

DP/23-2

経済財政分析ディスカッション・ペーパー

経済動向分析における家計簿アプリデータの更なる活用

小林 周平・鈴木 源一郎

Economic Research Bureau

CABINET OFFICE

内閣府政策統括官（経済財政分析担当）付

本稿は、政策統括官（経済財政分析担当）のスタッフ及び外部研究者による研究成果を取りまとめたもので、学界、研究機関等、関連する方々から幅広くコメントを頂くことを意図している。ただし、本稿の内容や意見は、執筆者個人に属するものである。

目次

1. はじめに.....	1
(1) 調査・分析の目的.....	1
(2) 家計簿アプリデータの特徴.....	1
(3) 令和4年度事業の概要と本稿の方向性.....	4
2. 分析.....	6
(1) 前処理.....	6
(2) 費目分類の正確性.....	6
(3) サンプルの代表性.....	7
(4) アプリへの口座連携の十分性.....	7
① 対処方法の検討.....	7
② 評価.....	9
(5) 日別の動向に関する評価.....	11
3. まとめ.....	14
補論1 日別データの周期性への対処について.....	16
補論2 家計簿アプリデータのその他の有用な性質について.....	17
参考文献.....	19
図表.....	20

経済動向分析における家計簿アプリデータの更なる活用*、†

小林 周平[‡]・鈴木 源一郎[§]

【要旨】

内閣府では、令和3年度以来、家計簿アプリデータを経済動向の把握や政策課題分析に活用する際の一般的な課題を明らかにするために、データの基本的な性質をベンチマーク指標との比較等を通じて確認してきた。本事業の第1報に当たる小林・鈴木（2022）では、家計簿アプリデータを活用する上での課題を、費目分類の正確性、サンプルの代表性、アプリへの口座連携の十分性の三点に整理した。

本稿は、これに続く第2報という位置づけであり、以下の二点が主な内容である。第一に、家計行動を把握するデータとしての精度向上を図る観点から家計簿アプリへの口座連携が不十分なサンプルが少なからず含まれている問題を軽減するための手法について検討を行った。その結果、消費支出や収入の頻度や内容、ログイン履歴を基に条件付け、口座連携の十分性が低いと考えられるサンプルを除外することで、公的統計と比較しても妥当な結果が得られることが明らかになった。第二に、高頻度データとして活用する観点から、日別の収入・消費支出の動向についても観察した。日別動向は月内や週内における周期性の影響が大きく、本稿ではこうした振れの除去、趨勢部分の抽出を、経済動向の把握に十分活用可能な程度まで行う手法を示すことができなかった。しかしながら、周期性の背景は実際の経済行動から説明可能なものであり、データの利用目的に応じた処理を施すことで有用なデータとなりうると指摘した。

本稿での成果は、小林・鈴木（2022）と併せて今後、家計簿アプリデータを経済動向把握や政策課題分析に活用していく上で基礎的な知見となることが期待される。

* 本稿は、「令和3年度『リアルタイムデータを活用した経済動向分析（家計簿アプリデータ活用）』」事業及び「令和4年度『家計簿アプリのビッグデータを一層活用した新たな分析手法の開発』」事業（いずれも株式会社三菱総合研究所への委託調査）において得られた主な成果を報告するものである。本事業には、内閣府からは筆者の他、石井達也前内閣府政策統括官（経済財政分析担当）付参事官（企画担当）、吉中孝内閣府政策統括官（経済財政分析担当）付参事官（企画担当）が参加した。

† 令和3年度事業においては、有識者によって構成される研究会を開催し貴重な意見をいただいた。研究会構成員は、東京大学大学院経済学研究科教授の渡辺努氏（座長）、京都大学経済研究所教授の宇南山卓氏、株式会社ニッセイ基礎研究所生活研究部上席研究員の久我尚子氏、国立情報学研究所情報社会関連研究系准教授の水野貴之氏。また、令和4年度事業においては、京都大学経済研究所教授の宇南山卓氏、慶應義塾大学経済学部教授の星野崇宏氏、横浜市立大学大学院国際マネジメント研究科教授の中園善行氏から有益なコメントをいただいた。また、本稿の作成において、内閣府政策統括官（経済財政分析担当）の村山裕氏、内閣府大臣官房審議官（経済財政分析担当）の松多秀一氏、堤雅彦氏から有益なコメントを頂いた。ここに記して感謝を申し上げる。ただし、本稿に残された誤りはいずれも筆者の責に帰すものである。本稿で示された見解は筆者の個人的なものであり、必ずしも内閣府の見解を示すものではない。

‡ 内閣府政策統括官（経済財政分析担当）付参事官（企画担当）付政策企画専門職

§ 内閣府政策統括官（経済財政分析担当）付参事官（総括担当）付参事官補佐

1.はじめに

(1) 調査・分析の目的

コロナ禍を経て、的確かつ迅速な経済動向の把握や政策課題の分析の重要性が改めて認識された。例えば、「骨太方針 2021」¹においては、「感染症等の社会経済のリアルタイムデータを迅速に収集し、分析能力を向上させ、きめ細やかな政策立案につなげる」こと、「コロナ克服・新時代開拓のための経済対策」²においては、「今後の効果的な支援策等の立案に資するよう、リアルタイムデータを活用した経済動向の分析を行う」こととされた。さらに、「骨太方針 2022」³においても、「デジタル技術も活用し、予算編成プロセスなどでEBPMに基づく意思決定を推進するなど、より機動的で柔軟な政策形成・評価を可能とする取組を進める」ことが盛り込まれた。

このように、オルタナティブデータ活用の機運は一層高まっている。中でも、家計簿アプリデータは、家計の収入・支出の動向を一体的に、高い頻度で把握できるデータであり、今後、さらなる活用が期待されている。一方で、我が国においては現状、家計簿アプリデータを経済動向の分析に活用する上での課題に関する研究事例が十分に蓄積されていない。こうした中、内閣府政策統括官（経済財政分析担当）では、家計簿アプリデータを経済動向の把握や政策課題の分析に活用する際の一般的な課題を明らかにするために、「令和3年度『リアルタイムデータを活用した経済動向分析(家計簿アプリデータ活用)』事業(以下「令和3年度事業」という。)を実施した。小林・鈴木(2022)においては、令和3年度事業の成果の一部が紹介されており、家計簿アプリデータの経済動向把握や政策課題分析に活用する際の有用性を指摘するとともに、残された課題の整理を行っている。

内閣府政策統括官（経済財政分析担当）においては、令和3年度事業において積み残された課題に対処し、家計簿アプリデータをより有効に活用するための基礎的な知見を提供すべく、「令和4年度『家計簿アプリのビッグデータを一層活用した新たな分析手法の開発』事業(以下「令和4年度事業」という。また、令和3年度事業と令和4年度事業を合わせて「本事業」という。)を実施した。本稿では、これまでに行った本事業の概要を紹介するとともに、今後、家計簿アプリデータを活用して経済動向や政策課題の分析を行う上でいくつかの重要な知見を明らかにする⁴。

(2) 家計簿アプリデータの特徴

家計簿アプリは、利用者が保有する銀行口座やクレジットカード、電子マネー等の金融サ

¹ 「経済財政運営と改革の基本方針 2021」(令和3年6月18日閣議決定)

² 令和3年11月19日閣議決定

³ 「経済財政運営と改革の基本方針 2022」(令和4年6月7日閣議決定)

⁴ 本事業においては、2章2節で述べる条件に該当した家計簿アプリ事業者(マネーツリー株式会社)のデータを用いて分析を行ったが、本事業の目的はあくまで家計簿アプリデータを扱う際の一般的な課題について分析、評価することである。

ービス（以下「口座」と総称する。）における取引記録を集約し、一つの画面に表示することで、家計の資産や出納管理をサポートするアプリケーションである。我が国においても複数事業者によって家計簿アプリサービスが提供されている。こうした家計簿アプリの運営を通じて蓄積される「家計簿アプリデータ」には、利用者が行った取引に関するフロー面のデータ（取引の摘要、取引日時、取引金額など）や、資産や負債などストック面のデータが含まれている。

家計簿アプリデータは、クレジットカード利用データや銀行口座データといった他のオルタナティブデータと同様、一般に1) 速報性が高い、2) (十分に普及していれば) サンプルが豊富、3) 記録が自動的で記入漏れ等のリスクが低いといった点で公的統計に対して優れているといえる。また、他のオルタナティブデータと比べると、4) もし、利用者が家計の持つすべての口座を家計簿アプリと連携していれば、家計収支や金融資産・負債の保有状況が網羅的に記録の対象となる。つまり、集計の対象が特定の口座への記録に限定されず、収支の双方について網羅性がありうるという点で優れているといえよう。

他方、公的統計や他のオルタナティブデータと比べたときの課題については以下の三点に整理される。まず、1) 家計簿アプリデータを経済動向分析等に活用するに当たっては、記録されている膨大な取引について経済学的に意味のある分類（支出であれば、消費支出、消費以外の支出、繰越などといった具合）を行う必要がある（「費目分類の正確性」に関する課題）。さらに、2) 他のオルタナティブデータと同様に、サンプルの分布が偏っていることが想定される（「サンプルの代表性」に関する課題）。また、3) すべての、もしくは家計の収支を捕捉するのに十分な数の口座を連携している利用者をいかに抽出するかという点も大きな課題である（「アプリへの口座連携の充分性」に関する課題）。

こうした課題に関する評価や対処方法の研究を行うために、内閣府では令和3年度事業を実施した。その中では、宇南山（2019）において紹介されているR I C Hプロジェクトの手法を参考に、家計簿アプリ利用者にアンケート調査を実施し、「属性付きデータセット」を構築している⁵。なお、アンケート調査は、アプリ内通知や家計簿アプリ事業者からのメール配信で参加を呼び掛けており、また、参加者には家計簿アプリに三つ以上の銀行口座又はクレジットカード（内、一つ以上は普通預金口座）、世帯の主たる収入の受取口座を紐づけていることを求めていることから、属性付きデータセットのサンプルは一定以上の利用密度で家計簿アプリを利用している利用者に限定されることになる⁶。令和3年度事業にお

⁵ 家計簿アプリデータには、利用者の年齢や世帯類型といった属性情報が十分に紐づいていない。そこで、別途アンケート調査を実施することで、属性情報等を取得し、これと照らし合わせながら分析を実施できる環境を構築した。本稿では、本事業において実施したアンケート調査の回答が紐づいたデータセットを「属性付きデータセット」、アンケート調査による結果が紐づいていない大多数の利用者のデータを含む、家計簿アプリ事業者が管理しているデータセットを「事業者管理データセット」と呼ぶ。

⁶ このほか参加者には、1) 家計簿アプリ事業者に蓄積されている取引記録を内閣府及び業務委託先における分析に利用することに同意すること、2) すでにアンケートに回答しているアカウントに登録されている口座と同一の口座を連携していないこと（例えば、ある人物や家計が複数のアカウントを保有してい

いては、属性付きデータセットについて詳細に分析するとともに、アンケート調査による結果が紐づいていない「生の」データセット（「事業者管理データセット」）を利用した分析も行った⁷。小林・鈴木（2022）は、ここでの成果の一部について紹介しており、手法面に改善の余地は大きいものの、属性付きデータセットを用いた集計結果については、総じてみれば、公的統計の収入や消費支出の動向をおおむね捕捉できると結論付けている（主な検証結果は図表1-2-1のとおり。）。その下で、小林・鈴木（2022）で確認した課題について、以下のとおり改めて整理する。

まず、費目分類の正確性について、収入は、少なくとも検証の中心とした勤労者世帯では公的統計の水準と比べて大きな違いはなく、季節性も捉えられていることから一定の精度で取引を識別できたとしている。また、消費支出については、水準が過大となっている可能性を否定できないものの、収入の水準と比べて極端な結果にはなっておらず、今後の改善次第では経済動向把握等に活用可能な指標となり得ると指摘している。

サンプルの代表性については、まず、世帯主が高齢者の世帯割合が我が国の世帯分布と比べて圧倒的に小さいため、総世帯ではなく、現役世代が中心である勤労者世帯を検証対象の中心とせざるをえないとしている。その上で、年齢については若年層、年間収入については高所得者、居住地域については関東地方にサンプルの属性が集中している。総じてみると、年間収入の偏りに起因して収入、消費支出等の集計時に生じるバイアスが最も大きく、高所得者層に利用者が比較的偏っていることによって、収入や消費支出の平均値が公的統計対比で高めになりやすいが、こうしたバイアスは、特に年齢や年間収入を軸としたウェイト・バックによって対処できると指摘している。

最後に家計簿アプリへの口座連携の十分性については、収入については有業人員数が複数になると世帯収入の内、家計簿アプリで捕捉可能な割合（捕捉率）が低下する傾向があること、消費支出については、有業人員数の増加に加え、世帯人員が多い場合に捕捉率が低下することを指摘している。また、口座連携の十分性が通時的に変化（典型的には口座連携数が次第に増加）している可能性についても確認しており、収入についてはこの影響が確認できる。消費支出については、これに加えてマクロ的なキャッシュレス化の進展の影響も現れているが、ATM引出しを考慮すれば影響は緩和されるとしている。

他方、事業者管理データセットについては、収入や消費支出の集計値を公的統計と比較、評価したところ、前年同期比や前期比のような変動については属性付きデータセットと同様におおむね再現できていると評価できるものの、口座連携の十分性が高く、家計簿アプリに世帯全体の経済行動を十分に反映させているサンプルが限定されていると考えられることから、特に水準でみると集計結果に大きな影響を及ぼしているとは指摘している。

る場合など）を求めている。

⁷ 令和3年度事業の実施体制やアンケート調査の概要などについては、小林・鈴木（2022）を参照されたい。

(3) 令和4年度事業の概要と本稿の方向性

家計簿アプリデータは、家計の収入・支出が網羅的かつ自動的に記録されうるという他のオルタナティブデータにはない特徴を有しており、小林・鈴木（2022）では、家計簿アプリデータに一定の有用性があることを明らかにした。一方で、実際に経済動向把握や政策課題分析に活用する上では、いくつかの対処すべき課題が残されている。

まず、属性付きデータセットの活用について考える。サンプルの代表性に関して残された問題としては、家計簿アプリ利用者が比較的少ない高齢者世帯に関して精度の高い集計を行えなかったことが挙げられるが、この点については、今後、ウェイト・バックを行うのに十分な程度に高齢者世帯において家計簿アプリが普及していくことを期待したい。また、費目分類の正確性については、消費支出について水準が過大となっている可能性を否定できないものの、小林・鈴木（2022）においてもすでに、取引明細（摘要欄）に注目して、住宅投資やローン返済とみられる取引を除く処理等を行っており、機械的な識別のみでは大幅な精度向上は難しい⁸。一方で、アプリへの口座連携の十分性については、データセットの中で世帯の主たる収入・支出を記録した口座を連携している利用者を特定、抽出し、平均的に口座連携の十分性を高める余地がまだ残されていると考えられる。また、事業者管理データセットによる分析について考えると、連携口座数でも、記録されている明細数でも、属性付きデータセットと比較して明らかに口座連携の十分性が大きく劣るサンプルが多いことが最も優先して対処すべき課題であるといえよう。

こうした問題認識を踏まえ、令和4年度事業では、まず、データセット中の口座連携の十分性が高い利用者を効率的かつ効果的に抽出する手法に関する研究に取り組んだ。さらに、こうして抽出したサンプルについて収入や消費支出の動向を日次単位で確認できるため、家計簿アプリデータを用いて月内や週内における特徴的な動向を観察することで、アプリデータ活用時の基礎的な知見を得ることを目指した。

なお、令和4年度事業は株式会社三菱総合研究所に委託して実施した。また、内閣府において定めていた要件⁹を満たす家計簿アプリ事業者であるマネーツリー株式会社¹⁰に一部の

⁸ 水準が上振れしている要因として、例えば、本来はその世帯の消費支出ではない立替払いや贈与等の支出も含めて消費支出と識別していることなどが挙げられる。

⁹ 要件は以下のとおり。

- ・個人を対象に、アカウント・アグリゲーション技術により複数の金融機関の口座の取引情報等を集約して、自動的に一つの画面に表示する技術によって、家計の資産管理を支援するアプリケーションを提供している事業者であること
- ・事業の開始が2015年1月1日以前であり、それ以降のデータを保有していること
- ・アンケート調査に際して、利用者に直接的にアプローチする手段を有すること
- ・利用者本人の許諾を前提に、事業者において管理する当該利用者の取引データを提供できること
- ・現在までのダウンロード数が100万以上であり、属性付きデータセットの作成業務において、有効な回答が1万程度、確保できる見込みがあること
- ・国際規格「ISO/IEC27001:2013」の認証を取得していること
- ・本事業への協力について承諾が得られており、契約締結後速やかに業務を開始できること
- ・その他、受注者が業務内容を行う上で、十分な体制を整えていること

¹⁰ マネーツリー株式会社が提供する個人資産管理アプリ「Moneytree」は、国内2,500以上の金融機関の

業務を再委託している。業務の内、属性付きデータセットによる分析については、内閣府及び株式会社三菱総合研究所において、事業者管理データセットによる分析については、マネーツリー株式会社において実施した。ここで属性付きデータセットは、紐づけられた家計簿アプリデータを本事業における分析に利用することについて個別に許諾を得た利用者のデータのみが含まれているほか、いずれのデータセットにも個人を直接的に識別可能なデータは含まれていない。

本稿は、本事業を通じて得られた成果の一部を報告するものであり、構成は以下のとおりである。2章で実際に行った分析についてその概要、3章で結論を報告する。

口座や電子マネー、ポイントカードなどのサービスと連携可能であり、一つの家計簿アプリアカウントにつき、50口座まで無料で登録できる。金融口座における受取、支払の動向について自動的に記録されるほか、現金による支払も手動で記録できる。さらに、受取、支払については自動的に「給料」、「食費」といった Moneytree 独自の費目分類「Moneytree Category List」に仕分けされる。また、性別、年齢を始め、属性情報については一切紐づけられていない。なお、同社によれば 2022 年 9 月時点の利用者数は 550 万人以上である。

2. 分析

(1) 前処理

本章では、データセット中の口座連携の十分性が高い利用者を効率的かつ効果的に抽出する手法について検討する。さらに、こうして抽出したサンプルについて収入や消費支出の動向を日次単位で確認して、月内や週内における特徴的な動向を観察することで、集計値への基礎的な理解を深める。

まず、本節ではこうした検討に先立って行った、異常値を取り除く処理（以下「前処理」という。）の手法及び収入や消費支出といったフロー指標の構築方法について説明する。前処理は、おおむね、小林・鈴木（2022）と同様に、Moneytree 独自の費目分類「Moneytree Category List」（以下「事業者カテゴリー」という。）やアンケート回答結果を基に、図表 2-1-1 の流れに沿って 3 ステップで実施した^{11,12}。

