

イントロダクション

2024年秋学期 応用計量経済学

講師：遠山 祐太

最終更新：2024-11-16

はじめに




自己紹介

- 経歴：
 - 仙台市生まれ、仙台一高卒
 - 京大経済→東大院修士（公共政策→経済学）→ノースウェスタン大学Ph.D. in Economics
 - 帰国後、2018年9月より早稲田大学政治経済学部准教授
 - その他のお仕事：UTEconアドバイザー
- 研究分野：実証産業組織論、応用計量経済学、エネルギー・環境経済学
- これまでの研究：
 - 企業合併の経済評価：現代・起亜自動車の事例
 - 電力産業における排出権取引制度の経済評価
 - その他：投票行動モデルの推定、ビッグデータを用いた特別定額給付金の分析

データ分析がますます大事に

ARTICLE | HARVARD BUSINESS REVIEW | OCTOBER 2012

Data Scientist: The Sexiest Job of the 21st Century

 PRINT  SHARE  EMAIL

Abstract

Key to the effective use of big data are the analytical professionals known as "data scientists," who can both manipulate large and unstructured data sources and create insights from them. Data scientists are difficult to hire and retain, but their skills will be necessary to any organization wishing to profit from big data.

Keywords: Big Data; Data Scientists; Business Analytics; Data and Data Sets; Mathematical Methods; Jobs and Positions

Format: Print

[READ NOW](#)

実際にデータ分析は役立つのか？

- Koning et al (2023) "Experimentation and Start-up Performance: Evidence from A/B Testing"
 - A/Bテストを導入することで、スタートアップIT企業のWeb訪問者数が1年で30-100%増加
- Berman and Israeli (2023) "The Value of Descriptive Analytics: Evidence from Online Retailers"
 - データダッシュボードを導入することで、収入が4-10%程度改善

政府・民間企業における活躍の場の拡大

- 政府・公的部門
 - データ分析を政策立案の基礎に用いる
 - 各種省庁・自治体
- テック企業
 - GAFA (Google, Amazon, Facebook, Apple), サイバーエージェント、アリババ、etc
 - 例：オンライン広告のオークション設計、価格付のための需要・価格弾力性を推定、施策の評価(A/Bテスト)
 - Athey and Luca (2018) “Economists (and Economics) in Tech Companies” [日本語要約](#)
- エコノミック・コンサルティング会社 (修士・博士以上)
 - 企業経営における各種問題について経済分析を提供
 - NERA、アリックスパートナーズ、東京大学エコノミックコンサルティング(UTEcon)

The Economist "Why economists are flocking to Silicon Valley" (September 7, 2022)

- 引用: *For big tech, meanwhile, economists offer skills that computer scientists and engineers often lack. They tend to have a good grasp of statistics, as well as a knack for **understanding how incentives affect human behaviour**. Most important, economists are adept at **designing experiments to identify causal relationships** between variables. Machine-learning engineers usually think in terms of prediction problems, ...*

本講義：経済学における因果推論を学ぶ

- 考え方(substance) :
 - データ分析において因果関係を検証する考え方を学ぶ。
 - 研究やジャーナリズムにおける因果関係について批判的に検証する見方を学ぶ
 - 経済学の実証手法の2大アプローチについて学ぶ：いわゆる「誘導形」と「構造推定」
- 手法(method)：各種計量手法・識別戦略を学ぶ
 - ランダム化比較試験 (Randomized controlled trial; RCT)
 - 回帰分析 (regression),
 - 差の差分法 (difference-in-differences; DID)
 - 操作変数法 (instrumental variable; IV)
 - 離散選択モデル (discrete choice models)
 - (回帰不連続デザイン, regression discontinuity design; RDD)
- 実践：R/Rstudioを用いたプログラミング (手を動かしつつ学ぶ)

講義の目標・その1(/ 3)

- 統計学・計量経済学を用いた**より良い実証分析**を行えるようになる。
 - 過去の履修者の中には、データサイエンスコンペ受賞者、政経論文コンクール受賞者など。

講義の目標・その2(／3)

- 統計学・計量経済学を用いたより良い実証分析を行えるようになる。
 - 過去の履修者の中には、データサイエンスコンペ受賞者、政経論文コンクール受賞者など。
- データ分析や因果関係に関する主張を**批判的かつ建設的に検証する力**を身につける。
 - 例：上に挙げた「XX受賞」は本当にこの授業のおかげなのか？
 - なお「建設的」に見るという点がより重要。「相関≠因果」は誰でも言える！

講義の目標・その3(/ 3)

