区分を基に、男女と年齢という基本的な属性で区分した人口規模による比推定を試みることにした。その際、比推定により結果がどう変化するのか、既存情報と比較してどのような特徴があるのかを基礎情報として明らかにすることとした。

3 比推定乗率の作成方法

第2回予備調査においては、2種類の比推定用乗率を作成することとし、①地域別（6区分）×男女別（2区分）×年齢10歳階級別（6区分）、②市町村の規模別（3区分）×東西地域別（2区分）²²×男女別（2区分）×年齢10歳階級別（6区分）のそれぞれについて、有効回答数当たり人口を計算した。

JGSS-2000では、第2回予備調査と比べて有効回答数が多いことにかんがみ²³、当初、上記①と②の区分をクロスさせた精緻なものを作成することとした。つまり、地域別（6区分）×市町村の規模別（3区分）に、男女別×年齢10歳階級別をクロスするという調査対象の抽出方法を踏まえた分割の仕方を想定した。しかし、これでは標本規模0となる箇所が出現し、比推定乗率の作成が困難になることから、市町村の

²² 東日本は北海道・東北、関東、中部、西日本は近畿、中国・四国、九州

²³ 有効回答数は、第2回予備調査が776、JGSS-2000が2,893

規模を大括りして市郡別とした。すなわち、地域別（6区分）×市郡別（2区分）×男女別（2区分）×年齢10歳階級別（6区分）²⁴に有効回答数当たり人口を計算し、比推定乗率とした。基準として用いた人口は、2000年10月1日国勢調査（総務庁統計局）の総人口である。

比推定乗率、すなわち、有効回答数当たり人口は次の範囲に分布している。

最小 1.5万（中国・四国・郡部70歳代・80歳代女性）

最大 8.6万（九州・市部30歳代男性）

最大／最小 5.8倍

また、有効回答数は1区分に最低3となっている。

4 比推定の方法

比推定乗率により試算した比推定の結果を単純合算した結果と比較すると、以下のようになる。

4.1 推定式

JGSS は確率比例抽出法を用いていることから、調査対象は等確率で抽出されている。単純無作為抽出法を想定して構成比を推定する方法は、標本規模を n 、 x_j と y_j は、標本に含まれる観測属性で、0又は1とすると、

$$\hat{p}_x = \frac{\sum_{j=1}^n x_j}{\sum_{j=1}^n y_j}$$

となる。

これに対して、作成した比推定乗率をもとに結果を推定する方法は、国勢調査の結果を踏まえた層化推定となり、具体的には以下のようになる。

層数を l 、層 i に属する母集団の大きさを N_i 、標本規模を n_i 、 x_{ij} と y_{ij} は、層 i に属する観測標本の属性で、0又は1とすると、

²⁴ 70～79歳と80～89歳は統合して1区分とした。また、年齢不詳は都道府県別・男女別に合計（市郡計）、市部についてそれぞれ按分し、郡部は合計から市部を引くことにより計算した。

$$\hat{p}_x = \frac{\sum_{i=1}^l \frac{N_i}{n_i} \sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^l \frac{N_i}{n_i} \sum_{j=1}^{n_i} y_{ij}}$$

と推定される。

このとき、 $w_i = \frac{N_i}{n_i}$ が比推定乗率である。

他に、母集団の大きさを N として、層化比例抽出法を想定した

$$\hat{p}_x = \frac{1}{N} \sum_{i=1}^l N_i \frac{\sum_{j=1}^{n_i} x_{ij}}{\sum_{j=1}^{n_i} y_{ij}}$$

が想定できるが、 $\sum_{i=1}^{n_i} y_{ij}$ が 0 になることがあるので実用的ではない。

4.2 推定結果の比較

地域、男女、年齢ごとに、JGSS-2000の個別データをそのまま合算した結果と、比推定の結果で属性別構成比をとり比較してみた。ここで取り上げた属性は、調査日の前後（2年以内）に実施された調査でJGSS-2000よりも有効回答数の多い調査（仮にこれを大規模調査と呼ぶことにする。）においても調査しているもので、データの妥当性を比較するための基礎資料として併せて大規模調査の結果による構成比を対比させた。

JGSS-2000では、調査対象を個人単位で抽出しているので、世帯、家族、住宅などの単位で分析をする場合には、使用する変数について、各個人の値を当該個人の同居家族数の逆数で加重する必要がある。これは、比推定乗率を使用しない場合でも留意すべき事項である。

- (1) 比推定により、当然のことではあるが、地域別×男女別×年齢別構成が国勢調査の結果と一致する。例えば、原データでは女性の割合が54.4%であるのに対し、比推定により51.5%と国勢調査の人口に一致するように変更される。また、20歳代は原データでは13.6%しか存在しないのに対し、18.2%に変更される。これにより、これらの属性の構成を現状に近づけた状態でその他の属性に着目した分析が可能となる。
- (2) 比推定により、上記の属性のみならず配偶関係、教育の程度、世帯属性、従業上の地位、産業、労働時間、電子メールの利用の有無、インターネットによるシ

ショッピング・バンキングの有無、世帯人員数、住宅の種類、住宅の広さについて属性別結果の多くが大規模調査の結果に近づく。これらの属性は地域、市郡、男女、年齢のいずれかと強い関係がある。

- (3) 逆に、地域、市郡、男女、年齢のいずれとも関係が弱い属性については、(2)よりも頻度は少ないものの、比推定によりむしろ現状と離れることもある。例えば、職業について、原データのままで現状よりも大きい事務職の割合は、20歳代後半前後の比較的若い人達で大きいのが、年齢を補正する比推定乗率によりこれらの世代が拡充されるので、現状から一層離れることとなる。
- (4) JGSS-2000では、第2回予備調査と比較して有効回答数が多いため、原データによる推定結果が国勢調査等による大規模調査の結果により近い値となっている。例えば、未婚者の割合は、国勢調査では22.6%であるのに対して、原データでは15.1%に過ぎない。しかし、第2回予備調査では13.3%であったので、それに比べれば国勢調査の結果に近づいている。したがって、比推定による変化の大きさも第2回予備調査のときと比較して大方小さくなっている。

4.3 標準誤差

一般に、標本誤差は非標本誤差より小さいので、標本誤差のみで精度を評価すべきではない。しかし、ここでは標本誤差について、単純無作為抽出を想定する推定と層化比推定との間で、標準誤差を計算し、比較した。

その際、単純無作為抽出法を想定した場合については、標準誤差は

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j, \quad \bar{y} = \frac{1}{n} \sum_{j=1}^n y_j \text{ として}$$

$$\sigma(\hat{p}_x) \approx \hat{p}_x \sqrt{n \left(\frac{\frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2}{\left(\sum_{j=1}^n x_j \right)^2} + \frac{\frac{1}{n-1} \sum_{j=1}^n (y_j - \bar{y})^2}{\left(\sum_{j=1}^n y_j \right)^2} - 2 \frac{\frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^n x_j \sum_{j=1}^n y_j} \right)}$$

とした。層別総合比推定を想定した場合については、標準誤差は

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}, \quad \bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij} \text{ として}$$

$$\sigma(\hat{p}_x) \approx \hat{p}_x \sqrt{\sum_{i=1}^l N_i(N_i - n_i) \frac{1}{n_i} \left(\frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 - 2 \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i) \right) \frac{1}{\left(\sum_{i=1}^l \frac{N_i}{n_i} \sum_{j=1}^{n_i} x_{ij} \right)^2 + \left(\sum_{i=1}^l \frac{N_i}{n_i} \sum_{j=1}^{n_i} y_{ij} \right)^2 - 2 \left(\sum_{i=1}^l \frac{N_i}{n_i} \sum_{j=1}^{n_i} x_{ij} \right) \left(\sum_{i=1}^l \frac{N_i}{n_i} \sum_{j=1}^{n_i} y_{ij} \right)}$$

とした。

ただし、世帯の構成を見る場合には世帯人員数を h_{ij} として

$$x_{ij} \rightarrow \frac{x_{ij}}{h_{ij}}, \quad y_{ij} \rightarrow \frac{y_{ij}}{h_{ij}}$$

として上式に代入した。

その結果は表 10 のとおりであり、層化の基準指標については層化比推定の標準誤差は消失したが、他の指標についてはほぼ同等の結果となった。

表 10 単純和と比推定の項目別標準誤差

個人の基礎的項目			個人の就業関連項目			世帯の項目					
	単純和	比推定		単純和	比推定		単純和	比推定			
地域	北海道・東北	0.6	0.0	就業 状態	従業	0.9	0.7	世帯 人員数	単身	1.2	1.2
	関東	0.8	0.0		休業	0.2	0.2		2人	1.0	1.0
	中部	0.7	0.0		失業	0.3	0.3		3人	0.8	0.8
	近畿	0.7	0.0		家事	0.7	0.6		4人	0.6	0.6
	中国・四国	0.6	0.0		学生	0.0	0.1		5人以上	0.5	0.5
	九州	0.6	0.0		その他	0.6	0.5		持ち家・戸建て	1.1	1.1
市町村 の規模	政令指定都市等	0.7	0.0	従業上 の地位	経営者・役員	0.6	0.6	住宅 の種類	持ち家・集合住宅	0.5	0.5
	その他の市	0.9	0.0		常雇	1.3	1.2		借家・戸建て	0.4	0.4
	町村	0.8	0.0		臨時雇・パート・アルバイト	1.2	1.1		借家・集合住宅	1.0	1.1
性別	男	0.9	0.0	企業 の従業者 数	派遣社員	0.3	0.3	住宅 の広さ	その他	0.2	0.2
	女	0.9	0.0		自営業主・内職	1.0	1.0		50㎡未満	1.0	1.1
年齢	20～29歳	0.6	0.0	産業	家族従業者	0.7	0.7	世帯の 年間 収入	50㎡以上100㎡未満	1.0	1.0
	30～39歳	0.7	0.0		1～29人	1.8	1.7		100㎡以上	1.1	1.1
	40～49歳	0.7	0.0		30～99人	1.5	1.4		100万円未満	1.1	1.2
	50～59歳	0.8	0.0		100～499人	1.5	1.4		100万円以上450万円未満	1.6	1.5
	60～69歳	0.7	0.0		500～999人	0.9	0.8		450万円以上1000万円未満	1.4	1.3
	70～89歳	0.7	0.0		1000人以上	1.5	1.3		1000万円以上	1.0	1.0
配偶 関係	有配偶	0.8	0.8	職業	官公庁	1.3	1.2				
	未婚	0.7	0.6		第1次産業	0.7	0.7				
	離死別	0.6	0.5		第2次産業	1.3	1.2				
教育 の程度	小学・中学	0.8	0.7	第3次産業	1.1	1.1					
	高校・旧中	0.9	1.0	専門的・技術的	1.0	1.0					
	短大・高専	0.6	0.6	管理的	0.5	0.4					
	大学・大学院	0.7	0.7	事務	1.2	1.2					
世帯主	0.9	0.7	販売	0.9	1.0						
年間 収入	100万円未満	1.3	1.1	サービス	0.8	0.8					
	100万円以上450万円未満	1.4	1.3	保安	0.3	0.3					
	450万円以上1000万円未満	1.4	1.2	農林漁業	0.7	0.7					
	1000万円以上	0.6	0.6	運輸・通信	0.6	0.6					
	電子メールの利用（世帯主）	1.7	1.7	生産工程・労務作業	1.3	1.3					
インターネットによるショッピング・バンキング	1.0	1.1	労働 時間	1～14時間	0.7	0.7					
子供数	0人	0.6	0.8	15～34時間	1.2	1.2					
	1人	0.8	0.9	35～42時間	1.3	1.3					
	2人	1.0	1.1	43～48時間	1.1	1.1					
	3人	0.9	0.9	49～59時間	1.0	1.0					
	4人	0.4	0.4	60時間以上	0.9	0.9					
	5人以上	0.3	0.3								
生活 の程度	上	0.1	0.2								
	中の上	0.6	0.6								
	中の中	0.9	1.0								
	中の下	0.9	0.9								
	下	0.5	0.5								

5 その後の対応

本研究の結果を踏まえ、JGSS においては比推定乗率が提供されるようになった。調査環境が厳しくなる中で、属性ごとの偏りは避けられないが、データセットにおける比推定乗率を用いて比推定を行うことにより、調査結果の改善を図ることが可能である。

このような、公的統計の結果を活用して民間統計の精度向上を図る研究は、他の統計調査や世論調査にも応用できると考える。幸い、日本においては国勢調査という精度が高く、基本属性による区分が可能なベンチマークが存在していることから、日本でこそ有益な研究であると考ええる。

第Ⅲ部 統計利用の高度化研究

第Ⅲ部においては、公的統計の利用の高度化に関する研究として、時系列又は地域別に実データを各種のモデルに適用することにより、経済や社会に関して得られた知見を示す。

第1章は、人口が経済に及ぼす影響として、1969年度から1998年度まで、各種の統計データから実質GDPをCobb-Douglas型で要因分解すると、供給については物的資本分配率が3割程度、人的資本分配率が4割程度、人口が3割程度となり、1988年度まで代入した場合とほとんど変化していないこと等から、バブルが崩壊した後も供給側から見た経済成長の構造には大きな変化がなかったことを示した。また、1970年から90年代後半までのOECD加盟国の統計データから、実質GDPをCobb-Douglas型で要因分解すると、供給については物的資本分配率が3割程度、人的資本分配率が2割程度、労働投入量分配率が5割程度となり、経済成長の先進国間格差の主要な要因は労働投入量であること、需要については物的資産分配率が5割程度、人的資産分配率が2割程度、人口分配率が3割程度となり、国間格差の主要な要因は物的資産であるとしている。さらに、同様の方法を1975年度から1988年度までと1975年度から1998年度までについて都道府県間に適用すると、いずれにおいても供給側、需要側とも物的資産の分配率は2割程度と低い一方で、供給における労働投入量、需要における人口の分配率は5割前後～6割弱と高く、都道府県間格差についてもマンパワーが大きな要因であることを実証した。

第2章は所得が出生に及ぼす影響として、JGSS-2002の個別データにおいて、生涯レベルの夫の所得、妻の賃金、就業率及び出生数を計算し、Butz and Ward Modelを適用すると、係数は同モデルの条件を満たしていることから、出生に対して所得は夫婦で同一方向に作用するのではなく、夫についてはプラスに、妻についてはマイナスに作用するという旧態の構造が依然として根付いていることを示した。

第3章は科学技術研究調査による企業の生産性の要因分析として、科学技術研究調査の個別データを生産関数に適用することにより、多くの産業において生産性に最も大きな影響を及ぼしている要因は労働力であること、また、研究投資の効果は産業によって異なり、化学系、機械系の製造業で高く、第3次産業では認められないことを示した。

第1章 人口が経済に及ぼす影響

日本の人口は、選挙区の区分や社会保障政策の基礎資料、各種推計の基礎データなど統計として重要な役割を果たしているが、それ自体経済成長に大きな影響を及ぼしている。一般的には消費を通して働く分配面へのマイナス効果が取り上げられる。例えば、開発途上国に見られるように、過度な人口は貧困に直結し、時に人口は経済成長の障壁になる。しかし、労働を通して働く生産面へのプラス効果もある。日本と類似の先進国に限れば、労働の原資としてのマンパワーが経済成長の1つの要因になっている²⁵。また、このような現象は、人口統計がある意味で経済統計の重要な要素にもなり得ることを示している。

一般に、経済成長とは長期間にわたる経済の実質生産額の増分を意味する。日本の経済成長は1960年から1990年の間に歴史的にも珍しいほど高い水準であったが、1990年代から衰え、それからの間を「失われた時代」と称することもある。経済成長の衰退には様々な要因が想定されるが、人口要因もその1つとして挙げられる。本章では、基本的なモデルをベースに、人口要因による経済成長に対する影響を時系列で、また、地域横断的に概観し、定量的な評価を試みることにする。特に、日本の人口の減少幅は大きくなるという見方が一般的であり、それが経済成長に対してどの程度になるのかを計量してから将来の経済成長の可能性を議論する必要があるものと考えられる。

一方で、1990年代にはそれ以前と比較して変わらない傾向もある。人口要因の中で、少子高齢化はそのような傾向の典型的な例である。本章には経済成長の要因の中で1990年前後で一貫しているものを抽出し、今後の経済成長について議論を深める際の基礎資料を提供しようという試みもある。

経済成長については、理論的には様々なモデルが考案されているが、それぞれ利用目的や条件が定められているので、モデルの適用に当たっては分析の目的と対象を確認しておく必要がある。統計はモデルの実証データとして有用であるが、分析の対象を正確に認識するためにも、実態を正確に表現する統計を入手する必要がある。

一方、経済成長に利用される統計としては、一般にはSNAやそれを元に作成された県民経済計算が用いられるが、これらは独自の概念を元に各種の推計が組み込まれているので、利用に当たっては様々な注意が必要になる。細部にこだわらず概観に留めることも重要な留意点の1つであるが、国、都道府県という大きな地域単位で、長期的に、産業区分を設けずに比較をしようとする分析には格好の材料であると考えられる。

²⁵ 例えば、収入が出生率に影響を及ぼすように、逆に経済的要因が人口要因に影響を及ぼす場合もある。人口と経済は相互依存関係にあるという見方が一般的である。

1990年代について、このような観点から分析を行った研究に清水誠[29]があり、以下でその研究を示す。本研究の基本的な方法は、Mankiw et al.[30]に倣っているが、本研究は、世界に先駆けて人口減少の局面を迎える日本を中心に、データの特徴に留意しつつ、時系列と地域間の両面で、かつ名目と実質の両方で試算を行うことにより要因間の構造を複眼的に把握するものである。

1 時系列分析

まず、地域を日本に限定し、1990年頃を境に経済成長と人口にどのような変化が起き、2000年以降にどのような変化が見込まれるのか、また、経済成長を人口要因などで分解して、どの要因が大きく寄与しているのかを概観することとする。要因分解に当たっては、人口要因以外は1990年頃を境に変化が起きていない要因に区分するよう努めることとする。

1.1 人口と労働時間に着目した供給面の評価

最初に、経済成長を生産力として供給面から捉え、1960年度から2000年度までの労働時間と人口の推移による変化を時系列で概観し、21世紀初頭におけるこれらの推移を推計した。

GDPは1国の生活水準を示す指標であるとともに生産力を示す指標でもある。GDPと人口の間には

$$\text{GDP} = \frac{\text{GDP}}{\text{人口}} \times \text{人口}$$

という関係がある。これは、生活水準として需要面から捉える場合には有用な分解の仕方であるが、近年のように、短時間就業の増加やワークシェアリングの導入など、就業形態が多様化している世の中では、労働時間（就業時間の総和）を主体とする

$$\text{GDP} = \frac{\text{GDP}}{\text{労働時間}} \times \frac{\text{労働時間}}{\text{就業者数}} \times \frac{\text{就業者数}}{\text{15歳以上人口}} \times \frac{\text{15歳以上人口}}{\text{人口}} \times \text{人口}$$

という分解の仕方のほうが、供給に関する議論の進展には役立つものとする。

実際、日本で1960年度からの推移を実質値で見ると、**図3**のように、

1人当り $\text{GDP} = \frac{\text{GDP}}{\text{人口}}$ は増加傾向で推移しているが、1990年前後で増分は著しく減少

している。その傾向を直線的と言うには無理がある。これに対して、1人当り GDP

を $\text{15歳以上人口比} = \frac{\text{15歳以上人口}}{\text{人口}}$ 、 $\text{就業率} = \frac{\text{就業者数}}{\text{15歳以上人口}}$ 、 $\text{平均就業時間} = \frac{\text{労働時間}}{\text{就業者数}}$

労働時間当りGDP = $\frac{\text{GDP}}{\text{労働時間}}$ の4要素に区分すると、次のように、ほぼ一貫した傾向

の合成による説明が可能になる。まず、15歳以上人口比は少子高齢化により増加傾向で推移している。また、就業率と平均就業時間は減少傾向で推移しており、1990年以降に趨勢的な変化があったとは言い難い（図4、図5）。図には掲げていないが、2001年1月に公表された労働力調査の結果を見ても、就業率と平均就業時間の減少傾向は継続している。ちなみに、1990年代に史上最長の景気拡大を遂げた米国では、就業率と平均就業時間は1983年以降一時的な減少を除いて増加傾向で推移している。このように、就業率と平均就業時間は経済成長を追跡する上で欠かせない要素を構成している。一方、日本の労働時間当りGDPは増加傾向で推移しており²⁶、しかも1990年以降の傾きの減少は、人口当りで見ると場合よりも小さい（図6）。

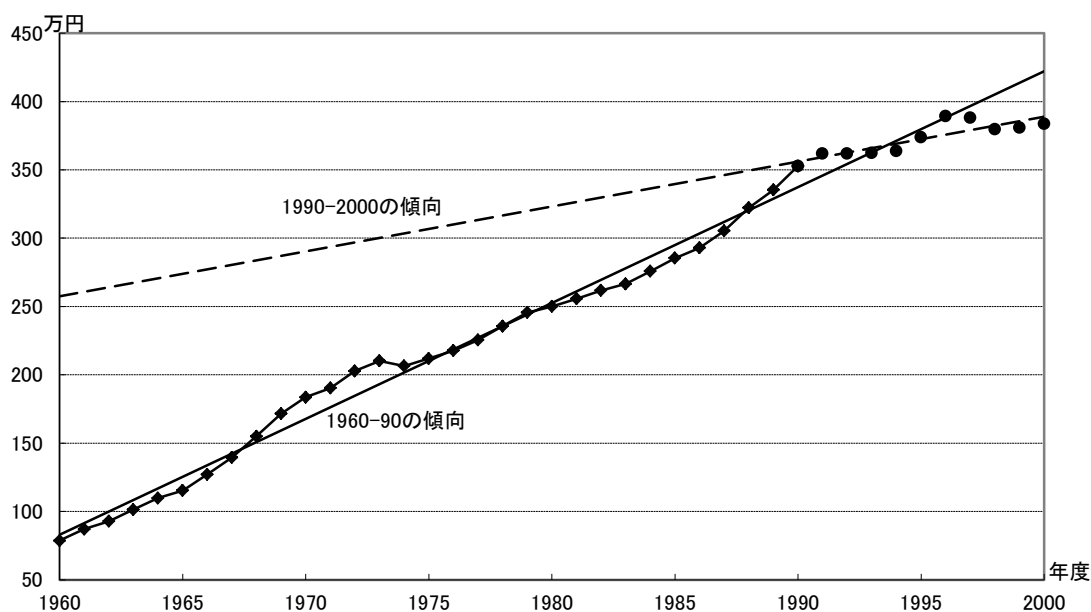


図3 1人当りGDP（実質）の推移

注 68SNAによる1990年基準値（以下同じ）

²⁶ GDPは68SNAによる実質値、人口は10月1日現在人口推計、労働時間は労働力調査による平均就業時間の年平均に従業者数の年平均を乗じた延週間就業時間、就業者数は労働力調査による年平均である。

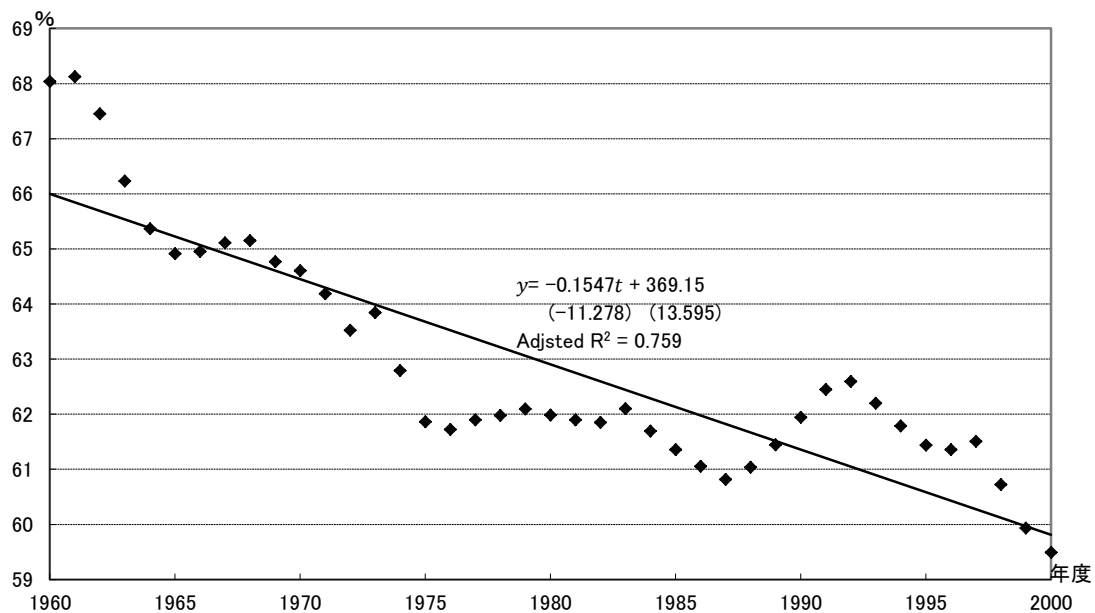


図4 就業率の推移

注 () は t 値である。(以下同じ)

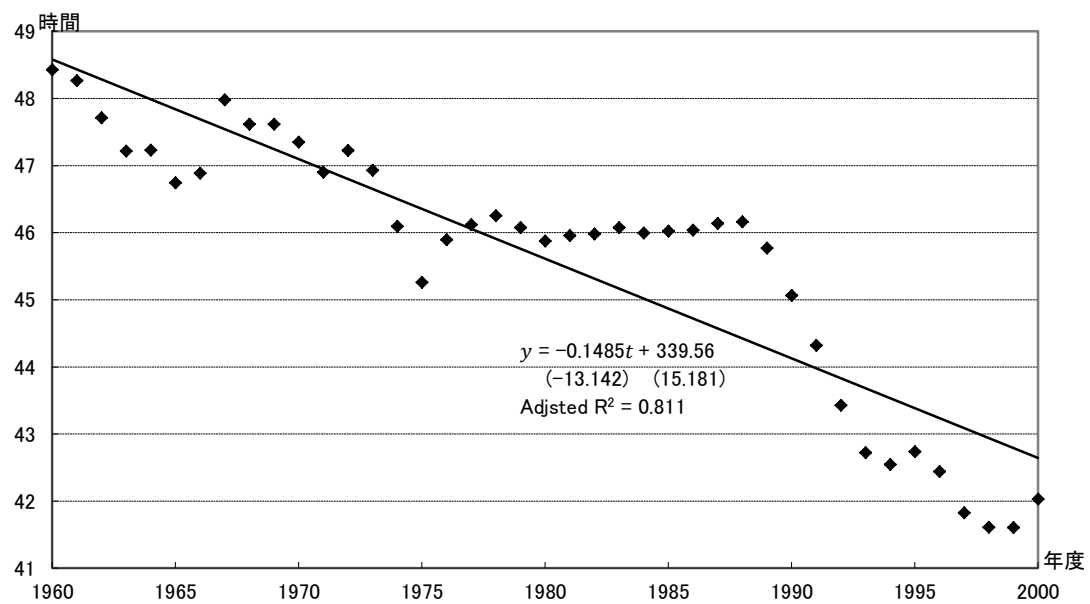


図5 週間平均就業時間の推移

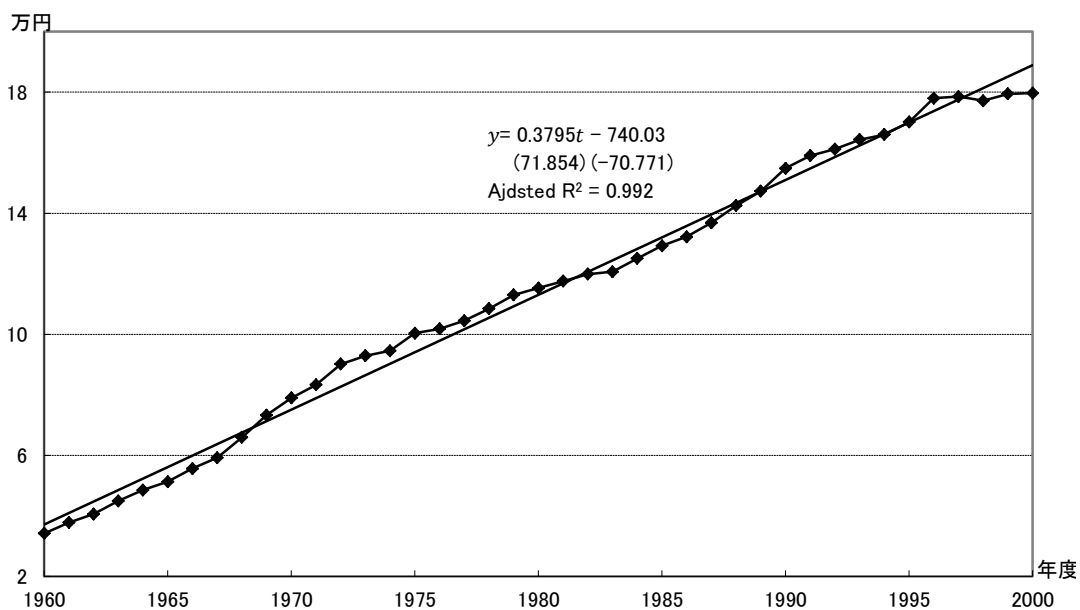


図6 労働時間当り GDP（実質）の推移

日本において、上記の4要素と人口の対前年度増加率の推移を見ると、図7のとおりである。これまでの経済成長の牽引役は、この中では技術進歩や能率向上などを含む労働時間当り GDPであったが、人口増加も第2次ベビーブーム期の1970年代前半には大きなシェアを占めていた。これに対して、就業率は景気好調期を除いておおむねマイナスに寄与し、長期的には減少傾向で推移してきた平均就業時間も1989年度以降マイナスに寄与する年度が多くなっている。

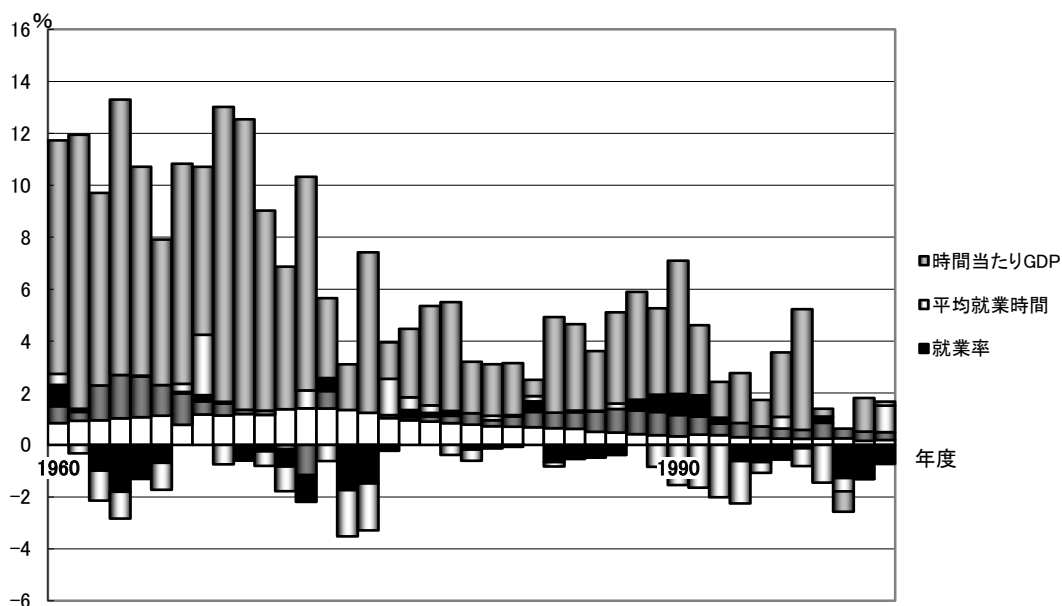


図7 GDP（実質）対前年度増加率の要因分解

注 1972年度の人口増加率は沖縄県を除いて計算

GDP を名目値で捉えるにはこれに GDP デフレータの増分を上乗せすれば良い。GDP デフレータは増加傾向で推移し、1960年度を100とした2000年度の指数は444であり、この間のGDP デフレータも直線近似することができる。しかし、その後のデフレに配慮すると、むしろ1995年を頂点とする曲線を当てはめるほうが良い。

人間は生まれると必ず1年に1歳年を取り、平均的には平均寿命（将来推計された水準）程度のところで死ぬことから、人口は経済と比べて将来推計しやすい面がある。経済成長についても、部分的にでも人口で説明できるのであれば、その分将来推計の可能性が増すこととなる。

2002年1月に公表された将来推計人口（中位推計）から計算した2025年度の人口減少率はマイナス0.5%である。先に記した要因分解式において、仮に、将来の人口と15歳以上人口比がこの推計結果に従い、就業率、平均就業時間及び労働時間当たりGDPが1960-2000年度のタイムトレンドに沿って推移するとすると、GDPの要因別増減は図8のようになる。これによると、2007年度に人口が増加から減少に転じ、しかもその減少率が拡大することにより、労働時間当たりGDPがこれまでどおりに増加し続けたとしても、高い経済成長を達成することが困難な様子が見えてくる。

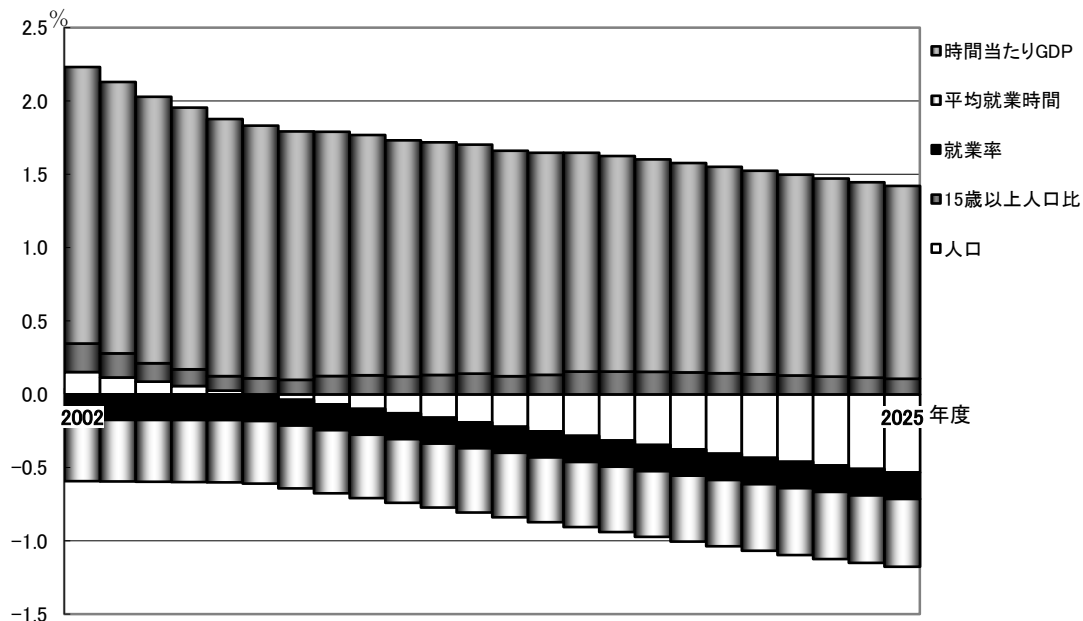


図 8 GDP（実質）対前年度増加率の要因分解 推定値

上記における人口減少率は、人口が直接経済成長に及ぼす影響であり、経済成長の他の4つの要因、すなわち、15歳以上人口比、就業率、就業時間、労働時間当たりGDPは人口と無関係に推移しているとは言えない点に注意が必要である。しかし、「失われた時代」とも呼ばれる時期でも、なお以前と同じ方向で推移していた傾向を抽出し、それらについては将来も継続するという見方は、将来推計の方法として、単純ではあるが合理的な手法であると考えられる。

1.2 人口要因と資産要因の区分

1.1 で紹介した経済成長の要因分解では、各変数が比較的単純な動きをしているために、過去の長期的動向を概観し、将来推計をする場合に便利であったが、各変数は必ずしも独立とは言えず、相互作用が存在することが前提となっている。これに対して、ここでは、ある程度それぞれが独立な要因に区分した上で、人口要因の占める割合について分析を行う。ただし、生産力として供給面から見る場合には、前節同様時間を加味した労働投入量で捉えることとする。

まず、供給面から経済水準の要因分解をするために、資本ストックを物的資本ストックと人的資本ストックに区別し、以下の Cobb-Douglas 型を想定する。

$$Y = K_p^\alpha H_p^\beta W^{1-\alpha-\beta} \quad 0 < \alpha < 1, 0 < \beta < 1, \alpha + \beta < 1$$

Y : 経済水準、

K_p : 物的資本ストック、

H_p : 人的資本ストック、

W : 労働投入量

具体的には、 Y は68SNAによるGDP（実質）、 K_p は68SNAによる純固定資産（実質）から住宅分を除いたもの、 W は労働力調査による延週間就業時間とする。ただし、いずれも資産について同一水準で遡りが可能な1969年度を基準として指数化しておくこととする。問題は人的資本ストックである。人的資本ストックは各個人が有する知識や技術などに基づく一般的な能力であり、その源は個人の教育、訓練、経験などである。これを量として把握することは困難であり、いくつかの方法が研究されているものの、人的資本ストックとして作成された指標は、物的資本ストックや労働投入量と比較して、定量化の方法により様々な動向を示し得るので、指標として限界がある。しかし、教育水準が高く、勤勉な人口は技術進歩を生み出す源泉であることを考えると、人的資本ストックの結果として技術進歩が発生するとの見方もできるので、経済成長の実態を表現するために有用な指標でもある。ここでは仮に、生産を向上させる能力の総量は専門的・技術的職業従事者数に比例すると考え、 H_p を労働力調査による専門的・技術的職業従事者数を指数化したものとしたが、実際には学歴、研究水準、教育費など様々な要因に連動している可能性が十分想定される。したがって、モデルの適用結果はかなり幅広く捉える必要がある。

上記のモデルの対数をとると、

$$\ln Y - \ln W = \alpha(\ln K_p - \ln W) + \beta(\ln H_p - \ln W)$$

となる。

一般に、経済成長理論の多くは、上記のように経済水準を供給側から捉え、資本、労働、技術進歩の成長の程度やこれらの分配率に注目するものである。しかし、経済成長を生活水準として捉える場合には、むしろ支出に着目し、需要側から消費や投資などの促進要因を考えることが重要である。生産によってもたらされた所得は何らかの物品やサービスに変換されてはじめて豊かさが享受されるものだからである。実際、経済成長についての一般的な議論は、生産性の向上手段よりも消費と投資の上昇手段に向きがちである。これには、生産性の向上は、結局は個々の産業や企業の技術進歩に依存する面が大きく、経済情勢は需給両面で相互に影響し合いながら変化するという前提の下で、需要側のほうが政策として一般的な検討がしやすいという理由もある。また、実証分析の対象となるGDPも、日本においては生産ではなく支出から計算されたものである。ただし、需要側からのアプローチにおいては、成長の構成要因が不明瞭であるという制約もある。このような制約に配慮すれば、1つの試算であるが、ここでは、需要としての経済水準の要因分解をするために、GDPが物的資産、人的資

産、人口の3要素により決定されるものとする。物的資産は、固定資産、土地資産、金融資産などであり、一般的にはこれらが多いほど消費や投資も多いという実態を踏まえて配慮することにした変数である²⁷。一方、人的資産は能力の総量であり、学歴が高い者ほど消費支出が多く、高学歴者を多く採用している企業ほど投資額が多いなどの想定により設定した変数である。モデルとしては、供給側から見る場合と同様に、以下のCobb-Douglas型を想定する。

$$Y = K_l^\alpha H_l^\beta N^{1-\alpha-\beta} \quad 0 < \alpha < 1, 0 < \beta < 1, \alpha + \beta < 1$$

Y : 経済水準、

K_l : 物的資産、

H_l : 人的資産、

N : 人口

具体的には、 K_l は68SNAによる国富（正味資産）を純固定資産デフレータにより実質化したもの²⁸、 H_l は国勢調査などから推定した大学卒業者数²⁹とする。ここでは仮に、人的資産は大学卒業者数に比例するとしたが、様々な要因が想定し得ると考えられるので、結果についてはかなり幅広に捉える必要がある。

この式の対数をとると、

$$\ln Y - \ln N = \alpha(\ln K_l - \ln N) + \beta(\ln H_l - \ln N)$$

となる。

上記で挙げた供給側モデル、需要側モデルに投入される独立変数並びに従属変数をすべて時系列で並べると図9のようになる。

²⁷ 例えば、家計については、1999年全国消費消費実態調査の結果の中で、貯蓄額で見ても、資産額で見ても、階級が上昇するほど消費支出金額が大きくなっている。また、2000年頃の法人企業統計調査の結果では、資本金額階級が上昇するほど企業当り設備投資額が大きくなっている。

²⁸ 実質化の方法として、国富においては純固定資産以外の土地資産などの占める割合が高いので、純固定資産デフレータを使用することについては異論があろう。ただし、ここでは、少なくとも1998年まで適用した結果については、例えば、GDPデフレータやCPIを使用するなど、方法を少し変更しても結果の大要に変化がないことを確認している。

²⁹ 西暦1桁0年については国勢調査の大学・大学院卒業者数+25歳以上在学者数とした。その他の年については各年学校基本調査の大学卒業者数を加算した。西暦1桁0年における実績値と10年前からの積み上げ値との差は10等分して各年加算した。

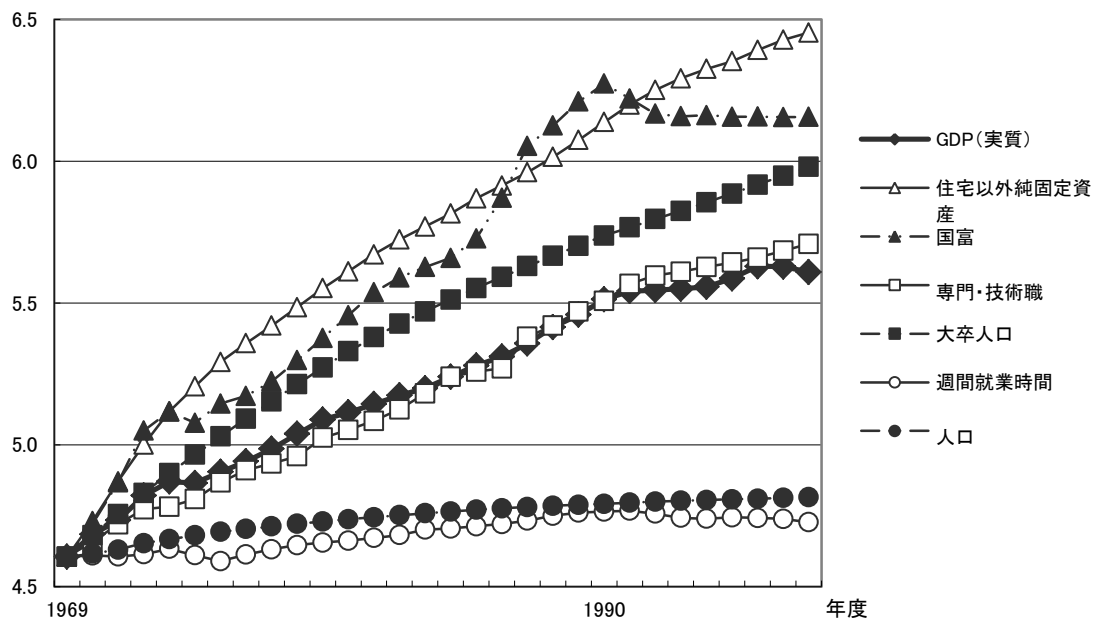


図9 各変数の推移 1969年度=100として指数化したものの対数値

図9に掲載した変数は、供給側、需要側のいずれから見ても物的資産要因（住宅以外の純固定資産、国富）の増分が最も大きく、人口要因（週間就業時間、人口）の増分が最も小さい。実質GDPの増分の大きさは両者の間に位置する。一方、人的資産要因の増分は物的資産要因と実質GDPの間である。これらのことから、物的資産要因が実質GDPの増分を拡張し、人的資産要因（専門的・技術的職業従事者数、大卒人口）もある程度の増分を促進する一方で、人口要因は増分を抑制する働きをしていることになる。実質GDPは1990年代に増分が停滞しているが、国富と週間就業時間が類似の傾向で推移している。いずれにしても、おおむね右肩上がりの要因が α 、 β 、 $1-\alpha-\beta$ という係数で程度を押さえながら実質GDPが合成されるモデルを想定していることになる。

上記の対数線形モデルを最小二乗法により当てはめた結果を表11に示す。表11は供給、需要について、データは少数になるが、バブル期の1988年度まで代入した場合と、バブル崩壊後の1998年度まで代入した場合の2通りについて比較した結果を示している。決定係数（自由度修正済）は高いが、本来は0であるはずの定数項が0より少し離れ、また1998年度までの需要については、Durbin-Watson比が小さくなっている。

最初にバブル期の1988年度までデータを代入した場合について概観すると、供給側の分配率は、物的資本ストックが3割程度、人的資本ストックが4割程度、人口が3割程度、需要側の分配率は、物的資産が2割程度、人的資産が3割程度、人口が4

割程度となる。つまり、供給、需要のいずれから見ても、3～4割程度は人口要因により説明できることとなる。なお、供給について、物的資本ストックの分配率は一般には3分の1程度とされており、本研究はそれに近い結果が得られたこととなる³⁰。人的資本ストックと労働投入量の分配率を区別して大きさを測定することは困難ではあるが、労働の対価のうち最低賃金が労働投入量に相当し、それに上乗せされた分が人的資本ストックに相当するとの見方ができる。しかし、その場合、最低賃金に満たない報酬で働く短時間就業者の位置付けが困難になる。そこで、ここでは逆に、国民全体の最低限の生活が保障される水準で労働投入量に相当する労働の対価の総額が決定されるという見方をすることにする。近年の老齢基礎年金の水準や個人住民税の課税最低限所得金額を見ると、最低限の生活は1人当り年間80万円程度で保障されるとみなされていると考えられる³¹。これに2000年10月における総人口を乗じて2000年度の国民所得で割ると、0.27程度となる。厳密にはこれを各年で計算して平均する必要があるが、表11の労働投入量の分配率はこれに近い水準である。

次にこれらを1969年度から1998年度まで代入した場合と比較する。供給側の分配率は、物的資本ストックが3割程度、人的資本ストックが4割程度、人口が3割程度で1988年度まで代入した場合とほとんど変化していない。つまり、バブルが崩壊した後も供給側からみた経済成長の構造には大きな変化がなかったことになる。他方、需要側の人口は10年前の4割程度から3割程度に分配率を減らし、その分人的資産が5割近くに増加している。これは、 β が1990年前後で変わらないとする仮定に無理があることを意味する。

³⁰ 実際、1969～1998年度の間、労働分配率（国民所得に対する雇用者所得と農林水産業に関する個人企業所得の割合）は平均69.0%であるが、これは人的資本ストックと労働投入量に対する分配率の合計に近い数字である。ただし、農林水産業以外の産業に関する個人企業所得の中にも労働の対価と言える部分があるので、実際にはこれらの分配率の合計は7割強であったと予想される。つまり、物的資本ストックへの分配は3割弱と予想される。

³¹ 老齢基礎年金は2001年9月現在で月額67,017円、個人住民税の課税最低限所得額は2001年8月現在において夫婦と子供2人世帯で年収325万円である。

表 11 適用結果（日本時系列）

説明変数	供給				需要			
	1969-88年度		1969-98年度		1969-88年度		1969-98年度	
	係数	t値	係数	t値	係数	t値	係数	t値
$\ln K - \ln W$	0.296	7.906	0.283	8.484	0.240	8.253	0.197	3.880
$\ln H - \ln W$	0.392	5.397	0.420	7.793	0.328	8.619	0.470	7.075
定数	0.012	1.408	0.013	1.371	0.008	1.462	-0.023	-1.855
Adjusted R ²	0.995		0.996		0.996		0.986	
D. W.	1.370		1.161		1.641		0.309	
α	0.30		0.28		0.24		0.20	
β	0.39		0.42		0.33		0.47	
$1-\alpha-\beta$	0.31		0.30		0.43		0.33	

供給について、技術進歩を人的資本ストックに取り込むことについては異論がある。一般的には、技術進歩として時間による指数関数を盛り込むことが多い。しかし、このような方法ではおおむね右肩上がりの指標の中に完全に右肩上がりの指標が加わるため、多重共線性が起き、安定的な計量実証結果は得られない。もし技術進歩を盛り込みたいのであれば、人的資本ストックが存在しないと仮定する $\beta = 0$ のモデルを適用することにより良好な結果を得ることができる。しかし、このことは、人的資本ストックで技術進歩を代替できるとする仮説を実証的に示したものと解釈することもできる。

需要について、物的資産はむしろ金融資産に限定するほうが妥当であるとの見方もある。手持ちの金銭が多ければ需要も強まるという考え方をする場合である。物的資産を金融資産（純固定資産デフレータで実質化したもの）に限定して同じモデルを当てはめると、分配率は物的資産（金融資産）が 2 割程度、人的資産が 4 割程度、人口要因が 4 割程度³²で比較的安定性が高い。ただし 1998 年度まで適用する場合には Durbin-Watson 比が 0.549 と小さい。

上記のような分解の仕方は、経済成長を独立な要因に区分できる反面、前節とは異なり将来推計が困難になる。というのは、適用結果を見るかぎり、少なくとも需要については 1990 年頃を境にこれらの分配率が変化している可能性があるからである。また、人口要因とは異なり資産要因の将来推計は困難である。ただし、分配率に安定性があれば人口要因の影響だけは試算することができる。仮に、人的資産の分配率が物的資産を金融資産に限定して 1998 年度まで適用した場合の 4 割程度で今後も推移するとすると、2025 年度に人口が前年よりも 0.5% 減少することにより経済成長は 0.2% 減少することになる。これを 1.1 の推計結果に当てはめて解釈すると、2025 年