(2) 費目分類の正確性

費目分類についても、事業者カテゴリーを基に、小林・鈴木（2022）と同様に実施した。まず、収入と支出の分類を図表 2-2-1 のとおり行う。その中から、実収入については、事業者カテゴリーにおける「給料」と「収入」の和と定義した（図表 2-2-2）。また、支出については Kaneda et al. (2021) を参考に、2 つの定義で消費支出を集計した（図表 2-2-3）。消費支出①は確実に消費であるとみなせる事業者カテゴリーの取引情報のみ（ただし、摘要欄の文言から住宅資産の取得やローンの返済と機械的に判定できる取引情報は除外）、消費支出②は、消費支出①に A T M 引出しのみを加えた指標である¹³。なお、小林・鈴木（2022）によれば、消費支出①と消費支出②を比較すると、消費支出②の方が概して安定的に推移していたことを踏まえ、消費支出②を中心に分析を行う¹⁴。

¹¹ 小林・鈴木（2022）からの変更点としては、以下の 2 点が挙げられる。

- ・小林・鈴木（2022）においては、住宅購入は消費支出に含まれないことを踏まえ、支出データから事業者カテゴリーが「住宅」かつ±1,000 万円以上の取引を除外していたが、本稿では「住宅」カテゴリーに含まれない場合でも、±1,000 万円以上の取引は資産の購入等に該当するケースが多いと考えられ、集計時のノイズにもなることから、すべて除外することとした。
- ・小林・鈴木（2022）においては、現金の受払による取引（利用者が逐一入力することで登録される。）を除外することはしなかったが、この場合、A T M 引出しと現金による支払が二重計上されていた可能性がある。本稿では後述するとおり、主に A T M 引出しを含む消費支出②を分析の対象とすることを踏まえ、現金の受払による取引をすべて除外することとした。なお、小林・鈴木（2022）の図表 3-1-36 で示したとおり、本データセットにおける現金の受払による消費支出の割合は消費支出の定義に依らずわずかである。

¹² 属性付きデータセットについて、こうした前処理を行った後のサンプル数は 6,425。

¹³ 小林・鈴木（2022）では、消費支出②に、さらに振替などの消費支出以外を含むと考えられる事業者カテゴリーを加えた消費支出③を構築していたが、消費支出③は他の定義と比べて公的統計との乖離度が水準でも変動でも大きいため、本稿では分析の対象としない。

¹⁴ この背景をより子細に考察すると、現金決済は家計簿アプリに記録されづらいため、消費支出①はキャッシュレス決済が全国的に普及した中でキャッシュレス決済による取引が増加したため上昇トレンドが観察された一方、消費支出②には A T M 引出しが含まれており、現金決済とキャッシュレス決済の双方を捉

(3) サンプルの代表性

属性付きデータセットについては、同時に行ったアンケート調査を基に家計簿アプリ利用者の属性情報を得ることができる。1章でも触れたように、年齢については若年層、年間収入については高所得者、居住地域については関東地方に集中している点が顕著であり、こうした分布上の特徴の中でも、年間収入の偏りに起因して収入・消費支出等の集計値で生じるバイアスが最も大きい。

また、事業者管理データセットについては属性情報が紐づいていないので、分布上の特徴について明確なことはわからない。ただし、以降の分析で口座連携の十分性が高い利用者を抽出していけば、属性付きデータセットと同様に、若年層、高年収、関東地方居住者といった属性への偏りが見出せると期待される¹⁵。

小林・鈴木(2022)を踏まえれば、いずれのデータセットについてもサンプルの偏りが収入や消費支出の前月比や前年同月比に与える影響は限定的と考えられるが、水準については、概してサンプルの偏りが上振れ要因になることに注意して結果を解釈する必要がある。

(4) アプリへの口座連携の十分性

① 対処方法の検討

小林・鈴木(2022)においては、主に属性付きデータセットを用いて、口座連携の十分性の低い世帯属性の特定や通時的な口座連携の十分性の変化の影響評価を目的に分析を行った。しかしながら、こうした口座連携の十分性のばらつきが収入や消費支出の集計値に及ぼす影響については、前月比や前年同月比といった変動への影響が限定的であることを指摘するにとどめ、水準として下振れしている可能性に有効な対処方法を見出すことができていない。そこで、本節では、データセットの中で世帯の主たる収入・支出を記録した口座を連携している利用者を抽出し、平均的に口座連携の十分性を高める手法について議論する¹⁶。具体的には、支出対象や頻度などの観点で「標準的な」家計の経済行動を想定して、そこから逸脱したサンプルを除外する方法を中心に検討していく¹⁷。図表2-4-1では、そ

えうするため、上昇トレンドが緩和されたと考えられる。

¹⁵ 属性付きデータセットに含まれる利用者は1章で確認したとおりアンケート参加条件を満たしている必要があることから、口座連携の十分性が相対的に高い利用者の割合が高いと想定されるため。

¹⁶ これまでみてきたように、家計簿アプリデータには口座連携の十分性以外にも費目分類の正確性やサンプルの代表性について課題があり、これらの影響を切り出して評価することは難しい。それゆえ、例えば各データセットの集計値が公的統計(総務省「家計調査」など)と近いからといってすなわち口座連携の十分性が高いという評価をすることはできないことに注意が必要である。

¹⁷ サンプル除外の条件を設ける際には、本来含めるべきサンプルを誤って除いてしまう可能性と、本来除くべきサンプルを誤って含めてしまう可能性を比較して検討する必要があるが、本稿では前者の誤りを最小化することを重視して条件を設定している。どういった基準が適切であるかは分析の目的によって都度検討されるべきであり、本稿で採用した手法は一つの考え方に過ぎない。また、公的統計等の分布に基づいて「本来の」分布に近づくように重みづけするといったような手法も考えられる。ただし、今回のように口座連携の十分性以外にも費目分類の正確性やサンプルの代表性について課題がある場合には何をもっ

の結果（2021年度における集計値）を整理している¹⁸。

まず、消費支出の頻度を検討する。家計簿アプリに消費支出が記録される頻度が非常に低いサンプルは、家計の経済行動を家計簿アプリに十分に記録していないサンプルであると考えることができ、その閾値を設定したい。ここで、ある程度独立して生計を営んでいる家計を考えると、週に1度も消費支出をしないことは想定しづらいので、この条件として週に1度以上の消費支出の記録を付すことも候補となる。しかし、小林・鈴木（2022）で確認したとおり、家計簿アプリでは都度の現金支払は記録されづらいが、源泉となるATM引出しから総額としては見積もることが可能である。そこで、本稿では一定の期間で均して家計の消費行動が十分に捕捉されているサンプルも集計の対象とすべきとの考えの下、分析期間（2019年4月以降）中、各月に1件以上、消費支出の取引が記録されていることを条件として加える。

次に、消費支出の対象を検討する。すなわち、どの家計でも確実に支出していると考えられる費目が家計簿アプリ上に記録されている利用者は口座連携の十分性が高いと想定される。本稿では、事業者カテゴリーの「水道光熱費」が記録されているかに注目したい。なお、家計簿アプリ上の事業者カテゴリーと公的統計における費目分類は対応しないことが多いが、光熱水道費に関しては、公的統計の「光熱・水道」とカバレッジが一致している。その上で、光熱水道費に注目するのは、光熱水道費を支出しない家計はほとんど存在せず¹⁹、また、光熱水道費は口座振替やクレジットカードによる支払が年代や地域を問わず普及していることから家計簿アプリに記録されやすく、もし、家計簿アプリ上に記録されていなければ、家計の経済行動が家計簿アプリで十分に記録されていないと考えられるためである。よって、本稿では、2021年度に光熱水道費が記録されていることも条件に加える。

さらに、収入の頻度・金額についても検討する²⁰。消費支出と同様に一定以上の頻度で収入が記録されているサンプルを抽出することとしたいが、収入を得る頻度は就業形態によって異なるため、特定の類型の世帯が除外されてしまうリスクに注意する必要がある。例えば、会社や官公庁に勤めている勤労者のみを考えれば、月1回以上、収入の取引が記録されていることが自然であるが、年金支給は2カ月に1回であるし、自営業主などであれば収入が得られる頻度はあまり規則的ではないかもしれない。また、失業などで一時的に収入が得られない場合もあるだろう。本稿では、まずは可能な限り多様な類型の世帯を分析の対象と

て本来の分布とするか議論の余地が大きく、本稿では議論しない。

¹⁸ すでに前処理において条件としている3口座以上の連携を4口座以上に変更することも考えられるが、例えば給与受取口座を含む3口座を連携していれば、当該家計におけるすべての経済行動を捕捉することも可能であり、本稿では連携口座数について追加的に条件を付すことはしなかった。さらに、一口に口座と言っても、普通預金口座から投資信託、電子マネー、ポイントカードと多様であり、この閾値をどの程度まで引き上げれば家計の経済行動の主たる部分を捕捉できるかは先験的に明らかでない。

¹⁹ 総務省「全国家計構造調査」において、月当たりの光熱・水道支出金額（2019年10月、11月）が0円の世帯は総世帯でみて約0.7%である。

²⁰ 消費支出は収入を源泉に行うものなので、収入が十分に捕捉されているサンプルについては消費支出についても捕捉可能性が高まると考えられる。

すべく、分析期間中の各年度²¹に1件以上という条件を採用した。ただし、収入取引の明細をみるとポイント還元など生活の主たる原資でないと考えられる取引が含まれている。そのため、ここでは金額が1,000円以上の収入取引という条件で判定した²²。

最後に、家計簿アプリが家計の資産や出納管理をサポートするアプリケーションであることに鑑みれば、起動ログが記録されている利用者は比較的、利用密度が高いと考えられる。本稿では、直近約1年間²³に起動ログが1回以上記録されていることを条件に付す²⁴。

以上でみてきた条件を付すと、事業者管理データセットについて小林・鈴木（2022）での報告値と比べて、消費支出②は18.9万円程度、また、収入は9.6万円程度水準が上昇する。以降では、これらの条件をベースラインとして付したときの口座連携の十分性について議論を進めていきたい。

② 評価

ベースラインの条件で抽出を行った結果について時系列推移を示すと図表2-4-2、図表2-4-3、図表2-4-4のとおりである。収入、消費支出①、消費支出②のいずれも水準としては小林・鈴木（2022）の結果から概して平行に上方シフトしている。また、消費支出②に関しては属性付きデータセットの集計値と事業者管理データセットの集計値の水準がかなり近くなった。

まず、属性付きデータセットについて詳細にみていきたい。属性付きデータセットには、別途アンケートで得られた年間収入額のデータが利用可能であり、収入については家計簿アプリから得られた年間収入とよく対応していれば、口座連携の十分性が高いと推測することが可能である。図表2-4-5（左図）は横軸にアンケートから得られた年間収入、縦軸に家計簿アプリから得られた年間収入をプロットしている。ベースライン条件で抽出を行うことで、傾向線から大きく下側に外れたサンプルが除外され、プロットの散らばりが小さく、また、傾向線そのものも上方にシフトした。さらに、45度線と傾向線との乖離は、口座連携の十分性が低い可能性があることのほかに、家計簿アプリから得られる年間収入が税や社会保険料が源泉徴収された後のいわゆる手取り収入であることによって説明できるが、前者の要因を切り出して評価するために公的統計との比較を行う。図表2-4-5（右図）は、年間収入十分位階級ごとに、横軸に年間収入、縦軸に公的統計については可処

²¹ 2019、2020、2021の各年度。

²² このほか、2020年、2021年には「特別定額給付金」や「子育て世帯臨時特別給付金」をはじめとする給付金が支給されており、収入取引にはこうした臨時、特別な受取も含まれる。しかしながら、こうした受取を逐一判別して除外することは作業上困難である上、ここに挙げた給付金については、世帯主の口座に振り込まれることとなっており、こうした取引を収入としてカウントしても、口座連携の十分性の高いサンプルを特定する上でのノイズになるとは考えづらい。

²³ 2022年1月以降、データを抽出した2022年11月4日まで。

²⁴ 属性付きデータセットはアンケート調査参加時に家計簿アプリを起動する必要があるため、すべてのサンプルについてこの条件が満たされている。

分所得²⁵、属性付きデータセットについては家計簿アプリから得られた年間収入をプロットしている。これによれば年間収入 700 万円あたりまでは公的統計と家計簿アプリのプロットがよく一致する。それ以上の年収帯については乖離が広がり²⁶、口座連携の充分性が低いことの影響が現れていると考えられるが、サンプル抽出を行うことでその影響は緩和される。

続いて消費支出についてみていく。図表 2-4-6 (左図) 及び図表 2-4-7 (左図) では、横軸に消費支出の決定要因として代表的な年間収入 (アンケート結果)、縦軸にそれぞれ消費支出①、消費支出②をプロットした。図表 2-4-5 (左図) で収入についてみたのと同様に、傾向線から大きく下側に外れたプロットの多くが除外され、散らばりが小さく、また傾向線が上方にシフトした。消費支出についてはアンケート結果がないため、口座連携の充分性については公的統計との比較によって評価を行うほかない。図表 2-4-6 (右図) 及び図表 2-4-7 (右図) では、図表 2-4-5 (右図) と同様に年間収入十分位階級の別に年間収入と消費支出をプロットしている。属性付きデータセットにおいては高所得者世帯割合の高さが集計結果に顕著な影響を及ぼしていたが、こうすることにより、所得階級をそろえた消費支出の水準について視覚的な評価が可能になる。ここで、消費支出の定義を振り返ると、消費支出①は現金支出が集計の対象でないなど、記録漏れが生じていると考えられること、消費支出②はその家計の消費支出として支出されないATM引出しも含まれていると考えられることから、「消費支出①<公的統計<消費支出②」となることが望ましいが、サンプル抽出を行うことでこうした姿に近づいたことが確かめられる。ただし、ここでは、どれだけ精緻にサンプルを抽出したとしても、平均的にみて口座が完全に連携されている状況はあり得ないとすれば、ここで示した結果において、「消費支出①<公的統計<消費支出②」という条件が満たされたのは、口座連携の不十分性による下振れが、公的統計と比べたときの費目分類の不正確さによる上振れによって相殺されただけに過ぎない事象とも考えられることに注意が必要である²⁷。とはいえ、もちろん小林・鈴木 (2022) で指摘したような世帯属性による口座連携の充分性の違いをよく認識しておく必要はあるが、実際に経済動向の把握や政策課題分析を行う実用面から評価すれば、サンプル抽出によってより実態に近いと考えられる集計値が得られるようになったといえる。

次に、事業者管理データセットについて考えたい。以上で行った属性付きデータセットに関する評価を要約すると、サンプル抽出を施すことで口座連携の充分性が向上し、その集計値は、我が国全体の収入や消費支出を捉える上で、よりもっともらしいものになったというものだった。これを踏まえると、事業者管理データセットのサンプルについて、デモグラフ

²⁵ ここでは、実収入から勤労所得税、個人住民税、社会保険料を控除した、手取り収入と対応する部分を可処分所得と呼ぶこととする。

²⁶ 小林・鈴木 (2022) の議論を踏まえれば、主に共働き世帯が増えることが影響していると考えられる。

²⁷ ただし、小林・鈴木 (2022) でも触れられているように、家計への調査に基づく公的統計が過小推計となっている可能性も指摘されているが、その点に係る検証は本稿の範囲を超えている。

イック、非デモグラフィックの双方で世帯属性が属性付きデータセットと大きく違わないという前提の下ではあるが、事業者管理データセットに属性付きデータセットと近い特徴が現れれば、事業者管理データセットについても適切なサンプル抽出の条件が設定できたという評価を下せることになる。図表2-4-8、図表2-4-9では、両データセットについて、サンプル抽出後の収入、消費支出の取引数や金額の分布を示している。これによれば、消費支出については事業者管理データセットの方が若干下振れしているが、分布としてはかなり似通っている。一方、収入については、事業者管理データセットについて収入200万円以下の層がかなり厚くなっており、改善の余地があり得る。ただし、いずれにしても小林・鈴木（2022）で報告したような、極端な分布の違いはみられず、口座連携の十分性としては大きく改善していると評価できる²⁸。

最後に前月比や前年同月比といった変動について確認したい（図表2-4-10、図表2-4-11、図表2-4-12）。公的統計との乖離をRMSEで評価すると、サンプル抽出条件を付し、口座連携の十分性を改善してもほとんど変化しない（図表2-4-13）。小林・鈴木（2022）でも指摘したが、変動だけを利用するのであれば口座連携の十分性について、あまり深刻な問題と捉えなくてもよいと考えられる。

（5）日別の動向に関する評価

これまで主に月次ベースで集計して評価を行ってきたが、家計簿アプリデータの利用価値の一つとして、より細かい時間的粒度で分析が可能なが挙げられる。本節では、日別の動向を観察し、その特徴を示す。また、ここまでは、家計簿アプリデータを精緻に集計していけば、究極的には公的統計と同様の水準・変動を示すはずとの前提で検討を進めてきたが、例えば、口座振込みによる支払やATM引出しは家計簿アプリ上の記録と実際の消費支出のタイミングがずれるので、年末年始や大型連休の周辺などで家計簿アプリデータ特有の季節性が生じている可能性がある。このように細かい時間的粒度でデータの特徴を把握することはより粗い時間的粒度で分析をする際にも重要となる。

始めに、家計簿アプリデータを日別に集計した結果を確認する²⁹（図表2-5-1）。収入、消費支出ともに周期的な変動が大きく、この結果だけでは経済動向の把握等を行うことは困難である。図表2-5-2は、過去3年間について月内の日別動向を平均したものであ

²⁸ サンプル抽出前の分布は小林・鈴木（2022）の図表3-2-3及び図表3-2-4を参照のこと。

²⁹ 収入や消費支出の変動を評価する際には、intensive margin（前期に収入（消費支出）があったサンプルについて今期の収入（消費支出）がどの程度増加しているか）と、extensive margin（収入（消費支出）のあったサンプル数が前期から今期にかけてどの程度増加しているか）を区別することが重要である。月次の動向をみる際には、収入（消費支出）の記録があるか（extensive margin）は、実際に収入（消費支出）があったかではなく口座連携の十分性の変化など、家計簿アプリが家計を捕捉する程度が変化したことの影響が大きいと考えられるため、月内に収入（消費支出）が1円以上あったサンプルを対象として集計を行っていたが、日次の動向をみる上では、その日に収入（消費支出）があるか（extensive margin）も重要であるため、すべてのサンプルを対象に集計を行っている。ただし、4節で検討したとおり、本稿で用いるサンプルは月に1件以上消費取引が記録されているサンプルに限られている。