- 統計学・計量経済学を用いたより良い実証分析を行えるようになる。
 - 過去の履修者の中には、データサイエンスコンペ受賞者、政経論文コンクール受賞者など。
- データ分析や因果関係に関する主張を批判的かつ建設的に検証する力を身につける。
 - 例：上に挙げた「XX受賞」は本当にこの授業のおかげなのか？
 - なお「建設的」に見るという点がより重要。「相関≠因果」は誰でも言える！
- 経済モデルを活用した「**構造推定アプローチ**」の考え方を学ぶ。
 - 例：消費者行動分析、需要・供給分析
 - 統計的因果推論手法(誘導形手法)と比較しての強み・弱みを学ぶ
 - (院レベルへの橋がけ)

本日の流れ

1. 因果とは何か？
2. 事例：コロナ禍のデータ分析。特に特別定額給付金の効果について。

因果とは何か？

経済学・政治学における因果関係に関する問い

- 因果関係：XがYを引き起こす・XがYに影響する
- 多くの問いが因果に関するもの！
 - 教育年数が一年増えると賃金はどれだけ上昇するか？
 - 商品の売れ行きにネット広告はどれだけ影響するか？
 - 企業の合併は商品価格を上昇させるか？
 - 民主主義は経済成長を引き起こすか？
 - 投票率の向上は大統領選で民主党に益するか？

理論モデル→因果関係

- 消費者の効用最大化：
 - 所得が増えれば消費も増える（他の変数を固定したとき）
 - 価格が上がれば（通常）消費は減る（他の変数を固定したとき）
- 教育への投資は所得を高める：
 - 高い人的資本
 - シグナリング効果
- 政府支出は乗数効果によってGDPを刺激する

理論からデータへ

- 理論と現実のギャップ? :
 - 本当に理論は現実に当てはまるのか?
 - 理論予測の符号(正・負)や大きさは如何ほどなのか?
- 現実の事象に関するデータを用いて、理論の検証や効果の測定を行う。
- 定量化されたモデルによって、理論及びデータ分析を意思決定に用いることができる。
 - 例：需要関数+消費者の選好パラメタ -> 需要予測・価格付け
 - 例：教育の効果 -> 教育支出の妥当性の検討・費用便益分析

データ分析の第一歩としての記述分析

- 分析はデータの**記述**から始まる。
 - 大卒者はその他に比べて時給が98%高い。
 - 現在の所得不平等は30年前に比べて大きい。
 - 医療制度改革後の医療費の伸びは緩やか。
 - 現在の航空券価格は合併以前より高い。
- 次なる問：発見したデータにおけるパターンを**どのように解釈するか？**
- **データで発見した相関を因果として解釈できるのか？**

キーポイント：相関は因果を意味するか？

- データを見たときにXとYが共に変化しているとしよう（**相関**）
- 例：
 - 1：警察官の多い都市は犯罪も多い。（正の相関）
 - 2：大学に進学した人の方が10%収入が多い。
- 問：
 - これは本当に「XはYを引き起こす」ことを意味するだろうか？
 - 相関の大きさは、因果効果の大きさとして解釈できるだろうか？

三つの可能性

1. 因果(知りたいもの)
2. 逆因果
3. 第三の要因：**交絡因子 (confounder)**

三つの可能性を図で

事例：アビガンはCovid-19の治療に有効か？

- Covid-19に罹患した著名人の中には、抗インフルエンザ薬「アビガン」が治療に有効だったと語る人もいます。
 - 石田純一さんの体験談
- これはアビガンがCovid-19に対するSilver bullet(切り札)であることを意味するだろうか？
- なぜ重要な問いか？
 - アビガンの副作用、公共支出の使い方、などなど

石田純一「一刻の猶予もない」でアビガン「大量投与」効いて平熱に コロナ禍中の沖縄行きも謝罪「非常にまずかった」

2020年4月23日 22時25分

ツイート

BI 1



石田純一



新型コロナウイルスに感染して入院中の俳優石田純一（66）が23日、文化放送の「斉藤一美 ニュースワイドSAKIDORI！」（月～金曜午後3時半）で病床から肉声を伝え、外出自粛要請が出る中で沖縄に行ったことを謝罪した。また一時症状が悪化したものの、抗インフルエンザ薬のアビガンを処方して回復したことも明かした。

石田は同番組の木曜コメンテーターを務める。22日に収録した電話インタビューを放送した。

10日に沖縄に渡り、経営する飲食店で打ち合わせをしたことに「非常にまずかった。反省しています」とかすれ気味の細い声で謝罪。11日、仕事関係者とゴルフをプレー中にだるさを感じたが、沖縄のホテルに13日まで滞在し帰京した。「ホテルにも大変ご迷惑をおかけし、沖縄の人たちに不快な思いをさせて

難しさ 1: 反実仮想 (counterfactual)

- **反事実的アウトカム (counterfactual outcome)** は観察されない。
- アビガンがなくても、患者は回復したかもしれない。

難しさ 2 : 内生的な処置の選択

- 治療でアビガンを投与された患者と投与されなかった患者のデータがあるとしよう。
- これらの人々の比較でアビガンの処置効果（treatment effect）がわかるだろうか？
- ポイント：アビガンを投与された患者とされなかった患者は、他の面でも異なるかもしれない。