³² 分配率は、物的資産、人的資産、人口要因の順に、1988 年度までなら 0.18、0.37、0.45、1998 年度までなら 0.24、0.35、0.41 となる。各係数の t 値は 1.1 以上である。

度に人口は 0.5%減少するが、15 歳以上人口比、就業率、平均就業時間、労働時間当り GDP に影響を及ぼすため、これらと併せて評価すると、全体で経済成長を 0.2%減少させることになる。

2 地域間比較

本節では、1.2 で適用したモデルに従って、先進国間で、また、日本における都道府県間で、経済力の格差を解明するために横断的な経済成長の構造分解を試みることとする。ただし、定常状態を仮定したモデルを使用し、資産要因については代理変数を採用する。

2.1 先進国間比較

当節では、先進国の間における経済力の格差を解明するために、1.2 同様 Cobb-Douglas 型の要因分解を試みることとする。

経済成長に関する研究の主要な目的は貧富の差の解明である。しかし、人口要因は経済成長に対して、貧しい地域ではマイナスに寄与し、豊かな地域ではプラスに寄与する傾向があるという点に配慮する必要がある。先進国は開発途上国に比べて数が少ないために、分析結果は開発途上国の状況に埋没してしまい、先進国での人口要因の影響とはまったく異なる結果が出ることも予想される。しかし、地球規模で先進国の占める GDP シェアは大きいので、先進国にデータを限定し、先進国特有の構造上の特徴を見出すことには意義があると考えられる。ただし、データを収集・比較できる範囲には限界があるので、ここでは OECD 加盟国に限定して行うこととする。

経済成長の国間横断的な要因分解については、Mankiw et al.[30]では、

$$Y = K^\alpha H^\beta (AL)^{1-\alpha-\beta} \quad 0 < \alpha < 1, 0 < \beta < 1, \alpha + \beta < 1, L = L(0)e^{nt}, A = A(0)e^{gt}$$

というモデルにおいて、対数を取り定常状態を仮定した

$$\ln Y - \ln L = \ln A(0) + gt + \frac{\alpha}{1-\alpha-\beta} \{\ln s_k - \ln(n+g+\delta)\} + \frac{\beta}{1-\alpha-\beta} \{\ln s_h - \ln(n+g+\delta)\}$$

を 1960 年から 1985 年までのデータに適用した事例を示している。ここで、 Y は実質 GDP、 L は World Bank による生産年齢人口、 s_k は投資率(GDP 当り)、 s_h は UNESCO による中学卒業者比率 (生産年齢人口当り)、 n は 1960 年から 1985 年までの生産年齢人口の年平均増加率であり、 s_k 及び s_h は推定期間中の平均値である。また、技術進歩率 g 、固定資本減耗率 δ はそれぞれ一定で、 $g + \delta = 0.05$ と仮定している。モデルは、非産油 98 か国、小国等を除く 75 か国については、いずれも $\alpha = 0.3$ 程度で適合するが、OECD 加盟 22 か国に限定すると、適合度が低くなるとしている。しかし、開発途上国については先進国と比較できる水準でデータを揃えることが難しい。もち

ろん、開発途上国であっても先進国並にデータを整理しているところもあるが、そのような国を集めても、その範囲の持つ意味を見出すことは困難である。

ここでは、対象国を OECD 加盟国に限定し、1990 年以降も含めたデータに対して、まず、供給面での国間格差を見るために、1.2 同様、技術進歩は人的ストックによって達成されるとした類似のモデル

$$Y = K^\alpha H_p^\beta (\gamma W)^{1-\alpha-\beta} \quad 0 < \alpha < 1, 0 < \beta < 1, \alpha + \beta < 1, W = e^{wt}$$

を想定する。

対数を取り定常状態を仮定すると、

$$\ln Y - \ln W = \ln \gamma + \frac{\alpha}{1-\alpha-\beta} \{\ln s_k - \ln(w+\delta)\} + \frac{\beta}{1-\alpha-\beta} \{\ln s_{hp} - \ln(w+\delta)\}$$

となる。ただし、 γ は定数とし、変数はできるだけ 1.2 で取り上げたモデルの変数の概念と一致させるために

Y : GDP (68SNA に基づく 1990 年基準ドル換算値【OECD】)、

s_k : 総固定資本形成 (68SNA に基づく 1990 年基準ドル換算値) の GDP 比 (%）、

s_{hp} : 専門的・技術的職業従事者数【ILO】の人口【OECD】比 (%）、

w : 労働時間 (就業者数【OECD】×平均就業時間) 増加率³³ (%)

とする。データの年次は、収集上の制約などにより一部古いものを採用せざるを得ないものの、1997 年を基本とし、推計開始年は韓国についてデータの遡及が可能な 1970 年とする。 s_k は 1970 年からの平均値とし、 δ は日本におけるその間の固定資本減耗率³⁴ (%) の平均値とする。

なお、上記のモデルは開発途上国を含める場合には適用が困難になるものとする。理由の 1 つは、固定資本減耗率を一定と仮定していることであり、もう 1 つは開発途上国を含めて同一の基準でデータを収集することが困難であるからである。また、対象を先進国に限定したからには、人的資本の代理変数についても Mankiw 等のモデルで採用された中学卒業者比率にこだわる必要はないことから、1.2 を踏まえて一般的にはもう少し高度な能力が必要とされる専門的・技術的職業従事者の比率とした。

³³ 労働時間増加率は就業者数増加率に年間平均就業時間増加率を乗じて計算した。年間平均就業時間は、基本的には OECD からデータを収集したが、そこから入手できない国については ILO、更に入手が不可能な国については World Bank からデータを収集した。OECD からデータを収集した国については増加率を 1996 年から 1999 年までの年平均増加率とした。その他の国については増加率を 1990 年から 1994 年平均値に対する 1995 年から 1999 年平均値の年平均増加率とした。就業時間は、国により対象となる産業や年齢が異なるので、結果を幅広に見る必要がある。

³⁴ 純固定資産 (名目値) と固定資本減耗額 (名目値) の和を分母とする固定資本減耗額 (名目値) の割合である。

2002年1月現在においてOECDを構成する30か国のうち、上記のデータをすべて揃えることができた24か国³⁵について上記モデルを適用した。総固定資本形成からは住宅に係る分を除外したいところだが、24か国すべてについてこのような指標を揃えることは困難であったため、やむを得ず含めることとした。

同様に、需要についても、

$$Y = K^\alpha H_l^\beta (\gamma N)^{1-\alpha-\beta} \quad 0 < \alpha < 1, 0 < \beta < 1, \alpha + \beta < 1, N = e^{nt}$$

と、対数を取り定常状態を仮定した

$$\ln Y - \ln N = \ln \gamma + \frac{\alpha}{1-\alpha-\beta} \{\ln s_k - \ln(n+\delta)\} + \frac{\beta}{1-\alpha-\beta} \{\ln s_{hl} - \ln(n+\delta)\}$$

を想定することとする。ただし、 s_{hl} は25～64歳大学卒業者³⁶比率(%)【OECD】、 n は1970年から1997年までの人口の年平均増加率とする。上記モデルはOECDの構成国のうち上記のデータをすべて揃えることができる26か国³⁷に適用した。なお、物的資産としては金融資産や土地資産を含めた概念に相当する変数を採用したいところであるが、比較に耐えうる状態で揃えることが困難なため、代理変数として総固定資本形成のGDP比を採用した。

モデルの適用結果を表12に示す。これによると、供給については物的資本分配率が3割程度、人的資本分配率が2割程度、労働投入量分配率が5割程度となり、経済成長の先進国間格差の主要な要因は労働投入量である。つまり、先進国の間では各自の労働時間の総量が供給面での格差を生む大きな要因となっている。需要については物的資産分配率が5割程度、人的資産分配率が2割程度、人口分配率が3割程度となり、国間格差の主要な要因は物的資産である。

³⁵ 除外国は1990年以前のデータを入手できないチェコ、ハンガリー、ポーランド及びスロバキアに加え、専門的・技術的職業従事者数を把握できないフランス及びルクセンブルグである。ドイツは1990年以前のデータに旧東ドイツを含めた。日本は1972年以前のデータから沖縄県を除外し、人口増加率の平均値は1973年を除外して計算した。

³⁶ ISCED (International Standard Classification of Education) において5A及び6に該当する者とした。

³⁷ 供給について対象とした24か国にフランス及びルクセンブルグを加えた。

表 12 適用結果 (OECD 構成国) 1990 年基準ドル換算値

説明変数	供給		需要	
	係数	t値	係数	t値
$\ln s_k - \ln(w + \delta)$	0.518	1.411	1.424	2.288
$\ln s_h - \ln(w + \delta)$	0.458	4.234	0.697	2.464
$\ln \gamma$	9.488	21.307	7.709	10.566
Adjusted R ²	0.591		0.294	
α	0.26		0.46	
β	0.23		0.22	
$1 - \alpha - \beta$	0.51		0.32	

表 13 はこれを購買力平価で見たものである。 β 及び $1 - \alpha - \beta$ を表 12 と比較すると、物価水準で調整をすることにより、供給面でも、需要面でも人的資産の分配率が減少し、人口要因の分配率が増加している。ここでの調整は、同じ物やサービスであれば同じとみなすことなので、これにより、誰が評価しても同じである人間の数や労働時間の寄与が強まり、一見しただけでは区別できない個々人の能力や経験の寄与が弱まるという結果は、納得できるものと考えられる。

表 13 適用結果 (OECD 構成国) 購買力平価

説明変数	供給		需要	
	係数	t値	係数	t値
$\ln s_k - \ln(w + \delta)$	0.391	1.534	1.445	4.331
$\ln s_h - \ln(w + \delta)$	0.157	2.244	0.329	2.350
$\ln \gamma$	8.066	25.727	6.394	16.128
Adjusted R ²	0.366		0.491	
α	0.25		0.52	
β	0.10		0.12	
$1 - \alpha - \beta$	0.65		0.36	

2.2 都道府県間比較

次に、都道府県の間における経済力の格差を解明するために、2.1 の考え方に沿って Cobb-Douglas 型の要因分解を試みることにする。

このため、先進国間で構造分解したときと同様に、まず、都道府県ごとの供給としての経済水準を

$$Y_p = K_p^\alpha H_p^\beta (\gamma W)^{1-\alpha-\beta} \quad 0 < \alpha < 1, 0 < \beta < 1, \alpha + \beta < 1, W = e^{wt}$$

とし、対数を取り定常状態を仮定した

$$\ln Y_p - \ln W = \ln \gamma + \frac{\alpha}{1-\alpha-\beta} \{\ln s_{kp} - \ln(w + \delta)\} + \frac{\beta}{1-\alpha-\beta} \{\ln s_{hp} - \ln(w + \delta)\}$$

で推定する。

具体的には

Y_p ：県内総支出（68SNAに基づく県民経済計算実質値）、

s_{kp} ：民間及び公共設備投資³⁸（68SNAに基づく県民経済計算名目値）の県内総支出（固定資本形成＋最終消費支出）（名目値）比³⁹（%）、

s_{hp} ：専門的・技術的職業従事者数の昼間人口比率⁴⁰（%）、

W ：労働時間⁴¹

とする。ただし、 s_{kp} は推定期間中の平均値とし、 w は推定期間中の開始年度に対する年平均増加率、 δ は全国における固定資本減耗率の推定期間中の平均値（%）とする。推定開始年度は同一水準で遡りが可能な1975年度とする。

経済成長を供給側から見る場合には、居住地よりも従業地で区分することが重要なので、総生産と設備投資は県内ベースで捉え、専門的・技術的職業従事者数や労働時間も従業地ベースで把握している。

モデルの適用に当たっては、特に、1990年前後で労働時間の増減に変化が起きている点に注意が必要である。すなわち、図10のとおり、1975年から1990年に労働時間の増加率が高い都道府県ほど1990年から2000年の増加率が低くなる傾向がある。このため、労働時間当り県内総生産も、バブル崩壊以前は、労働時間増加率が低い都道府県ほど高くなる傾向があったが、それ以降は労働時間増加率が高い都道府県ほど高くなる傾向がある⁴²。

³⁸ 1998年度における茨城県の公的設備投資額は公表されていないが、公的総固定資本形成に対する割合が1997年度と同じであると仮定した。

³⁹ 本来実質値の比で捉えるべきところであるが、1980年前後において実質値が把握されていない都道府県がいくつか存在したため、やむを得ず名目値の比で捉えることとした。 s_{kl} についても同様である。

⁴⁰ 国勢調査年ごとに計算し、1975年から推計年の次の国勢調査年までの平均値とした。

⁴¹ 都道府県ごとに直近国勢調査年における平均実労働時間（賃金構造基本統計調査）×12（月）×従業地就業者数（国勢調査）で計算した。ただし、全国計が推定年の延週間就業時間（労働力調査）×52（週）に一致するように等倍した。

⁴² 労働時間増加率と時間当り実質GDPの回帰分析による傾きは、1975年から1990年労働時間増加率×時間（1990年）当り実質GDP（1988年）ではマイナス1293（ t 値はマイナス3.414）、1990年から2000年労働時間増加率×時間（2000年）当り実質GDP（1998年）では287（ t 値は2.187）である。

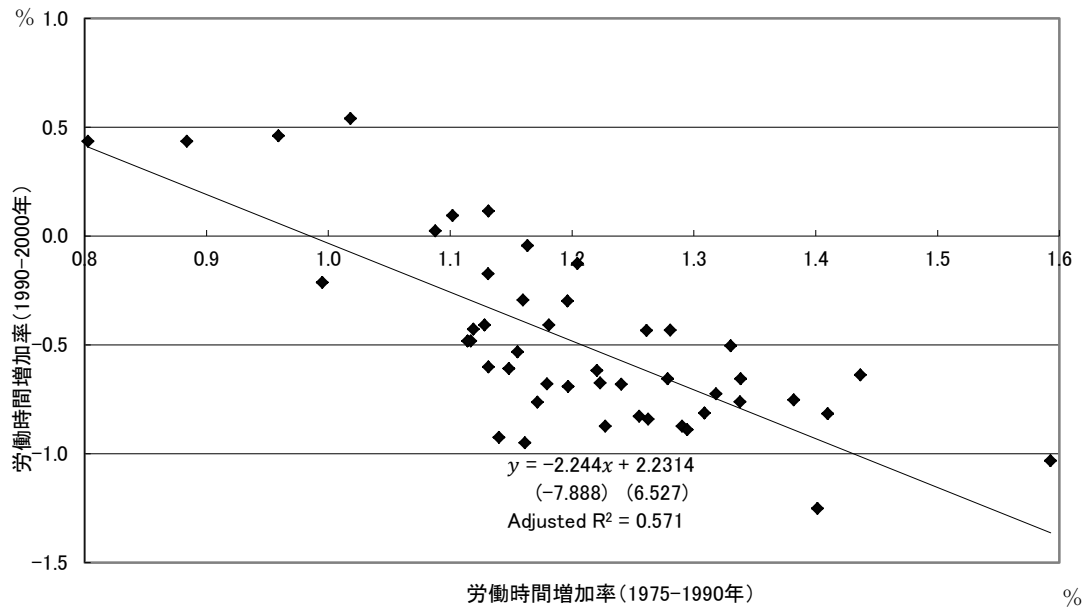


図 10 都道府県別労働時間の変化 年平均変化率

そこで、データを 1990 年前までにとどめる場合と、それ以降も含める場合にモデルを適用し、両者を比較した。すなわち、推定期間を 1975 年から 1988 年と 1975 年から 1998 年の 2 通りについてモデルを適用し、比較した。

同様に、各都道府県に対して需要側から見た経済水準を

$$Y_t = K_t^\alpha H_t^\beta (\gamma N)^{1-\alpha-\beta} \quad 0 < \alpha < 1, 0 < \beta < 1, \alpha + \beta < 1, N = e^{nt}$$

とし、対数を取り定常状態を仮定した

$$\ln Y_t - \ln N = \ln \gamma + \frac{\alpha}{1-\alpha-\beta} \{\ln s_{kl} - \ln(n+\delta)\} + \frac{\beta}{1-\alpha-\beta} \{\ln s_{hl} - \ln(n+\delta)\}$$

で推定する。

具体的には

Y_t : 県民総支出 (68SNA に基づく県民経済計算実質値)、

s_{kl} : 固定資本形成 (68SNA に基づく県民経済計算名目値) の県民総支出 (固定資本形成 + 最終消費支出 + 県外からの要素所得) (名目値) 比 (%)、

s_{hl} : 大学卒業者 (25 歳以上の在学者を含む。) の夜間人口比率⁴³ (%)、

とし、 s_{kl} は推定期間中の平均値、 n は人口の 1975 年からの年平均増加率 (%) とする。

⁴³ 西暦 0 年ごとに計算し、1980 年から推計年の次の国勢調査年までの平均値とした。

ここで、 Y_i を県内総支出ではなく県民総支出としたのは、昼間に就業や就学のために他県から来る通勤・通学者ではなく、夜間に在住する住民の観点から需要を測定しようとしているからである。特に、大都市近郊では県内ベースと県民ベースで指標値が大きく異なるので注意を要する。

モデルの適用結果は表 14 のとおりである。これによると、供給側、需要側とも物的資産の分配率は2割程度と小さい。大きな理由にはならないが、供給側では1.2の表 11と比較して推定開始年度が1975年度と6年遅いことが一因として考えられる⁴⁴。なお、供給側で物的資本ストックに相当する代理変数を民間設備投資に限定しても同様の結果⁴⁵が得られることから、経済水準の小さな都道府県に手厚く配分されている公的設備投資による分配率への影響は大きくないと考えられる。一方、供給における労働投入量、需要における人口の分配率は5割前後～6割弱と高く、都道府県間格差についてもマンパワーが大きな要因であることがうかがえる。

また、1988年度から1998年度の10年間の変化に着目すると、先程示したように、供給面で労働時間は経済水準に対して1990年頃を境にマイナスからプラスに影響するようになったが、その影響で分配率は増加している。これに対して、人的資本ストックの分配率は低下している。1990年以降の急速な情報化を考慮すれば、能力の地域差による地域経済に対する影響も小さくなっていると考えられるので、妥当な結果であると考えられる。他方、需要については1988年度までと1998年度までで分配率に大きな違いは見られない。

表 14 適用結果（都道府県）

説明変数	供給				需要			
	1988年度		1998年度		1988年度		1998年度	
	係数	t値	係数	t値	係数	t値	係数	t値
$\ln s_k - \ln(w + \delta)$	0.448	3.954	0.423	4.495	0.308	3.196	0.317	3.251
$\ln s_h - \ln(w + \delta)$	0.709	4.450	0.350	2.654	0.498	9.205	0.406	7.360
$\ln y$	7.765	59.595	7.714	74.464	14.481	98.529	14.523	102.120
Adjusted R ²	0.438		0.398		0.643		0.544	
α	0.21		0.24		0.17		0.18	
β	0.33		0.20		0.28		0.24	
$1 - \alpha - \beta$	0.46		0.56		0.55		0.58	

都道府県間比較においても先進国間の比較と同様、物価水準で調整をすることとした。表 15 は実質 GDP を前年に実施された全国物価統計調査の地域差指数で割って同

⁴⁴ 国民所得に対する労働分配率（雇用者所得と農林水産業に関する個人企業所得の割合）は1975-88年度平均で69.7%、1975年度から1998年度平均で70.6%であり、1969年度から1998年度平均の69.0%より少し高い。

⁴⁵ 物的資本の分配率は1988年度において0.20、1998年度において0.22である。各係数のt値は2.5以上であり、適合度は高い。

一モデルを適用した結果である。これによると、物価調整をしない場合（表 14）と比較して、供給面でも、需要面でも人的資産の分配率が小さく、人口要因の分配率が大きくなっている。一方、物的資産の分配率にはほとんど差がない。

表 15 適用結果（都道府県） 都道府県間の物価水準を調整

説明変数	供給				需要			
	1988年度		1998年度		1988年度		1998年度	
	係数	t値	係数	t値	係数	t値	係数	t値
$\ln S_k - \ln(w+\delta)$	0.435	4.261	0.413	5.062	0.283	2.609	0.279	3.137
$\ln S_h - \ln(w+\delta)$	0.546	3.801	0.163	1.425	0.419	7.890	0.306	5.691
$\ln y$	7.701	65.536	7.682	85.431	14.521	100.609	14.599	104.005
Adjusted R ²	0.416		0.388		0.567		0.412	
α	0.22		0.26		0.17		0.18	
β	0.28		0.10		0.25		0.19	
$1-\alpha-\beta$	0.50		0.63		0.59		0.63	

3 定量的評価の留意点及び課題

以上のように、経済水準における人口要因の分配率は比較的高く、人口や労働投入量は経済成長の水準や経済格差の要因に大きな影響を及ぼしている。

ただし残念なことに、統計上の制約により、これまでの定量的な評価についてはかなり幅広に見る必要がある。また、必要な統計を類似の基準で獲得できずに分析を断念したこともあった。例えば、先進国の経済格差についても、1990年頃を境に、前後で比較を試みようとしたが、教育水準や労働時間などの統計を同一ベースで得ることが困難な国がいくつかあり、統計を入手できる国だけでは数が足りずに回帰分析には耐えられないという問題があった。さらに、数年後に類似の分析を試みて結果を比較したいという希望があっても、SNAのベースが1968年から1993年に変わり、さらには2008年に変わり、2008SNAでは今回分析した時点まで遡れないという問題が発生する。このように、経済理論の実証可能性は、データの存否や性格に依存するため、実証分析においては分析のタイミングと限界の見極めが重要になる。

需要面の解釈は課題である。供給面においては労働分配率が人的資本ストックと労働投入量の分配率の合計であり、また、生活保障の観点から労働分配率のみを計測することも可能である。しかし、需要面においては、家計と企業からの支出が投資と消費に向けて混在している。物的資産は投資に影響し、人口要因は消費に影響するという割り切り方もあるが、それでも人的資産は投資と消費の両方に影響する。

また、供給側と需要側の需給ギャップの考え方と試算も課題である。

第2章 所得が出生に及ぼす影響

人口要因と経済要因の間には各種の関連がある。近年、日本を始めとするいくつかの先進国では、出生率の低下が社会問題となっており、それにはいくつかの経済的要因も影響を及ぼしているという見方がある。

代表的なモデルの1つに **Butz and Ward Model** がある。同モデルは、夫婦について夫の所得の増加は子どもに対する需要を増加させる所得効果を生むが、妻の賃金の上昇は子どもに対する需要を減少させる代替効果を生み、これが夫の所得による所得効果を上回るならば、その結果出生率は低下するとするものである。

このような現象は2つの視点で捉えることができる。1つは、所得と出産行動について平均的な状況を長期に概観したときにそのような構造が当てはまるのかという視点である。もう1つは、特定の時点において夫婦間にそのような構造があるのかという視点、すなわち、夫婦間に夫の所得が高いほど、妻の賃金が低いほど出生率が高くなる構造があるのかなどの視点である。

本章では、まず、前者の視点を踏まえて時系列データを想定して構築された同モデルの理論を簡単に振り返った上で、公的統計を利用し、日本について最近の時系列データによる検証を試みる。**Butz and Ward**[31]はこれを1948年から1975年の米国のデータに適用して理論どおりの結果を得ている。英国については **John Ermisch**[32]が1951年から1975年のデータを用いて係数が条件を満たすことを確認しているが、誤差項には系列相関が、説明変数には多重共線性が認められたとしている。我が国についても、**Ogawa and Mason**[33]が1966年から1984年のデータに適用して良好な結果を示し、**大沢真知子**[34]が1960年から1984年のデータに適用して女性の市場賃金の上昇による就業率の増大が出生を抑制する方向に働いていることを指摘する一方で、**大淵寛**[35]、**Hisakazu Kato**[36]などは係数の符号条件が満たされないとの結果を報告している。

本章の新規性は、むしろ後者の視点から個々の夫婦について、夫婦及び子どもの年齢、夫婦双方の所得及び就業状態が調査されているという **JGSS** の特色を生かし、同モデルをクロスセクションで再構築し、適用可能性を検討した点である。この際、各データを適正に比較するため、年齢と就業年数を用いてあらかじめ標準化しておくこととする。

このような分析は、最初に **清水誠**[37]で **JGSS-2000** のデータを用いて行ったが、その後、所得水準と出生率の低下が同時に進行した中で、**清水誠**[38]で **JGSS-2002** のデータで適合状況を確認した。本章は **清水誠**[38]の内容を示す。

1 時系列データの適用

Butz and Ward Model は、米国において景気が回復しても出生力が回復しない現象を説明するための時系列モデルであり、具体的には次の形をしている。

有配偶女性の就業率を K 、就業していない有配偶女性の出生率を B_1 、就業している有配偶女性の出生率を B_2 とすれば、有配偶女性の出生率 B は

$$B = (1 - K)B_1 + KB_2$$

と表される。ここで、 B_1 は夫の所得 Y_m と他の要因 X で決まり、 B_2 は両者に加えて妻の賃金 W_f によって決まるものとした。実用上はこれを全微分して得られる

$$\ln B = \gamma_0 + \gamma_1 K \ln Y_m + \gamma_2 \ln Y_m + \gamma_3 \ln W_f$$

が提案された。ただし、所得効果がプラス、代替効果はマイナスであることから、係数には

$$\gamma_1 + \gamma_2 > 0, \quad \gamma_2 > 0, \quad \gamma_3 < 0$$

という制約条件がある。

1.1 1971年から2000年における Butz and Ward Model の適合状況

同モデルについては1948年から1975年の米国のデータに適用した場合を始め様々な実証研究が行われ、適合性を示す結果と非適合性を示す結果が存在している。同モデルは時系列データの適用を前提としているが、もともと所得が上昇している局面を想定して構築されたものであることから、近年の日本のように平均所得が長期にわたり低下していた状況下では適合するとは限らない。清水誠[37]は、出生率が低下、夫の所得が増加する傾向にあった1971年から2000年の時系列データに同モデルは適合しないとされた。

1.2 1975年から2004年における Butz and Ward Model の適合状況

2000年以降も、しばらく同モデルの適合条件に沿わない傾向、すなわち、所得水準と出生率の低下が同時に進行するという傾向が続いた。そこで、これらのデータを含めた1975-2004年の日本の時系列データに同モデルを適用し、適合状況の変化をみてみることにする。

最初に、検証に用いるデータについて説明する。まず、夫の所得は、家計調査の勤労者世帯の世帯主の収入を年額換算したものとする。妻の賃金は、賃金構造基本統計調査から女性労働者に関するきまって支給する現金給与額（年額換算）に年間賞与其他特別給与額を加え、所定内実労働時間数に超過実労働時間数を加えたもの（年換算）で割ることにより計算した。実際には同じ女性でも、有配偶者とそれ以外とで平均賃金は異なるが、有配偶女性が有配偶者のみに該当する金額を見て出生行動を変え

ることではないので、市場単価としては一括りで計算しても問題はない。なお、夫の所得と妻の賃金はCPIにより実質化した。また、Butz and Ward Modelの適合度は、単位の取り方により変化するため、ここでは金銭の単位をすべて万円に統一して計算した。一方、妻の就業率は労働力調査の有配偶女性の就業率とした。次に、出生率は、日本において出生はほとんど有配偶者により行われるという事情にかんがみ、人口動態調査による出生児数を労働力調査による有配偶女性数で割ることにより、いわゆる普通出生率として計算した。

必要な変数には対数をとって、Butz and Ward Modelを適用した結果は表16のとおりである。これによると、Adjusted R²は高いものの、Durbin-Watson比が低く、 $\gamma_1 + \gamma_2 > 0$ という係数の符号条件は満たされず、夫の所得に係る係数のt値も小さい。

清水誠[37]では夫の所得の対数値と出生率の対数値の間には負の相関が見られたが、1975-2004のデータについては両者の間に正の相関が見られた。これは、出生率の低下傾向には変化がないものの、夫の所得については、データの範囲が新しくなるにつれて減少傾向が強まり、したがって、出生率との相関が強まるからである。

表16 時系列データにButz and Ward Modelを適用した結果

説明変数	ln B	
	偏回帰係数	t値
$K \ln Y_m$	-1.447	-2.599 **
$\ln Y_m$	1.180	1.043
$K \ln W_f$	-2.432	-3.642 ***
定数項	-8.452	-1.404
n	30	
Adjusted R ²	0.932	
F	132.855	***
D. W.	0.452	

注 *、**、***はそれぞれ10%、5%、1%基準で有意であることを示す。(他の表についても同じ)

2 JGSS-2002における夫婦データを用いた検証

このように、近年の日本の平均的状況を時系列で観測したときにButz and Ward Modelが適合するとは言えないが、夫の所得が高いほど出生率が高い、妻の賃金が高いほど出生率が低いという構造は現在の日本においても存在している可能性があり、それを実証するには同一時点における個々の夫婦の間での比較が必要である。つまり、同モデルは本来時系列データの適用が前提となっているが、夫の所得が高いほど出生力が強いのか、また、妻の市場賃金が高いほど出生力が弱いのかという問題は、ある時期、ある集団内の個々の夫婦間の相対的關係においても関心の対象となり得る。

2.1 検証の方法

しかし、一般に、ある時点において、同一集団内には若年者と高齢者が共存し、出生期に差し掛かったばかりの者とそれを終えた者がいる。また、所得についても就業期間の長い人ほど高いことが一般的であり、就業の有無についてもその時期にたまたま就業していることもあるので、これらを同じモデルにそのまま当てはめることは無意味である。時系列データに対して **Butz and Ward Model** を適用する際には、年齢や就業年数などの所得に影響を及ぼす要因を制御して比較するための工夫がなされるが、本研究のように同モデルを同一時点で空間的に適用する際にも、これらの要因を取り除く必要がある。つまり、個々のデータを時間ごとにまとめる場合には空間的格差をコントロールする必要があるのに対し、空間ごとにまとめる場合には時間的格差をコントロールする必要がある。そこで、調査時を断片的に捉えるのではなく、各個人（夫婦）の過去と将来の状況を予測し、生涯レベルでの比較を試みることにする。

まず、**Butz and Ward Model** を次のように再定義することとする。すなわち、有配偶女性が、一生の間に就業している期間の割合を K とし、就業していない間の出生率を B_1 、就業している間の出生率を B_2 とすれば、この女性の両期間を合わせた出生率 B は

$$B = (1 - K)B_1 + KB_2$$

と定義される。ここで、 B_1 は夫の所得に依存し、 K と B_2 は夫の所得と妻の賃金に依存するという仮定を設定すれば、**Butz and Ward Model** と同じ型になる。もちろん、長期的には夫の所得も妻の賃金も出生率も変動するが、これらは全体から推定される所得関数と出生関数に沿って動くこととする。また、妻の市場賃金は、学歴、産業、職業などにより異なり、また、就業年数などに応じて変化するものとし、妻が出生の決断をする際には、このような変動要因を加味して総合的に行うとの前提に立つこととする。

2.2 使用するデータセット

JGSS-2002 では、本人のみならず、配偶者がいる場合にはその年齢、学歴、就業状況等についても調査をしているので、ケースを夫婦に限定し、夫婦両面のデータをセットで利用することとする。利用に当たっては本人と配偶者の変数から性別により夫と妻の変数に組み替えた。ただし、配偶者については変数の種類が限定されているので、所得関数や出生関数の説明変数として組み込むことができる変数には制約があった。また、就業経歴については、現在の主な職に関する問が中心であり、前職や副業がある場合に、そこから得られる所得については情報が限られている。そこで、分析の前に、実態を把握できない状態が長期間発生しているケースをあらかじめ除外する

などの作業が必要であった⁴⁶。ただし、このような作業をしても、JGSS-2002 のデータだけでは過去の職歴と副業による影響を十分に調整できないことも確かである。したがって、本節の検証結果はかなり幅広に捉える必要がある。また、選択されたデータは、生涯推定用に構築したものであることから、例えば、夫が最初の子供の出生以降に転職をしていない夫婦だけになっているなどの偏りが存在するため、現状を把握するために平均やばらつきを計算するためのものではないことに注意が必要である。

変数について施した措置の概略は次のとおりである。まず、年間収入と企業の従業者数は階級の中央値に変換して数量化し⁴⁷、就業年数とともに無職（先週仕事をしていない人）のデータを0とした。また、学歴については小学・中学、高校・旧中、短大・高専、大学・大学院の4区分に、就労地位については経営者・役員、常用雇用、パート・アルバイト、派遣社員、自営業主、家族従業者、内職の7区分に、産業分類と職業分類は日本標準分類の大分類に近づけるようにコードを大括弧にした上で、分類別に該当者を1、非該当者を0としてダミー変数を作成した。地域（6区分）、市郡（3区分）についても同様のダミー変数を作成した。また、子ども数と子どもの出生年を含めこれらの変数に1つでも「わからない」としたケースと回答しないケースはあらかじめ対象夫婦から除外した。

こうして用意されたデータセットの該当ケースは180組（夫婦）になった。

2.3 所得・出生関数の作成方法

夫、妻の所得関数は、年間収入を被説明変数としてそれぞれ影響力があると考えられる変数を説明変数とする重回帰分析を適用することにより作成した。変数の選択にはステップワイズ法を用いた。選択前に投入した説明変数は、夫、妻それぞれについて、年齢、就業年数、企業の従業者数、最終学歴（ダミー）、就労地位（ダミー）、職業（ダミー）、産業（ダミー）、地域（ダミー）、市郡（ダミー）である。このうち、地

⁴⁶ 具体的には、次のとおりである。

最初の子どもを出産する時点が夫の現職の就業開始時期よりも早い25のケースを除外した。これは、子どもの出産時期に、夫は無職ではなく現職とは異なる職に就いて所得を得ていた可能性が高いので、それについての情報がなければ分析に支障が生じるからである。妻についても、現職の就業開始時期より早く出産している場合には同様の問題があるが、夫よりは無職であった可能性が高く、また、仮に就業していても所得は平均的には低いと想定されることから、無職と仮定しても生涯レベルの推定をすれば大きな問題は生じないとの判断をした。

さらに、妻の年齢を40歳未満かつ夫の年齢を50歳未満に限定した。これも、年齢が高くなれば現職の前に仕事をしている可能性が高いという理由によるが、併せて、できるだけ最近の出生事情を反映させること、世代間格差を避けることなどにも配慮した。

⁴⁷ 年間収入について、2300万円以上は実額のままである。企業の従業員数について、1万人以上は1万人とし、官公庁は夫・妻別に平均値を当てた。

域（ダミー）及び市郡（ダミー）は夫・妻ともに共通である。また、回答に伴う年齢や地域等の分布の偏りを是正するため、分析には比推定乗率を付与した。実際に選択された変数とその係数及び関数の適合度に関する情報は表 17 及び表 18 のとおりである。

特に、夫については、年齢が高く、就業年数が長く、学歴や役職の高い者ほど収入が高いという常識に沿った推定式が選択された。

表 17 夫年間収入の重回帰分析の結果

説明変数	偏回帰係数	t値
年齢	7.556	2.573**
就業年数	8.341	2.692***
企業規模	0.014	3.752***
学歴—大学・大学院 ダミー	121.097	4.630***
就労地位—経営者・役員 ダミー	221.561	3.174***
就労地位—自営業主 ダミー	233.211	4.077***
職業—管理的 ダミー	232.388	2.885***
定数項	80.136	1.011
n	180	
Adjusted R ²	0.441	
F	21.184***	

表 18 妻年間収入の重回帰分析の結果

説明変数	偏回帰係数	t値
就業年数	14.939	14.757 ***
企業規模	0.018	6.558 ***
就労地位—自営業主 ダミー	167.981	5.014 ***
就労地位—常用雇用者 ダミー	102.005	9.455 ***
産業—卸売業、小売業、飲食店 ダミー	77.811	6.012 ***
産業—サービス業 ダミー	76.883	5.515 ***
産業—公務 ダミー	177.629	5.960 ***
職業—専門的・技術的 ダミー	84.921	5.493 ***
職業—販売 ダミー	-64.390	-2.904 ***
職業—サービス職業 ダミー	-75.818	-3.849 ***
北海道・東北 ダミー	-46.690	-3.066 ***
13 大市 ダミー	18.778	2.646 ***
定数項	-4.767	-1.113
n	180	
Adjusted R ²	0.923	
F	180.505 ***	

一方、出生関数の作成にも上記と同じ変数を用いた。JGSS-2002 には、他にも例えば、住居の種類や広さ、同居人数など、出生に影響を及ぼすと考えられる変数が含まれているが、これらは子どもが増えれば変化するものなので、過去及び将来への延長

を想定している当関数の説明変数の候補には加えなかった。また、意識関連変数の中にも、例えば、理想の子ども数、結婚後の子どもの必要性など出生力に影響を及ぼす変数はあるが、夫婦の片側だけしか調査されていないので候補から除外した。実際に選択された変数とその係数及び関数の適合度に関する情報を表 19 に示す。

このような関数の作成は、平均的な傾向を抽出することにより、個々のデータが持つ特殊要因を取り除くという意味でも有意義であると考ええる。

表 19 子ども数の重回帰分析の結果

説明変数	偏回帰係数	t値
夫の就業年数	0.052	5.060 ***
妻の就労地位—常用雇用 ダミー	-0.642	-4.199 ***
夫の産業—農林水産業 ダミー	1.194	2.446 **
妻の産業—建設業 ダミー	0.876	2.534 **
北海道・東北 ダミー	-0.660	-2.486 **
関東 ダミー	-0.246	-2.076 **
定数項	0.966	7.178 ***
<i>n</i>	180	
Adjusted R ²	0.272	
F	12.144	***

2.4 生涯レベルの推定方法

ある時点 t の夫の所得を $Y(t)$ とすると、 $Y(t)$ のうち時間に依存する要素は年齢と就業期間だけであり、調査時点の年齢を t_0 、就業期間を w_0 、時間に依存しない部分を Z とすると、

$$Y(t) = \alpha t + \beta(t - t_0 + w_0) + Z$$

と表される。60歳まで働くとする、生涯レベルの所得は、

$$\int_{t_0 - w_0}^{60} \{\alpha t + \beta(t - t_0 + w_0) + Z\} dt$$

$$= (60 - t_0 + w_0) \left\{ \frac{1}{2} \alpha (60 + t_0 - w_0) + \frac{1}{2} \beta (60 - t_0 + w_0) + Z \right\}$$

となる。実用上は、所得は連続変数ではなく年ごとに変化するものとする。また、 α 及び β には重回帰分析の説明変数の係数を、 Z には推定値から逆算した値を適用することにより生涯レベルの推定値が求まる。妻の所得についても同様である。ただし、このような同一時点における各年齢の状況をすべて経験する生涯は実際には有り得ない。これらの推定は、個々人の生涯を正確に予想しようとするよりは、データ間の比較を行う以前の標準化のための措置であり、あくまで現在の職の状態が当該職の開始

時期から 60 歳まで維持されたときに生涯レベルでどの程度の所得が得られるかを求めたものに過ぎない⁴⁸。

妻の賃金については、上記で得られた生涯レベルの推定値を現職の開始時期から 60 歳までの期間に年平均 50 週×40 時間働くものとして単位時間に換算した。これも、実際の平均賃金を求めようとするのではなく、生涯レベルの所得を時間単位に換算したに過ぎない。このことは、期間が長いとその間に賃金は変動するので、子育ての代替としての捉え方は現実的には困難であり、むしろ就業期間中の子どもに対する需要は現職から推定される生涯レベルの所得を踏まえて決定されるとの見方に立つことを意味している。なお、金銭の単位はすべて万円に統一して計算した。

生涯レベルの子ども数も基本的には同じ方針で推定した。ただし、子ども数は、フローである所得とは異なり、ストックの変数なので、時間に関連のある部分を延長するだけで計算できる。説明変数の中で時間の流れに応じて変化するものは夫の就業年数だけなので、その係数に妻の現年齢から 50 歳になるまでの年数を掛けて推計値に加算した。

妻の就業率は 15 歳から 60 歳までの 45 年間に對する就業年数の割合とした。

2.5 結果の評価

2.4 で計算した生涯レベルの夫の所得、妻の賃金、就業率及び出生数を用いて Butz and Ward Model に適用した結果は表 20 のとおりである。これによると、係数は同モデルの条件を満たしている。これは、出生に対して所得は夫婦で同一方向に作用するのではなく、夫についてはプラスに、妻についてはマイナスに作用するという旧態の構造が依然として日本に根付いていることを物語っている。しかも、今回の結果は清水誠[37]で検証した結果と比べて t 値の絶対値、Adjusted R^2 などが高いことから、その構造はますます強まっていると考えられる。

時系列分析と夫婦間分析の結果を合わせると、出生に対する所得の影響が夫についてはプラスに、妻についてはマイナスに作用するという構造は、近年平均的には弱まっているものの、夫婦間で見た場合に依然として実在していることとなる。

特に、妻の就業の程度（就業率×賃金の対数値）と出生率（対数値）との間に負の関係があるという点については、清水誠[37]で検証したときと同じであり、さらに、

⁴⁸ この推定によると、現在無職の者は、現在を就業開始時期として 60 歳まで働くこととなるが、当推定ではそれにより得られる所得は僅かであるとの結果になる。Butz and Ward Model を適用するには、対数をとる必要があるが、真数として 0 以下を避けるという意味でも都合がよい推定となっている。一方、結婚以前に就業していた期間の所得は、すべて貯蓄されて結婚後の出生に影響を及ぼすとの想定になっている。

この関係は時系列分析においても多少時期を前後させても現れることから、妻の所得は出生にマイナスに影響するという構造は相当根が深いと考えられる。

表 20 Butz and Ward Model の検証

説明変数	ln B	
	偏回帰係数	t値
$K \ln Y_m$	-0.050	-4.971 ***
$\ln Y_m$	0.530	8.154 ***
$K \ln W_f$	-0.496	-6.516 ***
定数項	-4.886	-7.425 ***
n	180	
Adjusted R ²	0.415	
F	43.335	***

なお、本研究の推定方法は、例えば、時間に応じて変化する要因が直線的に変化すると仮定する、就業終了年齢を 60 歳に揃えるなど、便宜的にかなり単純化したものとなっている。また、JGSS では配偶者の情報に制約があることから、推定においても前職からの所得を無視するなど概念上の制約を図っている。したがって、今回の推定結果はある程度の幅を持って見るべきであり、今後も引き続きこのような検討を蓄積することにより、所得と出産行動との関係を少しずつでも明確にしていくことが重要であると考えられる。

第 3 章 科学技術研究調査による企業の生産性の要因分析

企業の生産性がどのような要因に左右されるのかという課題は、昔から各国様々なところで問われている重要なテーマである。企業の生産性の源泉は大きく分けて資本、労働力、技術進歩の 3 つの要因に区分される。第Ⅲ部の最後の章では、清水・玉熊[39]が、科学技術研究調査のデータを用い、技術進歩について研究開発に着目して 3 つの要因に区分する研究を行ったので、その一部を紹介することとする。分析手法としては、企業ごとの生産性を売上高（総売上高）とみなし、資本金、従業員数（従業員総数）、研究費（研究費総額）等から構成される生産関数を構築する方法を採用することとする。

最初に、科学技術研究調査のデータから企業の売上高と研究費の関係を概観する。次に、生産関数に関する先行研究を紹介し、2003 年から 2007 年までの科学技術研究調査の個別データを基に、産業別に生産関数を構築し、産業別特徴を明らかにする。

特に、科学技術研究調査が研究開発活動を明らかにする調査であることを踏まえ、研究費を無形資産とみなし、弾力性とタイムラグに着目する。

なお、科学技術研究調査では、毎年、従業者数等の人数及び資本金については3月31日現在、研究費や売上高などの経理項目については前年度の状況を調査している。

1 企業の研究投資効果

科学技術研究調査によると、日本における研究費のGDPに対する比率は増加傾向で推移し、2006年度には3.62%と、調査開始以来最高となった。これは諸外国と比較しても高い水準である。また、研究費については公的機関、大学等と比べて企業の割合が高く、年々増加傾向で推移している。日本は他国と比べても企業の割合が高く、OECD[40]によると、研究費の支出源別に見ると、ルクセンブルクに次いで2位、実施者別に見ると、ルクセンブルク、韓国に次いで3位となっている。このように、日本の研究における企業の役割は極めて大きなものとなっている。

ここでは、科学技術研究調査の個別データを再集計して、企業の研究費と売上高の関係を通して、研究費の投資効果を研究する。

もっとも、2007年科学技術研究調査によると、2006年度の企業の売上高に占める研究費の割合は1.4%に過ぎず、研究をしている企業に限定しても3.0%程度にとどまる。この僅かな比率の差が売上高の相異をもたらしているのかどうかを検証することが本研究の目的である。

今回使用した科学技術研究調査のデータは、2003年以降継続して調査している企業のうち、2007年調査で3月31日現在の資本金が10億円以上の企業で、研究費及び売上高が2002年度から2006年度のいずれかの年度で0になる企業並びに売上高に記入の必要がない企業を除いた1,472企業に係るいわゆるパネルデータである。

なお、2006年度において、企業の研究費の総額の89.9%は資本金10億円以上の企業によって占められている。

1.1 研究費と売上高の関係

まず、研究費と売上高の関係を論じる。

総じて、研究費が多いほど売上高も多い関係がある。しかし、1.2.1に示すとおり、研究費と売上高の関係を見ると、必ずしも強い関係が見られない。それは、研究は基礎研究から実用のための研究まで様々であり、また、研究が生産に結実するまでの期間は研究の分野、内容、性格等によって異なるからである。しかし、それらの条件をすべて制御することはできないので、最初に、科学技術研究調査産業小分類(38区分)

ごとに、研究費の支出と売上高の年度の対応にタイムラグを設け、両者の関係を見てみた。

研究費支出のタイムラグを1年としたとき、医薬品工業で強い相関が見られた。また、タイムラグを2年としたとき、鉄鋼業でも強い相関が見られた。しかし、詳細に検討すると、これらの結果は大きな企業の数社が外れ値として作用したことによるもので、研究費と売上高の関係を示すものとは言えない。すなわち、この種の分析に際しては、売上高と研究費の規模が大きい企業による影響について十分に配慮する必要がある。

1.2 対前年度変化率による階級区分

規模の効果については、前年度変化率に着目すればある程度除去することが可能である。そこで、ここでは研究費の対前年度変化率と売上高の対前年度変化率の関係を検討する。

研究費の対前年度変化率は、個々の企業の研究の内容と状況によって区々であり、年によって変動が激しいが、ある程度大括りにして比較をすればおよそその傾向が現れることが期待される。そこで、研究費の対前年度変化率について、まず、増加している企業と減少している企業に区分した上で、増加している企業については2倍以上とそれ未満に区分し、それぞれのグループごとに売上高の対前年度変化率の平均を比べることにした。なお、研究費に増減がない企業は、産業別には存在しない場合があるが、存在する場合には増加している企業と同じ区分に含めた。

1.2.1 全産業の研究費と売上高の関係

表21は、研究費の対前年度変化率(%)を100以上、0以上100未満、0未満の3つの階級に区分し、階級ごとに研究費を支出した年度以降の経過年別に売上高の対前年度変化率の平均を比較したものである。研究費の支出と売上高の年度が同一のとき、研究費の対前年度変化率が高ければ売上高の対前年度変化率も高いことが観察される。したがって、1年以内の極めて短期間に研究費が売上高に影響を及ぼしている可能性が認められる。しかし、研究費の対前年度変化率が高くても数年後の売上高の対前年度変化率が高くなるとは限らない。このように、全体として研究開発投資の効果は継続性が明確でない。

表 21 研究費の対前年度変化率階級別売上高の対前年度変化率の平均（全産業）

調査年度	研究費の対前年度変化率	売上高の対前年度変化率(%)の平均			
		0年目	1年目	2年目	3年目
2004	100%以上	14.2	9.1	2.3	6.9
	0以上100%未満	5.7	6.0	5.1	7.9
	0未満	4.6	9.1	6.2	7.3
2005	100%以上	13.5	9.4	6.7	
	0以上100%未満	8.7	5.4	7.3	
	0未満	5.9	5.5	7.9	
2006	100%以上	32.0	4.0		
	0以上100%未満	5.8	8.4		
	0未満	3.6	6.4		
2007	100%以上	21.8			
	0以上100%未満	8.9			
	0未満	4.8			

1.2.2 医薬品工業のケース

これを産業別に見ると、表 21-1 に示したように、事情は異なる。医薬品工業を例にとると、同一年度の 2004 年度と 2006 年度については全産業で観察された関係は認められない。ただし、2004 年については、研究費の対前年度変化率が高いほど 3 年後の売上高の対前年度変化率も高く、医薬品工業においては研究開発投資の効果が遅れて現れることが示唆される。

表 21-1 研究費の対前年度変化率階級別売上高の対前年度変化率の平均

（医薬品工業）

調査年度	研究費の対前年度変化率	売上高の対前年度変化率(%)の平均			
		0年目	1年目	2年目	3年目
2004	100%以上	0.0	7.6	4.2	4.5
	0以上100%未満	0.5	-0.4	4.3	2.5
	0未満	1.4	1.9	0.1	0.9
2005	100%以上	2.1	35.3	2.1	
	0以上100%未満	0.7	1.3	3.4	
	0未満	0.7	2.9	-0.3	
2006	100%以上	21.0	2.8		
	0以上100%未満	0.3	4.0		
	0未満	1.8	-1.5		
2007	100%以上	10.3			
	0以上100%未満	3.1			
	0未満	-0.9			

1.2.3 鉄鋼業のケース

鉄鋼業においては、表 21-2 に示したように、同一年度の研究開発投資が売上高に及ぼす効果が認められない。1 年後及び 2 年後に及ぼす効果については、2005 年度に

において認められるものの、他の年度においては認められないことから、研究開発投資の効果がタイムラグを伴って現れるとは言えない。

表 21-2 研究費の対前年度変化率階級別売上高の対前年度変化率の平均
(鉄鋼業)

