

全国消費実態調査と国勢調査の照合データを用いた分析

An Analysis Using a Dataset Linked between the Population Census and the National
Survey of Family Income and Expenditure

宇南山 卓
統計研究研修所客員教授
京都大学教授

UNAYAMA Takashi
SRTI Guest Professor
Professor, Kyoto University

令和 7 年 3 月
March 2025

総務省統計研究研修所
Statistical Research and Training Institute (SRTI)
Ministry of Internal Affairs and Communications

受理日：令和 7 年 2 月 17 日

本ペーパーは、総務省統計研究研修所の客員教授である執筆者が、その責任において行った統計研究の成果を取りまとめたものであり、その内容については、統計局または統計研究研修所の見解を表したものではない。本ペーパーの内容については、執筆者に問い合わせ願いたい（Email: unayama.takashi.4x@kyoto-u.ac.jp）。

本研究では、統計法(平成 19 年法律第 53 号)第 32 条の規定に基づき、国勢調査、家計調査及び全国消費実態調査に係る調査票情報を使用した。

全国消費実態調査と国勢調査の照合データを用いた分析¹

宇南山 卓²

概要

本研究では、国勢調査と全国消費実態調査を個票レベルで照合した新たなデータセットを構築し、学歴別の家計収支の状況を明らかにした。より高い学歴を持つことによって本人の収入水準が高くなることはよく知られているが、背後にある世帯の状況まで観察することはこれまで難しかった。家計収支に関する主要な統計である全国消費実態調査に学歴情報を付加することで、世帯全体の収支を観察できる大規模なデータが構築できた。より高い学歴によって得られる高い「学歴プレミアム」として計測すると、本人の収入で約4割、世帯の収入で約3割、消費で約2割となった。学歴プレミアムは生年がより後の若い世代ほど小さくなる傾向があり、特に1970年代に生まれた氷河期世代で小さくなっていた。学歴が高いほど世帯規模が縮小しており、世帯単位の消費の格差は縮小しているが、等価尺度を調整すると依然として学歴プレミアムは大きかった。

キーワード： 国勢調査 全国消費実態調査 データ照合 学歴プレミアム

¹ 研究遂行にあたり、根本和郎氏に多くの支援をいただいた。また、本稿で使用したデータは、総務省統計局との共同研究で構築されたものであり、データ構築にあたり塚田武重氏、佐藤朋彦氏にさまざまな協力をいただいた。ここに記して感謝したい。本研究の一部は科学研究費補助金の資金援助を受けている（基盤研究(A) 23H00053）。

² 京都大学経済研究所教授・統計研究研修所客員教授 Email: unayama.takashi.4x@kyoto-u.ac.jp

An Analysis Using a Dataset Linked between the Population Census and the National Survey of Family Income and Expenditure

UNAYAMA Takashi

Abstract

This study constructs a new dataset by linking individual-level data from the Population Census and the National Survey of Family Income and Expenditure (NSFIE) to reveal household income and expenditure patterns by educational attainment. While it is well known that individuals with higher education tend to have higher income levels, observing the underlying household conditions has been difficult. By adding educational information to the NSFIE, we created a large-scale micro dataset that contains education, income and expenditure. When measured as a "premium," higher education is associated with a 40% increase in individual income, a 30% increase in household income, and a 20% increase in household consumption. The education premium tends to be smaller for younger cohorts born in more recent years and was particularly reduced for the "lost generation," born in the 1970s. Households with higher education levels tended to be smaller in size, which has contributed to narrowing disparities in household consumption. However, when adjusting for equivalence scale, the education premium remained significant.

Keywords: Population Census, National Survey of Family Income and Expenditure, Data Matching, Premium for Higher Education

1. はじめに

本研究では、国勢調査と全国消費実態調査の調査票情報を個票レベルで照合したデータを利用し、学歴別の収入・支出・資産保有の動向を分析する。全国消費実態調査では世帯員の学歴に関する情報は調査していなかったが、そこに国勢調査で調査された学歴の情報を付加することで分析を可能にした。全国消費実態調査は家計調査とともに最も重要な世帯調査であり、収入・支出に関する包括的な統計である。その情報とともに学歴の情報が利用可能であれば、学歴が個人・世帯の経済活動に与える影響を多面的に分析できる。

国勢調査は日本に居住する全ての個人・世帯を調査しており、全国消費実態調査の調査世帯の大部分も調査対象となっているはずである。一方で、日本では国勢調査を含むあらゆる個人・世帯に関する調査において、氏名・住所・電話番号などの個人情報、マイナンバーのような識別子は収集していない。そのため、厳密な個人・世帯固有の情報に基づく照合はできない。それに対し、宇南山(2022)では、世帯員の年齢・性別や住居の状況によって世帯の同一性を確認する方法を提案し、一定程度信頼できる調査票情報の照合方法を確立した。ここでは、宇南山(2022)で構築されたデータのうち、2009年と2014年の全国消費実態調査に国勢調査を照合したデータを活用した。

全国消費実態調査の後継調査である全国家計構造調査では、2019年調査で学歴に関する調査項目が追加され、2024年調査でも継続して調査されている。その意味では、今後は学歴別の収入・消費は国勢調査と照合をせずとも観察可能であり、国勢調査との照合の必要はない。一方で、現時点では公式には2019年の横断面の情報だけが利用可能であるが、ここで2009年からの10年間の変化について分析できており、いち早く時系列的な分析ができた点と2009年・2014年まで遡及した比較を可能にした点、が本研究の貢献となる。

教育を受けることは、経済学的には人的資本への投資とみなすことができ、学歴は収入を規定する重要な個人の属性となる。こうした教育の効果の見方は、Becker(1975)を先駆けとしており、人的資本論とよばれる分野を形成している(赤林, 2012)。現在では、教育経済学の分野でも主流の考え方である(佐野, 2024)。ただし、学歴による収入の違いは、職業選択や消費行動など個人の経済活動全般に大きな影響を与える。しかし、これまで日本の家計に関する学歴別の総合的な分析は十分にされていない。それは、学歴を調査している公的統計に限られるからである。

学歴が個人の収入に与える影響に限定すれば、日本でも多くの研究の蓄積があり「教育のリターン」の計測分野として知られる。教育のリターンの計測は教育歴と個人の収入が同時に観察できるデータがあれば可能であり、たとえば厚生労働省の賃金構造基本統計調査が使われる¹。いわゆる「賃金センサス」とよばれる事業所を対象にした調査であり、雇用している労働者について賃金に加え職種・勤続年数・学歴が調査されている。労働者個人ベースで情報が収集されており、教育のリターンが推計可能である。一方で、事業所側の調査であるため、収入源は給与に限定されており、配偶者の収入など世帯単位での収入は分からず、消費に関する情報もない。

世帯を対象にした調査で学歴を調査しているものとして、総務省の社会生活基本調査が存在するが、こちらでは収入の詳細や消費に関する情報はない。その意味では、家計収支に関する詳細な情報とともに学歴が観察でき、3調査時点10年にわたる変化が分析できる本研究でのデータは貴重なものであり、学歴が経済活動に与える影響を分析する上で望ましいデータである。

¹ 賃金構造基本統計調査を用いて学歴別の所得を分析したものとして、たとえば川口(2011)を参照。

全国消費実態調査は、家計の収入・支出・資産保有について包括的に調査をしているが、いくつかの限界もある。特に問題となるのは、全国消費実態調査において消費の調査をしているのが、9・10・11月の3か月（全国家計構造調査については10・11月の2か月）のみという点である。消費や所得には明白な季節性があり、一部の品目は3か月（もしくは2か月）の調査では消費の全体像が把握できない。また、収入については、「年収・貯蓄等調査票」という調査票を用いて年額を調査しているが、そこでは税や社会保険料などの非消費支出についてはほとんど調査されていない²。そのため、可処分所得の状況も観察できない。

この年間収入・年間消費が観察できないという問題に対しは、宇南山・大野（2017；2018）で、家計調査・家計消費状況調査等を補助情報として活用し、個々の家計ベースで年間可処分所得と年間消費額を推計する手法が提案されている。本研究では、この宇南山・大野（2017；2018）の方法を踏襲し、2009年および2014年の全国消費実態調査、2019年の全国家計構造調査の調査結果をベースに年間の家計収支を推計して用いた。

国勢調査との照合手続きと年間の家計収支の推計がともに可能であった世帯は、おおむね全国消費実態調査の調査世帯の約半数となった。もとのデータと比べれば少ないものの、各調査時点で約2万世帯の観測数が確保できている。公的統計以外であれば家計側から学歴と家計収支を観察できるデータは存在するが、それらと比較しても大規模なものである³。

構築されたデータにおける世帯主の学歴分布は、生年によって水準が異なり、その動向は国勢調査で把握される学歴の分布と概ね一致していた。また、同一生年のグループでは3時点を通じてほぼ一定となっていた。こうした観察は付加された学歴情報が信頼できることを示唆している。

このデータを用いて、まず世帯主自身の年間収入が学歴によってどの程度異なるかを計測した。日本では年功賃金等の影響で年齢が高いほど収入が多い傾向が顕著であるため、年齢階級別に大学・大学院卒（以下、大学以上卒）と非大卒者の年間収入を比較した。その結果、大学以上卒の世帯主は非大卒の世帯主に比べ3割から4割収入が多いことが分かった（この差を、以下では学歴プレミアムとよぶ）。こうした素朴な比較では、潜在的な能力の違いなどが考慮できないため、（同一個人が）大学に進学した場合に得られる所得の増加分としての「教育のリターン」は正しく計測できないことが知られている（佐野，2024）。しかし、先行研究で推計された学歴プレミアムは教育年数1年あたり10%とされており、ここでの結果とおおむね一致している。このことは、ここでの学歴データ・年間収入のデータの妥当性を示している。

ただし、世帯主個人の収入と学歴の関係だけを観察するのであれば、ここで構築されたデータを用いる必然性はない。むしろ、労働時間や勤続年数などのより詳細な就労条件が利用可能でサンプルサイズも大きな賃金構造基本統計調査を用いることの方が望ましい。ここでのデータが重要となるのは、世帯主個人ではなく「世帯の」収入や消費が観察できる点である。そこで、配偶者の収入などを含めた世帯全体の収入についてより詳細な分析をしている。

世帯全体が観察できたことで明らかになった事実として、重要なことは配偶者の年間収入も世帯主の学歴と相関していることである。大学以上卒の世帯主の配偶者の就業率は非大卒と比べて高く、特に正

² 全国家計構造調査では、企業年金掛金（本人負担分のみ）、固定資産税、自動車税・軽自動車税・自動車重量税については調査している。

³ 公的統計以外では、たとえば大阪大学による『くらしと好みに関するアンケート調査』や日本版 General Social Surveys などが使われている（安井・佐野 2009；佐野・安井 2009）。

規雇用の割合が高い傾向がある。結果として、大学以上卒の世帯主の方が、非大卒よりも配偶者の収入が高くなり、学歴プレミアムを拡大させている。生年コーホート別に見れば、世帯主自身の年間収入における学歴プレミアムは縮小傾向にあるが、世帯全体で見れば高い学歴プレミアムが維持されていることが明らかになった。

消費についても、学歴との相関が確認されている。大学以上卒の世帯主を持つ世帯では、非大卒の世帯よりも約 20%程度高い消費水準となっている。消費における学歴プレミアムは、そのまま経済厚生の違いにつながる重要な情報である。収入に比べて小さいものの、学歴によって大きな差が生じていることが確認できた。

消費の時系列的変化をみると、2009 年から 2019 年にかけて全体的に減少していた。特に、高学歴層での消費減少幅が大きく、大学以上卒では 2009 年と比べて 2019 年の年間消費額が約 20%低下していた。この時系列的な変化は、主として世帯人員数の減少が寄与していた。高学歴世帯ほど少子化の影響が大きく、18 歳未満の世帯員が大幅に減少していることが示され、それが世帯の消費を抑制する要因になったと考えられる。実際、世帯規模の影響を調整するため、等価尺度を用いて調整して計算すると、学歴別の動向の差は小さくなり、消費の落ち込み自体も小さい。

こうした世帯主の学歴別に見た収入・消費の違いは、通常の公的統計では観察できないものであり、重要な情報である。一方で、学歴別の家計行動の違いが、ライフサイクル理論などの経済理論と整合的なものであるかは、今後検討する必要がある。

本稿の以下の構成は次のとおりである。まず、第 2 節では、宇南山・大野（2017；2018）に従い、全国消費実態調査・全国家計構造調査マイクロデータの構築方法を概観する。第 3 節では、宇南山（2022）に従い、国勢調査と全国消費実態調査の照合手順を見る。構築されたデータにおける学歴の状況や記述統計を見るのが第 4 節である。第 5 節と第 6 節では、学歴別の家計収支のうち、それぞれ収入・消費について分析している。第 7 節は、まとめと今後の展望を議論している。

2. 全国消費実態調査・全国家計構造調査を用いた年間収入と年間消費の推計

2.1 データとその課題

全国消費実態調査は、1959 年から 5 年に 1 度の頻度で実施されてきた家計収支・資産保有に関する調査である。調査方法・調査内容は家計収支の代表的な統計である家計調査に類似しているが、調査世帯数は 55,000 世帯程度と家計調査の 7 倍の規模である。2019 年調査から全国家計構造調査に名称が変更され調査内容も刷新されている。ここでは、2009 年、2014 年の全国消費実態調査および 2019 年全国家計構造調査を使用する。

全国消費実態調査・全国家計構造調査は、基本的に「世帯票」・「年収・貯蓄等調査票」・「家計簿」の 3 つの調査票が用いられる。全国家計構造調査に移行した 2019 年調査では、一部の世帯は家計簿を記録しない「簡易調査」の対象となっているが、ここでは家計簿までを回答している「基本調査」の対象のみを分析対象とする。また、全国消費実態調査では「耐久財等調査票」によって各種耐久財の保有状況の調査がされていたが、ここでは当該調査票の情報は使っていない。

全国消費実態調査・全国家計構造調査によって家計行動を分析するには、データの制約上の課題をいくつか解決しなければならない。特に重要なのは、家計収支の調査期間が 9・10・11 月の 3 か月（全国家計構造調査については 10・11 月の 2 か月）のみという点である。調査結果からだけでは、年間を通じた家

計収支は把握できない。

消費や所得には明白な季節性があり、消費行動は月によって大きく変動する。たとえば、授業料などの教育費支出は年度をまたぐ3月や4月にピークがあり、年2回のボーナスの時期には耐久財の消費が増加する。しかし、全国消費実態調査・全国家計構造調査では、年度末やボーナスの時期の消費は観察できない。教育費の支出は特定の年齢層に偏っており、それを無視すれば年齢別の消費パターンを的確に把握することはできない。

収入に関しては「年収・貯蓄等調査票」で年間収入が調査されており、季節性の問題は回避できる。しかし、「年収・貯蓄等調査票」では税や社会保険料などの非消費支出についてはほとんど調査されていない。勤労者世帯や無職世帯については、家計簿によって税や社会保険料も把握しているが、一年を通じた金額は把握できず、可処分所得が計算できない。家計行動を規定するのは、税などを控除する前の収入ではなく可処分所得であり、「年収・貯蓄等調査票」だけでは不十分なのである。

