

# 経済統計・政府統計の数理的基礎と応用-III<sup>1</sup>

川崎能典・国友直人・山本拓 共編

2014年1月

<sup>1</sup>この報告集は文部科学省・科学研究費プロジェクト「経済統計・政府統計の数理的基礎と応用」（2011年度～2014年度）が開催した研究集会における講演内容をまとめたものである。

# 前書き

本報告書は、日本学術振興会・科学研究プロジェクト「経済統計・政府統計の数理的基礎とその応用」（2011年度－2014年度、研究代表者：山本 拓）が、2014年1月31日（金）に東京大学小島ホールにおいて開催した2013年度の研究集会における講演内容をまとめたものである。

本プロジェクトの目的は、経済統計・政府統計における主要な課題の、技術的および制度的問題を、統計学的な立場から理論的・学術的に検討し、具体的解決策を提言することである。

経済統計、とりわけ政府統計は、経済・社会の動向を理解し、政策を実施、評価するためには不可欠な情報であることは言うまでもない。最近は evidence-based policy ということもよく言われ、政府統計の重要性は一般に広く認識されつつあると思われる。しかし、経済統計・政府統計への信頼性は、近年必ずしも増しているとは言えない状況である。経済社会の急激な変化に伴い、政府統計の質の確保が困難になりつつある。マクロ経済統計の側面では、GDP 統計などに代表されるマクロ公表系列の質と信頼性の問題、信頼性の高い将来人口の推計の問題、地域による経済情勢のばらつきの把握などの問題を挙げるができる。またミクロ経済データにおいては、統計調査をとりまくプライバシー意識の高まりから、調査精度の確保が難しくなりつつあるという問題や、情報開示と秘密保持の両立という匿名化問題などを挙げるができる。

新しい統計学的知見の導入に関しては、日本の政府統計部局が分散化されているために、これまでは、個別の担当部局あるいはその時々担当者に個別に招かれた研究者によって知見や助言が提供されることが多かった。政府統計を巡る重要な論点について、担当部局をまたいでその知見が共有されることは少なかったと思われる。またそれらの話題が広く研究者間で議論されることも少なかった。そのような意味で、経済統計・政府統計の技術的・制度的問題点を、統計学的立場から総括的に検討していくという本研究プロジェクトは、一つの新しい方向性を示している。

本プロジェクトの研究集会は、プロジェクトのメンバーと実際に経済統計・政府統計に作成者または利用者として携わっている方々との積極的な交流を目指している。したがって研究集会における研究報告は、メンバーと外部の方の報告が概ね半々になるように構成されている。

2011年度の第1回目の研究集会の特徴は、外部の報告者として、実際に主要な政府統計を作成されている担当者を招き、作成上のポイントや課題を報告して頂いたことである。2012年度の第2回目の研究集会の特徴は、外部の報告者としては地方政府にお

いて統計に関わっている方に、そのあり方や課題などについて報告して頂いたことにある。さらにマクロ経済統計の作成者および利用者としての立場からその問題点や改善の方向性についての報告を頂いた。これら2回の研究集会の報告は、それぞれ東京大学大学院経済学研究科附属・日本経済国際共同研究センター（CIRJE）研究報告書シリーズのCIRJE-R-10ならびにCIRJE-R-12にまとめられている。

今年度は第3回目の研究集会として、外部の報告者として雇用・失業統計、人口統計、ならびに生産性統計についての報告をして頂いた。プロジェクトのメンバーからは物価指数、ならびにデータの匿名化についての報告が行われた。さらに季節調整の様々な問題について、海外からの招聘研究者を含む外部の方から報告して頂くと同時に、プロジェクトのメンバーの報告が英語セッションとして行われた。

このような機会が情報交換ならびにお互いの刺激となり、経済統計・政府統計の今後の改善の一助となることを期待する次第である。

2014年2月

**編者**

# プログラム

## <セッション I>

Chair: 山本拓

10:00-10:40 「雇用・失業統計に関する国際統計基準の動向」 川崎茂(日本大学)

10:40-11:30 「JIP データベースにおける全要素生産性の計測と公的統計の課題」  
深尾京司(一橋大学)

11:30-12:10 「人口減少・高齢社会の進展と人口統計の課題」  
金子隆一(国立社会保障・人口問題研究所)

<昼休み>

## <セッション II>

Chair: 国友直人

13:30-14:10 「高齢者世帯の消費行動と物価指数」  
宇南山卓(財務省財務総合政策研究所)

14:10-14:50 「エビデンスに基づいた匿名化」 星野伸明(金沢大学)

## <セッション III (英語セッション) >

Chair: Yoshinori Kawasaki

15:00-15:40 “Benchmarking Problem And Robust Denton Method” Kengo Kato  
(東京大学)

15:40-16:20 “On Official Seasonal Adjustments in Japan” Makoto Takaoka  
(琉球大学)

16:20-17:00 “Seasonal Adjustment of CPS Labor Force Series during the Great  
Recession” Thomas D. Evans (U. S. Bureau of Labor Statistics)

2014年1月31日 研究集会  
「経済統計・政府統計の数理的基礎  
と応用」  
於：東京大学経済学部小島ホール

# 雇用・失業統計に関する 国際基準の動向

日本大学経済学部

川崎 茂

[kawasaki.shigeru@nihon-u.ac.jp](mailto:kawasaki.shigeru@nihon-u.ac.jp)

## 目次

1. はじめに
2. 我が国の雇用・失業統計
3. 現行の国際基準
4. 国際基準の改訂
5. 労働力調査における課題
6. まとめ

## 1. はじめに

- 昨年10月に開催された国際労働統計家会議において、雇用・失業統計の国際基準の変更が決定された。
- 新たな国際基準では、「仕事の形態」の概念の新設、雇用・失業の定義の明確化、潜在労働力人口の概念の新設などが行われた。
- この変更に伴い、日本の労働力調査においても対応が必要となる。
- この報告では、国際基準の変更の内容とその影響について紹介する。

(参考) ILO International Conference of Labour Statisticians "Resolution concerning statistics of work, employment and labour underutilization" (Oct. 2013)  
<http://www.ilo.org/global/statistics-and-databases/meetings-and-events/international-conference-of-labour-statisticians/19/lang--en/index.htm>

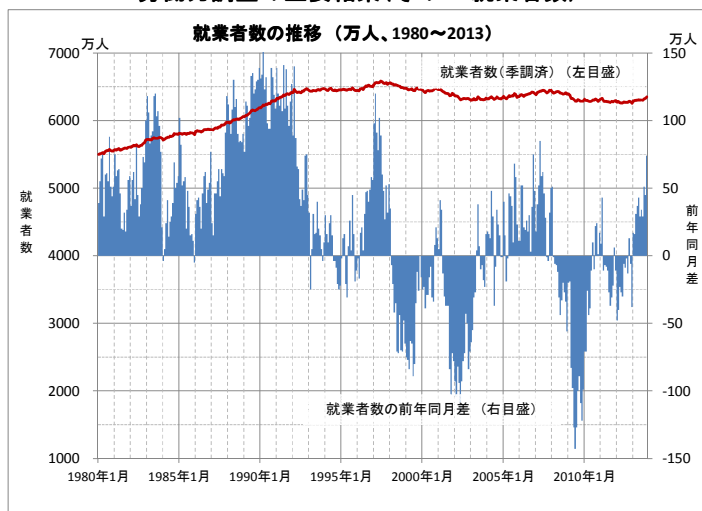
## 2. 我が国の雇用・失業統計 (1)

- 我が国の雇用・失業の実態を把握する最も基本的な統計調査として、労働力調査が1947年から毎月実施されている。
- 労働力調査では、全国約4万世帯、10万人(15歳以上)を2段抽出により選定・調査し、翌月末に結果を公表。
- 調査対象全員について、就業状態、産業、職業、週間就業時間、求職理由等を調査する(基礎調査票)ほか、約1/4には、就業・失業等のより詳細な実態について調査(特定調査票)。
- これらの統計は、ILOにおいて定められた国際統計基準に従って作成。
- 毎月公表される就業者数、完全失業者数等の動向は、景気指標として多くの利用者が注目。

(参考) 総務省統計局「労働力調査」 <http://www.stat.go.jp/data/roudou/index.htm>

## 2. 我が国の雇用・失業統計 (2)

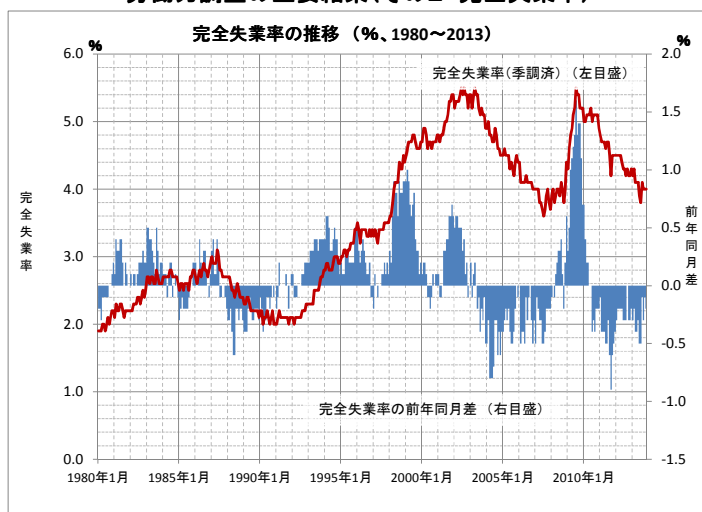
### 労働力調査の主要結果(その1 就業者数)



(出典) 総務省統計局「労働力調査」ホームページからデータをダウンロードして作成

## 2. 我が国の雇用・失業統計 (3)

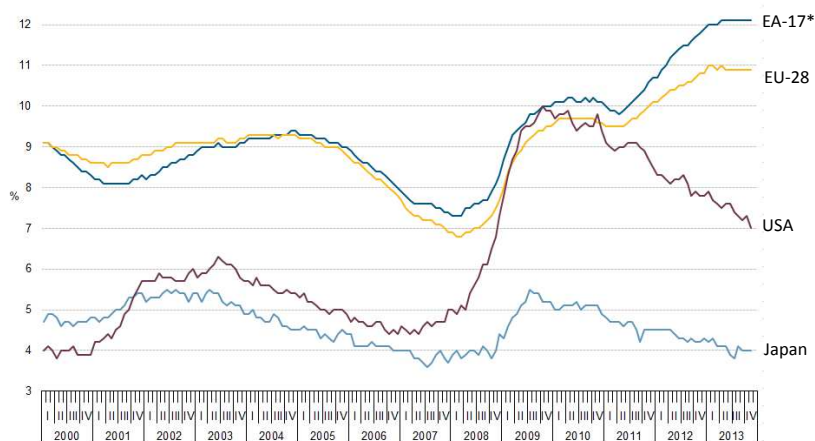
### 労働力調査の主要結果(その2 完全失業率)



(出典) 総務省統計局「労働力調査」ホームページからデータをダウンロードして作成

## 2. 我が国の雇用・失業統計 (4)

日本、アメリカ、EUの失業率 (2013年10月、季節調整値)



(出典) Eurostat "Unemployment Statistics" (\*) EA-17とは、通貨ユーロを採用している17か国のこと。  
[http://epp.eurostat.ec.europa.eu/statistics\\_explained/index.php/Unemployment\\_statistics](http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Unemployment_statistics)

## 3. 現行の国際基準 (1)

### (1) 統計の国際基準とは

- 公的統計の基本原則、概念・定義、分類、調査・推計の手法など様々な側面に関し、信頼性及び国際比較性の向上のため、国際機関と各国が協力して決定。
- 基準の多くは国連(統計部及び統計委員会)が中心となって検討及び決定。

(参考) United Nations Statistics Division: <http://unstats.un.org/unsd/default.htm>

- 労働関連統計の国際基準は、ILOのICLS\*において決議の後、理事会の同意により決定。

(参考) ICLS = International Conference of Labour Statisticians

1923年設置。各国の政府統計部局の統計専門家のほか、労働側、使用者側の代表等も出席。戦時を除き、約5年ごとに開催。2013年の第19回会議には約100ヶ国から参加。



## 3. 現行の国際基準 (2)

### (2) 雇用・失業統計に関する現行の国際基準

- 雇用・失業、労働時間、産業分類、職業分類、従業上の地位の概念・定義、不完全就業等について国際基準が存在。
- 今回の改訂に関連する主なものは次のとおり。
  - 1982年 Resolution concerning statistics of the economically active population, employment, unemployment and underemployment
  - 1998年 Resolution concerning the measurement of underemployment and inadequate employment situations
- 最も基本となる1982年の雇用・失業に関する基準は、経済活動人口、就業者、失業者等の概念を明確化したもの。
- 失業については、①仕事を持っていない、②現に就業が可能である、③実際に仕事を探していた の3要件すべてを満たす者として定義。
- 要件③は、「最近の特定の期間に、有給雇用又は自営業を探すために特別な手段を講じた」事実の有無により判定。
- 求職活動の期間に関しては、具体的な規定なし。

## 4. 国際基準の改訂(1)

### (1) ICLSの決議の概要

- 名称は、“Resolution concerning statistics of work, employment and labour underutilization”
- 近年の労働市場における新たな実態をよりの確に把握し、政策・分析のニーズにより活用できるものとなるよう見直し。主な内容は次のとおり。
  - 1) 新たに「仕事の形態(Forms of Work)」の概念を導入。  
2008SNAとの整合性に配慮し、「仕事」の範囲を幅広くとらえる。
  - 2) 労働力状態の分類・定義を見直し、明確化。  
雇用・失業のボーダーラインに近い状態の者の扱いを明確化する。  
例 Unpaid traineesの扱い、失業の要件における求職活動の期間の厳密化
  - 3) 未活用の労働力に関する統計の概念の新設。  
「時間関連不完全就業(Time Related Underemployment)」の定義を明確化。  
「潜在労働力(Potential Labour Force)」の概念と指標を規定。

## 4. 国際基準の改訂(2)

### (2) 「仕事の形態」の概念の新設

- ・仕事(Work)の概念を2008SNAの一般生産境界と一致させる。

Intended destination of production	for own final use		for use by others				
	Forms of work	Own-use production work		Employment (work for pay or profit)	Unpaid trainee work	Other work activities	Volunteer work
of services		of goods	in market and non-market units				in households producing goods services
Relation to 2008 SNA	Activities within the SNA production boundary						
	Activities inside the SNA General production boundary						

- ・仕事(Work)には、雇用・自営の仕事のほか、自家使用のための財・サービスの生産、無給訓練生、ボランティアなども含まれる。

## 4. 国際基準の改訂(3)

### (3) 労働力状態の分類の定義の見直し・明確化

#### ○ 休業者(就業者の内訳)

雇用の継続性(休暇中の給与・給付等の受取、職場への復帰など)が確保されている場合にのみ就業者とする。

#### ○ 失業者(求職活動期間)

失業者の要件は、①仕事を持っていない、②現に就業が可能である、③実際に仕事を探していたと従前どおりだが、③の職探しの期間は、従前の曖昧さを改め、過去4週間又は1ヶ月と定める。

就業内定のため職探しをしていなかった者、海外出稼ぎのため出国待ちの者は失業者とする。

## 4. 国際基準の改訂(4)

### (4) 時間関連不完全就業の定義の明確化

- 不完全就業については、これまで統一的な基準で調査されていないことから、新基準において明確化が図られた。
- 時間関連不完全就業とは、就業者のうち、次の条件をすべて満たす者とされた。特定の短い期間のうちに、
  - ①労働時間の増加を希望している
  - ②実際の労働時間が定められた下限を下回る
  - ③追加就業機会があれば就業可能である
- ①は、現在と同じ仕事でも、別の仕事でも可。
- ②は、フルタイム雇用者とパートタイム雇用者の労働時間の境界となる労働時間又は法制上の規範により定める。

## 4. 国際基準の改訂(5)

### (5) 潜在労働力の概念と指標 その1

- 潜在労働力を、非労働力人口のうち、失業の3要件の一部を満たしていないものとして定義。具体的には次の2種類。
  - ①最近の一定期間の間に求職活動を行った者で、現在は就業可能ではないが、しばらくの後に就業可能となる見込みの者
  - ②求職活動は行っていないが、就業を希望しており就業可能な状態の者
- ②のうち、労働市場の見通しから求職活動を控えている者を「求職意欲喪失者(Discouraged workers)」として区別することは有用。
- 労働力と潜在労働力の和を、拡張労働力(Extended Labour Force)と呼び、労働力の低利用(Labour Underutilization)を示す指標の分母人口として使用。

## 4. 国際基準の改訂(6)

### (6) 潜在労働力の概念と指標 その2

- 異なる経済事情や局面を分析するために、労働力の低利用の状態を表す次の指標がの潜在労働力が有用。

LU1: Unemployment Rate

[persons in unemployment/labour force] X 100

LU2: Combined Rate of Time-Related Underemployment and Unemployment

[(persons in time-related underemployment + persons in unemployment) / (labour force)] X 100

LU3: Combined Rate of Unemployment and Potential Labour Force

[(persons in unemployment + potential labour force) / (extended labour force)] X 100

LU4: Composite Measure of labour underutilization

[(persons in time-related underemployment + persons in unemployment + potential labour force) / (extended labour force) X 100

## 5. 我が国への影響 (1)

### (1) 国際基準の変更に伴う影響

- 休業者の扱い……ほとんど影響しない
- 失業者の求職期間……失業者数に影響が生じる
- 就業内定者の扱い……失業者数に影響が生じる
- 時間関連不完全就業の定義の明確化
  - ……調査事項の追加が必要
- 潜在労働力人口……調査事項の再検討が必要

## 5. 我が国への影響 (2)

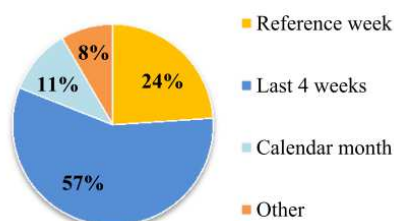
### (2) 失業者の求職期間

- 新基準では、過去4週間又は1ヶ月の求職活動の有無により失業を判定。日本では月末1週間の活動で判定。
- 日本と同様の国は世界の約1/4。
- 労働力調査2012年結果によれば、就業希望の非労働力人口のうち、過去1か月間に求職活動を行った者は56万人。

注意：特定調査票の情報を集計した詳細集計結果による。なお、この中には②現に就業可能という要件を満たさない者も含まれる。

(参考：完全失業者数=285万人)

Figure 16a: Reference period used for measuring active job search in LFS (%)



Source: ILO Department of Statistics "National practices in the measurement of the economically active population, employment, unemployment and time-related underemployment - Household-based sources" 2013

## 5. 我が国への影響 (3)

### (3) 就職内定者の扱い

- 新基準では、就業内定のため職探しをしていなかった者、海外出稼ぎのため出国待ちの者は失業者と扱われる。
- 日本では、就職内定者は原則として非労働力として扱われている。
- 日本で新基準によった場合、毎月の失業者数がどの程度増加するか正確な数字は不明。
- 仮に大学卒業(予定)者のうち就職内定者が、就職を待っている場合にはすべて失業者として扱われるとすれば、3月には約35万人が失業者に加わることになる。(学校基本調査(平成24年度)の就職者数による。)

## 5. 我が国への影響 (4)

### (4) 時間関連不完全就業者の定義の明確化

- 新基準では、時間関連不完全就業者を調査するために、次の要件を満たしていることを確認することが必要とされる。
  - ①労働時間の増加を希望している
  - ②実際の労働時間が定められた下限を下回る
  - ③追加就業機会があれば就業可能である
- ①については、労働力調査の特定調査票の質問A2「仕事時間についての希望はありますか」で把握可能。
- ②については、基礎調査票の質問8により月末1週間の労働時間を尋ねているので把握可能。
- ③については、既存の調査事項がないため、新たな設問が必要。また、質問・回答肢について十分な検討が必要。

## 5. 我が国への影響 (5)

### (5) 潜在労働力の把握

- 潜在労働力の把握には、非労働力人口に対して、①就業希望の有無、②求職活動を行ったか否か(調査週間よりも前の時期)、③就業可能な状態か否か、を調査することが必要。
- ①については、特定調査票のC1(収入になる仕事に就くことを希望しているか)により把握可能。
- ②については、同じくC4(この1年間に仕事を探したり会合の準備をしたことがあるか)により把握可能。
- ③については、同じくC5(今仕事があればすぐ就くことができるか)により把握可能。
- ただし、いずれの場合も、質問・回答肢の検討が必要。

## 5. 我が国への影響 (6)

### (6) (参考)現在の労働力統計から得られる潜在労働力の情報

- ・現在の労働力調査では、特定調査票において就業希望の有無、就業可能性等の事項が調査されていることから、潜在労働力についてはある程度把握が可能である。
- ・例えば、求職意欲喪失者については、非労働力人口のうち、①就業を希望し、②仕事がありそうにないために求職活動をしておらず、③仕事があればすぐ就け、かつ、④過去1年以内に求職活動を行ったことがある者がおおむね該当すると考えられる。その数は、平成24年には37万人と推定されている。

	男女計				
	非労働力人口	うち就業希望者	うち適当な仕事がありそうにない	うち仕事があればすぐつける	うち過去1年間に求職活動あり
平成16年	4340	529	190	73	48
17	4348	489	172	67	44
18	4353	479	153	60	40
19	4360	465	154	57	36
20	4388	454	149	55	35
21	4422	471	163	67	42
22	4446	467	165	69	46
* 23	4487	467	164	68	45
* 23	4571	471	165	68	45
* 24	4534	417	142	58	37
平成16年	52	-1	-16	-15	-9
17	8	-40	-18	-6	-4
18	5	-10	-19	-7	-4
19	7	-14	1	-3	-4
20	28	-11	-5	-2	-1
21	34	17	14	12	7
22	24	-4	2	2	4
* 23	41	0	-1	-1	-1
* 24	23	-54	-23	-10	-8

(出典)総務省統計局  
「労働力調査(平成24年詳細集計速報)表18」

## 6. まとめ

- ・雇用・失業統計の国際基準の変更に伴って、我が国の統計も適切な対応が必要。
- ・新たな調査事項・設問の設計に当たっては、国際基準に沿った正確な回答が得られるよう、試験調査等を行い、予め十分に確認することが必要。
- ・時系列の不連続が生じる可能性が高く、特に失業者数など注目度の高い系列については、各方面の利用者の意見も予め十分に把握することが必要。(当分の間、季節調整済系列の算出が困難となる可能性が高い。)

## おわり

本資料の作成に当たり、高岡信行氏、植松洋史氏  
(いずれも総務省統計局)から有用な情報とコメント  
をいただきました。ここに記して謝意を表します。



# JIPデータベースにおける全要素生産性の計測と公的統計の課題

「経済統計・政府統計の数理的基礎と応用」コンファレンス  
2014年1月31日、東京大学  
における報告用資料

深尾京司  
(一橋大学経済研究所)

1

## 問題意識

- 人口減少、高齢化、長期経済停滞の継続等が続く日本が豊かな生活を維持する上で、生産性上昇の重要性は、ますます高まりつつある。
- しかし、多くの先進国政府が生産性統計を整備する中で、日本政府の生産性統計作成は出遅れている。
- この報告では、まずJIPデータベースを使って、生産性統計から何が分かるかについて例示する。
- 次に、政府の生産性統計作成のために、最低限何を推計する必要があるかを示す。
- また、生産性計測の精度を上げるために何が必要かについて考えてみる。

2

# 報告の構成

1. 問題意識
2. 生産性計測で何が分かるか：JIPデータベースの経験
3. 政府による生産性計測：日本および海外の動向
4. 生産性計測の視点からみた政府統計の課題
5. 終わりに

3

## 2. 生産性計測で何が分かるか： JIPデータベースの経験

- 日本産業生産性データベース(Japan Industrial Productivity Database、以下ではJIPと略記)は、マクロ経済を108部門に分け、各部門別に全要素生産性(TFP)を推計するために必要な、資本サービス投入指数と資本コスト、質を考慮した労働投入指数と労働コスト、名目および実質の生産・中間投入、TFPの上昇率を計算した成長会計の結果、などの年次データから構成されている。
- 毎年更新し、公開。最新のJIP 2013は1970年から2010年をカバー。
- JIPは、KLEMSタイプのデータベース(労働の質を考慮。資本サービスを推計。中間投入を含め、投入はディビジア指数として作成)。また国民経済計算にほぼ準拠。
- JIPデータベースは元々内閣府経済社会総合研究所のプロジェクトとして推計を開始し、現在は経済産業研究所と一橋大学が改訂・更新を行っている。
- 慶應義塾大学のKEOデータベースは40数セクターで1955-2004年をカバーしているが公開されていない。マクロ経済については、アジア生産性機構のAPO Productivity DatabaseやConference BoardのTotal Economy Database、OECDがTFPを推計・公表。日本生産性本部でも従業者5人以上の事業所について鉱工業生産指数及び第三次産業活動指数を用いて月次の産業別労働生産性指数を作成している。

4

## 2. 生産性計測で何が分かるか： JIPデータベースの経験

### 実質賃金・労働生産性・労働分配率

- 一国全体の労働生産性は、労働時間当たりどれだけ実質国内総生産(GDP)が生産されるかで計測される。労働時間当たり実質GDP(2011年には名目で1時間4,300円)のうち、時間当たり実質労働コスト(1時間約2,500円)が労働に分配される。したがって単純化して言えば、**実質賃金の上昇率が労働生産性の上昇率を上回ると、労働分配率(上記の例では58%)が上昇していくことになる。**(なお、2010年の米国の労働時間当たり名目GDPは56.4ドル、時間当たり実質労働コストは32.9ドル)
- 労働分配率の上昇が続けば、資本収益率が下落し設備投資が減退するから、そのような賃金上昇は持続できない。これが賃金上昇を考えるにあたって労働生産性の動向に注目する主な理由。

5

### 実質賃金・労働生産性・労働分配率(続)

	1970-80	1980-90	1990-2000	2000-11
実質賃金率(時間当たり労働コスト／消費者物価)の上昇	59.2%	23.8%	16.8%	0.4%
労働生産性上昇	51.3%	45.4%	20.8%	16.4%

資料：JIPデータベース2012、消費者物価統計他。

- 70年代や90年代には、労働生産性上昇にほぼ見合うだけの実質賃金率引き上げが達成された。
- 80年代や2000年以降は労働生産性上昇に比べて実質賃金率の引き上げが格段に小さい。
- 特に2000年以降、労働生産性は16%向上したのに、**実質賃金率はほとんど上がらないという特異な現象が起きている。**

6

## なぜ労働生産性上昇が実質賃金上昇に結実しなかったか

- 労働分配率を、要素費用で表示した名目GDP(「雇用者報酬」「営業余剰・混合所得」「固定資本減耗」の合計)に占める総労働コスト(自営業主の労働所得分を含む)と定義すると、近似的には以下のような形に分解できる。

労働分配率

= 総労働コスト / 要素費用表示の名目GDP

≡ (時間当たり労働コスト / 消費者物価) × (消費者物価 / GDPデフレーター) ÷ (実質GDP / 総労働時間)

- ここでGDPデフレーターは名目GDPを実質GDPで割った値であり、日本で生産された財・サービスの価格変化を表す。この式の両辺にGDPデフレーター / 消費者物価と労働生産性を掛けると次式を得る。

実質賃金率 ÷ 労働生産性 × (GDPデフレーター / 消費者物価) × 労働分配率

7

## なぜ労働生産性上昇が実質賃金上昇に結実しなかったか(続)

	1970-80	1980-90	1990-2000	2000-11
実質賃金率(時間当たり労働コスト / 消費者物価)の上昇	59.2%	23.8%	16.8%	0.4%
労働生産性上昇	51.3%	45.4%	20.8%	16.4%
GDPデフレーター / 消費者物価の上昇	-12.5%	-4.4%	-5.7%	-11.5%
労働分配率(GDPベース)の変化	19.4%	-9.5%	3.1%	-2.9%

- GDPデフレーターが輸出分を含む国内生産された財・サービス全体の価格を反映するのに対し、消費者物価は国内消費の対象のみをカバーし、輸入財・サービスの価格上昇も反映する。
- GDPデフレーター / 消費者物価の下落は、日本の交易条件(輸出する財・サービスと輸入する財・サービスの相対価格)の悪化をかなりの程度反映している(この他、投資財価格の下落、消費者物価の作り方にもおそらく影響を受けている)。

8

## 労働分配率データ

英国誌『エコノミスト』も引用したOECD統計は、自営業主・家族労働者の労働コストは同じ属性の被雇用者の労働コストと同一と仮定しているため、過去の労働分配率を過大に推計している可能性が高い。

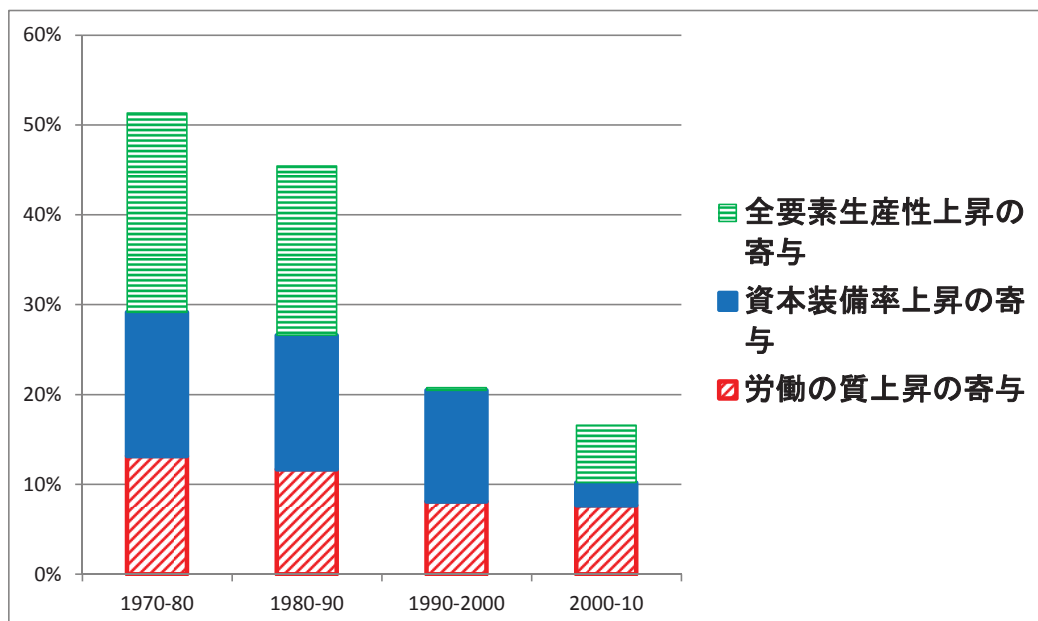


“Labour Pains, Workers’ Share of National Income,”  
*The Economist*, November 2, 2013.

9

## なぜ労働生産性上昇が減速したか

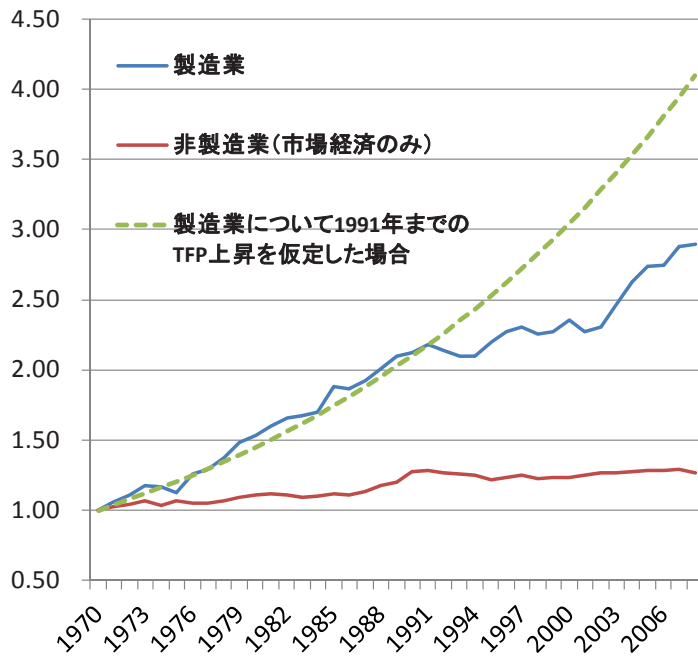
- 失われた20年において、全要素生産性上昇が停滞した。
- 2000年以降は、資本の低収益を反映して資本装備率上昇が停滞したことも大きい。
- 非正規雇用の拡大は、労働の質上昇を阻害している。



# どの産業で生産性上昇が停滞したか

- 90年代以降製造業で全要素生産性(TFP)上昇が減速。
- 非製造業では昔からTFP上昇が停滞。非製造業のTFP上昇加速も急務。
- 米国と異なり日本ではIT技術導入による非製造業のTFP上昇加速が起きなかった。

製造業と非製造業の全要素生産性水準の推移、1970-2008年(1970年=1)

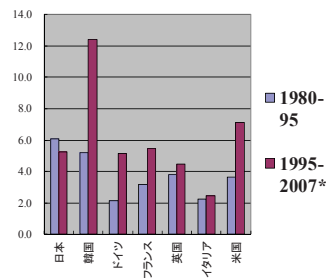


注: TFPは付加価値ベースの値。非製造業(市場経済のみ)は、住宅・分類不明を除いた値。  
資料: JIPデータベース2011

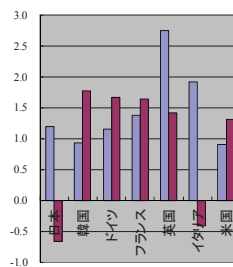
図 市場経済におけるTFP上昇:産業別・国別:1980-95年および1995-2007年\*

日本では、米国と異なり、電気機器以外の製造業や商業・運輸業のような、IT投入産業でTFP上昇が停滞。

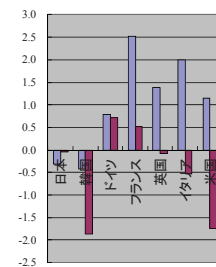
電気機器・郵便・通信



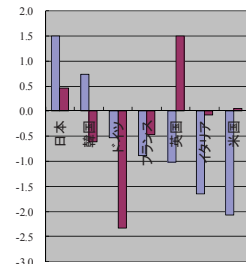
電気機器以外の製造業



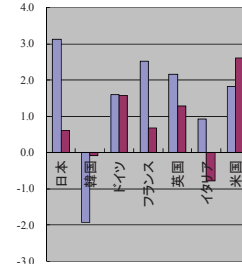
一次産業・公益・建設



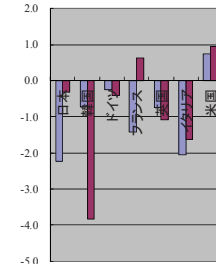
金融・対事業所サービス



商業・運輸業



対個人・社会サービス



出所: EU KLEMSデータベース2009年11月版  
日本は①995-2006年の値

# 報告の構成

1. 問題意識
2. 生産性計測で何が分かるか：JIPデータベースの経験
3. 政府による生産性計測：日本および海外の動向
4. 生産性計測の視点からみた政府統計の課題
5. 終わりに

13

## 3. 政府による生産性計測：日本および海外の動向

### 日本の現状

日本政府は加工統計として、労働生産性や全要素生産性統計を作成していない。

統計委員会基本計画部会が総務省の次期基本計画(2014-18年度)に関する諮問を受けて答申用に作成した『公的統計の整備に関する基本的な計画』修正案(2014年1月)では、「別表 1 経済関連統計の整備 (1) 国民経済計算の整備 エ一次統計等との連携強化」において、内閣府が「労働生産性及び全要素生産性指標を把握するための基礎情報の整備について検討を行う。」こととされている。

また、「本文 4 基本計画の推進 1 施策の効果的かつ効率的な実施」において、統計委員会が重点的に実施する事項として、「統計法施行状況審議や個別の諮問審議において把握した府省横断的な統計上の課題(欠測値補完、歪みの強い分布に関する推計の改善、サービスの質の計測に関する動向等)に関する研究や日本学術会議及び関連学会との連携強化方策について検討し、取組の推進を図る」と記述される見込み。

14

### 3. 政府による生産性計測：日本および海外の動向（続）

国際連合：

国連で合意したThe System of National Accounts, 2008では、労働生産性および全要素生産性に言及し、望ましい計測方法について記述。

米国政府：

労働統計局(BLS)が労働生産性の計測を続けてきた。BLSはまた日本を含む先進主要国の製造業の労働生産性(時間あたり実質付加価値)、時間あたり労働コスト、単位労働コスト(実質付加価値あたり労働コスト)を推計するプロジェクトInternational Labor Statistics Programを実施してきたが、オバマ政権の財政問題のため、民間の研究機関Conference Boardに引き継がれることになった。

国民経済計算を担当している経済分析局(BEA)は、National Economic Accountsの一部として、詳細な産業別労働、資本サービス投入と全要素生産性データを含むIntegrated GDP-Productivity Accountsを推計、公表するようになった([http://www.bea.gov/national/integrated\\_prod.htm](http://www.bea.gov/national/integrated_prod.htm))。長期遡及も実施している。

15

### 3. 政府による生産性計測：日本および海外の動向（続）

その他の先進諸国：

米国と同様に、オランダ、カナダ、スイス、ニュージーランド等の政府でも、国民経済計算と統合的な形でマクロおよび産業別の全要素生産性を計測する試みが進められ、統計局が全要素生産性を推計し一部統計の公表を始めている。

オランダ：<http://unstats.un.org/unsd/EconStatKB/KnowledgebaseArticle10109.aspx>

カナダ：<http://www.statcan.gc.ca/nea-cen/about-apropos/prod-eng.htm>

OECD：

OECDは加盟国政府等の協力の下、多くの加盟国のマクロ経済および産業別の労働、全要素生産性を公表している。

<http://stats.oecd.org/Index.aspx?DataSetCode=MFP>

産業別データは、韓国を含む約20カ国について掲載されているが、日本はマクロ経済のデータのみしか公表されていない。なお、OECDではSTANデータベースなど一部統計にJIPデータベースを利用。

16



### 3. 政府による生産性計測：日本および海外の動向（続）

政府・国際機関以外の国際連携活動

EU KLEMSプロジェクトおよびWorld Input-Output Databaseプロジェクト

欧州委員会の資金により、フローニンゲン大学等が中心になって。詳細な産業別の全要素生産性や国際産業連関表データを作成。その成果の多くはOECDにも継承された。日本の生産性関連データについては、JIPデータベースが提供した。

World KLEMSプロジェクト

KLEMSタイプの産業別全要素生産性データベース作成のための国際連携組織。ハーバード大学のDale Jorgenson教授がリーダー。2014年5月19-20日に東京で全体集会が開かれる予定（経済産業研究所主催、一橋大学、学習院大学等共催）。日本はJIPデータベース、中国は中国産業生産性（CIP）データベース（RIETI・一橋大学）が参加。

Asia KLEMSプロジェクト

アジアにおける生産性データ作成の連携組織。ADB、RIETI、韓国中央銀行、韓国生産性センター等が支援。宮川努学習院大学が現議長。

17

## 報告の構成

1. 問題意識
2. 生産性計測で何が分かるか：JIPデータベースの経験
3. 政府による生産性計測：日本および海外の動向
4. 生産性計測の視点からみた政府統計の課題
5. 終わりに

18

## 4. 生産性計測の視点からみた政府統計の課題

先行国並みのマクロ経済および産業別労働・全要素生産性計測のために最低限何が必要か

現在の国民経済計算では、就業者数・雇用者数、雇用者の労働時間数、雇用者所得を推計。

←なお、一上 響・原 尚子「日本の労働生産性に関するリアルタイムデータ分析」日本銀行ワーキングペーパーシリーズ、No.10-J-7 2010年3月 は国民経済計算を元に作成した労働生産性が、確報化、確各報化、基準改定等により、過去平均して上方改定される問題を指摘している。

- 労働生産性の計測には、自営業主・家族従業者の労働時間を推計する必要がある。
- 全要素生産性の計測や、OECD等に準拠した労働分配率の計測には、自営業主・家族従業者の労働所得の推計が必要。
- KLEMSタイプの全要素生産性を計測するためには、労働については更に、属性別（性、従業上の地位、最終学歴、年齢、産業）別の就業者数、労働時間、単位時間労働コストの情報が必要。
- 内閣府による時価資本ストックと資本コストの計測が進んでおり、資本サービスデータについては、順調に整備が行われているように思われる。

19

## 4. 生産性計測の視点からみた政府統計の課題（続）

より正確な労働・全要素生産性計測のための課題

### ①デフレーター・実質生産額データの改善

これは、生産性計測だけの問題ではない。経済成長率や物価上昇率の正確な計測においても重要。

- どの国でも実質生産額の変動は、対象とする経済活動により、(i)ダブルデフレーション、(ii)生産指数による推計、(iii)生産要素投入量の変化、等で計測されている。日本は(i)を重視して計測しており、比較的優等生だが、物価上昇率を正確に計測することの重要性はその分高い。また推計方法の改善、正確な国際比較等のためには、コモディティー・フロー・データの透明性改善や国際共同研究が重要。産業連関表接続表等と国民経済計算の整合性向上も課題。次期基本計画では、当局間の情報共有が盛り込まれる予定。
- CPIの連鎖型指数への移行

Jessie Handbury, Tsutomu Watanabe, David E. Weinstein, "How Much Do Official Price Indexes Tell Us About Inflation," NBER Working Paper, No. 19504, October 2013.

20

## 4. 生産性計測の視点からみた政府統計の課題(続)

より正確な労働・全要素生産性計測のための課題(続)

### ①デフレーター・実質生産額データの改善(続)

- 新製品導入による質の変化をどう計測するか:オーバーラップ方式の問題点、コスト把握の難しさ

日本銀行調査統計局「企業物価指数「電子部品・デバイス」の指数精度向上への取り組み－企業物価指数 2010 年基準改定に向けた先行事例－」*BOJ Reports and Research Papers*, 2010年7月。

- サービスの質の計測
- 非市場経済アウトプットの質の計測

藤澤美恵子「産出量法を用いた医療の国民経済計算推計に関する考察」一橋大学。

藤澤美恵子「国民経済計算における教育のアウトプット計測についての考察」国民経済計算、2012年12月。

- イノベーションによる経済厚生改善の計測

Robert J. Gordon, “Issues in Price Measurement in U.S. Economic Growth During the Early 20th Century” PPT for *Maddison Memorial Conference*, November 6-7, 2010, Amsterdam.

## 4. 生産性計測の視点からみた政府統計の課題(続)

より正確な労働・全要素生産性計測のための課題(続)

### ②絶対価格水準の国際比較

- 最終需要面から把握しようとするICPプロジェクトに加え中間財価格を把握する必要(ICOPプロジェクト等)
- アウトプットの質の国際比較(例えば小売業の場合、営業時間、消費者からの距離、等を見ても良いのか)

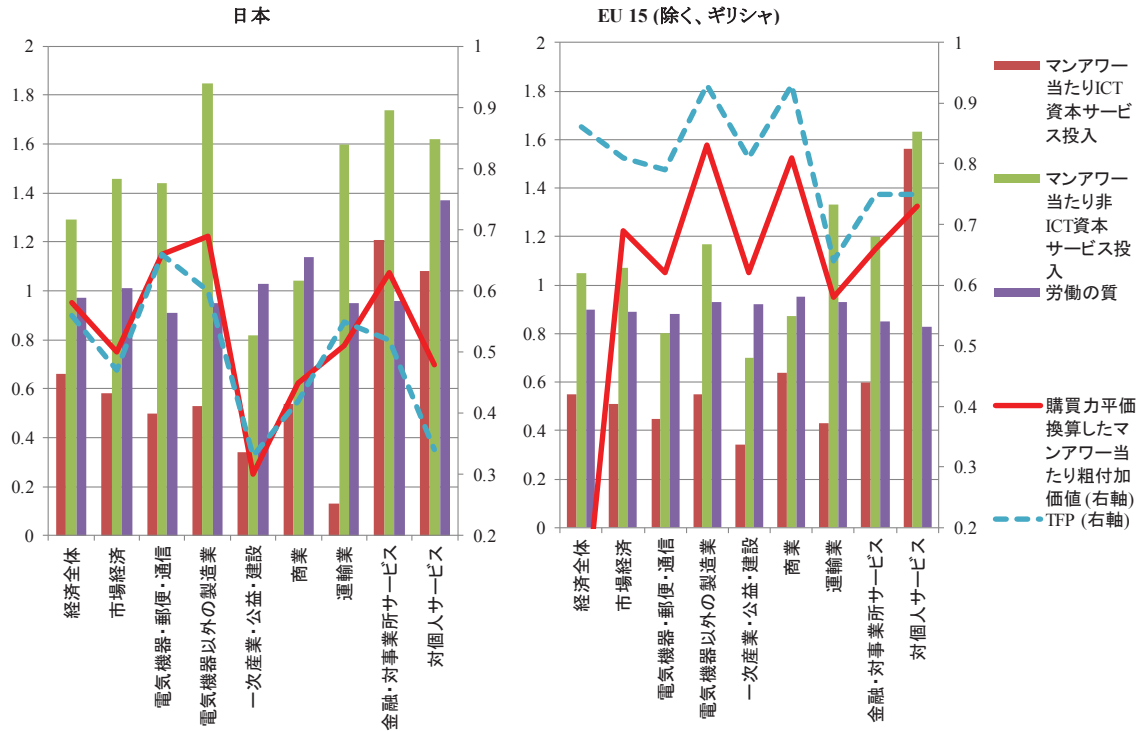
### ③中間投入の正確な把握

- Offshoring Bias問題
- 企業・事業所間の中間投入価格、生産物価格の違い

### ④長期遡及の必要性

- ⑤無形資産の計測: R&Dの資本化、企業内訓練の計測?

図12. 購買力平価換算した労働生産性・要素投入・TFP水準: 日本・EU・米国比較(2005, 米国=1)

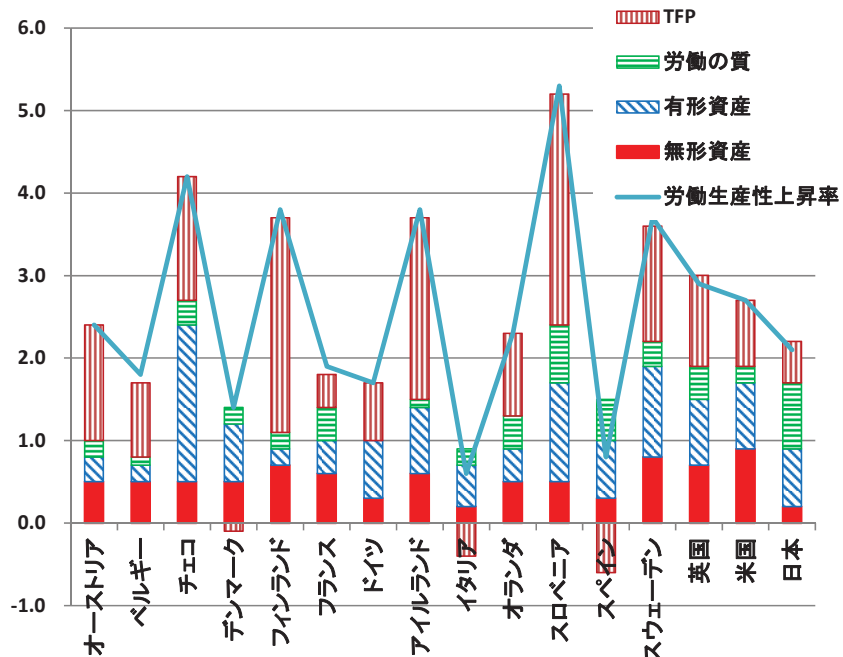


資料: Inklaar and Timmer (2008).

労働生産性上昇の源泉: 無形資産を考慮した成長会計分析: 1995-2007年 (年率, %)

日本は、サービス業における経済的競争力獲得のための無形資産投資(組織変革、Off-JT、広告宣伝等)が特に少ない。

中小企業を中心にICT投資やR&D、Off-JT等を支援することがおそらく重要。



資料: 日本は宮川・比佐 (2012)、海外諸国はCorrado et al. (2012)。

## 5. 終わりに

- 生産性の計測は、日本経済の長期停滞を理解する上でも重要。
- 公式の生産性統計が無い日本は、米国、オランダ、カナダ等に出遅れている。
- 今日では、KLEMSタイプの生産性統計が主流。国民経済計算と統合的な生産性統計を目指す国が多い。
- 日本政府が生産性統計を作成するためには、労働投入に関する推計が鍵。
- 正確な生産性計測のためには、デフレーター、中間投入、無形資産等に関する更なる知見の蓄積、国際協調が重要。

# 人口減少・高齢社会の進展と 人口統計の課題

国立社会保障・人口問題研究所  
 金子隆一

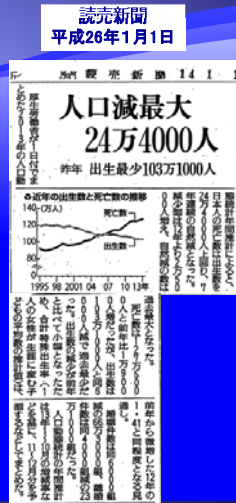
## I. 人口動向の概観

- ◆ 人口減少社会の到来
- ◆ 少子高齢化の進展

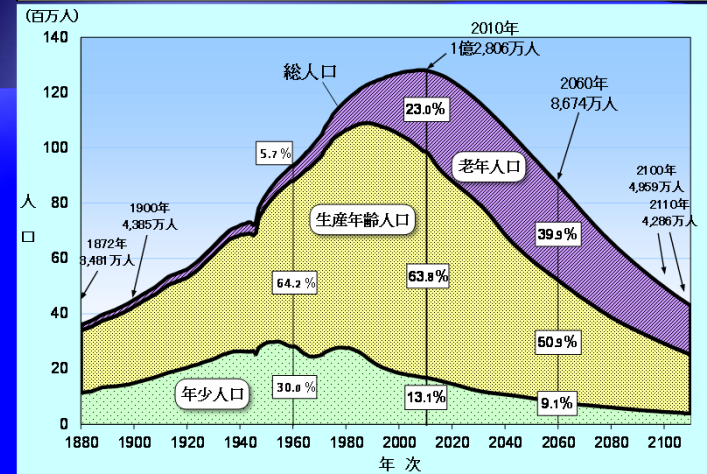
## II. 人口統計の課題と挑戦

- ◆ 人口現象の把握
- ◆ 人口統計の問題点

### 本格化する人口減少・・・

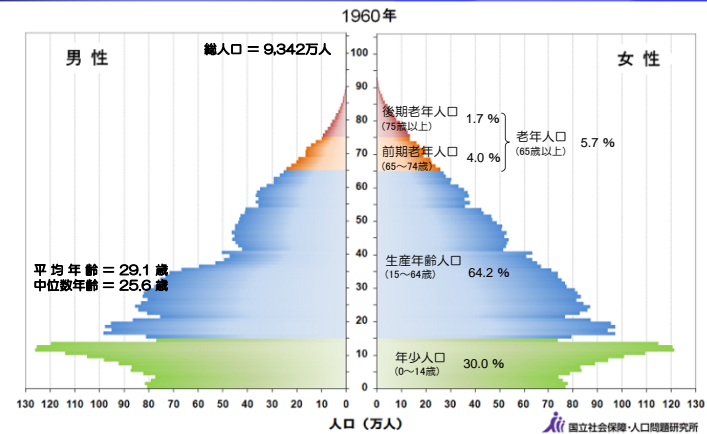


### 日本の人口推移(明治期～21世紀～2110年)



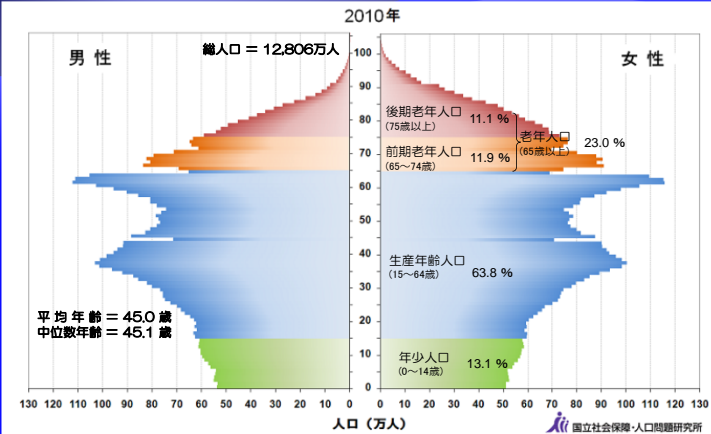
資料: 旧内閣統計局推計、総務省統計局「国勢調査」推計人口、国立社会保障・人口問題研究所「日本の将来推計人口」(平成24年1月推計(出生・死亡中位推計))

## 日本の人口ピラミッド 1960年



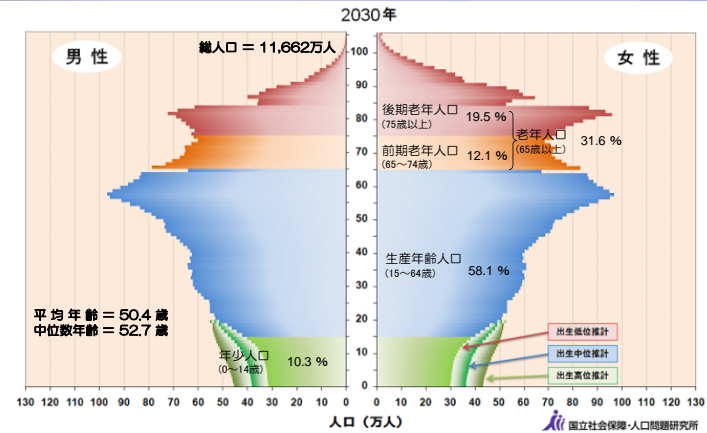
5

## 日本の人口ピラミッド 2010年



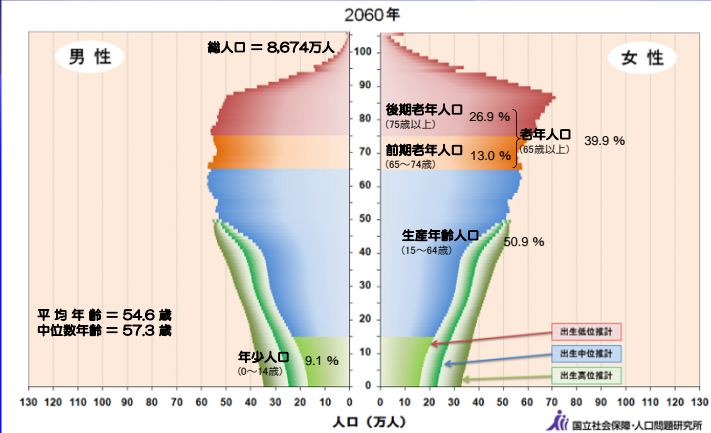
6

## 日本の人口ピラミッド 2030年



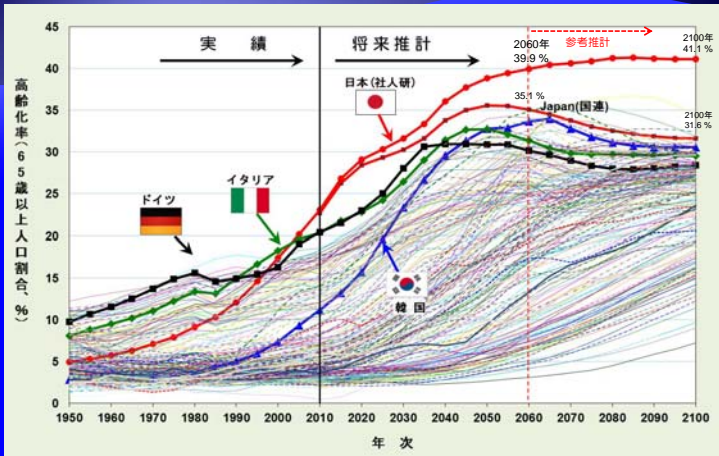
7

## 日本の人口ピラミッド 2060年



8

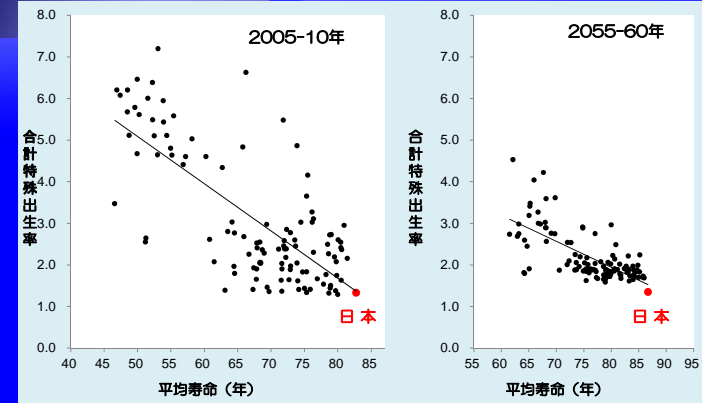
### 世界各国の高齢化率の推移(1950~2100年)



資料: United Nations, 2011, World Population Prospects: The 2010 Revision, 日本は総務省統計局「国勢調査」、国立社会保障・人口問題研究所「日本の将来推計人口(平成24年1月推計[出生中位・死亡中位推計])」

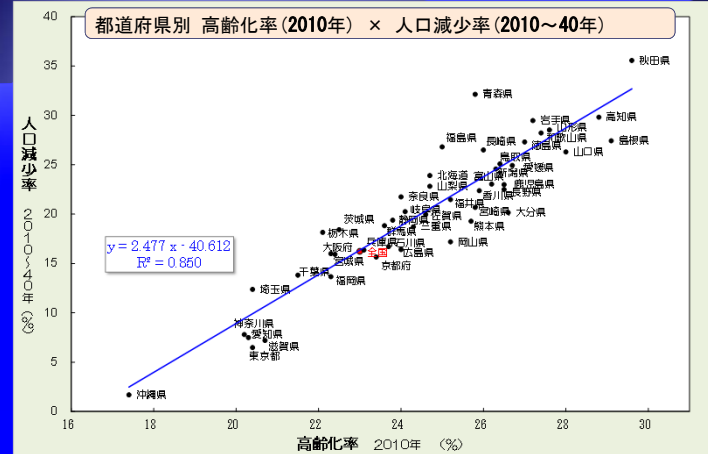
### 人口動態事情の国際比較 - わが国の特殊性

合計特殊出生率と平均寿命 (2005年人口1千万人以上の国)



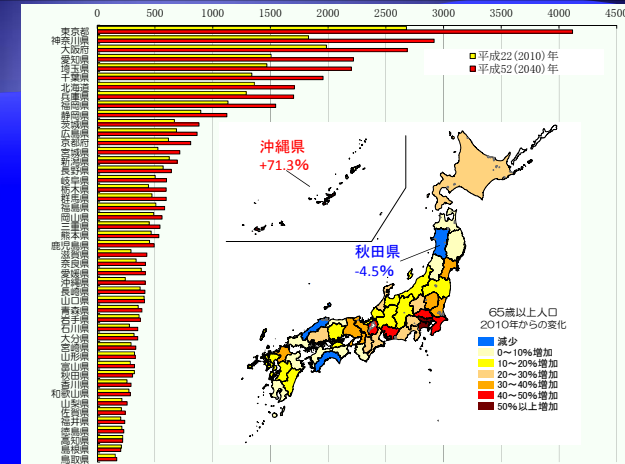
資料: United Nations (2011), World Population Prospects: The 2010 Revision, 総務省統計局「国勢調査」、国立社会保障・人口問題研究所「日本の将来推計人口(平成24年1月推計)」[出生中位・死亡中位推計]

### 高齢化が進んでいる地域では、今後の人口の減少が著しい



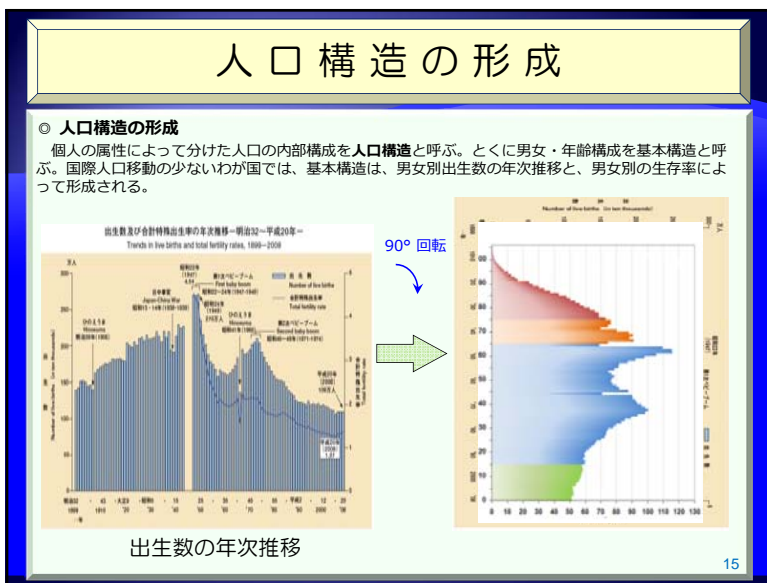
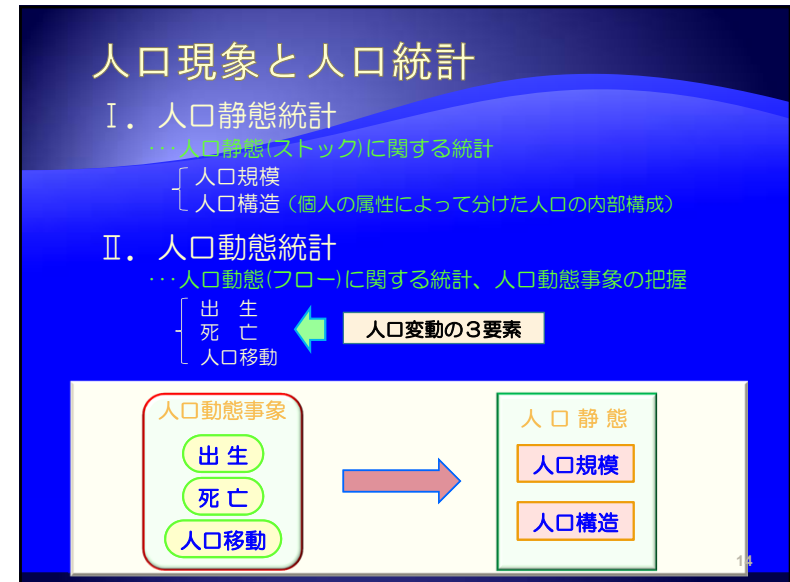
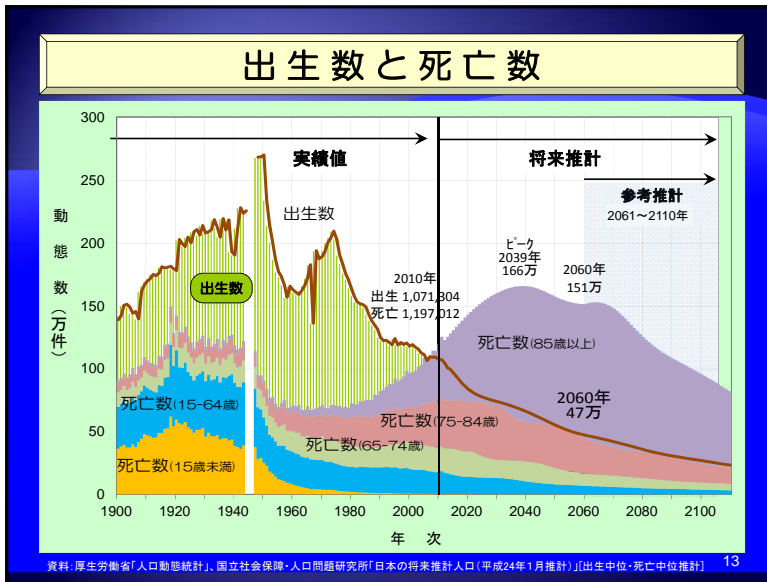
資料: 総務省統計局「国勢調査」、国立社会保障・人口問題研究所「日本の世帯数の将来推計(全国推計)平成25年1月推計」

### 都道府県別、65歳以上人口の変化: 2010年、2040年比較



資料: 総務省統計局「国勢調査」、国立社会保障・人口問題研究所「日本の世帯数の将来推計(全国推計)平成25年1月推計」





### 人口統計の体系

調査対象・形態による整理

調査形態	調査対象	人口静態			人口動態	
		日本人	在留外国人	海外邦人	自然動態 (出生・死亡)	社会動態 (人口移動)
全数調査 (全容把握)	実地調査	国勢調査			出生児数—国勢調査 (1950, 1960, 1970年のみ)	国勢調査 (前住地/現住地)
	行政記録	住民基本台帳・戸籍	在留外国人統計 (旧登録外国人)	海外在留邦人調査	人口動態調査 住民基本台帳による人口動態 母体保護統計	住民基本台帳 人口移動報告 出入国管理統計
標本調査 (詳細分析)	実地調査	国民生活基礎調査 住宅・土地統計調査等 (国勢調査 試験調査・事後調査)			21世紀縦断調査 出生動向基本調査等	人口移動調査
	行政記録	-			人口動態統計 特殊報告等	-

# 人口統計の課題 (その1)

## 人口静態統計

### 国勢調査

- 全数に対する唯一の 実地調査
- 現住地人口 → 実態の把握(実測)
- 実施間隔 5年 → 途中年は“推計”が必要
- 調査環境の変動 (個人情報意識の高まり) → 不詳増 (不在・オートロック等の増加)
- 財政負担 中

### 住民基本台帳

(戸籍)  
日本人登録

- 行政記録 → 継続性・安定性 (調査環境の影響小)
- 追加の財政負担 小 ← ICT 化有効
- 常住人口(登録人口) → 実態? (移動・登録タイムラグ) (消えた高齢者問題)
- 集計: 粗い = 男女・年齢5歳階級(1994年~)
- 対象: 日本人のみ } 平成26(2014)年より 変更!
- 期日: 3月末

# 人口統計の課題 動態率の分子・分母の不整合問題

## 合計特殊出生率における 不整合



合計特殊出生率 =  $\sum (\text{出生数 } x / \text{女性人口 } x) \dots$  年齢  $x$  について合計  
 都道府県別  
 分子 = 日本人出生数  
 分母 = (国勢調査年次) 日本人女性人口  $\dots$  実査人口  
 (非国勢調査年次) 総人口女性人口 (外国人を含む)  $\dots$  推計人口

資料 1

「住民基本台帳に基づく人口、人口動態及び世帯数」の調査項目、調査期日の変更等について

平成25年7月5日  
総務省自治行政局住民制度課

**現在** 住民基本台帳に基づく人口、人口動態及び世帯数  
 ○調査目的: 地方公共団体等において、住民基本台帳事務の参考とするため  
 ○調査項目: 下表参照 ○調査期日: 3月31日現在

**経緯** 「公的統計の整備に関する基本的な計画」(平成21年3月13日閣議決定)(関係要旨)  
 住民基本台帳に基づく人口等の調査に関して、地方公共団体の意見も聞きつつ、以下の検討を行う。  
 ①集計の充実(1才毎人口、世帯主との続柄別人口等)②作成時期(現行3月末)の見直し

**対応** 統計専門機関等から意見を聴取し、市区町村へ照会した結果を踏まえ、以下のとおり対応することとした。  
 ※ 市区町村には、平成22年度から逐次変更準備を周知し、現在、H25. 3. 31現在調査データを集計中

**1 調査項目** 市区町村の意見も踏まえ、主に外国人住民に関する項目を拡充(H25. 3. 31現在調査から)

※1 外国人住民(以下、「外国人」という。)については、平成24年7月から住民基本台帳法の適用対象となった。

※2 1才毎人口、世帯主との続柄別人口等の項目については、市区町村におけるニーズ、職員の事務負担、システム改修の面から市区町村の賛同が得られなかったため、追加しないこととした。

**2 調査期日** 市区町村の意見も踏まえ、現行(3月31日現在)から1月1日現在へ変更(H26. 1. 1現在調査から)

※ 移行時の調査結果公表に当たっては、基本項目については、H25. 1. 1現在数値との比較ができるよう工夫する。

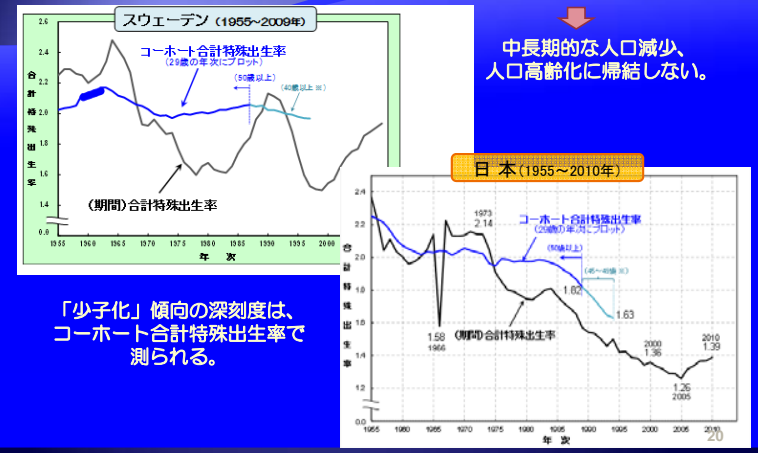
	変 更 前	変 更 後
調査期日	毎年3月31日現在 ※業務終了時現在 日本人の人口(男女別、市区町村別、5歳毎年齢別)	毎年1月1日現在 ※午前0時現在 同左
調査項目	— 日本人の人口動態(出生者数、死亡者数、転入者数、転出者数、左記に加え、国内・国外別の転入者・転出者数、帰化等別のその他記載数・消除数を追加 ※期間: 1/1~12/31) — 日本人の世帯数(日本人と外国人との世帯を含む)	外国人の人口(同上) 外国人の人口動態(同上) 外国人の世帯数、日本人と外国人との世帯数を各々集計 外国人のみの世帯数

(注)変更後の調査期日、調査項目となるのは、H26. 1. 1現在調査以降から(H25. 3. 31現在調査は、変更後の調査項目で実施。)

18

# 問題の所在 少子化の指標としての合計特殊出生率の問題

ピリオド合計特殊出生率の変動は、「少子化」の悪化/改善を意味しない



## 人口統計の課題 動態率の分子・分母の不整合問題

「人口動態統計」公表人口動態率に見られる不整合

$$\text{人口動態率} = (\text{動態事象数}) / (\text{リスク人口})$$

### ◆ 観察対象期間に関する 不整合

分子 = 年間 動態事象数 (日本人)

分母 = 国勢調査/推計人口による 10月1日 日本人人口

※ 人口減少下においては、分母が過小になる (率=過大)

e.g. 今後、婚姻率・出生率は過小、死亡率は過大の傾向

生命表はOK!  
(分母の平均人口を  
独自に推計)

### ◆ 地域別 人口動態率 における 不整合

分子 = 住民登録地別 動態事象数 (日本人)

分母 = 国勢調査による 現住地別 日本人人口

※ 転出届未提出者の多い地域は分子が過大になる

e.g. 出稼ぎの多い地域で、死亡率が過大に? (青森県)

21

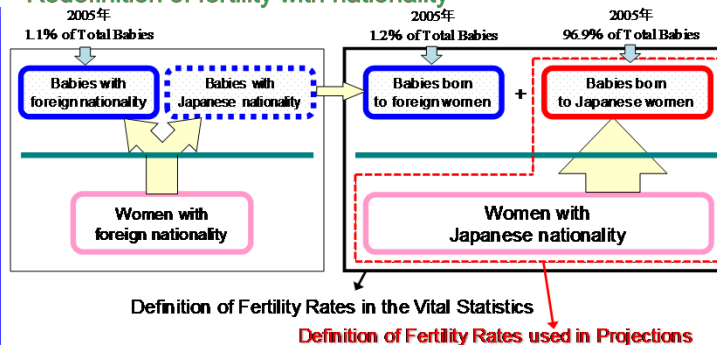
## 人口統計の課題 動態率の分子・分母の不整合問題

「人口動態統計」公表 出生率に見られる 不整合

$$\text{人口動態率} = (\text{動態事象数}) / (\text{リスク人口})$$

分子に他人口由来  
の事象が混在

### Redefinition of fertility with nationality



22

## 人口統計の課題 (その1)

( 人口統計の特殊性 )

○ なぜ、今「人口減少」「少子高齢化」が注目される?