調査年度	研究費の対前年度変化率	売上高の対前年度変化率(%)の平均			
		0年目	1年目	2年目	3年目
2004	100%以上	2.4	-1.0	-5.2	7.7
	0以上100%未満	10.0	19.8	8.4	16.8
	0未満	7.2	20.7	11.9	6.0
2005	100%以上	20.9	12.1	18.7	
	0以上100%未満	15.8	10.6	10.2	
	0未満	23.5	9.2	9.7	
2006	100%以上	10.6	-1.8		
	0以上100%未満	12.2	12.1		
	0未満	6.8	9.9		
2007	100%以上	7.7			
	0以上100%未満	10.7			
	0未満	10.7			

2 生産性の要因分析

第1節で企業の研究投資効果を概観したが、一部の産業で研究費の変化が売上高の変化に結び付いている可能性が認められるものの、総じて研究費だけで売上高に有意に影響するとは言えない。売上高は、研究費に加えて、資本や従業者数等、他の変数とも密接に関係しており、それらを総合して影響を計測することで説明力が強まる。ここでは、生産量と資本、労働等との関係を定式化した生産関数を構築し、研究開発投資に相当する変数の影響を計測することとする。

企業の生産性の分析において、生産関数として次の Cobb-Douglas 型を用いることが多い。

$$Q_t = T_t K_{t-1}^\alpha L_t^\beta \quad (3-1a)$$

ここで、 Q_t は当期の生産量、 T_t は当期の全要素生産性、 K_{t-1} は前期の資本ストック、 L_t は当期の労働力である。実証分析においては、入手できるデータの制約もあり、 Q_t には売上高、 K_{t-1} には建物、設備等の資産、 L_t には総労働時間又は従業者数が使用される。

なお、生産量の指標としては付加価値が用いられることが適当であるが、使用し得るデータの制約が強いため、ここでは売上高で代理している。

両辺の対数をとると、

$$\ln Q_t = \ln T_t + \alpha \ln K_{t-1} + \beta \ln L_t \quad (3-1b)$$

と表され、(3-1b)を推定することによりデータとして直接観察できない T_t を事後的に求めることができる。

科学技術の発展を踏まえた全要素生産性は

$$T_t = Ae^{\lambda t} R_{t-1}^{\gamma} \quad (3-2a)$$

と表される。 R_{t-1} は前期の研究開発ストックである。実証分析においては、 R_{t-1} として研究開発費、又はその累積費用が利用される。この場合、研究開発ストックは年を経るごとに有用性、あるいは産業財産権の有効期間の減少によって減価するので、何らかの減耗を考慮する必要がある。また、資本ストックには研究施設・設備が含まれており、労働力には研究者が含まれているので、厳密にはその部分が研究開発ストックと二重計算になっている。

(3-2a)の両辺の対数をとると、

$$\ln T_t = \ln A + \lambda t + \gamma \ln R_{t-1} \quad (3-2b)$$

となる。

他方、(3-2a)を(3-1a)に代入すると、

$$Q_t = Ae^{\lambda t} K_{t-1}^{\alpha} L_t^{\beta} R_{t-1}^{\gamma} \quad (3-3a)$$

となる。(3-3a)の両辺の対数をとると、

$$\ln Q_t = \ln A + \lambda t + \alpha \ln K_{t-1} + \beta \ln L_t + \gamma \ln R_{t-1} \quad (3-3b)$$

となる。 α 、 β 、 γ は、それぞれ資本ストック、労働力、研究開発ストックに対する生産量の弾力性を示すが、独立であるとは限らず、例えば、 $\alpha + \beta = 1$ 、 $\alpha + \beta + \gamma = 1$ などの条件が付されることもある。特に、 γ は研究開発投資効果の指標として限界収益率（研究開発弾力性）と呼ばれている。

データを用いた実証分析は、通常(3-3b)、又はその階差により行われるが、最初に(3-1b)から全要素生産性を計算し、(3-2b)から計算する方法もある。

(3-3a)で t の値を変えることによりデータの時点が変わるモデルは時系列モデルと呼ばれている。一方、(3-3a)で t を特定時点に固定したモデルはクロスセクションモデルと呼ばれている。また、資本ストックと研究開発ストックの時点を1期前でなく当期にしている場合もある。

時系列モデル(3-2a)の両辺を微分すると

$$\frac{\partial T_t}{\partial t} = \lambda T_t + \gamma \frac{T_t}{R_{t-1}} \frac{\partial R_{t-1}}{\partial t}$$

となる。すなわち、全要素生産性の成長率は

$$\frac{\frac{\partial T_t}{\partial t}}{T_t} = \lambda + \gamma \frac{\frac{\partial R_{t-1}}{\partial t}}{R_{t-1}}$$

と研究開発ストックの成長率で表すことができる。

また、(3-3a)を R_{t-1} で微分した

$$\frac{\partial Q_t}{\partial R_{t-1}} = \gamma \frac{Q_t}{R_{t-1}}$$

を研究開発収益率と呼ぶことがある。

上記のような生産関数を踏まえた研究は、これまで複数の国・機関の研究者によって数多く行われており、過去の実証分析の結果の一部は Congressional Budget Office, Congress of the United States[41]に詳しい。そこでは、研究費を個々の企業ごとに見ると、時系列のほうがクロスセクションよりも差が小さいことから、時系列モデルはクロスセクションモデルと比べて限界収益率が小さくなる傾向があると指摘されている。上記論文で紹介されている分析結果は、時系列モデルについては 1969 年から 2003 年までの間に行われた 12 件、クロスセクションモデルについては 1969 年から 1995 年までに行われた 19 件であり、限界収益率はそれぞれ、時系列モデルが 0.00～0.17、クロスセクションモデルが 0.05～0.56 の範囲で報告している。

クロスセクションモデルの結果の中には、日本企業を対象にした分析結果として、Mansfield が 1988 年に製造業 17 社を対象として限界収益率を 0.42 と計測し、Griliches and Mairesse が 1990 年に 1973 年から 1980 年における製造業 406 社を対象として限界収益率を 0.20～0.56 と計測したと記している。

Goto and Suzuki[42]も、研究開発データに関する企業アンケート調査の結果を用い、日本開発銀行の企業財務データ、日本銀行の物価指数、通商産業省の鉱工業指数、労働省の毎月勤労統計調査などと組み合わせて計測し、1976 年から 1984 年における限界収益率を約 0.4 と分析している。

以上の先行研究から、1960 年代から 1980 年代は、日本企業は、歴史的に見ても国際的に見ても経済的に好調であったことを反映し、研究費の限界収益率も比較的高い値となっていた。

1990 年代のバブル経済崩壊後には経済情勢の悪化に伴い限界収益率は小さくなっている。高橋通典[43]は、1991 年から 2000 年までの製造業 11 産業のパネルデータを用いて、0.1 と推計している。最近では、中野諭[44]が 1984 年から 2006 年の科学技術研究調査から、パネル化が可能で財務データとの接合が可能な 460 企業の個別データを用いて、4 つのケースで限界収益率を推計しており、製造業全体では、0.07～

0.12 であると報告している。Goto and Suzuki[42]では産業別に限界収益率から研究開発収益率も計算している。

研究開発ストックは、フローである研究費と1期前の研究開発ストックを用いて以下の式で推定される。研究開発支出額を E_t 、研究開発ストックを R_t 、研究開発支出が生産に影響を与えるまでのタイムラグを s 、減耗率を δ とすると、

$$R_t = E_{t-s} + (1 - \delta)R_{t-1}$$

と表される。これは設備投資と資本ストックとの間における恒久棚卸法の考え方である。ただし、研究開発ストックの初期値や減耗率は不明であり、何らかの方法に基づいて仮定されることが多い。

減耗率については、文部科学省が調査した「民間企業の研究活動に関する調査」による産業別特許収入期間又は特許の残存件数を利用する方法が一般的である。

2.1 生産関数の構築

清水・玉熊[39]では、科学技術研究調査の個別データを用いて、産業別に上記の先行研究に基づき生産関数を構築し、研究投資のタイムラグや研究費の限界収益率を推計している。先行研究では、タイムラグについてアンケート調査等による所与の値として研究が行われているが、清水・玉熊[39]では推計結果の評価をもとに判断している。

科学技術研究調査は、2002年に対象産業を第3次産業にまで広げるなどの調査の大幅な見直しが行われたが、今回使用した個別データは、2003年から2007年までの間において調査が行われた資本金が10億円以上で、研究費及び売上高がいずれかの年で0になる企業を除いた9,733企業のデータである。

第1節で用いたデータと異なる点は、まず、第1節では継続して存在するという条件を設けているのに対し、ここでは設けていないという点である。したがって、特定年だけ研究費が存在するような企業が含まれている。また、第1節では異なる時点の同一企業のデータを1つのデータとみなしているのに対し、ここでは時点ごとに異なる項目が別々に割り当てられた別のデータとみなしているという点も異なっている。したがって、例えば、2003年から2007年まで研究費が存在する企業について、第1節では2003年から2007年までの各年に研究費が存在する1つのデータとみなしているのに対し、第2節では、年ごとに5つの異なるデータが存在するとみなし、2003年のデータは調査年だけ、2004年のデータは調査年と前年、2005年のデータは調査年から2年前まで、2006年のデータは調査年から3年前まで、2007年のデータは調査年から4年前まで研究費が存在しているとみなしている。ただし、第2節では常にこれらのすべてのデータを用いるのではなく、分析によっては2007年についてのみ行

ったものも含んでいる。さらに、データの区分に用いる資本金額階級と産業分類について、第1節では調査結果を踏まえているのに対し、ここでは調査実施前の状況を反映させている。

生産関数において、各変数は絶対量よりデータ間の相対的大きさが重要である。このため、科学技術研究調査からは相当するデータが得られない資本ストックの代わりに、金額的には小さいものの、ある程度企業間の相対的大きさを近似していると想定される資本金を利用することにした。なお、資本金は、科学技術研究調査の標本設計の基準であり、毎年資本金10億円以上の企業と1億円以上10億円未満の企業のうち前回調査で研究をしている企業のみが悉皆調査となっていることから、パネルデータとして分析をする際には制約要因となる。

2.2 データの特性

本研究で利用したデータについて調査年度の売上高、資本金、従業者数、研究費を産業別に見ると、売上高に対する資本金が最も大きい産業は鉱業、売上高に対する従業者数が最も大きい産業はその他の事業サービス業である。また、売上高に対する研究費は学術研究機関で際立って大きく、次いで医薬品工業、精密機械工業と続いている。逆に最も小さい産業は放送業であった。

表22は、上記のデータの中で、2006年度に注目して、データの継続期間によって売上高や研究費の平均や最小値がどのように変化するかを見たものである。2006年度における1,998の調査データ全体では売上高の平均は1658億円であるが、1年前(2005年度)から継続しているという条件を加えることにより平均は1806億円に上昇する。さらに、2年前(2004年度)から継続しているという条件を加えると平均は1920億円に上昇し、3年前(2003年度)から継続しているという条件を加えると2004億円に、4年前(2002年度)から継続しているという条件を加えると2079億円に増加する。このように、過去から継続しているという条件の遡る時点が古いほど平均は上昇する傾向がある。また、このことは売上高の最小値や研究費の平均や最小値にも当てはまる。

上記のように、パネルデータを作成する際には、継続して成果を挙げているという条件により企業規模に関連する結果が上方にシフトするので、一定程度結果を幅広に捉えることが必要である。また、長期間でパネルデータを作成すると、上方シフトによる影響が強くなるので、パネルデータで分析を行うに際しては、分析の内容によって短期間に限定するなどの工夫が必要である。

表 22 2006 年度のデータにおける売上高及び研究費の平均値及び最小値

		総数	1年前 から継続	2年前 から継続	3年前 から継続	4年前 から継続
標本数		1998	1793	1628	1525	1433
売上高 (百万円)	平均値	165753	180555	191962	200409	207863
	最小値	41	41	44	517	517
研究費 (万円)	平均値	570134	625160	664440	695572	726329
	最小値	232	257	473	473	473

2.3 分析に用いた生産関数

今回分析に用いた生産関数は以下の2種類のモデルである。

(1) 時系列モデル

第1のモデル分析は(3-3a)を時系列モデルとして扱ったものである。

① 研究費単年型

$$Q_t = Ae^{\lambda t} K_{t-s}^\alpha L_t^\beta R_{t-s}^\gamma$$

過去における資本金と研究費の効果を見るためにタイムラグ（波及までの経過年数） s の値を0、1、2、3と変化させる。 $s=1$ のときは過去に研究実績のある典型的な時系列モデルになるが、今回は資本金と研究費の時間的波及効果を見るために、 s の値を可能な限り変化させてみることにした。もともと、 s の値が大きくなるにつれて選択できる t の種類が限られ、これにより時間係数 λ が有意でなくなる傾向がある。なお、 $t = (\text{西暦年の下1桁}) - 2$ とした。

なお、 t は2001年度から調査年度までの経過年数であり、売上高及び研究費については $(t+2001)$ 年度の値を意味することとなる。また、データの単位は、調査票の記入単位とし、売上高及び資本金は百万円、研究費は万円、従業者総数については人である。生産関数において、単位の相違は A の値に反映されるだけなので分析そのものに影響はない。

経理項目については本来実質値で計算すべきであるが、データを用いた期間が5年以内と短いことから、便宜上名目値のままデータを適用した。同一時点のすべてのデータに対してデフレーターを同一とみなすのであれば、デフレーターは A の値に反映されるだけなので本質的な影響はない。

② 研究費蓄積型

$$Q_t = Ae^{\lambda t} K_{t-s}^\alpha L_t^\beta \left(\sum_{u=1}^{t-s} R_u \right)^\gamma$$

研究費について、企業ごとに2002年度以降の研究費を累積してストックとみなしている。つまり、研究費単年型では減耗率を1、研究費蓄積型では減耗率を

0という極端なケースを想定していることになる。また、研究費蓄積型では研究費の初期値を $R_0 = 0$ と仮定していることとなる。なお、研究費蓄積型では 2002 年度から研究費が毎年存在しているデータだけを利用し、途中で又は途中まで研究費が存在しないデータは除外している。

(2) クロスセクションモデル

第2のモデルは(3-3a)をクロスセクションモデルとして扱ったものである。

① 研究費単年型

$$Q_s = Ae^{5\lambda} K_{5-s}^\alpha L_5^\beta R_{5-s}^\gamma$$

時系列モデルでは、 s の値が2や3で時間係数 λ が有意でなくなる場合が多くなるので、時間変数を除外した。 s の値は時系列モデルと同じである。また、データを2006年度調査に限定し $t=5$ とした。

② 研究費蓄積型

$$Q_s = Ae^{5\lambda} K_{5-s}^\alpha L_5^\beta \left(\sum_{u=1}^{5-s} R_u \right)^\gamma$$

同様に、クロスセクションモデルについても研究費蓄積型の分析を行った。

2.4 生産性の要因分析

各生産関数の両辺の対数を取り、データを適用した結果は以下のとおりである。

時系列モデル、クロスセクションモデルのいずれについても、研究費単年型、研究費蓄積型のいずれについても、全体として売上高は資本金、従業者数及び研究費（又はその累積）で説明が可能である。

また、時系列モデル、クロスセクションモデルのいずれについても、研究費単年型の場合にはタイムラグを2年、研究費蓄積型の場合にはタイムラグを0年とみなすと決定係数と多くの変数に係る t 値が最も高くなった。資本金、従業者数、研究費のうち売上高に対して最も強い影響を及ぼす変数は従業者数であり、従業者数は単独でも相当の説明力を有する。他方、研究費の影響は資本金よりも小さくなっている。

研究費の限界収益率を見ると、時系列モデルについては0.04~0.05、クロスセクションモデルについては0.03~0.05と、いずれもかなり低い水準であり、モデルによる差は大きくない。時系列モデルではタイムラグの大きさにかかわらず単年型と蓄積型の間で差はなかった。クロスセクションモデルではタイムラグが大きくなるほど大きくなる傾向が見られた。研究費を蓄積するとタイムラグが短いほど研究費の蓄積が増えるが、実際にはそれほどの影響が存在せず、それだけ限界収益率は小さくなる傾向がある。

これらの実証結果は表 23 に示した。表中*は 10%の水準で有意、**は 5%の水準で有意、***は 1%の水準で有意であることを示している。また、表中 n はデータ数、Adjusted R² は自由度調整済決定係数である。(以下同)

表 23 売上高の要因分解 全産業

(1)時系列モデル①研究費単年型												
説明変数	ラグ 0年			ラグ 1年			ラグ 2年			ラグ 3年		
	係数	t 値		係数	t 値		係数	t 値		係数	t 値	
定数	2.48	50.04	***	2.41	40.23	***	2.46	31.86	***	2.44	19.02	***
年	0.04	8.65	***	0.05	7.09	***	0.04	3.75	***	0.04	1.74	*
資本金(対数)	0.29	38.57	***	0.32	35.93	***	0.34	32.99	***	0.34	26.39	***
従業者数(対数)	0.77	90.08	***	0.74	73.88	***	0.71	59.82	***	0.71	47.03	***
研究費(対数)	0.04	7.20	***	0.04	6.39	***	0.04	5.28	***	0.05	5.12	***
n	9707			6998			4978			3227		
Adjusted R ²	0.79			0.79			0.80			0.79		

(1)時系列モデル②研究費蓄積型												
説明変数	ラグ 0年			ラグ 1年			ラグ 2年			ラグ 3年		
	係数	t 値		係数	t 値		係数	t 値		係数	t 値	
定数	2.46	50.65	***	2.42	41.94	***	2.50	33.69	***	2.59	20.55	***
年	0.03	6.22	***	0.03	4.51	***	0.02	1.36		0.00	-0.21	
資本金(対数)	0.31	39.78	***	0.33	36.23	***	0.34	32.57	***	0.34	26.04	***
従業者数(対数)	0.76	82.99	***	0.74	70.91	***	0.71	58.06	***	0.70	46.25	***
研究費(対数)	0.04	6.76	***	0.04	5.80	***	0.04	5.46	***	0.05	5.67	***
n	8272			6301			4666			3122		
Adjusted R ²	0.81			0.81			0.81			0.80		

(2)クロスセクションモデル①研究費単年型												
説明変数	ラグ 0年			ラグ 1年			ラグ 2年			ラグ 3年		
	係数	t 値		係数	t 値		係数	t 値		係数	t 値	
定数	2.61	23.21	***	2.59	22.32	***	2.69	23.61	***	2.66	23.39	***
資本金(対数)	0.29	16.08	***	0.32	17.17	***	0.35	19.01	***	0.33	18.01	***
従業者数(対数)	0.78	39.69	***	0.73	35.45	***	0.70	34.07	***	0.73	33.75	***
研究費(対数)	0.05	3.90	***	0.05	4.10	***	0.04	3.23	***	0.03	2.63	***
n	1998			1793			1674			1619		
Adjusted R ²	0.76			0.77			0.79			0.79		

(2)クロスセクションモデル②研究費蓄積型												
説明変数	ラグ 0年			ラグ 1年			ラグ 2年			ラグ 3年		
	係数	t 値		係数	t 値		係数	t 値		係数	t 値	
定数	2.67	21.47	***	2.62	21.28	***	2.60	21.75	***	2.57	21.67	***
資本金(対数)	0.34	18.15	***	0.36	18.68	***	0.36	18.71	***	0.33	17.48	***
従業者数(対数)	0.72	31.82	***	0.71	31.15	***	0.70	31.50	***	0.73	32.66	***
研究費(対数)	0.03	2.35	**	0.03	2.28	**	0.04	2.65	***	0.05	3.36	***
n	1433			1433			1468			1514		
Adjusted R ²	0.81			0.81			0.81			0.80		

これを産業別に見ると、38 産業のうち、農林水産業、鉱業、印刷業、放送業、新聞・出版・その他の情報通信業、運輸業、金融保険業、専門サービス業、学術研究機関、その他の事業サービス業の 10 産業で、対象が 10 に満たない数しか存在しないケースがあるので、これらの産業は分析に耐えられないものと判断した。

残り 28 産業のうち、電子応用・電気計測器工業についてはすべてのモデル・型において資本金に係る係数が 10%の水準で有意でなく、クロスセクションモデルにおける研究費蓄積型では研究費に係る係数も 10%の水準で有意でなかった。また、パルプ・紙工業、油脂・塗料工業についてはすべてのモデル・型において研究費に係る係数が 10%の水準で有意でなかった。

以下は産業別に見て各係数が10%の水準で有意な結果のみを掲載し、その状況をまとめたものである。根拠となる分析結果は清水・玉熊[39]に掲載している。

なお、定数項及び時間係数については、他の変数に大きな影響はないので、有意でなくても問題ないとした。

(1) 建設業

建設業については、モデル、研究費の蓄積の有無、ラグにかかわらず係数はすべて1%の水準で有意であり、研究費よりも従業者数の影響力が大きかった。また、ラグが1年及び2年のときに決定係数は最も高く、研究費の限界収益率は0.10~0.15であった。全体として、型間での研究費の限界収益率の差は大きくなかった。

(2) 食品工業

食品工業の場合、時系列モデルについては研究費の蓄積の有無、ラグにかかわらず係数はすべて1%の水準で有意であるが、クロスセクションモデルについては研究費単年型でラグが0年のときを除いて研究費が1%の水準で有意でなかった。従業者数の影響力が最も大きく、資本金の影響力も比較的大きかった。ラグが0年における研究費の限界収益率は、クロスセクションモデルー研究費単年型のみ0.18で、他は0.12であった。

(3) 繊維工業

繊維工業の場合、時系列モデルのラグが2年未満のみすべての変数について10%の水準で有意であり、時系列モデルのラグが2年以上及びクロスセクションモデルは資本金又は研究費に係る係数が10%の水準で有意でなかった。時系列モデルではラグが0年のときのみ研究費に係る係数が1%の水準で有意であり、研究費の限界収益率は研究費単年型が0.15、蓄積型が0.13であった。いずれにおいても従業者数の影響力が大きかった。

(4) 医薬品工業

医薬品工業は、クロスセクションモデルー研究費蓄積型のすべて、その他のモデルの一部で資本金又は研究費に係る係数が10%の水準で有意でなかった。また、資本金がマイナスに寄与するという特徴がある。研究費に係る係数が1%の水準で有意になるのは時系列モデルー研究費単年型のラグ0年、時系列モデルー研究費蓄積型のラグ0年及び3年、クロスセクションモデルー研究費単年型のラグ1年の4種類のみであり、それらの研究費の限界収益率は0.10~0.20であった。医薬品工業では、売上高に対して資本金がマイナスに寄与している分を従業者数の影響が極めて大きくなることにより補填している。

(5) 総合化学・化学繊維工業

総合化学・化学繊維工業については、モデル、研究費の蓄積の有無、ラグにかかわらず係数はすべて10%の水準で有意であり、ケースにより研究費の方が従業者数よりも影響力が大きかった。クロスセクションモデルー研究費単年型では、ラグが2年以下で従業者数が1%の水準で有意でなかった。ラグが3年のときの研究費の限界収益率は、時系列モデルで0.31又は0.32と、すべての産業の中で最も高くなるが、クロスセクションモデルでは0.21又は0.22であった。

(6) その他の化学工業

その他の化学工業の場合、クロスセクションモデルの一部で、資本金又は研究費に係る係数が10%の水準で有意でなかった。クロスセクションモデルで研究費の蓄積の有無にかかわらず有意になるのはラグが0年のときだけであった。ラグが0年のときの各モデルー型の研究費の限界収益率は0.19~0.22と比較的高いが、この場合にも最も影響力の大きい変数は従業者数であった。全体として、型間での研究費の限界収益率の差は大きくなかった。

(7) 石油製品・石炭製品工業

石油製品・石炭製品工業については、大半のモデルにおいて従業者数又は研究費に係る係数が10%の水準で有意でなく、時系列モデルー研究費蓄積型においてラグが2年以上のときにのみ各係数が10%の水準で有意であり、従業者数の影響力が大きかった。このうち、ラグが2年のときに決定係数と各係数の t 値が高く、研究費の限界収益率は0.28と比較的高かった。

(8) プラスチック製品工業

プラスチック製品工業の場合、クロスセクションモデルにおいてラグが3年以上のときに研究費に係る係数が10%の水準で有意でなかった。また、クロスセクションモデルは、いずれも研究費に係る係数は1%の水準で有意でなかった。

他方、時系列モデルについては、研究費の蓄積の有無、ラグにかかわらず係数はすべて1%の水準で有意であった。研究費単年型、蓄積型のいずれにおいてもラグが0年において決定係数とすべての係数に係る t 値が最も高くなり、いずれも従業者数の影響力が大きいが、研究費の限界収益率は0.22又は0.23と比較的高かった。全体として、型間での研究費の限界収益率の差は大きくなかった。

(9) ゴム製品工業

ゴム製品工業は時系列モデルのラグが0年以外は資本金に係る係数が10%の水準で有意でなかった。(クロスセクションモデルー研究費単年型ーラグ0年については研究費に係る係数も10%の水準で有意でなかった。)従業者数の影響が大きく、研究費の限界収益率は研究費単年型の場合には0.19、蓄積型の場合には0.16であった。

(10) 窯業

窯業の場合、モデル、研究費の蓄積の有無、ラグにかかわらずすべての変数について5%の水準で有意であり、多くのモデル型で従業者数の影響力が最も大きかったが、クロスセクションモデルー研究費蓄積型でラグが1年以内のケースで研究費の影響力が最も大きかった。また、時系列モデルにおいてはラグが0年のときに決定係数と各変数の t 値が最も高く、研究費の限界収益率は0.17又は0.19であった。他方、クロスセクションモデルにおいてはラグが2年のときに決定係数及び研究費の t 値が最も高く、研究費の限界収益率は0.29又は0.30と比較的高かった。

(11) 鉄鋼業

鉄鋼業の場合、時系列モデルにおいて研究費単年型でラグが1年以下又は研究費蓄積型でラグが0年のとき、クロスセクションモデルにおいて研究費単年型でラグが2年のときにのみすべての変数について10%の水準で有意な結果が得られた。他のケースについては研究費に係る係数が10%の水準で有意でなかった。時系列モデルにおいてラグが0年のときにのみ研究費に係る係数が5%の水準で有意であった。売上高に対しては従業者数の影響が大きく、研究費の限界収益率は0.06又は0.07と比較的低かった。

(12) 非鉄金属工業

非鉄金属工業の場合、大半のケースで研究費に係る係数が10%の水準で有意でなく、クロスセクションモデルー研究費単年型においてラグが2年のときにのみ10%の水準で有意な結果が得られたが、研究費に係る係数が5%の水準で有意でなく、研究費の効果が逆向きになっている。非鉄金属工業についても従業者数の影響力が最も大きくなっている。

(13) 金属製品工業

金属製品工業については、時系列モデルについては、研究費の蓄積の有無、ラグにかかわらず係数はすべて5%の水準で有意であった。クロスセクションモデルについては研究費に係る係数が10%の水準で有意でないケースが見られた。すべてのモデル、型で10%の水準で有意であるのはラグが1年のときだけで、研究費の限界収益率は0.09～0.19であった。金属製品工業についても従業者数の影響力が最も大きかった。

(14) 機械工業

機械工業については、モデル、研究費の蓄積の有無、ラグにかかわらずすべての変数について1%の水準で有意であり、従業者数の影響力が最も大きかった。また、いずれのモデル型においても、研究費の限界収益率は比較的高く、0.20～0.23と狭い範囲に収まっていた。型間での研究費の限界収益率の差は存在しなかった。

(15) その他の電気機械器具工業

その他の電気機械器具工業の場合、時系列モデルについてはすべてのケースにおいて各変数の係数は1%の水準で有意であった。クロスセクションモデルについては、研究費単年型でラグが1年以下において各変数の係数は10%の水準で有意であったが、その他のケースは研究費に係る係数が10%の水準で有意でなかった。また、クロスセクションモデルー研究費単年型で研究費に係る係数が5%の水準で有意であるのはラグが1年のときだけである。ラグが1年のときの研究費の限界収益率は0.14～0.16であった。いずれについても売上高に最も大きな影響を及ぼしているのは従業者数である。全体として、型間での研究費の限界収益率の差は大きくなかった。

(16) 情報通信機械器具工業

情報通信機械器具工業については、モデル、研究費の蓄積の有無、ラグにかかわらずすべての変数について10%の水準で有意であった。クロスセクションモデルー研究費蓄積型における研究費に係る t 値は比較的低い、その中ではラグが2年のときに最も高く、研究費に係る係数は5%の水準で有意になった。いずれについても売上高に最も大きな影響を及ぼしているのは従業者数であり、ラグが2年のときの研究費の限界収益率は0.15～0.18となっている。

(17) 電子部品・デバイス工業

電子部品・デバイス工業については、モデル、研究費の蓄積の有無、ラグにかかわらずすべての変数について10%の水準で有意であり、従業者数の影響力が最も大きかった。クロスセクションモデルー研究費蓄積型においては研究費に係る t 値が低い、その中ではラグが1年のときに最も高くなった。ラグが1年における研究費の限界収益率は0.18～0.26と比較的高かった。全体として、型間での研究費の限界収益率の差は大きくなかった。

(18) 自動車工業

自動車工業については、モデル、研究費の蓄積の有無、ラグにかかわらずすべての変数について5%の水準で有意であり、決定係数はどのケースもすべて0.92～0.93と説明力に大差はない。また、時系列モデルについてはラグが短いほど各係数の t 値が高くなるが、クロスセクションモデルについては総じてラグが長いほど t 値が高くなる傾向があった。自動車工業の場合も、研究費よりも従業者数の影響力が大きく、研究費の限界収益率は0.11～0.15であった。全体として、型間での研究費の限界収益率の差は大きくなかった。

(19) その他の輸送用機械工業

その他の輸送用機械工業の場合、資本金及び研究費に係る係数が10%の水準で有意であるケースは時系列モデルにおける研究費単年型でラグが2年以内又は研究費蓄

積型でラグが0年に限られた。時系列モデルでラグが0年における研究費の限界収益率は0.12又は0.15であった。この場合にも従業者数の影響力が最も大きかった。

(20) 精密機械工業

精密機械工業の場合、資本金及び研究費が有意であるケースは時系列モデルにおける研究費蓄積型でラグが0年又は3年に限られた。このうち、ラグが3年のときには資本金に係る係数が5%の水準で有意でない。ラグが0年における研究費の限界収益率は0.15であり、この場合にも従業者数の影響力が最も大きかった。

(21) その他の工業

その他の工業については、クロスセクションモデルー研究費蓄積型でラグが1年及び2年のときに資本金に係る係数が10%の水準で有意でないが、他のケースはすべての変数について10%の水準で有意であった。これらのケースにおいては、ラグが0年のときに多くの係数の t 値が最も高い。また、ラグが0年における研究費の限界収益率は0.18~0.24であった。その他の工業についても従業者数の影響が最も大きかった。全体として、型間での研究費の限界収益率の差は大きくなかった。

(22) 電気・ガス・熱供給・水道業

電気・ガス・熱供給・水道業については決定係数が高く、モデルの説明力が強い。それでもクロスセクションモデルについては、ラグが2年以上で研究費の係数が10%の水準で有意でない。その他のケースの中では、ラグが0年のときに多くの係数の t 値が最も高い。売上高に対する影響は従業者数で大きく、研究費の限界収益率は0.07又は0.08と比較的低い。型間での研究費の限界収益率の差は大きくなかった。

(23) ソフトウェア・情報処理業

ソフトウェア・情報処理業については、研究費に係る係数が10%の水準で有意でなくなる場合が多く、時系列モデルにおけるラグが1年以下のみすべての変数について10%の水準で有意であった。従業者数の影響力が極めて大きく、研究費の限界収益率は0.06~0.08と比較的低かった。

(24) 通信業

通信業については、資本金又は研究費に係る係数が10%の水準で有意でなくなる場合が多く、時系列モデルー研究費蓄積型、クロスセクションモデルー研究費単年型におけるラグが0年のときにのみすべての変数が有意であった。研究費がマイナスに影響し、従業者数の影響が極めて大きくなっているが、数少ない従業者数の多い大企業の影響を受けた結果である。

(25) 卸売業

卸売業については、資本金又は研究費に係る係数が10%の水準で有意でなくなるケースが多く、時系列モデルの一部のみすべての変数について10%の水準で有意であ

った。研究費がマイナスに寄与しているが決定係数が低い。従業者数が大きくプラスに寄与している。

2.5 研究投資の効果とタイムラグ

表 24 は、2.4 の結果から研究投資効果が認められた産業について、タイムラグと限界収益率をまとめたものである。タイムラグは、生産関数の適合度、すなわち全体の説明力や説明変数に係る t 値が最も高い期間又はその周辺とみなしたが、使用したデータが5年間に限られていることから、それよりも長期の波及期間が存在する場合には検出できていないことになる。

これを見ると、タイムラグは明確でない産業が多いものの、1年以内の産業が多く、2年を超える産業は総合化学・化学繊維工業のみである。国際レベルの企業間競争と技術革新が進展するなかで、研究依存型産業では研究開発の効果は早く現れ、また、蓄積による効果がほとんど見られなかったことにかんがみると、その寿命は短くなっている状況がうかがえる。このことは、企業における研究費の大半は基礎研究費や応用研究費でなく、開発研究費であることも関連している。

生産性に対しては、総じて労働力の効果が大きく、研究開発の限界収益率は高くても0.3前後までで、研究投資が生産に直結する産業は見られない。産業別に見ると、研究投資の効果が生産性に現れるのは、建設業、製造業など第2次産業であり、第3次産業については、電気・ガス・熱供給・水道業、ソフトウェア・情報処理業で効果が小さく、通信業、卸売業では効果が見られない。製造業の中では化学系、機械系で大きく、個別には総合化学・化学繊維工業で最も大きく、その他の化学工業、石油製品・石炭製品工業、プラスチック製品工業、窯業、機械工業、電子部品・デバイス工業などで比較的大きい。

さらに、研究費の蓄積の有無により限界収益率を比較したが、多くの場合、差は大きくなかった。

表 24 産業別研究投資のタイムラグと限界収益率

産業	ラグ	限界収益率
建設業	1～2年	0.1強
食品工業	0年	0.1～0.2
繊維工業	0年	0.1強
医薬品工業	不明	0.1～0.2
総合化学・化学繊維工業	3年	0.2強～0.3強
その他の化学工業	0年	0.2前後
石油製品・石炭製品工業	2年	0.3弱
プラスチック製品工業	0年	0.2強
ゴム製品工業	0年	0.2弱
窯業	不明	0.2弱～0.3弱
鉄鋼業	0年	0.1弱
金属製品工業	1年	0.1弱～0.2弱
機械工業	不明	0.2強
その他の電気機械器具工業	1年	0.1～0.2
情報通信機械器具工業	2年	0.2弱
電子部品・デバイス工業	1年	0.2弱～0.3弱
自動車工業	不明	0.1強
その他の輸送用機械工業	0年	0.1強
精密機械工業	0年	0.1～0.2
その他の工業	0年	0.2前後
電気・ガス・熱供給・水道業	0年	0.1弱
ソフトウェア・情報処理業	0～1年	0.1弱

2.6 生産性の要因分析に関する結論と課題

上記のことから導かれる主な結論は以下のとおりである。

第1に、研究投資は企業の生産性向上に結び付いており、その程度は、資本や労働力など生産性の向上に影響している他の要因との関係である程度推定することができる。

第2に、多くの産業において生産性に最も大きな影響を及ぼしている要因は労働力である。すなわち、同一割合であれば、研究開発費を増加させるよりも従業者数を増加させる方が売上高の増加に結び付きやすい。

第3に、研究投資の効果はストックとしての研究費の捉え方によって異なるが、それによる結果に大差はない。

第4に、研究投資の効果は産業によって異なり、化学系、機械系の製造業で大きく、第3次産業では認められない。

しかし、今回の分析は大括りに概略を把握するためのもので、今後、以下のような幾つかの点について検討が必要である。

第1に、科学技術研究調査においては資本ストックに相当する項目が把握されていないことから、今回の分析では資本金で代替したが、他の調査結果又はデータベース

とのマッチングにより資本ストックに近い変数を利用して分析を試みるものが課題である。

第2に、本来必要な変数であるストックとしての研究開発についてもより緻密に把握をすることが必要である。これを推計するには研究成果の減耗率の適切な評価が必要であり、それを把握するための手法が課題である。

第3に、今回投入した資本金、労働力、研究開発費には部分的に相互に重複する部分があるのでそれを除外する方法についての検討も課題である。

なお、科学技術研究調査については、2002年に調査の見直しが行われたことから、個別データとして同一条件で分析できる期間は限られていた。調査は今後も継続して行われ、調査データの蓄積が進むにつれて分析可能な期間も長くなっていくので、このような分析から、より有意義な結果が得られることが期待される。

第IV部 消費者物価指数の意義、作成技法と利用の高度化

これまで、第I部から第III部まで、公的統計の現代的意義、作成技法の高度化、利用の高度化全体について、それぞれ包括的に解説したが、これらの事情は、個々の統計により、また、個々の国により異なる。そこで、第IV部では、経済の体温計と呼ばれ経済・金融政策にも重要なCPIに焦点を当て、具体的な研究結果を示す。

第1章のボスキンレポートをめぐる議論とCPIへの影響では、米国CPIには上方バイアスが存在するとしたボスキンレポートとそれに対するBLS (Bureau of Labour Statistics) の考え方を詳述し、日本を含む各国と比較した米国CPIの特殊性を解説している。

第2章の日本のCPIの作成技法の高度化では、デフレ下において、価格と数量が並行して変化する場合が多く、Laspeyres指数に下方バイアスが生じる費目が多いことを実証している。また、財・サービス分類別に2001年から2005年までの毎月の指数及び前年同月比の標準偏差を比較すると、指数、前年同月比とも最大は米類を含む他の(生鮮商品以外の)農水畜産物であり、石油製品(ガソリン、灯油及びプロパンガス)が続いていることを示し、物価の基調を表す指標としてそれらを除いた「食料(酒類を除く)及びエネルギーを除く総合」が適切であるとしている。さらに、基準改定による変化率の変化を、指数を100にリセットする影響、品目ウエイトを変更する影響、品目の改廃等に区別している。加えて、家賃指数が借家の転出入により変動する問題に対し、小売物価統計調査の個別データを用いて家賃を床面積、建築年及び人口密度で説明する回帰式を利用して推計する方法、転出入前後の家賃の欠測値を補完する方法についての研究結果を示す。

第3章の物価の実感に関する定量的評価では、小売物価統計調査、CPI、「生活意識に関するアンケート調査」(日本銀行)のデータを基に、物価の上昇率に関する実感と指数との差は、主として①物価指数で消費者が日常意識することがない品質の調整をしていることによる影響、②実感が購入頻度の高い品目に影響されやすいこと、③実感が直近の変化に影響されやすいことで相当程度の説明が可能であることを実証している。

第1章 ボスキンレポートをめぐる議論とCPIへの影響

米国では、1990年代後半に入って政府作成のCPIをめぐる議論が活発に行われるようになった。とりわけ1996年12月に出されたボスキンレポートはその米国のみならず他の国でも注目された。

その後で日本でも、CPIをめぐる批判があり、その中にはボスキンレポートを根拠にしているものが見られる。米国と日本ではCPIの作成方法に違いがあるため、同レポートの議論がそのまま日本に当てはまるわけではないが、同レポートは、米国のCPIの在り方を様々な角度から論じたものとして、CPIに関する見方や技術的な論点を明らかにするという意味で示唆に富んでいる。

ボスキンレポートの概要、BLSの対応、日本のCPIへの適用可能性等については1999年から2000年の間にまとめた清水誠[45]～[50]が存在するが、第1章では、ボスキンレポートが報告されて20年が経過したことを受け、ボスキンレポートをめぐる議論とCPIへの影響を冷静に振り返ることとする。本章の内容は、清水誠[45]～[50]を要約し、2000年以降の状況等を加筆したものである。

ボスキンレポートの概要に触れている論文は数多く存在するが、それに対するBLSの対応と日米のCPIの作成方法の違いを踏まえた日本のCPIへの適用可能性に言及している論文は見られず、清水誠[45]～[50]の一連の研究が初めてのものである。

なお、本章で述べているボスキンレポートの内容は、Advisory Commission to Study the Consumer Price Index[51]から、それに対するBLSの対応の内容はBureau of Labor Statistics[52]から引用している。

1 ボスキンレポートの背景

米国においては、CPIの上昇率に上方バイアスが存在するという批判が長く存在していた。

1961年には、George Stiglerによる「物価統計検討委員会」がヒアリングを重ね、Subcommittee on Economic Statistics, Joint Economic Committee, Congress of the U.S.[53]にあるように、米国CPIと「真の」生計費指数の相違点を指摘し、BLSがCPIを生計費指数に近づけるための研究を開始するよう勧告したが、それをきっかけに、1990年代に議会が予算と社会保障生計費及び税指数を決定する際のCPIの役割に関心を持つようになった。

1990年代後半には政治の舞台でこの問題が定量的に取り上げられるほど加熱することとなったが、その背景には、米国の財政事情がある。

米国で CPI は、1970 年代初期以前は財政部門でほとんど利用されていなかったが、その後のインフレの進行に伴い、CPI を税率区分や社会保障費などと連動させる仕組みが作られた。これにより CPI の作成方法や精度は財政状況に大きな影響を与える要素となった。

1980 年代に入ると、税収は伸び悩む一方で歳出削減が計画どおりに進まず、財政赤字が急拡大した。同時に、個人貯蓄率の急速な低下に伴って、経常収支は急速に悪化し、「双子の赤字」の時代が到来した。このため、1990 年代半ばにさしかかると、財政赤字の削減が最重要課題の 1 つとされた。

このような中、1995 年 1 月にアラン・グリーンズパン連邦準備制度理事会議長が議会の予算委員会で、米国 CPI には上方バイアスがあり、このバイアスを是正すれば財政赤字の改善に役立つとの趣旨の発言をしたことがきっかけとなって、CPI の精度に関して検討するよう委員会に諮問することとされた。こうして発足したのが、ボスキン氏を含む 5 名から成る CPI 諮問委員会である。

ボスキンレポートは、米国議会の CPI 諮問委員会が、上院財政委員会に提出した報告であり、*Toward a More Accurate Measure of the Cost of Living*（「生計費のより正確な尺度を目指して」）と題されている。同諮問委員会の委員長がスタンフォード大学のマイケル・ボスキン教授であることからこう呼ばれた。ボスキン氏は、以前から経済統計の精度改善に熱心であり、ジョージ・H・W・ブッシュ元大統領のホワイトハウス経済諮問委員会の委員長を務めている間に、経済測定プログラムへの追加財源の措置に努力している。

このように、米国においては財政赤字の削減が最重要課題の 1 つとされている中で、米国 CPI が財政状況に大きな影響を与える指標であったことから、ボスキンレポートを理解するに当たっては、その内容には統計学的な視点だけでなく、政治的、財政的視点も織り込まれていることに留意する必要がある。

ちなみに日本でも財政に対する CPI の影響力は大きいですが、直接連動させる仕組みが作られている費目は米国ほど広範にわたっていない。まず、歳入において、CPI は税率区分に連動していないため、直接的影響を及ぼす費目は極めて限定されている。つまり、間接的に影響を及ぼす費目は多いとはいえ、直接法律等で規定されているのは、郵便料金のようにやや特殊な収入に限られている。また、歳出においても、法律等で社会保障費が物価スライドされることとなっているが、対象は年金、児童手当などに限定されており、医療保険給付費、老人福祉費、生活保護費などは対象外である。この他、日本の CPI は土地収用の価格の算定に用いるなどの規定はあるが、歳出全体から見れば米国ほど広く規定されているとは言えない。このように、日本では、CPI が

財政に及ぼす影響は副次的かつ複雑であり、単純に CPI の上昇率の大小だけで財政赤字の総量の変化を評価できるものではない。

このことは、つきつめれば米国でも同様であろう。CPI の作成方法を見直して上昇率を引き下げると、社会保障費の削減と税金の増額などにより一時的には財政赤字の縮小効果が見込めるが、消費や貯蓄の減少などを通じて、めぐりめぐって他の費目に影響が出ることもあるので、これらを含めて財政赤字がどの程度縮小するかの視点が必要である。

2 ボスキンレポートの概要

本節では、ボスキンレポートの全体像を簡単に解説する。

CPI 諮問委員会は、1995 年 9 月に中間報告を提出し、現行の米国 CPI は物価上昇率を年率 1.5 ポイント過大評価していると指摘した。その後、1996 年 12 月の最終報告では 1996 年 3 月に BLS が公表した CPI の改定予定を反映させるとともに、一部評価手法を改め、過大評価の大きさを 1.1 ポイントと改めた。

2.1 米国 CPI における上方バイアス

ボスキンレポートにおいて、CPI は生計費の変化を測定するものと定義した上で、次の 4 つのバイアスが存在すると指摘している。

(1) 代替バイアス

商品の代替に関するバイアスであり、消費者がより安価な代替商品を購入しても、それが CPI に反映されないというバイアスである。特に、現行 CPI が固定マーケットバスケット方式を前提としていることに起因するものである。

この 1 つの要因は指数算式の問題である。委員会は、本来なら最良指数⁴⁹を採用すべきと考えるが、現行 CPI は Laspeyres 指数を採用しており、品目間での代替効果が反映されないために、上位集計レベルで年率 0.15 ポイントの上方バイアスが発生しているという。

もう 1 つは品目内で各銘柄の価格を集計する時のバイアスである。下位集計レベルについて、幾何平均を真の生計費指数とみなすとすれば、年率 0.25 ポイントの上方バイアスが発生しているとしている。

(2) 新店舗バイアス

消費者がディスカウントストアなどの安売り店にシフトする行動が適切に反映されないことによるバイアスであり、年率 0.1 ポイントの上方バイアスがあるとしている。

⁴⁹ Superlative Index。ボスキンレポートでは Fisher 又は Törnqvist 指数としている。

(3) 品質バイアス

製品の品質改良が正確に評価されないことに起因するバイアスである。ボスキンレポートでは、CPIの作成に用いられる206の品目を27のカテゴリーに分類し、それぞれバイアスの大きさを推計している。このうち、光熱、家事用品、家事サービス、他の個人交通、公共交通、医療保険、娯楽サービス、煙草の8カテゴリーについてはバイアスが存在しないとしている。一方、食料（農産物を除く）、家賃、新車などの19カテゴリーについては上方バイアスが存在するとしている。

(4) 新製品バイアス

新しい製品が出現してもCPIにはすぐに反映されないことによるバイアスであり、(3)のバイアスと合せて年率0.61ポイントの上方バイアスになるとしている。

2.2 CPI バイアスの財政への影響

これらの4つのバイアスを合計すると、先述の1.1ポイントとなる。ボスキンレポートでは、バイアスの推計幅を0.8~1.6ポイントとしている。また、同レポートでは、CBO（議会予算局）の推計によれば、今後10年間で毎年1.1ポイントの上方バイアスがあるとすれば、2006年には1480億ドルの財政赤字が水増しされ、政府の負債は6910億ドル増加することとなるとしている。この赤字要因は、社会保障費、保健医療費、防衛費に次ぐ4番目に大きい要因である。さらに、この推計を延長すると、2008年までに財政赤字は2020億ドル増加し、また、政府債務は1兆700億ドル増加するとの結果も得られている。このようにCPIのバイアスが財政に大きな影響を与えるのは、米国においては、社会保障費を始めとする歳出の約3割がCPIに自動的に連動しているほか、歳入の45%前後を占める個人所得税の税率区分が物価にスライドして調整されるためである。

2.3 勧告

ボスキンレポートでは、これらの推計を踏まえた上で、16項目の勧告を出している。このうち、最初の12項目はBLSに対するもの、その他は大統領と議会に対するものとなっている。この中の最初の4項目は以下のとおりである。

- ①消費者物価を測定する目標として生計費指数を掲げるべきである。
- ②2つの指数を作成、公表すべきである。1つは生計費指数の考えに基づくものを月1回公表し、もう1つは指数に関する最新の情報や検討の成果を反映したものを年1回公表すべきである。年次指数は、過去に遡及し改定する。

(短期的勧告)

- ③月次指数には、マーケットバスケットの変化を反映させ、生計費指数の概念に適応したものとするため、Laspeyres 指数をやめ、最良指数の算式を採用すべきである。年次指数にも同様の算式を採用し、該当年以降のデータ、更新されたウエイト及び新製品の導入を反映させるべきである。
- ④下位集計レベル（品目レベル）の指数の算出には幾何平均を用いるべきである。

2.4 BLS の対応

BLS はボスキンレポートに対して、1997 年 6 月に **Measurement Issues in the Consumer Price Index** という報告書を書いており、さらに、1998 年 6 月には特に勧告への対応についてその改定版を出している。その中で BLS は、ボスキンレポートの指摘や勧告について、BLS の見解や対応方針を述べている。BLS は、ボスキンレポートの指摘には統計的な根拠の薄弱なものが多くあること、実務的に実行が困難なものも多いことなどを説明し、BLS としては CPI の基本的な枠組みを変更することについては慎重な姿勢を示しているが、CPI に最近の社会・経済動向をよりの確に反映させるために様々な改善や検討を行いつつあることも併せて説明している。

2.4.1 指数の目的と算式の選択

まず、ボスキンレポートの勧告①から④までは CPI の基本的な目的と算式に関するものであり、CPI を生計費指数と位置付けている。ボスキンレポートでは、生計費指数とは、2つの異なる価格体系において、一定の豊かさの水準（福祉、効用、生活水準ともいう。）を達成するために必要な最低限の支出を比較するものであり、通常、2時点間の比較として考えられると述べている。換言すれば、生計費指数とは、「消費者が当初と同じ豊かさの水準を得るには、支出を当初に比べていくら増やす（減らす）必要があるか?」を表したものであるということである。そのために最良指数の算式によって指数計算を行うべきだとしている。また、現行のような固定ウエイトの Laspeyres 指数は、真の生計費指数に比べて上方バイアスを有すると指摘している。これに対して BLS は、CPI 作成の目的には様々なものがあり、その中でも最も重要なものの1つが生計費指数の作成であるとして、この意味ではボスキンレポートの考えに賛同している。しかし、BLS としては、仮に毎月公表できるような真の生計費指数の作成方法が存在するのならそれを採用したいところであるが、生計費指数というのは理論的な概念であり、実用に耐える単一明快な万能の指数算式は現存しないことに留意する必要があるとして、この点ではボスキンレポートとは立場を異にしている。

