

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

財務省財務総合政策研究所 前研究員

多田 隼士

尾道市立大学経済情報学部准教授、
財務省財務総合政策研究所上席客員研究員

大野 太郎

一橋大学経済研究所准教授、
財務省財務総合政策研究所特別研究官

宇南山 卓

2016 年 2 月

本論文の内容は全て執筆者の個人的見解であり、財務省あるいは財務総合政策研究所の公式見解を示すものではありません。

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*

多田 隼士^{*1}

大野 太郎^{*2}

宇南山 卓^{*3}

概要

本稿では厚生労働省『国民生活基礎調査』（平成 22 年調査）の個票データを利用し、家計の社会保険料負担額（年金保険料・健康保険料・介護保険料・雇用保険料）の理論値をマイクロ・シミュレーションによって推計した。さらに、各世帯が報告した記入値と計算された理論値の比較を通じて、理論値の妥当性を考察した。

社会保険料負担額の推計のために、まず各世帯の加入する社会保険制度を推定した。収入に加え家計の報告した世帯主の職業を用いて推定することで、各世帯の報告する加入制度との一致率は高いものとなった。推定された加入制度を用いて計算された理論値は、実際の記入値と平均的にはほぼ乖離がなく、誤差の分布も対所得比で 7%程度の範囲に収まっていた。

本研究において社会保険料負担額の理論値作成に用いた変数は、総務省統計局の実施している『全国消費実態調査』でも利用可能な変数であり、推計に係る手法についても『全国消費実態調査』に適用可能である。

JEL 区分：C15, H24

キーワード：社会保険料、マイクロ・シミュレーション分析、記入値、理論値

* 本研究の一部は科学研究費助成事業（学術研究助成基金助成金、若手研究(B)、課題番号 26780176）、京都大学経済研究所共同利用・共同研究拠点平成 27 年度プロジェクト研究の成果であり、また厚生労働省『国民生活基礎調査』の調査票情報を利用して独自集計したものである。関係者各位に厚く御礼を申し上げる。本稿の作成にあたっては財務総研研究会の参加者から貴重なコメントを頂き、謝意を表する。なお、本稿の内容は著者らの個人的見解であり、著者らが所属する機関の公式見解を示すものではない。

^{*1} 財務省財務総合政策研究所 前研究員

^{*2} 尾道市立大学経済情報学部 准教授／財務省財務総合政策研究所 上席客員研究員

^{*3} 一橋大学経済研究所 准教授／財務省財務総合政策研究所 特別研究官

I. はじめに

近年、日本では家計マイクロ・データ（調査票情報）を用いた税制・社会保障分野の分析へのニーズが高まっている¹。社会保障制度の維持には、家計の税・保険料負担の見直しが必要であると考えられる中、税制・社会保障制度の変更による再分配効果の評価が重要な役割を果たす。こうした分析には、家計の非対称性を考慮したマイクロ・データを用いた精緻な理論値の推計が欠かせない。特に、少子高齢化の進む現在の日本においては、社会保障財源として消費税の役割が重要となっており、家計の税・社会保障負担の全体像を把握するためには消費関連の情報を持つマイクロ・データの重要性が高まっている²。

日本における代表的な消費関連のデータとして『家計調査』および『全国消費実態調査』が存在している。しかし、『家計調査』および『全国消費実態調査』では勤労者世帯・無職世帯以外の世帯（自営業者など）は税・社会保険料の負担額を調査していない。しかも、大野ほか(2015)は『家計調査』および『全国消費実態調査』において各家計の報告する社会保障負担が過少になっている可能性を指摘している。そのため、現状では、マイクロ・データを単純に用いただけでは、社会保障負担の全体像を把握できない状態である。

それに対し、田中ほか(2013)などの先行研究では、調査票に記載された世帯属性や各世帯員の所得などの情報を利用し、世帯の構造に現実の制度に当てはめた社会保障負担額を算出することでデータを補完して分析をしている。方法論的には、利用可能な変数から社会的な制度などを世帯ごとに適用して、新たな変数を仮想的に構築するマイクロ・シミュレーションの手法と同じであり、調査対象となっていない世帯や変数についても社会保障負担の状況を推計することができる。しかし、理論的に計算される推計値（以下では理論値と呼ぶ）は限られた情報から世帯ごとの税・社会保障負担という個別性の高い変数を推計することになるため、大きな測定誤差を含む可能性があり、その妥当性について検証が必要である。

そこで、本稿では厚生労働省『国民生活基礎調査』（平成 22 年調査）の個票データにもとづき、家計の社会保険に関する加入制度および社会保険料負担額（年金保険料・健康保険料・介護保険料・雇用保険料）の実際に報告した金額（以下では記入値と呼ぶ）と理論値の比較を通じて、理論値の妥当性を考察する。『国民生活基礎調査』の記入値と比較することで理論値の妥当性を評価するという手法は、大野ほか(2015)で所得税・住民税負担額について適用された手法である。田近・八塩(2006a)は所得税負担額について理論値と記入値を所得階層ごとに比較をしているが、比較対象となっている記入値の正確性に疑義がある。また、田中・四方(2012)は各種の税・保険料負担額について理論値とマクロ統計の比較をし

¹ 例えば、マイクロ・データを用いた税・社会保障の分析をした研究として田近・古谷(2003, 2005)、大竹・小原(2005)、田近・八塩(2006a, 2006b, 2008, 2010)、白石(2010)、土居・朴(2011)、八塩・長谷川(2009)、高山・白石(2011)、田中(2014)などを参照。

² 特に消費税負担について計測をした先行研究としては大竹・小原(2005)、八塩・長谷川(2009)、高山・白石(2010)、白石(2011)、高山・白石(2011)、田中(2014)などがある。

ているが、マクロに集計する際の手法に起因する違いが大きいと考えられ直接の比較が困難という問題があった。それに対し、大野ほか(2015)では、他の統計との比較によって『国民生活基礎調査』の記入値の妥当性を確認しており、理論値の妥当性を適切に評価できていると考えられる。

ただし、大野ほか(2015)で分析された所得税・住民税と、社会保険料では方法論的に大きな違いがある。それは、税と異なり社会保障の負担については加入制度の推定が必要になることである。個々人の加入している社会保険制度は、雇用状態などにより異なり、また制度によって負担額も大きく異なる。それに対し、田中・四方(2012)では、世帯員がどの社会保険制度に加入しているかを雇用者所得によって推定をしていたが、ここでは記入値による各世帯員の職業の情報も加味して推定をした。その結果、各世帯員の加入する社会保険制度を高い精度で推定することができた。

各世帯が加入している社会保険制度を特定することができれば、税と同様に制度によって規定される計算式を適用し、各世帯の社会保障負担額を計算することができる。推計された理論値は、実際の記入値と平均的にはほぼ乖離がなく、誤差の分布も対所得比で7%程度以内の範囲に収まっていた。これは大野ほか(2015)で計算された所得税・住民税負担額の誤差の分布がおおむね4%程度以内の範囲であったのと比較して大きな誤差である。しかし、個別の世帯でこの程度の誤差であれば、世帯属性毎のグルーピングなど集計をすれば十分に分析に耐えうる精度である。