- ・ 人口高齢化は、昭和30年代以降の将来推計人口で指摘
- ・ 人口減少は、1980年代以降の将来推計人口で指摘

→ 経済社会活動と人口変動のタイムスケールの違い

- ・ 人口変化は、ライフコース変容を通して進行(例:長寿化、少子化)
- ・ 変化に気がつかない & 潮流を押しとどめることは難しい

不可視性

不可逆性

( 人口変化は、「時計の短針」・・・動いて見えないが、着実にたゆみない )

人口現象の時間スケール

ライフコース (ライフサイクル) 的視点の重要性

23

## ライフコース分析

### 事象歴分析法

### Event history analysis

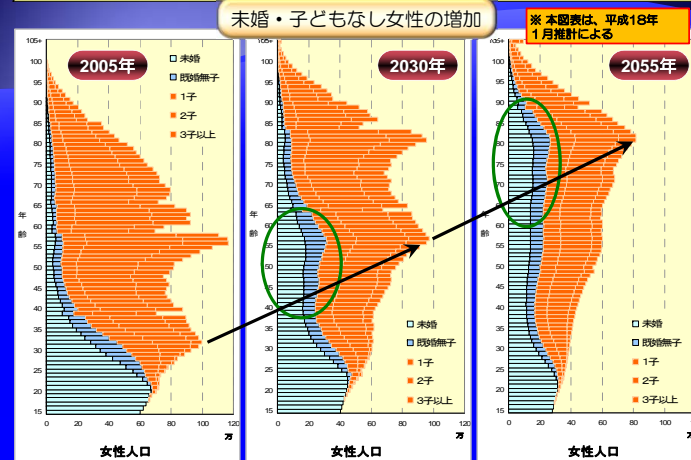
24

### 「将来推計人口」の描くライフコース

女性の出生力要素指標	実績値 1960年 生まれ	将来推計人口の出生仮定 1995年生まれ女性コホート		
		中位仮定	高位仮定	低位仮定
(1) 平均初婚年齢	25.7歳	28.2歳	27.9歳	28.5歳
(2) 生涯未婚率	9.4%	20.1%	14.7%	26.2%
(3) 夫婦完結出生児数	2.07人	1.74人	1.91人	1.57人
(4) 離死別再婚効果係数	0.962	0.938	0.937	0.938
子ども数	0人 (子なし割合)	17.5%	26.6%	44.8%
	1人 (一人子割合)	13.9%	18.2%	17.1%
	2人以上	68.6%	46.2%	38.1%
コホート合計特殊出生率 (日本人女性の出生に限定した率)	1.81	1.35 (1.30)	1.60 (1.53)	1.12 (1.15)

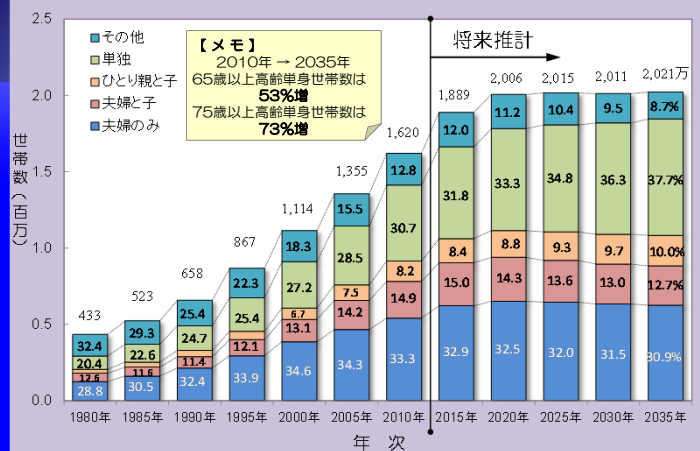
資料：国立社会保障・人口問題研究所「日本の将来推計人口（平成24年1月推計）」

### 「将来推計人口」の描くライフコース



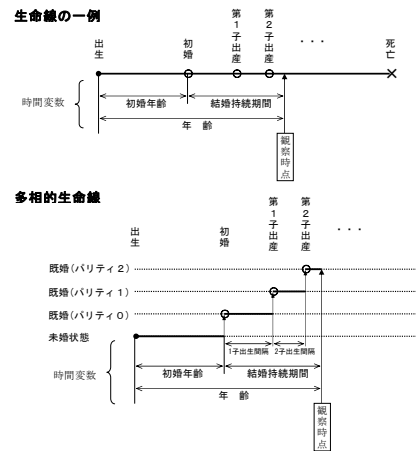
資料：国立社会保障・人口問題研究所「日本の将来推計人口（平成18年12月推計）」に基づき、講演者が作成（公表値ではない）。 26

### 高齢者世帯数および家族類型別構成の推移（1960～2030年）



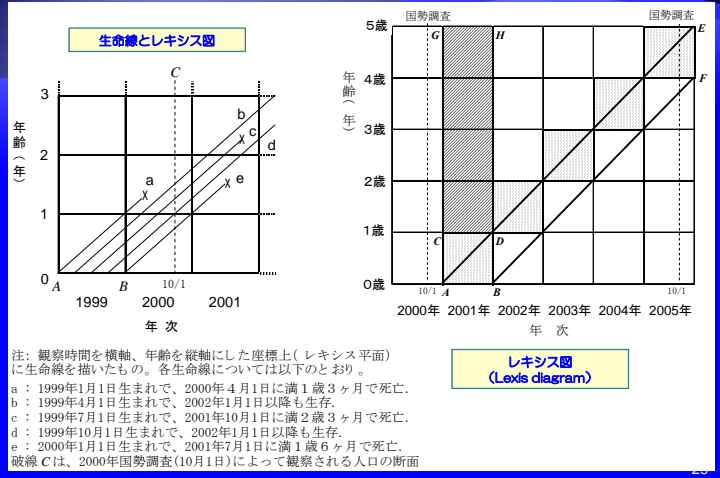
資料：総務省統計局「国勢調査」、国立社会保障・人口問題研究所「日本の世帯数の将来推計（全国推計）平成25年1月推計」 27

### ライフコース分析の枠組み

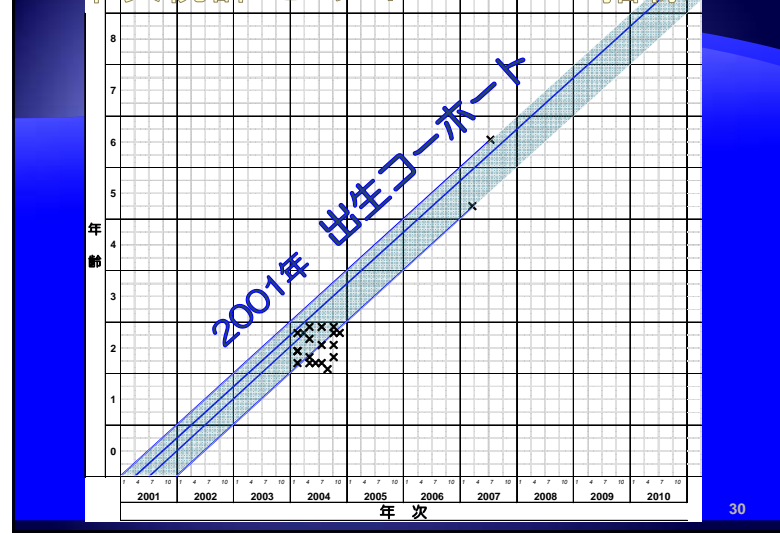


28

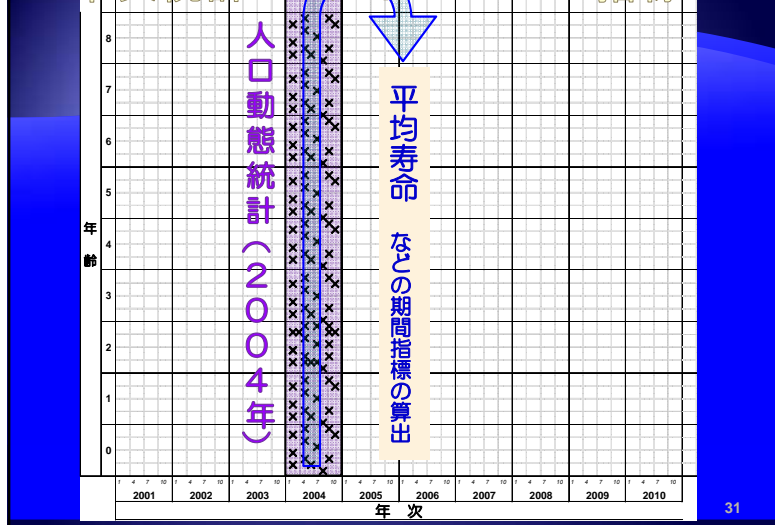
# 年次統計とライフコース指標



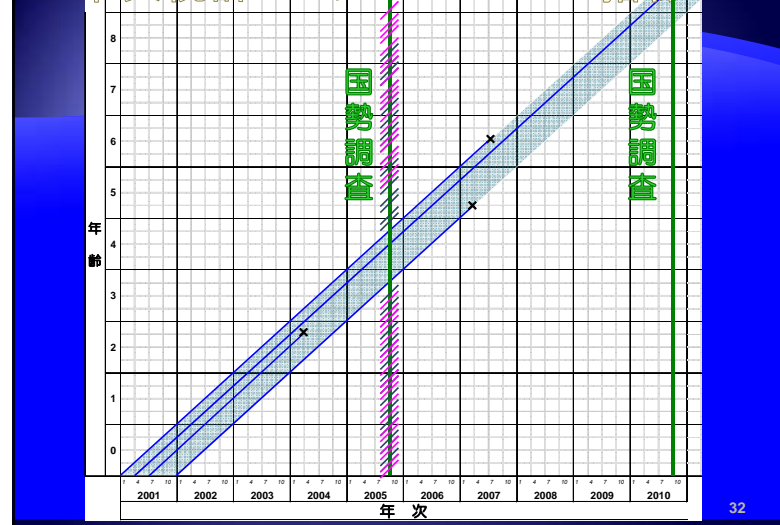
# 年次統計とライフコース指標



# 年次統計とライフコース指標



# 年次統計とライフコース指標



## Woman's Life-course Composition by Family Status : Recoded and Projected

Derived from Multistate Life Tables

Family Status	Birth year of woman's cohort								
	Recorded		Projected						
	1950	1955	1960	1965	1970	1975	1980	1985	1990
Life time distribution of woman by number of child ( at birth )									
Childless	18.4	17.7	20.8	24.8	31.4	35.5	36.4	37.9	38.1
Never married	13.6	11.2	12.9	14.4	17.9	21.7	23.6	24.3	24.3
Ever married	4.8	6.5	7.9	10.5	13.5	13.8	12.8	13.6	13.8
Only child	11.2	11.2	13.3	16.5	18.7	18.8	18.8	18.1	18.0
Two children	46.8	44.4	41.8	39.5	35.2	32.9	33.0	32.6	32.8
Three children	19.4	22.0	19.6	15.3	11.5	10.1	9.4	9.3	9.3
Four and more children	4.2	4.7	4.5	3.8	3.2	2.7	2.3	2.0	1.9
Net Reproduction Rate	87.5	90.0	84.5	76.3	66.3	61.2	59.6	58.1	57.9
No grandchild	22.2	21.2	25.6	31.6	41.2	46.8	48.1	50.0	50.2

Derived from Assumption

Life time proportion of woman ( without mortality effect = directly derived from fertility assumption )									
Never married	5.0	5.8	9.3	12.0	16.2	20.4	22.6	23.5	23.5
Childless	10.3	12.7	17.5	22.7	30.0	32.8	35.7	37.1	37.4
No grandchild	12.1	15.0	21.3	28.8	39.3	42.9	46.8	48.9	49.4

33

## 縦断調査・データ

Longitudinal survey / data

34

## 統計調査の連携

人口動態統計

ライフイベントに関する  
統計の新時代!

全数調査であり、代表性が高いが、調査項目に制約。

横断調査

代表性もあり、調査項目も豊富だが、因果立証に弱い。

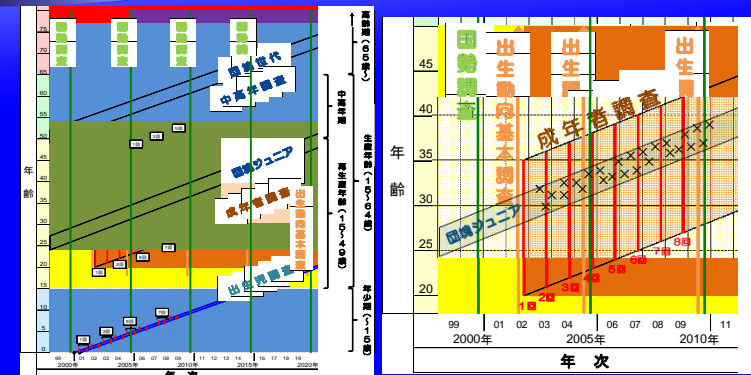
縦断調査

調査項目も豊富で、因果立証に強いが、代表性が弱い。

35

## 年次統計とライフコース指標

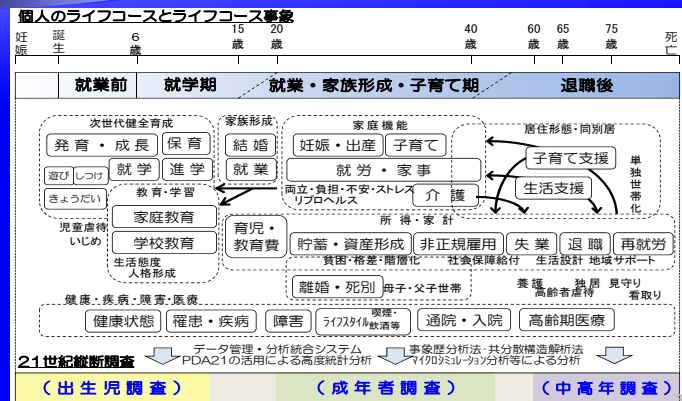
21世紀縦断調査、出生動向基本調査の例



36

# 統計調査の連携

国民のライフコースに寄り添い、有効な政策立案に必要な科学的知見をもたらす三つの21世紀縦断調査の例



21世紀縦断調査 (出生児調査) (成年者調査) (中高年調査)

# 分析用データベース

Database for scientific analyses

# 人口統計の国際的データベース

- ◆ 出生データベース
- ◆ マックス・プランク人口研究所
- ◆ 出生順位別 年齢別 出生率
- ◆ 期間観察
- ◆ コーホート観察

- ◆ 死亡データベース
- ◆ カリフォルニア大学バークレー校
- ◆ マックス・プランク人口研究所
- ◆ 男女 年齢別 死亡率 <生命表>
- ◆ 期間観察
- ◆ コーホート観察

# 人口統計のデータベース整備

# 人口統計のデータベース整備 JMDの紹介

# 人口統計のデータベース整備 JMDの紹介

# 高度な加工統計指標 の必要性

例：高齡化の再定義

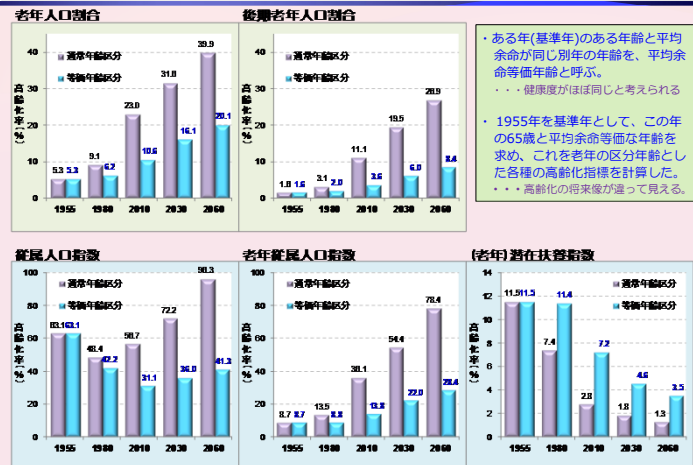
## 1955年基準 平均余命等価年齢

男性					女性				
年次	平均余命				年次	平均余命等価年齢 (1955年基準)			
	60歳	65歳	75歳	90歳		60歳	65歳	75歳	90歳
1955	15.0	11.8	7.0	2.9	1955	60.0	65.0	75.0	90.0
1980	18.3	14.6	8.3	3.2	1980	64.4	69.0	77.9	91.6
2010	22.7	18.7	11.4	4.2	2010	70.0	74.4	82.8	95.1
2030	24.7	20.6	13.1	5.1	2030	72.3	76.9	85.4	98.1
2060	26.5	22.3	14.6	5.9	2060	74.5	79.0	87.6	100.3

資料：2010年以前は「完全生命表」、2030年、2060年は「将来推計人口（平成24年1月推計・死亡中位仮定）」を用いて算出。



## 1955年基準 平均余命等価年齢による 高齢化指標



資料：2010年以前は「完全生命表」、2030年、2060年は「将来推計人口(平成24年1月推計・死亡中位仮定)」を用いて算出。

# データ集積システム

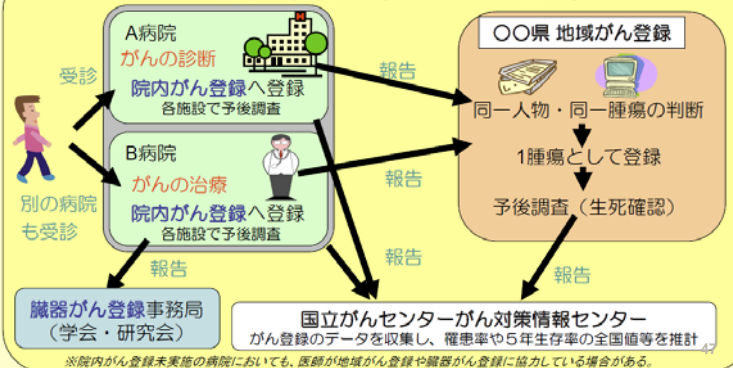
## Scientific data collection systems

## がん登録とは

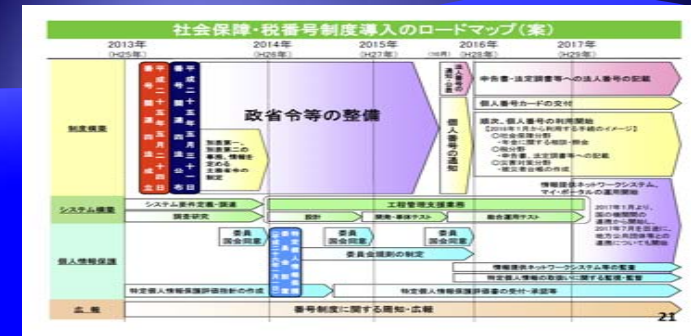
がん登録は、

がんの罹患や転帰その他の状況を登録・把握し、分析する仕組みであり、**がん罹患数・罹患率、がん生存率、治療効果の把握**など、**がん対策の基礎となるデータの把握のために必要なものである。**

＝がん登録の仕組み(3つのがん登録)＝

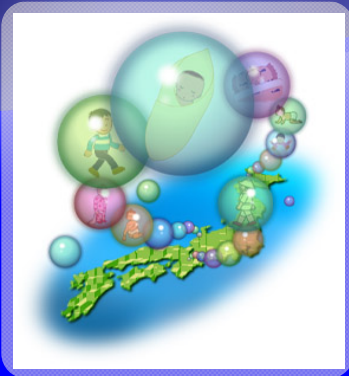


## 社会保障・税番号の導入と人口統計



統計委員会「公的統計の整備に関する基本的な計画の変更について」(平成25年10月30日閣議)

- 平成 25 年 5 月に法律が公布されたことを受け、平成 28(2016)年からの施行に向けた準備が行われているところであり、施行後 3 年を目途とした見直し条項により、最速での見直しは平成 31(2019)年頃となることと想定される。
- 利用範囲の拡大に関する番号法の見直しに併せて、統計作成における活用について検討する。



経済統計・政府統計の数理的基礎と応用  
**人口減少・高齢社会の進展と  
 人口統計の課題**

2014年 1月31日(金)  
 東京大学経済学部小島ホール2F  
 国立社会保障・人口問題研究所  
 金子隆一

### 意思決定構造の高齢化 (1955~2060年)

年次	総人口中の 有権者割合	有権者 平均年齢	有権者人口(20歳以上日本人)の年齢構成				
			有権者 「青年」率 (35歳未満)	有権者 「壮年」率 (35~64歳)	有権者 高齢化率 (65歳以上)	後期高齢率 (75歳以上)	
実績	1960年	59.7 %	41.5 歳	42.9 %	47.5 %	9.6 %	2.9 %
	1990年	73.1	47.2	27.0	56.5	16.5	6.6
	2010年	81.0	52.7	20.9	50.8	28.3	13.6
将来推計	2030年	83.4	57.5	16.8	45.6	37.6	23.3
	2050年	83.5	60.4	14.3	39.9	45.7	29.1
	2060年	83.6	61.5	14.2	39.1	46.7	31.7

(注) 有権者割合：総人口に占める20歳以上日本人人口の割合、有権者平均年齢：有権者の平均年齢。  
 有権者「青年」率：有権者総数に占める20~34歳の有権者数の割合。  
 有権者「壮年」率：有権者総数に占める35~64歳の有権者数の割合。  
 有権者高齢化率：有権者総数に占める65歳以上の有権者数の割合。  
 有権者後期高齢率：有権者総数に占める75歳以上の有権者数の割合。  
 資料：国立社会保障・人口問題研究所「日本の将来推計人口」(平成24年1月推計[出生中位・死亡中位推計])

## 人口変動のもたらす

## 経済社会の課題

## ライフコース

## 人口減少社会の処方箋

人口減少、高齢化にいかに対処するのか？

## 長寿化の真価

## (現在)人口推計の概要 (その1)

総務省統計局ウェブサイト  
<http://www.stat.go.jp/data/jinsui/1.htm>

### 1. 目的

「人口推計」は、国勢調査の実施間の時点における毎月、毎年的人口の状況を把握するために行うもの。

### 2. 内容

毎月1日現在 年齢(5歳階級), 男女別推計人口  
 毎年10月1日現在 全国, 年齢(各歳), 男女別推計人口、および  
 都道府県, 年齢(5歳階級), 男女別推計人口、その他に、補間補正人口

### 3. 方法

- 1) 国勢調査による人口を基礎に、その後の人口動向を他の人口関連資料から得て、毎月1日現在の人口を算出。
- 2) 推計人口 = 基準人口
  - + 自然動態(出生児数-死亡者数)
  - + 社会動態(入国者数-出国者数)
  - + 国籍の異動による純増(日本人について)
  - (+ 都道府県間転入者数-都道府県間転出者数)

※( )は、都道府県別人口について適用

53

## (現在)人口推計の概要 (その2)

総務省統計局ウェブサイト  
<http://www.stat.go.jp/data/jinsui/1.htm>

### 3) 算出に用いている資料

- ・ 出生児数及び死亡者数 …… 「人口動態統計」 (厚生労働省)
- ・ 出入国者数 …… 「出入国管理統計」 (法務省)
- ・ 国籍異動 …… 法務省資料
- ・ 都道府県間転入者数 …… 「住民基本台帳人口移動報告」 (総務省統計局)
- ・ その他 …… 「国勢調査」 (総務省統計局)
- …… 「在留外国人統計」 (法務省)
- …… 都道府県資料

全国人口の推移 (国勢調査結果による補間補正人口)

年次	各年 10月1日現在 総人口	人口増減 (前年10月～当年9月)							
		総数	自然動態			社会動態			補間 補正数
			出生児数	死亡者数	自然増減	入国者数	出国者数	社会増減	
平成12年	126,926	...	...	...	...	...	...	...	
13	127,316	390	1,185	966	219	19,266	19,120	146	25
14	127,486	170	1,176	981	195	16,321	16,372	-51	25
15	127,694	208	1,138	1,023	115	15,038	14,970	68	25
16	127,787	93	1,126	1,024	103	17,673	17,709	-35	25
17	127,768	-19	1,087	1,078	9	18,951	19,004	-53	25

注) 千人未満は四捨五入しているため、合計の数値と内訳の計は必ずしも一致しない。

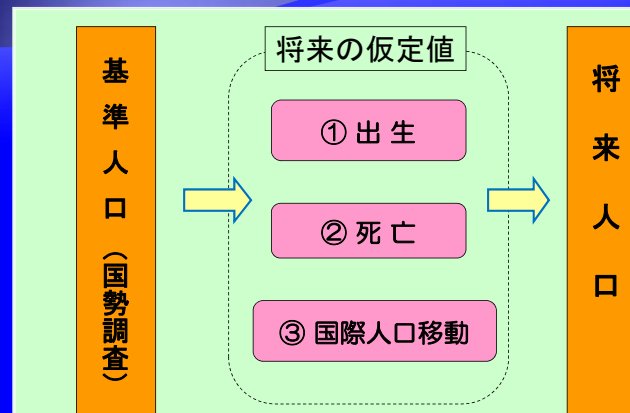
54

## 将来人口推計

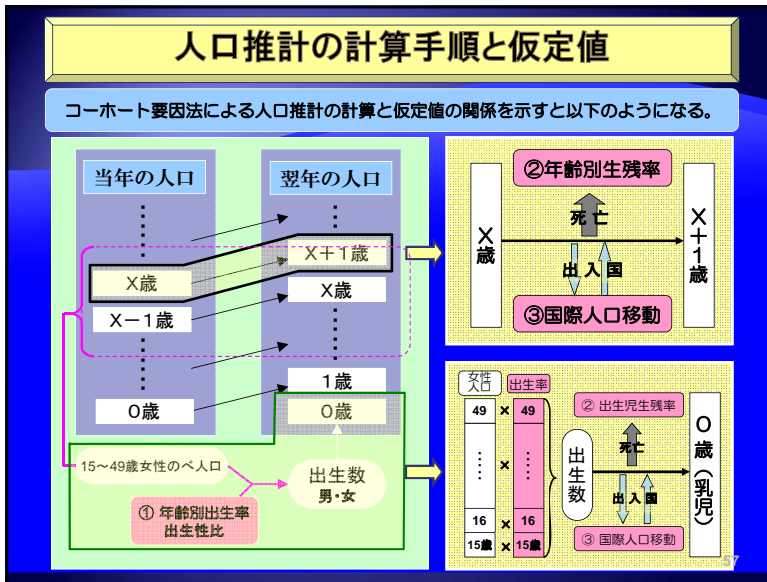
— 手法の概略 —

55

## 将来人口推計(要因法)

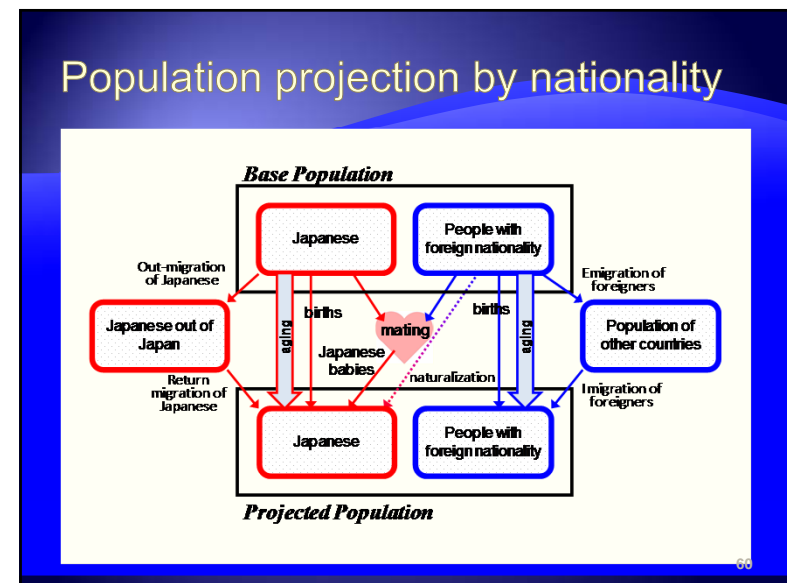


56

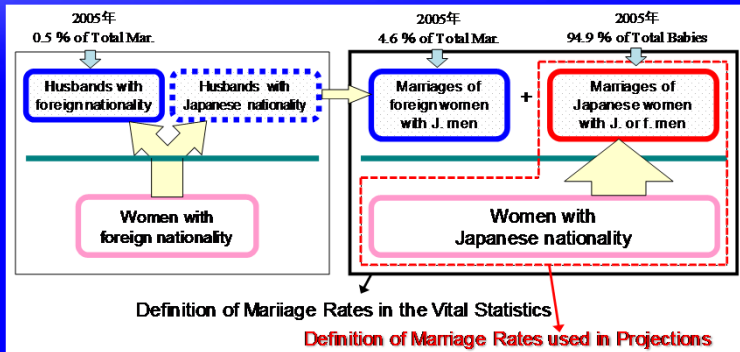


- ### 将来人口推計に必要なデータ
- 基準人口・・・男女年齢別人口（国勢調査）
  - 3つの仮定値
    - ① 出生  
（将来の）女性の年齢別出生率  
（将来の）出生性比
    - ② 死亡  
（将来の）男女年齢別生残率（将来生命表）
    - ③ 国際人口移動  
（将来の）男女年齢別国際純移動数（率）

- ### Basic strategies of the population projections
- ◆ **Objectivity and Neutrality**  
= To be *scientific* = Aim to use the best methods + best data
  - ◆ **Incorporation of life course changes**  
= To cope with "*the life course revolution*", i.e. the lowest-ever-fertility, the world highest longevity, and the globalizing movements of people.
  - ◆ **Internationalization of the population**  
= structured population model by nationality (Japanese or not)

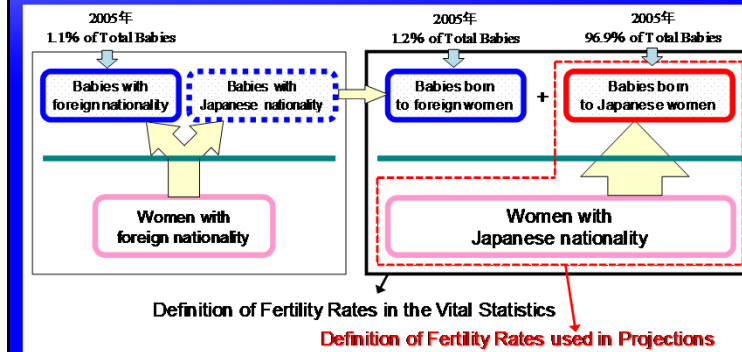


## Redefinition of nuptiality with nationality



61

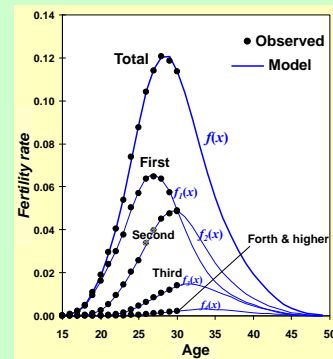
## Redefinition of fertility with nationality



62

## Age schedule model for fertility

A mathematical model is developed for birth-order, age-specific fertility rate of female cohort. The model has several parameters which represent behavioral traits of cohorts.



### Model of the age-specific fertility rate - The generalized log-gamma model -

The fertility rate ( $f$ ) of birth order ( $n$ ), age ( $x$ );

$$f_n(x) = C_n \cdot \gamma_n(x; u_n, b_n, \lambda_n)$$

$C_n$  = lifetime probability of birth,  $\gamma_n(x; u_n, b_n, \lambda_n)$  denotes birth,

$$\frac{|\lambda_n|}{b_n \Gamma(\lambda_n)} \left( \frac{1}{\lambda_n} \right)^{x-u_n} \exp \left[ \frac{1}{\lambda_n} \left( \frac{x-u_n}{b_n} \right) - \frac{1}{\lambda_n} \exp \left\{ \lambda_n \left( \frac{x-u_n}{b_n} \right) \right\} \right]$$

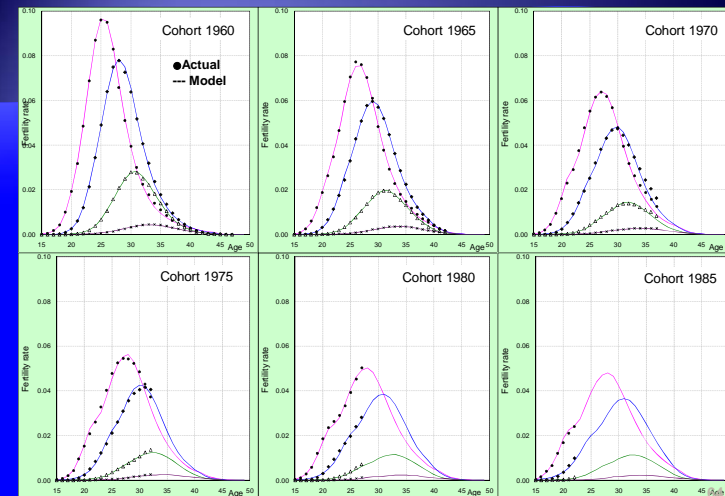
where  $\Gamma$ ,  $\exp$  the gamma and exponential function,  $C_n$ ,  $u_n$ ,  $b_n$ , and  $\lambda_n$  are parameters for  $n$ -th birth.

The cohort fertility schedule is given with the standard error pattern  $\varepsilon_n$  as

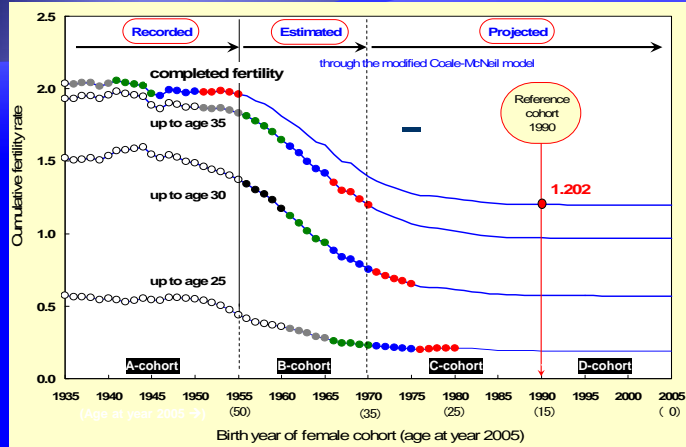
$$f(x) = \sum_{n=1}^4 C_n \cdot \left\{ \gamma_n(x; u_n, b_n, \lambda_n) + \varepsilon_n \left( \frac{x-u_n}{b_n} \right) \right\}$$

63

Figure 10 Actual and Modeled Cohort Fertility Rates by Birth Order



## Trends of Cohort Completed Fertility by Age : Recoded Estimated and Projected



65

## Model for the cohort completed fertility

$$CTFR = \sum_{n=1}^{k_n} C_n$$

$C_n = (1-\gamma) \cdot CEB_n^*(\theta_0) \cdot \kappa_n \cdot \delta_n$

$C_n$  : the eventual probability having the  $n$ -th birth

$\gamma$  : the eventual proportion never married

$CEB_n^*(\theta_0)$  : the completed number of children for a woman married at age  $\theta_0$

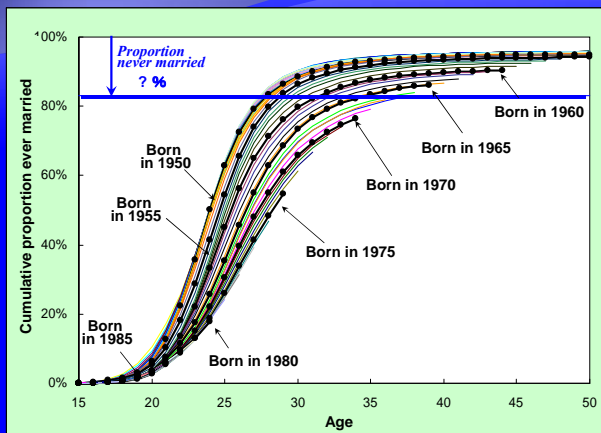
$\kappa_n$  : the coefficient of deviation from the standard marital fertility

$\delta_n$  : the coefficient of effects from divorce, bereavement and remarriage

$k_n$  : the highest birth order

66

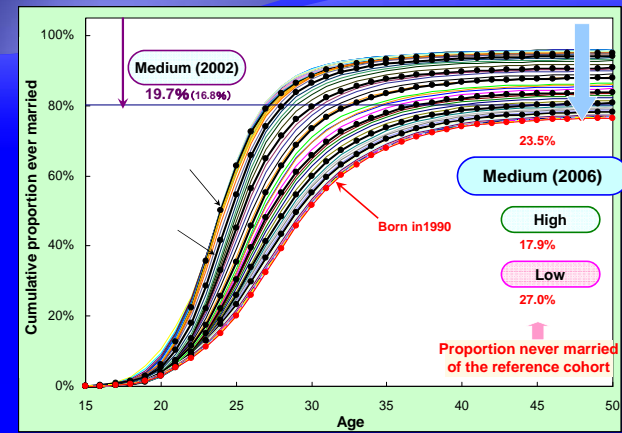
## The cumulative first marriage rates for female cohorts



Source: The Vital Statistics, (corrected for delayed registration).

67

## The cumulative first marriage rates for female cohorts



Source: The Vital Statistics, (corrected for delayed registration). NIPSSR(2006), Population Projection for Japan: 2006-2065.

## Model for the cohort completed fertility

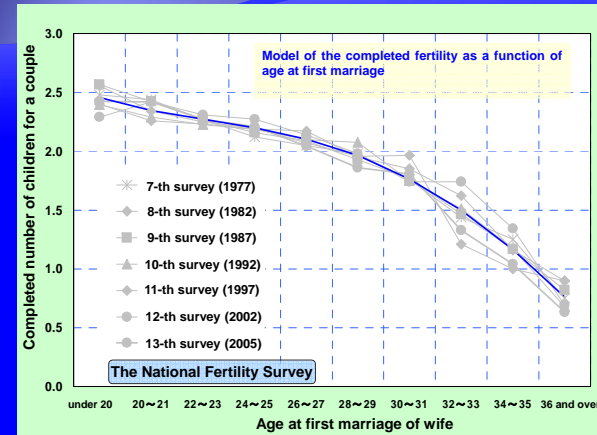
$$CTFR = \sum_{n=1}^{k_n} C_n$$

$C_n = (1-\gamma) \cdot CEB_n^*(\theta_0) \cdot \kappa_n \cdot \delta_n$

- $C_n$  : the eventual probability having the  $n$ -th birth
- $\gamma$  : the eventual proportion never married
- $CEB_n^*(\theta_0)$  : the completed number of children for a woman married at age  $\theta_0$
- $\kappa_n$  : the coefficient of deviation from the standard marital fertility
- $\delta_n$  : the coefficient of effects from divorce, bereavement and remarriage
- $k_n$  : the highest birth order

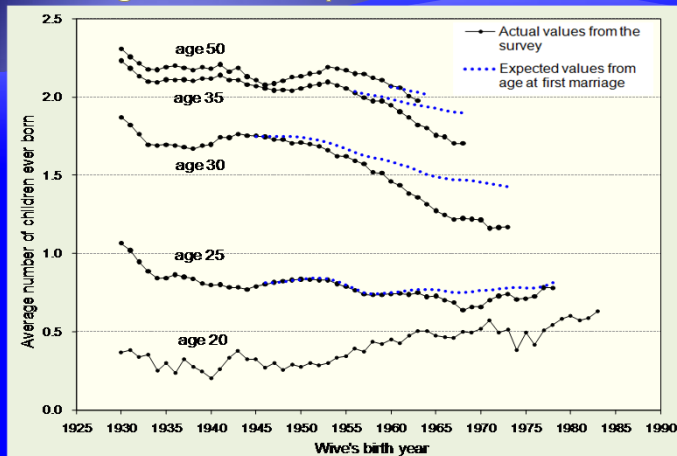
69

## The completed fertility depends on age at first marriage



70

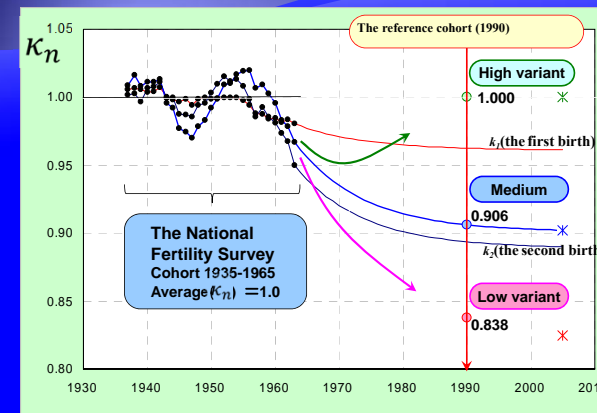
## Cohort Trends of CEB in Married Women by Age Actual & Expected from MAFM



Source: The National Fertility Survey (NIPPSR).

71

## The Coefficient of Deviation from the Standard Marital Fertility



72

## Model for the cohort completed fertility

$$CTFR = \sum_{n=1}^{k_n} C_n$$

$$C_n = (1-\gamma) \cdot CEB_n^*(\theta_0) \cdot \kappa_n \cdot \delta_n$$

Labels in diagram:  
 - Marriage prevalence:  $\gamma$   
 - Completed fertility for first married couples:  $CEB_n^*(\theta_0)$   
 - Effects of divorce etc. (residual):  $\delta_n$

- $C_n$  : the eventual probability having the  $n$ -th birth
- $\gamma$  : the eventual proportion never married
- $CEB_n^*(\theta_0)$  : the completed number of children for a woman married at age  $\theta_0$
- $\kappa_n$  : the coefficient of deviation from the standard marital fertility
- $\delta_n$  : the coefficient of effects from divorce, bereavement and remarriage
- $k_n$  : the highest birth order

73

## The Coefficient of Other Effects (divorce, bereavement and remarriage + residual)

Classification of marital status (The figures are those for cohort born in 1955)

Type of marital status (woman at age 50)		Women born in 1955	CEB	CEB ratio to first-married	
Single	Never married (n)	$\gamma$ (5.8%)	$C_n$ ( $\neq 0.00$ )	$R_n$ (0.00)	
	Divorced (d)	$P_d$ (11.8%)	$C_{dw}$ (1.58)	$R_{dw}$ (0.76)	
		Widowed (w)			$P_w$ (2.7%)
Married	Ever divorced / widowed	After divorced	$P_{rd}$ (5.5%)	$C_r$ (1.86)	$R_r$ (0.90)
		After widowed	$P_{rw}$ (0.2%)		
	Non-first married	Wife-remained	$P_{fr}$ (5.1%)	$C_{fr}$ (1.90)	$R_{fr}$ (0.92)
		Wife-first, husband-remarried (fr)			
	First married couple (ff)	$P_{ff}$ (68.9%)	$C_{ff}$ (2.07)	$R_{ff}$ (1.00)	

$\gamma$  : the proportion never married

$R_{dw} = C_{dw} / C_{ff}$  etc.

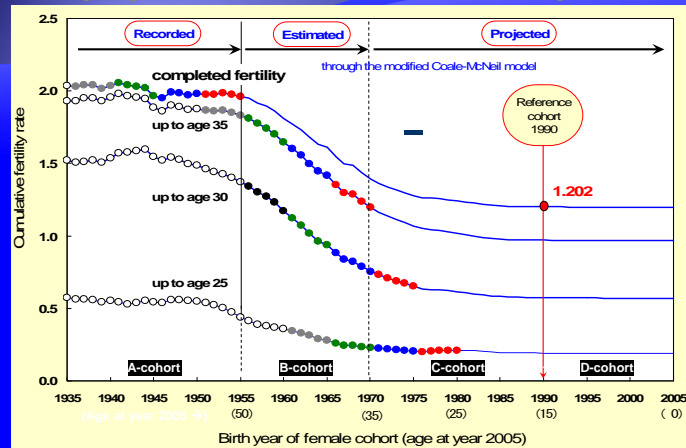
Record for cohort-1955  
 $\delta \approx 0.952$   
 ↓  
 Projected for cohort-1990  
**0.925**

Structure of  $\delta$

$$\delta = \frac{1}{1-\gamma} \{ P_{ff} + P_{fr} R_{fr} + P_r R_r + P_{dw} R_{dw} \}$$

74

## Trends of Cohort Completed Fertility by Age : Recoded Estimated and Projected



75

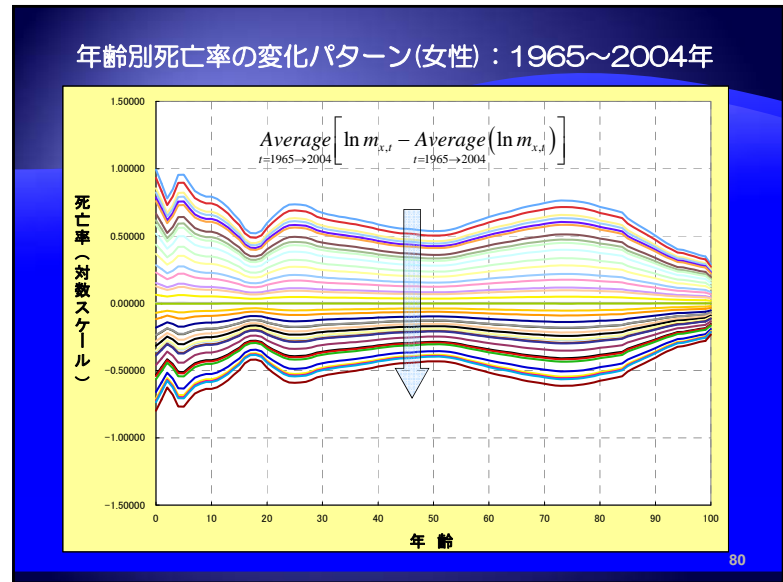
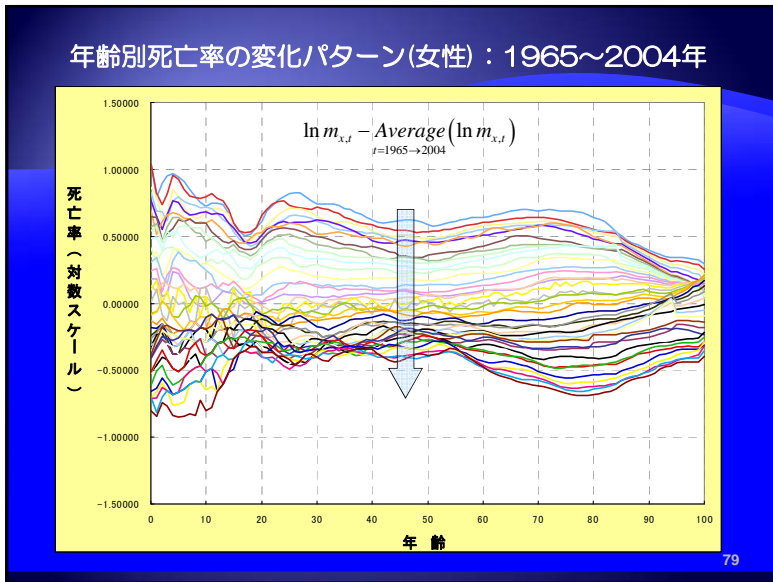
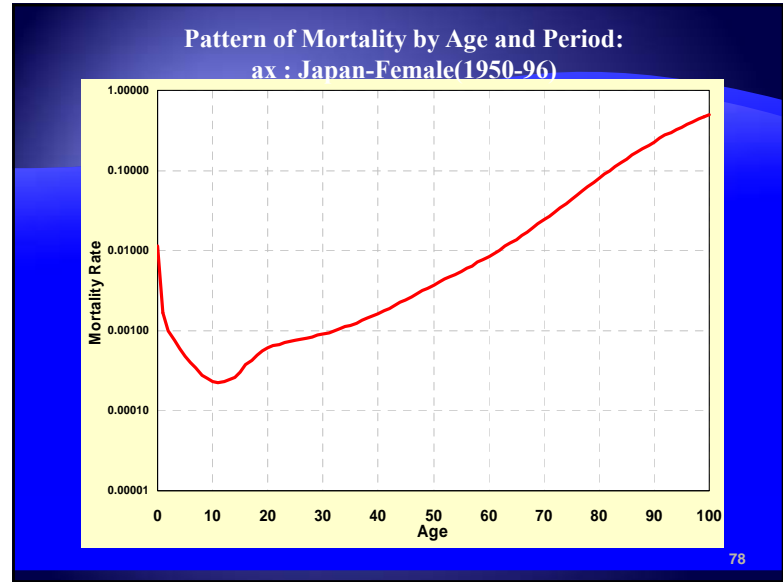
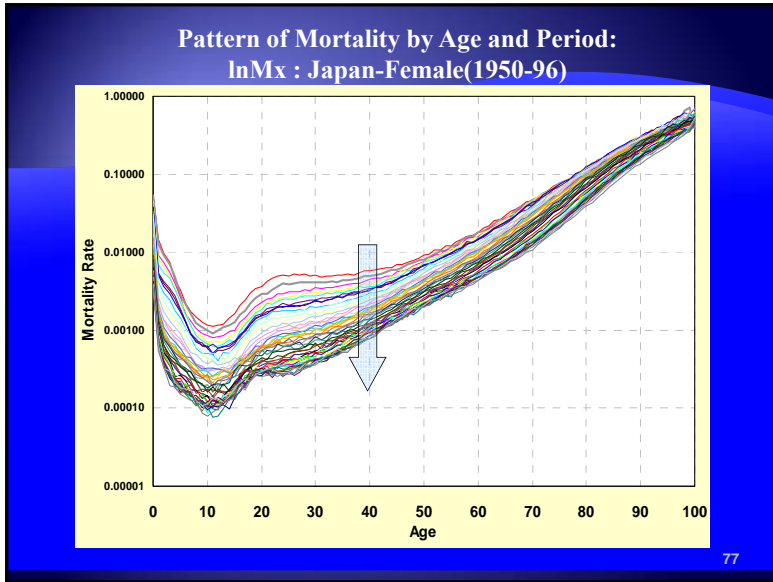
## 出生率要因の仮定値

女性の出生力要素指標	実績値 1955年 生まれ	将来推計人口の出生仮定 1990年生まれ女性コーホート			
		中位仮定	高位仮定	低位仮定	
(1) 平均初婚年齢	24.9	28.2	27.8	28.7	
(2) 生涯未婚率	5.8%	23.5%	17.9%	27.0%	
(3) 夫婦完結出生児数	2.16	1.70	1.91	1.52	
(4) 離死別再増効果係数	0.952	0.925	0.938	0.918	
子ども数	0人 (子なし割合)	12.7%	37.4%	28.6%	43.3%
	1人 (一人子割合)	11.8%	18.2%	15.4%	19.4%
	2人以上	75.6%	44.4%	55.9%	37.2%
コーホート合計特殊出生率 (日本人女性の出生に限定した率)	1.94	1.26 (1.20)	1.55 (1.47)	1.06 (1.02)	

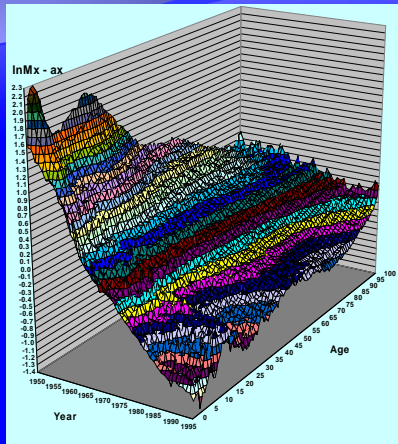
資料：国立社会保障・人口問題研究所「日本の将来推計人口(平成18年12月推計)」

76





### Residual Pattern of Mortality by Age and Period: lnM<sub>x</sub> - a<sub>x</sub> : Japan-Female(1950-96)



81

### 死亡スケジュール変化の表現のためのモデル

#### 生命表のリレーショナルモデルとリー・カーターモデル

リレーショナルモデルとは、死亡スケジュール変化の経験的パターンを用い、その変化を表現するパラメータと組み合わせて任意のスケジュールを表現する方法

平成14年、18年推計では、リー・カーターによって開発されたリレーショナルモデル(Lee-Carter Model)をベースにした方法により将来生命表を作成した。

$$\ln(m_{x,t}) = a_x + b_x \cdot k_t + \varepsilon_{x,t}$$

$\ln(m_{x,t})$  年次( $t$ )、年齢( $x$ )の死亡率の対数値

$a_x$  死亡率の標準年齢パターン(年齢別死亡率の年次「平均」)

$k_t$  死亡の一般的水準(死亡指数)

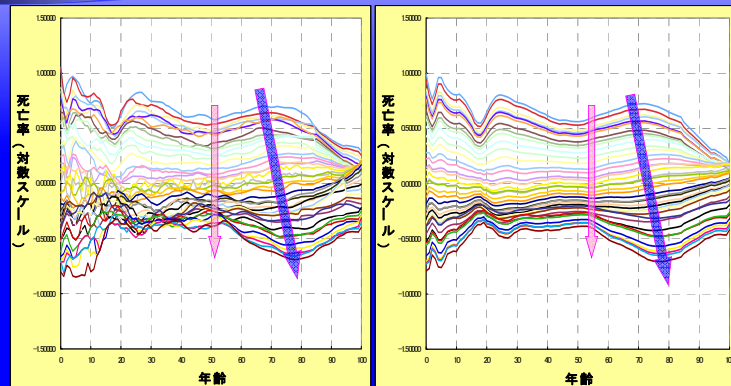
$b_x$  死亡率変化の標準年齢パターン

$\varepsilon_{x,t}$  残差(平均0)

82

### 死亡スケジュール変化の表現のためのモデル

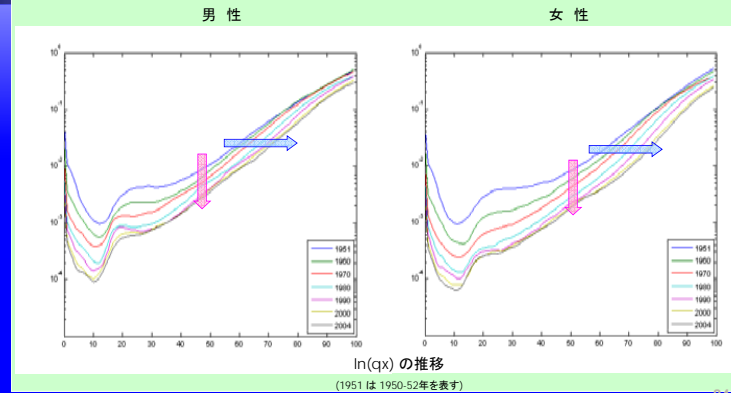
右の図(観測値)を第一、第二特異値からの推計に変えてみると、死亡率水準のピークが高年齢へシフトしていくような動きが加わることが観察される。



83

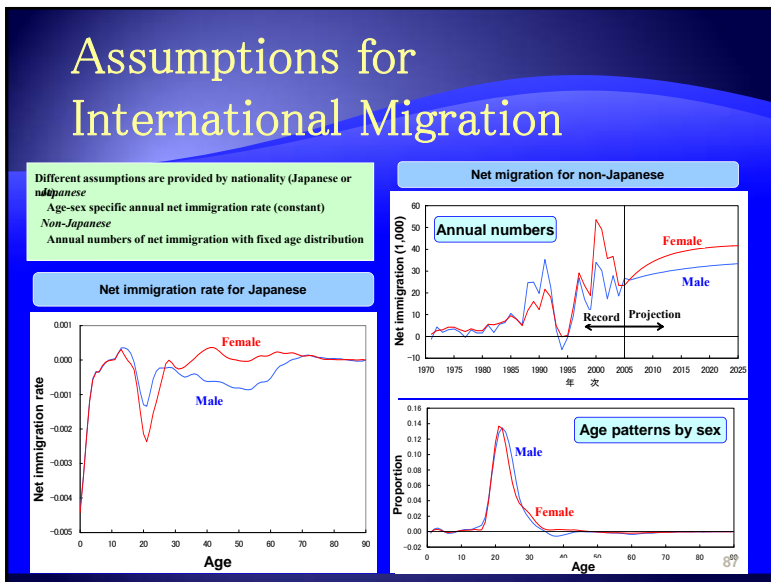
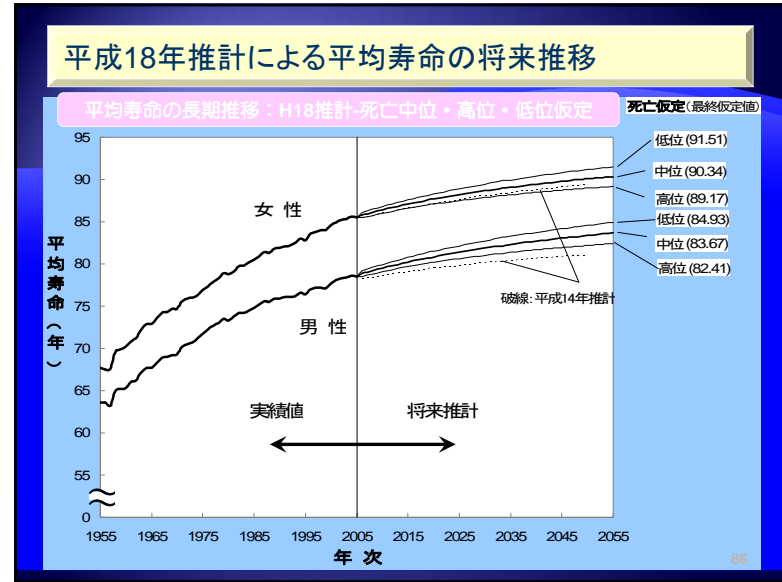
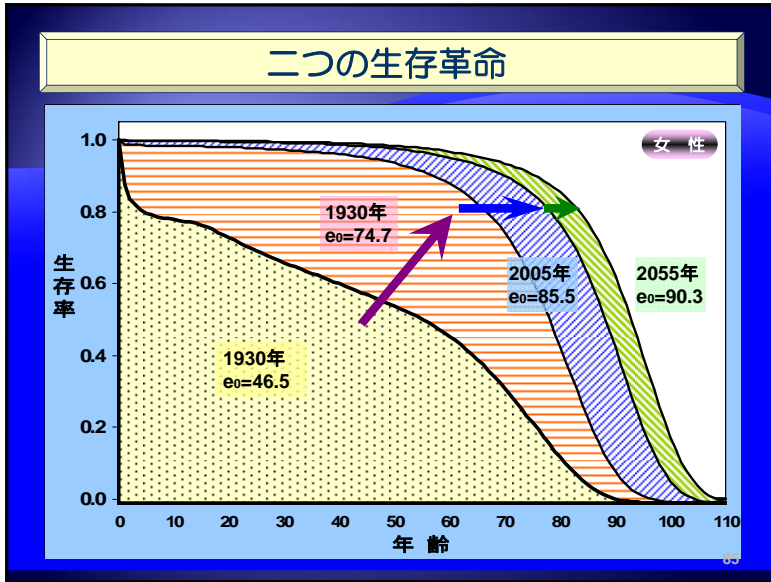
### 死亡スケジュール変化の表現のためのモデル

年齢別死亡率の高年齢における改善は、死亡率曲線が高年齢側にシフトする年齢シフトという死亡状況の変化と捉えることができる。



(1951は1950-52年を表す)

84

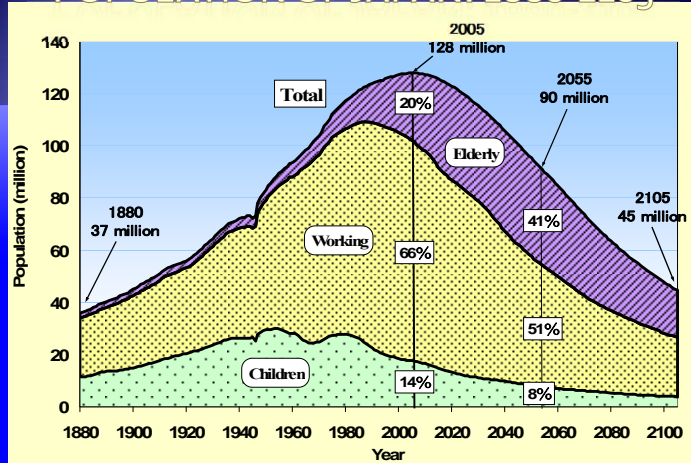


# 将来人口推計

## — 結果の概略 —

88

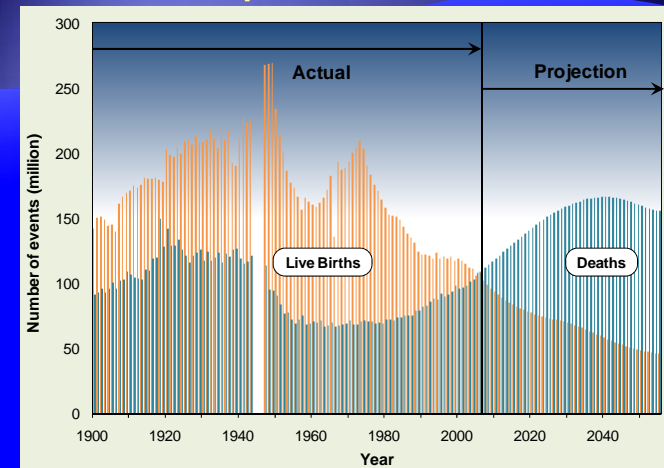
## POPULATION OF JAPAN: 1880-2105



Source: Ministry of Internal Affairs and Communications, Statistics Bureau, Census, NIPSSR(2006), Population Projection for Japan:2006-2055.