る³⁰。収入については25日が特異日となっており、これは給与の支払日が25日であることが多いためではないかと考えられる。次いで10日も高水準であるが、6月、12月に支払われる賞与の支給日に当たることが多いためと推測できる。消費支出についても確認すると、消費支出①、消費支出②ともに、25日から月初めにかけて高めの水準で推移し、特に27日は顕著に高い。総務省「家計調査」の日別支出では、月内で特段の傾向が確認されないことを踏まえると、家計調査の日別支出に含まれない月ぎめ払いの支出が25日以降に集中しているためと考えられる。実際、消費支出を事業者カテゴリーの別に確認すると、25日以降、とりわけ27日の増加分はほとんどが、家賃支払などを含む「住宅」で説明される（図表2-5-3）。また、ATM引出しは25日以降に増加する傾向にある。

次に、曜日別の平均値により週内の動向を確認するが、月内の動向と比べれば変動は小さい（図表2-5-4）。収入については、給料日が休日に当たった場合、休日を避けて前倒しして支給されることが多いため、平日、とりわけ金曜日の水準が高い。一方、消費支出については平日の支払が多い。家計調査の日別支出では土日に支払が増える傾向にあるが、月ぎめ払いの引き落とし日は週末を避けて行われることが影響していると考えられる。また、引き落とし日が休日に当たった場合は後ろ倒しして引き落とされる場合が多いため、平日、とりわけ月曜日の水準が高くなる。実際、事業者カテゴリーの別に確認すると、やはり「住宅」の変動が大きい（図表2-5-5）。また、ATM引出しも平日（特に月曜日、金曜日）に多く休日に少ないが、休日は手数料が高いためと考えられる。

収入は多くの場合月1回の支給だが、消費支出については、こうした周期性を取り除いて日次の動向を趨勢として確認できれば、政策や自然災害等が人々の行動に与える影響を評価する際に有用な情報になる。ここでは簡単な方法として、後方7日移動平均を取ることで特異的な動きの影響を均すとともに、週内の周期性を除去している（図表2-5-6（上図））。これをみると、月内の周期性に起因する波の影響が大きいのが、例えば、初めて緊急事態宣言が発出された2020年4月から5月にかけての消費支出の落込みが確認できる。さらに、図表2-5-6（下図）では、季節性³¹の影響を取り除くために前年同日比の動向を示している。家計調査の日別支出とは、カバレッジの違いもあるほか、月ぎめ払いの引き落とし日が毎年の曜日並びによってずれてしまうことなどから、ノイズは大きいものの、2020年4-5月の落込み³²のほか、2回目の緊急事態宣言が発出された2021年初めも消費支出の落込みが観察できる（図表2-5-7（1）、図表2-5-7（2））。

このように後方7日移動平均というような簡易でわかりやすい処理によって、一定程度、

³⁰ 以降においては主に2019~2022年度のデータを用いて周期性等の性質を確認するが、この間のデータは新型コロナウイルス感染拡大の影響から不規則な動向を示している可能性がある。本節で行った分析については今後、コロナ禍後のデータの蓄積を待って再度検証されるべきである。

³¹ ここでの季節性には月内の周期性も含まれる。

³² 2020年4月末から5月初めに、一時的に前年同期比が大幅なプラスになっているが、これは2019年の同時期は改元に伴う10連休となっており、月ぎめ払いの引き落としがなされなかったことで消費支出が減少したためであると考えられる。

高頻度の動向把握が可能になったものの、必ずしも経済動向の把握に十分活用可能な程度まで行う手法を示すことができなかった。しかしながら、周期性の背景は実際の経済行動から説明可能なものであり、データの利用目的に応じた処理を施すことで有用なデータとなりうるだろう³³。

³³こうした課題への対処方法について、補論1において追加的な議論を行う。

3. まとめ³⁴

本稿では、小林・鈴木（2022）で行った、家計簿アプリデータを経済動向分析や政策の効果分析に活用する上での課題や対処方法に関する議論を引き継いで、第一に、家計行動を把握するデータとしての精度向上を図る観点から家計簿アプリへの口座連携が不十分なサンプルが少なからず含まれている問題を軽減する手法について検討した。本稿では消費支出や収入の取引頻度や内容、ログイン履歴を基にサンプルを条件付け、口座連携の十分性が低いと考えられるサンプルを除外した。そうすることで、特に事業者管理データセットにおいて集計値の水準が大きく上昇し、消費支出については公的統計と比較しても妥当な結果が得られるようになった。本稿で採用した手法は一つの考え方に過ぎず、今後、事業性資金や家計内での振替など関心のある家計行動以外が記録されているサンプルの除外と並行して、家計行動を把握するのに十分な利用密度のあるサンプルをより精緻に特定する余地は大きい。ただし、本来含めるべきサンプルを誤って除いてしまう可能性を踏まえれば、これ以上の絞り込みの是非は分析の目的に応じて都度、検討されるべきであろう。例えば、分析の対象を勤労者世帯のみと割り切るのであれば、収入に関する条件をより厳しくすることも考えられるが、失業など家計の急変の影響を捉えたいのであれば不適切である。また、本稿では、特定の属性値に対して閾値を設定することによってサンプルを除外するアプローチを採ったが、あるいは、それらの属性値について、本来の分布に近づくように重みづけするといったような手法も考えられる。

第二に、収入、消費支出に関して日次、週次といった高頻度データとしての活用の可能性を検討した。コロナ禍において緊急事態宣言等の影響を探る観点から高頻度での動向把握が求められたことを踏まえれば、特に消費支出に関する高頻度指標への期待は高い。日別の収入、消費支出の動向について観察したところ、月内、週内の周期性の影響、特に消費支出について月ぎめの支払が特定日に集中する影響が大きいことが明らかになった。本稿ではこうした振れの除去、趨勢部分の抽出を、経済動向の把握に十分活用可能な程度まで行う手法を示すことができなかつたものの、周期性の背景は実際の経済行動から説明可能なものであり、データの利用目的に応じた処理を施すことで有用なデータとなることが期待される。

分析のデザイン全体に目を転じれば、令和3年度事業で実施したアンケート調査において、収入については世帯単位の値を聴取したことで、家計簿アプリデータから得られる情報との差分を確認しながら、どの程度世帯全体の動向を捕捉しているかを把握する（口座連携の十分性を議論する）手がかりとなった一方、消費支出については、アンケート形式で家計簿アプリ利用者から過去に遡って値を確認することが困難であるとの判断の下、世帯全体の消費支出額を尋ねなかつた。しかしながら、家計簿アプリの有用性を確認していく上では、こうした検証作業を行うことが望ましく、さらに、アンケート結果と替え規模アプリデータ

³⁴ 家計簿アプリデータ活用についての今後の展望等については、小林・鈴木（2022）も参照されたい。

を比較する中で蓄積される知見が、アンケート調査によって収集されることが多い公的統計の精度向上にも役立つことが展望される。

本事業を通じて家計簿アプリデータの有用性と課題が明らかになった。今後、同データが経済動向分析や政策の効果分析等に一層活用されることを期待したい。

補論1 日別データの周期性への対処について

2章5節で確認したように、消費支出の日別データには月内、週内の周期性があり、趨勢的な動向が隠されてしまっている。近年、高頻度のデータが広く活用されるようになりつつあることを受け、日次レベルのデータに対する季節調整についても研究の蓄積が進みつつある³⁵。しかしながら、家計簿アプリデータの周期性は主に口座からの自動引き落としによる月ぎめ払いの支払が特定日に集中しているためと考えられるが、引き落としは平日を選んで行われることが多い。そのため、曜日や祝日の並びに依存して、7日周期や1カ月周期ではない可変的な周期で大きめの変動が発生していることが対処を難しくしている。ここでは、家計簿アプリデータの振れへの対処を検討する出発点として、回帰分析を用いて集計値から直接的に周期性を取り除くことを試みたい。

具体的には、日々の消費支出に対する説明変数として、2章5節で確認した周期性に対応した日付ダミー（1～31日）、曜日ダミー（日曜日～土曜日）、さらに、祝日ダミーを加えるほか、月末ダミー、月末最後の平日ダミー、各日付後の最初の平日ダミー（例えば、27日のダミーは27日の後、最初の平日に1を取るダミー。1～31日のそれぞれについて構築。）を投入して回帰分析を行う。月末ダミー、月末最後の平日ダミー、各日付後の最初の平日ダミーは、月ぎめ払いの引き落とし日として想定されるスケジュールを踏まえ設定した。これらの説明変数で説明できる部分が周期性であるといえるため、被説明変数となっていたもとの集計値からこれらを取り除いた誤差項の動向が消費支出の趨勢を説明すると考えることができる。なお、推計期間は2019年4月から2021年3月までの3年間とし、2022年4月以降は推計期間内で得られた回帰係数を用いて調整を行った。

図表補1-1（上図）はこうしたプロセスによって得られた消費支出の趨勢部分について、水準をもとの消費支出と一致させるよう変換した上で示している。調整前の図表2-5-1と比べれば周期性の影響はかなり軽減されているものの、図表補1-1（中図）では、後方7日移動平均をとることで振れを均している³⁶。さらに、図表補1-1（下図）では季節性を取り除いて評価するために前年同日比を示した。単に後方7日移動平均を取った図表2-5-6と比べても、緊急事態宣言に伴う消費支出の落込みなどをよりクリアに確認できるようになった。一方で、同様の手順で周期性を調整した家計調査の日別支出の動向と比べても、カバレッジが異なることには注意する必要があるが、やはり不規則な振れは解消されておらず、季節調整に関する研究の進展に併せて取り組んでいく必要がある（図表補1-2（1）、図表補1-2（2））。

³⁵ 我が国における事例としては、有田（2022）など。

³⁶ 2章5節では曜日による周期性を取り除くことを目的に後方7日間移動平均をとったが、ここではすでに曜日に起因する周期性は取り除かれており、後方7日間移動平均をとったのは特異的な振れを均すためである。ゆえに、7日間であることに積極的な意味はない。

補論2 家計簿アプリデータのその他の有用な性質について

ここでは、家計簿アプリデータのもつ有用な性質のうち、本論で取り上げなかった点について二つ、簡単に確認しておきたい。なお、以下の議論にはすべて属性付きデータセットを用いた。

まず、家計簿アプリデータから、各家計の生活水準、暮らし向きに関する分析結果が得られる点である。一例として、ここでは、食費³⁷に注目して消費支出や収入に占める割合を確認したい。図表補2-1(1)は横軸に収入、縦軸に食費比率(左図は消費支出②に占める割合、右図は収入に占める割合)をプロットしている。図表補2-1(2)は公的統計と家計簿アプリを比較する形で収入十分位ごとのプロットとして示している³⁸。これによれば、消費支出②に占める割合、収入に占める割合ともに、家計簿アプリデータによる結果は公的統計による結果よりも低い。乖離が生じる要因としては、家計簿アプリデータと公的統計の費目分類の違い、自動的な費目分類がなされない現金支払の存在などが考えられる³⁹が、支出割合と収入水準の変化の関係は公的統計とおおむね同じである⁴⁰。このように、消費支出全体や収入全体に占める特定の支出項目の水準や推移について、所得階級別などの属性間の比較分析への活用が展望できる。

また、家計簿アプリデータは同一家計の収入・支出が長期間にわたって記録されるパネルデータであり、これは比較対象となる公的統計(例えば総務省「家計調査」の調査期間は最長6カ月)にはない性質である。図表補2-2(1)(左図)は、横軸にアンケートで得られた年間収入の2020年度から2021年度にかけての変化率、縦軸に家計簿アプリから得られた収入の変化率をプロット、図表補2-2(1)(右図)は、アンケートで得られた年間収入、家計簿アプリから得られた年間収入の変化率の分布を示している。また、図表補2-2(2)(左図)は、アンケートで得られた年間収入の変化率ごとの家計簿アプリから得られた年間収入の変化率の分布、図表補2-2(2)(右図)はその逆に家計簿アプリから得られた年間収入の変化率ごとのアンケートで得られた年間収入の変化率の分布を示している。いずれの図表からも、両者には一定の相関関係があることがうかがえ、パネルデータとして一定の有用性があることを示している。ただし、今回実施したアンケートでは、過去3年間の年間収入をまとめて尋ねたこともあり、約4割が変化率0%と回答していることから明らかのように、所得の変動が大きい場合には横ばいで回答した可能性が高

³⁷ 公的統計における「食費」とおおむね対応すると考えられる事業者カテゴリーにおける「食費」、「外食」、「居酒屋・バー」の合計。なお、家計の消費支出に占める食費の割合をエンゲル係数と呼び、一般に、所得が上昇するほどエンゲル係数が低下し、暮らし向きに余裕が増すと解釈される。

³⁸ 本来、公的統計は、家計簿アプリデータとそろえて手取り収入に基づく十分位階級を作成するべきであるが、集計表が公表されていないため、ここでは便宜的に年間収入の十分位階級に基づいて作成した。

³⁹ 事業者カテゴリーは基本的に品目ではなく、取引単位で付される。食費として支出した場合でも、「買い物」などといったカテゴリーとして分類されているケースも多い。また、こうした家計簿アプリデータ側の問題のほか、繰り返しになるが、家計への調査に基づく公的統計が過小推計となっている可能性も指摘されている。

⁴⁰ 図表補2-1(2)の右図では、家計簿アプリデータの第1十分位で食費比率が極端に高くなっているが、小遣い帳のような形で消費支出についてのみ記録している利用者による影響と考えられる。

い。家計簿アプリデータのパネルデータとしての有用性をより精緻に確認するためには、より信頼性の高い方法で家計簿アプリ利用者の実際の収入・支出を把握し、両者を突き合わせることを望ましく、そのアプローチを含め今後の検討課題となろう。

ここで取り上げた家計簿アプリデータの特性はいずれも、経済分析の幅を広げうる有用なものであり、引き続き精緻な検証や技術的な蓄積が期待される。

参考文献

- 有田帝馬（2022）「日次データへの季節調整の適用 一家計調査による実例」統計研究彙報、79、pp. 22-40
- 宇南山卓（2019）「家計の把握とR I C Hプロジェクト」経済研究、70(4)、pp. 331-356
- 小林周平・鈴木源一郎（2022）「経済動向分析における家計簿アプリデータの活用」経済財政分析ディスカッション・ペーパー、DP/22-3
- Kaneda, M., Kubota, S., & Tanaka, S. (2021) “Who spent their COVID-19 stimulus payment? Evidence from personal finance software in Japan” *The Japanese Economic Review*, 72(3), pp. 409-437

図表

(図表 1-2-1 家計簿アプリデータ分析に際しての課題とその検証結果)

	収入	消費支出
費目分類の 正確性	アンケート結果と比較 ウェイト・バック後に公的統計と比較 ⇒少なくとも勤労者世帯であれば水 準の妥当性、季節性などを確認	ウェイト・バック後に公的統計と比較 ⇒消費支出以外を捕捉している可能 性は残るが、改善次第で経済動向把 握等に活用可能 (水準を上振れさせる可能性を否 定できない)
サンプルの 代表性	アンケート結果を公的統計と比較 ⇒若年層、高所得者、関東地方在住者に偏り、高齢者は極めて少ない (年齢は下振れ要因、年間収入は上振れ要因、居住地域は若干の上振れ要因) ⇒高齢者は極めて少なく、属性内での代表性にも疑問 (勤労者世帯ベースでの比較が基本となる)	
口座連携の 十分性	アンケート結果と比較 ⇒複数有業者世帯で十分性が低下 (水準を下振れさせる) ⇒口座連携数増加影響あり (変化率に上方バイアス)	公的統計と比較 ⇒複数有業者世帯や世帯人員増で十 分性が低下 (水準を下振れさせる) ⇒口座連携数の増加やキャッシュレ ス化の影響はATM引出しを考慮 すれば緩和

(図表 2-1-1 前処理の実施)

前処理条件	属性付きデータセット	事業者管理データセット
ステップ 1 : 取引の除外		
1-1) 支出データから±1,000万円以上の取引を除外	アプリデータ	アプリデータ
1-2) 現金の受払による取引を除外	アプリデータ	アプリデータ
ステップ 2 : 対象者条件の精査		
2-1) Moneytree アプリ利用者限定	調査参加条件	アプリデータ
2-2) 金融機関登録が3以上ある利用者限定	調査参加条件	アプリデータ
2-3) 世帯の主たる収入受取口座がある利用者限定	調査参加条件	(実施せず)
2-4) 2020年以前のデータが存在する利用者限定	アプリデータ	アプリデータ
2-5) 個人事業資金の管理口座がない利用者限定	アンケート回答	アプリデータ ※1
ステップ 3 : 利用者(外れ値)の除外		
3-1) アプリデータにおいて、収入、消費支出、貯蓄、負債が上位0.5%の利用者を除外	アプリデータ	アプリデータ
3-2) アンケートの回答において、年間収入、貯蓄、負債が上位0.5%の利用者を除外	アンケート回答	(実施せず)
3-3) 収入、消費支出、貯蓄、負債の総額がマイナスの利用者を除外	アプリデータ	アプリデータ

※ このほか、事業者管理データセットについては、属性付きデータセットではみられなかった種類の口座や、コーポレート口座として登録されている口座を除外している。

※1 分析対象期間(2019年1月~2022年5月)に、事業者カテゴリー「営業費」として合計10万円以上支払われている場合には、事業資金の管理口座が含まれていると考慮して除外。

(図表 2-2-1 事業者カテゴリーの取扱い)

事業者カテゴリー	取扱い	
	符号が「+」(受取)	符号が「-」(支払)
給料	収入	収入でない ⇒明細として除外
収入		
ATM入金		
食費、光熱水道費、… (消費支出に分類)	支出 ⇒支払と相殺	支出
税金、ローン返済、… (消費支出以外に分類)		
ATM引出し	支出でない ⇒明細として除外	
振替	収入	支出
投資・金融サービス		
未定		

(図表 2-2-2 実収入の定義)

事業者カテゴリー	実収入
給料	○
収入	○
ATM入金	×
振替	×
投資・金融サービス	×

(図表 2-2-3 消費支出の定義)

事業者カテゴリー	消費支出①	消費支出②
食費、光熱水道費、… (消費支出に分類)	○※	○※
税金、ローン返済、… (消費支出以外に分類)	×	×
ATM引出し	×	○
振替	×	×
投資・金融サービス	×	×
未定	×	×