ここでの分析対象は『国民生活基礎調査』であるが、検証に用いた変数は『全国消費実態調査』でも利用可能な変数である。しかも、佐野・多田・山本(2015)、多田・三好(2015)では、『国民生活基礎調査』と『全国消費実態調査』は世帯属性や収入の情報については整合的であることが確認されている。つまり、ここでの理論値の推計に係る手法についても『全国消費実態調査』に適用可能であり、構築される理論値の妥当性は同程度の精度を持つと考えられる。

家計の所得段階における税・社会保険料負担を、消費段階における税負担と結びつけることができ、家計の税・社会保障負担の全体像を分析することが今後の課題である。その前提として、本稿で示した理論値作成の手法を『全国消費実態調査』に適用できること、及び理論値の精度も保証できること、を示したのが本稿の意義である。

以下、本稿の構成を述べる。まずⅡ節では使用するデータ、及び社会保険料負担額の推計方法について説明する。それを踏まえて、Ⅲ節では公的年金制度及び健康保険制度を対象に、加入制度に関する推定の妥当性について考察する。次にⅣ節では各種保険料負担額を対象に、記入値と理論値の比較を行い、双方の近似・乖離について考察する。最後にⅤ節で結論を述べる。

Ⅱ．使用データと推計方法

Ⅱ-1．使用したデータ

本節では使用データと理論値の推計方法について説明する。データは厚生労働省『国民生活基礎調査』（平成 22 年調査）の個票データ（世帯票・所得票）を使用した。同調査は 2010 年 6 月に実施されたもので、各世帯員について調査時点の属性（続柄、性別、出生年月³、就業状況等）や調査前年 1 年間（2009 年 1 月～12 月）の所得・社会保険料等を調査しており、社会保険料については医療保険料、年金保険料、介護保険料、その他（雇用保険料等）といった内訳も含まれている。本稿では、これら各世帯員の属性および収入に関する情報を用いて社会保険料の理論値を作成し、それらを社会保険料の記入値と比較することで理論値の妥当性を検証する。

サンプルは世帯票・所得票の双方に回答している世帯（26,115 世帯）のうち、年齢不詳者を含む世帯、社会保険料が不詳である者を含む世帯を除外した。分析に使用したサンプルサイズはそれぞれ最大で、医療保険 23,168 世帯、年金保険 24,115 世帯、介護保険 23,985 世帯、その他（雇用保険等）24,465 世帯、社会保険料合計で 23,100 世帯である。

Ⅱ-2．推計方法

加入制度の推定

社会保険料を推計するためにはまず、世帯員がどの社会保険制度に加入しているかを推定する必要がある。ここでは田中・四方（2012）に倣い、雇用者所得に注目して推定（以下、「収入等による分類」と呼ぶ）を行った。医療保険、年金保険、介護保険については雇用者所得が「パートタイム労働者・短時間労働者の平均賃金×30 時間×52 週」より多い場合に⁴、健康保険（被用者保険）・厚生年金の被保険者とした⁵。ただし、医療保険については 75 歳以上の者は全て後期高齢者医療制度の被保険者、介護保険については 65 歳以上の者は全て第 1 号被保険者とした。また、年間所得が 130 万円以下の者について、同居親族に健康保険被保険者がいる場合には健康保険の被扶養者であり、配偶者が厚生年金被保険者である場合には第 3 号被保険者であると推定した。そして、それ以外の者は全て国民健康保険・国民年金に加入しているものと見なすことにした。雇用保険については雇用者所得が「パートタイム労働者・短時間労働者の平均賃金×20 時間×52 週」より多い場合に雇用保険の被保険者とし⁶、それ以外の者は加入者ではないものと見なした。

本稿では上記に加えて、加入制度を推定する際に職業も加味した推定（以下、「職業も加味した分類」と呼ぶ）も行った。ここでは世帯票で回答している「勤めか自営かの別」の

³ 年齢は出生月に関わらず、1 月 1 日に 1 歳年を重ねるようにプログラムされている。また、属性と収入等の調査時点が異なるため、収入等は調査時点の年齢よりも 1 歳若い時の収入等であるものと見なしている。

⁴ 『賃金構造基本統計調査』における民営事業所・産業計・企業規模計・一般労働者区分の、短時間労働者の 1 時間当たり所定内給与額から算出した。

⁵ ただし、2009 年に 69 歳以下の者に限る。

⁶ ただし、2009 年に 65 歳以上の者は一律に雇用保険に加入していない者と見なした。

情報も判断材料に加えて、各世帯員の加入制度を推定した。「収入等による分類」との違いは次の2点である。第1に、健康保険（被用者保険）・厚生年金の被保険者であるかどうかを推定する際に、「雇用者所得が「パートタイム労働者・短時間労働者の平均賃金×30時間×52週」より多い」という条件に加えて、「勤めか自営かの別」が「会社・団体等の役員」「一般常雇者」「1月以上1年未満の契約の雇用者」のいずれかである者」という条件を追加した。なお、「勤めか自営かの別」が「日々又は1月未満の契約の雇用者」は含めていない。第2に、雇用保険の被保険者であるかどうかを判定する際に、「雇用者所得が「パートタイム労働者・短時間労働者の平均賃金×20時間×52週」より多い」という条件に変えて、「勤めか自営かの別」が「一般常雇者」「1月以上1年未満の契約の雇用者」のいずれかである者」という条件を用いている。ただし、各世帯員の「勤めか自営かの別」は2010年6月時点の状況であるのに対し、年間所得などは2009年のものであるため、必ずしも就業状況が同じとは限らない点については留意する必要がある。

医療保険料の推計

次に、各保険料の理論値の推計方法について説明する。医療保険のうち、国民健康保険の保険料は「旧ただし書き」方式（世帯の収入から公的年金控除等の必要経費と基礎控除を差し引いた段階の所得に対して賦課する方式）とし、『国民健康保険実態調査』から算出した所得割率、資産割、均等割、平等割の全国平均値を用いて世帯員ごとに計算した上で、それらを足し合わせて世帯全体の額を推計した⁷。ここでは国民健康保険料の賦課限度額、応益割に対する減額制度も考慮している⁸。

健康保険の保険料は、全国健康保険協会のホームページに掲載されている保険料率を用い、労使で折半するものと仮定して、年間所得に保険料率を乗じて算出した。なお、年間収入が標準報酬月額の上限額の12倍と標準賞与額の上限の2倍の和（以下、「賦課上限額」と呼ぶ）を超えた場合には、年間所得の代わりに賦課上限額を用いて健康保険の保険料を計算している。