89

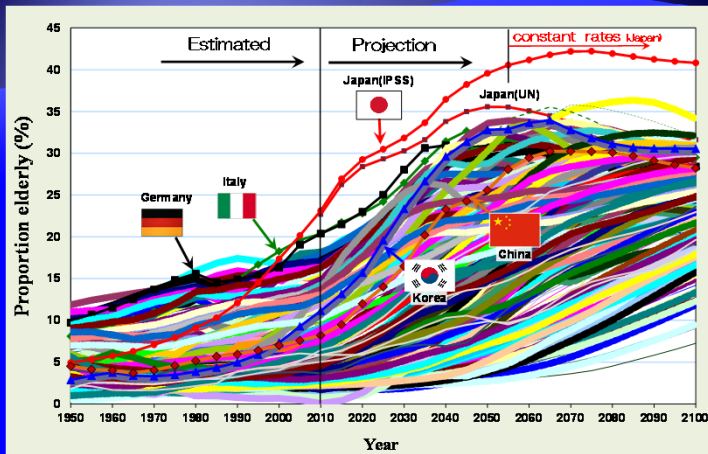
## Japanese Cross



Source: NIPSSR(2006)

90

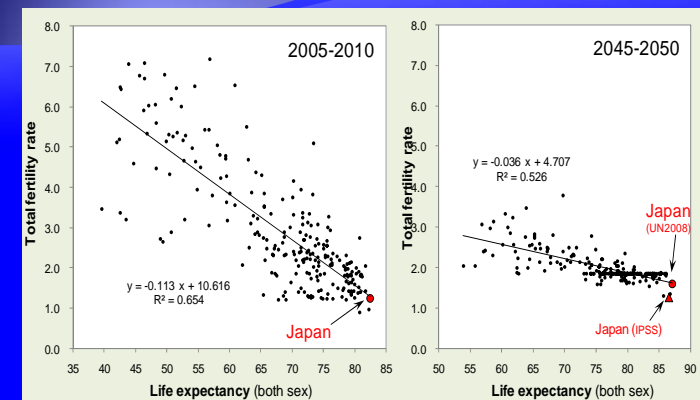
## Proportion Elderly: 1950-2100



Source: United Nations(2011), World Population Prospects: The 2010 Revision, NIPSSR(2006), Population Projection for Japan:2006-2055.

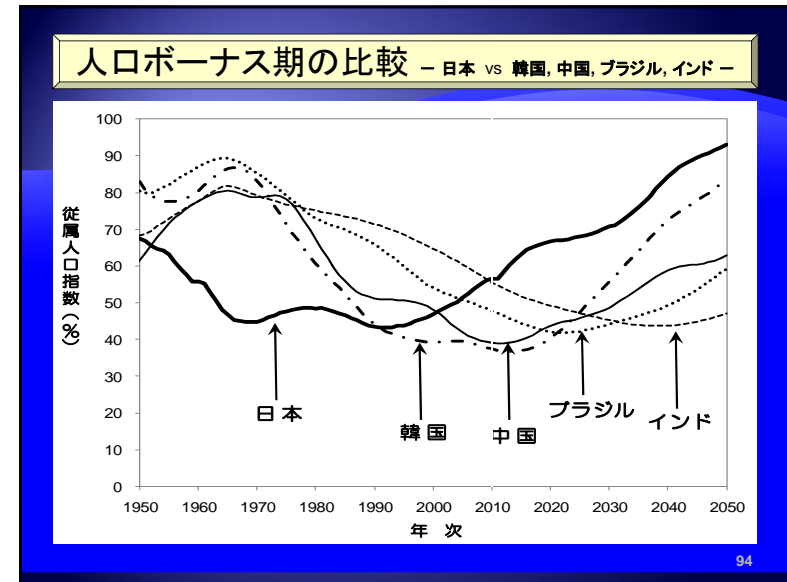
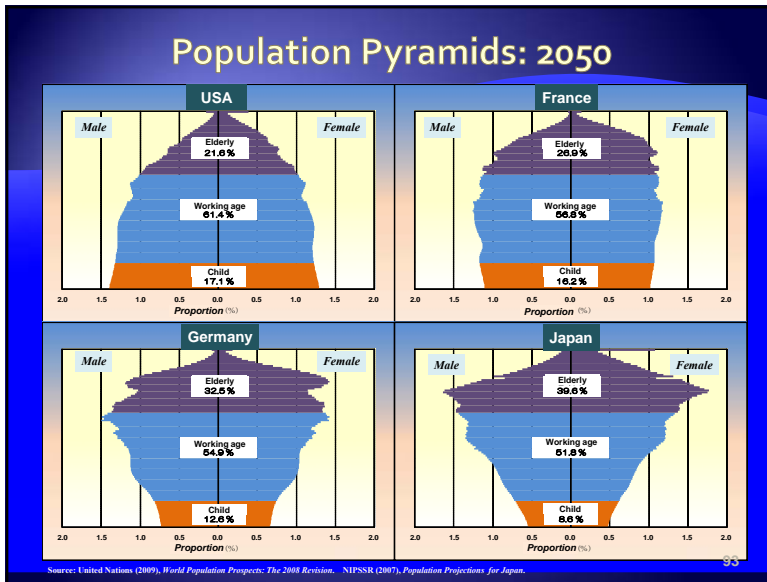
91

## Japan's Distinguishing Position: Present and Future



Source: United Nations (2009), NIPSSR(2006)

92



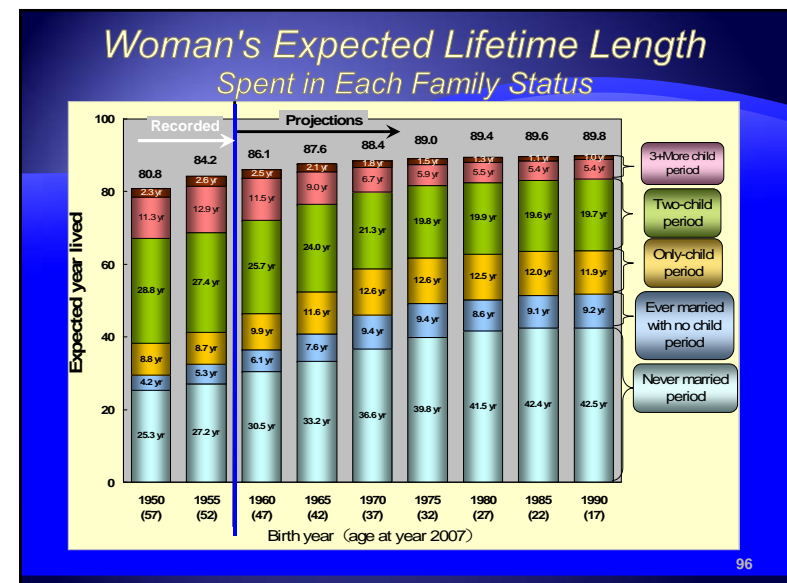
### Woman's Life-course Composition by Family Status : Recoded and Projected

Derived from Multistate Life Tables

Family Status	Birth year of woman's cohort								
	Recorded		Projected						
	1950	1955	1960	1965	1970	1975	1980	1985	1990
Life time distribution of woman by number of child (at birth)									
Childless	18.4	17.7	20.8	24.8	31.4	35.5	36.4	37.9	38.1
Never married	13.6	11.2	12.9	14.4	17.9	21.7	23.6	24.3	24.3
Ever married	4.8	6.5	7.9	10.5	13.5	13.8	12.8	13.6	13.8
Only child	11.2	11.2	13.3	16.5	18.7	18.8	18.8	18.1	18.0
Two children	46.8	44.4	41.8	39.5	35.2	32.9	33.0	32.6	32.8
Three children	19.4	22.0	19.6	15.3	11.5	10.1	9.4	9.3	9.3
Four and more children	4.2	4.7	4.5	3.8	3.2	2.7	2.3	2.0	1.9
Net Reproduction Rate	87.5	90.0	84.5	76.3	66.3	61.2	59.6	58.1	57.9
No grandchild	22.2	21.2	25.6	31.6	41.2	46.8	48.1	50.0	50.2

Derived from Assumption

Life time proportion of woman (without mortality effect = directly derived from fertility assumption)									
Never married	5.0	5.8	9.3	12.0	16.2	20.4	22.6	23.5	23.5
Childless	10.3	12.7	17.5	22.7	30.0	32.8	35.7	37.1	37.4
No grandchild	12.1	15.0	21.3	28.8	39.3	42.9	46.8	48.9	49.4



# 参考文献

## References

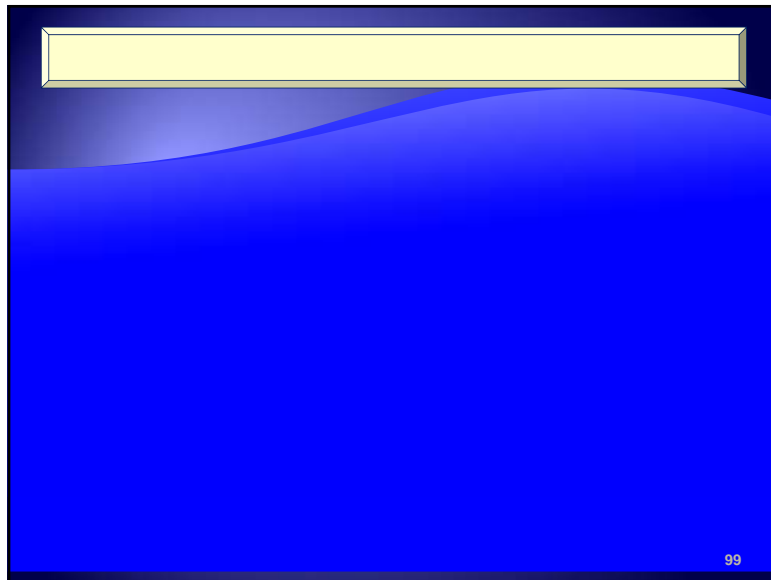
石川晃・佐々井司、2010。「行政記録に基づく人口統計の検証」『人口問題研究』第66巻第4号、pp.23-40。  
 石井太、2008。「近年のわが国の死亡動向に適合した将来生命表推計モデルの研究 一年齢シフトモデルの開発」『人口問題研究』第64巻第3号、pp.28-44。  
 岩澤美帆、2008。「初婚・離婚の動向と出生率への影響」『人口問題研究』第64巻第4号、pp.19-34。  
 金子隆一、「人口統計学の展開」『日本統計学会誌』第31巻、第3号、pp.345-377、日本統計学会（2001.12）  
 金子隆一、2009年、「将来人口推計における出生仮定設定の枠組みについて」『人口問題研究』国立社会保障・人口問題研究所 第65巻 第2号 pp.1-27。  
 金子隆一・三田房美「将来人口推計の基本的性質と手法的枠組みについて」『人口問題研究』第64巻第3号、pp.3-27。  
 国立社会保障・人口問題研究所編（2007）『日本の将来推計人口ー平成18年12月推計ー』厚生統計協会  
 国立社会保障・人口問題研究所編（2008）『日本の将来推計人口ー平成18年12月推計の解説および参考推計（条件付推計）ー』厚生統計協会

# 人口統計の体系

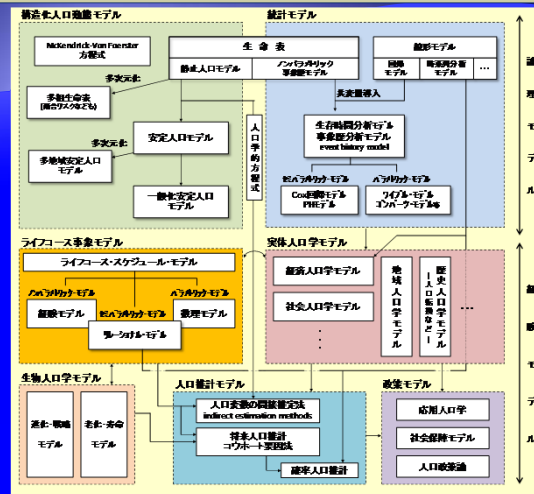
(人口統計資料集 2011年版)

- I. 人口および人口増加率
- II. 年齢別人口
- III. 人口動態率
- IV. 出生・家族計画
- V. 死亡・寿命
- VI. 結婚・離婚・配偶関係別人口
- VII. 世帯
- VIII. 労働力
- IX. 地域移動・地域分布
- X. 国籍別人口および国際移動
- XI. 教育
- XII. 都道府県別統計

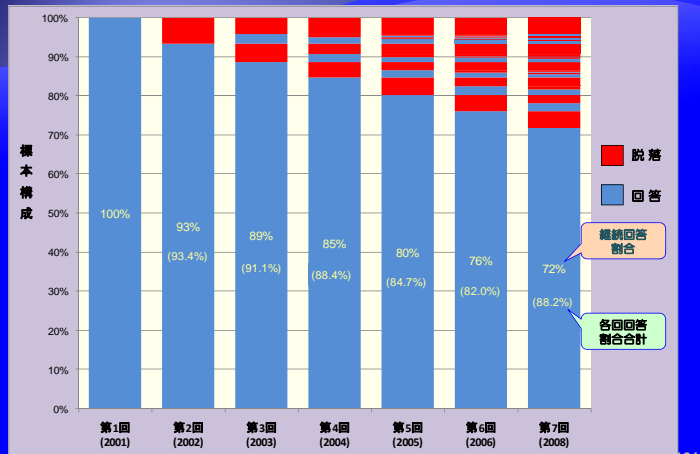
国立社会保障・人口問題研究所ウェブページ <http://www.ipss.go.jp/>



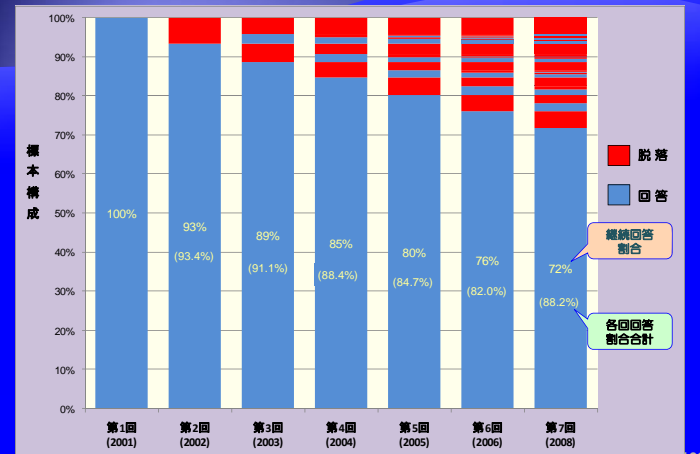
# 人口科学におけるモデルの体系



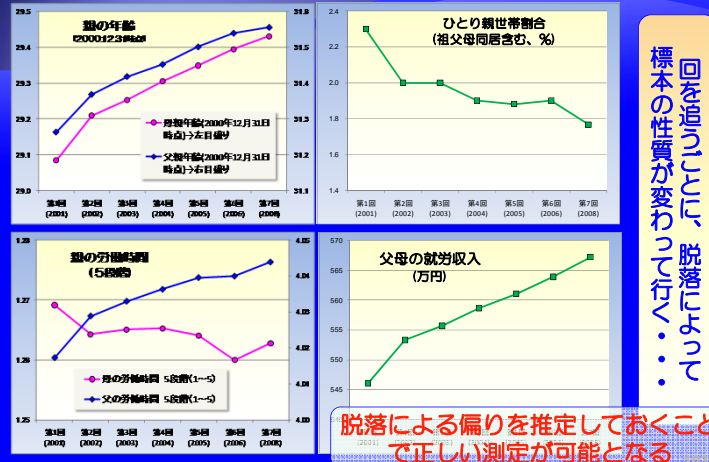
# 出生児縦断調査の脱落・復活状況



# 出生児縦断調査の脱落・復活状況



# 同一変数に対する脱落の影響の例



回答者の性質が変わって行くと、脱落によって...

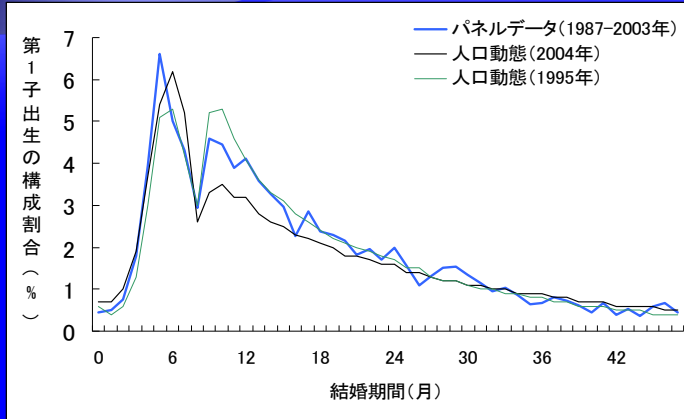
# 脱落の影響を受けた変数 (抜粋)

図表3 脱落の有無により第1回調査時の変数がとる値及び第1回調査との有意差の有無

変数対象	第1回		第2回		第3回		第4回		第5回	
	回答者	脱落者	回答者	脱落者	回答者	脱落者	回答者	脱落者	回答者	脱落者
母の年齢	31.26	31.37 ***	31.42 ***	31.45 ***	31.45 ***	31.51 ***	31.50 ***	31.57 ***	31.57 ***	31.57 ***
2000年12月31日時点での父親年齢	29.08	29.21 ***	29.25 ***	29.30 ***	29.31 ***	29.37 ***	29.35 ***	29.43 ***	29.43 ***	29.43 ***
2000年12月31日時点での母親年齢	96.6	97.0 ***	97.1 ***	97.2 ***	97.2 ***	97.3 ***	97.2 ***	97.4 ***	97.4 ***	97.4 ***
父母とも日本人	18.6	18.0 **	17.7 ***	17.5 ***	17.5 ***	17.2 ***	17.2 ***	16.9 ***	16.9 ***	16.9 ***
対象児を婚前妊娠していた	99.9	99.9	99.9 *	99.9 *	99.9 *	99.9 **	99.9 **	99.9 **	99.9 **	99.9 **
同居の状況(母)	97.7	97.9 ***	98.0 ***	98.1 ***	98.1 ***	98.2 ***	98.1 ***	98.2 ***	98.1 ***	98.2 ***
同居の状況(父)	6.4	6.3	6.2 *	6.1 *	6.1 *	6.0 **	6.1 **	6.1 **	5.97 ***	5.97 ***
兄弟姉妹の人数(母子込み)	0.69	0.68	0.68	0.68 *	0.68	0.68	0.68	0.68 *	0.68 *	0.68 *
核家族世帯	76.7	76.9	77.0	77.1 *	77.2 *	77.2 *	77.1 *	77.4 **	77.4 **	77.4 **
ひとり親世帯(祖父母同居含む)	2.3	2.0 ***	2.0	1.9 ***	1.9 ***	1.8 ***	1.9 ***	1.8 ***	1.8 ***	1.8 ***
妊娠出産に伴う課題・増援の有無	11.7	11.3 **	11.2 ***	11.1 ***	11.1 ***	11.0 ***	11.0 ***	10.8 ***	10.8 ***	10.8 ***
ふだんの保育者(母)	97.1	97.3 *	97.3 *	97.3 **	97.3 **	97.4 **	97.4 **	97.5 **	97.5 **	97.5 **
ふだんの保育者(父)	46.4	47.1 **	47.2 ***	47.5 ***	47.5 ***	47.9 ***	47.7 ***	48.2 ***	48.2 ***	48.2 ***
ふだんの保育者(保育所の保育士)	3.9	3.7	3.7 *	3.6 *	3.6 *	3.6 **	3.6 **	3.5 **	3.5 **	3.5 **
ふだんの保育者(親と保育士等)	2.0	2.0	1.9	1.9	1.9	1.9 *	1.9	1.9 *	1.9 *	1.9 *
保育士や保育ママやベビーシッターの利用	4.2	4.1	4.0 *	4.0 *	4.0 *	3.9 **	4.0 **	3.9 **	3.9 **	3.9 **
授乳は母乳のみ	21.0	21.5 *	21.6 **	21.8 ***	21.8 ***	21.7 ***	21.8 ***	22.1 ***	22.1 ***	22.1 ***
母乳を与えた期間(月)	4.48	4.55 ***	4.56 ***	4.58 ***	4.58 ***	4.61 ***	4.59 ***	4.63 ***	4.63 ***	4.63 ***
母の育児(食事の世話をする) 4段階(1~4) ※1が高頻度	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02 *	1.02 *	1.02 *
母の家事(食事をとく) 4段階(1~4) ※同上	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08 ***	1.08 ***	1.08 ***
母の家事(食卓の後片づけをする) 4段階(1~4) ※同上	1.08	1.08	1.08	1.07	1.07	1.07	1.07	1.07 **	1.07 **	1.07 **
母の家事(日常の買い物をする) 4段階(1~4) ※同上	1.17	1.16	1.16	1.16	1.16	1.16	1.16	1.16 *	1.16 *	1.16 *
父の育児(入浴させる) 4段階(1~4) ※同上	1.81	1.80	1.80	1.80 *	1.80 *	1.80 *	1.80 *	1.80 *	1.80 *	1.80 *
子を持ってよかったこと身近な人が喜んでくれた	78.1	78.4	78.5	78.6 *	78.5	78.7 **	78.5 *	78.5 *	78.5 *	78.5 *
子を持って負担に思うこと子育てによる身体の疲れが大きい	39.5	39.8	39.8	40.0 *	39.9	40.0 *	39.9	40.1 *	39.9	40.1 *
子を持って負担に思うこと子育てで出費がかさむ	34.7	34.4	34.2 *	34.2 *	34.1 **	34.0 **	34.0 **	34.0 **	34.0 **	34.0 **
子を持って負担に思うこと自分の自由な時間が持たない	55.2	55.6	55.7 *	55.8 **	55.8 **	56.0 **	56.0 **	56.0 **	56.0 **	56.0 **

## 結婚期間による第1子出生過程

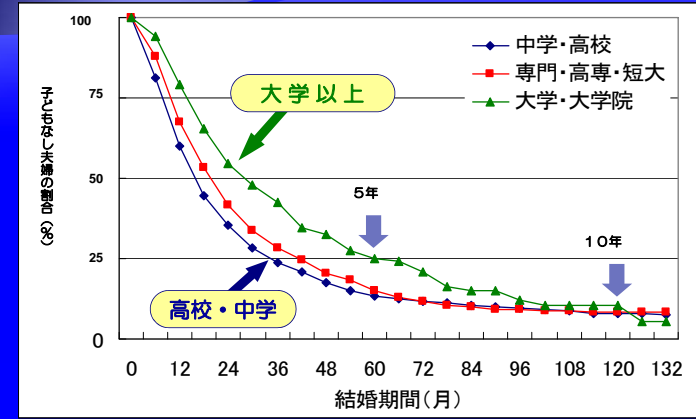
— 縦断調査と人口動態統計の比較 —



105

## 妻の年齢による第1子出生過程

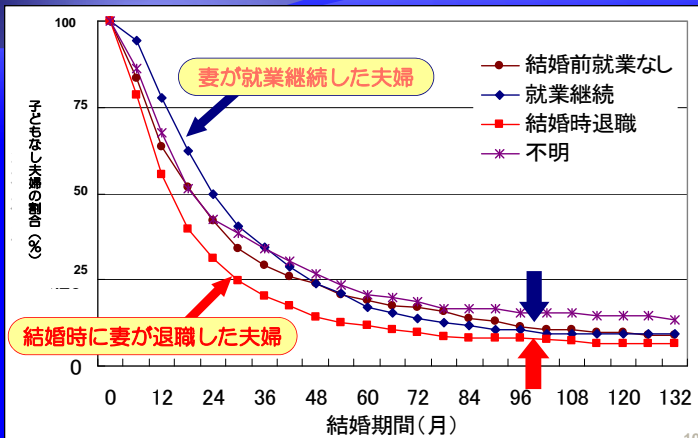
— 妻の教育程度の別 —



106

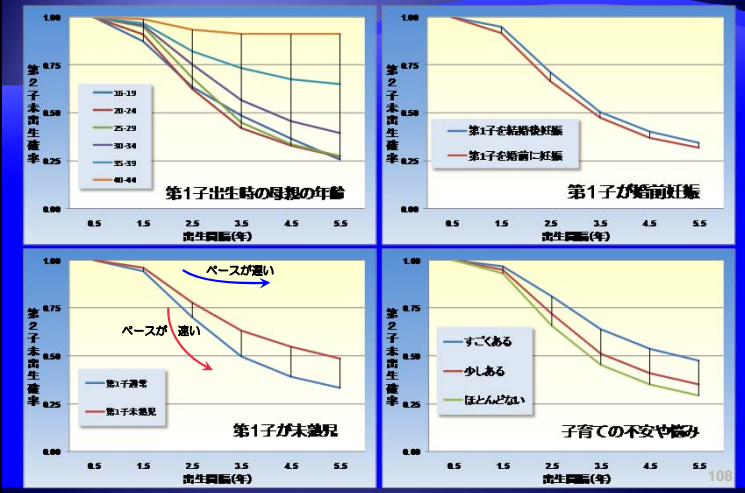
## 結婚期間による第1子出生過程

— 妻の就業継続状況の別 —



107

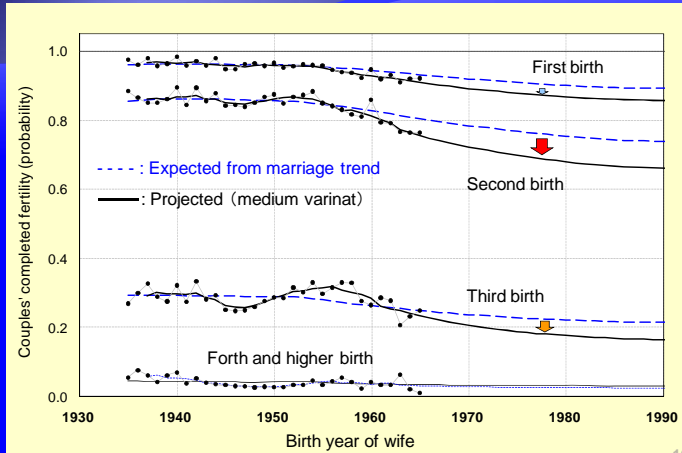
## 第2子の生み方に対する第1子の影響



108



### Cohort Trends of $CEB_n$ for Married Couple : Expected from Marriage Trend and Projected



109

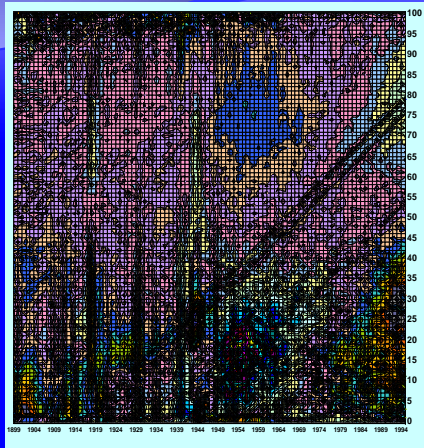
### Assumed Values for Measures of Fertility Components, and Fertility in Population Projection 2006 in Japan

Measures of Fertility Components	Observed: cohort born in 1955	Assumptions of Population Projection: Femal cohort born in 1990		
		Medium	High	Low
(1) Mean age at first marriage	24.9	28.2	27.8	28.7
(2) Proportion never married at age 50	5.8 %	23.5 %	17.9 %	27.0 %
(3) Couples' completed fertility	2.16	1.70	1.91	1.52
(4) Effect of divorce, widowhood and remarriage	0.952	0.925	0.938	0.918
<b>Children</b>				
0 (Childless Ratio)	12.7 %	37.4 %	28.6 %	43.3 %
1 (Only-Child Ratio)	11.8 %	18.2 %	15.4 %	19.4 %
2 and over	75.6 %	44.4 %	55.9 %	37.2 %
<b>Cohort Total Fertility Rate</b> (the rate only for birth from Japanese women)	1.94	1.26 (1.20)	1.55 (1.47)	1.06 (1.02)

Source: Ministry of Health, Labour and Welfare, Vital Statistics, NIPSSR(2006), Population Projection for Japan:2006-2055.

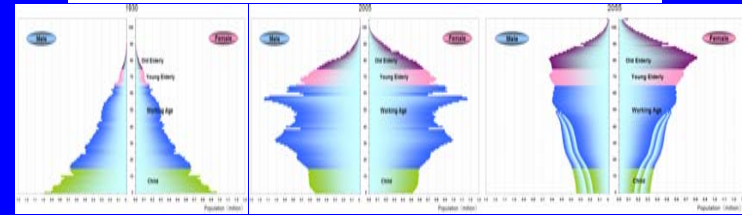
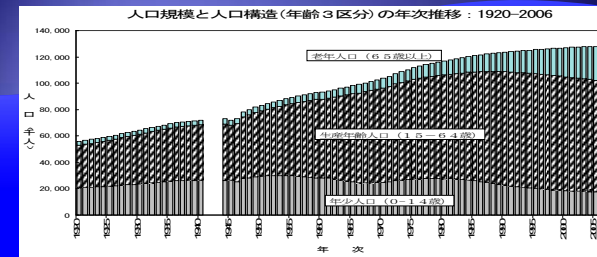
110

### Residual Pattern of Mortality by Age and Period: $\ln M_x - (ax + bx \cdot Pc1)$ : France-Female(1899-1995)



111

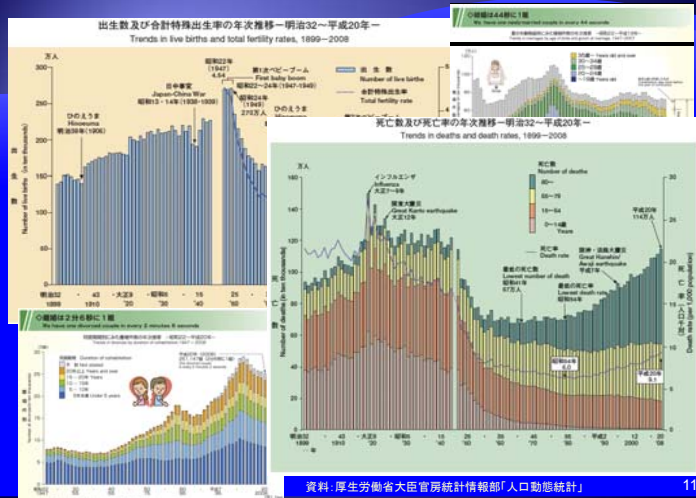
### 人口静態統計



資料: 総務省統計局「国勢調査」「推計人口」、国立社会保障・人口問題研究所「日本の将来推計人口(平成18年12月推計)」

112

# 人口動態事象と人口動態統計



# 人口統計学の基礎知識

## 人口動態事象

# 人口学的方程式

Demographic balancing equation

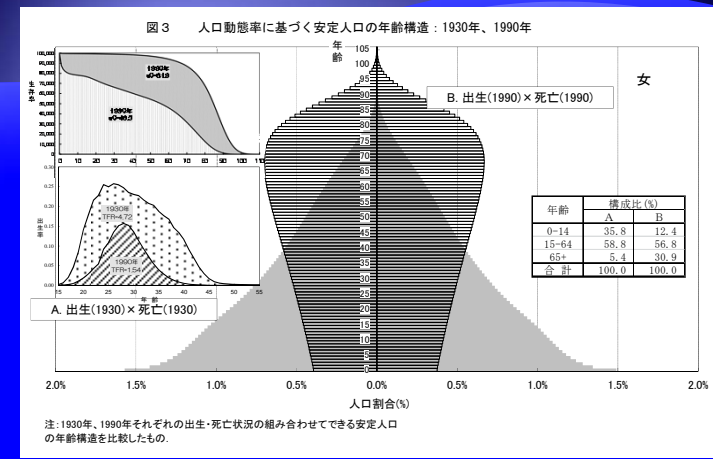
$$P_{t1} - P_{t0} = B_{t0,t1} - M_{t0,t1} + I_{t0,t1} - E_{t0,t1}$$

(期末人口) - (期首人口) = (出生数) - (死亡数)  
+ (転入者数) - (転出者数)

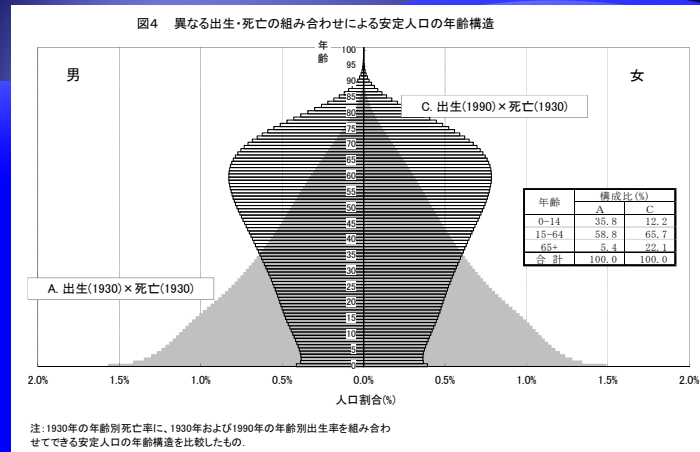
人口増加

- ① 人口静態と人口動態事象との関係を記述
- ② 実際に未知数を求める方程式として用いる  
e.g. 「人口推計」
- ③ 測定誤差の推定に用いる

# 安定人口モデルによる高齢化シミュレーション

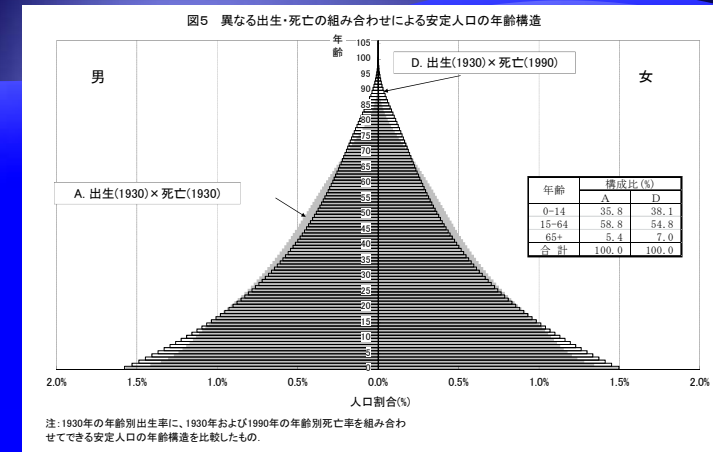


## 安定人口モデルによる高齢化シミュレーション



117

## 安定人口モデルによる高齢化シミュレーション



118

## 人口統計の体系 (その1)

(政府統計委員会2008年ワーキンググループ資料より)

### I. 人口静態に関する統計 (人口の規模と構造 (属性別構成) に関する統計)

1. 人口センサス (男女・年齢・配偶関係・居住地・就業状態などの属性別人口)
  - ① 「国勢調査」(総務省統計局)
2. 行政記録
  - ① 地域住民(日本人)・・・「住民基本台帳に基づく人口・人口動態及び世帯数」(総務省自治行政局)
  - ② 日本人(日本国籍を有する国民)・・・戸籍
  - ③ 外国人(日本に居住する者)・・・「在留外国人統計」(法務省)
  - ④ 海外に居住する日本人・・・「海外在留邦人調査統計」(外務省)
3. 加工統計
  - ① 「現在推計人口」(総務省統計局)
  - ② 「将来人口推計」(厚生労働省・研究所)
4. 標本調査
  - ① 「国民生活基礎調査」(厚生労働省)など

119

## 人口統計の体系 (その2)

(政府統計委員会2008年ワーキンググループ資料より)

### II. 人口動態に関する統計 (人口変動要因に関する統計)

- A. 自然動態 (出生・死亡とこれに直接関連する事象の統計)
  1. 人口センサス
    - ① 同居見法による推定出生統計・・・「国勢調査」
  2. 行政記録
    - ① 出生・死亡・婚姻・離婚・死産・・・「人口動態統計」(厚生労働省)
    - ② 人工妊娠中絶・・・「母体保護統計」(厚生労働省)
  3. 加工統計
    - ① 生存・死亡状況・・・「生命表」(厚生労働省)
  4. 標本調査
    - ① 縦断(パネル)調査・・・「21世紀縦断調査」(厚生労働省)
    - ② 結婚・出生調査・・・「出生動向基本調査」(厚生労働省・研究所)など

120

# 人口統計の体系 (その3)

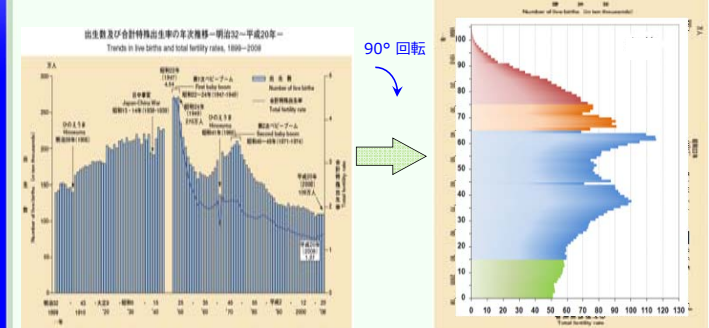
(政府統計委員会2008年ワーキンググループ資料より)

- II. 人口動態に関する統計 (人口変動要因に関する統計) . . . つづき
- B. 社会動態 (国内・国際人口移動/国籍変更)
- a. 国内人口移動
1. 人口センサス
    - ① 前従地/現住地による「国勢調査」
  2. 行政記録
    - ① 転入届・転出届による「住民基本台帳人口移動報告」(総務省統計局)
  3. 標本調査
    - ① 「人口移動調査」(厚生労働省・研究所)
- b. 国際人口移動
2. 行政記録
    - ① 「出入国管理統計」(法務省)
- c. 国籍異動
- ① 国籍異動(法務省資料)

# 人口構造の形成

## ◎ 人口構造の形成

個人の属性によって分けた人口の内部構成を人口構造と呼ぶ。とくに男女・年齢構成を基本構造と呼ぶ。国際人口移動の少ない国では、基本構造は、男女別出生数の年次推移と、男女別の生存率によって形成される。



出生数の年次推移

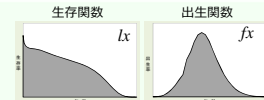
生存率を反映

# 構造化人口動態モデル

## The Structured Population Dynamics model

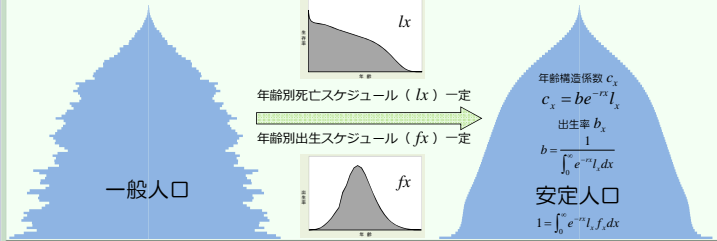
### ◎ 出生と死亡の年齢スケジュール

構造化人口モデルにおいては、出生と死亡の年齢スケジュールが重要な役割を果たす。ここでは死亡スケジュールを年齢別生存率 (生まれてからx歳まで生存する率) で表し、一方出生スケジュールを (女性の) 年齢別出生率 (各年齢での出生頻度) によって表す。



### ◎ 安定人口モデル stable population model

出生と死亡の年齢スケジュールが長期間一定だと、人口は初期の年齢構成に関係なく、当該スケジュールの組み合わせに固有な年齢構成に近づいて行く。この一意的に決まる年齢構成を持った理論人口を安定人口と呼ぶ。



# 人口現象の把握

人口動態事象と人口静態の機能的関係

## I. 人口静態統計

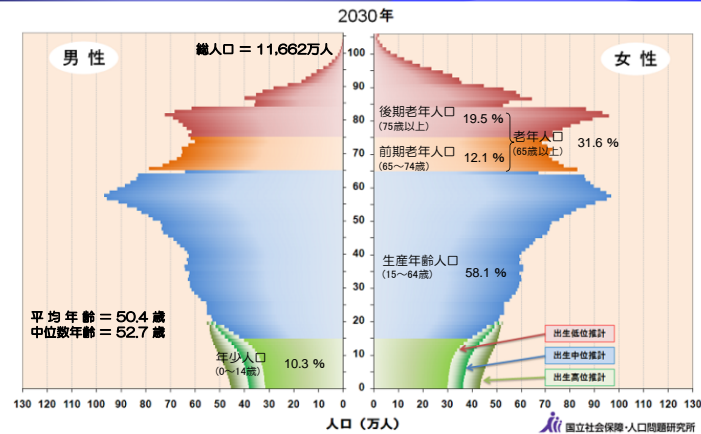
- 人口静態(ストック)に関する統計
  - 人口規模
  - 人口構造 (個人の属性によって分けた人口の内部構成)

## II. 人口動態統計

- 人口動態(フロー)に関する統計、人口動態事象の把握
    - 出生
    - 死亡
    - 人口移動
- 人口変動の3要素

事象 (フロー)	→	状態 (ストック)
人口動態事象	→	人口静態
出生・死亡・移動	→	人口規模
その他の事象 e.g. 結婚 就業・失業 ...	→	人口構造 e.g. 配偶関係構造 労働力人口構造 ...

## 日本の人口ピラミッド 2030年



125

## 人口統計システムの整備

- ・登録人口&人口調査システム
- ・スウェーデン方式



126

## 人口統計の課題 (その2)

### 人口動態統計

#### 人口動態調査

- ・全数に対する唯一の 実地調査
- ・現住地人口 → 実態の把握(実測)
- ・実施間隔 5年 → 途中年は“推計”が必要
- ・調査環境の変動 (不在・オートロック等の増加) → 不詳増
- ・財政負担 中

#### 住民基本台帳

(戸籍)  
日本人登録

- ・行政記録 → 継続性・安定性 (調査環境の影響 小)
  - ・追加の財政負担 小 ← ICT 化有効
  - ・常住人口(登録人口) → 実態? (移動・登録タイムラグ)  
(消えた高齢者問題)
  - ・集計：粗い = 男女・年齢5歳階級(1994年~)
  - ・対象：日本人のみ
  - ・集計期日：3月末
- 平成26(2014)年より  
変更!

127

## 人口統計の課題 (その1)

### 人口動態統計

#### 国勢調査 (総務省統計局)

#### 住民基本台帳

#### 在留外国人統計 旧 登録外国人統計 (法務省)

#### 海外在留邦人統計 (外務省)

#### 戸籍統計 (法務省)

#### 人口動態統計 (厚生労働省)

128

# 人口登録システム

Population registration system

129

# 高齢者世帯の消費行動と物価指数\*

宇南山 卓（財務総合政策研究所）

慶田 昌之（立正大学経済学部）

## 要 旨

本稿では、年齢別の消費者物価指数を構築した。年齢が異なる家計は、ライフサイクルの違いから異なる財・サービスを消費しており、その行動の違いは直面する物価の違いに反映される。さらに、年齢が異なると同じ財を購入するとしても、異なる店舗を利用すると考えられる。店舗が異なれば、提供される小売サービスも異なり、物価動向にも影響を与えると考えられる。ここでは、全国物価統計調査・全国消費実態調査の業態別の情報を用いて、年齢別の購入店舗の違いも考慮した年齢別物価指数を計測した。こうした消費行動の違いを、過去 20 年の物価動向に適用すると、30～39 歳の家計に比べ 70 歳以上の家計が経験したインフレ率は約 4%（年率で約 0.5%ポイント）高かった。

キーワード：高齢者、物価指数、物価スライド、消費行動

JEL classification: E31、D12

---

\* 本研究の一部は、経済産業研究所のプロジェクトの一部である。吉川洋東京大学教授、森川正之経済産業研究所副所長には貴重なコメントを頂いた。記して感謝したい。

## 1. はじめに

本稿では、年齢別の物価指数を構築した。年齢が異なる家計は、ライフサイクルの違いから異なる財・サービスを消費しており、その行動の違いは直面する物価の違いに反映される。ここでは、年齢別の選好の違いが過去 20 年程度の長期の物価動向にどのような影響を与えたかを観察する。

年齢別の消費行動の違いの最も重要な部分は、年齢によって消費する財・サービスが異なるという点である。たとえば、高齢になると健康状態が悪化することや、高齢者は相対的には IT 関連の変化に十分に対応できないなどの理由で、保健医療の支出シェアは高めに情報通信関係費の支出は少なめに出る。

こうした財・サービスごとの消費パターンの違いは、公式の消費者物価指数でも「世帯主の年齢階級別指数」として公表されている。しかし、ウエイトの違いは 10 大費目別までしか公表されておらず、長期的に比較可能な時系列にもなっていない。また、研究者レベルでは、北村(2010)が、世帯ごとに支出シェアを計測して、「家計別の物価指数」を計測しているが、必ずしも年齢別の物価指数が目的となっていない。ここでは、年齢別の物価指数に注目し、その時系列的な動向を観察する。

さらに、年齢別の消費行動の違いとして、購入先の業態の違いを考慮する。阿部・外木(2007)は、日経 POS データを用いて、同一商品であっても店舗間で価格の動きが大きく異なることを示している。これは、店舗間で付随する小売サービスが異なることを示唆している。また、Aguiar and Hurst (2007) はスキャナー・データを用いて高齢者の購買行動が若年者と異なることを示している。すなわち、消費者は年齢によって求める小売サービスに違いがあり、小売サービスに対する選好によって購入する店舗を決めていると考えられる。実際、年齢によって購入する店舗の業態は大きく異なる。

これまで、財別のシェアには多くの注意が払われてきたが、購買行動そのものについて考慮して物価指数を計測した研究はない。また、公式の消費者物価指数(CPI)は、品目ごとに「販売数量又は従業者規模等の大きい店舗」を調査対象としており、購買行動の変化については考慮していない。宇南山・慶田(2008)でも指摘されたように、購買行動は経済厚生に大きな影響を与える可能性があり、それを無視しては適切な物価指数の計測はできない。

年齢別の物価指数を計測することは、政策的な意義も大きい。その理由の一つは、公的年金の物価スライド制を適切に運用するために必須の情報となることである。Stephens and Unayama (2011) によれば、引退後の高齢者の所得の 90%以上は公的年金であり、年金支給額は高齢者の生活水準の最も重要な決定要因である。その年金の実質的な価値の維持は重要な政策課題である。そのために物価スライドを適切に運用することは必須であり、高齢者世帯の消費行動を正確に計測して物価に反映させることが必要である。

現実の物価スライドでは、基準となる「物価」は「総務省において作成する年平均の全国消費者物価指数」であり、若年世代も含めた「平均的な消費者が直面する物価」となっている。より適切なスライドには、高齢者の消費する財の種類や購入先を考慮して、高齢者が直面する物価が平均的な物価の動向とどれだけ違うかを計測する必要がある。

財別の支出シェアの違いと購入先業態の違いを考慮して、物価指数を構築した。その結果、高齢者の物価指数は、過去 20 年間の累積で、財別ウエイトの違いが原因で 1.5%、購入先業態別ウエイトの違いが原因で 0.5%、平均的な世帯の物価指数を上回っていた。これは、高齢者でウエイトの大きな保健医療の物価が上昇したこと、一般小売店やデパートといった高齢者が多く支出をする業態の相対価格が上昇



してきたことによってもたらされた。

本稿の構成は以下の通りである。第 2 節では年齢別の物価の比較をすることの理論的な背景を確認した。第 3 節では、年齢別の購入行動がどのように異なるのかが検討された。第 4 節では、これらの違いを考慮した物価指数の動向を示している。第 5 節は、まとめと政策インプリケーションを論じている。

## 2. 世帯属性別の物価指数

### 2.1. 直面する物価の違いと物価指数の違い

物価指数とは、多くの財・サービスの価格(価格ベクトル)を 1 つの値(スカラー)として表現したものである。ここでは、各財・サービスの価格がどのように集計され、物価として計測できるかを考える。

まず、消費者が効用最大化問題を解いていると考えれば、一定の効用水準  $\bar{u}$  を得るのに必要な支出額  $X$  は、支出関数  $e$  によって次のように書ける。

$$X = e(p, \bar{u}; \delta)$$

ただし、 $p$  は財・サービスの価格ベクトルで、 $\delta$  は効用関数のパラメータであり家計属性による選好の違いである。

通常の物価指数の議論は、さらに  $\delta$  を固定して、財・サービスの価格が変化した時の「真の物価」を次のように定義する (たとえば、Diewert (1981) を参照)。

$$P_1 = \frac{e(p_1, \bar{u}; \delta)}{e(p_0, \bar{u}; \delta)}$$

ただし、 $p_0$  および  $p_1$  はそれぞれ基準時点・比較時点での各財・サービスの価格ベクトルである。さらに、ラスパイレス指数によって計算される消費者物価指数(CPI)がその上限となることもよく知られている。すなわち、基準時点での各財・サービスの消費量ベクトルを  $q_0$  として、

$$CPI_1 \equiv \frac{p_1 \cdot q_0}{p_0 \cdot q_0} \geq \frac{e(p_1, \bar{u}; \delta)}{e(p_0, \bar{u}; \delta)}$$

となる。

一方で、概念的には支出関数から、同一の財・サービスの価格ベクトルのもとで直面している物価が高いのはどのような家計属性であるかを知ることができる。たとえば、 $\delta_0$  を基準となる家計の効用パラメータ、 $\delta_1$  を比較対象の家計の効用パラメータとすれば、

$$R = \frac{e(p, \bar{u}; \delta_1)}{e(p, \bar{u}; \delta_0)}$$

ただし、 $R$  は同じ効用水準を得るのに必要な費用の相対的な高さを示す指数であり、比較対象の属性を持つ世帯が基準となる属性を持つ世帯とを比べたものである。たとえば、肉が安くて魚が高いような状況で、若年者は肉が好きで高齢者は魚が好きであれば、同じ効用水準を得ようとする高齢者ほど多くの支出をする必要があるというイメージの比較である。

こうした世帯属性ごとの「同一の効用水準を得るために必要な支出額」を計測することは、政策的には重要な役割がある。たとえば、同じ金額でも世帯属性ごとに実質的な価値（その補助金によって得られる効用水準）が異なるのであれば、支給する補助金を調整することができる。また、異なる家計属性を持つ世帯間で効用水準を比較するのも利用することができる。

しかし、残念ながら、効用そのものが観察可能ではないため、異なる効用関数を持つ家計間で物価を比較することは強い仮定を置かない限り不可能である。異なる家計属性を持つ世帯が、同じ価格ベクトルの下で、異なる行動をとっていたとして、それが効用水準の違いによるものなのか効用パラメータの違いによるものなのかを識別することは不可能なのである。さらに言えば、そもそも効用とは序数的な概念であり、異なる選好を持つ主体間では効用水準の比較は基本的に不可能なのである。言い換えれば、どのようなデータが利用できるにしても、識別は不可能なのである。

そこで、家計属性ごとに直面する物価の絶対水準を比較することではなく、「家計属性ごとの物価の変化」を比較することを考える。すなわち、世帯属性が  $k$  である世帯の消費者物価指数を

$$P_1^k \equiv \frac{e(p_1, \bar{u}; \delta_k)}{e(p_0, \bar{u}; \delta_k)} \quad (1)$$

と定義するのである。

この指数によって、たとえば、肉の価格も魚の価格も変化した時に若年者と高齢者のどちらが直面する物価がより大きく上昇したかが比較可能になる。もちろん、物価上昇率が高いからといって、高い物価に直面しているとは限らない。しかし、ある時点での状態を所与とすれば、その後の変化は計測できるのである。この「家計間の物価の絶対水準は比較が不可能」であることと、「家計属性間の CPI の違いによって直面する物価の変化の違いが比較できる」ということの区別は、以下の議論を通じて十分に注意する必要がある。

## 2.2. 年齢階級別の物価指数

ここまで、「家計属性の違い」と一般的な表現を用いてきたが、現実に属性別の物価指数が計測できるケースは限定される<sup>1</sup>。属性別の指数が意味を持つのは、選好が安定的に異なる家計のグループが存在し、さらにその違いが公式統計等で観察可能でなければならない。ここでは、家計属性の違いのうち、特に年齢別の違いに注目する。

年齢別に注目する理由の 1 つは、先験的に消費する財・サービスが年齢によって大きく異なることで

---

<sup>1</sup> 北村(2010)は、世帯ごとに支出シェアを計測しているが、それが選好の違いによってもたらされているかは検証していない。

ある。高齢者は、平均的には健康状態が良くないため保健医療の支出シェアが大きい。また、子育て世代では極めて大きな支出シェアを占める教育費も、まだ子供のいない世代や子育てを終えた世代にとっては大きなウェイトを持たない。さらに、高齢者は相対的には IT 関連の変化に十分に対応できていないため、通信機器やパソコンなどに対する支出が小さいなど、消費パターンが異なることは十分に予想できる。

また、年齢別のインフレ率の計測が、世代別の再分配政策に重要な含意を持つことも理由の 1 つとなる。その典型が、「年金給付水準の実質価値の維持を自動的にはかるため、物価上昇分だけ年金額をスライドさせる制度 (牛丸, 1996)」である物価スライド制度である。民間保険では回避が難しい「将来のインフレによる年金額の目減りリスク」に対応するための制度であり、高齢者にだけ影響のある制度である。それにもかかわらず、物価スライドの基準となる「物価」は、「総務省において作成する年平均の全国消費者物価指数」であり、若年世代も含めた「平均的な消費者が直面する物価」である。もし年齢別に直面するインフレ率が異なるなら、年金受給者の直面するインフレ率で物価スライドを実施すべきである。その基礎となる年齢別の物価指数の計測はここでの目的の一つである。

年齢別の物価指数を計測するということは、概念的には(3)式を計測することになるが、通常の消費者物価指数の手法を踏襲して、

$$CPI_1^k \equiv \frac{p_1 \cdot q_0^k}{p_0 \cdot q_0^k} \quad (4)$$

を計測することとする。すなわち、家計属性別に基準時点での消費ベクトル  $q_0^k$  を作成し、それをウェイトとしてラスパイレス指数を計算するだけである。

総務省統計局が公表する公式の消費者物価指数でも、「世帯主の年齢階級別指数」が公表されている。これは、基準年の家計調査に基づき年齢別の支出シェアを計算し、ウェイト作成しているのである。その意味では、年齢別の選好の違いを財・サービスへの支出の違いだけで把握しているのである。

それに対し、ここでは同じ財・サービスへの支出であっても、購入先が違えば異なる消費行動であるとみなしている。阿部・外木(2007)は、日経 POS データを用いて、同一商品であっても店舗間で価格の動きが大きく異なることを示している。市場が十分に競争的であるなら、異なる価格が併存する(価格の高い店舗が存在できる)のは、店舗間で付随する小売サービスが異なるからだと解釈せざるを得ない。つまり、財・サービスの価格の違いとは小売サービスの違いなのである。Aguiar and Hurst (2007) はスキャナー・データを用いて高齢者の購買行動が若年者と異なることを示している。すなわち、消費者は年齢によって求める小売サービスに違いがあり、小売サービスに対する選好によって購入する店舗を決定していると考えられる。

さらに、店舗によってサービスが異なれば価格動向にも違いが生まれると考えられるため、その違いも定量的に把握する必要がある。ここでは、店舗ごとの小売サービスの違いを「業態」としてとらえ、業態ごとの価格の動向と消費者の購入行動を観察する。すなわち、総務省統計局の年齢別物価指数では、財・サービスが  $N$  種類あれば、 $p$  は  $N$  次元のベクトルとなっていたが、ここでは  $K$  個の購入先業態それぞれで異なる小売サービスが付加されると考え  $K \times N$  次元のベクトルとして基準となる消費ベクトルを構築する。