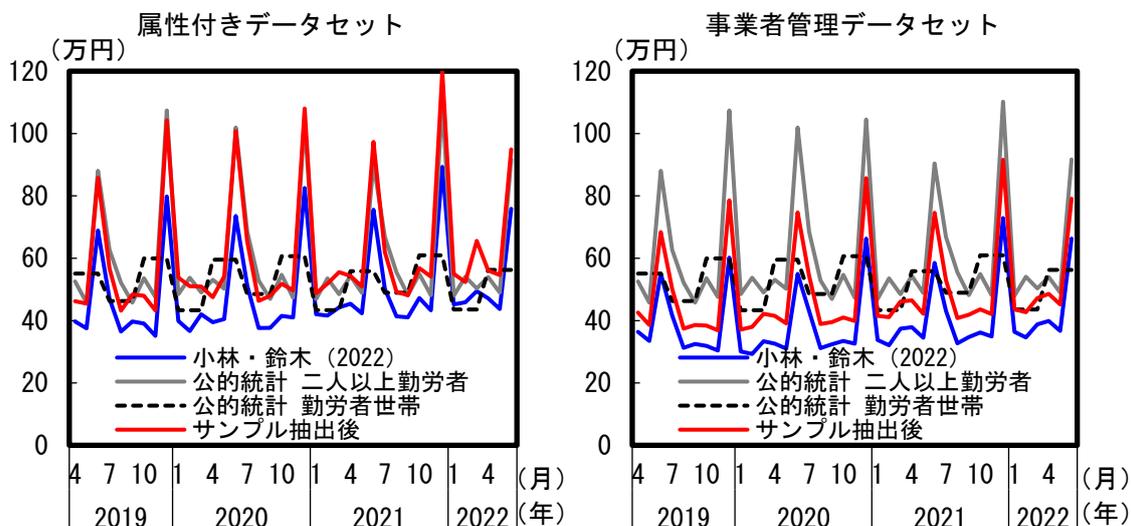
※ 摘要欄の文言から住宅資産の取得やローンの返済と機械的に判定できる取引情報は除外している。

(図表 2-4-1 家計簿アプリデータの集計値)

サンプル抽出条件				2021 年度平均値 (万円)					
				収入		消費支出①		消費支出②	
消費取引	水道 光熱費	収入取引	ロゲイン 履歴	属性付き	事業者 管理	属性付き	事業者 管理	属性付き	事業者 管理
各月 1 件以上	—	—	—	60.3	45.4	30.9	20.6	36.5	29.8
—	年度 1 円以上	—	—	62.5	47.1	31.9	22.3	37.5	32.2
各月 1 件以上	年度 1 円以上	—	—	63.4	49.8	33.3	24.9	39.0	35.8
各月 1 件以上	年度 1 円以上	各年度 1 件以上	—	63.8	50.7	33.5	26.9	39.4	39.7
各月 1 件以上	年度 1 円以上	各年度 1 件以上	直近 1 年 1 回以上	—	50.8	—	26.6	—	38.8
小林・鈴木 (2022) での報告値 (条件なし)				51.3	41.3	28.0	—	32.9	19.8
公的統計 (総世帯)				—		24.9			
公的統計 (勤労者世帯)				可処分 : 42.7 実収入 : 52.3		27.2			

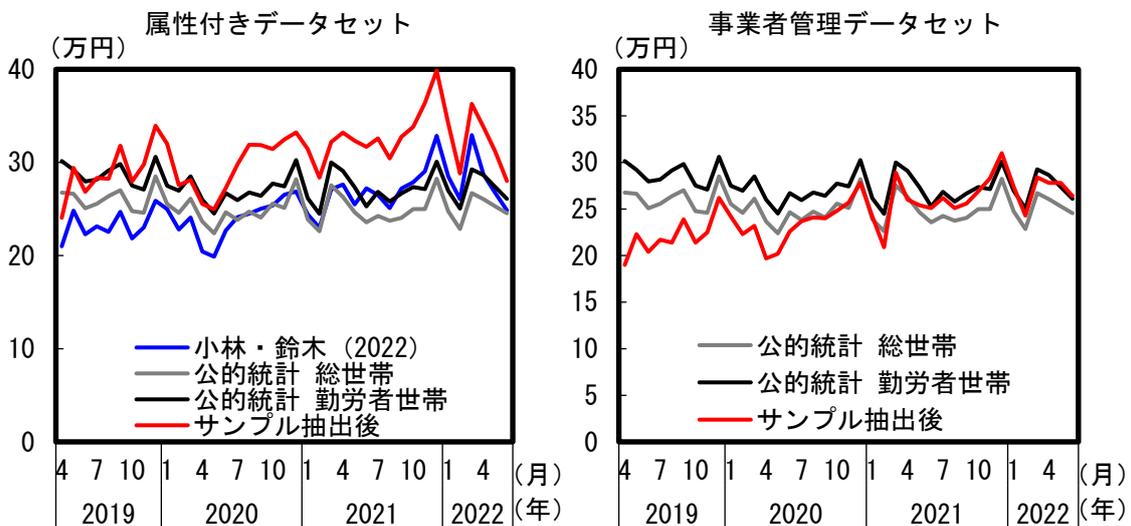
(備考) 公的統計は、収入については総務省「家計調査」、消費支出については総務省「世帯消費動向指数」により作成。なお、総世帯の収入は公表されていない。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において収入や消費支出が記録されているサンプルのみを集計の対象にしている。また、小林・鈴木 (2022) においては、事業者管理データセットについて消費支出①を集計していない。

(図表 2-4-2 収入の推移)



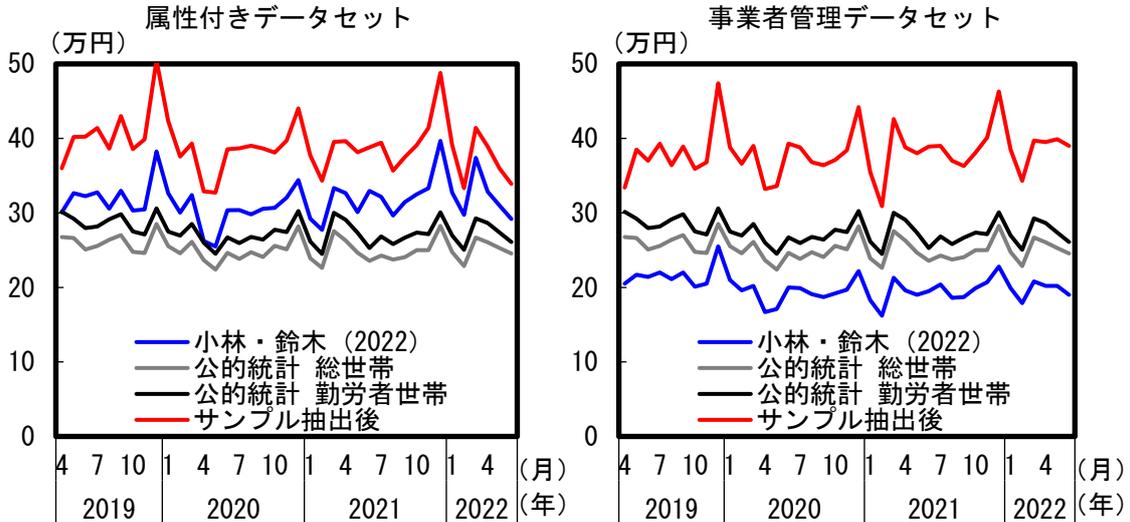
(備考) 公的統計は、総務省「家計調査」により作成。なお、勤労者世帯の収入は四半期単位で公表されているため、その月が含まれる四半期の公表値を示している。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において収入が記録されているサンプルのみを集計の対象にしている。

(図表 2-4-3 消費支出①の推移)



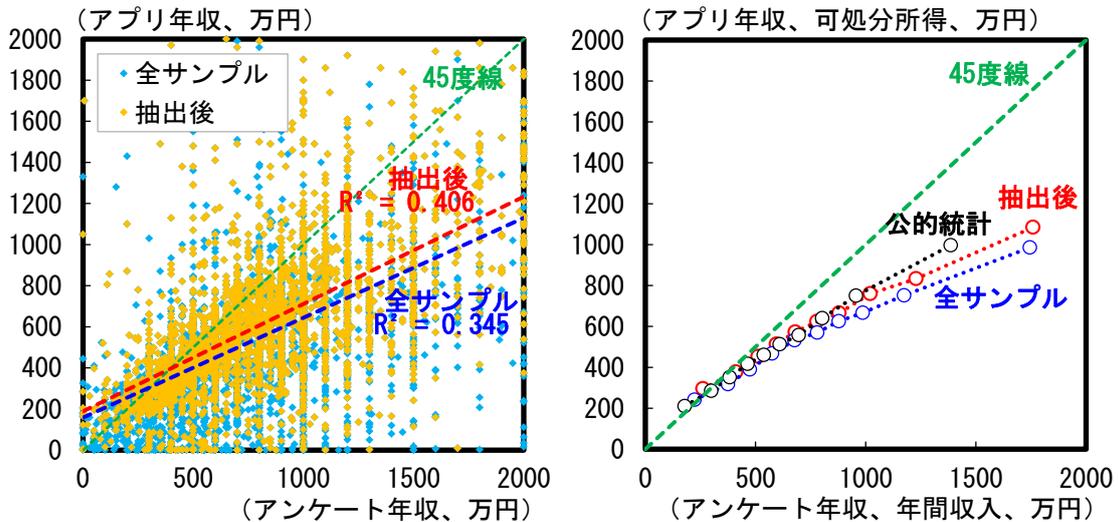
(備考) 公的統計は、総務省「世帯消費動向指数」により作成。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において消費支出①が記録されているサンプルのみを集計の対象にしている。また、小林・鈴木(2022)においては、事業者管理データセットについて消費支出①を集計していない。

(図表 2-4-4 消費支出②の推移)



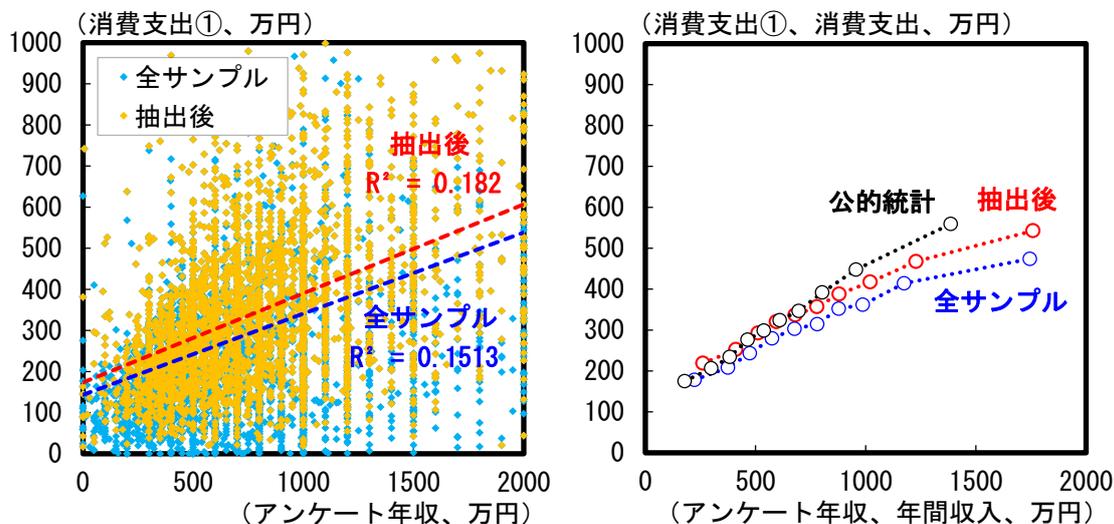
(備考) 公的統計は、総務省「世帯消費動向指数」により作成。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において消費支出②が記録されているサンプルのみを集計の対象にしている。

(図表 2-4-5 アプリから得られた収入とアンケートで得られた収入)



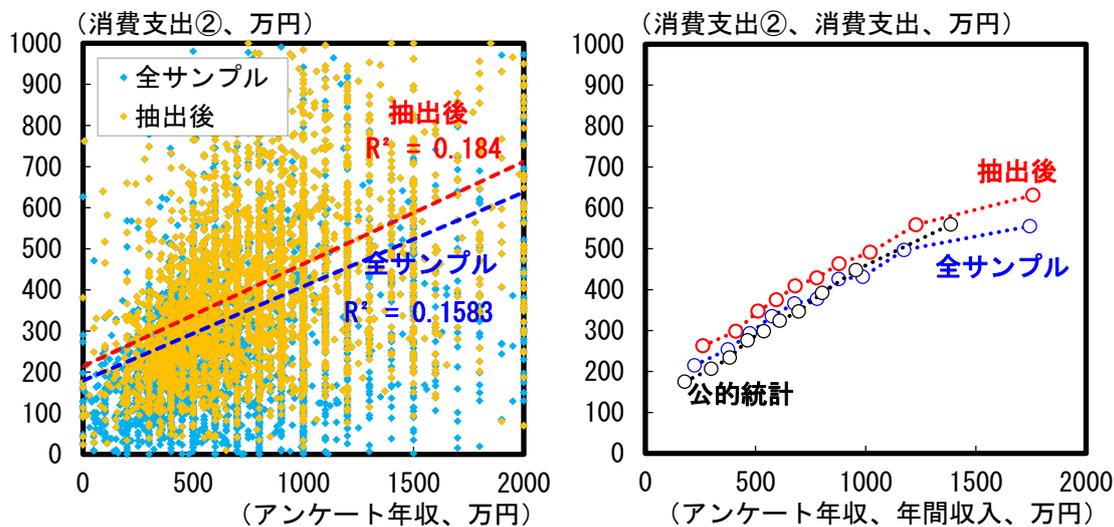
(備考) 公的統計は、総務省「家計調査」により作成。2021年の結果から作成しており、横軸は年間収入、縦軸は可処分所得（実収入から勤労所得税、個人住民税、社会保険料を控除した、手取り収入と対応する部分）をプロットしている。属性付きデータセットに関しては、内閣府及び株式会社三菱総合研究所作成。横軸はアンケートから得られた2021年度の年間収入、縦軸は家計簿アプリから得られた2021年の収入をプロットしている。右図は公的統計、属性付きデータセットいずれも、年間収入十分位の集計値をプロットしている。

(図表 2-4-6 年間収入と消費支出①)



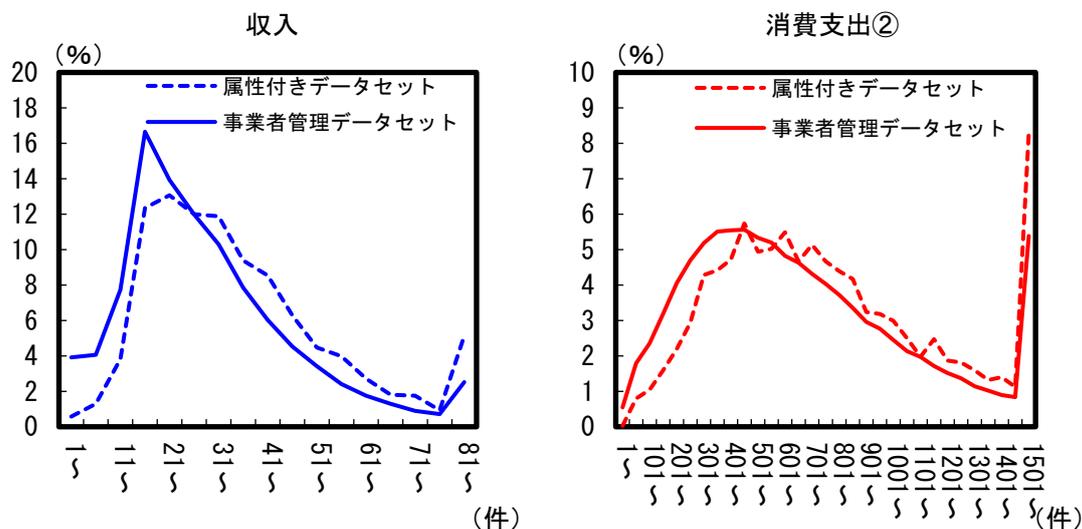
(備考) 公的統計は、総務省「家計調査」、「世帯消費動向指数」により作成。2021年の結果から作成しており、横軸は年間収入、縦軸は家計調査の消費支出を世帯消費動向指数の水準で補正してプロットしている。属性付きデータセットに関しては、内閣府及び株式会社三菱総合研究所作成。横軸はアンケートから得られた2021年度の年間収入、縦軸は家計簿アプリから得られた2021年の消費支出①をプロットしている。右図は公的統計、属性付きデータセットいずれも、年間収入十分位の集計値をプロットしている。

(図表 2-4-7 年間収入と消費支出②)



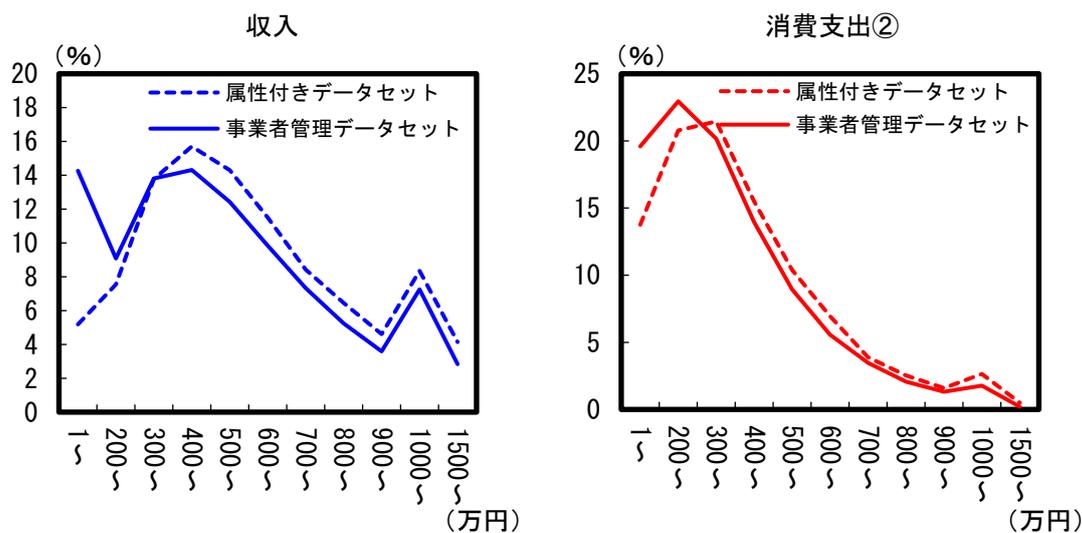
(備考) 公的統計は、総務省「家計調査」、「世帯消費動向指数」により作成。2021年の結果から作成しており、横軸は年間収入、縦軸は家計調査の消費支出を世帯消費動向指数の水準で補正してプロットしている。属性付きデータセットに関しては、内閣府及び株式会社三菱総合研究所作成。横軸はアンケートから得られた2021年度の年間収入、縦軸は家計簿アプリから得られた2021年の消費支出②をプロットしている。右図は公的統計、属性付きデータセットいずれも、年間収入十分位の集計値をプロットしている。

(図表 2-4-8 年間の取引数)