後期高齢者医療保険の保険料は、厚生労働省のホームページに掲載されている所得割率及び均等割の全国平均を用いて各世帯員の保険料を算出した。ここでは後期高齢者医療保険の保険料の賦課限度額、所得割額および均等割の減額制度も考慮している⁹。

⁷ 国民健康保険料の計算方法には、本稿で採用した4方式に加え、3方式（所得割、均等割、平等割）、2方式（所得割、均等割）がある。『平成25年国民健康保険事業年報』によると、4方式を採用している保険者は1,130、3方式は520、2方式は62となっているが、各世帯の住む自治体でどの方式を採用しているかは識別できないため、本稿ではどの自治体も4方式を採用しているものと見なした。

⁸ 国民健康保険の保険料については、被保険者の保険料負担能力に応じて付加される応能分（所得割、資産割）と、受益に応じて等しく被保険者に賦課される応益分（被保険者均等割、世帯別平等割）から構成されるが、低所得等の事情のある被保険者については、応益分を軽減（7割、5割、2割）する制度を設けている。

⁹ これらの計算結果から医療保険料の理論値については、年金保険料、介護保険料、雇用保険料と異なり、世帯全体の負担額がゼロである世帯は存在しない。

年金保険料の推計

年金保険のうち、国民年金の保険料は定額保険料であり、免除制度（全額、4分の3、半額、4分の1）について該当する者は全て申請すると仮定して算出している。

厚生年金の保険料は、日本年金機構のホームページに掲載されている第1種被保険者の保険料率を用い、労使で折半するものと仮定して算出した。健康保険の場合と同様に、年間所得が賦課上限額を超えた場合には賦課上限額を算定基礎としている。

介護保険料の推計

介護保険の被保険者のうち、第1号被保険者の保険料は厚生労働省のホームページに掲載されている各都道府県の平均保険料基準額を用いて、本人及び世帯所得ごとに6段階で各世帯員の保険料を算出した。

第2号被保険者のうち、国民健康保険に加入している者については『国民健康保険実態調査』から算出した介護給付分の所得割率、資産割、均等割、平等割の全国平均値を用いて世帯員ごとに算出した。保険料の賦課限度額、応益割に対する減額制度も考慮している。

健康保険に加入している者については全国健康保険協会のホームページに掲載されている保険料率を用い、労使で折半するものとして算出した。標準報酬月額と標準賞与額の上限も考慮している。

雇用保険料の推計

雇用保険の保険料率は、一般の事業に関する労働者の負担率を当てはめて算出した。

Ⅲ. 加入制度に関するモデルの妥当性

本節では個人（世帯員）ベースで、調査票に記載されている社会保険の加入制度とⅡ節の手法により推定した加入制度を比較し、推定した加入制度の妥当性について検証する。こうした作業を行う理由としては、社会保険料負担額の推定を正しく行うにあたって、まず加入制度の推定を正しく行う必要が出てくるためである。本稿の検証方法は以下のとおりである。『国民生活基礎調査』では世帯員ごとに、「医療保険の加入状況」「公的年金の加入状況」について調査している（以下、「調査票に基づく分類」と呼ぶ）。ここでは「調査票に基づく分類」を真の加入制度と見なし、これとの比較を通じて、世帯員ごとに推定した「収入等による分類」や「職業も加味した分類」がどの程度妥当性を有しているかを検証する。

表1は「調査票に基づく分類」を左に、「収入等による分類」「職業も加味した分類」を上にとり、制度ごとにどの程度一致しているかを示している。「調査票に基づく分類」と推定した加入制度が一致している部分には網掛けをつけて表記した。「調査票に基づく分類」

と一致している割合は概ね 70%以上の水準を示している。ただし、「調査票に基づく分類」で国保・組合に加入していると回答している者のうち、「収入等による分類」でも一致した者は 42.2%と低い。また、「調査票に基づく分類」で国民年金第 1 号被保険者であると回答している者のうち、「収入等による分類」でも一致した者は 62.5%と低い。これは一定の雇用者所得がある場合に、実際は国民健康保険加入者や国民年金第 1 号被保険者であるが、推定では被用者保険・家族や国民年金第 3 号被保険者と判定されているためである。これに対して現在の職業も判断材料に加えると、「職業も加味した分類」では国保・組合の一致率が 70.8%に、国民年金第 1 号被保険者の一致率が 79.7%に上昇する。

表 1 推計した加入制度と「調査票に基づく分類」との比較

医療保険料											
調査票に基づく分類	収入等による分類					職業も加味した分類					各制度に加入する人数のシェア
	国保	被用・本人	被用・家族	後期	全体	国保	被用・本人	被用・家族	後期	全体	
国保・市町村	75.8%	11.9%	12.3%	0.0%	100.0%	88.4%	4.7%	6.8%	0.0%	100.0%	27.2%
国保・組合	42.2%	29.8%	28.0%	0.0%	100.0%	70.8%	12.9%	16.3%	0.0%	100.0%	1.5%
被用者保険・本人	14.7%	77.0%	8.2%	0.0%	100.0%	18.7%	74.1%	7.2%	0.0%	100.0%	29.9%
被用者保険・家族	15.2%	1.9%	82.9%	0.0%	100.0%	20.7%	0.4%	78.9%	0.0%	100.0%	26.5%
後期高齢者医療	7.0%	0.2%	0.0%	92.8%	100.0%	7.1%	0.1%	0.0%	92.8%	100.0%	13.0%
不詳	72.5%	11.6%	16.0%	0.0%	100.0%	81.9%	6.7%	11.3%	0.0%	100.0%	1.2%

年金保険料											
調査票に基づく分類	収入等による分類					職業も加味した分類					各制度に加入する人数のシェア
	1号	2号	3号	加入せず	全体	1号	2号	3号	加入せず	全体	
国民年金第1号	62.5%	21.9%	10.5%	5.2%	100.0%	79.7%	8.9%	5.7%	5.8%	100.0%	11.6%
国民年金第2号	14.3%	77.2%	4.2%	4.3%	100.0%	16.7%	74.8%	3.7%	4.9%	100.0%	29.3%
国民年金第3号	8.9%	4.5%	86.6%	0.0%	100.0%	15.4%	0.8%	83.8%	0.0%	100.0%	7.8%
加入していない	4.1%	4.8%	0.6%	90.6%	100.0%	4.7%	2.4%	0.4%	92.5%	100.0%	37.9%
不詳	78.6%	5.6%	4.0%	11.9%	100.0%	83.3%	2.4%	2.4%	11.9%	100.0%	0.2%

(データ出所)『国民生活基礎調査』(平成 22 年調査)