### 3. 年齢別の購買行動の違い

#### 3.1. データ

ここでは、年齢別に購入先業態別がどれほど異なり、業態ごとの価格がどの程度異なるかを検証する。業態ごとの価格は、日本全体に対する代表性を持つ統計として、全国物価統計調査の「業態別小売価格」の情報を用いることができる。全国物価統計調査とは、5年に一度実施されていた価格の構造調査であり、地域間・店舗間・銘柄間の価格差を把握する統計である。毎月実施され消費者物価指数の基礎統計となる小売物価統計調査と同じ総務省統計局が作成しており、調査対象となる市町村数は4倍程度である。ここでは、1987年調査から2007年調査の5調査分のデータを用いた<sup>2</sup>。

全国物価統計調査の分類に従い、ここでは業態を「一般小売店」・「スーパー」・「量販専門店」・「コンビニ」・「デパート」・「生協」・「その他」の7つに分類した<sup>3</sup>。ただし、1997・2002・2007年の調査では上記の7業態に加えて「ディスカウントストア」調査されているが、ディスカウントストアは量販専門店に類似した業態であると考え、量販専門店とディスカウントストアの単純平均を「量販専門店」の価格として利用した。また、量販専門店・生協は1997年調査に導入されていたため、それ以前は欠損となっている<sup>4</sup>。

もちろん、すべての品目の業態別価格が意味を持つわけではなく、実際に全国物価統計調査でも調査されない品目もある。その理由は大きく分けて3つあり、業態で分けることができないもの（家賃、電気代、上下水道料など）、業態を定義するのが困難なもの（外食などのサービス関連品目）、価格が全国一律で業態間に差がないもの（教科書・学習参考教材、書籍・他の印刷物、たばこ）である。これらの品目については、業態間で価格差ないものとした。

一方、各家計がどのような業態で財・サービスを購入しているかについては、全国消費実態調査を用いた。全国消費実態調査は、5年に一度実施される家計の所得・消費の構造統計であり、毎回約5万5千世帯が調査される。毎月調査されている家計調査よりも詳細な集計が公表されており、品目別・購入先業態別の支出金額が利用可能である。

全国消費実態調査も、全国物価統計調査と同じ総務省統計局の調査であり、業態の区分はほぼ同じである。ただし、単身世帯を含む全世帯ベースで、年齢別・購入先別の集計表は中分類までだけが利用可能であるため、ここでも中分類を最小の財・サービスの最小分類単位とした。

#### 3.2. 業態と価格

まず、業態別にどの程度の価格差があるのかを示すために、中分類別・業態別価格指数を作成する。このために、全国物価統計調査で調査された品目別・業態別の価格を品目別に平均が1となるような指数に変換し、その指数を消費者物価指数の品目別ウエイトで集計し、中分類別・業態別指数を作成し

<sup>2</sup> 全国物価統計調査は、2007年調査を最後に小売物価統計調査の一部となっている。

<sup>3</sup> 1997年調査から、大規模店舗と小規模店舗で別集計となっているが、ここでは両者の単純平均をその業態の価格とした。

<sup>4</sup> 欠損した部分は最も高い業態と同じ価格としている。

た<sup>5</sup>。

表1は、2007年の全国物価統計調査に基づく中分類別・業態別の価格指数である。基本的に、全国物価統計調査では品目ごとに銘柄や属性を細かく指定して調査しており、ほぼ同一の財の価格が調査されている。それにもかかわらず、多くの品目でコンビニ・デパートでは価格が高く、スーパー・量販店の価格が低い傾向がある。一般小売店は、特に食料品を中心に平均価格を下回っているが、スーパー等よりは高めである。

ただし、最も価格差が大きい大分類「被服および履物」に含まれる費目については、価格差の少なくとも一部が、業態ごとの財の品質差である可能性が否定できない。衣類などは、完全に同一の商品を調査するのが困難なためである。たとえば、「洋服」は「スーパー」が0.52に対して、「デパート」が1.87と3倍程度の価格差があるが、スーパーとデパートの洋服が同一商品とは考えられない。ただし、衣類等については販売している業態自体が付加価値を生んでいる可能性もあり、小売サービスの差が価格に反映されていると解釈することもできる。この問題については、次節で考察する。

### 3.3. 年齢と購入先業態

表2は、全国消費実態調査から得られる年齢別の購入先業態別の支出シェアを示している。特に消費行動が異なりやすい食料(酒類・外食を除く)、家庭用耐久財、教養娯楽耐久財の購入先業態別の支出シェアを示している。

最も顕著な傾向は、高齢世帯になるほど「一般小売店」で購入する割合が、平均的な世帯よりも高いことである。ここでは全ては示していないが、「食料」ではすべての中分類項目で「一般小売店」での購入割合が平均的な世帯よりも高く、「乳卵類」では10.6パーセント・ポイントの差がある。その一方で、「スーパー」で購入する割合がすべての「食料」に含まれる中分類で低い。すなわち、高齢者世帯では食料を「スーパー」で購入せず、「一般小売店」で購入していることが分かる。

食料については、単身者の多い30歳未満の家計では「コンビニ」のシェアが高い。これは、中分類レベルでも顕著である。30歳以上の家計についても、年齢が高いほどコンビニの割合が低いことは明らかであり、年齢別の購入先業態の違いを象徴するものとなっている。

一方で、耐久財については、65歳以上の世帯では、「家庭用耐久財」では9.4パーセント・ポイント、「教養娯楽耐久財」については16.7パーセント・ポイント、平均的な世帯よりも「一般小売店」での購入割合が多い。これらの財では、若年層が「量販専門店」での購入割合が高く、「家庭用耐久財」では6.4パーセント・ポイント、「教養娯楽耐久財」については13.1パーセント・ポイントの差がでている。つまり、耐久財については、高齢者世帯では「量販専門店」で購入せず、「一般小売店」で購入する傾向があると言える。

このような購入手続きの違いは、財に対するアクセスのしやすさに影響を受けたものと解釈される。たとえば、高齢者では一般小売店での購入が他の世代に比較して多い。これは自動車などの交通手段に制約があり、住居に近い一般小売店を利用している結果であるかもしれない。また、デパートなどの利用が多いのは、数少ない外出の機会を効率的に利用してまとめて購入活動をしている可能性がある。また、

---

<sup>5</sup> 1987年と1992年は平均価格が報告されていないため、全業態の平均価格が計算されていないため、各業態の欠損値を除く単純平均を平均価格とした。

30歳未満の若年者では単身者も多く、通常的时间外で買い物をする必要があり、コンビニの利用頻度が高いと考えられる。データとしては業態別の支出シェアの違いが、店舗への物理的・時間的なアクセスのしやすさの影響で決まっているなら、それこそ「小売サービスの違い」であり、価格差の源泉であると考えられる。

## 4. 年齢別の物価指数

### 4.1. 財別・業態別・年齢別のウエイトの構築

ここでの最初の目的は、年齢別に「財別・業態別」の消費ベクトル（すなわち、(4)式における $q_0^k$ ）を構築することである。これは、通常CPIで「ウエイト」を作成することに相当するが、年齢別に構築しなければならない点と、購入先業態別にも分割する必要がある点で異なる<sup>6</sup>。

具体的な手順としては、まず財別にウエイトを分割しそれをさらに業態別に分割するという手順で構築する。すでに上で述べたように、購入先業態別の支出額は中分類レベルまで利用可能であるので、財別にも中分類レベルでウエイトを作成する。

年齢別・財別の支出額については、2004年全国消費実態調査の家計収支編「世帯主の年齢階級別支出金額」から取った。公式の年齢別物価指数の元データである「消費者物価指数年報」の「世帯主の年齢階級別ウエイト」を使うこともできるが、消費の10大費目だけが公表されており、購入先業態別データと中分類レベルでは接続ができないため、ここでは全国消費実態調査を用いた。

図1は、ここで構築された年齢別・財別の支出シェアである。この図より、多くの世帯が無職となる60歳以降に、消費の構造に大きな変化が発生することが分かる。高齢者世帯は、「食費」と「光熱・水道」に対する支出が若年者世帯よりも数%ずつ多い。また、「保健医療」のウエイトは、60歳未満の世帯では3%前後であるのに対し、60歳以上では6%と倍増する。それに対し、「教育費」のウエイトは、40～49歳階級でピークの10%程度であるのに対し、60歳以上の世帯では実質的に0%である。さらに、「交通・通信」は、60歳大きな断絶があるわけではないが、高齢者世帯になるほどシェアが低い。これは、携帯電話の通信料が、高齢者層ではそれほど大きくないことに起因する。図1にはないが、中分類レベルでの詳細な内訳を見ると、例えば、「食料」のうち「肉類」に対する支出は若年層で多く、「魚介類」に対する支出は高齢者層で多い。また、「被服および履物」に含まれる「洋服」に対する支出は若年層で多く、老年層で少ない。

この財別のウエイトを、購入先業態別に分解する。全国消費実態調査では年齢階級別・購入先業態別の支出金額も公表しており、その割合を財別シェアにかけることで分解する。すでに表2で見たように、一般に65歳以上の世帯では「一般小売店」で購入する割合が高く、「スーパー」や「量販専門店」の割

<sup>6</sup> ただし、世帯主の年齢で分類されていることには注意が必要である。世帯調査によるデータで年齢別の消費を観察する場合、世帯は世帯主の年齢で分類せざるを得ない。しかし、たとえば3世代同居のケースでは、祖父・祖母が世帯主であれば若年世代の消費が高齢者の消費として、逆に息子・娘が世帯主であれば高齢者の消費が若年世代の消費に見なされてしまう。つまり、同居高齢者の消費行動が別居高齢者と大きく異なる場合には、世帯主が65歳以上の世帯だけを見ても「代表的な高齢者」を観察することはできない。ここでは、世帯主が65歳以上の世帯の世帯員は平均で1.95人でありそのうち65歳以上の世帯員が1.45人であり、世帯主が65歳以下の世帯の平均世帯員数2.65人のうち65歳以上の世帯員は0.57人であることから、世帯主の年齢で分類することの影響は小さい。

合は低い。

この財別のシェアと購入先業態別のシェアの 2 つをかけ合わせ、財別・購入先業態別にウェイトを構築した。ただし、購入先に分割できない中分類項目(家賃・光熱水道・サービスなど)は、購入先別の分割はしなかった。その意味で、年齢別の消費パターンに大きな違いがある、医療保健と教育については業態の違いが反映されていないことには注意が必要である。

#### 4.2.財別・業態別の価格指数の構築

基準となる消費ベクトルが財別・購入先業態別になっているため、もちろん価格ベクトル(すなわち(4)式の  $p$ ) も財別・業態別に構築する必要がある。価格ベクトルについては、基準時点と比較時点のデータが必要なため、時系列方向と業態別のクロスセクション方向で別のデータを用いた。時系列的には、全購入先業態の平均の財別価格指数を作成し、各時点での各購入先業態の価格の平均価格との比をかけることで購入先業態別の時系列比較可能な価格指数を構築するのである。

時系列方向には、総務省統計局の公表する消費者物価指数(CPI;2005 年基準)を利用した。CPI の価格データは小売物価統計調査で収集されるが、その調査店舗は「調査品目ごとに販売数量又は従業者規模等の大きい店舗の順に」選定されており、業態別の価格動向は考慮されていない。ただし、実際には主要な購入先業態は含まれるように調査されており、時系列的な動向は全購入先業態の平均とみなすことができると考えられる。

一方、クロスセクション方向には、全国物価統計調査で調査された品目別・業態別価格を業態別の平均が 100 となるように換算したものをを用いる。全物価統計調査は、すでに見たように、1987 年から 2007 年までの 5 年毎に調査されており、ここでも購入先業態別の指数を 5 年毎に作成する。また、1987 年と 1992 年の全国物価統計調査では、全業態の平均価格が計算されていないため、各業態の欠損値を除く単純平均を平均価格とした。さらに、大規模店舗と小規模店舗に分けて価格が調査されている場合には、単純平均をその業態の価格とした。

#### 4.3.年齢別の価格指数の構築

上で構築した財別・購入先業態別の消費ベクトルをウェイトとして、財別・購入先業態別の価格ベクトルを指数化した年齢別の物価指数を示すのが表 3 であり、メインの結果である。使用した消費者物価指数が 2005 年のものであり、2005 年が 100 となるように基準化されているが、計算された指数を定数倍することで 1987 年を 100 となるように換算している。

第 1 列の「平均」は、財別のウェイトも購入先業態のウェイトも全世帯平均のものをを用いて計算された結果である。平均の物価指数は 2007 年に 108.50 で、20 年間で 8.5%の物価上昇があったことになる。公式の CPI は、1987 年を 100 とすれば 2007 年は 113.08 であり、同時期に約 13%の物価上昇となっている。これは、ここでの計算の元になるウェイトが、2004 年の業態別支出シェアで固定されている影響であり、相対的に価格が上昇した業態ほど 2004 年時点での支出シェアは低いことを意味している。

その横からの各列、30 歳未満、30~39 歳、・・・が年齢別の指数であり、一番右側の列は 65 歳以上の世帯を再集計したものである。65 歳以上の世帯の物価指数は、1987 年を 100 として、2007 年が 110.5

であり、平均的な世帯と比較すると 2.0%ポイント高い。また、30 歳未満を除くと、年齢が高いほどインフレ率も高くなっている。言い換えれば、高齢になるほど高いインフレを経験してきたのである。

この乖離がどのように発生したかを見ているのが、表 4-1 と表 4-2 である。表 3 は、財別ウェイトも購入先業態別ウェイトも年齢別のものを用いており、財別ウェイト・購入先業態別ウェイトを全年齢の平均シェアを用いた物価指数との差は、2 つのウェイトの差の合計となっているのである。表 4-1 及び表 4-2 では、それぞれ財別ウェイトのみ、購入先業態別ウェイトのみを年齢別のものとして、他方は全年齢の平均を用いている。もちろん、2 つのウェイトのクロス項も存在するが、1 次近似としてはこの分解が意味を持つ。

表 4-1 の財別ウェイトのみ年齢別としたものは、家計は財に関しては若年者と異なる選好を持つが、購入先の小売サービスに対しては同一の選好を持つと仮定したケースとみなすこともできる。上で見たように、年齢ごとの財別ウェイトは、10 大費目の教育・保健医療および IT 関連の耐久財で大きく異なっていた。保健医療のウェイトは 60 歳未満の世帯はで 3%前後であるのに対し、60 歳以上では 6%と倍増する。一方で、保健医療は制度的な要因によって特に価格上昇率が高い。また、IT 関連の耐久財は大幅に価格が低下したが、高齢者の支出シェアは低い。こうした要因によって、高齢者の経験したインフレ率は平均よりも約 1.5%ポイント高くなっている。一方で、財別に物価上昇率の高かった保健医療・教育のどちらも支出シェアが小さく、IT 関連の支出が多い 30 歳代は、平均よりも 1.8%インフレ率が低かった。

表 4-2 は、財別ウェイトは平均の支出シェアを用いるが、購入先業態別は年齢ごとのウェイトを用いた指数を作成し、平均との差をとったものである。これは、高齢者が若年者と財に対しての選好は同じで、小売サービスに対する選好だけが異なるケースに該当する。購入先業態別のウェイトが 65 歳以上では、平均よりも約 0.5%ポイント高くなっている。これは、高齢者が選好する一般小売店のような伝統的な業態が、量販専門店やスーパーと比較して物価の上昇率が高かったことを示している。年齢別に見れば、若年者ほどインフレ率の低い業態で購入していたことが分かる。

## 5. まとめとインプリケーション

本稿では、高齢者と若年層の消費行動の違いを考慮することで、年齢ごとに直面する物価がどのように異なるかを検討した。高齢者と若年層では、支出する財も異なり購入先も異なり、同一の消費行動をとるわけではない。支出する財の違いを考慮すれば、過去 20 年間の高齢者の直面する物価は、平均的な消費者の直面する物価よりも 1.5%程度高い率で上昇していた。また、購買行動の違いによって、0.5%程度高い率の物価上昇に直面していた。

財別の支出シェアで大きな違いが生まれたのは、IT 革命によってハイテク家電やパソコンなどの電子部品集約的な財に対するシェアの違いが原因であった。これらの製品では品質向上が大きく、物価指数が大幅に低下した一方で、高齢者これらの IT 関連の消費のシェアが若年層よりも低かった。言い換えれば、高齢者が IT 革命の恩恵を若年層ほどには受けることができなかったために生じた差である。一方で、購買行動による違いは、ディスカウント店の一般化の影響である。過去 20 年に日本では紳士服や家電の量販専門店、ドラッグストアなどが急激にシェアを伸ばした。一方で、高齢者は、価格の低下幅の小さかった、いわゆる「一般小売店」や「デパート」などで購入することが多かった。すなわち、高齢者は、



IT 革命のみならず「流通革命」にも十分に対応できなかったのである。

補足として、ここで財別・購入先別の物価指数を構築するのに、基本的にラスパイレス指数算式を用いていることには注意が必要である。すなわち、ウェイトとして、2004 年の全国消費実態調査の支出シェア・購入先別支出シェアを固定して用いている。こうした特定の消費パターンや購買パターンに基づいた物価指数は、財の選択や購入先の選択といった代替行動は考慮していないため、経済学的に適切な「直面する物価の違い」と解釈することはできない。

これは、物価指数の上方バイアスの一部であり、代替バイアスとして知られる問題である。物価指数の上方バイアスとは、現実に計測される CPI は、効用水準を一定に保つために必要な支出を示す「真の物価指数」に比べ、インフレ率が高めに出るという性質である。Unayama (2004; 2008)は、日本の公式統計の作成手順や物価の動向から、年当たりのバイアスを 0.07%程度と推計している。すなわち、物価指数の上方バイアスのうち少なくとも「上位代替バイアス」はそれほど大きくない。

さらに、より一般に、CPI による物価スライドの問題点を考えるためには、物価指数の上方バイアスについてより詳細に検討することは必要である。アメリカでは、1996 年に「ボスキンレポート」と呼ばれる消費者物価に関する諮問委員会の答申(Boskin, et al., 1996)が出されており、年 1.1%の上方バイアスがあることが指摘されている。このバイアスは、基本的に毎年発生するものであり、1.1%のバイアスを持てば、10 年後には年金額が「実質的な価値を維持する水準」よりも 11%過剰になることを意味する。言い換えれば、CPI を適正化すると年金支給を大幅に抑制できることを意味する。<sup>7</sup>

ただし、日本の CPI の上方バイアスは、アメリカほどは大きくないと考えられる。白塚(1998)は、ボスキンレポートに準じて日本でのバイアスの大きさを計測し、年 0.9%程度と結論付けていた。それに対し、総務省統計局(1999)では、アメリカでの問題が必ずしも日本には当てはまらないことを指摘し、CPI の上方バイアスが十分に小さいと反論している。実際、バイアスの半分以上が品質調整によって生じているが、もともと価格調査の銘柄管理の方法が日米で異なっており、アメリカでの問題の大部分は日本では生じていなかった。さらに、2000 年には品質調整の問題が特に重要となるパソコンとデジタルカメラについては、より洗練された手法であるヘドニック法が導入されている。また、サービス関連についても品質調整の問題が多く指摘されている。本稿では、「真の物価指数」と「計測された物価指数」の差という技術的な面については分析の対象とはしなかったが、結論には大きな影響はないと考えられる。

現在、物価指数は、年金の物価スライド制の基準としてではなく、デフレーションを背景とした金融政策の文脈でも注目されている。物価指数の政策的な位置づけがこれまでになく高まっており、経済学的な観点から物価指数とは何かを明らかにし、正確に把握することは重要な課題であろう。

---

<sup>7</sup> ボスキンレポートに対する批判や、統計作成当局の対応については Gordon (2006)を参照。

<参考文献>

- 牛丸 聡 (1996) 『公的年金の財政方式』、東洋経済新報社.
- 宇南山 卓・慶田 昌之(2008)、「流通業における規制緩和の効果: 少子高齢化社会へのインプリケーション」、RIETI Discussion Paper Series 08-J -047
- 北村 行伸(2008)「家計別物価指数の構築と分析」『金融研究』第 27 巻第 3 号、2008 年 8 月、pp. 91-150.
- 白塚 重典 (1998) 『物価の経済分析』東京大学出版会
- 総務省統計局(1999)「消費者物価指数の精度について」総務省統計局ホームページ「消費者物価指数に関する Q&A」(<http://www.stat.go.jp/data/cpi/3.htm>)
- Aguiar、 M. and E. Hurst (2007) “Lifecycle Prices and Production、” *American Economic Review*、 Vol.97、 No.5、 pp.1533-1559.
- Boskin、 M.、 et al. (1996) "Toward a More Accurate Measure of Cost of Living: Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index." Washington DC: U.S. Govt. Print Office.
- Cutler、 David、 and Ernest R. Berndt (2001) *Medical Care Output and Productivity*. NBER Studies in Income and Wealth vol 62. University of Chicago Press.
- Diewert, W. E. (1981) “The Economic Theory of Index Numbers: A Survey.” A. Deaton ed. *Essays in the theory and measurement of consumer behaviour*. pp.163-208; Cambridge University Press: London.
- Gordon、 R. J. (2006) "The Boskin Commission Report: A Retrospective One Decade Later、" *International Productivity Monitor* (Centre for the Study of Living Standards)、 vol. 12、 pages 7-22、 Spring.
- Stephens、 M. and T. Unayama (2011) “The Consumption Response to Seasonal Income: Evidence from Japanese Public Pension Benefits" Forthcoming in *the American Economic Journal: Applied Economics*.
- Triplet、 Jack E. (1999) *Measuring the Prices of Medical Treatments*、 The Brookings Institution.
- Triplet、 Jack E.、 and Barry P. Bosworth (2004) *Productivity in the U.S. Service Sector: New Sources of Economic Growth*、 The Brookings Institution.
- Unayama、 T.、 (2004) "Upward Bias in the Consumer Price Index Under the Zero Inflation Economy、" *Economics Letters*、 vol. 85、 pp. 139-144.
- Unayama、 T.、 (2008) "The Demand System and the Substitution Bias in the CPI: Evidence from the Japanese Household Survey Data、" *Applied Economics*、 vol. 40、 pp. 1795-1806.

図1 年齢階級別の財別支出シェア(10大費目)

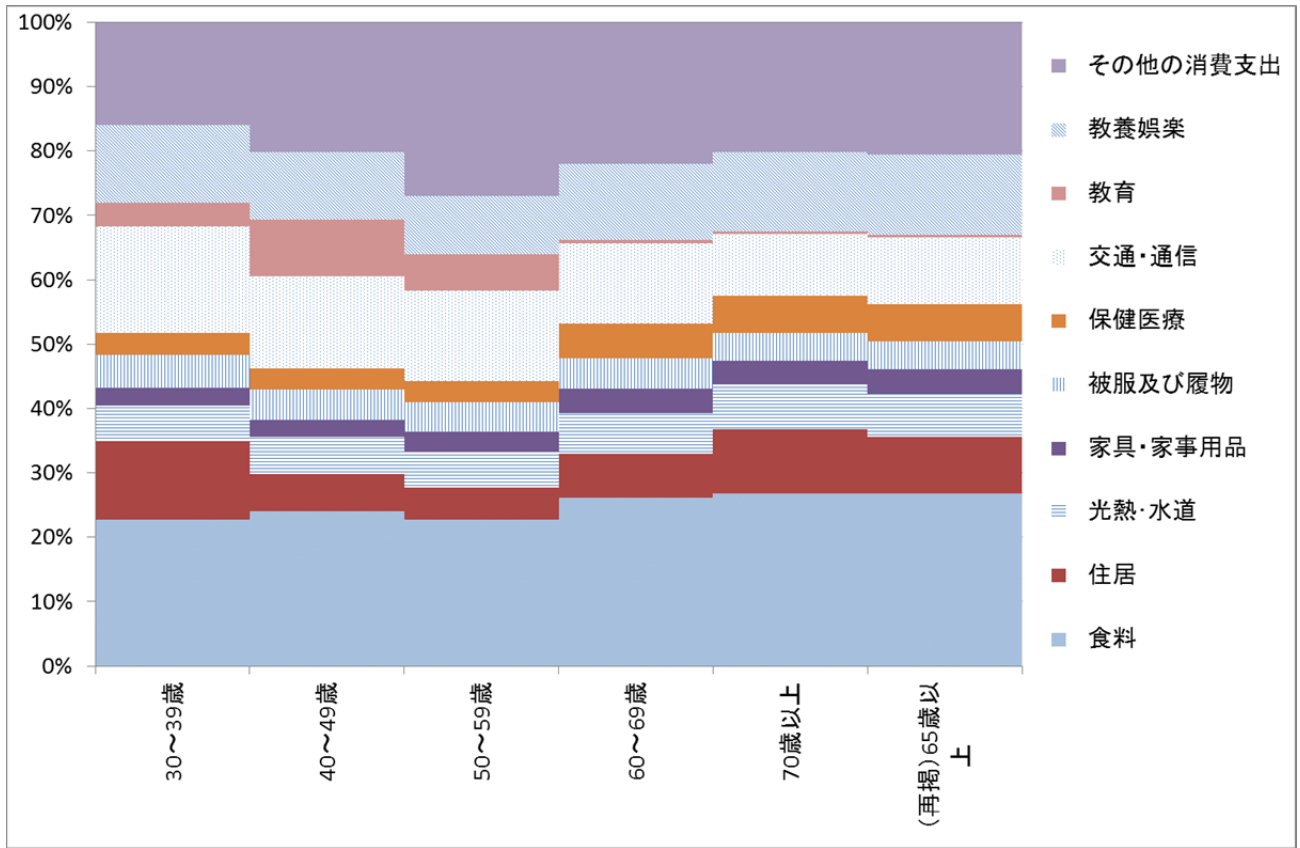
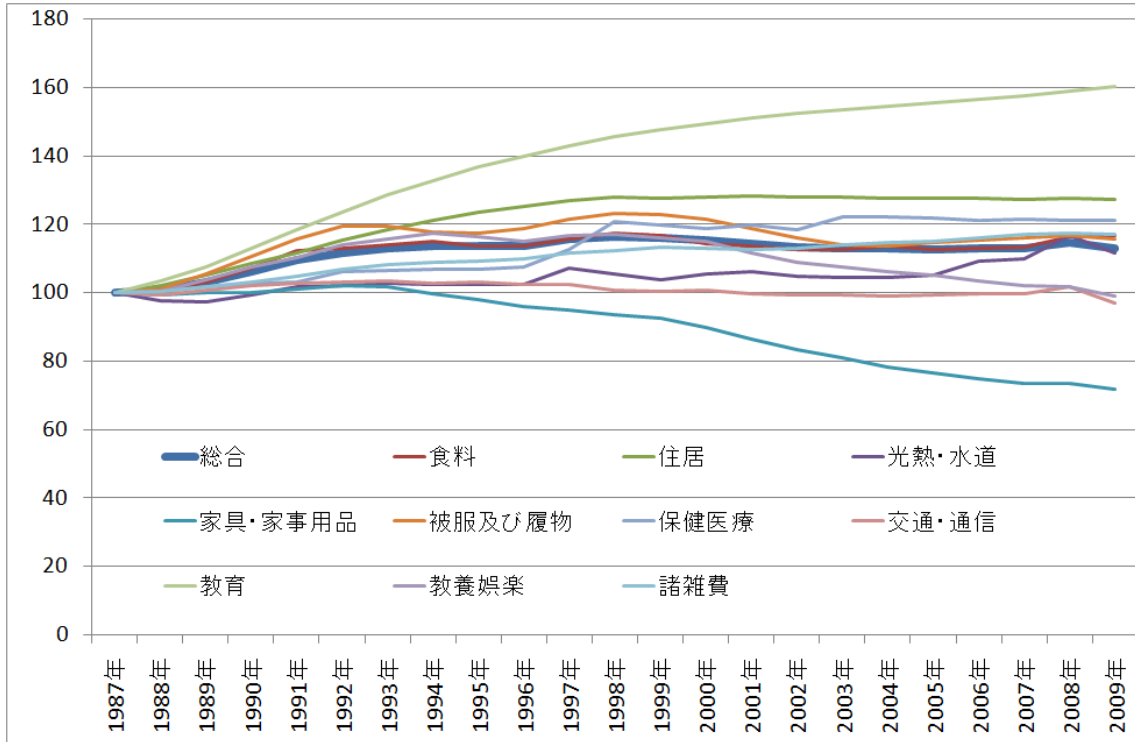


図2 財別の消費者物価指数の動向(2005年基準)



(消費者物価指数年報より筆者作成・1987年=100に換算)

表1 2007年 費目別業態別価格指数

		一般小売店	スーパー	量販専門店	コンビニ	デパート	生協	その他
食料	穀類	0.970	0.991	0.899	1.132	1.131	0.974	0.968
	魚介類	1.030	0.965	0.925	1.027	1.351	1.000	1.040
	肉類	1.014	0.988	0.947	0.962	1.276	0.978	0.977
	乳卵類	1.020	0.970	0.865	1.083	1.086	0.997	0.993
	野菜・海藻	0.997	0.991	0.830	1.127	1.351	1.029	0.892
	果物	0.984	0.993	0.934	1.182	1.375	1.041	0.837
	油脂・調味料	1.039	0.983	0.826	1.068	1.264	0.979	1.041
	菓子類	1.058	0.927	0.912	1.013	1.260	0.942	1.061
	調理食品	1.017	0.970	0.972	1.063	1.271	0.976	0.981
	飲料	1.013	0.948	0.880	1.149	1.144	0.954	1.092
酒類	1.014	0.962	0.931	1.061	1.088	0.954	1.033	
光熱・水道	ガス代	1.003	1.000	0.982	1.000	1.000	0.985	0.998
家具・家事用品	家庭用耐久財	1.022	0.953	0.974	1.000	1.090	0.984	0.907
	室内装備品	1.158	0.695	0.595	1.000	1.768	0.801	0.841
	寝具類	0.995	0.774	0.661	1.000	2.257	0.854	1.471
	家事雑貨	0.965	0.873	0.862	1.053	1.699	0.850	0.967
	家事用消耗品	0.991	1.005	0.946	1.030	1.104	1.005	1.045
被服および履物	洋服	1.195	0.521	0.605	1.000	1.867	0.633	0.895
	シャツ・セーター類	1.135	0.541	0.495	0.738	2.329	0.600	0.774
	下着類	1.062	0.864	0.760	1.085	1.704	0.889	0.844
	履物	1.248	0.706	0.533	1.000	1.548	0.652	0.605
	他の被服類	1.151	0.975	0.657	0.885	2.039	0.762	0.783
保健医療	医薬品・健康保持用摂取品	1.001	0.988	0.991	1.003	1.063	1.012	1.015
	保健医療用品・器具	0.953	0.821	0.817	1.033	1.265	0.925	1.026
交通・通信	自動車等関係費	1.000	0.996	0.996	1.000	1.002	0.996	0.979
教養娯楽	教養娯楽用耐久財	1.020	1.006	0.976	1.000	1.039	1.020	1.000
	教養娯楽用品	1.011	0.944	0.870	1.059	1.278	0.997	0.981
	教養娯楽サービス	1.109	0.924	1.022	1.020	1.513	0.767	1.058
諸雑費	理美容用品	0.991	1.018	0.970	1.013	1.056	1.020	1.023
	身の回り用品	1.107	0.843	0.799	0.958	1.576	1.178	1.233

表 2 年齢階級別の業態別支出シェア

大分類	中分類		一般小売店	スーパー	量販専門店	コンビニ	デパート	生協	その他
食料	(酒類、外食を除いたすべて)	平均	0.163	0.560	0.025	0.041	0.051	0.091	0.069
		30歳未満	0.128	0.575	0.034	0.134	0.033	0.054	0.042
		30～39歳	0.117	0.597	0.036	0.067	0.036	0.101	0.045
		40～49歳	0.121	0.601	0.030	0.048	0.033	0.118	0.049
		50～59歳	0.153	0.576	0.025	0.030	0.046	0.099	0.071
		60～69歳	0.190	0.538	0.020	0.021	0.064	0.081	0.085
		70歳以上	0.222	0.512	0.017	0.021	0.077	0.068	0.083
		65歳以上	0.213	0.521	0.018	0.020	0.074	0.070	0.085
家事家具用品	家庭用耐久財	平均	0.359	0.057	0.383	0.000	0.047	0.022	0.131
		30歳未満	0.261	0.038	0.524	0.000	0.065	0.008	0.104
		30～39歳	0.190	0.052	0.495	0.000	0.059	0.007	0.198
		40～49歳	0.301	0.058	0.431	0.000	0.031	0.020	0.159
		50～59歳	0.375	0.067	0.391	0.000	0.042	0.025	0.100
		60～69歳	0.379	0.049	0.342	0.000	0.062	0.032	0.135
		70歳以上	0.486	0.063	0.293	0.000	0.034	0.013	0.112
		65歳以上	0.453	0.059	0.319	0.000	0.039	0.026	0.105
教養娯楽	教養娯楽用耐久財	平均	0.289	0.026	0.534	0.000	0.016	0.007	0.127
		30歳未満	0.165	0.028	0.615	0.000	0.017	0.031	0.145
		30～39歳	0.242	0.017	0.507	0.000	0.028	0.002	0.203
		40～49歳	0.263	0.017	0.598	0.000	0.006	0.003	0.113
		50～59歳	0.273	0.031	0.577	0.000	0.014	0.007	0.097
		60～69歳	0.362	0.026	0.487	0.000	0.017	0.001	0.106
		70歳以上	0.434	0.050	0.409	0.000	0.014	0.022	0.070
		65歳以上	0.456	0.047	0.403	0.001	0.022	0.014	0.059

表 3-1 財のウエイトと購買行動を考慮した年齢階級別物価指数

	年齢別・財ウエイト + 年齢別・購入先業態別ウエイト (1987年=100)							
	平均	30歳未満	30～39歳	40～49歳	50～59歳	60～69歳	70歳以上	65歳以上
1992年	108.9	108.7	108.0	108.8	108.8	109.2	109.9	109.7
1997年	111.0	110.8	109.3	110.7	110.9	111.3	113.0	112.5
2002年	107.8	108.0	105.9	107.5	107.8	108.0	109.9	109.4
2007年	108.5	108.4	106.1	107.8	108.7	109.1	111.0	110.5

表 4-1 財ウエイトのみを年齢別にした場合の物価指数の乖離幅

	年齢別・財ウエイト + 全年齢平均・購入先業態別ウエイト (1987年=100)						
	30歳未満	30～39歳	40～49歳	50～59歳	60～69歳	70歳以上	65歳以上
1992年	-0.04	-0.76	-0.07	-0.09	0.18	0.84	0.66
1997年	-0.35	-1.59	-0.23	-0.04	0.26	1.87	1.40
2002年	0.02	-1.80	-0.05	0.10	0.08	1.73	1.27
2007年	-0.35	-2.25	-0.17	0.24	0.35	1.87	1.49

注) 財ウエイトを各年齢ごとのものを使った場合の乖離幅(パーセントポイント)

表 4-2 購買行動のみを年齢別にした場合の物価指数の乖離幅

	全年齢平均・財ウエイト + 年齢別・購入先業態別ウエイト (1987年=100)						
	30歳未満	30～39歳	40～49歳	50～59歳	60～69歳	70歳以上	65歳以上
1992年	-0.10	-0.11	-0.06	0.00	0.10	0.19	0.18
1997年	0.14	-0.11	-0.10	-0.03	0.02	0.15	0.13
2002年	0.10	-0.10	-0.25	-0.07	0.11	0.36	0.27
2007年	0.26	-0.17	-0.54	-0.08	0.20	0.64	0.49

注) 購入先業態別ウエイトを各年齢ごとのものを使った場合の乖離幅(パーセントポイント)



## 高齢者世帯の消費行動と物価指数

宇南山 卓(財務省財務総合政策研究所)

慶田 昌之(立正大学)

## 年齢別の物価の比較

- 年齢別の物価指数の構築:
  - 年齢別の消費する財の違い
  - 年齢別の購入先の違い
    - 店舗形態による価格の違い
    - 年齢による購入先店舗形態の違い
  - = 全国消費実態調査・全国物価統計調査で店舗形態別の物価が観察可能
- 年齢別物価指数の重要性
  - 公的年金の実質化に不可欠
  - 年齢別の経済厚生 of 計測
  - 学術的な関心
    - Amble and Stewart (1994): 高齢者ほど高い物価上昇率に直面
    - Goda, Shoven, and Slavov (2011) 医療費について高齢者ほど物価上昇率が高い

## 真の物価指数と定義と計測

- 家計の支出最小化問題

$$X = e(p, \bar{u}; \delta)$$

- X: 家計支出
- p: 価格ベクトル
- $\bar{u}$ : 効用水準
- $\delta$ : "taste" パラメータ

- 真の物価指数の定義

$$P_1 = \frac{e(p_1, \bar{u}; \delta)}{e(p_0, \bar{u}; \delta)}$$

- 物価指数とは、異なる物価ベクトルの下で、同じ効用水準を得るのに必要な支出額の比率
  - 物価指数には必ず「基準」が存在する
- 物価指数の性質については、例えば Diewert (1981)を参照

## ここでやっていないこと: 選好と物価

- 概念的には「同一価格ベクトルのもとで、異なる選好を持つ家計が、同じ効用水準を得るのに必要な支出額」によって直面する物価の比較が可能であるように見える

$$R = \frac{e(p, \bar{u}; \delta_1)}{e(p, \bar{u}; \delta_0)}$$

- 選好が異なるため、同じ価格ベクトル(すべての財の価格が同一)であっても、直面する「物価」が異なる可能性がある
  - たとえば、高齢者と若者では、同じ支出でどちらがより高い効用が得られるか？
- しかし、この比較は理論的にも実証的にも不可能
  - 理論的には、序数的効用の仮定より個人間での効用水準の比較は不可能
  - (基数的効用を仮定したとしても)実証的に どの識別は不可能
    - 消費行動が異なっても、その理由が「選好が異なる」からなのか「効用水準が異なる」からなのかは区別できない

## 選好とインフレ率

- 物価水準ではなく、物価の上昇率であれば、選好が異なる個人間で比較が可能

$$P_1^k \equiv \frac{e(p_1, \bar{u}; \delta_k)}{e(p_0, \bar{u}; \delta_k)}$$

- 選好のパラメータは固定
- もし  $P_1^k > P_1^l$  であれば
  - 選好が“ $k$ ”であるような家計は、選好が“ $l$ ”である家計より直面するインフレ率が低い
  - 選好が“ $k$ ”であるような家計が高い物価水準に直面しているとは限らない

## 年齢別の物価指数の構築

- 構築されるべき真の物価指数は

$$P_1^k \equiv \frac{e(p_1, \bar{u}; \delta_k)}{e(p_0, \bar{u}; \delta_k)}$$

- 需要システムの推計が必要
- 実務的に標準的な物価指数の作成方法
  - Laspeyres 指数の構築
  - 真の物価指数の上限となる

$$CPI_1^k \equiv \frac{p_1 \cdot q_0^k}{p_0 \cdot q_0^k} \geq \frac{e(p_1, \bar{u}; \delta_k)}{e(p_0, \bar{u}; \delta_k)}$$

## ここでやったこと

- 以下の物価指数を構築:

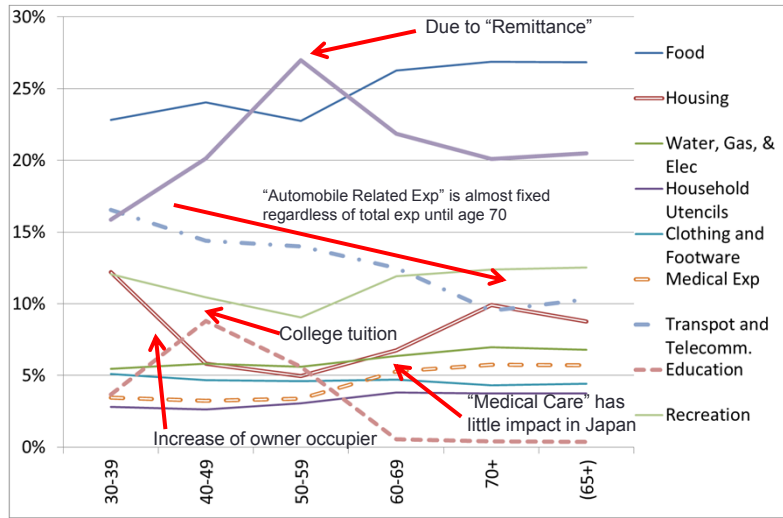
$$CPI_t^k \equiv \sum_i w_{i0}^k \left( \sum_s w_{is0}^k \frac{p_{ist}}{p_{is0}} \right)$$

- 価格データ
  - 小売物価統計調査
    - 全業態の平均的な価格の動向
    - CPIの元データ
    - 月次統計
  - 全国物価統計調査
    - 業態間の価格差をとらえることのできる統計**
    - 5年に一度 (1987, 1992, 1997, 2002, and 2007→廃止)
- ウエイトデータ
  - 全国消費実態調査
    - 財別・業態別・年齢別の支出額が分かる
  - 5年に一度 (ここでは2004年のデータを利用)

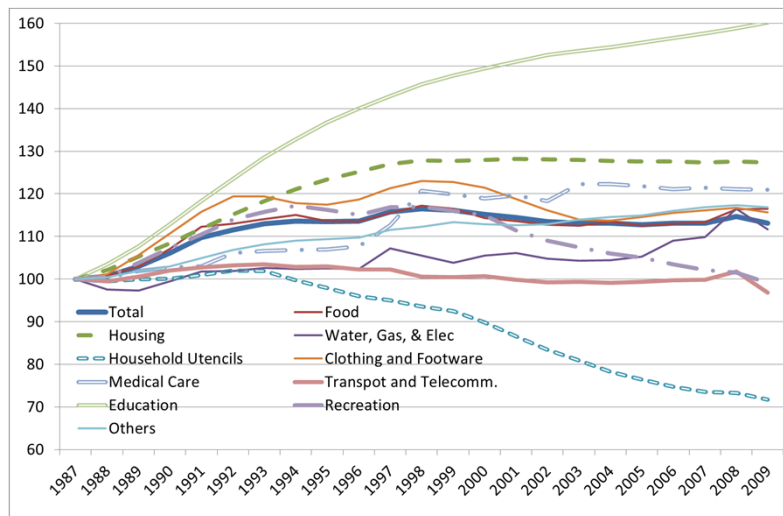
## 年齢階層別の「嗜好」の違い

- 高齢者と若年者では消費行動が大きく異なる.
  - 高齢者は若年者とは異なる財を購入する.
    - 価格は財によって異なる動きをする
    - 通常は「嗜好」とは個々の違いを指す
      - 総務省統計局も「年齢階級別」の物価指数を公表している
  - 高齢者は若年者と異なる店舗で購入する.**
    - 店舗によって価格の動きは異なる.
    - この論文の貢献!**
- 年齢別・財別・購入先業態別の支出額・価格動向を把握
  - 日本のデータであれば可能

## 年齢による消費パターンの違い



## 財別のインフレ率



## 購入先業態別に関する選好

- 業態によって価格水準が異なる
  - Scannerデータを用いた検証(for example, Abe and Tonogi, 2007)
  - ただし小売店舗全体に対する代表性はない
- 異なる業態は異なるサービスを提供.
  - 一般小売店(Mom-Pop store): カスタマイズされたサービス
  - スーパー: 食料品については安価で画一的なサービス
  - ディスカウントストア: 特定の品目が安価だが市街立地も多い
  - 百貨店: 多種多様な商品を高い品質で提供
  - コンビニ: 営業時間が長いが一般に割高
  - 生協: メンバーシップが必要だが比較的安価で宅配などもある
- 消費者は価格とサービスを比較して購入先を選択
  - 基本的には異なる業態での商品は「異なる財」と考えるべき

## 年齢による購入先業態の違い

They sometime offer delivery service

		Retail Store	Supermarket	Discount Store	Convenience Store	Department Store	Cooperative Store	Others
Food (Excl. "Alcohol" and "Eating Out")	Ave	16%	56%	3%	4%	5%	9%	7%
	30-	13%	58%	3%	13%	3%	5%	4%
	30-39	12%	60%	4%	7%	4%	10%	5%
	40-49	12%	60%	3%	5%	3%	12%	5%
	50-59	15%	58%	3%	3%	5%	10%	7%
	60-69	19%	54%	2%	2%	6%	8%	8%
	70+	22%	51%	2%	2%	8%	7%	8%
	(65+)	21%	52%	2%	2%	7%	7%	8%
Recreational Durable Goods	Ave	29%	3%	53%	0%	2%	1%	13%
	30-	17%	3%	62%	0%	2%	3%	15%
	30-39	24%	2%	51%	0%	3%	0%	20%
	40-49	26%	2%	60%	0%	1%	0%	11%
	50-59	27%	3%	58%	0%	1%	1%	10%
	60-69	36%	3%	49%	0%	2%	0%	11%
	70+	43%	5%	41%	0%	1%	2%	7%
	(65+)	46%	5%	40%	0%	2%	1%	6%

Don't like a "mass sales" outlet?  
 Prefer more "traditional" or customized outlet?

Online Shopping?

## 業態による価格水準の違い

- 業態による価格水準の違い

	Retail Store	Supermarket	Discount Store	Convenience Store	Department Store	Cooperative Store	Others
Meat	101.4	98.8	94.7	96.2	127.6	97.8	97.7
Snacks	105.8	92.7	91.2	101.3	126.0	94.2	106.1
Kitchen Items	99.1	100.5	94.6	103.0	110.4	100.5	104.5
Household Appliances	102.2	95.3	97.4	NA	109.0	98.4	90.7
Recreational Durables	102.0	100.6	97.6	NA	103.9	102.0	NA
Shampoo and Cosmetics	99.1	101.8	97.0	101.3	105.6	102.0	NA

↑ "Cheapest" outlet
↑ "Most expensive" outlet

- ここでは「固定ウェイト」指数を構築しており、業態ごとのサービスに変化がなければ、サービスが異なること自体は問題にならない
- 業態による価格水準ではなく「価格動向」が重要

## 業態によるインフレ率の違い

- どの業態が最も価格が上昇したか？

$$P_{st} \equiv \sum_i \bar{w}_{i0} \frac{p_{ist}}{p_{is0}}$$

	1992	1997	2002	2007
Ave.	108.9	111.0	107.8	108.5
Retail Store	110.6	111.7	108.4	109.5
Supermarket	110.0	111.2	106.0	105.0
Discount Store	105.1	108.3	104.4	100.8
Convenience Store	107.5	112.3	108.6	108.3
Department Store	110.5	113.0	112.7	119.8
Coop	104.5	109.4	107.2	104.9
Others	108.0	107.5	104.3	102.1

## メインの結果：年齢別の物価指数

	Age Specific Goods-Outlet Weights (1987=100)							
	Ave	30-	30-39	40-49	50-59	60-69	70+	(65+)
1992	108.9	108.7	108.0	108.8	108.8	109.2	109.9	109.7
1997	111.0	110.8	109.3	110.7	110.9	111.3	113.0	112.5
2002	107.8	108.0	105.9	107.5	107.8	108.0	109.9	109.4
2007	108.5	108.4	106.1	107.8	108.7	109.1	111.0	110.5

2 percent point more  
(0.2 percent point per year between 1987-97)

The older experienced slightly smaller deflation

## 直面するインフレ率の差の要因分解

- 構築される指数

$$CPI_t^k \equiv \sum_i w_{i0}^k \left( \sum_s w_{iso}^k \frac{p_{ist}}{p_{is0}} \right)$$

- 消費される財だけが異なる場合の指数

$$\widehat{CPI}_t^k \equiv \sum_i w_{i0}^k \left( \sum_s \bar{w}_{iso}^k \frac{p_{ist}}{p_{is0}} \right)$$

- 購入する業態は年齢によって差がないケース
- 購入する業態だけが異なる場合の指数

$$\widetilde{CPI}_t^k \equiv \sum_i \bar{w}_{i0}^k \left( \sum_s w_{iso}^k \frac{p_{ist}}{p_{is0}} \right)$$

- 消費する財は年齢によって差がないケース



## 要因分解の結果

Differences are mainly caused by consumption pattern

	Age Specific Goods Weight (Deviation from the Average)						
	30歳未満	30～39歳	40～49歳	50～59歳	60～69歳	70歳以上	65歳以上
1992年	-0.04	-0.76	-0.07	-0.09	0.18	0.84	0.66
1997年	-0.35	-1.59	-0.23	-0.04	0.26	1.87	1.40
2002年	0.02	-1.80	-0.05	0.10	0.08	1.73	1.27
2007年	-0.35	-2.25	-0.17	0.24	0.35	1.87	1.49

	Age Specific Outlet Weight (Deviation from the Average)						
	30歳未満	30～39歳	40～49歳	50～59歳	60～69歳	70歳以上	65歳以上
1992年	-0.10	-0.11	-0.06	0.00	0.10	0.19	0.18
1997年	0.14	-0.11	-0.10	-0.03	0.02	0.15	0.13
2002年	0.10	-0.10	-0.25	-0.07	0.11	0.36	0.27
2007年	0.26	-0.17	-0.54	-0.08	0.20	0.64	0.49

The outlet that the older prefer experienced higher inflation rate

## 結論

- 高齢世帯ほど直面したインフレ率が高かった
  - 1987-97の年平均で0.2パーセントポイントほど公式のCPIよりも高い
- 乖離の75%程度は消費する品目の違いで説明できる
  - 高齢者は「住居」および「保健医療」に多く支出をしている
  - ただし、これは「バイアス」ではなく経済構造の変化による結果
- 高齢者が利用する購入先形態はより大きく物価が上昇した
  - 一般小売店や百貨店などが相対的に物価上昇率が高かった
  - 若者の支出が多いコンビニは相対的に物価上昇率は低かった

## 補足

- Boskin et al. (1996)が指摘するCPIのバイアスについては考慮していない
  - Laspeyres 指数は真の物価上昇率を過大評価
  - バイアスの大きさが年齢によって異なる可能性はある
- 教育費と保健医療は、物価上昇率が高い
  - 年齢によって支出シェアが大きく異なる財でもある
  - 品質調整が困難な財でもある
  - 品質調整の影響は年齢によって異なる

# エビデンスに基づいた匿名化

星野 伸明\*

平成 26 年 1 月 21 日

## Evidence Based Anonymization

Nobuaki Hoshino\*

### 概要

匿名データや個人情報、は、個体識別が可能か否かで法律上区別される。しかしこの区別の方法は不明確で、改善のための明示的議論の対象になっていない。従って本論文は、個体識別可能性の判定方法を明確化する。このような判定に関する既存研究は、個体識別可能性の定量評価について閾値を定める理論を欠く。この点について本論文では、個体識別が起きていないという観測可能な事実に基づいて閾値を推定する。また部分的にしか観測されず定量評価できない情報も、事例間で等しいか否かという判断しやすい方法で利用する。このような観測に基づいて意思決定する態度は、エビデンスに基づいた匿名化と呼ぶのがふさわしい。この立場から、匿名データ審査体制の改善点が指摘できる。

Japan Law discriminates Anonymized Data or personal information by discerning that a specific individual is identifiable. The state of being identifiable, however, is not defined, and thus we can not explicitly improve the evaluation of identifiability. Therefore the present article explicates a method to decide whether given data are identifiable or not. The existing literature on this issue lacks the theory of deciding the critical value of measured identifiability; we estimate it based on a fact that identification has not been observed. Also partially observed factors are compared in our decision, which is called evidence based anonymization. This theory leads to institutional improvements on Anonymized Data.

キーワード: 母集団一意, プライバシー, 統計的開示制限.

---

\*金沢大学経済学類, 〒 920-0927, 石川県金沢市角間町, E-mail: hoshino@kenroku.kanazawa-u.ac.jp

# 1 はじめに

匿名データは、平成 21 年度に四調査（全国消費実態調査、社会生活基本調査、就業構造基本調査、住宅・土地統計調査）から提供が開始された。平成 23 年度末現在、国民生活基礎調査や労働力調査の匿名データ提供も決まっている。新しい制度がこのように実績を重ねてきたことは喜ばしい。ただ今後は実績という経験を活かし、制度を継続的に改善する道筋をつけるべきである。特に利用者からのデータ改善要求に応える必要がある。

匿名データは元の個票を変換（匿名化）して作られる。例えば全国消費実態調査等の匿名データでは、15 歳から 84 歳までの年齢を 5 歳階級別に変換している。また地域情報は「3 大都市圏」及び「その他の地域」の 2 区分に変換している。このような変換により、各歳別の分析や詳細な地域別分析は不可能となる。データ分析において、匿名化は明らかに望ましくない。故に匿名化の緩和は利用者の典型的な要求である。

しかし全ての匿名化を外せるわけではない。匿名データの定義（統計法第 2 条第 12 項）を引用すると、「一般の利用に供することを目的として調査票情報を特定の個人又は法人その他の団体の識別（他の情報との照合による識別を含む。）ができないように加工したもの」である。元の個票（調査票情報）はこの定義を満たすように匿名化されなければ、匿名データとして提供不可能<sup>1</sup>である。従って匿名化は、個体識別が不可能な範囲で少ない方がよい。つまり匿名データの改善の多くは、個体識別が可能か否かという判断を必要とする。

この判断について、総務省政策統括官（統計基準担当）（2011）による「匿名データの作成・提供に係るガイドライン」（以下、ガイドライン）には、審査用資料として「チェックリスト」を作成することが定められている。そして「チェックリストに記載された内容等を基に」、「匿名化処理の妥当性等に係る審査を実施する」とある。参考として世帯調査のチェックリスト（H23/3/28 改正版）の要約を付録 A に収めた。チェックリストは個体識別に関係する要因を記載しているはずである。しかしその使い方は説明されていない。

結局「一律に匿名化の基準を設定することは困難」なので「一橋大学における匿名標本データの試行的提供の事例<sup>2</sup>及び諸外国の統計機関における同様の提供の事例等を参考に」匿名化せよとガイドラインは書く。同様とはどのような事例で、それをいかに参考にしたらよいか。この点についての判断は審査担当者の見識に委ねられている。個体識別可能と不可能の区別は、不明確である。

この区別の明確化、精密化は匿名データに関してだけの課題ではない。いわゆる個人情報保護法において個人情報とは「生存する個人に関する情報であつて、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの（他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む。）」と定義される（第 2 条）。このように個体識別が可能か否かを区分の基準とする例は外国法<sup>3</sup>でも見られる。いか

<sup>1</sup> 本論文において統計法改正は手段として除外する。しかし個体識別性をデータ提供の基準とするのは必ずしも望ましくない。例えば個体識別されてもデータが悪用されなければよいという主張は妥当かもしれない。

<sup>2</sup> 試行的提供の詳細については山口（2008）を見よ。

<sup>3</sup> 例えば U.S. Privacy Act など。U.S. Office of Federal Statistical Policy and Standards（1978, pp.3-5）の解説を見よ。

に匿名化すれば個体識別が不可能かという問題は普遍的である。

ところがこの基本的な問題がないがしろにされている。匿名化についての多くの研究は個体識別の危険性（開示リスク）の測り方は定める。しかし開示リスクの目標値について、せいぜい危険選好に応じてデータの分析価値（有用性）とバランスさせよ<sup>4</sup> としか言わない。つまり現実の匿名化が許容範囲か否かは、実務家の価値判断の問題とされる。このような態度は匿名化の技術開発には好都合である。しかし結果として実務家は途方に暮れ、判断を何らかの権威に投げているのが現状であろう。

現行法の下での匿名化の実務的判断について、客観化の余地が残っていると著者は考える。本論文では諸要因と個体識別行為の関係をモデル化し、個体識別が可能という状態を要因と関係づける。このようなモデル化は、開示リスクを具体的に定めることに他ならない。従って、評価される開示リスクの目標値が問題となる。この点について本論文は、個体識別が不可能な状態を過去の事例を基に決める方法を提案する。過去に公開されたデータについて個体識別が観測されていないとすれば、その事実は個体識別不可能ということについて情報を持っている。故に個体識別が観測されること確率モデルを構成し、既公開の匿名データ等を匿名化の程度判断についての統計的証拠に転ずる。

このような理論なくして、明確な個体識別性の審査はあり得ない。そもそもチェックリストの記載事項は、個体識別と理論的に関係する要因であるべきだ。またリストの使い方も理論が定める。それから本論文の理論は観測結果と関係を持ち、実証の対象である。いかなる理論も実証を経ることで継続的に改善される。従って本論文は、個体識別に関する統計的証拠—エビデンスに基づいた匿名化 (Evidence Based Anonymization, EBA) を主張する。

本論文の構成は以下の通りである。2章は全体で、個体識別が可能か否かの判定方式を明らかにする。まず2.1節において、個体識別の観測と可能性の関係を確率モデルで表す。次に2.2節では、個体識別行為を所与の要因についての確率モデルで表す。2.3節では、個体識別の要因について計量可能な方式を考察する。ここまでの議論で、情報不足により定性評価で満足せざるをえない個体識別の要因があることになる。2.4節ではそのような要因を明らかにする。2.5節では、個体識別の行為と観測の関係について考察する。3章では2章の理論を基に、匿名データ審査体制の改善点を指摘する。

## 2 個体識別可能性の判定方式

### 2.1 個体識別の観測モデル

個体識別が可能か否かは、明らかにデータの表現に依存する。ここで匿名化による表現の変化は滑らかだが、個体識別が不可能と可能の差は不連続である。これをモデル化する場合、データ表現の適当な実数特性値が閾値を超えれば個体識別が可能とみなすのが定石であろう。本論文で

---

<sup>4</sup>例えば Duncan et al. (2001) や Domingo-Ferrer and Torra (2001) など。伊藤 (2012) のサーベイを参照せよ。

もこのように考え、個体識別の難易度とみなせる特性値に注目する。この難易度が閾値より高ければ個体識別が不可能とみなすのは自然<sup>5</sup>である。

具体的な難易度測度の設計については後で考察するとして、モデル化をすすめよう。難易度  $\delta$  を引数とする関数  $f$  は個体識別が可能なら 1、不可能なら 0 を返すとする。つまり閾値が  $\alpha$  として

$$f(\delta) = \begin{cases} 1 & \delta < \alpha \text{ の場合} \\ 0 & \delta \geq \alpha \text{ の場合} \end{cases} \quad (1)$$

ということになる。もし所与のファイルが個体識別可能か否か判定したいなら、その難易度を求めて  $\alpha$  と大小を比較すればよい。ただ  $\alpha$  が既知となるような難易度測度の設計は難しい。後述されるように、事実上観測可能でない個体識別の要因が残るはずだ。では  $\alpha$  が未知の場合にどうすべきか。

まず  $\alpha$  が推定可能か考えてみよう。統計的に未知母数  $\alpha$  を推定するには、観測値が必要になる。しかし個体識別が可能か否かは、観測されることではない。観測可能な事実は、個体識別が起きたか否かである。モデルを用いて説明しよう。確率変数  $X$  が 1 なら個体識別が観測され、0 なら観測されないこととする。個体識別が不可能なら必ず  $X = 0$  である。個体識別が可能の場合、難易度  $\delta$  に依存する確率  $p(\delta)$  で識別が観測されると考えよう。すなわち  $\Pr(X = 1; \delta < \alpha) = p(\delta)$ ,  $\Pr(X = 0; \delta < \alpha) = 1 - p(\delta)$  とする。個体識別は決して起きないと考えれば危機管理にならないので、 $p(\delta)$  は正と想定するべきだ。

このような状況で閾値が共通する  $n$  件の事例が存在するとしよう。 $i, i = 1, 2, \dots, n$ , 番目について観測されるのは、少なくとも難易度  $\delta_i$  と識別の有無  $x_i$  である。単純化のため  $\delta_1 < \delta_2 < \dots < \delta_n$  としよう。この中で個体識別が観測された ( $x_i = 1$  となる  $i$  が存在する) 場合は  $\delta_i < \alpha$  と分かる上、実務的に重要でない。故に個体識別がこれまで<sup>6</sup> 起きていない (全ての  $i$  について  $x_i = 0$ ) として考察を続ける。この場合モデルの尤度  $\ell$  は、 $\delta_i < \alpha \leq \delta_{i+1}$  の時  $\ell(\alpha) = \prod_{j=1}^i (1 - p(\delta_j))$  となる。そして全ての  $\delta$  について  $0 < p(\delta) < 1$  なら、 $\alpha$  の最尤推定値  $\hat{\alpha}$  は  $\delta_1$  以下である。つまり過去の事例で個体識別が観測されていなければ、その最も低い難易度  $\delta_1$  以下と閾値  $\alpha$  は推定される。このように  $\hat{\alpha}$  が一意に定まらないのは、情報が無いので区別出来ないことを意味する。

情報を補うため  $\alpha$  の事前分布を用いることは考えられるが、それよりも観測情報の増加を工夫する方が健全だろう。つまり  $X$  を  $\{0, 1\}$  の二値とするのではなく、 $\alpha$  と  $\delta$  の距離に依存する連続量と出来ればよい。これは治験薬の用量反応関係の推測と考え方が同じである。薬の臨床試験では、人体に決定的な悪影響を及ぼしてはならない。このような制約下では、薬剤の投与を少量から始めて徐々に増やし、危険な兆候が見られれば中止する。生死の二値ではなく、投与量が死亡の閾値と近いことを示す兆候 (心拍や呼吸の異常等) を観測するのである。

<sup>5</sup>一般目的汎用ファイル (Public Use File, PUF) の作成においても整合的な考え方である。星野 (2010) の考察の通り、PUF は匿名データよりも強い匿名性を必要とする。しかし統計法には個体識別が可能か不可能かという二区分しか存在せず、PUF と匿名データを区別できない。ところが個体識別の難易度という概念を用いれば、個体識別が不可能という状態の中で匿名データと PUF を区別出来る。

<sup>6</sup>データの公開直後に識別が起きなくても、ある程度後で識別が起きることはあり得る。閾値の推定をする時点に依存して各  $p(\delta_i)$  は変化するかもしれない。しかし  $0 < p(\delta_i) < 1$  なら  $\hat{\alpha} \leq \delta_1$  という結論は変わらない。

問題は、 $\alpha$  と  $\delta$  の距離に依存する観測可能な事象として何をを用いるかである。例えば個体を識別できたと誰かが誤って主張することなどが、個体識別発生の兆候として考えられる。この場合、個体識別を試す気にさえならない水準よりは難易度が下がっていることがわかる。一般に1件の重大事故の陰には300件のヒヤリ・ハットが起きているという。匿名データ提供に対する社会的反応が、警鐘になる可能性はある。また攻撃者の動機を考えることで、ある程度  $p(\delta)$  を定めることができるかもしれない。2.5節でそのような考察を部分的に行うが、実務への反映は拙速と思われる。

では現時点で、新しく公開するファイルの難易度  $\delta_{n+1}$  をどのように決めたらよいか。一つの考え方は  $\delta_{n+1} = \delta_1$  とすることだろう。強い仮定を置かずに推測出来るのは、難易度の閾値が  $\delta_1$  以下ということまでである。これは  $\delta_1$  未満の難易度について個体識別が不可能な証拠がないということだ。個体識別が可能になるという過誤の可能性<sup>7</sup>を考えれば、慎重な判断は正当化されるだろう。また  $\delta_{n+1} = \delta_1$  として難易度  $\delta_1$  における観測が蓄積されることは、将来的に意味を持つ。同難易度の複数のファイルについて個体識別が観測されなければ、その難易度が個体識別不可能な確率は高まる。また観測情報が増えないと、閾値との近さについて確かな判断は出来ない。

