

SSJDA-RPS 62

January, 2018

Social Science Japan Data Archive Research Paper Series No. 62

Labour, Education, Lifestyle and Family in Southern Europe and East Asia

‘Open Call Research Project Type Secondary Data Analysis Programme 2015’
Joint Usage and Research Center Programmes at the Center for Social Research and
Data Archives, Institute of Social Science, The University of Tokyo
Sponsored by the Ministry of Education, Culture, Sports, Science and Technology

文部科学省 共同利用・共同研究拠点事業
社会調査・データアーカイブ共同利用・共同研究拠点

東アジアと南欧諸国における労働，教育，ライフスタイル，
および家族にかんする比較研究

2015 年度課題公募型二次分析研究会 研究成果報告書
東京大学社会科学研究所 附属社会調査・データアーカイブ研究センター
2018 年 1 月

Contents

<i>Preface</i>	1
序文	3
PART I LONGITUDINAL DATA ANALYSIS	
女性の労働参加からみる南欧と東アジア諸国の類似点と相違点 福田隆巳	7
結婚・出産タイミングの日韓比較 戸田淳仁・福田隆巳	41
母親の就業状況と家計の教育費支出——韓国労働パネルデータを用いて 李秀眞	51
PART II MICRODATA AVAILABILITY	
Are Data Public Goods or Intellectual Properties? Dan SASAKI	63
A Note on Microdata Availability in Japan: Accessibility and Practical Information Rui FUKUDA	77
APPENDICES	
Activity Records	101
Research Workshop Poster	102

Preface

Currently, many countries are suffering from aging populations, and are thus facing the difficulty of maintaining social security systems. This is particularly the case for Southern European and East Asian countries where family ties play an important role in society; there are clear similarities in terms of demographic trends in these countries.

However, differences and similarities across those countries have not been examined from comparative perspectives using longitudinal microdata. This study aims to clarify similarities and differences between Southern European and East Asian countries, i.e. Spain, Italy, Japan and Korea, in order to contribute to current theories as well as to policies for existing social systems.

Based upon this idea a research group was formed. In April 2015, the group applied to a programme entitled ‘Open Call Research Project Type Secondary Data Analysis Programme’ of the Centre for Social Research and Data Archives (CSRDA) at Institute of Social Science, University of Tokyo. Applicants for the programme had to form a study group, and suggest a topic and select appropriate datasets from the data archive of the CSRDA for the study. For accepted study groups, the CSRDA provided microdata sets, facilities and places to hold group meetings. After receiving approval from the CSRDA, the group was supported by the institution. Research group meetings were held three times and then several papers were presented at a workshop at the end of the first fiscal year.

The study uses longitudinal microdata sets and focuses on topics such as labour, education, family formation patterns, lifestyles, and the linkage between the family and the individual’s behaviour in Southern Europe and East Asia. The approach has been made from diverse perspectives, such as labour economics, population economics, political economy, home economics, and sociology. Using longitudinal data allows for a focus on the individual’s behaviour from various aspects.

Apart from empirical research based upon microdata analysis, the group also studied microdata capabilities as a topic itself in the countries above. Therefore, this report will also be significant for the improvement of reproducibility of studies.

Finally, five papers were reported under the responsibility of each author. Due to time constraints, however, most of the papers have been written practically within one year, and therefore the project could only report provisional results. There are three papers on longitudinal data analyses in Part I, reported in Japanese, and two papers on microdata

availability in Part II, reported in English. Nevertheless, as mentioned, the main aim of the project was to provide research ideas and information on research infrastructure. Deeper and more sophisticated studies may come at a later date, from those involved in the project or otherwise.

Special thanks should be given to two advisors of the study group who provided extremely helpful comments and supports. Ryuichi Tanaka, gave detailed comments on microdata analyses and was the discussant for the group's workshop, and Noboru Koshizuka, also provided important feedback on microdata availability research as well as giving a lecture on 'Open Data' and social innovation from computer science perspectives that motivated the group.

January, 2018

Contributing authors

序文

近年、多くの先進経済諸国が人口の高齢化に直面し、また同時に、社会保障制度の維持の困難に直面している。人口学的な動態については、とりわけ南欧諸国と東アジア諸国に同質性がみられ、背景としての「家族」のあり方と社会の変化との関連性の存在が指摘される。他方で、両地域の異同につき、個人レベルの縦断的データを用いた分析、かつ国際比較的な視点からは明らかにされていない点が多いといえる。

学術的・政策的な含意を得るためにも、両地域における社会の様相のより厳密な把握が必要となると考えられる。本研究では、東アジア諸国と南欧諸国——日・韓・西・伊——における、労働、教育、家族形成、ライフスタイルおよびそれらと家族との関連のあり方について各国の縦断的データを用いて探究する。

縦断的データを用いることによって、より精密に多角的な視点から個人の行動パターンの追求が可能となる。加えて、国際比較的な視点から各国の異同を相対化することにより、多方面からの日本への政策的含意を得られることが期待できる。本研究では、主として日本の「働き方とライフスタイルに関する全国調査」データを利用し、他国の縦断的マイクロデータも併用しながら、実証分析を行う。

また、本研究ではマイクロデータ分析と併せて、社会科学の再現性の向上に貢献することを意図し、各国におけるマイクロデータの利用可能性についての俯瞰的な研究を行う。マイクロデータの利用可能性そのものを、学術研究の水準を左右する重要な研究基盤として捉えるべき課題と考えたためである。

そのような研究関心の下、本研究会が発足・運営された。しかしながら、本報告書の発行にいたるまでに数多くの制約に直面したことから、本報告書は暫定的なアイデア・研究基盤の紹介をするものにとどまっている点にご留意いただきたい。マイクロデータ分析においては、研究会の掲げた目標とは乖離が生ずるが、全4カ国分のデータを利用した比較研究は必須要件とせず、最低1カ国分のデータ分析を行うことのみを条件に課した。時間の制約および各国データのハンドリングの煩雑さの理由による。本研究会の活動報告としては、計3回の研究集会の開催、および1回の中間成果報告会の開催となる。

最終的に、著者各人の執筆責任の下、計5本の論文が集まり、それらが本報告書に収録されている。本報告書は、縦断的データを用いた実証研究の第I部、およびマイクロデータの利用可能性を俯瞰的に考察した第II部の全2部構成となっている。第I部は日本語論文が3本、第II部は英語論文が2本収録されている。繰り返しになるが、本研究は基礎的なアイデア・研究基盤の紹介にとどまっていることから、それらを活かした今後のさらなる研究に期待したい。

最後に、数多くの方々にご支援いただいたお蔭で、本報告書の発行に至ること

ができたことに感謝申し上げたい。なかでも、研究会全体としてお世話になった2名のアドバイザーである、田中隆一先生および越塚登先生に謝辞を述べたい。田中先生は、本研究会メンバーの実証論文へ実践的なアドバイスを下さり、また、中間成果報告会ではコメンテーターを務め有益なコメントを下さった。越塚先生は、第1回研究集会にて「オープンデータ」と社会的イノベーションの実際例についての報告を担当し研究会メンバーに研究課題の分析枠組みや視点のアイデアを下さり、また、データの利用可能性論文へ計算機科学の視点から有益なコメントを下さった。本報告書の発行に至ることができたのは、お二方のお力添えによるところが大きい。心より感謝申し上げます。

2018年1月

著者一同

PART I LONGITUDINAL DATA ANALYSIS

女性の労働参加からみる南欧と東アジア諸国の類似点と相違点

福田隆巳

(東京大学大学院経済学研究科)

2017年3月

本稿では、男性稼ぎ主型社会である南欧と東アジア諸国における、女性の就業行動に縦断的な視点から着目した。記述的なマイクロデータ分析により、両地域の女性の労働力状態・労働時間の遷移、労働時間の分散を概観し、各国におけるそれらの多様性を明らかにした。南欧諸国では女性の労働市場行動の硬直性および学歴差が明瞭に観察され、韓国では無業とフルタイム就業との間の出入りが多い傾向が観察された。就業女性の労働時間の恒常的異質性・分散の分析では、日本の女性間の異質性が大きく表れ、同異質性は観察可能な家族形成行動により説明される度合いが大きいことが確認された。他方、南欧諸国では家族形成による同異質性の説明可能性は低く、そのことは学歴や家族形成が就業選択の場面でより作用するためだということが示唆された。本研究は、労働市場の硬直性・短時間就業の機会構造と女性の就業決定・労働時間決定との関連の探求可能性を示唆し、今後のさらなる研究のための基礎資料を提供した。

1 はじめに

本稿は、主として女性の労働参加に着目しながら、記述的に縦断的データの傾向を概観し、南欧と東アジア諸国の類似点と相違点を確認することを目的とする。

近年、多くの国々が少子高齢化という人口学的な課題に直面し、社会保障制度の維持の困難に直面していることから、女性の労働参加を探求することの重要性はとみに増加している。先進経済諸国の中には、女性の労働市場参加の面で劇的な変化を遂げた国がいくつかあり、この変化が「受動的な女性から能動的な女性への変化 (from passive actors to active participants)」と形容されることもある (Goldin 2006)。他方で、キャリア中盤における女性の労働市場参加の低水準が継続し、同時に仕事と家族生活との調和の面で多くの困難を有する国々がある。これら各国間の差異の背景として、従前より社会に存在する性別役割分業観の違いが、様々な面での女性の地位の変化タイミングに影響している可能性が窺いしれる (Feyrer et al. 2008)。

本研究では、家族政策が脆弱で福祉の拠り所として家族を軸に据える、南欧諸国と東アジア諸国における女性の就業行動の特徴を把握することを試みる (各国の社会支出については Appendix A を参照されたい)。関連する社会規範について、例えば Alesina and Giuliano (2014) では、諸制度や労働市場パフォーマンスなど社会の多くの側面と家族紐帯意識の強さとの相関が指摘されている。文化的信念や社会規範は、前近代社会から根付く、時間を通

じて安定的なものであり (Alesina et al. 2013, Alesina et al. 2014), 上の世代から次世代へと継承されるものである (Giuliano 2010, Alesina and Giuliano 2014: 207-210, Alesina et al. 2014, Giuliano 2015, Hwang 2016). さらにいえば, 文化的信念や社会規範は, それ自体に基づく制度や政策によって自己強化され, 経路依存性を有すると考えられる (Alesina and Giuliano 2014: 208, Giuliano 2015: 51). すなわち, 政策の内生性の存在が示唆される.

他方で, 外生的とみなされることが多い, 技術革新などの力があつた場合でも, 長いあいだ社会に醸成され根付いてきた社会的態度に影響を受ける領域——典型的には女性へのケア・家庭責任の帰属——は, すぐには変化しない (Feyrer et al. 2008: 4). 個人や企業的意思決定は, 社会規範やそれに基づく帰属意識に規定されているためである (Akerlof and Kranton 2000, Juhn and McCue 2017: 201).

各国の制度的文脈と女性の労働参加との関連を研究する分野に福祉国家研究がある. その中には, 各国間の制度や諸パフォーマンスの差異を, 主として労働者のスキル特性に着目しながら考察する議論がある——それらは, しばしば福祉国家の起源について, 企業を中心に据えアプローチを試みる (Estévez-Abe et al. (2001) の資本主義の多様性論および後続の研究群). 各国間の差異を諸制度の補完性と関連づける視点が軸にある研究であるが, そもそも各国間の諸制度の違いはそれぞれの社会規範を介して作られているため, 国際比較でみた各国間の諸々の差異は過大に示されうる点に慎重となる必要がある (Olivetti and Petrongolo 2017: 206, Juhn and McCue 2017: 200-201, Keck and Saraceno 2013: 307).

他に, 各国の女性の労働参加ライフサイクル曲線の変化と福祉国家の変化とを関連づけて論じる興味深い研究がある (深沢 2000). 同研究では, 横断的データから各国女性の年齢別労働力曲線の経年変化が概観され, 時代ごとの同曲線の形の変化と社会政策の変化とが関連づけて考察されている. 残念ながら, 同研究は横断的な視点の統計に基づくものであるが, ライフサイクルや同一集団の変化を論じたい場合には同一集団を追跡する縦断的な視点が必要となる (縦断的な視点については Appendix B を参照されたい).

福祉国家研究の中には, 各国をいくつかの型に類型化することだけに終始してしまうものも多くある中, 本研究の分析視座に示唆に富む, 含蓄のある研究もある. 経済発展した国々を「生活保障」の視点から 3 類型に区分した Osawa (2011: 23-24) によれば, 80 年代の南欧諸国と東アジア諸国は「男性稼ぎ主型」に類型される. この型に入る国々は, 男性の——より正確には壮年常用雇用男性の——雇用を保護するため労働市場が規制され, 男性が家族を養うに足る「家族給」を稼得できるような社会となっている. 社会保険制度は男性の失職リスクに対応するために作られ, 女性や子どもは男性の扶養家族としての保護を受ける. 女性は子育てや家事に全時間を割くことが期待され, 並行してケアや家族責任に対する公的支援やサービスは限られ, 加えて東アジアや南欧諸国では第 3 セクターも未発達である. このような男性稼ぎ主型社会は, 社会経済の変化——技術進歩, 産業構造の変化, 女性の高学歴化や労働参加の進展など——により, 家族形成に難が生ずると考えられる (Feyrer

et al. 2008, Hwang 2016).

本研究の分析対象の、南欧と東アジア4カ国の2000年代における労働市場の傾向をワンショット統計から一瞥し、男性稼ぎ主型社会経済内での多様性を示すと以下のようなになる(値はOECD Stat., OECD Database. 雇用保護は2008年, その他は2010年の値). まず, 15歳から64歳の就業者に占める自営業者の割合は, 韓国, イタリア, スペイン, 日本の順に高い(各々28.8%, 25.6%, 16.8%, 12.2%). 趨勢的には各国とも同比率が減少傾向ではあるが, 韓国とイタリアでは今日でも比較的雇用労働以外の就業機会が多くある.

雇用者に目を向けると, 25-54歳女性の雇用者に占める週30時間未満の短時間就労者割合は, イタリア, 日本, スペイン, 韓国, 日本の順に高い(各々31.6%, 30.2%, 20.2%, 12.2%). 日本とイタリアでは女性がパート就労をする機会が比較的が多い. 補足をすると, フルタイム雇用男性の労働時間は, スペインとイタリアは欧州の中では長い部類に位置づけられるが(Keck and Saraceno 2013: 300), 日本と韓国では欧州以上にフルタイム雇用男性の長時間労働の問題が存在することはよく知られている.

雇用保護の度合いをみると, 2008年版総合指標の値は, スペイン, イタリア, 韓国, 日本の順に高い(それぞれ3.11, 2.58, 2.13, 1.73). 南欧社会は同指標でみる限り, 趨勢でいえば雇用保護を弱める傾向にあるが, 硬直的な労働市場であることがわかる¹⁾. 関連する点として, スペインとイタリアは非労働力女性が多く, また, 女性就業率の学歴差が大きい社会となっている(Keck and Saraceno 2013)²⁾.

本稿では, マクロおよびマイクロデータを記述的に用い, スペイン, イタリア, 韓国, 日本における女性の行動を観察していく. 記述的に, 縦断的視点からデータの傾向をみていき, 南欧と東アジア諸国の類似点と相違点を確認する. 当然のことながら, 縦断的な視点を貫徹するには諸人口集団の追跡が可能な情報量の豊富な, 中期マイクロデータが蓄積され利用可能であることが必要となるため, その点で大きな制約は残る.

本研究は, 各国の実態に迫った, より深い, 今後の理論・実証研究のための基礎資料としての貢献を果たすものである. 本稿は5節で構成される. 第2節では縦断マクロ統計の整理を通じて事例を検討し, 第3節では分析に用いるマイクロデータを紹介する. 第4節では女性の就業行動の記述的な分析を行い, そして, 最終節では暫定的な小括および今後探求すべき課題を議論する.

2 女性の出生行動と労働市場参加

本節では, 縦断的なマクロ統計から, スペイン, イタリア, 日本, 韓国における女性の出生行動と労働市場行動をみていく. 先述のように, 一方で, 女性の労働参加パフォーマンスの面で「革命的」に大きな変化を遂げたとされる国々もあれば(Goldin 2006), いま一方で, そうではない国々がある. 今日の女性の就業行動——およびそれを取り巻く社会規範——

の起源や生成過程を考えた場合、各国間の違いは前近代の農業のあり方に由来すると指摘する研究がある (Alesina et al. 2013, Giuliano 2014, Giuliano 2015)³⁾。本稿では、戦後世代のみのデータを扱い、上の国々の女性の意思決定を観察していく。

2.1 戦後世代女性のライフサイクルの進展

横断的な指標である合計特殊出生率 (Total Fertility Rates: TFR) でみた場合 (非図示)、近年、多くの国々が同値の低下を経験し、経済的に発展した国々の中ではとりわけ南欧と東アジアでその傾向が著しい。同値が経年的にやや回復傾向を示す国々もあるが、南欧と東アジア諸国は低位水準を維持している。さらにいえば、下図の通り、より精緻な縦断的な指標でみた場合にも同諸国において、晩産化のみならず、同一生年コホート女性が生涯で最終的に産む平均子ども数を意味するコホート完結出生率 (Cohort Completed Fertility Rates: CFR) の低下が示され、少産化傾向が観察される。前者と後者は各々、いわゆる the timing/tempo effect および the quantum effect であり、両効果とも観察されるということである。

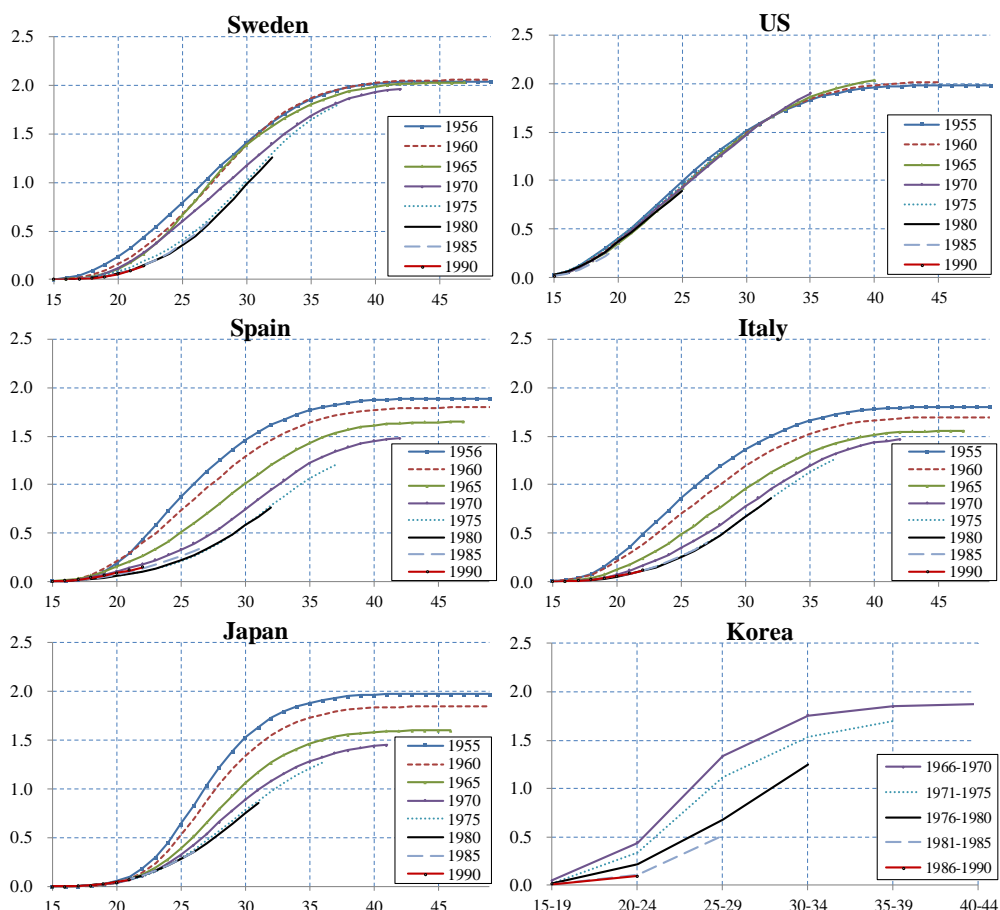


図1 生年コホート別の加齢に伴う累積出生率の推移

出所：EUROSTAT, Centers for Disease Control and Prevention, 国立社会保障・人口問題研究所「人口統計資料集 2014」, STATISTICS KOREA; Vital Statistics (2015年3月取得)

注：アメリカと日本以外は、横断統計を加工し (スライド法)、生年コホート別縦断統計を作成

図 1 は、各国女性の生年コーホート別の年齢別粗出生率を積み上げ累積出生率を示したものである。生年世代は 1955 年生まれから 1990 年生まれまでを捉えている（韓国のみデータの制約上、より短い期間の世代）。図には、南欧と東アジア諸国の特徴を把握するため、スウェーデンとアメリカも含めている。図示されているのは、各生年世代コーホートの各年齢時点での平均的な子ども数を意味する値である。

例えば、（新しい世代はまだ観察されていないが）40 代以降おおむね安定的に推移する同値から、その世代の女性が最終的にもつ平均的な子ども数を把握できる。アメリカおよびスウェーデンでは、どの生年コーホートも最終的（40 歳あたりを超えた時点）には、平均子ども数が 2 人程度に収斂する様子が窺える。一方で、出産タイミングについては両国で異なる傾向にある。戦後世代に限定した図から、アメリカの場合ほどの生年世代も同じようなタイミングで子どもを産むが（曲線間の乖離が少ない）、スウェーデンの場合は生年コーホート間で出産タイミングのピーク（曲線が比較的急勾配で上がる年齢）が異なることがわかる。スウェーデンは新しい世代でやや晩産化が生じているようだが、最終的に持つ子ども数は 2 人近くいるという面で、アメリカ、スウェーデン両国ともに同様である。

他方で、南欧および東アジア諸国では、晩産化のみならず少産化傾向が観察される。具体的には、上の生年世代に比べ新しい世代は曲線が上昇する年齢が右に移動しており（晩産化）、そして累積出生率が安定的に推移する水準が下方に移動している（少産化）。南欧と東アジア諸国の新しい世代は、40 歳を超えた時点での累積出生率が 1.5 前後となっており、一人の女性が生涯に産む最終的な子ども数が従前世代の 2 人前後から速いスピードで減少している（とりわけ日本と韓国でこの傾向が著しい）。さらにいうと、南欧と東アジア諸国では、出生行動の変化は労働市場行動の変化と連関している。

図 2 は、生年コーホート別の加齢に伴う女性の労働力率の推移を示している。生年世代は 1955 年生まれから 1984 年生まれまでを捉えている。大まかな傾向をみると、図より、第 1 に、各国間で女性のライフサイクル労働のパターン（曲線の形）がばらつくことが確認できる。第 2 に、アメリカおよびスウェーデンでは、女性のライフサイクル労働力率水準の生年コーホート間の差が比較的小さい一方で、スペイン、イタリア、韓国、日本は比較的に同差が大きいことがわかる。第 3 に、各国内において、生年コーホート間で労働力水準が異なる場合であっても、ライフサイクル労働参加のパターン（曲線の形）は世代を超えて概ね安定的であることが観察される。日本と韓国の特徴は 30 代前半で窪みが生じる M 字型がはっきりしている点である。全体のおおまかな傾向は以上の通りである。

さらに図 2 を観察していくと⁴⁾、まず截然たる傾向としては、スペイン、韓国、日本では明瞭に、イタリアではやや、新しい生年世代ほど若年期の曲線の労働市場参加の水準が高い傾向が窺える。とりわけスペイン女性の生年世代間の差が大きく、加えて、30 代半ば以降に曲線の水準が収斂する日韓とは異なり、壮年期を通じて生年世代間の水準差が持続する点で特徴的でもある。新しい生年世代ほど労働参加率の水準が高いのは、新しい時代ほど実

質賃金が上昇する一方、所得効果が減少し代替効果がそれを上回る現象が組み合わさったという背景が考えられる（Costa 1998, Blau and Kahn 2007, Goldin and Mitchell 2017）。

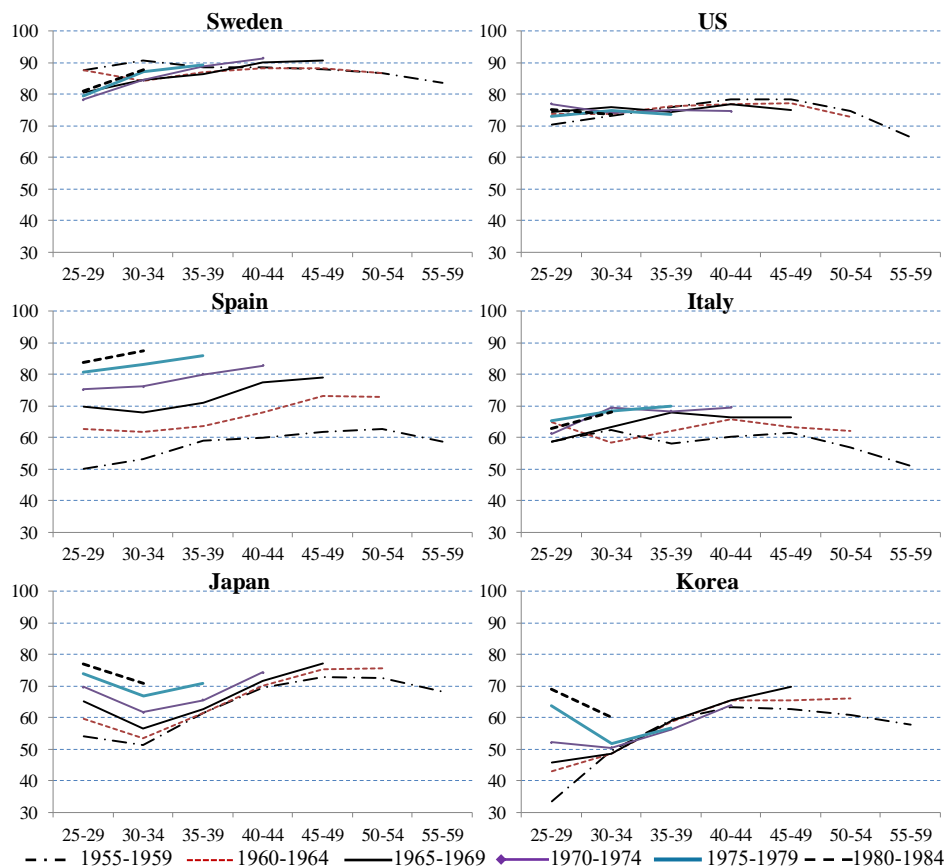


図2 生年コーホート別の加齢に伴う労働参加率の推移

出所：OECD Employment Database（2015年3月取得）

注：(1) 値は労働力率（(就業人口+失業人口)÷人口）で、就業人口は自営業者・家族従業者を含む。(2) 横断統計を加工し（スライド法）、生年コーホート別縦断統計を作成

韓国とスペインは、最も古い1955-1959年生まれの世代は逆U字型の曲線であったが世代を追うごとに、韓国は継続的にM字型に、スペインは1965-1969年世代のみでややM字型を経験しつつも以後の世代ではキャリア中盤の窪みがなくなる傾向、となっていることがわかる。4カ国共通の傾向としては、高齢期の女性の労働参加の増加傾向が概ね窺える。具体的には、図2では、59歳までの数値しか載せていない点や高齢期が観察されていない世代が多い点などに留意は必要だが、高齢期の曲線の裾野がやや広がる傾向が全体として観察される。

以上のように、戦後の生年世代において、平均的な女性の労働ライフサイクルのパターンの各国間・各国内での類似点と相違点が確認できた。他方で、南欧および東アジアのいずれの国も共通で、大小の差はあれ、新しい生年コーホート女性ほど（とりわけ若年期の）労働参加率の水準が上昇傾向にあることが観察された——したがって、潜在的には高齢期の労働

働市場参加の増加も示唆される。

女性間の異質性という観点でいえば、図 2 では、キャリア中盤時点で平均的に比較的少ない子どもを持ち多く労働市場参加する女性の世代、およびその逆の性質の世代が顕在しており、生年コーホート間の異質性が窺える⁵⁾。同時に、世代内でみればより多くの女性が就業している新しい世代は従前の世代より比較的同質的だといえる。上の 4 カ国でも多かれ少なかれ、女性の労働参加の拡大が観察されることから、女性間の異質性は減少傾向にあると推察できる。

2006 年のクロスセクションデータに基づく観察ではあるが、スペインとイタリアは両国とも 25-49 歳女性の非労働力割合が欧州の中で高く、イタリアについては無子女性および有子女性のいずれのサブサンプル別にみても欧州で最大の非労働力率を示す (Keck and Saraceno 2013: 309, Figure 1)。労働市場の比較的な硬直性を加味すると、両国では、就業者と非就業者が比較的に固定的であり女性間の異質性が高い国々だということが示唆されるが、近年の女性の高学歴化や労働参加の増加で趨勢的には同異質性は減少傾向にあるのかもしれない。

若年期女性の労働参加率の水準の上昇については、家事と関連する技術の変化、社会規範の変化 (就業する母親から子への継承)、避妊薬の登場による出生コントロール、近隣からの学習効果 (身近な就業女性の観察) などによって説明が可能となる (Giuliano 2014, Feyrer 2008)。しかし、図 1 および図 2 から容易に想像できるように、一部の国々では女性の労働市場参加と出生行動とは明確な負の共変動関係にある。

2.2 女性の家族形成と労働市場参加の共存の困難

図 3 は、各国の 1950 年代半ば生まれと 1970 年代前半生まれの、2 つの生年コーホート間のキャリア中盤時点 (30 代半ば) での累積出生率と労働参加率の差分を示したものである。データの制約上、戦後の比較的短い生年差の 2 世代間比較である点を留意されたい。縦軸は、上の世代に比べ新しい世代の平均子ども数何%変化しているか、横軸は上の世代に比べ若い世代の労働参加率が何%ポイント変化しているかが示されている。なお、新しい世代ではまだ観察されていない年齢のデータが多いという制約上、本来示すべき 40 歳以上時点での数値の比較を扱えない点を留意されたい。

図 3 をみると、アメリカは縦横両軸共に変化に乏しく安定的で、北欧諸国はそれに次いで変化が小さい。他方、他の国々、とりわけ南欧と東アジア諸国の変化が大きい。これらの地域では、上の世代に比べ新しい世代の女性の 30-34 歳時点での労働参加率が増加している一方で、35 歳時点での子ども数が大きく減少していることがわかる。これらの地域では、女性の労働参加と子ども数とは負の共変動関係にあるといえる。

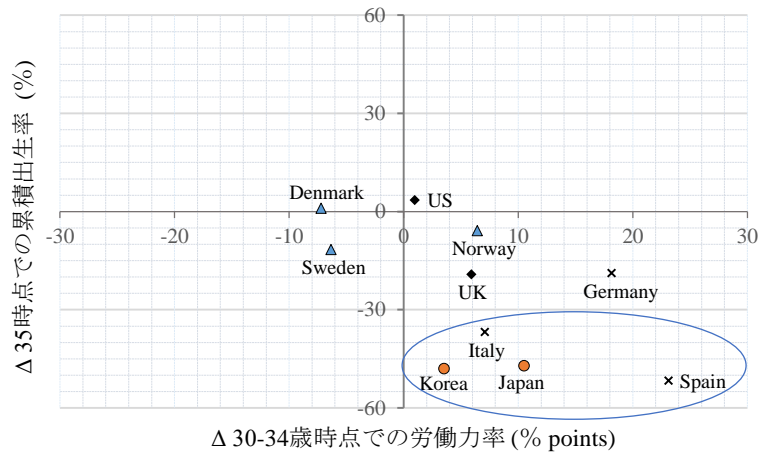


図3 壮年期時点での累積出生率および労働力率の2世代間の差分

出所：出生率の値は図1，労働力率の値は図2と同様のもの

注：(1) 縦軸は、35歳時点での累積出生率（平均子ども数）の2つの生年世代間の差（韓国と西ドイツのみ30-34歳時点での値）。差分の計算に用いた生年コーホートは、ノルウェイ、スウェーデン、イタリア、スペイン、イギリスは1957年と1972年、日本とアメリカは1955年と1972年、西ドイツは1956-1960年と1971-1975年、韓国は1966-1970年と1976-1980年。(2) 横軸は、2つの世代間の30-34歳時点での労働力率の差。差分の計算に用いた生年コーホートは全ての国で1955-1959年と1970-1974年であり、韓国のみ1965-1969年と1970-1974年。(3) 便宜的に、◇はアングロサクソン諸国、×は大陸欧州諸国、○は東アジア諸国、△は北欧諸国という類型ごとにマーカーがグループ分けされている

ところで、図3より、母親ペナルティの存在が示唆される。図3では変化に乏しいアメリカ女性ではあるが、同国女性の生年別ライフサイクル賃金プロファイルを観察した研究では、母親ペナルティという形で、出産と就業にかんする困難が残存・増強していることが示されている（Juhn and McCue 2017）。また、30代半ばで生じた賃金曲線の下落は高齢期にいたるまで従前の軌道に回復することは難しいとされている。一方で、新しい生年世代では学歴・稼働力が高い女性ほど結婚しやすく、多くの子どもを持ちやすいという正のセレクション傾向が強まっている——マイクロデータ分析による個体の個別効果を考慮に入れる前後の分析結果の考察より——という指摘もされる（Juhn and McCue 2017: 197-198）⁶⁾。

日本や韓国では同セレクション傾向はみられず、むしろ高学歴女性が結婚しにくく子どもを産みにくい傾向が観察されている（Hwang 2016）。結婚・出産選択のセレクション傾向や夫婦のソーティング・パターンの変容は、本稿の研究範囲を超えるものであるが、女性の労働を扱う際にとりわけ留意しておくべき重要な視点である。いずれにしても、アメリカにおいてさえも頑健な女性の家族形成ペナルティの存在が観察されており、男性稼ぎ主型社会である南欧や東アジア諸国ではなおその傾向は強いと考えられる。