次に(加入制度に関して)「調査票に基づく分類」「収入等による分類」「職業も加味した分類」それぞれのケースの下で、保険料負担額に関する記入値と理論値の乖離がどれくらい生じるのかについて見ていく。乖離の指標としては、乖離の年間所得に対する割合(以下、「乖離率」と呼ぶ)を用いる。表 2 は医療保険料と年金保険料について、「調査票に基づく分類」「収入等による分類」「職業も加味した分類」それぞれのケースで乖離率を算出した結果を示している¹⁰。なお、ここでは記入値および理論値が共にゼロである世帯をサンプルから除いている(以下の分析でも同様である)¹¹。医療保険料および年金保険料ともに、各分類における乖離率の分布は平均、散らばりいずれの点からも概ね近い結果となってい

¹⁰ 『国民生活基礎調査』では介護保険の加入状況や、雇用保険の加入状況について調査する項目が無い
ため、介護保険および雇用保険において「調査票に基づく分類」は存在しない。

¹¹ 記入値および理論値全てがゼロであるサンプル数については付表 1 を参照されたい。

る。保険料負担額に関する理論値計算の過程で、記入値と理論値の間に乖離が生じる要因として、第1に加入制度が一致できていないこと、第2に（加入制度は一致しているが）負担額の水準が一致できていないことが挙げられる。しかし、表2の結果が示すのは加入制度の推定に伴う要因からは乖離率をほとんど変化させないという点である。したがって、保険料負担額に関する乖離の発生要因として重要なものは加入制度の推定に伴う要素ではなく、むしろ負担額水準の推定に伴う要素（例えば理論値の計算方法や記入値の誤記入など）に求めていく必要がある。

表2 各分類の乖離率

加入制度の 推定方法	医療保険料			年金保険料		
	調査票に 基づく分類	収入等 による分類	職業も加味 した分類	調査票に 基づく分類	収入等 による分類	職業も加味 した分類
平均	-0.42%	-0.48%	-0.83%	1.78%	1.41%	1.65%
5パーセン タイル値	-5.93%	-5.94%	-6.28%	-7.68%	-7.68%	-8.57%
95パーセン タイル値	6.70%	6.74%	6.52%	7.05%	6.60%	7.13%
観測値数	49,033	49,033	49,033	23,330	23,330	23,330

（データ出所）『国民生活基礎調査』（平成22年調査）

（注1）サンプルからは記入値または加入状況が不詳の場合、記入値および理論値が共にゼロである場合を除いている。

（注2）乖離率＝（記入値－理論値）／世帯所得

IV. 保険料負担額に関するモデルの妥当性

IV-1. 世帯ベースでみた各保険料の乖離の状況

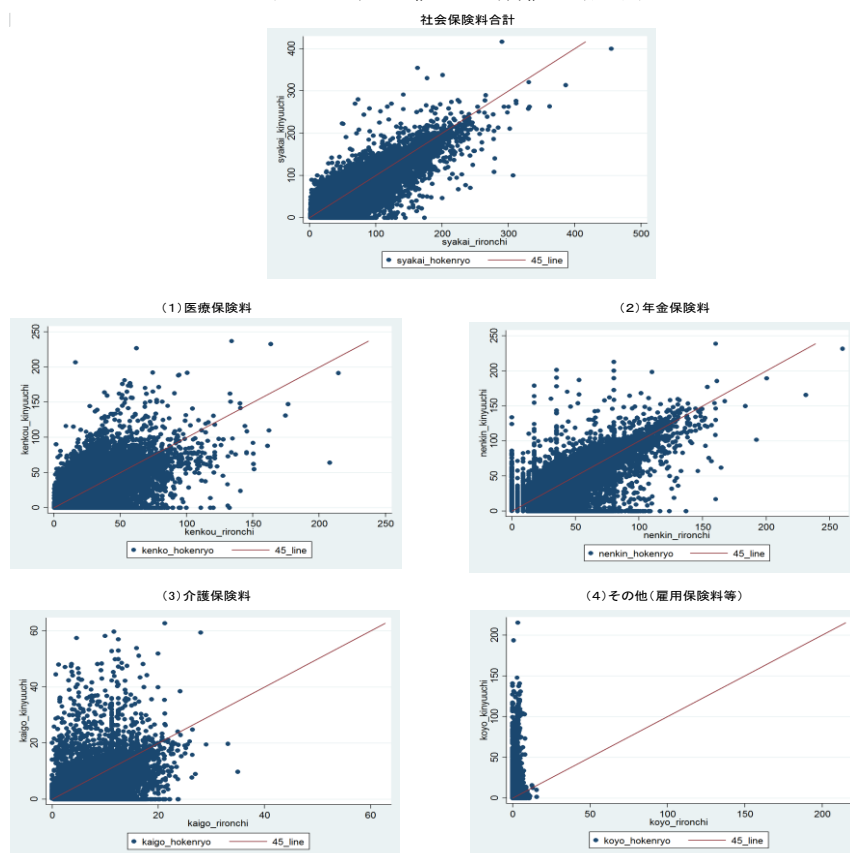
本節では世帯員全員の社会保険料を合算した世帯ベースで、保険料負担額に関する記入値と理論値の近似・乖離について考察する。ここでは、表1のように国保・組合や国民年金第1号被保険者を中心として加入制度の一致性が向上した点を重視し、加入制度の推定については「職業も加味した分類」のケースを採用する¹²。

まずは記入値と理論値の乖離の状況を視覚的に捉えてみる。図1は縦軸に記入値を、横軸に理論値をとり、社会保険料合計、医療保険料、年金保険料、介護保険料、その他（雇用保険料等）における記入値と理論値を散布図で示したものである。実線は45°線を表し

¹² 「職業も加味した分類」に基づいて計算した個人ベースでは、世帯主が他の世帯員の保険料をまとめて支払っているケースがあり、そのため乖離率が大きくなる可能性もある。しかしこうした問題は世帯ベースで見るとときには解消されると考えられる。

ている。散布図の点が 45° 線上にある場合は記入値と理論値が完全に一致しており、また 45° 線よりも上にある場合は記入値の方が大きく、 45° 線よりも下にある場合は理論値の方が大きいことを示している。はじめに社会保険料合計に注目すると、 45° 線から乖離した点も散見されるが、分布は概ね 45° 線の近くにあり、また右上がりの形状を持っており、記入値と理論値の一致した世帯が比較的多いことが分かる。医療保険料や年金保険料も社会保険料合計と同様に概ね 45° 線に近く右上がりの形状を持っているが、ただし記入値がゼロであるが理論値はゼロでない世帯や、反対に理論値がゼロであるが記入値はゼロでない世帯が一定程度存在している。介護保険料は概ね 45° 線の周りに分布しているものの、記入値が理論値よりも大きい乖離のある世帯がやや多く存在する。また年金保険料と異なり、記入値がゼロであるが理論値はゼロでない世帯は存在するが、反対に理論値がゼロであるが記入値はゼロでない世帯は見られない。その他（雇用保険料等）は 45° 線から乖離して分布していると言わざるをえない。記入値では世帯合計の保険料が年間 150 万円近くまで連続的に存在しているが、理論値では最も高い世帯で 30 万円程度であり、記入値が理論値よりもかなり大きい。