このように個体識別の難易度という概念を用いれば、個体識別が可能か否かの判断において過去の事例を統合して利用できる。しかし例えば国や公開時期が違う事例において、個体識別が可能となる難易度の閾値  $\alpha$  は同じだろうか。

閾値  $\alpha$  が共通する事例の範囲は  $\delta$  の具体型に依存する。個体識別に関する要因を全て勘定する理想的な  $\delta$  を用いる場合、全事例で閾値が共通するとみなしてよい。逆に閾値の変動は、 $\delta$  が考慮しない要因の変化から生ずる。データ表現の実数特性値として導入した  $\delta$  だが、その他の要因を算出に用いるべきかもしれない。以下では良い  $\delta$  の構成を考察しよう。

## 2.2 個体識別行為の確率モデル

前節では個体識別の難易度に依存して個体識別が確率的に観測されるモデルを考察した。具体的に個体識別の難易度を定めるには、個体識別行為をモデル化する必要がある。本節ではそのようなモデルを構成し、個体識別の難易度の定式化をすすめる。

個体識別行為の確率モデルに関する先行研究としては、英国国勢調査匿名化標本の開示リスクを評価した Marsh et al. (1991)、及びこの論文を再考した Dale and Elliot (2001) が挙げられる。Marsh 等は個体識別が起きる条件を具体的に挙げ、それらが満たされる確率を個別に評価することで、個体識別が起きる確率を計算しようとした。まずこの試みを検討しよう。

最もあり得る個体識別の形態は、識別を試みる者(「攻撃者」)が素性を知る個体を(匿名化された)公開ファイルの中に見つけることだと言われている。このような行為を竹村(1997)は「順攻撃」と呼ぶ。一方、公開ファイル中の特定個体を母集団に探す行為は「逆攻撃」と呼ばれる。この区別が意味を持つのは、攻撃者が探して素性を知ることが出来る個体群が、(攻撃用ファイルの)

<sup>7</sup>個体識別が可能であるにも関わらず識別が起きない ( $p(\delta) < 1$ ) と過誤が生ずる。故に過誤の可能性を減らすには、 $\delta_1$  の例だけでも攻撃実験をするなどして識別が可能が確認することも役立つ。2.5節の議論も参照のこと。

素性を知る個体群と異なる場合である。これがどのような場合に異なるかは後で考察するとして、順攻撃の考察を続ける。

既知の個体について攻撃者が知る属性(「キー変数」)を並べたファイルを「攻撃用ファイル」と呼べば、「攻撃」とは公開ファイルと攻撃用ファイルでキー変数が同じレコードを探すことと言える。しかしそのような個体が見つかったとしても、それは母集団に複数存在する属性が同じ個体のうちの一でしかないかもしれない。故に統計当局は、母集団に一しか存在しない個体(「母集団一意<sup>8</sup>」)を公開ファイルの中に攻撃者が発見することを警戒しなければならないとされている。

このような背景の下、Marsh 等は個体識別が実際に起きる確率を以下のように分解する。

$$\Pr(\text{識別が実際に起きる}) = \Pr(\text{識別が起きる} \mid \text{識別を試みる}) \Pr(\text{識別を試みる}) \quad (2)$$

更に識別を試みた時にそれが成功する事態は、以下の4つの条件が成立する場合だという。

- (a) 攻撃用ファイルと公開ファイルのキー変数が同じ(時点や分類の)基準で記録されている。
- (b) 公開ファイルに個体が含まれている。
- (c) 個体が母集団一意である。
- (d) 個体が母集団一意と確証出来る。

これらの条件が満たされる事象をそれぞれ  $a$  から  $d$  と書けば

$$\Pr(\text{識別が起きる} \mid \text{識別を試みる}) = \Pr(a) \Pr(b|a) \Pr(c|a, b) \Pr(d|a, b, c) \quad (3)$$

ということになる。右辺の確率を個別に評価できれば、左辺の確率が求められる。

このような分解により、個体識別という漠然とした行為は直感的に解釈できる事象の積となる。(3)式の分解で鍵となる母集団一意概念は、この種の議論では珍しく非専門家でも理解が可能であり、よく知られている。

ただし母集団一意は特殊な匿名化<sup>9</sup>において無意味な場合がある。また普通の匿名化手法だけ用いるとしても、個体識別が論理的に可能なのは母集団一意に限らない。母集団二意の個体も、自レコードが分かるなら、もう片方の個体のレコードが識別できる。そして三意以下でも、個体間で結託すれば識別できる。しかし母集団一意数と二意以下の珍しい個体数は経験的に比例する。故に母集団一意数を管理すれば、二意以下の識別可能性も同時におさえられる。母集団一意は実際の値に意味があるというより、管理対象のリスク測度として分かり易い点が望ましい。

また母集団一意数は、匿名化の程度についてある種の単調性を持つ。主に使われる匿名化手法は「再符号化<sup>10</sup>」と呼ばれ、個体属性を粗く分類する。再符号化で分類を併合してより粗い分類に

<sup>8</sup>公開ファイル中で所与のキー変数の組み合わせの条件を満たす個体数が1の場合、そのような個体は「標本一意」と呼ばれる。標本一意でも母集団一意とは限らないが、母集団一意なら標本一意である。

<sup>9</sup>Hoshino (2009, Section 6) で議論したとおり、匿名化された表現が互いに排他的でないとは母集団一意は一意にならない。通常用いられる大域的再符号化なら、表現は互いに排他的となる。

<sup>10</sup>トップコーディングや削除 (suppression) も再符号化の特殊ケースである。



変換する場合、母集団一意数は非増加である。つまり匿名化が施されて直感的により安全なデータの母集団一意数は、匿名化が施される前のデータの母集団一意数より多いことはない。EBA では匿名化事例に順序をつけるので、このように直感と矛盾しない順序を得られる方法が重要である。

母集団一意数は使用に異論<sup>11</sup>もあるが、以上の議論のような望ましい性質を持つ。Marsh 等の分解を活かして個体識別の難易度を構成出来ないだろうか。

問題は、Marsh 等が分解した要因 ( $\Pr(\text{識別を試みる})$ ,  $\Pr(a)$ ,  $\Pr(b|a)$ ,  $\Pr(c|a, b)$ ,  $\Pr(d|a, b, c)$ ) はそれぞれ評価できるとは限らないことである。まず  $\Pr(\text{識別を試みる})$  の評価は難しいと Marsh 等も認めており、定性的に議論<sup>12</sup>した上で「識別を試みた例を知らないので識別を試みる確率の最良の推定値は経験からゼロ」と述べている。また  $\Pr(d|a, b, c)$  についても分からないので、「非常に多くのキー変数について事前情報が無いはずなのでゼロと信ずるが 0.001 と仮定」している。これでは数値評価が出来ているとは言えない。これらの確率評価は出来るとしても膨大な情報を必要とする。現実的に Marsh 等の方法では、リスク測度 (2) と (3) のいずれも数値評価できない要因を抱える。

もちろん  $\Pr(\text{識別が起きる} | \text{識別を試みる})$  の条件付き確率の積による分解は一意ではない。故にこれを全て評価できる要因の積に書ければ、評価出来ない問題は解決する。しかしそのような分解は不可能であろう。最大の困難は識別が観測されないことである。個票の公開で先行する海外でも識別が起きないように匿名化しているので、ほとんど観測されない事象<sup>13</sup>の確率の推定を強いられる。それにも関わらず、個体識別という事象を細かい要因の積に分解すれば、要因毎に十分な観測数が得られるはずがない。目的の事象の観測に限られる以上、要素の分解に依存したアプローチは、どこかで情報不足の壁に阻まれるであろう。個体識別について数値評価出来ない要因の存在を前提とするべきである。

実は数値評価する要因  $y_1$  と評価しない要因  $y_2$  が分かっても、ある程度は個体識別の難易度を相対比較できる。数値評価した結果  $g(y_1)$  について、個体識別の難易度  $\delta$  が  $h(g(y_1), y_2)$  と書けるとしよう。ここで  $y_2$  は数値評価できないので、難易度関数  $h$  の具体型は分からない。しかし  $h$  は以下のような単調性を持つとする。

$$g(y'_1) \geq g(y_1) \Rightarrow h(g(y'_1), y_2) \geq h(g(y_1), y_2) \quad (4)$$

つまり事例  $(y'_1, y_2)$  と  $(y_1, y_2)$  では、数値評価部が低いほうが難易度が低いということである。単調性 (4) さえ成り立てば、 $y_2$  が共通する複数の事例から、最も個体識別の難易度が低いものを選ぶ。そして 1 節のように  $n + 1$  番目の新しい匿名データを公開するとして、数値評価値  $g$  を  $y_2$  が共通する過去最低の事例に合わせればよい。このように  $g$  の達成目標値は  $y_2$  に依存して決まる。

<sup>11</sup>例えば同じ母集団一意でも、似た個体が居ない方が目立って識別し易いだろう。故に母集団一意のレコードの中で、似た属性の個体が多いか少ないかで開示リスクを変える考え方を「レコードレベルリスク」と呼ぶ。例えばより低次元の周辺分割表で一意になる個体の方が危険とみなす “Special Unique” は比較的計算しやすい (Elliot et al., 1998)。このような議論は一理あるが、リスク管理の対象として複雑な測度は望ましくない。また匿名化の程度についての単調性が崩れるかもしれない。

<sup>12</sup>曰く  $\Pr(\text{識別が起きる} | \text{識別を試みる})$  が減少すれば  $\Pr(\text{識別を試みる})$  も減る。またデータの観測と公開の時点が離れば  $\Pr(\text{識別を試みる})$  も減る、等。

<sup>13</sup>開示制限を失敗して個体識別が可能な例は Sweeney (2002) が報告している。

次に Marsh 等の方法と 1 節の個体識別モデルとの関係を整理しよう。まず個体識別が可能ということは、識別を試したときに識別が起きる確率が正ということと同じである。故に個体識別が可能かの判断は、(3) 式が正かの判断と同じである。そして識別が実際に起きた場合に必ず観測されるなら、(2) 式の  $\Pr(\text{識別が実際に起きる})$  は、1 節の  $p(\delta)$  と同じ概念となる。しかし攻撃者が識別に成功しても、黙っていれば観測されるか分からない。故に識別が実際に起きることと観測されることは区別した方がよいかもしれない。この議論は 2.5 節へ先送りする。

結局 (3) 式の右辺の要素のどれかが 0 なら、個体識別が不可能と言える。しかし公開される母集団一意が皆無になるのは例外的で、普通は  $\Pr(a, b, c)$  は正となる。(3) 式の右辺を書き換えると

$$\Pr(\text{識別が起きる} \mid \text{識別を試みる}) = \Pr(a, b, c) \Pr(d|a, b, c) \quad (5)$$

であり、 $\Pr(d|a, b, c)$  が 0 なら個体識別が不可能と考えられよう。つまり Marsh 等の枠組みにおいて通常の場合、個体識別が可能か否かは  $\Pr(d|a, b, c)$  が 0 か否かという問題に縮退する。しかし Marsh 等は  $\Pr(d|a, b, c)$  の評価に失敗している。

我々の議論に沿って  $\Pr(d|a, b, c)$  が 0 か否かの判別方式を構成しよう。これまでの議論では、個体識別の難易度  $\delta$  が (1) 式のように閾値  $\alpha$  未満なら個体識別が可能ということであった。また  $\delta$  は (4) 式の関数  $h$  で表されると考えていたので、

$$\delta = h(g(\mathbf{y}_1), \mathbf{y}_2) < \alpha \Rightarrow \Pr(d|a, b, c) > 0 \quad (6)$$

とすればこれまでの議論と整合する。つまり個体識別の難易度  $\delta$  が閾値  $\alpha$  を下回れば、個体識別が可能ということである。

このように考えると、関数  $h$  は  $\Pr(d|a, b, c)$  が正という判定とできるだけ直接関係するのが望ましい。そして事象  $(a, b, c)$  が条件の確率を判定するなら、 $h$  は  $(a, b, c)$  を要因とするべきだろう。これを基準化して  $-\Pr(a, b, c) = g(\mathbf{y}_1)$  とすれば、 $g$  が (4) 式の単調性を満たして都合がよい。何故なら確率  $\Pr(a, b, c)$  は、正確に表現されて公開される母集団一意数と比例する。そして正確に表現されて公開される母集団一意数の増加は、母集団一意の確証をより容易にすると考えられる。故にあとは  $\Pr(a, b, c)$  が数値評価可能であれば、その評価値に基づいて匿名化を管理できる。

これまでの議論では  $\Pr(a, b, c)$  の意味が曖昧だったが、計算方法を定めれば概念は限定される。また関数  $g$  の具体型と必要な情報  $\mathbf{y}_1$  も、計算方法に依存して定まる。そして  $\mathbf{y}_1$  が決まらなければ、 $\mathbf{y}_2$  も定まらない。これらは理論モデルとは異なる次元の問題なので、節を改めて考察しよう。

### 2.3 匿名性の計測—実質と下限

EBA において匿名性の評価値を相対比較する際、評価手法のゆれは望ましくない。また出来るだけ多くの事例を統計的証拠として用いるには評価が名人芸であってはならず、形式的な手続きでなければならない。そのように匿名性の計算手法は具体的に定めておくべきである。本節では前節のモデルに沿って匿名性の評価値  $g(\mathbf{y}_1) = -\Pr(a, b, c)$  の計算を考察する。

匿名性の評価をする際、実質か下限のいずれを求めるのか意識的でなければならない。実質とは実際の攻撃者の能力に合わせた評価という意味であり、下限とは統計当局と同じ情報を持つ「最強」の攻撃者を想定するということである。

両者の違いを母集団一意数を例にとって説明しよう。母集団一意数は、キー変数群の多元分割表における度数1のセル数と形容することも出来る。この母集団一意を計算する多元分割表で、各変数の区分（カテゴリー分類）は公開データの区分と一致させるのが常識的である。ただ攻撃用情報の精度が公開データより粗ければ、公開データの区分方法で算出した母集団一意は、攻撃者にとっての母集団一意にならない。例えば攻撃者が五歳階級のデータしかもっていなければ、各歳別でデータが公表されていても階級内で識別できない。故に実質的な母集団一意数の評価では、公開表現と攻撃用情報の粗い方に各変数の区分を合わせる。常識的な方法では公開表現の方が攻撃用情報より常に粗いので、最強の攻撃者が想定されている。

実質的な匿名性評価では、現実の攻撃者の能力を知る必要がある。そして攻撃者の能力を知るための情報収集体制については、Elliot et al. (2010) の重要な議論が存在する。この議論は2.4節で紹介するが、そのような情報の完全な収集は資源の制約等から無理であろう。つまり実質的な匿名性評価の問題は、必ずしも評価に必要な情報を得られないことである。

部分的な情報から実際にありそうな攻撃方法（シナリオ）を推定し、匿名性を評価することはできる。このようなシナリオ依存のリスク評価は、例えばPaas (1988) が採用している。しかし想定した攻撃者より強い攻撃者が存在した場合、個体識別の可能性は管理されない。

一方、下限の匿名性は後で確認するように、公開データ表現とその元データから評価する。これらの情報は統計当局にとって常に入手可能であり、シナリオ選択に起因する評価のゆれが起きない。また最強の攻撃者より弱い攻撃者についても、個体識別の可能性は（過剰だが）管理できる。ただ問題は、過去の事例における個体識別の有無が現実の攻撃者の能力を反映していることである。

この問題を一般的に考えよう。匿名性の数値評価値の要因  $y_1 = (e_1, e_2)$  について  $e_1$  は公開データ表現とその元データと考える。そして  $e_2$  は、必ずしも観測されない攻撃者の能力とする。 $e_2$  が観測されるとして、実質的な匿名性の数値評価値が  $g(e_1, e_2)$  で表される。一方、最強の攻撃者にとっての匿名性の数値評価値を

$$\inf_{e_2} g(e_1, e_2) =: \underline{g}(e_1)$$

で表そう。ここで  $\underline{g}$  を用いて過去の事例で計算した匿名性の最低数値評価値を  $\underline{\gamma}_1$  と書く。新規に公開するデータの匿名性を  $\underline{g}$  で計算して  $\underline{\gamma}_1$  としてよいだろうか。

所与の  $y_2$  について、匿名性の数値評価値  $g$  が  $\beta$  未満なら個体識別が可能としよう。過去最低の実質的な匿名性  $\gamma_1$  は  $\underline{\gamma}_1$  以上である。故に閾値  $\beta \leq \gamma_1$  が正しいとしても、 $\underline{\gamma}_1 < \beta$  となる場合があり得る。このとき新規に公開するデータの実質的な匿名性が例えば  $\underline{\gamma}_1$  と等しければ、個体識別は可能となってしまう。

このような事態は、匿名性の実質と下限の差が変動する場合に起こりうる。新規に公開するケー

スについて匿名性の要因を  $(e'_1, e'_2)$  と書く。ただし  $\inf_{e'_2} g(e'_1, e'_2) = \underline{\gamma}_1$  となるように匿名化がなされているとしよう。そして  $\underline{g}$  を用いて評価した過去最低の匿名性のケースの要因を  $(e_1, e_2)$  と書く。つまり  $\inf_{e'_2} g(e'_1, e'_2) = \inf_{e_2} g(e_1, e_2)$  が成立している。ここで匿名性の実質と下限の差を過去のケースは  $c = g(e_1, e_2) - \inf_{e_2} g(e_1, e_2)$ 、新規のケースは  $c' = g(e'_1, e'_2) - \inf_{e'_2} g(e'_1, e'_2)$  で表す。攻撃者の能力が向上して  $c > c'$  の時、新規ケースの実質的匿名性は過去最低の実質的匿名性を下回る。そして  $\beta = g(e_1, e_2)$  なら、新規ケースの実質的匿名性は  $g(e'_1, e'_2) = \underline{\gamma}_1 + c' < \beta$  であり、過去のケースでは不可能だった個体識別が可能となる。

なお上の考察で  $c$  は下限評価の歪みを含む。下限評価の真値からのずれは、一定なら問題にならないことは重要だ。つまり実質  $g$  と下限  $\underline{g}$  の差  $c$  がケース毎に変化しなければ、過去最低の匿名性の下限  $\underline{\gamma}_1$  を与えるケースでは実質的な匿名性も過去最低になる。そして  $\underline{\gamma}_1 > \beta - c$  なら  $\gamma_1 > \beta$  なので、 $\underline{g}$  を用いて匿名化の程度を決めれば実質も管理される。言い換えれば、EBA は匿名化を相対比較するので、匿名性の絶対値に意味は無いということである。

匿名性の実質と下限の差  $c$  が一定という重要な条件を満たすには、たとえ歪んでいても同じ方法で測ることが重要である。また  $e_2$  は無視できず、変化を確認するべきだ。しかし情報  $e_2$  は入手性に問題があるので、 $y_2$  の一部として定性評価するしかないだろう。

このような前提で、匿名性の数値評価は  $\underline{g}(e_1)$  を用いるのが望ましい。つまり公開データ表現とその元データから匿名性の下限を求めるということである。そのように  $\Pr(a, b, c)$  が計算できるか要素毎に確認しよう。

$\Pr(a)$  の評価 Marsh 等は誤分類や誤記が公開ファイルと攻撃用ファイルのキー変数で起きていない確率を  $\Pr(a)$  とした。1981年の英国センサスの事後調査 (Post Enumeration Survey) で求めた変数の誤分類率を参照して、5つのキー変数が全て正確に分類されている割合は0.8程度と見積もられている。この場合に誤分類が公開ファイルと攻撃用ファイルのキー変数で独立に起きていなら、 $\Pr(a) = 0.8^2 = 0.64$  である。なおキー変数が増えれば、全てのキー変数が正確に分類されている確率は減少する。しかし母集団一意数は増えることになる。

実際には、公開ファイルと攻撃用ファイルで調査時点の差や変数の定義の差も存在するだろう。これらの差は  $\Pr(a)$  を低下させる。1971年の英国センサスの1年後に再調査した結果、同じ職業だった人の割合が61%でしかない例を Marsh 等は挙げている。1991年の英国センサスについては、Dale and Elliot (2001) が各キー変数が経時変化する程度を調べている。ただ Dale and Elliot も述べているように、本気の攻撃者は特定の調査が数年後に公開されることを見込み、同時点に調査した攻撃用ファイルを準備しておくだろう。この場合は、調査時点や変数の定義の差に多くの保護効果を期待出来ない。このように攻撃のシナリオに依存して、 $\Pr(a)$  はかなり変化する。

我々は匿名性の下限を評価したいので、最強の攻撃者を想定する。このシナリオでは、匿名化される前のキー変数が全て攻撃者にばれていると考える。この場合キー変数の精度を評価するに

は、匿名化されていない元ファイル<sup>14</sup>と公開ファイルのキー変数を比較する。そして近いレコードが同個体（のペア）と判定し、正しく判定された割合<sup>15</sup>を  $\Pr(a)$  と考える。このような手法は開示リスク評価によく用いられるので、研究蓄積が利用可能である。例えば伊藤他 (2009) を見よ。なお我々のシナリオでは、公開ファイルと攻撃用ファイルで調査時点の差は存在しない。また両者で定義の差は、匿名化によるもののみである。そして元ファイルのキー変数がどれほど誤分類されていたとしても、個体と正しく対応可能である。

このようなシナリオの非現実性は、現実には用意可能な攻撃用データの質と量に依存する。この情報が  $e_2$  であり、 $y_2$  の一部と考える。一方  $y_1$  は元ファイルと公開キー変数データだが、これらと比較し、正確にマッチされたレコードの割合が  $\Pr(a)$  として計算可能であった。注意すべきなのは、毎回同じ方法でマッチさせることである。

$\Pr(b|a)$  の評価 Marsh 等の議論で確率  $\Pr(b|a)$  は、公開個体数が母集団サイズにしめる割合である。例えば 1991 年の英国センサス匿名化標本では 2% となる。全数調査から等確率でサブサンプリングした公開ファイルなら、個体は等確率で公開ファイルに含まれる。その場合に Marsh 等の方法は妥当である。

しかし現実の標本調査では不等確率の複雑な抽出が行われる。また一部の個体について、調査されたか否かを攻撃者が知っているかもしれない。例えば集落抽出を行う調査では、被調査者は隣家も調査されたと推測できる。従って一般に真の  $\Pr(b|a)$  は個体毎に異なる。

ただ我々は個体毎（いわゆるレコードレベル）の確率評価をしているのではなく、ファイルレベルの評価が目的である。ファイルレベルでは公開の平均的な可能性を評価すると考えて、Marsh 等の方法を用いることにしよう。

$\Pr(c|a, b)$  の評価 母集団一意数が母集団サイズにしめる割合を求めればよい。なお本節冒頭で議論したように、母集団一意数を計算するための変数の区分は公開表現に従うべきである。また世帯単位のファイルでは、世帯毎の固まりを「レコード」として母集団一意を計算するのが筋である。具体的には、世帯員のレコードを年齢順に連結したまとまりを一レコードとして扱えば良い。この場合、世帯人数が異なればレコード長も異なる。

ただし全数調査でない限り、母集団一意数は推定しなければならない。そして星野 (2003) で説明したように、母集団一意数の推定は単純ではない。

Marsh 等は英国センサスの全数データが使えなかったため、イタリアのセンサスデータで母集団一意を数えて外挿している。キー変数が 8 つで 10 万人レベルの地域区分を公開するとして、

<sup>14</sup> 攻撃者は補定、エディットのルールを知らないはずなので、補定等を施す前のデータを元ファイルとする方が現実に近いかもしれない。ただそのようなデータが常に利用可能とは限らない。相対比較可能性を考えれば、補定等を施した後のデータを元としてよいだろう。実質と下限の差があるとしても、補定等の割合が小さかったり調査毎に大きく変動しない場合は無視できる。

<sup>15</sup> 何を分母とするかは議論の余地がある。本当に評価したいのは、母集団一意レコードについてのキー変数の精度である。しかし全数調査でないと、母集団一意のレコードを決めるのは難しい。そして評価の歪みより方法の変動を避けたいので、標本一意数を分母とするのが一案である。近さの計算方法によるが、一意にペア相手が見つかるレコード数と標本一意数はほぼ同じである。なお分母が 0 の場合は  $\Pr(a) = 0$  とみなして差し支えないだろう。

$\Pr(c|a, b)$  は 2.4%程度とされた。なおこの値は世帯単位ではなく個人単位で評価されている。1991年の英国センサスデータについては、Dale and Elliot (2001) によるとキー変数が7つで12万人レベルの地域区分を公開する前提で、 $\Pr(c|a, b)$  は 4.8%であった。

母集団一意数評価は Marsh 等の時代に比べてかなり進歩しており、(母集団サイズが所与で)公開ファイルの情報だけから推定できる。しかし評価手法による結果の違いが大きいため、同一手法によって評価することの重要性も大きい。

幅広い母集団について一意数の推定精度をルーチンワークとして確保するには、ピットマンモデル(付録 B を参照のこと)の使用を推奨する。この方法においてデータは、無限母集団すなわちピットマン分布からの標本とみなされる。そして母集団<sup>16</sup>も同一無限母集団からの標本とみなすので、データからピットマン分布の母数を最尤推定し、推定値の下で母集団一意数の挙動を求める。より具体的には、付録 C の手順書を参照されたい。

開示リスクを評価するファイルのレコード数は、母集団個体数のせいぜい一割程度であろう。この場合に安定的な母集団一意数の推定量は、全てバイアス<sup>17</sup>を持つ。手順書の推定量も例外でなく、おそらく過大に一意数を推定する。しかし既に考察したように、バイアスは一定であれば問題にならない。

なお特定のモデルと決めつけるよりも、モデル集合からデータに良く当てはまるモデルを選択し、そのモデルで一意数を推定する方が正確になる。しかしモデル集合の空間をうまく張らないと、リスク評価値がぶれる。また経験的に多くの場合、ピットマンモデルが選択<sup>18</sup>される。故に手間や精度及び様々な結果の整合性を勘案すれば、母集団一意数は常にピットマンモデルによって推定するのが最善と思われる。

一点つけ加えておくと、母集団一意数の推定改善にセルの番地情報を使うアプローチはあり得る。しかし大規模かつ疎な分割表では絶対的に情報が不足しているので、うまくいかないであろう。またそのようなアプローチは高度なモデリングが要求され、開示リスク評価の試行錯誤にも向かない。故に実務への採用は難しいはずだ。

このように  $\Pr(a, b, c)$  の下限評価に必要なのは、

$$e_1 = (\text{元ファイル, 公開ファイルのキー変数, 母集団サイズ})$$

である。これらの情報が数値評価の対象となり、 $y_2$  には含まれないと考えるべきだろう。次節では数値評価しない要因  $y_2$  を確定しよう。

<sup>16</sup>一部が観測されている現実の母集団について推定するのではなく、同サイズの母集団を新たに発生させる場合の挙動が推定される。

<sup>17</sup>有限母集団から非復元単純無作為抽出する場合、一意数の不偏推定量は一意に存在する。しかしこの不偏推定量は標準誤差が大きく、標本抽出率がかなり高くないと実用に耐えない。そして一意な不偏推定量なので、推定を安定させるためのいかなる工夫もバイアスを生む。

<sup>18</sup>裾の長いモデルとして代表的な負の二項分布は、統計の開示制限の分野ではポアソン=ガンマモデルとして知られている。このモデルは基本的に広義のピットマンモデルの特殊ケース( $\alpha \leq 0$  に対応)である。故にピットマンモデルのデータへのあてはまりは、基本的にポアソン=ガンマモデルを下回らない。そしてポアソン=ガンマモデルによる母集団一意数の推定値は、必ず Pitman モデルの推定値より(かなり)小さくなると思って良い。

## 2.4 定性評価の要因

本節では個体識別について数値評価しない要因  $y_2$  を定める。これまでの議論より、我々は (6) 式に基づいて母集団一意の確証の可能性  $\Pr(d|a, b, c)$  を判断するので、 $h$  の引数  $y_2$  は母集団一意の確証にかかる要因である。そして前節では  $e_2$ 、すなわち攻撃用データの質と量が、 $y_2$  の一部ということであった。

Marsh 等は母集団一意の確証手法として、全数名簿と公衆の目の利用<sup>19</sup> を検討している。全数名簿の利用とは、職業人名簿等で母集団一意が分かる場合を指す。特定の条件を満たす集団について全数の名簿があれば、その集団内の一意<sup>20</sup> は母集団でも (特定の条件を満たす) 一意である。そのような個体について、Marsh 等は特に強い匿名化を求めている。それから公衆の目とは、珍しくて目立つ個体が有名な場合を言う。例えば職業が現職の首相であれば、母集団一意を確証可能である。昨今ではソーシャルネットワークの拡大により、公衆の目は無視できないように思う。

全数名簿が利用出来たり、属性が公衆に知られていたりする個体については、詳しい個人情報 が社会に流通しているということだ。母集団一意の確証可能性及び実質的な  $\Pr(a, b, c)$  は、そのような個人情報 の環境に依存するだろう。個人情報環境を知るため、Elliot et al. (2010) は (i) アクセス制限付きデータベースの調査項目、(ii) 公知の個体データの形態、(iii) ネットショッピング等での web 上データ収集項目、(iv) 商業データベースの情報、(v) 個人情報の収集実験結果、(vi) 情報保有組織における個人情報の取り扱い慣行、(vii) ソーシャルネットワークでの個体データの形態、を調べることを提案している。またそこで現れる様々な変数間の関係を、ツリー構造を用いて記録することとしている。これらの要因は調査できたとしても、定量評価は (識別成功が希なので) 難しい。ただこれらについての理解から、現実的な攻撃用データとして

$$e_2 = (\text{外部データに含まれる個体数、変数の種類、精度})$$

を想定するべきだろう。なお世帯データの  $e_2$  は、事業所データの  $e_2$  と明らかに異なる。従って世帯データの匿名化事例は、事業所データの匿名化のエビデンスとして直接使えないということになる。個人情報環境は個体単位 (個人、世帯あるいは事業所等) 毎に集約するべきだ。

外部データに含まれる個体数は、実質的に攻撃可能な母集団一意数と比例するだろう。なお外部データに含まれる個体数増加の効果は、サブサンプリングにより  $\Pr(b|a)$  を下げれば打ち消すことができる。公開個体率  $\Pr(b|a)$  も攻撃可能な母集団一意数と比例すると考えられるからだ。

外部データの変数の種類は、キー変数の決定に用いる。前節ではキー変数が所与であったが、実際はキー変数を選択しなければリスク評価が出来ない。そしてキー変数の選定基準の揺れは避けた方がよい。これを念頭におき、過去の事例で用いたキー変数の種類は、キー変数の選択で考慮すべきである。キー変数に相当すると判断した根拠の外部データの状況が変わらなければ、同じ変数はキーとして用いなければならない。根拠が変われば、キー変数も変えるべきだろう。なお

<sup>19</sup>他に母集団一意の確率を統計モデルで求めることを挙げているが、それでは母集団一意の確証にならない。Dale and Elliot (2001) による Marsh 等の議論の再評価でも、統計的推測は母集団一意の確証として扱われていない。

<sup>20</sup>全数調査において低次元クロス集計の結果の度数が 1 と分かるような場合も該当する。

Elliot et al. (2011) がキー変数の選択基準を考察している。彼らの議論では、変数のアクセス容易性を定性評価した上でキー変数が選択される。

外部データの変数の精度は、 $\Pr(a)$  と  $\Pr(c|a, b)$  の実質的な値と関係する。なお変数の精度上昇の効果は、匿名化を強く施せば無効化出来る。何故なら匿名化で定まるデータの粗さ以上に変数の精度が上昇しても、開示リスクは変化しない。

$e_2$  以外の定性評価要因として、匿名化の「曖昧さ」を検討しておこう。ここでは匿名化に用いたデータ変換  $m$  の形を攻撃者が完全には知らない場合を曖昧と呼ぶ。例えば米国センサスマイクロデータのように、スワッピングが施されているがその割合やスワップ相手の選択方法などが未公開な状態は曖昧である。他方、労働力調査等の匿名データでは、符号表を読むことで匿名化が施されている変数や程度が完全に分かる。この状態は曖昧ではない。

曖昧さは余り研究されておらず、その効果<sup>21</sup>に定説はない。一つの理由として、計算機科学では曖昧さによる安全性を認めないことが挙げられる。その前提で設計した匿名化は統計当局が隠した情報が漏れても<sup>22</sup> 破られないので、保守的と言える。しかしこのような態度は最強の攻撃者を想定することと同じである。従って下限と実質の差の問題が起きる。

曖昧さが母集団一意の確証に影響する例を挙げよう。年齢と性別の二キー変数について、元ファイルが  $\{(110, M), (120, F)\}$ 、公開ファイルが  $\{(120, M), (110, F)\}$  だとする。年齢をスワップしたと考えれば第一レコード同士が同一個体であり、性別をスワップしたと考えれば元ファイルの第一レコードと公開ファイルの第二レコードが同一個体となる。この場合は  $m$  について何も知らないと、公開ファイルのレコードが元ファイルのどちらのレコードか分からない。ところが年齢変数に適当なノイズを付加したという情報が有れば、元ファイルと公開ファイルで同一個体のレコードが判明する。そして年齢が 120 歳の母集団一意な個体は、公開ファイルの第二レコードと確証される。

このように母集団一意の実質的確証可能性は、曖昧さの程度に依存するかもしれない。故に曖昧さは匿名化設計の一部として、明示的に考察した方がよい。現実には、攪乱的手法の詳細を公開する程度を  $y_2$  の一部として管理するということになる。なお補定やエディットの母数を明らかにしないことは、曖昧と同じことになる。

ここまでの議論で、既存の情報に基づく順攻撃による個体識別はある程度管理されるだろう。しかし情報が追加できるなら、これまでの枠組みでは管理されない事態が起きる。例えばあるレコードの識別が既存情報から確証できないにせよ、可能性が高いとしよう。この場合に追加の情報を詐取などすれば、確証できるかもしれない。追加情報を想定しての匿名化はあり得るが、詐取の可能性を際限なく考慮すると、有用なファイルの提供は不可能だろう。それよりも追加情報取得の可能性を低く保つ工夫をする方がよい。

<sup>21</sup> 特定の曖昧さの効果は、例えば以下のように評価できる。保守的な攻撃者なら、曖昧な部分に自分に不利な事前分布を入れる。このように評価される攻撃の難易度と真の難易度の差が、曖昧さの効果である。

<sup>22</sup> 関係者による情報漏洩だけ考慮すれば良いわけではない。攻撃者が攪乱の母数を推定できる可能性がある。例えば匿名データと匿名化されていないデータの分析結果を比較することで、攪乱の率の見当をつけられるかもしれない。攻撃者本人が 33 条申請による目的外使用でデータを手に入れなくても、他人が書いた論文や公の集計表が比較対象になり得る。



重要な追加情報を得るには、当該個体に接触する必要があるのではないかと。そして接触するには、広い意味<sup>23</sup>での個体の位置（住所、職場など定期的に訪れる場所、電話番号等）を知らなければならぬ。故にそのような接触可能性に係る条件で母集団一意な個体は、そうでない母集団一意よりも追加情報を得やすいので、確証の可能性が上がる。また逆攻撃は、接触可能な範囲で行われる。従って広い意味での位置情報の精度は、一定以上にならないように管理するべきである。

本節の議論をまとめておこう。定性評価の対象  $y_2$  として過去の事例と比較されるのは、以下の3要因である。

1. 同種の母集団についての  $e_2$  : 民間データベース等に含まれる個体数、変数の種類、精度
2. 匿名化の曖昧さ
3. 接触を可能とする情報の精度

本節で考察したように、匿名化手法の変更により攻撃者の能力向上を無効化できる場合がある。故に  $y_2$  が過去と同じかどうかの判断は、データ表現にある程度依存してしまう。匿名性の数値評価値  $g$  を変えるために匿名化手法を変更すると、定性評価も変わるかもしれないことは注意すべきである。

## 2.5 識別を試みる確率の決定要因

これまでの議論で後回しにされた、個体識別が実際に起きる確率と観測される確率  $p(\delta)$  の差について本節では議論する。つまり  $p(\delta) = \Pr(a, b, c, d) \cdot \Pr(\text{識別を公表するつもりで試みる})$  という関係が (2) 式の関係  $\Pr(\text{識別が実際に起きる}) = \Pr(a, b, c, d) \Pr(\text{識別を試みる})$  と違うかもしれないので、識別を公開すること、公開しないことについて要因の考察を行う。

識別を試みるという意味決定は、識別成功の損得や容易性に依存すると考えられる。Marsh et al. (1991) が指摘するように、 $\Pr(\text{識別が起きる} | \text{識別を試みる}) = \Pr(a, b, c, d)$  の減少は  $\Pr(\text{識別を試みる})$  を減少させるだろう。他に Elliot et al. (2010) は、もっともらしい攻撃シナリオの考察こそが、 $\Pr(\text{識別を試みる})$  の妥当なモデル化につながると主張している。

攻撃者が真に識別を成功させた場合、その事実を公表して得られる利益と、識別を隠して得る利益がある。まず識別成功を公表した場合、攻撃者は有名になるだろう。そして識別された個体は情報の漏洩を知ることになり、識別によって入手した情報を用いた詐欺、ストーキング等は難しくなる。そのように識別で得た情報を実用するには、識別成功は公表しない方がよい。また識別成功を公表すれば法的、社会的制裁の対象<sup>24</sup> になるかもしれない。故に例えば商業目的なら識別成功を公表せず、攻撃者は精度の良いマーケティングの利益を享受するだろう。

<sup>23</sup>狭義の地理情報が強力なキー変数であることは良く知られている。

<sup>24</sup>匿名データの利用者については、統計法第43条第2項に「当該匿名データをその提供を受けた目的以外の目的のために自ら利用し、又は提供してはならない」とある。個体識別の成功を公表することは（識別目的でのデータ提供は行われぬので）本条に違反するが、直ちに罰則が適用されるわけではない。匿名データの利用者についての罰則は「匿名データを、自己又は第三者の不正な利益を図る目的で提供し、又は盗用した者」に対して「五十万円以下の罰金に処する」（61条3項）とだけ定められている。例えば匿名データの不備を指摘するための個体識別の公表は不正な利

このように考えれば、公開ファイルが含む実用（隠れて悪用）可能な情報が多ければ、 $\Pr(\text{識別を公表するつもりで試みる})$ を増加させる。また識別を試みるという事象は識別を公表するつもりで試みる事象を包含するので、 $\Pr(\text{識別を試みる})$ も増加する。

ただファイルが実用可能な情報を含まなくても、識別成功の公表により統計当局の面目を失わせ有名になることを、魅力的に感じる人間が居ないとは言えない。故に $\Pr(\text{識別を公表するつもりで試みる})$ は正のはずで $\Pr(\text{識別を試みる}) = 0$ にはならない。しかし実用の帰結は多様なのに対し、識別成功の公表は帰結が同じである。実用できない情報は公表することによってしか利益を得られないので、攻撃の誘因として全て等価ということになる。

では実用可能な情報とは何か。多様な犯罪を想像して判断するしかないが、匿名化によって実用性は変えられることを指摘しておく。例えば病歴という情報は、削除したり罹患時期を区間表示したりすることで、実用困難にできる。多くの統計調査は適切に匿名化すれば実用可能な情報を含まず、攻撃の誘因は識別成功の公表による利益のみとなる。そしてこの場合 $p(\delta) \doteq \Pr(\text{識別が実際に起きる})$ と考えて良いはずだ。

なお調査客体が秘密にしたい調査項目（変数）を「センシティブ変数」と呼ぶ。秘密でない情報は保護に値しないので、実用を妨げる目的での匿名化の対象は、センシティブ変数の一部と考えられる。重要なのは、センシティブか否かは調査客体の主観に依存<sup>25</sup>するということである。故にセンシティブの程度は、攻撃の動機付けの程度と必ずしも一致しない。

### 3 おわりに—匿名データの審査体制について

最初にこれまでの議論をまとめる。個体識別が不可能かつ有用なデータを統計的根拠に基づいて作成する手順は以下ようになる。

1.  $y_2$  が共通する過去の事例をリストアップする。
2. それらの事例について  $\Pr(a, b, c)$  の下限  $\underline{g}$  をそれぞれ計算する。
3. その中で最も高い  $\underline{g}$  を  $g^*$  と書く。
4. データの匿名性の評価値が  $g^*$  となるように匿名化する。

なお匿名性の評価値が同じになる複数のデータ表現では、データの有用性が高いものを選びたい。本稿で有用性の評価は議論しないが、例えば星野 (2010) は基本的な考え方を説明している。以下ではこのような立場から、望ましい匿名データの審査体制を考察する。

益を図る目的と必ずしも言えないので、罰則は適用できないのではないかと。なお 33 条申請によって調査票情報を手に入れた者が個人又は法人の秘密を漏らした場合は「二年以下の懲役又は百万円以下の罰金」(57 条 2 項 3 号)、自己又は第三者の不正な利益を図る目的で提供又は盗用した場合は「一年以下の懲役又は五十万円以下の罰金」(59 条 2 項)、と罰に差がつけられている。ところが匿名データの利用者が（個体識別によって入手した）秘密を漏らした場合の罰則規定はなく、そのような事態を統計法は想定していないように思われる。

<sup>25</sup>全ての調査客体の判断を聞くのは非現実的なので、リスクの評価者がセンシティブ変数を定める際に保守的であれ、ということになる。

まず審査用資料(チェックリスト)は、蓄積して参照するものだということをはっきりさせておきたい。EBAは過去の経験をエビデンスとして用いる。故に過去のチェックリストを、経験の要約として用いたい。この事情は、将来的に $\alpha$ と $\delta$ の関係をモデル化する場合も変わらない。

従ってチェックリストの記載事項は、事例の十分統計量であるべきだ。つまり個体識別が可能か否かの判断に用いる情報を(過)不足なく記入するということである。このような観点から、現行のチェックリストは日本の制度にふさわしいだろうか。世帯調査のチェックリスト(H23/3/28 改正版)についてのみ、改善できる点を指摘したい。

チェックリストに記載すべき項目で漏れているのは、まず $\Pr(a, b, c)$ の下限 $g$ である。付録の手順書に従えばそれほど計算に手間がかかるとは思えず、匿名化表現の要約として費用対効果が高い情報と考える。またキー変数の情報を持つ部分集団の全数名簿は、母集団一意の確証について重大な影響がある。故に質問項目として特に欄を設けるべきである。そのような名簿が存在するなら、名称、部分集団の種類、個体数、含むキー変数の種類、精度を記述させるとよい。他に狭義の地理情報については記入欄が存在するが、接触を可能とするような広義の地理情報の有無を確認するべきだ。

それからチェックリストに存在する項目で、記入の焦点をしぼるべき箇所がある。まず「マイクロデータを特定できる可能性のある外部ファイル」の存在を記入することになっているが、どのようなファイルが該当するのか明確化するべきである。具体的には、匿名データが含む変数と同じ情報(これがキー変数ということである)をもつ外部ファイルの有無を問うべきだ。そしてそのファイルの名称、含むキー変数の種類、精度、及び個体数を分けて記述してもらう方がよい。また「秘密の情報」(センシティブ変数)のうち、「特に秘匿する必要性の高い調査項目」の有無を聞いているが、必要性の意味を明白にしたほうがよい。個体識別の可能性を制限するための必要性ではなく、実用性を限るための匿名化の必要性を聞かなければならない。またチェックリストには「誤差(ノイズ)」を聞く項目が存在する。誤差の付加は「攪乱」手法の例なのだが、用語の問題は別にして、この項目には匿名化の曖昧さを評価するための情報を記入させるべきだ。具体的には、攪乱手法のパラメータと、その公開方針を分けて書かせるということになる。

チェックリストに記入される情報の使われ方は、説明書を用意するべきであろう。現行のチェックリストも冒頭で匿名化の考え方などが書かれているが、やや説明不足に見える。個体識別可能性の判定方式を明示すれば、焦点がずれたチェックリスト記入の恐れは減る。

## Acknowledgements

本研究は科学研究費及び統計数理研究所の共同研究経費の補助を受けている。以下の付録B,Cは星野(2012)の一部を改訂したものである。

## 付録

### A 世帯調査のチェックリスト（H23/3/28 改正版）要約

1. 地理的情報
  - (a) 地理情報のレベル、加工の有無
  - (b) 地理情報以外の地理的情報の有無
  - (c) 地域分析用の地理情報提供の有無
  - (d) 特定の種類の施設の情報の有無
2. 世帯の識別情報
  - (a) 世帯のキー変数
  - (b) キー変数への匿名化及び分布
  - (c) 世帯のまとめりへの匿名化の有無
3. 個人の識別情報
  - (a) 個人のキー変数
  - (b) キー変数への匿名化及び分布
4. 攪乱の有無
5. サブサンプリングの有無
6. 外部の情報
  - (a) 個人・世帯の特定に使える外部情報の存在
  - (b) 母集団情報として利用している情報
7. その他
  - (a) データの並び順についての匿名化措置
  - (b) サンプル情報により特定の地域や集団であることが明らかになる可能性
  - (c) センシティブ変数への匿名化
  - (d) 提供時期と調査時点との差
  - (e) その他の匿名化処理の有無

## B ピットマンモデルについて

自然数  $n \in \mathbb{N} := \{1, 2, 3, \dots\}$  を自然数の和で表す事を分割と呼ぶ。この和の中で自然数  $i$  が足される回数を  $s_i$  で表せば、 $\mathbf{s}_n := (s_1, s_2, \dots, s_n)$  は (順序無しの) 分割を表す。非負整数の集合を  $\mathbb{N}_0$  で表すと、 $n$  の全ての分割の集合は  $\mathcal{S}_n := \{\mathbf{s}_n : s_i \in \mathbb{N}_0, i \in \{1, 2, \dots, n\}, \sum_{i=1}^n i s_i = n\}$  で表される。この集合上の分布が自然数の確率分割である。以下では  $u := \sum_{i=1}^n s_i$  とする。

Pitman 分布 (Pitman, 1995) は自然数の確率分割であり、母数  $0 \leq \alpha < 1, \theta > -\alpha$  について確率関数は以下のように書ける。

$$p(s_1, s_2, \dots, s_n) = n! \frac{\theta^{[u:\alpha]}}{\theta^{[n]}} \prod_{j=1}^n \left( \frac{(1-\alpha)^{[j-1]}}{j!} \right)^{s_j} \frac{1}{s_j!}, \quad \mathbf{s}_n \in \mathcal{S}_n, \quad (7)$$

ただし  $\theta^{[u:\alpha]} = \theta(\theta + \alpha) \cdots (\theta + (u-1)\alpha)$ ,  $\theta^{[n]} = \theta(\theta + 1) \cdots (\theta + n - 1)$  である。

個票データとの対応を述べておこう。匿名化の程度を定めることで、キー変数に関する分割表が出来る。分割表の情報のうち度数を、第  $j$  セルについて  $f_j$  と書く。ただしセル総数が  $J$  として  $j \in \{1, 2, \dots, J\}$  である。ここで  $i = 1, 2, \dots, n$  について度数  $i$  のセルの数を  $s_i$  と表す。つまり指示関数  $1(\cdot)$  を使えば、 $s_i = \sum_{j=1}^J 1(f_j = i)$  である。例えば  $s_1$  は標本で一意なレコード数となる。このように作られる  $\mathbf{s}_n$  を「寸法指標」と呼び、 $n \ll J$  なら Pitman 分布の標本とみなせる。母集団サイズ  $\tilde{n}$  も  $J$  よりかなり小さいなら、確率変数  $S_{\tilde{n}} := (S_1, S_2, \dots, S_{\tilde{n}})$  が Pitman 分布に従う場合、 $S_1$  で母集団一意数の挙動が表せる。

経験ベイズ的に母集団一意数推定の論理を説明すると、以下の通りになる。 $S_{\tilde{n}}$  の事前分布が Pitman 分布であり、その母数  $(\alpha, \theta)$  は超母数である。超母数はデータ  $\mathbf{s}_n$  により (最尤) 推定される。推定したい母数は  $S_1$  であり、 $(S_2, S_3, \dots, S_{\tilde{n}})$  は局外母数である。母数  $S_1$  の周辺分布については Hoshino (2012, Theorem 3) を見よ。

Pitman 分布に従う  $S_n$  の任意の周辺階乗モメントは、Yamato and Sibuya (2000) が与えている。特に

$$E(S_i) = \frac{(1-\alpha)^{[i-1]} n^{(i)}}{i!} \theta \left( \frac{(\theta + \alpha)^{[n-i]}}{\theta^{[n]}} \right), \quad (8)$$

である。

超母数の最尤推定量を  $(\hat{\alpha}, \hat{\theta})$  と書けば、(8) 式に  $\alpha = \hat{\alpha}, \theta = \hat{\theta}, n = \tilde{n}, i = 1$  を代入して母集団一意数の点推定量が得られる。すなわち

$$\hat{S}_1 = \tilde{n} \frac{(\hat{\theta} + \hat{\alpha})(\hat{\theta} + \hat{\alpha} + 1) \cdots (\hat{\theta} + \hat{\alpha} + \tilde{n} - 2)}{(\hat{\theta} + 1)(\hat{\theta} + 2) \cdots (\hat{\theta} + \tilde{n} - 1)}. \quad (9)$$

なお Hoshino (2001, Proposition 3) によれば、 $\alpha \geq 0$  について  $\lim_{n \rightarrow \infty} E(S_1)/E(U_n) = \alpha$  であ

る。ただし  $U_n := \sum_{i=1}^n S_i$  は度数が 0 でないセルの総数なので、母集団で空でないセルのうち一意のセル数の割合は  $\alpha$  と解釈出来る。

次にフィッシャー情報量を確認しておこう。まず対数尤度関数を

$$L(\alpha, \theta) = \sum_{i=1}^{u-1} \log(\theta + i\alpha) - \sum_{i=1}^{n-1} \log(\theta + i) + s_1 + \sum_{i=2}^n s_i \sum_{j=1}^{i-1} \log(j - \alpha) + \text{Const.} \quad (10)$$

で表す。二次の微分係数は

$$\frac{\partial^2 L(\alpha, \theta)}{\partial \theta^2} = - \sum_{i=1}^{u-1} \frac{1}{(\theta + i\alpha)^2} + \sum_{i=1}^{n-1} \frac{1}{(\theta + i)^2}, \quad (11)$$

$$\frac{\partial^2 L(\alpha, \theta)}{\partial \alpha^2} = - \sum_{i=1}^{u-1} \frac{i^2}{(\theta + i\alpha)^2} - \sum_{i=2}^n s_i \sum_{j=1}^{i-1} \frac{1}{(j - \alpha)^2} < 0, \quad (12)$$

$$\frac{\partial^2 L(\alpha, \theta)}{\partial \theta \partial \alpha} = - \sum_{i=1}^{u-1} \frac{i}{(i\alpha + \theta)^2} < 0 \quad (13)$$

である。(13) 式より  $\hat{\alpha}$  と  $\hat{\theta}$  は負の相関を持つ。情報量はこれらの式について  $u$  を  $U_n$  に、 $s_i$  を  $S_i$  に置き換えて期待値をとる。E( $S_i$ ) は (8) 式で与えられているので、あとは

$$P(U_n = u) = \frac{\theta^{[u:\alpha]}}{\theta^{[n]}} (-1)^{n-u} C(n, u, \alpha) \alpha^{-u}, \quad u \in \{1, 2, \dots, n\}. \quad (14)$$

を利用して数値的に評価できる。 $C(\cdot, \cdot, \cdot)$  は C-ナンバーと呼ばれ、一般化されたスターリング数である。C-ナンバーについては Charalambides and Sing (1988) を参照のこと。Sibuya and Yamato (2001, Proposition 5) がフィッシャー情報量行列のオーダーを評価しており、 $n \rightarrow \infty$  の時  $I_{\theta\theta} = O(1)$ ,  $I_{\theta\alpha} = O(\log n)$ ,  $I_{\alpha\alpha} = O(n^\alpha)$  である。特に  $\theta$  の推定精度は悪い。

## C 母集団一意数の推定手順

以下では標本サイズを  $n$ 、母集団サイズを  $\tilde{n}$  と記す。

1. 評価するキー変数とその精度を決める。
2. 決められたキー変数全てについてクロス集計する。つまり (高次元の) 分割表を作り、各セルに所属するレコード数 (度数) を数える。
  - セル総数  $J$  は、全てのキー変数のカテゴリー数の積である。連続変数でも現実には有限個の表現しかとらず、その表現の数をカテゴリー数と考える。

- 第  $j$  セルの度数を  $f_j, j = 1, 2, \dots, J$ , と書く。以下の結果はインデクス  $j$  の付け方に依存しない。

3. 空でないセルの度数の度数 (寸法指標) を数える。

- $i = 1, 2, \dots, n$  について度数  $i$  のセルの数を  $s_i$  と表す。つまり指示関数  $1(\cdot)$  を使えば、 $s_i = \sum_{j=1}^J 1(f_j = i)$  である。
- 最大のセルの度数が  $m$  ならば、 $m < i$  について  $s_i = 0$  である。

4. データを生成した構造 (確率分布) を推定する。

- 現実の母集団を無限母集団 (超母集団) からの標本とみなす。この場合、手元の標本から超母集団の分布を推定すれば、母集団の挙動も推定される。
- 超母集団の分布として広義の Pitman モデルを仮定し、その母数を最尤推定する。
- Pitman モデルは 2 母数  $(\alpha, \theta)$  を持ち、 $\alpha$  が負の場合と正の場合で分けて考えた方がよい。どちらの場合も  $u = \sum_{i=1}^n s_i, n = \sum_{i=1}^n i s_i$  である。
  - $0 \leq \alpha < 1, \theta > -\alpha$  について Pitman モデルの確率関数は (7) 式で表される。
  - $\alpha < 0$  の場合は (7) 式で  $\theta = -J\alpha$  とおき、さらに  $-\alpha = \gamma$  とおく。すると一母数の確率関数を得る：

$$p(s_1, s_2, \dots, s_n) = \frac{n! J! \Gamma(J\gamma)}{\Gamma(J\gamma + n)} \prod_{i=0}^n \left( \frac{\Gamma(\gamma + i)}{\Gamma(\gamma) i!} \right)^{s_i} \frac{1}{s_i!}. \quad (15)$$

ここで  $\gamma > 0$  であり、 $s_0 = J - u$  である。

- モデル (7) を「(狭義の)Pitman モデル」と呼ぶ。モデル (15) を「多項ディリクレモデル」と呼ぶ。本来は AIC 等によりデータ依存でいずれかをモデル選択するのが良いが、ここでは簡易的な選択基準を示す：
  - 母集団サイズ  $\tilde{n}$  が総セル数  $J$  より大の場合、多項ディリクレモデルを用いる。
  - その他の場合は Pitman モデルを用いるが、尤度の最大化に失敗する (繰り返し計算が収束しない) 場合、多項ディリクレモデルを用いる。なお初期値をランダムに変えていると、そのうち収束することもある。正確に判断するには、尤度の等高線図を見れば良い。
- 狭義の Pitman モデルの最尤推定は以下のように行えば良い。
  - 対数尤度関数は (10) 式で与えられている。

(b) 最尤推定量は以下の同時方程式の解である .

$$\frac{\partial L(\alpha, \theta)}{\partial \theta} = \sum_{i=1}^{u-1} \frac{1}{\theta + i\alpha} - \sum_{i=1}^{n-1} \frac{1}{\theta + i} = 0,$$

$$\frac{\partial L(\alpha, \theta)}{\partial \alpha} = \sum_{i=1}^{u-1} \frac{i}{\theta + i\alpha} - \sum_{i=2}^n s_i \sum_{j=1}^{i-1} \frac{1}{j - \alpha} = 0.$$

- (c)  $L(\alpha, \theta)$  の最大化は汎用最大化ルーチン (R の `optim()` 関数等) に任せても良いだろう。
- (d) 最尤推定値を自前で評価するなら、二次の微分係数 (11),(12),(13) 式を用いたニュートン=ラフソン法が適当である。
- (e)  $c = s_1(s_1 - 1)/s_2$  として、以下の近似的なモメント推定量を得る。これらをニュートン=ラフソン法の初期値として使うことが考えられる。

$$\hat{\theta} = \frac{nuc - s_1(n-1)(2u+c)}{2s_1u + s_1c - nc}, \quad \hat{\alpha} = \frac{\hat{\theta}(s_1 - n) + (n-1)s_1}{nu},$$

● 多項ディリクレモデルの最尤推定は以下のように行えば良い。

(a) 対数尤度関数は定数を除いて

$$L(\gamma) = - \sum_{i=0}^{n-1} \log(J\gamma + i) + \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \log(\gamma + j).$$

(b) 最尤推定値は尤度方程式

$$\frac{dL(\gamma)}{d\gamma} = - \sum_{i=0}^{n-1} \frac{J}{J\gamma + i} + \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \frac{1}{\gamma + j} = 0$$

の解である。

- (c)  $L(\gamma)$  の最大化は汎用最大化ルーチン (R の `optimize()` 関数等) に任せても良いだろう。
- (d) 最尤推定値を自前で評価するなら、二次の微分係数

$$\frac{d^2L(\gamma)}{d\gamma^2} = \sum_{i=0}^{n-1} \frac{J^2}{(J\gamma + i)^2} - \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \frac{1}{(\gamma + j)^2}$$

を用いたニュートン=ラフソン法が適当である。



(e) 尤度関数は単峰であり、それほど初期値に依存せず最大化が可能である。ただ最尤推定値が無限大に発散する事はある事であり得て、それは確率関数が等確率  $J$  項分布である事を意味する。また最尤推定値が 0 の場合、狭義の Pitman モデルの方が適切と思われる。

- 狭義の Pitman モデルと多項ディリクレモデルの境界 ( $\alpha = 0$ ) のモデルを Ewens モデルという。Ewens モデルの確率関数は以下の通り：

$$p(s_1, s_2, \dots, s_n) = n! \frac{\theta^n}{\theta^{[n]}} \prod_{j=1}^n \left(\frac{1}{j}\right)^{s_j} \frac{1}{s_j!}. \quad (16)$$

- 同じデータについて Ewens モデルの最尤推定値を  $\hat{\theta}_E$  と書き、Pitman モデルの最尤推定値を  $(\hat{\alpha}, \hat{\theta}_P)$  と書く。もし  $\hat{\alpha} > 0$  ならば  $\hat{\theta}_E > \hat{\theta}_P$ .
- 上の結果は Pitman モデルのチェックに使える。また最尤推定の繰り返し計算の範囲を限定できる。
- Ewens モデルの尤度関数は単峰であり、最大化は容易である。

5. 同定されたデータ構造の下で母集団一意数の推定値  $\hat{S}_1$  を求める。

- (a) 狭義の Pitman モデルの場合、母数の最尤推定値を  $\hat{\alpha}, \hat{\theta}$  と書けば (9) で推定される。
- (b) 多項ディリクレモデルの場合、母数の最尤推定値を  $\hat{\gamma}$  と書けば

$$\hat{S}_1 = \tilde{n}(J-1)\hat{\gamma} \frac{((J-1)\hat{\gamma}+1)((J-1)\hat{\gamma}+2)\cdots((J-1)\hat{\gamma}+\tilde{n}-2)}{(J\hat{\gamma}+1)(J\hat{\gamma}+2)\cdots(J\hat{\gamma}+\tilde{n}-1)}.$$

- これらの推定値はモデルの下での度数 1 のセル数の期待値である。
- 注 1) Ewens モデルの母集団一意数推定式は、Pitman モデルの推定式に  $\hat{\alpha} = 0$  を代入して得られる。
- 注 2) 等確率  $J$  項分布の母集団一意数推定値は  $\tilde{n}(1 - 1/J)^{\tilde{n}-1}$  である。

## 参考文献

- [1] Charalambides, C.A. and Singh, J. (1988) A Review of the Stirling Numbers, Their Generalizations and Statistical Applications. *Communications in Statistics, Theor. Meth.*, **17**, 2533–2595.
- [2] Dale, A. and Elliot, M. (2001) Proposal for 2001 Samples of Anonymized Records: An Assessment of Disclosure Risk. *Journal of the Royal Statistical Society, Series A*, **164**, 427–447.

- [3] Domingo-Ferrer, J. and Torra, V. (2001) A Quantitative Comparison of Disclosure Control Methods for Microdata. *Confidentiality, Disclosure, and Data Access: Theory and Practical Application for Statistical Agencies*, Doyle et al. (Eds.), Elsevier, Amsterdam, 111-133.
- [4] Duncan, G., Keller-McNulty, S.A. and Stokes, S.L. (2001) Disclosure Risk vs. Data Utility: The R-U Confidentiality Map. Technical Report 121, National Institute of Statistical Sciences, Durham, North Carolina.
- [5] Elliot, M. J., Skinner, C. J., and Dale, A. (1998) Special Uniques, Random Uniques, and Sticky Populations: Some Counterintuitive Effects of Geographical Detail on Disclosure Risk. *Research in Official Statistics*, **1**, 53–67.
- [6] Elliot, M., Lomax, S., Mackey, E. and Purdam, K. (2010) Data Environment Analysis and the Key Variable Mapping System. *Privacy in Statistical Databases*, Domingo-Ferrer, J. and Magkos, E. (Eds.), LNCS 6344, 138–147, Springer-Verlag, Berlin Heidelberg.
- [7] Elliot, M., Mackey, E. and Purdam, K. (2011) Formalizing the Selection of Key Variables in Disclosure Risk. *Int. Statistical Inst.: Proceedings of the 58th World Statistical Congress*, 2777–2784.
- [8] Hoshino, N. (2001) Applying Pitman’s Sampling Formula to Microdata Disclosure Risk Assessment, *Journal of Official Statistics*, **17**, 499–520.
- [9] 星野伸明 (2003) 「超母集団モデルによる個票開示リスク評価」, *統計数理*, **51**, 297–319.
- [10] Hoshino, N. (2009) The Quasi-multinomial Distribution as a Tool for Disclosure Risk Assessment, *Journal of Official Statistics*, **25**, 269–291.
- [11] 星野伸明 (2010) 「公的統計マイクロデータ提供制度の課題」, *日本統計学会誌*, **40**, 23–45.
- [12] 星野伸明 (2012) 「公的統計の開示リスク評価—労働力調査の論点」, 『*経済統計・政府統計の数理的基礎と応用-I*』, 国友直人・山本拓共編, CIRJE 研究報告書シリーズ, CIRJE-R-10, 40–56.
- [13] Hoshino, N. (2012) On the Marginals of a Random Partitioning Distribution. 研究集会「数理統計学の沃野」予稿集, 78–86.
- [14] 伊藤伸介 (2012) 「政府統計マイクロデータの提供における匿名化措置—イギリス統計法における法制度的措置と攪乱的手法の適用可能性を中心に—」, *明海大学経済学論集*, **24**, 1–14.
- [15] 伊藤伸介・磯部祥子・秋山裕美 (2009) 「秘匿性の評価方法に関する実証研究—全国消費実態調査のマイクロアグリゲートデータを用いて—」, *統計センター製表技術参考資料*, **11**, 12–14.