以下の分析では、パネルデータを用いて、南欧と東アジア諸国における女性の労働市場行動の基礎的な分析を行い、その特徴を把握する。なお、データ期間が短いという制約上、生年コーホート別の分析は行わず、女性全体のデータ期間内の就業行動の変化に着目する。女性間の異質性を直接扱う分析では、家族形成行動との関連にも着目した分析を行う。以下、労働力状態の遷移、労働時間の遷移、および労働時間の恒常的・短期変動的分散をみていく。

3 用いるデータと変数

3.1 データ

本稿の主たる分析には、スペインおよびイタリアの European Community Household Panel Survey の 1994-2001 年データ (ECHP-ESP・ECHP-ITA)、日本の Japanese Life Course Panel Survey の 2007-2011 年データ (JLPS)、韓国の Korean Labor and Income Panel Study の 1998-2005 年データ (KLIPS) を用いる。日本のデータは 20-34 歳の若年 (JLPS-Y) および 35-40 歳の壮年 (JLPS-M) データを結合して用いる。韓国のデータは、欧州のデータ期間に合わせるため、利用可能なデータ期間よりも短い 8 年間分のデータのみを用いる。本稿では、暫定的・便宜的に、利用期間全てに回答しているサンプルに限定したデータを構築し分析に用いる (調査概要やサンプル特性については Appendix C1 を参照されたい)。

なお、4 節の状態移動の分析の解釈と関連するが、各データで観察期間の時代背景が異なり、マクロ経済トレンドも異なる点を留意されたい (マクロ経済トレンドについては Appendix C2 を参照されたい)。本稿の着眼点とは異なるが、利用するデータのクロスセクション時系列統計の傾向を簡単に記す (図表・数値は割愛)。スペインとイタリアは景気拡大期に向かうデータ期間であり、男女ともに就業率が経年的に増加傾向にある。日本は国際金融危機を挟んだデータ期間であり、同イベントでとりわけ男性の就業率が低下する傾向にある。韓国は IMF 経済危機直後のデータ期間であり、就業率でみると男性は経年的に増加傾向、女性は比較的安定的な推移を示す。本稿の分析視角の範囲を超えるが、とりわけ東アジアのデータは構造変化を内包している可能性がある点を留意されたい。

3.2 サンプルの制限

分析対象は、壮年の男女であり、年齢はスペインおよびイタリアは 25-49 歳、日本と韓国は 23-49 歳の個人である (対象年齢の違いは前者の離学年齢が遅いことによる)。データは、先に述べたように便宜的に、全年観測が可能な者のみを残している (balanced panel data)。したがって、ECHP データの例では、観測初年 1994 年に 42 歳の者は最終年 2001 年に 49 歳になるサンプル構造となっている。この結果、観測数 (個体数) は、スペイン、イタリア、日本、韓国でそれぞれ 20,424 (2,553)、26,744 (3,343)、12,630 (2,526)、22,488 (2,811) となっている。出生年でいえば、ECHP データでは 1952 年から 1969 年生まれ、KLIPS データでは 1953 年から 1975 年生まれ、JLPS データでは 1962 年から 1984 年生まれの個体がサンプルとなった。分析では自営業・家族従業者は含めず、設問無回答も除かれ、そして異常値は除去するため、上で述べたサンプルサイズがさらに低減したものが次節以降の分析対象となる。本稿では欠損値への値の割当 (imputation) や重み付けなどの方法は用いない。

3.3 変数の構成

本稿では、簡単化のため、基礎的な変数のみを扱う。分析に際しては、就業行動と労働時間に着目する。全体を通して主に学歴差に着目し、4.3節では他の共変量も扱っていく。以下に変数構成を簡潔に記す。

結果変数として、2つの質的変数を扱う。ひとつは、就業しているかどうか、いまひとつは労働力かどうか、である。量的変数としては、残業時間を含む週あたり労働時間を用いる。ただし、日本と韓国では主たる職での週あたり労働時間であるのに対し、スペインとイタリアでは副業も含めた週あたり労働時間である点に留意されたい。

共変量にかんして記す。学歴については、International Standard Classification of Education (ISCED) に基づいた3類型を用いる。値の0-2は中等教育の第2段階未満相当(≈中学校以下)、3は中等教育の第2段階相当(≈高校)、4-7は中等教育後の教育以上相当(≈専門・大学以上)を意味する。学歴変数の値は、観察期間中に变化しうるが、本稿ではStephens (2002: 533-534) に倣い、固定値の変数としてある。具体的には、スペイン、イタリア、韓国の場合は、個人ファイルデータで最も新しい調査年に観察された、欠損値ではない学歴の値を個人の学歴の固定値としている。日本のみ、初年度調査のみで学歴が訊かれているため、初年度調査の学歴の値を時間不変の学歴変数の値としてある。ただし、スペインとイタリアでは卒業した最高学歴の変数であるのに対し、日本と韓国では最終通学歴(卒業・中退の区別なし)の変数となっている点を留意されたい。

4.3節の労働時間の分散にかんするモデルでのみ、年齢、子どもの有無、婚姻状況、調査時点、居住地域の変数も共変量として扱う。子どもの有無については、ECHPおよびKLIPSデータには個人ファイルには同変数がないため、世帯ファイルからそれに準ずる情報を結合した。その結果、ECHPでは世帯内の16歳未満の子どもの有無、KLIPSでは世帯内の高校生以下の子どもの有無⁷⁾、JLPSでは同居する12歳以下の自身の子の有無の変数となっている。婚姻状況は、いずれの国も、現在結婚しているかどうかの変数である。居住地域は、スペインは17の自治州、イタリアは11の地域ブロック、日本は10の地域ブロック、韓国は16の第一級行政区画(済州島を除く)となっている。

4 縦断データでみる女性の労働市場行動の多様性

本節では、女性の就業状態と労働時間の遷移割合に着目し、次に、女性の労働時間の分散に着目した基礎的な分析を行う。本節を通して、個人の行動の状態移動および状態依存の両方の側面の重要性が強調される。

本分析に用いるデータの全サンプルの基礎的な分布についてみていく(標本平均の一覧についてはAppendix C1の表5を参照されたい)。女性の就業率はスペイン、イタリア、日本、韓国それぞれで約41%、47%、72%、42%となっている。男性の就業率はいずれの国も

女性のそれより高くなっている（いずれの国も男性は 84%を超える値を示す）．女性の労働参加率と就業率の%ポイント差（i.e. 失業割合）をみると、スペイン、イタリア、日本、韓国それぞれで 15%、12%、5%、3%ポイントとなっている（韓国以外は、同数値が、男性よりも女性の方が大きい）．女性の週当たり平均労働時間をみると、スペイン、イタリア、日本、韓国それぞれで約 35 時間、34 時間、33 時間、49 時間となっている．以下では、各期の状態移動の傾向に着目する．

4.1 労働力状態の遷移

本項では、女性の就業／労働力状態の遷移を観察していく．図 4 は、男女別・学歴別に時間を通じた就業／労働力状態の移動割合を示した図である．図では、 t 期から $t+1$ 期にかけて同じ状態にとどまる人の割合が示されている．具体的には、非就業、非労働力、就業、労働力状態に継続的にとどまる割合が男女別・学歴別に示されている．

例えば、図 4・パネル A の左の図から、スペイン女性全体（学歴計）でみると、 t 期に非就業の女性の約 90%が $t+1$ 期も非就業にとどまることが示されている．非就業にとどまる女性の割合の各国の値をみると、イタリア、スペイン、韓国、日本の順に高い（それぞれ 93.13%、89.43%、85.04%、79.34%）．本分析データの観察期間内においては、日本女性が最も非就業から就業に状態移動していることを意味する．学歴差をみると、スペインおよびイタリア女性の場合は、学歴水準が低いほど、非就業にとどまる割合が高いことも併せて確認できる．

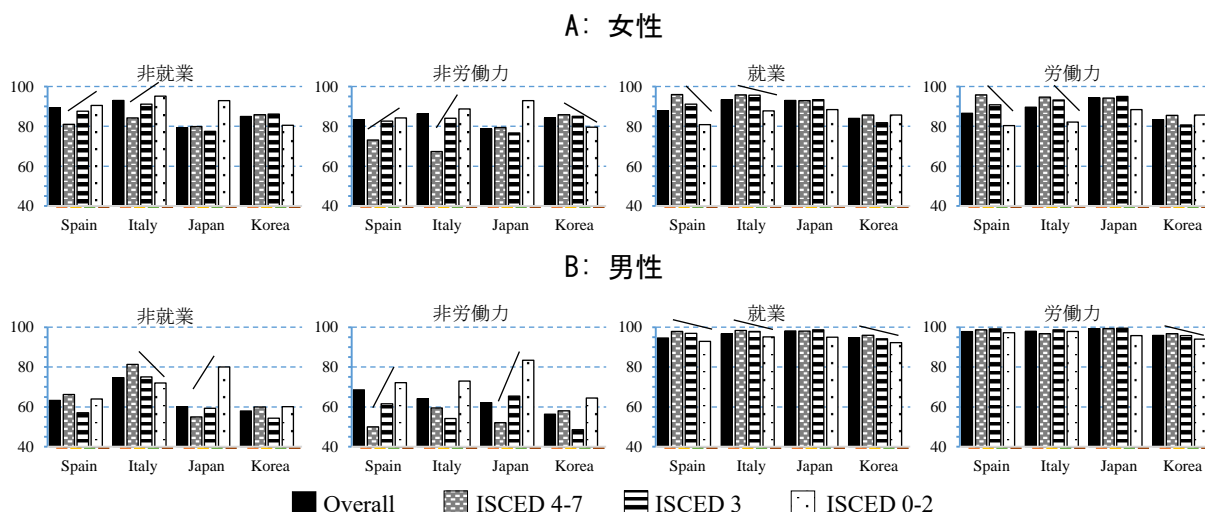


図 4 男女別・学歴別の 2 期連続で同じ労働力状態にとどまる割合

出所：ECHP-ESP 1994-2001, ECHP-ITA 1994-2001, JLPS 2007-2011, KLIPS 1998-2005 を用いて算出

注： t 期から $t+1$ 期にかけて状態不変の人の割合が示されている．学歴との関係が線形傾向の場合に斜線を付している．International Standard Classification of Education (ISCED) 0-2~中学校以下、3~高校、4-7~専門・大学以上

次に、図4・パネルAの左から3番目の図より、就業状態が継続する女性全体（学歴計）の割合に着目すると、イタリア、日本、スペイン、韓国の順に高い（それぞれ93.39%、93.06%、87.85%、84.03%）。本データの観察期間内においては、韓国女性がもっとも退職・失職をしやすいことを意味する。学歴差をみると、スペインおよびイタリア女性の場合は、学歴水準が低いほど就業状態にとどまる割合が低いことも併せて確認できる。先の結果と併せてみると、南欧社会では、学歴水準が高い女性の方が失職・退職する割合が相対的に低く、就業継続する割合が高いということである。

続いて、労働力状態の継続割合をみていく。先にみた就業状態の数値では、就業したくてもできていない女性を捉えられないが、ここでは求職中の者（失業者）も労働力に含まれるため、雇用者側の選好の影響が比較的弱まった統計であると考えられる。図4・パネルAの左から2番目の図より、非労働力状態が継続する女性全体（学歴計）の割合に着目すると、イタリア、韓国、スペイン、日本の順に高く（それぞれ86.31%、84.26%、83.47%、78.75%）、図4・パネルAの一番右の図から労働力状態が継続する割合をみると、日本、イタリア、スペイン、韓国の順に高い（それぞれ94.39%、89.63%、86.63%、83.48%）。学歴差に着目すると、非労働力状態継続、労働力状態継続ともにスペインとイタリアでより明瞭な学歴差が示される（学歴水準が高い方が前者になりやすく後者になりやすい）。また、イタリアはISCED4-7（高等教育学歴）女性の非就業状態にとどまる割合と非労働力状態にとどまる割合との差が他国より顕著に大きい点でも特徴的である（i.e. 失業女性が多い）。

男性の結果について詳述はしないが、全ての国で、非就業や非労働力にとどまる割合は女性のそれよりかなり低く、就業や労働力状態にとどまる割合は女性のそれより高い。加えて、男性の場合、各国間の差異が少ない。総じて、男性稼ぎ主色が濃い社会は、男性の就業行動で共通する面が多いといえる。以下では、女性に限定して労働時間の移動および分散に着目していく。

4.2 労働時間の遷移

本項では、女性の週当たり労働時間の遷移に着目した分析を行う。用いる労働時間区分はBooth and van Ours（2013）の5類型を、単純化のため、大括りに3類型化したものである。区分は0時間、1-32時間、33時間以上の3区分となっており、実質的には、それぞれ無業、パート就業、フルタイム就業相当を想定している。

まず、データ全体の基礎的な標本平均を確認していく（非図示）。まず、各国共通で比較的に割合が高い週労働時間0時間の女性は、スペイン、イタリア、日本、韓国でそれぞれ約60%、54%、31%、65%となっている。同じくいずれの国でも割合が比較的に高い、週労働時間33時間以上の女性は、それぞれの国で約29%、32%、38%、33%となっている。それぞれ残りの約12%、14%、31%、3%が週労働時間1-32時間である。ただし、週33時間以上の区分をより細かくし全体の傾向をみた場合（非図示）、スペインとイタリアは33-40時

間への分布，日韓は41時間以上への分布が比較的が多い点を留意されたい。

学歴別にみると（非図示），ISCED4-7（高等教育学歴）の女性は全体（学歴計）よりも0時間への分布が少ない．同教育水準の女性は，各国共通で週労働時間33時間以上の割合が全体（学歴計）よりも高く，スペイン，イタリア，日本，韓国でそれぞれ約54%，35%，41%，41%となっており，週労働時間1-32時間の割合は，それぞれで約20%，31%，29%，3%となっている．

続いて，静態的な傾向ではなく，労働時間区分間の出入りをみていく．図5は，t期の労働時間ごとのt+1期の労働時間を示し，例えば一番左の図ではt期に0労働時間の女性がt+1期にどの労働時間に移動したかが，学歴別に棒グラフで示されている．その他の労働時間の傾向もみると4カ国とも全体（学歴計）で，t期に0時間，1-32時間，33時間以上いずれの場合も2期連続で同じ状態にとどまる傾向，なかでも0時間と33時間以上の状態依存傾向が強いことがわかる．以下では労働時間の遷移の特徴的な傾向についてのみ記述する．

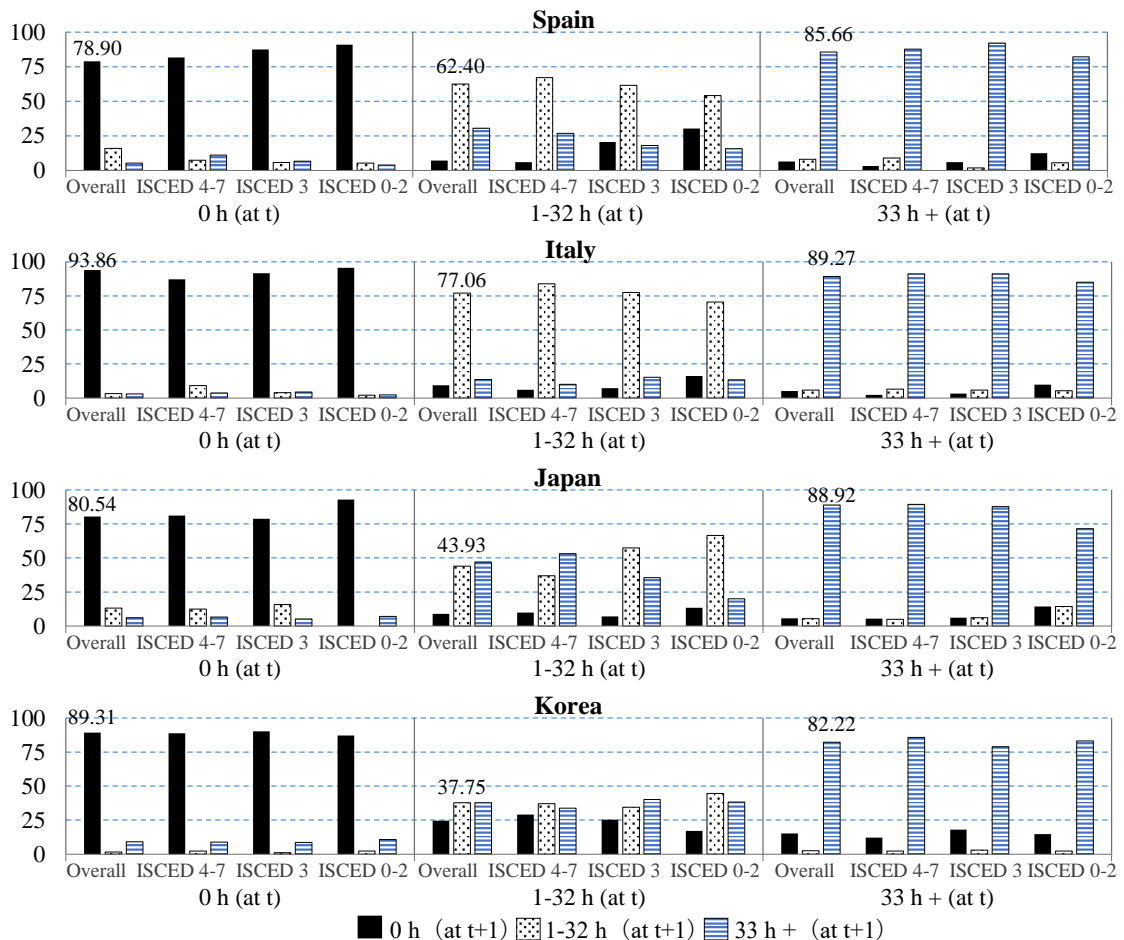


図5 t期の労働時間別・学歴別のt+1期にかけての労働時間の移動割合（女性）

出所：ECHP-ESP 1994-2001，ECHP-ITA 1994-2001，JLPS 2007-2011，KLIPS 1998-2005 を用いて算出
 注：t期からt+1期にかけての各労働時間への移動割合の棒グラフ．図中棒グラフ上の数値は，全体（学歴計）の2期間を通して同じ労働時間にとどまる人の割合．International Standard Classification of Education (ISCED) 0-2~中学校以下，3~高校，4-7~専門・大学以上

0 時間に 2 期連続でとどまる女性について学歴差をみていくと、韓国以外では、学歴水準が低いほど 0 時間にとどまりやすい傾向が窺える（スペインとイタリアでは明瞭に、日本では概ねその様である）。0 時間からの移動割合をみると、全体（学歴計）の傾向として、スペインと日本で、1-32 時間に移動する割合が比較的に高い（それぞれ 15.82%、13.29%）。イタリアは、0 時間にとどまる割合が最も高く（93.86%）、他の時間への移動が少ない。韓国については、イタリアに次いで 0 時間にとどまる割合が高いが（89.31%）、他方で、33 時間以上に移動する割合が 9.08%と 4 カ国の中で最大となっている点で特異である。

1-32 時間に 2 期連続でとどまる傾向については、スペインとイタリアで、学歴水準が高いほど同時間にとどまりやすい一方で、日本では学歴水準が低いほど同時間にとどまりやすいことがわかる。1-32 時間が日本では文字通りのパート労働を表すのに対し、労働時間が比較的短い欧州ではフルタイム就業も混在していることが関連していると考えられる。

全体（学歴計）の 1-32 時間から 33 時間以上への移動割合をみると、スペイン、イタリア、日本、韓国で、それぞれ 30.54%、13.56%、47.00%、37.76%となっており、イタリア以外では 33 時間以上に移動する割合がある程度存在することがわかる。1-32 時間から 33 時間以上への移動の学歴差をみると、学歴水準が高いほど同割合が高い傾向がスペインおよび日本で観察され（イタリア、韓国では概ね逆の傾向）、とりわけ日本でその傾向が顕著である。

上記結果で日本についての明瞭な特徴は、学歴水準が高いほど 1-32 時間にとどまりにくく、33 時間以上に移動しやすい点だといえよう。日本では、高学歴女性の労働時間の面に限り、1-32 時間と 33 時間以上は連続的なものである可能性が示唆され、例えば、高学歴のパート就業女性が（雇用形態などを変えずに）より基幹的に同一職場で長時間就業するようになる場合などが考えられる。

33 時間以上に 2 期連続でとどまる全体（学歴計）の傾向は、いずれの国も 8 割を超える高い数値となっている。学歴差をみると、日本では学歴水準が高いほど同時間にとどまりやすいことがわかる。全体（学歴計）の 33 時間以上から 0 時間への移動割合をみると、各国 6%前後の中、韓国は 15.2%と 4 カ国中で最大となっている。0 時間への移動の学歴差をみると、韓国以外では、ISCED 0-2（初等教育卒）の同割合が比較的に高いことが窺える。

以上から、まずもって、いずれの国も労働時間の状態依存傾向が強いことが確認された。また、どの時間の移動割合も、イタリアおよびスペイン並びに日本では比較的明瞭な学歴差が示され、各国共通に低学歴者の劣位が示唆された。労働力状態の遷移では学歴差が示されなかった日本ではあるが、働く女性の労働時間の遷移については学歴差が観察された。

各国で低学歴者の相対的劣位——とりわけ労働市場からの退出可能性という意味での——が観察される背景としては、女性個人の人的資本や選好の差、世帯要因との関連やパートナーとのソーティングのメカニズム、産業・労働市場構造や企業側の選好などの理由が考えられる。いずれにしても、労働市場参加のセレクション傾向の強弱が見え隠れする結果である。

韓国については、他の国に比べ、0 時間と 33 時間以上との間の移動割合が高く示されていた。このことは、パート就業の機会が同国では乏しく、フルタイム就業か無業かの二者択一の側面が強いという背景が関連していると考えられる（各国の短時間就労の推移は Appendix D を参照されたい）。

以上、労働時間の遷移の分析から、各国の女性の働き方を観察した。次節では、パネルデータの特質を引き続き活かしつつ、女性の労働時間の分散について、多面的な説明可能性の各国間比較を行っていく。

4.3 労働時間の分散の恒常的側面と短期変動的側面

本項では、パネルデータの特質を可能な限り活かすため、Lillard and Willis (1978) に倣ったシンプルな方法を利用し、女性の労働時間の分散の恒常的および短期変動的側面に着目していく。本計量経済学モデルの包括的な紹介は Arellano (2003: ch.3) を参照されたい。本項では、以下の変量効果モデル (a random effects linear model/variance-components model) の推定式を考える。

$$y_{it} = \beta'x_{it} + \eta_i + v_{it}, \text{ with } v_{it} = \rho v_{it-1} + \gamma_{it} \quad (i=1, 2, \dots, N; t=1, 2, \dots, T)$$

y_{it} は結果変数、 x_{it} は観察可能な共変量、 η_i は確率変数である個体固有の観察期間を通じて不変の個別効果、 v_{it} は個体内で時間可変の誤差項である。攪乱項は自己回帰系列相関モデルを想定し (AR(1) model), $e|\rho| < 1$ とする⁸⁾。ここで、 η_i と v_{it} とが相互に独立で、加えて、両者が x_{it} とも独立であると仮定すると、以下のような、結果変数の誤差項の分散が得られる。

$$\text{Var}(\eta_i + v_{it}) = \text{Var}(\eta_i) + \text{Var}(v_{it}) = \sigma_\eta^2 + \sigma_v^2 \quad (i=1, 2, \dots, N; t=1, 2, \dots, T)$$

誤差項は、観察期間を通じて一定の「個体間の恒常的な分散 (inter-person/permanent variance component)」 σ_η^2 、および個体内で観察期間を通じて確率的に変動する「個体内の短期変動的な分散 (intra-person/transitory variance component)」 σ_v^2 、へと分解が可能となり、パネルデータで得られる情報を最大限に活かすことが可能となる。

本分析で用いる変数は、簡略化のため、とてもシンプルなものに限定している。結果変数として、週あたり労働時間の対数値、予測変数として、学歴 (3 類型)、婚姻状況 (結婚しているかどうか)、子どもの有無、を用いる。ただし、労働時間を対数化しているため就業者に限定される点、および、学歴変数は観察期間を通じて不変のため個体内変動の分散に対して説明力を有さないことを留意されたい。

本分析では、分散構造の変化をみるため逐次的に変数を追加していく。(i) 調査時点変数のみ、(ii) 学歴変数の追加、(iii) 家族形成変数⁹⁾、年齢の追加、(iv) 居住地域、および年齢以外の全変数と調査時点変数との交互作用項の追加、の 4 段階のモデルである。以下、残

差構造の推定結果をみていく。

表 1 は女性の労働時間の残差構造パラメータの推定結果である。表の $\hat{\sigma}_\eta^2$ と $\hat{\sigma}_v^2$ は各々恒常的分散と短期変動的な分散を示す。推定列 (3), (8), (13), (18) の $\hat{\sigma}_\eta^2/(\hat{\sigma}_\eta^2 + \hat{\sigma}_v^2)$ は全分散に占める時間不変の個体間の恒常的異質性 (heterogeneity that remain constant over time/permanent variance component) の割合を示し, 残りの割合は個体内で確率的に変動する分散 (transitory variance component) を意味する。

表 1 逐次的モデルにおける残差構造パラメータ推定結果 (恒常要因と短期変動要因)

(i) Null model						
	Obs.	(1) $\hat{\sigma}_\eta^2$	(2) $\hat{\sigma}_v^2$	(3) $\frac{(1)}{(1)+(2)}$	(4) $\hat{\rho}$	(5) $\hat{\gamma}$
Spain	4,074	0.1553	0.0764	0.6703	0.5102	0.0565
Italy	5,721	0.0779	0.0277	0.7376	0.4506	0.0221
Japan	4,613	0.2246	0.0806	0.7359	0.3571	0.0703
Korea	3,162	0.0487	0.0755	0.3918	0.3066	0.0684
(ii) +Education model						
	Obs.	(6) $\hat{\sigma}_\eta^2$	(7) $\hat{\sigma}_v^2$	(8) $\frac{(6)}{(6)+(7)}$	(9) $\hat{\rho}$	(10) $\hat{\gamma}$
Spain	4,074	0.1462	0.0764	0.6568	0.5102	0.0565
Italy	5,721	0.0759	0.0277	0.7324	0.4506	0.0221
Japan	4,613	0.2219	0.0806	0.7336	0.3483	0.0708
Korea	3,162	0.0443	0.0756	0.3698	0.3066	0.0685
(iii) +Family formation model						
	Obs.	(11) $\hat{\sigma}_\eta^2$	(12) $\hat{\sigma}_v^2$	(13) $\frac{(11)}{(11)+(12)}$	(14) $\hat{\rho}$	(15) $\hat{\gamma}$
Spain	4,074	0.1421	0.0762	0.6509	0.5064	0.0566
Italy	5,721	0.0744	0.0276	0.7293	0.4498	0.0220
Japan	4,613	0.1776	0.0806	0.6878	0.3429	0.0711
Korea	3,162	0.0439	0.0751	0.3689	0.3037	0.0682
(iv) Full model						
	Obs.	(16) $\hat{\sigma}_\eta^2$	(17) $\hat{\sigma}_v^2$	(18) $\frac{(16)}{(16)+(17)}$	(19) $\hat{\rho}$	(20) $\hat{\gamma}$
Spain	4,074	0.1348	0.0738	0.6463	0.4967	0.0556
Italy	5,688	0.0722	0.0273	0.7258	0.4472	0.0218
Japan	4,613	0.1717	0.0798	0.6826	0.3410	0.0706
Korea	3,162	0.0400	0.0715	0.3585	0.3139	0.0645

出所：ECHP-ESP 1994-2001, ECHP-ITA 1994-2001, JLPS 2007-2011, KLIPS 1998-2005 を用いて算出

注：(1) ‘Null model’は調査時点変数を共変量に含む。‘+Education model’は、学歴と調査時点、‘+Family formation model’は、学歴、調査時点、婚姻状況、子どもの有無、年齢、および年齢二乗、‘Full model’は、学歴、調査時点、婚姻状況、子どもの有無、年齢、年齢二乗、居住地域、および(年齢を除く)それら全ての変数と調査時点との交互作用項、を共変量に含む。(2) 小数点以下の数字の関係で表内の計算の数値が一致しない場合がある。(3) $\hat{\gamma}$ (純粋な確率要因) は短期変動要因と系列相関の項の推定値を用いて、 $\hat{\gamma} = \hat{\sigma}_\eta^2 \times (1 - \hat{\rho}^2)$ で算出。(4) ‘Full model’におけるイタリアの観測数の減少は居住地域変数の欠損値による

例えば、Null model (正確には調査時点変数のみを含むモデル) では、スペイン、イタリア、日本、韓国において、それぞれ女性の労働時間の全分散の約 67.03%、73.76%、73.59%、39.18%が個体間で一定の恒常的分散であり、残りのそれぞれ約 32.97%、26.24%、26.41%、

60.82%は個体内での短期変動分散であることが確認できる¹⁰⁾。ただし、短期変動分散のうち、各個体で純粋に確率的な分散の割合を計算 ($\hat{\gamma}/(\hat{\sigma}_\eta^2 + \hat{\sigma}_v^2)$) すると、各々順に約 24.38%, 20.93%, 23.03%, 55.07%となり、残りの差分ポイントのそれぞれ約 8.59%, 5.31%, 3.38%, 5.75%は系列相関によるものであることがわかる。

観察期間を通じて女性個体間で固定的な労働時間の異質性は、イタリア、日本、スペインでは、全分散の中で大きな割合を占めることからとても貴重な情報といえよう。これらの個体間で観察期間を通じて不変の異質性——観察可能な要因も観察されざる要因も含む——は、単純な within-group 推定では捨象される点に留意されたい。

次に、Full model をみていく。同モデル——学歴、調査時点、年齢、年齢二乗、婚姻状況、子どもの有無、居住地域、および年齢を除くそれら全ての変数と調査時点との交互作用項を共変量に含む——では、労働時間の全分散に占める観察期間を通じて女性個体間で一定の恒常的異質性は、スペイン、イタリア、日本、韓国でそれぞれ約 64.6%, 72.6%, 68.3%, 35.9%となっており、いずれの国も先の Null model に比べ同割合がやや減少している。

以上の通り、恒常的な要因が全分散に占める大きさを確認できた。続いて、各共変量を逐次的に追加した場合に、労働時間の恒常的分散と短期変動的分散とがどう変動するかをみていく。

表 2 は、表 1 に示されている値を用いて、逐次的なモデルで労働時間の各種分散がどの程度変動するかを計算したものである。図 6 は、表 2 の数値を棒グラフにしたものである。

表 2 逐次的モデルにおける恒常分散と短期変動分散の変化割合 (Ref: Null model)

個体間の恒常的分散 $\hat{\sigma}_\eta^2$ の変化 (%)			
	(ii) +Education model $\left(1 - \frac{(6)}{(1)}\right) \times 100$	(iii) +Family formation model $\left(1 - \frac{(11)}{(1)}\right) \times 100$	(iv) Full model $\left(1 - \frac{(16)}{(1)}\right) \times 100$
Spain	5.881	8.543	13.229
Italy	2.633	4.571	7.387
Japan	1.232	20.946	23.575
Korea	8.889	9.744	17.828
個体内の短期変動分散 $\hat{\sigma}_v^2$ の変化 (%)			
	(ii) +Education model $\left(1 - \frac{(7)}{(2)}\right) \times 100$	(iii) +Family formation model $\left(1 - \frac{(12)}{(2)}\right) \times 100$	(iv) Full model $\left(1 - \frac{(17)}{(2)}\right) \times 100$
Spain	—	0.284	3.448
Italy	—	0.421	1.663
Japan	—	-0.026	0.947
Korea	—	0.510	5.291

出所：ECHP-ESP 1994-2001, ECHP-ITA 1994-2001, JLPS 2007-2011, KLIPS 1998-2005 を用いて算出
注：(1) 学歴の変数は調査観察期間を通じて変化しないため個体内の分散を説明しない(説明力 ≈ 0)。(2) 小数点以下の数字の関係で表内の計算の数値が一致しない場合がある。(3) ‘+Education model’は、学歴と調査時点、‘+Family formation model’は、学歴、調査時点、婚姻状況、子どもの有無、年齢、および年齢二乗、‘Full model’は、学歴、調査時点、婚姻状況、子どもの有無、年齢、年齢二乗、居住地域、および(年齢を除く)それら全ての変数と調査時点との交互作用項、を共変量に含む

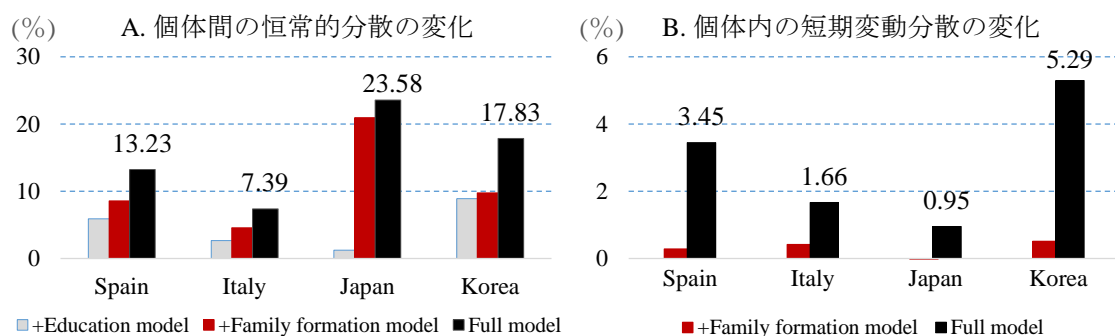


図6 逐次的モデル（表2）に基づく女性の労働時間の分散の変化割合

注：図中棒グラフ上の数値は，‘Full model’の場合の説明力の数値

労働時間の恒常的異質性

まず，図6・パネルAの+Education modelより，個体間の恒常的な異質性に対して学歴がどの程度説明力を持つかをみていく。スペイン，イタリア，日本，韓国それぞれで，労働時間の恒久的異質性の約5.9%，2.6%，1.2%，8.9%が観察可能な時間不変の学歴変数で説明されることが示されている。とりわけスペインと韓国で比較的に高い説明力を有している。両国では学歴間で女性の労働時間がばらつくということである。