2.2 年間の家計収支の把握方法

この全国消費実態調査・全国家計構造調査の家計収支データが持つ問題に対し、宇南山・大野（2017；2018）では、さまざまな補助的な情報を使い、マイクロシミュレーションとよばれる手法を用いて、個別家計ごとの年間収入・年間非消費支出・年間消費支出を推計している。ここでは、その手法を援用することで、全国消費実態調査・全国家計構造調査の家計収支データを補正した。

手法の詳細については、宇南山（2023）第14章を参照することとして、ここでは基本的な考え方のみを紹介する。2009年および2014年のデータについては、宇南山・大野（2018）ですでに分析されており、その結果を利用している。2019年についても、同様の計算を応用している。

まず、年間非消費支出については、マイクロシミュレーションの手法を用いて推計する。マイクロシミュレーションとは、世帯構成や税制などの制度情報を用いて、所得税・社会保険料などの非消費支出を世帯ごとに推計する手法である。税・社会保険料は客観的で一律のルールで決まっており、世帯の置かれた状況を正しく把握できれば正確に推計可能である。

全国消費実態調査・全国家計構造調査では、年間収入を世帯主・配偶者・その他の世帯員それぞれについて別に把握できる。その他の世帯員はさらに65歳以上と65歳未満に分けられている。実際には、各世帯には3人以上の収入のある世帯員がいるケース（世帯票で就業しているとされる世帯員）は多くないため、実質的には個人レベルでの収入が把握可能である⁴。この推計された個人所得に基づき、世帯内の最高所得者を「世帯主」とした。この「世帯主」をベースに、世帯主との続柄、年齢、職業、収入などに関する状況から、配偶者の状況・扶養関係などを推定する。

社会保険料については、加入する社会保険制度を特定する必要があるが、それについては世帯主の雇用形態と所得に応じて加入制度を推定した。さらに、各世帯員の所得と扶養関係に基づき、それぞれの加入制度を推定した。現実には適用される社会保険制度は（特に配偶者について）、労働時間が基準の一つになっている。しかし、2014年調査までは労働時間の情報はないため、代理変数として所得を用いて特定している。たとえば、公的年金保険料の加入制度については、「勤め先からの年間収入」が一定額（「短時間労働者の平均賃金×30時間×52週」など）よりも多い世帯員を厚生年金加入者（第2号被保険者）とみなしている。

⁴ その他の世帯員に複数の就業者が該当した場合、性別・年齢などの情報から収入を按分した。

このように特定された個人所得、扶養関係、加入制度に対し、税制・社会保険料の計算式を適用し、税・社会保険料を計算した。税額を計算する際には所得控除、基礎控除・配偶者（特別）控除・扶養控除・老年者控除・社会保険料控除について考慮した。もちろん、住宅ローン減税や障害者控除のように調査で把握されていない属性で決まる特例的な取扱いが存在しており、推計値は不完全なものである。

一方、年間消費については、利用可能な3か月分（2019年は2か月分）の情報から年間消費額を推計する。そのためには、家計簿に記録された消費データの2つの課題を解決する必要がある。第1に、消費の季節性の問題である。消費には強い季節性があり、調査対象期間の月平均消費額を単純に12倍することでは年間消費額を正しく推計できない。たとえば、年末年始や夏休みなどのイベントに伴う消費、ボーナスの支給時に購入される耐久財などへの支出、授業料や入学金など3・4月に集中する教育費関連支出、などが観察できない。第2に、自由記入の家計簿で支出を記録させることに起因する、支出の記入漏れの問題である。宇南山(2023)第11章でまとめられているように、家計簿での支出調査には、特に高額支出を伴う品目において過小申告の可能性が高い。

これらの問題に対応するために、消費支出を次の3つに分類し、それぞれ別の推計方法を採用した。

- 季節性の小さな品目
- 季節性の大きな品目
- 家計簿によるバイアスが大きい品目
- 教育費

季節性の小さな品目とは、毎月定額を支払うような品目である。このような品目であれば、家計簿によるバイアスの問題は小さく、全国消費実態調査・全国家計構造調査での支出額を信頼することができる。たとえば、家賃・通信量・教養娯楽サービス（インターネット接続料やウェブサービス、新聞購読料など）が該当する。これらの品目については、月平均の支出額を単純に12倍して年間支出額を推計した。

一方、第2の類型は、家計簿によるバイアスの問題はないが、月ごとに支出額が変動する品目である。食費、光熱費など金額的にも大きく一般的な支出である。こうした品目については、家計調査を用いて、計算される年間消費額と9・10・11月の平均消費額の比率（以下、年間換算比率とよぶ）を計算し、全国消費実態調査・全国家計構造調査における月平均支出額にかけて補正した。

ただし、年間換算比率は、収入や年齢階級・居住エリアなどによって異なると考えられる。たとえば、北海道では11月にはすでに暖房費が必要であるのに対し、沖縄の11月は一年でもっとも光熱費がかからない月の一つとなる。そこで、品目ごとに、年間収入・年齢・居住都道府県などの世帯属性ごとのグループを作り、異なる年間換算比率を計算した。

第3の類型は、家計簿での記入漏れが発生しやすい高額消費であり、具体的には自動車の購入や冠婚葬祭に関する支出（結婚式・葬式の費用）などである。こうした高額消費は2つの理由で、一般的な支出に対して適用した年間換算比率での補正が機能しない。第1に、家計簿での記入漏れが発生しやすいことが知られており、家計調査を用いた年間支出額がそもそも過小である可能性がある。第2に、高額消費は支出の発生頻度が低いため、全国消費実態調査・全国家計構造調査では支出がゼロになっている可能性が高い。ゼロは何倍してもゼロであり、年間換算比率をかけても補正とならない。

こうした問題に対応するために、家計消費状況調査を用いて補正した。家計消費状況調査は、自由記入

の家計簿とは異なり、調査品目のリストが与えられるプリコード式の調査票を用いており、高額支出を適切に把握しているとされる。年間を通じて実施される月次の調査であり、年間平均支出額を把握することができる。調査対象は50品目程度であるが、高額支出を中心に選択されており、全国消費実態調査・全国家計構造調査の品目と対応づけが可能な設定であるため、ここでの補正に適した補助情報である。

この家計消費状況調査の2009年、2014年、2019年のデータを用いて、年間換算比率を計算したのと同様に、所得階級・年齢階級・居住地域などによる属性グループごとに年間平均支出額を計算する。家計調査を用いた季節性の調整と異なり、この第3の類型については、全国消費実態調査・全国家計構造調査の月平均支出額の代わりに、グループ別年間平均支出額をそのまま当該グループに属する全世帯の支出額として代入することで補正した。この方法では、同一グループ内別世帯で同一金額が代入されることになり、その意味で全国消費実態調査・全国家計構造調査の情報を捨てることになる。しかし、上でも述べたように、多くの世帯が支出額ゼロであり、比率による補正は困難なため不可避な処理である。

最後に、教育費については、世帯票の情報と家計消費状況調査から推計した。教育費は第3の類型と同じく、家計簿上の支出額がゼロになっているケースが多く存在する。たとえば、3月に1年分の授業料を支払う大学生のいる世帯などでは、9・10・11月の授業料支出はゼロになる。一方で、高額消費などのように、所得階級・年齢階級などで世帯を分類しても負担している教育費は適切に推計できない。そこで、まず世帯票の情報から、世帯員のうち通学中の者の学校の種類（小中高大）および国公立私立の別ごとに人数をカウントし、次に家計消費状況調査から各類型の一人当たりの授業料の年額を計算し、両者をかけて合計したものを年間の教育費とした。

3. 国勢調査と全国消費実態調査の照合

3.1 国勢調査の世帯員情報と統計の照合

本研究の目的の一つは、学歴別の家計行動を明らかにすることである。Becker(1975)以来の人的資本投資の理論でも指摘されているように、より多くの教育を受けることでより高い収入を得ることができる。学歴が高いと相対的には大企業で働く可能性が高くなり、企業内での出世の可能性も高くなり、平均的には所得水準が高くなるイメージである。高い収入は生涯を通じて利用できる「生涯可処分リソース」を増やすため、生涯を通じた平均の消費水準を引き上げる。言い換えれば、学歴が高いほど消費の水準も高くなる。こうした教育の効果が収入・支出両面にどのような影響を与えるかは、経済格差の分析、教育支援の立案、マクロ経済動態の解析、などさまざまな側面で重要な課題と密接に関連する実証的な分析テーマである。

しかし、これまで全国消費実態調査では（類似の統計である家計調査でも）、世帯員の「学歴」については調査がされてこなかった。個人・世帯の属性については、多くの情報が調査されており、さまざまな分析が可能であるが、学歴については忌避感が強いとされてきたこともあり、家計収支調査の中では調査されてこなかったのである。それに対し、後継調査である2019年の全国家計構造調査では、世帯票の調査項目として学歴が追加された。この点は大きな進展であり、2024年の調査でも引き続き調査されており、今後はさまざまな分析が可能となると期待されるが、現状では1回だけの調査結果が利用可能で時系列での比較等はできない。

この問題に対し本研究では、宇南山(2022)で確立された全国消費実態調査の調査世帯と国勢調査の調査世帯を照合するという手法を適用することで、学歴の利用できる年次を過去に遡及して拡張した。国

勢調査のうち西暦の末尾が 0 の年に実施される大規模調査では、各世帯員について学歴が調査される。この情報を全国消費実態調査の所得や消費の情報と一体的に使えば、過去の年次についても学歴別の集計が可能となる。ここでは、宇南山（2022）で照合された 2010 年の国勢調査と 2009 年・2014 年全国消費実態調査を組み合わせたデータを使う。

宇南山（2022）では総務省統計局内の内部情報を活用することで両統計を照合している。その詳細な手法については宇南山（2022）を参照することとして、ここではその概略だけを紹介する。国勢調査は国内の個人に関する全数調査であり、国外との転出・転入がない限り、全国消費実態調査の調査世帯も調査対象になるはずである。そこで、全国消費実態調査の調査対象世帯を国勢調査から見つければ、国勢調査の変数を追加できるのである。

問題は、2 つの統計で共通して利用できる統一の世帯識別子が存在していないことである。国際的に見れば氏名や社会保障番号などで複数の統計を照合できるケースはあるが、日本では個人識別子を調査している事例はない。それに対し、宇南山（2022）では、世帯を構成する世帯員の年齢・性別を用いて同一世帯を識別する手法をとっている。

3.2 照合の手順

データを利用するにあたり、まず宇南山（2022）で示された国勢調査と全国消費実態調査の照合方法について概観する。その基本的な考え方は、世帯員の年齢・性別や住居の状況によって世帯の同一性を確認するというものである。もちろん年齢や性別などの情報は個人固有の情報ではないため、単純にそれらの属性が一致したとしても同一世帯と断言することはできない。しかし、複数の世帯員の情報や住居の情報が全て一致するケースはそれほど多くない。たとえば、「夫婦子供 2 人の 4 人世帯で、夫の年齢が 48 歳、妻の年齢が 47 歳、第 1 子が女で 13 歳、第 2 子が男で 10 歳、80 平米の集合住宅の持家に住む世帯」と特定すれば、それほど多くの該当世帯が存在するわけではない。

とはいえ、5000 万世帯の日本全体で見れば同一の属性を持つ世帯は複数存在するのが一般的である。そこで、さらに地域を限定して同一属性の世帯の存在を確認できれば世帯の同定が可能となる。宇南山（2022）では、総務省統計局内の内部情報を使うことで全国消費実態調査の調査単位区と国勢調査の調査区を対応づけることで、この照合手法を実行可能とした⁵。

全国消費実態調査では国勢調査をサンプリングフレームとしており、近隣する国勢調査の調査区を 2 つ合わせて「単位区」として設定し、その調査区から原則として 12 世帯が抽出され調査対象となっていた。国勢調査の調査区はおおむね 50 世帯が含まれるように設定されており、全国消費実態調査の調査区には 100 世帯程度存在することになる。両統計の調査区を対応づけられれば、国勢調査の調査世帯 100 世帯程度の中から全国消費実態調査の 12 世帯の調査世帯を特定できればよいのである。その範囲であれば、世帯員の年齢・性別、住居の状況程度で同一世帯を識別できる可能性が高い。

実際の照合の手順では、全国消費実態調査と国勢調査では、同じ調査項目でも調査方法が異なるケースに対応が必要である。たとえば、全国消費実態調査では「住居の延べ床面積」を実数で調査しているのに対し、国勢調査では 10 平米の幅を持つ階級値で調査している。両調査で一致しているかを確認するには、全国消費実態調査の数値を国勢調査の分類に合わせて階級値とした。また、国勢調査では生まれ年月が

⁵ 総務省統計局に保存された行政文書であり、全国消費実態調査と国勢調査の対応関係を示したものである。調査世帯の特定を避けるため、一般には公開されない情報である。

調査されるのに対し、全国消費実態調査では満年齢が調査されているため、国勢調査を全国消費実態調査の実施時点での満年齢に変換した。

一般に、情報が多いほど照合はしやすくなるが、時点によって変化する可能性がある変数は使えない。たとえば、世帯員の就学状況や就業状況も両統計で調査されるが、卒業や退職などが発生すれば、異なる時点で調査している2つの統計では一致しない。逆に、属性の変化のために、異なる世帯を同一とみなしてしまう可能性も否定できない。そのため、年齢・性別といった個人にとって不変な変数のみを使用した。もちろん、比較対象の世帯を調査区内に限定しても、世帯員の情報や住居の情報が同一の世帯が存在する可能性はあり、その場合には世帯の照合は不可能とした。

属性が同じでも同一世帯と判断できないケースがある一方で、全ての属性が一致しなくとも同一世帯である可能性もある。たとえば、調査世帯の世帯員の一部が転居したり死亡したりするケースである。たとえば、子供の一人が世帯から離脱しても、残った世帯員の情報が一致すれば「同一世帯」とみなすことができる。この場合、どこまで世帯員の入れ替わりを許容するかによって、同一世帯の判定は異なる。

宇南山(2022)では、こうしたケースに配慮しつつ照合方法を検討している。その結果によれば、全国消費実態調査の調査世帯のうち国勢調査と照合できたケースは、単純に全ての属性が一致する世帯のみを「同一」とみなした場合には約3割、一定の記入誤差や転居の可能性などまでを考慮すれば約5割となったことが報告されている。また、照合できなかった世帯には単身世帯がやや多く、持家世帯の方が照合できる可能性がやや高かったことが指摘されている。一方で、職業や地域による大きな違いはなく、照合できるかどうかには強い偏りがある状況ではなかった。

3.3 世帯の照合と世帯員の照合

宇南山(2022)では基本的に世帯単位での照合が試行されたが、国勢調査を照合する目的は世帯員に関する情報を付加することである。一方で、宇南山(2022)における照合手法では、必ずしも全世帯員が同一であることは「同一世帯とみなす基準」となっていない。そのため、2つの統計の調査世帯が同一と照合されたとしても、世帯員レベルで2つの調査の結果を統合できるわけではない。その意味では、同一世帯であることが確認できるだけでなく、世帯員レベルでどれだけ照合できるかがデータの価値を左右する。