ボスキンレポートでは、指数計算における品目を合算して上位レベルの指数を算出する算式として、比較時のウェイトを用いた Törnqvist 型の算式を用いるべきであるとしている。しかし、BLS は、この方式ではウェイト作成に時間がかかり指数の公表が著しく遅れるため、仮にこれを適用するとすれば近似方式を使用せざるを得ないが、この近似方式では下方バイアスが生じる傾向にあることや、毎年事後的な改訂が必要となり利用上不便であること、結果の解釈が難しいこと、などの問題を指摘している。このため、BLS は、指数算式の変更については慎重に対応すべきであるとして、今後も引き続き Laspeyres 型を維持しつつ、指数の数式、対象品目及び作成方法など、CPI の理論的前提、概念、範囲などについて十分な情報を提供していくこととしている。

固定ウェイト算式に起因する上方バイアスの問題は、すでに述べた代替バイアスの問題の 1 つである。ボスキンレポートでは、最良指数の算式による指数を真の生計費指数とみなし、その算式による試算値と現行の Laspeyres 型の CPI の値と比較した結果、現行 CPI が年率 0.15 ポイント高いとして、これを指数算式に起因する上方バイアスの推計値としている。この点について BLS は、指数算式の性格上このような差が生じるのは当然のこととして、試算結果を妥当であると認めている。

ボスキンレポートでは、指数算式に関するもう 1 つの勧告として、品目レベルでの合算に幾何平均を用いるべきと指摘している。(勧告④) この点について BLS は、同じ層(品目)の中で商品間に代替性がある場合には、現行の算術平均よりも幾何平均のほうが生計費指数のより良い近似となっているとしている。そして、1997 年 4 月分から幾何平均の算式による試算値を試験的に公表し、1999 年 1 月分からは正式な指数として公表している。BLS の推計によると、幾何平均の採用により CPI 上昇率は年率 0.2 ポイント減少するとの結果が得られている。

2.4.2 その他のバイアス

ボスキンレポートでは、代替バイアス以外にも、既存の各種研究結果を基にして各バイアス要因の大きさを推計しており、それらについて BLS は見解を述べている。

まず、新店舗バイアスについては、BLS の Reinsdorf がかつて行った研究結果と、Lebow ほかによるその延長推計の結果に基づいている。BLS は、この推計の問題点として、①Reinsdorf の研究はわずか 2 年間という特定の短い期間のデータに基づいていること、②同研究の推計結果は標準誤差が大きいこと、新店舗バイアスはゼロである可能性もあること、③同研究の対象となった食品とガソリンの状況が他の品目にも妥当するかどうか疑問であること、を指摘している。そして、このバイアスの推計結果は信頼性が低く、むしろ他の要因を考慮すると、この結果は過大となっていると指摘している。

次に、品質・新製品バイアスについて、ボスキンレポートでは年率 0.6 ポイントと、上方バイアス全体の推計値の半分以上を占めているとしている。これに対して BLS は、まず、すでに CPI には品質変化に対する調整方法をかなり適用して効果が上がっていること、また、新製品を CPI に反映させる手法を確立していることを述べている。その上で、ボスキンレポートにおける品質・新製品バイアスの推計の妥当性についてコメントしている。同レポートでは、CPI の品目を 27 分野に分けてバイアスを推計しているが、BLS は、①これらの推計は裏付けとなるデータがないこと、②推計方法が具体的に示されておらず、多くの場合には直観的判断 (judgement) に基づいていること、③必ずしも代表性のない偏ったデータを基に推計していること、などの理由から、同レポートの推計結果は信頼できないと指摘している。また、BLS は、2 分野について自ら行った推計の方法と結果を紹介し、これらの分野については同レポートの推計結果が過大であるとしている。さらに、品質・新製品バイアスに関して同レポートに特段の提言がなされていないことから、BLS は、この問題への対処方法についてエコノミストの間でコンセンサスがないと考えられると指摘している。

なお、新製品バイアスに関連して、BLS は、1999 年 1 月から、品目のサンプリングローテーションを改定し、従来は調査区ごとに見て 5 年に 1 回しか新しい品目の設定が行われなかったところを、一部のカテゴリーについてはより高頻度で行うようにした。これによって、CPI には従来以上に新製品が反映されるようになってきているという。

3 上位代替バイアス

代替バイアスには、品目（同一の類とみなされる商品のグループ）の間での代替効果が反映されないために発生するバイアスと、品目内の銘柄（ブランドや規格等の詳細まで規定した個別の商品）の間の代替効果が反映されないためのバイアスがあるが、本節では、ボスキンレポートをめぐる議論の中で、上位代替バイアス (upper-level bias) とも呼ばれる前者に関する議論を紹介することとする。同バイアスは指数算式の問題でもあるので、同レポートの勧告でも取り上げられた生計費指数に関する議論にも触れることとする。また、これらの議論に対する理解を深めるため、指数理論に詳しい専門家の最近の見方や考え方も併せて紹介することとする。

3.1 ボスキンレポートの主張

3.1.1 生計費指数の作成

ボスキンレポートでは、勧告の 1 番目の項目として、消費者物価を測定することの目標は、生計費指数を作成することとすべきであるとしている。一般に消費者は、あ

る商品の価格が上がった場合に、その商品の代わりに相対的に低価格の商品を求めることによって支出パターンを変える。ボスキンレポートの主張は、指数作成において相対価格の変動の結果起こる代替を考慮に入れるべきであるということである。

3.1.2 最良指数の採用

ボスキンレポートでは、最良指数は生計費指数を十分近似したものであるので、現行の固定ウェイトの指数算式に代えて最良指数 (superlative index) を用いるべきであるとしている。

最良指数の概念はカナダの経済学者 Erwin W. Diewert[54]が定義したものである。最良指数の計算においては、基準時と比較時の価格と数量のデータが必要である。

最良指数の計算方法はいくつかあるが、ボスキンレポートでは、その代表例として2種類のもので紹介されている。1つは、基準時ウェイトによる指数と比較時ウェイトによる指数の幾何平均を用いるものであり、Fisherの理想指数と呼ばれる。この算式は、商務省経済分析局の国民所得・生産勘定の計算に用いられているという。もう1つの方法は、基準時と比較時のウェイトの平均値を使って価格比の加重幾何平均を用いるものであり、Törnqvist 指数と呼ばれる。

この2つの最良指数の算式は、消費者が合理的な行動をとるという前提の下では、固定バスケットに基づく算式よりも生計費指数をより良く近似したものとされている。

このような背景を踏まえ、ボスキンレポートでは、より具体的には、年次指数については trailing Törnqvist 指数に移行すべきであるとしている。これは、加重幾何平均の算式によるものであり、ウェイトとして、データの利用可能な直近の2ないし3年の平均の支出構成比を使用するものであり、このウェイトは毎年更新される。これを実現するために、BLSは、従来のように10年に1度ウェイトの改定を行うのではなく、より頻繁に改定を行うことが必要となる。

3.1.3 上位代替バイアスの評価

ボスキンレポートでは、CPIの上位代替バイアスを評価するために、代替バイアスが小さいと考えられる Törnqvist 指数を基準としてCPIを比較した結果、CPIには年率0.15ポイントの上方バイアスがあると結論付けている。この評価に当たって、CPI諮問委員会が既存のいくつかの研究結果を調べたところ、多くの結果は0.2ないし0.25ポイントの上方バイアスを示しているとのことである。しかし、BLS自らの研究における試算値があり、これによると、1988年から1995年間の平均的な差は約0.15ポイントとの結果が得られているので、これを採用したとのことである。

3.2 BLSの主張

3.2.1 生計費指数の作成について

消費者物価の測定は生計費指数の作成を目的とすべきであるとのボスキンレポートの主張に関して、BLSは、CPI作成の目的には様々なものがあり、その中でも最も重要なものの1つが生計費指数の作成であるとの立場を取っており、ボスキンレポートの考えに一応は賛同している。しかし、BLSは、仮に毎月公表できるような真の生計費指数の作成方法が存在するならば、それを採用したいところであるが、生計費指数というものは理論上の概念であり、すべての目的に合った、実用に耐える明解な単一の指数算式は現存しないとしている。より具体的に言えば、生計費指数とは、マーケットバスケットに捉われることなく、異なる時点において個々の消費者が一定の生活水準を維持するのに必要な費用を比較するものであるため、実際の指数作成に当たっては、例えば、対象を都市の住民に限定するなど、追加的な想定が必要である。仮に理論的に優れていると考えられる生計費指数の算式が存在するとしても、十分に検証されていない新たな算式に性急に移行するとすれば、統計体系への信頼感が失われるなど、得るものよりも失うもののほうが大きいとして、慎重な姿勢を示している。

BLSは、今後も最良指数の算式による実験的な指数の試算や研究を行うとともに、公表している指数の概念や理論的前提など、指数の利用における留意事項などの情報を提供していくこととしている。

3.2.2 最良指数の採用について

CPIの計算には最良指数の算式を使うべきであるというボスキンレポートの主張に対して、BLSは、最良指数を作成するには比較時のマーケットバスケットの推定が必要となるが、消費者の支出に関するデータの収集及び加工には時間がかかることから、指数の公表に遅れが出るとしている。

また、最良指数のウエイトに必要な支出データは、年次系列でしか十分な精度を得られず、月次指数は作成できないとしている。

最良指数を迅速に作成するための方法として、近似的な算式を使う方法も考えられるが、仮に比較時のウエイトを用いたTörnqvist型の近似式を用いると、生計費の増加を系統的に過小評価してしまうとしている。このような近似式としては、他にもCES (Constant Elasticity of Substitution) 指数に基づく方法などが提案されているが、この方法では、時期によっては最良指数を過大又は過小評価することがあり、その場合には事後的に改定値を公表すべきかどうかという問題が生じる。

CES指数による近似については次のような問題もある。品目の分類の中には、リンゴ、バナナ、オレンジなどの生鮮果物のように、相互に代替可能性の高い類もあれば、

外科医、歯科医、眼科医などの医療サービスのよう、相互代替可能性がほとんどない類もある。CES 指数は全項目に等しい単一の代替パラメータを用いているので、消費者による代替の行動が類の間で異なると考えられる条件下でこのような算式が有効といえるのかどうか研究を重ねる必要があるとしている。

さらに、このような近似的な算式に基づく指数については、データの利用者にその意味をわかりやすく説明することが困難と考えられる。このような理由から、BLS は、公式の CPI に最良指数又はその近似式を拙速に導入することに否定的である。

3.2.3 上位代替バイアスの評価について

BLS は、現在の CPI では、計算の基礎となるマーケットバスケットを基準時におけるものに固定しているため、比較時点では相対的に高価になって消費が低下したような商品のウェイトが最新の実態よりも高目となることによって代替バイアスが発生すること、また、最良指数では、消費支出の変化に合わせた調整がおおむね行われるので、この種のバイアスを避けることができることは認めている。年間 0.15 ポイントとされた上位代替バイアスは、最良指数と Laspeyres 指数を比較した BLS の研究結果を採用したものであり、BLS もこの種のバイアスの大きさと性格については同意している。

なお、2002 年から BLS は CPI の上位代替バイアスの大きさを推定するための指標として、Törnqvist 型の月次連鎖指数を公表した。

3.3 指数算式に関する理論と見解

上記のように、CPI 諮問委員会と BLS との間には指数算式について様々な議論があるが、指数算式の問題は米国だけの問題ではない。実際、指数算式とその背景にある生計費指数については、これまでも多くの国や国際機関で研究、検討、議論が重ねられてきており、委員会と BLS の見解について理解を深めるためにも、同分野の専門家の最近の見方や考え方を見ておくことは有意義であると考えられる。そこで、これらのいくつかを紹介することとする。

3.3.1 生計費指数の背景

生計費指数の概念は 1920 年代の Konüs[55] 以来の長い歴史があるが、CPI の基本概念が生計費指数の理念によることについては、統計作成機関の間に国際的合意は形成されていない。米国、オランダ、スウェーデンなどいくつかの国では生計費指数の理念が実際に CPI の計算を行うための概念的枠組みとされている一方で、オーストラリアや英国では生計費指数は CPI の枠組みとして位置付けられていない。生計費指数の

枠組みは、欧州消費者物価調和指数 HICP (Harmonized Indexes of Consumer Prices) にも導入されていない。他に、カナダのように態度が明確でない国もある。生計費指数の対極には、CPI は「純粋な価格変化」を捉えるための指数という考え方がある。日本の CPI は、総務省統計局の資料に「世帯が購入する商品とサービスの種類や購入数量の変化に伴う世帯の生活費の変化を測定するものではない。」との記述にあるとおり、生計費指数とは位置付けられていない。

3.3.2 生計費指数の問題点

CPI を研究する学者たちの間でも CPI の基本概念を生計費指数とすることについて、賛否両論がある。

これについては、1999 年に開催された Ottawa Group 会合という CPI に関する有識者の非公式な場においても「異なる目的にかなう物価指数の異なる概念」と題して、当分野の研究者や専門家の間で様々な議論が行われた。これらの議論の素材となる論文から生計費指数を取り巻く論点が見出されたのでいくつか紹介することとする。

元 ILO 統計局長で、1989 年版 CPI マニュアルの筆者である Ralph Turvey^[56]は、公式の CPI を生計費指数とすることに対して、実用性の観点から反対の立場をとっている。彼の考えでは、生計費指数の概念は、消費者が従前と同じ生活水準を維持するために必要な費用という一意的に決定することの難しい仮想的な概念に基づいているため、これを実用的な指数として設計し、作成していくことには無理があるとしている。生計費指数を近似するために CPI を作成すると、帰属計算を行う項目で不都合が生じる⁵⁰などの問題があり、むしろ、実用的な観点からは、CPI を単純に基準時における消費バスケットをウェイトにした支出の変化と定義するほうが良いとしている。Turvey がこのような考え方をとった理由をいくつか挙げると次のとおりである。

- ・ 生計費指数では消費者の嗜好が複数の消費者で同じで変わらないという仮定を置いているが、消費者の嗜好は多様で、かつ、常に変化しており、この仮定には無理がある。
- ・ 同理論では個々の消費者は独立であると仮定しているが、この仮定も成り立たない。例えば、新製品が市場に現れた場合、最初は売れ行きは遅く、その後に加速し、最後に飽和状態に達するというように、時間とともに売れ行きは変化する。また、価格とウェイトの基準期間は一致しない場合もあり⁵¹、そのような場合、ウェイトの

⁵⁰ 例えば、帰属家賃の価格が変化すると帰属家賃の一部は収入でもあるため、価格変化とともに収入も変化することになり、価格変化の前と条件が異なる。

⁵¹ 日本では価格とウェイトの基準期間が一致しているが、欧米では一致していないことが普通である。これは経常的な家計調査が実施されていないためである。

基準期間中に消費した数量は価格の基準期間における価格で消費者の効用を最大にするとは言えない。

- ・ 消費者が購入する前に期待していた効用と、実際に購入した後の効用は異なり、どちらが生計費指数において適切であるのか明確でない。
- ・ 最良指数によって生計費を測定するのが可能となるのは、価格体系以外のすべての条件が同じ場合に限られる。しかし、通常、CPIのウェイトの元となる消費を行っているのは消費者の集団であり、その中の人は死んだり、成長したり、退職したりするため、構成は常に変化しているし、ファッションや収入が変化して嗜好が変化するなど、消費を取り巻く条件が同じことはあり得ない。代替効果は支出パターンの変化に影響を及ぼす要因の1つでしかない。
- ・ ある消費者が効用を維持するための支出額は、他の消費者の消費行動の変化に依存しているので、すべての消費者を単純に合算するような計算方法を採用できない。

このような Turvey の見方に対して、最良指数の理論を展開してきた Diewert[57] は、どちらかという生計費指数の概念の導入に肯定的であり、Turvey の主張は指数理論に対する経済的アプローチを軽視しているとした上で、Turvey が指摘した問題点にコメントをしている。以下では、特に、指数算式の問題に関係の深い箇所についてのみ簡単に触れることとする。

まず、基準時と比較時の間に人が年齢を重ねたり出生死亡で人間の構成が変わったりすることによって、人口集団の選好が変化することは確かだが、その変化は連続的に徐々に起こるものであり、それによる影響は代替効果の大きさに比べて小さいと考えている。また、すべての消費者に当てはまるただ1つの指数が存在しないという問題は、理論そのものの問題ではなく理論の応用の問題であり、収入に帰属家賃や富の純増を加味して計算するか、あるいは人口属性によって複数の生計費指数を作成することによって対処すべきであるとしている。

3.3.3 各種指数の算式の比較

各種の指数算式の特徴については次のような見解がある。

まず、Turvey[56]によると、最良指数は、比較時点におけるウェイトが必要だが、データ収集に時間がかかるために最新情報をリアルタイムに得るのが難しく、したがって、月次系列には適用できず、また、個々の消費者にとってどの指数算式が最も実用的かの答えが導かれずとしている。

一方 Diewert[58]は、Fisher の理想指数と Törnqvist 指数は伝統的な固定ウェイトの Paasche 指数や Laspeyres 指数よりも理論的には優れた生計費指数であるとしている。

る。また、最良指数は、2次項までのテイラー近似で近い値をとることから、理論的には同程度に優れているとしている。

また、商務省エコノミストの経歴を持つ Jack Triplett[59]は、両指数を比較すると、実用的には、Fisher の理想指数のほうが幾分計算しやすいとしている。Fisher の理想指数には解釈もしやすいという利点もあり、また、その内訳である Laspeyres 指数と Paasche 指数に分解することにより指数の動きを機械的に理解することができ、このような分解は価格と数量の分析にも資するとしている。さらに、Fisher の理想指数は要素逆転テストを満たすが、Törnqvist 指数は要素逆転テストを満たさないことも指摘している。

他方、Jörgen Dalén[60]は、Fisher の理想指数と Törnqvist 指数は月次系列では使用できないが、年次系列として連鎖的に使用することができ、Laspeyres 指数のようにより単純な式を評価するものさしでもあるとしている。また、最新時点のウェイトがわからなくても、2つの商品間に交差価格弾力性が0に近ければ、固定ウェイト算術平均が生計費指数の近似値となり、交差価格弾力性が大きければ、固定ウェイト幾何平均が生計費指数の近似値となるとしている。さらに、指数理論家の間で共通認識が得られているような単一の算式体系や、すべての利点を備えた算式は存在しないので、各指数算式はトレードオフの状況にあり、これらの有用性については次のような見解があるとしている。

- ・ ある指数算式が形式的なテストで条件を満たさない場合でも、実際にはその式は近似的にみて問題にならない場合もある。このことは最良指数に多く見られるが、十分研究されているとは言えない。
- ・ 実用上の重要な基準は合算の一貫性である。Törnqvist 指数は、合算の一貫性を満たさないが、同指数のウェイトは合計が1になるように分解することができ、これらを使って指数を段階的に合算することができる。一方、Fisher の理想指数は明確に定義されたウェイトで階層的に積み上げ計算することができない。このような視点から見れば、ウェイトに部分指数を掛けてから加える Edgeworth 指数は「ほとんど最良」の指数の中では最も優れているといえる。
- ・ 他に実用上重要な評価基準としてゼロ価格の許容性がある。Törnqvist 指数を含め、幾何平均指数はすべて比較時点のいずれかで価格がゼロになると、ゼロか無限大になってしまう。価格が極端に変わると、対応する集計関数が2階微分可能な関数を近似するという最良指数の特徴は無意味なものとなる。

3.4 生計費指数と指数算式の留意点

生計費指数の概念は、理論的には興味深い概念であるが、現実に応用しようとする、強い制約のある概念である。例えば、この概念は、全国民の消費行動を合理的な消費行動をとる1人の消費者に十分近似できるという仮定のもとで成り立つとされる。しかし、実際には、同じサービスを購入しても大人と子供とで料金が異なる場合もあるし、テレビを購入したばかりの者とテレビが故障した者とは購入意欲が異なるなど、明らかに個人個人の嗜好や消費行動は異なるので、そのような仮定には少なくとも部分的には無理がある。

また、最良指数には実用上はいくつか問題があり、本研究で示した議論や論文においてはとりわけ次の2点が強調されていた。1つは最新時点の情報を素早く取り込むことが困難なことであり、最良指数は年次系列では有用でも、月次系列には適用できない理由の1つとなっている。もう1つは、生計費の概念が広く多岐にわたることから、それを的確に、かつ実用的に数量化することが困難なことである。特に、月次系列については、価格に季節変動がある場合に最良指数は必ずしも適当とは言えない⁵²など、問題点が多い。BLSが生計費指数の概念を重視しつつもLaspeyres指数でCPIを公表しているのはこれらの問題を回避できないからである。このような問題に対しては様々な検討が重ねられているが、打開策を見出すまでにはかなりの時間を要すると考えられる。

ボスキンレポートでは、生計費指数の概念に基づいたCPIを求めており、生計費指数を最もよく近似する指数算式という観点から、最良指数が最も望ましいとしている。これに対してBLSは、CPIが生計費指数を目指すことについては一面同意しつつも、生計費という概念が主観的であることや、最良指数が前提としている条件が現実には成立していないことなどの問題点を指摘し、実用的な観点から、公式のCPIとして最良指数の算式を採用することについては慎重な態度をとっている。

同様に、Ottawa Groupにおける指数算式をめぐる議論でも、理論面を優先するか、実用面を優先するかといった立場の違いによって見解の相違が生じている。実際、ボスキンレポートと同様の立場をとる者は、Diewertなど理論家やエコノミスト系統の人に多く、逆にBLSと同様の立場をとる者は、Turveyのように政府統計の作成実務に深い関わりを持つ統計家に多い。

⁵² 例えば、旅行のオンシーズンには、旅行費は高くなるが旅行者数も多くなる。基準時と比較時の季節が異なる場合、Laspeyres指数よりもPaasche指数が高くなる逆転現象が生じ得る。また、季節的に出回りが無くなる品目の扱いも理論上明確でない。

4 下位代替バイアス

本節では、代替バイアスのうち、品目内の銘柄（ブランドや規格等の詳細まで規定した個別の商品）の間の代替効果が CPI に反映されないことに起因するバイアス（下位代替バイアス）に関する議論を紹介することとする。

4.1 ボスキンレポートの主張

ボスキンレポートの勧告では、下位集計レベル（品目レベル）の指数の計算には幾何平均を用いるべきであるとしている。また、幾何平均を採用することは、上位集計レベルにおける Törnqvist 指数への移行と併せて、生計費の測定という米国 CPI の目指すべき本来の役割に向けた改善であるとしている。

ボスキンレポートでは幾何平均で計算した指数の利点をいくつか強調している。その1つは、算術平均指数は時点逆転テストを満たさないが、幾何平均指数はそれを満たすことである。2つ目は、幾何平均指数は、基準時と比較時における価格の相関係数が小さければ、生計費指数に対してほぼ不偏であることである。幾何平均指数は、このような性質を持つために、下位集計レベルにおける様々な手法を比較するための基準として活用されるとしている。

さらに、カナダ統計局を始め、指数の計算方法を算術平均から幾何平均に移行する統計作成機関は増えていると述べている。

同レポートでは、このような見解を示した上で、下位集計レベルにおいて、幾何平均指数を真の生計費指数とみなし、算術平均で計算した現行 CPI には下位代替バイアスと呼ばれる年率 0.25 ポイントの上方バイアスがあると推計している。

この推計は Brent R. Moulton と Karin E. Smedley の研究に基づいている。それによると、1992年6月から1992年12月までの期間、住居費を除く支出について算術平均指数と幾何平均指数を計算すると、前者が 0.49 ポイント高かったという。しかし、その後、BLS は 1995年1月と1996年6月と2度にわたり指数の計算方法を改定した。これらの改定によってバイアスが 0.24 ポイント縮小したと評価されることから、幾何平均指数に対する上方バイアスは 0.25 ポイントと推計された。

なお、ここで BLS による計算方法の改定を手短に紹介しておく。1995年1月には、食料品（外食を除く）に調整指数（seasoned index）が採用されたほか、持ち家の帰属家賃と処方薬についての計算方法が改善された。（バイアスが年率 0.14 ポイント縮小）また、1996年6月から調整指数の適用が食料品以外の品目（住居費を除く）に拡大された。（バイアスが年率 0.10 ポイント縮小）調整指数についての詳しい説明は省略するが、代替商品の基準時の価格を推計するに際して、従来の方法を改良したものが調整指数と呼ばれている。

4.2 BLS の主張

このようなボスキンレポートの主張に対して、BLS は、同じ層（品目）の中で商品間に代替性があるという条件のもとで、現行の算術平均指数よりも幾何平均指数のほうが生計費指数のより良い近似となっていると、条件付きながら賛同している。しかし、一方で、BLS はこの条件を重視しており、幾何平均指数にバイアスが存在しないのは代替弾力性が1の場合に限られ、仮に代替弾力性が0ならば、算術平均で計算した調整指数のほうが価格変化を正確に表し、幾何平均指数にはむしろ下方バイアスが発生するとしている。ここで代替弾力性とは、相対価格に対する相対需要量の弾力性の意味である。

また、ボスキンレポートにおいて、幾何平均指数は、基準時と比較時における価格の相関係数が小さければ、生計費指数に対してほぼ不偏であるとしていることについて、BLS は、それが成立するには、効用関数が代替弾力性一定を前提としたものとなること、価格が定常分布に従っていることなどの条件が前提となるが、そのことが述べられていないとしている。その他、BLS は、品目内のすべての商品の価格が同一方向にシフトしているという条件も前提条件として必要だが、そもそも異質な商品が多い品目ではこのような現象はあり得ないし、同質商品からなる品目でも商品の生産工程の違いにより価格変化の方向は違ってくるとしている⁵³。

BLS は、仮にこれらの前提がすべて満たされるのであれば、1995年から導入した調整指数でも、代替弾力性の大きさにかかわらず生計費指数の推計値になるとしている。

他方、ボスキンレポートで指摘された下位代替バイアスの大きさについては、BLS は大きすぎるとし、幾何平均を採用すればCPI上昇率は年率約0.2ポイント下がるとの計算結果を得たとしている。

4.3 幾何平均指数の作成状況

CPIの基本系列の計算に幾何平均を用いる可能性については、BLSはCPI諮問委員会発足以前の1993年にすでに研究し始めている。そして、BLSは1997年4月からCPIの全品目について幾何平均で計算した試験的指数を作成した。その後、1999年1月からは総人口の87%を占める都市消費者を対象とした指数（CPI-U）と総人口の

⁵³ 米国では、CPIの品目数が少なく、品目内の商品の価格が異なる場合をかなり許容している面がある。これに対し、他の国では、品目数を多くし、品目内の商品の価格の動きが同じ傾向になるようにしている点に留意が必要である。

32%を占める都市労働者を対象とした指数（CPI-W）の計算において幾何平均の算式を採用することとした。

ただし、これらの指数の計算には、すべての品目に幾何平均が用いられているのではなく、いくつかの品目では従来どおり算術平均が用いられている。

その1つは住宅サービスに関するいくつかの品目である。BLSによると、これは、住宅ストックの総量が大きく変化しないので消費者が家賃の変化に応じて住宅を自由に変えることができないからである。2つ目は公共サービスに関するいくつかの品目である。BLSは、公共サービスの多くは本来政府と独占会社が供給していることに加え、多くの場合、消費者は住所を変えなければ供給者を代替できないからの理由を述べている。また、これまでの研究結果によると、少なくとも短期間では代替弾力性が低いとしている。

3つ目は医療サービスに関するいくつかの品目である。やはり、病院及び専門医療サービスにおいて代替弾力性が低いとしている。

その他、輸送、登録、免許、自動車税などでも継続して算術平均が用いられている。

これらを除くと、幾何平均を採用している品目は、CPI-Uにおいては消費支出全体の61%になる。

BLSは、幾何平均は一定の条件下において優れているとしながらも、結局はCPIに採用したことについて、次のように説明している。BLSは、1993年以来、消費者の代替行動を様々なデータにより分析してきたが、一部の品目では明らかな代替行動が見られたものの、データの制約などから代替行動の存否や大きさについて決定的な結論は得られなかった。しかし、幾何平均の算式を採用すれば、多くの品目について調整指数のような複雑な手法を考慮する必要のないことや、幾何平均指数には算術平均指数にはない良い性質があることなどから、代替弾力性の低いと考えられる一部の品目を除き、この算式を導入した。

4.4 下位集計レベルの指数算式の採用条件

ボスキンレポートが主張する米国CPIの品目内の集計における幾何平均の採用に対して、BLSは代替弾力性が高いとみられる品目に限定はしているものの、基本的には移行に踏み切ったとみなせる。ボスキンレポートでも指摘しているとおり、品目内の集計における算式を算術平均から幾何平均に移行している国が目立つが、BLSもこのような世界的な動向に従ったものと見ることができる。

米国における上記の一連の議論の中では、品目内の集計において幾何平均を採用することはCPIを生計費指数に近づける立場で議論されているが、別の見方からすると、この議論はCPIを生計費指数とするか否かという議論とは必ずしも関連するもので

はない。公式の CPI を生計費指数とすることに反対の立場をとっている Turvey も、品目内の集計における幾何平均の採用については、生計費指数の概念とは切り離して検討すべきであるとしている。実際、近年、米国の他にもカナダ、フランス、オーストラリアなどのいくつかの国では生計費指数の議論とは別に、品目内の集計における指数算式を一部幾何平均に移行している。

しかし、品目内の集計における指数算式の妥当性は、代替弾力性の大きさだけでなく、各商品の価格変化の方向が同じか異なるか、店舗や銘柄の選定方法及び価格の飛跳ね（price bouncing）⁵⁴の有無などに関連しているとみられる。品目内の各商品の価格変化が同じとみなせるのであれば、代替弾力性の大きさに関係なく、算術平均でも問題がない可能性があり、品目を適切に区分しているならば、むしろ他の要因の方が重要になる。OECD の Carl Obst[61]は、同質性が高く、価格競争が激しく、特売が頻繁に行われる場合には、調和平均が適切であり、幾何平均でも上方バイアスが生じると述べている。これは、価格の飛跳ねの存在がバイアスや指数算式に関係することを意味している。CPI マニュアルと呼ばれる ILO, IMF, OECD, UNECE, Eurostat, The World Bank[62]では、幾何平均の持つ一般的な注意事項として、調査データの中に 0 があれば、他のデータの値が何であれ、その幾何平均は 0 となり、極度の物価下落に敏感で、それを使うときには、個別の価格比に仮に 10 及び 0.1 というような上方及び下方修正を施すことが必要になると記載している。

幾何平均に移行した国では、それぞれの国の事情があるとみられる。例えば、米国の場合、代替弾力性の大きさを判断材料としたというよりも、①店舗によって調査銘柄が異なり、同質とはみなせないことから、従前は、店舗ごとに価格比をとってから算術平均していたことでバイアスが生じていたこと、②今後、総合指数を連鎖指数に近い方式に移行するためドリフトの問題があることなどにより、代替弾力性がある程度あるとみられる品目について幾何平均に移行したとみられる。また、英国は、銘柄の選定は非確率的であるが、店舗を売上比例確率抽出していることなどの理由で幾何平均を採用しているようである。スイスでは、2000 年改訂で幾何平均に移行することを表明しているが、同国では、①調査銘柄の規定が緩く、各店舗で実際に価格を調査している商品を同質とはみなせないために、価格比の算術平均を指数算式として採用していたこと、②さらに、連鎖指数方式に移行するため、指数算式を変更しないとドリフトが生じる可能性があることを理由に挙げている。

なお、HICP の作成方法を検討する過程においても、下位集計レベルの指数算式について検討が行われ、価格比の算術平均は上方バイアスが発生するおそれがあるとし

⁵⁴ 各店舗がバラバラに特売を繰り返すことなどで価格が上下する現象

て不適當とされたが、価格の算術平均の比と価格の幾何平均の比は優劣が判定できないため、いずれを採用しても良いとされている。

下位集計レベルの指数算式については、まだ研究すべき余地が残されているとみられ、Obst[61]が述べているように、POS データの利用などによる研究の進展が望まれる。

5 新店舗バイアス

本節では、消費者がディスカウントストアなどの安売り店にシフトする行動が反映されないことによるバイアス（新店舗バイアス）に関する議論を紹介する。

5.1 ボスキンレポートの主張

ボスキンレポートでは、近年消費者の購入店舗が変化し、スーパーマーケットやディスカウントストアなどのシェアが拡大しているにもかかわらず、CPI には特にディスカウントストアなどの安売り店にシフトする消費者の行動が適切に反映されていないとしている。

同レポートでは、このために発生するバイアスを新店舗バイアスと呼び、CPI には年率 0.1 ポイントの上方バイアスがあると推計している。

この推計は、BLS の Marshall B. Reinsdorf が行った研究に基づいている。その研究では、CPI の価格調査において調査店舗の交代によって調査対象から外れる店舗と新たに調査対象となる店舗との間で価格水準を比較した。その結果によると、食料及びガソリンのいくつかの品目について、その差は年率 0.25 ポイントであった。また、David E. Lebow、John M. Roberts 及び David J. Stockton の研究によると、新店舗バイアスによって影響を受ける品目は CPI 対象品目の消費量全体の 40%に相当すると推計している。この 2つの結果から、ボスキンレポートでは、新店舗バイアスを年率 0.1 ポイントと推計した。

そして、これを改善するには、毎年消費者の行動をよりの確に把握して新店舗を速やかに取り込むことが必要であり、このためには、現在センサス局が収集している商品レベルの出荷データを BLS が使用すべきであるとしている。

米国 CPI における店舗の抽出確率は、POPS（Point of Purchase Survey、購入地点調査）をもとに、地域ごと、品目ごとに定められており、支出総額に占める店舗のシェアに比例するように決定される。この方法によると、店舗抽出時には消費者の購買先を反映した調査が行われることになるが、その後で、例えば、ディスカウントストアなどが登場して相対的に支出総額を増やしても、それを反映させることはできない。

なお、店舗の抽出は、米国では5年に1回のローテーション方式で行われており、毎年全体の約20%に店舗の抽出替えがあり、したがって、新店舗が抽出される可能性は5年に1回に過ぎないとしている。

5.2 BLSの主張

このようなボスキンレポートの主張に対して、BLSは、新店舗バイアスが生じるおそれがあることは認めているものの、ボスキンレポートの推計結果は信頼性が低く、過大となっていると指摘している。

BLSは、その根拠としてReinsdorfの研究に関し次のような制約を指摘している。①同研究は1980年代後半のわずか2年間という特定の短い期間のデータに基づいており、この期間のデータが長期的傾向を適切に代表しているとは言えない。②同研究の推計結果は標準誤差が大きいため、新店舗バイアスはゼロである可能性もある。③同研究の対象となった食品とガソリンの状況が他の品目にも妥当するかどうか疑問である。

BLSはさらに、Reinsdorfの研究を解釈する前提にも問題があることを指摘している。その1つは、この推計では、低価格で販売している店舗は、より高い価格で販売している他の店舗と同じ品質のサービスを提供しているとの前提に立っていることが指摘されている。ディスカウントストアなど低価格販売の店舗は、店員の商品知識、製品の多様性、品質保証、立地条件などサービス面でのコスト削減によって低価格を実現している場合が多い。仮にサービス水準の低い安売り店が登場した場合、既存店においても競合によりサービス水準が低下し、結果として価格も低下することも考えられるが、その場合に品質調整もしないでそのまま指数を計算すると、上方のバイアスではなく、下方バイアスが発生するとしている。したがって、0.25ポイントという数字は新店舗バイアスの上限と見るべきであり、Reinsdorf自身もそう解釈していることを指摘している。

もう1つの問題として、Reinsdorfの研究に見られる価格低下は必ずしも低価格店舗の参入だけが原因ではないことが指摘されている。BLSは1978年以来、店舗の更新と同時に調査対象となる銘柄も併せて更新してきた。消費者の選択が安価な商品にシフトしたのであれば、そのような選択はCPIの調査対象銘柄に反映され、その結果、新たに調査対象となった銘柄の平均価格は下がることになる。したがって、Reinsdorfの研究で得られた新旧店舗の価格差は、新しい店舗の出現の影響だけではなく、銘柄の交代の影響も受けているので、新調査店舗のほうが旧調査店舗よりも価格水準が低いからと言って、それがすべてディスカウントストアなどの出現による影響とみなすべきではないと指摘している。BLSは、さらに、Reinsdorfのより最近の研究結果を

紹介して、新店舗バイアスの数字は年率 0.04 ポイント程度としている。これによると、製品の銘柄の更新がほとんど行われなかった時代である 1948 年から 1963 年までと 1967 年から 1976 年までにおいて食料品の平均価格の推移を CPI と比較すると、年率 0.1 ポイント程度の差だったという。当時の BLS は、店舗を変更しても調査対象となる銘柄は基本的には変更しなかったため、新店舗の登場による価格差としては、この数字のほうが、ボスキンレポートで使用された年率 0.25 ポイントよりも適切と考えられると指摘している。仮にこの数字を前提とすれば、CPI 全体における新店舗バイアスはその約 40%、すなわち、年率 0.04 ポイントとごくわずかになる。

BLS は、新店舗バイアスについてこのように反論しつつも、POPS を CATI (Computer Assisted Telephone Interviewing) 方式により効率化し、多頻度で実施することが可能となったことを受け、店舗交代の早期化を図ることとした。これにより 1999 年 1 月から店舗の抽出方法を変更し、調査品目を 16 のグループに分け、グループごとに定められた周期で全国一斉に店舗交替を行い、一部のグループについてはより高頻度で新店舗の設定を行えるようにした。

なお、ボスキンレポートで指摘されている、「センサス局が収集している商品レベルの出荷データを使用すべき」という点については、BLS の見解や対応方針の記述は見当たらない。

5.3 新店舗の価格反映の留意点

新店舗バイアスは調査手法に依存している面が大きい。ボスキンレポートではこの大きさを年率 0.1 ポイントと他の誤差要因よりも小さく見込んでおり、しかもこの推計の信頼度はあまり高くないようである。BLS は、この推計は過大であると指摘した上で、さらに、最近の改定で一部店舗の抽出替えの頻度を上げたことから新店舗バイアスは大きな問題ではないと考えているようである。

昨今のいわゆる「価格破壊」と呼ばれる動きの中で、ディスカウントストアなどの新しい店舗の価格動向を適切に CPI に反映することは重要な課題であろう。しかし、これらを適切に CPI に反映するにはいくつかの注意が必要である。まず、新しい店舗が相対的に高いシェアを有しているか否かを確認することが必要なのは当然であろう。また、新規店舗の扱う商品が在来店舗の商品と同等とみなせるかどうか、確認が必要である。例えば、衣料品のようにファッション性やブランドなどの違いが重要な品質差と認識されるなど、異なる店舗では取扱商品の品質が微妙に異なる場合もある。このような場合に価格低下が生じているとすれば、それはディスカウントストアの出現という要素だけではなく、品質変化に伴う要素も含まれていることに留意すべきであろう。仮に商品自体の品質が店舗の間で差がないとした場合に、店舗の差による価格

差をどこまでが店舗の提供する付加的サービスの品質の違いによるものとするかは見解が分かれるところである。例えば、コンビニエンスストアとディスカウントストアとでは、利便性などの観点からすると後者のほうがサービスを省いていると考えられ、それに伴って価格も後者が安いのであろう。しかし、都心と郊外の同規模のスーパーとで価格に違いがある場合、場所の違いを品質と見るかどうかは、個々の消費者の機会費用を考えた上で判断しなければならない。また、都心部のスーパーと郊外のスーパーを比較する場合、地域全体の消費者の平均的評価として考察する必要がある。立地条件が良い店舗は、より多くの消費者にとって利便性が高いという意味で、提供するサービスの質が高いと見る必要がある。このように、店舗間の価格差を品質差と見ようとする、何を以て品質とみなすかという問題に帰着する。品質としてどのような要素を考慮に入れるかという点について、さらに、議論が必要である。

また、店舗選定の議論においては、時間的変化と価格水準の問題を混同しないよう留意すべきである。仮に店舗の価格差が品質差を反映したものとみなせるならば、店舗間の価格差自体は価格水準の問題であるので、価格の時間的変化を捉えることを目的とする CPI の問題とは必ずしも言い切れない。より具体的に言えば、例えば、ディスカウントストアと一般のスーパーとでは価格水準に差があるとしても、両者の価格水準はほぼ並行的に変動しているケースが多いと考えられる。もちろん、両者の価格水準の関係が一定でないことがあれば、個々の商品ごとに価格動向は異なり得るし、価格の変動する時期にもずれがあると考えられるが、中長期的に見た動向はおおむね並行的となることは十分考えられることである。仮にこのような関係が成立しているならば、仮に新しい店舗への交代が多少遅れたとしても、CPI の動きには見かけほど大きな影響を与えないことも考えられる。いずれにしても、このような店舗間の相対的な価格動向が実際にどのようなになっているか、把握し分析して見る必要がある。

6 品質・新製品バイアスをめぐる議論

ボスキンレポートにおいて米国 CPI に存在すると指摘された 4 種類のバイアスのうち、製品（商品）の品質改良が正確に評価されないことに起因するバイアス（品質バイアス）と、新しい製品が出現しても CPI にはすぐに反映されないことによるバイアス（新製品バイアス⁵⁵）に関する議論を紹介することとする。なお、議論の中で、こ

⁵⁵ ボスキンレポートでは、新製品バイアスを新製品発売後の価格低下が指数に反映されないことによるバイアスとしているが、実際には、新製品が技術進歩などによって消費者が購入できる程度にまで価格を下げて販売することが可能になったことにより発売されたとみなし、発売前の仮想的な価格（Hicks の demand reservation price）からの低下による消費者の余剰が指数に反映されないことも新製品バイアスに含めて議論している。

これらの2つのバイアスは併せて論じられていることが多いので、ここでも両者を併せて品質・新製品バイアスと呼ぶことにする。

6.1 ボスキンレポートの主張

6.1.1 品質・新製品バイアスの評価

ボスキンレポートでは、米国 CPI には品質・新製品バイアスが存在し、1996 年において年率 0.61 ポイントの上方バイアスになるとしている。また、同バイアスが発生する理由としては、一般に、製品の多くは何年も前の同種の製品より改良されていること、新製品は比較的高い価格で導入され、製造技術の改良と販売量の増加により価格の低下と品質の向上が可能になるので、指数への新製品の組み込みが遅れると、この間の新製品の価格の動きを取り込むことができなくなることを挙げている。そして、同レポートでは、品質・新製品バイアスの大きさと方向は品目類ごとに異なり、同バイアスについては分類ごとに生計費指数の定義に沿って議論すべきであるとして CPI の作成に用いられる 206 の品目を 27 に分類し、それぞれ同バイアスの大きさを推計している。このうち、光熱、家事用品、家事サービス、他の個人交通、公共交通、医療保険、娯楽サービス、煙草の 8 分類についてはバイアスが存在しないとしている。一方、食料（農産物を除く）、家賃、新車などの 19 分類については上方バイアスが存在するとしている。

これらの概要は次のとおりであるが、後で BLS の主張の中でも説明するとおり、裏付けとなるデータや判断の根拠が明らかでないものが多い。

ア 食料及び飲料

近年、CPI には季節に限定されずに製品を選択できるようになったこと、スーパーマーケットで必要な製品を 1 度に手に入れられるようになったこと、などの品揃えと利便性の拡大が価格に反映されていないことにより上方バイアスが存在する。上方バイアスを 1967 年から 1996 年までの 30 年間で推計したところ、食料（農産物を除く）で年率 0.30 ポイント、生鮮果物・野菜で年率 0.60 ポイント、外食で年率 0.30 ポイント、アルコール飲料で年率 0.15 ポイントとなる。

イ 住居

まず、家賃については、単位面積当たりの家賃を尺度としているが、備え付けの電気器具、冷暖房施設、浴室の配管などの品質向上を看過したことにより、1976 年から 1996 年までの 21 年間で年率 0.25 ポイントの上方バイアスが存在する。光熱については、石炭及び薪からその他の燃料への移行がおおむね 1970 年代初めまでに完了していることから、特段バイアスは存在しない。一方、電話やケーブルテレビなど他の住宅設備については、電話における音質と利便性の向上が加味されていないことによ

り、年率 1.50 ポイントの上方バイアスが存在していることから、同分類における電話のウェイト (2/3) を乗じて年率 1.00 ポイントの上方バイアスが存在する。また、電気器具については、Robert J. Gordon 等が推定した個人消費支出 PCE (Personal Consumption Expenditures) デフレーターの上方バイアスを基に計算すると、1994 年から 1996 年までの 3 年間で年率 5.60 ポイントもの上方バイアスが存在する。さらに、他の家具にも年率 0.30 ポイントの上方バイアスが存在する。なお、品質変化が頻繁に起こる電気器具については、CPI のウェイトは 0.80% と小さすぎるので、これを 1995 年における PCE のウェイト (2.03%) まで引き上げるべきである。

ウ 被服

平均的な都市の消費者が購入する実用的な被服の品質変化は頻繁には起こらないとみなした上で、1965 年から 1993 年まで同一の被服に対して Robert J. Gordon が通信販売カタログから作成した指数と比較すると、1985 年から 1993 年の間に CPI には年率 1.92 ポイントの上方バイアスが存在することになるが、それをやや低めに見積もると、1985 年から 1996 年の間に年率 1.00 ポイントの上方バイアスが存在する。

エ 交通

自動車のような耐久財については、持ち家の帰属家賃と同様、その製品の価格ではなく、その利用により得られるサービスの価格に変更すべきであると勧告する。まず、新車については、CPI には自動車購入時における実質利子支払額と原価償却費からなるユーザーコストが反映されていない。前者は常に価格の 10.0% であるのに対し、後者は 1970 年から 1993 年の間に技術水準の向上により価格の 10.0% から 6.7% に減少した。これにより、ユーザーコスト全体としては、この間に 20.0% から 16.7% に年率 0.79 ポイントずつ減少した。もっとも、ユーザーコストの減少率はこの間一定ではない。Gordon の研究によると、自動車の耐久期間は 1970 年から 1983 年の間には年率 2.1 ポイント、1983 年から 1993 年の間には年率 1.3 ポイントで増加していたので、これをもとに 0.79 ポイントを区分し、1970 年から 1983 年の間には年率 0.96 ポイント、1983 年から 1993 年の間には年率 0.59 ポイントでユーザーコストは減少したと推計することができ、結局 1983 年から 1996 年の間に CPI には年率 0.59 ポイントの上方バイアスが存在する。これは 1987 年以降の中古車についても言えることである。また、自動車燃料については、1984 年以降、クレジットカードでガソリン給油機を利用できるようになったことを踏まえ、年率 0.25 ポイントの上方バイアスが存在する。

オ 医療

まず、処方薬については、Ernst R. Berndt 等の調査研究と Zvi Griliches 及び Iain Cockburn による関連調査研究を踏まえ、1970 年から 1995 年までの間は年率 3.00 ポイントの上方バイアスが存在していたが、1995 年 7 月から CPI において同等な一般

薬が出現したときの価格低下を品質低下とみなさなくなったことを受け、年率 2.00 ポイントの上方バイアスに縮小した。また、処方薬以外の薬及び医薬品には年率 1.00 ポイントの上方バイアスが存在する。さらに、診療報酬と入院費用からなる医療サービスについては、David M. Cutler 等の研究によると、1983 年から 1994 年まで心臓発作治療で患者の支払額のみから計算した指数が延命に必要な費用を含めて計算した指数と比較して、年率 4.40 ポイントの上方バイアスを有する。この結果は、Matthew D. Shapiro 及び David W. Wilcox が、1969 年から 1993 年までの間に白内障手術については、年率 4.57 ポイントの上方バイアスが存在するとした研究結果と類似している。これらを控えめに評価すると、医療サービスには年率 3.00 ポイントの上方バイアスが存在する。なお、医療についても電気器具と同様、品質変化が激しいにもかかわらず現行 CPI のウェイトが 7.36% と小さ過ぎるのは問題である。CPI の対象は、サービス利用者の支払額だけに限定されているが、使用者拠出金を含めた総医療請求額にすべきであり、国民所得生産勘定によれば 16% を占めることになる。

カ 娯楽

娯楽については、年率 15 ポイント近くにも及ぶ電子玩具の上方バイアスと、その他の玩具のほとんど 0 のバイアスを平均して上方バイアスの大きさは年率 2.00 ポイントとなる。また、身の回り品には、煙草、トイレ用品、化粧品など品質・新製品バイアスがほとんど存在しない製品がある一方で、小さな電気器具も含まれており、これには大きな電気器具と同程度のバイアス（年率 3.00 ポイント）が含まれていることから、身の回り品全体では年率 0.90 ポイントの上方バイアスが存在する。教育費を含む他の支出については、約 10 分の 1 を占める個人金融サービスにおいて、ATM や一括現金管理勘定の普及などの技術革新により、上方バイアスが年率 2.00 ポイント存在することから、全体では年率 0.20 ポイントの上方バイアスが存在する。

6.1.2 品質・新製品バイアスの縮小方策の提言

ボスキンレポートでは、品質・新製品バイアスの大きさをこのように評価した上で、これを縮小するには、まず、BLS が用いている品質調整方法のうち直接比較法、除去法、リンク法、コスト推定法の使用を制限し、BLS が被服に用いているヘドニック回帰法をテレビやパソコンなど新モデルが継続的に導入されている製品にも拡張すべきであるとしている。また、毎年店舗と製品の抽出替えをして新製品の特定と指数への取り込みを早めることも重要であると述べている。