次に，家族形成についての変数を追加した+Family formation modelモデルでは，スペイン，イタリア，日本，韓国それぞれの説明力は，8.54%，4.57%，20.95%，9.74%となっている。各国とも+Education modelに比べ説明力が向上しているが，とりわけ日本での説明力の高さが顕著である。すなわち日本では，就業女性間で，結婚している人と結婚していない人との間や子を持つ人と子を持たない人との間の労働時間の異質性が大きいのである。イタリアとスペインの同数値が比較的に低く，働く女性間の労働時間の恒久的異質性が家族形成の有無によって説明される度合いが低いことが示される。

居住地変数などを追加した最終的なFull modelが有する個体間で恒常的な分散への説明力は，スペイン，イタリア，日本，韓国それぞれで約13.23%，7.39%，23.58%，17.8%となっており，本分析に用いた観測可能な極めてシンプルな変数で，各国の女性の労働時間についての個体間の異質性のある程度の部分が説明可能だという結果が得られた。

以上，女性の労働時間の恒久的異質性をみてきた。全分散に占める同異質性はスペイン，イタリア，日本，韓国それぞれで約67.03%，73.76%，73.59%，39.18%，と各国大きな値を示していた。標本平均でみると日本女性の就業率が比較的に高いデータであることから（Appendix C1；表5），他国に比して日本女性間の同異質性が低いことが想定され（本稿注5を参照されたい），実際に日本のそれについては学歴による説明可能性は低かった。しかし，日本女性の労働時間においては，まずもって上の数値の通り同異質性は少なくはなく，また，同異質性の家族形成による説明可能性の高さが特徴的であった。日本女性は，家族形成の有無による労働時間の個体間の差が大きいということである（the intensive margin）。

他方、既存研究の指摘や 4.1 節図 4 の労働力状態の移動図からも示唆されるように、スペイン、イタリアの女性は就業選択と学歴とに正のセレクション傾向がみられる点で特徴的であるが、それでもなお就業女性間で労働時間の学歴間の差が大きい傾向が観察された。また、両国（および韓国）は家族形成の有無による女性の労働時間の異質性の説明可能性が低かったことから、家族形成が労働時間選択とではなく、就業選択の部分と関連していると推察される（the extensive margin）。

最終モデルから示唆されたのは、+Family formation model から Full model にかけて比較的に大きく説明力が高まっている、スペインと韓国は女性の労働時間についての地域間のばらつきが大きい国だという点である¹¹⁾。

労働時間の短期変動分散

最後に、図 6・パネル B より、女性個体内の労働時間の分散の短期変動面をみていく。本図は、労働時間との共変動による説明力——労働時間の変動と各要因の変動との関係——を捉えた指標であり、先の個体間の恒常的な異質性とは異なる面をみている。パネル B の棒グラフの水準はパネル A よりも低いが、一般に、個体内短期変動分散の説明可能性は個体間の恒常的な差異のそれよりも低く、上のような水準値でとりわけ低いわけではない点をあげておきたい。

図をみると、学歴は観察期間を通じて変化しないため、労働時間の短期変動分散への説明力を有さない（+Education model の説明力 ≈ 0 ）。+Family formation model をみていくと、わずかではあるが、スペイン、イタリア、韓国で労働時間の短期変動的側面を説明する（それぞれ説明力は 0.28%、0.42%、0.51%）。より具体的には、同一女性の家族形成行動と労働時間とがある程度共変動するようである。

そして、Full model の説明力は、スペイン、イタリア、日本、韓国それぞれで約 3.4%、1.7%、0.9%、5.3%となっている。韓国とスペインでは、Full model において説明力が比較的大きく向上しているが、同モデルで追加される地域間移動との共変動で労働時間への説明力が大きくなっているといえる。

以上、本項では観察可能な変数を用いて、女性の労働時間の恒常的な異質性と短期変動的分散の変動を分析した。分析に用いた観測可能な極めてシンプルな変数だけでも、女性の労働時間について多面的にある程度説明可能だという結果が得られた。

分析の結果示されていたのは、日本については、家族形成と女性の労働時間の恒常的な異質性との関連が他国に比して大きいこと、スペインと韓国については同異質性と地域および学歴との関連が比較的大きいことであった。

短期変動的分散をみても、スペインと韓国は居住地域を追加したモデルで説明力が向上していたことから、地域の差が女性の労働時間と比較的強く連関する国々だといえる。地域間移動と女性の労働時間との共変動を想定した場合、スペインと韓国では地域移動に際し

働く女性が労働時間を調整する様子が窺え、日本やイタリアでは地域移動しても変動しない就労（パート就労など）を地域移動以前からしている可能性が推察できる。短時間就労の機会が関連していると考えられる。

日本女性については、労働力状態・就業選択と学歴との関連は比較的小さい一方で、家族形成の有無による労働時間の恒常的な個体間の差が大きい点が特徴的であった。また、スペイン、イタリアは高学歴女性ほど就業選択・就業継続をしやすい傾向にあったが、就業女性の労働時間をみてもなお学歴差が大きく存在していた。しかし、家族形成と労働時間の個体間の恒常的な差異との連関は小さかったことから、南欧では学歴や家族形成が労働時間選択よりも就業選択の面でより強く作用していることが示唆された。

5 小括と今後の探求可能性

本稿では、縦断的な視点により、男性稼ぎ主型社会経済である南欧諸国と東アジア諸国のデータを用いて、女性の行動の類似点と相違点を観察した。縦断マクロ統計から、同諸国において女性の労働市場参加と出生行動とが負の共変動を示すことが確認され、両者の両立の困難が示唆された。

縦断的マイクロデータを用いた分析では、スペイン、イタリア、日本、韓国の女性の就業行動および労働時間に着目した。データ観察期間内での状態遷移をみた分析では、南欧諸国で就業状態の移動が少ないことが示され、また、状態遷移割合の学歴差が明瞭に示されていた。労働時間の遷移をみた分析でも東アジア諸国で比較的活発な動きがあった。データ観察期間の時代効果に留意は必要だが、南欧社会における状態依存が強い傾向は、厳格な雇用保護などの制度的背景が関連していると考えられる（各国の雇用保護については Appendix D を参照されたい）。とはいえ、その様な制度は、元来の社会規範に基づいて作られている可能性の考慮も必要となる（Alesina et al. 2015）。

韓国についてはフルタイム労働と無業との間の出入りが比較的に多く、この点はパート労働の機会が限定されているという背景によるものだと考えられる。総じて、マクロな静態的指標から示唆される点をマイクロ動的な側面から検証できたという点で、本研究は基礎資料提供の役割を果たしたといえ、したがって概観的作業にも意義が認められよう。

また、パネルデータの特性を活かし、労働時間についての個体間の恒常的分散および個体内の短期変動的な分散を観察したところ、日本においては、結婚や出産にかんする変数による就業女性間の労働時間の恒常的異質性の説明力が高く、家族形成と就業女性の労働時間との関連が大きいことが確認された。南欧諸国では、同異質性に対する家族形成の説明力は低く、労働市場の硬直性や短時間就業機会構造の背景が関連し、家族形成は労働時間選択ではなくむしろ就業選択の場面で影響を有すると推察される。

本研究を通して、各国間および各国内における社会経済の異同を捉える視点のみならず、

女性個体間の恒常的分散および個体内の短期変動的な分散の両者を可視化し、それぞれの側面を探究する可能性を示した点は、本研究の意義の一つであるといえる。特筆すべき点として、各国のパネルデータでは、女性個体間での恒常的な異質性の存在の大きさが示されていた。個体間の異質性は無視できない存在で、それ自体の説明可能性の追求の意義は極めて大きいであろう。異質な個体・集団が生成されソーティングされるメカニズム、動学的に意思決定をする多様な個体・集団を規定する社会規範や制度についての探求がなされること、社会の多様性を理解するため、より必要なことではなかろうか¹²⁾。

本研究で扱った4カ国は、マクロ統計では類似したパフォーマンスを示すものの、女性の労働参加からみる内実は多様であった。しかし、人口学的変化と社会保障の維持の困難という同様の課題を抱える国々であり、多角的な視点からのさらなる研究が必要となろう（各国の高齢化と社会保障費については Appendix E を参照されたい）。

なかんずく、類似点が多い南欧諸国内で、同じ EU 指令下でありながら変容に分岐が生じている点は興味深い。マクロ統計でみて、多方面で近年とみに変化しているのがスペインであるが——男性稼ぎ主型社会であるにもかかわらず女性の労働参加の拡大および若い生年世代では出生率の回復も観察される——、とりわけ同国のさらなる研究は他国への含意を得るという意味でも期待される^{13) 14)}。

最後に、各国間比較をするに際して慎重となるべき点を挙げておきたい。第一に、各国間比較をすると制度やパフォーマンスなど諸々の差異が過大に示される可能性がある点である（Juhn and McCue 2017: 200）。例えば、家族政策のジェンダー格差への効果をワンショット・マクロ統計に基づき国際比較（推論）する場合、両変数に影響する社会規範を介して、過大に政策効果が示されると考えられる。関連して、各国内の政策変化についてのマイクロデータに基づく評価では、多くの場合、より弱い政策効果の結果が得られている。社会規範は糊着的で、外生的な刺激があってもすぐには変化しない側面に気を配ることが必要となる——古くはオグバーンの文化遅滞論 *cultural lag theory* として知られるように。

第二に、各国異なるようで実態としての機能代替が存在する点がある点である。例えば、欧州の解雇規制とアメリカにおける差別禁止法例（Estlund 2013）、戦後欧州諸国の福祉諸政策と日本における企業福祉・公共事業（Estévez-Abe 2008）、および北欧諸国の公的セクター雇用による女性就労の拡大とアングロサクソン諸国における男女の混淆を伴ったサービス経済部門での女性就労の拡大（深沢 2000: 60-61）、のそれぞれが部分的に機能的等価物となっていたとの指摘がある。そのため、それらを見過ごさないためにも各社会経済をよく理解した、各国内の実態に迫った緊密な研究が必要となる。

そして、社会経済の動態を捉えるに際しては、プラグマティズム的に、変化の側面のみならず本質的な不変の側面——そしてそれがなぜ生き残るかも含め——も同時に捉える努力をしない場合、浅い研究を再生産するだけの可能性があるのではなかろうか。その意味で、今後の女性の労働研究は、今日の労働モデルの原型を詳細に描く経済史研究（谷本 2003 年

ど)、政策の実態的側面や個人の側からの意味付けを探求する社会政策研究、および多様なデータに基づく多様な集団の意思決定分析に長けた労働経済学研究などに期待が持たれる。

[注]

- 1) 90年代後半における無期雇用男性の勤続年数の構成をみても、欧州の中で、スペインとイタリアは「10年以上」および「3-9年」の割合が高い (Toharia 2005: 81, Gráfico II.8).
- 2) 女性の労働市場行動と関連する各国の家族政策の財政支出については、本稿 Appendix A の図 8, 南欧諸国の子育て支援策の展開については、西岡 (2003) を参照されたい。同様に関連する各国の税制面について補足的に記すと (鎌倉 2009), 本稿利用データ期間のイタリア, 日本, 韓国は個人単位課税, スペインは個人か世帯 (合算非分割) 単位課税かが選択可能となっている。イタリアは税控除, 日韓は所得控除の形で配偶者への控除・減免が存在する。社会保障面については、南欧諸国および日本は年金給付費に占める (受給者が多くの場合女性である) 遺族年金の大きさが国際的に顕著である (大沢 2013: 365)。
- 3) 例えば, Alesina et al. (2013) や Giuliano (2015) では, 鋤鋤農業社会が焼畑農業社会よりも性別分業の度合いが強く, そしてその前近代の土地の農業のあり方が, 今日の性別分業および文化的信念・社会規範と関連することを実証的に示している。分析には, 前近代の民俗学的な地理データと今日の国レベルの統計データとを結合したものを用いている (分析の詳細は Alesina et al. (2013) や Giuliano (2015: 42-43) を参照されたい)。より具体的な結果としては, 鋤鋤を使う社会であった国ほど, 現在の女性の労働参加, 起業家活動, 政治参加, および合計特殊出生率が低位水準にあるという。
- 4) Goldin and Mitchell (2017) は, (移民を除く) 図 2 と同様のアメリカ女性の生年コーホート別ライフサイクル労働力率を戦前世代を含め分析し, より微細な視点から多くの知見を得ている。アメリカ女性の労働市場参加にかんして観察されるのは, 生年世代別の曲線の水準については, ①新しい生年世代になるほど若年期および高齢期の労働参加率の水準が高い傾向, 曲線の形については, ②古い世代では逆 U 字型が示されていたものの新しい世代では 30 代半ばに窪みが生じている傾向 (sagging middles), また同時に, ③新しい世代ほど高齢期の曲線の下降が緩やかになり裾野が広がる傾向 (expanding tops), である。背景の理由としては, 新しい世代の女性は従前世代の女性よりも, 高学歴化を経て若年期の就労が増加し, 出産タイミングが遅れ (晩産化), 高齢期の労働参加が増加する行動が存在するという。具体的には, 女性のライフサイクル労働力曲線の形として 30 代半ばで窪み (いわゆる M 字型曲線) が生じはじめているのは, 若年期の労働市場参加が増加したことおよび出産・育児がいつの時代も女性のキャリアを妨げることが関連する。しかし他方で, 若年期のより多くの教育投資および就労経験は, 生涯キャリアを通しての労働市場参加を増加させるとされている。
- 5) ここでいう異質性は, 例えば (Goldin and Mitchell 2017: 171), 集計データで女性のキャリアを通じた就業率が 60% という場合, 全女性がある期間の 6 割を働いている場合と, 6 割のみ

の女性がフルタイム就業し残りの4割が全く就業しない場合との両方があり、前者は同質な女性群、後者は異質な女性群という意味合いである。Goldin and Mitchell (2017: 173, Figure 3, Panel B) は、アメリカ女性の生年世代別の、25歳から34歳にかけての10年間の8割以上就業する女性および同期間の2割未満しか就業しない女性の比率の推移を示し、趨勢的に前者は増加、後者は減少しており、したがって、女性の労働参加の進展に伴い(新しい世代ほど)女性間の異質性が減少することを明らかにしている。

- 6) 同研究では、母親ペナルティだけではなく女性の結婚ペナルティ(結婚後に賃金が下落する傾向)についても分析されており、後者は新しい世代ほど減少傾向にあることが確認されている。稼働力が高くキャリア継続志向の男女のソーティングが強まっていることが背景にあるためだという。さらに、分析対象の最も新しい生年コーホート(1966-1975年生まれ)では、大卒学歴・無子の女性には結婚プレミアム(結婚後に賃金が上昇する傾向)さえ観察されている(Juhn and McCue 2017: 197, Table 3)。
- 7) KLIPSの第2回調査では、世帯における高校生以下の子どもの数を訊いていない。そのため、本変数のみ、前年(調査初年度)の同情報を第2回調査の情報として代入して用いた(2期連続で同じ値となる)。第1回調査から第2回調査にかけて子どもの数が変動した世帯の情報は不正確となるが、第2回調査まるまるの情報を失うことを回避する次善の策である。
- 8) そのため、より正確にいうと、本モデルにおける個体内で観察期間を通じて確率的に変動する短期変動的な分散(v_{it})は、全個体に共通の自己回帰系列相関の項(the first-order serial correlation component common to all individuals)および各個体で純粋に確率的な項(a purely random component $\gamma_{it}(\sim i.i.d.N(0, \sigma_v^2))$)とに分解することができる。
- 9) 分析では簡単化のため、婚姻状況と子どもの有無を家族形成変数とし、両者を含めた場合の分散の変動を示す。単純化のためではあるが、扱う4カ国は非嫡出子が少ない国々であり、結婚と出産との関係については前者が後者の必要条件、または両者は同時決定と考えられる。
- 10) 本分析では、働く女性間および働く女性内の経年的な労働時間の分散を扱っている。具体例を示すため、いま週当たり40時間程度働く女性と、週15時間程度働く女性がいると想定する。両者の労働時間は毎週変動するが、平均的にはそれぞれ40時間と15時間である。両者の固定的な差異(平均的な労働時間水準の差異)を表すのが、個体間の恒常的な分散である。長時間働く個体と短時間のみ働く個体のため両者は異質な個体といえる。他方、各個体の労働時間は毎週変動する(前者は40時間が42時間、後者は15時間が13時間など)。その同じ個体内での労働時間の変動を表すのが、個体内の短期変動の分散である。改めて留意点として強調しておきたいのは、本分析は対数労働時間を用いているため非就業者は含まず、したがって、逐次的モデルの表1および表2でみているのは、働く女性の間での労働時間の分散の説明可能性である点である(就業選択のサンプルセレクションの問題は残される)。
- 11) Full modelでは居住地域以外にも変数が追加されているが、居住地域のみを追加した場合(非図示)とで説明力は大きく変わらない。ただし、非図示の男性版の同様の図では、スペイン

とイタリアで+Family formation model から Full model にかけて説明力が大きく高まる傾向がみられ、日韓ではその傾向はみられないため、労働時間の地域間のばらつきがジェンダー非対称の可能性が窺える。

- 12) そのためには、日本ではないものねだりに等しいが、良質なマイクロデータ——初回調査回収率が5割以上の中長期の世帯追跡調査データおよび企業内の縦断的人事データ、加えて、個人間・個人-組織間のマッチング可能性および業務統計との接続可能性の担保——の利用可能性が待たれる。
- 13) スペインの変容の背景としては、無期雇用の解雇規制を緩和しつつ——例えば本稿 Appendix D の図 11 を参照——新型の無期雇用推奨策を展開し (Toharia 2005)、雇用創出を促進したのみならず、ジェンダー関連の諸政策が展開されてきたことが関連するであろう (Léon 2011)。中でもとりわけ、公的サービス給付である「保育・就学前教育」支出の拡大とともに、3-6 歳未満の子の就学率は 2009 年には EU 内でフランスに次ぐ 99.1% という高い水準となった (Secretaría General Técnica 2012: 17)。女性の自己投資による高学歴化が進む中、同政策の意義は大きい。労働政策および家族政策の組み合わせが要となることが示唆される事例である。
- 14) ただし、欧州では家族政策と女性の短時間就業とがセットで女性の労働市場参加が増加しているが、そのことが女性の昇進機会の制約にも繋がっている可能性を、企業側の視点を加味した場合に注意する必要がある (Blau and Kahn 2013, Goldin and Mitchell 2017: 179, Juhn and McCue 2017: 200)。翻って、元来より女性の短時間就業が社会にどれほど根付いているかが、欧州各国内での女性の本質的な地位変化の分岐と関連する可能性が示唆される。

[謝辞]

本研究に用いたパネルデータは、スペイン統計局 (Instituto Nacional de Estadística)、イタリア統計局 (Istituto Nazionale di Statistica)、韓国労働研究院 (Korea Labor Institute)、および東京大学社会科学研究所附属社会調査・データアーカイブ研究センター (Center for Social Research and Data Archives at Institute of Social Science, University of Tokyo) より提供頂いた。また、本稿の執筆にあたり、田中隆一、大沢真理、佐藤香、角能、各氏より有益な助言を頂いた (敬称略)。記して感謝申し上げたい。いうまでもなく、本研究における誤りの責任は全て著者に帰する。

[文献]

- Akerlof George A., Kranton Rachel E., 2000, “Economics and Identity.” *Quarterly Journal of Economics*, 115: 715-753.
- Alesina Alberto, Giuliano Paola, Nunn Nathan, 2013, “On the Origins of Gender Roles: Women and the Plough.” *Quarterly Journal of Economics*, 128: 469-530.
- Alesina Alberto, Giuliano Paola, 2014, “Family Ties.” *Handbook of Economic Growth*, 2A: 177-215.
- Alesina Alberto, Algan Yan, Cahuc Pierre, Giuliano Paola, 2015, “Family Values and the Regulation of Labor.”

- Journal of the European Economic Association*, 13(4): 599-630.
- Arellano Manuel, 2003, *Panel Data Econometrics*. Oxford University Press.
- Blau Francine D., Kahn Lawrence M., 2007, "Changes in the Labor Supply Behavior of Married Women: 1980-2000." *Journal of Labor Economics*, 25(3): 393-438.
- Blau Francine D., Kahn Lawrence M., 2013, "Female Labor Supply: Why is the US Falling Behind?" *NBER Working Paper*.
- Booth Alison L., van Ours Jan C., 2013, "Part-Time Jobs: What Women Want?" *Journal of Population Economics*, 26:263-283.
- Costa Dora L., 1998, "The Unequal Work Day: A Long-Term View." *American Economic Review*, 88(2): 330-334.
- Estlund Cynthia J., 2013, "Individual Employee Rights at Work." In Eds. Frege Carola, Kelly John, *Comparative Employment Relations in the Global Economy*. Routledge, pp. 29-48.
- Estévez-Abe Margarita, Iversen Torben, Soskice David, 2001, "Social Protection and the Formation of Skills: A Reinterpretation of the Welfare State." In Eds. Hall Peter A., Soskice David, *Varieties of Capitalism: The Institutional Foundation of Comparative Advantage*. Cambridge University Press, pp. 145-183.
- Estévez-Abe Margarita, 2008, "Rashomon: The Japanese Welfare State in a Comparative Perspective." In *Welfare and Capitalism in Postwar Japan*. Cambridge University Press, pp. 19-50.
- EPUNet Team, 2004, *ECHP User Guide*, Euro Panel Users Network.
- Feyrer James, Sacerdote Bruce, Stern Ariel D., 2008, "Will the Stork Return to Europe and Japan? Understanding Fertility within Developed Nations." *Journal of Economic Perspectives*, 22(3): 3-22.
- 深沢和子, 2000, 「女性労働と社会政策」, 木本喜美子・深沢和子編『現代日本の女性労働とジェンダー 新たな視角からの接近』ミネルヴァ書房, pp. 52-85.
- Giuliano Paola, Alesina Alberto, 2010, "The Power of Family." *Journal of Economic Growth*, 15: 93-125.
- Giuliano Paola, 2010, "Ties that Matter: Cultural Norms and Family Formation in Western Europe." In Eds. Brown Clair, Eichengreen Barry J., Reich Michael, *Labor in the Era of Globalization*. Cambridge University Press, pp. 211-238.
- Giuliano Paola, 2014, "Female Labour Force Participation: Persistence and Evolution." In Eds. Durlauf Steven N., Blume Lawrence E., *New Palgrave Dictionary of Economics*. Palgrave Macmillan.
- Giuliano Paola, 2015, "The Role of Women in Society: From Preindustrial to Modern Times." *CESifo Economic Studies*, 61(1): 33-52.
- Goldin Claudia, 2006, "The Quiet Revolution that Transformed Women's Employment, Education, and Family." *American Economic Review*, 96: 1-26.
- Goldin Claudia, Mitchell Joshua, 2017, "The New Life Cycle of Women's Employment: Disappearing Humps, Sagging Middles, Expanding Tops." *Journal of Economic Perspectives*, 31(1): 161-182.
- Hwang Jisoo, 2016, "Housewife, "Gold Miss," and Equal: The Evolution of Educated Women's Role in Asia and

- the U.S.” *Journal of Population Economics*, 29: 529-570.
- Juhn Chinhui, McCue Kristin, 2017, “Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap across Cohorts.” *Journal of Economic Perspectives*, 31(1): 183-204.
- 鎌倉治子, 2009, 「諸外国の課税単位と基礎的な人的控除——給付付き税額控除を視野に入れて」『レファレンス』2009(11): 103-130.
- Keck Wolfgang, Saraceno Chihara, 2013, “The Impact of Different Social-Policy Frameworks on Social Inequalities among Women in the European Union: The Labour-Market Participation of Mothers.” *Social Politics*, 20(3): 297-328.
- Lillard Lee A., Willis Robert, 1978, “Dynamic Aspects of Earning Mobility.” *Econometrica*, 46(5): 985-1012.
- Marta Guillén A., León Margarita eds., 2011, *The Spanish Welfare State in European Context*, Routledge.
- Nam Jaeryang, 2009, *An Overview and Prospects of the KLIPS*. Online slides.
- 西岡八郎, 2003, 「南欧諸国の低出生率と子育て支援策の展開」『人口問題研究』59(3): 43-61.
- Olivetti Claudia, Petrongolo Barbara, 2017, “The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries.” *Journal of Economic Perspectives*, 31(1): 205-230.
- Osawa Mari, 2011, *Social Security in Contemporary Japan*. Routledge.
- 大沢真理, 2013, 『生活保障のガバナンス——ジェンダーとお金の流れで読み解く』有斐閣.
- Peracchi Franco, 2002, “The European Community Household Panel: A Review.” *Empirical Economics*, 27: 63-90.
- Secretaría General Técnica: Ministerio de Educación, Cultura y Deporte, 2012, *Datos y Cifras: Curso Escolar 2012/2013*, A.G. Rupem S. Coop.
- Stephens Melvin, 2002, “Worker Displacement and the Added Worker Effect.” *Journal of Labor Economics*, 20(3): 504-537.
- 谷本雅之, 2003, 「近代日本の女性労働と『小経営』」, 氏家幹人・谷本雅之・桜井由幾・長野ひろ子編『日本近代国家の成立とジェンダー』柏書房, pp. 144-187.
- Toharia Cortés Luis ed., 2005, *El Problema de la Temporalidad en España: Un Diagnostico*. Ministerio de Trabajo y Asuntos Sociales.
- Venn Danielle, 2009, “Legislation, Collective Bargaining and Enforcement: Updating the OECD Employment Protection Indicators.” *OECD Social, Employment and Migration Working Papers*.

Appendices

Appendix A 部門別社会支出および家族関連支出細目の推移

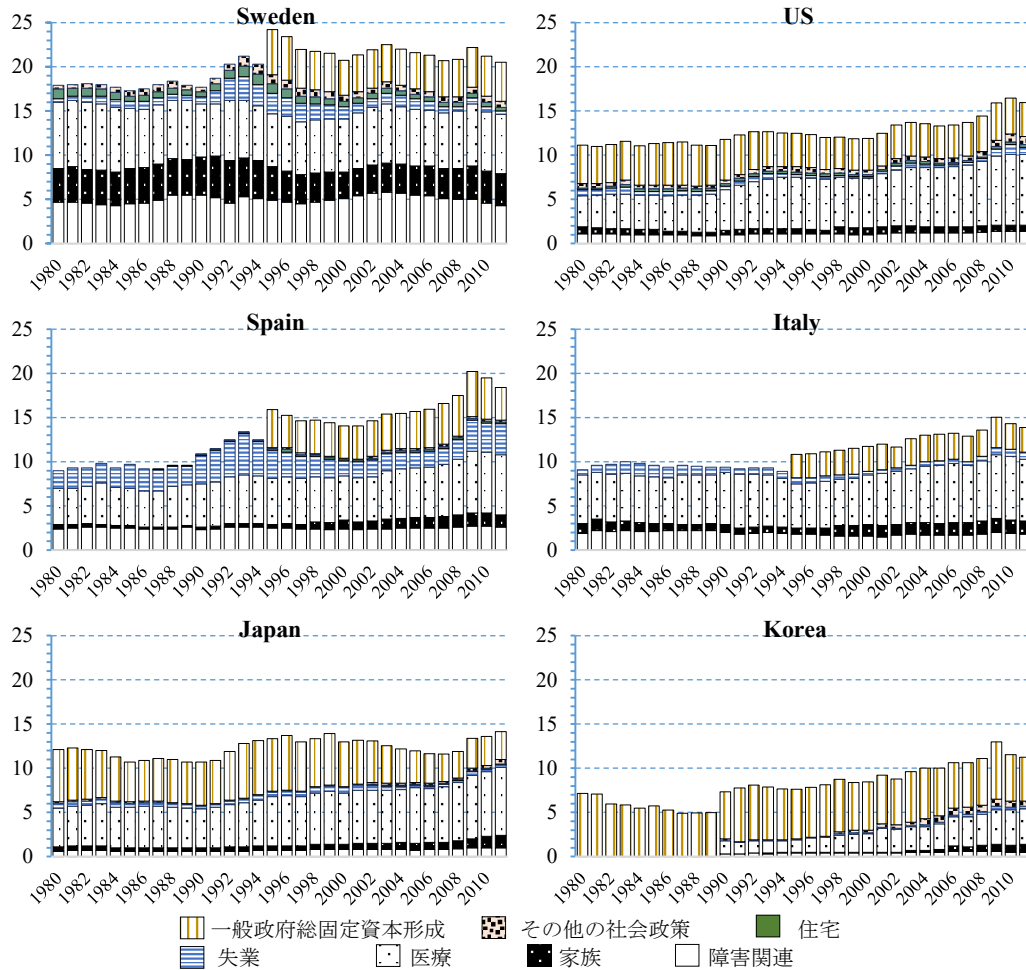


図7 対GDP比部門別社会支出の推移

出所：社会支出は OECD Social Expenditure Database, 一般政府総固定資本形成（GFCF）は National Accounts at a Glance 2015, および国民経済計算確報（2016年3月取得）

注：欧州諸国の一般政府総固定資本形成は95年以降，韓国の部門別社会支出は90年以降の値のみがアクセス可能．戦後日本で公共事業（一般政府総固定資本形成）が公的福祉の機能代替を果たしたという議論を考慮し同数値を追加

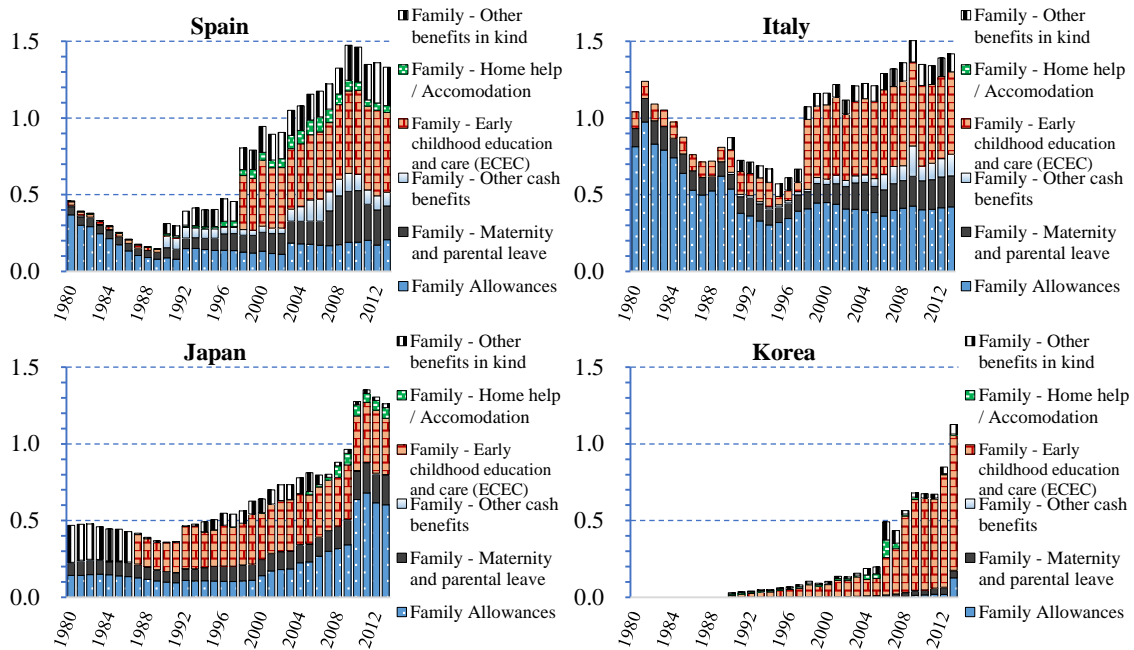


図8 対GDP比家族関連支出・細目別の推移

出所：OECD Social Expenditure Database（2016年3月取得）

Appendix B 横断時系列統計のスライド法加工による疑似・生年コーホートデータの作成

一般に、横断的統計を用いて個人のライフサイクル／キャリア——例えば、ある年の年齢階層別の労働参加率を用いて、個人の加齢に伴う労働参加の傾向——について論じる場合には、注意が必要となる。これは Age-Generation-Period/Age-Period-Cohort の識別問題と関連し、ワンショットデータで推論される年齢の効果が、加齢効果によるものなのかコーホート効果によるものなのか、それとも時代効果によるものなのかの識別が難しいということがある（例えば、Goldin and Mitchell（2017: 162）を参照されたい）。ライフサイクルについて論じる際に最低限必要となる、スライド法による横断統計の加工手順の具体例を、以下の①~③の表を用いて簡単に記す。

①スペイン女性の年齢階層別労働参加率（1984-2014年）

	25-29	30-34	35-39	40-44	45-49	50-54	55-59
1984	50.0	37.3	31.4	29.1	27.7	25.1	22.3
1989	62.8	53.0	44.6	38.5	33.0	27.8	23.5
1994	69.8	61.9	59.0	52.0	41.5	32.8	23.6
1999	75.3	68.0	63.6	59.8	50.2	38.5	27.1
2004	80.7	76.1	70.9	67.9	61.7	49.9	34.2
2009	83.9	83.1	80.0	77.4	73.1	62.8	47.9
2014	84.3	87.4	86.0	82.8	78.9	72.7	58.7

②上の非網掛部分を削除／スライドし同一集団追跡

	25-29	30-34	35-39	40-44	45-49	50-54	55-59
1984	50.0	53.0	59.0	59.8	61.7	62.8	58.7
1989	62.8	61.9	63.6	67.9	73.1	72.7	
1994	69.8	68.0	70.9	77.4	78.9		
1999	75.3	76.1	80.0	82.8			
2004	80.7	83.1	86.0				
2009	83.9	87.4					
2014	84.3						

①の表は OECD ウェブ上統計抄録より入手した各年の年齢階層別スペイン女性の労働参加率の表である。年齢階層の幅が5歳であり、かつ年度幅も5年毎であることから、斜めの網掛け部分は同じ集団を追跡していることになる。具体的には、1984年に25-29歳の女性は1989年には30-34歳になる。そのため、②の表のように、①の表の網掛け部分の上を全て削除すれば、同じ集団を追跡した表へと変更される。最後に、③の表のように、全ての行につき、年度から年齢の引き算

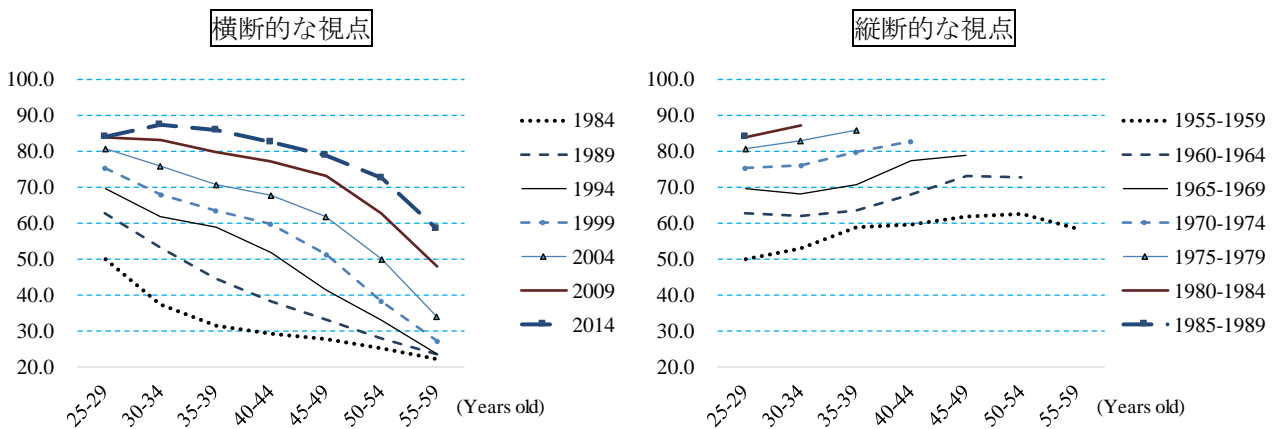
③年度から年齢を引き生年を算出

	25-29	30-34	35-39	40-44	45-49	50-54	55-59
1955-1959	50.0	53.0	59.0	59.8	61.7	62.8	58.7
1960-1964	62.8	61.9	63.6	67.9	73.1	72.7	
1965-1969	69.8	68.0	70.9	77.4	78.9		
1970-1974	75.3	76.1	80.0	82.8			
1975-1979	80.7	83.1	86.0				
1980-1984	83.9	87.4					
1985-1989	84.3						

を行っていくことで出生年を算出する. 例えば $1984-29=1955$ となる. ①の表と③の表を図にしたものが下図である. 左の図で女性のライフサイクルを推論した場合, 加齢に伴い労働参加率が減少するという誤った傾向を

読み取ってしまう. 他方で, 右の図のように同一世代の追跡を行うと, 50歳前を頂点とした右上がりの曲線が示される. スペインのように生年世代ごとの変化(コーホート効果)が大きい国のワンショット統計の扱いは注意を要することが含意される.