家計収支の分析において、世帯の分類に世帯主の属性を用いることが多い。たとえば、世帯には「年齢」という概念は存在しないが、通常は世帯主の年齢を用いて世帯を分類している。これは、収入や消費がライフステージにより大きく異なるため、その違いを反映するための単純化である。その意味では、2つの統計の照合に際しても、世帯主について情報が付加できるかどうかは重要である。

ただし、世帯主がどの世帯員であるかは調査によって異なる。国勢調査では、世帯主は基本的に各世帯の判断によっており、収入の多少や住民基本台帳の届出等とは必ずしも一致しない。全国消費実態調査では「世帯の家計の主たる収入を得ている人」と定義される。ここでは、家計収支分析により適した全国消費実態調査における世帯主を基準として、世帯員の照合をする。

同様に、全国消費実態調査の世帯主に対応する配偶者についても国勢調査の情報を付加できるかを確認する。もちろん、単身世帯には配偶者が存在せず、親などとの同居の未婚者、ひとり親世帯などにもいない。しかし、配偶者の経済行動は、世帯単位で決定される消費には大きな影響を与えるため、配偶者の状況も確認できるようにした。

4. 構築されたデータ

4.1 年間収支の推定と国勢調査との照合

今回の分析では、宇南山(2022)で構築された、2009年および2014年の全国消費実態調査に2010年の国勢調査の結果を照合したデータを用いている。また、2019年全国家計構造調査では、調査の中で学歴が調査されており同等の情報が使える。そこで、2009年、2014年、2019年の3回の調査を使用する。

Table 1は、第2節で述べたデータの処理をし、第3節で述べた国勢調査との照合ができたサンプルについて、観測数を調査年ごとにまとめたものである。最初の「A. 調査世帯数」とは、データ処理前の元データのサンプルサイズである。全国消費実態調査は集計表によって対象世帯数が変動するが、ここで示したのは2009年および2014年は「総資産集計」の対象である。単身世帯と二人以上の世帯をあわせた観測数は、2009年が52,787世帯、2014年が51,919世帯となっている。全国家計構造調査となった2019年調査については、簡易調査の導入があり家計簿まで記入する世帯は若干減少しており40,290世帯である。ただし、世帯数は同時期の家計調査の調査世帯（特別調査）を追加した分も含む。

この調査世帯に対し、第2節で概観した年間の収支を計算する処理をした後の調査世帯数が「B. 年間収支計算可能世帯数」である。年間の収支を計算するには、世帯員の構成と年間収入の構造が一致している必要がある。たとえば、世帯票では世帯主の配偶者がいないはずの世帯で、年間収入のうち「配偶者の収入」が報告されているケースなどでは、税や社会保険料などを整合的に計算することができない。また、世帯員の一部で年齢が就学状況などと不整合なケースなども非消費支出や消費支出が計算できない原因となる。Aとの差を見ることで、こうしたイレギュラーな世帯の割合が計算できるが、おおむね5%程であり、それほど深刻な問題ではないことが分かる。

この年間収支が計算できる世帯のうち、国勢調査との照合ができた世帯のサンプルサイズを示したのが「C. 国勢調査と照合できた世帯数」である。国勢調査との照合の結果全体は宇南山(2022)で詳細に報告されているが、年間収支の計算は国勢調査との照合とは独立なため、Cの結果は宇南山(2022)での報告数よりやや少ない。最終的に、年間収支が計算できた世帯のうち国勢調査と照合できた世帯の割合は5割前後となった。

Table 1 データの加工とサンプルサイズ

	2009 年	2014 年	2019 年
	全国消費実態調査		全国家計構造調査
A. 調査世帯数	52,787	51,919	40,290
B. 年間収支計算可能世帯数	49,738	48,845	36,191
C. 国勢調査と照合できた世帯数	27,520	22,647	---

(注) 調査世帯数は、2009年と2014年は家計総資産編の調査世帯数、2019年は家計総合集計（全国単身世帯収支実態調査の調査世帯も含む）。

この年間収支が計算可能であり、なおかつ国勢調査と照合ができた世帯のうち、世帯主・配偶者に該当する世帯員が国勢調査で特定できるかを確認したものがTable 2である。全国消費実態調査の世帯主の年

齢と性別と一致する世帯員が国勢調査で唯一に存在していれば、世帯主が特定可能と判断する⁶。第一行目の「A. 合計世帯数」とは Table 1 の C であり、世帯としては同一とみなされた世帯の合計数である。

そのうち全国消費実態調査の世帯主と同一と思われる世帯員が国勢調査の該当世帯内で特定できた世帯の数を示すのが「B. 国勢調査に唯一の該当者」の行である。その比率は約 9 割を超えており、世帯が照合できればおおむね世帯主の情報は利用可能であることが分かる。同一世帯内に世帯主と年齢・性別が一致する世帯員が複数存在するケース（「C. 国勢調査に複数の該当あり」）もあったが、極めて少数であり無視できる。また、照合された全国消費実態調査の世帯主のうち、大部分は国勢調査においても世帯主とされている（B-2）。照合にあたっては世帯員の年齢や性別は用いたが、世帯員間の関係については利用していない。それにもかかわらず、世帯主が世帯主に対応していることは、国勢調査においては世帯主の定義が世帯の選択に任されているが実質的には「主たる収入を得ている者」が世帯主として選択されていることを示唆する。

また、Table 2 では、配偶者がどの程度特定できるかも示している。そもそも「A. 合計世帯数」のうち、（全国消費実態調査の世帯における）配偶者が存在する世帯の数を示したのが「A-1 うち配偶者あり」である。照合された世帯のうち（全国消費実態調査を基準として）配偶者のいる世帯の割合は 8 割程度である。配偶者が存在しないケースの一つの典型が単身世帯となるが、2009 年時点でもその比率は 20%を超えている。それにもかかわらず配偶者のいない世帯が 1-2 割にとどまるのは、ここでは単純な調査世帯数をカウントしているからである。全国消費実態調査では、原則として、単身世帯は調査世帯の 12 分の 1 程度になる。加えて、宇南山(2022)によれば、単身世帯が国勢調査と照合できる可能性はやや低くなる。そのため世帯数で見た単身世帯の割合は 5%程度であり、ここでの結果と矛盾しない。二人以上の世帯であっても親と同居する未婚者やひとり親世帯では配偶者がいないため、配偶者がいない世帯の割合が 1-2 割になるのは妥当な水準である。

Table 2 全国消費実態調査と国勢調査の世帯主と配偶者

	2009		2014	
	世帯主	配偶者	世帯主	配偶者
A.合計世帯数	27,520 (100%)		22,647(100%)	
A-1 うち配偶者あり	23,182(84%)	23,182 (100%)	18,388(81%)	18,388(100%)
B.国勢調査に唯一の該当者	25,901 (94%)	22,461 (97%)	21,143 (93%)	17,779 (97%)
B-2 うち国勢調査の世帯主	24,421 (89%)	---	19,254 (85%)	---
C.国勢調査に複数の該当あり	17 (0%)	5 (0%)	17 (0%)	10 (0%)
D.国勢調査に該当者なし	1,602 (6%)	826 (3%)	1,487 (7%)	818 (3%)
E.世帯主・配偶者特定可能	25,197		20,433	

配偶者が存在することを前提として、全国消費実態調査における世帯主の配偶者と同じ属性を持つ世帯員が特定できるかどうかを示したのが「配偶者」の列である。配偶者のいる世帯「A-1 うち配偶者あり」を 100%として「B. 国勢調査に唯一の該当者」を特定できたケースは、2009 年も 2014 年も 97%と

⁶ 国勢調査との世帯単位での照合では世帯員の年齢に±1 歳の誤差を認めていたため、ここでも同様に年齢に±1 歳の差があっても同一人物とみなす。

なっている。世帯主と同様に複数の世帯員が該当するケースがわずかに存在するが、無視できる。つまり、配偶者のいる世帯では、ほぼ全ての世帯で国勢調査での情報が付加できるのである。

Table 2 の結果は、世帯主と配偶者をそれぞれ別に特定していたが、その両者が特定できるサンプルをカウントしたのが E である。以下では、この E のサンプルをメインの分析対象とする。もとの全国消費実態調査の調査世帯との比率で言えば 4-5 割にとどまっており、サンプル選択の問題は一定の配慮が必要である。たとえば、世帯の照合が正しいとすれば、全国消費実態調査の時点で存在する世帯員が国勢調査の時点で存在しないこと自体に一定の情報が含まれている。国勢調査より時点が前の 2009 年の場合なら、離婚などで配偶者が世帯から脱落したケースが考えられ、逆に 2014 年なら結婚により新たに世帯に加わったケースが考えられる。しかし、E を分析対象とすれば、こうした世帯構成の変化については分析しないことを暗黙に前提としている。

4.2 世帯主の学歴別の分布

世帯主・配偶者まで特定できると、国勢調査の情報を補完した全国消費実態調査の分析が可能となる。ここでは、補完された情報がどのようなものであるかを確認する。Table 3 は、国勢調査が照合された世帯 (Table 2 の E) の世帯主の学歴の分布を示したものである。また、比較のため 2019 年全国家計構造調査の調査結果についても載せている。さらに、学歴分布の妥当性を考察するために、2010 年および 2020 年の国勢調査の結果（男性・卒業者）も示している。

Table 3 照合の結果と世帯主の情報

	2009 年	2014 年	2019 年	2010 年	2020 年
	全国消費実態調査		全国家計構造調査	国勢調査	
世帯主の学歴が観察可能	24,964 (100%)	19,483 (100%)	31,758 (100%)	卒業者 (100%)	卒業者 (100%)
中学校卒	2,736 (12%)	2,068 (10%)	2,879 (8%)	18%	13%
高校卒	11,113 (45%)	9,064 (45%)	13,494 (37%)	45%	43%
短大・高専等卒	2,175 (10%)	1,869 (11%)	4,099 (14%)	8%	9%
大学以上卒	7,940 (33%)	6,482 (34%)	11,286 (41%)	29%	35%
世帯主の学歴不詳など	1,233	950	4,433	---	---
計	25,197	20,433	36,191		

(注) 各学歴のシェア（カッコ内）は、抽出率調整済みである。国勢調査は、15 歳以上の男性「卒業者」のうち卒業した学校が「不詳」を除くシェアである。

学歴分布は 2009 年から 2014 年の期間ではほぼ不変である一方、2019 年には高校卒以下の学歴シェアが大きく低下している。高校卒以上の学歴のシェアが増加するという結果は、過去 50 年ほど継続して大学進学率は上昇してきているという事実と整合的である。一方で、国勢調査全体の結果と比べると、中学校卒の割合がやや低く、大学以上卒の割合がやや高くなっている。

その一つの理由として、ここで観察しているのが「世帯主」であることが挙げられる。世帯主は「主たる収入を得ている者」であるため、たとえば 3 世代同居の世帯では第一世代の高齢者ではなく、相対的に若い第二世代が世帯主になると考えられる。高学歴化が進む環境では若い世代ほど大学以上卒である比率が高く、高齢者を含む 15 歳以上の卒業者全体よりは大学以上卒の比率が高くなるのである。

ここでのデータにおける学歴情報の妥当性を示すために、国勢調査との比較に加え、世帯主の出生年によるコーホートを作成し、コーホート別の大学以上卒の割合を計算したものが Figure 1 である。日本ではいったん労働市場に入ると学歴はほとんど変化しないため、大学を卒業した 20 歳代後半以降の学歴分布は変化しない。大学進学率の上昇のため世代によって大卒者の割合は異なるが、同一生年のコーホートのデータを追跡すれば時点を通じてほぼ一定となると考えられるのである。

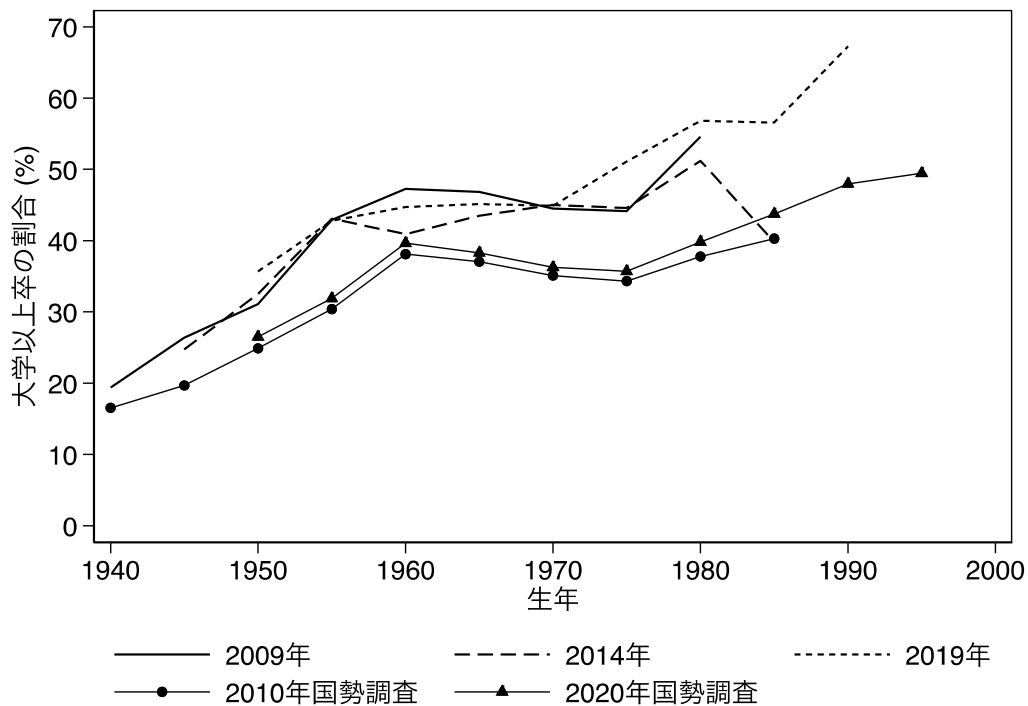
Figure 1 では、出生コーホートごと大学以上卒の割合を 2009 年、2014 年、2019 年の 3 時点それぞれを線グラフとして示している。この各線はおおむね重なり合っており、同じコーホートなら時点によらず同じ大学以上卒の割合となっている。これは、補完された学歴の情報がある程度信頼できることを確認するものであり、2019 年に観察された高学歴化は世帯主の世代交代によるものと推定できる。

一方で、国勢調査の結果との比較では、コーホートをコントロールしても国勢調査がやや低い水準となっている。特に相対的に若い世代になるほど乖離が大きくなっており、5%から 10%の差が生じている。この傾向は 2009 年および 2014 年の国勢調査と照合した年次に限らず、2019 年でも同様の性質が確認できる。言い換えれば、国勢調査との乖離は宇南山(2022)による照合手順によって発生したのではなく、全国消費実態調査・全国家計構造調査の調査世帯が持つ偏りと考えられる。

理由として考えられるのは、たとえば、全国消費実態調査や全国家計構造調査は国勢調査に比べると調査負担の大きな調査であるため相対的に学歴の低い世帯では調査への協力度合いが低くなった可能性がある。特に単身世帯の方が調査への協力度合いが低いことが知られているが、調査への協力が学歴と関係しているのであれば、若年者ほど大学以上卒業者の割合が高い理由にはなる。