6.2 BLSの主張

6.2.1 品質・新製品バイアスの評価の妥当性

ボスキンレポートで上方バイアス全体の推計値の半分以上を占めているとしている品質・新製品バイアスの推計結果について、BLSは、①推計方法が不明確で、多くの場合には直観的判断(judgement)に基づいていること、②前提に誤りがあること、③推計の基になったデータに必ずしも代表性がないこと、などの理由から信頼できないとしている。また、詳細は後述するが、BLSは、生鮮果物・野菜と自動車燃料の2分野について自ら行った推計の方法と結果を紹介し、これらの分野については同レポートの推計結果が過大であるとしている。

さらに、同レポートにおいてはCPIに生じる下方バイアスについて体系的に考察していない点が問題であるとしている。つまり、航空旅行や教育の質の低下などサービス水準の悪化に対する不満を聞くことは多いが、同レポートではサービス水準の改善ばかりが取り上げられていると批判している。また、価格差を過度に品質差とみなして調整すると指数に下方バイアスが生じるおそれもあるが、同レポートではほとんどこの問題に配慮していないとしている。

ここからは、このようなBLSの反論を分野ごとに見た具体的内容である。

ア 推計方法が不明確としている分野

最初に、食料(農産物を除く)、生鮮果物・野菜、外食、アルコール飲料、他の住宅設備、他の家具、自動車燃料、処方薬以外の薬及び医薬品、他の支出の9分野については、ボスキンレポートで指摘された品質・新製品バイアスの寄与度は0.11ポイントになるが、これらの分野については、同レポートで研究の根拠が明示されていないことが問題であるとしている。

このうち生鮮果物・野菜について、BLSは、新製品バイアスの大きさは総支出額に占める新製品の支出額の割合を価格弾力性の2倍で割った値に相当するとするJerry Hausmanの推定式を根拠に反論している。

生鮮果物については、同レポートでは、出回り季節の長期化により生鮮果物には30年間で20ポイントの上方バイアスが発生していると主張しているが、仮に生鮮果物の価格弾力性の大きさが1であるとすると、Hausmanの推計式から逆算して、生鮮果物総支出額に占める出回り季節の長期化による増額分の割合は、1996年までの30年間で40ポイントにもなるとしている。ところが、1995年において、バナナ、リンゴなど季節とはおよそ無縁の製品だけで支出額の61%を占めることから、その他の製品を合せても40ポイントもの支出額の割合の増分は望み薄であり、結局そもそも同レポートの主張が誤りであったことになるとしている。

他方、生鮮野菜については、同レポートの主張に沿って支出額の増額分の割合を Hausman の推定式から逆算すると、1972 年から 1995 年までの 24 年間で 29 ポイントになるが、これは BLS が個々の製品を用いて計算した 27 ポイントとほぼ等しい値であると述べている。しかしながら、BLS の食料分野の専門家である Bill Cook によると、生鮮野菜で出回り季節の長期化はほぼ 1985 年以前に終わっていることから、最近の生鮮野菜の支出額の増加は、健康指向を受けた生鮮野菜への嗜好のシフトによるものであり、新製品バイアスとは直接関係しないとしている。

次に、自動車燃料について BLS が反論した内容を要約すると次のとおりである。クレジットカードでガソリン給油機を利用することによる利便性の向上と時間の節約が考慮されていないことにより、年率 0.25 ポイントの上方バイアスが存在するとする同レポートの主張に対して、クレジットカードでガソリン給油機を利用することによる時間の節約を金額換算すると、ガソリン 1 ガロン当たり約 0.45 ポイントと推定されるが、このような支払い方をする運転手の割合は 1996 年において 25% に過ぎない。したがって、これによる上方バイアスの大きさは年率 $0.45 \times 0.25 = 0.11$ ポイントに過ぎない。

イ 前提に誤りがあるとする分野

家賃、被服、新車及び中古車の 4 分類については、同レポートで指摘された上方バイアスの寄与度は 0.16 ポイントになるが、これらの分野については、いずれも同レポートがその要因として掲げている前提に誤りがあるとしている。

家賃については、一般に家賃が家屋の広さに比例して変化するものではないと述べた上で、そもそも平均的家屋の広さは、1976 年から 1993 年の間に、同レポートで述べられているように 20% 増加したという前提は誤りであり、Residential Energy Consumption Survey 等のデータを用いた最近の研究では、家屋の広さの増加分は約 6% と見るべきであるとしている。そして、この数値を用いれば、家賃指数についてはもはや上方バイアスは存在しないことになるとしている。

被服については、同レポートでは 1 年以上品質変化が見られない製品について、カタログを使用して作成した指数と比較して年率 1.00 ポイントの上方バイアスが存在するとしているが、一社のカタログから結論を導くこと自体問題であるとしている。また、被服の価格変化は、特に、ファッションの変化に連動しやすいので、品質変化が見られない製品に限定して計算すると下方バイアスが発生するおそれがあるとしている。

新車及び中古車については、まず、その利用により得られるサービスの価格に変更すべきであるとの勧告に対しては、その精神を尊重しているものの、ユーザーコストの推定に必要なデータが欠如していること、指数の作成に耐え得る賃貸市場が発達し

ていないことを理由に、実用的でないとして一蹴している。また、耐久期間の増分が加味されていないという同レポートの指摘に対して、現行 CPI においてはすでに製造コストのデータを使って耐久期間に関する品質調整が行われているとしている。

ウ 推計の基になったデータに代表性がないとする分野

残りの6つの分野（電気器具、処方薬、専門医療サービス、入院関連サービス、娯楽、身の回り品）については、同レポートで指摘された上方バイアスの寄与度は0.34ポイントにもなるが、これらの分析では研究がまだ進んでおらず、バイアスを推計するに当たって特殊な製品の価格変化のバイアスを利用していることが問題であるとしている。

例えば、医療サービスについて、同レポートでは特に心臓発作治療と白内障治療という進歩が著しいサービスで代表させているが、医療サービスには、歯科、眼鏡、眼科、精神療法、足病治療、指圧療法、物理療法のような多様なサービスが混在しているとしている。医療サービスの価格変化には、寿命の伸長や苦痛の除去など評価が困難な品質変化が含まれており、BLS はこれらの品質変化について今後研究を進めていくとしている。

一方、電気器具については、特に、パソコン、テレビ、ビデオなどにヘドニック回帰法を用いた研究結果を基に CPI に系統的バイアスが存在するとする批判が多いことを認めた上で、BLS もいくつかの製品に対して同法の適用を検討していると述べている。

6.2.2 品質・新製品バイアスの縮小方策

ボスキンレポートが品質・新製品バイアスを縮小するために BLS が対応すべきと主張している事項について、BLS は、まず、品質バイアスの面から、すでに CPI には住居等を除く全体の約 70%に品質変化に対する調整方法を適用して効果が上がっており、実際、これらの分野において 1995 年の品質調整の効果は 1.7 ポイントにも達しているとしている。しかし、品質調整方法としてヘドニック回帰法の適用を拡大すべきであるとする主張に対しては、BLS は賛成しており、今後、耐久消費財の生産価格や特性についてデータ収集をするための予算を確保できれば、同手法を積極的に拡大している。

一方、新製品バイアスについては、POPS（購入地点調査）により店舗の抽出は5年に1回のローテーション方式で行われ、毎年全体の約 20%に店舗と製品の抽出替えがあることから、新製品を CPI に反映させる仕組みは確立していると述べている。なお、BLS は、1999 年 1 月から、品目のサンプリングローテーションを改定し、従来は調査区ごとに5年に1回しか新しい品目の設定が行われなかったところを、一部の

品目類についてはより高頻度で行うようにした。これによって、CPIには従来以上に新製品が反映されるようになってきているとしている。

6.3 品質バイアスをめぐる議論

CPIを考えるに当たって、品質とは何か、また、品質変化とは何かを定義することは非常に難しい問題である。Obst[61]は、品質変化を推計する最大の難点は、製品の特性が変化した時の消費者の効用に対する効果を推計するところにあり、その理由は、消費者の効用が嗜好の変化などによっても変化し、代替効果も考慮する必要があるためと述べている。品質調整法には様々なものがあるが、ここではボスキンレポートでも取り上げられたヘドニック回帰法、コスト推定法、リンク法及び直接比較法について各国等の適用状況、問題点、議論などを概観することとする。

6.3.1 ヘドニック回帰法について

ボスキンレポートが品質調整法としての適用の拡大を主張しているヘドニック回帰法は、1939年にAndrew T. Court[63]が最初に提言した方法で、1990年代まではそれ程普及していたとは言えない。それは、コストなど実用面で大きな制約があることも理由となっている。

ヘドニック回帰法の実際の適用方法には大きく分けて2種類ある。ヘドニック回帰式に時間ダミー変数を加えるなどして回帰式から直接価格指数を推計する方法と、新旧モデルの交替時にヘドニック回帰式を用いて新旧モデルの品質差を推計する方法である。前者の方法は、ヘドニック回帰式の算出に必要な多数のデータがリアルタイムで必要になるため、ほとんどの場合実用的でなく、後者の方法が用いられる場合がほとんどであろう。しかし、説明変数となる特性同志に相関があるため、多重共線性が生じるおそれがあり、推計結果が必ずしも安定しているとは限らないという問題がある。また、少なくとも後者の方法では以下のような問題が生じる。

- ①新製品の属性に照らして独立変数を追加しようとしても新製品には過去の価格が存在しない。
- ②新しい機能を搭載している製品がすでに存在しているものの、まだ数少ない場合、ヘドニック回帰式にその機能に関する独立変数を追加すると、偏回帰係数が不安定になるため推計結果が安定しないことがある。例えば、その機能が高価格のモデルのみに搭載されているような場合、偏回帰係数が高目に推計されてしまうおそれもある。

③パソコンのハードディスク容量のように、品質の向上が速く、相対的に価格低下が大きい特性がある場合、古い時点の回帰式を使うと、新旧モデルの品質差を過大評価するおそれがある。

BLSのFixler et al.[64]は、品質調整の手法としてヘドニック回帰法には、製品の特性が定量的であること、製品の相対的特性が急激には変化しないこと、価格変化はすべて需要の変化に依存することの3つの条件が不可欠であるとしている。BLSは、このような前提を踏まえつつ、同法の適用を1991年に被服から開始し、近年少しずつ適用範囲を拡張している。まず、1998年にパソコンに適用し、これにより同年のパソコンの指数が約6.5ポイント低下したとしている。また、1999年には、テレビにも適用し、1993年8月から1997年8月の間にテレビの指数は毎年約0.1ポイント低下したとしている。

ただし、BLSは、ヘドニック回帰法の適用は、実務的負担を伴い、実際、パソコンに同法を適用するために683種のデータを高い頻度で収集する必要があったとしている。なお、BLSは、1999年4月からの同法適用拡大の研究を支援する資金を受け、約2,500種ものデータを2回に分けて収集した⁵⁶。

ヘドニック回帰法には、技術的にもコスト面でも問題があるにもかかわらず、米国が適用範囲の拡大を図っていたのは、後述するように、米国では、新旧モデルについて同時点で価格を調査することができないことが動機になっていたようである。

また、ヘドニック回帰法は、国際的にみても少しずつ普及してきており、最近、カナダでもパソコン及びテレビに同法が適用されたほか、フランスにおける皿洗い機、フィンランドにおける住宅や家賃など同法を採用する国が増えているが、国によって積極性は異なるようである。

自動車については、かなり早くから多くの検討が行われたが、フィンランドのように部分的に適用している国はあるものの、自動車に同法を完全な意味で採用している国はない。日本でもこれまでに研究が行われたが採用には至らなかった。自動車への採用が困難である理由として、Triplet[65]は、効用関数に影響を及ぼす自動車の特性は非常に複雑でモデル化が困難であり、したがって、独立変数の特定が困難であることを挙げている。医療などのサービスについても、Fixler et al.[64]は、同法の有用性を認めつつも、推定やデータ収集において困難に直面することになるので、同法の適用にはかなりの年月を要すると見ている。

⁵⁶ 具体的には、電話、ビデオデッキ、DVD プレイヤー、ビデオカメラ、冷蔵庫、電子レンジ、洗濯機、乾燥機ごとにデータを収集し、品質変化の計測の妥当性、有効なヘドニックモデル、品質調整に必要な独立変数について研究した。

6.3.2 コスト推定法について

ボスキンレポートでコスト推定法と呼んでいる品質調整法は、製造コストから新旧モデルの品質差を評価する手法であり、米国CPIでは自動車に用いられている。自動車に関する品質調整法としては、他に旧モデルでオプションであった装備が新モデルで標準装備になった場合に当該装備のオプション価格から品質差を評価するオプションコスト法⁵⁷が知られている。

ボスキンレポートでは、コスト推定法について生産の初期段階における品質向上が見過ごされやすいとした上で、特に、自動車に同法を用いることについては、自動車は生産過程における投入コスト以上の品質向上を遂げるので妥当でないとしている。旧モデルについてはオプションの、新モデルでは標準装備に組み込まれた装備の価格を基に行われることが多いが、追加仕様が旧モデルのオプションとして入手可能でない場合には同法の適用は困難であるとしている。また、コスト推定法とオプションコスト法を明確に区別し、前者については、新しい生産技術の導入で、旧モデルよりも性能の良い新モデルがより安価なコストで生産される場合があり、やはり問題があると指摘している。また、ILO事務局は、特に、製品が全国各地や世界各国で作成されている場合には、コストに関するデータ収集は困難であると述べている。このような問題があるため、自動車についてコスト推定法を採用している国は、米国だけのようである。

さらに、コスト推定法には、CPIにおいて品質は消費者の効用の大小で判断する必要があるので、製造コストの大小は必ずしも品質の良し悪しに関係しないという問題もある。米国CPIでは、自動車価格を製造業者から収集した生産コストに関する情報を用いて品質調整しており、BLSは、1967年から1998年までは燃費など排出状況の改善を品質改善とみなしてきたが、1999年以降は価格上昇とみなすように方針変更し、これにより自動車やガソリンの指数は上方にシフトすることとなった。なお、特に、自動車の公害防止装置については、ボスキンレポートでもこれに要する費用を品質向上ではなく価格上昇とみなすべきであるとしている。

この問題は、環境を改善することも消費者全体として見ると効用があると判断するかどうかにかかる論争的なものである。生計費指数の概念をCPIの基本的枠組みとしていない国では、例えば、燃費の良いモデルがより高い価格で販売されているならば、消費者が燃費の良さを品質とみなしていると考えられるであろう。米国では、生計費指数の概念をCPIの概念的枠組みとしているものの、「条件付きでない生計費指数」概念では、環境も含めて消費者の効用を考える必要があり、自動車について燃費など環境に

⁵⁷ 通常、オプション価格の半分を品質差とみなす。

関する特性を品質評価から除外していることは、米国では、「条件付きでない生計費指数」概念ではなく環境が一定という前提の下で消費者の効用を考える「条件付きの生計費指数」概念を採用していることを意味する。

6.3.3 リンク法及び直接比較法について

リンク法は新旧モデルの同時点の価格差が品質差を反映しているとみなせる場合、直接比較法は新旧モデルがほぼ同質であるとみなせる場合に適用される品質調整法である。後者は、品質調整を行っていないとも言えるかもしれないが、価格調査するモデルを交替する時に、旧モデルとほぼ同質のモデルを選択するという調査銘柄の選定方法も広い意味で品質調整法に含まれる。これらの手法は、コスト推定法やヘドニック回帰法など他の品質調整法と比較して、単純であるが故にわかりやすく、コストなどの実用面で優れ、適用範囲が広いので、多くの国で最も良く使われている。

リンク法に関して厳密に述べると、新旧モデルの同時点の価格を同一店舗において調査できる場合にはoverlap法と呼ばれている。新旧モデルを同時点で調査できない場合には、何らかの方法で過去の旧モデルの価格を調整する必要があり、幾つかの手法が考案されている。そのうち、価格変化が類似しているとみられる他の製品の価格変化を用いて過去の旧モデルの価格を外挿するimputed-overlap法（linking法）が一般的であろう。このほか、価格変化が無かったとみなして過去の旧モデルの価格をそのまま用いるlink-to-show-no-change法などがある。

リンク法については批判的な議論も存在するが、リンク法を評価するに当たってはoverlap法と他の手法を区別し、実際にどの手法が適用されているのかを理解しておく必要がある。これは、各国が採用している調査銘柄の抽出・交替方法などに関わる問題でもある。

BLSは、ボスキンレポートの推計結果に反論しつつも、ヘドニック回帰法の適用拡大には賛同しているが、これには調査銘柄の選定の仕方に米国特有の事情があることに留意する必要がある。つまり、米国では、調査銘柄の選定に確率抽出法を用いているので、店舗ごとに調査銘柄が異なっており、調査銘柄の管理を中央で行うことができない。このため、調査していたモデルが店頭からなくなってから他のモデルに変更するので、新旧モデルの価格を重複して調査することが通常できない。したがって、過去の旧モデルの価格を他の店舗における価格の変化で外挿するimputed-overlap法が用いられているが、この外挿方法の妥当性が問題となる。Fixler et al.[64]は、overlap法は、最も単純で最も透明度の高い品質調整法であるが、米国では実際にはほとんど適用できないのに対し、imputed-overlap法を用いると、パソコンでは品質向上分を過小評価し、被服では過大評価するおそれがあり、これが米国CPIでヘドニック

ク回帰法の適用を拡大する動機になっていると述べている。代表的なモデルを調査する非確率的な方法を採用している国の中にも、調査銘柄の規定を緩くし、実際に調査するモデルの選定などをほとんど調査員の判断に委ねている国では米国と同様の問題がある。これに関連して、Astin and Sellwood[66]は、link-to-show-no-change法は、欧州でもよく使われているようであるが、この方法ではバイアスが生じると指摘しており、HICPの作成過程においても不相当とされている。

日本のCPIでは、調査銘柄の交替をあらかじめ中央で定めて全国一斉に行っており、overlap法が適用できる。そして、モデル変更時の品質調整法として、製品ごとに、数量リンク法、価格リンク法（overlap法）、直接比較法から適切と判断されるものを選択して品質調整を行っている。

リンク法及び直接比較法の適用に伴うバイアスの特徴については次のような見解がある。

overlap法を使用する際に、旧モデルがその価格を維持したまま市場から姿を消す場合など、品質変化の直接的な判断を避けるためにCPIに上方バイアスが発生する可能性があるが、最も良く指摘されている問題は、新モデルを発売したときに、新モデルの品質が旧モデルよりも良くなってはいるが、あるいは良くなっているように見せ、実際には価格の上昇分ほど向上せずに値上げが生じるhidden price hikeの問題であろう。この場合、overlap法を適用すると下方バイアスが生じる。

他方、直接比較法では、価格差は品質差の影響を受けていないとみなすので、新旧モデル間の品質差が価格差に反映されている場合に直接比較法を適用すると上方バイアスが発生する。

このようなことから、通常、各国では、個々の調査銘柄の交替について新旧モデルを吟味し、overlap法と直接比較法のうちより適切なものを選択している。

6.4 品質調整法の比較

ボスキンレポートは品質・新製品バイアスについて小数点2桁まで出して、大きなバイアスが生じていると主張している点は、最大の話題であったが、BLSが指摘しているように、その根拠が明確でないことは、レポート全体に対する信頼をゆるがす要因になっている。また、同レポートでは品質改善ばかりに焦点が当てられ品質悪化はそう起きる現象ではないとし、分野ごとの記述の中でも、過去には下方バイアスが存在していたが現在は解消されていると述べている箇所が散見するが、明確な根拠は示されていない。

リンク法は、新旧モデルの品質差を直接評価しないインプリシットな方法のため、品質調整をしていないとの誤解や、米国のようにoverlap法をほとんど適用できない国

の問題と混同している批判が見られるが、価格差は品質差を反映しているという前提はヘドニック回帰法でも同じであり、この前提を否定するのであれば、何を尺度として品質を評価すべきなのか判断基準自体がほとんど存在しなくなる。一方、ヘドニック回帰法は、ほとんどの場合、新旧モデルの品質差の推計に適用され、前述したような特有の問題が生じるおそれがある。すべての品目に共通してどんな場合でも適用できる1つの品質調整法は無いと見るべきであり、個々の品目、個々のケースについて、その品目の性質、データの利用可能性等を踏まえて適切な品質調整法を検討していく必要がある。

なお、リンク法に関連して、米国CPIやHICPなど短期間の特売価格を除外していない場合に問題になるのは、特売価格を価格リンク法に用いることができない点である。欧米諸国においては、日本に比べて短期間の特売が少ないかもしれないが、短期間の特売価格を除外していないために日本よりもoverlap法の適用に制約があることも考えられる。

7 新製品バイアスをめぐる議論

新製品バイアスの議論は、新製品の登場により発生する消費者余剰がCPIに反映されないことによるものと、新製品のCPIへの採用の遅れから新製品発売後の価格低下がCPIに反映されないことによるものの2種類がある。また、新製品はこれまで存在しなかった品目に該当する場合と既存の品目に含まれる場合とに区別される。

7.1 消費者余剰について

新製品の導入により発生する消費者余剰については、1940年にJohn R. Hicks[67]が、発売前の新製品の価格を、新製品の数量が0から増加し価格が減少したと仮定することにより推定が可能であるとした。ボスキンレポートの中には消費者余剰について明示的に取り上げている箇所が見当たらないが、バイアスの推計の一部にこの種の考え方が含まれているとみられる。

新製品の登場によって生じる消費者余剰を留保価格⁵⁸から推計するという考え方に対する見解をいくつか紹介すると、まずBLSのBrent R. Moulton[68]は、消費者余剰は純然たる新製品の効果であるが、新製品が導入されると品質改善と価格低下が伴うので、新製品バイアスを品質バイアスと代替バイアスから明確に分離することは困難であると述べている。Obst[61]も、留保価格を推計することは現時点ではできないとし、問題となる品目は限られること、新製品の登場による消費者余剰は小さいと主張

⁵⁸ 発売前に想定される数量が0のときの価格

する経済学者もいるとも述べている。UNECE顧問のPeter Hill[69]は、留保価格は、消費者が存在と特性を認知していない全く新しい製品に対して定めることはできず、新製品が登場して十分時間が経過した時期の嗜好からでしか定義できないとし、仮に消費者が以前から知識を有していたならば、すでにその製品が発売され、新製品とは言えなかったかもしれないこと、留保価格からの低下がいつ起きたのか明確でないといった問題があると指摘している。したがって、留保価格の基本的な問題は、推計の実際的な難しさというよりも、より根本的な概念上の問題であると述べ、環境変化など非価格的要素も考慮した広い意味での生計費指数には、このような全く新しい製品による消費者余剰を含めるとしても、CPIの対象範囲としては、基準時と比較時の両時点で入手可能か、どちらかの時点で入手可能で他の時点で製造可能な製品に限定することを提案している。

ボスキンレポートでは新製品には必ず価値があると述べているが、消費者余剰を論じる際に、新製品の価格が低下し普及するに伴って陳腐化し、消費者の価値観が変化していくことも問題となる。Moulton[68]は、電灯が無人状態で使用されるように、新製品の価格が極端に低下すると利用価値も低下することもあると述べている。つまり消費している量に比べて消費者の効用は大きくなるということである。これは、ヘドニック回帰法による品質調整の箇所ですべて述べたように、パソコンの内臓ハードディスクの容量が急速に増加しても販売価格は変化せず、ハードディスクの性能が向上してもその相対的価値が低下していく現象とも共通している。

また、消費者余剰を論じる際には、比較的高い価格で市場に登場し、価格低下とともに普及していくものを想定しているとみられるが、新製品の中にはこれに該当しないものもある⁵⁹。

7.2 新製品のCPIへの採用の遅れについて

新製品発売後の価格低下がCPIに反映されないことについては、Obst[61]も、ボスキンレポートと類似の主張であり、新製品は可能な限り速やかにCPIに採用すべきであり、Laspeyres指数の制約はあるものの、バスケットとウエイトの定期的な改訂が潜在的なバイアスを縮減するとしている。

⁵⁹ 例えば、発毛剤などは、技術進歩で製造コストが下がり発売可能になったというよりも、偶然効能が発見され発売可能となった製品である。また、デジタルカメラは以前から高価な製品は存在していたが消費者向けではなかった。ところが、あるメーカーが低品質低価格の製品を販売してみたところ、一定の需要があることがわかり、消費者側も次第にデジタルカメラを認知し、性能に対する要求を強めたため、その後、高機能で高価な製品に対する需要も発生した。発泡酒なども発売してみても低品質でも商売になる程一定の需要のあることが判明した製品である。

BLSは、ボスキンレポートをめぐる議論の直後の1998年にCPIの基準年を約10年ぶりに改定し、パソコンなどのウエイトを大幅に増加し、携帯電話など新製品を数多くCPIに採用した。この定期的改定により少なくとも一時的には新製品バイアスは縮小することとなった。BLSは、2002年以降は2年ごとにウエイトを改訂しており、また、POPSを改善し、全国一斉に新しい品目を導入できるようにし、一部品目類について抽出替えの頻度を高めたことは同バイアスの改善に大きく寄与する。

新品目のCPIへの採用基準については、HICPでは、消費支出に占める割合が1/1000以上になった品目を採用することとしている。一方、日本のCPIでは、1/10000を目安として5年ごとに定期的に採用品目の改廃とウエイトの改訂を行っている。

以上のように、いずれの国においても、実際に新品目を導入するのは、ある程度その品目への支出割合が高くなってからであるが、新製品をできるだけ早くCPIに取り込むべきとの主張には発売直後から調査すべきであるという意味が含まれている。しかし、この意味では、おそらく異論があろう。1つの問題は、発売された新製品が必ず普及するとは限らないため、ある程度普及したことが確認できないとCPIの品目に採用できないことである⁶⁰。仮に新製品が普及することがわかっていたとしても、発売当初の購入者は、比較的所得が高く、新しもの好きで、まだ高価で他人が購入していないものを得ることで満足するいわゆるヴェブレン効果が働くので、一般的消費者の購買行動と言えないこともある。さらに、新製品だけ十分普及していない段階から特別に調査することは、代表的な製品を調査する選定方法や確率的に抽出する選定方法と相容れない面もある。

7.3 新製品バイアスに関する留意点

新製品をできるだけ早くCPIに取り込むべきとの主張には発売直後から調査すべきであるという意味が含まれているが、発売された新製品が必ず普及するとは限らないため、ある程度普及したことが確認できないとCPIの品目に採用できないという問題がある。いずれの国においても、実際に新品目を導入するのは、ある程度その品目への支出割合が高くなってからである。現時点では、一定の基準を設け、その基準を超え、ある程度普及したと判断される製品についてできる限り早くCPIに取り入れることが妥当な目標と言える。

⁶⁰ 例えば、DAT（デジタルオーディオテープ）やDATレコーダーは期待されながらも結局普及しなかった。パソコンは、1970年代に登場し、最初は性能が低く、一種の玩具として限られたマニア向けに販売され、その後のワープロ専用機やゲーム専用機の登場で思うように普及しなかった。現在のように用途が広がったのは性能の向上だけでなく、インターネットの一般利用の解禁などが要因になっている。

なお、新製品に関する問題は国際的にも随所で取り上げられているが、その反対の側面である旧製品が市場から消えたために生じる消費者の損失に関する研究が見られない。例えば、家電製品は製品の入れ替わりが頻繁で、新製品が登場すると、旧製品は店頭から消えていくが、旧製品は消費者の人气がなくなって店頭から消えるとは限らず、新製品の生産のために旧製品の生産が中止されるという事情がある。仮に旧製品も並行して販売されたならば価格は当然新製品（旧製品よりも性能が高い場合）よりも低くする必要があるものの一定の需要があるかもしれない。したがって、新製品の登場による消費者の余剰を考慮するのであれば、旧製品が市場から消えることによる消費者の損失も考慮する必要があるだろう。

8 ボスキンレポートの妥当性

これまでボスキンレポートとそれに対する BLS の主張を中心に近年の米国 CPI をめぐる様々な議論を紹介してきた。同レポートについては、米国 CPI の上昇率に合計すると年率 1.1 ポイントも上方バイアスが存在するというセンセーショナルな推計だけが一人歩きしている感があるが、本章では、その計算根拠が希薄であること、また、同レポートの主張は、実際には有効とは言えない面があること、国際的には特殊ともいえる米国の背景事情や CPI の作成方法と切り離せないことを説明してきた。

ボスキンレポートが米国 CPI に対する改善方策として提示した考え方のうち BLS が賛同している箇所であっても、あくまで米国 CPI を改善する指針として提示されたものであり、背景事情、作成目的、作成方法などが異なる他の国の CPI に対しては意味のない場合もある。同レポートをめぐる議論を日本の CPI に当てはめて検討する際にも、このことには十分留意する必要がある。そこで、本節では、米国と日本の事情を対比させながら、同レポートが米国 CPI に対する改善方法として提示した事項の妥当性、留意点などを述べることとする。

8.1 CPI の概念的枠組み

米国 CPI については、BLS も生計費指数の理念を CPI の計算を行うための概念的枠組みとしているが、CPI の基本概念については統計作成機関の間に国際的合意が形成されていないのが現状である。

CPI の概念的枠組みについて米国、オランダ、スウェーデンなどいくつかの国では生計費指数の理念が CPI の概念的枠組みとされている一方で、オーストラリアや英国では生計費指数は CPI の枠組みとして位置付けられていない。日本の CPI も、家計の消費構造を一定のものに固定し、これに要する費用が物価の変動によってどう変化するかを指数値で示したものであるとされており、生計費指数の理念に基づいていな

い。ただし、2003年にILOの国際労働統計家会議で採択されたCPIに関する決議では、CPIは、「消費者物価の時間的な変化を計測することを目的としている。このことは、同一の品質及び同様な属性の財・サービスの固定された買い物かごを消費者が購入する費用を計測することによって行うことができる。」とした上で、「また、同一の生活水準（効用又は福祉水準）を達成する費用に対する物価の変化の影響を計測することを目的とすることもできる。この概念は生計費指数（COLI）と呼ばれる。」として生計費指数の役割を含めている。

生計費指数の理念は、理論的には興味深い概念であるが、全国民の消費行動は効用という単一の次元尺度に関して極めて合理的な消費行動をとる1人の消費者に十分近似できるという仮定は強い制約である。実際には、明らかに個個人の嗜好や消費行動は異なるので、そのような仮定には少なくとも部分的には無理がある。生計費指数論者のTriplet[70]は、消費者一人一人が必ずしも合理的な行動をとっていないとしても全国民の消費行動を実証分析してみると生計費指数論に基づくモデルが当てはまるとしている。しかし、これは、本来の仮定は正しくないが結果的に見ると正しいということになる。生計費指数の理念が公式の統計指標の基礎理論として認知されるには、より透明度の高い説明が必要である。

8.2 上位集計レベルの指数算式

ボスキンレポートでは、生計費指数の理念がCPIの計算を行うための概念的枠組みであるとの前提に立った上で、最良指数の採用が望ましいと主張している。基準時と比較時のバスケットの何らかの平均をバスケットとした指数算式を考えると、最良指数の一種になることから、生計費指数の理念とは関係なく最良指数が最適であるとの主張もあり、生計費指数の理念の妥当性と最良指数の妥当性は必ずしもイコールとは言えないが、比較年のウェイトが必要な最良指数には実用上の制約があることから、速報性を要する公式のCPIで最良指数を採用できるものではない。

生計費指数の理念がCPIの概念的枠組みである米国においても、最良指数には実用上はいくつか問題があり、本章で紹介した議論や論文においてはとりわけ次の2点が強調されていた。1つは最新時点の情報を素早く取り込むことが困難なことであり、最良指数は年次系列では有用でも、月次系列には適用できない理由の1つとなっている。もう1つは、生計費の概念が広く多岐にわたることから、それを的確に、かつ実用的に数量化することが困難なことである。BLSが生計費指数の概念を重視しつつもLaspeyres指数でCPIを公表しているのはこれらの問題を回避できないからである。このような問題に対しては様々な検討が重ねられているが、打開策を見出すまでにはかなりの時間を要すると考えられる。

このため、基準年を毎年変更することで代替バイアスをある程度抑える Laspeyres 連鎖指数の採用を目指している国もある。米国も基準年をより高い頻度で改定し、連鎖指数を導入した。日本では 1975 年から年平均の連鎖指数を参考系列として公表しているが、公式 CPI との差は小さく、多くの利用者の利便性を考えると、簡明で使い易い Laspeyres 指数を維持する意味は十分ある。

8.3 下位集計レベルの指数算式

ボスキンレポートの勧告では、下位集計レベルの指数算式を幾何平均に移行すべきとしており、実際、米国、カナダ、フランス、オーストラリアなどは、下位集計レベルにおける指数算式を一部幾何平均に移行した。しかし、それには指数の作成方法に起因する事情がある。

米国においては、同一製品でも店舗によって調査銘柄が異なるため同質とはみなせず、各店舗の価格の算術平均に意味がないことから、従前は価格比をとってから算術平均していたことでバイアスが生じていた。また、今後、総合指数を連鎖指数に近い方式に移行するためドリフトの問題⁶¹を避ける必要があった。このようなことから、代替弾力性がある程度あるとみられる品目について幾何平均に移行したとみられる。英国においては、銘柄の選定は非確率的であるが、店舗を売上比例確率抽出していることなどの理由で幾何平均を採用している。

一方、日本では、店舗及び銘柄とも代表的なものを選定する非確率的抽出法を採用するとともに、調査銘柄は、各店舗とも同質とみなせる製品の価格を取集するよう規定している。このような店舗・銘柄の抽出方法をとる場合、算術平均と幾何平均の差はかなり小さくなるとの指摘がある。また、同じ算術平均でも各店舗の価格の算術平均を計算してから基準時の平均価格との比率をとる手順を採用しているので、従前の米国で生じたようなバイアスの問題を避けることができる。HICP の作成方法を検討する過程においても、価格比の算術平均は上方バイアスが発生するおそれがあるとして不適当とされたが、価格の算術平均の比と価格の幾何平均の比は優劣が判定できないため、いずれを採用しても良いとされている。

したがって、日本においては下位集計レベルの指数算式として従来通り算術平均を用いても大きな問題はないが、例え大差はないにしても、品目ごとに見て算術平均と幾何平均のいずれが優れているかについてはまだ研究すべき余地が残されているとみられるので、Obst[61]が述べているように、POS データの利用などによる研究の進展が望まれる。

⁶¹ 価格が上昇と下落を繰り返している品目がある場合、指数が高目になる現象

8.4 新店舗の取扱

ディスカウントストアなどの新しい店舗の価格動向を適切にCPIに反映すべきであるというボスキンレポートの主張について、米国では最近の改定で一部店舗の抽出替えの頻度を上げている。米国のように代表性のない店舗も確率的に抽出している場合には抽出替えの頻度が重要であるが、日本のように各調査地区において代表的な店舗を選定して調査する方法を採っている場合には、抽出替えの頻度よりも、調査店舗を交替するときの新旧店舗間における同一製品の価格の接続方法が重要である。それは、平均的に見ると、各店舗の同一製品の価格はほぼ並行して変動しているとみられるので、調査店舗の代表性を維持している限り、価格の変化を正確に捉えているとみられるからである。

店舗の抽出替えを早め、まだ代表性を維持している調査店舗から他の店舗に交替しようとするときに、両店舗間の同一製品の価格差を無視して調整せずに比較するのは問題があろう。このため、米国では、店舗の抽出替えが行われると新旧店舗間の同一製品の価格差は両店舗の付加的サービスの品質差とみなしてリンク（overlap法）している。しかし、新旧店舗の同一製品の価格差をすべて品質差とみなすことには議論があろう。米国では、調査店舗を交替するとき、各店舗の価格を取集する銘柄も抽出替えしていることから、同一銘柄の価格を比較することができないことも、リンクする理由になっている。日本では、調査店舗が代表性を失って新店舗に交替するとき、サービス品目及び一部の財を除いてリンクせずに直接比較している。これは、代表性が失われたことで消費者が旧店舗の価格に対して新店舗の価格よりも割高と感じていたと判断できるからである。

店舗間の品質差の評価方法については、国際的にも定まった見解がないようである。米国では代表性のない店舗も確率的に抽出しているのに対して、日本ではそのようなことをしていないので、日本では米国に比べて議論の焦点が絞りやすいとは言え、この点について更に研究の進展が望まれる。

8.5 品質調整法

ボスキンレポートでは、米国CPIの品質調整法としてヘドニック回帰法を推奨しているが、ヘドニック回帰法には、コスト面も含めて問題点があり、万能とは言えない。

米国や日本など多くの国で用いられているリンク法については批判的な議論も存在するが、リンク法を評価するに当たってはoverlap法と他の手法を区別し、実際にどの手法が適用されているのかを理解しておく必要がある。これは、各国が採用している調査銘柄の抽出・交替方法などに関わる問題でもある。

米国特有の事情について振り返ると、米国では調査銘柄の選定にも確率抽出法を用いているので、店舗ごとに同一品目の調査銘柄が異なっており、調査銘柄の管理を中央で行うことができない。このため、調査していたモデルが店頭からなくなってから他のモデルに変更するので、通常新旧モデルの価格をオーバーラップして調査することができない。

日本では、調査銘柄の交替をあらかじめ中央で定めて全国一斉に行っているので overlap 法が採れる。このため、リンクに関して米国のような問題は少ないが、overlap 法、直接比較法など適切な品質調整法の選択について不断の検討を続ける必要がある。この際、各品目に対してできるだけ広い範囲の中から適切な品質調整法を選択していく姿勢が重要と考える。ヘドニック回帰法については各国で実証研究が次第に蓄積されてきているが、事後的な比較検証の有効な手段となり得ても実用面でまだ多くの検討が必要であり、実用性を考えると簡明さも重要である。

8.6 新製品の取込み

ボスキンレポートにおける主張の 1 つである、新製品をできるだけ早く CPI に取り込むべきとの点については、米国 CPI では基本的には約 10 年に 1 度しかウエイトを改正しないという状況を前提としている。日本の CPI では、1/10000 を目安として 5 年ごとに定期的に採用品目の改廃とウエイトの改訂を行い、中間年においても急速に普及して一定割合を超えた新製品を取り込んでいることから、日本では意味が弱まる。

また、新製品をできるだけ早く CPI に取り込むべきとの主張には発売直後から調査すべきであるという意味が含まれているが、これについては発売された新製品が必ず普及するとは限らないなどの問題や異論もあり、現段階では、一定の基準を設け、その基準を超え、ある程度普及したと判断される製品についてできる限り早く CPI に取り入れることが重要である。

9 21 世紀におけるボスキンレポートの影響

ボスキンレポートは 1996 年の公表前後においては各方面に衝撃を与えたものの、当の米国では、次第に存在感を薄めてきている。この理由について、Triplet は、BLS が CPI を改善したことのほかに、大統領と議会との間で結ばれた予算に関する協定に CPI の作成方法の見直しに関する規定が存在しないことを挙げている。しかし、米国で政治問題から CPI がはずされたのは、統計学的視点を重視したというよりは、財政状況が好調であったことと、社会保障費の削減と税金の増額に結びつくような施策を実施することは困難であるという、政治的な事情による。

その後、2000年にGAO（General Accounting Office）はボスキンレポートの更新版 *General Accounting Office[71]Consumer Price Index: Update of Boskin Commission's Estimate of Bias* を出している。その中で、CPIについてBLSが7種の改善を実施したことから、バイアスは縮小し、委員会の構成員による推定では、年率0.73~0.9ポイントであるとしている。

2002年には、国家統計委員会が招集したパネルによるいわゆるCNSTAT（Committee on National Statistics）[72]の中で生計費指数の概念、方法、その他の課題が提示されたが、これはBLSがボスキンレポートに対する回答を委託したものである。しかし、CPIの測定概念が生計費指数かについては議論が収束しなかった。

ボスキンレポートが報告されて10年が経過すると、BLSの中に同レポートを評価する見解が現れた。BLSは、ボスキンレポートは経済学者の間で物価に対する関心を強める効果があり、BLSと外部有識者との間の関係を強める働きをしたとしている。ボスキンレポートは政治的な背景事情のもとで提出されたものではあるが、統計が政策に大きな影響を及ぼすことを統計利用者のみならず統計作成者にも強く印象付けるものであった。今となっては肯定的に評価すると、政策のエビデンスとしての統計の重要性を実感させる報告であったと言える。

ボスキンレポートの影響は波のように米国外に広がったが、総じて冷静に受け止められた。オーストラリアでは、ボスキンレポートの公表時期が偶然にも1960年以來のCPIの見直し時期に重なったとして見直しをするための参考情報として活用した。Keith Woolford[73]によると、1997年の見直しで、対象人口を勤労者世帯のみから全世界帯にする等により29%から64%に拡大した。また、下位代替バイアスを縮小するために下位品目レベルの算式を原則として価格比の算術平均から幾何平均に変更した。さらに、品質バイアスを縮小するために、それまで実地で価格を記録していたパソコンや、専門家の評価に依存していた新車に対しヘドニック法を適用した。

欧州には、ボスキンレポートを受け、一部各国のCPIにバイアスが存在するとする者も現れたが、全体として大きな議論に発展しなかったようである。もともと欧州では、CPIの作成方法が米国よりも日本に近いという事情があった。しかし、2002年に通貨が統合されたことにより、CPI上昇よりも物価上昇の実感が大きく上回り、世論の中にCPI上昇は物価上昇を過小評価しているとの批判が生じた。Paolo del Giovane and Roberto Sabbatini[74]によると、イタリアでは、1ユーロが公式には1,936リラに換算されたにもかかわらず1,000リラで換算されたとの誤解が生じ、民間研究機関からの物価上昇の推定も高かったため、実感がCPI上昇を大幅に上回り、それが縮小するのは2004年になってからであった。このように、CPIに対する批判が生じることは

あったが、懸念されたバイアスの方向が逆向きであったこともあり、ボスキンレポートとの関連は強くはなかった。

Johnson et al.[75]によると、2000年代前半には、当の米国においても、Charles R. Hultenや財界からCPIに下方バイアスが存在しているという見方が現れ、2004年には、ボスキン委員会の委員であるRobert Gordonでさえ、CPIには以前から家賃と衣料の扱いに伴う下方バイアスが存在することを示唆する2本の論文を書いている。

多くの国に当てはまる現象として、20世紀後半のように物価上昇率が高い時代にはCPIに上方バイアスが存在するとの批判が起きやすく、21世紀のように物価の変化率が小さな局面ではCPIに下方バイアスが存在するとの批判が起きやすい。

第2章 日本のCPIの作成技法の高度化

次に、近年の日本のCPIをめぐる事情について説明する。

日本では、21世紀に入り、長期間CPIの変化率がマイナス傾向で推移するという先進国では異例の事態が生じた。このため、日本では、CPIに下方バイアスを懸念する見方はあっても、米国のように上方バイアスの存在を指摘する者は少なかった。

しかし、CPIの精度そのものよりも、これにより、賃金や生産額の低下などの社会・経済の問題が起き、やがてデフレ問題に発展し、その解決は重要な政策課題となった。特に、金融政策において、量的緩和政策の導入とその解除、物価上昇目標などとの関連で、CPIの動きに対する関心が強まった。CPIはまさに経済の体温計として、政策の発動と評価の両面から注目されたが、下落傾向とはいえゼロに近い水準であることから、ゼロをはさんでプラスかマイナスかの小さな動きが注視された。

その中でも、2005年基準改定において、物価の基調として注目されていた生鮮食品を除く総合の変化率が改定前より0.6ポイント下方に修正されたことが関係者に大きな衝撃として受け止められ、その要因が追究された。

そこで、本章では、2005年基準改定を中心に、その内容、改定前後の変化率の差の要因等、2005年前後における日本のCPIの作成技法の高度化についての実証研究を示す。

具体的には、2005年基準改定では、総世帯ウエイトに基づく月次指数、生鮮食品を除く総合の月次指数、食料（酒類を除く）及びエネルギー除く総合を新たに公表することになったが、そこに至るまでに行った様々な実証研究を示す。他方、指数算式については大きな変更をしなかったが、実証研究を重ねた結果の措置であるから、それについても説明する。これらは2005年基準改定において研究、分析又は反映した点

であり、一部を除き、Makoto Shimizu[76]、Makoto Shimizu[77]などに概要を記載している。

さらに、2005年前後に、長期的な課題として家賃指数の欠測値補完の研究を行ったので、それらの研究についても以下に示す。

本論文で示す研究以外にも、世帯類型別の指数の作成について Sakashita and Shimizu[78]、携帯電話指数の作成方法について Makoto Shimizu[79]などにまとめているので、適宜参照されたい。

1 総世帯ベース月次指数の作成

最初に、参考系列として 2000 年基準から年単位で公表されている総世帯ベース指数の月次化について論じる。

CPI の主系列の算定に用いるウェイトは、2人以上の世帯の消費支出を基に計算している。しかし、CPI の算定に用いる価格は、原則として単身者が購入する商品等の価格を含めて平均したものである。このため、価格とウェイトの概念に不整合が生じている。また、CPI の利用の仕方としては、年金の物価スライドを始め、2人以上の世帯のみならず単身者も含めた対象に向けたものとなっている。さらに、家計調査においても単身者を含めた調査結果を公表している。したがって、CPI のウェイトも単身者を含めた総世帯ベースに変更することが必要である。

諸外国においても、ウェイトの算定対象を2人以上の世帯に限定している国は、韓国を除き見当たらない。

特に、日本において単身世帯は増加傾向で推移し、2015年国勢調査では世帯全体の34.5%を占めており、もはや少数として特別扱いできる対象ではない。単身世帯は今後も増加することが想定されることから、CPI についてもそれらを含んだウェイトを基に作成することが必要であると考えられる。

この問題は我が国固有の問題であることから、これらに関する研究は、筆者の知限り Makoto Shimizu[76]及び Makoto Shimizu[77]以外に存在しない。

1.1 変更による影響

参考系列の総世帯ベース指数は、主系列（2人以上世帯ベース指数）と大差はなく、総合指数の前年比は、主系列と比べて 2001 年においては 0.1 ポイント低いものの、2002 年及び 2003 年においては同じである。したがって、総世帯ベースへの変更によ

る年金の物価スライドや物価連動債などの重要な利用目的に対する影響は小さい。また、GDP デフレーターには個々の品目単位で使用されるため、影響はほとんどない⁶²。

1.2 月次指数の計算方法

CPI のウェイトは基準年全体で作成されるので、計算方法としては、総世帯ベース指数についても年次指数のウェイトをそのまま適用すれば良い。ただし、月別にウェイトを定めている生鮮食品について何らかの推計が必要である。家計調査の単身者に係る結果が四半期単位でしか公表されていないことから、生鮮食品に係る支出の四半期における月ごとの構成が、単身者については2人以上世帯と同じであると仮定して総世帯ベースの月次指数を計算すると、主系列と大差はない。

試算結果の実用可能性が確認されたことから、2005年基準から参考系列として総世帯ベースの月次指数を作成・公表することになった。

1.3 指数の接続

主系列を総世帯ベース指数に変更する場合、過去の主系列との接続の仕方について検討が必要である。ただし、総世帯ベース指数への変更は、持家の帰属家賃のように大きな品目を追加するものではなく、ウェイトの分割の仕方に関する変更なので、そのまま接続するという方法が妥当である。実際、従来も基準改定のたびに品目改廃、支出割合の変化等に伴いウェイトの構成を変えているが、それらによる指数への影響のほうが（総世帯ベースに変更することよりも）大きかったにもかかわらずそのまま接続している。

2 上位指数算式の検討

上位指数算式の検討に関する実証研究結果を示す。

通常、上位指数算式において、Laspeyres 指数は価格と数量が反比例するという状況の中で上方バイアスを持ちやすいと言われている。日本の CPI では、Laspeyres 指数を用いているので、例えば、価格が下落し数量が増加している電気冷蔵庫や電気洗濯機などの家事用耐久財などで上方バイアスが発生する。この場合、比較時における品目ごとの購入量を固定する Paasche 指数で計算すると Laspeyres 指数よりも低くなり、下方バイアスが発生する。実態は Laspeyres 指数と Paasche 指数の間になること

⁶² 現在主系列では市町村別指数を合算するときには2人以上世帯数（国勢調査）を用いているが、総世帯ベース指数を計算する場合には総世帯数を用いるため、若干の違いが生じる可能性がある。

が想定される。また、算術平均型でなく幾何平均型の指数を作ると、そちらのほうが最良指数に近づくとされている。

しかし、Laspeyres 指数で必ずしも上方バイアスが発生するとは限らない。供給よりも需要が先行する場合などに購入量と価格が並行して変化する場合があるからである。例えば、BSE 問題が発生した直後は牛肉の購入量が減ったので供給側は価格を下げざるを得なかった。外国パック旅行は休みの時期に需要が増大することから、供給側は価格を上げる。固定電話は携帯電話に需要がシフトし、通話機の価格も下げざるを得ない。価格下落は固定電話のほうが大きいにもかかわらず、購入量は携帯電話にシフトしている。このような場合、Laspeyres 指数は下方バイアスを含み、Paasche 指数よりも低くなる。また、幾何平均型で指数を計算すると最良指数よりも離れる。

分類ごとに上方バイアスと下方バイアスが発生し、総合では相殺してバイアスが小さくなることもある。このため、公式指数の実態との乖離については、分類ごとに連鎖指数や固定型の Paasche 指数などと比較しながら注視することが必要である。

なお、Paasche 指数の対象となる品目は、基準時に選ばれたものであり、比較時において購入量が多いものから選定したものではない点に注意が必要である。

2.1 インフレ時及びデフレ時の実データの比較

価格と数量の関係を分類すると、図 11 のとおり、それぞれの上昇、下降の種類によって、第 I 象限、第 II 象限、第 III 象限、第 IV 象限とに分かれる。一般的には網掛けで示した第 II 象限と第 IV 象限が良く起こる状況である。