- [16] Marsh, C., Skinner, C., Arber, S., Penhale, P., Openshaw, S., Hobcraft, J., Lievesley, D. and Walford, N. (1991) The Case for a Sample of Anonymized Records from the 1991 Census. *Journal of the Royal Statistical Society, Series A*, **154**, 305–340.
- [17] Paass, G. (1988) Disclosure Risk and Disclosure Avoidance for Microdata. *Journal of Business and Economic Statistics*, **6**, 487–500.
- [18] Pitman, J. (1995) Exchangeable and Partially Exchangeable Random Partitions. *Probability Theory and Related Fields*, **102**, 145–158.
- [19] Sibuya, M. and Yamato, H. (2001) Pitman’s Model of Random Partitions. 数理解析研究所講究録, **1240**, 64–73.
- [20] 総務省政策統括官（統計基準担当）(2011). 「匿名データの作成・提供に係るガイドライン（平成23年3月28日改正版）」
- [21] Sweeney, L. (2002)  $k$ -Anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, **10**, 557–570.
- [22] 竹村彰通 (1997) 「個票データ開示の理論」, 科学研究費補助金（課題番号 08209102）報告書, 2–25.
- [23] U.S. Office of Federal Statistical Policy and Standards (1978). *Report on Statistical Disclosure and Disclosure Avoidance Techniques*. Statistical Policy Working Paper 2, U.S. Department of Commerce, Washington DC.
- [24] Yamato, H. and Sibuya, M. (2000). Moments of Some Statistics of Pitman Sampling Formula. *Bulletin of Informatics and Cybernetics*, **32**, 1–10.

# ベンチマーク問題とロバスト・デントン法

加藤賢悟

2014年1月20日

## 概要

月次・四半期データと年次データとの非整合性を調整する操作をベンチマーキング (benchmarking) と呼び、特に官庁統計の作成において標準的に行われている。本稿では、ベンチマーキング手法として最も広く使用されているであろうデントン (Denton) 法を解説するとともに、新しいベンチマーキング法、ロバスト・デントン法 (robust Denton method), を提案する。また、提案する手法が実務上望ましい性質を持つことを議論し、実データを用いた比較を行う。

## 1 はじめに

官庁統計などにおいて、しばしば、同種のデータに対して集計頻度が異なる複数の系列 (月次、四半期、年次など) が作成されることがある。典型的には、集計頻度の低いデータ (例えば年次データ) は、集計頻度の高いデータ (例えば月次や四半期データ) よりも精度の高いデータとみなされることが多い。ここで問題となるのは、例えば、四半期ごとに集計されたデータの合計が、年次データと必ずしも一致しない、ということが起こりうるという点である。日本の GDP 統計を例にとると、「速報」として四半期データが公表され、しばらくして「確報」として年次データが公表されるが、実は「速報」と「確報」には、推計手法・基礎総計などに大きな違いがあり、従って、「速報」の四半期データを合計しても、「確報」の年次データに一致しない (得能, 2007)。このような非整合を調整するために、通常、より精度の高いとされる年次データを「ベンチマーク (benchmark)」として、月次・四半期データを年次データと整合的になるように調整することが行われ、そのような調整のことを「ベンチマーキング (benchmarking)」と呼ぶ。

ベンチマーク問題をより正確に述べると、いま、1 年を  $s$  期に分け、 $n$  年のデータが利用可能であるとする ( $s$  と  $n$  は正の整数とする)。四半期データを扱う場合は、 $s = 4$  であり、月次データを扱う場合は、 $s = 12$  である。  $T = ns$  とおく。各  $t (= 1, \dots, T)$  期にお

いて、データ  $I_t$  が利用可能であり ( $I_t$  は月次・四半期データを想定している)、同時に、各  $k (= 1, \dots, n)$  年において、年次データ  $A_k$  が利用可能であるとする。このとき、ある  $k = 1, \dots, n$  において、

$$\sum_{i=1}^s I_{s(k-1)+i} \neq A_k,$$

なる非整合が生じることをベンチマーク問題と呼び、原系列  $I_1, \dots, I_T$  を適当に調整して、

$$\sum_{i=1}^s X_{s(k-1)+i} = A_k, \quad k = 1, \dots, n, \quad (1)$$

なる調整済み系列  $X_1, \dots, X_T$  を作成することをベンチマーキングと呼ぶ。

ベンチマーキング法として最も広く使用されている手法は、デントン (Denton) 法であろう (Denton, 1971)\*<sup>1</sup>。デントン法は例えば、日本の官庁統計において季節調整法として標準的に使用されている X-12ARIMA にオプションとして装備されている (X-12ARIMA に関しては、国友, 2004, を参照)。また、日本の GDP 統計の作成においても、従来はより素朴な (しかし問題点も指摘されている) プロ・ラタ (Pro-Rata) 法が使用されてきたが、最近では、デントン法が使用されるようになってきている (内閣府, 2011)。デントン法の正確な定義は 2 節で与えるが、直感的には、ベンチマーク制約 (1) の下で、短期の変動をなるべく保存するように調整済み系列を作成する一つの手法と理解することができる。

本稿の目的はデントン法を紹介するとともに、デントン法の一つの代替法として、ロバスト・デントン法なるベンチマーキング法を提案することである。提案するロバスト・デントン法は、デントン法の目的関数 (2 次関数) を絶対値関数に置き換える、という一見単純なものであるが、次の特徴を持つことが示される (2 節を参照)：

- 最適化問題が線形計画法として定式化できる。そのため、追加的な (線形) 制約を入れて最適化を実行する、といったことなどが容易にできる。
- ロバスト・デントン法によって作成された調整済み系列  $X_1, \dots, X_T$  に対して、 $X_t \neq 0, t = 1, \dots, T$  なら、傾き  $X_t/X_{t-1}$  のうち、少なくとも  $(T - n)$  個は原系列の傾き  $I_t/I_{t-1}$  に一致する ( $I_t \neq 0, k = 1, \dots, T$  は仮定している)。

ベンチマーク制約 (1) は  $X_1, \dots, X_T$  に関して線形の制約であり、従って、目的関数を線形 (絶対値関数) にすることは最適化の観点から自然な発想と思われるが、いままでこの

---

\*<sup>1</sup> 本稿のデントン法とは、修正デントン法 (modified Denton method)、ないし、比例デントン法 (Denton proportional method) とも呼ばれる。デントン法には様々なバリエーションがあるが、本稿では、2 節で定義する最も標準的なものを考察する。

ような提案は(少なくとも筆者の知る限り)なされてこなかったようである。言うまでもなく、今では、多くの標準的な数値計算ソフトウェア上で線形計画法を解くパッケージが利用可能であり、従って、手軽にロバスト・デントン法を実行することができる\*2。2番目の特徴は、ロバスト・デントン法が短期の変動をかなりの部分厳密(exact)に保存することを保証しており、デントン法にはない実務上の利点といえる。ここで、ロバスト・デントン法は、デントン法と類似の目的関数を最小化していることから、(プロ・ラタ法とは異なり)年次が変わるところでも変動が原系列のそれとなるべく乖離しないように調整している、ということに注意すべきである。

ベンチマーク問題はデータ集計に関する基本的な話題であり、古くから研究が行われている。本稿では歴史的な考察は控えるが、ベンチマーキング法を体系的に紹介している文献として、Bloem et al. (2001) と Dagum and Cholette (2006) を挙げておく。なお、本稿では主に経済時系列におけるベンチマーク問題を念頭に置いているが、近年、小地域推定におけるベンチマーク問題が統計学の文献において関心を集めている。この方面への参考文献として、久保川 (2012) を挙げておく。

本稿では、線形計画法の一般論をしばしば用いる。本稿で用いる程度の一般論は、大抵の教科書ならカバーしているものであるが、線形計画法に詳しくない読者を想定して、必要な事実を付録にまとめておく。また、線形計画法の日本語での標準的な入門書として、今野 (1987) を挙げておく。

記号：本稿で用いる記号は標準的なものであるが、いくつかあらかじめ説明しておく。ベクトル・行列は太字で表し、ベクトルは列ベクトルとする。転置は'で表すとする。正の整数  $m$  に対して、 $\mathbf{0}_m, \mathbf{1}_m$  はそれぞれ 0 と 1 を  $m$  個並べたベクトルとする。また、 $m \times m$  の単位行列は  $\mathbf{I}_m$  と書く(原系列  $I_1, \dots, I_T$  と混同しないこと)。

## 2 デントン法とロバスト・デントン法

本節では、まずデントン法を紹介し、続いて提案するロバスト・デントン法を紹介する。記号は1節のものを踏襲するとする。また、

$$I_t \neq 0, t = 1, \dots, T$$

を以降特に言及することなく常に仮定する。

---

\*2 例えば、Matlab 上では linprog, R 上では lpSolve などが利用可能である。

## 2.1 デントン法

デントン法は、目的関数

$$\sum_{t=2}^T \left( \frac{X_t}{I_t} - \frac{X_{t-1}}{I_{t-1}} \right)^2 \quad (2)$$

をベンチマーク制約 (1) の下で最小化する  $X_1, \dots, X_T$  を調整済み系列として用いる手法である。ベクトル・行列を用いてデントン法の最適化問題を記述してみよう。いま、 $\mathbf{x} = (X_1, \dots, X_T)'$ ,  $\mathbf{b} = (A_1, \dots, A_n)'$  とし、 $(T-1) \times T$  行列  $\mathbf{P}$  と、 $n \times T$  行列  $\mathbf{Q}$  を

$$\mathbf{P} = \begin{pmatrix} -1/I_1 & 1/I_2 & 0 & 0 & \dots & 0 \\ 0 & -1/I_2 & 1/I_3 & 0 & \dots & 0 \\ \vdots & & & \ddots & & \\ 0 & \dots & \dots & 0 & -1/I_{T-1} & 1/I_T \end{pmatrix},$$

$$\mathbf{Q} = \mathbf{I}_n \otimes \mathbf{1}'_s = \begin{pmatrix} 1 & \dots & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & 0 & \dots & 0 \\ \vdots & & & & & \ddots & & & \\ 0 & \dots & 0 & 0 & \dots & 0 & 1 & \dots & 1 \end{pmatrix}$$

で定める。ただし、 $\otimes$  はクロネッカー積を表す<sup>\*3</sup>。すると、デントン法の最適化問題は

$$(D-1) \quad \min_{\mathbf{x} \in \mathbb{R}^T} \mathbf{x}' \mathbf{P}' \mathbf{P} \mathbf{x} \\ \text{s.t. } \mathbf{Q} \mathbf{x} = \mathbf{b}.$$

と書き直すことができる。この問題は 2 次計画問題であるが、国友・川崎 (2011) で示唆されているように、 $n > 1, s > 1$  のとき、ある技術的な仮定 (下記 (4) 式) の下で、問題 (D-1) は一意な最適解を持ち、かつ最適解は明示的な表現を持つ。

以下、デントン法の最適解を、国友・川崎 (2011), 定理 1 に従って明示的に導出してみ

<sup>\*3</sup>  $m \times n$  行列  $\mathbf{A} = (a_{ij})$  と  $k \times l$  行列  $\mathbf{B}$  のクロネッカー積は  $mk \times nl$  行列

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11} \mathbf{B} & \dots & a_{1n} \mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1} \mathbf{B} & \dots & a_{mn} \mathbf{B} \end{pmatrix}$$

で定義される。

よう。まず，上で述べた「ある技術的な仮定」を正確に述べる。  $s \times (s - 1)$  行列

$$\mathbf{J}_{s-1} = \begin{pmatrix} \mathbf{I}_{s-1} \\ -\mathbf{1}'_{s-1} \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \\ -1 & \dots & -1 \end{pmatrix} \quad (3)$$

を用意し，

$$\begin{pmatrix} I_1 \\ \vdots \\ I_T \end{pmatrix} \notin \text{Im}(\mathbf{I}_n \otimes \mathbf{J}_{s-1}) \quad (4)$$

を仮定する。ただし，右辺は，  $T \times (T - n)$  行列  $\mathbf{I}_n \otimes \mathbf{J}_{s-1}$  の  $(T - n)$  個の列ベクトルが張る，  $\mathbb{R}^T$  の部分ベクトル空間である。仮定 (4) は，例えば原系列  $I_1, \dots, I_T$  がすべて非負ならみだされるので，経済時系列への応用を考える場合，制約的な仮定ではないと思われる。仮定 (4) の下で，国友・川崎 (2011)，定理 1 の主張を示してみよう\*4。

基本的なアイデアは，デントン法の制約付き最小化問題を，変数  $\mathbf{x}$  に適当な正則変換をかけることで，  $(T - n)$  変数の制約なしの最小化問題に帰着させることである。いま，  $\mathbf{B} = \mathbf{P}'\mathbf{P}$  とおき，  $T \times T$  行列

$$\mathbf{R} = \begin{pmatrix} \mathbf{Q} \\ \mathbf{I}_n \otimes [\mathbf{I}_{s-1} \ \mathbf{0}_{s-1}] \end{pmatrix}$$

を考える。簡単に確認できるように，この行列  $\mathbf{R}$  は正則であり，

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \mathbf{R}\mathbf{x}, \quad (\mathbf{y}_1 \in \mathbb{R}^n, \mathbf{y}_2 \in \mathbb{R}^{T-n}),$$

と変換すると，  $\mathbf{Q}\mathbf{x} = \mathbf{b} \Leftrightarrow \mathbf{y}_1 = \mathbf{b}$  であるから，(D-1) は  $(T - n)$  変数の制約なしの最小化問題

$$(D-2) \quad \min_{\mathbf{y}_2 \in \mathbb{R}^{T-n}} (\mathbf{b}', \mathbf{y}_2')' \mathbf{R}'^{-1} \mathbf{B} \mathbf{R}^{-1} \begin{pmatrix} \mathbf{b} \\ \mathbf{y}_2 \end{pmatrix}$$

に帰着される。さらに，

$$\mathbf{R}^{-1} = [\mathbf{R}_{(1)} \ \mathbf{R}_{(2)}], \quad \mathbf{R}_{(1)} : T \times n, \quad \mathbf{R}_{(2)} : T \times (T - n)$$

と分割すると，(D-2) の目的関数は

$$\mathbf{y}_2' [\mathbf{R}'_{(2)} \mathbf{B} \mathbf{R}_{(2)}] \mathbf{y}_2 + 2\mathbf{y}_2' [\mathbf{R}'_{(2)} \mathbf{B} \mathbf{R}_{(1)}] \mathbf{b} + \mathbf{b}' [\mathbf{R}'_{(1)} \mathbf{B} \mathbf{R}_{(1)}] \mathbf{b}$$

\*4 細かい点であるが，国友・川崎 (2011)，定理 1 では仮定 (4) が抜けていることに注意する。仮定 (4) が成立しないとき，(5) 式において行列  $\mathbf{R}'_{(2)} \mathbf{B} \mathbf{R}_{(2)}$  が正則にならないので，最適解の一意性が成立しない。



と書けるから、これを最小にする  $\mathbf{y}_2$  は

$$\mathbf{y}_2^* = -[\mathbf{R}'_{(2)}\mathbf{B}\mathbf{R}_{(2)}]^{-1}[\mathbf{R}'_{(2)}\mathbf{B}\mathbf{R}_{(1)}]\mathbf{b}$$

で一意に与えられる。ここで、仮定 (4) の下で右辺の逆行列は存在する (後述)。もともとの変数に直すと、(D-1) の最適解が、

$$\mathbf{x}^* = \mathbf{R}^{-1}\mathbf{y}^* = \left( \mathbf{I}_T - \mathbf{R}_{(2)}[\mathbf{R}'_{(2)}\mathbf{B}\mathbf{R}_{(2)}]^{-1}\mathbf{R}'_{(2)}\mathbf{B} \right) \mathbf{R}_{(1)}\mathbf{b} \quad (5)$$

で一意に与えられることがわかる。さらに、国友・川崎 (2011)、補題 1 より、 $s \times 1$  ベクトル  $\mathbf{e}_s = (0, \dots, 0, 1)'$ 、と (3) 式で与えた  $s \times (s-1)$  行列  $\mathbf{J}_{s-1}$  を用いると、 $\mathbf{R}_{(1)}$  と  $\mathbf{R}_{(2)}$  は

$$\mathbf{R}_{(1)} = \mathbf{I}_n \otimes \mathbf{e}_s, \quad \mathbf{R}_{(2)} = \mathbf{I}_n \otimes \mathbf{J}_{s-1}, \quad (6)$$

と明示的に表現できる ( $\mathbf{R}[\mathbf{I}_n \otimes \mathbf{e}_s \quad \mathbf{I}_n \otimes \mathbf{J}_{s-1}] = \mathbf{I}_T$  を確認すればよい)。 (6) 式と (5) 式を合わせて、デントン法の最適解が完全に明示的に計算できることが確認できた。

最後に  $(T-n) \times (T-n)$  行列  $\mathbf{R}'_{(2)}\mathbf{B}\mathbf{R}_{(2)}$  が正則であることを確認しよう。 $\mathbf{R}'_{(2)}\mathbf{B}\mathbf{R}_{(2)} = (\mathbf{P}\mathbf{R}_{(2)})'(\mathbf{P}\mathbf{R}_{(2)})$  なので、行列  $\mathbf{P}\mathbf{R}_{(2)}$  の定める  $\mathbb{R}^{T-n}$  から  $\mathbb{R}^T$  への線形写像が単射であること、すなわち、

$$\text{Ker}(\mathbf{P}\mathbf{R}_{(2)}) = \{\mathbf{0}_{T-n}\} \quad (7)$$

を確認すればよい。  $\text{rank}\mathbf{P} = T-1$  と次元定理から、  $\dim(\text{Ker}\mathbf{P}) = 1$  であり、従って、簡単な計算から、

$$\text{Ker}\mathbf{P} = \left\{ c \begin{pmatrix} I_1 \\ \vdots \\ I_T \end{pmatrix} : c \in \mathbb{R} \right\}$$

であることがわかる。この事実と仮定 (4) より、(7) 式が従うことがわかる。実際は、もう少し強い主張

$$(4) \Leftrightarrow \mathbf{R}'_{(2)}\mathbf{B}\mathbf{R}_{(2)} \text{ が正則}$$

も従う。以上の事実をまとめて、次の補題を得る。

**補題 1** (国友・川崎 (2011)、定理 1 の修正)。  $s > 1, n > 1$  とする。仮定 (4) の下で、デントン法は一意的な最適解 (5) を持つ。ここで、  $T \times n$  行列  $\mathbf{R}_{(1)}$  と  $T \times (T-n)$  行列  $\mathbf{R}_{(2)}$  は (6) 式で与えられる。

デントン法の直感的な意味として、ベンチマーク制約 (1) の下で、なるべく

$$\frac{X_t}{I_t} \approx \frac{X_{t-1}}{I_{t-1}}, \text{ すなわち, } \frac{X_t}{X_{t-1}} \approx \frac{I_t}{I_{t-1}},$$

となるように調整済み系列を作成する手法と解釈することができる。つまり、デントン法はベンチマーク制約 (1) の下で、短期的な変動 ( $X_t/X_{t-1}$ ) をなるべく原系列のそれ ( $I_t/I_{t-1}$ ) に近いように調整する手法と解釈できる。GDP 統計を例にとると、GDP そのものの値 ( $X_t$ ) より、成長率

$$\frac{X_t - X_{t-1}}{X_{t-1}} = \frac{X_t}{X_{t-1}} - 1$$

に興味があるケースが多いので、短期の変動をなるべく保存するということは実務上望ましい性質といえるであろう。

**注意 1.** Denton (1971) はもともと、(2) ではなく、原系列の初期値  $I_0$  が利用可能として、目的関数

$$\sum_{t=1}^n \left( \frac{X_t}{I_t} - \frac{X_{t-1}}{I_{t-1}} \right)^2,$$

をベンチマーク制約 (1) と初期値制約  $X_0 = I_0$  の下で最小化することを提案した。本節で紹介したデントン法は、正確には Cholette (1984) による修正デントン法である。Cholette (1984) は初期値制約が調整済み系列にゆがみを生じさせることを指摘し、初期制約を取り除き、 $t = 1$  を目的関数の和から取り除いた修正法を提案した。

## 2.2 ロバスト・デントン法

本稿で提案するロバスト・デントン法は、デントン法の変形と考えることができる。形式的には、デントン法の目的関数を絶対値関数に変えた

$$\sum_{t=2}^T \left| \frac{X_t}{I_t} - \frac{X_{t-1}}{I_{t-1}} \right| \quad (8)$$

をベンチマーク制約 (1) の下で最小化する  $X_1, \dots, X_T$  を調整済み系列として用いる手法と定義される。ロバスト・デントン法も、デントン法と同様に、ベンチマーク制約 (1) の下で、短期的な変動 ( $X_t/X_{t-1}$ ) をなるべく原系列のそれ ( $I_t/I_{t-1}$ ) に近いように調整する手法と解釈できる。ただし、ロバスト・デントン法は詳細においてデントン法とは異なる性質を持つので、その点を見てみよう。

まず、ロバスト・デントン法の最適化問題は、線形制約の下で線形関数を最小化する問題なので、線形計画問題として定式化できる。前節で導入したベクトル・行列を用いると、ロバスト・デントン法の最適化問題は、次の線形計画問題として定式化できる：

$$\begin{aligned}
 \text{(RD-1)} \quad & \min_{\mathbf{u}, \mathbf{v} \in \mathbb{R}^{T-1}, \mathbf{x} \in \mathbb{R}^T} \mathbf{1}'_{T-1} \mathbf{u} + \mathbf{1}'_{T-1} \mathbf{v} \\
 \text{s.t.} \quad & \mathbf{u} - \mathbf{v} = \mathbf{P}\mathbf{x}, \quad \mathbf{Q}\mathbf{x} = \mathbf{b}, \\
 & \mathbf{u} \geq \mathbf{0}_{T-1}, \quad \mathbf{v} \geq \mathbf{0}_{T-1}.
 \end{aligned}$$

ただし、不等式  $\mathbf{u} \geq \mathbf{0}_{T-1}, \mathbf{v} \geq \mathbf{0}_{T-1}$  は成分ごとの不等式と了解する。さらに、

$$\mathbf{w} = (\mathbf{u}', \mathbf{v}', \mathbf{x}')', \quad \mathbf{c} = (\mathbf{1}'_{T-1}, \mathbf{1}'_{T-1}, \mathbf{0}'_T)', \quad \mathbf{A} = \begin{pmatrix} \mathbf{I}_{T-1} & -\mathbf{I}_{T-1} & \mathbf{P} \\ \mathbf{O} & \mathbf{O} & \mathbf{Q} \end{pmatrix}$$

とおくと、線形計画問題 (RD-1) は

$$\begin{aligned}
 \text{(RD-2)} \quad & \min_{\mathbf{w} = (\mathbf{u}', \mathbf{v}', \mathbf{x}')' \in \mathbb{R}^{3T-2}} \mathbf{c}'\mathbf{w} \\
 \text{s.t.} \quad & \mathbf{A}\mathbf{w} = (\mathbf{0}'_{T-1}, \mathbf{b}')', \quad \mathbf{u} \geq \mathbf{0}_{T-1}, \quad \mathbf{v} \geq \mathbf{0}_{T-1}
 \end{aligned}$$

と書き直すことができる。制約行列  $\mathbf{A}$  のサイズは  $(T-1+n) \times (3T-2)$  であり、容易にわかるように、

$$\text{rank} \mathbf{A} = T-1+n$$

である。(RD-2) において、集合

$$\{\mathbf{w} = (\mathbf{u}', \mathbf{v}', \mathbf{x}')' \in \mathbb{R}^{3T-2} : \mathbf{A}\mathbf{w} = (\mathbf{0}'_{T-1}, \mathbf{b}')', \quad \mathbf{u} \geq \mathbf{0}_{T-1}, \quad \mathbf{v} \geq \mathbf{0}_{T-1}\}$$

は実行可能集合と呼ばれるが、実行可能集合は自明に非空であり、また、目的関数は実行可能集合上で下に有界である(目的関数は非負であるから)。従って、線形計画法の一般論から、次の補題が従う。

**補題 2.** 線形計画問題 (RD-2) は最適解を持つ。従って、ロバスト・デントン法は必ず最適解を持つ。

ロバスト・デントン法は、デントン法とは異なり、最適解の明示的な表現は得られないが、最適化問題が線形計画問題として定式化できるので、その実行は容易である。また、ロバスト・デントン法は、デントン法にはない、実務上興味深い性質を持つ。この点を見ていこう。いま、(RD-2) を単体法を用いて解くことを考え、最適な基底解  $\mathbf{w}^*$  が得られたとする。 $\mathbf{w}^*$  の  $\mathbf{u}$  成分、 $\mathbf{v}$  成分、 $\mathbf{x}$  成分をそれぞれ  $\mathbf{u}^*, \mathbf{v}^*, \mathbf{x}^*$  とおく。いま、 $\mathbf{x}^*$  の各成

分は非ゼロであると仮定すると、 $\mathbf{u}^*, \mathbf{v}^*$  に含まれる変数のうち基底変数となるものはちょうど  $(n-1)$  個である。ここで、

$$\mathbf{x}^* = (X_1^*, \dots, X_T^*)', \mathbf{u}^* = (u_1^*, \dots, u_{T-1}^*)', \mathbf{v}^* = (v_1^*, \dots, v_{T-1}^*)'$$

と書いて、 $u_t^*$  と  $v_t^*$  がともに非基底変数であるとする、 $u_t^* = v_t^* = 0$  であって、線形制約  $\mathbf{A}\mathbf{w} = (\mathbf{0}'_{T-1}, \mathbf{b}')'$  の第  $t$  式に注目すると、

$$-\frac{X_{t-1}^*}{I_{t-1}} + \frac{X_t^*}{I_t} = 0 \Leftrightarrow \frac{X_t^*}{X_{t-1}^*} = \frac{I_t}{I_{t-1}}$$

となることがわかる。 $u_t^*$  と  $v_t^*$  がともに非基底変数であるような  $t$  はちょうど  $(T-n)$  個あるから ( $u_t^*$  と  $v_t^*$  が同時に基底変数になることはないから)、次の補題が示された。

**補題 3.** (RD-2) の基底解に対応するロバスト・デントン法の最適解を  $X_1^*, \dots, X_T^*$  とする。いま、 $X_t^* \neq 0, t = 1, \dots, T$  ならば、少なくとも  $(T-n)$  個の  $t \in \{2, \dots, T\}$  に対して、

$$\frac{X_t^*}{X_{t-1}^*} = \frac{I_t}{I_{t-1}},$$

が成立する。

**注意 2.** 補題 3 はロバスト・デントン法が短期の変動をかなりの部分厳密 (exact) に保存すること示している。このような性質はデントン法にはないものであり、実務上利点があるものと思われる。ここで、ロバスト・デントン法は、デントン法と類似の目的関数を最小化していることから、(プロ・ラタ法とは異なり) 年次が変わるところでも変動が原系列のそれとなるべく乖離しないように調整している、ということに注意すべきである。

**注意 3.** デントン法と同様に、目的関数 (8) に非負のウェイト  $d_t \geq 0, t = 2, \dots, T$  をつけて、

$$\sum_{t=2}^T d_t \left| \frac{X_t}{I_t} - \frac{X_{t-1}}{I_{t-1}} \right|,$$

をベンチマーク制約 (1) の下で最小化することも考えられる。ただし、ウェイトの選び方は自明な問題ではない。本稿では (簡単のため) 一貫して定数ウェイト  $d_t \equiv 1$  を考える。

**注意 4.** ロバスト・デントン法という名称は、ロバスト統計から拝借したものである。しばしば指摘されるように、線形回帰モデルの係数ベクトルの推定問題において、最小二乗推定値は誤差項の外れ値に大きく影響される。外れ値に大きく影響されないような推定法 (そのような推定法をしばしばロバストな推定法と呼ぶ) の一つとして、最小絶対偏差推定

(least absolute deviation estimation) が知られおり，これは最小二乗推定の目的関数を絶対値関数に置き換えた推定法である（最小絶対偏差推定については Koenker, 2005, が参考になる）。

以下，ロバスト・デントン法に関して，補助的な議論を行う。

### 2.2.1 原系列が非負であるとき

特に経済時系列の場合，原系列が非負であることが多い．そのような場合，調整済み系列も非負であることが望まれる．目的関数 (8) をベンチマーク制約 (1) と非負制約  $X_t \geq 0, t = 1, \dots, T$  の下で最小化する  $X_1, \dots, X_T$  を調整済み系列として用いる手法を便宜的に非負制約付きロバスト・デントン法と呼ぶことにする．非負制約付きロバスト・デントン法の最適化問題は，(RD-2) において非負制約  $\mathbf{x} \geq \mathbf{0}_T$  を加えた次の線形計画問題に帰着する\*5：

$$(RD-3) \quad \min_{\mathbf{w}=(\mathbf{u}', \mathbf{v}', \mathbf{x}')' \in \mathbb{R}^{3T-2}} \mathbf{c}'\mathbf{w}$$

$$\text{s.t. } \mathbf{A}\mathbf{w} = (\mathbf{0}'_{T-1}, \mathbf{b}')', \mathbf{w} \geq \mathbf{0}_{3T-2}.$$

補題 2 と 3 は次のように変更される．証明は，補題 2 と 3 の証明よりほとんど自明であるので，省略する。

**補題 4.** 線形計画問題 (RD-3) は最適解を持つ．従って，非負制約付きロバスト・デントン法は必ず最適解を持つ．

**補題 5.** (RD-3) の基底解に対応する非負制約付きロバスト・デントン法の最適解を  $X_1^*, \dots, X_T^*$  とする．いま， $X_t^* > 0, t = 1, \dots, T$  ならば，少なくとも  $(T - n)$  個の  $t \in \{2, \dots, T\}$  に対して，

$$\frac{X_t^*}{X_{t-1}^*} = \frac{I_t}{I_{t-1}}$$

が成立する\*6．

**注意 5.** もちろん，実務上，非負制約を明示的に入れなくても，自動的に非負制約がみたされる場合も多いであろう．その様な場合，非負制約を入れるか入れないかは本質的では

\*5 むしろ (RD-3) の方が標準形である分，(RD-2) よりある意味で簡単である，ともいえる。

\*6 あくまでも数学的な注意であるが， $X_t^* > 0, t = 1, \dots, T$  という仮定から，右辺が負の値になるケースは自動的に排除されている．補題では  $I_t > 0, t = 1, \dots, T$  とは仮定していないことに注意せよ（もちろん非負制約付きロバスト・デントン法を用いるのは， $I_t > 0, t = 1, \dots, T$  となるときであろうから，この注意はあくまで数学的な注意である）。

ない。しかしながら、重要であるのは、ロバスト・デントン法が、必要ならば非負制約を容易に計算に組み込むことができる、というオプションを与えていることにある。例えば、デントン法では非負制約を明示的に入れることはそれほど自明でないことと思われる。

## 2.2.2 双対問題

線形計画問題 (RD-2) と (RD-3) を解くには、対応する双対問題を解けばよいことは線形計画法の一般論の教えるところである。本節では、これら線形計画問題の双対問題を見てみる。簡単のため、(RD-3) の双対問題を見てみる：

$$\begin{aligned} \text{(Dual of RD-3)} \quad & \max_{z \in \mathbb{R}^{T-1+n}} (\mathbf{0}'_{T-1}, \mathbf{b}')z \\ & \text{s.t. } \mathbf{A}'z \leq \mathbf{c}. \end{aligned}$$

ここで、 $z = (z'_1, z'_2)'$ ,  $z_1 \in \mathbb{R}^{T-1}$ ,  $z_2 \in \mathbb{R}^n$  と分割すると、

$$\mathbf{A}'z = \begin{pmatrix} \mathbf{I}_{T-1} & \mathbf{O} \\ -\mathbf{I}_{T-1} & \mathbf{O} \\ \mathbf{P}' & \mathbf{Q}' \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ -z_1 \\ \mathbf{P}'z_1 + \mathbf{Q}'z_2 \end{pmatrix}$$

であるから、

$$\mathbf{A}'z \leq \mathbf{c} \Leftrightarrow -\mathbf{1}_{T-1} \leq z_1 \leq \mathbf{1}_{T-1}, \mathbf{P}'z_1 + \mathbf{Q}'z_2 \leq \mathbf{0}_T.$$

従って、結局、(Dual of RD-3) は

$$\begin{aligned} & \max_{z=(z'_1, z'_2)' \in \mathbb{R}^{T-1+n}} \mathbf{b}'z_2 \\ & \text{s.t. } \mathbf{P}'z_1 + \mathbf{Q}'z_2 \leq \mathbf{0}_T, -\mathbf{1}_{T-1} \leq z_1 \leq \mathbf{1}_{T-1} \end{aligned}$$

と書き直すことができる。  $z$  を  $-z$  と取り換えると、

$$\begin{aligned} & \min_{z=(z'_1, z'_2)' \in \mathbb{R}^{T-1+n}} \mathbf{b}'z_2 \\ & \text{s.t. } \mathbf{P}'z_1 + \mathbf{Q}'z_2 \geq \mathbf{0}_T, -\mathbf{1}_{T-1} \leq z_1 \leq \mathbf{1}_{T-1} \end{aligned}$$

とも書き直すことができる。

## 付録 A 線形計画法速習

本節では、線形計画法に詳しくない読者を想定して、本稿で用いた線形計画法の基本的な事実を (多くは証明抜きで) 紹介する。日本語での標準的な線形計画法の教科書として

今野 (1987) を挙げておく。線形計画問題とは、目的関数と制約が変数に関して線形であるような最適化問題である。  $n$  変数と  $m$  制約を持つ標準形の線形計画問題

$$(P-1) \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}'\mathbf{x} \\ \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$$

を考えてみよう。ここで、 $\mathbf{c}, \mathbf{A}, \mathbf{b}$  はそれぞれ、 $n \times 1$  ベクトル、 $m \times n$  行列、 $m \times 1$  ベクトルである。 $\mathbf{c}, \mathbf{A}, \mathbf{b}$  はすべて既知である。一般性を失うことなく

$$\text{rank} \mathbf{A} = m \tag{9}$$

を仮定する ( $\text{rank} \mathbf{A} < m$  ということは線形制約の中に“無駄な”ものがあるということであり、そういうものはあらかじめ排除しておく、ということである)。不等式の線形制約や、 $\mathbf{x}$  (の一部) に非負制約がない場合も、適当なスラック変数を導入することで、上記の標準形に変形できるので (例えば、 $\mathbf{x}$  に非負制約がないときは、 $\mathbf{x} = \mathbf{x}_+ - \mathbf{x}_-$ ,  $\mathbf{x}_+ \geq \mathbf{0}, \mathbf{x}_- \geq \mathbf{0}$  として、新しく非負変数  $\mathbf{x}_+$  と  $\mathbf{x}_-$  を導入すればよい)、以下、問題 (P-1) を考える。

1. 基本定理. 問題 (P-1) において、制約をみたす変数の集合

$$S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

を実行可能集合と呼ぶ。 $S$  が空集合でないとき、(P-1) は実行可能と呼び、 $S$  に属するベクトルを実行可能解と呼ぶ。次の定理は基本的である。

**定理 1.** (P-1) が実行可能で目的関数が  $S$  上で下に有界ならば、最適解が存在する。

この定理より、線形計画問題 (P-1) に対して、次の 3 ケースのみが起こりうるようになる。

- (i) (P-1) は実行不可能である、つまり、 $S$  は空集合である。
- (ii) (P-1) は実行可能であるが、 $S$  上でいくらでも目的関数を小さくできる。
- (iii) (P-1) は実行可能で、最適解が存在する。

2. 単体法. 以下、(P-1) は少なくとも実行可能と仮定する。ところで、(9) より  $m \leq n$  であるが、 $m = n$  のときは実行可能集合が高々 1 点となるのでおもしろくない。よって、少なくとも  $m < n$  を仮定する。 $\mathbf{A}$  から任意に  $m$  列を選んでそれらを並べた  $m \times m$  行列を  $\mathbf{B}$  と書き、便宜的に  $\mathbf{A}$  を

$$\mathbf{A} = [\mathbf{B} \ \mathbf{N}]$$

と分割する．ここで， $\mathbf{N}$  は  $m \times (n - m)$  行列である．この分割に対応して，ベクトル  $\mathbf{c}$  と  $\mathbf{x}$  を

$$\mathbf{c} = (\mathbf{c}'_B, \mathbf{c}'_N)', \quad \mathbf{x} = (\mathbf{x}'_B, \mathbf{x}'_N)'$$

と分割する．いま，行列  $B$  は正則であると仮定すると，

$$\mathbf{Ax} = \mathbf{b} \Leftrightarrow \mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N$$

であるから，問題 (P-1) は

$$\begin{aligned} \text{(P-2)} \quad & \min_{\mathbf{x}=(\mathbf{x}'_B, \mathbf{x}'_N)' \in \mathbb{R}^n} \mathbf{c}'_B \mathbf{B}^{-1}\mathbf{b} + [\mathbf{c}'_N - \mathbf{N}'(\mathbf{B}')^{-1}\mathbf{c}'_B]'\mathbf{x}_N \\ \text{s.t.} \quad & \mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N, \quad \mathbf{x}_B \geq \mathbf{0}, \mathbf{x}_N \geq \mathbf{0} \end{aligned}$$

と変形できる．このとき， $B$  を基底行列， $N$  を非基底行列，変数  $\mathbf{x}_B$  を基底変数，変数  $\mathbf{x}_N$  を非基底変数，

$$\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}, \quad \mathbf{x}_N = \mathbf{0} \tag{10}$$

を基底解と呼ぶ．基底解が実行可能であるとき，すなわち，

$$\mathbf{B}\mathbf{b}^{-1} \geq \mathbf{0}$$

のとき，(10) を実行可能基底解と呼ぶ．実は (P-1) を解くためには，基底解の中から最適解を見つければよい．

**定理 2.** (P-1) が最適解をもつならば，基底解の中に最適解が存在する．

最適な基底解をうまく探すアルゴリズムとして，単体法 (simplex method) が知られている．ところで，(10) が実行可能であって，(P-2) において

$$\bar{\mathbf{c}} := \mathbf{c}'_N - \mathbf{N}'(\mathbf{B}')^{-1}\mathbf{c}'_B \geq \mathbf{0}$$

であるなら， $\mathbf{x}_N$  の各成分を 0 から増加させると，目的関数の値が増加するので，基底解 (10) が最適解となる．単体法とは，初期実行可能基底解から出発して， $\bar{\mathbf{c}} \geq \mathbf{0}$  となる基底変数の組  $\mathbf{x}_B$  を見つけるアルゴリズムと理解できる．具体的には，以下のステップからなる．

### 単体法

Step 1. 実行可能基底解 (10) を入力する．

Step 2.  $\bar{\mathbf{c}} \geq \mathbf{0}$  なら現在の基底解が最適解．そうでないなら，(P-2) において負の係数を持つ非基底変数  $x_s$  を一つ選ぶ．



Step 3. 非基底変数のうち,  $x_s$  だけ増加させて, 最初に 0 となる基底変数  $x_r$  を求める (基底変数と非基底変数の間には  $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N$  なる関係は成立していることに注意). このような  $x_r$  が存在するときは,  $x_s$  を新しく基底変数に,  $x_r$  を新しく非基底変数に入れ, Step 2 に戻る. そうでないときは, いくらでも  $x_s$  を大きくすることで, いくらでも目的関数を小さくすることができるので, 最適解は存在しない.

Step 3 において基底変数の入れ替えをピボットという. ピボットによって  $\mathbf{B}$  も変わるが, 正則であることは保たれる. 単体法はピボット規則の分だけ任意性があるが, Bland の最小添え字規則なるピボット規則を用いると, 必ず有限回で収束することが示される.

ところで, 幾何学的な観点からは, 単体法は多面集合である実行可能集合の頂点の中から最適解を探すアルゴリズムと解釈できる. 単体法とは異なるアプローチとして, 実行可能集合の内部を経由して最適解を探す内点法 (interior point method) も知られているが, 本節では触れない. 今野 (1987) の 14 章, または小島ほか (2001) を参照せよ.

3. 双対理論. (P-1) の双対問題とは

$$(D-1) \max_{\mathbf{y} \in \mathbb{R}^m} \mathbf{b}'\mathbf{y}$$

$$\text{s.t. } \mathbf{A}'\mathbf{y} \leq \mathbf{c}$$

なる線形計画問題である. もとの問題 (P-1) のことを主問題と呼ぶ. ところで,  $\mathbf{x}$  と  $\mathbf{y}$  をそれぞれ (P-1) と (D-1) の実行可能解とすると,

$$\mathbf{b}'\mathbf{y} = (\mathbf{A}\mathbf{x})'\mathbf{y} = (\mathbf{A}\mathbf{y})'\mathbf{x} \leq \mathbf{c}'\mathbf{x}$$

が成り立つ. ここで, 最後の不等式は,  $\mathbf{A}\mathbf{y} \leq \mathbf{c}$  と  $\mathbf{x} \geq \mathbf{0}$  から従う.

**定理 3 (弱双対定理).**  $\mathbf{x}$  と  $\mathbf{y}$  をそれぞれ (P-1) と (D-1) の実行可能解とすると,  $\mathbf{b}'\mathbf{y} \leq \mathbf{c}'\mathbf{x}$ .

いま, 単体法を実行して, (P-1) の最適な基底解  $\mathbf{x}^*$  が得られたとし, 対応する基底行列と非基底行列をそれぞれ  $\mathbf{B}$  と  $\mathbf{N}$  と書くことにする. このとき,  $\mathbf{c}_N \geq \mathbf{N}'(\mathbf{B}')^{-1}\mathbf{c}_B$  となっていることに注意する. ここで,

$$\mathbf{y}^* = (\mathbf{B}')^{-1}\mathbf{c}_B$$

とおくと,

$$\mathbf{A}'\mathbf{y}^* = \mathbf{A}'(\mathbf{B}')^{-1}\mathbf{c}_B = \begin{pmatrix} \mathbf{B}' \\ \mathbf{N}' \end{pmatrix} (\mathbf{B}')^{-1}\mathbf{c}_B = \begin{pmatrix} \mathbf{c}_B \\ \mathbf{N}'(\mathbf{B}')^{-1}\mathbf{c}_B \end{pmatrix} \leq \mathbf{c}$$

であるから、 $\mathbf{y}^*$  は (D-1) の実行可能解であり、

$$\mathbf{b}'\mathbf{y}^* = \mathbf{c}'_B \mathbf{B}^{-1} \mathbf{b} = \mathbf{c}'\mathbf{x}^*$$

であるから、弱双対定理より、 $\mathbf{y}^*$  は (D-1) の最適解となることがわかる。以上の議論をまとめると、次の定理を得る。

**定理 4.** (P-1) の最適な基底解  $\mathbf{x}^*$  が得られたとし、対応する基底行列  $\mathbf{B}$  と書く。すると、 $\mathbf{y}^* = (\mathbf{B}')^{-1} \mathbf{c}_B$  は (D-1) の最適解となる。

とくに、(P-1) が最適解を持てば最適な基底解をもつので、(D-1) の最適解が存在し、(P-1) と (D-1) の最適値は一致する。また、(D-1) をスラック変数を用いて標準系に変形し、双対問題を考えると、主問題 (P-1) に一致することが確認できる（つまり、双対問題の双対問題は主問題）、双対問題 (D-1) が最適解を持てば、主問題 (P-1) の最適解が存在し、(P-1) と (D-1) の最適値は一致する。

## 参考文献

- Bloem, A.M., Dippelsman, R.J. and Maehle, N.O. (2001). *Quarterly National Accounts Manual: Concepts, Data Sources and Compilation*. International Monetary Fund.
- Cholette, P.A. (1984). Adjusting sub-annual series to yearly benchmarks. *Survey Methodology* **10** 35-49.
- Dagum, E.B. and Cholette, P.A. *Benchmarking, Temporal Distributions, and Reconciliation Methods for Time Series*. Lecture Notes in Statistics, Springer.
- Denton, F. (1971). Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization. *J. Amer. Stat. Assoc.* **66** 99-102.
- Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.
- 久保川達也 (2012). 小地域推定におけるベンチマーク問題と制約付き経験ベイズ推定の誤差評価. CIRJE-R-12 「経済統計・政府統計の数理的基礎と応用-II」(国友直人・山本拓 共編) 所収 pp. 143-146.
- 国友直人編 (2004). 「解説 X-12ARIMA」 CIRJE-R-1.
- 国友直人・川崎能典 (2011). ベンチマーク問題と経済時系列. CIRJE-J-234.
- 小島政和, 土谷隆, 水野眞治, 矢部博. (2001). 「内点法」朝倉書店.
- 今野浩 (1987). 「線形計画法」日科技連.

得能雅之 (2007). GDP の値はなぜ改定されるのか～速報・確報の作成方法について. 季刊国民経済計算 No.134 1-6.

内閣府 (2011). GDP 統計の在り方の検討結果 (短期的課題) について. <http://www.esri.cao.go.jp/jp/sna/seibi/suikai/suikai.html> より利用可能.

# ON OFFICIAL SEASONAL ADJUSTMENTS IN JAPAN

Makoto Takaoka

University of the Ryukyus

## INTRODUCTION

- Many of Japan's official statistics are seasonally adjusted using the X-12-ARIMA software suite developed by the United States Census Bureau.
- Today's talk is about the relationship between **seasonally adjusted series revisions** and **X-12-ARIMA model selection** in Japan.

# OUTLINE

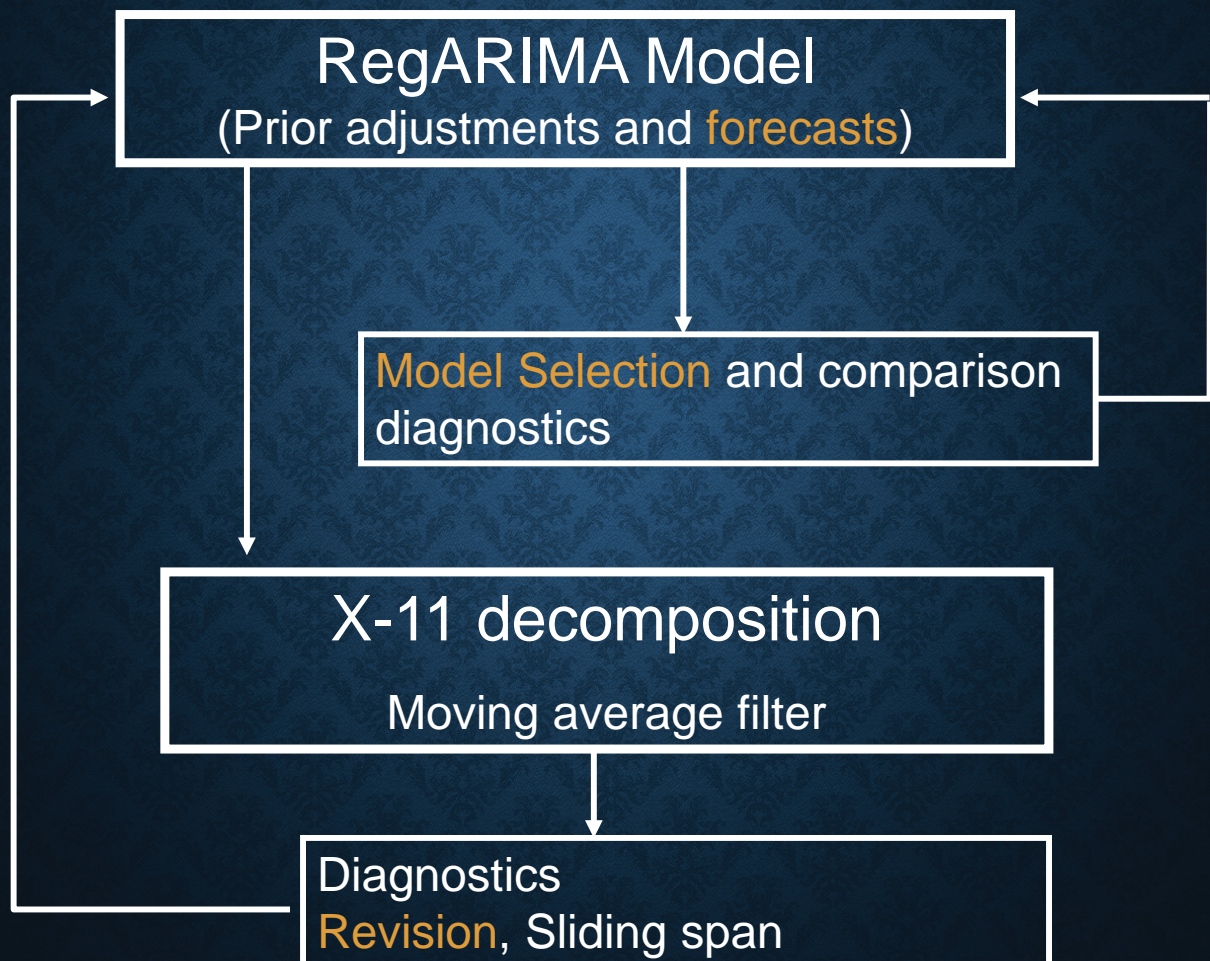
- Basic X-12-ARIMA concepts
- Seasonal adjustment revisions
- Overview of current practices in Japan
  - The case of Financial Statements Statistics of Corporations by Industry (Ministry of Finance)
- Some simulation results
- Concluding remarks

## X-12-ARIMA

- Seasonal adjustments in X-12-ARIMA proceed in the following order
  - (1) Prior adjustments: RegARIMA model
  - (2) X-11 decomposition
  - (3) Diagnostics

# REGARIMA MODEL

- Detects and adjusts for outliers in order to improve the forecasts and seasonal adjustments
- Estimates missing data
- Detects and directly estimates trading day effects and other effects
- **Extends the series with forecasts**



# STABILITY OF SEASONAL ADJUSTMENTS

Revisions associated with continuous seasonal adjustments over a period of years.

$Y_t$  : Unadjusted series

$A_{t|s}$  : Seasonally adjusted value at time  $t$  given observations up to and including at time  $s$

$R_{t|s}$  : Percent change at time  $t$  given observations up to and including at time  $s$

$$R_{s|t} = \frac{A_{s|t} - A_{s-1|t}}{A_{s-1|t}} \times 100 \quad (s \leq t)$$

$R_{t|t}$  : Concurrent percent change

## STABILITY OF THE SEASONAL ADJUSTMENT(2)

- If the differences among  $R_{t|t}, R_{t|t+1}, R_{t|t+2}, \dots$  are small, the seasonal adjustment is stable.

# SEASONAL ADJUSTMENT REVISIONS

- All forecasts generated from the estimated RegARIMA model contain statistical uncertainties.
- If, as a result, the forecasts are close to the observed value, the seasonal adjustment revision is small.
- Revisions are inevitable within the X-12-ARIMA framework, but choosing the most appropriate settings can reduce them.

## REVISION HISTORY FUNCTION OF X-12-ARIMA

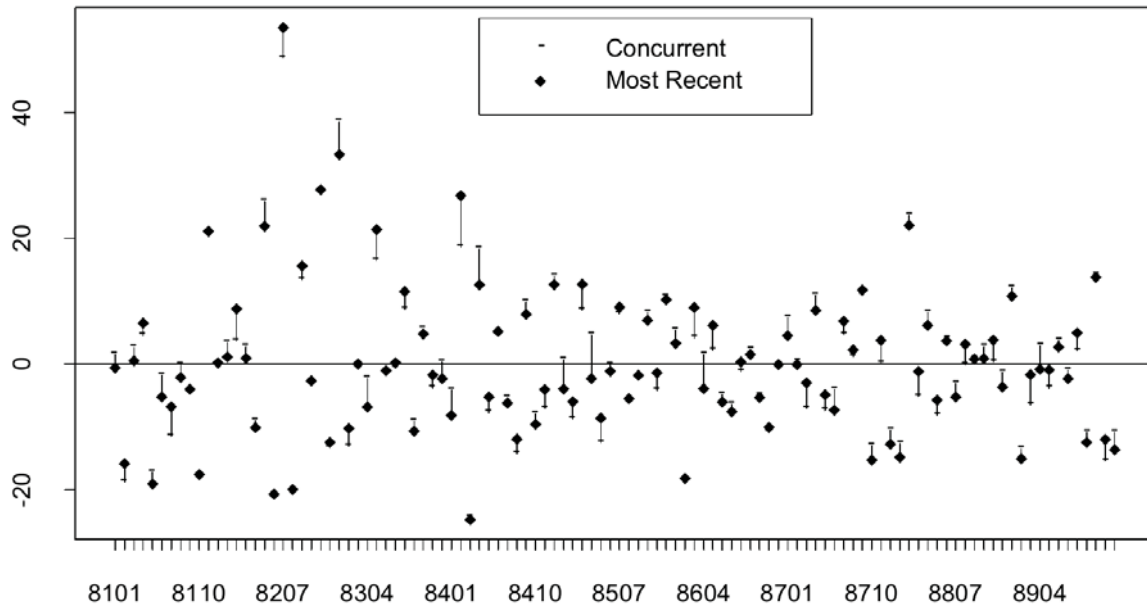
- X-12-ARIMA includes two types of stability diagnostics, sliding spans and **revision histories**

$R_{t|N} - R_{t|t}$  ( $t = 1 \dots N$ ) : **Revision from the concurrent and the most recent percent change for time  $t$**

Revision function built into X-12-ARIMA can be used to select an option with high stability.



(a) Seasonal Adjustment Changes from the Previous Month.



## PROBLEMS IN PRACTICE

- In practice, it is not always easy to integrate X-12-ARIMA revision history results into settings.
- If the statistical properties of the data are not stable, setting changes are often necessary.
- Capturing the statistical properties around the most recent data is important for short-term predictions.

## PROBLEMS IN PRACTICE (2)

- In the X-12-ARIMA revision history, revisions due to model reselection are not considered.
- In situations where data increases sequentially, it is desirable to check models periodically.
- Most agencies let the software choose the options, and then fix the settings for production.

## OFFICIAL STATISTICS IN JAPAN

- In major official statistics released by Japan, X-12-ARIMA spec-files are published.
- In some statistics, the revision policy (model selection policy) is also published.
- In several official statistics released by Japan, AIC is used for model selection in X-12-ARIMA.

**National Accounts, Financial Statements Statistics of Corporations, etc.**

Information of statistics whose revision policy is published are summarized in these three tables.  
As shown in these tables, model selection policy and treatment of outliers seem to be different little by little.

Ministry and Agency	Title	Software	ARIMA Model	Outlier
<b>Cabinet Office</b>	Explanation of Orders Received for Machinery	X-12-ARIMA	(1)AIC, (2)MAPR	Auto Detection(AO,TC
	Consumer Confidence Survey	X-12-ARIMA	(1)AIC, (2)MAPR	Auto Detection(AO,TC
	Indexes of Business Conditions	X-12-ARIMA	(1)AIC, (2)MAPR	Auto Detection(AO,TC
	Quarterly Estimates of GDP	X-12-ARIMA	AIC (chose among 81 models {(010)(010)-(212)(212)})	Examine individually
<b>Ministry of Internal Affairs and Communications (MIC)</b>	Labour Force Survey	X-12-ARIMA	AIC (chose among 81 models {(010)(010)-(212)(212)})	AO,LS, Ramp
	Family Income and Expenditure Survey	X-12-ARIMA	AIC (chose among 81 models {(010)(010)-(212)(212)})	Auto Detection
	Consumer Price Index	X-12-ARIMA	AIC	Auto Detection

Ministry and Agency	Title	Software	ARIMA Model	Outlier
<b>Ministry of Finance</b>	Financial Statements Statistics of Corporations by Industry	X-12-ARIMA	Penalized AIC (chose among 81 models {(010)(010)-(212)(212)})	AO, Ramp
	Trade Statistics	X-12-ARIMA	AIC (chose among 81 models {(010)(010)-(212)(212)})	AO, Ramp
	Balance of Payments	X-12-ARIMA	AIC (chose among 81 models {(010)(010)-(212)(212)})	Auto Detection
<b>Minister of Economy, Trade and Industry</b>	Indices of Industrial Production	X-12-ARIMA	AIC (chose among 57 models)	
	Indices of Tertiary Industry Activity	X-12-ARIMA	BIC (chose among 57 models)	
	Indices of All Industry Activity	X-12-ARIMA	BIC (chose among 57 models)	

Ministry and Agency	Title	Software	ARIMA Model	Outlier
Ministry of Land, Infrastructure, Transport and Tourism	Building Starts	X-12-ARIMA	AIC (chose among 81 models {(010)(010)-(212)(212)})	Auto Detection(AO,TC)
	Current Survey on Orders Received for Construction	X-12-ARIMA	AUTOMDL, AIC	
Bank of Japan	Money Stock	X-12-ARIMA	ACF, PACF, AIC, Ljung-Box Q-test	Auto Detection(AO,TC) LS, Ramp
	Monetary Base	X-12-ARIMA	ACF, PACF, AIC, Ljung-Box Q-test	Auto Detection(AO,TC) LS, Ramp
	Real Exports and Real Imports	X-12-ARIMA	ACF, PACF, AIC, Ljung-Box Q-test	Auto Detection(AO,TC) LS, Ramp

## FINANCIAL STATEMENTS STATISTICS OF CORPORATIONS BY INDUSTRY (MINISTRY OF FINANCE)

(Reference) Percent changes (seasonally adjusted, Sales)

Classification	2012					2013
	Jul.-Sep.	Oct.-Dec.	Jan.-Mar.	Apr.-Jun.	Jul.-Sep.	
All Industries	-0.3	-1.3	0.5	0.8	0.9	
Manufacturing	-1.7	-1.3	0.4	-0.5	1.6	
Non-Manufacturing	0.3	-1.3	0.5	1.4	0.6	

Notes 1 : Percent change is compared with the previous quarter.

2 : All Industries and Non-Manufacturing don't include Finance and Insurance.

(Reference) Percent changes (seasonally adjusted, Ordinary profits)

Classification	2012					2013
	Jul.-Sep.	Oct.-Dec.	Jan.-Mar.	Apr.-Jun.	Jul.-Sep.	
All Industries	1.1	5.0	7.0	9.6	-1.6	
Manufacturing	4.2	15.4	17.2	10.3	-5.2	
Non-Manufacturing	-0.2	0.4	1.8	9.2	0.4	

Notes 1 : Percent change is compared with the previous quarter.

2 : All Industries and Non-Manufacturing don't include Finance and Insurance.

(Reference) Percent changes (seasonally adjusted, excluding investment in software)

Classification	2012					2013
	Jul.-Sep.	Oct.-Dec.	Jan.-Mar.	Apr.-Jun.	Jul.-Sep.	
All Industries	-2.2	0.7	0.1	2.7	-0.5	
Manufacturing	-3.4	-4.1	-0.9	-1.0	-0.7	
Non-Manufacturing	-1.5	3.5	0.7	4.6	-0.4	

Notes 1 : Percent change is compared with the previous quarter.

2 : All Industries and Non-Manufacturing don't include Finance and Insurance.

# SEASONAL ADJUSTMENT POLICY

- Seasonal adjustment practices are applied
  - Models are kept fixed for one year, but their parameters are re-estimated in each calculation round.
  - Models used are checked once each year.
  - **The model that has the lowest AIC is selected from 81 candidates.**

Candidates						
	AR	D	MA	ARs	SD	SMA
Order	0 - 2	1	0 - 2	0 - 2	1	0 - 2

$$3 \times 3 \times 3 \times 3 = 81$$

	Model	$R_{t-4 t}$ 10/7-9	$R_{t-3 t}$ 10/10-12	$R_{t-2 t}$ 11/1-3	$R_{t-1 t}$ 11/4-6	$R_{t t}$ 11/7-9
Sales (Manufacturing)	CURRENT	-6.62	0.54	1.71	-8.25	4.85
	NEW	-6.89	0.40	2.43	-8.45	4.43
Sales (Non-manufacturing)	CURRENT	-9.82	1.84	-0.50	-3.42	0.72
	NEW	-10.42	2.49	-0.77	-3.00	-0.24
Ordinary profits (Manufacturing)	CURRENT	2.85	0.67	-4.71	-11.93	-0.78
	NEW	3.08	0.32	-8.17	-7.88	-1.12
Ordinary profits (Non-manufacturing)	CURRENT	-5.27	12.21	-7.77	-11.01	5.75
	NEW	-4.86	12.00	-7.48	-11.65	6.51
Capital investment (Manufacturing)	CURRENT	4.50	4.06	0.67	-9.77	2.99
	NEW	5.05	3.45	1.54	-10.67	3.82
Capital investment (Non-manufacturing)	CURRENT	-1.22	-2.90	-3.60	-4.32	-5.84
	NEW	-1.22	-2.90	-3.60	-4.32	-5.84

# PENALTY OF REVISION

$$AR_m^{(k)} = \frac{1}{m} \sum_{j=0}^{m-1} \left| R_{t-j|t}^c - R_{t-j|t}^{(k)} \right| \quad (1 \leq k \leq K)$$

$AR_m^{(k)}$  is the **average revision** of model  $k$  percent change .  
The model that minimizes  $AR_m^{(k)}$  is the current model.