こうした点で、日本の学歴分布と比較すると、一定の偏りがある可能性は排除できない。しかし、全体として見ると、学歴の分布はおおむね正確に把握できており、コーホート別に見ても安定的な結果となっている点でも信頼性は十分に高い。学歴不詳のケースを除いても 2014 年で 19,483 世帯、2009 年では 24,964 世帯となっており、学歴の情報がある家計収支のデータとしては公的統計以外を含めても最大規模のサンプルである。2019 年の結果とも組み合わせることで、10 年という比較的長い時系列で比較可能なデータであり利用価値は十分に高いと評価できる。

Figure 1 出生コホート別大学以上卒の比率



4.3 学歴情報と平均の家計収支

構築されたデータの学歴情報には一定の妥当性が確認できたが、元の全国消費実態調査のサンプルのうち 4 割程度になっており、無視できない世帯が分析から除外されることになる。国勢調査との照合ができなかったことによる脱落は、必ずしもランダムなサンプル選択ではないため、一定の偏りを持つ可能性がある。

そこで、家計収支を分析するデータとしての適切さを評価するため、Table 4 では構築されたデータの記述統計を全国消費実態調査・全国家計構造の公式統計等と比較している。最初の列の「A. 公式統計」とは、総務省統計局で公表されている調査結果である⁷。それに対し、宇南山・大野(2017; 2018)による年間収支の計算が可能なサンプルが「B. 年間収支計算可能」であり、世帯員の情報が不完全だったり不整合だったりする世帯を除外したサンプルである。一番右の「C. 世帯主の学歴が利用可能」がここで最終的に構築されたデータであり、Table 3 のうち世帯主の学歴が不詳の世帯を除いたものである。

年間の収支を計算する際に 1 割程度のサンプルが落とされ、国勢調査と照合する際に約半数になることは、ここまでに見てきた通りである。また、2019 年の B と C の差は、世帯主の学歴が利用できるかどうかの差である。全国家計構造調査でも約 1 割程度は学歴を適切に回答していないことが分かる。

まず、各サンプル世帯の構成について比較すると、A と B のサンプルでは世帯人員数から持ち家率まで、おおむね同等の結果となっている。脱落サンプルの比率も高くないため、年間の収支を計算するだけでは大きなサンプル選択の問題は生じていないことが分かる。

⁷ 2019 年については全国家計構造調査の家計総合集計の「家計収支に関する結果」に基づく結果であるが、年間収入のみは「年間収入・資産分布等に関する結果」を用いており厳密には集計対象が異なる。

Table 4 構築されたデータの記述統計

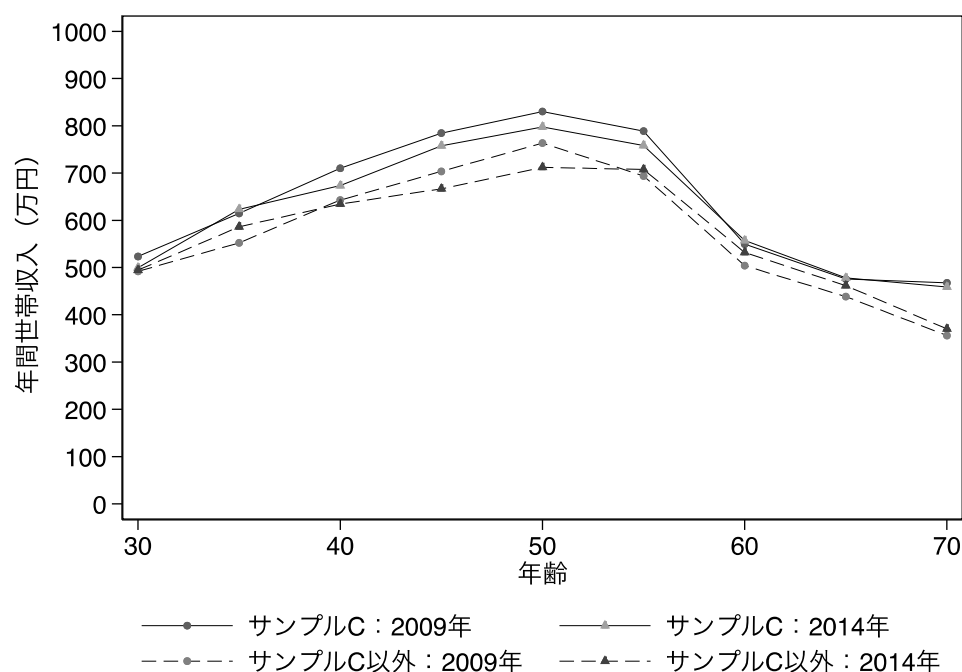
	A. 公式統計			B. 年間収支計算可能			C. 世帯主の学歴が利用可能		
	2009	2014	2019	2009	2014	2019	2009	2014	2019
集計世帯数	55,089	54,208	40,290	49,738	48,845	36,191	23,964	19,483	31,758
世帯人員（人）	2.48	2.39	2.28	2.52	2.38	2.26	2.71	2.54	2.24
18 歳未満人員（人）	0.44	0.41	0.36	0.47	0.40	0.37	0.53	0.42	0.36
65 歳以上人員（人）	0.59	0.67	0.64	0.60	0.67	0.59	0.53	0.68	0.57
有業人員（人）	1.17	1.14	1.20	1.16	1.13	1.21	1.28	1.21	1.21
世帯主の年齢（歳）	55.5	57.7	56.6	56.6	57.7	55.8	54.9	58.6	55.1
持ち家率(現住居)（%）	71.7%	75.6%	68.7%	74.2%	75.5%	68.0%	79.9%	85.9%	66.7%
年間収入（千円）	5,532	5,331	5,568	5,539	5,339	5,622	6,193	5,959	5,629
消費支出（円）	263,439	254,402	237,091	267,001	255,911	238,705	296,054	283,049	237,887

それに対し、B. のサンプルのうち国勢調査との照合ができ世帯主の情報が得られたサンプルに限定した C. は、照合できなかった世帯も含む A. および B. とやや性質が異なる。B と C のサンプルを比較すると、まず平均的には世帯人員数が 0.2 人（1 割程度）多い。2009 年は 18 歳未満が多く 65 歳以上の世帯員が少ない結果になっているものの、主たる違いはその中間である 18 歳から 64 歳までの世帯員の人数である。これは、独立して世帯から離脱する可能性の高い子供や死亡する可能性の高い高齢者を除外した「世帯」を比較した照合手順を使ったためであり、サンプル選択のバイアスと言える。また、持ち家率が高めになるのも、持ち家世帯の方が照合のための情報が多く、国勢調査と照合できる可能性がより高かった結果である。また、持ち家世帯の方が、転居などで照合できなくなる可能性が低いことも理由の一つである。

こうした違いは世帯属性にとどまらず、年間収入や消費支出にも一定の差がある。公式統計との比較が可能な年間収入について見ると、サンプル A と B ではほとんど差がないが、サンプル C は B よりも平均の年間収入が 2009 年も 2014 年も 60 万円程度高い。サンプル C がサンプル B から選択された約半分の世帯であることを考慮すれば、サンプル B のうちサンプル C に含まれる世帯とそれ以外の世帯では、サンプル B とサンプル C の違いの 2 倍の差があることになる。つまり、国勢調査との照合ができ世帯主の学歴が分かる世帯（サンプル C）とそれ以外の世帯では、年間収入に 100 万円程度の差が存在する。

このサンプル B に含まれる世帯のうちサンプル C の世帯とそれ以外の世帯に分けて、年間収入を比較しているのが Figure 2 である。違いの発生原因を明らかにするために、世帯主の年齢別に集計している。この図からもサンプル C の世帯はそれ以外の世帯に比べ年間収入が 100 万円程度高めになっていることが分かるが、さらに違いが年齢によらずに生じていることが分かる。年間収入は世帯主の年齢ごとに大きな差はあるが、世帯主の平均年齢はサンプル C とその他の世帯で大きくは異ならず、年間収入の差は年齢構成で生じているわけではないことが示唆される。

Figure 2 国勢調査との照合と世帯の年間収入



同様に、平均の消費支出も、サンプル B と比較してサンプル C の世帯は月額 2 万 5 千円ほど多くなっている。すなわち、サンプル C とそれ以外の世帯を比較すれば月額 5 万円もの差があることになる。年間収入とも合わせて考えると、国勢調査と照合ができた世帯は、それ以外の世帯に比べ一般に富裕な世帯であることになる。

照合手順として、富裕層ほど国勢調査と照合しやすくなる理由はないと考えられるが、上で見た世帯属性との関係はこの差の一部を説明できる。すなわち、18 歳以上 65 歳未満の「現役」の年齢層が多いほど国勢調査との照合可能性が高くなることから、サンプル C の方がそれ以外の世帯よりも有業人員が多くなっている。有業人員が多いほど、世帯の年間収入が高くなるのは自然である。また、照合できる可能性が高い持ち家世帯は、一般に賃貸住宅の世帯より所得水準が高いと考えられる。ただし、こうした差を調整して比較しても（持ち家世帯同士の比較、有業人員を揃えた比較など）、国勢調査と照合できる世帯の方が収入も消費も多くなる傾向は残った。残念ながら、これ以上の原因については解明できなかった。

こうした点から、理由は明らかではないが、国勢調査との照合によって構築されたここでのデータは、一定のバイアスがあるサンプルである可能性は高い。ただし、もしこの偏りが学歴によらず発生しているなら（学歴グループごとに相対的に収入の多い世帯がサンプルになっているなら）、時点内での学歴別の比較には大きな影響はない。一方で、2019 年では国勢調査との照合に関するサンプル選択の問題はないため、2009 年・2014 年の結果と比較する際には注意が必要である。

5. 学歴別の個人と世帯の収入

5.1 個人の年間収入

この節以降は、構築されたデータを用いて、学歴別の家計収支を比較する。世帯主の学歴の情報が重要となる理由の一つは、学歴別に経済的な豊かさが大きく異なるからである。教育を受けることで、より高度な職業に就くことができるなど、所得獲得能力が高まる。この教育の経済的なメリットの大きさは「教

育のリターン」もしくは「学歴プレミアム」とよばれる。学歴と収入の情報があればその大きさを計測でき、労働経済学や教育経済学で関心の高い変数である。

学歴プレミアムの計測は、再分配政策を考察する際にも重要な意味を持つ。世帯間で大きな収入差が観察されるが、その大きな部分を学歴プレミアムによって説明できる。その意味では、学歴の差は収入差を発生させる主要な要因となる。一方で、学歴プレミアムは教育投資の結果で、必ずしも是正されるべき「格差」とは言えない。再分配政策によって、どのような世帯間の収入差を、どの程度是正するかを考える上で学歴プレミアムの大きさは必須の情報となる。

そこで、前節までに構築された世帯主の学歴の情報を用いて、学歴別の所得の差がどの程度存在するかを明らかにする。まず、単純に学歴別の世帯主個人の年間収入を計算したものが Table 5 である。この表では、Table 3 で学歴が観察できた世帯を対象にした集計であり、前節で述べた国勢調査を照合することで発生すると考えられるバイアスを含む平均であることには注意が必要である。

Table 5 によれば、どの調査年でも学歴が高いほど平均年間収入が高いことが示されている。ただし、高校卒と短大・高専等卒の差はそれほど大きくない。現在では高校進学率はほぼ 100%に近く、特に若年世帯では中学校卒のシェアは小さい。また、世帯主の多くは男性であり、短大に進学することはまれであり、短大・高専等卒のシェアも小さい。

中学校卒と大学以上卒の平均の差が 350 万円なのに対し、標準偏差が各学歴で 250-450 万円ほどとなっている。各グループの観測数は 2000 以上あることから、平均の標準偏差はおおむね 5 万円~10 万円であり、学歴ごとの収入差は統計的にも有意であることは明らかである。一方で、この大きな標準偏差から、個人の年間収入の差のうち学歴によって説明できる部分はそれほど大きくないことが分かる。

また、この Table 5 を横に見ることで、学歴別の平均年間収入の時系列的な推移を見ることができる。この表からは学歴によらず、10-15%ほど年間収入が低下していることが分かる。ただし、この結果が必ずしも「調査時点の収入環境」を表すとは限らない。サンプル選択のバイアスを無視するとして、学歴を固定した収入水準の低下の一部は年齢構成の変化である可能性が高い。特に、この Table 5 では年齢がコントロールされていないため、時点間の比較は困難である。

学歴とともに所得の最も重要な決定要因の一つとなるのが年齢であるが、日本では高齢化が進む中で世帯主の年齢構成も変化している。現役世代については、年功賃金などにより年齢が高いほど収入も増加する傾向があるが、引退をすれば収入は大きく低下する。その意味では、高齢化が世帯主の収入に与える影響は自明ではないが、団塊の世代がほぼ労働市場から退出しつつある 2019 年の時点では高齢化が収入の低下をもたらしている可能性が高い。

この年齢による影響を考慮するため、年齢別・学歴別の世帯主の年間収入をプロットしたものが Figure 3 である。ここでは、大学以上卒とそれ以外（非大卒者）の比較をしていく。2009 年、2014 年、2019 年の 3 時点、学歴別の 6 本の折れ線グラフが描かれている。もし学歴・年齢をコントロールして収入に変化がないのであれば、3 時点の折れ線グラフは重なる。一方で、同じ学歴でも収入が低下していくような状況では、折れ線グラフは下にシフトするのである。

実際、60 歳未満の現役世代に注目すると、大学以上卒では 2009 年から 2014 年にかけて収入が 40 万円前後低下している。一方、2014 年から 2019 年にかけては大きな落ち込みはない。それに対し、非大卒ではむしろ 2014 年から 2019 年にかけて、多くの年齢層で 40 万円程度の落ち込みが観察されている。結果として 2014 年には学歴による収入差は一時的に縮小したが、2019 年には再び拡大したことになる。

Table 5 世帯主の学歴別の世帯主の年間収入（万円）

		2009	2014	2019
中学校卒	平均	296.8	281.0	256.9
	標準偏差	246.2	242.7	269.6
	観察数	2,736	2,068	2,879
高校卒	平均	418.7	393.7	358.2
	標準偏差	286.9	282.6	244.0
	観察数	11,113	9,064	13,494
短大・高専等卒	平均	452.9	427.2	399.7
	標準偏差	276.4	308.2	245.8
	観察数	2,175	1,869	4,099
大学以上卒	平均	650.1	602.9	561.2
	標準偏差	432.7	441.1	389.9
	観察数	7,940	6,482	11,286