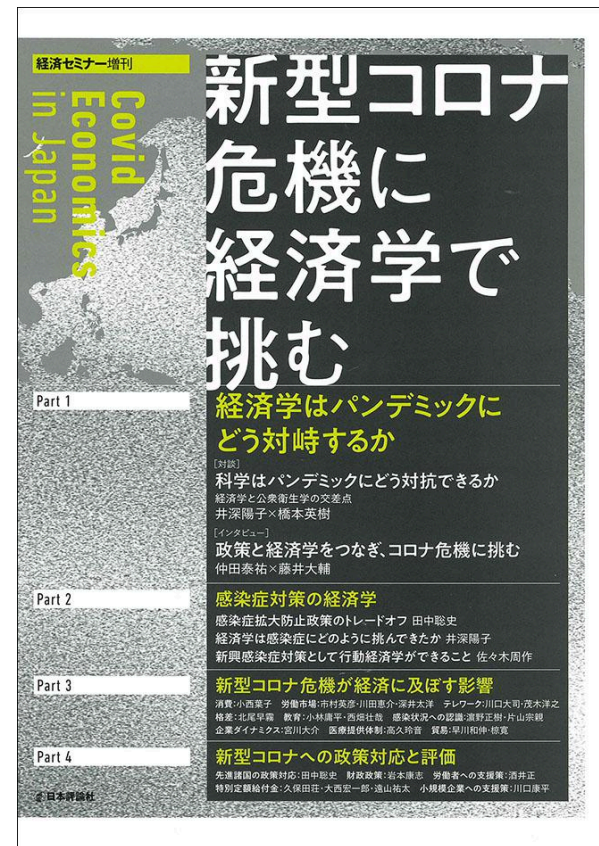
解決策：ランダム化比較試験

- 患者を集め、処置を受ける患者を**無作為に（randomly）**選ぶ。
 - 処置群（treatment group）と統制群（control group）
- 処置群と対照群の結果を比較し、処置の効果を確認する。
- ランダム化により、**両群は平均的には似通っており、唯一の差が処置の有無になる。**
- 2020年以降、様々な機関がCovid-19に対するアビガンの臨床試験を行ったところ、その効果は認められなかった。

事例：コロナ禍のデータ分析

背景：新型コロナ危機における経済・実証分析

- 新型コロナ危機における経済分析
 - マクロ+SEIRモデル
 - ナウキャストイング：米国, JCB
 - 政策評価(例：各種給付金・GoTO)
- 背景
 - 速報性の重視 (Covid Economics)
 - オルタナティブデータの広がり(GPS、クレジットカード、などなど)
- 経済セミナー別冊「新型コロナ危機に経済学で挑む」に多くの例

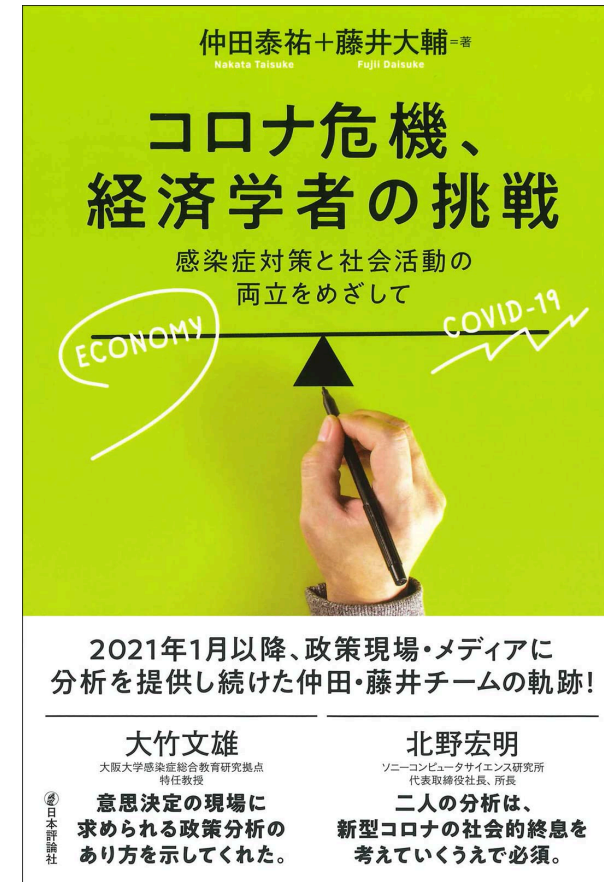


政策課題分析の難しさに関する私見

- データ・課題を所与としての分析
 - ランダム化比較試験は(ほぼほぼ)できない
 - 自然実験であるとは限らない。
 - 必要な情報が全て手に入るとは限らない。
- 難点を認識し、「何がどこまで言えるか(限界・留保点)」を把握しつつ、分析を行う。
- そのためには、理想的な状況における分析方法、各種ツールの利点・欠点の理解が重要。