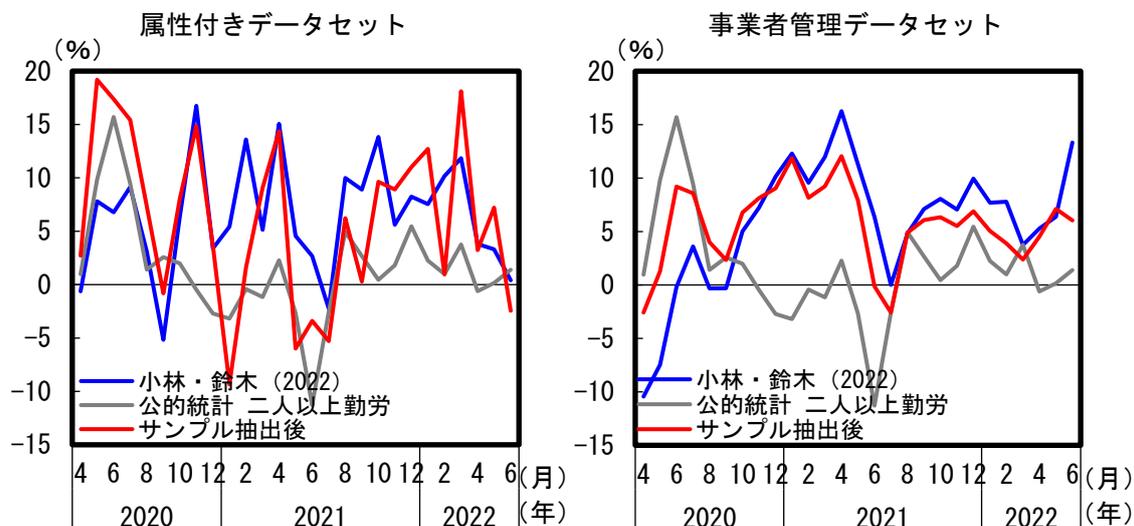
(備考) 属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。いずれも 2021 年度内の集計値。

(図表 2-4-9 年間の取引金額)



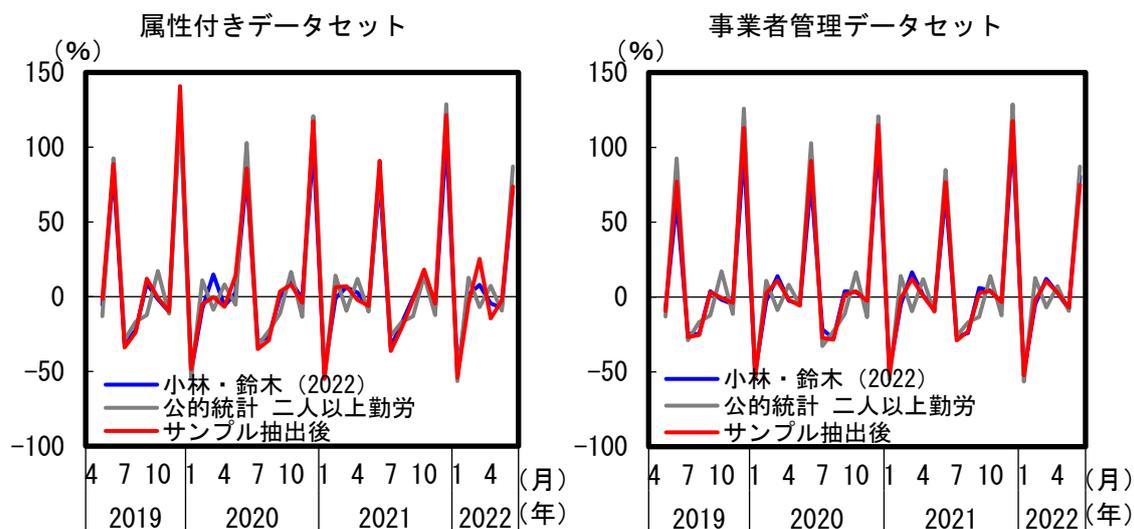
(備考) 属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。いずれも 2021 年度内の集計値。

(図表 2-4-10 (1) 収入の変動 (前年同月比))



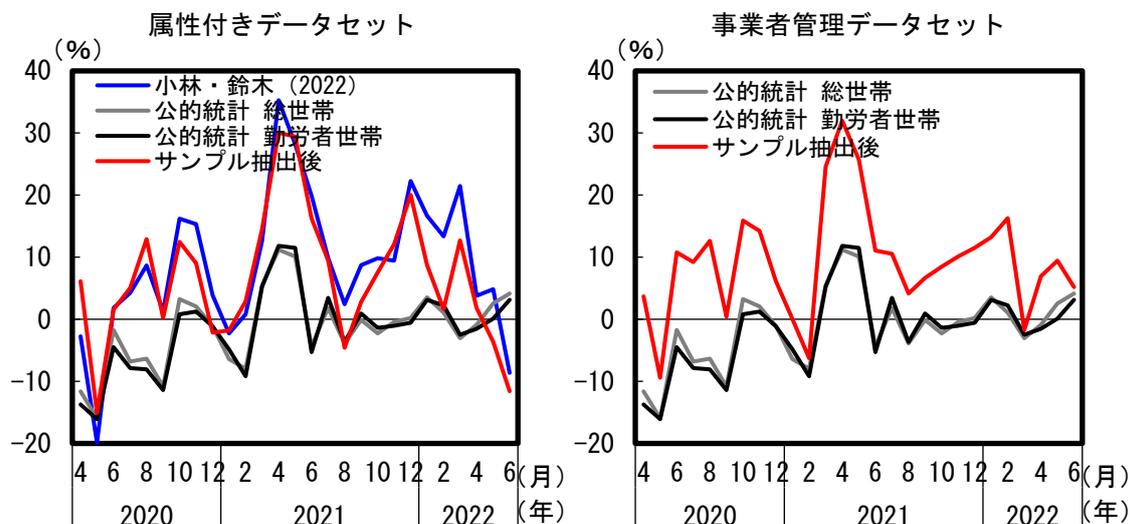
(備考) 公的統計は、総務省「家計調査」により作成。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において収入が記録されているサンプルのみを集計の対象にしている。

(図表 2-4-10 (2) 収入の変動 (前月比))



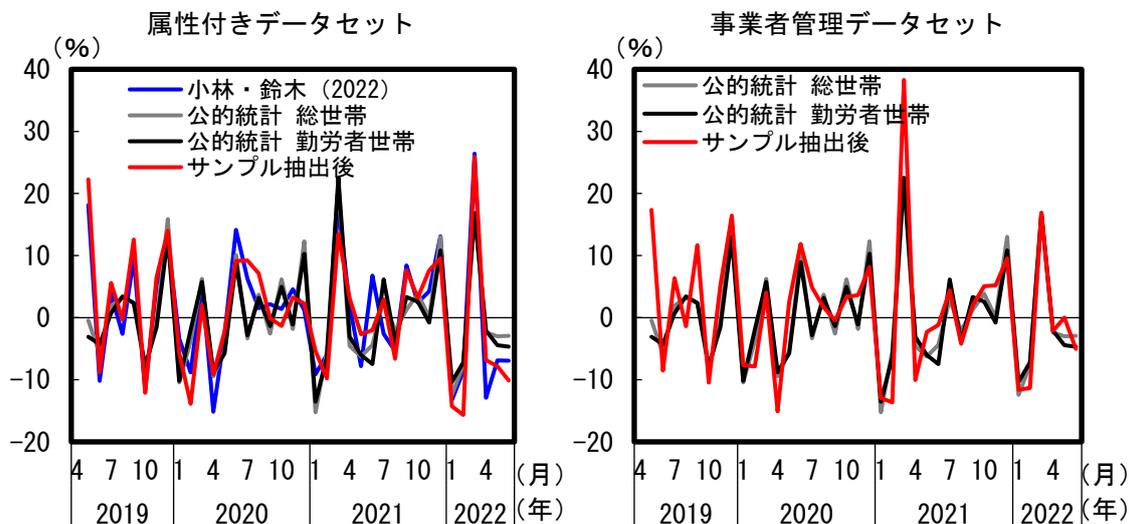
(備考) 公的統計は、総務省「家計調査」により作成。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において収入が記録されているサンプルのみを集計の対象にしている。

(図表 2-4-11 (1) 消費支出①の変動 (前年同月比))



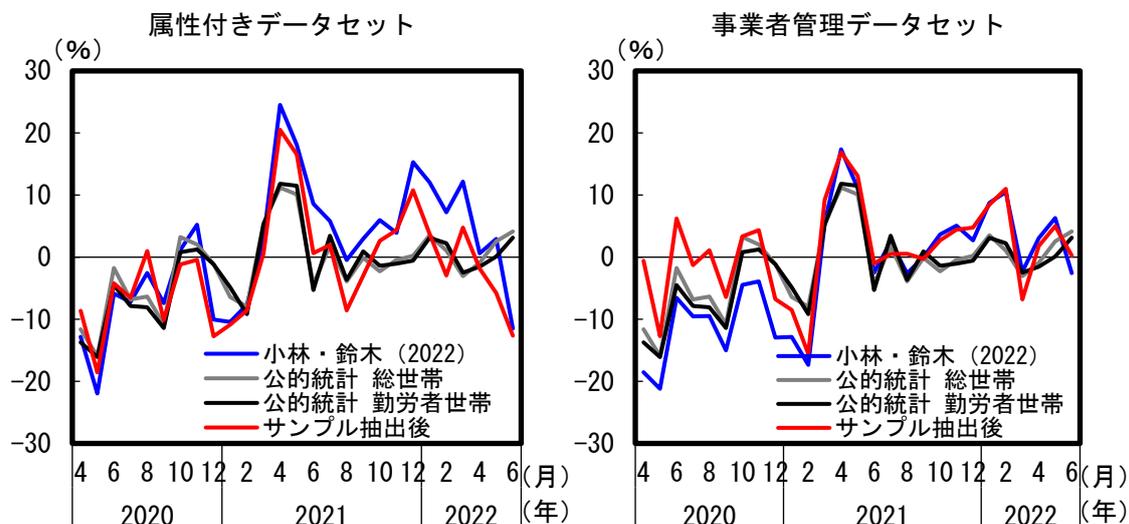
(備考) 公的統計は、総務省「世帯消費動向指数」により作成。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において消費支出①が記録されているサンプルのみを集計の対象にしている。また、小林・鈴木(2022)においては、事業者管理データセットについて消費支出①を集計していない。

(図表 2-4-11 (2) 消費支出①の変動 (前月比))



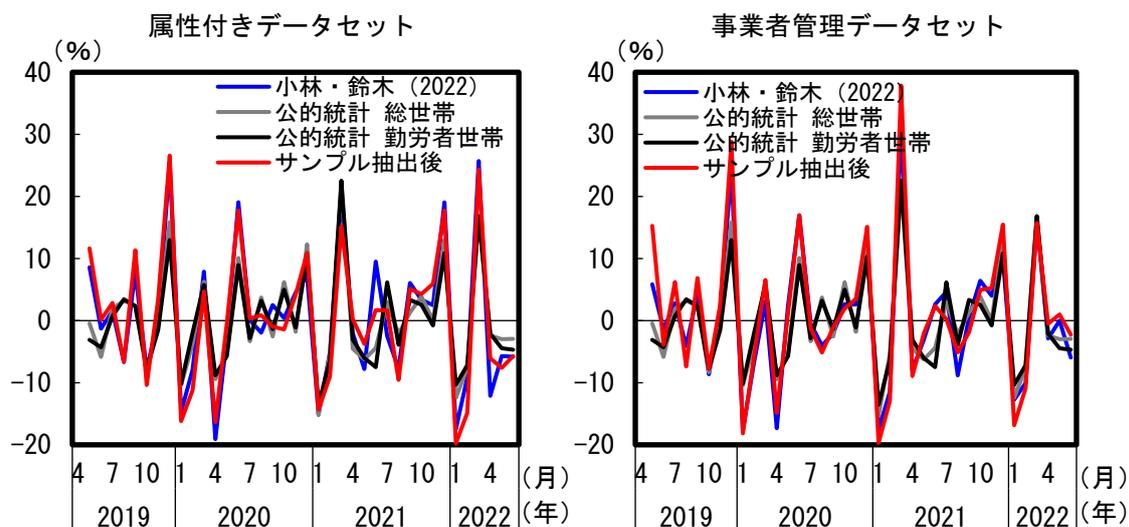
(備考) 公的統計は、総務省「世帯消費動向指数」により作成。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において消費支出①が記録されているサンプルのみを集計の対象にしている。また、小林・鈴木(2022)においては、事業者管理データセットについて消費支出①を集計していない。

(図表 2-4-12 (1) 消費支出②の変動 (前年同月比))



(備考) 公的統計は、総務省「世帯消費動向指数」により作成。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において消費支出②が記録されているサンプルのみを集計の対象にしている。

(図表 2-4-12 (2) 消費支出②の変動 (前月比))



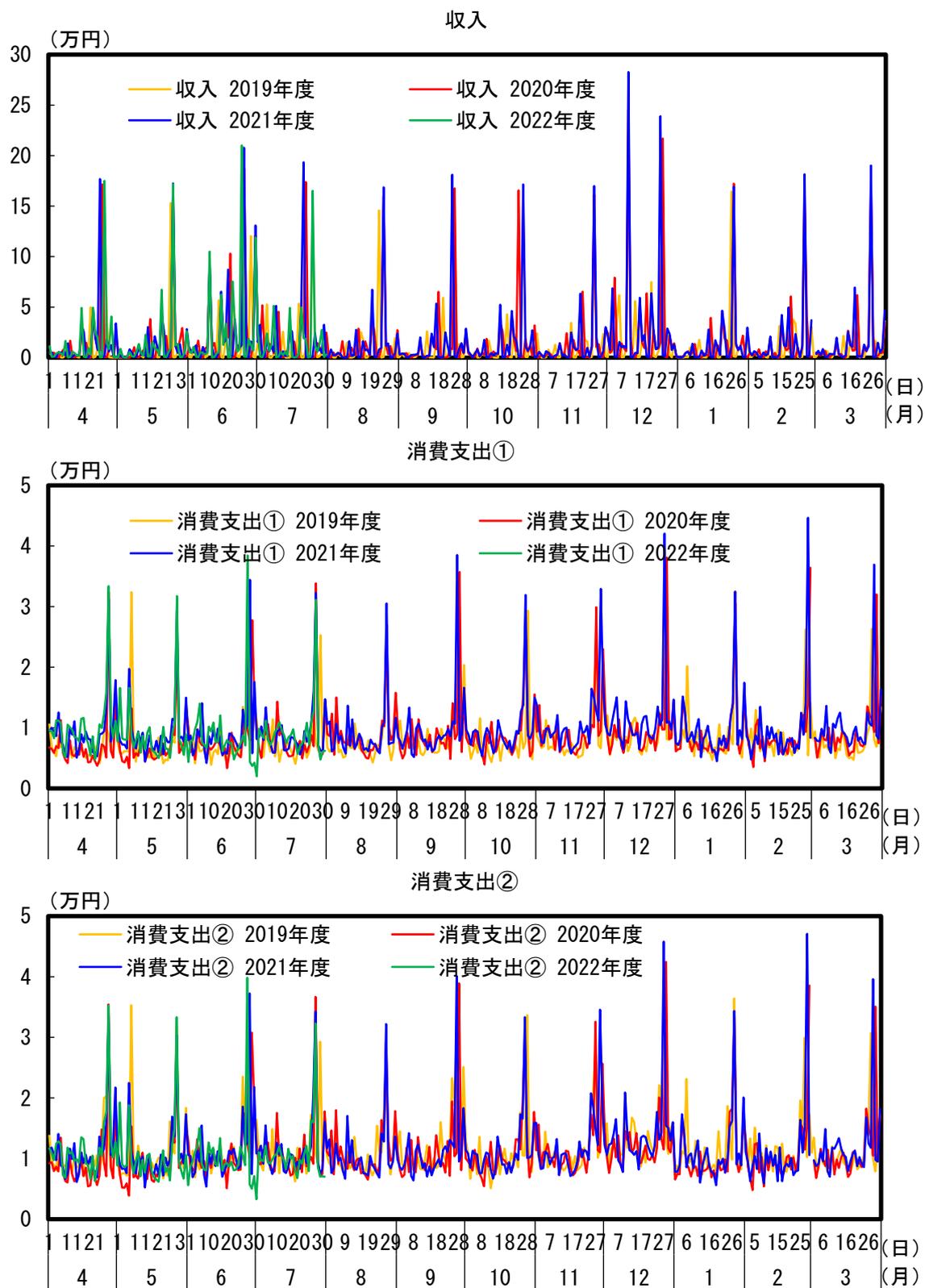
(備考) 公的統計は、総務省「世帯消費動向指数」により作成。属性付きデータセットは、内閣府及び株式会社三菱総合研究所作成。事業者管理データセットは、マネーツリー株式会社作成。なお、各月において消費支出②が記録されているサンプルのみを集計の対象にしている。

(図表 2-4-13 収入の変動・消費支出の変動の公的統計に対するRMS E)

			属性付きデータセット		事業者管理データセット	
費目	比較対象	変動	小林・鈴木 (2022)	抽出後	小林・鈴木 (2022)	抽出後
収入	二人以上 勤労者世帯	前年同月比	0.0793	0.0727	0.0981	0.0690
		前月比	0.1199	0.1262	0.1450	0.1095
消費支出①	総世帯	前年同月比	0.1311	0.1178	—	0.1192
		前月比	0.0632	0.0702	—	0.0568
	勤労者世帯	前年同月比	0.1335	0.1209	—	0.1239
		前月比	0.0665	0.0716	—	0.0604
消費支出②	総世帯	前年同月比	0.0758	0.0625	0.0546	0.0517
		前月比	0.0576	0.0584	0.0473	0.0614
	勤労者世帯	前年同月比	0.0749	0.0619	0.0510	0.0572
		前月比	0.0637	0.0640	0.0522	0.0690