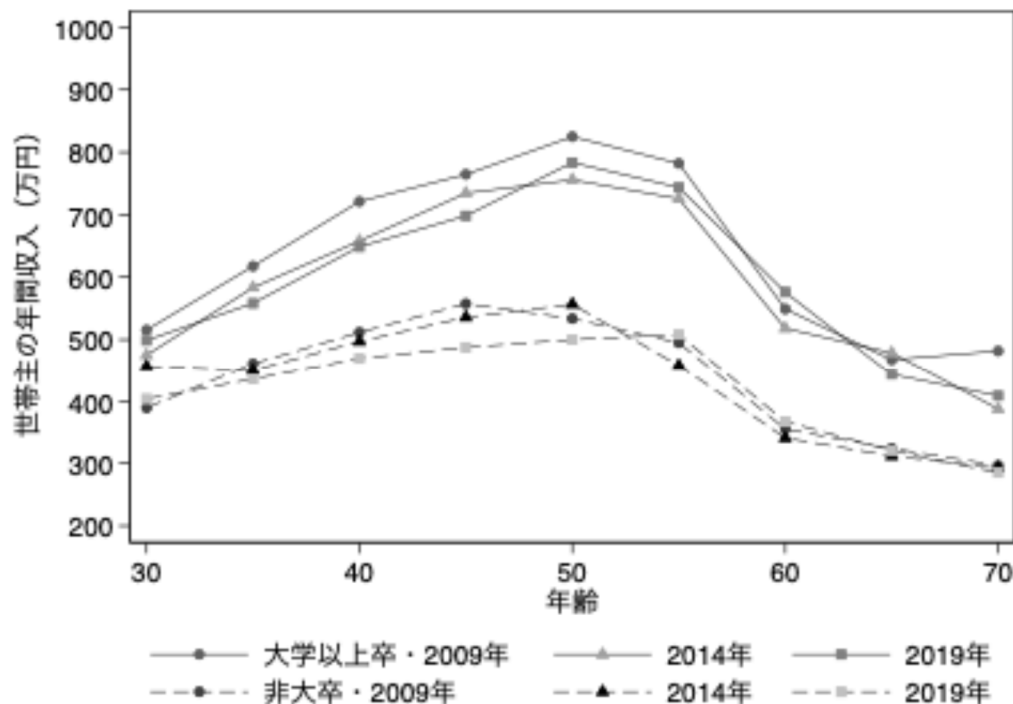
学歴とともに所得の最も重要な決定要因の一つとなるのが年齢であるが、日本では高齢化が進む中で世帯主の年齢構成も変化している。現役世代については、年功賃金などにより年齢が高いほど収入も増加する傾向があるが、引退をすれば収入は大きく低下する。その意味では、高齢化が世帯主の収入に与える影響は自明ではないが、団塊の世代がほぼ労働市場から退出しつつある 2019 年の時点では高齢化が収入の低下をもたらしている可能性が高い。

この年齢による影響を考慮するため、年齢別・学歴別の世帯主の年間収入をプロットしたものが Figure 3 である。ここでは、大学以上卒とそれ以外（非大卒者）の比較をしていく。2009 年、2014 年、2019 年の 3 時点、学歴別の 6 本の折れ線グラフが描かれている。もし学歴・年齢をコントロールして収入に変化がないのであれば、3 時点の折れ線グラフは重なる。一方で、同じ学歴でも収入が低下していくような状況では、折れ線グラフは下にシフトするのである。

実際、まず 60 歳未満の現役世代に注目すると、大学以上卒では 2009 年から 2014 年にかけて収入が 40 万円前後低下していることが分かる。一方で、2014 年から 2019 年にかけては大きな落ち込みはない。それに対し、非大卒ではむしろ 2014 年から 2019 年にかけて、多くの年齢層で 40 万円程度の落ち込みが観察されている。結果として 2014 年には学歴による収入差は一時的に縮小したが、2019 年には再び拡大したことになる。

引退後の 60 歳以上については、60 歳から 70 歳にかけて緩やかに減少しているが、どちらの学歴区分でも 3 時点の比較では大きな変化はない。学歴プレミアムを賃金構造基本統計調査で推計した川口 (2011) では、収入の指標として「賃金」で計測していたため、サンプルは 59 歳以下に限定すべきとしていた。ここでは、個人の公的年金受給なども含めた「世帯主の年間収入」を使っているため、この引退後の学歴プレミアムの状況まで計測できているのである。

Figure 3 年齢別・学歴別の世帯主の年間収入



この年齢をコントロールした学歴別の収入差のデータを用いると、労働経済学の分析における学歴の効果を計測する方法である「ミンサー型賃金関数」の推計ができる⁸。ミンサー型賃金関数とは、「労働者の時間当たり賃金率の自然対数値」を「教育年数と学卒後経過年数である潜在経験年数」の関数で表現したものである（川口，2011）。労働経済学者は潜在経験年数の効果により注目をし、教育経済学者は教育年数に注目するという違いはあるが、分析方法として極めて一般的なものである。

川口（2011）はミンサー型賃金関数を推計するにあたり、サンプルは59歳以下の労働者に限定し、対数賃金率を被説明変数とし、教育については学歴ダミーを用いて、学歴ごとに異なる年齢プロファイルを許容すべきとしている。ここでは、2009年・2014年には労働時間のデータがないこと、賃金以外の収入についても情報があることから、学歴別の世帯主の年間収入を大学以上卒ダミーなどに回帰する。また、学歴の変数については、大学以上卒と非大卒の2カテゴリーとした。年によって学歴プレミアムが変化する可能性を考慮して、年別に別のダミーを用いた。

収入は地域、産業、職種などによって異なるが、それらの要因は学歴との相関がある内生変数である。ここでは、職業選択などの影響も「学歴間での年間収入の差」とみなし、それらの要因をコントロールしない。一方で、景気動向、学歴別の年功賃金、男女間賃金格差をコントロールするため、時点ダミー、性別ダミー、学歴×年齢階級ダミー（5歳刻み）を説明変数に加えている。

全国消費実態調査と全国家計構造調査のデータをプールして、ミンサー型賃金方程式の推計結果を示したのがTable 6である。ここでの結果から、大学以上卒の学歴を持つ世帯主は非大卒の世帯主に比べ25-37%年間収入が高いことがわかる。また、引退世代と現役世代に分けるため年齢を65歳で区切ってみても結果はほとんど変化しない。また、世帯主の年間収入全体を被説明変数としても、世帯主の勤め先

⁸ ミンサー型賃金関数の推計については、その成り立ちから論じた赤林(2012)、日本のデータへの応用を論じた川口(2011)を参照。

の収入を被説明変数としても、全年齢および65歳未満ではほとんど結果は変化しない。すなわち、学歴プレミアムは被雇用者としての収入だけでなく、事業所得などにも存在することが示唆されている。世帯主の年間収入で、65歳未満と比べ65歳以上について学歴プレミアムが大きくなっていることは、公的年金等で学歴による差が大きくなっていることが示唆される。

ここでの結果は、たとえば佐野・安井(2009)で教育年数が1年延びるごとに時間当たり賃金が7-8%上昇するとした結果とおおむね整合的な数字である。労働時間はコントロールされていないが、世帯主を対象としているため、ほぼフルタイムで働く個人が対象となり時間当たり換算しないことの影響は小さいと考えられる。この推計結果は、Figure 3での観察を定量的に評価した自然な拡張であり、得られた結果も興味深いものである。学歴による収入の差が先行研究と一致している点で、ここで構築されたデータの妥当性を示す証拠にもなる。

Table 6 学歴プレミアムの推計

	Log(世帯主の年間収入)			Log(世帯主の勤め先収入)		
	全年齢	65歳未満	65歳以上	全年齢	65歳未満	65歳以上
大学以上卒ダミー 2009年	0.33 (0.02)	0.33 (0.02)	0.37 (0.02)	0.33 (0.02)	0.32 (0.02)	0.52 (0.06)
大学以上卒ダミー 2014年	0.26 (0.02)	0.25 (0.02)	0.31 (0.02)	0.25 (0.03)	0.25 (0.02)	0.38 (0.05)
大学以上卒ダミー 2019年	0.31 (0.02)	0.31 (0.02)	0.33 (0.02)	0.31 (0.02)	0.32 (0.02)	0.33 (0.05)
観測数	75,137	48,555	26,582	48,486	42,572	5,914

(注) 年間収入の対数を、大学以上卒ダミー、大学以上卒ダミー×年齢階級(5歳刻み)ダミー、性別ダミー・年ダミー、で回帰した結果。

そこで、さらに3つの時点が観察できるというデータの利点を活用して、学歴プレミアムが世代によってどのように推移してきたかを分析する。日本ではいったん学校教育を終了して労働市場に入ると追加的な教育を受ける機会はほとんどないため、学歴とは個人固有の属性となる。一方で、大学進学率の高まりによって、大学以上卒の希少性は低下しており、プレミアムも縮小している可能性がある。この学歴ダミーの世代による差を検証するために、大学以上卒ダミーの係数が生年コーホート別に異なることを許容した推計をする。

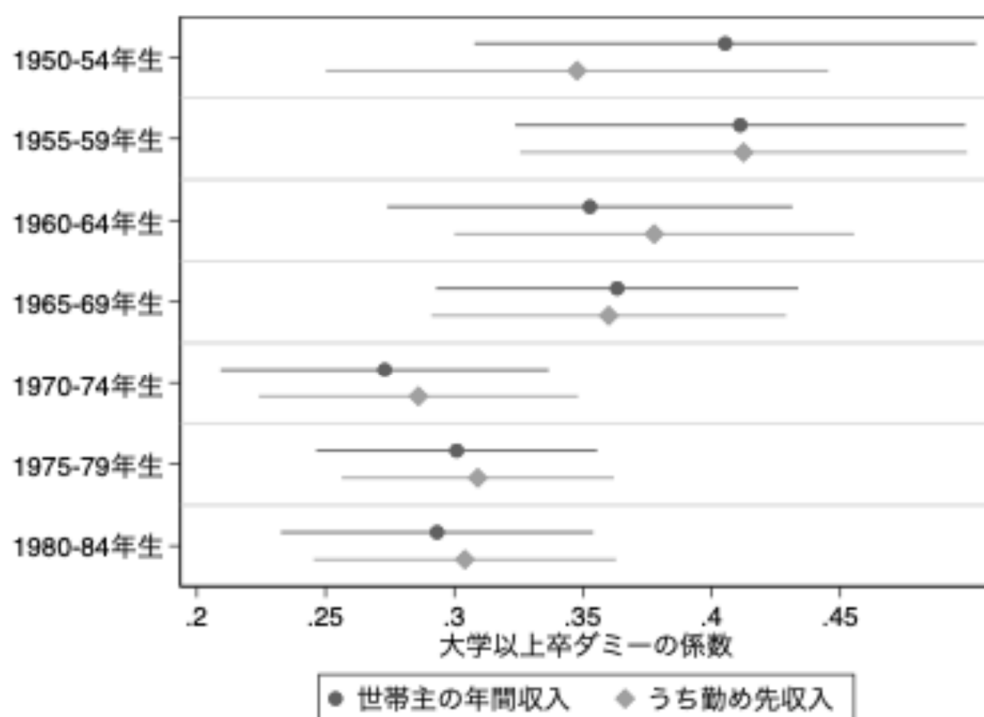
この推計は、識別条件がTable 6と異なるものである。全世代で共通の学歴プレミアムが存在し、それが時点とともに変化することを許容したのがTable 6の推計であった。それに対し、学歴プレミアムが時点を通じて一定であるが、生年コーホートごとに異なることを許容することに変更するのである。日本では新卒採用時点での賃金スケジュールが長期に渡り影響を与えられると考えられるため、おおむね同時期に労働市場に入る同一生年の個人間での賃金差は長期に渡り一定と考えられる。その意味では、時点で可変のプレミアムを許容するより、世代間で可変のプレミアムを許容する方が適切な識別条件となる。

Figure 4は、生年コーホート別の大学以上卒ダミーの係数をプロットしたものであり、世代別の学歴

プレミアムを示したものである。マークの部分が点推定値であり、左右の棒で 95%信頼区間を示している。この推計結果から、学歴プレミアムが 1970 年生の前後で大きく低下してきていることが分かる。学歴プレミアムは、一貫して 30%以上あるが、いわゆる氷河期世代で 5%ほど減少しており、統計的にも有意な落ち込みとなっている。

Figure 1 で見たように、大学以上卒の割合が増加したのは 1950 年から 1960 年に生まれた世代であり、1970 年代に生まれた世代ではほぼ横ばいになっていた。それにもかかわらず学歴プレミアムが低下したのは、大学以上卒の希少性が下がったことよりも、大学卒業時に就職状況が極めて悪かったことが原因と考えられる。1970 年代以降の世代が労働市場に入った 2000 年前後から、大学以上卒の労働者でも非正規雇用となるなど、学歴が十分に活かせない経済状況になったことが影響している⁹。

Figure 4 出生コホート別の学歴プレミアム



5.2 世帯の年間収入

世帯主の学歴が本人の収入に与える影響は、教育が経済的な便益とどのように関係しているかを明らかにする点で重要な観察である。一方で、多くの経済活動は世帯単位でされており、個人の所得だけでは各個人の経済状況が決まるものではない。夫婦について考えれば、世帯主と配偶者間で社会的な分業をして一方が所得を得るケースもあれば、夫婦共働きで両者が収入を得るケースもある。世帯主の収入が低いほど配偶者の収入が高くなるならば、世帯単位での収入差は小さくなる。

公的統計を使い個人と世帯の収入を学歴別に観察することは、2019 年全国家計構造調査が実施されるまで困難であった。詳細な収入に関する情報を調査している世帯調査は、全国消費実態調査以外、家計調査、国民生活基礎調査があるが、これらは学歴については情報を収集していない。事業所側の調査では、

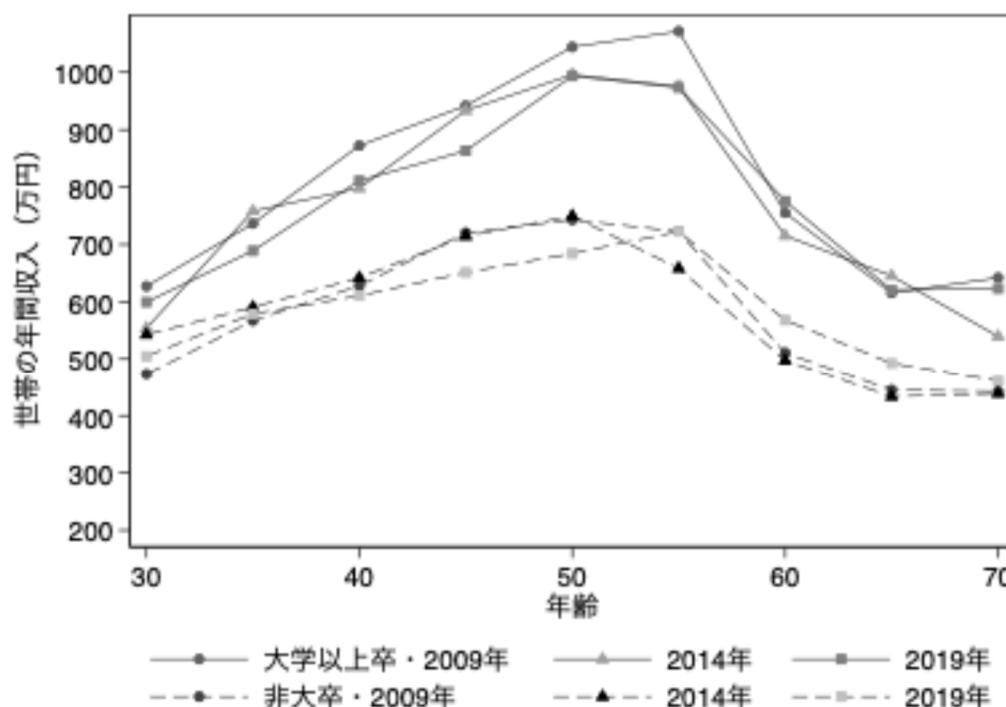
⁹ 氷河期世代の就労については、太田・玄田・近藤(2007)、近藤(2024)などを参照。