図 1 記入値と理論値の散布図



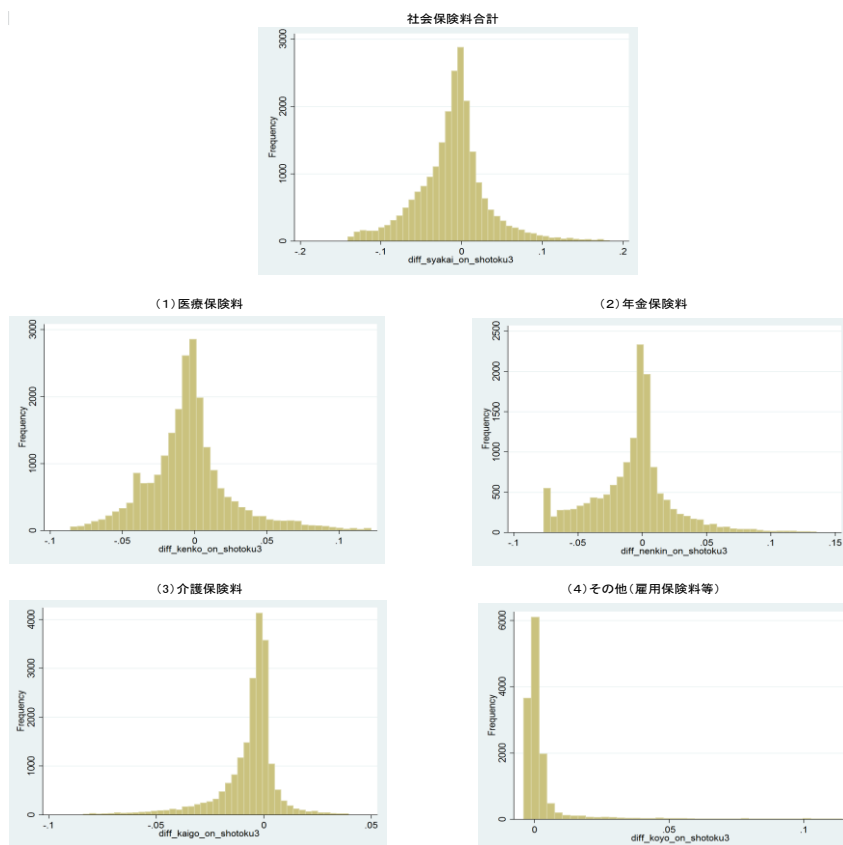
(データ出所)『国民生活基礎調査』(平成 22 年調査)

(注 1) 縦軸は記入値、横軸は理論値を示す。金額の単位は万円(年額)である。

(注 2) サンプルからは記入値不詳である場合、記入値および理論値が共にゼロである場合を除いている。

上述の散布図ではひとつの点にどれだけの世帯が含まれているかについては識別できず、また外れ値の分布状況によって印象が大きく変わりうる。そこで、次に記入値と理論値の乖離率に焦点をあて、平均や散らばりを含めた乖離の状況を考察する。図2は横軸に乖離率を、縦軸に度数を取り、社会保険料合計、医療保険料、年金保険料、介護保険料、その他（雇用保険料等）それぞれについて、記入値と理論値の乖離をヒストグラムで示したものである¹³。ここでは「乖離率＝（記入値－理論値）／世帯所得」として定義し、乖離率がゼロよりも大きい場合は記入値が理論値よりも大きく、乖離率がゼロよりも小さい場合は理論値が記入値よりも大きいことを示している。図2から、いずれの保険料でもゼロ付近で最も頻度が高くなっているが、左側の裾がやや厚く、理論値が記入値よりも大きい世帯がやや多いことを示している。特に医療保険料および年金保険料では左側の裾に度数の突出した箇所が見られる。

図2 乖離率に関する分布：ヒストグラム



（データ出所）『国民生活基礎調査』（平成22年調査）

（注1）縦軸は度数、横軸は乖離率を示す。乖離率＝（記入値－理論値）／世帯所得。

（注2）サンプルからは記入値不詳である場合、記入値および理論値が共にゼロである場合を除いている。

また、上下1％をカットしている。

¹³ 極端なサンプルの影響を排除するため、ヒストグラムの作成に当たっては乖離率の上下1％をカットしている。

IV－２．乖離の発生要因に関する考察

ここでは乖離の発生要因について考察する。記入値が理論値よりも大きい場合（乖離がプラスの場合）、記入値が過大もしくは理論値が過小に評価されている可能性がある。

一方、理論値が記入値よりも大きい場合（乖離がマイナスの場合）、記入値が過小もしくは理論値が過大に評価されている可能性がある。図２では社会保険料合計、医療保険料、年金保険料、介護保険料それぞれについて乖離の分布は左裾で度数が高く、また特に医療保険料や年金保険料では左裾で突出した箇所が見られた。これも理論値が記入値よりも大きい場合に該当する。以下ではこれらが生じたいくつかの背景について考察したい。

加入制度に関する推定の課題

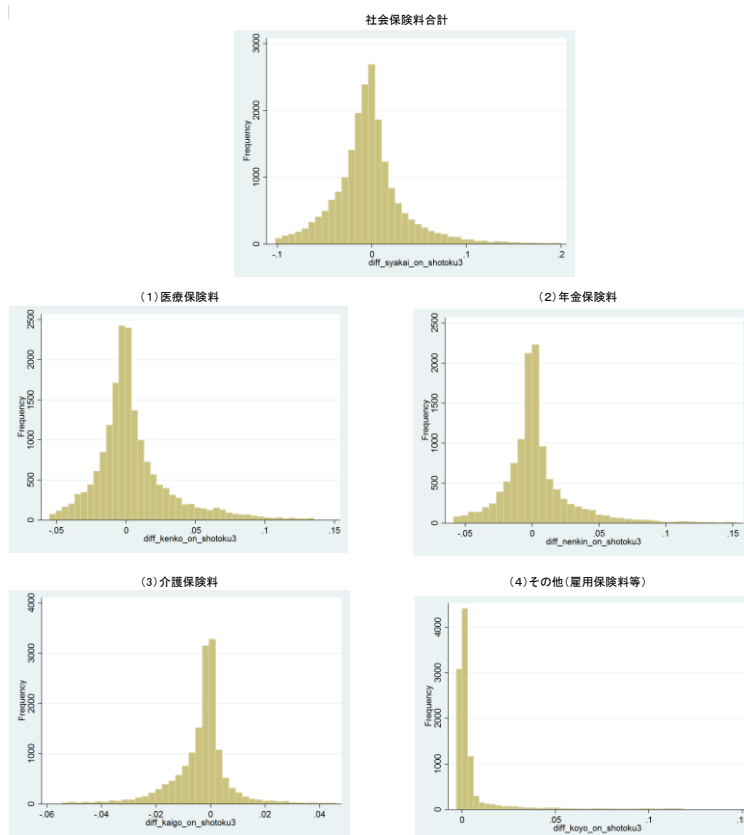
記入値が理論値よりも大きくなる一つの要因としては、加入制度に関する推定の失敗も挙げられる。再び表１を確認すると、「職業も加味した分類」において「調査票に基づく分類」との不一致が高いケースがいくつかある。例えば、実際は医療保険で被用者保険に加入しているにも関わらず国民健康保険と推定している場合（18.7%）や、実際は国民年金第２号であるにも関わらず第１号と推定している場合（16.7%）などである。こうしたケースでは、実際は高い保険料を納めているにも関わらず、相対的に低い保険料負担額を推定してしまう可能性がある。