Goldin and Mitchell (2017) によれば, 近年の米国女性の労働市場参加においては, コーホート効果は減少し加齢・ライフサイクル効果の重要性が増している.



Appendix C1 各国調査の概要, 回収率・サンプル摩耗, サブサンプル

European Community Household Panel (ECHP) 調査は, 欧州統計局 (European Union Statistical Office: EUROSTAT) により, 欧州各国で 1994 年から 2001 年までの計 8 年間にかけて実施された世帯追跡調査である. 途中参加国も含め計 15 カ国が参加した大規模な調査である (Peracchi 2002, EPUNet Team 2004). スペインとイタリアの実査はそれぞれの国家統計局が担った. 日本の「働き方とライフスタイルに関する全国調査」(Japanese Life Course Panel Surveys: JLPS) は, 東京大学社会科学研究所が 2007 年より実施している個人追跡調査である. 韓国の Korean Labor and Income Panel Study (KLIPS) は 1998 年より韓国労働研究院が実施してる世帯追跡調査である.

JLPS の公開データは若年 (youth) と壮年 (middle-aged) のものにとわかれているが, 実質的な調査は両者とも同じ枠組みのため, 重みづけせずに両者を結合することができる. したがって, 本稿の利用データは, 2007 年時点で 20-40 歳のサンプルとなる. ECHP と KLIPS データについては, 個人データファイルに必要な世帯情報を結合して用いる. KLIPS データは継続中の調査であるが, ECHP とデータ期間をそろえるため, 1998-2005 年の 8 年間のみのデータを用いる.

なお、本稿では便宜的に全ての国のデータにつき、利用するデータ期間全てに回答しているサンプルに限定したデータを用いる。その結果、それぞれのデータの対象サンプル/全サンプル（後者に対する前者の比率）は、ECHP-ESPは2,553/10,998（23.21%）、ECHP-ITAは3,343/11,944（27.99%）、JLPSは2,526/4233（59.67%）、KLIPSは2,822/10743（26.17%）となった。回収率、サブサンプル、および基礎統計については表3、表4、および表5を参照されたい。なお、分析では対象に諸制約を設け、変数の欠損値処理も行うことから、標本規模はさらに縮減する。

表3 各調査の初年度回収率・継続回答率

	ECHP-ESP	ECHP-ITA	JLPS-Y	JLPS-M	KLIPS
調査初年度の標本規模 (回収調査票数)	7,206	7,115	3,367	1,433	5,000
調査初年度の回収率 (回収調査票数/配布数×100)	67.00%	90.70%	34.46%	40.38%	71.67%
8年目時点での調査協力継続率 (調査8年目の回収調査票数/調査初年度の回収調査票数×100)	76.14%	78.65%	66.29%	75.72%	77.74%

出所：調査初年度の回収率はPeracchi（2002:74）、Nam（2009:9）、JLPS ウェブページより。その他の値はECHP-ESP 1994-2001、ECHP-ITA 1994-2001、JLPS 2007-2011、KLIPS 1998-2005を用いて算出

注：(1) 回収率は回収調査票数/調査票配布者数で算出（転居・長期不在・住所不明・その他（死亡、病气、障害など）は分母から除かれている）。サンプリング設計の違いにより、JLPSは個人票に基づく回収率、その他は世帯票に基づく回収率が示されている。(2) 最下行では、JLPSのみ利用可能データ期間の制約から5年目時点の継続回答率が示され、KLIPSは初回調査サンプルのみを用いて（追加サンプルを含めず）8年目時点の継続回答率を計算

表4 バランスド・パネルデータにおけるサブサンプル

	ECHP-ESP		ECHP-ITA		JLPS		KLIPS	
	調査参加状況		調査参加状況		調査参加状況		調査参加状況	
	一年以上	全年	一年以上	全年	一年以上	全年	一年以上	全年
全データ期間	1994-2001 (8年)		1994-2001 (8年)		2007-2011 (5年)		1998-2012 (15年)	
男女計	22,341	7,877 (35.26%)	21,892	9,704 (44.33%)	4,800	2,797 (58.27%)	22,480	4,286 (19.07%)
男性	10,905	3,647 (33.44%)	10,740	4,692 (43.69%)	2,365	1,219 (51.54%)	11,107	1,849 (16.65%)
女性	11,436	4,230 (36.99%)	11,152	5,012 (44.94%)	2,435	1,578 (64.80%)	11,373	2,437 (21.43%)
最初の8年	1994-2001		1994-2001		2007-2011 (5年)		1998-2005	
男女計	22,341	7,877 (35.26%)	21,892	9,704 (44.33%)	4,800	2,797 (58.27%)	16,806	6,183 (36.79%)
男性	10,905	3,647 (33.44%)	10,740	4,692 (43.69%)	2,365	1,219 (51.54%)	8,321	2,747 (33.01%)
女性	11,436	4,230 (36.99%)	11,152	5,012 (44.94%)	2,435	1,578 (64.80%)	8,485	3,436 (40.49%)
内25(23)-49歳	1994-2001		1994-2001		2007-2011 (5年)		1998-2005	
男女計	10,998	2,553 (23.21%)	11,944	3,343 (27.99%)	4,233	2,526 (59.67%)	10,743	2,811 (26.17%)
男性	5,542	1,204 (21.73%)	5,967	1,620 (27.15%)	2,091	1,107 (52.94%)	5,380	1,348 (25.06%)
女性	5,456	1,349 (24.73%)	5,977	1,723 (28.83%)	2,142	1,419 (66.25%)	5,363	1,463 (27.28%)

出所：ECHP-ESP 1994-2001、ECHP-ITA 1994-2001、JLPS 2007-2011、KLIPS 1998-2005を用いて算出

注：(1) 括弧内の比率は（一度でも調査に参加した）全調査者数に占める全調査継続回答者数の割合。(2) JLPSは若年と壮年データを結合したもの。(3) 網掛け部分は、スペインとイタリアの25-49歳、および日本と韓国の23-49歳のサンプルのなかで全調査年継続回答したサンプルの情報であり、同サンプルを本稿の分析に利用

表 5 基本統計

ECHP-ESP	女性					男性				
	Obs.	Mean	S.D.	Min.	Max.	Obs.	Mean	S.D.	Min.	Max.
年齢	9,963	37.027	5.667	25	49	7,802	36.939	5.604	25	49
高等教育(ISCED4-7)	9,963	0.181	0.385	0	1	7,802	0.177	0.381	0	1
既婚	9,963	0.787	0.410	0	1	7,802	0.753	0.432	0	1
有子	9,963	0.686	0.464	0	1	7,802	0.650	0.477	0	1
就業	9,963	0.412	0.492	0	1	7,802	0.845	0.362	0	1
労働力	9,963	0.560	0.496	0	1	7,802	0.942	0.233	0	1
フルタイム就業	4,074	0.797	0.403	0	1	6,549	0.976	0.152	0	1
週労働時間(全体)	9,957	14.253	18.441	0	96	7,794	36.330	18.103	0	96
週労働時間(就業者)	4,074	34.835	10.682	2	96	6,549	43.237	9.558	1	96
週労働時間対数	4,074	3.474	0.460	0.693	4.564	6,549	3.740	0.250	0.000	4.564

ECHP-ITA	女性					男性				
	Obs.	Mean	S.D.	Min.	Max.	Obs.	Mean	S.D.	Min.	Max.
年齢	12,512	37.086	5.585	25	49	9,656	36.798	5.776	25	49
高等教育(ISCED4-7)	12,512	0.109	0.312	0	1	9,656	0.127	0.334	0	1
既婚	12,512	0.796	0.403	0	1	9,656	0.692	0.462	0	1
有子	12,512	0.664	0.472	0	1	9,656	0.589	0.492	0	1
就業	12,510	0.474	0.499	0	1	9,654	0.845	0.362	0	1
労働力	12,510	0.594	0.491	0	1	9,654	0.943	0.232	0	1
フルタイム就業	5,721	0.754	0.431	0	1	7,922	0.952	0.215	0	1
週労働時間(全体)	12,393	15.586	17.798	0	72	9,517	33.362	16.406	0	90
週労働時間(就業者)	5,721	33.762	8.514	1	72	7,922	40.079	7.356	2	90
週労働時間対数	5,721	3.477	0.319	0.000	4.277	7,922	3.671	0.220	0.693	4.500

JLPS	女性					男性				
	Obs.	Mean	S.D.	Min.	Max.	Obs.	Mean	S.D.	Min.	Max.
年齢	6,846	34.199	5.190	23	44	5,098	34.215	4.962	23	44
高等教育(ISCED4-7)	6,841	0.711	0.453	0	1	5,088	0.687	0.464	0	1
既婚	6,846	0.643	0.479	0	1	5,098	0.579	0.494	0	1
有子	6,846	0.482	0.500	0	1	5,098	0.419	0.493	0	1
就業	6,812	0.715	0.451	0	1	5,074	0.951	0.216	0	1
労働力	6,812	0.763	0.425	0	1	5,074	0.976	0.152	0	1
フルタイム就業	4,618	0.562	0.496	0	1	4,670	0.769	0.422	0	1
週労働時間(全体)	6,593	23.176	21.128	0	135	4,943	42.476	22.428	0	144
週労働時間(就業者)	4,618	33.089	17.584	1	135	4,670	44.960	20.511	1	144
週労働時間対数	4,618	3.286	0.753	0.000	4.905	4,670	3.648	0.645	0.000	4.970

KLIPS	女性					男性				
	Obs.	Mean	S.D.	Min.	Max.	Obs.	Mean	S.D.	Min.	Max.
年齢	9,458	36.415	6.044	23	49	8,008	36.241	6.080	23	49
高等教育(ISCED4-7)	9,458	0.255	0.436	0	1	8,008	0.431	0.495	0	1
既婚	9,458	0.871	0.336	0	1	8,008	0.714	0.452	0	1
有子	8,261	0.747	0.435	0	1	6,974	0.678	0.467	0	1
就業	9,458	0.421	0.494	0	1	8,008	0.847	0.360	0	1
労働力	9,458	0.451	0.498	0	1	8,008	0.898	0.302	0	1
フルタイム就業	3,162	0.331	0.471	0	1	5,725	0.811	0.391	0	1
週労働時間(全体)	8,639	17.951	25.057	0	140	6,947	44.794	23.659	0	129
週労働時間(就業者)	3,162	49.044	13.795	4	140	5,725	54.355	12.628	3	129
週労働時間対数	3,162	3.842	0.353	1.386	4.942	5,725	3.968	0.245	1.099	4.860

出所：ECHP-ESP 1994-2001, ECHP-ITA 1994-2001, JLPS 2007-2011, KLIPS 1998-2005 を用いて算出
 注：(1) 自営業・家族従業者は含まれず。(2) フルタイム就業は、スペインとイタリアでは週当たり 30 時間以上、日本と韓国は週当たり 35 時間以上就業を指す（本変数は分析には利用せず）

Appendix C2 各調査データ期間におけるマクロ経済変動

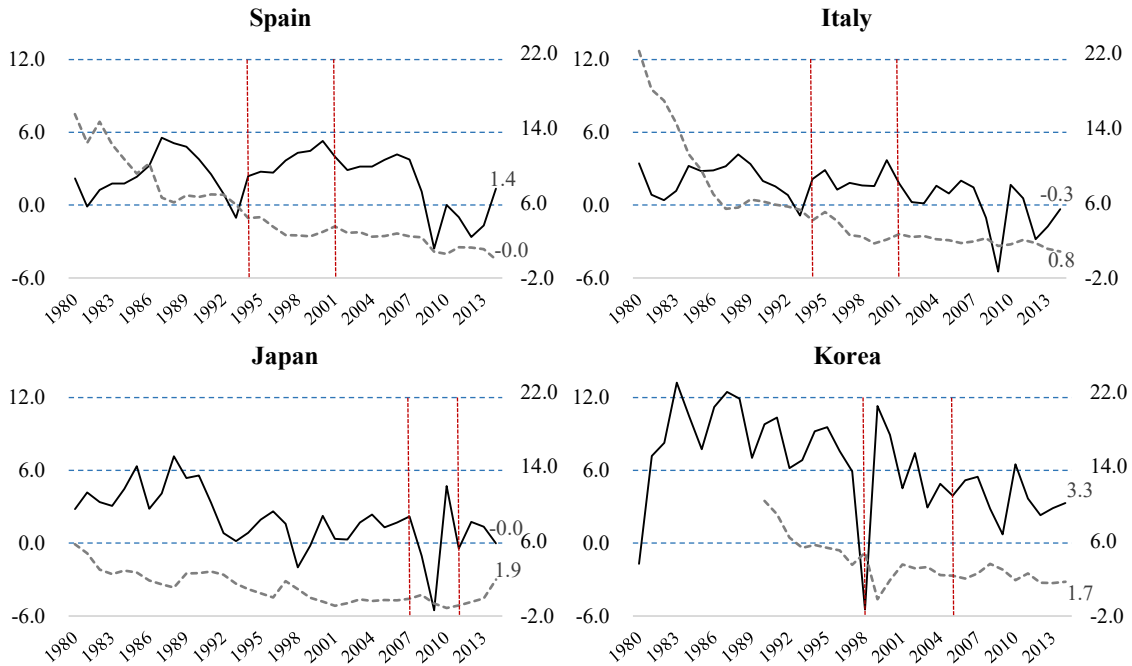
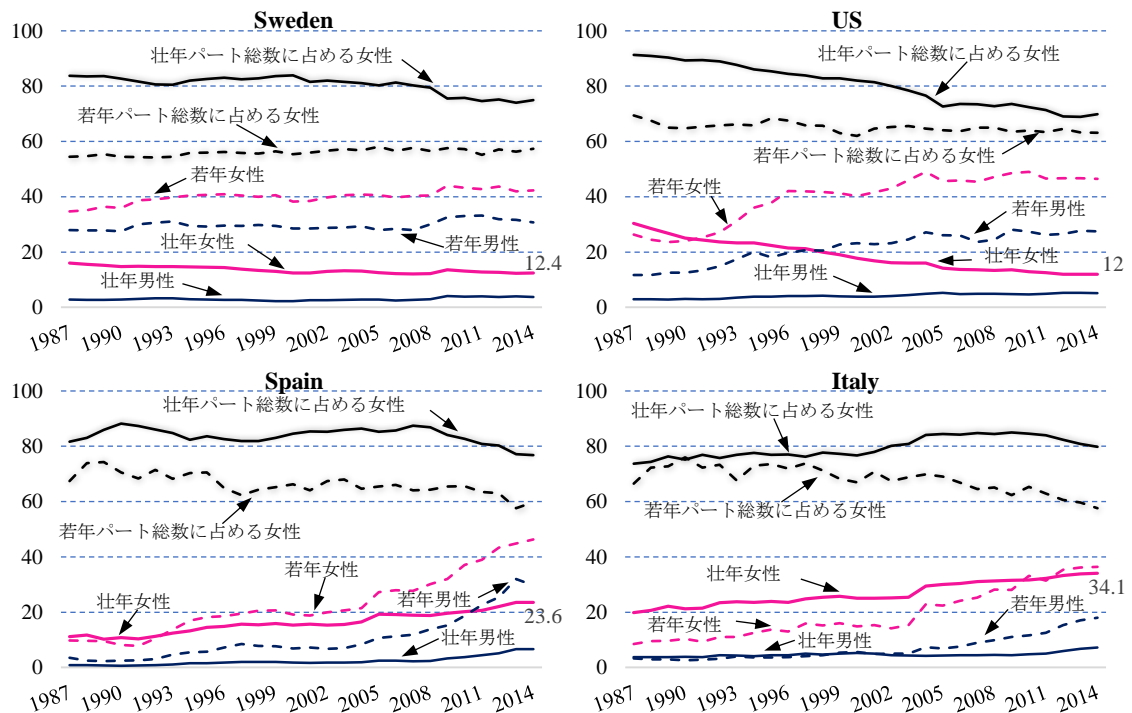


図9 経済成長率と物価の推移

出所：OECD Stat（国内総生産・消費者物価指数）（2016年3月取得）

注：(1) 直線と破線は、各々前年比のGDP成長率、消費者物価指数（食品・エネルギーを除く）を示し、それぞれ左目盛、右目盛りに対応する。2014年時点の両値の数値が図中に示されている。(2) 縦の破線は、本研究で用いるマイクロデータの観察期間

Appendix D 各国における労働市場特性（短時間就労・雇用保護）



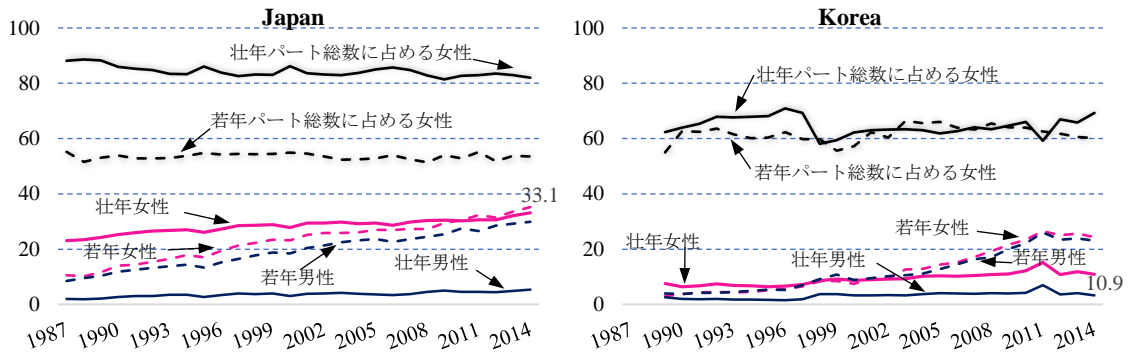


図 10 男女別・年齢層別の雇用者に占めるパートタイム労働者比率の推移

出所：OECD Employment Database（2016年3月取得）

注：(1) 若年は15-24歳，壮年は25-54歳の年齢層を指す。(2) パートタイム労働者は，主たる仕事における，週当たり30時間未満就業の労働者を指す。利用可能なOECD統計上，日本と韓国のみ自営業者を含む統計値。(3) 総数に占める女性は，男女計パートタイム労働者数に占める女性労働者の割合。(4) 壮年女性のパート比率の2014年時点の値のみ図中に示されている

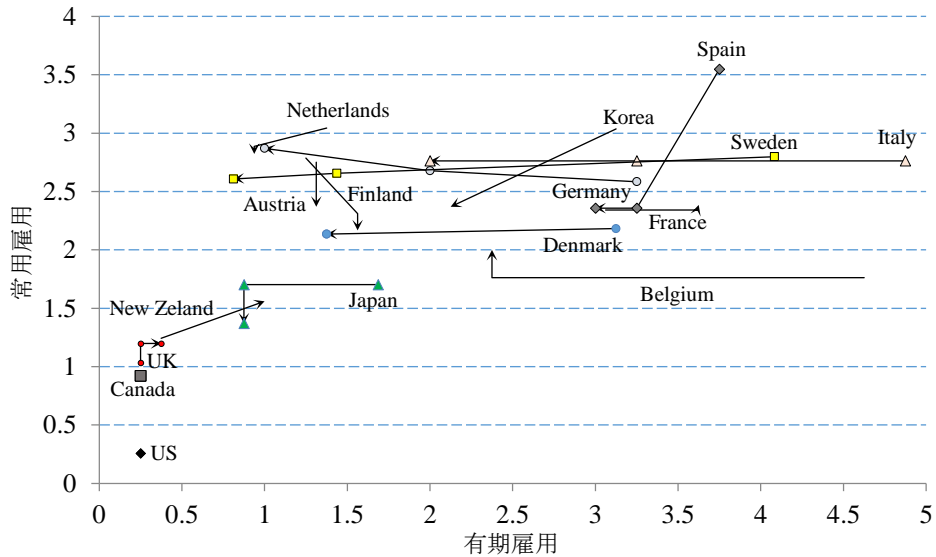


図 11 各国の雇用形態別の雇用保護の推移（1990年・2000年・2010年）

出所：OECD Employment Database（2016年3月取得）

注：常用雇用の保護は解雇手続きの困難性や補償の程度などで測られ，有期雇用の保護は契約の締結条件や更新可能性および派遣業の規制の程度などで測られ，各々0~6の値をとる数値指標（Venn 2009）

Appendix E 各国における高齢人口と社会保障費の変動

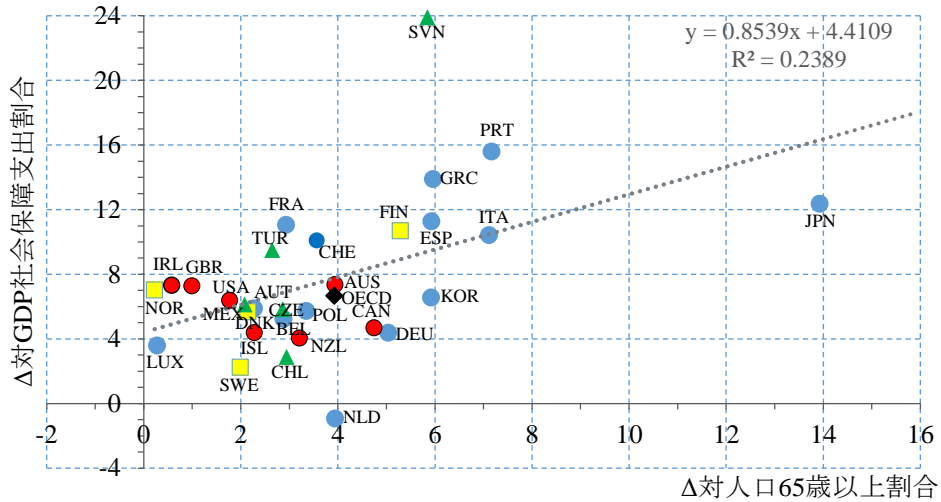


図 12 各国における高齢人口と社会保障費の 1980 年から 2010 年にかけての変化

出所：OECD Stat. (2016 年 9 月取得)

注：(1) 筆者算出の 2 期間の差分%ポイント (2010 年－1980 年) の値。(2) ただし、Mexico (MEX) は 85 年からの変化の値で、Chile (CHL)、Czech Republic (CZE)、Iceland (ISL)、Korea (KOR)、Poland (POL)、Slovenia (SVN) は 90 年からの変化の値。利用可能期間がこれらの国よりも短い国は図に含まれず。(3) 便宜的に、赤丸はアングロサクソン諸国、青丸は大陸欧州諸国＋東アジア諸国、四角は北欧諸国、三角は東欧諸国＋新興国、という類型ごとにマーカーがグループ分けされている。(4) 社会保障給付費は (public and mandatory private) 社会支出の対 GDP 比の全項目合計値

結婚・出産タイミングの日韓比較

戸田淳仁

福田隆巳

(リクルートワークス研究所)

(東京大学大学院経済学研究科)

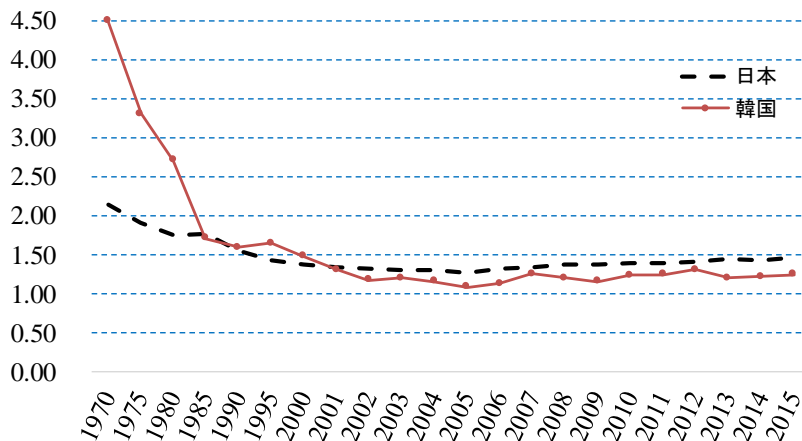
2017年3月

本稿では、日本と韓国の家計パネルデータを用いて、出生コホート別の結婚行動・出産行動の日韓比較を行った。記述的な分析の結果、結婚のタイミングについては、日韓ともに若い出生コホートになるにつれ晩婚化・非婚化が進んでいたが、韓国ではより速いスピードでそれが生じていた。また、既婚女性の出産タイミングについては、日本では出生コホート間の大きな変化は示されず、韓国においては若い出生コホートほど出産をしない傾向にあることが示された。したがって、少子化は、日本では晩婚化・非婚化による現象で、韓国では晩婚化・非婚化および少産化による現象、ということが示唆された。

1 はじめに

本稿の目的は、女性の結婚・出産のタイミングについて、パネルデータを用いて日本と韓国の類似性と差異を整理することである。

日本と韓国はともに少子化という人口学的課題を抱えている。図1は日本と韓国の合計特殊出生率の推移を示したものである。図より、両国とも2000年代には人口置換水準の2.1を大きく下回る少子化が観察される。日本は、2005年に合計特殊出生率が1.26まで低下した後に上昇傾向に転じ、直近の2015年では1.45まで回復している。韓国は、2005年に1.08までに落ち込んだ後は上下変動を繰り返し2012年には1.30まで回復したが、直近の2015年では1.24となっている。



出所) 国立社会保障・人口問題研究所「人口統計資料集」

図1 日本と韓国の合計特殊出生率の推移

本稿では、パネルデータを用いて結婚行動・出生行動の分析を行う。パネル調査を活用することにより、同一個人の子供や出産のタイミングを追跡できるため、少子化の要因が晩婚化によって起こっているのか、既婚女性の出生力によって起こっているのかを把握することができる。集計データを用いても、既婚者割合や子どもを持っている人の割合を把握することはできるが、集計データとくにクロスセクションデータにおいては、コーホート効果と時代効果を合わせた結果しか表すことができない。個票データでかつ同一個人を追跡するパネル調査であれば、既婚者に限定して結婚してからの出産タイミングも把握することが可能となる。著者のうち一人はこれまで、日本の家計経済研究所が実施した「消費生活に関するパネル調査」を用いて、日本女性の結婚・出産のタイミングや就業行動の特徴をデータから観察してきたが（戸田，2015），同様の手法を韓国のパネルデータにも適用することによって、日韓比較を行いたい。戸田（2015）の繰り返しになるが、本稿の特徴でもある出生コーホートごとの差異に注目する意義を示す。

第1に、晩婚化（未婚化），少子化といった女性の結婚行動，出産行動の変化がある。厚生労働省「出生に関する統計」（2010年）によると、女性の平均初婚年齢は、1975年に24.7歳、1985年に25.5歳と緩やかに上昇した。その後1997年で26.6歳と12年間で1歳上昇し、2003年に27.6歳と、6年間で1歳上昇するペースとなっている。これに関連して少子化の進展も進んでいる。少子化の要因として晩婚化や非婚化が与える影響を指摘する研究や（例えば、岩澤2002），それ以外の要因を指摘する研究もあるが（例えば、山口2009），いずれにせよこの20年間で女性の結婚・出産行動に大きな変化が確認されている。

第2に、非正規雇用の増大を中心とした労働環境・就業行動の変化がある。総務省統計局「労働力調査」によると、女性について、役員を除く雇用者に占める非正規雇用の比率は1990年に38.1%であったものが2003年には50.6%と半数を超え、2012年は54.5%となっている。韓国についても、日本より複合的に非正規雇用の定義がされるため留意は必要だが、80年代末以降、非正規雇用が増加している（有田，2016）。日本では、酒井・樋口（2005）により、フリーター経験者ほど結婚や出産タイミングが遅れることが実証されている。非正規雇用の拡大は家族形成に影響し、本稿の分析では厳密な検証はできないが、若年世代はまさに労働環境の変化の影響を受けている可能性が高い。

このように結婚，出産にまつわる環境が変化している中で、世代によってどのような差異がみられるのか、あるいはみられないのかといった点は今後の研究を進めていくうえでも重要である。特に日本においては少子化が課題とされ、その文脈の一環として女性活躍推進が政策的課題として議論されている。データの傾向を素朴に観察し現状把握をすることは重要であろう。そこで本稿では世代ごとに結婚，出産の動向を概観し、世代間の差違について検討することで今後の研究の方向性について議論したい¹⁾。

2 使用するデータ

本稿で使用するデータは日本の「消費生活に関するパネル調査」(Japanese Panel Survey of Consumers: JPSC) と韓国の「労働・所得パネル調査」(Korean Labor and Income Panel Study: KLIPS) データである。

2.1 日本のデータ

家計経済研究所「消費生活に関するパネル調査」は、1993年に24～34歳の若年層の女性(コホートA)1500名を全国規模で抽出し、留置回収法で調査を行っている。毎年11月に調査を実施し、1997年からは24～27歳(コホートB)、2003年からは24～29歳(コホートC)、2008年からは24～28歳(コホートD)が追加された。分析には1993年から2014年までの22年間のパネルデータを利用する。なお、韓国のデータとあわせるため、コホートCまでを活用している。

本稿で使用するデータと、出生コホートの定義などについては表1にまとめてある。出生コホートは4種類に分けている。ただし、パネル調査のコホートAは、調査対象として10歳の幅があるのでコホートを2つ分けている。また、各出生コホートの最終学歴もまとめている²⁾。大学・大学院の構成比をみると、コホート1(1959～63年生まれ)では10.9%、コホート2(1964～69年生まれ)は12.8%であり、若い世代ほど同構成比が上昇していることがわかる。なお、コホート1～コホート4という呼び名は韓国のデータでも同年の出生年の女性を対象としている。

表1 日本データのコーホート別の基本統計量

出生 コホート名	出生年	パネル調査における コホート名	最終学歴			調査開始年 のサンプル サイズ
			中等教育	短大・ 高専	大学・ 大学院	
コホート1	1959～63年生まれ	コホートA(1993年～)	68.4%	20.6%	10.9%	802
コホート2	1964～69年生まれ	コホートA(1993年～)	67.2%	20.0%	12.8%	871
コホート3	1970～73年生まれ	コホートB(1997年～)	60.2%	22.2%	17.6%	582
コホート4	1974～79年生まれ	コホートC(2003年～)	54.7%	23.3%	22.0%	884

2.2 韓国のデータ

韓国労働研究院「労働・所得パネル調査」は、1998年より韓国全国規模(済州島除く)で世帯を抽出し、15歳以上の世帯員を調査対象とした毎年質問紙調査を行うパネル調査である。初回収率(回収数/総アタック数×100)は71.76%となっており、2009年に新規サンプルが追加されている。分析には1998年から2012年までの15年間のパネルデータを利用する。

本稿で使用する韓国のデータと、出生コホートの定義などについては表2にまとめてある。先述のように、出生コホートは日本と同一の出生期間で区分けしている。最終通学

歴の学歴構成をみると、大学・大学院の構成比については、コーホート1（1959～63年生まれ）は11.1%、コーホート2（1964～69年生まれ）は21.7%であり、若い世代ほど同構成比が上昇していることが確認できる。コーホート1～4の違いをみると、若い世代ほど高等教育学歴割合が高くなる傾向は日本と同様であるが、本稿のデータでは、韓国の方がその傾向が著しいことがわかる。

表2 韓国データのコーホート別の基本統計量

出生 コーホート名	出生年	最終学歴			調査開始年 のサンプル サイズ
		中等教育	短大・ 高専	大学・ 大学院	
コーホート1	1959～63年生まれ	82.2%	6.7%	11.1%	659
コーホート2	1964～69年生まれ	69.1%	9.2%	21.7%	841
コーホート3	1970～73年生まれ	59.1%	15.8%	25.1%	500
コーホート4	1974～79年生まれ	36.9%	23.4%	39.7%	836