$A_{s|t}^c$  : Seasonally adjusted value calculated using the current model

$R_{s|t}^c$  : Percent change of  $A_{s|t}^c$

$$R_{s|t}^c = \frac{A_{s|t}^c - A_{s-1|t}^c}{A_{s-1|t}^c} \times 100 \quad (s \leq t)$$

$K$  : Number of candidate models

$A_{s|t}^{(k)}$  ( $k = 1, \dots, K$ ) : Seasonally adjusted value calculated using model  $k$

$R_{s|t}^{(k)}$  : Percent change of  $A_{s|t}^{(k)}$

$$R_{s|t}^{(k)} = \frac{A_{s|t}^{(k)} - A_{s-1|t}^{(k)}}{A_{s-1|t}^{(k)}} \times 100 \quad (s \leq t)$$

# PENALIZED MODEL SELECTION PROCEDURE

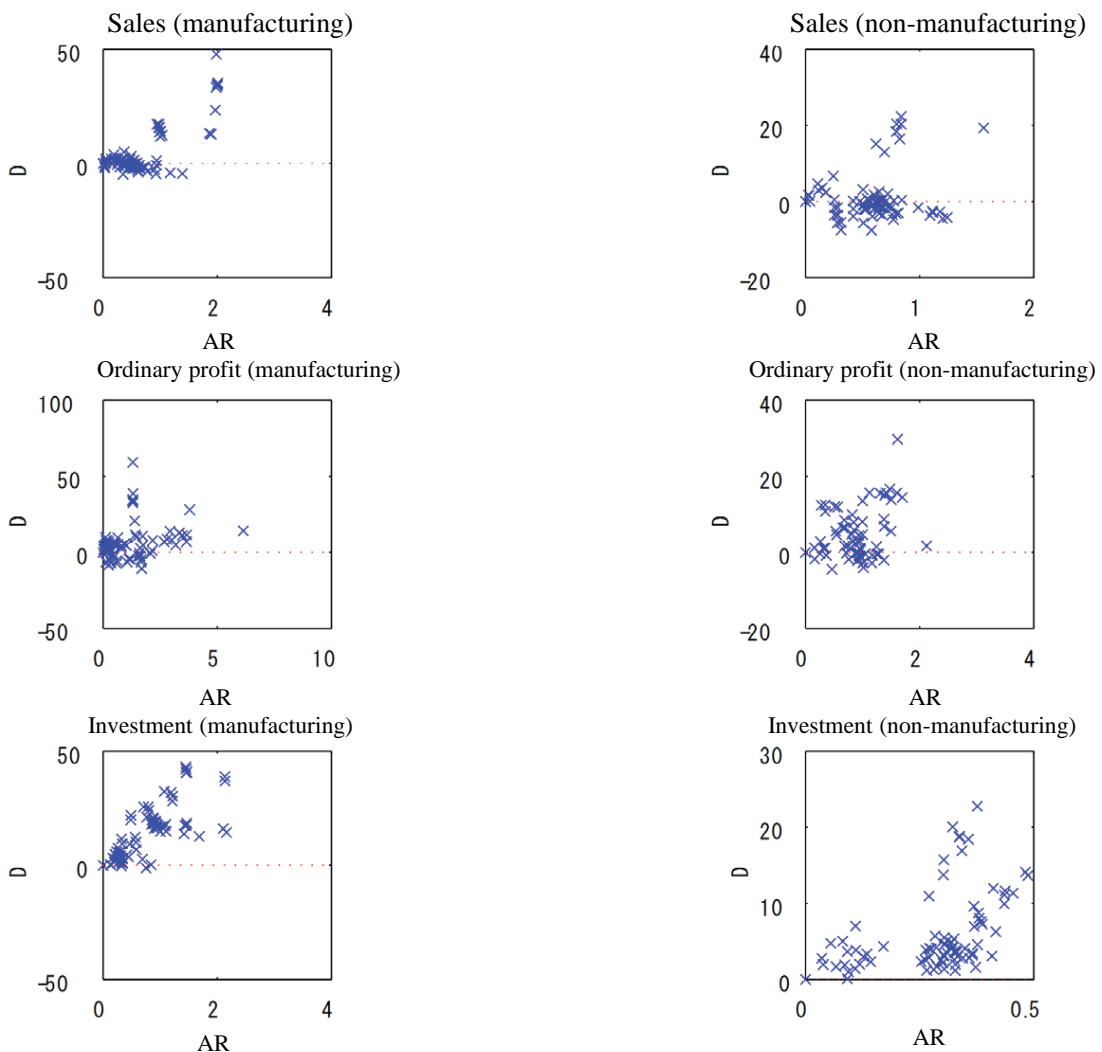
For threshold value  $\alpha > 0$ , the best model is given as

$$\hat{k} = \arg \min_k AIC^{(k)} \quad \text{subject to} \quad SR_m^{(k)} \leq \alpha,$$

where  $AIC^{(k)}$  is  $AIC$  of the model  $k$

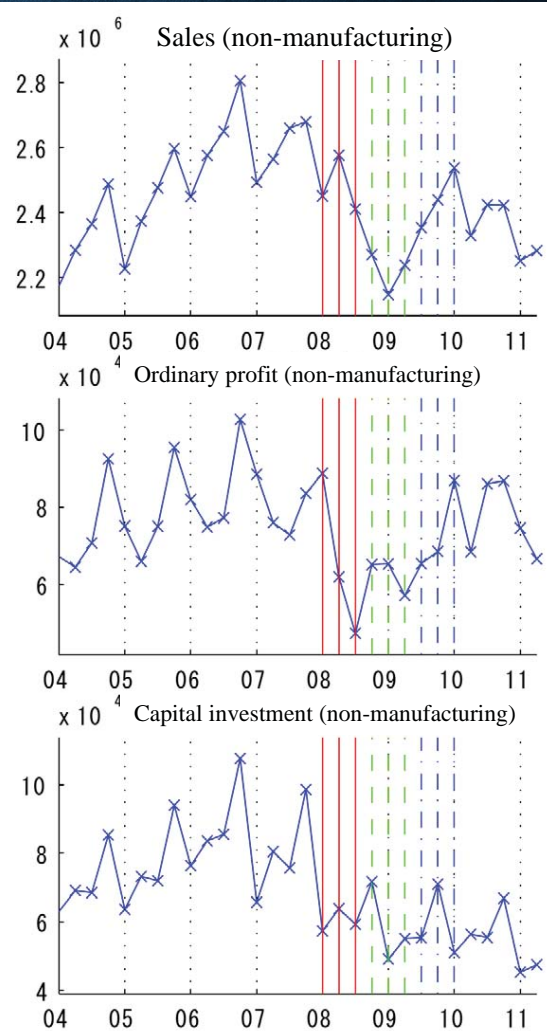
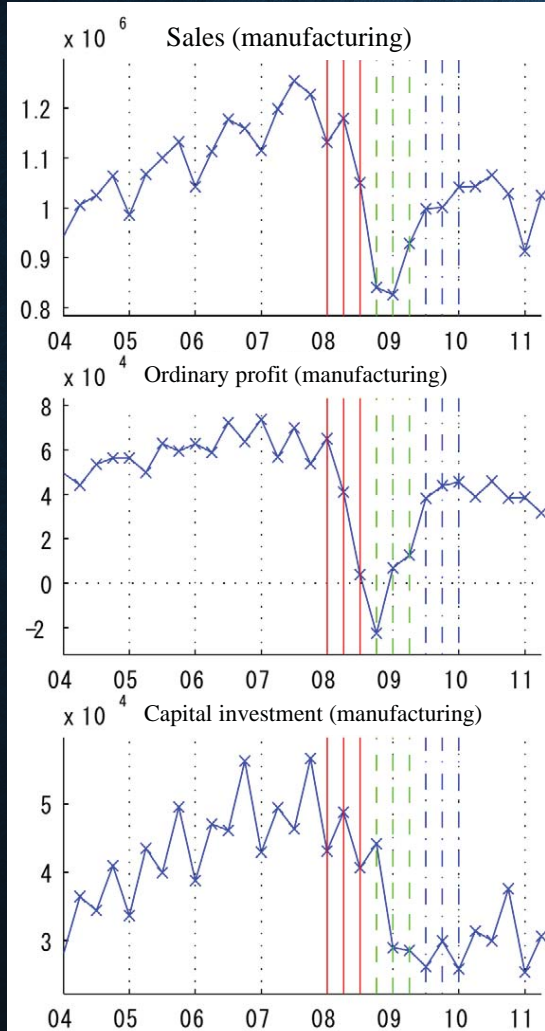
- (1) Calculate  $AIC$  and  $SR$  of all candidate models.
- (2) Exclude models with  $SR$  that is larger than  $\alpha$ .
- (3) Chose the model that minimizes  $AIC$  form the rest of the models.

$$D = AIC^{(k)} - AIC^C$$



# OUTLIER SETTINGS

- **Impact of the collapse of Lehman Brothers**
  - Between 2008 to 2009, a large decline occurred.
  - After a major drop, several series recovered rapidly.
- Most other agencies use the automatic outlier selection procedure.



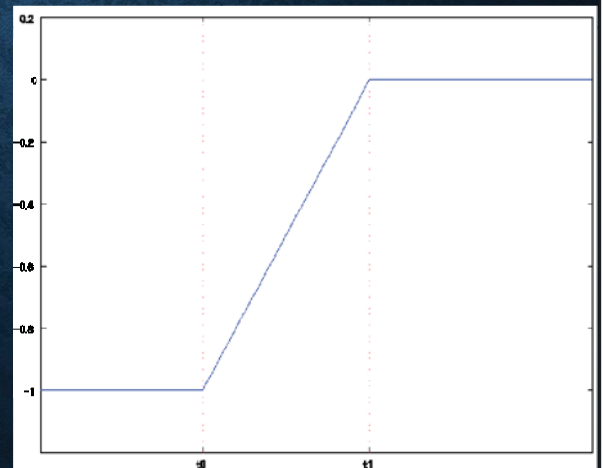


# X-12-ARIMA OUTLIER

## Ramps

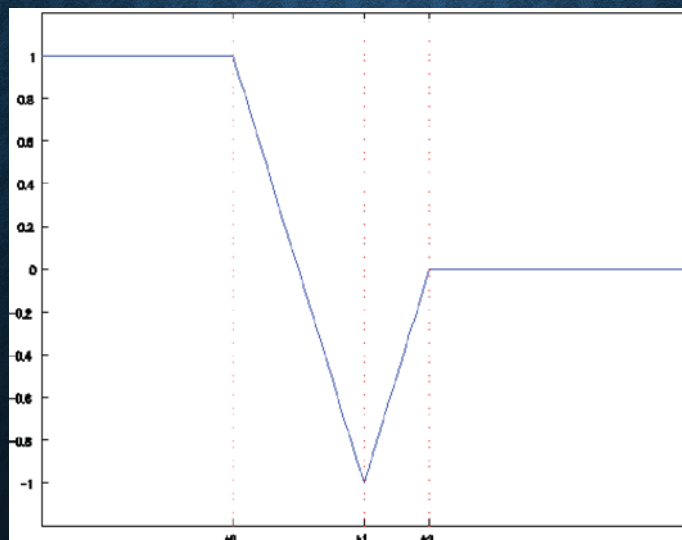
- Series start at one level, slowly shift to others, and continue on at their new levels.
- Ramps have a start date ( $t_0$ ) and an end date ( $t_1$ ).

$$rp_{[t_0, t_1]}(t) = \begin{cases} -1 & t \leq t_0 \\ \frac{t - t_0}{t_1 - t_0} - 1 & t_0 < t < t_1 \\ 0 & t \geq t_1 \end{cases}$$



## RAMP COMBINATION

$$rp_{[t_0, t_1, t_2]}(t) = \alpha_1 rp_{[t_0, t_1]}(t) + \alpha_2 rp_{[t_1, t_2]}(t)$$



# CHANGE POINT DETECTION

Pick out change point candidates from the graph

$t_0$  : 2008.4-6   2008.7 – 9   2008.10 – 12

$t_1$  : 2009.1-3   2009.4 – 6   2009.7 – 9

$t_2$  : 2009.10-12   2010.1 – 3   2010.4 – 6

Three types of dummy variables

(1) No Ramps   1

(2)  $rp[t_0, t_1]$     $3 * 3 = 9$

(3)  $rp[t_0, t_1, t_2]$     $3*3*3 = 27$

All combinations :  $(1+9+27)*81=2997$

## RESULTS OF PENALIZED MODEL SELECTION

	Model	10/7-9	10/10-12	11/1-3	11/4-6	11/7-9
Sales (Manufacturing)	Current	-6.62	0.54	1.71	-8.25	4.85
	$\alpha = 1.0$	-7.97	1.37	2.11	-7.99	3.10
	$\alpha = 2.0$	-6.75	-1.24	-0.86	-3.68	4.03
	$\alpha = 3.0$	-6.75	-1.24	-0.86	-3.68	4.03
	$\alpha = \infty$	-6.75	-1.24	-0.86	-3.68	4.03
Sales (Non-manufacturing)	Current	-9.82	1.84	-0.50	-3.42	0.72
	$\alpha = 1.0$	-10.20	1.82	-2.00	-1.50	-0.20
	$\alpha = 2.0$	-10.48	1.96	-2.67	-0.41	-0.83
	$\alpha = 3.0$	-10.48	1.96	-2.67	-0.41	-0.83
	$\alpha = \infty$	-10.48	1.96	-2.67	-0.41	-0.83

## RESULTS OF PENALIZED MODEL SELECTION (2)

	Model	10/7-9	10/10-12	11/1-3	11/4-6	11/7-9
Ordinary profit (Manufacturing)	Current	2.85	0.67	-4.71	-11.93	-0.78
	$\alpha = 1.0$	2.41	2.48	-5.62	-12.76	0.03
	$\alpha = 2.0$	4.52	-3.77	-4.88	-10.30	0.12
	$\alpha = 3.0$	4.88	-4.58	-6.61	-7.91	0.59
	$\alpha = \infty$	8.08	0.62	-16.56	-5.28	3.87
Ordinary profit (Non-manufacturing)	Current	-5.27	12.21	-7.77	-11.01	5.75
	$\alpha = 1.0$	-5.04	10.44	-7.32	-10.50	6.34
	$\alpha = 2.0$	-6.06	11.27	-9.15	-9.57	7.31
	$\alpha = 3.0$	-6.06	11.27	-9.15	-9.57	7.31
	$\alpha = \infty$	-6.06	11.27	-9.15	-9.57	7.31

## RESULTS OF PENALIZED MODEL SELECTION (3)

	Model	10/7-9	10/10-12	11/1-3	11/4-6	11/7-9
Capital investment (Manufacturing)	Current	4.50	4.06	0.67	-9.77	2.99
	$\alpha = 1.0$	4.32	3.17	0.62	-8.54	2.33
	$\alpha = 2.0$	2.45	3.26	1.02	-7.25	0.57
	$\alpha = 3.0$	2.45	3.26	1.02	-7.25	0.57
	$\alpha = \infty$	2.45	3.26	1.02	-7.25	0.57
Capital investment (Non-manufacturing)	Current	-1.22	-2.90	-3.60	-4.32	-5.84
	$\alpha = 1.0$	-0.90	-3.43	-3.38	-4.34	-5.24
	$\alpha = 2.0$	-0.90	-3.43	-3.38	-4.34	-5.24
	$\alpha = 3.0$	-0.90	-3.43	-3.38	-4.34	-5.24
	$\alpha = \infty$	-0.90	-3.43	-3.38	-4.34	-5.24

# MODEL USED FOR THE PUBLICATION OF THE 2011 FOURTH QUARTER

	$\alpha$	ARIMA	$t_0$	$t_1$	$t_2$
Sales (Manufacturing)	1.00	(2 1 2)(2 1 0)	08/4-6	09/4-6	10/1-3
Sales (Non-manufacturing)	1.00	(0 1 2)(1 1 0)	08/7-9	09/1-3	10/4-6
Ordinary profits (Manufacturing)	2.00	(0 1 0)(0 1 1)	08/7-9	09/1-3	10/1-3
Ordinary profits (Non-manufacturing)	1.00	(2 1 2)(1 1 1)	08/4-6	09/1-3	10/4-6
Capital investment (Manufacturing)	1.00	(1 1 2)(1 1 2)	08/7-9	09/1-3	09/10-12
Capital investment (Non-manufacturing)	1.00	(1 1 0)(1 1 2)	08/4-6	09/4-6	10/4-6

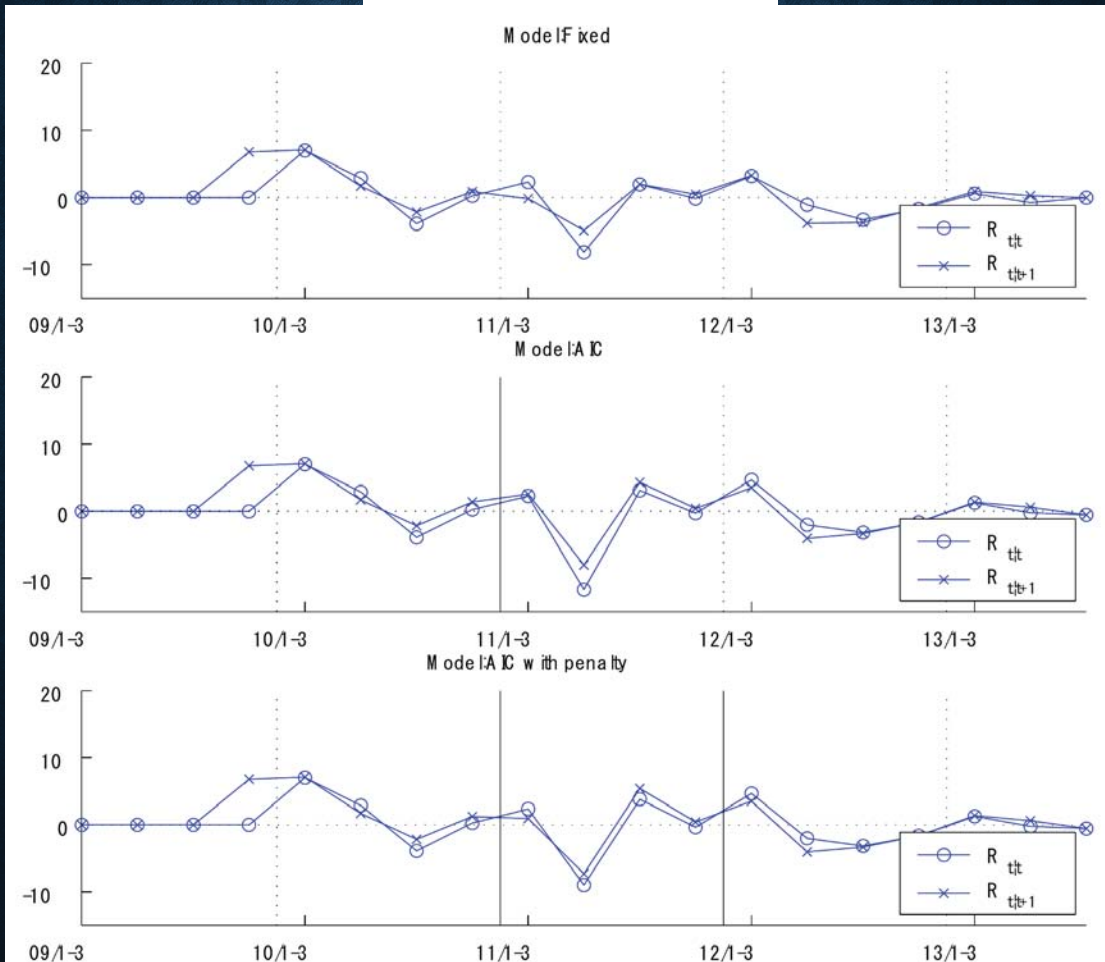
## SIMULATION SETTINGS

Compare selection procedure for three models:

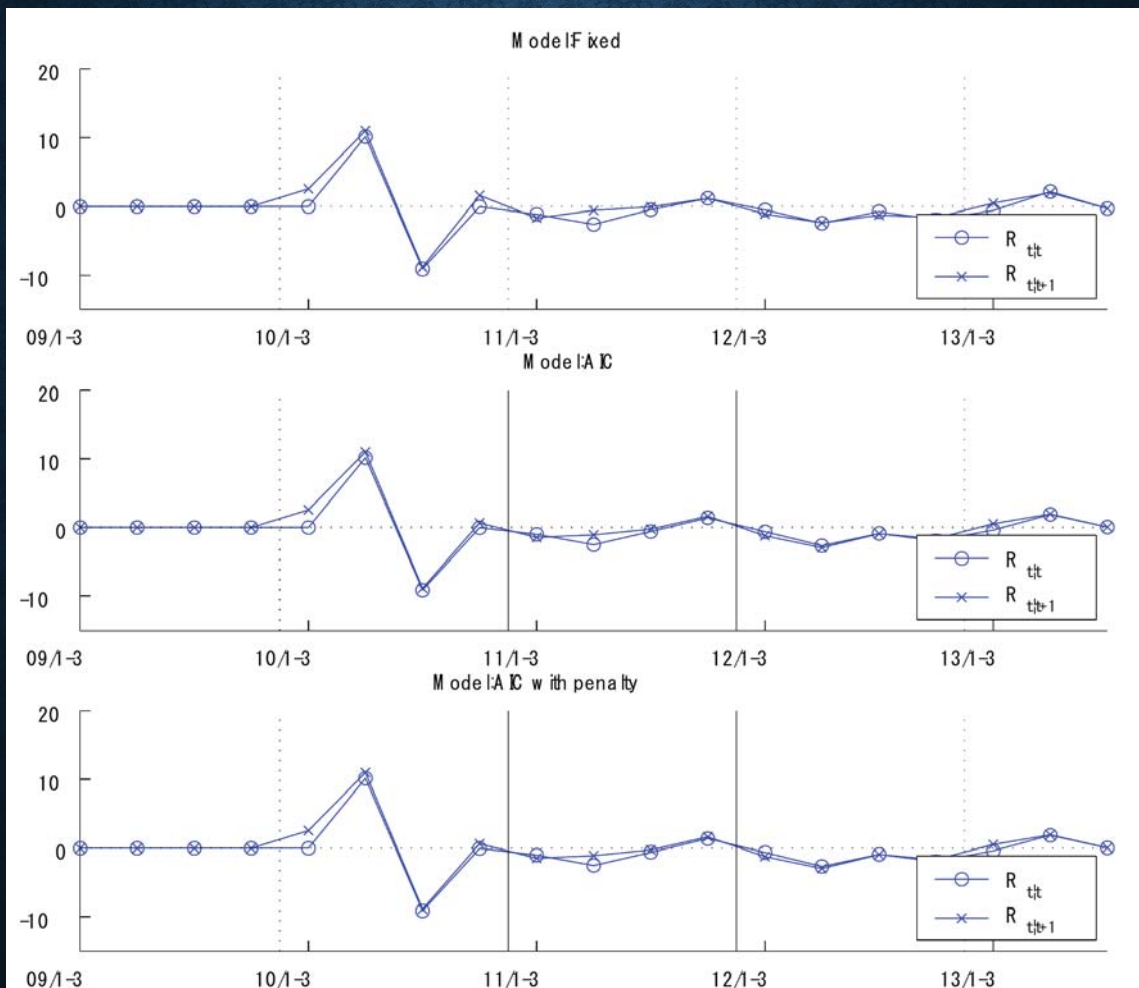
- (1) Fixed model
- (2) Select by AIC
- (3) Penalized selection

- Assume that the data are going to be added sequentially, by period, starting from the 1st quarter of 2009.
- Model selection is performed every four periods.
- In the estimates for each series, a ramp-dummy corresponding to the collapse of Lehman Brothers is used.

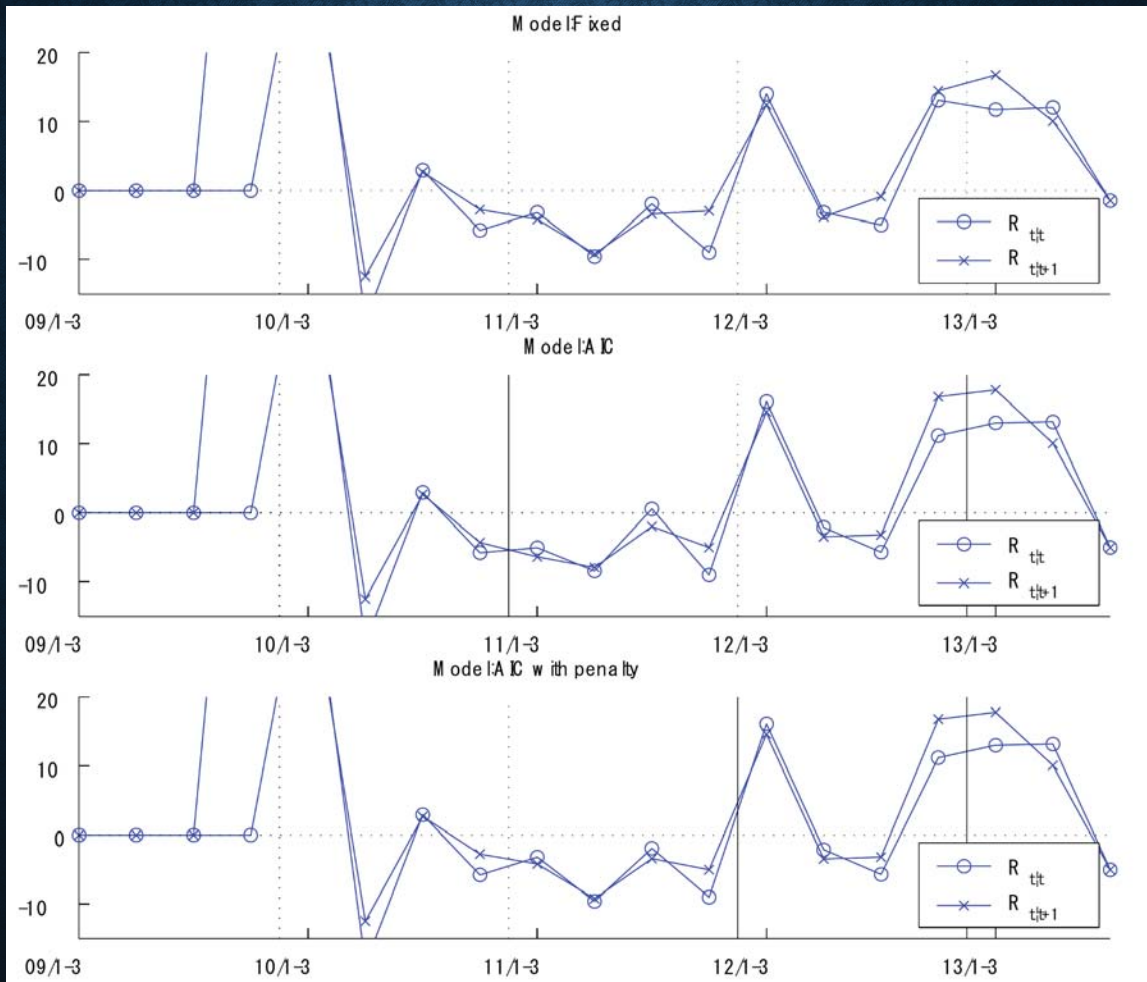
### Sales (manufacturing)



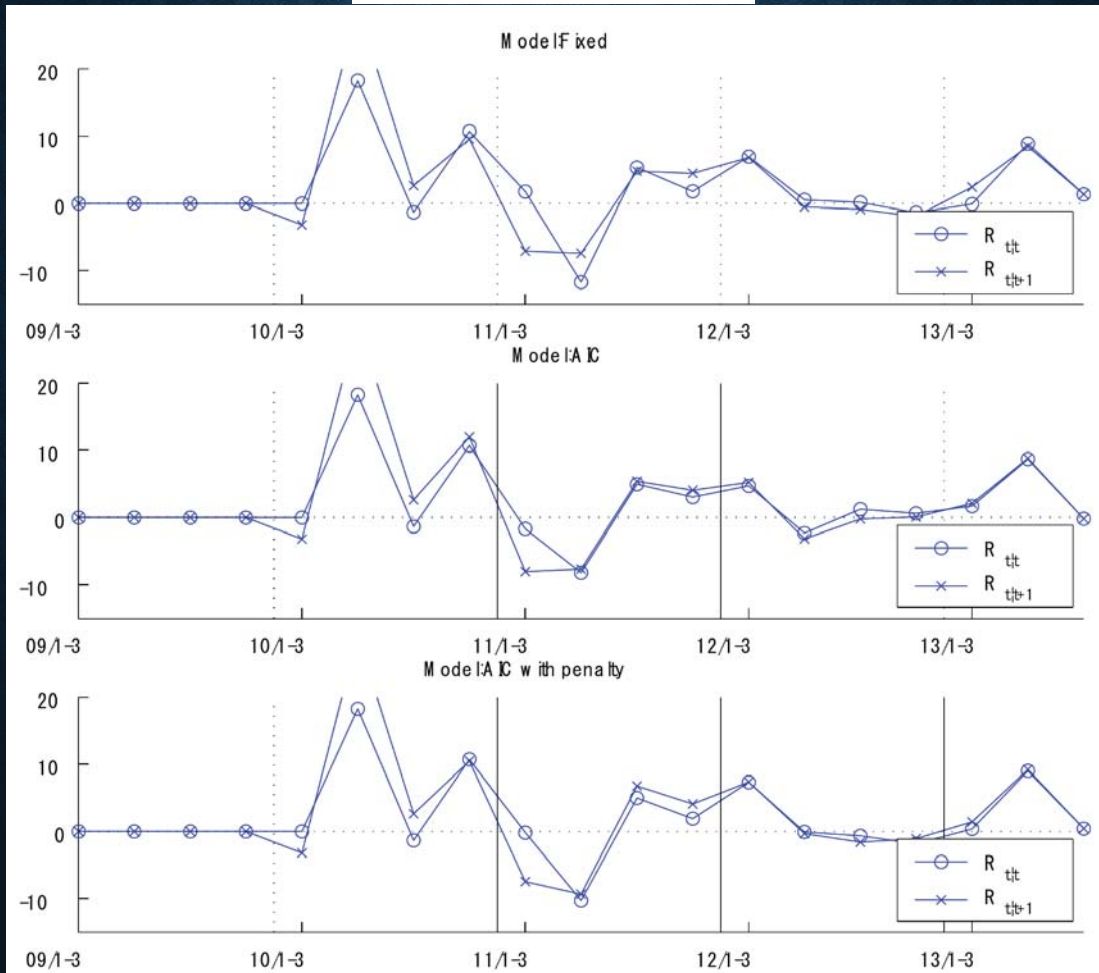
### Sales (non-manufacturing)



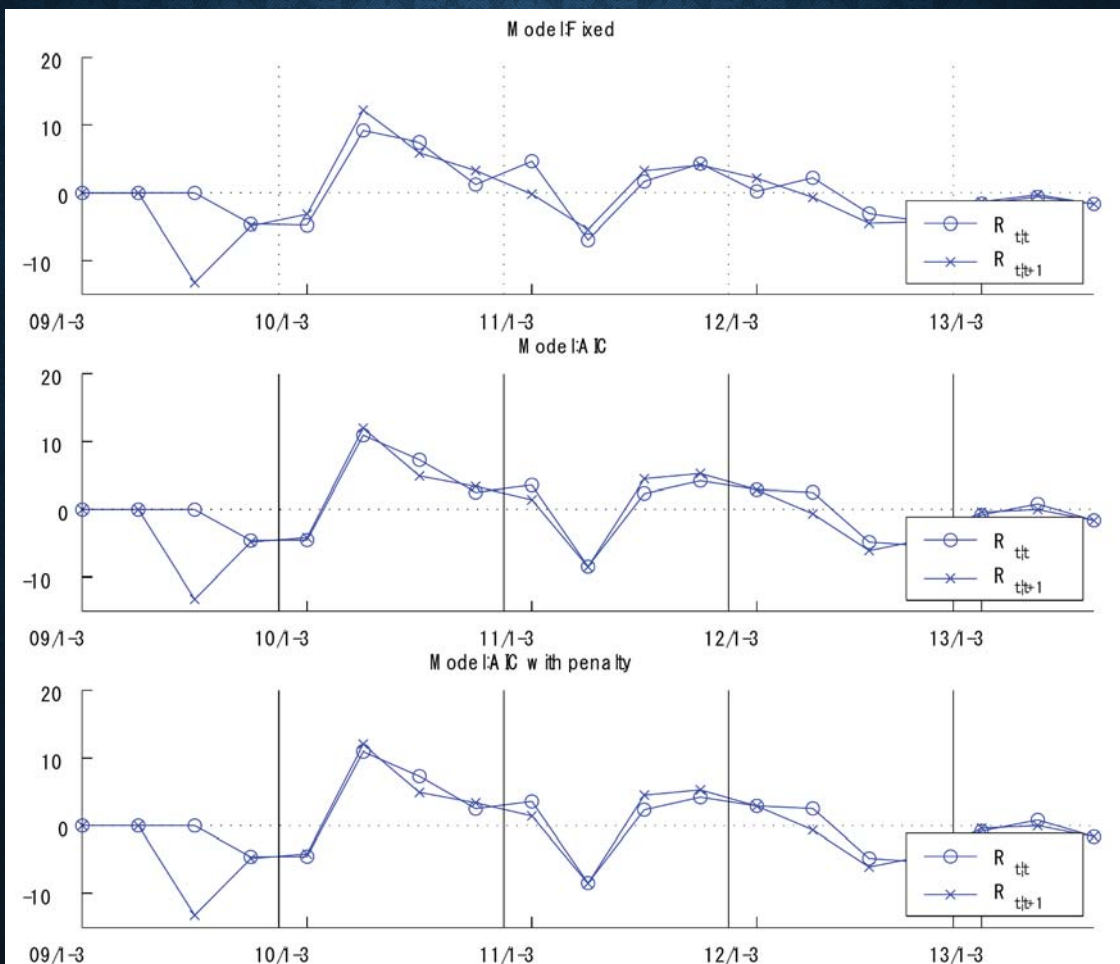
Ordinary profit (manufacturing)



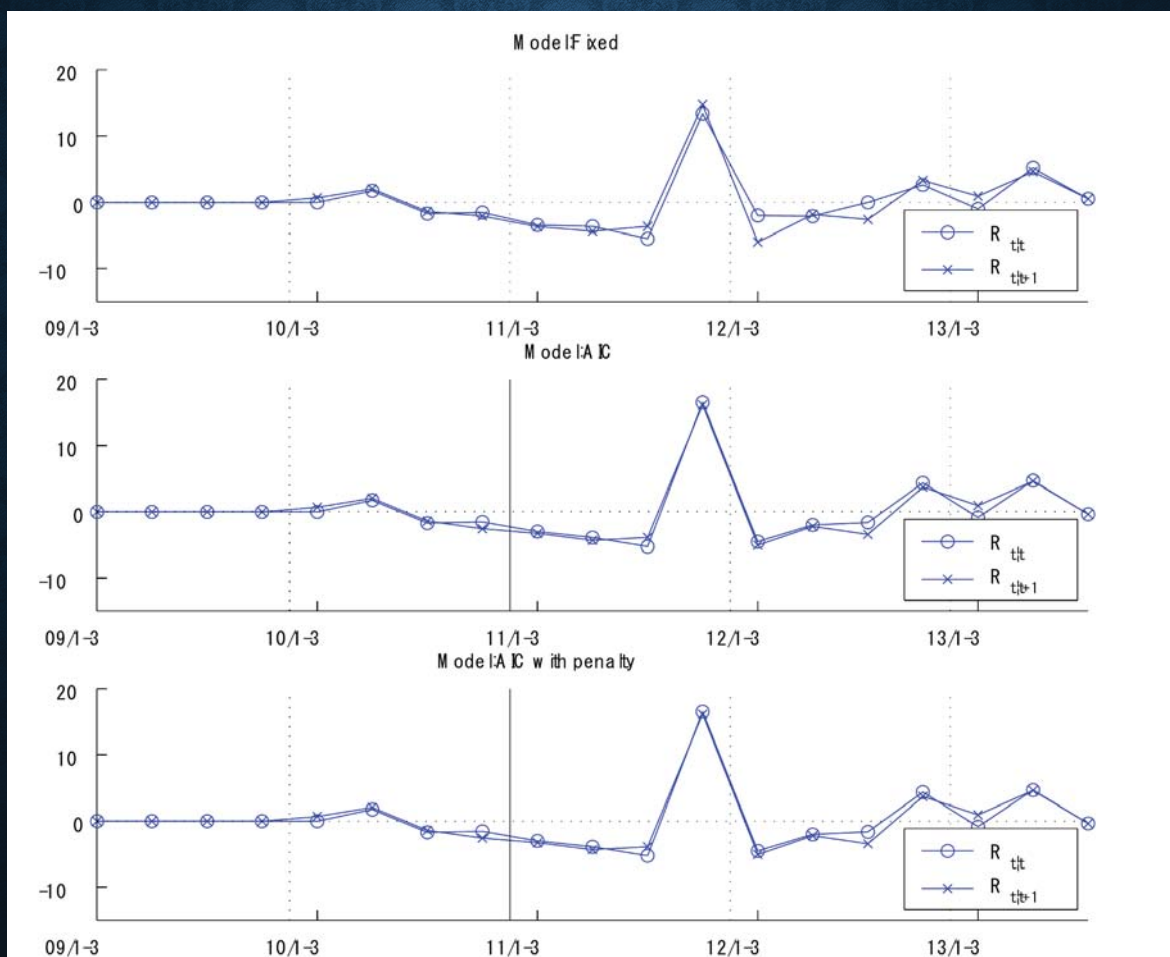
Ordinary profit (non-manufacturing)



### Investment (manufacturing)



### Investment (non-manufacturing)



# AVERAGE REVISION

The average of  $|R_{t|t+1} - R_{t|t}|$

	Sales (Manufacturing)	Sales (Non- manufacturing)	Ordinary profits (Manufacturing)	Ordinary profits (Non- manufacturing)	Capital investment (Manufacturing)	Capital investment (Non- manufacturing)
Fixed	1.135	0.578	6.673	2.165	1.944	0.839
AIC	1.108	0.474	6.787	1.656	1.596	0.518
AIC with penalty	1.077	0.474	6.786	1.772	1.596	0.518

## CONCLUDING REMARKS

- **Penalized Model Selection Procedure**
  - The penalized model selection procedure adopted in the Financial Statements Statistics of Corporations was explained.
- **Simulation results**
  - In general, the re-selection of models will reduce the revisions required for the seasonally adjusted series.
  - The impact of the penalized selection method was not clearly observed.
- **Balance Between Accuracy and Stability**
  - It is possible that a trade-off occurs between the model optimality and the seasonal adjustment stability.
  - More detailed examination of this possibility may be required.



# REFERENCES

- [1] Akaike, H. (1973), "Information Theory and an Extension of the Likelihood Principle," in the *Second International Symposium on Information Theory*, eds. B. N. Petrov and F. Czaki, Budapest: Akademia Kiado, 267-287.
- [2] Findley, D. F., B. C. Monsell, W. R. Bell, M. C. Otto and B. C. Chen (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program," *Journal of Business and Economic Statistics*, 16, 127-176 (with Discussion).

## Seasonal Adjustment of CPS Labor Force Series During the Great Recession<sup>1</sup>

Thomas D. Evans and Richard B. Tiller, Bureau of Labor Statistics  
Bureau of Labor Statistics, Statistical Methods Staff,  
2 Massachusetts Ave. NE, Room 4985, Washington, DC 20212

Key Words: Time Series; Level Shifts; Ramps

### Introduction

The Current Population Survey (CPS) is a monthly household survey that collects information on labor force characteristics for the United States. The seasonally adjusted monthly levels of employment, unemployment and unemployment rate are key indicators of the health of the economy. Recent economic turbulence related to the Great Recession has raised public interest in how the CPS series are seasonally adjusted. The most widely used approaches to seasonal adjustment apply weighted moving averages to the original data to separate the seasonal from the nonseasonal components of the data. These weights may be derived from a model fit to each series or from a non-parametric method using pre-determined filters. These methods have been successful because of their flexibility in accommodating stochastic changes in seasonal patterns. The presence of stochastic seasonality, however, creates conceptual and statistical difficulties in separating seasonal from the nonseasonal components. The difficulties are compounded during periods of rapid economic change.

The national CPS data are adjusted using X-12-ARIMA (Findley, et al., 1998) which is a nonparametric approach. This method has a long history of development and refinement and is currently used by statistical agencies throughout the world. This paper reviews how well this method was designed to handle changes in trend-cycle and seasonality during both economic expansions and contractions, what tools are available to adjust for major economic shocks, and how the method has performed in the last recession.

The original X-11 filtering algorithm that performs seasonal adjustment remains at the core of the X-12 process (Shiskin, Young, and Musgrave, 1967). The X-11 part of X-12 has a broad range of seasonal and trend-cycle filters. The symmetric Henderson trend filter reproduces third-degree polynomials within the span of the filter which can be as short as one to two years. Thus, the trend filter is flexible enough to follow rapid changes in growth rates as well as quickly occurring turning points in the trend-cycle. Towards the end of the series less adaptive asymmetric filters must be used. Although the one-sided Henderson filter is less flexible, it can still track a local linear trend and with ARIMA forecasting so more adaptive filters are produced for the end points (Dagum 1983).

A well-known problem with seasonal adjustment is that moving averages are highly vulnerable to sudden strong atypical changes or outliers. For example, sudden shocks to the trend-cycle can be absorbed into the seasonal factor estimates and erroneously removed from the seasonally adjusted series. A second potential source of distortion is a break in the seasonal pattern which some analysts argue is likely to occur during recessions. Either type of shock may lead to distorted seasonally adjusted estimates in later years following the recession.

The best way to detect and correct for sudden disturbances is with prior information on their source, time of occurrence, magnitude, and duration. The original X-11 developers added an option for users to

---

<sup>1</sup> Disclaimer: Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

supply adjustment factors to the original data to remove the effects of shocks prior to seasonal adjustment. Users, however, had to develop these prior adjustments outside of the software. This is hardly feasible to do when there are hundreds or thousands of series to seasonally adjust. With the addition of the RegARIMA modeling capabilities in X-12 (Findley, et al., 1998), it became standard practice to utilize the automatic outlier detection routine prior to the seasonal adjustment step. RegARIMA also allows users to specify their own outlier models.

Currently, there are 144 CPS national series that are directly seasonally adjusted. One hundred and thirty-nine of these series are released monthly and the others quarterly. Many more are indirectly seasonally adjusted based on the directly-adjusted series. See Tiller and Evans (2013) for more details.

The focus of this paper is on eight CPS series that make up the national unemployment rate:

- Employment (EM) ages 16-19 by gender
- Employment ages 20+ by gender
- Unemployment (UN) ages 16-19 by gender
- Unemployment ages 20+ by gender

The eight series are directly seasonally adjusted and then used to derive the seasonally adjusted unemployment rate (UR). The official national seasonally adjusted unemployment rate is plotted over time in Figure 1. The trend for the unemployment rate (not officially published) appears in Figure 2. Both series rise sharply in the 2008-2009 period. Commentary from Wall Street analysts and other members of the press suggested that the sharp rise (fall) in UN (EM) from the fourth quarter of 2008 to the first quarter of 2009 was absorbed into the seasonal component and this caused seasonal adjustment to overestimate recent economic growth in those quarters (a decline in UN or a rise in EM) and underestimate in the second and third quarters.

### **Testing for Trend-Cycle and Seasonal Breaks with RegARIMA Models**

To investigate the claims of distortions to the seasonal adjustment process, we use RegARIMA models to test for the presence of recession effects due to trend-cycle and seasonal breaks. For seasonal breaks we also use a graphical analysis of X-11 seasonal-irregular (SI) sub-plots by month which will be described in more detail below.

The RegARIMA model may be represented as,

$$y_t = X_t' \beta + Z_t$$

where  $X_t$  is a vector of fixed regressors and  $Z_t$  follows a seasonal ARIMA model. The parameters of the ARIMA model implicitly determine the stochastic properties of seasonality as well as trend-cycle and other components of the series without having to specify specific models for these components. The variables of the regression part of the model provide the basis for testing for the presence of exogenous effects, such as trend and seasonal breaks that we wish to estimate and remove from the series prior to seasonal adjustment to prevent distortions to the X-11 seasonal adjustment process.

Trend breaks can be handled by modeling level shifts (LS), temporary changes (TC), and ramps which are briefly described below and the form of the regression variables for testing for each effect are shown in Table 1. These variables are part of a large set of built-in regressors available in RegARIMA.

LS: A permanent abrupt shift in level.

TC: An abrupt change in level followed by a gradual return to normal.

Ramp: A change to a new level with a user specified start ( $t_0$ ) and end date ( $t_1$ ) and a fixed rate of change per period ( $1/(t_1 - t_0)$ ).

To test for the presence of seasonal breaks we use the partial change in regime test which is another predefined option available in RegARIMA. An exogenous shift in the seasonal factors occurring prior to the change point ( $t_0$ ) is modeled in terms of the  $s-1$  seasonal dummy variables ( $M_{j,t}$ ) shown in Table 1. This effect is referred to as a partial change of regime since it is estimated for only the early span and set to zero for the complementary span (U.S. Census Bureau, 2013).

LSs and TCs are usually detected during the automatic outlier detection option of RegARIMA. This option assumes no prior knowledge of the timing or type of outliers. The procedure identifies additive outliers (AOs), level, and TCs. The setting of the critical value is always an issue since it involves multiple testing (type 1 errors are inflated) and because the outliers may have no economic explanation. To reduce the number of spurious outliers, critical values for the outlier T-values (nonstandard distributions) are adjusted for the number of observations (see Findley, et al., 1998).

**Table 1: RegARIMA Outlier Variables**

Regression Effect	Variable definition
Additive	$X_t^{AO} = \begin{cases} 1 & t = t_0 \\ 0 & t \neq t_0 \end{cases}$
Level shift	$X_t^{LS} = \begin{cases} -1 & t < t_0 \\ 0 & t \geq t_0 \end{cases}$
Temporary change	$X_t^{TC} = \begin{cases} 0 & t < t_0 \\ \alpha^{t-t_0} & t \geq t_0 \end{cases}$
Temporary Ramp	$X_t^{Rp} = \begin{cases} -1 & t \leq t_0 \\ (t - t_0)/(t_1 - t_0) - 1 & t_0 < t < t_1 \\ 0 & t \geq t_1 \end{cases}$
Partial change of regime in seasonality	$X_{j,t}^E = \begin{cases} M_{j,t} & t < t_0 \\ 0 & t \geq t_0 \end{cases}; M_{j,t} = \begin{cases} 1 & t = j + ks \\ -1 & t = s \\ 0 & t \neq j + ks \end{cases}$ for $j = 1, \dots, s-1; k = 0, 1, 2, 3, \dots; s = 12$

Since the automatic outlier identification procedure is run at the end of each year, we make no attempt to identify LSs in real time. During the recession period, we identified two LSs: May 2008 for (UN) teen males and UN teen females. Note that the original t-value for the teen females LS was originally only 3.8, even though the default critical value  $\cong 4.0$  (it varies by series length). Occasionally, an outlier with a slightly lower T-value may be added to a series if it is believed that doing so will be helpful or if the effect is known in advance. Should we reduce the critical value during a recession period? An examination of the “almost” critical outliers identified by the program confirmed our suspicion that lowering the default critical value would have introduced too many spurious outliers.

Table 2 shows the RegARIMA estimates for the two level shifts. Both LSs have large relative level increases. Notice that the LS for UN teen females is no longer as significant as additional data lowered the t-values further below the critical value. An obvious question is whether the LSs for UN teens affected the seasonally adjusted values? The surprising answer is no, but they did have strong effects on the trend and irregular components. If we just examined the seasonal factors, we might conclude that no recession-related LS effects occurred. This point is clearly seen in Figures 3 and 4. However, it is also important to examine the trend and irregular components for interpreting what happened. The trend in Figure 5 shows a different story in which the break is clearly noticeable. Without accounting for an LS, the trend over smoothes the recession effects, and the irregular component (Figure 6) compensates for the LS. The plot for UN teen females is in Figure 7.

**Table 2: RegARIMA Estimates for Level Shifts**

Level Shifts from Automatic Outlier Detection						
Series	Date	Coef (exp) <sup>2</sup>	T-Value	AICC <sup>3</sup> w/LS	AICC w/o LS	AICC Difference
UN M 16-19	May 2008	1.26	4.4	10,343	10,360	-16.5
UN F 16-19	May 2008	1.17	3.1	10,236	10,243	-6.8

If the LSs were ignored for the UN teen series, the large upward shifts in May would be interpreted as a random deviation from normal as opposed to a fundamental level shift. Data users would see no noticeable effect on the seasonally adjusted series since the seasonal patterns are not affected by the LS.

After testing for LSs, we moved to experimenting with ramps. Ramps have not been used for CPS series and are not commonly used in other applications of seasonal adjustment. Examples of the use of ramps are given by Buszuwski and Scott (1993), Maravall and Perez (2011), and (Lytras and Bell 2013).

The specification of a ramp is not automated and requires the user to specify beginning and end points for the adjustment period. This leads to the problem of how to select  $t_0$  and  $t_1$ . Since there is not much guidance without prior information, we can visually look for segments of series where the change appears to be relatively constant. Another problem is that in this process it can be fairly easy to achieve “better” fits but they may be spurious. We try to determine this by utilizing the AICC goodness-of-fit criterion.

Test results appear in Tables 3 and 4 below. In Table 3, the exponentiated coefficients show small effects even though the T-values appear significant. Revisions between the concurrent and most recent seasonally adjusted values with and without ramps (Table 3) show little differences, yet the AICC tests seem to indicate a preference for including the ramps as regressors in the model. As the testing is not conclusive, we examine plots. The seasonally adjusted results for EM 20+ with and without ramps are plotted in Figure 8. As the coefficient value close to one implies, the addition of the ramp makes almost no difference. Using the three series with tested ramps to derive the national UR in Figure 9 also shows only very small differences.

**Table3: Ramp Results**

Series	Ramp Start	Ramp End	Ramp Coef (exp)	T-Value
EM M 20+	Nov 2008	Mar 2009	0.99	-5.66
UN M 20+	Apr 2008	Feb 2009	1.06	4.92
UN F 20+	Apr 2008	Feb 2009	1.04	4.79

<sup>2</sup> Coefficients in the tables are exponentiated from logs to levels.

<sup>3</sup> See U.S. Census Bureau (2013) for a detailed explanation of the AICC and model-selection criteria in RegARIMA. Models with minimum AICC are usually preferred.

**Table 4: Seasonal Adjustment Revisions due to Ramps and AICC Comparisons**

Series	SA Revision Medians, Average Absolute %		AICC	AICC w/Ramp	AICC Difference
	Official	With Ramp			
EM M 20+	0.07	0.08	11,511.45	11,484.39	-27.06
UN M 20+	0.78	0.81	6,822.91	6,805.60	-17.32
UN F 20+	0.83	0.81	6,732.70	6,713.66	-19.04

The partial change of regime test for a break in the seasonal pattern is shown in Table 5. While we tested other months, October 2008 makes sense as a potential breakpoint, yet the AICC criterion rejects the presence of a deterministic break for all 8 series.

**Table 5: AICC Results for Seasonal Breaks for October 2008**

Series	AICC		
	No Break	With Break	Difference
EM M 16-19	10,698.7	10,711.3	12.7
EM F 16-19	10,672.0	10,685.4	13.3
EM M 20+	11,511.5	11,517.7	6.2
EM F 20+	11,525.6	11,527.9	2.2
UN M 16-19	10,343.3	10,348.6	5.3
UN F 16-19	10,235.7	10,257.7	22.0
UN M 20+	6,822.9	6,829.9	7.0
UN F 20+	6,732.7	6,748.4	15.7

We further explore the possibility of a seasonal break with a visual examination of X-11 SI sub-plots for each series. The SI is the de-trended series produced in the X-11 part of the procedure. The trend-cycle is estimated by the Henderson filter and removed from the original series. The seasonal factors are later derived by smoothing the SIs by month, usually with a 3x5 moving average. These SI sub-plots by months are useful for identifying potential breaks in seasonality. However, our examination found no apparent breaks for any of the eight series. An example of an SI sub-plot is shown in Figure 10.

### Performance of Post-Recession Seasonal Adjustment

The NBER dates the Great Recession as lasting from December 2007 to June 2009. We use the period January 2008-December 2009 as a more relevant recession period for the labor market since the unemployment rate doubled and then began to decline around January 2010. Largely because NBER dating procedures depend heavily on real product and income measures, it is not unusual for labor market recovery to lag the end of NBER designated recessions.

In order to examine the effect of the recession data (January 2008-December 2009) on the January 2010-March 2012 period for the national UR, we treated the recession data as missing. When observations are set to zero, RegARIMA automatically adds AO regressors to produce forecasts based on pre-recession data. The original series with forecasted values for the recession period was seasonally adjusted with a January 2010 level shift added to the model. Since unemployment is still much higher in 2012 compared to pre-recession levels, we use the term “post-recession” advisedly.

The officially seasonally adjusted national UR series is plotted against the seasonally adjusted series with recession data replaced with forecasts in Figure 12. Note how the seasonally adjusted series with the recession data forecasted ignores the actual rise due to the recession. Starting in 2010, the two series

become very close again which indicates that the recession data does not affect the post-recession seasonal adjustment much.

In November 2010, an upward blip occurs in the seasonally adjusted national UR series (see Figure 12). Various sources have attributed this movement to seasonal bias (for example, see Zentner, et al., 2011). Actually, the blip is due to variation in the irregular component that was eliminated by the trend filter. In short, we see no evidence of seasonal bias during the post-recession period in question.

## Summary

Our overall analysis of the performance of the seasonal adjustment for the CPS national unemployment rate is that there is no evidence of trend or seasonal breaks that biased the adjustments during the post-recession period. The major findings are:

- Six of the eight series analyzed need no outlier adjustments during the recession.
- Recessionary level shifts were identified in the normal way for UN male and female teenagers. This reduced the irregular variation but had little effect on the seasonally adjusted series.
- Fitting ramps models had little effect.
- Removing the recession data had little effect.
- The standard X-11 symmetric filters appear to perform well during and after the 2008-2009 recession.

## References

- Buszuwski, J., and Scott, S. (1993), "Some Issues in Seasonal Adjustment when Modeling Interventions," in JSM Proceedings, Business and Economics Section, 208-213.
- Dagum, E. (1983), *The X-11-ARIMA Seasonal Adjustment Method*, Ottawa: Statistics Canada.
- Dagum, E., and Morry, M. (1985), "Seasonal Adjustment of Labour Force Series during Recession and Non-Recession Periods," *Survey Methodology*, XI, no. 2.
- Findley, D., Monsell, B., Bell, W., Otto, M., and Chen, B. (1998), "New capabilities and methods of the X-12-ARIMA seasonal adjustment program," *Journal of Business and Economic Statistics* **16**, 127--177 (with discussion). Available online at <http://www.census.gov/ts/papers/jbes98.pdf>.
- Lytras, D., and Bell, W. (2013) "Modeling Recession Effects and the Consequences on Seasonal Adjustment," presented at the 2013 Joint Statistical Meetings.
- Maravall, A., and Perez, D. (2011), "Applying and Interpreting Model-Based Seasonal Adjustment. The Euro-Area Industrial Production Series," Bank of Spain. Available online at <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/11/Fich/dt1116e.pdf>.
- Shiskin, J, Young, A., and Musgrave, J. (1967), "The X-11 Variant of the Census Method II Seasonal Adjustment Program," Bureau of the Census, Technical Paper 15. Available online at <http://www.census.gov/ts/papers/ShiskinYoungMusgrave1967.pdf>.
- Tiller, R., and Evans, T. (2013), "Methodology for Seasonally Adjusting National Household Labor Force Series with Revisions for 2013," Bureau of Labor Statistics CPS Technical Documentation. Available online at <http://www.bls.gov/cps/cpsrs2013.pdf>.

U.S. Census Bureau (2013), *X-13ARIMA-SEATS Reference Manual* (Version 1.0), Washington, DC: Author. Available online at <http://www.census.gov/ts/x13as/docX13AS.pdf>.

Zentner, E., Amemiya, A., and Greenberg, J. (2011), “Stronger Data Ahead: Explanation and Implication,” Nomura Global Weekly Economic Monitor. Available online at <https://www.bea.gov/about/pdf/Stronger%20Data%20Ahead.pdf>.