インフレ期の 1995 年から 1998 年、デフレ期の 2000 年から 2003 年を比較するため、象限ごとにウエイトを合算した結果を表 25 に示す。ウエイトが最も多い象限はいずれも第 IV 象限であり、インフレ時には半数以上がこの空間の中に入っている。しかし、デフレ期になると、第 III 象限の割合がかなり高くなり、33.8%を占める。

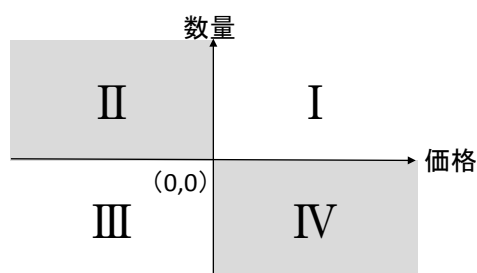


図11 指数変化の分類

表25 ウェイト構成比(%)

	基準年後3年後の変化	1995年	2000年
分類	合計	100.0	100.0
I	価格変化 ≥ 0 、数量変化 ≥ 0	19.6	7.4
II	価格変化 < 0 、数量変化 ≥ 0	12.6	19.8
III	価格変化 < 0 、数量変化 < 0	9.8	33.8
IV	価格変化 ≥ 0 、数量変化 < 0	58.0	38.9

出典 Makoto Shimizu(2004)“Arithmetic Formula of CPI Based on Changing Consumption Pattern in Japan”、国際コンファレンス「我が国SNAの次期整備に向けて一より精確な計測方法、より包括的な勘定を目指して一」、内閣府経済社会総合研究所

表 26 はインフレ時において Laspeyres 指数が Paasche 指数よりも小さくなる費目だけを集めたものである。穀類で見ると、Laspeyres 指数が 97.2、Paasche 指数が 97.3 となっている。幾何平均にすると Laspeyres 指数が 97.1、Paasche 指数が 97.4 となり、その幅が広がっている。

また、楽器については、ピアノの数量が増加して価格も増加しているため、Paasche 指数のほうが Laspeyres 指数よりも、算術平均型では 0.6、幾何平均型では 0.7 も大きくなっている。

表 27 はデフレ時における状況である。一般家具と自動車等維持費はインフレ時と同じ項目であるが、その他に保健医療サービス、通信、教育など、費目の数は少ないが広範囲に及んでいる。保健医療サービスについては代替効果が見られない品目が数多く存在するにもかかわらず、指数算式に影響が出ている。教育は 10 大費目の 1 つなので、全体に対する影響は大きい。

一例として通信を見ると、主な要因は固定電話通信料であり、数量は減っているが、価格もかなり下がっており、その影響で指数算式に逆転現象が起きている。算術平均型では Laspeyres 指数が 92.4、Paasche 指数が 93.0 であり、Paasche 指数のほうが 0.6 も大きくなっている。これを幾何平均型にすると Laspeyres 指数で 92.2、Paasche 指数で 93.2 となり、Paasche 指数のほうが 1.0 も大きくなっている。

表 26 1998 年において Laspeyres 指数が Paasche 指数を下回る費目 1995 年=100

品目	算術平均型			幾何平均型		
	Laspeyres	Paasche	Fisher	Laspeyres	Paasche	Törnqvist
穀類	97.2	97.3	97.2	97.1	97.4	97.2
パン	101.6	101.7	101.6	101.7	101.7	101.7
めん類	102.3	102.4	102.3	102.3	102.4	102.3
生鮮肉	107.2	107.3	107.2	107.2	107.3	107.2
加工肉	103.5	103.6	103.5	103.5	103.6	103.5
乾物・海藻	101.3	101.4	101.3	101.4	101.4	101.4
大豆加工品	102.6	102.7	102.7	102.6	102.7	102.7
油脂	91.5	91.6	91.5	91.4	91.6	91.5
主食的調理食品	103.7	103.8	103.7	103.7	103.8	103.7
工事その他のサービス	105.4	105.5	105.4	105.3	105.5	105.4
一般家具	100.3	100.7	100.5	100.2	100.8	100.5
ティッシュ・トイレット	94.2	94.3	94.2	94.1	94.3	94.2
和服	103.0	103.3	103.1	103.0	103.3	103.2
靴下	102.7	102.8	102.8	102.6	102.9	102.7
自動車等維持	91.3	91.6	91.5	91.0	92.0	91.5
授業料等	106.7	106.8	106.7	106.7	106.8	106.8
楽器	103.0	103.6	103.3	103.0	103.7	103.3
月謝類	104.9	105.0	104.9	104.9	105.0	104.9
かばん類	101.3	101.4	101.3	101.2	101.4	101.3

出典 表 25 と同じ

表 27 2003 年において Laspeyres 指数が Paasche 指数を下回る費目 2000 年=100

品目	算術平均型			幾何平均型		
	Laspeyres	Paasche	Fisher	Laspeyres	Paasche	Törnqvist
酒類	97.6	97.7	97.6	97.6	97.7	97.7
一般家具	94.9	95.0	95.0	94.9	95.0	95.0
男性シャツ・セータ類	96.7	96.9	96.8	96.6	96.9	96.8
婦人シャツ・セータ類	87.9	88.0	87.9	87.8	88.1	87.9
婦人下着類	99.5	99.6	99.5	99.5	99.6	99.5
保健医療サービス	110.1	110.2	110.1	110.0	110.2	110.1
自動車等維持費	101.2	101.3	101.3	101.1	101.4	101.3
通信	92.4	93.0	92.7	92.2	93.2	92.7
教育	102.7	102.8	102.7	102.7	102.8	102.8
他の娯楽用品	94.3	94.4	94.4	94.2	94.5	94.3

出典 表 25 と同じ

2.2 適用すべき指数算式

2.1 で得られた知見をまとめると、デフレ下においては、価格と数量が並行して変化する場合が多いことから、Laspeyres 指数に下方バイアスが生じる費目が多くなる。また、Laspeyres 指数について、算術平均型より幾何平均型のほうが最良指数に近くなるとは限らない。

このような状況を踏まえると、CPIの上位指数算式には、実用性と時系列の観点でメリットの多い算術平均型のLaspeyres指数を引き続き適用することが妥当であると考える。

なお、本節の内容は、2005年に開催された国際コンファレンス「我が国SNAの次期整備に向けて—より精確な計測方法、より包括的な勘定を目指して—」（内閣府経済社会総合研究所）で筆者が発表した内容の一部である⁶³。世界的にはデフレが珍しい時代において、価格と数量が並行して変化する場合が多いという傾向は日本独自の現象であった。

3 月次連鎖指数の作成

本節では、参考系列として1975年基準から年単位で公表されている連鎖指数の月次化について説明する。

固定型指数算式のバイアスを少しでも緩和し実態に近い水準を計算するために、基準年の間の年についても前年の家計調査の結果から支出割合をウェイトとして指数を毎年作成し基準年から掛け合わせた連鎖型のLaspeyres指数が参考系列として公表されている。2004年において公式指数と連鎖指数を比較すると、総合では0.3ポイント、前年比で0.1ポイントだけ公式指数が高くなっている。生鮮食品を除く総合で見ると0.3、前年比で0.2ポイントだけ公式指数が高くなっている。さらに、これを10大費目別に見ると、差は教養娯楽で大きく、前年比で0.5ポイント公式指数が大きくなっている。教養娯楽には価格下落が著しいパソコンやプリンタなどが含まれていることから、購入数量の変化が指数全体に大きな影響を及ぼすからである。

実態との乖離について今後はよりきめ細かな分析ができるように、2005年基準改定では、年1回公表していた連鎖指数のうち変動の激しい生鮮食品を除く総合について毎月公表することとした。

しかし、連鎖指数では品目の改廃が反映されていない点に注意が必要である。また、連鎖指数には加法整合性がない点、例えば、内訳の合計が全体に一致するとは限らない点が不便な点である。例えば、2004年において教養娯楽の指数を教養娯楽用耐久財、教養娯楽用品、書籍・他の印刷物、教養娯楽サービスから加重平均して計算すると、表28に示すとおり、連鎖指数については2000年ウェイトを用いても2004年ウェイトを用いても直接計算したものとは一致しない。

⁶³ ラスパイレス指数に下方バイアスが発生する例については清水誠[80]に記載

	ウエイト		2004年の指数		
	2000年	2004年	公式	連鎖	
教養娯楽用耐久財	386052	328143	50.4	45.8	
教養娯楽用品	837826	744571	93.3	93.1	
書籍・他の印刷物	521365	515068	101.0	100.9	
教養娯楽サービス	1977647	1908468	97.6	97.7	
合計（教養娯楽）	公表結果	3722890	3496250	92.2	90.6
	2000年ウエイトに基づく試算			92.2	91.7
	2004年ウエイトに基づく試算				92.3

出典 清水誠（2006）「消費者物価指数の実際」『経済セミナー』612，日本評論社

諸外国では、特に、欧州で連鎖指数を主系列に掲げている国が目立つが、基準年が古いことが多いことから、固定型では実態との乖離が大きくなってしまおうという事情を踏まえた措置である。

連鎖指数に関する先行研究は数多く、米国では、2002年から新たに連鎖指数を導入したことを受け、Cage et al.[81]が指数算式や接続方法の検討状況をまとめている。本節の内容は、日本のCPIについて月次の連鎖指数の作成・公表に向けたMakoto Shimizu[76]及びMakoto Shimizu[77]の研究を基にしている。接続方法の検討に当たりウエイト等の詳細な情報が必要なため、先行研究は見当たらない。

3.1 月次連鎖指数の作成方法

主系列として公表されている固定型指数と比べると、月次連鎖指数の品目別指数はまったく同じである。その意味では、品目ベースではすでに連鎖指数を毎月公表していることになる。総合（生鮮除く）、中分類指数、小分類指数は連鎖の手法により異なるものとなる。

月次の連鎖指数については様々な作成方法が存在するが、2005年基準から新たに公表する毎月の生鮮食品を除く総合については、以下の3つの方法について試算した。

①前月接続

毎月前月の指数で接続させる連環的方法である。

②12月接続

前年12月の指数で接続させる方法であり、日本銀行が作成する企業物価指数で採用されている。

③年平均接続

前年年平均の指数で接続させる方法であり、参考系列として公表している年次連鎖指数と整合的である。

①、②、③のいずれについてもウエイトは各年年平均を使用する。ただし、違いはウエイトによる加重平均をどのタイミングで行うかによるものである。

②12月接続と③年平均接続は①前月接続や公表指数と比べて毎月の変動が小さくなる傾向がある。したがって、季節性が除外されやすい。しかし、そもそも公表系列に現れる季節性が適切かという問題もある。公表系列には生鮮食品以外にも季節性が現れる品目が含まれるが、生鮮食品に倣って月単位の変動ウエイトにするという発想もある。

$$I_{y,m}^L = \frac{1}{12} \sum_{m=1}^{12} I_{y,m}^L$$

$$I_{y,m}^{Mon-Mon} = I_{2000,12}^L P_{2001,1} \cdots P_{y,m-1} P_{y,m} \quad P_{y,m} = \sum_{i=1}^n s_{i,y-1} \frac{I_{y,m}^L}{I_{y,m-1}^L} \quad (m \neq 1)$$

$$I_{y,1}^{Mon-Mon} = I_{2000,12}^L P_{2001,1} \cdots P_{y-1,12} P_{y,1} \quad P_{y,1} = \sum_{i=1}^n s_{i,y-1} \frac{I_{y,1}^L}{I_{y-1,12}^L}$$

$$I_{y,m}^{Dec-Mon} = I_{2000,12}^L P_{2001,12} \cdots P_{y-1,12} P_{y,m} \quad P_{y,m} = \sum_{i=1}^n s_{i,y-1} \frac{I_{y,m}^L}{I_{y-1,12}^L}$$

$$I_{y,m}^{Yea-Mon} = I_{2000}^L P_{2001} \cdots P_{y-1} P_{y,m} \quad P_y = \sum_{i=1}^n s_{i,y-1} \frac{I_y^L}{I_{y-1}^L}$$

$$P_{y,m} = \sum_{i=1}^n s_{i,y-1} \frac{I_{y,m}^L}{I_{y-1}^L}$$

3.2 月次連鎖指数の評価方法

2000年から2004年までのデータに上記の3つの手法を適用すると図12のとおりである。

まず、①前月接続の結果は公表指数と大きく異なる。また、衣料品を中心に公表指数より高目の数値が出ることから、結果として実態を反映しているとは言えない。

②12月接続と③年平均接続にはほとんど違いはない。ただし、生鮮野菜を除く野菜・海藻については②12月接続が公表指数より高目にシフトするのに対し、③年平均接続は低目にシフトするなど、双方で結果が異なる分類もある。

加法性を満たすという点からも①毎月連鎖よりも②12月接続と③年平均接続のほうが優れている。ただし、この点から評価をすると、完全に加法性を満たしている公表指数が最も優れていることになる。

③年平均接続を年平均するとすでに別途公表している年次連鎖指数と算式としては一致する。また、③年平均接続は生鮮食品についても計算が可能である。このような実用性の観点からみれば③年平均接続は有用である。

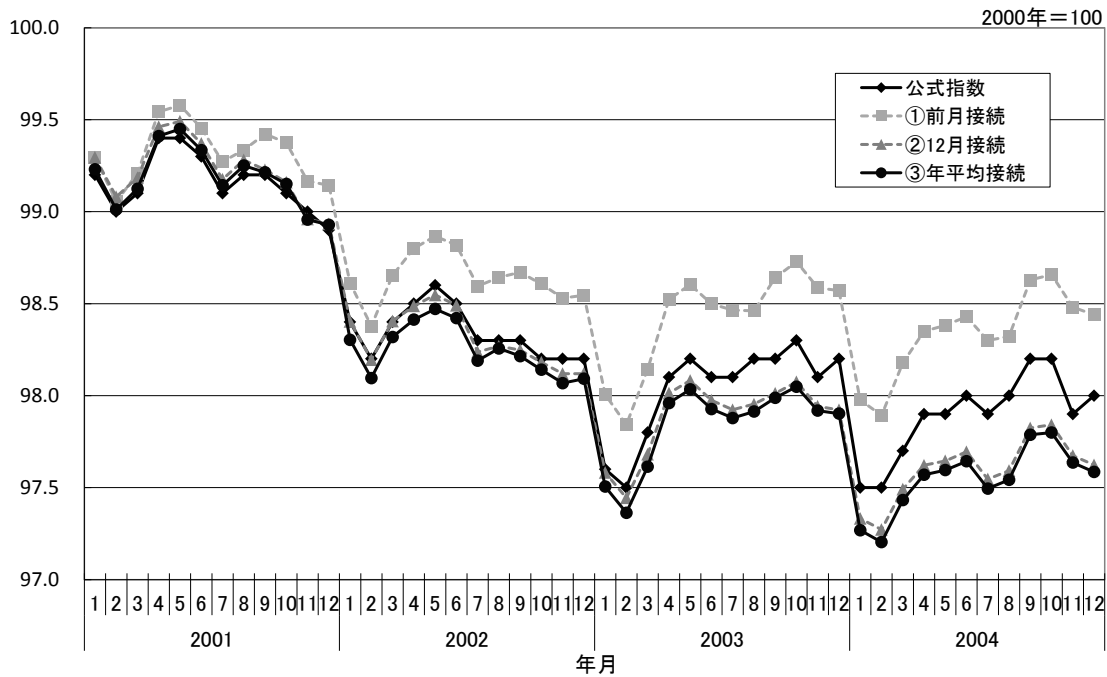


図 12 生鮮食品を除く総合指数の試算

3.3 2005 年基準改定以降の状況

上記を踏まえ、2005 年基準改定では、新たに参考系列として公表される生鮮食品を除く総合の月次連鎖指数の接続方法を③年平均接続とした。

その後も連鎖指数は Laspeyres 指数による総合指数のバイアスを推定するメルクマールになっており、生鮮食品を除く総合について前年比を見ると、2010 年に公式指数で -1.0% のところ、連鎖指数では -1.3% と 0.3 ポイント下回っている。

しかし、2010 年基準改定後の 2015 年の公式指数は 0.5% であるのに対し、連鎖指数では 0.6% と逆に連鎖指数のほうが 0.1% 上回っている。主な要因は、基準時点を前年とした連環指数を算出する際、連鎖時点を指数の下落の大きかった公立高校授業料の品目指数を 100 に戻した後、2014 年 4 月の高等学校等就学支援金制度改正により公立高校授業料の前年比が上昇したことによる影響である。このように、連鎖指数は時々の個別事情が大きく反映されることがあるので結果を評価する際には注意が必要である。

4 コア指数の作成

2005 年基準から新たに食料及びエネルギーを除く総合指数が公表されることになったが、その理由と状況は以下のとおりである。

CPIについては物価の趨勢を把握するための指標として、天候等一時的な状況に左右されやすい生鮮食品を除外した指数を作成・公表している。

2005年前後、物価が下落か上昇かといった微妙な動向が注目される中で、特に、金融政策の判断材料としての物価の趨勢についての議論が多方面で展開されていた。特に、原油価格が高騰している中で、石油製品は物価の最大の上昇要因であったが、石油製品は一時的要因に左右されやすいことから、それを除外して指数を作成すべきという主張があった。他方、食料についても、生鮮食品の他に、米価格が天候の影響を受けやすく、実際前年同月比で見たときに2003年における冷夏で10月から1年間高騰し、その後1年間は反動で下落するという動きがあったため、それを除外すべきという主張があった。

また、諸外国では、金融政策について物価の趨勢となる指標を判断している機関は中央銀行であるが、統計作成機関がオプションとして食料とエネルギーを幅広く除外した指数を作成しているところが多く、OECDもそのような指数をコア指数として各国比較している。

このようなことから、日本も食料とエネルギーを幅広く除外した指数を公表することとした。当時原油価格は高騰しており、また、このような指数は、通常公表されている数値から作成することが可能であったことから、基準改定を待たずに2005年12月から公表数値から計算した「食料（酒類を除く）及びエネルギーを除く総合」を作成・公表した。

本節では、それらの状況について記載した Makoto Shimizu[77]の一部と清水・永井[82]の研究を示す。

コア指数については中央銀行を中心に数多くの先行研究が存在している。日本で採用した除外法に限っても、例えば、Vega and Wynne[83]でEUについて分類ごとに指数や変動率の標準偏差を計算・比較している。Makoto Shimizu[77]及び清水・永井[82]には、それらを2001～2005年の日本のCPIに適用した分析を含んでいる。

4.1 コア指数の作成状況

物価の不規則な変動を取り除いた基調的な動向を把握するための指標はコア指数と呼ばれることがあり、各国・地域で作成されている。「コア」の意味は、英語では通常 core inflation であり、underlying inflation とほとんど同じ意味である。また、コア指数は、統計作成機関が「コア指数」として公表しているのではなく、中央銀行がそう呼び利用している場合が多い。

しかし、統計作成機関としても、食料又はその一部やエネルギーなどを除いた指数、変動率の激しい品目を除外した指数等、中央銀行、エコノミスト等の利用者がコア指

数として利用できるような指数を作成・公表している国・地域が多い。しかも、米国、フランス等ではそのような指数を複数作成・公表している。

CPI マニュアルにおいても、経済分析のために一時的要因を除いたコア指数を構築することが望まれている。ただし、対象品目を絞り過ぎて代表性を喪失しないようにすることが必要であるとしている。

コア指数を作成する方法はいくつかあるが、大きくは特定品目を除く方法と統計的処理を施す方法に区別される。以下ではこれらの概要を紹介する。

4.1.1 特定品目を除く方法

総合指数からいくつかの品目を除外して作成する方法は除外法と呼ばれる。ここでは各国・地域で用いられる除外法のうち、毎回常にあらかじめ定められた特定の同じ品目を除外する方法を紹介することとする。

除外法を採用している国・地域では、統計作成機関が食料、エネルギー又はそれらの両方を除外した指数を作成・公表している場合が多い。

除外品目を分野ごとに細かく見ると、以下のとおりである。

(ア) 食料について

食料については、米国、EU を始め、多くの国・地域の統計作成機関で生鮮食品のみならず加工食品や外食などを幅広く除外した指数を作成している。また、米国、EU のほか、ドイツ、メキシコ、ブラジル等の中央銀行が用いる指数も、食料を幅広く除いたものである。

OECD でも、標準的コア指数として国際比較する際に、日本を含め食料を幅広く除外したものを用いている。

これに対して、中央銀行が食料のうち生鮮食品のみを除外した指数を用いている国は、イタリア、タイ等少数である。また、オーストラリアの中央銀行が用いる指数の除外対象は野菜及び果物と更に狭いが、このような国は例外的である。ただし、英国、ドイツ等では、中央銀行が用いなくても統計作成機関が食料について生鮮（季節）食品のみを除外した指数を作成している。また、英国、EU 等ではオプションとして非加工食品を除く指数を作成している。

なお、日本と韓国以外の国においては、「食料」と言う場合に 93SNA で示された COICOP の分類に従い酒類が除外される。このため、食料を除外している国・地域でも酒類を除外せずに含めている場合がほとんどである。

(イ) エネルギーについて

エネルギーについては、需要面との関係で様々な議論が存在するが、天候や戦争など経済の趨勢とは異なる一時的な要因で変動しやすい面があることから、米国、EU

を始め、多くの国・地域の統計作成機関がエネルギーを除外した指数を作成している。また、米国、EU のほか、ドイツ、イタリア、メキシコ、ブラジル等の中央銀行が用いる指数も、エネルギーを除いたものである。ほとんどの国・地域でエネルギーの対象はガソリン、燃料油、電気代及びガス代となっている。ただし、オーストラリアはガソリンだけ、韓国では石油製品だけを除外した指数を金融政策に用いている。

日本においては、自動車のガソリン使用量が米国と比べて少ないこともあり、エネルギーに関する指数を試算してもほとんど変動しない時代が続いた。しかし、2004年5月以降、原油価格が上昇したことから、その分を一時的要因とみなして除外した物価動向を把握することに対するニーズが強まった。2005年基準改定の計画についても、石油製品等のエネルギーを除外した指数を作成すべきであるという指摘があった。

(ウ) その他について

食料やエネルギーのほかに、税金、住宅金利の支払額、輸入品、政府料金、政府統制価格などを除外している国・地域がある。中央銀行では、ニュージーランドが金利の支払額を除外したもの、スウェーデンが住宅金利の支払額を除外し間接税及び補助金による影響を調整した指数を用いている。また、フランス及びEUの統計作成機関は、煙草を除外した指数を作成している。

ドイツを始めこれらを食料やエネルギーとともに除外している国・地域は多い。中央銀行では、ノルウェーは税調整を施した上でエネルギーを除外した指数、EUは食料とエネルギーに加え税金を除外した指数を用いている。

なお、除外分野という観点から統計作成機関が作成している系列を見ると、「持家の帰属家賃を除く総合」のように他にも数多くの種類が存在するが、これらがすべてコア指数を作成するという意図のもとで除外されていないという点に注意が必要である。

4.1.2 統計的処理を施す方法

上記のようなあらかじめ定められた品目を除外するのではなく、毎回統計的処理を施す方法もある。

例えば、変動率が平均から標準偏差の数倍以上離れた品目を除外した指数、変動率の大きな品目を一定割合だけ除外した指数 (trimmed means) はその典型である。さらに、極端な場合として、この割合を 50% にして変化率の中央値を用いた指数 (weighted median) がある。それらは除外法の一つである。カナダの中央銀行は、変動率の大きな 8 品目を除外した上で残りの品目について間接税による影響を調整したものをコア指数としている。

また、変動率にその大きさに反比例するようなウェイトを付与して平均する方法（volatility weighted means）、通常のウェイトを変動率の大きさに割ったものをウェイトとして指数を作成する方法（double weighted means）もある。

他に、移動平均法を用いる方法もあるが、移動平均法には系列の両端で欠損値が生じるという問題がある。

なお、趨勢や傾向を見るときに季節調整値を利用するときがあるが、季節調整値は、12か月を周期とする変動を除いたものなので、趨勢・傾向変動の他にその時々の変動値を含んだものとなっている。

さらに、毎月の指数の変化を経済モデルの組み合わせで表現する方法もある。この方法には、政策との関連をつかみやすいという長所があるが、概念がわかりにくい、最新データがモデルそのものに影響する等の欠点もあるので、実用されることはほとんどない。

4.2 コア指数の条件

これまでに紹介したような特定品目を除外した指数や統計的処理を施した指数は数多く試算することができるが、これらをコア指数と呼ぶにはいくつかの条件があるとされている。ここでは、このようなコア指数に必要な主な条件について紹介することとする。

最初の条件として適時性が挙げられる。これは、CPIが公表されると同時に公表されるということであり、事後のデータが揃ってから試算が可能という方法では不適当ということである。

次に信頼性が必要である。これは、金融政策を担当する中央銀行から独立した機関で確定されることにより保たれるとされる。

さらに、理解のしやすさも必要である。これは、国民にわかりやすいということであり、この観点を踏まえると、除外法は優れた方法と言えるが、この場合にも除外対象を個別に列挙するよりもある程度の幅のある分類で括るほうが良いということになる。

透明性が重要であるという主張もある。これは、公表されたデータを用いて誰もが計算できるという意味である。

ここまではコア指数に限らずCPIそのものがある程度備えるべき条件でもある。コア指数特有の条件としては予測可能性がある。これは、予測すべき指標が総合指数であるならば、コア指数は総合指数を予測できるということであり、少なくとも総合指数そのものよりも予測可能性が高いということである。予測可能性を評価するには、コア指数から推定される予測値と総合指数の実績値との乖離を検査する方法等がある。

これらについては数多くの実証分析があるが、結果はデータの採用期間に依存するという点に注意が必要である。いくつかの国ではインフレターゲットとしてコア指数を使用しているが、その場合にはそれ自体の予測可能性が問われる。

不偏性も重要である。これは総合指数からの偏りが無いことであり、統計的検定を満たしていることが必要である。

他にも、コア指数の変化率が異常値と関係があるかどうかを検査するための Granger 因果性、総合指数とコア指数の関係が持続的であるかを検査するための共分散等の統計的検定がなされることもある。

逆に、条件を満たす指標が複数ある場合にコア指数は1つに限定しなければならないという決まりはなく、また、時代によって適切なコア指数が変わることもある。諸外国には、コア指数の条件を満たす指標を複数提供することによって利用者の便宜を図る統計作成機関も存在する。

なお、本節で述べた条件は実用上の視点を踏まえたものであるが、コア指数については、もちろん、理論的基礎が必要であるとする主張もある。ただし、コア指数は最近理論を欠如したまま実用面で多くの計算方法が開発されてきたという実情がある。

4.3 日本の状況

日本では、生鮮食品の変動が極めて不規則なことから、物価動向の基調を表す指標としてこれまで「生鮮食品を除く総合」を公表してきたところである。「生鮮食品を除く総合」は、上記のコア指数の条件をかなりの程度満たしており、実際これを「コア指数」と呼ぶ利用者も存在する。また、日本は先進国の中では生鮮食品、特に、生鮮魚介のウエイトが高いという特徴に配慮が必要である。このため、物価指数全体が天候による影響を受けやすく、逆に生鮮食品を除くと毎月の不規則な大きな変動を除外することができる。

しかし、通常物価指数で傾向を捉える際には前年同月比を見る場合が多く、少し長い期間で見れば他にも天候等で変動しやすい品目がいくつか存在している。

財・サービス分類別に 2001 年から 2005 年までの毎月の指数及び前年同月比の標準偏差を比較すると、表 29 のとおり、指数、前年同月比とも最大は米類を含む他の農水畜産物であり、石油製品（ガソリン、灯油及びプロパンガス）が続いている。これらは生鮮商品（生鮮食品に生鮮肉、鶏卵及び切り花を加えたもの）より大きい。なお、前月比の標準偏差が最大になるのは繊維製品であるが、これは主に季節性によるものであり、不規則性が強いという意味ではない。

表 29 月次指数の標準偏差（2001～2005 年）

			2000年=100			
分類		ウエイト (1万分比)	標準偏差			
			指数	前月比(%)	前年同月比(%)	
財	農水畜産物	生鮮商品	671	3.4	2.8	5.1
		他の農水畜産物	104	6.9	1.8	10.5
	工業製品	食料工業製品	1,373	1.1	0.1	0.4
		繊維製品	564	3.3	3.3	1.4
		石油製品	301	6.5	1.0	4.8
		他の工業製品	1,414	2.7	0.1	0.8
		電気・都市ガス・水道	482	1.6	0.3	1.2
		出版物	169	0.0	0.2	0.2
サービス	公共サービス	公営・公団公社家賃	48	0.6	0.2	0.6
		家事関連サービス	301	1.2	0.4	1.9
		医療・福祉関連サービス	203	4.1	1.1	3.8
		運輸・通信関連サービス	565	1.6	0.7	1.5
		教育関連サービス	60	2.1	0.4	0.2
		教養娯楽関連サービス	118	0.8	0.4	0.9
	一般サービス	外食	620	0.5	0.2	0.6
		民営家賃	300	0.4	0.0	0.1
		持家の帰属家賃	1,360	0.2	0.1	0.3
		家事関連サービス	482	0.3	0.1	0.1
		医療・福祉関連サービス	32	1.2	0.2	0.8
		教育関連サービス	328	1.0	0.2	0.3
		教養娯楽関連サービス	506	2.3	2.1	0.6

他方、先述のとおり、多くの国で、物価動向の基調を表す指標として食料とエネルギーを幅広く除外したものが作成されている。

以上のことから、「生鮮食品を除く総合」に加え、「食料（酒類を除く）及びエネルギーを除く総合」を新たに公表することとした。この場合のエネルギーは石油製品と電気代及び都市ガス代である。

ただし、この新たな指標は物価動向の基調を表す指標のオプションの1つであり、CPI そのものを「総合」とし、また、趨勢を見るための指標として「生鮮食品を除く総合」を前面に掲げる従来のスタンスに変更はない。

なお、「食料（酒類を除く）及びエネルギーを除く総合」は、指数の対象範囲に注目して「米国型コア指数」と呼ばれる場合もあるが、日本と米国では生活が異なるため、除外した品目、対象となる品目は完全に一致しないという点に注意が必要である。例えば、米国には羊肉や中古車という品目があるが、日本にはない。逆に日本にはだいたい餅や外国バック旅行という品目があるが米国にはない。また、米国では「柑橘類」、「掛け時計、ランプ及び装飾品」等、日本の品目をいくつか大括りにした品目があることに加え、品目を含む分類が両国で異なる場合もあること等から、同じ分類であっても、両国で細部を比較する際には注意が必要である。

4.4 「食料（酒類を除く）及びエネルギーを除く総合」の動向

最後に「食料（酒類を除く）及びエネルギーを除く総合」の長期的傾向を「総合」や「生鮮食品を除く総合」と比較しながら概観することとする。

遡ること第1次石油危機の1974年には「生鮮食品を除く総合」が前年よりも23%近く上昇したのに対し、「食料（酒類を除く）及びエネルギーを除く総合」は20%弱の上昇に留まっている。それでも、その時代の特徴として、「石油危機」と呼びながら石油製品の高騰に留まらず他の製品・サービスの価格も高騰したという点が挙げられる。また、第2次石油危機の1980年には、第1次石油危機の教訓を受けた物価対策が先行したことから、「生鮮食品を除く総合」は7.5%の上昇に留まったが、「食料（酒類を除く）及びエネルギーを除く総合」の上昇幅は更に低く7%を下回っていた。その後の原油価格の高騰が激しくない時代には、いずれの変化幅も小さくなったこともあり、両者の差も小さかった。

しかし、2005年前後に原油価格が高騰していることから、「食料（酒類を除く）及びエネルギーを除く総合」は、前年同月比で「生鮮食品を除く総合」を下回っている。2000年基準で前年同月比を見ると、「生鮮食品を除く総合」は2006年に入り毎月0.5%以上で推移しているが、「食料（酒類を除く）及びエネルギーを除く総合」は0.1%又は0.2%で推移している。

また、両指標の関係は地域によって異なる。CPIの特徴として、毎月東京都区部について全国の先行指標として中旬速報値を公表している。しかし、「生鮮食品を除く総合」に石油製品が含まれていることから、前年同月比で見ると、自動車等の利用が比較的少ない東京都区部では、2004年9月以降毎月全国よりも低くなっている。これを「食料（酒類を除く）及びエネルギーを除く総合」で見ると、両者の間の大小関係は一定ではなくなる。

5 新旧指数の差の要因分解

基準改定に関する最後の節として、基準改定により同一年月の指数⁶⁴や前年同期（月）比が変化したことを受け、前年同期比を中心にその程度及び要因を説明することとする。2005年基準改定については、Makoto Shimizu[77]で事前の試算を行い、清水・永井[84]及びMakoto Shimizu[85]で事後の分析をしている。海外を含めて基準改定がこれほど注目されたことは、それまでなかったことから、そこで記述したような新旧基準の差に関する詳細な要因分解は初めての試みであった。本節では、それらの研究内容を示すとともに、その後の状況を簡単に付記することとする。

⁶⁴ 同一年の平均を100にして比較した場合であっても変化するという意味である。

2005年基準指数は2005年1月から作成されているので、指数についてはその月から、前年同月比については2006年1月から存在する。また、2000年基準指数は2006年12月まで作成されている。

2006年7月までにおいて新旧基準に基づく指数及び前年同月比を比べると、2005年基準指数のほうが概して低くなった。しかし、その大きさは品目ごとに、月によって異なることから、基準改定による変化を評価するには、総合指数に近い括りで数か月まとめて比較をすることが必要である。2006年1～6月平均について生鮮食品を除く総合の前年同期比で比較してみると、2005年基準は0.0%と、2000年基準の0.6%より0.6ポイント低くなった。この要因は、(1) 指数を100にリセットすることによる影響、(2) ウェイトを変更することによる影響、(3) チューハイ、カーナビゲーション、DVDレコーダーなど34品目を追加し、ミシン、鉛筆、ビデオテープなど48品目を整理統合した影響などが考えられる。

過去に遡り、現行系列と同一概念で主系列を作成し始めた1985年基準以降について、基準年の翌年1～6月における生鮮食品を除く総合の前年同期比の基準改定による下げ幅を見ると、当時は2005年基準改定の0.6ポイントが最大であった。また、微妙な物価の変化が注目される中で、変化の仕方が0.0からマイナス0.6に変わるといふ衝撃とも受け取れる改定であった。

そこで、本研究では、従来の改定時には行っていなかった定量的な要因分解を初めて行った。

5.1 要因分解の考え方

前年からの変化の要因を見るために品目ごとに寄与度を分解すると次のように表すことができる。

品目Aの寄与度

= 品目Aの前年比

$$\times \frac{\text{前年の品目Aの指数}}{\text{前年の生鮮食品を除く総合の指数}} \quad (5-1)$$

$$\times \frac{\text{品目Aのウェイト}}{\text{生鮮食品を除く総合のウェイト}} \quad (5-2)$$

生鮮食品を除く総合の前年からの変化は、そこに含まれるすべての品目について上式の寄与度を合算したものになる。

基準年翌年において基準が改定されると、

- (1) 基準時点を変更することに伴い(5-1)の分母・分子がともに100にリセットされることによる影響

(2) (5-2)の分母・分子のウェイトが 2000 年から 2005 年に変化することによる影響

(3) その他の影響

により生鮮食品を除く総合の前年同期比は変化する。

その他の影響として考えられる大きな要因は、品目内のウェイトの変更と品目の改廃による影響である。

航空運賃や電気代などの料金関係品目等については、価格の変動を的確に指数に反映させるため、業務統計などの資料を用いてモデル式により月々の価格指数を算出している。これらの品目においては、基準年の間で料金体系や世帯での利用形態等が変化したことを受け、より適切な指数を推定するためにモデル式における料金の区分とそのウェイトを変更したことによる影響がある。

また、モデル式による品目でなくても、基準年の間で市町村間の消費金額の分布や消費者が購入する商品の種類が変化しているため、それに合わせて市町村別ウェイトや商品の種類ごとのウェイトも変更したが、それによる影響も存在する。

5.2 2005 年基準改定への適用

ここでは、基準改定による改定幅の目安として、生鮮食品を除く総合の 2006 年 1～6 月平均の前年同期比の旧基準値と新基準値の差に注目することとする⁶⁵。旧基準における当指数に含まれるすべての品目の寄与度の合計は 0.52（生鮮食品を除く総合の前年同期比を直接計算すると 0.62⁶⁶）である。

寄与度を上記の 3 つの要因ごとに区分すると表 30 及び表 31 のようになる。

(1) 指数を 100 にリセットする影響

品目ごとに(5-1)式の分母・分子をともに 100 にリセットすると⁶⁷、プラス幅は 0.52 が 0.29 となり 0.23 縮小する。これは、パソコンやプリンタなど IT 関連品目が大幅に下落し、2000 年基準指数が小さくなっていることが主な要因である。ちなみに 2005

⁶⁵ 寄与度の分解式は、最小単位の指数、すなわち、月次の市町村別指数に当てはめればそのまま適合するが、半期や年次、又は広範囲に及ぶ指数であれば、指数の合算を最初に行うことになるので、各品目の前年同期比は新旧基準で一致するとは限らない。もちろん、計算結果には計算過程の四捨五入による影響も含まれる。したがって、本節の試算結果は幅広に見る必要がある。

⁶⁶ この差は指数を小数 1 位まで四捨五入してから半期平均や前年同期比を計算していることによる。

⁶⁷ 当指数は 1～6 月の値であることから、基準を 2005 年にしても厳密には①の分母・分子は 100 になるとは限らない。しかし、いずれも 100 に近い値（90 以上 114 未満、全体の 4 分の 3 程度が小数点以下の四捨五入により 100）になることから、便宜上 100 であると想定して要因分解を行った。

年1～6月における2000年基準のパソコン（ノート型）の指数は17.4であることから、これが100にリセットされることによって、マイナス幅を約6倍に拡大させることになる。

（2）品目ウェイトを変更する影響

（5-2）式の分母・分子のウェイトを消費金額の変化に沿って2000年から2005年に変更すると⁶⁸、プラス幅は0.52が0.64となり0.12拡大する。一般にCPIのようなLaspeyres指数にはウェイトを固定することによる上方バイアスが発生すると考えられているが、この結果によると、ウェイトを固定したことによるバイアスは下方バイアスであったことになる。これを品目ごとに見ると、ウェイトの改定による上昇幅は灯油で最も大きく0.06となっている。灯油については近年の原油価格の高騰を受け、指数が30.5%も上昇したが、他のエネルギーへの代替が進まず、消費量の抑制につながっていないことが改定幅の引き上げ要因になった。石油製品としてはガソリン（レギュラー）も0.04上昇するなど、全体としては0.10の上昇となった。また、通所介護料は2005年に食費に対する補助がなくなったことにより値上がりしたが、需要は減少しないため、0.03上昇した。逆に、下落要因については、携帯電話通信料などがあるが、その幅が0.02以上に達する品目は存在しなかった。このように、品目ごとに見ると、価格変化が需要増減に結びつかない消費構造がうかがえる。

（1）と（2）の2つの要因は、明確に区別することができず、基準改定に伴い双方同時に発生している点に注意が必要である。（5-1）式の分母・分子をともに100にリセットし、（5-2）式の分母・分子のウェイトを2000年から2005年に変更することにより（この操作は連鎖指数の前年同期比を計算することと同じ）、（1）と（2）の両方を評価すると、プラス幅は0.52が0.41となり-0.11変化する⁶⁹。

（3）その他の影響

上記（1）及び（2）は、品目を2000年基準に固定したまま2005年の消費割合を反映させた場合に、前年同期比にどの程度の変化があるかを見たものである。ここでは、これを前提にした上で更に2005年基準の品目の改廃を行うと前年同期比にどの程度の変化があるかを概観することとする。

（1）及び（2）以外の要因はすべて（3）であるとすると、（3）に相当する下落幅は全体の下落幅-0.52から上記の-0.11を引いた-0.41に相当する。ここでは、

⁶⁸ 2000年基準ウェイトの品目区分に従い2005年家計調査の結果を踏まえて換算したウェイトであり、連鎖指数や地域差指数を計算するとき用いられるウェイトと同じ方法で作成したものである。したがって、2005年基準ウェイトとは異なる。

⁶⁹ Makoto Shimizu[77]では、2005年までのデータを基に影響を見ているが、それによると（1）により-0.23、（2）により0.08、（1）&（2）により-0.10であり、傾向としては今回の結果と類似している。

この-0.41 に相当する下落幅を要因分解した結果を表 32 に示す。品目の改廃は、単に当該品目が追加、削除又は統合されるのみならず、継続品目、すなわち、新旧基準間で同一の品目のウエイトや指数にも変化を及ぼすという点に注意が必要である。新旧基準の比較において、旧基準については（1）及び（2）で評価した品目を 2000 年基準に固定したまま 2005 年に換算したウエイト及び指数、新基準については基準改定での品目の追加及び整理統合等も加味し新たに計算したウエイト及び指数を意味する。

まず、品目ウエイトとして同一時点の消費金額から作成する場合であっても、追加品目と整理統合品目があるため、継続品目に割り当てられるウエイトは相対的に変化する。したがって、継続品目であっても、2005 年基準のウエイトは、2000 年基準に沿って 2005 年において換算されたウエイトと異なる。

また、2005 年において換算されたウエイトは、家計調査の品目を細分した品目については、2000 年基準の配分率に従って区分しているため、2005 年における消費の実態を踏まえた 2005 年基準の配分率とは異なっている。⁷⁰ 品目ごとのウエイトは市町村単位で計算された後で全国に合算されるが、これらの配分率は市町村単位で決められる。

このような品目内のウエイトの相違は、当該品目のウエイトのみならず指数にも影響するので、当該品目の指数も異なっている。

ア モデル式の改定

まず、継続品目のうちモデル式を改定することによる影響を評価することとする。新旧基準ともにモデル式を用いている品目⁷¹について寄与度の合計から（1）及び（2）による影響を差し引いた分がすべてモデル式を改定したことによる影響であるとする。その幅は-0.12 となる⁷²。

下落幅が最大となる品目は移動電話通信料である。移動電話通信料の寄与度から（1）及び（2）による影響を差し引くと-0.14 となる。移動電話通信料では、家計調査のデータより、1 支払い当たりの金額の分布を作成し、それを基に価格を調査する利用パターン（通話時間及び通信量の組み合わせ）を決めている。移動電話の利用が増大し、世帯での支出金額分布が 2000 年に比べ、2005 年に 1 支払い当たりの金額が高い方にシフトしたため、調査価格も多く利用するパターンにシフトした。2005

⁷⁰ 配分率は基準年の家計調査の調査個票を特別集計した結果と業務統計を基に設定している。

⁷¹ 新旧基準で品目名が同じものに限る。旧基準の「通所介護料」については新基準では「介護料」なので、品目の改廃による影響に分類した。

⁷² 先述したとおり、新基準ウエイトは 2000 年基準に沿って 2005 年に換算したものと必ずしも一致しないので、算式だけの影響ではない点に注意が必要である。

年 11 月に一部の社で通信量の多いプランで大幅な値下げがあったために、2005 年基準の方が値下げの影響を強く受けたことが主な要因である。

モデル式を改定した品目は、必ずしもプラス幅の縮小に寄与しているものばかりでなく、外国パック旅行などはプラス幅の拡大に寄与した。

イ その他の品目内ウエイトの変更

このような品目内ウエイトの変更はモデル式だけに限ったことではない。モデル式を用いていない継続品目について（1）及び（2）による影響を差し引くと -0.14 となる。

これには様々な要因が存在する。まず、品目内の市町村別ウエイトを改定したことが挙げられる。持家の帰属家賃（木造中住宅）では、価格（家賃）が下落している市町村でウエイトが増加し、価格（家賃）が上昇している市町村でウエイトが減少したために、全国で見ると -0.05 となっている。

また、品目内の商品の市場における出回りが変化したためにウエイトを改定した影響もある。例えば、カメラについては、フィルム式からデジタル式への移行が急速に進んだことを受け、2000 年基準では両者を加重平均して指数を作成していたのに対し、2005 年基準ではデジタル式だけを対象とすることとした。デジタルカメラの価格は急速に下落しているので、カメラ全体では -0.01 に相当する。

ウ 品目の改廃による影響

品目の改廃による影響は、残りの -0.15 に含まれるが、この幅は新旧基準で品目名が一致しない品目の寄与度を合算したものから（1）及び（2）の影響を除外した数値であり、いわば残りの影響をまとめたものに過ぎない。これから品目の改廃による影響だけを取り出すのは困難である。CPI の品目は突然発生・消失するのではなく、追加品目は過去に存在していた品目が切り出されて組み合わせられて誕生し、逆に整理統合品目は他の品目に分割・合併されるという形をとるが、その流れを明確に把握することが困難だからである。例えば、2000 年基準の果物缶詰（もも）は果物缶詰（みかん）に整理統合されて 2005 年基準で果物缶詰となるが、2005 年基準で果物缶詰（もも）に相当する部分がどの程度とみなすかによって評価が変わる。

しかし、いくつかの追加品目については、切り出しのもととなる品目が推定できるので、追加したことによる影響を評価することができる。例えば、追加品目のテレビ（薄型）については、同じウエイトだけブラウン管テレビが調査されていた場合と比べた影響を試算すると -0.05 となる。DVD レコーダー、録画用 DVD についても、同じウエイトだけビデオテープレコーダー、ビデオテープが調査されていた場合と比べ

ると⁷³、それぞれ -0.02 、 -0.01 となる。移動電話機については、同じウエイトだけ固定電話機が調査されていた場合と比べても -0.00 とほとんど影響がない。

(3)には、他にも様々な要因が存在するが、それだけを取り出して評価することは困難である。

5.3 2000年基準改定への適用

次に、同様の比較を1995年基準から2000年基準に改定したときに適用してみることにとする。

まず、1995年基準の2001年1～6月における生鮮食品を除く総合に含まれるすべての品目の寄与度の合計は -0.54 （生鮮食品を除く総合の前年比を直接計算すると -0.59 ）である。

(1) 指数を100にリセットする影響

上式①の分母・分子をともに100にリセットすると、マイナス幅は -0.54 が -0.67 となり変化幅は -0.13 である。

(2) 品目ウエイトを変更する影響

上式②の分母・分子のウエイトを2005年から2000年に変更すると、マイナス幅 -0.54 は変わらない。

(1)と(2)の両方を評価すると、マイナス幅は -0.54 が -0.68 となり変化幅は -0.14 である。

(3) その他の影響

その他の影響は、全体の下落幅 -0.29 からこの -0.14 を引いた -0.15 に相当する。

ア まず、新旧ともにモデル式を用いている品目の寄与度から(1)及び(2)による影響を差し引いた分がすべてモデル式による影響であると仮定すると、その影響は 0.05 となる。

イ その他の品目のうち、継続品目について(1)及び(2)による影響を差し引いた分を計算すると 0.02 となる。

ウ 残りは -0.22 であるが、追加品目による影響としては、パソコン（デスクトップ型、ノート型）を導入したことによる影響が -0.19 あった。

ただし、これらを基準間で比較する際には、モデル式の適用範囲や改廃品目のウエイトが各基準間で異なっているという点に配慮が必要である。

(4) 2005年基準改定との比較

⁷³ ビデオテープレコーダー及びビデオテープは2005年基準改定の整理統合品目なので2000年基準の前年同期比を用いて試算した。

表 30、表 31 及び表 32 によれば、今回の改訂幅を 5 年前と比較すると、改定幅は（ $-0.52 - -0.29 =$ ） -0.23 拡大したが、その内訳を区分すると、指数を 100 にリセットすることで（ $-0.23 - -0.13 =$ ） -0.10 、モデル式による影響で（ $-0.12 - 0.05 =$ ） -0.17 、その他の継続品目内ウエイトの変更等で（ $-0.14 - 0.02 =$ ） -0.16 拡大し、品目ウエイトを基準年に換算することにより（ $0.12 - 0.00 =$ ） 0.12 、品目の改廃等で（ $-0.15 - -0.22 =$ ） 0.07 縮小した。

表 30 各種試算値 生鮮食品を除く総合 1～6月平均の旧基準指数の前年同期比(%)

年	基準年	各品目の寄与度の合計	(1) 前年指数を 100 にした場合	(2) ウエイトを前年のものにして計算した場合	(1) & (2) 前年指数を 100 にしてウエイトを前年のものにして計算した場合
2001	1995	-0.54	-0.67	-0.54	-0.68
2006	2000	0.52	0.29	0.64	0.41

表 31 生鮮食品を除く総合 基準年翌年 1～6月平均の前年同期比(%)への影響

基準年	計	(1) 前年指数を 100 にした場合	(2) ウエイトを前年のものにして計算した場合	(1) & (2) 前年指数を 100 にしてウエイトを前年のものにして計算した場合	(3) その他
2000	-0.29	-0.13	0.00	-0.14	-0.15
2005	-0.52	-0.23	0.12	-0.11	-0.41

表 32 生鮮食品を除く総合 基準年翌年 1～6月平均の前年同期比(%)への影響

表 31(3)の要因分解

基準年	計	ア モデル式のウエイト等	イ その他の継続品目内ウエイト等	ウ 品目改廃等
2000	-0.15	0.05	0.02	-0.22
2005	-0.41	-0.12	-0.14	-0.15

5.4 その他の指数の差について

2006 年 1～6 月平均の前年同期比を総合指数で見ると、2005 年基準は 0.0% であり、2000 年基準の 0.5% より 0.5 ポイント低かった。また、食料（酒類除く）及びエネルギーを除く総合で見ると、2005 年基準は -0.5% であり、2000 年基準の 0.1% より 0.6 ポイント低かった。食料（酒類除く）及びエネルギーを除く総合は、総合指数や生鮮食品を除く総合と比べて価格下落の激しい品目に対するウエイトが相対的に高くなるので新旧基準の差が大きくなりやすく、1 月には 0.8 ポイントにも達していたが、半期全体では 0.6 ポイントに留まった。