3 結婚のタイミングに関する分析

図2は、出生コーホート別に未婚率の推移を表したものである³⁾。大まかな傾向としては、日韓ともに若い出生コーホートになるにつれて、未婚率が高まっていることが示されるが、韓国の方がより未婚率が上昇していることがうかがえる。例えば、25歳時点に限定してみると、日本の未婚率は、コーホート1が43.6%であるが、コーホート2は57.2%、コーホート3は61.4%、コーホート4は67.2%と上昇している。それに対し、韓国の25歳の未婚率はコーホート1が46.0%と日本とほぼ同水準、コーホート2（51.1%）、コーホート3（61.9%）も日本と同様であるが、コーホート4は86.9%と日本の同じコーホートと比べて20ポイント程度高くなっている。

同様に、30歳時点に限定してみると、日本の未婚率は、コーホート1が16.9%、コーホート2が28.2%、コーホート3が33.3%、コーホート4が38.2%と徐々に上昇している。それに対し、韓国ではコーホート1では23.6%と日本の同じコーホートよりも7ポイント高く、コーホート2は22.7%と韓国のコーホート1と遜色ないが、コーホート3は40.7%、コーホート4は56.6%と、若いコーホートになるにつれて未婚率が大きく上昇していることが観察できる。

未婚率については日本と韓国ともに、若いコーホートになるにつれて上昇している傾向にあったが、韓国のコーホート4においてはそれ以前のコーホートに比べ未婚率が著しく上昇している点が特徴的であった。

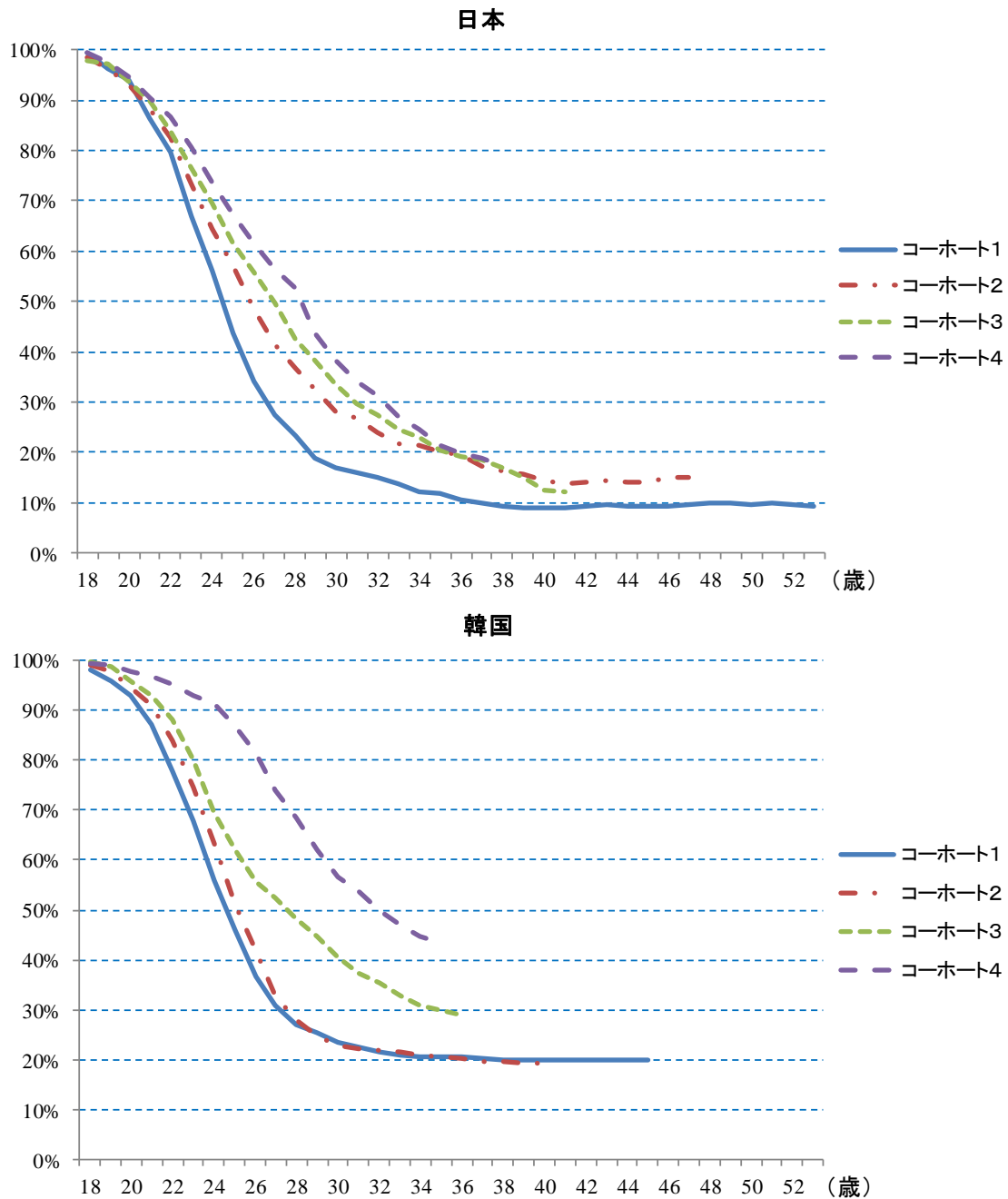


図2 出生コホート別・未婚率の推移 (全サンプル)

4 出産のタイミングに関する分析

次に、出産のタイミングについてみていこう。図3はコホート別に子どものいない女性の割合の推移を示したものである⁴⁾。グラフの推移をみる限り、日韓ともに若いコホートになるにつれて、子どものいない女性の割合が高まっているが、韓国の方がその傾向が強いことがわかる。例えば、25歳時点に限定するとコホート1では日本は62.4%、韓国は49.1%と韓国の方が既に13ポイントほど低い。コホート2では日本は69.4%、韓国

は 62.2%と日韓の差は少し縮まり、コーホート 3 では日本は 82.9%、韓国は 70.5%と日本の方が子どものいない女性の割合が高くなり逆転する。しかしコーホート 4 では日本は 76.8%、韓国は 93.2%となっており、韓国の方が高くなっている。韓国の少子化のスピードはすでにコーホート 1 の段階で日本を上回っていたが、コーホート 4 ではほとんどの女性は子どもがいないことがわかる。

また、30歳時点に限定をしてコーホート 1 からコーホート 4 の子どものいない女性の割合の数値を順に並べてみると、日本は、各々 24.4%、37.7%、44.6%、50.9%と上昇しているが、韓国は、各々 10.8%、17.6%、53.9%、89.0%と大きく上昇している。コーホート 1 とコーホート 2 では、韓国よりも日本において子どものいない女性の割合が高いが、コーホート 3、コーホート 4 においては、韓国の方が同割合が高い。韓国での少子化が日本以上に急激に進んでいることがうかがえる。

こうした動向が、晩婚化による影響なのか、有配偶女性の出生力の低下によるものか、それとも両方によるものかを把握するため、有配偶女性に限定して、結婚してからの年数ごとに子どものいない女性の割合を算出した。この結果が図 4 に示されている。

日本と韓国ともに、コーホートが若くなるにつれて、子どものいない女性の割合が上昇しているが、上昇の度合いは日本の緩やかな状況に比べて、韓国は大きく上昇している。とりわけ、コーホート 3 とコーホート 4 での上昇幅が大きいといえる。

結婚 5 年目に限定して子どものいない女性の割合をみると、コーホート 1 では日本は 9.2%、韓国は 4.5%とわずかではあるが韓国の方が子どものいない女性の割合が低い。過去のコーホートでは韓国よりも日本の結婚 5 年目時点での女性の出生力は低かったといえる。コーホート 2 では、日本は同値が 11.4%、韓国は 10.7%とそれほど遜色はない。コーホート 3 では、日本の同値が 20.0%であるのに対し、韓国は 50.0%と大きく上昇している。コーホート 4 では、日本の同値が 12.8%とコーホート 3 より低下しているのに対し、韓国は 92.0%とコーホート 3 より大きく上昇している。

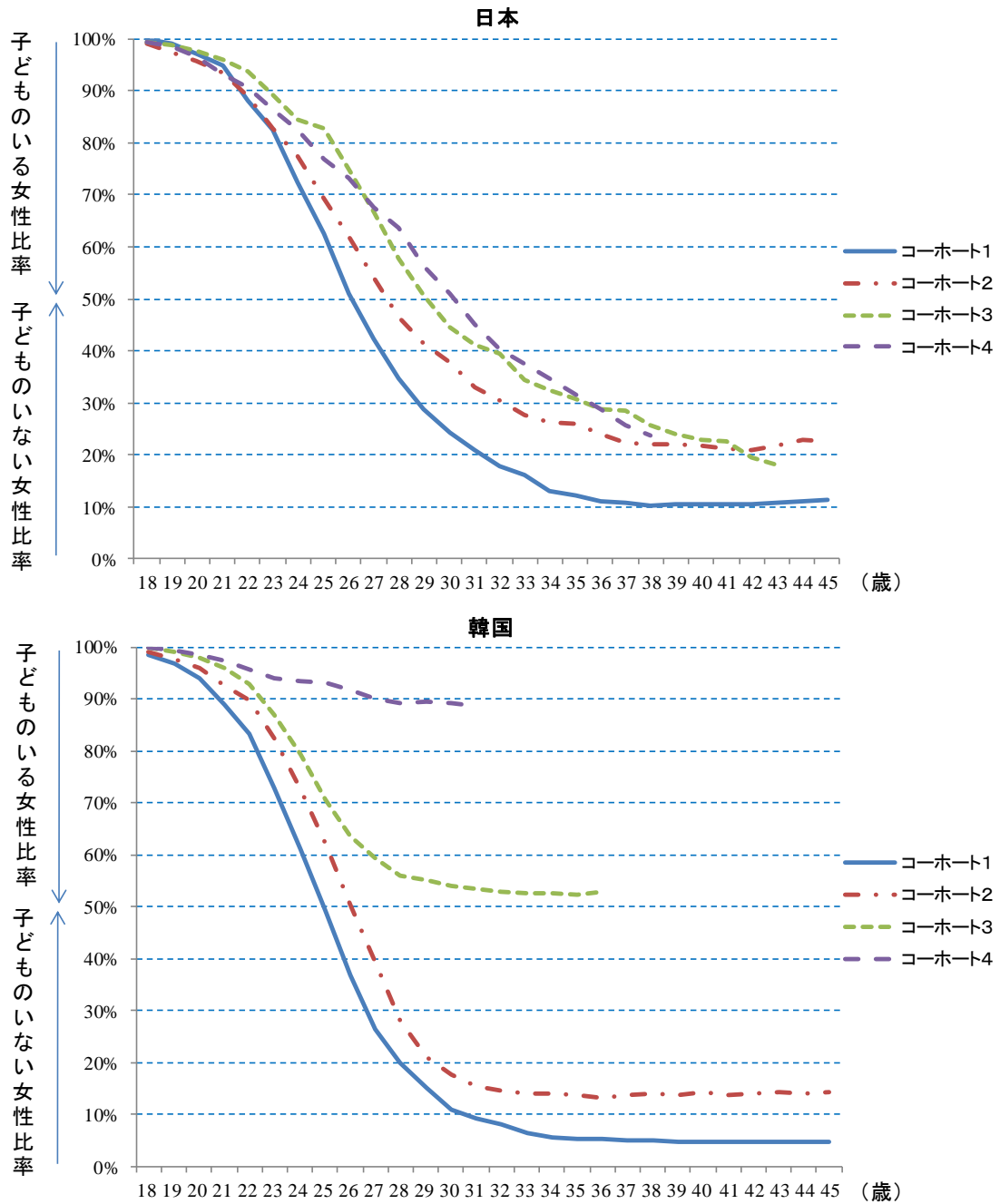


図3 出生コホート別・子どものいない女性の割合の推移（全サンプル）

この結果をみる限り、日本ではやや有配偶者の出生力が低下しているが、韓国に比べて大きく低下しているとはいえない。そのため、日本では主に晩婚化・非婚化によって少子化が起きているといえる。他方の韓国では、有配偶者の出生力が若いコホートになるにしたがって大きく低下しているため、晩婚化・非婚化および有配偶女性の出生力の低下が同時に少子化を生じさせているといえる。

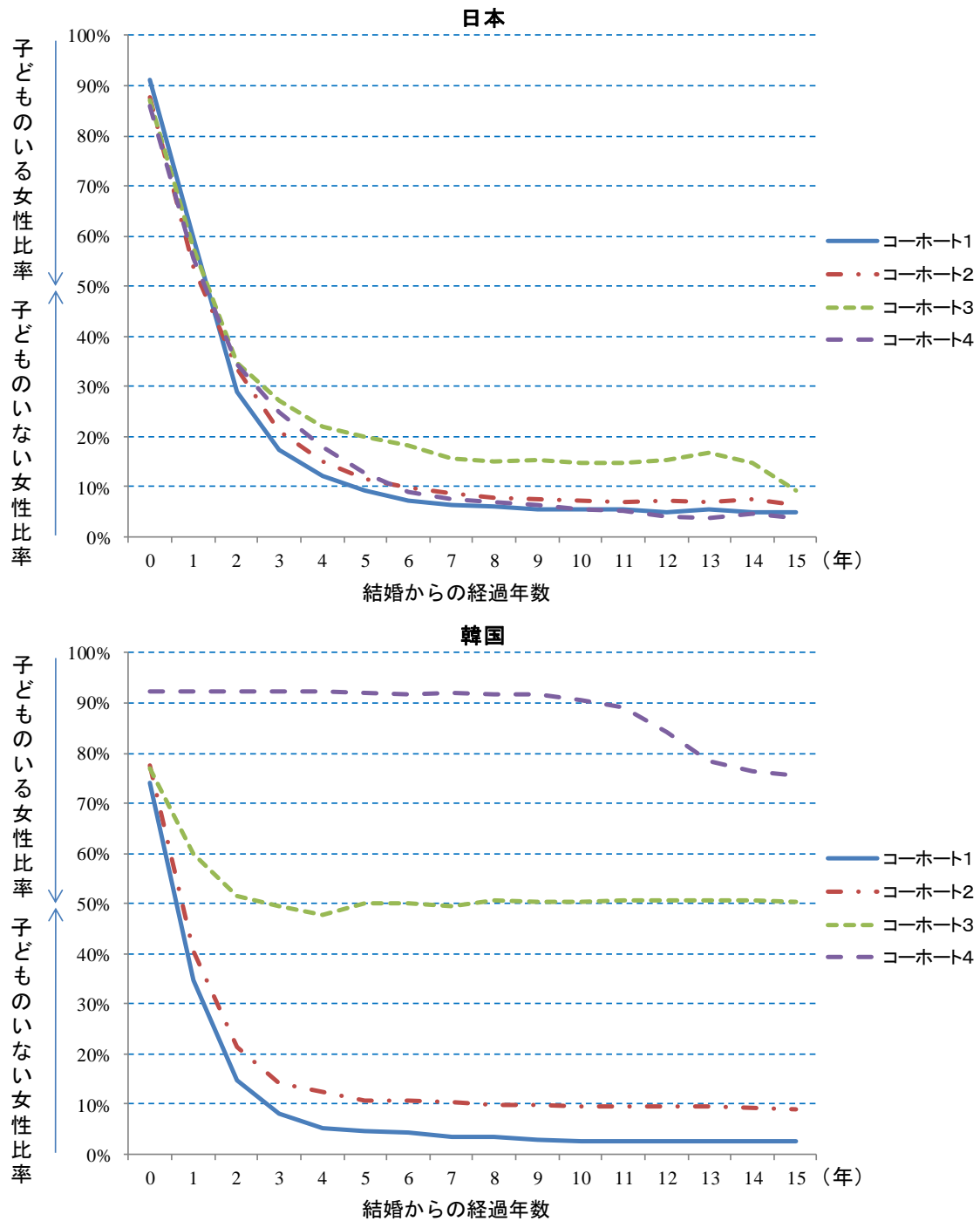


図4 出生コホート別・子どものいない女性の割合の推移
(有配偶者, 結婚からの経過年数)

5 むすびにかえて

本稿では、日本と韓国の家計パネルデータを用いて、出生コホート別の結婚行動・出生行動の日韓比較を行った。記述的な分析の結果、結婚のタイミングについては、日韓ともに若い出生コホートになるにつれて未婚者の割合が増えていたが、韓国ではより速い

スピードで同割合が増えていた。また、既婚者の出産タイミングについては、日本では出生コーホート間の大きな変化は示されなかったが、韓国においては若い出生コーホートほど晩産化・無子化の傾向にあることが示された。本稿の分析から、日韓両国において少子化が生じている背景の異なるメカニズムが確認された。具体的には、日本の少子化は非婚化・晩婚化による現象であるのに対し、韓国の少子化は晩婚化・非婚化および少産化による現象であることを記述的な分析から明らかにした。異なる少子化の背景メカニズムの存在は、両国において理論および政策の適用可能性が異なることを含意する。

以上の結果をふまえると、今後は以下の研究をさらに進めていく必要があるだろう。第1に、結婚と出産は同時に意思決定されている可能性があり、同時決定を意識した動学モデルによって結婚と出産行動を分析する必要がある。とりわけ日本において、子どもを作り育てたいために結婚をすると考えられているとすると、育児を困難にする要因を緩和させるだけでなく、結婚しない／できないようにしている要因を除去しない限り少子化を緩和させることができないであろう。今後の分析の範囲としては、家族形成行動のみならず就業行動も含めることが必要となる。第2に、日韓の出生世代別の既婚女性の出産タイミングに差が生じている要因を検討することである。ひとつの暫定的な仮説として、日本では出産意欲の高い人ほど結婚をする可能性があるが、韓国ではその可能性は低く、むしろ結婚してダブルインカムによって生計を維持することを志向する傾向が日本よりも強い、といった点があげられる。この点は今後の検討課題としたい。

[注]

- 1) 本稿のように、「消費生活に関するパネル調査」を用いて、結婚、出産、就業の動向を調べた研究として、樋口・阿部（1999）がある。
- 2) 学歴の変数は、最終通学歴を表しており、中退の場合はその学歴を卒業していない。
- 3) 日本の「消費生活に関するパネル調査」では、調査対象者は最低でも24歳以上となる。それ以前の婚姻状況については、調査対象初年度において既婚である者に対して、いつ結婚をしたかを訊いているためその情報をもとに、18歳からの未婚率を算出している。離別、死別の場合は未婚者ではない、としている。韓国の「労働・所得パネル調査」では、初婚の時期を訊いているため、その情報をもとに18歳からの未婚率を算出している。
- 4) 日本のデータでは、調査対象期間以前の出産状況については、調査開始年時点での長子の年齢より逆算して、出産したタイミングを算出した。韓国のデータでは、初産時期をパネル調査の初年度で訊いているためその情報を基本として用いた。初回時点で子どものいない女性については、パネル調査期間中に世帯内の高校生以下の子どもの数が増えたか否かで長子の出産有無とタイミングを把握した。

〔謝辞〕

本研究では、公益財団法人家計経済研究所「消費生活に関するパネル調査」、および韓国労働研究院「労働・所得パネル調査」の個票データを利用した。記して感謝申し上げたい。

〔参考文献〕

有田伸，2016，『就業機会と報酬格差の社会学——非正規雇用・社会階層の日韓比較』東京大学出版会。

岩澤美帆，2002，「近年の期間 TFR 変動における結婚行動および夫婦の出生行動の変化の寄与について」『人口問題研究』58(3), 15-44.

酒井正・樋口美雄，2005，「フリーターのその後——就業・所得・結婚・出産」『日本労働研究雑誌』535:29-41.

戸田淳仁，2015，「新規追加コーホートの結婚・出産のタイミング，就業行動比較」『季刊家計経済研究』108, 18-26.

樋口美雄・阿部正浩，1999，「経済変動と女性の結婚・出産・就業のタイミング」樋口美雄・岩田正美編『パネルデータからみた現代女性 結婚・出産・消費・貯蓄』東洋経済新報社。

山ロー男，2009，『ワークライフバランス 実証と政策提言』日本経済新聞出版社。

母親の就業状況と家計の教育費支出 ——韓国労働パネルデータを用いて——

李秀眞

(弘前大学教育学部)

2017年12月

本研究では、有配偶状態にあり、すべての子どもの年齢が満6歳から満18歳以下である有配偶女性を対象に、就業状況と教育費支出の関係を明らかにした。その結果は以下のとおりである。第1に、母親の就業状況別の家計の消費支出費目の支出額を比較したところ、本研究で注目する私的教育費は、未就業の場合、就業している場合よりやや高いものの、母親の就業状況別の差が認められなかった。一方、公的教育費および子どもへの小遣いにおいては、就業状況別の差がみられ、就業している場合に支出額が多いことが確認された。第2に、有配偶女性の就業状況が家計の消費支出のうち、私的教育費に与える影響を分析した結果、未就業に比べて就業している場合に私的教育費の金額が少ないことが明らかになった。今回の分析では、一時点での就業状況に注目しているため、今後の分析では母親の就業選択と私的教育費との双方向的な関係を考慮にいった分析が必要である。

1 研究の背景

1.1 研究の目的

家計に占める教育費の割合が高いことは、韓国の教育環境の特徴の一つとして挙げられている。家計に占める教育費の増加に対し、教育費以外の支出を節約したり、貯蓄を崩したりして対応するという研究がある (Yoo, 2010)。主な収入元である夫の収入のみならず妻の収入が家計の教育費支出に影響を与えることや、共働き世帯の教育費支出が他の世帯に比べて高いとの報告もある (都村, 2006)。これらの背景から教育費を賄うためには収入源の増加が必要となり、既婚女性が労働市場に参加することによって教育費支出に対応しようとする家計行動が予想される。

本研究では、既婚女性の就業状況と教育費支出の関係を明らかにすることを目的とする。また、教育費に関しては、公的教育費と私的教育費に分けて考えることができるが、公的教育費は学校教育費等を含むものとし、また、私的教育費は塾、個別指導、放課後教室等を含むものとする。

2 先行研究

2.1 私的教育費に影響を与える要因

ここでは、教育費のうち、学校教育費等を除く、私的教育費支出に影響を与える要因について検討する。

まず、母親の就業については統一された結果は得られていない。Kim (2008) では、母親の就業形態は私的教育費に有意な影響を与えないとの結果が得られた。一方、Seok・Kim and Noh (2012) の研究では、女性の経済活動類型を4つに分類し、それぞれが私的教育費にどのように関係しているのかを分析した。その結果、女性が経済活動をしていない場合に、他の経済活動類型に比べて私的教育費をより多く支出することを明らかにしている。

他方、母親の学歴と就業形態を同時に考慮した Kim (2008) の研究では、学歴が低い母親の場合、就業状態にいることは、私的教育費を増加させる要因であるが、学歴が高い母親に関しては、就業していない場合に私的教育費が増えると指摘している。Kim (2008) は、これらの現象を高学歴の母親の教育熱が母親自身の就業を委縮させる可能性があるとして解釈している。

親の学歴に関しては、一貫して父親あるいは母親の学歴が高いほど、私的教育費の支出が高いことが確認されている (Kim, 2008 ; Kim and Lee, 2009 ; Seok・Kim and Noh, 2012 ; Song, 2012)。特に、Kim (2008) は、世帯所得をコントロールしても、親の学歴によって私的教育費が増加することを明らかにし、親の学歴と教育に対する期待すなわち教育熱と関連があると解釈している。Kim and Lee (2009) では、有配偶女性の学歴が2年制大学以上である場合、また、配偶者の学歴が2年制大学以上である場合、私的教育費が高いことが確認された。

世帯所得の増加も私的教育費の増加につながることも多数の研究で検証された (Seok・Kim and Noh, 2012 ; Song, 2012)。また、母親の勤労所得が多いほど私的教育費支出が増加するとの結果もある。

子ども関連要因のうち、子ども数に関しては、各学年別の子ども数が多いほど、私的教育費が増加するとの研究があるが (Kim, 2008)、中学生、小学生、高等学生の順に、子ども数による私教育費の差が大きいことも示されている。他方、子ども数が多いほど私的教育費は減少するとの研究 (Seok・Kim and Noh, 2012 ; Song, 2012) もある。

3 データおよび分析方法

3.1 データ

実証分析のために韓国労働パネル調査データを用いる。韓国労働パネル調査は世帯の所得と消費の状態に関する情報と個人の経済活動参加および労働市場移動過程等の情報を提供している。本研究は、有配偶女性の就業状況と家計の私的教育費の関係を検証するため、女性の就業状況および家計の所得・消費を中心に標本を抽出した。韓国労働パネル調査は、調査年度一年前 (t-1 期) を基準に世帯所得および個人の勤労所得、世帯の消費支出の情報を提供している。所得および家計の私的教育費支出に対する情報を得るために、本研究では、第15次年度 (2012年度) と第16次年度 (2013年度) のデータを統合して使用する。

本研究での分析対象は、韓国労働パネル調査の第15次年度時点で有配偶状態にあり、す

べての子どもの年齢が満6歳から満18歳以下である、有配偶女性787名を対象とする。

3.2 変数の測定

変数の測定に関しては、表1に示した。従属変数として私的教育費を設定したが、家計の消費支出費目のうち、私的教育費の名目金額を用いる。独立変数としては、第15次年度の就業状態を用いるが、第15次年度において未就業状態であった場合には0の値を、就業状態であった場合には1の値を与える。

その他の変数として、中学生子どもの有無、子ども数、本人の学歴、世帯の総勤労所得、本人の勤労所得を用いる。

表1 変数の測定

項目	変数の測定	
従属変数	家計の消費支出(生活費)のうち、私的教育費の名目金額	
独立変数	第15次年度の就業状態	未就業=0 就業=1
	中学生子ども有無	なし=0 あり=1
	子ども数	
	本人の学歴	高卒以下=0
		大卒以上=1
	世帯の総勤労所得(年間平均)	連続変数として利用
本人の勤労所得(月平均)	連続変数として利用	

4 分析結果

4.1 基礎分析

4.1.1 分析対象者の一般的属性

分析対象者の一般的属性を表2に示した。学歴をみると、夫は高卒以下が45.0%、大卒以上が55.0%であり、妻においては、高卒以下54.1%、大卒以上45.9%である。世帯の年間平均収入をみると、3,600万ウォン未満は31.3%、6,000万ウォン未満は40.1%、6,000万ウォン以上は28.7%を占めている。妻の月平均収入は300万ウォン未満が33.5%、500万ウォン未満が38.2%、500万ウォン以上が28.3%を占めている。妻の就業状況は、未就業が46.1%、就業が53.9%である。また、妻の年齢は30代が38.9%、40代58.6%、50代が2.5%を占めている。子どもに関してみると、子どもが1人いる場合は21.5%、子どもが2人いる場合は70.4%、3人以上いる場合は8.1%を占めている。子どもの学校段階をみてみると、第1子の場合、未就学児3.9%、小学生39.1%、中学生28.5%、高校生28.5%である。第2子の学校段階をみると、未就学児13.3%、小学生55.2%、中学生25.7%、高校生5.8%である。第3子の学校段階は、未就学児15.6%、小学生73.4%、中学生10.9%である。

表 2 研究対象者の一般的属性

		度数	%			度数	%
夫学歴	高卒以下	354	45.0	子ども数	1人	169	21.5
	大卒以上	433	55.0		2人	554	70.4
妻学歴	高卒以下	426	54.1		3人以上	64	8.1
	大卒以上	361	45.9	第1子の学校段階	未就学	31	3.9
世帯の総勤労所得 (年間平均)	3600万ウォン未満	236	31.3		小学生	308	39.1
	6000万ウォン未満	302	40.1		中学生	224	28.5
	6000万ウォン以上	216	28.7		高校生	224	28.5
妻の月平均 収入	300万ウォン未満	248	33.5	第2子の学校段階	未就学	82	13.3
	500万ウォン未満	283	38.2		小学生	341	55.2
	500万ウォン以上	210	28.3		中学生	159	25.7
妻の就業状況	未就業	363	46.1		高校生	36	5.8
	就業	424	53.9	第3子の学校段階	未就学	10	15.6
妻の年齢	30代	306	38.9		小学生	47	73.4
	40代	461	58.6		中学生	7	10.9
	50代	20	2.5				

4.1.2 母親の就業状況および学歴別にみる私的教育費支出

本節では、母親の就業状況と学歴別に私的教育費にどれくらい支出しているのかを確認する。図示はしないが、母親の就業状況と年間平均の世帯収入との関係についてみると、未就業の場合は4,661万ウォン、就業の場合5,889万ウォンであり、母親の就業状況別に年間平均の世帯収入に違いがあることが確認され、この差は統計的にも有意であった。

母親の就業状況および学歴と私的教育費との関係を図1に示した。未就業の場合は平均56万ウォン、就業の場合は平均52万ウォンの私的教育費支出があることが確認できるが、その差は統計的に有意ではない。母親の学歴別にみた場合、高卒以下が44万ウォン、大卒以上が66万ウォンであり、その差は統計的にも有意であった。

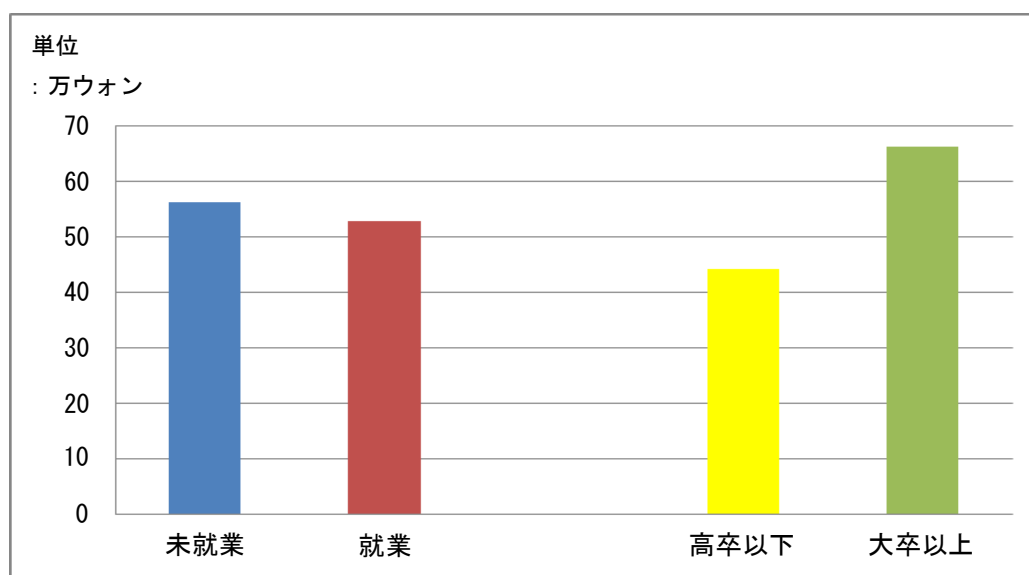


図1 母親の就業状況・学歴と私的教育費との関係

注) 資料：韓国労働パネル第15次年度(2012年度)と16次年度(2013年度)のデータ

4.1.3 母親の勤労所得と私的教育費支出との関係

図2は母親の勤労所得と私的教育費支出との関係を示したものであるが、母親の勤労所得と家計の私的教育費支出には平行な関係がみられた。

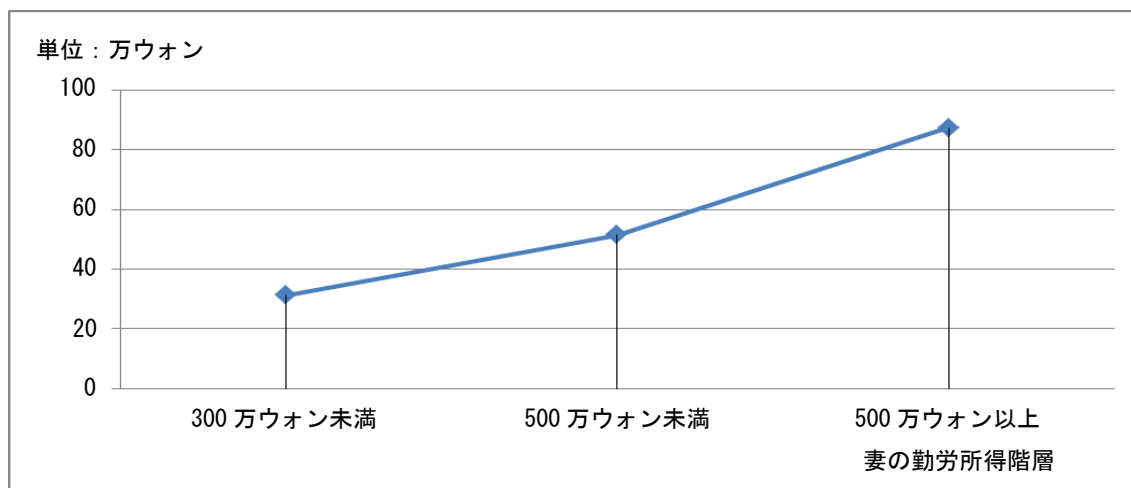


図2 母親の勤労所得と私的教育費支出との関係

注) 資料は図1と同様

4.1.4 母親の就業状況別の家計の消費支出

表3 母親の就業状況別家計の消費支出

	母親の就業状況別 消費費目別平均生活費 単位：万ウォン(SD)		t値
	未就業	就業	
食費	49.52(20.65)	50.37(20.98)	-0.561
外食費	11.81(8.73)	15.44(13.16)	-4.388 ***
公的教育費	12.45(18.78)	15.63(21.34)	-2.159 *
私的教育費	56.13(47.55)	52.76(45.82)	0.991
車両維持費	27.64(17.02)	33.13(22.19)	-3.773 ***
住居費	21.42(10.54)	23.54(15.18)	-2.198 *
衣服費	7.82(5.55)	9.59(7.13)	-3.765 ***
慶弔費	6.79(6.98)	7.89(7.17)	-2.115 *
保健医療費	5.24(12.83)	4.68(9.54)	0.690
教養娯楽費	8.94(11.35)	9.74(10.00)	-1.025
耐久財	3.87(19.49)	3.40(13.53)	0.382
通信費	15.94(5.62)	17.98(7.19)	-4.286 ***
子ども小遣い	5.38(7.01)	7.13(7.66)	-3.265 **
その他世帯員の小遣い	24.42(18.49)	29.87(19.96)	-3.877 ***