記入値ゼロの世帯

理論値が記入値よりも大きくなる要因として、理論値はゼロではないが記入値がゼロとなっている世帯（以下、「記入値のみがゼロの世帯」と呼ぶ）が一定の影響を及ぼしている可能性がある。これまでの考察ではサンプルから「記入値および理論値が共にゼロである世帯」を除いているが、ここでは「記入値のみがゼロの世帯」を除いてみる。図３はサンプルから「記入値のみがゼロの世帯」を除いた場合における乖離率の分布を示している¹⁴。その結果、社会保険料合計、医療保険料、年金保険料は概ねゼロを中心として左右対称になり、図２のような分布の左裾における厚みや突出箇所が解消されている。なお、介護保険料においてはやや左側の裾における厚みが残っている。

¹⁴ 理論値がゼロではなく記入値はゼロであるサンプル数については付表１を参照されたい。

図3 乖離率に関する分布：ヒストグラム（記入値ゼロを除く）



（データ出所）『国民生活基礎調査』（平成 22 年調査）

（注 1）縦軸は度数、横軸は乖離率を示す。乖離率＝（記入値－理論値）／世帯所得。

（注 2）サンプルからは記入値が不詳である場合、記入値がゼロである場合を除いている。

また、上下 1 % をカットしている。

次に、「記入値のみがゼロの世帯」を除いた場合における乖離率を記述統計で確認する。表 3 は社会保険料合計および各種保険料について、乖離率の平均、5 パーセンタイル値、95 パーセンタイル値、観測値数を示している。社会保険料合計を見ると、平均はほぼゼロで左右対称であり、散らばりは対所得比 7% 程度である。ここで大野ほか(2015)と同様に所得税および住民税に関する乖離率を計測し、社会保険料の乖離について税の場合と比較してみたい。表 4 はサンプルから記入値が不詳である世帯、記入値がゼロである世帯を除いて、社会保険料および税に関する乖離率の分布を示している。所得税や住民税における乖離の分布は平均がゼロ、散らばりが対所得比 3～4 % 程度であることを踏まえると、社会保険料における乖離は散らばりがやや大きい結果と言える。このほか、医療保険料や年金保険料は、5 パーセンタイル値の変化からも示されるように、サンプルから「記入値のみがゼロの世帯」を除く前には記入値よりも理論値が大きい世帯が多かったが（表 2）、除いた後はそうした記入値よりも理論値が大きい世帯は大きく減少している（表 3）。

表 3 乖離率に関する分布：記述統計

	社会保険料合計	医療保険料	年金保険料	介護保険料	その他保険料	医療＋介護
平均	0.04%	0.63%	0.73%	-0.35%	0.78%	-0.11%
5パーセンタイル値	-6.64%	-3.47%	-3.68%	-2.60%	-0.17%	-5.49%
95パーセンタイル値	7.41%	6.74%	5.84%	1.31%	4.97%	6.45%
観測値数	20,398	17,978	11,890	15,685	10,505	19,770

(データ出所)『国民生活基礎調査』(平成 22 年調査)

(注) サンプルからは記入値が不詳である場合、記入値がゼロである場合を除いている。

表 4 乖離率に関する分布：記述統計 (税との比較)

	社会保険料合計(再掲)	医療保険料(再掲)	年金保険料(再掲)	税	所得税	住民税
平均	0.04%	0.63%	0.73%	0.10%	0.46%	-0.29%
5パーセンタイル値	-6.64%	-3.47%	-3.68%	-4.69%	-2.59%	-3.00%
95パーセンタイル値	7.41%	6.74%	5.84%	6.24%	4.45%	2.77%
観測値数	20,398	17,978	11,890	15,324	15,008	16,597

(データ出所)『国民生活基礎調査』(平成 22 年調査)

(注 1) サンプルからは記入値が不詳である世帯、記入値がゼロである世帯を除いている。

(注 2) 税は所得税と住民税の和を示す。

こうした記入値ゼロをもたらす理由としてはいくつかの要素が考えられる。例えば、税で記入値がゼロである場合、その理由は主に調査票における負担額の記入漏れと考えられる。これに対して社会保険料で記入値がゼロである場合、その理由は調査票における記入漏れのほか、免除や未納といった点も考えられる。これらの理由を明確に識別することは難しいが、大野ほか(2015)の結果から税でも記入漏れが一定程度あることを考慮すると、社会保険料において乖離が相対的に大きいのは免除や未納の存在が大きな影響を与えている可能性がある。

また考察から得られる示唆として、モデルのサンプル・サイズは約 20,000 世帯であることから、乖離率の標本平均の平均ゼロ、分散もほぼゼロ ($=0.07/\sqrt{20000}$) となる。それゆえ、マイクロ・シミュレーション分析などにおける理論値は集計したマクロの値 (例えば保険料収入の合計) については、バイアスがほとんどなく、誤差もほとんどない推計値をもたらすものであり、こうした点は評価できる。

内訳の記入ミス

「記入値のみがゼロの世帯」は単に記入漏れや未納ばかりが原因とは限らず、調査票における内訳の記入ミス（医療保険料、年金保険料、介護保険料、その他（雇用保険料等）に関する負担額の記入ミス）によっても生じているかもしれない。『国民生活基礎調査』における調査票の記入に際して、社会保険料合計については源泉徴収票・確定申告書・住民税の納税通知書からの転記を求める一方、内訳については給与明細書、納入（税）通知書、年金振込通知書等を基に計算させる手法を採っている。記入方法については調査票上で具体的かつ丁寧に説明しているものの、回答者が全員、社会保険料を正確に把握しているとは限らず、また内訳については転記ではなく計算することが必要であるため、内訳を正しく記載できていない調査票も幾らか存在するのではないかと推測される。

特に識別することが難しいのは、医療保険料と介護保険料の違いである。表3からは医療保険料では記入値が理論値よりも大きい世帯が多く、介護保険料では記入値よりも理論値が大きい世帯が多くなっており、このことは介護保険料もまとめて医療保険料の欄に誤記入している可能性も考えられる。表3では医療保険料と介護保険料の和についても乖離率を表記しているが、「医療＋介護」の欄からは平均がゼロに近づいている。こうした結果は、医療保険料と介護保険料を合算してどちらか一方の欄に記入している世帯が一定程度存在していることを示している。