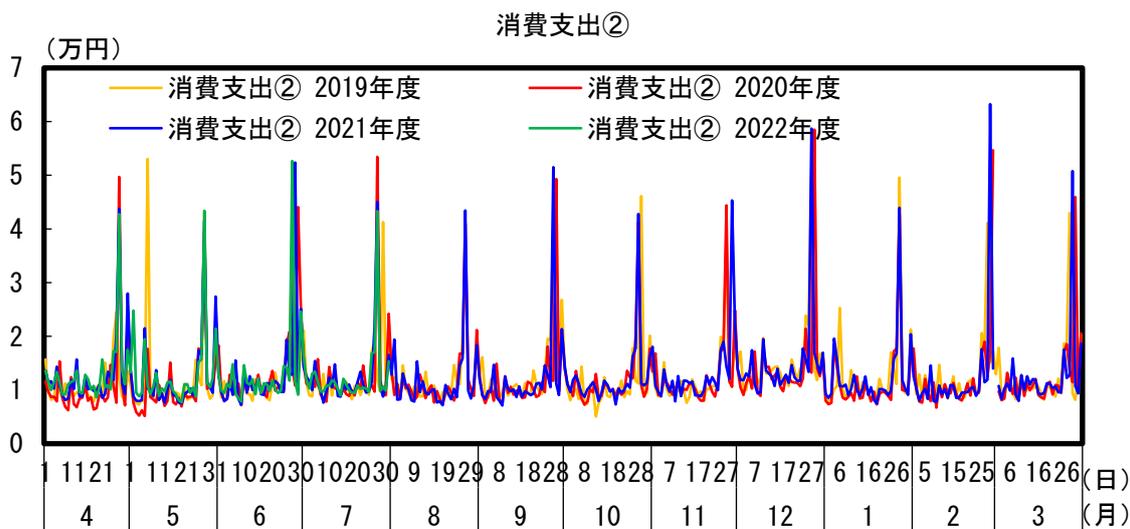
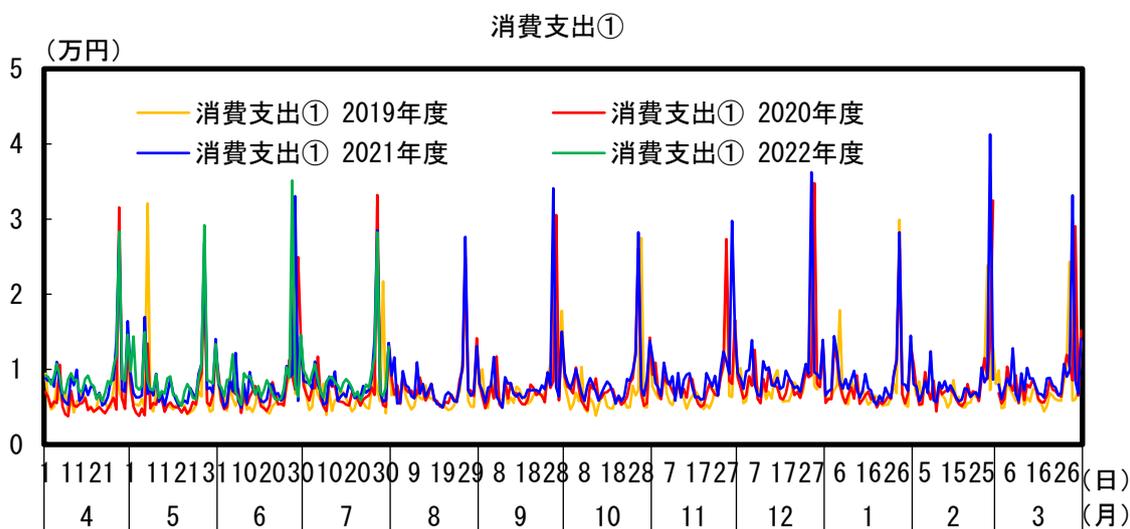
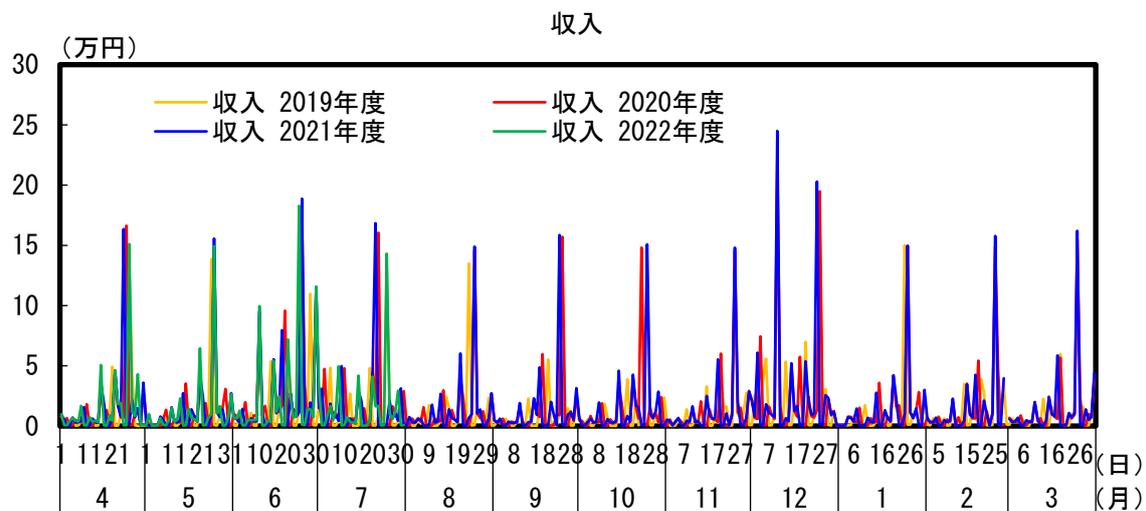
(備考) 図表 2-4-10~図表 2-4-12 における集計結果について、サンプルのすべての期間を用いて集計した。なお、小林・鈴木(2022)においては、事業者管理データセットについて消費支出①を集計していない。

(図表 2-5-1 (1) 日次の動向 (属性付きデータセット))



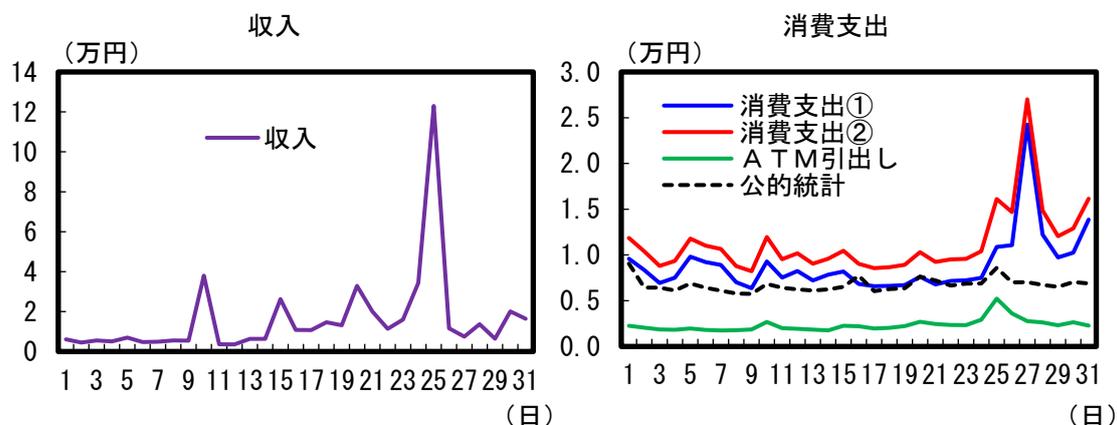
(備考) 内閣府及び株式会社三菱総合研究所作成。その日に取引が記録されていないサンプルを含め、すべてのサンプルを対象に集計の対象としている。

(図表 2-5-1 (2) 日次の動向 (事業者管理データセット))



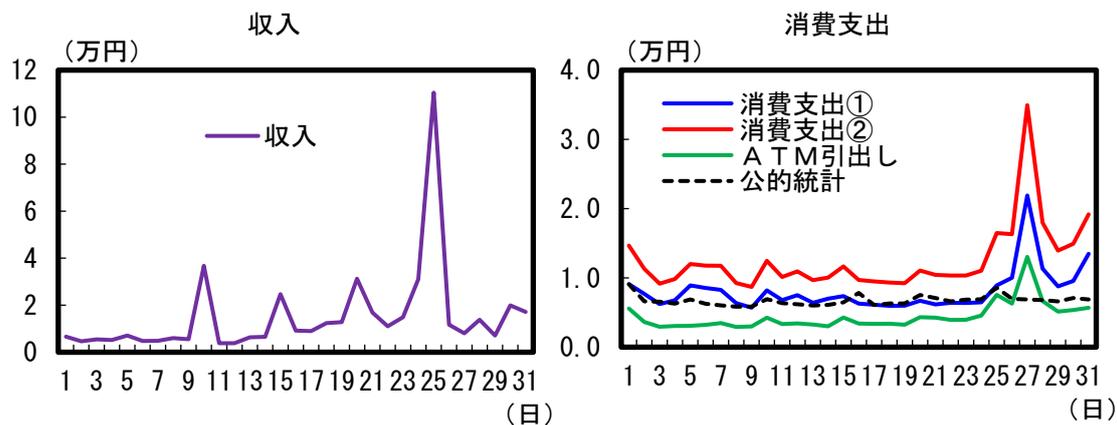
(備考) マネーツリー株式会社作成。その日に取引が記録されていないサンプルを含め、すべてのサンプルを対象に集計の対象としている。

(図表 2-5-2 (1) 月内の周期性 (属性付きデータセット))



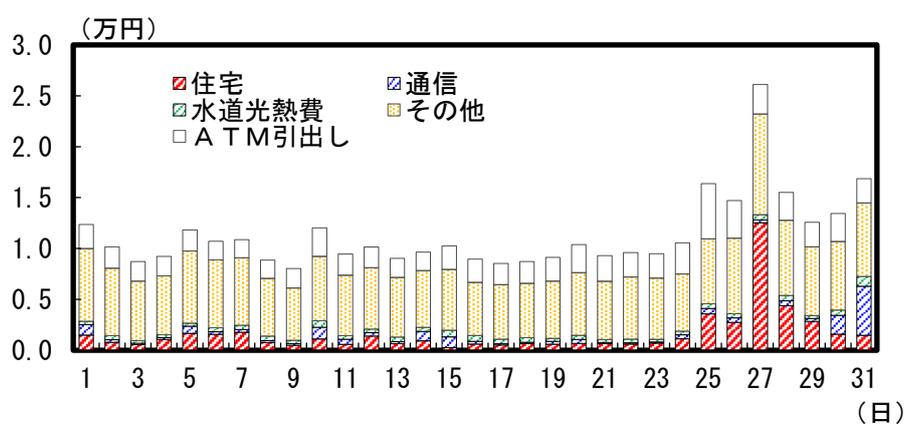
(備考) 内閣府及び株式会社三菱総合研究所作成。2019~2021 年度の平均。

(図表 2-5-2 (2) 月内の周期性 (事業者管理データセット))



(備考) マネーツリー株式会社作成。2019~2021 年度の平均。

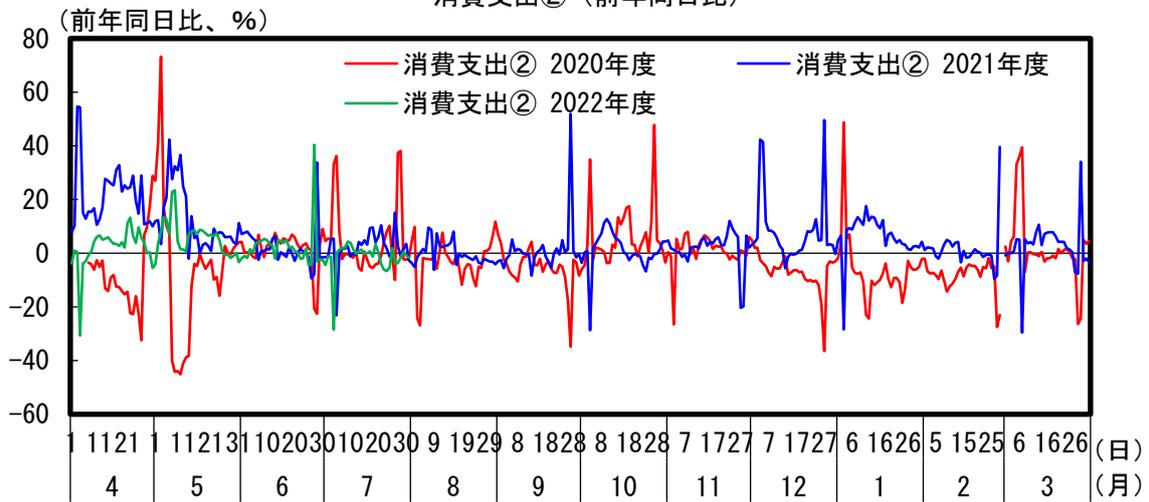
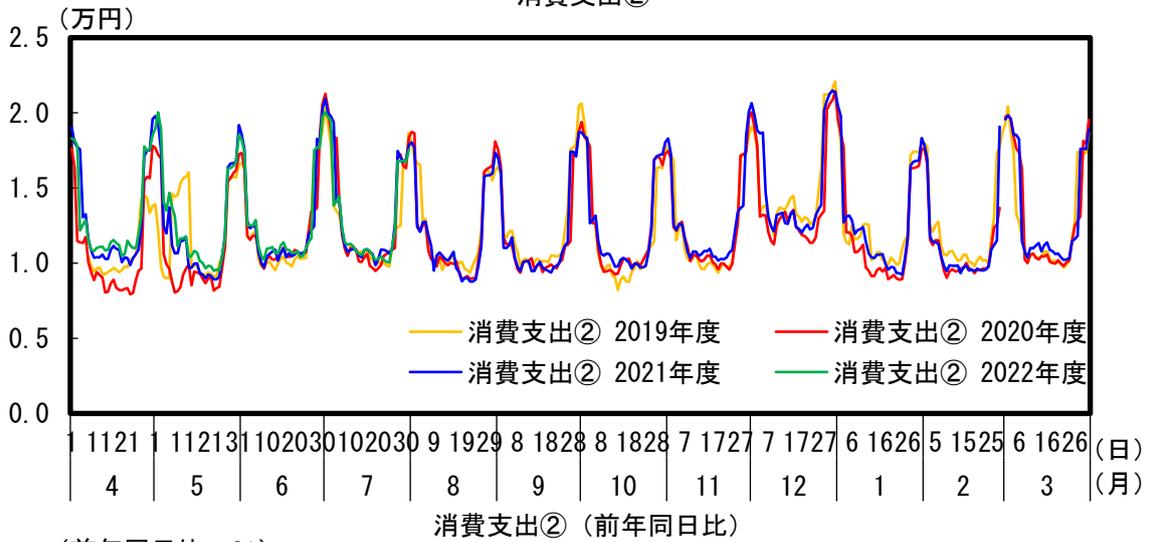
(図表 2-5-3 事業者カテゴリー別の月内の周期性 (属性付きデータセット))



(備考) 内閣府及び株式会社三菱総合研究所作成。2019~2021 年度の平均。

(図表 2-5-6 日次の動向 (事業者管理データセット・後方7日間移動平均))