(参考) リアルタイム・コロナ分析

- 仲田・藤井チーム(東京大学)は2020年末以降、日本のコロナ感染と経済活動に関する分析をリアルタイムでの発信を続けてきた。
- Site:
<https://covid19outputjapan.github.io/JP/index.html>
- 書籍の中では、時間・労力・データ制約の中で政策形成に資する情報を如何に発信していくか、という試みに関する裏舞台が描かれている。



テーマ：特別定額給付金の消費喚起効果

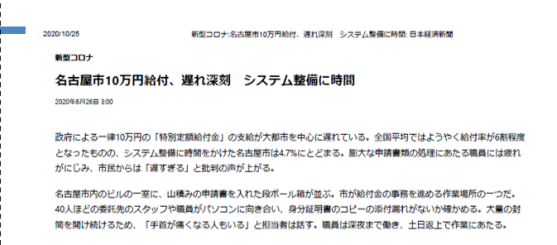
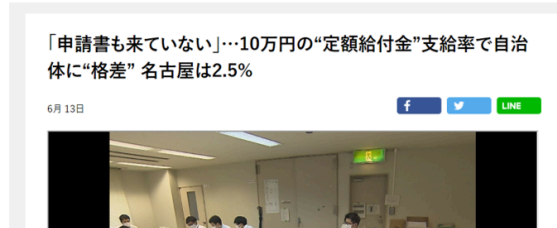
- Kubota, Onishi, and Toyama (2021) "Consumption responses to COVID-19 payments: Evidence from a natural experiment and bank account data"
- 問：10万円の特別定額給付金がどれだけ消費に回ったか検証した
- モチベーション：
 - 政策：特別定額給付金の政策効果検証。
 - 学問：所得のうちどれだけが消費に回るかという限界消費性向（Marginal Propensity to Consume, MPC）はマクロ経済学上で理論・実証両面で重要なコンセプト。特に近年は異質的個人モデルで注目される。
 - 経験：銀行口座データがどれだけ正確性・速報性の面で研究・政策貢献できるか。
- アプローチ：銀行口座の匿名データ+自然実験

制度の経緯と概要

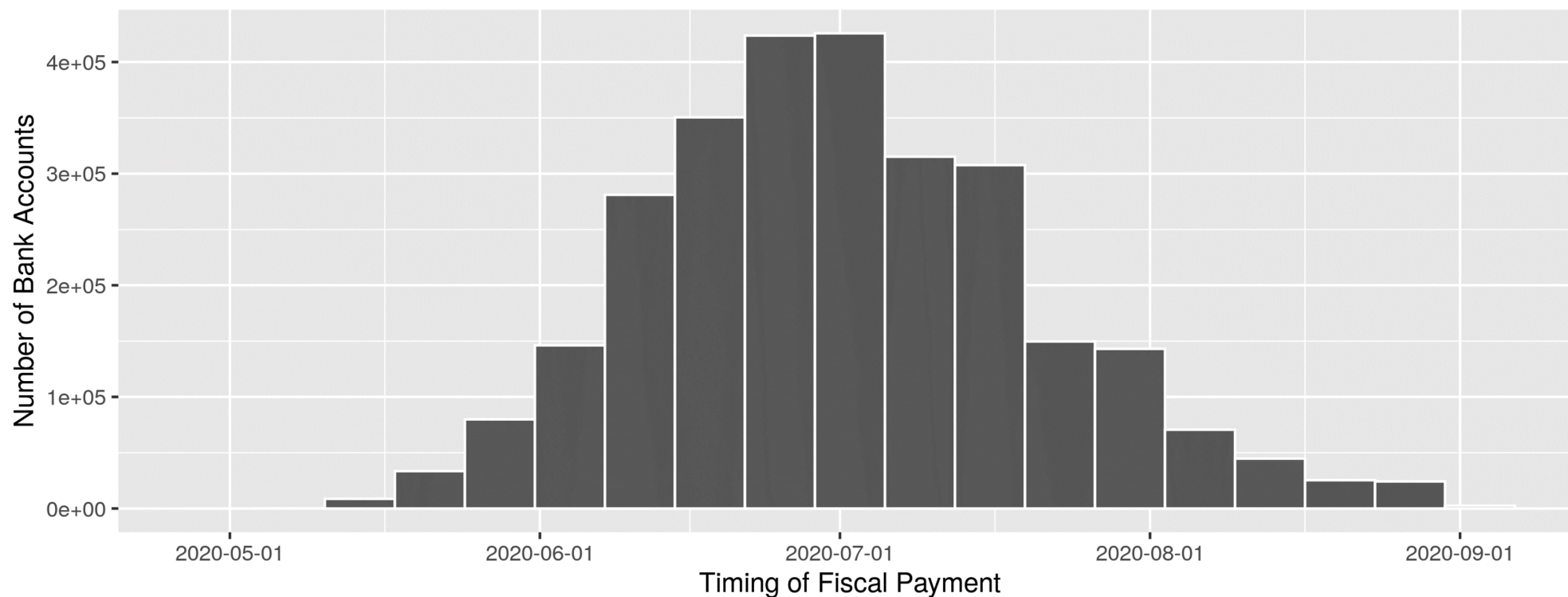
- 経緯：
 - 2020年4月3日：安倍前首相、コロナで収入が大きく減少した世帯に30万円給付を表明
 - 2020年4月17日：1人あたり一律10万円へと変更される。
 - 2020年4月末より自治体において申請が開始された。
- 制度の概要
 - 一人あたり一律10万円が支給される。
 - 申請は住民票における世帯単位。
 - 自治体から申請書類が来た後に、郵送・オンライン申請を行う。
 - 申請が受領された後に、世帯ごとの総額が申請した銀行口座にまとめて振り込まれる。