また、同様のことがその他（雇用保険料等）においても生じているかもしれない。図1で示したように、その他（雇用保険料等）の記入値は（雇用保険料から算出された）理論値よりもはるかに大きい値をとる傾向にある。このことは社会保険料合計のうち、内訳の分からないものが文字どおり「その他」として割り振られている可能性もある。表5はその他（雇用保険料等）の記入値が上位10%である世帯に限定して、各社会保険料に関する乖離率の分布を示したものである。その他（雇用保険料等）の分布は平均でプラス4%強の値をとる一方、その他以外の分布は平均でマイナス3%強の値をとり、社会保険料合計の分布は平均でゼロに近い値をとる。こうした結果は、その他（雇用保険料等）にまとめて記入している世帯が一定程度存在していることを示している。

表5 その他（雇用保険料等）の上位10%の乖離率

	その他保険料	医療＋年金＋介護	社会保険料合計
平均	4.12%	-3.20%	0.92%
5パーセンタイル値	0.60%	-10.19%	-3.90%
95パーセンタイル値	11.11%	2.79%	7.66%
観測値数	995	995	995

（データ出所）『国民生活基礎調査』（平成22年調査）

（注）サンプルはその他（雇用保険料等）が上位10%である世帯に限定している。

なお、サンプルからは記入値が不詳である場合、記入値がゼロである場合を除いている。

こうした記入値における記入漏れや未納の存在、内訳の記入ミスによる影響は、全ての社会保険料について記載のある世帯に限定することである程度取り除かれる。表6は医療保険料、年金保険料、介護保険料、その他（雇用保険料等）それぞれにおいて記入値がゼロでない世帯に限定して、これまでと同様に乖離率の分布を確認したものである。社会保険料合計に関する乖離の分布を確認すると、平均はほぼゼロであり、また散らばりは対所得比5%程度とこれまでの計測結果よりも小さくなっており、記入値と理論値に関する乖離は記入値側の問題による影響も大きいことが分かる。

表6 全ての社会保険料に記載がある世帯の乖離率

	社会保険料合計	医療保険料	年金保険料	介護保険料	その他保険料
平均	0.14%	-0.14%	0.07%	-0.09%	0.29%
5パーセンタイル値	-4.30%	-2.85%	-3.38%	-1.03%	-0.18%
95パーセンタイル値	5.60%	3.03%	4.07%	0.69%	1.64%
観測値数	6,436	6,436	6,436	6,436	6,436

（データ出所）『国民生活基礎調査』（平成22年調査）

（注）サンプルは全ての社会保険料に記載がある世帯に限定している。

なお、サンプルからは記入値が不詳である場合、記入値がゼロである場合を除いている。

その他の要因（理論値の課題など）

表6のように、全ての社会保険料について記載のある世帯に限定しても乖離は存在し、内訳を確認すると医療保険料および介護保険料では乖離率がマイナス、年金保険料および雇用保険料では乖離率がプラスとなっている。これらの背景には理論値側の問題による影響も含まれる。

年金保険料について乖離率が上位1%である世帯の記入値と理論値を比較すると、理論値がゼロである世帯が半分以上を占める。本稿では社会保険料の理論値を推定する際に保険料減額の要件を満たす者は全て減額制度を適用するものとして扱ったが、実際には減額制度を利用していない者もいるかもしれない。例えば国民健康保険料では前年所得を考慮した上で、予め減額された額で納付書が送付されてくる。これに対して、年金保険料については軽減措置を受けるためには改めて申請が必要になるため、本来は軽減措置の対象であるが申請をせずに全額支払っている世帯が存在している可能性もある。

医療保険料および介護保険料については、本稿で用いた算出方法は単純すぎるかもしれない。国民健康保険および後期高齢者医療保険では保険料を算出する際に全国一律の保険料率を用いたが、保険料率は自治体ごとに異なることから、保険料負担額を正確に算出できていないこともある。また、本稿では保険料の軽減措置については考慮しているが、減免制度については考慮していない。減免制度とは災害や貧困等の理由により国民健康保険

料等の減額・免除を認める制度であるが、条例で定められるものであるために自治体ごとに異なる。また、健康保険について保険料率は協会けんぽの料率を適用しているが、大企業に勤務する者が加入する健康保険組合では保険料率や負担割合などで独自の運用を行っている場合もある。

その他保険料については、内訳の記入ミスを勘案してもまだ記入値が理論値よりも大きい。業種による保険料率の違いに加え、記入値に雇用主負担分が含まれている可能性や、雇用保険以外で反映すべきものなどについてもさらに検討する余地がある。

VI. おわりに

本稿では『国民生活基礎調査』（平成 22 年調査）の個票データにもとづき、家計の社会保険料負担額（年金保険料・健康保険料・介護保険料・雇用保険料）を推計し（理論値）、それと実際に報告した金額（記入値）の比較を通じて、理論値の妥当性を考察した。

『国民生活基礎調査』の記入値と比較することで理論値の妥当性を評価する手法は、大野ほか(2015)で所得税・住民税負担額について適用されたものであるが、社会保険料では方法論的に大きな違いがあり、それは税と異なり社会保障の負担について加入制度の推定が必要になることである。田中・四方(2012)では世帯員がどの社会保険制度に加入しているかを雇用者所得によって推定していたが、本稿では収入に加え家計の報告した世帯主の職業を用いて推計することで、各世帯の報告する加入制度との一致率は高いものとなった。

各世帯が加入する社会保険制度を特定することができれば、税と同様に制度によって規定される計算式を適用し、各世帯の社会保障負担額を計算することができる。推計された理論値は、実際の記入値と平均的にはほぼ乖離がなく、誤差の分布も対所得比で 7%程度の範囲に収まっていた。これは大野ほか(2015)で計算された所得税・住民税負担額の誤差の分布が概ね 4%程度以内の範囲であったのと比較して大きな誤差である。しかし、個別の世帯でこの程度の誤差であれば、世帯属性毎のグルーピングなど集計をすれば十分に分析に耐えうる精度である。

本稿の分析対象は『国民生活基礎調査』であるが、検証に用いた変数は『全国消費実態調査』でも利用可能な変数である。佐野・多田・山本(2015)や多田・三好(2015)では、『国民生活基礎調査』と『全国消費実態調査』は世帯属性や収入の情報については整合的であることが確認されており、それゆえ本稿の手法は『全国消費実態調査』にも適用可能であり、構築される理論値の妥当性は同程度の精度を持つと考えられる。現在の日本において消費税を含めた家計の税・社会保障負担の全体像を把握するためには消費の情報が豊富な全国消費実態調査等のマイクロ・データに基づいた分析が欠かせないが、その前提として本稿の家計の社会保険料負担額に係る理論値作成の手法は『全国消費実態調査』にも適用可能であり、理論値の精度も保証できることを示したのが本稿の意義である。