消費支出②

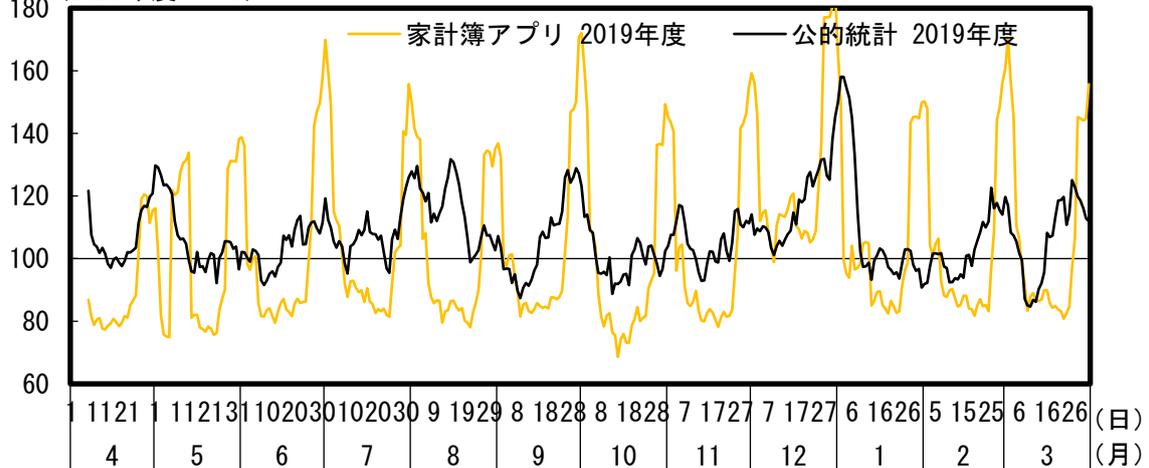


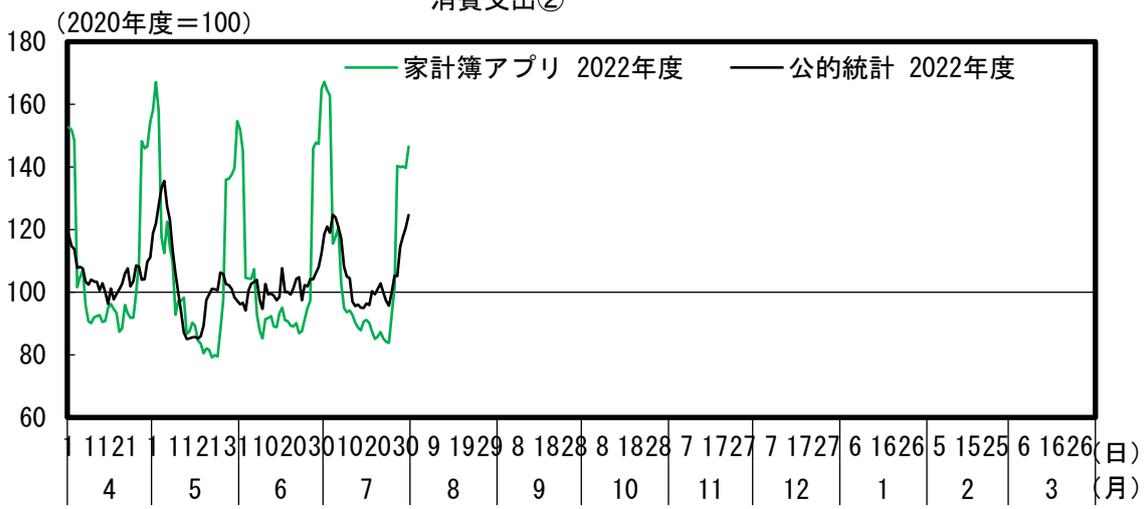
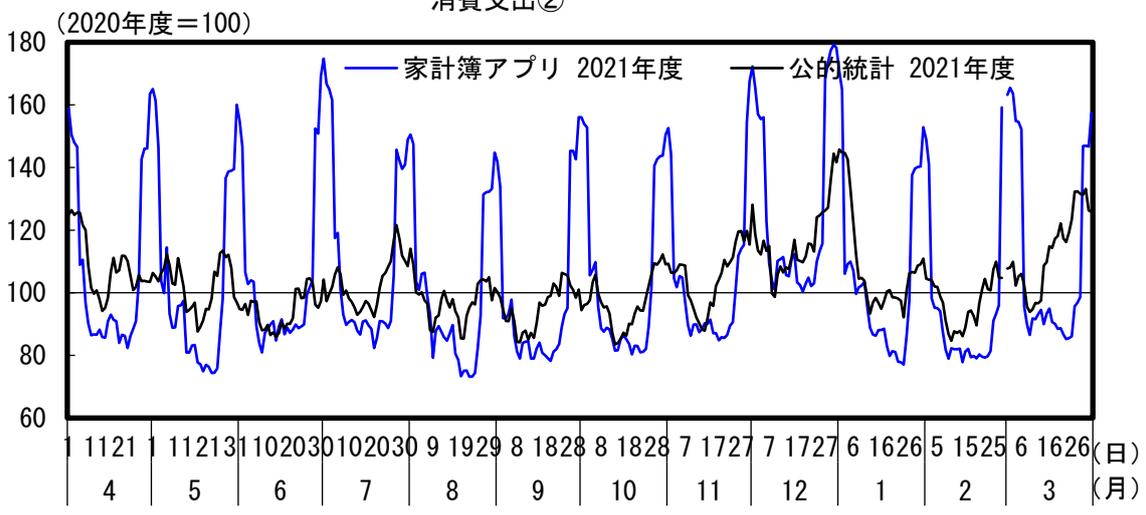
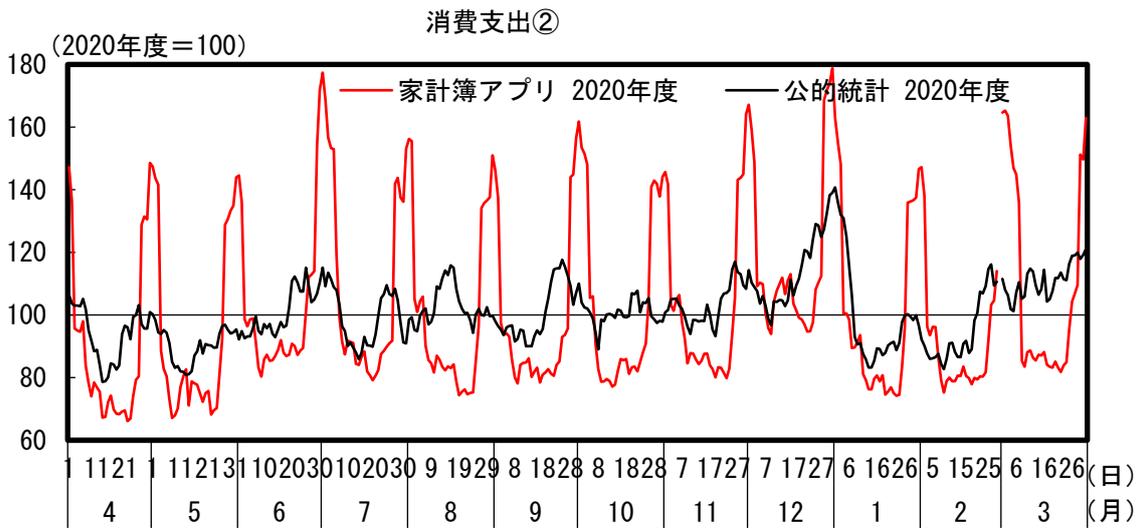
(備考) 図表 2-5-1 を基に作成。

(図表 2-5-7 (1) 公的統計との比較 (後方7日間移動平均・指数))

消費支出②

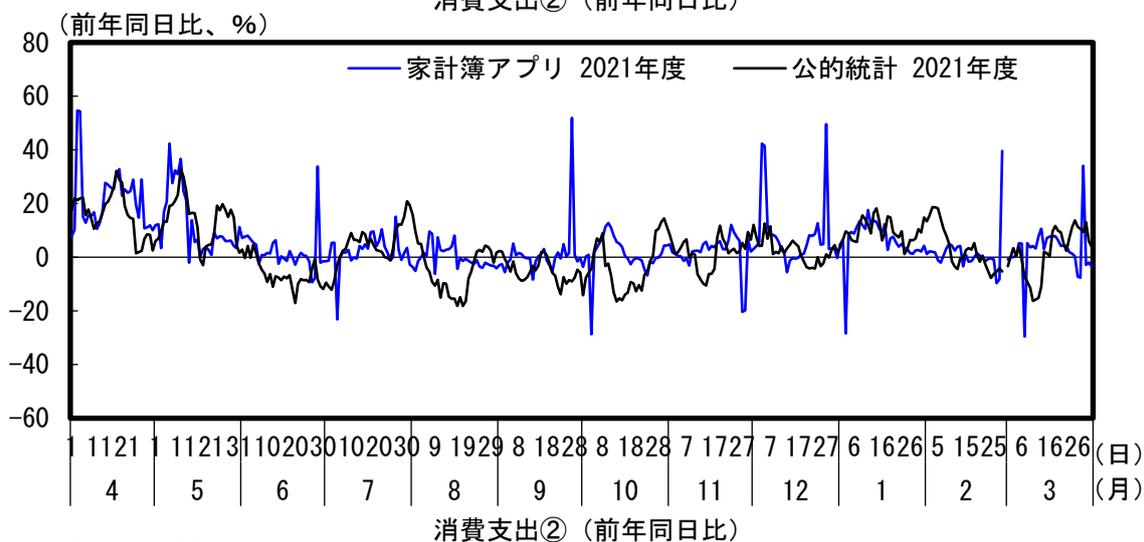
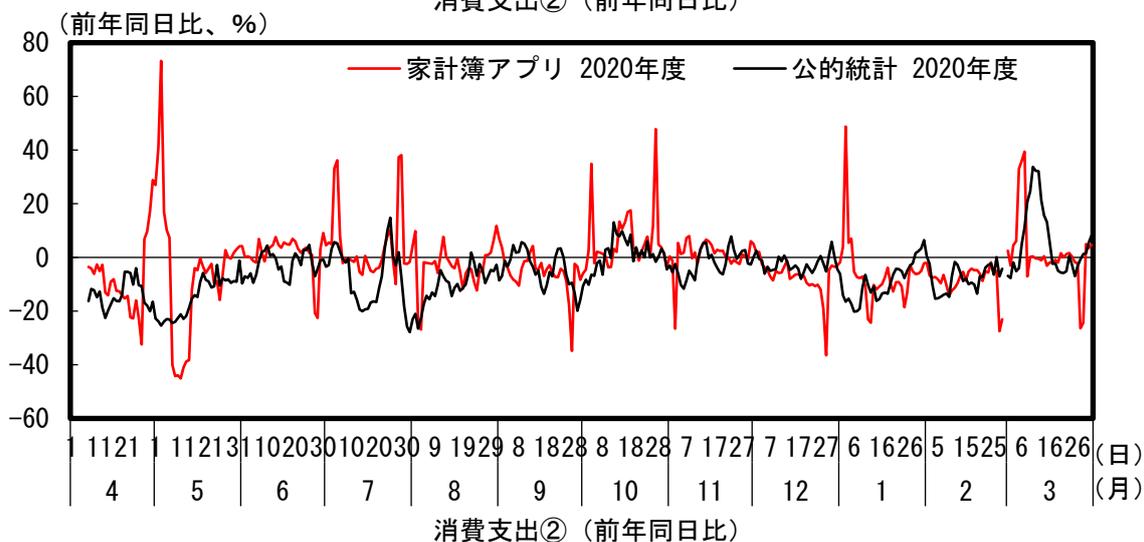
(2020年度=100)





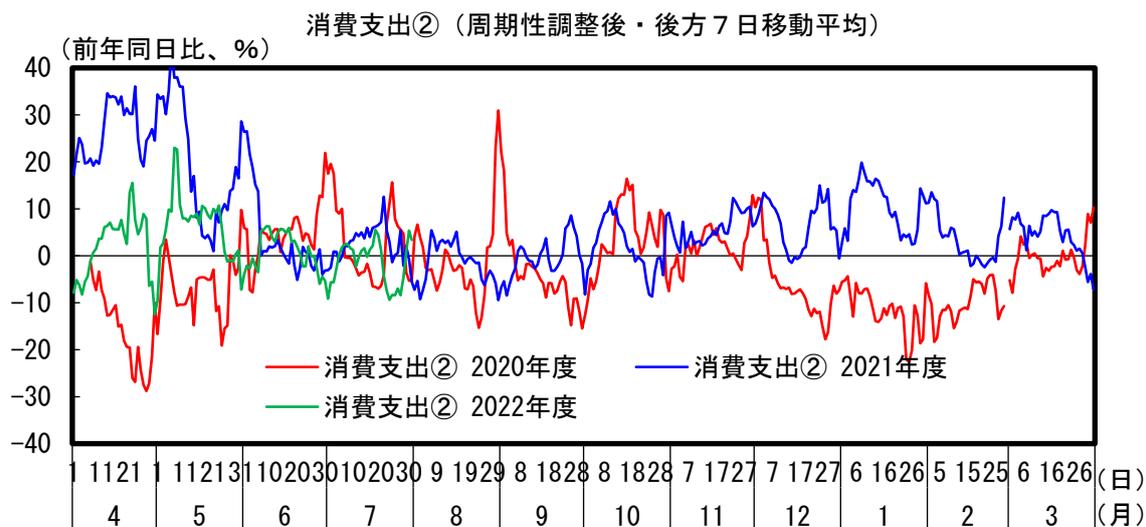
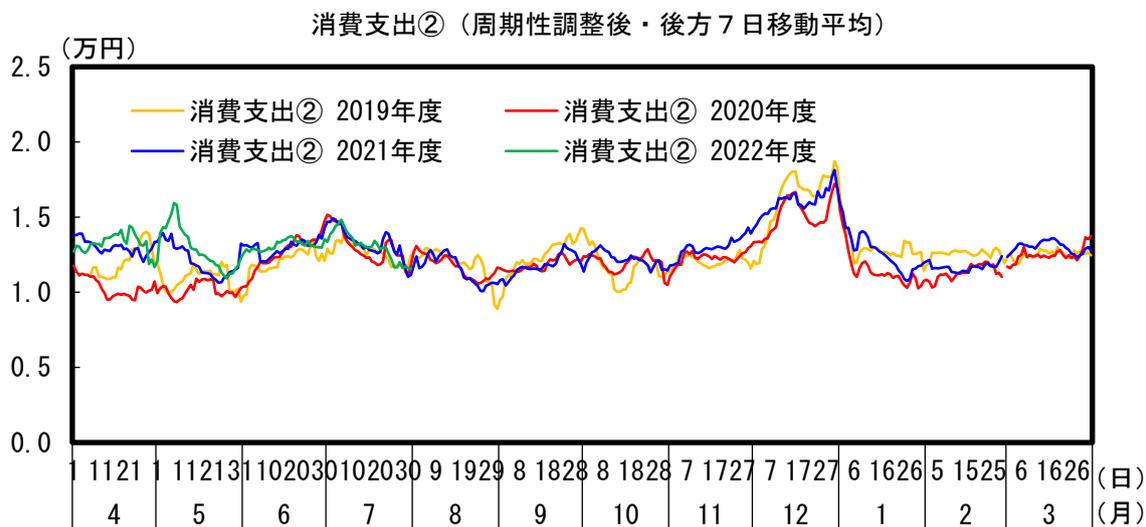
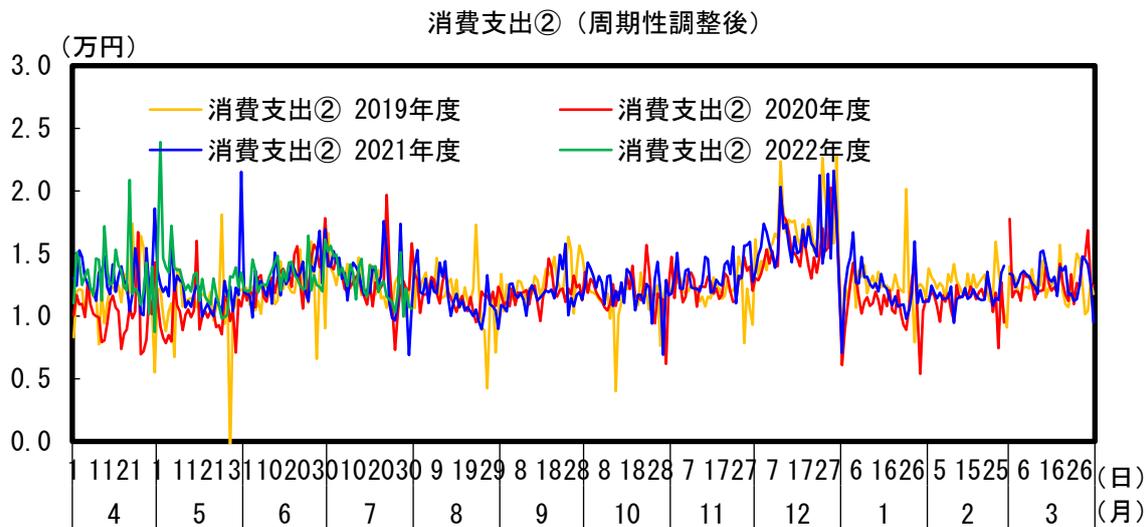
(図表 2-5-7 (2) 公的統計との比較 (後方7日間移動平均・前年同日比))

消費支出② (前年同日比)



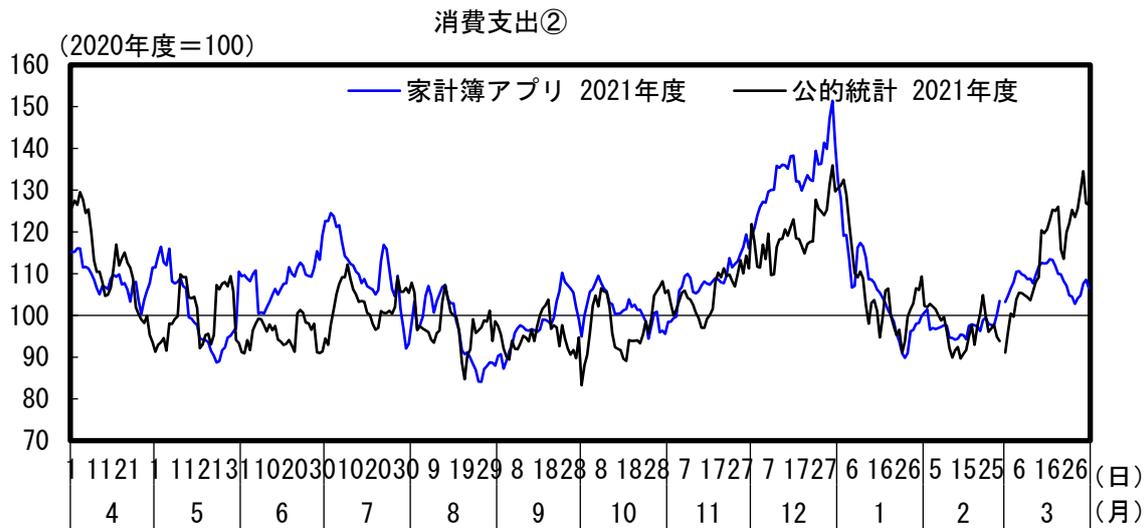
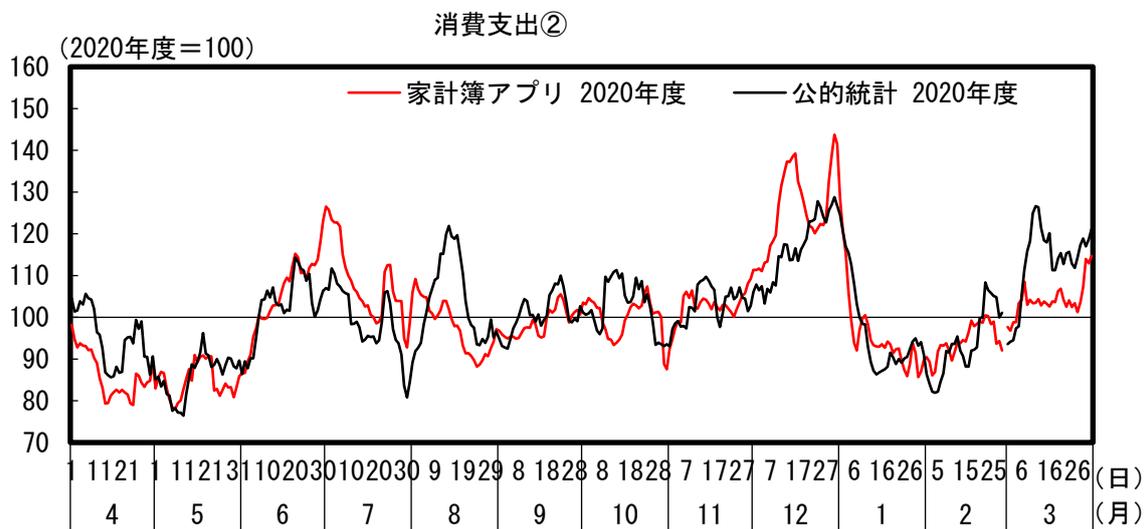
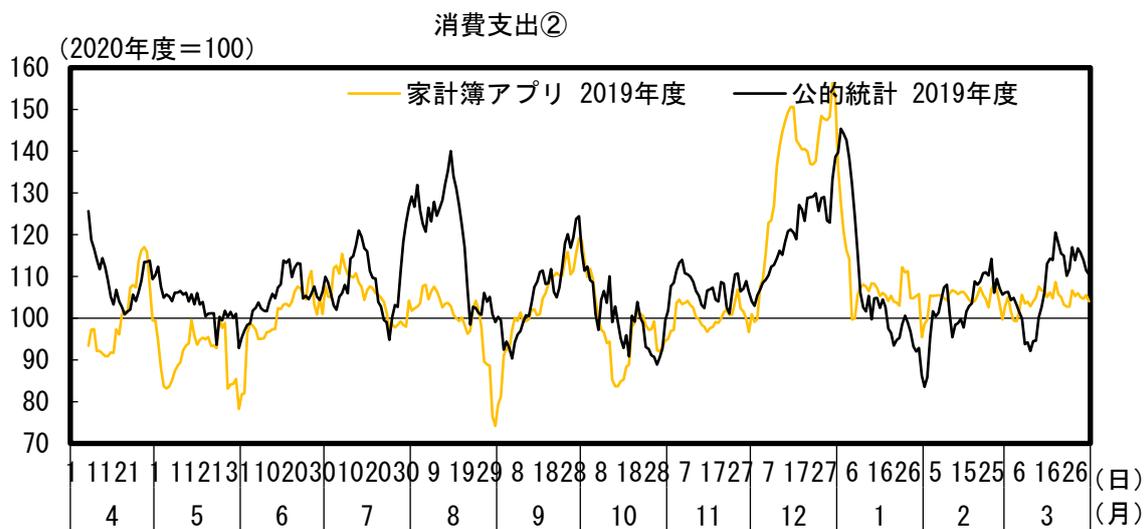
(備考) 公的統計は、総務省「家計調査」により作成。二人以上勤労者世帯の結果。家計簿アプリデータは、図表 2-5-1 を基に作成。