特徴：支給のタイミングのバラツキ

- 申請・支給のタイミングの地域差及び全体的な遅れ
 - 6月中旬時点で札幌市ではほぼ配布が終わっていたが、名古屋市では給付開始。
 - 同一自治体で同じ日に申請しても、処理に時間がかかり支給日に大きな差が出た。
 - オンライン申請や、書類不備によって、給付日がさらに不確実になった。



振り込みのタイミング(銀行データより)



データ：みずほ銀行における口座取引の匿名化データ

- みずほ銀行と早稲田大学の学術交流協定により利用可能となったデータ
- 期間：2019年1月から2020年8月
- サンプル口座：口座の取引情報から特別定額給付金を受領した世帯を識別
 - 振込金額(10万円単位)、摘要情報(「トクベツ」や「テイガク」など)
 - 最終的な口座数：約280万口座(世帯)
 - サンプルサイズ：約2億 (週次一口座レベル)
- 変数：
 - 支出指標：(1)各週における口座からの出金額, (2)ATMからの出金額(現金)
 - **給付金の支給日(週)**
 - 家計属性：口座保有者の年齢、世帯人数(給付金額から)、資産額、預金額、給与額

おまけ：研究のイメージ



「マグロ」を捌く難しさ：国会答弁より

- 質問書
- 答弁書

衆議院議員中谷一馬君提出一人につき十萬円の特別定額給付金を再給付することに関する質問に対する答弁書

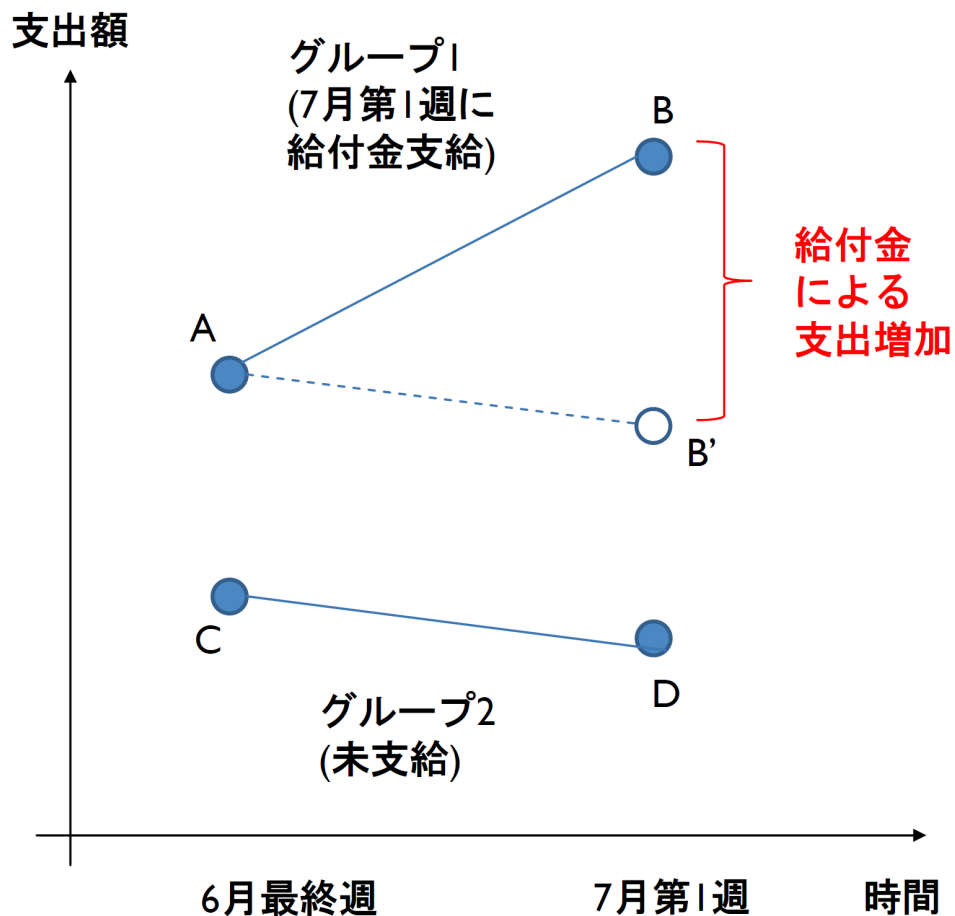
一及び二について

特別定額給付金による経済や国民生活への効果については、先の答弁書（令和二年十月二日内閣衆質二〇二第二号）二についてでお答えしたとおり、特別定額給付金が「新型コロナウイルス感染症緊急経済対策」（令和二年四月二十日閣議決定。以下「緊急経済対策」という。）の一連の施策の一つであり、特別定額給付金以外にも様々な施策を講じていることに加え、「新型コロナウイルス感染症緊急事態宣言」の解除による経済や国民生活への影響も生じていた中で、特別定額給付金の効果のみを抽出することは技術的な困難を伴うと考えられる。なお、御指摘の総務省の家計調査においては、令和二年六月分及び七月分の「家庭用耐久財」や「教養娯楽用耐久財」への支出が前年同月に比べて伸びており、特別定額給付金による一定の効果があった可能性も示唆される。

識別戦略

- 「観察データ」であるがゆえの課題：
 - 全ての人々が給付を受けるため、厳密な意味で対照群(コントロール・グループ)は存在しない。
 - 多くの交絡因子：コロナ禍によるマクロ経済の悪化、家計により異なる経済状況、など。
- 「ランダム化比較対象実験」が行えない状況において、どのように因果効果を推定するか？
- 解決策：差の差分法 (Difference-in-differences)による推定
 - ポイント：**支給タイミングが異なる家計**を上手に比較する。
 - 支給のタイミングがランダムであったらろうという点。

差の差分法のアイデア



- 2グループ
 - 1 : 7月第1週に給付金受け取り(支出額AとB)
 - 2 : それ以降に給付金受け取り(支出額CとD)
- 給付金の効果は？
- 前後比較AB : 給付金以外の影響も含むかも？
- 交差比較BD : そもそもグループ1の支出大きい
- **差の差分法(Difference-in-differences)**
- 「もしグループ1が給付金を7月第1週にもらわなかった場合」の支出をB'と推定する。
 - グループ2と同様の変化を想定する。(B'-A=D-C)
- **給付金の効果はBB'となる！**
 - $BB' = (B-A) - (D-C)$