注 *p<.05 **p<.01 ***p<.001

表3は、母親の就業状況別の家計の消費支出を示したものである。食費、外食費、公的教育費、車両維持費、住居費、衣服費、慶弔費、通信費、子どもへの小遣い、その他世帯員への小遣いにおいて、就業状況別の差がみられた。これらの費目に関しては、10個の費目ともに就業している場合に消費支出が多いことが確認された。一方、前節でも述べたが、母親の就業状況別の私的教育費の差は認められなかった。また、保健医療費、耐久財においては未就業である場合に消費支出が若干多い傾向にあったが、就業状況別の差は認められなかった。

表4 所得階層別家計の消費支出

	所得階層別 消費費目別平均生活費 単位: 万ウォン(SD)			F値
	世帯所得<下>	世帯所得<中>	世帯所得<上>	
食費	43.47	48.47	59.34	37.42 ***
外食費	8.24	13.70	19.91	68.63 ***
公的教育費	10.83	13.93	18.30	7.84 ***
私的教育費	30.87	47.65	89.42	124.60 ***
車両維持費	18.50	30.64	43.52	114.71 ***
住居費	20.37	21.55	26.32	13.21 ***
衣服費	5.64	8.28	12.88	87.56 ***
慶弔費	4.17	7.04	11.43	70.45 ***
保健医療費	3.58	4.32	7.35	7.320 ***
教養娯楽費	5.85	8.47	14.61	44.77 ***
耐久財	1.49	3.23	6.52	5.40 **
通信費	14.19	17.06	20.09	52.01 ***
子ども小遣い	5.16	5.99	8.10	9.60 ***
その他世帯員の小遣い	16.61	26.07	41.08	118.23 ***

注 **p<.01 ***p<.001

表4は、所得階層別の家計の消費支出を示したものである。世帯の所得階層別の分析を実施するにあたって、世帯所得が<下>のグループ、世帯所得が<中>のグループ、世帯所得が<上>のグループに分類するが、それぞれ3,600万ウォン未満、3,600万ウォン以上6,000万ウォン未満、6,000万ウォン以上と定義する。

分析の結果、食費、外食費、公的教育費、私的教育費、車両維持費、住居費、衣服費、慶弔費、保健医療費、教養娯楽費、耐久財、通信費、子ども小遣い、その他世帯員の小遣いにおいて、所得階層別の差がみられた。これらの費目に関しては、世帯所得<下>、世帯所得<中>、世帯所得<上>グループの順に消費支出が多いことが確認された。特に、私的教育費に注目してみると、世帯所得<中>のグループの場合47.65万ウォンであるが、世帯所得<上>のグループの場合89.42万ウォンであり、所得階層別の差がもっとも大きいことが確

認できた。

表5は、公的教育費、私的教育費、子ども小遣いを取り上げて、母親の就業状況と所得階層別の消費支出を算出したものである。3つの消費費目においては、就業している場合と未就業の場合ともに、世帯所得<下>、世帯所得<中>、世帯所得<上>グループの順に各費目の支出金額が多いことが確認できた。特に、母親が未就業の場合、所得階層別の私的教育費の差が大きいが、世帯所得<上>のグループにおいての私的教育費の消費支出は113.39万ウォンと、母親が就業している世帯所得<上>のグループの79.99万ウォンよりも多い。

表5 母親の就業状況と所得階層別 家計の消費支出

	母親の就業状況と所得階層別 消費費目別平均生活費 単位:万ウォン(SD)			F値
	就業			
	世帯所得<下>	世帯所得<中>	世帯所得<上>	
公的教育費	13.20	13.69	19.10	3.35 *
私的教育費	25.24	42.28	79.99	62.8 ***
子ども小遣い	6.24	6.46	8.35	3.25 *
	未就業			F値
	世帯所得<下>	世帯所得<中>	世帯所得<上>	
	公的教育費	9.32	14.21	
私的教育費	34.47	53.79	113.39	88.90 ***
子ども小遣い	4.47	5.45	7.46	3.93 *

注 **p<.01 ***p<.001

4.1.5 子ども数と子どもの学校段階別 平均私的教育費

表6には、子ども数と子どもの学校段階別にみた月平均の私的教育費を示した。子どもが1人の場合、中学生43.49万ウォン、小学生36.5万ウォン、未就学児30.38万ウォン、高校生25.00万ウォンの順であった。子どもが2人の場合、第1子が中学生、第2子が未就学児の場合にもっとも私的教育費支出が多く、その金額は83.50万ウォンであり、続いて、第1子が中学生、第2子が小学生の場合が67.47万ウォンで金額が多い。また、第1子が高校生で、第2子が中学生の場合に61.84万ウォンの順であった。一方、第1子が高校生、第2子が小学生の場合に48.09万ウォンでもっとも少ない結果であった。中学生がいる世帯において私的教育費が多いことが推察される。

子どもが3人の場合をみると、第1子、第2子ともに小学生、第3子は未就学児の場合に77.00万ウォンで私的教育費の支出が最も多かった。その次に多いのは、第1子が高校

生，第2子が中学生，第3子が小学生の場合で金額としては，72.00万ウォンであった。

表6 子ども数と子どもの学校段階別 平均私的教育費

子ども数と子どもの学校段階別の平均 私的教育費		一月当たり 平均教育費 単位:万ウォン/SD	頻度(単位:人)
子ども1人	未就学	30.38(29.08)	24
	小学生	36.5(26.49)	71
	中学生	43.49(52.41)	39
	高校生	25.0(35.36)	26
子ども2人	第1子、第2子ともに未就学	59.67(51.37)	6
	第1子は小学生、第2子は未就学	58.63(36.59)	68
	第1子、第2子ともに小学生	54.74(35.65)	146
	第1子は中学生、第2子は未就学	83.50(73.03)	6
	第1子は中学生、第2子は小学生	67.47(50.00)	123
	第1子、第2子ともに中学生	58.19(48.20)	27
	第1子は高校生、第2子は小学生	48.09(41.17)	34
	第1子は高校生、第2子は中学生	61.84(62.36)	103
	第1子、第2子ともに高校生	54.91(58.95)	23
子ども3人	第1子、第2子ともに小学生、第3子は未就学	77.00(59.39)	5
	第1子、第2子、第3子ともに小学生	47.00(26.26)	7
	第1子、第2子ともに中学生、第3子は小学生	63.57(55.95)	7
	第1子、第2子は高校生、第3子は小学生	52.86(62.11)	7
	第1子は中学生、第2子、第3子は小学生	45.00(50.76)	10
	第1子は高校生、第2子は中学生、第3子は小学生	72.00(45.68)	12

*注:該当セルの人数が5以上の場合のみ示す

4.2 推定結果

4.2.1 既婚女性の就業形態が私的教育費に与える影響

本節では，既婚女性の就業形態が家計の消費支出のうち，私的教育費に与える影響を分析した。既婚女性の就業状況は，第15次年度の就業状況が未就業なのか，就業なのかによってダミー変数化したものである。

分析結果をみると，未就業に比べて就業している場合に私的教育費の金額が少ない。これらの傾向は世帯の所得階層別の分析結果と同様であった。子ども関連変数の影響をみると，子どものうち中学生がいる場合，また，子ども数が多いほど私的教育費支出が高いことが確認された。中学生がいる世帯において私的教育費が多いという前節の内容を裏付ける結果であった。世帯の総勤労所得と妻の勤労所得が多くなるにつれて，家計の私的教育費は増えるという結果であった。その他，学歴が高卒以下より大卒以上の場合に私的教育費の支出が多いことが確認できた。

表 7 推定結果

家計の教育費	全体	
	Coef.	Std. Err
【既婚女性の就業状態】		
就業(ref: 未就業)	-15.106	2.857 ***
【子ども関連変数】		
中学生子ども有無	11.291	2.900 ***
子ども数	11.399	2.740 ***
【統制変数】		
学歴 (ref: 高卒以下)		
大卒以上	10.512	2.936 ***
世帯の総勤労所得	.005	.001 ***
妻の勤労所得	.047	.014 ***
cons		†
F	67.96***	
R ²	.358	
Adjusted R ²	.352	
N	740	

注) ***p<.001

5 結論

本研究では、有配偶状態にあり、すべての子どもの年齢が満6歳から満18歳以下である有配偶女性を対象に、就業状況と教育費支出の関係を明らかにした。その結果をまとめ、考察を行う。

第1に、母親の就業状況別の家計の消費支出費目の支出額を比較したところ、本研究で注目する私的教育費は、未就業の場合が就業の場合よりやや多いものの、母親の就業状況別の差がみとめられなかった。一方、公的教育費および子どもへの小遣いにおいては、就業状況別の差がみられ、就業している場合に消費支出が多いことが確認された。

第2に、有配偶女性の就業状況が家計の消費支出のうち、私的教育費に与える影響を分析した結果、未就業に比べて就業している場合に私的教育費の金額が少ない。これらの傾向は世帯の所得階層別の分析結果と同様である。

第3に、母親の就業状況と所得階層別に私的教育費の消費支出を算出したところ、就業している場合と未就業の場合ともに、世帯所得<下>、世帯所得<中>、世帯所得<上>

グループの順に私的教育費の支出金額が多いことが確認できた。特に、母親が未就業の場合、所得階層別の私的教育費の差が大きく、母親が就業している世帯所得<上>のグループよりも支出金額が多いことが特徴であった。

以上の結果を踏まえると、女性の経済活動特性が、非経済活動である場合に、他の経済活動類型に比べて私的教育費をより多く支出することが明らかになったという Seok ら (2012) に一致する結果であった。また、学歴をコントロールしても、未就業の場合の私的教育費が高いという点で、学歴が高い母親に関しては、就業していない場合に私的教育費が高いと指摘した、Kim (2008) の研究とも一致する結果が得られた。

ただ、一時点での就業状況を基準にしているため、私的教育費の支出が多いから、それを補うために就業選択をしたのか、就業しているから私的教育費により支出することができたかの明らかにすることには至っていない。

就業選択と私的養育費の関係においては、双方向性が考えられるため、これらの情報を利用可能なパネルデータの利点を生かした分析は今後の課題としたい。

[謝辞]

二次分析に当たり、韓国労働研究院から「韓国労働パネル、15次～16次年度」の個票データの提供を受けました。

[参考文献]

- Kim D. L. (2008). “The Effects of Children in School on Married Women’s Labor Supply,” *Korean Journal of Labour Economics*, 31(2): 73-102. (in Korean)
- Kim Y. O. and Lee S. H. (2009). 「子ども教育と既婚女性の経歴断絶」, 『第1回 女性家族パネル 学術シンポジウム 発表論文』, 1-19.
- Seok J. E.・Kim B. K and Nho H. J. (2012). “Effect of Married Women’s Economic Activity Type on the Private Education Expenditure for Children in Korea,” *Korean Social Security Studies*, 28(3): 87-120. (in Korean)
- Song H. J. (2012). “Study on Trend and Determinants of Domestic Households’ Debt - Analysis of Micro-Level Data,” *Korea Review of Applied Economics*, 14(3): 51-78. (in Korean)
- Yoo K. W. (2010). “The Impact of Education Expenditure on Household Saving and Consumption Behaviors in Korea,” *Journal of Consumer Studies*, 21(4): 135-160. (in Korean)
- 都村聞人 (2006). 「教育費負担に影響を及ぼす諸要因—JGSS-2002 データによる分析」, *JGSS Research Series*, 2: 135-148.

PART II MICRODATA AVAILABILITY

Are data public goods or intellectual properties?

Dan SASAKI

Institute of Social Science

The University of Tokyo

dsasaki@iss.u-tokyo.ac.jp

March 2016

Abstract : Are data public goods, or intellectual properties? Some data sets are publicly available free of charge, whilst others are commercially provided for a fee, typically by a monopolist. In addition, almost by definition, data are intermediate inputs utilised in research publications which, again, are either publicly or monopolistically delivered to the end audience. In this paper we characterise data provision and its influence on the end audience, using an upstream-downstream two-tier market model, highlighting those efficiency and welfare issues associated with this two-tier market structure, including double marginalisation.

Keywords : Deadweight loss, quality, general interest, price discrimination.

JEL classifications : D42, H41, L12, L17.

Acknowledgements : The author is most indebted to Rui Fukuda for advice and encouragement, and also to Shinsuke Ito and Ryuichi Tanaka for detailed comments and discussions, without which this project would not have materialised. The author remains solely responsible for any imperfections in this paper.

1 Introduction

ARE DATA public goods, or intellectual properties? Some data sets are publicly available free of charge, whilst others are commercially provided for a fee, typically by a monopolist. Probably the most spontaneous and commonplace question here is how much deadweight loss the said monopoly may incur.

To shed light on this seemingly straightforward question, it is critically important to account for the fact that, almost by definition, data are intermediate inputs utilised in research publications which, again, are either publicly or monopolistically delivered to the end audience. This implies that the aforesaid question may involve issues such as double marginalisation.

In standard situations, marginalisation in the upstream market affects *marginal costs* in the downstream market. In the case of data and research publications, however, the subscription fee imposed for data use is a lump-sum payment per paper produced. It thereby constitutes a *fixed cost* downstream, not a marginal cost. To this extent, the presence or the absence of such fees might appear as if neutral to the resultant welfare efficiency.

However, the presence of a fee may deter the entry, i.e., the production of some potential research papers altogether. This is our case of double marginalisation. Accordingly, welfare and efficiency implications are distinct from those in the standard form of double marginalisation, especially in the following two fronts. Firstly, upstream pricing remains neutral to welfare efficiency unless it affects downstream entry decisions. Secondly and consequently, *price discrimination* upstream tends to help minimise deadweight loss, as the upstream monopolist harbours an incentive to discriminate so as to bind the *participation constraint* of every trader downstream.

Preceding literature on double marginalisation and price discrimination has been magnificently rich. Studies on double marginalisation, pioneered by Spengler (1950), indicate that market power exercised in successive markets, i.e., upstream and downstream, serve to compound deadweight losses. Vertical integration is recommended as a solution. The basic working of these studies is that mark-ups upstream constitute marginal costs downstream. Meanwhile, “the vast majority of legal and other policy

disputes over the price discrimination concern input markets, not final good markets” (Yoshida, 2000). Therefore, the upstream price discrimination would charge high (resp., low) prices to those downstreamers whose individual demand is high (ditto), thus discouraging enthusiastic downstreamers in exchange for encouraging unenthusiastic ones. This adverse selection aggravates allocative inefficiency, noted by Katz (1987), DeGraba (1990), *inter alia*. This whole strand of literature is incisively summarised and rephrased in terms of cost pass-through in twin papers by Adachi and Ebina (2014). Nearly all of these studies agree that price discrimination upstream, also referred to as *wholesale price discrimination*, is detrimental to welfare efficiency. Comparatively rare exceptions are found in Inderst and Shaffer (2009), and also in Inderst and Valletti (2009), who construct theory counterexemplifying the aforesaid conventional wisdom, wherein heterogeneous willingness-to-pay amongst downstreamers reflects their cost-efficiency, so that wholesale price discrimination may foster efficiency through production substitution from cost-inefficient downstreamers to cost-efficient ones.

These traditional ways of modelling double marginalisation and wholesale price discrimination, reflecting upstream pricing onto downstream marginal costs, would imply a scenario unlikely in the data vending industry, such that the data subscription price be charged not per data set, but per copy of the research paper sold utilising the data set. The key feature in the data vending industry, that the fee charged upstream constitutes a fixed cost rather than a variable cost downstream, shall be emphasised in what follows.

Section 2 lays out modelling primitives. Section 3 anchors a benchmark with a single representative data user who also acts as a monopolistic supplier in the downstream market for research papers, and compares the double marginalisation outcome against the aforesaid unlikely hypothetical scenario, to highlight the difference between our result and traditional results. Section 4 introduces heterogeneous downstreamers, or viewed alternatively, randomly drawn types of data users which, however, remains unobservable to the monopolistic data vendor upstream, so that there is no scope for upstream price discrimination. The results are then compared against those in section 5 where the types are observable, so that data vending prices can be discriminatory according to user types. A brief summary of results is provided in section 6,

followed by an intuitive discussion in section 7.

2 Preliminaries

2.1 Two-tier markets

Throughout this paper, our model involves two markets: the upstream market for data, and the downstream market for research papers utilising the data. These two markets are served by three tiers of traders: the data provider, the research author who is the data user, and the end audience.

In the upstream market, the data provider supplies a data set which, once collated, requires no marginal costs to supply. The researcher, as the user of the data set, has a two-dimensional type (b, q) where b is the breadth of the audience and q is the quality of the paper produced. For simplicity we remain agnostic whether these types are attributed to the researcher, to the paper, or to the audience, focusing on no more than one representative paper authored by the researcher.

In the downstream market, the researcher brings one research paper to the audience whose inverse demand for the paper is given by $(1 - x/b)q$ where x denotes the readership. Graphically, this gives a linear demand with price-intercept q and quantity-intercept b .

2.2 Supply structures

Altogether we contemplate $2 \times 2 = 4$ alternative scenarios, depending upon whether the data set is publicly provided free of charge or privately sold for a fee, and whether the research paper is published open-access or via commercial monopoly.

To economise our taxonomical burden, we first sort these four cases along the following dichotomy.

Free paper : The research paper is published open-access, free of charge, entailing

a gross surplus of

$$\int_{x=0}^b \left(1 - \frac{x}{b}\right) q dx = \frac{bq}{2}$$

in the downstream market. We assume that the researcher's payoff from the downstream market is proportional to this gross surplus.

Fee paper : The research paper is sold to maximise the revenue $(1 - x/b)qx$ which is maximised upto $bq/4$ at $x = b/2$, where the gross surplus

$$\int_{x=0}^{b/2} \left(1 - \frac{x}{b}\right) q dx = \frac{3bq}{8}$$

suffers a 25-percent reduction from the previous free paper case. ♣

Once again, we assume that the researcher's payoff from the downstream market is proportional either to the revenue or to the gross surplus.

In either case, the researcher's payoff from the downstream market is proportional to bq .

Next, we taxonomise the upstream market configuration.

Free data : The data set is publicly accessible without fees. Therefore researchers of any type b, q can afford to utilise it.

Fee data : Access to the data set is subject to a subscription fee, set by a monopolistic data vendor. The researcher, who is the subscriber to the data set, is willing to pay proportionate to bq .

The fee-charging data vendor sets the subscription fee so as to maximise its revenues. Therefore, the fee is a function (a functional to be precise) of the distribution of bq .

2.3 Welfare efficiency evaluation

We assess the economic efficiency of outcomes based upon the gross utility (measured in willingness to pay) of the end audience. This is because the objective functions of the data vendor and of the research author vary, making it difficult to rank the efficiency of allocative outcomes across $2 \times 2 = 4$ configurations.

These outcomes are then to be compared with the standard case of two-tier markets and double marginalisation, where the price paid in the upstream market is not a lump sum but a standard linear price per x , which then constitutes marginal costs in the downstream market. This exercise serves to highlight the difference between our results and the standard double marginalisation results, and thus how data should be supplied as opposed to other standard intermediate inputs.

Our comparative analysis can be taxonomically illustrated as below.

	Free publication	Fee publication
Free data	No marginalisation	Single marginalisation
Fee data	Single marginalisation	Double marginalisation
(Data fee per copy)		(Standard double marginalisation)

3 Benchmark

We commence our analysis with the simple case where the type distribution is degenerate at $b = q = 1$, so that the research paper faces the inverse demand deterministically and identically equal to $1 - x$.

Free paper publication reaches the entirety of the end audience, entailing the downstream gross surplus

$$\int_{x=0}^1 (1 - x) dx = \frac{1}{2}.$$

Fee paper publication, on the other hand, charges the monopoly (i.e., revenue-maximal) price of $1/2$, resulting in the gross surplus

$$\int_{x=0}^{1/2} (1 - x) dx = \frac{3}{8}.$$

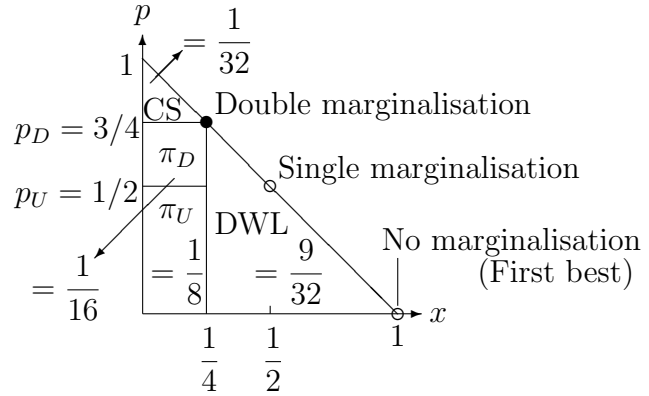
In either case, if a data subscription fee is charged, it is a simple transfer from the research author to the data vendor, thus welfare neutral.

The above observations are now contrasted against the standard format of double marginalisation where the price charged in the upstream market constitutes marginal costs in the downstream market. This corresponds, in our context, to the somewhat unlikely situation where the data usage fee is charged proportional to readership.

Straightforward computation confirms that the monopoly subscription fee is $x/2$, reflected in the downstream marginal cost of $1/2$, entailing a readership of $x = 1/4$ and thus the gross downstream surplus

$$\int_{x=0}^{1/4} (1-x) dx = \frac{7}{32}$$

from monopoly fee publication (as illustrated in the right figure).



In all, our benchmark results are summarised as follows.

	Free publication	Fee publication
Free data	No marginalisation TS = $\frac{1}{2}$, DWL = 0	Single marginalisation TS = $\frac{3}{8}$, DWL = $\frac{1}{8}$
Fee data (Neutral to allocation)	Single No marginalisation TS = $\frac{1}{2}$, DWL = 0	Double Single marginalisation TS = $\frac{3}{8}$, DWL = $\frac{1}{8}$
(Data fee per copy)		(Standard double marginalisation) TS = $\frac{7}{32}$, DWL = $\frac{9}{32}$

4 Complete information

We now contemplate the following two alternative type distributions, both resulting in the same distribution of bq , to compare the results.

[Q] Quality types : Researchers (or research papers) differ in q but not in b .

We continue to assume $b = 1$ for simplicity, whilst q is distributed uniformly between 0 and 2, so is bq .

[B] Breadth types : Researchers (ditto) vary in b but not in q . We assume

$q = 1$, whereas b , and thus bq also, to be uniformly distributed between 0 and 2.

In the current section only, we tentatively envision that the type (b, q) is observable *before* any subscription fee is charged. In other words, the subscription fee can be discriminated depending upon the type.

This inevitably implies that the pricing of data subscription is subject to the *participation constraint*, so as to enable the publication of any type. The following results ensue.

- In [Q], free paper publication attains the first-best gross downstream surplus

$$\frac{1}{2} \int_{q=0}^2 \int_{x=0}^1 (1-x)q \, dx \, dq = \frac{1}{2},$$

whilst fee-charging publication entails

$$\frac{1}{2} \int_{q=0}^2 \int_{x=0}^{1/2} (1-x)q \, dx \, dq = \frac{3}{8}.$$

- Similarly in [B], free publication scores

$$\frac{1}{2} \int_{b=0}^2 \int_{x=0}^b \left(1 - \frac{x}{b}\right) dx \, db = \frac{1}{2},$$

compared to the gross downstream surplus from fee publication

$$\frac{1}{2} \int_{b=0}^2 \int_{x=0}^{b/2} \left(1 - \frac{x}{b}\right) dx \, db = \frac{3}{8}.$$

In words, even when the type distribution is non-degenerate, insofar as the type is observable before the determination of the data subscription fee, the effect of the fee is neutralised and hence no double marginalisation materialises.

Counterfactually, should data subscription be priced per readership as in standard cases of double marginalisation, the monopoly fee would be $qx/2$, translated into downstream marginal costs $q/2$, implying the following.

- In [Q], monopoly fee publication earns a readership of $x = 1/4$ and the gross downstream surplus

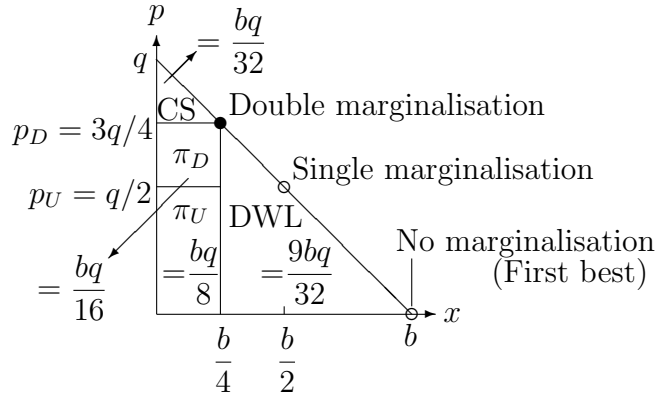
$$\frac{1}{2} \int_{q=0}^2 \int_{x=0}^{1/4} (1-x)q \, dx \, dq = \frac{7}{32}. \quad (\diamond)$$

- In [B], monopoly publication involves a readership of $x = b/4$ and the gross downstream surplus

$$\frac{1}{2} \int_{b=0}^2 \int_{x=0}^{b/4} \left(1 - \frac{x}{b}\right) dx \, db = \frac{7}{32}. \quad (\heartsuit)$$

The standard double marginalisation results remain unchanged in essence from the benchmark case (as illustrated in the right diagram).

Hence under complete information on data user types, our comparative results are summarised below.



	Free publication	Fee publication
Free data	No marginalisation $E[TS] = \frac{1}{2}, DWL = 0$	Single marginalisation $E[TS] = \frac{3}{8}, E[DWL] = \frac{1}{8}$
Fee data (Neutral to allocation)	Single No marginalisation $E[TS] = \frac{1}{2}, DWL = 0$	Double Single marginalisation $E[TS] = \frac{3}{8}, E[DWL] = \frac{1}{8}$
(Data fee per copy)		(Standard double marginalisation) $E[TS] = \frac{7}{32}, E[DWL] = \frac{9}{32}$

5 Incomplete information

Pertaining to those two alternative settings [Q] and [B] introduced in section 4, we now enlist the realistic assumption that, unlike previously, the type remains unobservable at the time of data subscription pricing. Therefore, price discrimination is no longer feasible.

The distribution of bq is identical between [Q] and [B], distributed uniformly between 0 and 2. In either scenario, the revenue-maximal data subscription charge is what exactly the upper half of the said distribution, i.e., $1 \leq bq \leq 2$, can afford. $\langle \spadesuit \rangle$

Free data, free paper : Obviously, this ideal case incurs no deadweight loss in either [Q] or [B]. The readership includes all of the end audience. The efficiency measure in [Q] amounts to

$$\frac{1}{2} \int_{q=0}^2 \int_{x=0}^1 (1-x)q \, dx \, dq = \frac{1}{2},$$

and in [B],

$$\frac{1}{2} \int_{b=0}^2 \int_{x=0}^b \left(1 - \frac{x}{b}\right) \, dx \, db = \frac{1}{2}.$$

Free data, fee paper : By $\langle \clubsuit \rangle$, the fee paper, as opposed to the free paper, incurs a constant 25% deadweight loss whether in [Q] or in [B], entailing an efficiency measure of

$$\frac{1}{2} \int_{q=0}^2 \int_{x=0}^{1/2} (1-x)q \, dx \, dq = \frac{3}{8}$$

or

$$\frac{1}{2} \int_{b=0}^2 \int_{x=0}^{b/2} \left(1 - \frac{x}{b}\right) dx \, db = \frac{3}{8}$$

respectively.

Fee data, free paper : By $\langle \spadesuit \rangle$, fee data, as opposed to free data, incurs a constant 25% deadweight loss whether in [Q] or in [B], entailing the efficiency equal to either

$$\frac{1}{2} \int_{q=1}^2 \int_{x=0}^1 (1-x)q \, dx \, dq = \frac{3}{8},$$

or

$$\frac{1}{2} \int_{b=1}^2 \int_{x=0}^b \left(1 - \frac{x}{b}\right) dx \, db = \frac{3}{8}.$$

Fee data, fee paper : By $\langle \clubsuit \rangle$, fee paper, as opposed to free paper, incurs a constant 25% deadweight loss whether in [Q] or in [B]. Namely,

$$\frac{1}{2} \int_{q=1}^2 \int_{x=0}^{1/2} (1-x)q \, dx \, dq = \frac{9}{32}$$

or

$$\frac{1}{2} \int_{b=1}^2 \int_{x=0}^{b/2} \left(1 - \frac{x}{b}\right) dx \, db = \frac{9}{32}.$$

Unlike previously, double marginalisation now materialises. Incomplete information disables data subscription fee discrimination, resulting in entry deterrence to some types.

We then hypothesise the counterfactual wherein data subscription is charged per readership.

- In [Q], the monopoly publisher maximises the nett profit $((1-x)q - p)x$, that is

$$x = \begin{cases} \frac{1}{2} \left(1 - \frac{p}{q}\right) & p < q \leq 2 \\ 0 & 0 \leq q \leq p \end{cases}$$

anticipating which the upstream vendor charges a price p per readership to maximise the revenue

$$\frac{1}{2} \int_{q=p}^2 \frac{1}{2} \left(1 - \frac{p}{q}\right) p dq = \frac{p}{2} - \frac{p^2}{4} (1 + \ln 2 - \ln p)$$

via the first-order condition

$$\frac{1}{2} \int_{q=p}^2 \left(\frac{1}{2} - \frac{p}{q}\right) dq = \frac{p}{2} \left(\frac{1}{p} - \frac{1}{2} - \ln 2 + \ln p\right) = 0$$

satisfied at $p = .569336 \dots$, where the gross surplus obtains

$$\begin{aligned} \frac{1}{2} \int_{q=p}^2 \int_{x=0}^{(1-p/q)/2} (1-x)q dx dq &= \frac{1}{2} \int_{q=p}^2 \left(\frac{1}{2} \left(1 - \frac{p}{q}\right) - \frac{1}{8} \left(1 - \frac{p}{q}\right)^2\right) q dq = \\ &= \frac{1}{16} \int_{q=p}^2 \left(3q - 2p - \frac{p^2}{q}\right) dq = \frac{1}{16} \left(6 - 4p + \frac{p^2}{2} \left(\frac{1}{2} + \ln 2 - \ln p\right)\right) = \\ &= \frac{1}{16} \left(6 - 4p + \frac{p^2}{2} \cdot \frac{1}{p}\right) = \frac{3}{8} - \frac{7p}{32} = .250457 \dots \end{aligned}$$

which is higher than that in (\diamond).

- In [B], even in the previous complete information case there was no scope for price discrimination, the upstream vendor charging $x/2$. Under incomplete information, the outcome remains unchanged from (\heartsuit).

Comparative results are illustrated in the following table.

	Free publication	Fee publication
Free data	No marginalisation $E[\text{TS}] = \frac{1}{2}, \text{DWL} = 0$	Single marginalisation $E[\text{TS}] = \frac{3}{8}, E[\text{DWL}] = \frac{1}{8}$
Fee data	Single marginalisation $E[\text{TS}] = \frac{3}{8}, E[\text{DWL}] = \frac{1}{8}$	Double marginalisation $E[\text{TS}] = \frac{9}{32}, E[\text{DWL}] = \frac{7}{32}$
(Data fee per copy)		(Standard double marginalisation) [B]: $E[\text{TS}] = \frac{7}{32}, E[\text{DWL}] = \frac{9}{32}$ [Q]: $E[\text{TS}] \approx 0.25, E[\text{DWL}] \approx 0.25$

6 Summative results

- Under complete information on downstream demand, the data subscription fee incurs no loss of efficiency.

Therefore, no double marginalisation arises when data and the resultant research publication are both monopoly priced.

- Under incomplete information, monopoly pricing of data incurs inefficiency.

When the downstream market is also monopoly, there arises double marginalisation.

However, the loss of efficiency is smaller than in the standard form of double marginalisation where the upstream price contributes to the downstream *marginal* cost (as opposed to the downstream *fixed* cost in our case).

- Complete information on q , the end consumers' willingness to pay, enables the upstream vendor to price discriminate, which serves to *eliminate* double marginalisation in our model, unlike in the usual settings where price discrimination aggravates allocative inefficiency.

7 Concluding discussion

In this paper we have launched a brief, simplified analysis on data provision, as an industry with markedly distinctive characteristics. Data, as an intermediate input, implies a two-tier market structure and the inevitable issue of double marginalisation. It is also noteworthy that data provision tends to be priced per data set, rather than per end-readership of the final output utilising the data, so that the price contributes only to fixed costs, not to marginal costs, in the downstream market.

Our findings suggest that, whilst it is always best that all data be provided open-access, even if some are provided monopolistically the resultant deadweight loss may not be quite as devastating as in standard cases where the upstream prices are reflected on the downstream marginal, not fixed, costs. In this sense, our double marginalisation simply consists of two consecutive occurrences of single marginalisation, unlike in standard cases where price mark-ups compound.

Furthermore, price discrimination upstream *hurts*, i.e., aggravates allocative inefficiency in standard cases, whilst it *helps*, i.e., alleviates inefficiency in our case. Namely, should data be distributed for fees, it serves the public interest better if

fees are discriminatory. Intuitively, this is because these fees are non-distortionary to downstream pricing, hence the only relevant effect is that discriminatory fees can accommodate all participation constraints in the downstream market, whilst nondiscriminatory fees deter the entry of some types. Price discrimination in this sense seems rarely encouraged, its benefit being largely unrecognised.