参考文献

- Miyazaki, T., and Y. Kitamura(2014), “Redistributive Effects of Income Tax Rates and Tax Base 1984-2009: Evidence from Japanese Tax Reforms”, Discussion Paper Series A No.610, Institute of Economic Research, Hitotsubashi University
- 阿部彩(2000)「社会保険料の逆進性が世代内所得不平等度にもたらす影響」『季刊社会保障研究』 36(1), pp.67-80
- 大石亜希子(2006)「所得格差の動向とその問題点」, 貝塚啓明・財務総合政策研究所 (編著)『経済格差の研究: 日本の分配構造を読み解く』 中央経済社
- 大竹文雄・小原美紀(2005)「消費税は本当に逆進的か—負担の「公平性」を考える」『論座』 第127号, pp.44-51
- 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.40-58
- 大野太郎・中澤正彦・松田和也・菊田和晃・増田知子(2014)「家計の税・保険料負担: 『全国消費実態調査』を用いた計測」『フィナンシャル・レビュー』 第118号, pp.77-94
- 大野太郎・中澤正彦・三好向洋・松尾浩平・松田和也・片岡拓也・高見澤有一・蜂須賀圭史・増田知子(2013)「家計の税・保険料負担: 『全国消費実態調査』『家計調査』『国民生活基礎調査』の比較」PRI Discussion Paper Series No.13A-07, 財務省財務総合政策研究所
- 小塩隆士(2009)「社会保障と税制による再分配効果」, 国立社会保障・人口問題研究所編『社会保障財源の効果分析』 東京大学出版会
- 小塩隆士・浦川邦夫(2008)「2000年代前半の貧困化傾向と再分配政策」『季刊社会保障研究』 44(3), pp.278-289
- 北村行伸・宮崎毅(2013)『税制改革のミクロ実証分析: 家計経済からみた所得税・消費税』 岩波書店
- 佐野晋平・多田隼士・山本学 (2015)「世帯調査の方法と調査世帯の性質—世帯構成、年収、学歴に関する比較—」『フィナンシャル・レビュー』 第122号, pp.4-24
- 白石浩介(2010)「給付つき税額控除による所得保障」『会計検査研究』 42, pp.11-28
- 白石浩介(2011)「消費税の負担水準と逆進性」日本財政学会第68回大会報告論文
- 高山憲之・白石浩介(2010)「わが国世帯における消費税の負担水準」, 一橋大学経済研究所世代間問題研究機構ディスカッション・ペーパー, CIS-PIE DP No.491
- 高山憲之・白石浩介(2011)「給付つき税額控除による消費税負担の軽減」, 一橋大学経済研究所世代間問題研究機構ディスカッション・ペーパー, CIS-PIE DP No.503
- 多田隼士・三好向洋 (2015)「家計収入の把握」『フィナンシャル・レビュー』第122号, pp.25-39
- 田近栄治・古谷泉生(2003)「税制改革のマイクロ・シミュレーション分析」, 小野善康ほか(編)『現代経済学の潮流2003』 第7章, 東洋経済新報社

- 田近栄治・古谷泉生(2005)「年金課税の実態と改革のマイクロ・シミュレーション分析」『経済研究』56(4), pp.304-316
- 田近栄治・八塩裕之(2006a)「日本の所得税・住民税負担の実態とその改革について」, 貝塚啓明・財務省財務総合政策研究所(編)『経済格差の研究: 日本の分配構造を読み解く』, 中央経済社, 第7章
- 田近栄治・八塩裕之(2006b)「税制を通じた所得再分配: 所得控除にかわる税額控除の活用」, 小塩隆士・田近栄治・府川哲夫(編)『日本の所得分配: 格差拡大と政策の役割』, 東京大学出版会, 第4章
- 田近栄治・八塩裕之(2008)「所得税改革: 税額控除による税と社会保険料負担の一体調整」『季刊社会保障研究』44(3), pp.291-306
- 田近栄治・八塩裕之(2010)「税収の確保と格差の是正: 給付付き税額控除制度の導入」, 土居丈朗(編)『日本の税をどう見直すか』, 日本経済新聞出版社, 第2章
- 田中聡一郎(2014)「消費税の低所得者対策の効果分析」, 日本財政学会第71回大会報告論文
- 田中聡一郎・四方理人(2012)「マイクロシミュレーションによる税・社会保険料の推計」, ソシオネットワーク戦略ディスカッションペーパーシリーズ第25号, 関西大学ソシオネットワーク戦略研究機構
- 田中聡一郎・四方理人・駒村康平(2013)「高齢者の税・社会保障負担の分析: 『全国消費実態調査』の個票データを用いて」, 『フィナンシャル・レビュー』第115号, pp.117-133
- 田中秀明(2010)「税・社会保険料の負担と社会保障給付の構造: 税制と社会保障制度の一体改革に向けて」, 一橋大学経済研究所世代間問題研究機構ディスカッション・ペーパー, CIS-PIE DP No.481
- 土居丈朗・朴寶美(2013)「所得税制改革が家計に与える影響: 平成23年度税制改正大綱に関するマイクロ・シミュレーション」KEIO/KYOTO GLOBAL COE DISCUSSION PAPER SERIES DP2011-001
- 府川哲夫(2006)「世帯の変化と所得分配」, 小塩隆士・田近栄治・府川哲夫(編著)『日本の所得分配: 格差拡大と政策の役割』, 東京大学出版会
- 八塩裕之・長谷川裕一(2009)「わが国家計の消費税負担の実態について」『経済分析』182号, pp.25-47
- 矢田晴那(2010)「政策分析ツールとしてのマイクロ・シミュレーションの研究」, PRI Discussion Paper Series No.10A-04, 財務省財務総合政策研究所

付表 1 サンプル数

	社会保険料合計	(1)医療保険	(2)年金保険	(3)介護保険	(4)その他(雇用保険等)	税	(A)所得税	(B)住民税
全サンプル	26,115	26,115	26,115	26,115	26,115	26,115	26,115	26,115
うち、年齢不詳者を含む世帯	26	26	26	26	26	26	26	26
うち、社会保険料・税額不詳者を含む世帯	2,989	2,921	1,974	2,104	1,624	4,886	3,394	3,086
本稿の分析対象	23,100	23,168	24,115	23,985	24,465	21,203	22,695	23,003
うち、①記入値ゼロ、理論値ゼロの世帯	0	0	9,196	3,576	10,542	4,082	5,199	4,278
うち、②記入値ゼロ、理論値ゼロでない世帯	2,702	5,190	3,029	4,724	3,418	1,797	2,488	2,128
うち、③理論値ゼロ、記入値ゼロでない世帯	0	0	485	62	1,644	1,324	1,022	1,197
うち、④記入値、理論値が共にゼロでない世帯	20,398	17,978	11,405	15,623	8,861	14,000	13,986	15,400

(データ出所)『国民生活基礎調査』(平成 22 年調査)