賃金構造基本統計調査が学歴と給与を調査しているが、収入の源泉が給与収入に限定されるため個人単位でしか把握できず、世帯に関する情報はない。その意味では、ここで構築されたデータによって3時点10年間について学歴と収入の関係が観察できることは重要な意味を持つ。

世帯の収入に注目するために、まずFigure 3と同様に年齢別の年間収入を示したのがFigure 5である。ただし、世帯主の学歴を使う一方で、世帯主本人の年間収入ではなく世帯全体の合計の年間収入をプロットしたものである。世帯主以外の収入も含まれるため、教育による「人的資本」の増加分だけでなく、配偶者の選択や世帯の構成の選択の影響も含むものである。

世帯主の年間収入は世帯の年間収入の一部であるため、Figure 5の各線はFigure 3よりも上方に位置している。学歴によらず、40歳代から50歳代にかけてより急激に収入が伸びており、世帯主の年間収入と比べても大学以上卒と非大卒との差は若干大きくなっている。

Figure 5 世帯主の年齢別・学歴別の世帯の年間収入



この30歳台から50歳代にかけての伸びは、主として配偶者の収入が寄与している。配偶者の収入は世帯主の収入に次ぐ世帯の重要な収入源であるが、その水準も世帯主の学歴と相関を持つ。ポイントとなるのは、配偶者の収入が、配偶者自身の学歴だけでなく、世帯主の学歴別と相関を持つ点である。世帯主と配偶者は別の個人であり、人的資本論の観点からは世帯主の学歴が影響を与える構造は存在しない。しかし、世帯主の学歴が高いほど配偶者の学歴も高い傾向があるため、両者に相関が生まれると考えられる。

そこで、Table 7では世帯主の学歴別の配偶者の年間収入を示している。配偶者の年間収入は100万円を少し上回る程度で推移しており、世帯主の年間収入の20-30%程度である。世帯主の学歴が高いほど年間収入が高い傾向が見られるが、Table 5で見た世帯主自身の年間収入に比べて差は小さい。ただし、近年になるほど差が大きくなる傾向があり、2009年では中学校卒が104万円に対し大学以上卒で116万円

と 1 割程度しか差がなかったが、2019 年には 112 万円に対し 149 万円と 3 割の差になっている。すなわち、高学歴層ほど配偶者の収入が大きく伸びている。

Table 7 世帯主の学歴別の配偶者の年間収入（万円）

		2009	2014	2019
中学校卒	平均	104.0	108.7	111.6
	標準偏差	111.0	106.6	96.5
	観察数	2,262	1,606	1,779
高校卒	平均	105.3	114.6	128.2
	標準偏差	146.6	139.4	134.3
	観察数	9,535	7,539	9,496
短大・高専等卒	平均	112.4	123.6	152.3
	標準偏差	167.4	166.7	165.1
	観察数	1,753	1,432	2,374
大学以上卒	平均	115.9	125.8	148.7
	標準偏差	187.0	182.3	193.9
	観察数	7,464	6,006	8,553

Table 8 世帯主の学歴別の配偶者の就業率

		2009	2014	2019
中学校卒	正規雇用	7%	5%	6%
	非正規雇用	20%	20%	23%
	計（その他含む）	43%	36%	37%
高校卒	正規雇用	11%	11%	12%
	非正規雇用	27%	28%	31%
	計（その他含む）	48%	47%	51%
短大・高専等卒	正規雇用	14%	15%	21%
	非正規雇用	28%	30%	36%
	計（その他含む）	53%	55%	66%
大学以上卒	正規雇用	14%	14%	19%
	非正規雇用	23%	26%	31%
	計（その他含む）	44%	47%	56%

世帯主の学歴による配偶者の年間収入の違いは、配偶者が就業しているかの違いによって生じている。世帯主の学歴が高いほど配偶者の学歴も高い傾向があると考えられ、より賃金率が高い可能性も考えられるが、就業をしていなくても所得をゼロとみなしているため、就業の有無がより大きな影響を持つ。

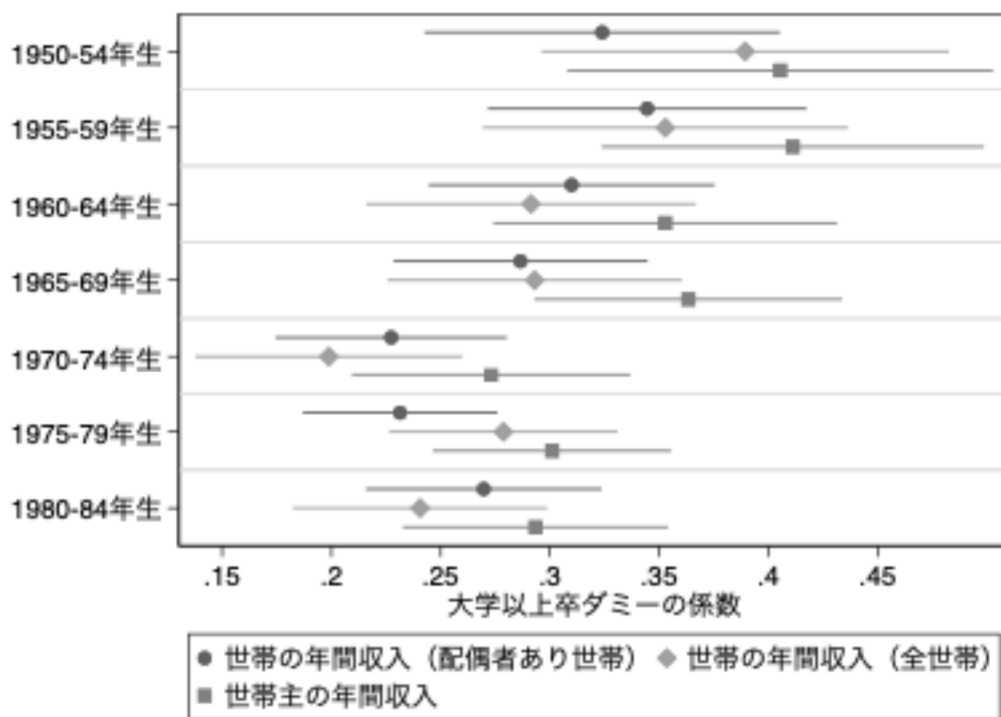
そこで、Table 8 では、世帯主の学歴別に配偶者の就業状態を示した。就業状態は、正規雇用、非正規

雇用（パート・派遣など）、その他就業（自営業・家族従業者など）の3類型に分類した。2009年では、世帯主の学歴別の配偶者の就業率に大きな差はなかったが、2019年には高学歴ほど就業率が高くなる傾向があり、これが配偶者の収入の差を生んでいる。

2009年時点では世帯主が高校卒の方が大学以上卒よりも配偶者の就業率が高いなど就業率そのものの差は小さいが、高学歴な世帯主の配偶者ほど正規雇用の割合が高い。Table 8では示していないが、たとえば中学校卒の場合では13%が家族従業者であるなど、就業をしていたとしても収入が少ないと考えられるケースが多い。これが就業率に差がないにもかかわらず、年間収入に差があった原因である。

こうした配偶者の就業状態の違いは、配偶者の学歴の違いでもたらされていると考えられる。世帯主と同様に配偶者についても高学歴化が進んでいるが、その変化は世帯主の学歴によって異なる。非大卒の世帯主のうち、配偶者が大学以上卒である割合は、2009年に4.1%、2014年に4.6%、2019年に6.6%とほとんど変化していない。それに対し、大学以上卒の世帯主の場合、31.4%、33.5%、39.0%と急速に上昇している。すなわち、高学歴な世帯主は、より高学歴な配偶者を持つことができ、その配偶者はより高い確率で正規雇用として働き、より高い収入を得ているのである。こうした配偶者の学歴分布の変化により、2009年時点では世帯主の学歴によって大きな差がなかった配偶者の収入が2019年には比較的大きくなったのである。この夫婦の学歴の組み合わせの変化は、世帯主の学歴別の「世帯の」年間収入の差を拡大させる方向に働く。

Figure 6 出生コホート別の学歴プレミアム（世帯の年間収入）



配偶者の収入を含めた「世帯の年間収入」でみた場合、学歴プレミアムが生年コホート別にどのように変化してきたかを Figure 4 と同様の推計式で見たのが、Figure 6 である。ここでは配偶者のいる世帯のみに限定し、世帯の年間収入を世帯主の年齢・性別のダミーおよび時点ダミーを入れ、生年×大学以上

卒ダミーで回帰した係数を示している。比較のために世帯主の年間収入を被説明変数とした学歴プレミアムも（配偶者のいる世帯に限定して）計算して掲載している。

世帯主の年間収入については、Figure 4 とほぼ変わらず、1970 年以前世代と以後の世代で学歴プレミアムが大きく異なることが確認できる。一方で、世帯全体の年間収入については、異なる動向を示している。1970 年以前生まれの世帯では、もともと世帯全体の年間収入は世帯主の年間収入より学歴プレミアムは小さい。これは、配偶者が就業している割合は低く、配偶者の学歴の差も小さいために、世帯全体の収入で見れば差が小さかったことを示している。一方で、1970 年以後に生まれた世代では、いったん世帯主の年間収入でのプレミアムの縮小を反映して、世帯の年間収入での学歴プレミアムは縮小した。しかし、1975 年以降生まれでは、世帯の年間収入での学歴プレミアムが急速に拡大傾向にある。これは、夫婦の学歴の組み合わせの変化度によって、配偶者の収入が学歴プレミアムを縮小させる傾向が弱まったことを示唆している。

世代別の学歴プレミアムは世帯単位でも縮小傾向にあるが、世帯主の年間収入に比べて縮小のペースは大幅に小さい。しかも、その源泉も異なり、かつては世帯主自身の年間収入が多いことで得られていた学歴プレミアムが、最近の世代では配偶者の年間収入が高いという形で得られるようになったことを意味する。世帯主の学歴が世帯主自身の収入を引き上げる効果は人的資本の差として理解でき、労働経済学的な分析の対象である。一方、高学歴な世帯主ほど高学歴な配偶者を持つことができ、世帯としてより高い収入が得られることは教育そのものの効果とは言えない。配偶者の収入による収入の違いに注目し再分配政策を評価するのであれば、結婚の意思決定など家族の経済学に基づく考察が必要である。

5.3 学歴別の非消費支出と可処分所得

ここまで、学歴が高いほど収入が多いことを確認してきた。より多くの教育を受けることで経済的なメリットがあるという点で、学歴ダミーで計測される学歴プレミアムは教育投資に対するリターンであり、一面では所得格差を生み出す源泉とも言える。この学歴プレミアムは、累進課税や収入に応じた社会保険料の負担により一部は再分配されることになる。

その再分配の状況については、全国消費実態調査では、税・社会保険料など世帯の自由にならない支出や借入利子などを非消費支出とよび、その金額を基本的に家計簿で調査している。しかし、そのままでは、ボーナスに対する課税などが反映できないため、第2節で述べたように、本研究では年間収入や世帯属性に基づきマイクロシミュレーションの手法で年間の非消費支出額を推計した。また、年間の非消費支出の合計だけでなく、税や社会保険料などの内訳も推計した¹⁰。

推計された税・社会保険料等の負担額を、年別・学歴別に示したのが Table 9 である。合計の非消費支出を見ると、大学以上卒の世帯は高校卒と比べて 1.8 倍程度支払っていることが分かる。世帯全体での収入は、世帯規模の縮小や引退世代の増加などで減少傾向にあったが、非消費支出は平均的にはほぼ横ばいである。これは厚生年金保険料の改定などを反映したもので、非消費支出の負担の増加を表す。

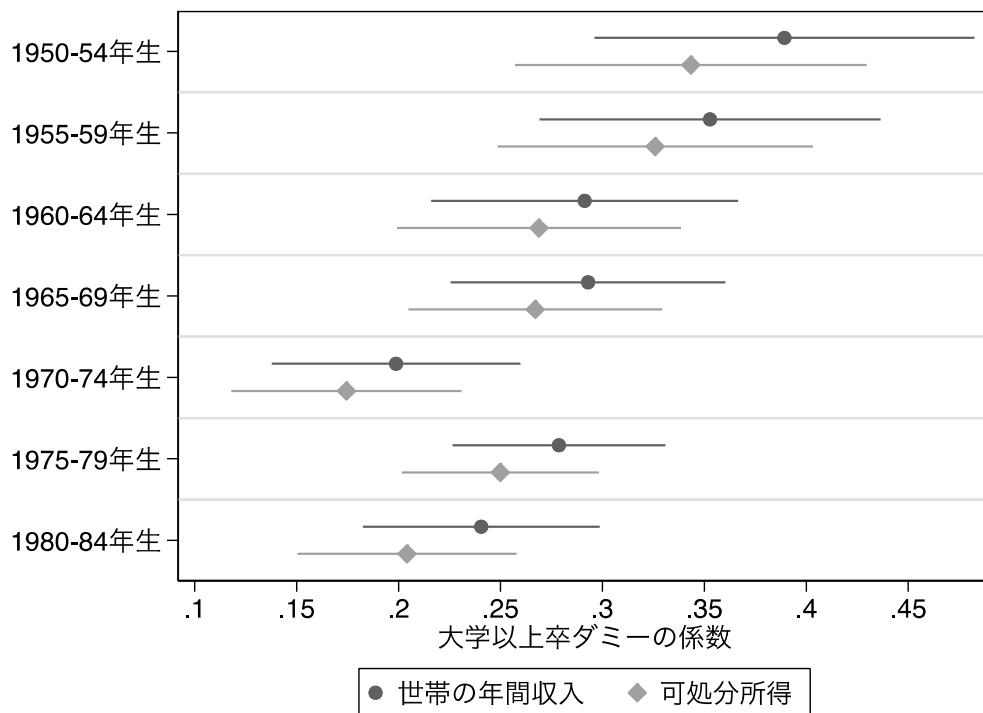
また非消費支出の内訳で見ると、特に累進課税の構造を持つ所得税では、その比率は 2.5 倍から 3 倍近くになっている。ただし、高学歴世帯においては公的年金保険料の負担が特に大きくなっているが、これは高学歴者ほど被雇用者となる可能性が高いことを反映している。

¹⁰ マイクロシミュレーションによる非消費支出の推計の妥当性については、大野他(2015)および多田(2016)を参照。

Table 9 世帯主の学歴別の非消費支出（万円）

		非消費 支出計	所得税	地方税	公的年金 保険料	健康 保険料	介護 保険料	他の非消 費支出
2009	中学卒	57.9	9.2	11.9	12.7	16.5	5.9	1.7
	高校卒	91.0	15.5	19.8	25.4	22.3	5.2	2.7
	短大・高専等卒	100.5	17.2	22.1	30.9	23.0	3.8	3.4
	大学以上卒	162.3	39.0	37.0	45.5	31.4	5.1	4.3
2014	中学卒	60.5	9.3	11.5	12.9	17.8	7.9	1.1
	高校卒	91.6	15.2	18.5	24.7	23.7	7.3	2.1
	短大・高専等卒	103.8	18.3	21.5	30.7	25.0	5.9	2.4
	大学以上卒	163.8	39.0	35.1	43.8	34.7	7.7	3.5
2019	中学卒	59.2	10.4	10.8	12.5	16.1	8.8	0.6
	高校卒	86.5	12.6	17.0	25.0	22.7	7.9	1.4
	短大・高専等卒	101.6	15.4	20.7	34.0	24.6	5.1	1.8
	大学以上卒	153.9	34.0	32.3	45.9	32.4	6.2	3.0