Figure 1: Official National Seasonally Adjusted Unemployment Rate

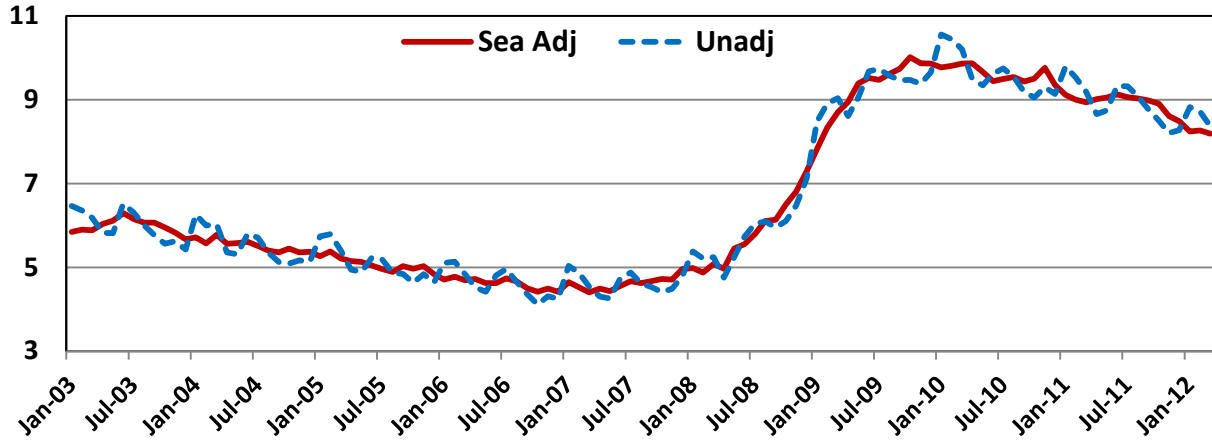


Figure 2: National Unemployment Rate Trend (not published)

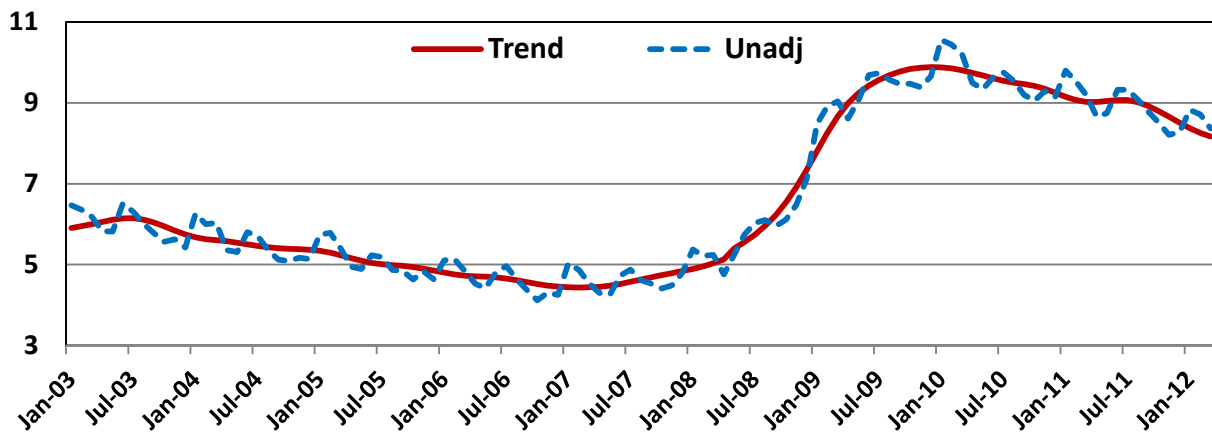
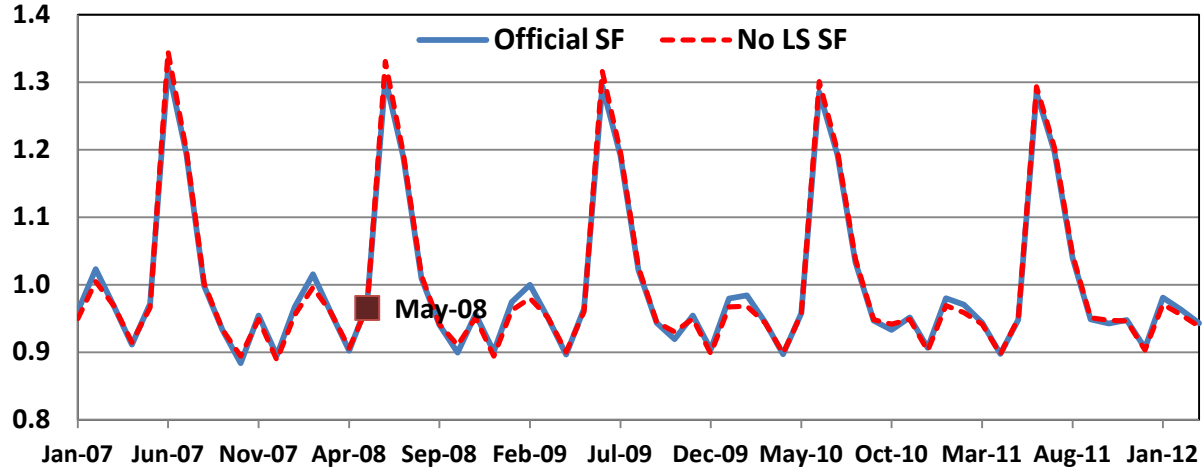
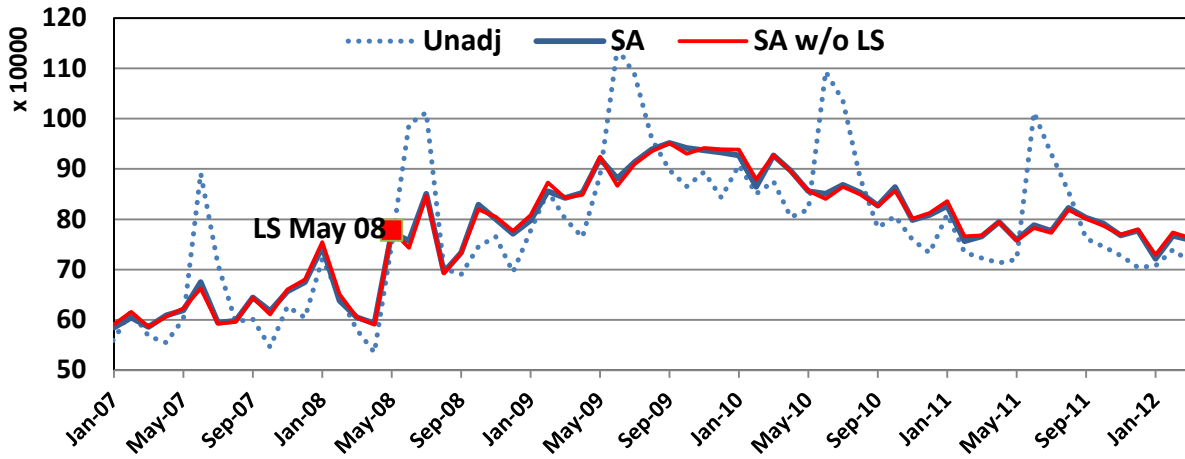


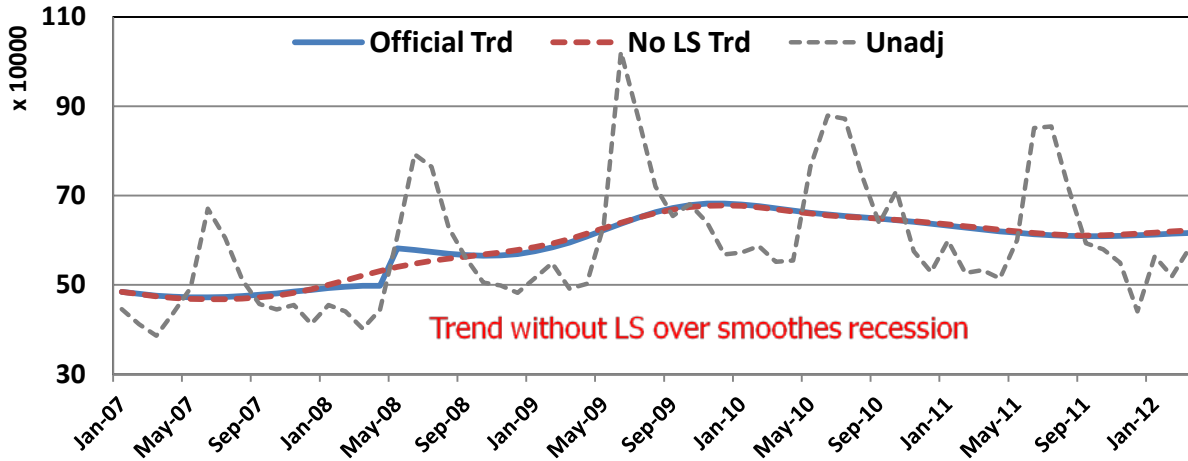
Figure 3: Level Shift Effect on UN Male Teen Multiplicative Seasonal Factors



**Figure 4: Level Shift Effect on UN Male Teen Seasonally Adjusted Series**



**Figure 5: Level Shift Effect on UN Male Teen Trend Component**



**Figure 6: Level Shift Effect on UN Male Teen Multiplicative Irregular Component**

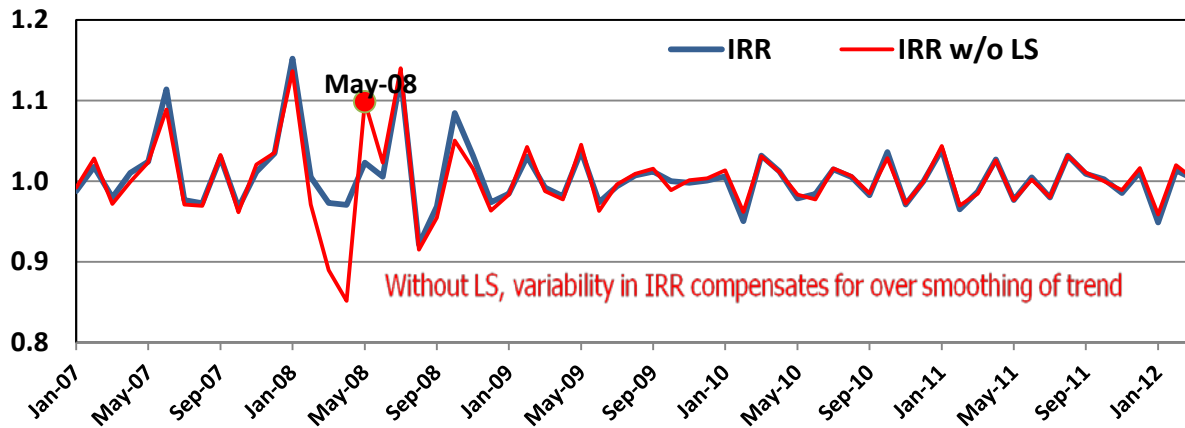


Figure 7: Level Shift Effect on UN Female Teen Seasonally Adjusted Series

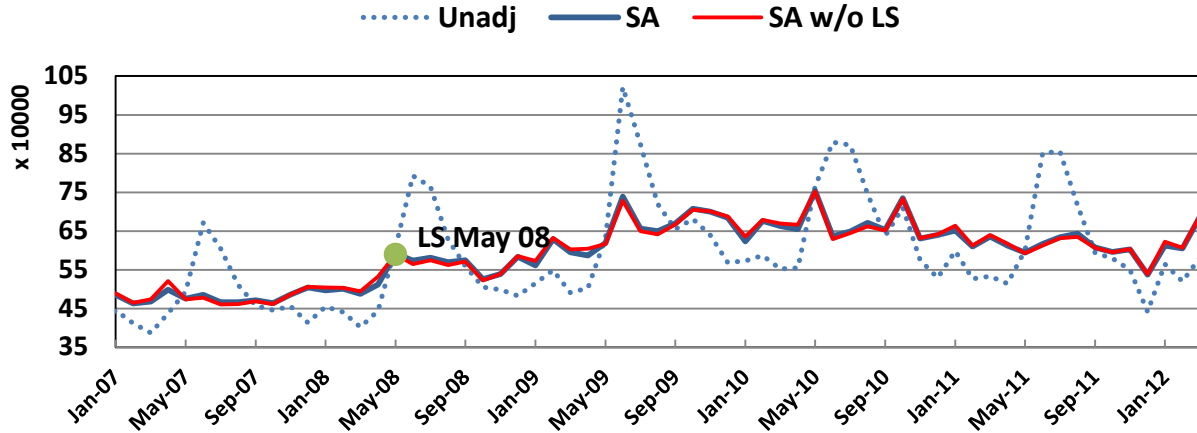


Figure 8: Seasonally Adjusted EM M 20+ with and without Ramps

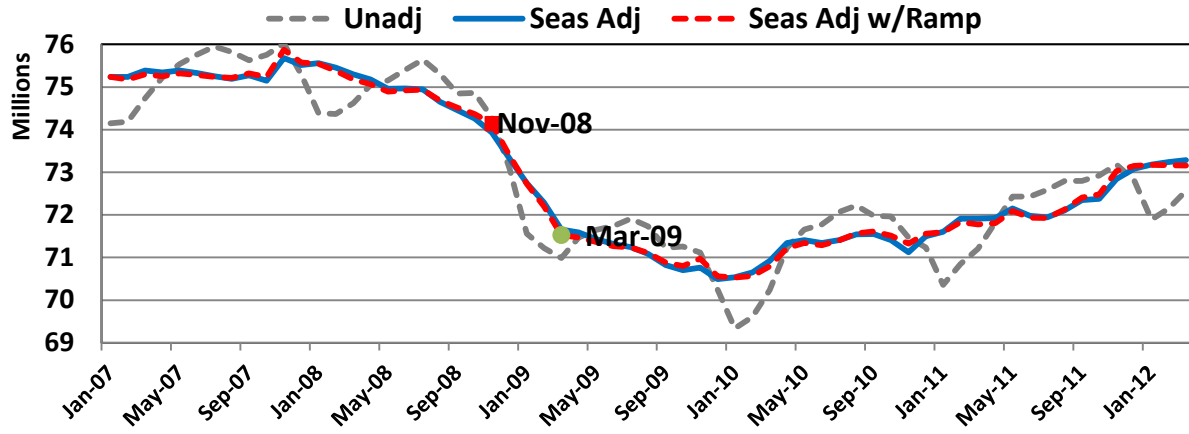
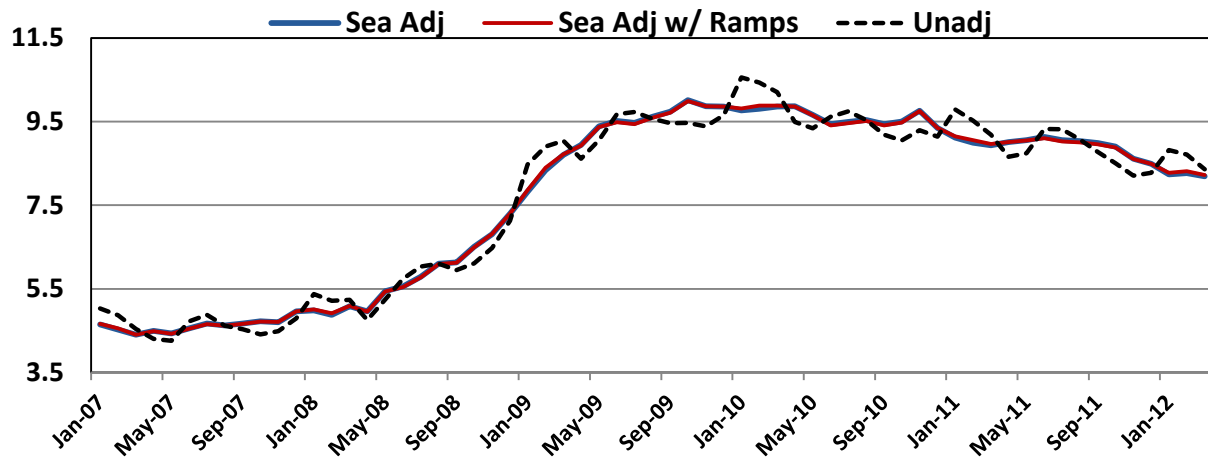
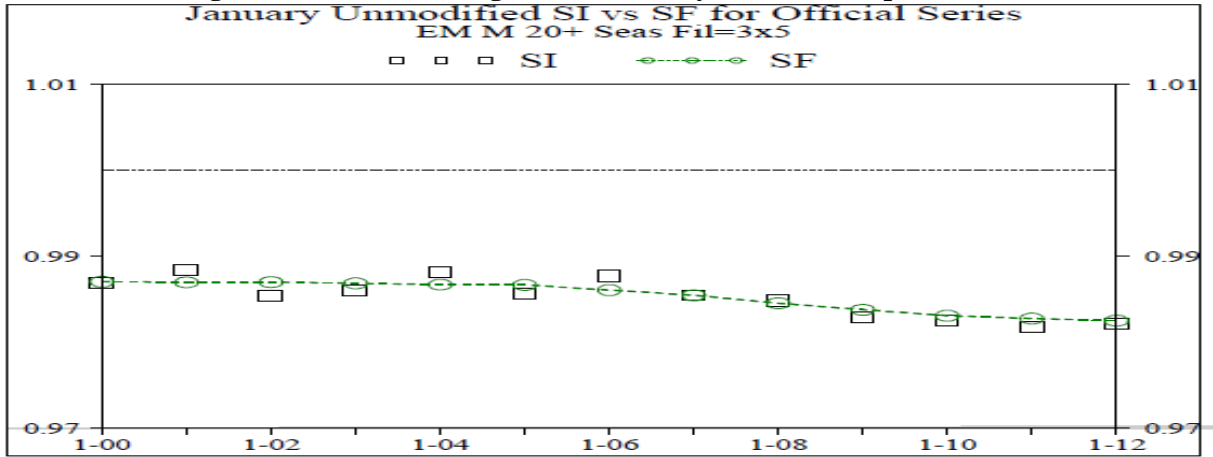


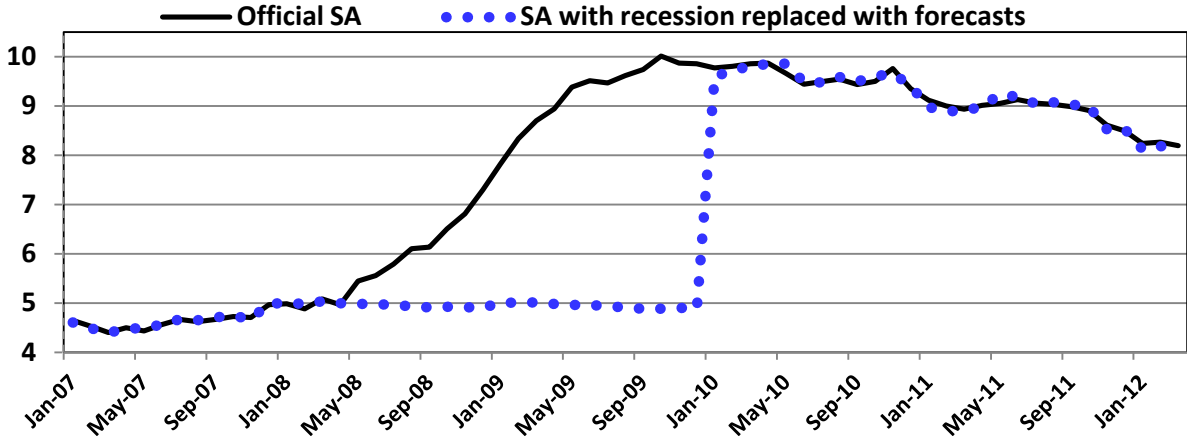
Figure 9: Effects of Modeling Ramps on National UR



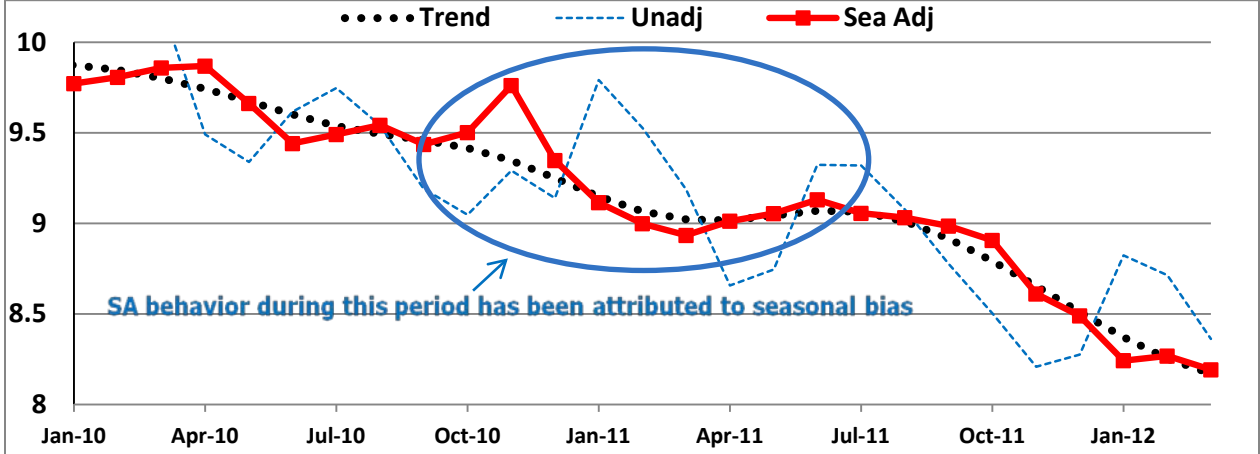
**Figure 10: Seasonal-Irregular Sub-Plot by Month Example**



**Figure 11: Seas Adj National UR with Recession Data (2008-2009) Removed**



**Figure 12: November 2010 UR Blip due to Irregular**



# Seasonal Adjustment of CPS Labor Force Series during the Great Recession

---

Thomas Evans  
Richard Tiller

Exploration of Mathematical Aspects of  
Economic/Official Statistics Conference

Tokyo

January 31, 2014



**BLS**

U.S. BUREAU OF LABOR STATISTICS  
1 2 5 Y E A R S

[www.bls.gov](http://www.bls.gov)

---

## Outline

---

- Quick Review of SA Method  
(extended versions of X-11 family)
- Seasonal Adjustment during recession
- Outlier Issues related to  
2008-09 recession
- Results for key U.S. CPS series
- Summary
- New Research

# X-12 (X-13ARIMA-SEATS)

---

## RegARIMA Models

Automatic Outlier detection, other pre-adjustments, forecasts, ...



## Seasonal Adjustment

### Basic X-11

(with enhancements)



## Diagnostics



## Seasonal Adjustment during recession

---

- Are standard procedures used to estimate seasonal factors during non-recession periods applicable to recession periods?
  - Do the standard trend & seasonal filters adequately reflect recessions?
  - Does the standard automated outlier detection need to be modified?
    - Reduce critical values for outlier identification
    - Introduce user-specified outlier types

# Structural Change associated with recessions

---

- Trend breaks
- Seasonal breaks

## Types of Trend Breaks

---

- Level shift (LS)—permanent abrupt shift in level occurring from a single month onward
- Temporary change (TC)—abrupt change in level followed by gradual return to normal
- Ramp—a change to a new level with a start & end date and a fixed rate of change per period

# Detection of Seasonal Breaks

---

- Analysis of SI plots for discontinuities
- Partial test of regime change for seasonal breaks with RegArima

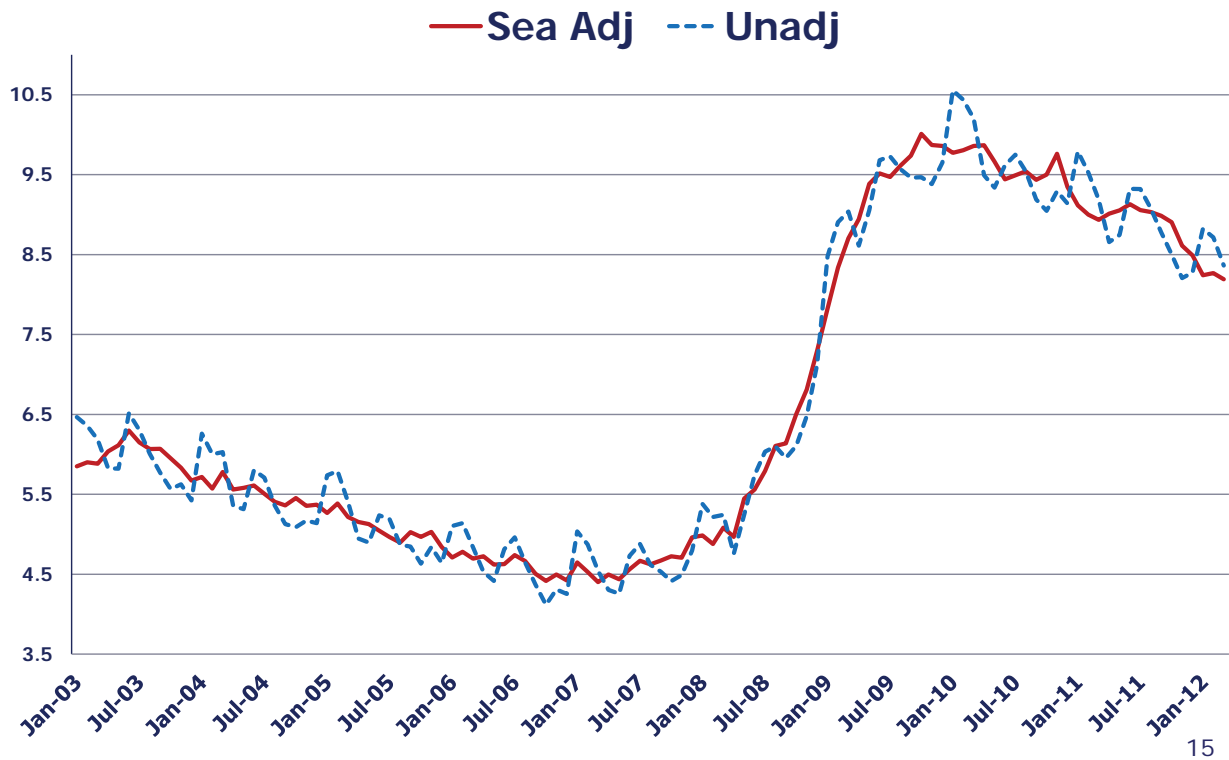
# National CPS Series

---

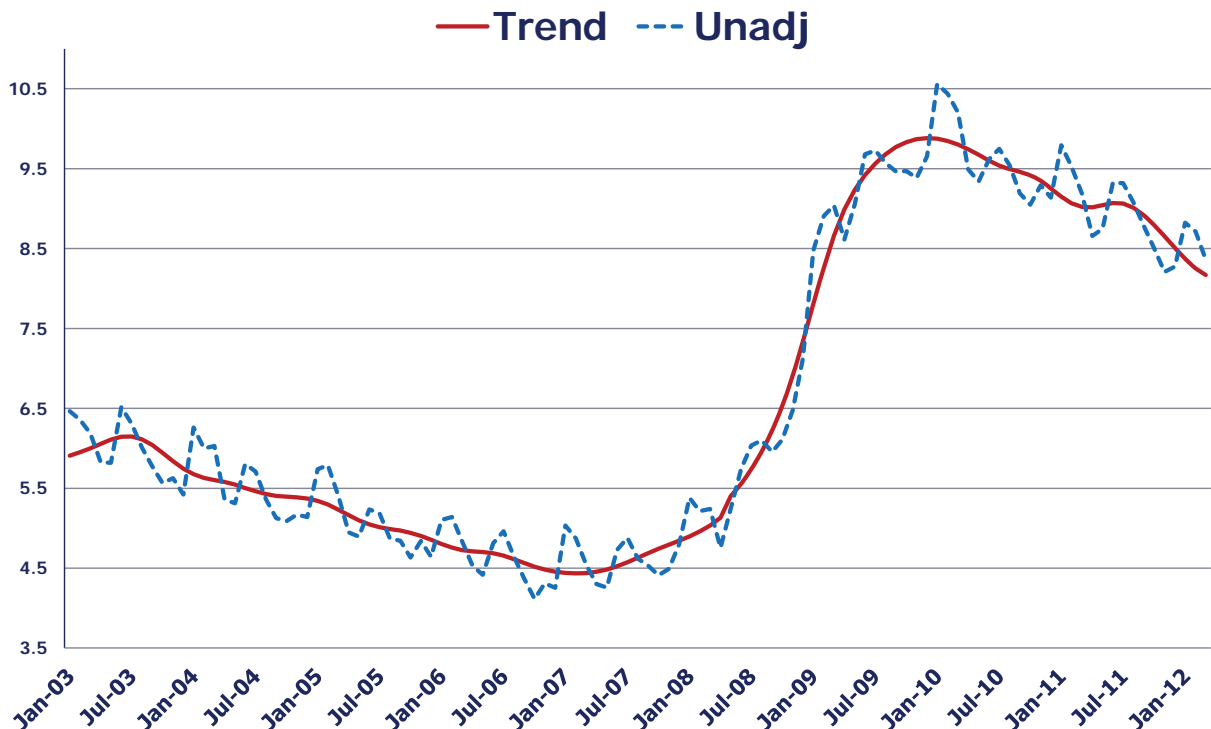
- Eight series: UN and EM adults and teens by gender
- These 8 series are used to derive the national unemployment rate



# National SA Unemployment Rate (official)



# National Trend Unemployment Rate (not published)



# Outlier Issues related to 2008-09 recession

---

- Were there recession related breaks in trend and/or seasonal patterns that caused distortions to the seasonal adjustment process in the post-recession period?
- Was the sharp rise (fall) in unemp (emp) from the 4th quarter of 2008 into the 1st quarter of 2009 absorbed into the seasonal component?
  - ▶ Did this cause seasonal adjustment to over estimate recent economic growth in 4<sup>th</sup> & 1<sup>st</sup> quarters (decline in unemp, increase in emp.) & underestimate in 2<sup>nd</sup> and 3<sup>rd</sup> quarters (increase in unemp, decrease in emp.) ?

# Types of Diagnostic tests for recession effects

---

- Trend breaks:
  - ▶ Automatic outlier detection
  - ▶ Experiment with ramps
- Seasonal Breaks
  - ▶ Examine SI sub-plots by month for discontinuities during recession period
  - ▶ Partial change of regime regression test with change point = Oct 2008

# Results for Level Shifts

- Automatic outlier identification run at the end of each year, no attempt to identify LS in real time
  - ▶ use default critical value  $\cong 4.0$
- Identified 2 LS's during recession period:
  - ▶ May 2008 LS for UN teen male
  - ▶ May 2008 LS for UN teen female (almost critical value =3.8)
- Reduce critical value?
  - ▶ too many spurious outliers, examine "almost" critical outliers for borderline cases
  - ▶ prefer intervention analysis based on prior information, need justification, avoid fishing exercises



# RegArima estimates for Level Shifts

Level Shifts from Automatic Outlier Detection						
Series	Date	Coef (exp)	T-value	AICC w/LS	AICC w/o LS	Dif
UN M 16-19	May 08	1.26	4.4	10,343	10,360	-16.5
UN F 16-19	May 08	1.17	3.1	10,236	10,243	-6.8

UN F 16-19 LS identified with data where t-stat almost met the critical value. Additional data lowered t-value further below critical value

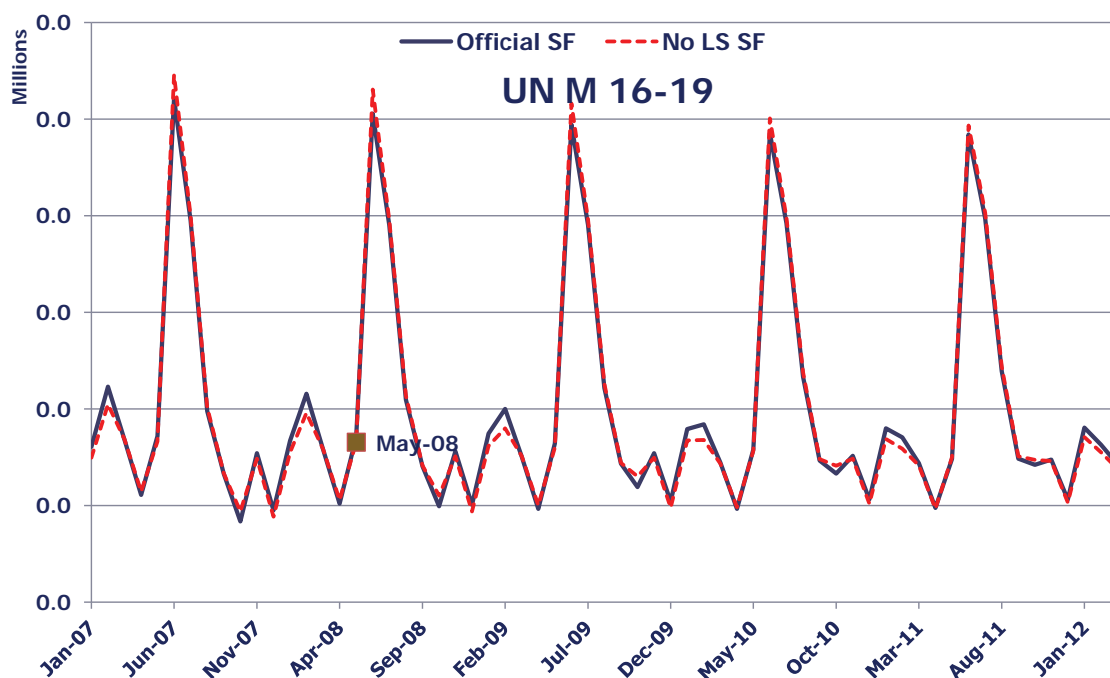
Large relative level increases



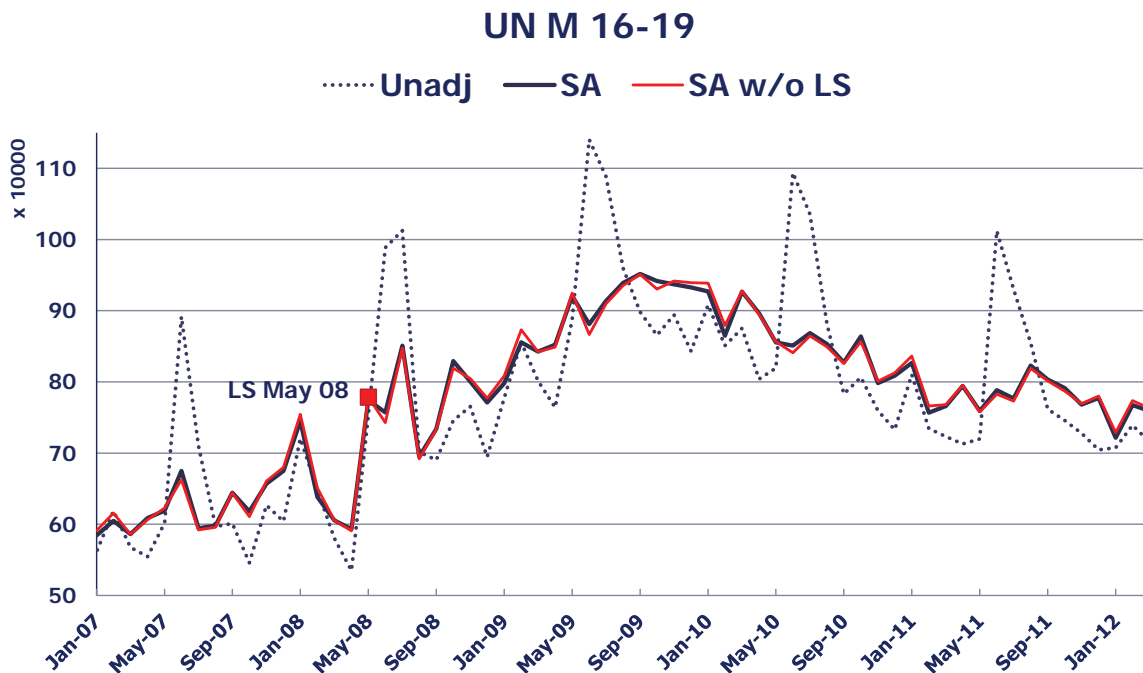
# Did LS's for Unemp teenagers affect SA?

- Surprisingly no, but major effects on trend & irregular components
- If we just examined the seasonal factors would conclude no recession related LS effect occurred
- Trend & irregular components important for interpreting what happen

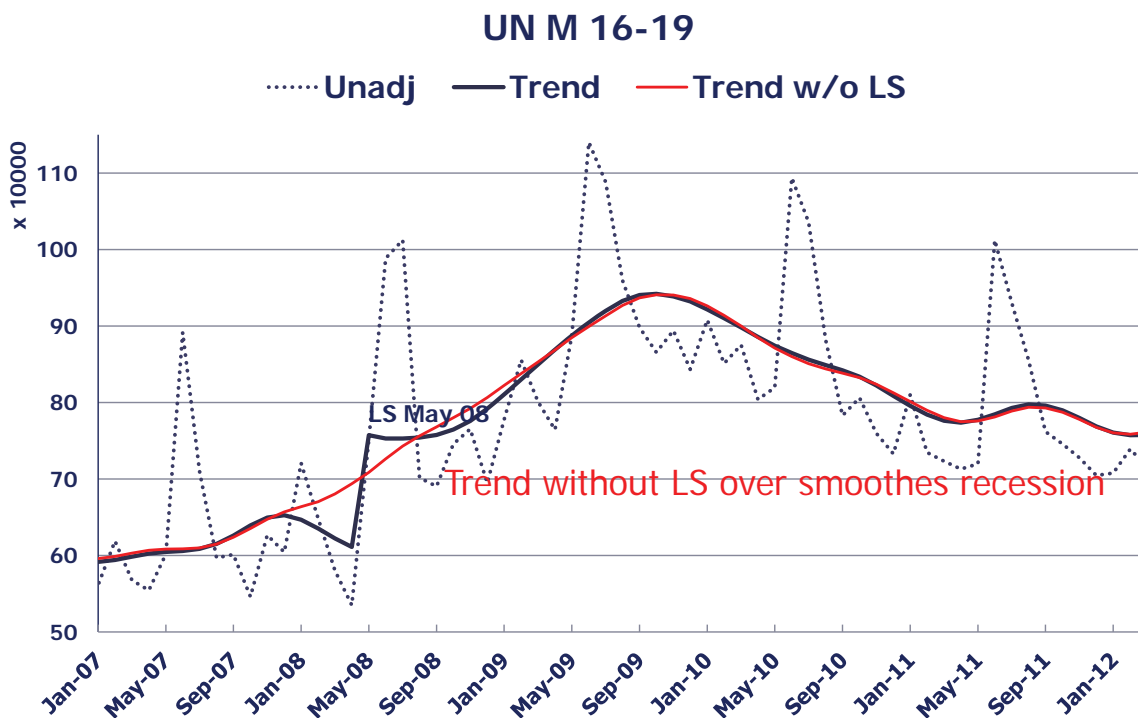
## LS had little affect on seasonal factors for male teens...



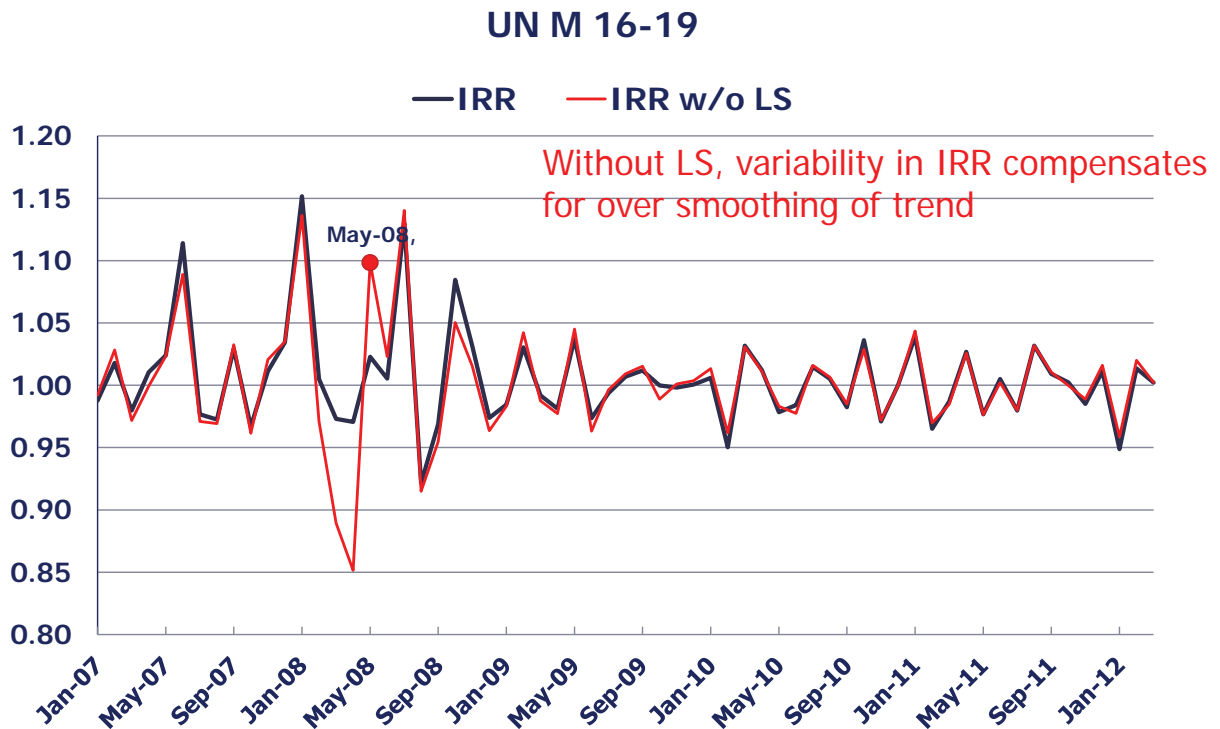
# ...which means little affect on SA series...



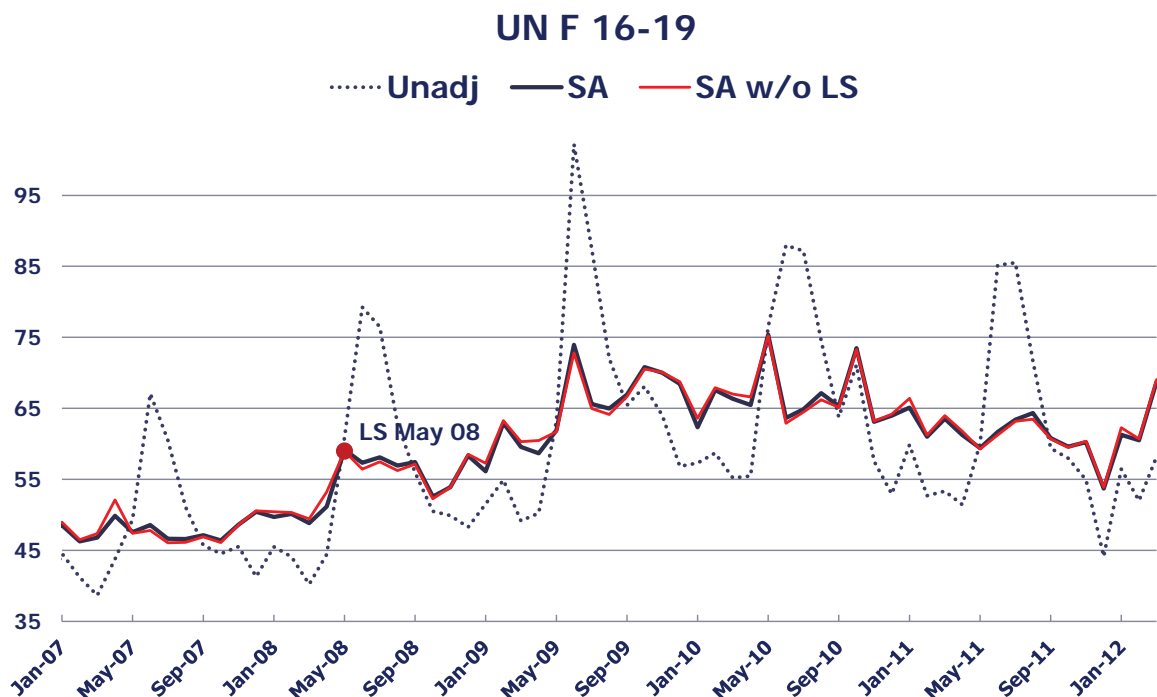
# ... trend clearly shows LS...



## ...variability in IRR causes steep May08 rise in SA w/o LS



## Same result for Female teens but less clear case for LS



# Summary on Level Shifts

---

- Normal automatic outlier identification found recession related LS's for 2 out of 8 series
- If LS's ignored for unemp teens, large upward shift in May would be interpreted as a random deviation from normal as opposed to a fundamental level shift
- Data users would see no noticeable effect on the SA series since seasonal patterns not affected by LS



27

# Experiment with Ramps

---

- Ramps:
  - ▶ Not used for CPS series
  - ▶ Early interest in ramps for BLS price indexes (Buszuwski and Scott, 1993)
  - ▶ Used in selected BLS producer and consumer price indexes
  - ▶ Recent interest in ramps (U.S. Bureau of Census, Bank of Spain)



28

# Ramps

---

$$RP_t = \lambda x_t, \quad x_t = \begin{cases} 0 & t_0 \leq t \\ \frac{t - t_0}{t_1 - t_0} & t_0 < t \leq t_1 \\ 1 & t_1 < t \end{cases}$$

- Constant rate of change ( $1/(t_1 - t_0)$ ) over interval  $t_0 : t_1$  to a new level ( $\lambda$ )
- Not automated, requires user to specify beginning & end points of adjustment period

## How to select $t_0$ & $t_1$ ?

---

- Not much guidance in absence of prior information
- Visually look for segments of series where change appears to be relatively constant
- Fairly easy to achieve “better” fits but are they spurious?



# Results for CPS Data

## Ramps

Series	Ramp Start	Ramp End	Exp Ramp Coef	T-Value
EM M 20+	2008.Nov	2009.Mar	0.99	-5.66
UN M 20+	2008.Apr	2009.Feb	1.06	4.92
UN F 20+	2008.Apr	2009.Feb	1.04	4.79

Small effects

## Revision Medians Avg Abs %

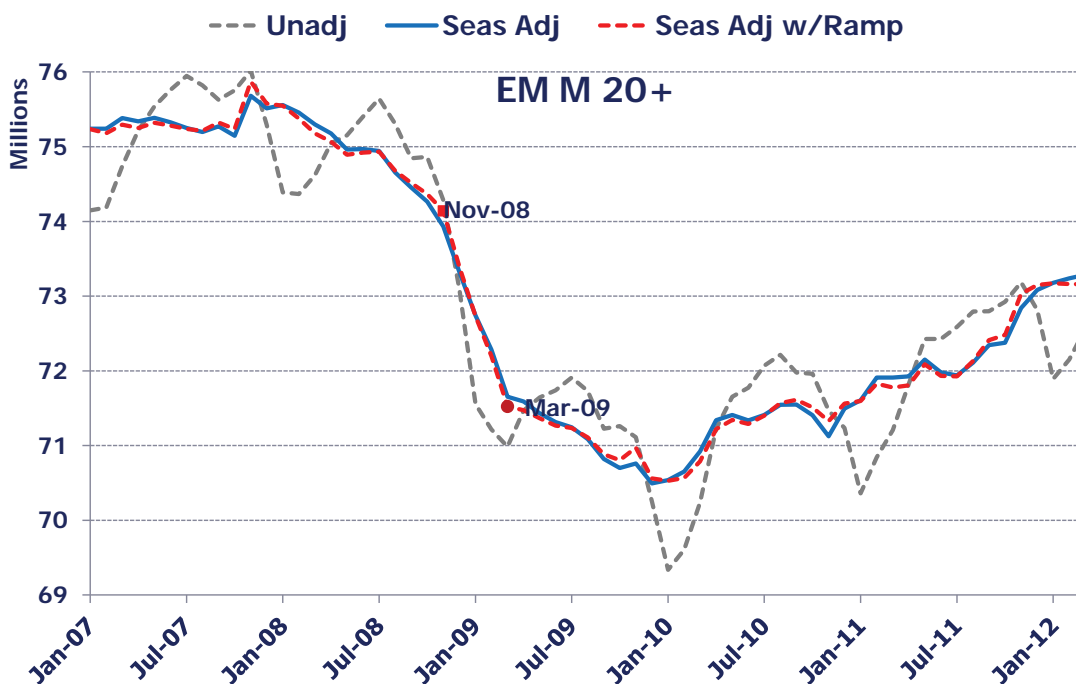
	Official	With Ramps	AICC Official	AICC w/ Ramp	AICC Dif
EM M 20+	0.07	0.08	11,511.45	11,484.39	-27.06
UN M 20+	0.78	0.81	6,822.91	6,805.60	-17.32
UN F 20+	0.83	0.81	6,732.70	6,713.66	-19.04

Pass minimum AICC criterion?

31

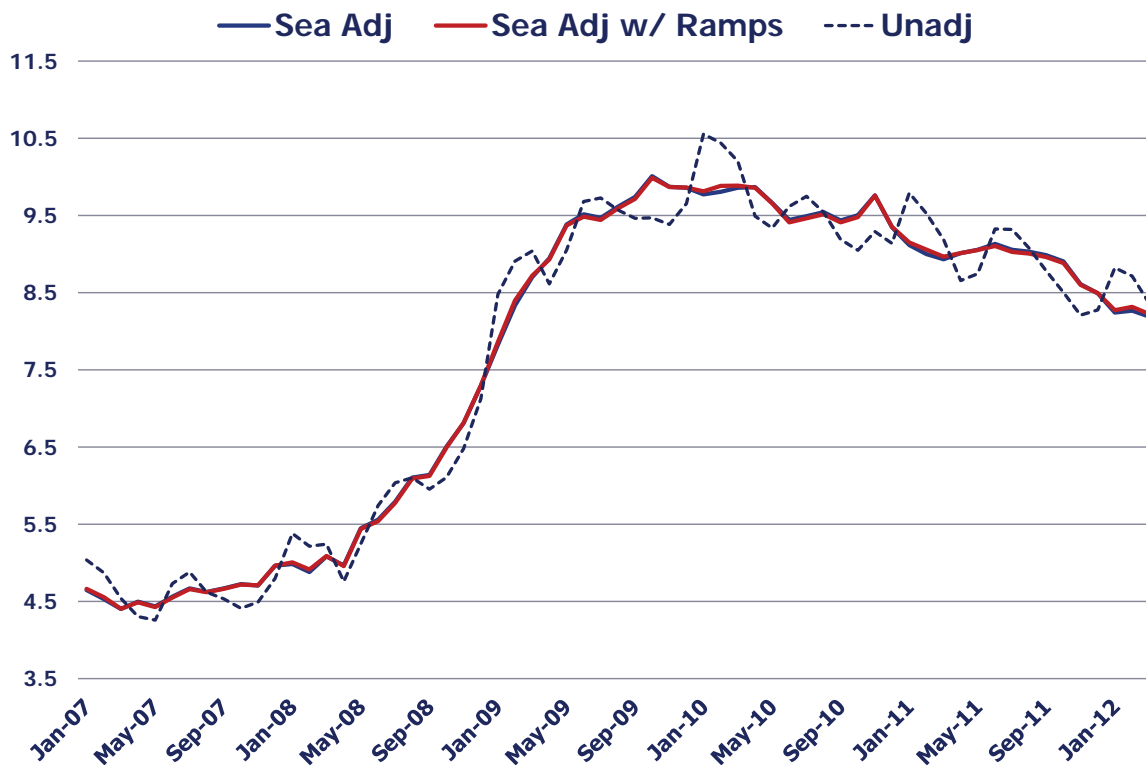


## Ramp coefficients imply small effect, illustrated with EM M 20+



32

# Modeling ramps had little effect on National UR



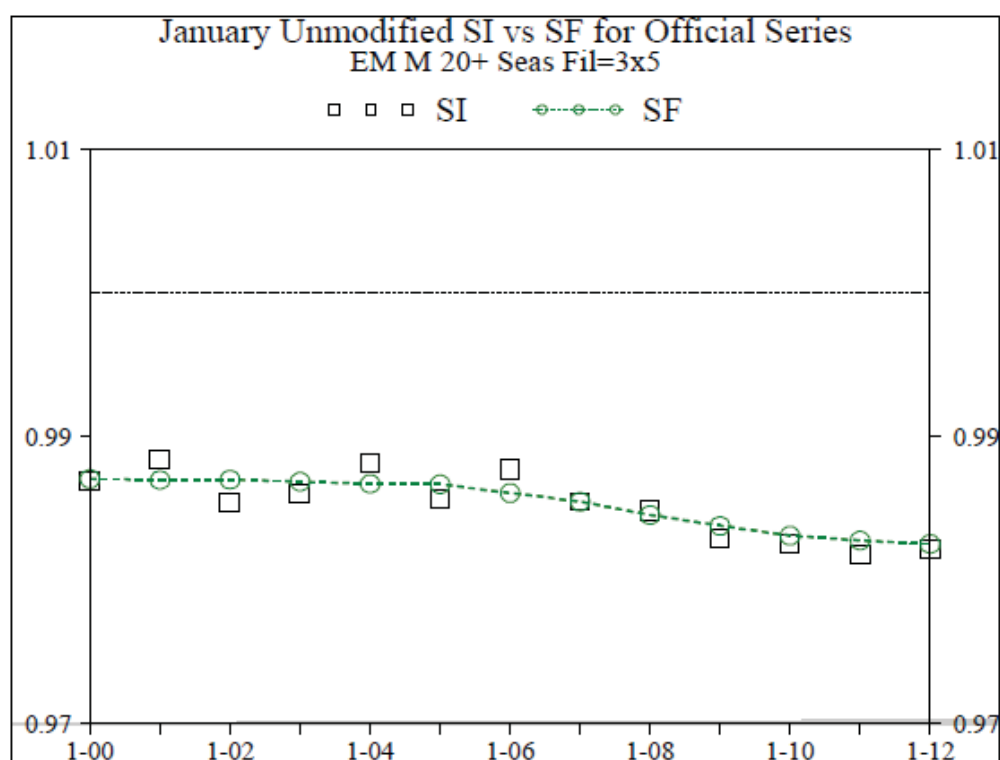
## Evidence on Seasonal breaks

- Seasonal Breaks  
Visual examination of SIs show no apparent breaks
- AICC test rejects presence of deterministic seasonal break in all 8 series, breakpoint = Oct 2008

# Seasonal-irregular (SI)

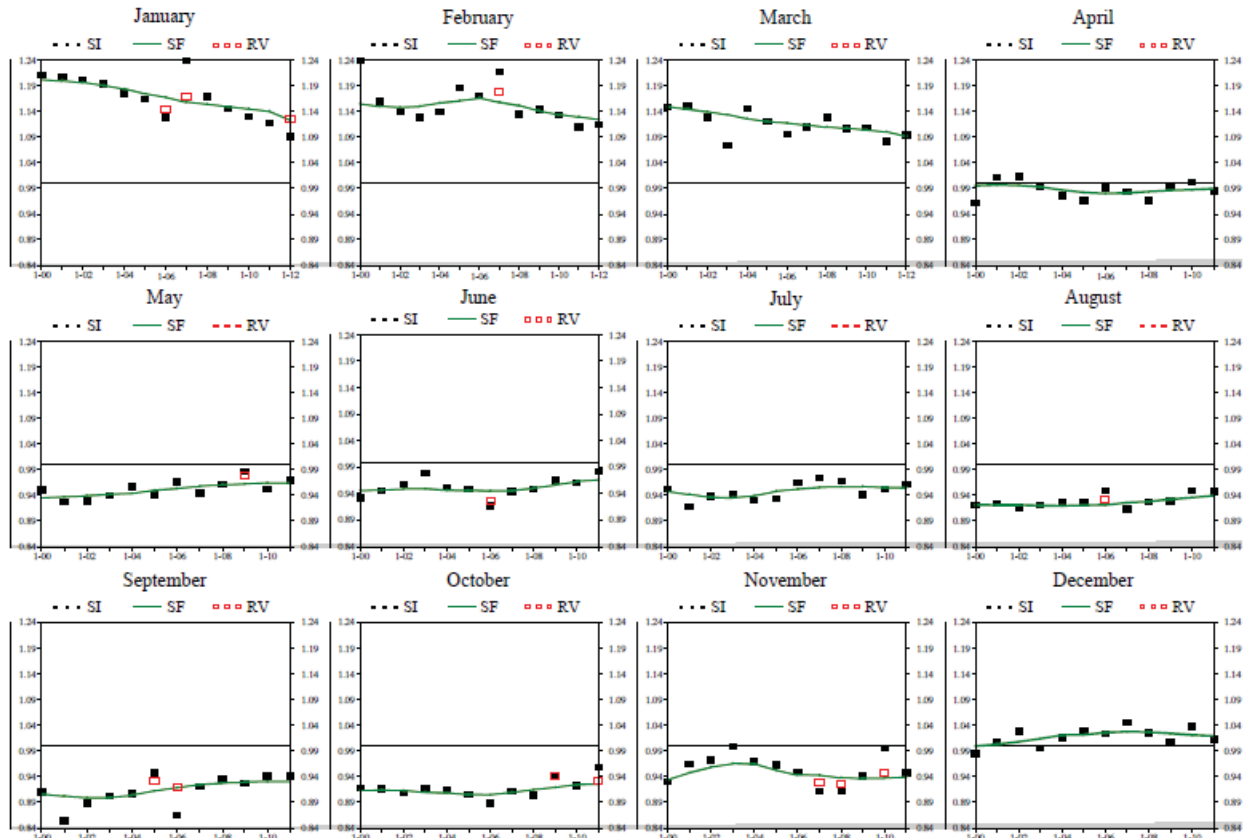
- De-trended series produced by X-11
  - ▶ Estimate & remove trend-cycle from original series using Henderson filter
  - ▶ seasonal factors are derived by smoothing SI's by month usually with a 3x5 MA
- SI sub-plots by month useful for identifying potential breaks in seasonality—found no evidence of breaks

## Seasonal-Irregular sub-plots by month



## UN M 20+ Monthly Unmodified SI Sub-Plots for Official Series

RV = SI Replacement Values  
SF = Seasonal Factors (3x5 MA of SI)



# AICC criterion rejects Seasonal break in all 8 series

Series	AICC		
	Official	w/Break	Dif
EM M 16-19	10,698.7	10,711.3	12.7
EM F 16-19	10,672.0	10,685.4	13.3
EM M 20+	11,511.5	11,517.7	6.2
EM F 20+	11,525.6	11,527.9	2.2
UN M 16-19	10,343.3	10,348.6	5.3
UN M 16-19	10,235.7	10,257.7	22.0
UN M 20+	6,822.9	6,829.9	7.0
UN F 20+	6,732.7	6,748.4	15.7

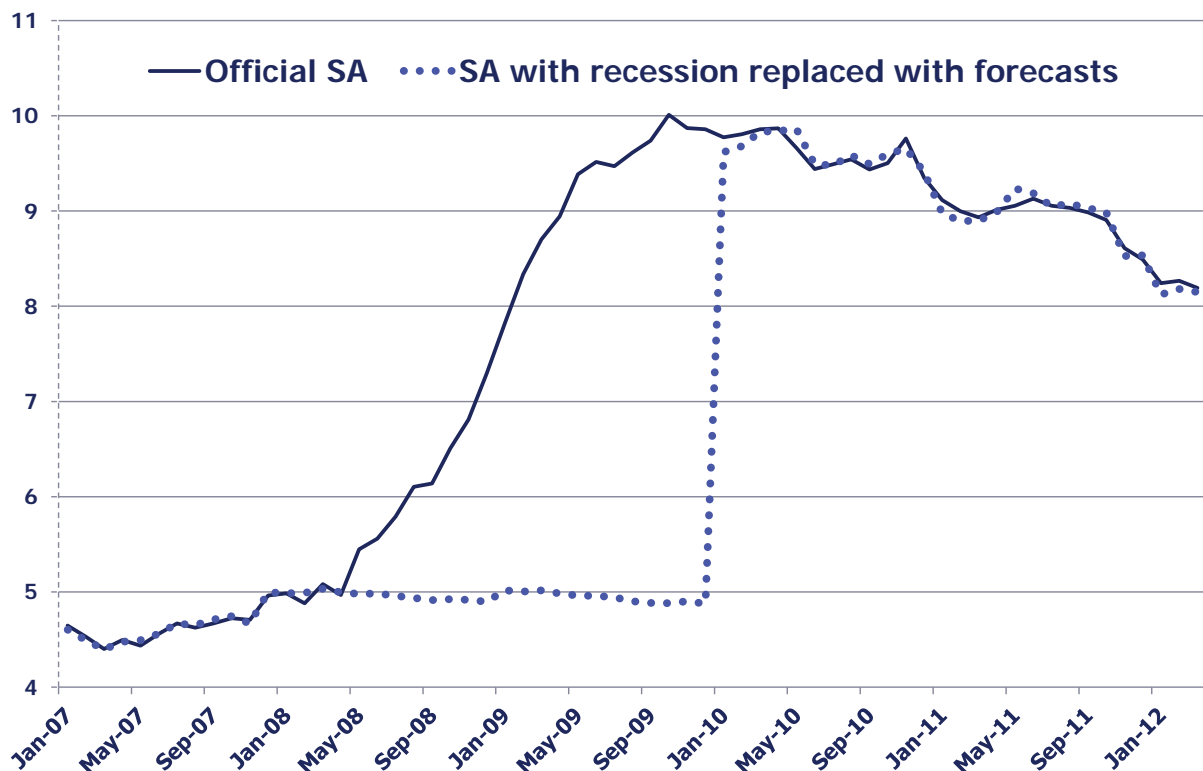


breakpoint = Oct 2008

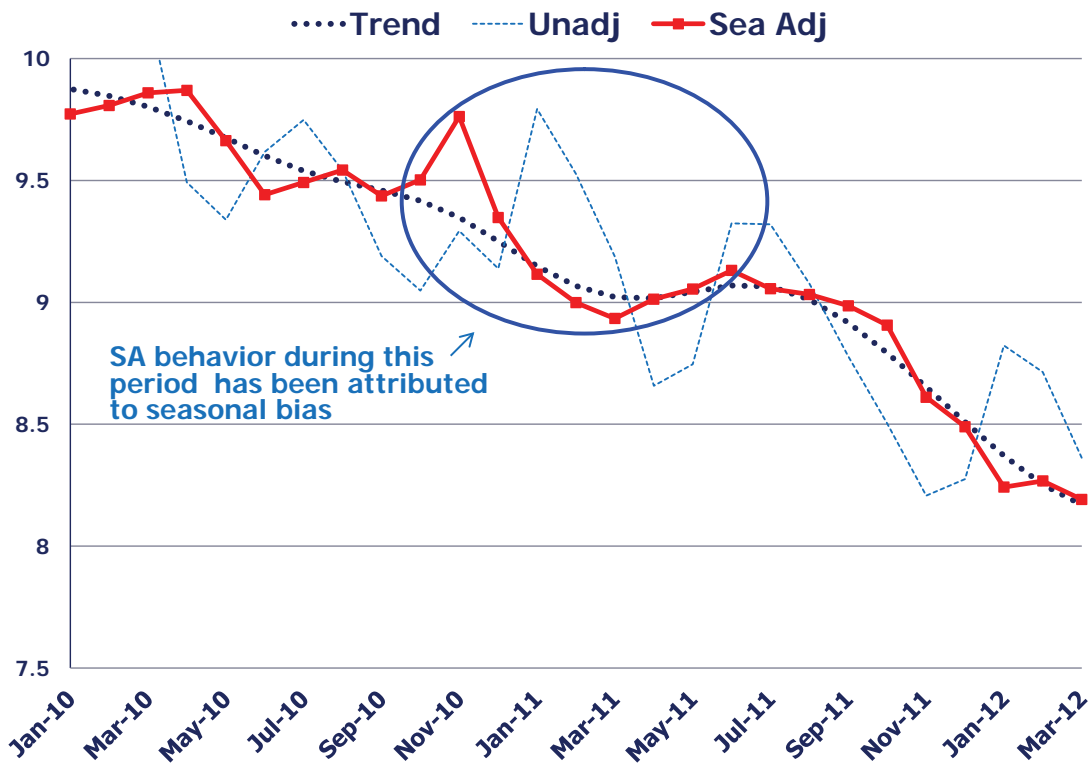
# Performance of Post-Recession Seasonal Adjustment

- Examine by treating the recession data as missing (Jan 2008-Dec 2009) when seasonally adjusting Jan 2010-Mar 2012
  - ▶ Set recession values as missing, RegARIMA automatically adds AO regressors to produce forecasts based on pre-recession data
  - ▶ Original series with forecasted values for 2008-09 seasonally adjusted with Jan2010 level shift added
- Not very insightful exercise since unemployment still much higher than pre-recession levels

## SA UR with recession data removed (recession period 2008-09)



## Nov 2010 blip due to IRR (not seasonal bias) eliminated by trend filter



## Summary

- Six of eight series needed no outlier adjustments during recession
- Recessionary level shifts identified in normal way for unemp. male & female teenagers, reduced IRR variation, little effect on SA series
- Fitting ramp models had little effect
- Removing recession data had little effect
- Standard X-11 symmetric filters appear to perform well during & after 2008-09 recession

# New Research

## Seasonal Adjustment and the Great Recession: Implications for Statistical Agencies

Richard B. Tiller and Thomas D. Evans  
Bureau of Labor Statistics

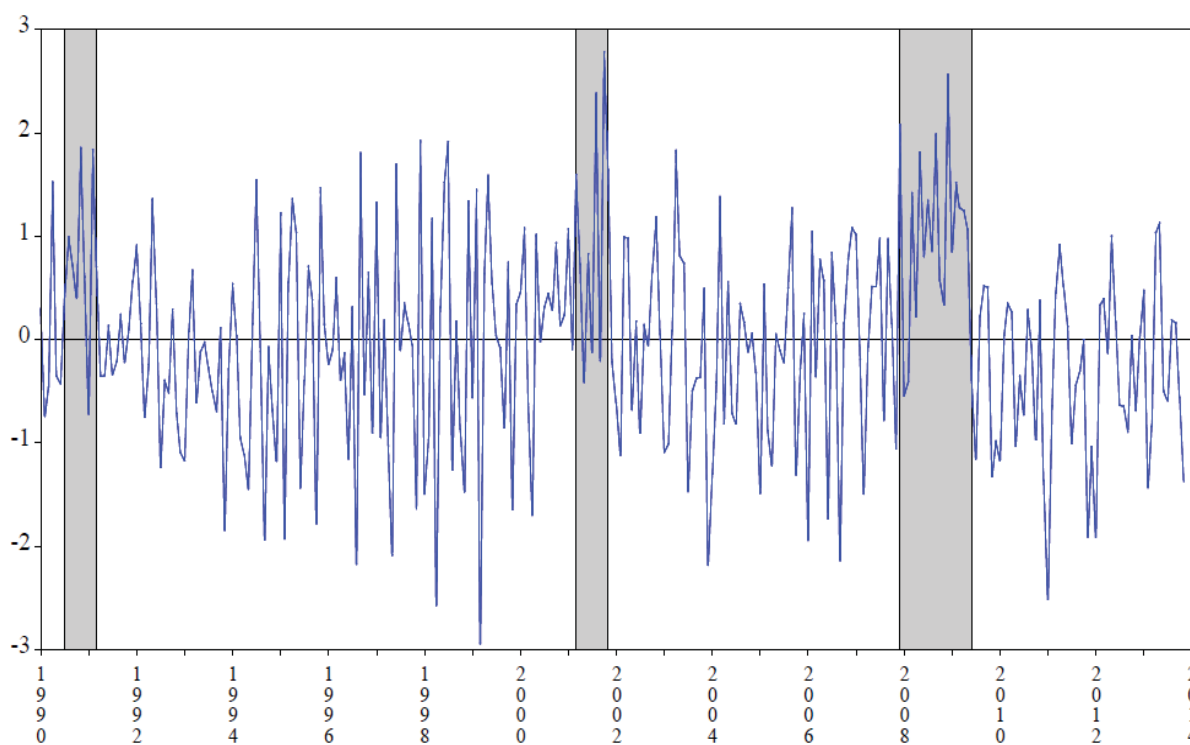
to be presented at:

2014 Joint Statistical Meetings in Boston



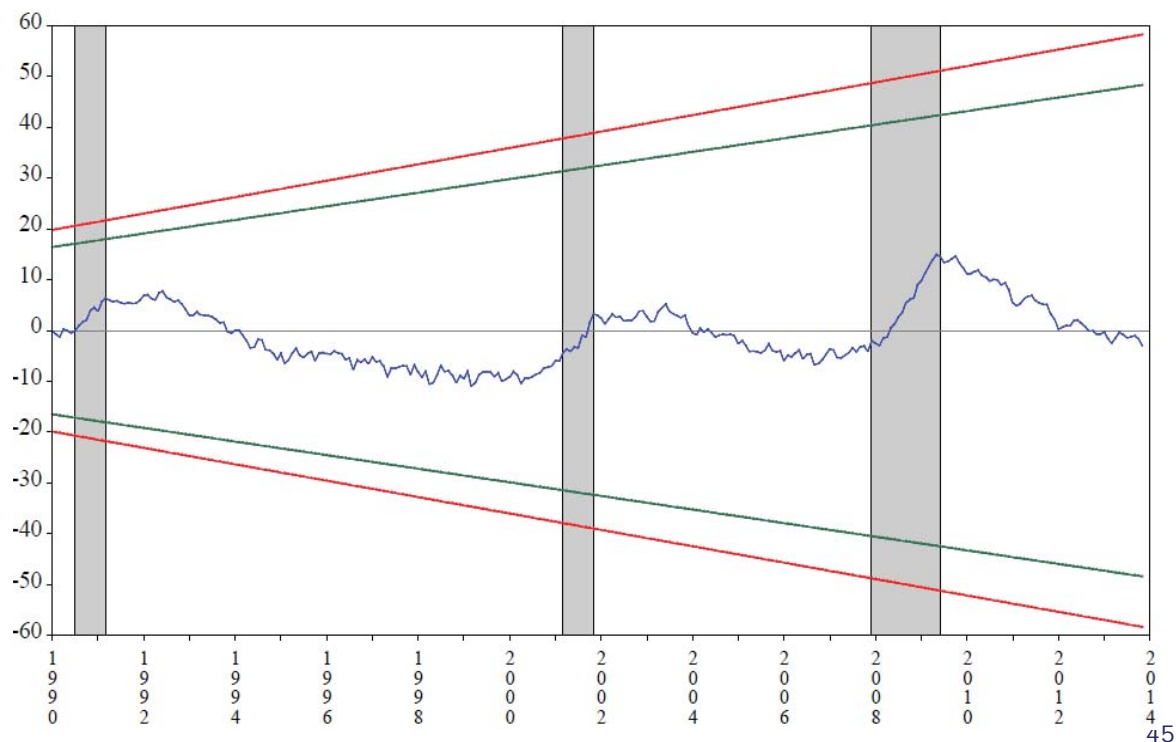
43

## UN M 20+ Standardized ARIMA Residuals



# UN M 20+

## Cusum of Standardized Residuals



## Contact Information

**Thomas Evans**

1-202-691-6354

[evans.thomas@bls.gov](mailto:evans.thomas@bls.gov)



[www.bls.gov](http://www.bls.gov)