他方、東京都区部について新旧基準の差を見ると、2005年基準のほうが総合では0.4ポイント、生鮮食品を除く総合では0.5ポイント、食料（酒類除く）及びエネルギーを除く総合では0.5ポイント低くなったが、いずれも全国より改定幅は小さかった。これは、石油製品や教養娯楽用耐久財など新旧基準差に強い影響を及ぼす品目のウエイトが、東京都区部では全国よりも小さい等の事情によるものである。

5.5 2010年以降の基準改定

今回の基準改定による前年同月比の改定幅は指数や月により異なるものの、総合指数に近い括りで見ると0.5ポイント前後と、物価の変化がプラスかマイナスかという微妙な時期においては相当大きなものを感じられた。

消費者の行動形態を過去と比較すると、価格変化が発生した場合、近年では品目間での代替行動は減少したものの、品目内で下落率の高い商品への消費者の選択が急速に進んでいる様子がうかがえる。このような消費行動の変化は通常のLaspeyres式には現れないが、品目ごとの指数の変化に大きな影響を及ぼしている。

2000年～2005年の間に現れたような消費行動の変化は今後も続くと考えられるので、その後の基準改定においても指数や前年同期比のある程度下落は避けられない。特に、2005年基準改定でテレビ（薄型）やDVDレコーダーなど下落率の高い品目を追加したことから、指数を100にしたときの前年同期比の下方シフトは避けられないと考える。

しかし、2005年基準改定により、生鮮食品を除く総合については消費行動をより迅速に反映する連鎖指数を毎月作成することとなったことから、少なくとも(1)と(2)に相当する分については、改定直前においても変化幅を推定することが可能となる。

その後、2010年基準改定、2015年基準改定と、5年ごとに基準が改正された。生鮮食品を除く総合について、基準年の翌年の1～6月平均の前年同期比の差を新旧基準で比較すると、表33に示したように2010年基準改定時には-0.8ポイントと、2005年基準改定時よりも下落幅が大きくなったが、2015年基準改定時には新旧基準で同じであった。いずれにおいても、中間年見直しの本格適用などにより新旧基準差の要因分解は2005年基準改定時ほど単純には表現できないものだが、新基準の指数の作成方法を事前により詳細に公表することにより、新旧基準差については市場で冷静に受け止められているように見える。

表 33 基準年の翌年の 1～6 月平均 (%)

基準年	総合			生鮮食品を除く総合		
	旧基準	新基準	差	旧基準	新基準	差
1985	1.1	1.2	0.1	1.3	1.3	0.0
1990	3.8	3.5	-0.3	3.3	3.1	-0.2
1995	0.0	-0.1	-0.1	0.2	0.1	-0.1
2000	-0.3	-0.6	-0.3	-0.6	-0.9	-0.3
2005	0.5	0.0	-0.5	0.6	0.0	-0.6
2010	0.1	-0.5	-0.6	0.2	-0.6	-0.8
2015	-0.2	-0.1	0.1	-0.2	-0.2	0.0

6 家賃指数の回帰式による推定方法の検討

2015 年基準の CPI について、家賃は 17.8% のウェイトを占めており、CPI 全体の動向に極めて重要な要因となっている。このうち、小売物価統計調査の家賃調査による実際の借家の家賃は 2.8% のみで、その他の大半は持家の帰属家賃であるが、借家の家賃の指数が帰属家賃の指数にそのまま適用されるので、家賃調査の結果は CPI 全体を左右するものとなっている。

家賃調査のうち民営家賃については、全国約 170 の調査市町村から家賃調査地区を選定し、居住者のいるすべての民営借家約 26,000 世帯を調査対象としている。家賃調査地区は 3 群に分け、各群地区を 3 か月ごとに⁷⁴調査している。

CPI における住宅関連指数の作成はどの国でも重要な課題となっている。このため、家賃を住宅に関する様々な要因で回帰分析する研究が各国で行われており、Hoffmann and Kurz[86]は、西ドイツにおいて家賃をヘドニック法やマッチドモデル法で品質調整する方法を試みている。

日本でも、1970 年から CPI に持家の帰属家賃が導入されたが、その計算過程で住宅・土地統計調査の個別データを用いて民営家賃（設備専用）を被説明変数、延べ床面積、建築時期等を説明変数として回帰分析している。しかし、小売物価統計調査の個別データを用いて家賃そのものを品質調整する研究は行われていなかった。

Makoto Shimizu[87]では、小売物価統計調査の家賃調査の個別データを用いて家賃の品質調整方法を検討し、東京圏⁷⁵において毎月の家賃、床面積、地価、人口変化率、昼間人口、1970 年以前の住居割合、耐震工事率などから回帰式を構築するとともに、持家について帰属家賃方式の別の選択肢として、住宅購入費用を平均する Acquisition 方式の試算を行い、2006 年に開催された Ottawa Group 会合で発表した。

その後、久我真理子[88]は大阪圏⁷⁶で同様の回帰式を構築している。

⁷⁴ 開始月は 1 月、4 月、7 月、10 月

⁷⁵ 東京都、埼玉県、千葉県、神奈川県

⁷⁶ 大阪府、京都府、兵庫県、奈良県

さらに、Makoto Shimizu[89]では、家賃調査で2008年1月から全面適用された新家賃調査地区において新たに建築年を調査し始めたことから、対象を全国に広げ、2009年に開催されたOttawa Group 会合での意見を踏まえ、説明変数を床面積、建築年、人口密度に絞って回帰式を構築し、それを用いて家賃指数の推定と欠測値補完を行った。Makoto Shimizu [89]は、日本において家賃を品質調整した研究であるとともに、回帰式を欠測値補完に適用した点に新規性がある。

以下では、Makoto Shimizu [89]の研究結果を、ヘドニック法の概要を含めて示す。

6.1 家賃指数の問題点

CPI では、家賃指数の問題点として、住居区別に市町村内の調査対象が少ない場合に、借家の転出入により指数が大きく変動することがある。特に、単位面積当たりの家賃が極端に大きい又は小さい借家に転出入がある場合、この影響は大きい。

この問題を解決するために、全国の家賃指数を利用し、品質を一定にすることが困難な小標本地域の家賃指数の改善を目指すこととする。このため、家賃を床面積、建築年及び人口密度で説明する回帰式を利用して、家賃を推定する方法及び転出入前後の家賃の欠測値を補完する方法について、転出入前後の家賃の欠測値を保合により補完する方法と比較しながら検討する。

6.2 回帰式による品質調整の方法

一般に、ある商品の価格が様々な機能から成るとみなし、これらを説明変数とする回帰式で商品価格を推定する方法は、ヘドニック法と呼ばれている。これは、消費者が個々の機能に対しどの程度の価値を見出しているかを明らかにする方法である。CPI のヘドニック法の概要は清水・永井[90]に記載しており、以下はそれを要約し、最近の状況を踏まえて加筆したものである。

6.2.1 ヘドニック法の適用の条件と方法

CPI は、ILO 決議にもあるとおり⁷⁷対象商品の品質を同じにした上で価格変化を測定するものであるが、品質の調整には様々な困難が伴う。品質調整法の中で、ヘドニ

⁷⁷ 第17回国際労働統計家会議（2003年）におけるCPIに関する決議第2条には、「CPIは消費者の価格の変化を測定するものである。これはある年又は他の特別な期間における世帯の代表的支出として選ばれた製品について同一の品質及び同様の特性を備えた財及びサービスから成る固定されたバスケットを購入する費用を測定することによって成し遂げられる。」との記述がある。また、第66条には「品質変化が発見されれば、指数ができるだけ純粋な価格変化を反映するように、価格に品質調整がなされなければならない。」とされている。

ック法は価格と機能の関係を明示的に捉える方法であり、2000年基準から商品の種類が多様で変化が著しく、機能と価格の関係を把握しやすいパソコンなどに適用している。それらは、技術革新が著しく、市場の製品サイクルが極めて短いので、従来の価格取集法では同質の製品を継続的に調査することが困難である。また、製品サイクルが極めて短いために新旧の銘柄の価格を同時点で調査できるとは限らないため、オーバーラップ法を用いた指数の作成は困難である。このような場合には、ヘドニック法による品質調整が有効であると考えられる。

ヘドニック法による品質調整を行うためには、相当数の価格データの収集や特性情報の把握が必要であり、CPIでは以下の条件が当てはまる場合にヘドニック法を採用している。

- ① 必要な価格等データの収集・利用が可能であること
- ② 品質を決定する特性情報の収集・利用が可能であること
- ③ ヘドニック回帰モデルの当てはまり等が良いこと

以上の条件を勘案し、パソコン（ノート型）、パソコン（デスクトップ型）及びカメラ⁷⁸については全国の主要な家電量販店のPOS情報による全機種の特性と販売価格及び販売数量を用いてヘドニック指数（ヘドニック法による価格指数）を作成している。

具体的な計算方法は以下のとおりであり、タイムダミー法と呼ばれている。

- ①パソコン（デスクトップ型）、パソコン（ノート型）及びデジタルカメラのそれぞれについて、各機種の平均販売価格を被説明変数、ハードディスクの容量や実装メモリ容量、有効画素数など各機種の特性及び販売時点等を説明変数とする以下の片対数型の回帰モデルを設定する。

連続する2か月（ $t, t+1$ ）に関する重回帰式

$$\ln p_s = \alpha_t + \beta_t \delta_{s,t} + \sum_k \gamma_{t,k} x_k$$

$s = t, t+1$: 時点

p_s : 販売価格

$\delta_{s,t} = 0(s=t), 1(s=t+1)$: 販売時点ダミー変数

x_k : 特性 k に関する数量

$\alpha_t, \beta_t, \gamma_{t,k}$: 偏回帰係数

⁷⁸ 2003年1月から「カメラ」のうちデジタルカメラについてヘドニック法を適用。2005年以降の基準では、「カメラ」の指数をデジタルカメラのみから計算

②全国で連続する2か月に販売された全機種⁷⁹について、①の回帰モデルにより、各機種の総販売台数をウェイトとして回帰計算を行い、販売時点ダミー変数の偏回帰係数 β_t を求める。

③②で求めた販売時点ダミーの偏回帰係数から、前月を基準時とする連環指数を求める。

$$I_{t,t+1} = e^{\beta_t} \times 100$$

④③で求めた連環指数を基準時から連乗して当月の指数を求める。

なお、説明変数は1年以内の頻度で見直している。

また、パソコン用プリンタについては、POS情報を用いて上記の方法で毎月指数を作成するのではなく、通常の品目と同様に、小売物価統計調査において売れ筋の銘柄（機種）を指定して収集された価格を用いて毎月の指数を計算しているが、銘柄又はその細部規定を変更する際には、多くの場合、別に収集した各種製品の機能及び価格情報を基にヘドニック法により品質調整を行っている。

このように、主にヘドニック法を用いて品質調整を行っている品目は、上記4品目だけであり、これらの価格が下落傾向にあることから、ウェイトにして1万分で計算すると、2000年基準で61、2005年基準で46、2010年基準で41、2015年基準で28と、縮小傾向にある。

6.2.2 諸外国の状況

パソコンなどは、諸外国でも、技術革新が著しく、市場の製品サイクルが極めて短いため、新旧の銘柄の価格を同時点で調査できるとは限らない。このため、オーバーラップ法を用いた指数の作成は困難であることなど同様の問題を抱えている。

これらの品目について、ドイツを始めとする多くの国でヘドニック法が用いられており、また、ヘドニック法はCPIマニュアルでも推奨されている国際的に見て標準的な手法である。

米国及びカナダでは、通常連続する2か月で存在する機種の価格比を幾何平均で連鎖する方法を用いて指数を作成しているが、新たな機種の前月価格は把握できないのでヘドニック法を用いて前月価格を推計している。

日本では、パソコンとデジタルカメラについて、直接、品質調整後価格の変化率を推計するタイムダミー法を用いているが、タイムダミー法の適用に当たっては、価格や特性に関する充実したデータが必要とされており、POSデータの利用によってこの手

⁷⁹ 前月まで販売されていた機種及び当月から販売された機種を含む。

法を採ることが可能となっている。諸外国では充実したデータが存在しないために、この手法を採ることができず、日本のこの手法は高い評価を得ている。

2003年に、米国でパソコンの品質調整法が毎月更新される250～300の変数に関するデータベースを基に属性（特性）ごとの価値（費用）を評価し、これを用いて品質調整を行うアトリビュート（attribute）法に変更された。ヘドニック法は各特性の価値がわからない場合に、回帰分析によって間接的に推計するものであるが、アトリビュート法はその属性の価値を直接的に把握するものであるため、推計にまつわる「わかりにくさ」がないという点では優れている。しかし、供給側の価値の低下がそのまま需要側に反映するという保証はない。

米国では、アトリビュート法への切り替え後にパソコン周辺機器の価格の下落幅が縮小しているが、これは切り替えが原因ではない点に注意が必要である。実際、2004年4～9月におけるアトリビュート法による前年同期比の下落幅は、ヘドニック法と比べて大きくなっていったとの結果が得られている。

一方で、2003年から英国では、パソコンなどの品質調整法がコスト評価法からヘドニック法に変更されており、アトリビュート法への変更という動きが世界的に起きているわけではない。

各国のパソコン関連指数の前年同月比の動きを見ると、国によって動きや変化幅はばらばらであるが、全体的な傾向として、近年下落傾向に歯止めがかかっている。

また、諸外国の中には、パソコン、カメラ関連のみならずヘドニック法を一層広範に用いている国もある。ドイツではテレビ、洗濯機及び中古車に、米国では多数の情報通信関連製品の他、家賃、被服、大学教科書等にヘドニック法を用いている。

6.3 家賃に関する回帰式による推定手法の意味

一般的なヘドニック法を概観したところで、再び家賃に関する検討に話を戻すことにする。

本節で試算する家賃の回帰式による推定手法は、特定の時点に着目して全国の個々の借家ごとの家賃を被説明変数とする回帰式で説明変数ごとの影響度を定め、地域別家賃指数を家賃の実測値でなく回帰式の推定値により計算するものであり、すなわち、パソコン用プリンタと同じ方法である。回帰式による推定において、家賃は品質と品質評価から成るものとする。品質は説明変数であり、品質の評価は説明変数の係数である。つまり、回帰式による推定は、地域ごとに、個々の借家ごとに異質な住宅の家賃を住宅の属性ごとの評価を一定にし、同質にした上で比較する品質調整である。

ここでの品質調整は地域間（クロスセクション）の品質調整である。CPIの変動は、品質を一定に保ったときの物価変動、すなわち、時系列間の品質調整を行った物価変

動を表すが、個々の借家について回帰式による推定結果を時系列で見て品質に相当する説明変数の係数が一定であれば、時系列間でも品質を調整したものとみなされる。

この考え方によると、調査対象が少ない地域において転出入により平均家賃が変わっても品質の変化を除いた品質の評価の変化だけの説明が可能である。また、これらの推定値は住宅区分ごとに全国における平均的水準を表すので、異常値を除外することにより家賃指数が変動するという問題を解決する手段にもなる。

以上のことから、家賃の実データについて上記の回帰式の推定値で置き換える、欠測値を補完する等により検討を進めることとする。

6.4 家賃の回帰分析

具体的には、家賃を被説明変数として、床面積、建築年及び人口密度を説明変数とする対数線形回帰式で表すことについて定量的評価を行うこととする。ここで用いたデータは、家賃、床面積、建築年については 2008 年における小売物価統計調査の家賃調査の個別データ⁸⁰、人口密度については 2005 年国勢調査の結果を全国都道府県市区町村面積調等による可住地面積で割った値⁸¹である。

6.4.1 回帰式の概要

本研究で推定した回帰式は以下のとおりである。

$$r_k = e^a \prod_{i=1}^3 x_{ik}^{b_i} \quad (6-1a)$$

ただし、各変数の意味は次のとおりである。

r_k …家賃

x_{1k} …床面積

x_{2k} …建築年

x_{3k} …人口密度

両辺の対数をとると、

$$\ln r_k = a + \sum_{i=1}^3 b_i \ln x_{ik} \quad (6-1b)$$

⁸⁰ 小売物価統計調査の家賃調査では 2008 年 10～12 月に奥州市、石巻市、平戸市及び八戸市について調査地区を拡大しており、Makoto Shimizu[89]では拡大分のデータも合わせて利用した。しかし、CPI は拡大前のデータのみを利用して作成されているので、拡大分を除いて分析を行った。なお、拡大分については建築年が調査されていないため、回帰式に対する影響はない。

⁸¹ 原則として市町村ごとに計算しているが、東京 23 区については個別に計算、甲府市及び富士河口湖町については合わせて計算している。

となる。(6-1b)を最小二乗法によりデータに当てはめ、住宅の区分(4区分)別・月別にその係数を、表 34-1、表 34-2、表 34-3、表 34-4 に示した。

表 34-1 によれば、木造小住宅⁸²においては、自由度調整済決定係数は 0.65~0.67 と 4 区分の中では最も高く、各変数の t 値は 2 桁、係数には月別に見て大きな違いはない。

表 34-1 木造小住宅の回帰式 (2008 年)

月	定数		床面積		建築年		人口密度		n	Adjusted R ²
	係数	t 値	係数	t 値	係数	t 値	係数	t 値		
1月	-262.54	-42.41	0.35	15.81	35.33	43.23	0.27	41.00	1804	0.66
2月	-266.57	-42.60	0.35	15.62	35.86	43.41	0.27	40.37	1773	0.66
3月	-265.21	-42.06	0.35	15.54	35.67	42.85	0.27	39.75	1712	0.67
4月	-264.12	-41.84	0.35	15.65	35.53	42.63	0.27	39.76	1697	0.67
5月	-257.52	-39.98	0.34	14.59	34.67	40.76	0.28	39.41	1658	0.66
6月	-255.89	-39.47	0.34	14.41	34.45	40.23	0.28	39.05	1641	0.65
7月	-250.23	-38.23	0.34	14.19	33.70	39.00	0.28	39.10	1622	0.65
8月	-248.18	-38.33	0.32	13.41	33.46	39.14	0.28	39.23	1619	0.65
9月	-246.93	-38.11	0.31	13.14	33.29	38.91	0.28	38.96	1598	0.65
10月	-247.58	-38.05	0.31	13.09	33.38	38.85	0.28	38.84	1586	0.65
11月	-246.94	-37.78	0.32	13.27	33.29	38.56	0.28	38.59	1573	0.65
12月	-245.79	-37.43	0.33	13.63	33.13	38.21	0.27	38.39	1544	0.65

表 34-2 によれば、木造中住宅においては、自由度調整済決定係数は 0.61 又は 0.62 であり、各変数の t 値は 2 桁、係数には月別に見て大きな違いはない。

表 34-2 木造中住宅の回帰式 (2008 年)

月	定数		床面積		建築年		人口密度		n	Adjusted R ²
	係数	t 値	係数	t 値	係数	t 値	係数	t 値		
1月	-211.39	-61.33	0.42	35.40	28.60	62.94	0.20	60.19	5524	0.61
2月	-210.74	-61.64	0.41	35.50	28.52	63.28	0.19	60.44	5434	0.62
3月	-209.69	-60.50	0.41	34.80	28.39	62.14	0.19	60.13	5307	0.61
4月	-208.75	-60.34	0.41	35.15	28.26	61.96	0.19	60.12	5257	0.62
5月	-203.81	-58.82	0.43	36.22	27.60	60.43	0.19	59.46	5207	0.61
6月	-203.99	-58.80	0.43	35.98	27.62	60.40	0.19	59.24	5169	0.61
7月	-203.38	-58.51	0.43	36.44	27.53	60.10	0.19	59.50	5142	0.61
8月	-201.55	-57.68	0.43	35.98	27.30	59.26	0.19	59.03	5084	0.61
9月	-202.61	-57.67	0.43	35.60	27.44	59.25	0.19	58.88	5055	0.61
10月	-202.37	-57.54	0.43	35.82	27.40	59.12	0.19	58.82	5024	0.61
11月	-201.85	-57.11	0.43	35.70	27.34	58.67	0.19	58.32	4980	0.61
12月	-203.05	-57.26	0.42	35.30	27.50	58.83	0.19	58.31	4944	0.61

表 34-3 によれば、非木造小住宅においては、自由度調整済決定係数は 0.58 と 4 区分の中では最も低い、各変数の t 値は 2 桁、係数には月別に見て大きな違いはない。

⁸² 小住宅は床面積 30 m²未満、中住宅は 30 m²以上である。

表 34-3 非木造小住宅の回帰式（2008年）

月	定数		床面積		建築年		人口密度		n	Adjusted R ²
	係数	t値	係数	t値	係数	t値	係数	t値		
1月	-115.54	-22.60	0.65	44.63	15.67	23.22	0.27	68.60	5160	0.58
2月	-106.79	-20.79	0.65	44.70	14.52	21.42	0.27	68.74	5115	0.58
3月	-108.48	-20.90	0.65	43.90	14.75	21.52	0.27	67.63	4995	0.58
4月	-108.62	-20.79	0.65	43.78	14.76	21.41	0.27	67.18	4921	0.58
5月	-108.44	-20.60	0.65	43.20	14.74	21.21	0.27	66.54	4799	0.58
6月	-110.46	-21.02	0.64	42.20	15.02	21.64	0.27	66.33	4736	0.58
7月	-110.98	-20.78	0.66	43.49	15.06	21.36	0.27	65.06	4738	0.58
8月	-120.86	-22.41	0.65	42.33	16.38	22.99	0.27	64.81	4712	0.58
9月	-119.01	-21.85	0.66	42.96	16.13	22.42	0.27	64.36	4675	0.58
10月	-117.71	-21.47	0.66	42.91	15.95	22.04	0.27	64.00	4636	0.58
11月	-114.43	-20.55	0.68	43.60	15.50	21.08	0.27	63.78	4600	0.58
12月	-117.96	-21.06	0.67	43.38	15.96	21.58	0.27	63.54	4554	0.58

表 34-4 によれば、非木造中住宅においては、自由度調整済決定係数は 0.61 又は 0.62 であり、各変数のt値は2又は3桁、係数には月別に見て大きな違いはない。

表 34-4 非木造中住宅の回帰式（2008年）

月	定数		床面積		建築年		人口密度		n	Adjusted R ²
	係数	t値	係数	t値	係数	t値	係数	t値		
1月	-150.11	-42.71	0.66	82.39	20.20	43.63	0.27	100.85	11228	0.61
2月	-148.15	-41.68	0.66	81.41	19.94	42.59	0.27	99.40	10987	0.61
3月	-140.92	-39.50	0.67	82.74	18.97	40.36	0.27	99.32	10818	0.61
4月	-140.66	-39.07	0.67	82.29	18.93	39.93	0.28	98.78	10722	0.61
5月	-141.66	-39.46	0.67	82.29	19.06	40.32	0.28	98.70	10650	0.61
6月	-144.43	-40.27	0.67	82.16	19.42	41.12	0.28	97.94	10608	0.61
7月	-144.31	-40.25	0.68	82.33	19.41	41.10	0.28	97.58	10567	0.62
8月	-145.73	-40.39	0.68	82.28	19.59	41.23	0.27	96.87	10472	0.62
9月	-147.04	-40.66	0.68	81.68	19.77	41.51	0.27	96.61	10430	0.62
10月	-147.16	-40.45	0.67	80.47	19.79	41.31	0.27	96.20	10360	0.61
11月	-147.28	-40.38	0.67	79.90	19.81	41.24	0.27	95.63	10298	0.61
12月	-147.83	-40.22	0.67	79.05	19.88	41.07	0.27	94.79	10214	0.61

これらの実証結果によると、床面積の係数は非木造住宅のほうが木造住宅より大きく、建築年の係数は木造住宅のほうが非木造住宅より大きい。相対的に見て、木造住宅は住宅の新しさ、非木造住宅は住宅の広さの影響を受けやすいことがわかる。他方、人口密度の係数は区分間で大きな違いはない。

推定された回帰式は、いずれも説明変数の数が少ないにもかかわらず（自由度調整済）決定係数が 0.58～0.67 と、CPI の帰属家賃のウェイトを換算する回帰式の決定係数 (0.53～0.59) と比べて低くないことから、家賃に影響を及ぼす主な要因は床面積、建築年、人口密度、すなわち、住宅の広さ、新しさ、地域の過密度であるとみなすことができる。説明変数の係数も安定している。したがって、これらの回帰式には、地域別家賃の推定への適用可能性がある。

6.4.2 タイムダミー法の適用可能性

CPIにおいては、パソコン及びカメラについて品質を調整した指数の作成方法としてタイムダミー変数を用いた回帰式が適用されている。そこで、今回も同様にタイムダミーを加えた回帰式についても検討を行った。しかし、個々の調査対象で見た場合、家賃については、毎月のように変動するIT関連製品と異なり、何か月かの間一定であることが多いためにタイムダミー変数のt値が有意でなかった。

6.5 地域別家賃指数の推定手法の検討

ここでは、これまでに検討した回帰式を使って、推定値を用いた欠測値の補完、実データの置き換え等により、保合による欠測値の補完法と比較しながら、小標本に依存する家賃指数の過大な変化という問題を解決するための方法について検討する。

6.5.1 転出入の割合

家賃の品質調整を検討する前に、転居に伴う小売物価統計調査の欠測値の発生割合を概観することとする。小売物価統計調査における家賃指数に問題が生じるような転出入は、その前後に空室又は建築物の喪失等が発生するために家賃の支払いがない又は家賃を把握できない場合である。このような転出入の割合を2008年について住宅区分ごとに推定すると表35のとおりである。転出入の割合が最も小さいのは木造中住宅であり、転入率は13.4%、転出率は17.0%である。逆に最も大きいのは非木造小住宅であり、転入率は21.9%、転出率は25.2%である。

表35 2008年における推定転出入率(%)

	木造小	木造中	非木造小	非木造中
転入	15.6	13.4	21.9	19.2
転出	23.9	17.0	25.2	21.9

注

小売物価統計調査において住宅の区分ごとに市町村別に毎月の標本規模当たりの転出入借家の割合を求め、市町村ウエイトで合算し、12か月分を乗じて年間値に換算

6.5.2 地域別家賃指数の推定のための回帰式の利用

6.4で導いた回帰式は、各変数が有意で、決定係数も高いので、地域別家賃指数を推定するために利用することができる。

回帰式による推定は、家賃調査地区内における住宅の新築、空室への入居、家賃調査地区外への転居、家賃調査地区の交替等により転出入が発生するようなときに、転入前及び転出後の家賃を推定し、家賃の品質を調整して本質的な価格変化を取り出すために有用であると考えられる。また、全国水準の構造に基づいて調整された家賃データ

は、通常、異常値を含んでいる実際の家賃データよりも転出入に伴う欠測値による影響が小さくなる。

6.5.3 比較の条件及び手法

地域別家賃指数の推定手法として CPI への適用を想定するときには次の2つの条件が必要である。1つ目は、毎月小売物価統計調査を実施した後に短期間で公表しなければならない CPI の作成前に必要なデータをすべて入手できることである。2つ目は、毎月データ自体が変わっても同じ手法を用いることができることである。

これらの2つの条件を満たすものとしてここでは次の3つの手法について検討した。

①保合

保合で欠測値を補完する。すなわち、新たに家賃を把握することができるようになった借家の前月の家賃を当月の家賃で保合する。同様に、家賃を把握することができなくなった借家の家賃を前月の家賃で保合する。

②回帰式補完

回帰式による推定値で欠測値を補完する。すなわち、新たに家賃を把握することができるようになった借家の前月の家賃を回帰式による推定値で補完する。同様に、家賃を把握することができなくなった借家の家賃を回帰式による推定値で補完する。

③回帰式推定

すべての家賃に回帰式による推定値を適用する。すなわち、すべての家賃を回帰式による推定値で置き換える。

CPI は、都道府県庁所在市及び政令指定都市以外の市町村については、個別に公表されず、10の地域に合わせた区分で公表されているが、ここでは、市町村別値を加算したものを総合的な評価とすることによりこれらの手法を比較することにする。市町村別値は、補完値を含め、総家賃を総床面積で割ることにより計算する。

本研究の試算では指数そのものは計算しなかったが、欠測値補完法、すなわち、①及び②については、補完値はその後の家賃に継続的に適用するのではなく、転出入月の前月比の計算のみに適用し、指数の計算は前月比を連鎖することにより行うべきであると考えられる。というのは、ひとたび推定された同一データを継続的に適用すると、指数が過度に安定してしまうからである。CPI マニュアルにおいても、最終的な観測値の保合はできるだけ避けるべきであり、ごく限られた期間に限定すべきであるとされている。また、小売物価統計調査の家賃調査地区の設定替えのような標本の大規模な入れ替えを行う際に同一データの継続的適用を一斉に遮断すると、新旧指数の間に段差が現れるおそれがある。

6.5.4 各手法の家賃推定値

最初に、各推定値が欠測値を補完していない CPI の原数値としての床面積当たり家賃と比較してどの程度離れているかを推定バイアスとして符号と大きさを評価する。推定バイアスの符号が正負両方に均等に現れるのであれば推定値に偏りが存在しないことになり、大きさが小さければ原数値に近い推定をしていることになる。

表 36 は、各手法による市町村別月別床面積当たり家賃の推定値と原数値との差を加重平均したものであるが⁸³、保合と回帰式補完については転出入いずれについても符号は正負両方の値をとり、偏りは見られない。また、保合のほうが回帰式補完よりも大きくなりやすい。例えば、いずれも7月に非木造小住宅の区分で最大となるが、保合では1㎡当たり8円であるのに対し、回帰式補完では6円である。このように通常回帰式補完のほうが原数値に近い推定値となる。

他方、回帰式推定による推定値は一貫して原数値から計算した値よりも低くなりがちである。これは、異常に高い値が低く推定されるためである。

表 36 保合、回帰式補完、回帰式推定による月別床面積当たり家賃に関する原数値との差の加重平均（2008年、1㎡当たり円）

月	木造小					木造中					非木造小					非木造中								
	転入		転出			Hed.	転入		転出			Hed.	転入		転出			Hed.	転入		転出			Hed.
	Car.	Reg.	Car.	Reg.	Car.		Reg.	Car.	Reg.	Car.	Reg.		Car.	Reg.	Car.	Reg.	Car.		Reg.	Car.	Reg.			
1月	2	1	-1	-2	-74	-1	-2	-2	-1	-49	-2	1	0	2	-25	-1	0	0	0	-74				
2月	0	0	-1	-5	-82	1	0	-1	-2	-49	-1	2	-2	-1	-24	1	1	0	-4	-77				
3月	0	1	2	-2	-85	-1	0	1	1	-49	1	0	0	0	-28	0	-1	1	-2	-82				
4月	1	4	2	-1	-82	1	1	0	-1	-47	1	1	0	0	-28	0	-1	0	-1	-82				
5月	1	0	2	-3	-95	1	0	0	-1	-51	-1	1	1	2	-27	0	0	0	-1	-81				
6月	1	2	5	3	-95	1	0	0	-1	-52	-1	2	0	1	-27	0	0	0	0	-80				
7月	0	0	-5	-3	-97	1	0	0	0	-51	0	2	8	6	-24	0	0	1	-1	-81				
8月	2	0	3	2	-93	0	-1	2	0	-50	-2	1	0	-2	-28	0	0	1	0	-80				
9月	0	1	0	-3	-95	0	0	0	0	-49	1	1	0	0	-26	-2	-2	0	0	-79				
10月	4	2	0	1	-97	0	0	1	0	-49	1	1	1	0	-28	0	0	-1	-2	-79				
11月	3	0	2	-1	-96	-1	0	-1	-1	-50	0	1	3	-2	-28	0	0	0	0	-79				
12月	-	-	1	0	-91	-	-	0	1	-48	-	-	1	-1	-31	-	-	0	-1	-79				

注

転入は新規に標本となった借家について前月の欠測値を補完する場合

転出は新規に標本から除外された借家について除外された月の欠測値を補完する場合

Car. は保合、Reg. は回帰式補完、Hed. は回帰式推定による推定値で原数値を置換

差は当該月の総床面積当たり総家賃を市町村ウエイトで加重平均した値の原数値による値との差

6.5.5 各手法の推定家賃の前月比

⁸³ 1～3月について、CPIは一部旧調査地区のデータを用いて作成されているので新調査地区のデータのみを用いている本節の結果はCPIの原数値と異なる。なお、旧調査地区においては建築年が調査されていないため、今回の回帰式が適用できない。

一般に、CPIにおいて変化は実際の金額よりも重要である。ここでは床面積当たり家賃の前月比に注目して上記①～③の評価を行う。

毎月公表されるという実務的な観点に配慮すると、保合と回帰式補完において、借家の転出入に伴う欠測値の補完は、家賃そのものではなく前月比の計算に対してのみ用いるべきである。というのは、転入者の情報を事前に把握することはできないことから、転入者の前月の家賃の反映は当月に行わざるを得ないが、前月の家賃は前月に公表せざるを得ないので事後修正になってしまうのに対し、当月の前月比であれば公表前に反映できるからである。

最初の評価のポイントは、市町村別床面積当たり家賃の過度な変動の抑制効果である。しかし、指数の異常値を調整すると個々の市町村別値は安定するものの全国平均に推定バイアスが生じる。すなわち、全国の推定値が原数値から乖離するおそれがある。そこで、第2の評価のポイントは市町村別値の調整に伴う推定バイアスが全国レベルで小さくなることとする。前者は推定値の前月比について市町村ウエイトで加重した標準偏差で評価することとする。後者は推定値と原数値の間での前月比の差について市町村ウエイトで加重平均することにより評価することとする。

前者の評価の結果として、表 37 は、2008 年において、原数値、保合、回帰式補完及び回帰式推定について住宅の区分ごとに月別総床面積当たり総家賃の前月比の標準偏差を月間で平均したものである。

保合は転出入借家の家賃を借家が存在しているときと同じとみなすことにより安定性を増強することになる。実際、表 38 によると、保合の結果は原数値による結果よりも小さくなる。他方、回帰式補完の結果は原数値による結果よりも必ずしも小さくなるとは限らない。

回帰式推定の結果は全国水準で基本的な変数に基づいて家賃を全面的に調整したものであるため、原数値による結果よりもはるかに小さくなっている。以上のことから、回帰式推定により最も安定的な家賃指数が得られることになる。

表 37 2008 年における原数値、保合、回帰式補完及び回帰式推定に関する住宅区分月別床面積当たり家賃の前月比の標準偏差の月間平均（%ポイント）

	欠測値補完法	木造小	木造中	非木造小	非木造中
原数値	—	1.7	1.1	1.7	0.7
転入	保合	1.6	0.9	1.4	0.5
	回帰式	2.0	1.1	1.7	0.7
転出	保合	0.8	0.7	0.7	0.5
	回帰式	1.7	1.0	1.5	0.8
回帰式推定	—	0.3	0.2	0.2	0.1

注

転入は新規に標本となった借家について前月の欠測値を補完する場合

転出は新規に標本から除外された借家について除外された月の欠測値を補完する場合

標準偏差は、月別市町村間総床面積当たり総家賃の変化の標準偏差であり、2008年1月から12月までの間で平均

保合と回帰式補完の前月比は、転入については当該月の原数値を前月の推定値で割った値であり、転出については当該月の推定値を前月の原数値で割った値

回帰式推定の前月比は当該月の推定値を前月の推定値で割った値

後者の評価の結果として、表 38 は、市町村別総床面積当たり総家賃の前月比を加重平均した値の原数値から見た推定値のバイアスを示したものである。それによると、推定バイアスは 0.9%ポイントを下回る程度のものであった。また、推定バイアスは各手法とも正負両方の符号をとるため、12 か月で平均すると絶対値にして 0.1%ポイント以下、年率に換算すると 1.2%ポイント以下であった。特に、帰属家賃としてのウエイトが大きい木造中住宅については、毎月絶対値で見て 0.3%ポイント以下、年率にして 0.3%ポイント以下と小さかった。

さらに、回帰式補完及び保合の推定バイアスは回帰式推定よりも小さくなる傾向がある。例えば、回帰式補完、保合については、いずれも7月に非木造小住宅の転出で推定バイアスは最大で、それぞれ 0.36%ポイント、0.43%ポイントとなり、回帰式推定については5月に木造小住宅で-0.84%ポイントと最大幅となる。しかし、推定バイアスの大きさの順位は、状況によって必ずしもこの傾向に従わない。

表 38 2008 年における保合、回帰式補完及び回帰式推定に関する床面積当たり家賃の前月比の推定バイアスの加重平均値（%ポイント）

月	木造小					木造中					非木造小					非木造中				
	転入		転出			転入		転出			転入		転出			転入		転出		
	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.
1月	0.03	-0.01	-0.07	-0.12	-0.27	-0.03	-0.01	-0.12	-0.03	0.05	-0.13	0.04	-0.01	0.09	0.04	0.08	0.05	0.01	0.00	0.06
2月	-0.19	-0.10	-0.02	-0.14	-0.46	0.04	0.15	-0.14	-0.19	-0.07	0.08	-0.04	-0.08	-0.06	0.03	0.06	-0.02	-0.04	-0.23	-0.29
3月	0.03	-0.01	0.09	-0.10	-0.13	-0.08	0.00	0.07	0.06	-0.05	0.01	-0.08	0.03	0.03	-0.21	-0.06	-0.07	0.05	-0.17	-0.43
4月	0.01	-0.04	0.11	-0.01	0.18	0.10	-0.01	0.00	-0.03	0.19	-0.06	-0.01	0.02	0.02	-0.05	-0.02	0.07	0.05	-0.04	0.00
5月	-0.09	-0.23	0.14	-0.13	-0.84	-0.10	-0.11	-0.03	-0.04	-0.21	-0.04	-0.08	0.05	0.18	-0.01	0.00	0.03	0.03	-0.06	0.03
6月	-0.03	0.02	0.19	0.09	-0.02	-0.07	0.02	0.06	-0.02	-0.06	0.08	-0.06	0.02	0.04	-0.03	0.03	0.04	-0.01	-0.02	0.02
7月	-0.03	-0.10	-0.18	-0.13	-0.08	-0.06	-0.03	-0.03	0.03	0.01	0.03	-0.10	0.43	0.36	0.31	0.02	0.01	0.04	-0.03	-0.01
8月	-0.01	0.01	0.10	0.11	0.19	-0.03	-0.01	0.09	0.00	0.08	0.00	-0.11	0.02	-0.03	-0.20	0.01	0.01	0.03	0.00	0.05
9月	-0.07	0.01	-0.01	-0.17	-0.08	0.01	0.04	0.04	0.02	0.10	0.08	-0.04	0.01	0.03	0.09	0.00	0.02	0.04	0.02	0.05
10月	-0.03	-0.07	-0.06	0.02	-0.15	-0.03	-0.04	0.06	0.07	0.02	-0.02	-0.04	0.03	0.02	-0.07	0.09	0.07	-0.04	-0.09	0.03
11月	-0.09	0.08	0.16	0.00	0.13	-0.02	-0.01	-0.04	-0.04	-0.10	-0.01	-0.06	0.16	-0.07	0.01	0.04	0.02	-0.01	0.00	0.02
12月	-0.10	0.00	0.02	-0.05	0.38	0.00	-0.01	0.05	0.07	0.11	-0.03	-0.04	0.04	0.00	-0.12	0.01	0.00	-0.02	0.04	0.01
平均	-0.05	-0.04	0.04	-0.05	-0.10	-0.02	0.00	0.00	-0.01	0.01	0.00	-0.05	0.06	0.05	-0.02	0.02	0.02	0.01	-0.05	-0.04
年率	-0.58	-0.42	0.48	-0.64	-1.14	-0.26	0.00	0.01	-0.10	0.06	-0.01	-0.61	0.72	0.58	-0.20	0.25	0.22	0.11	-0.65	-0.45

注

転入は新規に標本となった借家について前月の欠測値を補完する場合

転出は新規に標本から除外された借家について除外された月の欠測値を補完する場合

Car. は保合、Reg. は回帰式補完、Hed. は回帰式推定による推定値で原数値を置換

保合と回帰式補完の前月比は、転入については当該月の原数値を前月の推定値で割った値であり、転出については当該月の推定値を前月の原数値で割った値

回帰式推定の前月比は当該月の推定値を前月の推定値で割った値

表の数値は、当該月において総床面積当たり総家賃の前月比を市町村ウエイトで加重平均した値の原数値による値との差

最後に、これらの推定バイアスが CPI に及ぼす影響を試算する。

表 39 は、CPI における住宅区分ごとの民営家賃としてのウエイト及び帰属家賃としてのウエイトで合算することによりこれらの CPI の変化率に対する推定バイアスの影響力を推定したものである⁸⁴。表 39 において各月の数値の正負の符号には特定の傾向が見られないので、住宅区分間で合算すると全体としては小さくなる場合が多い。年率に換算すると、転入の保合が最も大きな幅になるが、-0.03%ポイントに過ぎない。欠測値補完法については転入及び転出を合わせて行うことが多いので、それらを合算すると、保合と回帰式補完は-0.02%ポイントと、0.00%ポイントの回帰式推定よりも推定バイアスが大きくなった。しかし、回帰式推定の推定バイアスは、各住宅区分では大きいものの、それらの正負が相殺して全体として小さくなったものである。欠測値補完法の転入と転出の双方に配慮し適用する場合の指数への反映手法に

⁸⁴ 2005 年基準 CPI に占める家賃のウエイトは民営家賃が 2.8%、帰属家賃が 14.2%、合わせて 17.0%を占めている。特に、木造中住宅は 0.8%の民営家賃を 10.9%の帰属家賃に適用しているため影響力が大きい。CPI 変化率に対する影響力は、厳密にはウエイトのみならず住宅区分ごとの家賃指数の大きさにも依存するが、住宅区分別家賃指数は 100 に近いことから、今回の試算ではウエイトのみによる合算で近似した。

については、今後検討すべき点が残されているので、両者の推定バイアスを単純に加算した結果はあくまで参考値に過ぎない。

表 39 2008 年における保合、回帰式補完及び回帰式推定に関する床面積当たり家賃の前月比の推定バイアスの CPI 変化率への影響力 (%ポイント)

月	転入		転出		転入+転出		Hed.
	Car.	Reg.	Car.	Reg.	Car.	Reg.	
1月	0.00	0.00	-0.01	0.00	-0.02	0.00	0.01
2月	0.00	0.02	-0.02	-0.02	-0.01	-0.01	-0.01
3月	-0.01	0.00	0.01	0.01	0.00	0.01	-0.01
4月	0.01	0.00	0.00	0.00	0.01	0.00	0.02
5月	-0.01	-0.01	0.00	0.00	-0.01	-0.02	-0.03
6月	-0.01	0.00	0.01	0.00	0.00	0.00	-0.01
7月	-0.01	0.00	0.00	0.01	-0.01	0.00	0.00
8月	0.00	0.00	0.01	0.00	0.01	0.00	0.01
9月	0.00	0.00	0.01	0.00	0.01	0.01	0.01
10月	0.00	0.00	0.01	0.01	0.00	0.00	0.00
11月	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01
12月	0.00	0.00	0.01	0.01	0.01	0.01	0.01
平均	0.00	0.00	0.00	0.00	0.00	0.00	0.00
年率	-0.03	0.00	0.01	-0.01	-0.02	-0.02	0.00

注

転入は新規に標本となった借家について前月の欠測値を補完する場合

転出は新規に標本から除外された借家について除外された月の欠測値を補完する場合

Car. は保合、Reg. は回帰式補完、Hed. は回帰式推定による推定値で原数値を置換

保合と回帰式補完の前月比は、転入については当該月の原数値を前月の推定値で割った値であり、転出については当該月の推定値を前月の原数値で割った値

回帰式推定の前月比は当該月の推定値を前月の推定値で割った値

表の数値は、当該月において総床面積当たり総家賃の前月比を市町村ウエイトで加重平均した値の原数値による値との差を CPI における住宅区分ごとの民営家賃としてのウエイト及び帰属家賃としてのウエイトで合算

表 39 の結果から、本研究で検討しているような家賃の推定により全国 CPI 総合の動きに深刻な影響が出ることはない結論する。つまり、これらの手法を使うと、家賃指数について住宅区分別・市町村別の変動を抑制しながら全国値への影響を小さくすることが可能である。

また、これらの手法のうち、回帰式推定の推定値は、市町村別指数の過剰な変動を抑制するという点では優れているが、原数値に対する推定バイアスは最も大きくなる。しかし、各月の原数値との相違は、個々の調査対象間の相違と比べて小さく、かつ長期的傾向は見られない。

全体として、家賃指数の安定性の観点を踏まえると、最も良い手法は回帰式推定、次いで保合である。しかし、推定バイアスの観点を踏まえると、最良な手法を見分け

ることが難しい。また、区分別に見ると、単一の手法で安定性と原数値との近似性の両方を満たすのは困難である。

なお、CPI は比較時価格を基準時価格で割ってから市町村ウエイトで合算して作成されているので、厳密に CPI に対する影響を試算するのであれば基準時価格で割ってから上記の比較をするべきである。その際には、基準時価格についても同様の推定をした上で試算をすべきであるが、建築年については 2008 年以降しか把握できないため、本研究では簡略化した対応ではあるが、床面積当たり家賃の前月比の比較で代替した。

6.6 全体評価と今後の課題

床面積、建築年、人口密度による回帰式は個々の家賃の推定手法としての適用可能性がある。当回帰式は、欠測値の補完法として利用することができ、また、すべての借家の家賃の推定に適用することもできる。

本研究で検討した欠測値の補完方法と回帰式推定法については、どれを採用しても、地域によって、月によっては家賃指数がしばしば公表値と異なるものの、全国値、年平均では原数値による結果と比べて大きな違いはないという試算結果が得られた。

総務省統計局では、2010 年基準改定から本研究で指摘している問題に対し、実務にも配慮し、保合を導入することにした。

しかし、今後も推定の改善に向けて、より詳細かつ長期のデータで課題を具体化することが必要である。特に、建築年については 2008 年以降に把握できる分しか存在しないので、今後のデータの蓄積を待ってから改めて試算をすることが必要である。また、指数の推定精度の向上のために回帰式の説明変数として他の変数についても検討すべきである。さらに、欠測値補完法における転入と転出の組み合わせ手法を検討するとともに、回帰式推定法と欠測値補完法の組み合わせ等、各種組み合わせについても試算をすることが望ましい。

7 IMF/ROSC 対応

これまでは特定のトピックごとに CPI を作成する立場から作成技法の高度化への分析・取組を説明してきた。本章の最後に、まとめとして、海外ではこのような分析・取組を経て発展してきた日本の CPI をどう評価しているのかについて、先般 IMF が日本に対して行った国際基準・規範状況に関する評価報告書（ROSC）を簡単に紹介する。本節は清水・永井[91]に基づく。

7.1 ROSC とは

1990年代のメキシコ金融危機やアジア通貨危機を背景に、IMFにおいて、国際金融危機を予防する観点から、世界経済の効果的な調査・監視及び潜在的な問題の把握が議論され、今後の世界的な金融危機を予防するためのプログラム（ROSC : Report on the Observance of Standards and Codes）の作成が提唱された。

ROSCには①データ提供、②金融・財政政策の透明性、③会計の透明性、④銀行行政（監視・管理）、⑤安全保障、⑥保険、⑦決済体系（賃金体系）、⑧マネーロンダリング（不正金融）の対処、⑨コーポレートガバナンス（企業統治機構）、⑩勘定体系、⑪会計監査、⑫債務超過と債権の12の分野があるが、このうち①のデータ提供に関するものがデータモジュール ROSC（データ ROSC）と呼ばれるものである。

データ ROSC は、IMF 特別データ公表基準（SDDS）に照らし、各国のデータ公表慣行の概要を評価するとともに IMF データ品質評価フレームワーク（DQAF）に基づく6つの主要マクロデータセットに対する品質評価を行うものである。この6つの主要マクロデータセットは、国民経済計算、CPI、生産者物価指数、政府財政統計、金融統計及び国際収支統計である。

このデータ ROSC のための審査の受入れ及び審査結果の公表については各国の任意となっているが、2006年までにフランス、イタリア、カナダ、スウェーデン、韓国、ドイツについての結果が公表されている。

2006年3月に公表された日本のデータ ROSC については、2003年11月に IMF から日本に対してミッション受け入れが打診された。その後、統計作成部局による事前の自己診断書の提出、IMF の求めに応じた各統計のユーザーに対するアンケート調査の実施及びその回答提出を経て、IMF は2005年9月12日～28日に対日ミッションを実施した。このミッションでは IMF から専門家で構成されたミッションチームが来日し、3回の全体会合に加え、それぞれのデータセットについて、詳細な情報の聴取及び意見交換が行われた。また、この間各統計のユーザーとの会合も開催された。なお、ミッションチームの物価統計の担当は Mick Silver であった。

IMF は、一般に公開されている情報、対日ミッションでの統計作成部局からの提出情報、ユーザーに対するアンケート調査結果等から評価報告書の暫定版を作成した。その後、この暫定版について日本の統計作成部局と何度か協議し、報告書の確定版を作成した。さらに、評価報告書に対して日本の統計作成部局からの回答を取り寄せ、日本の同意の下、2006年3月にこの回答を併せた評価報告書を公表した。

7.2 ROSCにおける評価基準

IMFは、各国のデータ品質を国際比較・評価するための具体的かつ詳細な基準としてDQAF（Data Quality Assessment Framework）を策定し、これをデータROSCにおける評価手法として組み込んでいる。

DQAFはデータの収集、処理及び普及・公表の様々な品質側面を包括的に網羅しており、次元、要素、指標という3つの段階式構造から成っている。第1のレベル（次元）は、品質の要件に加え、品質の5つの側面、すなわち、規範性の保証、方法の健全性、正確性・信頼性、実用性、利用しやすさの6つから成っている。これらの6つの次元がそれぞれいくつかの要素、それらの要素がいくつかの指標から成るといった構造となっている。それぞれの次元の内容については以下のとおりである。