Figure 7 出生コホート別の学歴プレミアム（可処分所得）



世帯の年間収入から非消費支出を差し引き、最終的に残る可処分所得でどれほどのプレミアムを得られるかを推計したものが **Figure 7** である。これまでと同様に、世帯主の生年×世帯主の学歴ダミーを使って推計したものである。世帯主自身の収入で見た **Figure 4** では、推計された学歴プレミアムが 25-40% であった。配偶者のいる世帯に限定すれば世帯全体の年間収入での学歴プレミアムは 20-30% に減少して

いたが、配偶者のいない世帯も含めた全体では学歴プレミアムは依然として 25・40%あった (Figure 6)。それらと比較すると、税・社会保険料による再分配がされた年間可処分所得での学歴プレミアムを比べると 15・35%まで縮小している。教育を受けるかどうかは、この税・社会保険料の負担増を考慮した可処分所得で比較するのが適切である。

6. 学歴別の消費と貯蓄

6.1 学歴別の消費

ここまで学歴別の年間収入の動向を見てきたが、この節では支出側である「消費」について観察する。家計消費に関する公的統計のうち世帯主の学歴の情報を利用できるのは 2019 年の全国家計構造調査だけである。その意味で、時系列で学歴別の消費を観察できるのは本研究の大きな貢献となる。一方で、第 4.3 節で見たように、国勢調査と照合できた世帯は平均的には収入・消費ともに高い水準の世帯である可能性が否定できない。そのため、2009 年・2014 年の結果と 2019 年の結果を比較する際には注意が必要となる。

Table 10 世帯主の学歴別の年間消費（万円）

		2009	2014	2019
年間消費支出 (万円/年)	中学卒	338.6	321.0	285.1
	高校卒	414.8	390.0	337.0
	短大・高専等卒	430.3	405.8	334.2
	大学以上卒	512.8	487.3	384.5
うち帰属家賃 (万円/年)	中学卒	63.4	62.2	55.3
	高校卒	72.6	70.7	61.6
	短大・高専等卒	68.5	75.6	55.9
	大学以上卒	82.8	83.3	61.0
消費支出 × 12 (万円/年)	中学卒	263.5	248.3	221.7
	高校卒	332.0	317.1	268.3
	短大・高専等卒	345.6	321.0	267.5
	大学以上卒	421.1	402.9	318.7

まず、推計された年間消費額を学歴別に見たものが Table 10 である。どの調査年においても学歴が高いほど消費水準が高くなっていることが確認でき、たとえば高校卒と比較して大学以上卒は 2 割程度消費水準が高い。Table 5 で見た世帯主の年間収入では高校卒と比較して大学以上卒は 1.5 倍程度であったのと比較すればその比率は低い。

全国消費実態調査では原則として 9・10・11 月の 3 か月（全国家計構造調査では 10・11 月の 2 か月）の支出が家計簿で調査されるだけである。そこで、第 2 節で述べたように、本論文では宇南山・大野 (2017; 2018) の手法を応用して年間支出額を推計している。年間消費の推計による影響を示すために、Table 10

では調査結果の公式の「消費支出」を単純に 12 倍したのも掲載している。

推計された年間消費は、この公式統計と比較する 50 万円以上高い水準となっている。その違いの最大の要因は、推計された年間消費支出には帰属家賃が含まれていることである。年間消費の推計にあたり、家計簿による調査によって生じるバイアスを調整したり、教育費の支払い時期を考慮したりして推計しているが、それらの調整の効果は少なくとも平均で見れば大きくない。

推計された年間消費支出でも公式の消費支出でも、時系列的には消費水準が低下する傾向がある。2009 年から 2014 年にかけては、どの学歴層でも約 5%の落ち込みとなっていたが、そこから 2014 年から 2019 年にかけては学歴が高いほど落ち込みが大きくなっており、落ち込み幅は高校卒で 15%、大学以上卒では 20%にもなっている。

6.2 消費の変化と世帯規模

この消費水準の低下の大部分は、世帯の規模の縮小によるものと考えられる。消費は世帯単位で計測されるため、世帯人員が多ければ消費額が多いのは当然である。日本では平均の世帯人員数は低下傾向にあり、消費もそれに合わせて縮小すると考えられる。

世帯規模の縮小の効果を調整するために、学歴別の世帯人員数を示したものが Table 11 である。平均の世帯人員の低下傾向はすでに Table 4 で示したが、それを学歴別に見たものである。「A. 世帯人員数」は単純な世帯人員数の平均である。2009 年時点では学歴が高いほど世帯人員が多く、高校卒に比べて大学以上卒は 1 割ほど世帯人員が多かった。それに対し、2019 年になると、学歴による世帯人員数の差はほとんどなくなり、高校卒の世帯人員数がむしろ多くなっている。

この世帯全体の人員数の内訳として 18 歳未満の世帯員数を示したのが「B. 世帯人員数うち 18 歳未満」である。世帯人員数の学歴別の動向の差の約半分が、この 18 歳未満の世帯人員数で説明できる。もともと学歴が高いほど 18 歳未満の人員が多くその傾向は 2019 年でも維持されるが、その差は大幅に小さくなっている。18 歳未満の世帯人員数を子供の数と解釈をすれば、高学歴世帯ほど少子化が急速に進んだとみなせる。

こうした世帯人員数の変化が消費に与える影響をコントロールするために、世帯規模を示す指標として等価尺度を計算したのが「C. 等価尺度」である。等価尺度とは、世帯の規模の経済（世帯の人数が多いほど一人あたりの消費が少なくても同じ厚生水準が得られる効果）を考慮して世帯の規模を示す指標である。概念的には大人 1 人が単独世帯の何人分の消費が必要かを示すものとされている。計算方法にはさまざまな手法が存在しているが、ここでは OECD で使われる等価尺度に準じた方法を使った。具体的には、世帯主を 1 人として、追加の 18 歳以上の世帯員 1 人を 0.7 人、18 歳未満の世帯員を 0.5 人として計算する方法である¹¹。たとえば、夫婦と（18 歳未満の）子供 2 人の世帯であれば、大人一人暮らし換算 2.7 人分（ $=1+0.7+2*0.5$ ）とカウントされる。等価尺度の考え方とは、上記の 4 人家族が大人一人暮らしの 2.7 倍の消費をすると同程度の生活水準が確保できると考えることであり、異なる世帯構成でも消費水準が比較できるのである。

計算された等価尺度は、世帯人員数よりも低下幅はやや小さくなっているものの、全体として減少傾向であり、特に高学歴層で大きく落ち込んでいる点では共通している。この等価尺度を用いて、年間消費支出を「大人 1 人あたり」に換算したものが「D. 大人 1 人換算年間消費支出」である。この指標を

¹¹ OECD の等価尺度では大人と子供を 14 歳以上と未満で区別しているが、ここでは 18 歳以上・未満を境界とした。

使えば、学歴別および時系列的な世帯人員の違いを考慮した上で消費水準が比較できるのである。

この指標で見ると、学歴が高いほど消費が多く、大学以上卒は高校卒より 15%ほど消費水準が高い。**Table 5** でみた世帯主の年間収入が高校卒 419 万円に対し大学以上卒が 650 万円であったことを考慮すると、収入に比べ消費の差は小さいことが分かる。落ち込み幅は大きく縮小している。この等価尺度調整済みの年間消費支出で見ると、むしろ学歴が低いほど消費の落ち込みは大きく、2009 年から 2019 年にかけて、中学卒は 12%消費が低下しているのに対し、大学以上卒の落ち込みは 8%である。

Table 11 世帯主の学歴別の世帯人員と消費

		2009	2014	2019
A. 世帯人員 (人)	中学卒	2.33	2.21	2.18
	高校卒	2.64	2.45	2.30
	短大・高専等卒	2.66	2.48	2.13
	大学以上卒	2.93	2.77	2.24
B. 世帯人員 うち 18 歳未満	中学卒	.20	.14	.18
	高校卒	.44	.33	.32
	短大・高専等卒	.67	.53	.41
	大学以上卒	.71	.57	.42
C. 等価尺度 (大人・人)	中学卒	1.89	1.82	1.79
	高校卒	2.06	1.95	1.85
	短大・高専等卒	2.03	1.93	1.71
	大学以上卒	2.21	2.12	1.78
D. 大人 1 人換算 年間消費支出 (万円/年)	中学卒	191.1	188.8	167.9
	高校卒	217.8	218.7	195.3
	短大・高専等卒	236.6	233.3	211.9
	大学以上卒	247.4	244.5	227.7

この学歴別の消費が年齢とともにどのように変化するかを観察する。収入と同様に消費もライフステージにより水準が大きく異なるため、年齢別の分析は不可欠である。そこで年間収入と同様に、大学以上卒と非大卒の 2 つのグループでの、年齢別の消費動向を示したのが **Figure 8** である。ここでは推計された年間消費を等価尺度で調整したものを消費の尺度としている。

Figure 8 年齢別・学歴別の等価尺度調整済年間消費

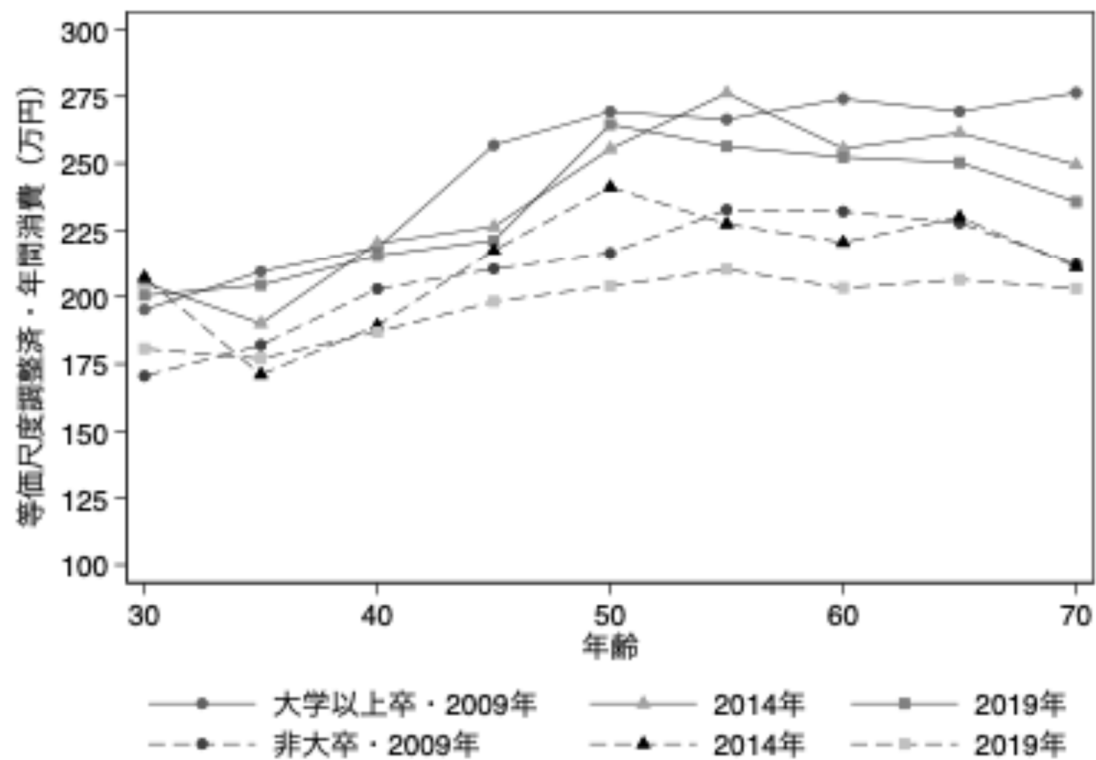


Figure 9 出生コホート別の学歴プレミアム（消費指標）

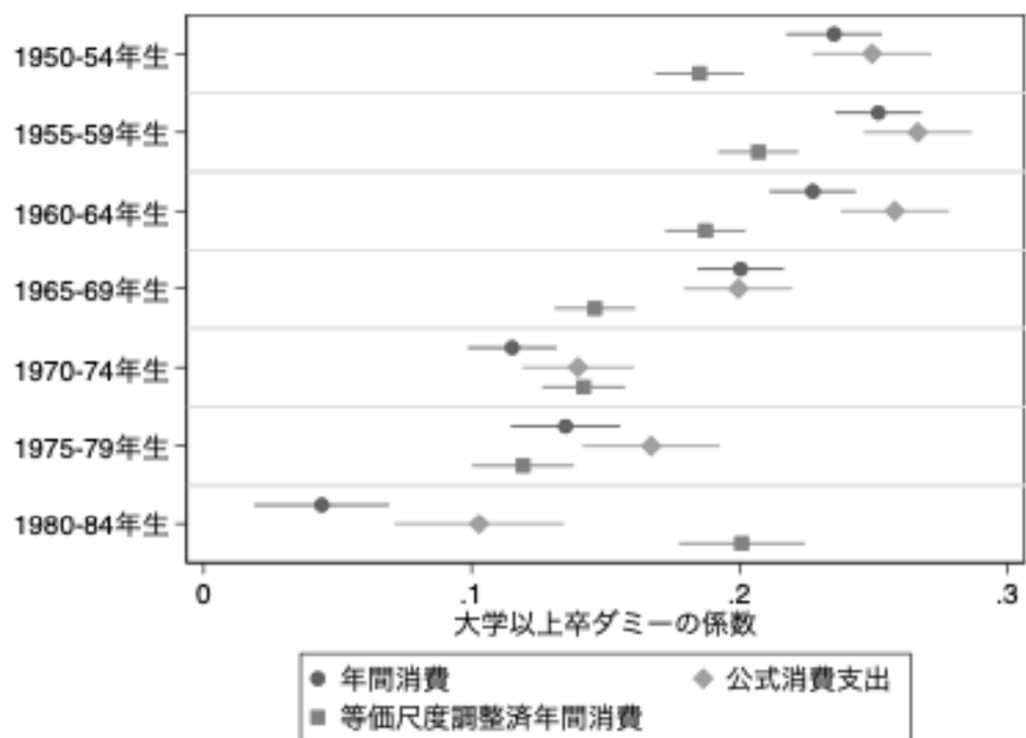


Figure 5 で見た年齢別の世帯の年間収入は 50 歳前後をピークとする山型であったが、消費はおおむね右上がりの形状をしている。また学歴別に見れば、どの時点で見てもどの年齢で見ても、世帯主が大学以上卒の世帯は非大卒世帯よりも消費水準が高くなっている。その差は年間収入と比べると小さいが、年齢とともに拡大していく傾向がある。30 歳代から 40 歳代の若年世帯では 25 万円程度（平均の 10%程度）にとどまる一方で、60 歳代以降は 50 万円程度（平均の 20%程度）になる。平均で見られた 2019 年の消費が落ち込みは非大卒の 50 歳代以降の中高齢層で大きい。