In the interest of simplicity and brevity, this paper has deliberately confined its attention to one and only one aspect of the data provision industry, which accommodates monopolistic intellectual property demanded as an intermediate input. In fact, the industry can also be characterised by other, perhaps equally important, distinctive features, such as a two-sided platform mediating between data collation and data vending. Furthermore, not every data vendor may be monopolistic in that multiple data sets may be close substitutes for some purposes. It is also practicable that the same data set be made available under a variety of differentiated terms and conditions, amongst which users may self-select, in which case a contract theory model may be desired. Last but not least, the public interest may be represented not entirely by total surplus in its market sense, but also in part by informational externalities caused through the very nature of the data used and its use *per se*, including ethical values such as privacy protection. Future research in such directions shall be highly desired.

Bibliography

- Adachi, T. and Ebina, T., “Double marginalisation and cost pass-through: Weyl-Fabinger and Cowan meet Spengler and Bresnahan-Reiss,” *Economics Letters* 122 (2014), 170 - 175.
- Adachi, T. and Ebina, T., “Cost pass-through and inverse demand curvature in vertical relationships with upstream and downstream competition,” *Economics Letters* 124 (2014), 465 - 468.
- DeGraba, P., “Input market price discrimination and the choice of technology,” *American Economic Review* 80 (1990), 1246 - 1253.
- Inderst, R. and Shaffer, G., “Market power, price discrimination, and allocative efficiency in intermediate-goods markets,” *RAND Journal of Economics* 40 (2009), 658 - 672.
- Inderst, R. and Valletti, T., “Price discrimination in input markets,” *RAND Journal of Economics* 40 (2009), 1 - 19.
- Katz, M.L., “The welfare effects of third-degree price discrimination in intermediate good markets,” *American Economic Review* 77 (1987), 154 - 167.
- Spengler, J.J., “Vertical integration and antitrust policy,” *Journal of Political Economy* 58 (1950), 347 - 352.
- Yoshida, Y., “Third-degree price discrimination in input markets: output and welfare,” *American Economic Review* 90 (2000), 240 - 246.

A note on microdata availability in Japan: accessibility and practical information

Rui Fukuda

Graduate School of Economics

The University of Tokyo

fukuda.rui@gmail.com

March 2017

Abstract In this paper an overview and examination of existing social survey data archiving systems and longitudinal data in Japan is provided. Official statistics are also briefly described. The aim of the study is to provide information on the availability of social scientific microdata in Japan. Practical information, such as how to access microdata and the accessibility for non-Japanese speaking researchers is also discussed.

Keywords Accessibility to microdata, social survey data archives, longitudinal data, official statistics, diverse data in society

JEL Classification C81, C83, D00

1 Introduction

The accessibility and types of microdata sets available strongly affect the quality of studies in social sciences. Microdata sets are typically records of individuals' answers to questionnaire surveys, and are constructed in the form of rows (individuals) and columns (answered values to the questions by the individuals).

For official statistics in Japan, the use of official statistical microdata sets for research purposes was previously highly controlled under the Statistics Act No.18 (1946). This limited the number of empirical studies based upon official statistical microdata in Japan (Kitamura and Kinoshita 2013). However, with the amended Statistics Act No.53 (2007) being put into effect in April 2009, the accessibility to official statistical microdata for research purposes has improved.

Furthermore, the accumulation of social survey data is advancing in several data archives, which also denote the improvement of reproducibility of studies. It appears that circumstances for social sciences research are changing for the better in Japan. In addition, panel survey data, which is typically individual data collected over time, is also accumulating in Japan. This data is very useful for analysing individuals' decision making, and can contribute to a better understanding of the dynamics of Japanese society. It should be noted that since the mid-2000s, possibly due to changing attitudes towards privacy after the introduction of the Act on the Protection of Personal Information in 2003, response rates of social surveys have been decreasing. This is a concern in terms of *unit nonresponse* and *item nonresponse*, as there is an effect the quality of collected data (Meyer et al. 2015). Despite this, the overawing positive of improved microdata availability should mean higher quality institutions and policymaking in Japan. There is an academic trend of conducting empirical research using microdata sets—the 'microfoundations of macroeconomics'—, and this increases understanding of people's decision making and diversity in society, which can enhance policymaking. According to Ichimura (2013), there are many reasons why microdata sets are important. First, in order to analyse model parameters on the decision making of individuals, it is essential to use microdata

sets. Second, need control over data processing to conduct analysis on topics of interest, meaning microdata sets are needed. In other words, the researcher needs control over aggregating data and constructing variables. Third, to analyse natural experiments, microdata is indispensable. In particular, for studies that attempt to implement policy evaluations, combinations of detailed information are required and thus microdata sets are essential. With a deeper understanding of diversity in our society, more appropriate institutions can be produced based on microdata analysis.

Several studies are starting to use microlevel administrative rich data sets, for instance in the US (Einav and Levin 2014). The Social Security and Tax Number System, which assigns an ID number to each resident for tax and other financial purposes, was introduced in Japan in January 2016. In the first year, the system was only utilised for the combined management of pensions and unemployment insurance, but from 2017 onwards social security information has also been included as part of the combined management. Unfortunately, public provision of associated microdata data has not been discussed in any detail. Recently, the Japanese government made a declaration, entitled ‘Declaration to be the World’s Most Advanced IT Nation,’ and is focusing on the effective use of information technology. Because of information and communication technology (ICT) improvement, diverse types of data are being generated and are accumulating. However, these share the same challenges as traditional data, which is constructed in the form of rows and columns based upon questionnaire surveys as aforementioned, in terms of data confidentiality (Heffetz and Ligett 2014, Ito 2016). Although we have to be careful about myopic thinking of ‘data-driven innovation,’ undoubtedly the importance of new type data is increasing day by day. Currently, most of the new type data is still underdeveloped in terms of research purpose usage, i.e. data accessibility and data format standardisation. However, in the near future the combination of traditional and new type data will be an important issue in economics and social sciences.

This paper provides background information on social sciences microdata infrastructure, particularly on social survey data and official statistical data in Japan. The study discusses the availability of Japanese microdata, and therefore increases the possibility of data usage, particularly for non-Japanese researchers interested in Japan. While the paper only introduces microdata provision structures of social surveys and official statistics and metadata on representative panel surveys in Japan, information presented in the paper may open new possibilities for studies that investigate behavioural patterns in Japan and studies that conduct comparative analysis including Japan.

The organisation of the paper is as follows. In section 2, changing trends on new type data in Japan is briefly presented. Section 3 introduces microdata archives of social surveys with practical information on accessibility. Section 4 describes the main longitudinal surveys. Section 5 presents the outline of the official statistics system with practical information on data accessibility. The final section provides a summary and the conclusions of the paper. Please mind that words coloured blue and in bold style, which appear after section 2, are those with hyperlinks to trustworthy webpages (accessed in March, 2017).

2 Data in society

In recent times, diverse types of data, such as system log data, multimedia data, and even sensor data, is being collected and analysed. In economic studies, on-line auction transaction data and retail store and online point-of-sales system data is currently often analysed. In labour economics, the evaluation system of the online labour market is utilised (Pallais 2014). In the US, there is the Billion Prices Project, which provides a Consumer Price Index (CPI) every month based on price information from online retailers (Einav and Levin 2014). There are some

practical applications for utilising new type data in Japan too: for example, there is an online library book inventory system called Calil, which provides real-time information from libraries across Japan, and a transfer guide search system based on location information, entitled Jorudan. Data volumes are increasing, pushing information technology advancement. Today, terms like ‘big data,’ ‘internet of things,’ and ‘open data’ are often heard in common discussion. It seems that anticipation for ‘data-driven innovation’ is growing day by day (OECD 2015).

Based on this background, the Japanese government made a declaration entitled ‘Declaration to be the World’s Most Advanced IT Nation,’ stating that “from the perspective of contributing to the use of big data we will encourage making public data available to the private sector (open data) including information in the possession of the government such as geospatial information, remote sensing data from satellites [...] and statistical data. [...] it will be possible to create new services and new public-private sector collaboration services, leading to the creation of a society that fosters innovation in corporate activities, consumer conduct, and social lifestyles” (Strategic Headquarters for the Promotion of an Advanced IT Network Society 2015: 13-14). Regarding the open data system, the Portal Site of Official Statistics of Japan entitled **e-Stat** provides ready-made aggregated statistics of many official and general statistics from the Japanese government in spreadsheet and CSV formats. Recently, based on that data, collaborations of public and private sector are emerging in a service entitled jSTAT MAP (Akou 2016).

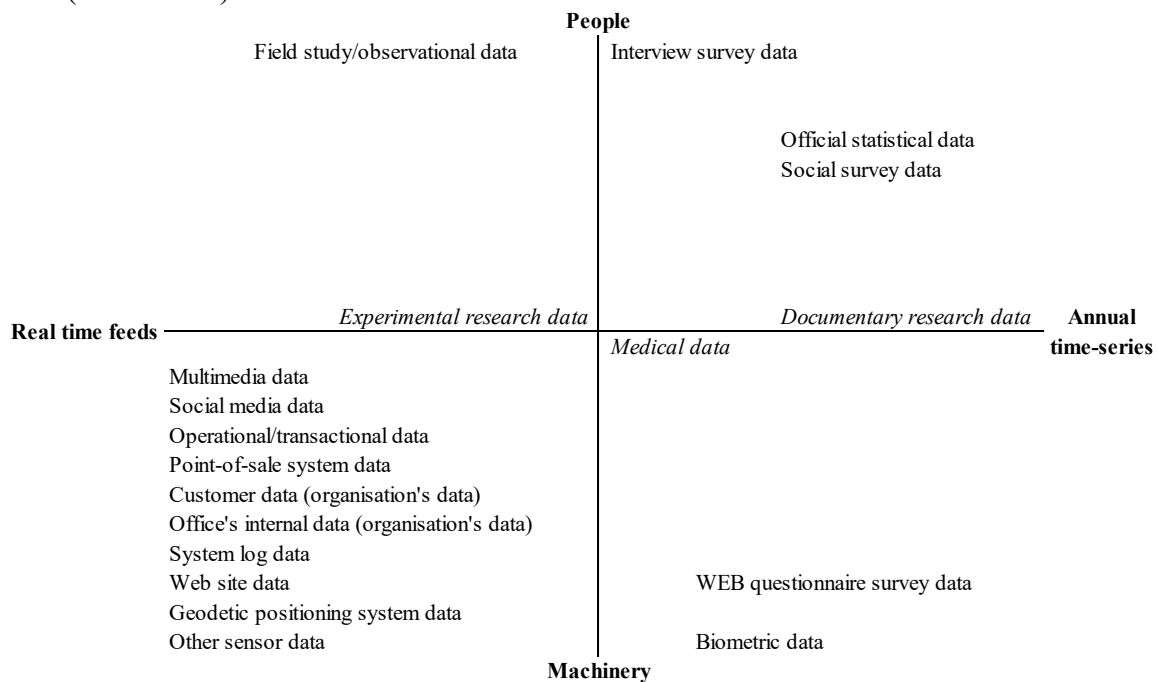


Figure 1: Typology of people’s behaviour concerned data by data collector and frequency dimensions

Source: Constructed by the author

Notes: The horizontal axis shows data frequency and the vertical axis shows the data collector/creator. On data frequency, all data are supposed to be continuous/repeated ones (not one-shot only). Data in *italics* indicate that those data may be in between the borderline of the vertical axis (can be collected by both people and machines)

The new existence of diverse types of data cannot be ignored, and as such, we attempt to categorise and identify the relative position of these data types. Figure 1 denotes data relating to people’s behaviours, including data that potentially can be, but is currently not used in social sciences. The horizontal axis indicates data frequency and the vertical axis denotes the source. As we can see in the figure, the new type data, such as GPS data and social media data, is in the bottom left corner, indicating that they are collected by machines in real time. However, whilst they hold a variety of information, they are still underdeveloped in terms of common accessibility and data format standardisation. From a technical perspective, the movement to

“reduce the dimensionality of large-scale, unstructured data is becoming a crucial challenge in empirical economic research” (Einav and Levin 2014).

In contrast, old type data, such as official and social survey data, is located in the upper-right, showing it is mainly collected by humans, typically annually. Data volumes may be smaller compared to new type data, but they contain rich information appropriate for social sciences analyses. Owing to the data provision format, such as CSV or other standardised formats, the data is easily manageable and is used very often for secondary analyses and usage. Data in this category is mostly provided for free, or with a small handling charge, and is used in many fields in social sciences. This paper focuses mainly on this old type of data—conventional household survey data—although we also consider the importance of new type data as well. Old type data and new type data may be combined in the near future for the purpose of economic and social sciences analysis. Whilst this is an interesting area, it is not the focus of this paper.

It should be also noted that ‘big data’ itself does not necessarily solve all the data problems in the field. Therefore, we should be careful not to expect too much from it. “Data mining of personal data can help increase welfare, lower search costs, and reduce economic inefficiencies; at the same time, it can be a source of losses, economic inequalities, power of imbalances between those who hold the data and those whose data is controlled” (Acquisti 2014: 77). The data requires “a sophisticated understanding of both big data predictive modelling and economic theory” (Einav and Levin 2014). This data needs to be appropriately handled by humans rather than machines, in order to get the most out of it.

3 Social survey data archiving system

In this section, we introduce the Center for Social Research and Data Archives (CSRDA) at Institute of Social Science, University of Tokyo. The CSRDA manages the Social Science Japan Data Archive (SSJDA), which is the central social sciences microdata archive in Japan. There are also several other data archives aside from the SSJDA, such as Japan Institute for Labour Policy and Training (JILPT) Data Archive and Rikkyo University Data Archive (RUDA). However, these contain only a small number of data sets and do not provide comprehensive information in English. Thus, only the SSJDA, as the largest data archive in Japan, is the focus of this paper. The information mainly comes from webpages and the brochure of the CSRDA.

First, data archives are highly significant for society. For example, data archiving systems are important for (1) the improvement of the reproducibility of studies, (2) the reuse of data through secondary data analysis, (3) placing a reduced burden on respondents and samples, and (4) the improvement of the quality of social surveys and microdata (Sato 2006, Sato and Sato 2006). Topic (1) is the most basic work of data archives in order to retain scientific aspects of social sciences. On topic (2), since the reuse of data indicates effective usages of existing data, this implies the reduction (avoidance) of similar surveys which are meaningless, i.e. researchers do not necessarily need to conduct the same kind of surveys which have already been done. This implies a lower burden on respondents, topic (3), since respondents would not need to answer new surveys of the same kind. Regarding topic (4), since existing microdata is accessible from data archives, researchers would refer to previous surveys before conducting new surveys and would care about ‘comparability’ (e.g. frameworks, questionnaires, and wordings). Similarly, researchers would consider the release of their new data to the archives, and would therefore pay more attention to creating reliable data. The product is higher quality social surveys and microdata. These are some potential reasons why social survey data archives are significant.

3.1 The Center for Social Research and Data Archives (CSRDA)

The CSRDA consists of four divisions, shown below. The Research Infrastructure Division manages the Social Science Japan Data Archive (SSJDA). The SSJDA used to be carried out by the Information Center for Social Science Research on Japan, which was established in 1996. However, in 2009 the Information Center was reorganised and renamed as the CSRDA.

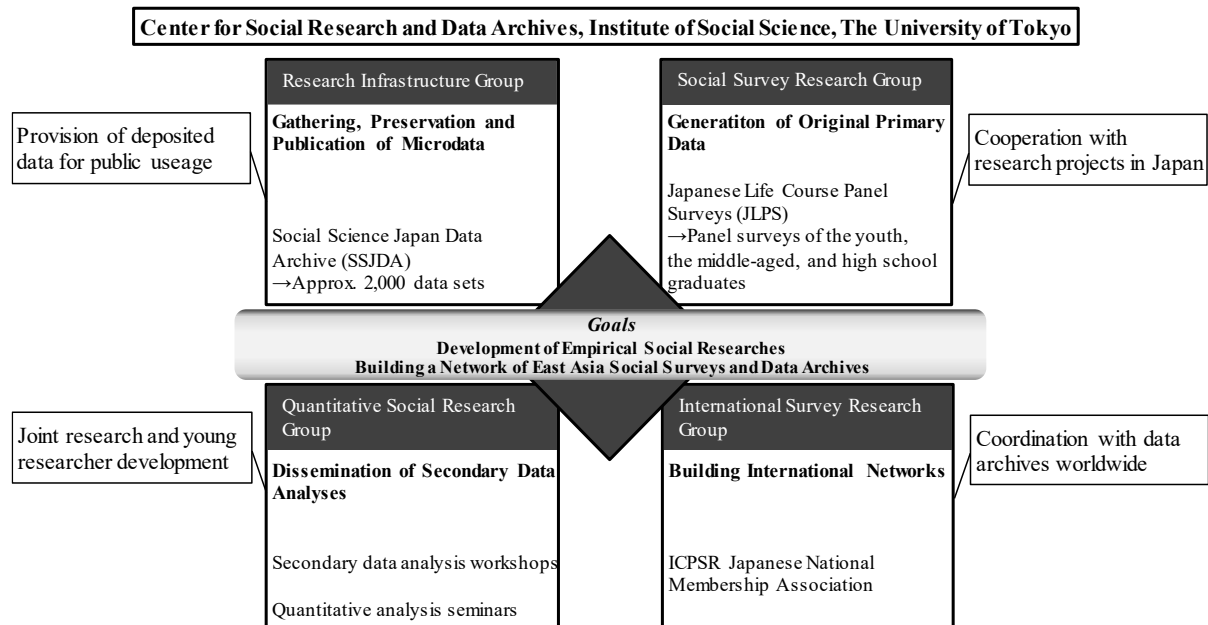


Figure 2: Divisions, duties, and goals of the CSRDA

Source: Brochure of the CSRDA 2015 (online), with minor modifications by the author

The Research Infrastructure Group manages the SSJDA as mentioned above. The Social Survey Research Group engages in activities to build original survey data and make that data publicly available through the SSJDA. The group cooperated in the Japanese General Social Surveys (JGSS) from 1998 through to 2012 and today engages in the Japanese Life Course Panel Surveys (JLPS). The Quantitative Social Research Group holds Secondary Data Analysis Workshops and Quantitative Analysis Seminars every year in order to encourage researchers to make the best use of the survey data maintained in the SSJDA. The International Survey Research Group was established in 2009, and the activities of this group aim at “internationalising” the CSRDA. For instance, the group has been part of the International Federation of Data Organisations (IFDO) since 2001, administering the Inter-university Consortium for Political and Social Research (ICPSR) Japanese National Membership Association as its hub organisation since 1999. It has also held international workshops and seminars related to data archive activities, and has cooperated in international comparative social surveys in East Asia. Please refer to Maeda (2012) for more details on the activities of the CSRDA and SSJDA.

3.2 Data provision structure and the functions of the SSJDA

The SSJDA is the largest social science data archive in Japan, acting as the base of data infrastructure for social sciences in Japan.

The processes of microdata provision by the SSJDA are summarised in Figure 3. First, data depositors store their social survey microdata sets at the SSJDA. After receiving social survey data (microdata and row data) from depositors, the SSJDA verifies the primary analysis tables

submitted, conducts data cleaning if needed, implements adjustment for metadata, and applies anonymisation techniques before releasing the microdata sets. There are two ways to access to published microdata sets in the SSJDA: usage application through the **SSJDA direct download system** webpage or remote analyses through **NESSTAR**. All aggregated data information for preserved microdata sets is available online as reports of descriptive analysis tables; however, this is only in Japanese. The format of microdata sets provided by the SSJDA direct download system is only in STATA data format (.dta) or SPSS data format (.sav) with variable and value labels in Japanese. There is also NESSTAR, which began in January 2014, and provides direct access to several microdata sets through the system. The system allows for cross tabulations, correlation analysis, and even simple regression analysis.

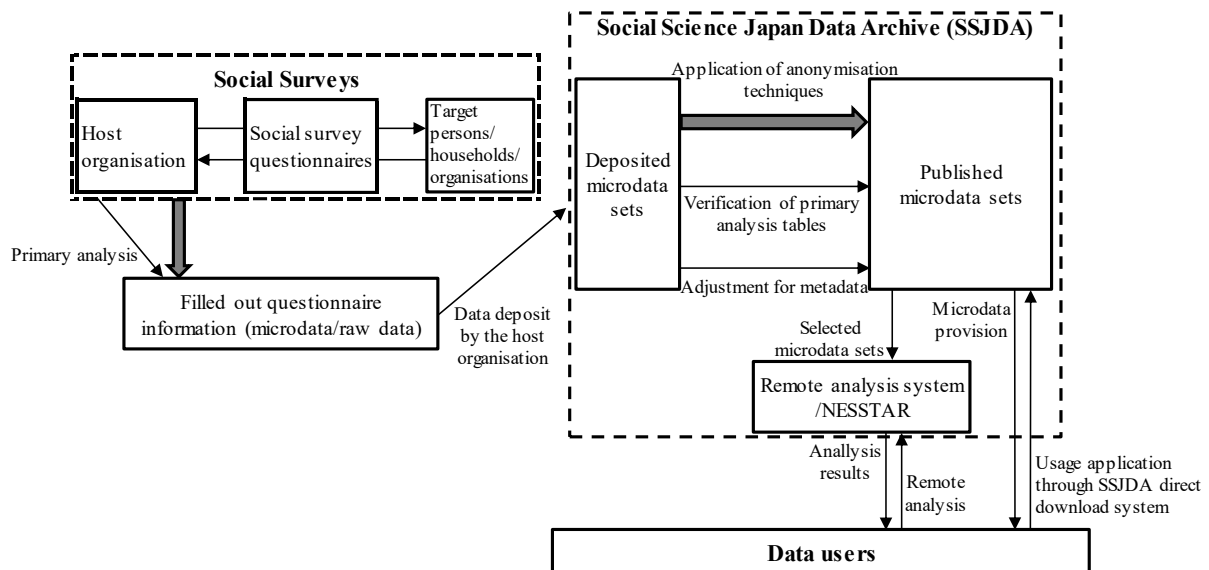


Figure 3: Provision forms and processes of social survey data in the SSJDA

Source: Constructed by the author

Note: Finer arrows indicate proceeding flows and thicker arrows denote technical data transformation

If the data users publish any articles or reports using the data accessed through the SSJDA direct download system or remote analyses through NESSTAR, they are obliged to mention the SSJDA. For data accessed through the SSJDA direct download system, i.e. downloaded microdata sets, two hard copies of any published materials must be sent to the CSRDA. One of the hard copies is distributed to the data depositor and the other is stored in the SSJDA.

Representative microdata sets maintained in the SSJDA are as follows. For longitudinal data sets, there are the Japanese Life Course Panel Surveys (JLPS) conducted by Institute of Social Science, University of Tokyo. For large scale repeated cross-section surveys, the Japanese General Social Surveys (JGSS) maintains a survey aimed at assessing social attitudes and behaviours of people in Japan. While we mainly focus on personal social survey data, business data also exists in the SSJDA.

According to SSJDA operation statistics, in 1998, the first year of the organisation, the number of accessible (preserved) microdata sets was 276, data usage applications was 11, and published articles using the provided data was only 2. Thereafter the usage of the SSJDA advanced, and in 2016 the numbers were 2,018, 1,088, and 191 respectively. In 2016, the usage count of online searching system was 140,206. This highlights how the SSJDA is playing an increasingly important role as a data provider for empirical social sciences in Japan.

3.3 Accessing released microdata sets in the SSJDA

Procedures for accessing SSJDA microdata sets for academic purposes are as follows. Microdata sets can be searched through **Data Search System** online, with various information on each data set, including survey abstracts, questionnaires and reports of descriptive statistics. Only abstracts of the surveys can be found in English, with the rest of the information in Japanese. After finding the desired microdata sets, a usage application must have been completed in the SSJDA direct download system. First-time users require an ID and password by registering to the system (for instructions, please consult the webpage of **New User Registration**). After the registration, data usage applications can be made through the 'My Page' section. Information required includes the social or academic significance of the study that the applicant is planning to conduct using the microdata sets, main variables to be used, expected outcome, and the plan for the publication of the study. After submitting the data usage application, the applicant will receive an e-mail approval notification for data usage. This usually takes two weeks or more. Once this approval is received, all microdata sets can be downloaded from the 'My Page' section. Microdata must be used within one year, and if the applicant makes a publication (papers or reports) using the data, they are required to send two hard copies of the published materials to the CSRDA as mentioned above. For details of conditions of data usage, please consult the **Agreement (Pledge)** webpage. Currently, most of the questionnaires, reports of descriptive statistics and microdata sets (variable and value labels) from the SSJDA are provided only in Japanese.

4 Social scientific longitudinal survey data

4.1 Longitudinal survey

This section presents longitudinal surveys in Japan. All panel surveys presented in this section are based on questionnaire surveys which are conducted every year (or every two years) and track the same persons over time. There are several advantages of panel data analysis compared to cross section data analysis, such as improved inference accuracy and the possibility to study data dynamics. Panel data allows us to analyse individuals' behaviours and decision making. The accumulation of panel data in Japan is slowly advancing.

4.2 Representative panel surveys

This subsection gives information on representative Japanese panel surveys for social sciences provided in English, with some practical information related to microdata sets provision. Most of the descriptions of the surveys rely on webpages of each survey with minor modifications. Basic information on panel surveys is summarised in Table 1. The contents summarised in the table are: topics related to sampling, response rate, and the availability of geographical or prefectural variables, which are important for policy evaluation. For further details of the surveys and rules on usage, please consult the webpages of principal organisations by clicking on each section header below, and contact the organisations if needed.

All panel survey data introduced hereafter is available for research purpose usage, including for non-Japanese speakers. Several principal organisations require exchanges in Japanese and provide questionnaire information only in Japanese, although non-Japanese can access this data with the help of a Japanese speaking research assistant. All microdata sets are either free of charge or have a small handling charge.

This subsection presents abstracts and metadata of representative panel surveys in Japan,

practical information on data provision, hyperlinks to questionnaires of the surveys, and examples of research articles using each panel data set.

■ **Japanese Panel Survey of Consumers (JPSC), 1993-, Institute for Research on Household Economics**

The Japanese Panel Survey of Consumers (JPSC) has been implemented every year since 1993 by the Institute for Research on Household Economics (IRHE). The objective of the survey is to examine changing lifestyles of young women, with a particular focus on income, expenditures, savings, work patterns and family relationships. The initial target group was young women aged between 24 and 34 who were selected from across Japan in October 1, 1993, entitled cohort A. Thereafter fresh samples were added several times. Cohort B, consisting of women aged between 24 and 27, cohort C, consisting of women aged between 24 and 29, cohort D, consisting of women aged between 24 and 28, and cohort E, consisting of women aged between 24 and 28, were added respectively in 1997, 2003, 2008 and 2013. The sample sizes and response rates were 1,500 and 41.4%, 500 and 34.3%, 836 and 28.4%, 636 and 30.5% respectively. For details of the survey, please see Sakaguchi (2014). This data can be procured through the **application form** on the website. For data provision by the IRHE, microdata sets are provided in CSV format on a CD-ROM sent in the post once data usage approval has been received.

⇒ Questionnaire for **the 16th year (2008)**

⇒ Example of a journal article using the JPSC data: Kohara M (2010) The response of Japanese wives' labor supply to husbands' job loss. *Journal of Population Economics* 23: 1133-1149.

Table 1: Abstracts of main longitudinal surveys in social sciences in Japan

Survey title	Time period	Host organisation	Sampling unit	Sampling method	Area	Mode of questionnaire collection	Sample size in the first year	Response rate in the first year	Geographical (prefectural) variable
Japanese Panel Survey of Consumers (JPSC)	1993-	Institute for Research on Household Economics (IRHE)	Females aged 24-34	Stratified random sampling	All over Japan	Home delivery (left and collected)	1,500	41.40%	Available after having one publication using the data
Longitudinal Survey of Newborns in the 21st Century	2001-	Ministry of Health, Labour and Welfare (MHLW)	Newborns in specific time periods	Stratified random sampling	All over Japan	Mailing	47,010	87.76%	Prefectural and municipal variables are in the data
Longitudinal Survey of Adults in the 21st Century	2002-	Ministry of Health, Labour and Welfare (MHLW)	Females and males aged 20-34	Stratified random sampling	All over Japan	Home delivery (left and collected) and mailing	27,893	82.80%	Prefectural and municipal variables are in the data
Longitudinal Survey of Middle-aged and Elderly Persons	2005-	Ministry of Health, Labour and Welfare (MHLW)	Females and males aged 50-59	Stratified random sampling	All over Japan	Home delivery (left and collected) and mailing	34,240	83.76%	Prefectural and municipal variables are in the data
Preference Parameters Study (PPS)	2003-	Institute of Social and Economic Research (ISER), Osaka University	Females and males aged 20-69	Stratified random sampling	All over Japan	Home delivery (left and collected)	1,418	70.90%	Available with an extra usage application
Keio Household Panel Survey (KHPS), Japan Household Panel Survey (JHPS)	2004-2009-	Panel Data Research Center (PDRC), Keio University	Females and males aged 20-69, and aged 20 or more years old	Stratified random sampling	All over Japan	Home delivery (left and collected)	4,005 4,022	KHPS: 41.88% JHPS: 41.68%	Available with an extra usage application
Japanese Life Course Panel Survey (JLPS)	2007-	Institute of Social Science (ISS), University of Tokyo	Females and males aged 20-34, and aged 35-40	Stratified random sampling	All over Japan	Mailing and collected through home delivery	Youth: 3,367 Middle-aged: 1,433	Youth: 34.46% Middle-aged: 40.38%	Available with an extra usage application
Japanese Study of Aging and Retirement (JSTAR)	2007-	Research Institute of Economy, Trade and Industry (RIETI), Hitotsubashi University, and University of Tokyo	Females and males aged 50-75	Stratified random sampling	Five municipalities in the eastern area of Japan (the first year)	Mailing and home delivery (collected) as well as computer-assisted personal interviewing	4,163	59.09%	Available with an extra usage application (although it is required to access the datasets via the RIETI remote control system)
National Family Research of Japan (NFRJ)	2009-2013	Japan Society of Family Sociology and NFRJ Committee	Females and males aged 28-72	Stratified random sampling	All over Japan	Home delivery (left and collected) and mailing	1,879	19.99%	Prefectural variable is in the data

Source: Webpages of each survey and information collected by the author through contacting the principal investigators/organisations

Notes: Response rates are defined as the ratio of number of the respondents for the interview to the target number for interview (excluding those cases where contact could not be made due relocation, long-term absence, ineligibility, unknown address, or other reasons, such as death, illness, or disabilities). The regional variable (variable of 47 prefectures in Japan) of each data set is mostly available with an extra usage application. However, the permission of usage depends on the judgement by each principal organisation

- **Longitudinal Survey of Newborns in the 21st Century, 2001-, Ministry of Health, Labour and Welfare**
- **Longitudinal Survey of Adults in the 21st Century, 2002-, Ministry of Health, Labour and Welfare**
- **Longitudinal Survey of Middle-aged and Elderly Persons, 2005-, Ministry of Health, Labour and Welfare**

The Longitudinal Survey of New-borns in the 21st Century was launched by the Ministry of Health, Labour and Welfare (MHLW) in 2001 and has been conducted every year since. The objective of the survey is to obtain basic information for making, planning and implementing policies for the ministry. The target is all babies born between January 10 and 17, 2001 and between July 10 and 17, 2001 across Japan. Fresh samples were added to the survey in 2010. Respondents were normally parents of the target babies. Contents of the survey include information on households, housing conditions, paternal and maternal labour activities and so forth. The sample size and response rate in the first year were 47,010 and 87.76% respectively.

There is also the Longitudinal Survey of Adults in the 21st Century, which has been implemented every year since its first implementation in 2002 by the MHLW. The sample of the survey is males and females (as well as their spouses) all over the country who were within the age range of 20 to 34 years old at the end of October, 2002. Fresh samples were added to the survey in 2012. Survey topics include type of employment, situation of support system for combination of work and child-care, marriage status, preference on marriage and children, and so forth. The sample size and response rate in the first year were 27,893 and 82.80% respectively.

Finally, there is the Longitudinal Survey of Middle-aged and Elderly Persons, which has been conducted annually since 2005 by the MHLW. The sample is nationally representative, taking data on middle-aged and elderly men and women who were 50-59 years old at the end of October, 2005. The contents of the survey include family situation, health status, housing and household finances etc. The sample size and response rate in the first year were 34,240 and 83.76% respectively.

Since the surveys are sponsored by a national ministry, data usage rules are relatively strict. Practically, the applicant must be involved in a project at a Japanese national university or must be financially supported from a Grant-In-Aid for Scientific Research or a Health and Labour Sciences Research Grant. Also, the applicant must assure the security of the microdata sets provided and the place of data handling on the application form. All data usage procedural details, variable and value information are provided in Japanese, and the data only can be analysed when the applicant is in Japan. The data is free of charge, although the applicant is expected to cover shipping costs. For applicants living outside Japan, microdata sets are provided to the principal organisation of the research project, which should be located in Japan, or to the applicant when they come to Japan.

- ⇒ Questionnaires for the first year (only in Japanese): **New borns, adults, middle aged and elderly**
- ⇒ Example of a journal article using the 21C New-borns data: Griffen AS et al. (2015) Fertility and maternal labor supply in Japan: conflicting policy goals? *Journal of the Japanese and International Economies* 38: 52–72.

- **Preference Parameters Study (PPS), 2003-, Institute of Social and Economic Research at Osaka University**

The Preference Parameters Study (PPS) of the Institute of Social and Economic Research (ISER) at Osaka University has been conducted annually since 2003. The sample is men and women aged between 20-69 years old who were selected from across Japan in 2003. The main aim of the survey is to calculate parameters of preferences defining the utility function, including time preferences, risk aversion, habit formation, and externalities. Fresh samples

were added from respondents to the surveys in 2004, 2006 and 2009 (sizes of added sample were 3,161, 1,396, and 6,181 respectively). The sample size and response rate in the first year were 1,418 and 71.1% respectively. The PPS survey is relatively short compared with other surveys: in the first year, it was only 17 pages including the cover of the questionnaire. That may explain the high response rate for the survey. In addition to the PPS survey in Japan, harmonised surveys are conducted in the US, urban and rural areas of China, and urban and rural areas of India (starting years differ). Please see the [application form](#) webpage for usage details. After receiving data usage approval, the ISER sends the applicant a URL via email providing the data download site, where the data sets are downloadable in CSV and STATA formats.