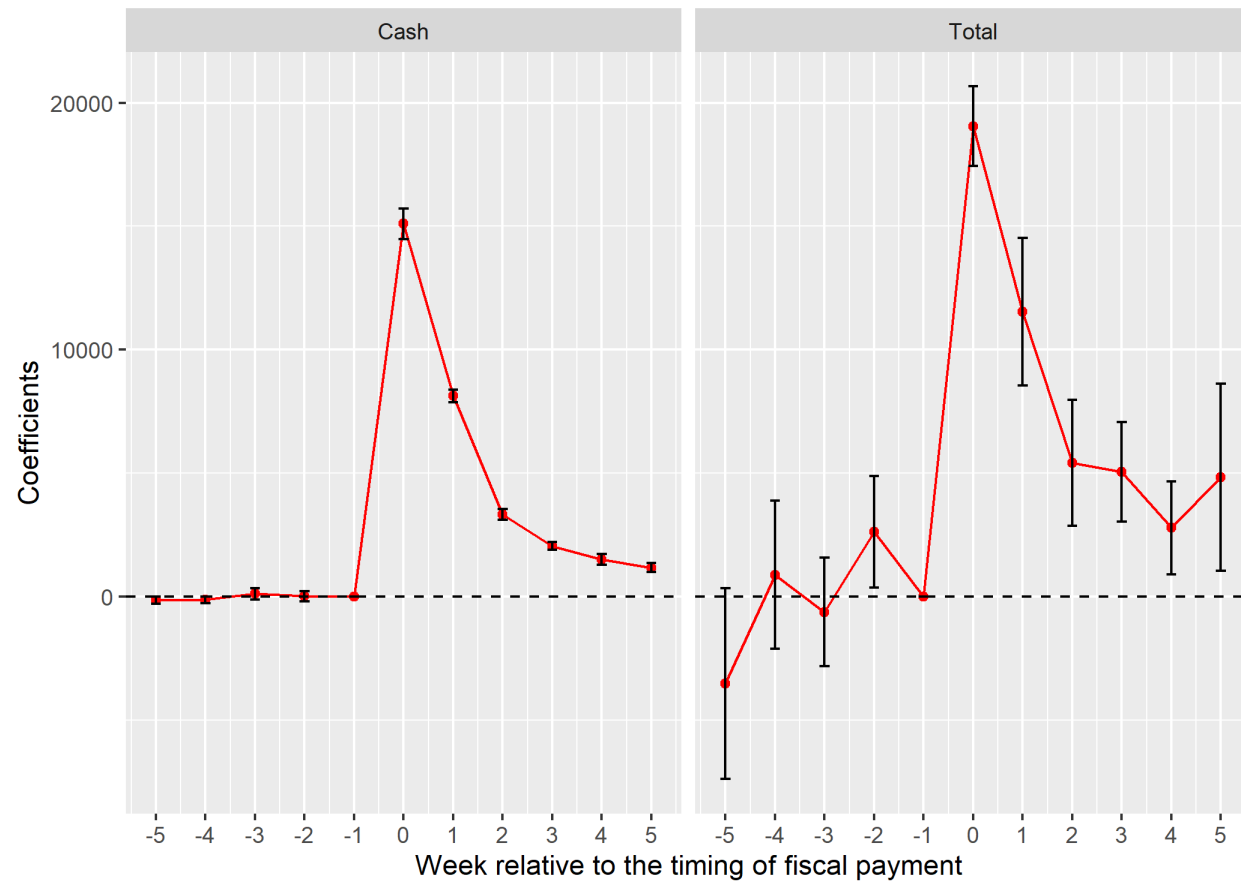
(今後の授業で詳しく) イベントスタディ型のDID

- 回帰式

$$y_{itw} = \alpha_i + \alpha_{iw} + \alpha_{tpw} + \sum_{k=-5}^5 \gamma^k D_{itw}^k + u_{itw}$$

- $t(= 2019, 2020)$ 年の w 週における家計 i
 - y_{itw} : 消費の指標 (全出金、ATM現金)