この等価尺度を調整した消費についても学歴プレミアムの世代差を推計したものが Figure 9 である。Figure 4・Figure 6・Figure 7 と同様に、被説明変数を等価尺度調整済の消費として、性別ダミー・学歴×年齢ダミー・時点ダミーに加え、生年×学歴ダミーで回帰した結果である。比較のために、等価尺度を調整していない年間消費および公式の月平均消費支出を説明変数としたものも示した。

消費における学歴プレミアムは、ここまで見てきた収入におけるプレミアムに比べれば小さい。再分配をした後の可処分所得でも大学以上卒のプレミアムは 30%程度存在していたが、年間消費・公式消費でも最大で 25%程度にとどまる。等価尺度を調整したプレミアムで見れば、最大でも 20%前後である。

一方で、世代による違いを見ると、年間収入の学歴プレミアムと同様に 1970 年代前後の世代で大きく低下したことがある。等価尺度を調整しない年間消費については、最も若い 1980-84 年生の世代では大学以上卒と非大卒では消費 5%ほどしか変わらないまでになっている。この学歴プレミアムの縮小は、すでに見たように、高学歴層における急速な世帯規模の縮小で生じたものであり、等価尺度調整済みの年間消費で見ればプレミアムの大きさは 1970 年代に生まれた世代よりむしろ大きくなっている。

6.3 学歴別の貯蓄率

ここまで見てきた年間の可処分所得と消費のデータを用いることで、学歴別の貯蓄率を計算することができる。貯蓄とは家計の可処分所得のうち消費されなかった部分であり、それを可処分所得で割ることで貯蓄率が計算できる。ここでも利用した、宇南山・大野(2017; 2018)による全国消費実態調査を使った年間の家計収支の補正方法は、マクロ的な貯蓄率と概ね整合的な推移をすることが確認されており、そのベースで貯蓄率が学歴別に計算できることになる¹²。

年間消費が年間可処分所得を上回ると、家計は資産を取り崩して消費することになり、貯蓄率はマイナスになる。ただし、年間の可処分所得がマイナスの場合には貯蓄率は適切に定義できないため、ここでの計算からは除外した¹³。また、学歴別・年齢階級別の貯蓄率の「平均」を計算する際に、各グループに該当する各世帯の貯蓄率を単純に平均することは適切な分析ではない。各世帯は少なくとも一定の消費をしており、たとえば 99 パーセンタイル点（消費の下位 1%）でも 100 万円を超えている。一方で、年間可処分所得はゼロに近いケースが多く存在しており、そうした世帯の貯蓄「率」は非常に大きな値のマイナスとなる。定義によって貯蓄率は 1 を超えることはない（消費はマイナスならないため）ため、単純に平均をすれば年間可処分所得の小さな世帯の結果が強く反映されてしまう。

この問題に対応するために、各グループの平均を計算するにあたり、各世帯の可処分所得をウェイトとして用いて計算した。これは、実質的に各グループの平均可処分所得と平均年間消費支出から貯蓄率を

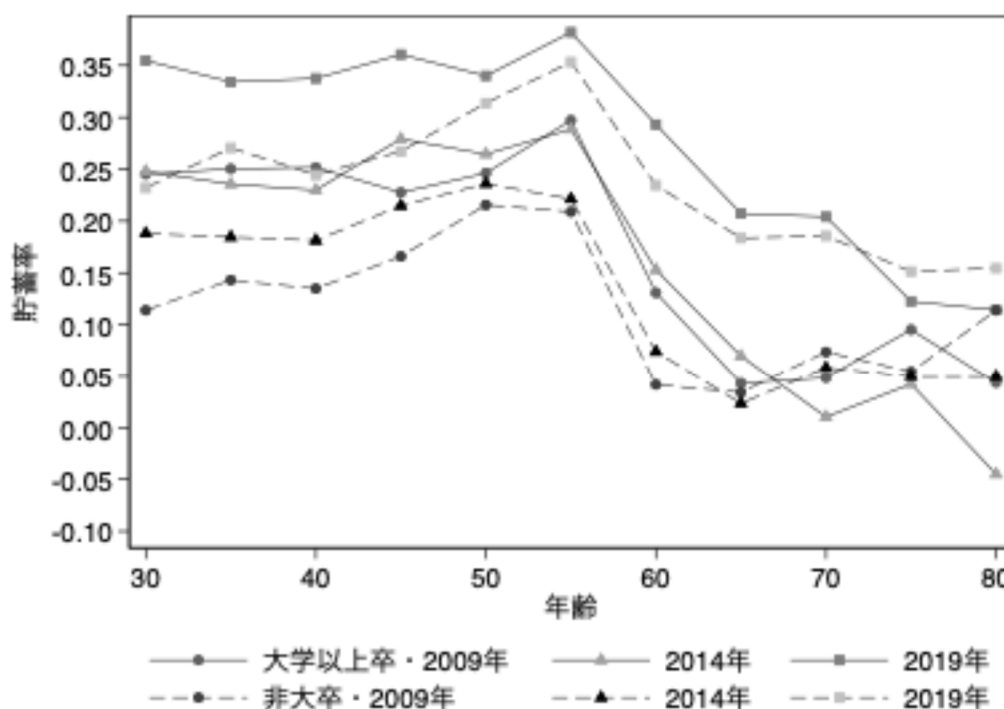
¹² 世帯調査とマクロ統計の貯蓄率の動向に関する最近の動向までカバーした研究として、宇南山・米田(2018)を参照。

¹³ 可処分所得がマイナスになったのは、2009 年で 59 世帯、2014 年で 29 世帯、2019 年で 12 世帯と少数であり、無視して分析しても大きな影響はない。

計算するのと数学的には同値である。

計算された貯蓄率を年齢別・学歴別に示したのが Figure 10 である。これは宇南山・大野（2018）において示された「年齢階級別の貯蓄率の変化（図 9）」を学歴別に描いたものに相当する。これまで見てきたように、学歴別に収入・消費のライフサイクルパターンは異なっており、結果としての貯蓄率のライフパターンも異なる可能性がある。

Figure 10 年齢別・学歴別の貯蓄率



年齢別の動向を見ると、大学以上卒・非大卒によらず、また調査年次によらず、貯蓄率は60歳まで高めに推移し、それ以降大きく低下する傾向が見られる。これは、宇南山・大野（2018）で示された年齢階級別の貯蓄率の動向と共通しており、貯蓄のライフサイクル理論とも整合的な動きである¹⁴。

学歴別の違いに注目すれば、貯蓄率の低下する60歳代前半までは大学以上卒の方が非大卒よりも貯蓄率は高い。ライフサイクル理論によれば、老後の所得の減少に備えることが貯蓄をする最大の理由となる。すでに Figure 5 で見たように、大学以上卒の方が現役時点での所得の伸びが大きく、引退後との所得差が大きい。引退後の所得の落ち込みに備える必要性が大学以上卒の方が高いために貯蓄率も高くなると考えられるのである。

60歳代後半以降は、学歴別の貯蓄率の差はほとんどなくなりゼロ近傍となるが、この観察もライフサイクル理論と整合的である。60歳代後半以降では収入の大部分は公的年金であるが、公的年金は生涯を通じた保険料負担に応じて決まるため、生涯を通じた平均の年間収入に近い水準となる。ライフサイクル理論によれば、家計が生涯効用を最大化しているなら、消費は生涯所得を平準化するように決定される。すなわち、いずれも「生涯所得」を人生の長さで割った水準となり、学歴によって生涯所得が異なるとしても貯蓄率には大きな差はないことになる。

¹⁴ ライフサイクル理論に基づく貯蓄率の動向の分析については宇南山(2023)第14章を参照。

ただし、2019 年については、その水準がゼロを大きく上回る水準となっている。時系列的に見れば、2009 年と 2014 年の違いは大きくなく、特に大学以上卒はほぼ同じプロファイルである。それに対し、2019 年は学歴によらず全年齢層で貯蓄率がおおむね 10%程度上昇している。

この 2019 年の貯蓄率の上昇は、Table 10 で示した消費の落ち込みによって発生している。Figure 8 で示したように世帯人員数をコントロールすれば消費はそれほど変化していないが、Table 11 で示したように世帯人員が大きく減少したために世帯あたりの消費は大きく落ち込んでいた。

国民経済計算（SNA）における家計貯蓄率は、2014 年に-1.3 パーセントであったが 2019 年は 2.9%と上昇している。宇南山・米田(2018)や宇南山(2023)第 13 章で分析されているように、SNA と世帯調査である家計調査や全国消費実態調査などで計算される貯蓄率には概念的な違いも多く単純な比較はできない。しかし、ここで観察された貯蓄率の動向は方向性としては不自然ではない。

7. まとめと展望

本研究では、全国消費実態調査・全国家計構造調査の家計簿の情報を、家計消費状況調査や家計調査などの補助情報を用いて補完し年間での家計収支が観察できるデータセット作成した。さらに、国勢調査と全国消費実態調査を個票レベルでの照合をした結果を適用することで、学歴別の家計収支の状況を総合的に分析できるデータセットを作成した。

まず、学歴の高い世帯主のいる世帯においては、世帯主自身の収入を中心に高い世帯収入を得ており、さらに消費水準も高くなるプレミアムを享受していることを明確に示された。学歴別の家計収支の差を定量的に把握したことは、本研究の重要な貢献である。

大学以上卒の世帯主では、中学卒や高校卒と比較して可処分所得が顕著に高く、その差異は夫婦双方の学歴が高いほど拡大する傾向が観察された。高学歴層の夫婦が生み出す世帯収入の相乗効果は、学歴による所得格差を世帯レベルでさらに強める要因になっていると考えられる。結果としては配偶者を含めた世帯全体の収入水準では、高学歴層の優位がなお維持されており、夫婦双方の高学歴がもたらす高収入の累積効果が学歴格差の持続を支えていると考えられる。この発見は、経済格差を分析する際に、家族の経済学など世帯の構成の決定についても分析することが不可欠であることを示唆している。

こうした学歴プレミアムは世代によって一様ではなく、若年世代においては以前より縮小傾向を示した。特に、1970 年代に生まれた、いわゆる「氷河期世代」での学歴プレミアムの落ち込みは顕著であった。背景として、大学進学率の上昇に伴う高学歴の相対的な希少性の低下や、非正規雇用の増加による雇用形態の変化が大きく影響していると推察される。教育を受ける価値が、経済状況に応じて変化すること、その差が長期に渡り持続し、家計収支全体に影響を与えることを確認したことは、先行研究の分析を補強するものである。

今後は、本研究で示された学歴別の家計収支の状況および貯蓄行動が、経済学の標準的なライフサイクル理論とどの程度合致しているかについては、検証をしていくことが必要である。現状では、学歴による差の年齢別の推移や時点による違いを明らかにしただけであり、十分に解釈ができていない状況ではない。しかし、生涯可処分リソースの異なると考えられる学歴の異なる世帯の所得・消費のプロファイルが利用できることは、理論の実証をする上では非常に大きな情報と言える。

また、本研究の結果は、結婚に関する意思決定の経済学的な側面での重要性を示唆するものである。配偶者の選択は、個人の経済構成に大きな影響を与える。一方で、特に女性は結婚後に労働市場を退出する

ケースが多く、女性の人的資本は必ずしも有効活用されていない。近年、夫婦の学歴の組み合わせの「同類婚」傾向が強まっており、学歴ごとの就業の意思決定なども興味深い研究テーマになる。

データの側面から言えば、2024 年の全国家計構造調査の結果を追加することが今後重要となる。2019 年調査と同様に、世帯員ごとの学歴を調査しており、国勢調査との照合のような手続きをせずとも本研究と同様に分析が可能である。世代による学歴プレミアムの動向などを分析するにはより長期の時系列が利用できることが望ましく、新たなデータが追加されることの意義は大きい。

本研究を通じて、既存の統計を照合して新たなデータを構築することの意義が明らかになった。特に、個人が識別できない世帯調査においても、一定の妥当性を持つ照合が可能であることが分かった。国勢調査は全数調査であり、原理的には任意の標本調査と照合することが可能である。また、国勢調査同士の照合も検討の価値がある。こうした統計間の照合を可能とするような情報（調査間での調査区の対応関係など）を保存することは、現状では制度的に困難であるが、今後の議論に期待したい。

参考文献

- [1] 赤林英夫（2012）「人的資本理論」『日本労働研究雑誌』第 54 巻 4 号 pp. 8-11.
- [2] 宇南山卓（2009）「SNA と家計調査における貯蓄率の乖離 -日本の貯蓄率低下の要因」RIETI Discussion Paper Series 10-J-003
- [3] 宇南山卓（2022）「全国消費実態調査と国勢調査の照合について」統計研究研修所・共同研究リサーチペーパー第 56 号(<https://www.stat.go.jp/training/2kenkyu/research/pdf/resear56.pdf>)
- [4] 宇南山卓（2023）『現代日本の消費分析—ライフサイクル理論の現在地』慶應義塾大学出版会, 532 頁.
- [5] 宇南山卓・大野太郎（2017）「貯蓄率の低下は高齢化が原因か？」『経済研究』第 68 巻 3 号 pp. 222-236.
- [6] 宇南山卓・大野太郎（2018）「日本の世帯属性別貯蓄率の動向について」RIETI Discussion Paper 17-J-035.
- [7] 宇南山卓・米田泰隆（2018）「日本の「家計調査」と「国民経済計算（SNA）」における家計貯蓄率の乖離—1994 年から 2015 年における日本の家計貯蓄率低下要因—」『フィナンシャル・レビュー』第 134 号 pp. 191-205.
- [8] 大野太郎・中澤正彦・菊田和晃・山本学（2015）「家計の税・社会保険料の比較」『フィナンシャル・レビュー』第 122 号, pp. 50-58
- [9] 太田聰一・玄田有史・近藤絢子（2007）「溶けない氷河」『日本労働研究雑誌』第 569 号 pp. 4-16.
- [10] 川口大司（2011）「ミンサー型賃金関数の日本の労働市場への適用」阿部顕三・大垣昌夫・小川一夫・田淵隆俊編『現代経済学の潮流』第 3 章 67-98.
- [11] 近藤絢子（2024）『就職氷河期世代—データで読み解く所得・家族形成・格差』中公新書 192 頁.
- [12] 佐野晋平（2024）『教育投資の経済学』日経文庫 日本経済新聞出版社 272 頁.
- [13] 佐野晋平・安井健悟（2009）「日本における教育のリターンの推計」『国民経済雑誌』第 200 巻第 5 号 pp. 71-86.
- [14] 多田隼士・大野太郎・宇南山卓（2016）「マイクロ・データを用いた社会保険料の推計とその妥当性の検証」PRI Discussion Paper Series No. 16A-02

- [15] 安井健悟・佐野晋平 (2009) 「教育が賃金にもたらす因果的な効果について—手法のサーヴェイと新たな推定」『日本労働研究雑誌』, No. 558, pp. 16-33.
- [16] Becker, Gary (1975) *Human capital: A theoretical and empirical analysis, with special reference to education*, 2nd Ed. University of Chicago Press.