⇒ Questionnaires for **all currently available years**

⇒ Example of a journal article using the PPS data: Hirano K, Wright JH (2017) Forecasting with model uncertainty: representations and risk reduction. *Econometrica* 85 (2): 617-643.

■ **Keio Household Panel Survey (KHPS), 2004-, Panel Data Research Center at Keio University**
■ **Japan Household Panel Survey (JHPS), 2009-, Panel Data Research Center at Keio University**

The Keio Household Panel Survey (KHPS) has been implemented annually by the Panel Data Research Center (PDRC) at Keio University since 2004. The KHPS is the first nationwide household panel survey where the subjects are selected via individual-based random sampling. The sample is on men and women aged 20-69 years old as of the end of January, 2004. The main objective of the survey is to examine employment behaviours, consumption, income, and housing. The survey covers a wide range of topics, such as employment information, time use, poverty trends, and inter-household transfers of real assets. Fresh samples 1,419 and 1,012 were added in 2007 and 2012 respectively. The sample size and response rate in the first year were 4,005 and 41.88% respectively.

The Japan Household Panel Survey (JHPS) has been conducted annually by the PDRC since 2009 in parallel with the KHPS. The sample is on males and females aged 20 years old or more as of the end of January, 2009. In addition to topics focused on in the KHPS, the JHPS also covers topics such as education and health. The sample size and response rate in the first year were 4,022 and 41.68% respectively. The JHPS data is currently included in the Luxembourg Income Study (LIS) Database as the Japanese representative data. The PDRC has also started a new panel survey entitled Japan Child Panel Survey (JCPS), which has been conducted every year since 2010 as a follow-on survey to the KHPS and JHPS, in order to examine the association between changes in society and child rearing. For usage application, please see the [Acquiring the Data](#) webpage. After receiving data usage approval, microdata sets are available from a special data download site, Data Management System, where the data sets are downloadable in CSV and TXT formats.

⇒ Questionnaires for **all currently available years**

⇒ Example of a journal article using the KHPS data: Seko M et al. (2012) Residential mobility decisions in Japan: effects of housing equity constraints and income shocks under the recourse loan system. *Journal of Real Estate Finance and Economics* 45 (1): 63-87.

■ **Japanese Life Course Panel Survey (JLPS), 2007-, Center for Social Research and Data Archives at University of Tokyo**

The Japanese Life Course Panel Surveys (JLPS) were conducted by the Center for Social Research and Data Archives (CSRDA) at Institute of Social Science, University of Tokyo. The surveys have been conducted annually since 2007 and fresh samples were added in 2012. The

JLPS contains two different surveys: youth and middle-aged group surveys. The sample groups are male and female residents in Japan aged between 20 and 34, and between 35 and 40 as of the end of December 2006, for the youth and middle-aged panels respectively. For the surveys, a letter was sent to the targets in the first year, stating that they would also be requested to participate in follow-up surveys. If they agreed to take part in the first survey but refused to participate in follow up surveys, they were not asked to participate. This may have contributed to relatively low response rates. Another reason may be that younger target groups normally have lower response rates. Sample sizes for the first year were 3,367 and 1,433, and response rates were 34.46% and 40.38% for the youth and middle-aged panels respectively. It should be noted that practically the youth and middle-aged panel data were collected in the same framework, and for that reason the two data can be appended without using any weights.

In addition, the CSRDA has also been conducting two longitudinal surveys entitled Japanese Life Course Panel Survey of the High School Graduates since 2004, and Parent and Child Longitudinal Survey on School Life and Future Prospects since 2015. However, the webpages are not yet provided in English. Microdata sets of the JLPS can be accessed through the SSJDA. After receiving data usage approval, microdata sets are provided in SPSS format and can be downloaded from the webpage ('My Page' section) of the SSJDA direct download system.

⇒ Questionnaires for all currently available years (only in Japanese): **Youth, middle-aged**

⇒ Example of a journal article using the JLPS data: Ishida H (2013) The transition to adulthood among Japanese youths: understanding courtship in Japan. *Annals of the American Academy of Political and Social Science* 646: 86-106.

■ **Japanese Study of Aging and Retirement (JSTAR), 2007-, Research Institute of Economy, Trade and Industry, Hitotsubashi University, and University of Tokyo**

The Japanese Study of Aging and Retirement (JSTAR) is a survey of elderly people, which was jointly launched by the Research Institute of Economy, Trade and Industry (RIETI) and Hitotsubashi University in 2007, with University of Tokyo joining from 2009 onwards. The survey is conducted biennially, and aims to collect longitudinal data on the lives and health of elderly people in Japan. The target is males and females (and their spouses) aged 50 to 75 years old at the first half of 2007. The survey is designed to ensure comparability with preceding surveys such as the Health and Retirement Study (HRS) in the US, the Survey of Health, Aging and Retirement in Europe (SHARE) in continental Europe, and the English Longitudinal Study of Aging (ELSA) in the UK. The JSTAR is therefore a globally comparable Japanese panel survey data of elderly people. The sample size and response rate in the first year were 4,163 and 59.09% respectively. For details of efforts to achieve higher response rates and higher data quality, please refer to Ichimura et al. (2009: 23-25, 37). Procedures on data provision and data formats differ depending on the anonymisation level of the microdata sets. Furthermore, since the surveys are sponsored by a national ministry, data usage rules are relatively strict (please refer to the homepage of JSTAR for details). Applications for usage can be made via the webpage **application form** (please consult also **application form with sample descriptions**). After receiving data usage approval, (anonymised) microdata sets are available from a special data download site of the JSTAR in CSV and STATA formats.

⇒ Questionnaires for **all currently available years**

⇒ Example of a journal article using the JSTAR data: Fujii M et al. (2014) Self-rated health status of Japanese and Europeans in later life: evidence from JSTAR and SHARE. *Japanese Economic Review* 65 (4): 483–498.

■National Family Research of Japan (NFRJ), 2009-2013 (completed survey), Japan Society of Family Sociology and NFRJ Committee

The National Family Research of Japan (NFRJ) was originally launched as repeated cross section surveys since 1999 by the Japan Society of Family Sociology and NFRJ Committee in order to examine family structure in Japan. The official names of the NFRJ surveys, which are on the webpage, are based on the fiscal year and not on the year the survey was conducted. However, for simplicity, in this paper we consider the survey conducted years for the description, as per other surveys. In the NFRJ, many topics relating to family, such as division of labour within the family, attitudes toward traditional beliefs, inter-household transfers and the contents of daily conversations are included. It is a large scale nationwide survey which was implemented in 1999, 2004, and 2009 (every five years). The target group consists of female and male residents in Japan aged between 28 and 77 at the end of December 1998, aged between 28 and 77 at the end of December 2003, and aged between 28 and 72 at the end of December 2008 for 1999, 2004 and 2009 respectively. The sample sizes (response rates) were 6,985 (66.52%), 6,302 (63.0%), and 5,203 (55.35%) respectively.

A letter was sent to respondents of the 2009 survey (cross section survey respondents), asking them to participate in a survey that would involve follow-ups. People agreeing to take part in the cross-section survey but not the follow up were not asked to participate in the panel survey. The panel survey has been implemented annually through 2013 (completed survey). Sample size in the first year was 1,879 and the estimated response rate is 19.99%. The response rate in the first year was comparatively low, although, sample attrition trends in the data can possibly be identified using the observables in the data, i.e. based upon a comparison with all the 5,203 respondents for the 2009 cross-section survey. Retention rates were comparatively high, due to the abovementioned reason, and sample sizes amounted to 1,622 for the second year, 1,555 for the third year, 1,515 for the fourth year, and 1,594 for the last year. Both cross-sectional and longitudinal data of the NFRJ can be accessed through the SSJDA. After receiving data usage approval, microdata sets are provided in SPSS format downloadable from the webpage ('My Page' section) of the SSJDA direct download system.

⇒ Questionnaires for **all currently available years**

⇒ Example of a journal article using the NFRJ-98 data: Wakabayashi M, Horioka CY (2009) Is the eldest son different? The residential choice of siblings in Japan. *Japan and the World Economy* 21 (4): 337-348.

5 Official statistical microdata

This section briefly presents official statistics in Japan, related statistical laws, and accessibility to anonymised microdata sets for research purposes. In April 2009, the revision of the former Statistics Act (No.18 of 1946) was completed and the amended Statistics Act (No.53 of 2007) was fully put into force. For an outline of the Statistics Act, please see the webpage of the **Ministry of Internal Affairs and Communications**. Due to this act, accessibility to official statistics microdata for research purposes has improved. Even with this improvement, microdata availability is limited compared to other countries (for microdata sets availability in Western countries, please see **Appendix B**).

Japanese official statistics have developed along with social and economic transitions (Yoshizoe 2009). For example, considering the trends of Fundamental Statistics (formerly known as Designated Statistics), denoting statistics that receive approval as important statistics for the country by the Minister of Internal Affairs and Communications, we can observe social needs in post-war Japan (Yohizoe 2009: 10). For chronological information on official statistics in post-war Japan, please see **Appendix A**. Historically, the Japanese statistics system has been

functionally-distributed/de-centralised, and as such it was difficult to adjust the overall plan of official statistics across Japan. However, with the recent efforts of researchers and bureaucrats, cooperation amongst ministries and the openness of statistics are becoming central issues.

5.1 Legal framework for academic purpose usage

This subsection briefly explains legal aspects on the usage of official statistics concerned research purposes. Hereafter article numbers denote articles of the Statistics Act.

First, it should be noted that unlike social surveys it is in principle prohibited to provide or utilise the information of official statistical surveys for purposes other than the original purpose of the surveys (article No.40 ‘Restriction on Use of Questionnaire Information, etc.’). Normally, official statistical surveys are conducted only to construct official statistics (primary use). Therefore, the idea of official statistics usage for research purposes is not based on ‘secondary analysis’ as in social surveys, but on ‘secondary use’. Articles from No.33 to No.38 (Chapter III Utilization and Provision of Questionnaire Information, etc.) relate to the secondary use of official statistics. In the case that secondary use of official statistics is in the public interest as equivalent to the production of statistics (article No.33 ‘Provision of Questionnaire Information’) or is recognised as contributing to the development of academic research (article No.35 and No.36 ‘Provision of Anonymized Data’), use of anonymised microdata is permitted (and is considered to be the lending rather than transfer of the anonymised microdata).

Ultimately, one of the most important topics in the Act is the protection of personal information of the respondents. The emphasis is on the confidentiality of the surveys, not the specific protection of personal information. The legal aspects of official statistical microdata provision relating to personal information protection rely on article No.40. The full text of the act is available online in English—please refer to the webpage of the **Ministry of Justice, Japan**. Anonymisation techniques which are applied to official statistics are indispensable for the provision of microdata sets of official statistics (Ito 2015). As for procedures of data anonymisation, the head of an administrative organisation which conducted statistical surveys hears the opinions of the Statistics Commission in advance of the application of anonymisation techniques to microdata (article No.35). There are broadly two data anonymisation methods; one is non-perturbative masking methods, which do not modify the values in the cells, and the other is perturbative masking methods, which do modify the values in the cells. Examples of the former are sampling and resampling, recoding, and top and bottom coding. For the latter, there is swapping, noise addition, and micro-aggregation (for concrete explanations of the anonymisation techniques, please refer to Domingo-Ferrer (2014)). Anonymisation methods applied to Japanese official statistical data are mainly non-perturbative masking methods, although swapping, a perturbative masking method, is applied to the Population Census data as well (Ito 2015: 467). There are no generalised anonymisation methods, and appropriate techniques are selected for each data set.

5.2 Diverse forms of data provision

In this subsection, we introduce the data provision structure of official statistics and present information on accessing anonymised microdata sets.

There are four ways to access to official statistics. Figure 4 shows flows of provision forms of official statistics. Apart from the provision of non-anonymised microdata sets (based on article No.33) and anonymised microdata sets (based on article No.35 and No.36), official statistics can be accessed through ordering tailor-made aggregated statistics tables. Currently 26 survey data sets are available, with a charge from the National Statistics Center (based on article No.34), and through **eStat** online.

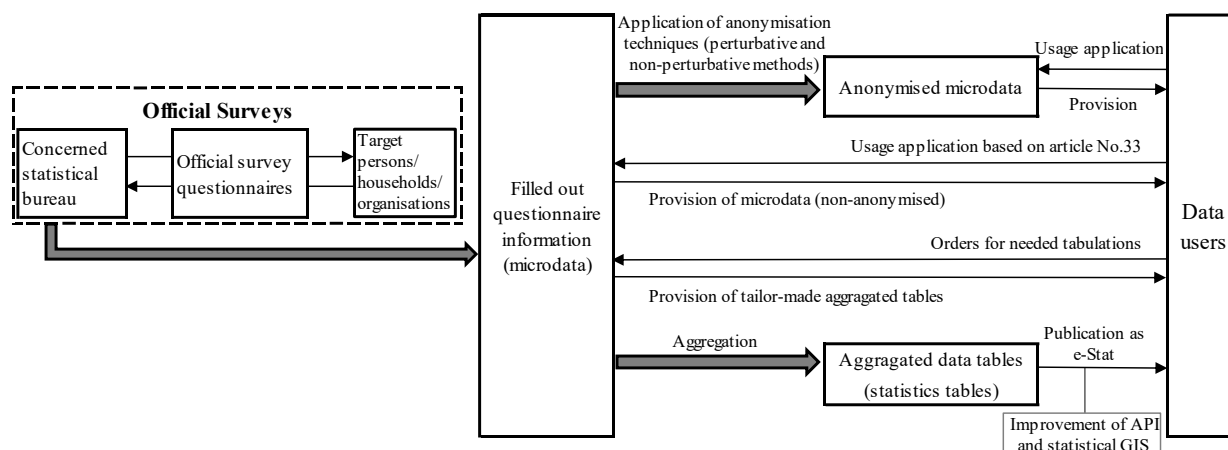


Figure 4: Provision forms and processes of official statistics data in Japan

Source: Ito (2016: 840), with minor modifications by the author

Notes: Finer arrows indicate proceeding flows and thicker arrows denote technical data transformation

Whilst the use of non-anonymised microdata sets is difficult for researchers living outside Japan, the reality is that accessibility of this data is limited even for Japanese-based researchers. Only information on the provision of anonymised-microdata sets are introduced hereinafter. Titles (years) of anonymised microdata sets of official statistics currently accessible are noted below.

- Population Census (2000, 2005)
- National Survey of Family Income and Expenditure (1989, 1994, 1999, 2004)
- Survey on Time Use and Leisure Activities (1991, 1996, 2001, 2006)
- Employment Status Survey (1992, 1997, 2002, 2007)
- Housing and Land Survey (1993, 1998, 2003)
- Labour Force Survey (1989 Jan. - 2012 Dec.)
- Comprehensive Survey of Living Conditions (1998, 2001, 2004, 2007, 2010)

Comprehensive Survey of Living Conditions is a survey implemented by the Ministry of Health, Labour and Welfare (MHLW). The other surveys are conducted by the Ministry of Internal Affairs and Communications (MIC). These MIC surveys can be accessed by making an application to the **National Statistics Center** or one of its satellite organisations noted below (Makita 2012). However, only the Center itself has an English webpage.

- The Research Center for Information and Statistics of Social Science (RCISSS) attached to the Institute of Economic Research (IER), Hitotsubashi University
- Kobe University Microdata Archive (KUMA)
- Microdata Information Provision Unit attached to the Japan Statistics Research Institute (JSRI), Hosei University
- On-site Research Office attached to the Research Organisation of Information and Systems (ROIS), Transdisciplinary Research Integration Center, Institute of Statistical Mathematics

The On-site Research Office attached to the Research Organisation of Information and Systems (ROIS) provides a suitable research environment and data analysis terminals for the sole purpose of researchers who have received usage approval for (non-anonymised) microdata sets of official statistics based upon the article No.33.

5.3 Accessing anonymised microdata sets

Procedures to access the above-mentioned six data sets, apart from Comprehensive Survey of Living Conditions, for academic purposes at the National Statistics Center are as follows. For details please consult the webpage of the **National Statistics Center**.

First, before submitting an application form, the researcher must contact (via e-mail) the center and make clear their desired data sets and explain briefly their research plans. The researcher then needs to send a filled-out application form and a photocopy of their ID to the center via post. For applicants living outside Japan, they must assure the security of the microdata sets and the place of data handling; detailed specifications are required on the application form. The application form is only in Japanese. When the application approval is sent (which may take some time) a paper pledge form is included. At this stage, the applicant has to remit handling charges (approximately JPY10,000 (\approx USD100) per file) and return a signed copy of the pledging document. Microdata sets are provided in CSV format on a CD-ROM and sent to the applicant via mail. For applicants living outside Japan, they are required to visit the center to hold a hearing there; in principle, this should be within one year of receiving the approval for data usage. After using the data, which must be within 3 years, the applicant must return the CD-ROM, delete any copies of the data, and submit a report of microdata usage. The application procedure is complicated and can take at least one month to complete. However, microdata sets are theoretically open to researchers interested in Japan from all over the world.

6 Summary and discussion

In this paper, we presented information of social survey archives, metadata on representative panel surveys, and official statistics in Japan. It is clear that data accessibility is increasing. Data accessibility for non-Japanese researchers is also improving, with increasing amounts of information in English. This is important, since delays in the provision of English information may limit the inclusion of Japan in sophisticated empirical studies based on microdata analysis.

Social survey data and panel survey data, which is mainly collected as social surveys, have steadily accumulated as well. However, more needs to be done to ensure this information is in English. This burden lies not only with the CSRDA, but also with the data archives which do not even provide webpages in English.

Regarding official statistics, there are some challenges to catch up to the level of other countries' statistical systems. The Japanese government appears to be working on overcoming these challenges, although not directly on microdata provision. In particular, "Making Public Data Available to the Private Sector (Open Data)" is included as a goal of the government's plan. As one of the key performance indicators (KPI) for this, "Number of data sets included in catalogues (number of registrations of machine-readable file format data, number of registrations of data in foreign languages, number of accesses, number of downloads" is listed (Strategic Headquarters for the Promotion of an Advanced IT Network Society 2015: 14-16). Currently, on-site safe centers or remote access to microdata sets have not yet been introduced in Japan. This kind of microdata provision still retains the security of data sets and is therefore often utilised in Western countries, such as in UK, Canada (Ito 2016), and Germany (Nakamura 2014). In order to conduct 'trial and error type' analysis, or exploratory research, which is a significant analysis type in social science research, it is very important to improve ways of accessing non-anonymised microdata sets. Currently, only anonymised microdata sets are available at a broad level for analysis (Makita 2012). Ways of providing data need to develop together with the diverse needs of data users (Ito 2016). Improved provision as well as a

statistical data archive are being considered by the Statistical Bureau.

Apart from official statistical data, microlevel administrative rich data sets and personnel management microdata of private firms, are sought to be increasingly accessible for research purposes in order to facilitate bottom-up innovations. Considering methods linking administrative microdata to social survey microdata would also be useful in improving data quality (Meyer et al. 2015: 222). Investigating data provision standards and anonymisation techniques will therefore be prominent topics.

In this paper, background information on social sciences microdata and official statistical microdata infrastructure was provided, in order to facilitate the use of Japanese microdata. There was discussion on the need for microdata sets to be accessible, even for researchers with only a limited knowledge of Japan or Japanese. Improved access for this type of researcher can encourage greater interest and understanding of Japan and its society. In order to conduct meaningful microdata analysis, it is necessary to understand any relevant policies in place and the theory of how the data *should* evolve before investigating how it actually changes. Even for exploratory research or data-driven research, setting up target groups for the policy of interest is essential, which requires an understanding of the policy. Greater microdata provision therefore causes academics to think more about the intent of policymaking, which is certainly a positive (Ichimura 2013: 7). These considerations also relate to open data. It can be argued that ‘just (by) making data open (leaving data to the public), the public will utilise it appropriately to satisfy their needs’. This idea extends to microdata availability. Just by making microdata sets accessible, researchers are able to utilise the data sets from diverse perspectives, bringing benefits for Japan but potentially also globally. Wider availability of Japanese microdata is therefore anticipated.

Acknowledgments

The author would like to give special thanks to Noboru Koshizuka, Shinsuke Ito, Naoki Makita, Hiroki Sato, Kaoru Sato, Dan Sasaki, Yuji Genda, Seung-Gyu Sim, Kieran Hull, Koji Takahashi, Taiyo Fukai, as well as Maria Dolores Morales Varas, who provided extremely helpful comments and discussions. The author would also like to thank participants of the Secondary Data Analysis Workshop 2015 at University of Tokyo. Any errors in this paper are my own.

References

- Acquisti A (2014) The economics and behavioural economics of privacy. In *Privacy, Big Data, and the Public Good: Frameworks for Engagements*. Eds. Lane J et al., Cambridge University Press, 76-95.
- Brandt M (2013) Improvement of access to European microdata. Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality, presentation materials.
- Comparative Perspectives Database (2015) *Harmonized Codebook + Data Dictionary: Comparative Perspectives on Precarious Employment Database*. Online reports.
- Domingo-Ferrer J (2014) Data anonymization: a tutorial. Interdisciplinary Workshop on Data Privacy 2014, presentation materials.
- Einav L, Levin J (2014) Economics in the age of big data. *Science* 346(6210).
- Fujii M, Oshio T, Shimizutani S (2014) Self-rated health status of Japanese and Europeans in later life: evidence from JSTAR and SHARE. *Japanese Economic Review* 65(4): 483-498.
- Griffen AS, Nakamuro M, Inui T (2015) Fertility and maternal labor supply in Japan: conflicting policy goals? *Journal of the Japanese and International Economies* 38: 52-72.
- Heffetz O, Ligett K (2014) Privacy and data-based research. *Journal of Economic Perspectives* 28(2): 75-98.
- Hirano K, Wright JH (2017) Forecasting with model uncertainty: representations and risk reduction. *Econometrica* 85(2): 617-643.
- Ichimura D, Franconi L (2013) Istat experience on releasing multiple microdata files stemming from the same survey. Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality, presentation materials.
- Ichimura H, Shimizutani S, Hashimoto H (2009) Japanese Study of Aging and Retirement: JSTAR first results

- 2009 report. *RIETI Discussion Paper Series* 09-E-047.
- Ishida H (2013) The transition to adulthood among Japanese youths: understanding courtship in Japan. *Annals of the American Academy of Political and Social Science* 646: 86-106.
- Jimenez-Martin S, Labeaga JM (2001) The Spanish system of social surveys. EuReporting Working Paper No. 22.
- Kohara M (2010) The response of Japanese wives' labor supply to husbands' job loss. *Journal of Population Economics* 23: 1133-1149.
- Maeda Y (2012) Data preservation and dissemination at Social Science Japan Data Archive. International Workshop on Social Science Data Preservation and Dissemination in East Asia, presentation materials.
- Meyer BD, Mok KC, Sullivan JX (2015) Household surveys in crisis. *Journal of Economic Perspectives* 28(2): 199-226.
- OECD (2015) *Data Driven Innovation: Big Data for Growth and Well-Being*. OECD Publications.
- Office for National Statistics (1999) *Tracking People: A Guide to Longitudinal Social Sources*. London: Office for National Statistics.
- Peracchi F, Viviano E (2001) The Italian system of social surveys. EuReporting Working Paper No. 21.
- Pallais A (2014) Inefficient hiring in entry-level labor markets. *American Economic Review* 104(11): 3565-3599.
- Sakaguchi N (2014) The Japanese Panel Survey of Consumers. The 20th International Panel Data Conference, presentation materials.
- Seko M, Sumita K, Naoi M (2012) Residential mobility decisions in Japan: effects of housing equity constraints and income shocks under the recourse loan system. *Journal of Real Estate Finance and Economics* 45(1): 63-87.
- Strategic Headquarters for the Promotion of an Advanced IT Network Society (2015) *Declaration to be the World's Most Advanced IT Nation*. Prime Minister of Japan and His Cabinet.
- Wakabayashi M, Horioka CY (2009) Is the eldest son different? The residential choice of siblings in Japan. *Japan and the World Economy* 21(4): 337-348.
- In Japanese**
- Akou T (2016) Dissemination of official statistics usage: effective uses of official statistics data in a big data era. *Statistics* 2016(1): 29-35.
- Ichimura H (2013) Why microdata? *Statistics* 2013 (4): 2-9.
- Ito S (2015) Anonymisation of official statistics: notes on the preparation of the base of personal data usage. *Annual Report of the Institute of Economic Research at Chuo University* 46: 457-478.
- Ito S (2016) Data sharing for official statistics: the current situation and future outlook in Japan. *Journal of Information Processing and Management* 58(11): 836-843.
- Kitamura Y, Kinoshita K (2013) Discussion on methods of encouraging empirical studies in making social sciences a 'true' member of the sciences. *Statistics* 2013(4): 10-16.
- Makita N (2012) The current situation and challenges of microdata usage of official statistics. *Statistics* 2012(3): 9-16.
- Nakamura H (2014) The secondary use of official statistics data through remote systems in European countries: data provision methods which are independent from the ethics of the users. *Advances in Social Research* 12: 38-45.
- Sato H (2006) Applications and challenges for the SSJDA. *Estrela* 147: 2-5.
- Sato T, Sato H (2006) The role of data archives and the current situation of the SSJDA: maintaining and reproducing empirical studies. *Japanese Journal of Labour Studies* 551: 42-54.
- Yoshizoe Y (2009) Prospects and challenges for the effective use of official statistics: statistics as common properties of the nation. *Estrela* 181: 9-17.

Appendix A: Official statistics appendix

Table 2: Designated statistics by ministries in charge under the previous Statistics Act in 2008

	1940s	1950s	1960s	1970s					
Ministry of Internal Affairs and Communications	1 Population Census	1947	30 Labour Force Survey	1950	108 National Survey of Prices	1967	114 Survey on Time Use and Leisure Activities	1976	
	2 Establishment and Enterprise Census of Japan	1947	35 Retail Price Survey	1950					
	14 Housing and Land Survey	1948	56 Family Income and Expenditure Survey	1952					
			57 Unincorporated Enterprise Survey	1952					
			61 Survey of Research and Development	1953					
			76 Fact-finding Survey on Salary of Local Government Employees	1954					
			87 Employment Status Survey	1956					
			97 National Survey of Family Income and Expenditure	1959					
Ministry of Finance								110 Financial Statements Statistics of Corporations by Industry	1970
National Tax Agency									
Ministry of Education, Culture, Sports, Science and Technology	13 School Basic Survey	1948	62 School Teachers Survey	1953					
	15 School Health Survey	1948	83 Social Education Survey	1955					
Ministry of Health, Labour and Welfare	5 Vital Statistics	1947	48 Statistics of Production by Pharmaceutical Industry	1952					
	7 Monthly Labour Survey	1947	65 Survey of Medical Institutions	1953					
			66 Patient Survey	1953					
			94 Basic Survey on Wage Structure	1958					
Ministry of Agriculture, Forestry and Fisheries	26 Census of Agriculture and Forestry	1949	33 Statistics on Milk and Dairy Products	1950					
			37 Statistics on Crops	1950					
			54 Statistics on Marine Fishery Production	1952					
			67 Census of Fisheries	1953					
			69 Statistics on Lumber	1953					
			40 Survey of Mineral Resources Reserves	1950	99 Survey on Motor Vehicle Transport	1960	113 Survey of Selected Service Industries	1973	
Ministry of Economy, Trade and Industry	11 Indices of Industrial Production	1947	43 Current Survey of Production Concerning Gas Industry	1951	103 Survey on Coastwise Vessel Transport	1963			
	23 Census of Commerce	1949	51 Current Survey of Petroleum Products Supply and Demand	1952					
			64 Monthly Report on the Current Survey of Commerce	1953					
Ministry of Land, Infrastructure and Transport	6 Survey on Port and Harbour	1947	32 Statistics on Building Construction Started	1950					
	28 Statistical Survey on Vessels and Seamen	1947	71 Survey on Current Rolling Stock Production	1954					
	29 Survey on Shipbuilding and Engineering	1949	84 Statistics on Construction Works	1955					
			90 Survey on Seamen's Labour	1957					

Continued

	1980s	1990s	2000s
Ministry of Internal Affairs and Communications			<i>122</i> Economic Census for Business Frame 2008
Ministry of Finance			
National Tax Agency			
Ministry of Education, Culture, Sports, Science and Technology			
Ministry of Health, Labour and Welfare	<i>116</i> Comprehensive Survey of Living Conditions 1986		
Ministry of Agriculture, Forestry and Fisheries		<i>119</i> Statistics on Farm Management and Economy 1994	
Ministry of Economy, Trade and Industry	<i>115</i> Current Survey of Energy Consumption in the Selected Industries 1980	<i>118</i> Basic Survey of Japanese Business Structure and Activities 1992	<i>121</i> Corporations Survey on Land and Buildings 1998
Ministry of Land, Infrastructure and Transport			

Source: Yoshizoe (2009: 10; Table 1), names of statistics in English come from the webpage of **Ministry of Internal Affairs and Communications, 'Outline of Japanese Official Statistics (Appendix)'**

Notes: Numbers to the left of survey names represent designated numbers. Numbers to the right of the survey name represent the designated year. Note, the designated number does not exist in the current Statistics Act. As such, official statistics under the former Statistics Act are listed in the table, to be easier to identify. Except for 'Establishment and Enterprise Census of Japan' and 'Statistical Survey on Vessels and Seamen,' all surveys in the table are treated as Fundamental Surveys/Statistics in the amended Statistics Act. Currently 56 statistics (surveys) in total are designated as Fundamental Statistics

Appendix B: Microdata sets availability in Western countries

In Europe, the availability of social scientific microdata sets is much higher than Japan (Brandt 2013). UK has nearly 50 years of history of social scientific longitudinal surveys (for metadata on longitudinal surveys in UK, please see the Office for National Statistics 1999). Southern European countries also have better microdata accessibility than Japan. In Spain, anonymised-microdata sets of official surveys are downloadable for free from the **National Statistics Institute of Spain** webpage. A variety of data sets, such as microdata sets for the Population Census, Time Use Survey, and Labour Force Survey are downloadable in ASCII (.asc) format. Therefore, variable rows have to be read by the user using statistical software (please notice that codebooks are mostly in Spanish). For further information on microdata availability in Spain, including private-sector data archives, please refer to Jimenez-Martin and Labeaga (2001). For Italy, public use files (anonymised-microdata sets) of official surveys are downloadable in TXT, STATA, SAS, and R formats free of charge from the **Italian National Institute of Statistics** webpage (website registration is required). For further information on official statistical microdata availability in Italy and sample characteristics of the microdata provided, please refer to Ichim and Franconi (2013) and Peracchi and Viviano (2001), respectively.

There are also several microdata resources which focus on inter-country comparisons. Comparative Perspective Database (2015) provides a harmonised codebook that summarises information on substantially identical variables in different survey microdata sets which have different variable names. The codebook consists of a subset of variable information from the Household, Income and Labour Dynamics in Australia (HILDA), the Canadian Labour Force Survey (CA LFS), the Survey of Labour and Income Dynamics (SLID), the European Labour Force Survey (EU LFS), the European Community Household Survey (ECHP), the European Union Statistics on Income and Living Conditions (EU-SILC), and the United States Current Population Survey (CPS).

The **Ohio State University User Package for the Cross-National Equivalent File (CNEF)** website provides harmonised longitudinal microdata sets of different countries. Specifically, microdata files on the website “provide a set of constructed variables (for example pre- and post-government income and United States and international household equivalence weights) that are not directly available on the original surveys”. The Cross-National Equivalent Files (CNEF) consists of a subset of variables from the British Household Panel Survey (BHPS), the Household, Income and Labour Dynamics in Australia (HILDA), the Korean Labor and Income Panel Study (KLIPS), the Panel Study of Income Dynamics (PSID), the Russia Longitudinal Monitoring Survey (RLMS), the Swiss Household Panel (SHP), the Survey of Labour and Income Dynamics (SLID), German Socio-Economic Panel (SOEP). The CNEF data is available after application approval (for instructions, please consult the **Access Procedures** webpage).

APPENDICES

Activity Records

Study title:

Labour, Education, Lifestyle and Family in Southern Europe and East Asia

Participants and sub-groups:

Seven participants (three faculties, a researcher, three graduate students)

Two sub-groups (microdata analysis group, microdata availability research group)

Data sets of the CSRDA data archive for the study:

The Japanese Life Course Panel Survey (youth and middle-aged), 2007-2011

Research meetings:

2015.06.30 (at Seminar Room 704, Administration Bureau Bldg.2, U-Tokyo)

2015.08.04 (at Conference Center Room 549, Akamon General Research Bldg., U-Tokyo)

2015.10.02 (Online meeting)

Secondary Data Analysis Workshop:

2016.02.04 (at Conference Center Room 549, Akamon General Research Bldg., U-Tokyo)

CSRDA Secondary Data Analysis Programme 2015 Workshop
Labour, Education, Lifestyle and Family in Southern Europe and East Asia

Date: Thursday, February 4, 2016

Venue: Conference Center Room, 5th floor, Akamon General Research Bldg., The University of Tokyo

13:30-15:00

Session 1: Panel Data Analysis Discussant: Ryuichi Tanaka (University of Tokyo)

Rui Fukuda (University of Tokyo)

“An Overview of Female Labour Supply Trends in Southern Europe and East Asia: Descriptive Analyses of the ECHP, JLPS and KLIPS Data”

Masayuki Kanai (Senshu University)

“Intergenerational Comparisons of the Effect of Family Relations on Life Satisfaction”

Taiyo Fukai (University of Tokyo)

“Who Uses Childcare Centers? Evidence from the JLPS”

15:00-16:30

Session 2: Microdata Availability Discussant: Shinsuke Ito (Chuo University)

Dan Sasaki (University of Tokyo)

“Are Data Public Goods or Intellectual Properties?”

Rui Fukuda (University of Tokyo)

“A Note on Microdata Availability in Japan”

Eunji Hwang (University of Tokyo)

“An Introduction of Korea’s Major Social Surveys: The System and Practice for Enhancing Accessibility”