 - $D_{itw}^k = 1\{T_i - w = k\}$,
 - T_i 家計 i が給付金を支給した週 (トリートメントのタイミング)
- 各種固定効果
 - α_i : 個人固定効果
 - α_{iw} : 個人かつ週固定効果 (季節性)
 - α_{tpw} : 時間かつ県固定効果 (地域特有のマクロショック)

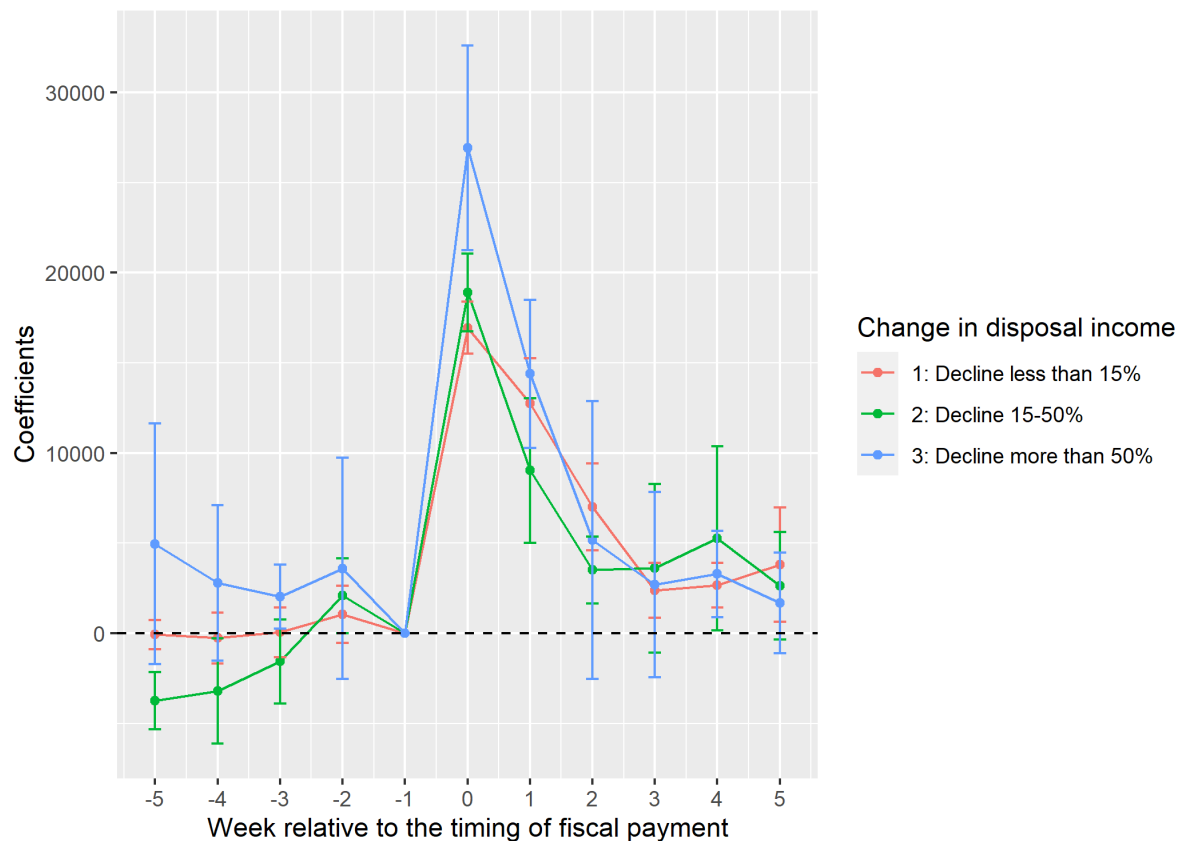
推定結果：全サンプル (左：ATM、右：全出金)