(図表補 1 - 1 日次の動向 (事業者管理データセット・周期性調整後))

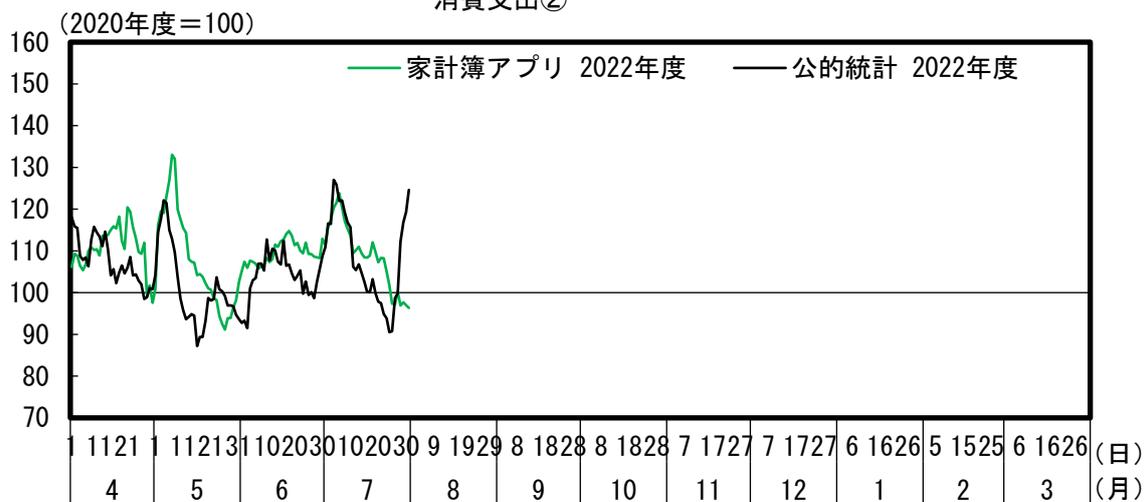


(備考) 図表 2 - 5 - 1 を基に作成。

(図表補1-2(1) 公的統計との比較(周期性調整後・指数))



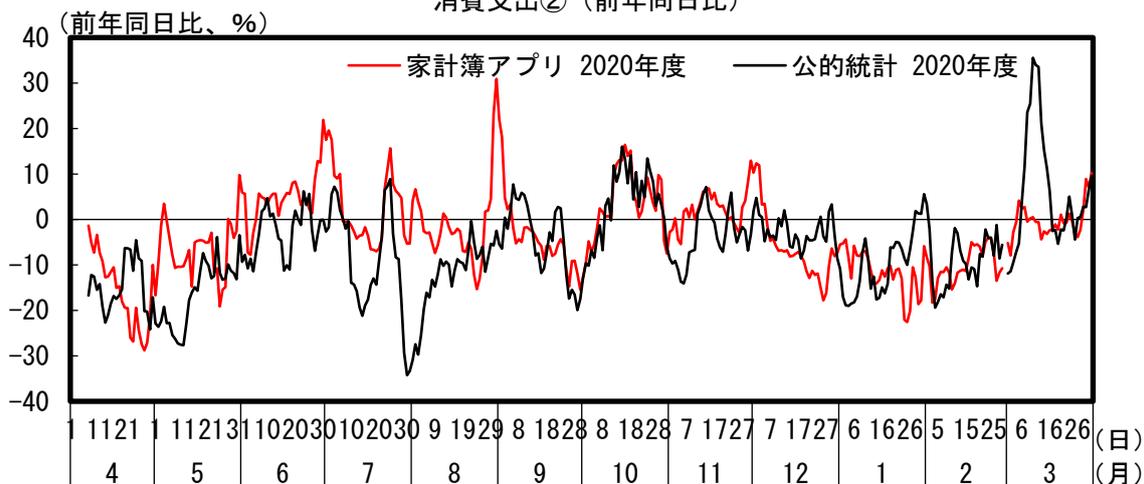
消費支出②



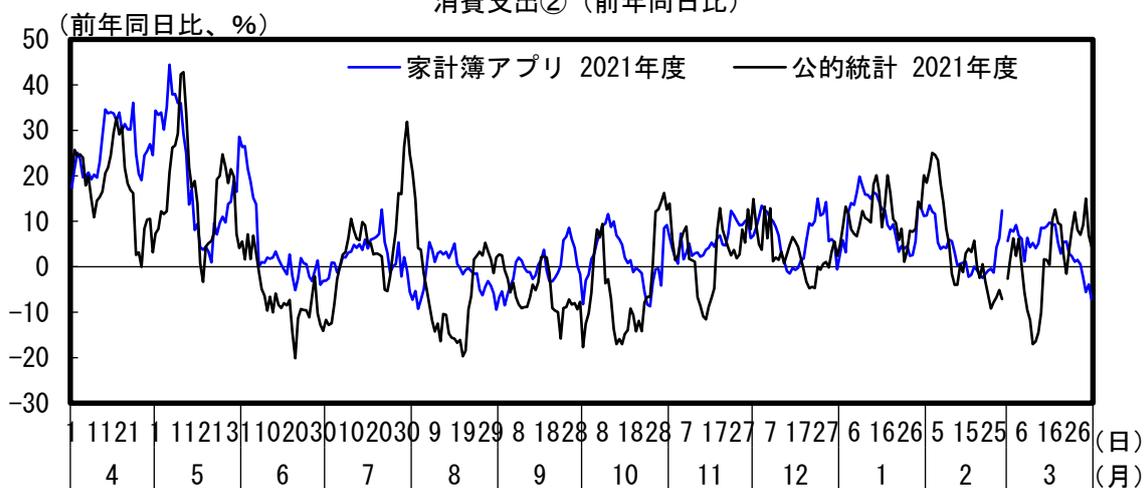
(備考) 公的統計は、総務省「家計調査」により作成。二人以上勤労者世帯の結果について周期性調整を施した。家計簿アプリデータは、図表補1-1を基に作成。

(図表補1-2 (2) 公的統計との比較 (周期性調整後・前年同日比))

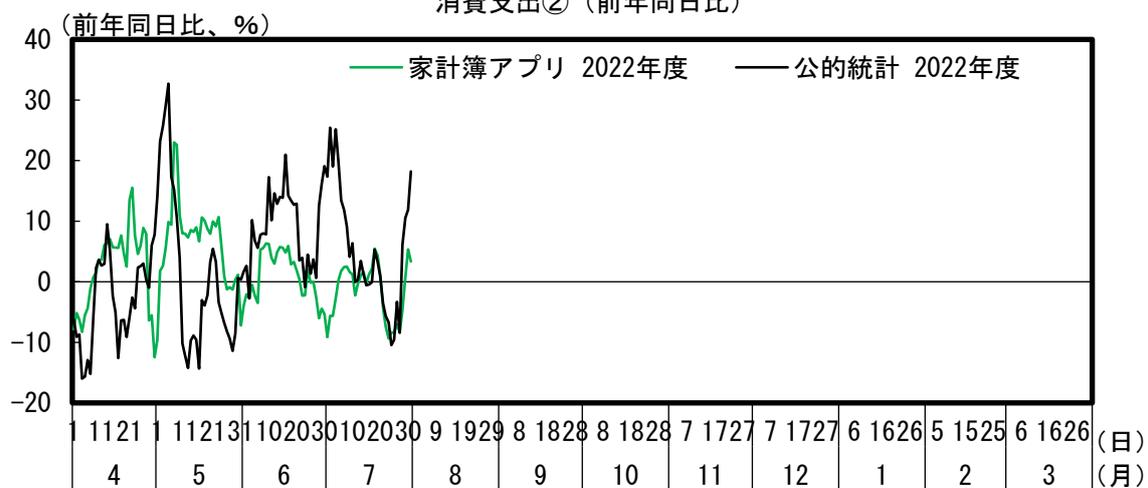
消費支出② (前年同日比)



消費支出② (前年同日比)

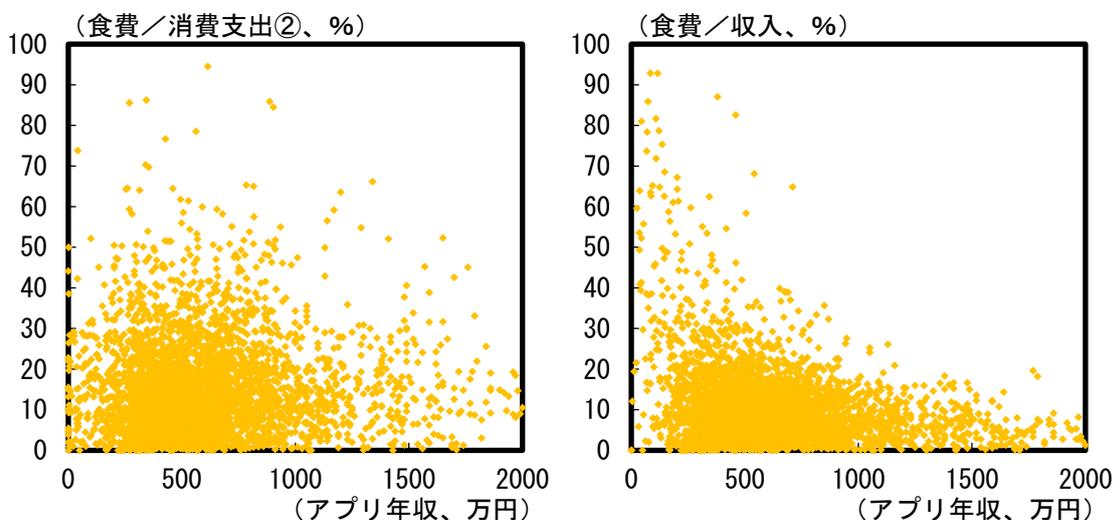


消費支出②（前年同日比）



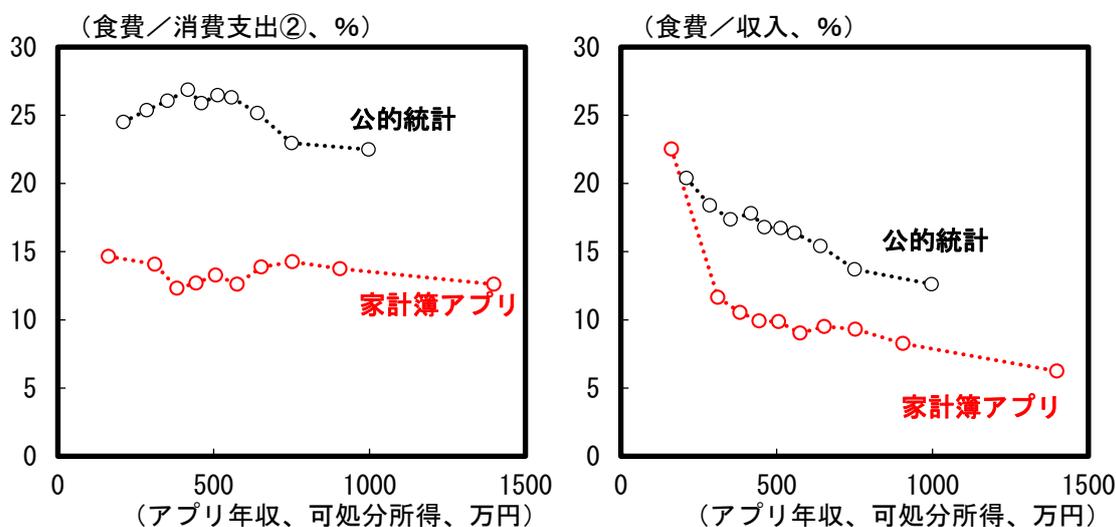
（備考）公的統計は、総務省「家計調査」により作成。二人以上勤労者世帯の結果について周期性調整を施した。家計簿アプリデータは、図表補1-1を基に作成。

(図表補 2-1 (1) 家計における食費の占める割合 (プロット))



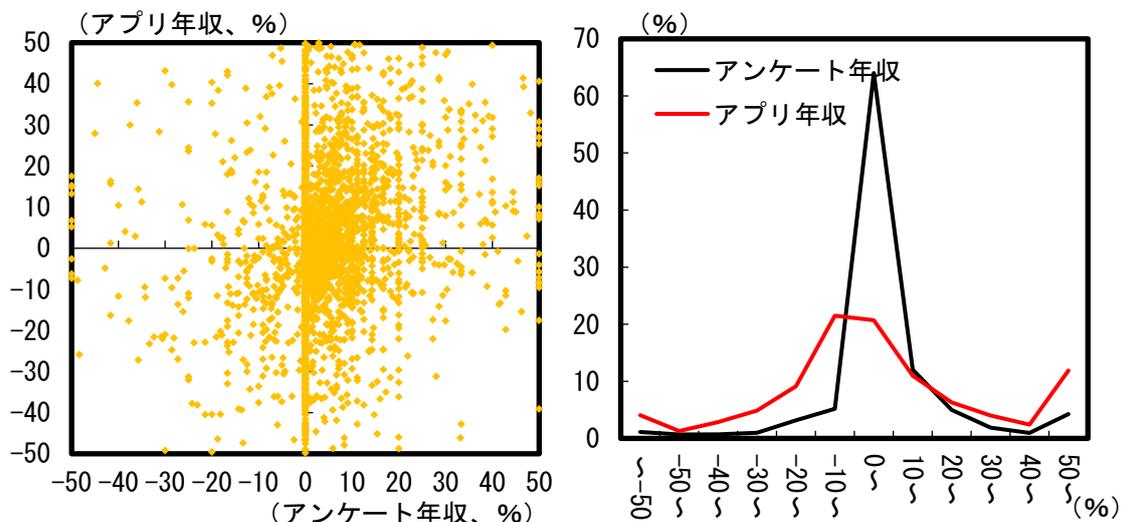
(備考) 内閣府及び株式会社三菱総合研究所作成。横軸は、左右図いずれとも家計アプリから得られた2021年の収入、縦軸は、左図は、家計簿アプリから得られた2021年の消費支出②に占める食費(事業者カテゴリーにおける「食費」、「外食」、「居酒屋・バー」の和)比率、右図は、家計簿アプリから得られた、2021年の収入に占める食費比率をプロットしている。

(図表補 2-1 (2) 家計における食費の占める割合 (公的統計との比較))



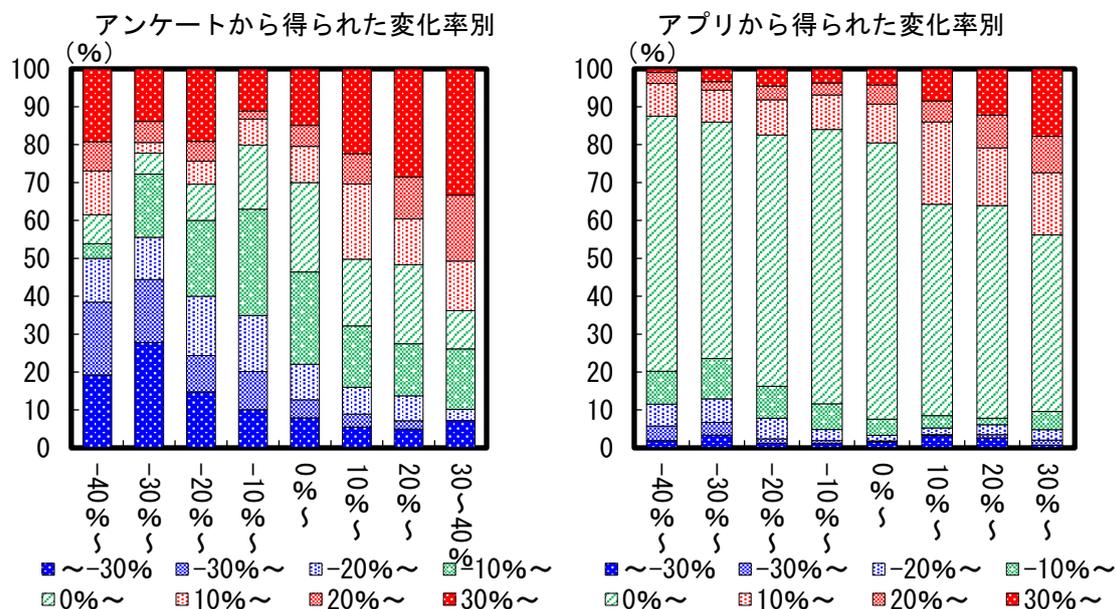
(備考) 公的統計は、総務省「家計調査」、「世帯消費動向指数」により作成。2021年の結果から作成しており、横軸は、左右図いずれとも可処分所得(実収入から勤労所得税、個人住民税、社会保険料を控除した、手取り収入と対応する部分)、縦軸は、左図は消費支出に占める食費の割合、右図は可処分所得に占める食費の割合をプロットしている。消費支出、食費については、家計調査の計数を世帯消費動向指数の水準で補正している。家計簿アプリデータは、内閣府及び株式会社三菱総合研究所作成。横軸は、左右図いずれとも家計簿アプリから取得した2021年の収入、縦軸は、左図は、家計簿アプリから得られた2021年の消費支出②に占める食費(事業者カテゴリーにおける「食費」、「外食」、「居酒屋・バー」の和)比率、右図は、家計簿アプリから得られた、2021年の収入に占める食費比率をプロットしている。いずれも、年間収入(家計簿アプリデータは家計簿アプリデータから取得した収入)十分位の集計値をプロットしている。

(図表補 2-2 (1)) アンケートから得られた収入とアプリから取得した収入の変化率



(備考) 内閣府及び株式会社三菱総合研究所作成。左図は、横軸はアンケートから得られた年間収入の2020年度から2021年度への変化率、縦軸は家計簿アプリから得られた年間収入の2020年度から2021年度への変化率をプロットしている。右図はアンケートから得られた年間収入の2020年度から2021年度への変化率、家計簿アプリから得られた年間収入の2020年度から2021年度への変化率の分布を示している。

(図表補 2-2 (2)) アンケートから得られた収入とアプリから取得した収入の変化率



(備考) 内閣府及び株式会社三菱総合研究所作成。左図は、アンケートから得られた年間収入の2020年度から2021年度への変化率別にみた家計簿アプリから得られた年間収入の2020年度から2021年度への変化率の分布、右図は、家計簿アプリから得られた年間収入の2020年度から2021年度への変化率別にみたアンケートから得られた年間収入の2020年度から2021年度への変化率の分布を示している。