0. 品質の要件

統計の品質の要件あるいは制度的前提条件である。

1. 規範性の保証

統計データの収集、作成、普及・公表における客観性の原則に関するものである。この次元には、政策及び実施における専門性、透明性、倫理基準を確保するための制度的調整が含まれる。

2. 方法の健全性

統計作成のための方法論の基盤は信頼できる健全なものであるべきで、それは国際的に認められている基準、指針、良い慣行に従うことによって達成できるという考え方である。この次元は、データセットによって異なっている。

3. 正確性・信頼性

統計データが経済の実態を十分に表すべきという考え方である。この次元は、データ固有のものである。

4. 実用性

時期を逸することなく適切な周期をもって統計データを公表すること、当該統計データとその他の統計データに一貫性があること、当該統計は定期的な見直し方針に従うこと、に関連するものである。

5. 利用しやすさ

データ及びメタデータが明確で理解しやすい方法で表示され、公平かつ容易に入手できること、メタデータは最新かつ適切であること、迅速で充実した支援サービスが提供可能であること、に関連するものである。

なお、ROSCでは、各国の評価を行う中で、使用する評価基準（DQAF）等について修正を加えているので、評価時点の異なる国同士の評価を単純に比較することはできない。今回の日本の評価は2003年7月版のDQAFに基づくものである。

7.3 CPI に対する評価の概要

上記の6つの次元ごとの各要素と、日本の6つのデータセットについて、これらの評価基準に基づき要素ごとに5段階で行われた評価は、評価報告書に添付された表40に示した表に要約されている。

日本のCPIは概ねDQAFの基準を満たしており、「2. 方法の健全性 2.2 範囲」と「3. 正確性・信頼性 3.3 統計技術」の2項目についてわずかに逸脱が認められるに過ぎないとされた。

この他、ROSCには、技術的な提言内容と統計作成部局による回答を併せて公表しているが、それについては清水・永井[91]に記した。

7.4 今後の課題

今回はCPIを含めた6つの指標について評価が行われたが、それらの指標の中には国民経済計算や企業物価指数などCPIとも関連の深い統計が含まれている。

日本ではデフレの時代と呼ばれて久しいが、政策との関係で、物価については今後も様々な議論が展開していくことが想定される。その際、CPIとは何かを改めて問われることとなろう。これまで、国内・国際いずれにおいても消費者物価と生産者物価は別々に議論されてきたが、政策課題の解決に向けて、経済全体を見据え、両者の位置付けや関係についてより深い考察が必要であると考えられる。

表 40 日本：データ品質評価フレームワーク（DQAF）2003年7月版 - 結果概要

記号：O=遵守、LO=概ね遵守、LNO=概ね未遵守、NO=未遵守、NA=該当なし							
次元/要素	データ・セット	国民経済計算	消費者物価指数	企業物価指数	政府財政統計	金融統計	国際収支統計
0. 品質の要件							
0.1	法的・制度的環境	O	O	O	LO	O	LO
0.2	資源	LNO	O	O	O	O	O
0.3	関連性	O	O	O	O	O	O
0.4	その他のクオリティ管理	O	O	O	O	O	O
1. 規範性の保証							
1.1	専門性	O	O	O	O	O	O
1.2	透明性	O	O	O	O	O	O
1.3	倫理基準	O	O	O	O	O	O
2. 方法の健全性							
2.1	概念と定義	O	O	LO	O	LO	O
2.2	範囲	O	LO	O	LNO	LO	O
2.3	分類／セクター分け	O	O	O	O	LO	LO
2.4	記録の基礎	O	O	O	O	LO	O
3. 正確性・信頼性							
3.1	元データ	LO	O	LO	LO	O	LO
3.2	元データの評価	LO	O	O	O	O	O
3.3	統計技術	O	LO	LNO	O	O	O
3.4	中間データと統計アウトプットとの評価・確認	O	O	O	O	O	O
3.5	見直し研究	LNO	O	O	LO	O	O
4. 実用性							
4.1	周期性と適時性	O	O	O	LNO	O	O
4.2	一貫性	O	O	O	LO	O	O
4.3	改定方針と実施	O	O	O	O	O	LNO
5. アクセス可能性							
5.1	データへのアクセス	O	O	O	O	O	O
5.2	メタデータへのアクセス	O	O	O	O	O	O
5.3	利用者への支援	O	O	O	O	O	O

遵守（O）：現在の実施状況は、重大な欠陥もなく、国際的に承認された統計事例であるDQAFの基準に基本的に合致し達成している。

概ね遵守（LO）：いくつかの逸脱はあるものの、それらは当局のDQAFを遵守する能力に疑問を抱くほど重大なものではない。

概ね未遵守（LNO）：重大な逸脱があり、当局はDQAFを遵守するために対策を講じる必要がある。

未遵守（NO）：ほとんどのDQAFの基準は達成されていない。

適用なし（NA）：統計的な基準がその国の環境に適用されない場合に例外的に使用する。

第3章 物価の実感に関する定量的評価

日本では、物価上昇の著しい高度成長期に、消費者を中心に物価指数の上昇率が実感を下回っているのではないかという疑問が生じ、多方面で注目された。それに伴い、物価上昇の実感を確認するための意識調査が幾度となく実施され、物価指数との差の要因分析の蓄積が進展した。永山貞則[92]は、物価指数と生活実感との差は、①指数の構成を知らないために起こる単純な誤解、②平均のウェイトと個人の消費構成との差、③購入頻度の差、④生活費との混同から生じているとしている。辻村江太郎[93]は、需要曲線と供給曲線をもとに、所得増加に伴う需要曲線の変位幅が大きい、つまり、消費者が消費量の増加を多く期待する商品ほど結果としては価格が上昇しやすく、希望した数量と実現する数量とのギャップが大きくなりやすいとしている。三瀧信邦[94]は、生活実感と統計のズレという批判は、実在と統計のズレの批判と、個々の統計単位と統計集団との関係、例えば、平均値の持つ意味の取り違えとがミックスして出される批判であるとし、平均値が実在の社会的集団（家計集団）に対する物価上昇の実態を少しでも正しく反映するように改善することは大いに意味のあることであるとしている。北山直樹[95]は、値上がりの激しい品目から指数を作成することだけが実感と称するものに合わせる唯一の方法であると述べている⁸⁵。中村隆英[96]も、生活用品のうち、購入頻度の高いものであって、しかも、値上がりの激しい数品目ないし十数品目を捉えてその上昇率を平均すれば、ほぼ生活「実感」に近づくのではないだろうか述べている。また、中村隆英[97]⁸⁶は、指数と実感の乖離について、消費の数量の変化と物価の変化の両方から生ずる消費支出額の変化をすべて物価の責任と考えること、消費内容の質的向上を忘れてしまうという傾向を指摘している。大槻義雄[98]は、CPI 上昇率と「物価上昇率に関する世論調査」（総理府広報室）の結果を比べ、個人個人の実感による物価上昇率の平均値は、実感の物価の動きとかなり

⁸⁵ 正確には、次のとおりである。「指数採用 428 品目のうちから、前年同月に比べ最も上がり方の激しい 20 品目を取って物価指数を作れば、指数の上昇率は 80%を超える。また、50 品目に増やしても 60%を超えている。しかし、428 品目に増やせば、25%に落ち着くのである。このような数字は、実感と称される値によく合っている。そして、今までの経験では、これ以外に「実感」と称するものに合わせる方法がない。購入頻度、必需品（所得弾性値階級）等、いろいろな商品を分けてみても、10%の差も説明できない。」

⁸⁶ 中村隆英[97]では、CPI が「実感」に合わないという批判が生じる理由として、①家計の日常生活に直結した物価を対象にしていること、②年々の労働組合の春闘など、賃金上昇の決定に直接結びついているために、問題が政治的な側面から取り上げられ、新聞雑誌にも大きく報道される場合が多いことを挙げている。

パラレルに動くこと、個人個人の実感は、1年前と比べるよりは、より短期的な最近の物価トレンドに敏感なようであること、生活費の増加感と物価の上昇率の間には強い相関があり、生活費の増加と物価の上昇を混同して意識する傾向が強いこと、物価上昇の実感は、身近な費目の動向で形成されやすいこと等を見出している。

その後、バブル経済が崩壊し物価の変動が小さくなった 1990 年代中頃から物価に関するそのような疑義は沈静化していた。

しかし、21 世紀に入り、ユーロ通貨が流通した EU でそのような問題が取り上げられ、また、日本でも 2005 年頃から物価が再上昇したことから、物価上昇が低いのではないかという疑義が再び話題となる機会が増えた。

そもそも物価に関する実感は、消費者一人一人の独自の意識に依存するものであるのに対し、CPI は消費者全体の消費構造を踏まえて価格変化を総合するものなので、両者の変化率に違いがあるのは当然である。

本章では、そのような違いの理由をできるだけわかりやすく表現するために、21 世紀に入ってからの物価の意識をめぐる事情とともに、日本における実感に関する意識調査による結果、CPI の基礎データによる実感物価変化率の試算等を通じ、可能な限り定量化して紹介することとする。これにより、物価に関する指数と実感は、相容れないものでなく、ある程度相互に整合しているものであり、概念を調整すれば両者ともに類似の動きをするものであることを示す。

本章は、指数と実感の乖離となる要因を①品質調整による影響、②購入頻度による影響、③直近の物価変動による影響としているが、これらの要因自体はすでに冒頭に掲げた先行論文でも指摘されている。本章の特徴は、これらの要因を考え方にとどまらず緻密に定量化した点にあり、特に、①品質調整による影響は品目ごとに銘柄単位で調整を行っている。本章の内容は、UNECE/ILO 会合で発表した Makoto Shimizu[99]と、それに諸外国とその後の日本の状況等を加えた清水誠[100]を加筆・修正したものである。

1 EU の状況

EU では、1999 年 1 月に決済通貨としてのユーロが導入され、2002 年 1 月からユーロ通貨が流通し始めた。2000 年以降の HICP の動きを見ると、毎年 2%台前半の上昇となっており、通貨流通による物価への影響は全体として大きくはない。これを品目別に見ると、上昇率が大きな品目は食料品、外食、エネルギー、煙草等の日常品であり、このうち、エネルギーと煙草は通貨流通とはまったく異なる要因で価格が上昇している。

しかし、通貨流通後、消費者による物価変動の実感が指数を大きく上回るとする指摘が数多くなされた。EU では、このよう指摘を踏まえて実感と指数の相違について各方面で分析が行われたが、最も主流となる解釈は、意識は購入頻度が高い商品の影響を受けやすいが、これらの商品の価格上昇率が高いので、消費者の実感は指数を上回るとするものであった。しかし、EC などはそのような解釈ではすべてを説明できないとしている。

この問題は、2005年にOECDで開催された *Inflation Measures: Too High – Too Low – Internationally Comparable?* というセミナーにおける「公式 CPI の信頼性の維持・確立」というテーマの中で取り上げられ、詳しい事情が紹介された。

Aucremanne et al.[101]は、実感が指数を大きく上回る現象について EC が実施する意識調査 *The Joint Harmonised EU Programme of Business and Consumer Surveys*⁸⁷に基づく実感物価変化率と HICP 変化率との差によって示した。

両者の乖離は、EU 各国においてユーロ通貨流通直後の 2002 年 1 月に始まり、2003 年 1 月にピークとなり、その後は縮小したが、ドイツ以外の国ではまだ存在しているとのことであった。他方、ユーロ通貨を導入していないデンマーク、スウェーデン及び英国にはそのような現象は見られないとされ⁸⁸、そのこと自体が EU における両者の乖離が通貨流通の影響によるものであるとの 1 つの根拠とされた。

この EU の意識調査では、12 か月前からの物価変化に関する意識を、数値で聞くのではなく、「かなり上昇」、「適度に上昇」、「わずかに上昇」、「ほとんど同じ」、「下落」の 5 階級に区分して聞き、階級ごとに変化の程度に応じた加重得点を与え⁸⁹、回答者の割合を乗じて合算した“balance statistic”として集約した指標を作成している。実感物価変化率はこのような方法により求められた指標なので、HICP の変化率とは尺度が異なるが、2001 年以前は各国で両者が似た動きをしていたためにそれらを合致させると、2002 年以降に多くの国で実感物価変化率が HICP 変化率を大きく上回ってしまうのである。

Aucremanne et al.[102]は、国ごとに“balance statistic”と HICP 変化率の差が定常かどうかについて単位根検定を行っている。検定に際しては、“balance statistic”をそのまま用いるのではなく、期間内で標準偏差を計算し、HICP 又は CPI の変化率の標準

⁸⁷調査は EU 内約 125,000 の企業と約 40,000 の消費者を対象に毎月実施している。消費者に対しては物価変動の認識に関する調査項目が含まれている。

⁸⁸ デンマーク及びスウェーデンでは実感物価変化率が HICP 変化率とほとんど同じである。英国では両者の間に乖離が見られるが、一貫して前者が後者を上回るのでなく、時期によって上下が逆転する。

⁸⁹ 上から順に 1 点、0.5 点、0 点、-0.5 点、-1 点となっている。なお、5 階級の他に「わからない」という階級がある。

偏差との比により標準化している。その結果、ユーロ導入国⁹⁰では、2001年以前に限定すると定常となる国が幾つか存在するが、2002年以降を含めるとすべての国で非定常となることが確認されたとのことである。

また、イタリアの Giovane and Sabbatini[74]は、両者の乖離の要因を、①通貨流通後に購入頻度の高い製品が購入頻度の低い製品と比べて相対的に価格上昇したこと、②人々が価格下落よりも価格上昇に強い影響を受けたこと、③個人が経験する物価変動はその人の消費バスケットを構成する価格変化を反映すること、④メディアが価格上昇並びに実感及び指数の対比を強調したこと、⑤耐久財のように通貨流通前に購入した商品について記憶が曖昧なことであると要約していた。

他方、CPIの基礎データから実感指数を表現しようという試みもある。フランスの Dominique Guédès[103]によると、フランスでは、過去の価格に関する意識について生鮮食品を除く食料、エネルギー、一部のサービス等で説明するモデルを構築しているとのことであった。

セミナーでは、これらの発表に対して、実感と指数の乖離は、①政府への信頼の喪失につながる、②通貨政策にも連動する、③生活必需品には新品目が含まれるので指数は変動しやすい、④メディアの役割は非常に重要である等の意見が述べられた。また、実感には心理的な要因が含まれるが、科学的な分析、すなわち、各種データを用いたこの分野の実証分析が重要であるとの意見もあった。

セミナー開催後のEUにおける統計を見ても実感がHICPの変化率を上回っているという状況に変化はない。乖離が特に大きな国はベルギー、ギリシア、フランス及びフィンランドである。

ユーロ通貨導入による物価の問題は、広くEUの人々の間に浸透している。2006年9月にECが実施した意識調査Flash Eurobarometer⁹¹では、ユーロ通貨導入が価格上昇効果をもたらしたとする者が93%も占めている。また、ユーロ通貨導入を大体において損失と考えている人達を対象にその実例を聞いているが、価格上昇を挙げる者の割合が81%に達しており、他の実例を挙げる者がそれぞれ2割を下回る中で際立っている。

2 物価安定下における日本の状況

日本では、2007年後半からの原油価格等の高騰を受け、食料品、ガソリン、光熱費等の価格が上昇し、CPIも上昇した。しかし、日常生活でそれらの価格が大幅に上昇

⁹⁰ 2001年12月以前に意識調査が行われていないルクセンブルクは除かれている。

⁹¹ 調査対象はユーロ通貨を導入した12か国における12,000強の住民である。

した割に CPI の上昇の程度が低いという指摘もあった。このような背景のもとで、日本における 2007 年前後の CPI と実感を概観すると以下のようなになる。

2.1 CPI の動き

日本の CPI は、歴史的にも国際的にも安定しており、前年（同月）比はゼロ近傍で推移しているが、その内訳を見ると、すべての商品の物価が安定している訳ではない。

2005 年基準 CPI には 584 品目存在するが、2007 年の前年比を品目別に見ると、そのうち 312 品目がマイナス 1% 超プラス 1% 以下となっている。しかし、図 13 に示すように、プラス 10% 超が 7 品目、マイナス 10% 以下も 13 品目存在しており、品目間の前年比の標準偏差は 4.2 ポイントと決して小さくはない。

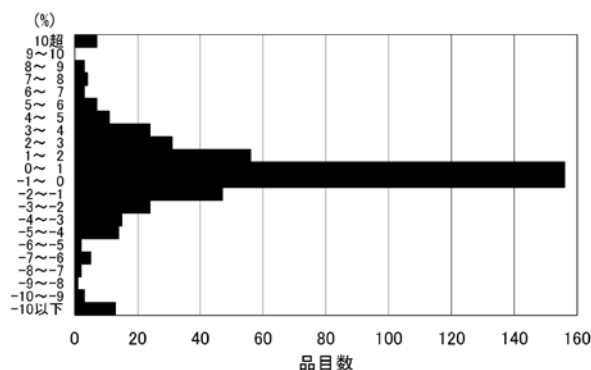


図 13 2007 年における CPI 品目別前年比

このような多様な動きを示す各種商品の中で、プラス変化の商品を重視するか、マイナス変化の商品を重視するかによって物価変動に関する意識が異なるものとなる。

なお、品目別前年比の分布にはおおむね上下対称であるという特徴がある。このような形は、2007 年の分布に限って見られるものでなく、1990 年代中旬以降における物価の変化が小さい時代に共通に当てはまるものである。これに対して、高度成長期には、品目別分布は物価上昇率が高い方に裾野が広がる非対称分布であった。

2.2 物価の実感と CPI

物価変動の実感については、日本銀行情報サービス局が四半期ごとに 4,000 人を対象に「生活意識に関するアンケート調査」を実施しており、その中には、EU のように物価の変化の程度を 5 段階に区分⁹²して聞いているだけでなく、「1 年前に比べ、

⁹² EU と異なり、上昇・下落面ともに 2 階級に区分されており、「かなり上がった」、「少し上がった」、「ほとんど変わらない」、「少し下がった」、「かなり下がった」の 5 階級からなる。

現在の『物価』は何%程度変わったと思いますか」という変化率そのものを聞く問がある。それについて 2007 年 12 月の結果を見ると、**図 14** に示すように、平均で 6.4%、中央値で 5.0%と、いずれも CPI の前年同月比 (0.7%) よりはるかに大きい⁹³。

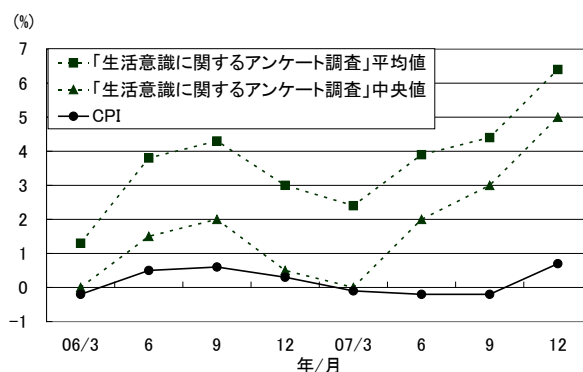


図 14 物価に関する意識と CPI の前年同月比

実感が一人一人異なるとはいえ、全体として物価指数とこれだけ差が生じるということは、最近の物価変動について実感が指数を上回る共通した要因があるものと推測される。

このような要因としては、大きく分けて、①物価指数で消費者が日常意識することがない品質の調整をしていることによる影響、②実感が購入頻度の高い品目に影響されやすいこと、③実感が直近の変化に影響されやすいことの3つが想定されるが、以下では、それらの区分ごとに定量的にどの程度の説明が可能なかを述べることにする。

2.3 品質調整による影響

1つ目は、物価指数では ILO 決議に沿って品質調整をすることが前提となっているが、日常の買い物の現場では品質に配慮することなく物価を実感することが多いことに伴うものである。CPI では、通常、①商品の出回りの変化を踏まえて銘柄を変更する場合、②価格収集している店舗が変化する場合、③同一銘柄で商品が変化する場合に品質調整を行っている。しかし、そのような計算上の処理は物価指数には必要不可欠であっても、商品の品質は関心を持って観察しなければ意識できないものことから、市場における主力商品の値上げの本質が品質向上だとすれば、価格変化に関

⁹³ CPI の価格の調査日は、原則として、12 日を含む週の水曜日、木曜日又は金曜日のいずれかであるのに対し、「生活意識に関するアンケート調査」では調査期間は調査月の前月中旬からの約 1 か月間と異なる点に注意が必要である。

する日常の意識は、むしろ品目ごとに価格をそのまま調査し平均している小売物価統計調査の結果から直接計算した値に近いことが推測される。

値段に大きな変化がなくても、メーカーの販売戦略で消費者が気付かないように品質を下げる実質値上げ、IT 関連製品のように多くの消費者が意識しない性能が向上することによる実質値下げもある。

特に、近年注目されている品質調整による影響は、実感よりも下落幅を拡大する方向に働く場合に顕著である。例えば、パソコンの場合、容量、機能、アプリケーション、アクセサリ等が多様であるために、完全に同一の仕様、形式の価格変化を追跡することは困難である。パソコン(ノート型)の価格変化を平均価格の変化で見ると、2001年以降毎年、前年と比べてマイナス20%~プラス40%の範囲で推移している⁹⁴。しかし、それは時々主力モデルの間で価格を比べた結果に過ぎない。CPIでは、ヘドニック法を利用して、価格変動から品質向上分を除外した指数が作成されている。一般に、品質調整に伴う指数の下落幅の拡大はこのような製造技術の進歩が著しい耐久消費財で顕著である。

そこで、まず、小売物価統計調査の結果から直接品質調整をしない仮想的な指数を試算して⁹⁵、それがCPIとどの程度相違するのかをみてみることにした。なお、消費者はすべての品質調整をすぐには意識できないかというとはそうではなく、日常の商品で数量に変化が生じることがあれば⁹⁶、それを加味して物価を意識することになる。そこで、品質調整をしない指数の試算に当たっては、数量調整によるものを除外せずに含めて計算した。

図15は、1996年から2007年までの年平均値について品質調整をしない指数とCPIの間で前年比を比べたものである。ほとんどの年で品質調整をしない指数のほうがCPIよりも前年比が大きくなっている。

⁹⁴ 小売物価統計調査の東京都区部平均値による。

⁹⁵ この試算は、東京都区部に限定して行った。これは、全国では地域ごとに銘柄が異なる場合があり、試算を行うのが難しいためである。ただし、ウエイトは全国のものを用いている。将来的には、地域を限定せずにすべての銘柄について品質調整をしない価格をCPIと同一方式で合算して指数を試算することも考えられる。

⁹⁶ 例えば、チョコレートの大きさや紙おむつの枚数が変わるような、数量が変わることによる価格変動は、売り場で消費者が自然に意識できるものなので、品質の変化であっても実感に反映されるものとみなした。

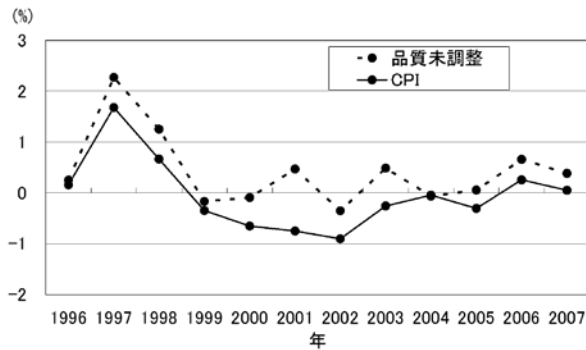


図 15 品質調整の有無による指数の前年比の比較

このような指数作成における品質調整による影響は、高度成長期にはあまり見られなかった現象である。

2.4 購入頻度による影響

第2の相違は、人々の実感は購入頻度の高い商品に左右されやすいという心理的印象に起因するものである。物価変動の実感は購入頻度に依存しやすいが、CPIは購入頻度でなく購入金額に基づいて計算される。

実際、CPIについて購入頻度が年9回以上の品目と9回未満の品目に区分して指数を作成し、両者の前年同月比を比較すると、図16のように、2006年1月以降毎月、前者は後者より一貫して高かった。この結果を見ると、人々の印象が購入頻度の高い品目に依存しやすいのであれば、物価変動の実感がCPIより高くなる。

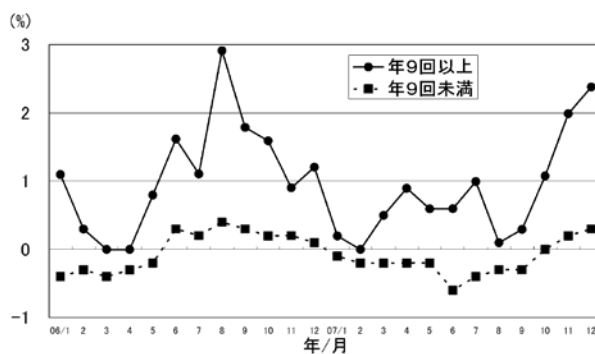


図 16 購入頻度別 CPI の前年同月比

図 17 は、1996年から2007年までの購入頻度年9回以上の品目に関する指数とCPIの間で前年比を比較したものである。これによると、2003年以降においては、2005

年を除き⁹⁷1ポイント程度後者が高くなっている。これは、価格下落の著しい耐久消費財の購入頻度が低いという事情を反映している。

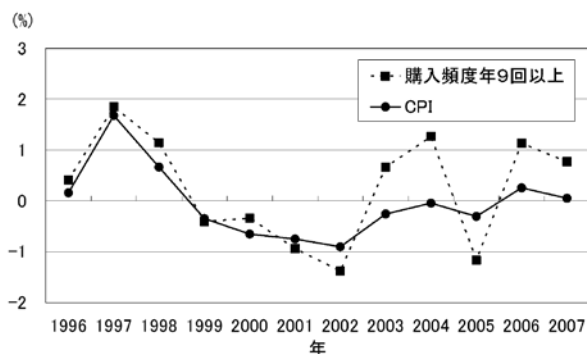


図 17 購入頻度による指数の前年比の比較

さらに、品質調整による影響と購入頻度による影響の両方の効果を同時に反映させるために、購入頻度の高い品目について品質調整をしない指数を作成し、図 18 で CPI との間で前年比を比較する。2000 年以前は両方の指数の間で差はほとんど見られなかったが、2001 年以降は 2005 年を除き購入頻度の高い品目について品質調整をしない指数のほうが前年比は高かった。特に、2007 年は CPI の前年比が 0.0%であったのに対し、購入頻度の高い品目について品質調整をしない指数は 1.7%であり、2006 年と比べると両者の差は拡大した。図 17 と比較すると、品質を調整しないことにより多くの年でグラフは少しだけ上方にシフトしている。これは、購入頻度の高い品目における品質調整の影響は小さいが、それでも平均すると指数を下方にシフトする傾向があることを示している。

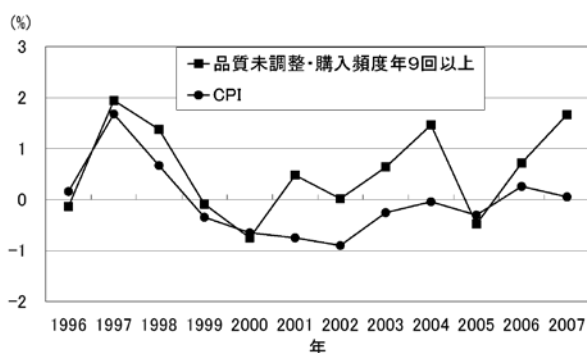


図 18 品質調整の有無及び購入頻度による指数の前年比の比較

⁹⁷ 2005 年は、生鮮野菜が前年の 2004 年の天候不順により高騰したことの反動で下落した。

2.5 直近の物価変動による影響

実感が指数の変化率を上回る要因は、物価と生活費の混同を含め、他にもいくつか存在するが、特に、大きな要因として、人々の実感は直近の物価の動きに左右されやすいという点がある。例えば、同じ1年の中の出来事であっても、直近1か月間の動きは、1年（12か月）前から11か月前までの1か月間の動きと比べて印象に残りやすいという現象である。この現象を少しでも反映させるために、購入頻度の高い品目について品質調整をしないだけでなく、直近6か月間の動き⁹⁸を年率換算してみることにする。

図19は、このように試算した指数を「生活意識に関するアンケート調査」の平均値と中央値のグラフに重ねてみたものであるが、特に、直近6か月間の動きを年率換算した指数は、意識調査の中央値と近い動きを示している。

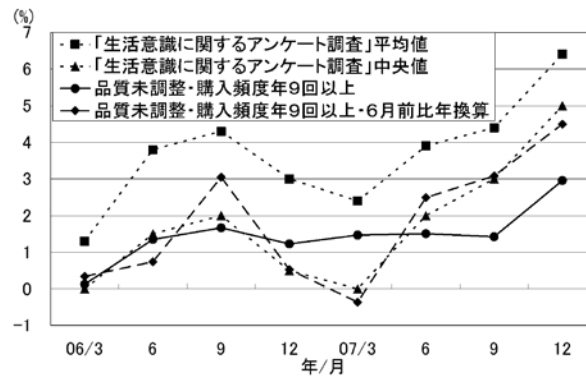


図19 物価に関する意識とCPIの基礎データから試算した指数の前年同月比

2.6 実感と指数の差の要因

以上のことから、物価の上昇率に関する実感と指数の差は、主として次の3つの要因、すなわち、①品質調整、②購入頻度、③直近の物価変動で相当程度の説明が可能である。また、CPIに利用されているデータを再編集することにより、これらの要因と実感物価変化率についてある程度の定量化が可能である。

しかし、上記の分析は簡素なものであり、今後、実感物価変化率を試算するに当たって、品質調整の除外の範囲や方法、直近の物価変動を反映する際の計算方法や季節変動の扱い等、更に検討すべき課題が存在している。

⁹⁸ 季節性が強いために欠価格となる月がある生鮮果物を除外して計算した。

2.7 2008年の状況

2008年に入ってから原油価格の高騰に起因する物価の上昇は深刻化した。また、ガソリン価格が暫定税率の期限切れにより4月だけ暫定的に下落し、消費者が短期間に物価の急激な上下動を経験したことなどにより、物価に関する意識はより直近の出来事に左右されやすくなった。

2008年は21世紀に入ってから物価上昇の実感が最も高くなった年である⁹⁹。2008年6月における「生活意識に関するアンケート調査」の中央値は10.4%に達した。他方、このときのCPIは前年同月比で2.0%の上昇、品目を購入頻度年9回以上に限定すると5.5%の上昇、さらに、購入頻度年9回以上の指数について6か月前からの変化を年率換算すると6.9%、3か月前からの変化を年率換算すると10.4%にも達する。これらに品質調整による影響を除外した試算はしていないが、それだけ高い物価上昇の実感があっても、実感による変化率は、実感に影響すると推察される条件をもとにCPIの基礎データを用いて説明できる範囲に収まっている。

3 EUと日本の共通点

ユーロ通貨流通後のEUと2008年の日本においては、いずれも①技術革新を始めとする品質向上に対する指数計算上の品質調整は消費者の実感に反映されにくい、②食料品やエネルギーなど購入頻度の高い品目で価格が上昇した等、共通に見られる現象が生じた。上述したように、これらはいずれも物価の変化率の実感が指数の変化率を上回りやすい環境であったと言える。

このような動きは、最近、他の国・地域でも見られる現象なので、今後も物価の実感をめぐる動きや分析について、定量的な比較を行い、国際レベルで情報の交換・共有を図っていくことが重要であると考えられる。

⁹⁹ 2014年にCPIの前年同月比が3%を上回る月が存在したが、「生活意識に関するアンケート調査」の中央値は6%を下回る水準であった。

おわりに

国民が意思決定を行うための情報基盤であるという公的統計の意義は、世界的には遅れていたとはいえ、統計法の全面改正から 10 年を経て日本でも定着してきたようである。研究においてはもちろんのこと、経営、医療、教育、スポーツなど様々な分野で統計の作成・利用が進むにつれて、政策のための統計利用にも変化が現れた。情報通信技術が発達し、データ量が増大し、それを利用するソフトウェアが進歩するにつれて、データの作成・利用が容易になっている。そのような中で、最初に政策目標を定め、後からそれを評価するための統計データを探すという動きが広がりつつある。典型事例は SDGs である。以前であれば、まず統計調査が存在し、その目的に沿った方式で結果が利用され、それ以外の利用の仕方を排除する傾向があった。しかし、ビッグデータの時代において、データの源は統計調査とは限らず、行政記録、地理情報、民間データなど様々なものが想定され、むしろ政策目的に応じてそれらを組み合わせるといった対応が一般化してきた。データが存在しない部分があれば、そこだけデータを入手する、他のデータで推定するなど、方法論における対処は当然かつ迅速になされるのが前提になっている。

本論文で論じた研究は、その具体例を示すという意味でも一定の役割を果たしたと考える。しかし、今後に向けた課題もいくつか存在しており、必要に応じて各章・各節の最後に記載した。本論文の最後にそれらに共通する課題に触れたい。

第 I 部と第 II 部で説明した内容については、最新の情報とデータを用いた実績が存在している。例えば、第 II 部第 1 章の季節調整法については 2015 年（度）までのデータを用いた実証分析となっている。また、第 II 部第 2 章の国勢調査を用いた JGSS の比推定乗率について、筆者自身は 2000 年までのデータで検証をしているが、その後も、JGSS が実施される都度、他の研究者が引き継ぎ、時々の最新の国勢調査の結果を用いて同様の検証を積み重ねている。しかし、第 III 部と第 IV 部については、結果の概要に変化がないことが想定されるが、大半が 2000 年代までのデータを用いた研究に留まっている。このため、2010 年代のデータを用いて研究を更新することが課題である。

そのためにはデータが存在することが要件となる。その意味では、2015 年国勢調査の結果が次々と公表されており、また、2006 年 12 月から国民経済計算が 2008SNA に準拠することになるので、間もなく研究更新の環境が整うことになる。もっとも、本論文で利用したデータは、更新されたデータが必ずしも入手できるとは限らず、調査が中断されてしまった項目も存在する。調査項目は必要性和回答負担のバランスで決まるので、今後この分野の研究の必要性を強調することが重要であると考えられる。

最後に、関連情報と個別データの提供を始め、本論文の研究一つ一つにご協力いただいた方々、また、このような形で一つの論文にまとめるに当たりご助言をいただいた方々に心より御礼を申し上げます。

参考文献

第 I 部第 1 章

- [1]清水誠 (2001)「官庁統計の意義と利用」『日本統計学会誌』31-3, 日本統計学会, p277-304
- [2]Makoto Shimizu (2009) “New Statistical System for Official Statistics in Japan”, 57th Session of the ISI
- [3]永山貞則 (1992)「開かれた政府統計の在り方」『日本統計学会誌』21-3, 日本統計学会, p277-282
- [4]大屋祐雪 (1993)「官庁統計制度と統計調査の現状—序章」, 官庁統計制度と統計調査の現状, 『日本統計学会誌』22-3, 日本統計学会, p615-620
- [5]竹内啓 (2001)「統計と民主主義」『ESTRELA』87, (財)統計情報研究開発センター, p36-39

第 I 部第 2 章

- [6]United Nations (2014) “Fundamental Principles of Official Statistics”, Resolution Adopted by the General Assembly
- [7]UNECE (2011) “Using Administrative and Secondary Sources for Official Statistics : A Handbook of Principles and Practices”
- [8]UNECE (2009) “Principles and Guidelines on Confidentiality Aspects of Data Integration Undertaken for Statistical or Related Research Purposes”
- [9]EU, OECD (2007) “Eurostat-OECD Manual on Business Demography Statistics”
- [10]UNECE (2015) “Guidelines on Statistical Business Registers”
- [11]Eurostat (2013) “ESS Guidelines on Revision Policy for PEEIs”
- [12]Eurostat (2009) “ESS Guidelines on Seasonal Adjustment”
- [13]Eurostat (2015) “ESS Guidelines on Seasonal Adjustment”
- [14]CES (2007) “Managing Statistical Confidentiality & Microdata Access: Principles and Guidelines of Good Practice”
- [15]UNSC, CES (2000) “Evaluating Efficiency of Statistical Data Editing: General Framework”
- [16]UNECE (2015) “Making Data Meaningful, Part 1, A Guide to Writing Stories about Numbers”

第 II 部第 1 章

- [17]清水誠 (1988)「季節調整法センサス局法 II, X11-ARIMA の適用」『統計局研究彙報』47, 総務庁統計局・統計センター, p29-86
- [18]木村武 (1996)「最新移動平均型季節調整法「X-12-ARIMA」について」『金融研究』15-2, 日本銀行金融研究所, p95-150
- [19]経済指標部会季節調整法検討小委員会 (1997)『季節調整法検討小委員会報告書』
- [20]上田聖 (1998)「季節調整プログラム X-12-ARIMA の安定性と適切性向上のためのオプション選択」『統計局研究彙報』56, 総務庁統計局・統計センター, p1-44
- [21]奥本佳伸 (2000)「季節調整法の比較研究, センサス局法 X-12-ARIMA の我が国経済統計への適用」『経済分析』政策研究の視点シリーズ 17, 経済企画庁経済研究所
- [22]清水誠 (2005)「統計局における時系列データの季節性と季節調整法」『統計研究彙報』63, 総務省統計研修所, p1-20
- [23]清水誠 (2010)「個人企業経済調査の季節調整法の検討」『統計研究彙報』67, 総務省統計研修所, p1-22
- [24]北川源四郎 (1986)「時系列の分解—プログラム DECOMP の紹介—」『統計数理』34-2, p255-271
- [25]Time Series Research Staff, Center for Statistical Research and Methodology, U.S. Census Bureau (2016) “X-13ARIMA-SEATS Reference Manual Version 1.1”

- [26] Víctor Gómez, Agustín Maravall (1996) “Programs TRAMO and SEATS, Instructions for the User”, Banco de España

第Ⅱ部第2章

- [27] 清水誠 (2001) 「JGSS 第2回予備調査 データセットのウエイトについて」, SSJ Data Archive Research Paper Series 14, 東京大学社会科学研究所
[28] 清水誠 (2002) 「5. データセットのウエイトについて」, I. JGSS-2000 調査概要, 『日本版 General Social Surveys JGSS-2000 基礎集計表・コードブック』東京大学社会科学研究所 資料 第19集, 東京大学社会科学研究所, p15-24

第Ⅲ部第1章

- [29] 清水誠 (2002) 「人口が経済成長に及ぼす影響 —90年代の定量的評価—」『社会科学研究』53-5, 東京大学社会科学研究所, p97-117
[30] N. Gregory Mankiw, David Romer, David N. Weil (1992) “A Contribution to the Empirics of Economic Growth”, *The Quarterly Journal of Economics* 107-2, p407-437

第Ⅲ部第2章

- [31] William P. Butz, Michael P. Ward (1979) “The Emergence of Countercyclical U.S. Fertility”, *American Economic Review* 69-3, p318-328
[32] John Ermisch (1979) “The Relevance of the ‘Easterlin Hypothesis’ and the ‘New Home Economics’ to Fertility Movements in Great Britain”, *Population Studies* 33-1, p39-58
[33] Naohiro Ogawa, Andrew Mason (1986) “An Economic Analysis of Recent Fertility in Japan: An Application of the Butz-Ward Model” 『人口学研究』9, p5-15
[34] 大沢真知子 (1993) 『経済変化と女子労働 日米の比較研究』, 日本経済評論社
[35] 大淵寛 (1988) 『出生力の経済学』, 中央大学出版部
[36] Hisakazu Kato (1997) “Time Series Analysis of Fertility Change in Postwar Japan” 『人口学研究』20, p23-35
[37] 清水誠 (2000) 「所得が出生に及ぼす影響 —JGSS-2000へのButz and Ward Modelの適用—」『日本版 General Social Surveys 研究論文集 JGSS-2000 で見た日本人の意識と行動』, 東京大学社会科学研究所, p149-158
[38] 清水誠 (2008) 「1章 所得と出産行動」『日本人の意識と行動 日本版総合的社会調査 JGSS による分析』谷岡一郎・仁田道夫・岩井紀子 編, 東京大学出版会, p47-57

第Ⅲ部第3章

- [39] 清水誠・玉熊亜友美 (2009) 「科学技術研究調査による企業の生産性の要因分析」『統計研究彙報』66, 総務省統計研修所, p1-31
[40] OECD (2007) “OECD Science, Technology and Industry Scoreboard 2007, Innovation and Performance in the Global Economy”
[41] Congressional Budget Office, Congress of the United States (2005) “R&D and Productivity Growth”, Congressional Budget Office Background Paper
[42] Akira Goto, Kazuyuki Suzuki (1989) “R&D Capital, Rate of Return on R&D Investment and Spillover of R&D in Japanese Manufacturing Industries”, *The Review of Economics and Statistics*, 71-4, The MIT Press, p555-564
[43] 高橋通典 (2004) 「90年代以降の企業の研究開発動向」『調査』63, 日本政策投資銀行
[44] 中野諭 (2008) 「企業レベルのR&Dストックと全要素生産性の計測 —『科学技術研究調査』個票データによる観察—」『科学技術統計応用調査研究報告書』丹羽富士雄, 中野諭, 伊藤万里, (財)政策科学研究所, p41-90

第IV部第1章

- [45]清水誠 (1999)「米国 CPI をめぐる議論(1)」『統計』50-11, (財)日本統計協会, p75-82
- [46]清水誠 (1999)「米国 CPI をめぐる議論(2)」『統計』50-12, (財)日本統計協会, p93-101
- [47]清水誠 (2000)「米国 CPI をめぐる議論(3)」『統計』51-1, (財)日本統計協会, p59-66
- [48]清水誠 (2000)「米国 CPI をめぐる議論(4)」『統計』51-2, (財)日本統計協会, p65-71
- [49]清水誠 (2000)「米国 CPI をめぐる議論(5)」『統計』51-3, (財)日本統計協会, p64-71
- [50]清水誠 (2000)「米国 CPI をめぐる議論(6)」『統計』51-4, (財)日本統計協会, p72-76
- [51]Advisory Commission to Study the Consumer Price Index (1996) “Toward a More Accurate Measure of the Cost of Living; Final Report”
- [52]Bureau of Labor Statistics (1997) “Measurement Issues in the Consumer Price Index”
- [53]Subcommittee on Economic Statistics, Joint Economic Committee, Congress of the U.S. (1961) “Government Price Statistics Report”
- [54]Erwin W. Diewert (1976) “Exact and Superlative Index Numbers”, *Journal of Econometrics* 4, p115-145
- [55]A. A. Konüs (1924) “The Problem of the True Index of the Cost of Living”, translated in *Econometrica* 7 (1939) , p10-29
- [56]Ralph Turvey (1999) “True Cost of Living Indexes”, 5th Ottawa Group Meeting
- [57]Erwin W. Diewert (1999) “Comments on: True Cost of Living Indexes by Ralph Turvey”, 5th Ottawa Group Meeting
- [58]Erwin W. Diewert (1999) “The Consumer Price Index and Index Number Purpose”, 5th Ottawa Group Meeting
- [59]Jack E. Triplett (1992) “Economic Theory and BEA’s Alternative Quantity and Price Indexes”, *Survey of Current Business* 72-4, p49-52
- [60]Jörgen Dalén (1999) “Some Issues in CPI Construction”, 5th Ottawa Group Meeting
- [61]Carl Obst (1999) “A Review of Bias in the CPI”, Joint UNECE/ILO Meeting
- [62]ILO, IMF, OECD, UNECE, Eurostat, The World Bank (2004) “Consumer Price Index Manual: Theory and Practice”
- [63]Andrew T. Court (1939) “Hedonic Price Indexes with Automotive Examples”, in “The Dynamics of Automobile Demand”, General Motors, p98-119
- [64]Dennis Fixler, Charles Fortuna, John Greenlees, Walter Lane, U.S. Bureau of Labor Statistics (1999) “The Use of Hedonic Regressions to Handle Quality Change: The Experience in the U.S. CPI”, 5th Ottawa Group Meeting
- [65]Jack E. Triplett (1991) “Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy”, Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth
- [66]John A. Astin, Don J. Sellwood (1997) “Harmonization in the European Union: A Review of Some Technical Issues”, 3rd Ottawa Group Meeting
- [67]John R. Hicks (1940) “The Valuation of the Social Income”, *Economica* 7-26, p105-124
- [68]Brent R. Moulton (1996) “Bias in the Consumer Price Index: What is the Evidence?”, *The Journal of Economic Perspectives*, 10-4, p159-177
- [69]Peter Hill (1999) “Inflation, the Cost of Living and the Domain of a Consumer Price Index”, Joint UNECE/ILO Meeting
- [70]Jack E. Triplett “Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?”, 5th Ottawa Group Meeting
- [71]General Accounting Office (2000) “Consumer Price Index: Update of Boskin Commission’s Estimate of Bias”
- [72]CNSTAT (2002) “At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes”, *National Academy Press*

- [73]Keith Woolford (2005) “The Boskin Report and Price Measurement in Australia”, Inflation Measures: Too High - Too Low - Internationally Comparable?, OECD Seminar
- [74]Paolo del Giovane, Roberto Sabbatini (2005) “The Introduction of the Euro and the Divergence between Officially Measured and Perceived Inflation : The Case of Italy”, Inflation Measures: Too High - Too Low - Internationally Comparable?, OECD Seminar
- [75]David S. Johnson, Stephen B. Reed, Kenneth J. Stewart (2006) “Price Measurement in the United States: A Decade after the Boskin Report”, Monthly Labor Review

第Ⅳ部第2章

- [76]Makoto Shimizu (2005) “Recent Methodological Developments of the CPI in Japan”, Inflation Measures: Too High - Too Low - Internationally Comparable?, OECD Seminar
- [77]Makoto Shimizu (2006) “Methodological Developments in the 2005 Revision of the CPI in Japan”, Joint UNECE/ILO Meeting
- [78]Nobuyuki Sakashita, Makoto Shimizu (2004) “CPI and Household Income Expenditure under Deflationary Trend”, 8th Ottawa Group Meeting
- [79]Makoto Shimizu (2007) “Telecommunication Indices in the Japanese CPI”, 10th Ottawa Group Meeting
- [80]清水誠 (2006) 「消費者物価指数の実際」『経済セミナー』612, 日本評論社, p50-56
- [81]Robert Cage, John Greenlees, Patrick Jackman (2003) “Introducing the Chained Consumer Price Index”, 7th Ottawa Group Meeting
- [82]清水誠・永井恵子 (2006) 「CPI に関する取組 2005～06 (1) —食料とエネルギーの取扱について—」『統計』57-9, (財) 日本統計協会, p84-90
- [83]Juan-Luis Vega, Mark A. Wynne (2001) “An Evaluation of Some Measures of Core Inflation for the EURO Area”, Working Paper Series 53, European Central Bank
- [84]清水誠・永井恵子 (2006) 「CPI に関する取組 2005～06 (2) —平成 17 年基準改定における新旧指数の差について—」『統計』57-10, (財) 日本統計協会, p76-81
- [85]Makoto Shimizu (2007) “Quantitative Evaluation of the 2005 Revision of the CPI in Japan”, 56th Session of the ISI
- [86]Johannes Hoffmann, Claudia Kurz (2002) “Rent Indices for Housing in West Germany 1985 to 1998”, Working Paper Series 116, European Central Bank
- [87]Makoto Shimizu (2006) “Review of the Prices of Rents and Owner-occupied Houses in Japan”, 9th Ottawa Group Meeting
- [88]久我真真理子 (2007) 「小売物価統計調査における家賃調査に関する一考察」『統計研究彙報』64, 総務省統計研修所, p17-26
- [89]Makoto Shimizu (2009) “Review for the Estimation with Regression Equations for Rent Indices in Japan”, 11th Ottawa Group Meeting
- [90]清水誠・永井恵子 (2006) 「CPI に関する取組 2005～06 (3) —ヘドニック法について—」『統計』57-11, (財) 日本統計協会, p74-80
- [91]清水誠・永井恵子 (2006) 「CPI に関する取組 2005～06 (最終回) —IMF における国際基準, 規範の遵守状況に関する報告書 (ROSC) について—」『統計』57-12, (財) 日本統計協会, p84-89

第Ⅳ部第3章

- [92]永山貞則 (1963) 『物価と家計』, 統計新書 10, 一粒社
- [93]辻村江太郎 (1969) 「物価指数と生活実感」『統計』20-2, (財) 日本統計協会, p1-6
- [94]三瀧信邦 (1974) 「統計・実感・実在 —消費者物価指数をめぐる—」『統計』25-8, (財) 日本統計協会, p55-56

- [95]北山直樹 (1975) 「消費者物価指数とインフレーション — 指数と実感の周辺—」『統計』 26-2, (財) 日本統計協会, p1-6
- [96]中村隆英 (1976) 「物価指数と「生活実感」」『統計』 27-6, (財) 日本統計協会, p55-56
- [97]中村隆英 (1981) 「指数と実感」『統計』 32-10, (財) 日本統計協会, p6-11
- [98]大槻義雄 (1983) 「物価の実感上昇率に関する考察 — 内閣総理大臣官房広報室世論調査の結果から—」『統計局研究彙報』 40, 総理府統計局, p1-27
- [99]Makoto Shimizu (2008) “Empirical Analysis of the Difference between Measured and Perceived Inflation in Japan”, Joint UNECE/ILO Meeting
- [100]清水誠 (2008) 「最近の物価の実感に関する定量的評価」『統計』 59-10, (財) 日本統計協会, p2-9
- [101]L. Aucremanne, M. Collin, E. Dhyne (2005) “Is There a Discrepancy between Measured and Perceived Inflation in the Euro Area Countries Since the Euro Cash Changeover?”, Inflation Measures: Too High - Too Low - Internationally Comparable?, OECD Seminar
- [102]L. Aucremanne, M. Collin, T. Stragier (2007) “Assessing the Gap between Observed and Perceived Inflation in the Euro Area: Is the Credibility of the HICP at Stake? ”, *NBB Working Paper* No.112
- [103]Dominique Guédès (2005) “Dispute around Inflation Since Euro Changeover in France”, Inflation Measures: Too High - Too Low - Internationally Comparable?, OECD Seminar