- バーは95%信頼区間

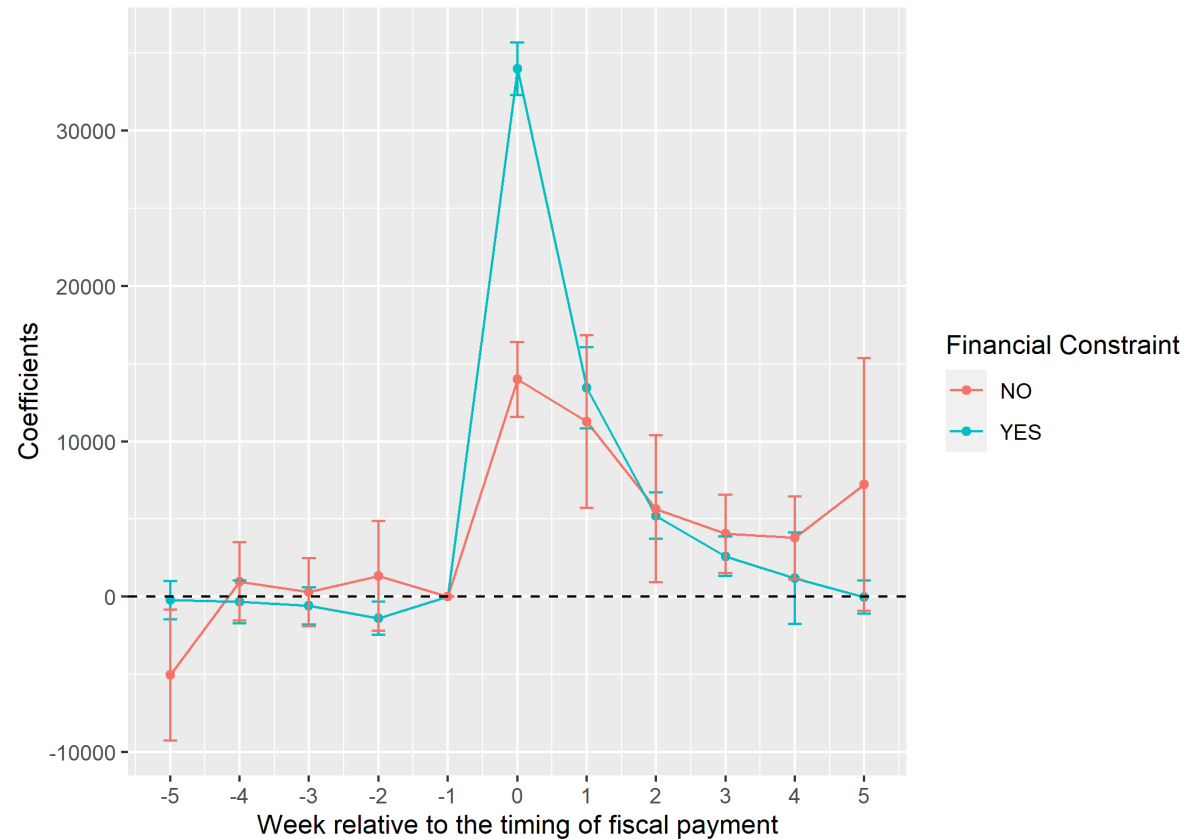
異質性：新型コロナ禍前後の所得変化 (指標：全出金)

- 定義：2020年4・5月における前年給与からの減少率



異質性 : Financial constraint (指標 : 全出金)

- Liquidity constraint: 前月末における口座残高が月収よりも低いか否か。



ディスカッション

- 本論文の限界：消費指標が口座からの出金であり、本当に消費に繋がっているか？
- 特別定額給付金に関するその他研究：
 - Kaneda, Kubota, and Tanaka (2021): マネーフォワードの個票データ。DID。
 - 宇南山・古村・服部(2021): 「家計調査」の都市別集計データ。他2つとは異なる識別戦略。
- 消費効果の推定値は異なるが、データ・分析デザインの差で説明可能。
 - Kubota et al: 31-49% (ATM、全出金)
 - Kaneda et al: 6-27% (消費、消費+ATM、全ての支出)
 - 宇南山 et al: 10%
- 研究・分析を積み重ねることが重要！