

令和 5 年 6 月 6 日現在

機関番号：13901

研究種目：挑戦的研究（萌芽）

研究期間：2020～2022

課題番号：20K20756

研究課題名（和文）非対照実験を用いた政策評価・形成プロセスの戦略的設計

研究課題名（英文）Strategic Design for Evaluating and Formulating Policies Using Non-randomized Trials

研究代表者

田村 彌（Tamura, Wataru）

名古屋大学・経済学研究科・准教授

研究者番号：60711950

交付決定額（研究期間全体）：（直接経費） 1,300,000円

研究成果の概要（和文）：本研究では、無作為化の実施可能性に制約がある状況における実証実験デザインについて理論的な分析を行った。特に、単純な群間比較がセレクションバイアスの影響を受けやすい場合に焦点を当て、理論モデルの解析とシミュレーションを行った。モデルの仮定が満たされている場合、二段階の実証実験を用いて規模を拡大した場合の処置効果が推定できることを示し、費用便益分析を行った。結果の頑健性についてパラメータの分布と実験参加者の行動バイアスそれぞれの観点から検証し、理論的分析を現実に応用する際の問題を整理した。

研究成果の学術的意義や社会的意義

エビデンスに基づく意思決定は現代の政策形成や経営においてますます重要となっている。しかし、質の高いエビデンスを得るための重要な手段である無作為化は、技術的な制約や倫理的な理由から必ずしも利用できるわけではない。この課題に対して、本研究では無作為化が難しい状況でもデータを有効に活用するための理論を提供している。同時に、本研究は小規模または中規模の実証実験から大規模な展開時の政策効果を推定するときに生じる規模効果の問題に対処するための新たな手法を提案している。本研究は特に予算や人員が限られている小規模な組織や公平性を問われる公的組織のデータ活用に貢献する可能性がある。

研究成果の概要（英文）：In this study, we conducted a theoretical analysis of experimental design in situations where randomization is not feasible. In particular, we focused on cases where a simple between-group comparison is susceptible to selection bias, and we both analyzed and simulated our theoretical model. We showed that when the model assumptions are satisfied, the treatment effect at scale is estimated using a two-round experiment method, and conducted cost benefit analysis. We examined the robustness of our results with respect to both the distribution of parameters and the behavioral biases of the experimental participants, and categorized the challenges associated with applying theoretical analysis to real-world situations.

研究分野：理論経済学

キーワード：実験設計 インセンティブ 無作為化

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

エビデンスに基づく政策立案(EBPM)は具体的なデータや研究結果を基に政策を立案し、効果を評価するというアプローチであり、特にランダム化比較試験(RCT)による因果効果のエビデンスが重視されている。これは、RCTを用いた社会実験または類似の実証研究によって推定される「プログラムの因果効果」を高品質なエビデンスと見なす学術的な視点に由来する(Dufulo et al., 2008)。計量分析手法を利用した事後的な政策評価は従来主流だったが、近年では、政策立案・形成の前に、開発、医療、教育などの分野で実験的な手法を使ったプログラム評価や理論検証が一般化しつつある。しかしながら、RCTは理想的な評価手法とされる一方で、適切なランダム化の実施可能性や政策効果のスケラビリティ(同一の政策が異なる実施規模で同様の効果をもたらす能力)に関して問題点が残されている(Davis et al., 2017; Kowalski, 2022)。

2. 研究の目的

本研究の目的は、ランダム化実験が適用できない状況におけるプログラムの試験的实施、評価、及びスケリングの過程を新たなデザイン問題として定式化・分析することである。特に新規性があるのは、政策決定者が実験計画やスケリング過程に戦略的に介入するモデルを考察する点で、以下の3つを特徴とする：

1. 非対照試験を通じてデータを収集し、政策評価を行う。これはランダム化試験が不可能な場合に対応する手法である。
2. プログラム事業者が戦略的に実験計画への介入を行う段階的な意思決定プロセスを分析する。
3. 費用対効果の最適な戦略をデザインするため、プログラム評価の精度と実証実験・政策実行にかかる費用の間のトレードオフを考慮する。

本研究の意義は、一方で学術における科学的政策評価と、他方で経営や政策形成における現実のEBPMとのギャップを埋めること、そして統計・計量理論と制度設計理論を応用した新たなデザイン問題を考察することである。具体的には、ランダム化が困難な状況でもデータを活用し、科学的な統計分析と現実の政策過程や企業戦略を融合する理論を提案することを目指す。また、大規模展開時の政策効果を小・中規模の実証実験から推定する新たな調整手法を探索し、スケリング問題への対処方法を提案する。

3. 研究の方法

本研究では、実証実験のプロセスをデザイン問題として理論的に定式化すること、ランダム化が行えない状況下での限界や利用可能な統計量を理論的に特徴づけること、シミュレーションを行うこと、実験計画を分析することの4つを計画し、実行した。

定式化：因果推論において標準的なRubin(1974)モデルに基づきプログラム評価の目的や状況設定をモデル化する。データを所与とする統計・計量理論とは異なり、事業者がデータを生成する最適デザイン問題を定義する。Banerjee et al. (2020)では、実験者が被験者の処置割り当てを完全にコントロールできると仮定し、実験者の目的関数を最大化する(ランダム化を含む)処置割り当てルールを考えている。それに対して、本研究では、実験者が処置割り当てをコントロールできないため、被験者のインセンティブを通じて割り当てに介入する点が特徴となる。また、介入方法については、メカニズムデザインのように一般的な枠組みで特徴づけるアプローチと、何らかのパラメータについて費用便益分析や比較静学を行う古典的なアプローチが考えられるが、本研究では後者のアプローチを採用した。

分析：戦略的な介入操作を利用した非対照試験はバイアスのない推定が行えるのか、どのような条件が必要か、といった推定精度に関する分析を行った。その後、実験費用と推定精度のトレードオフを分析した。

シミュレーション：理論的な分析で得られた知見について、シミュレーションによって定量的な分析を行った。最も費用対効果の高い制度を特徴づけた。

実験計画：現実の環境で本研究の妥当性や応用可能性を検証するための実験計画を策定した。理論モデルでは省略された要素が実際の運用において問題になるかを考察した。

4. 研究成果

この研究の理論的部分は「Using Uncontrolled Trials with Selection Bias and Scaling up」というタイトルとして研究会等で発表している。

基本モデルでは、実験者が2回にわたり実証実験を実施する設定を考えている。ある処置につい

て、各個人は潜在的な結果 (potential outcome) つまり、処置がある場合の結果と処置がない場合の結果 が定義されている。処置がない場合の母集団全体の平均的な結果は公知と仮定し、処置した場合の母集団全体の平均的な結果を推定することが実験者の目的となる。実証実験を実施し、そこに参加した被験者のみサンプルとして「処置を受けた場合の結果」が判明する。実証実験の対象が母集団からランダムに抽出できる場合、処置群の平均的な結果と公知である対照群の平均的な結果を比較することで、母集団全体の処置効果についてバイアスのない推定ができる。本研究では、個人が自発的に参加を希望する場合のみ実験対象となる設定を考える。これにより、単純な平均的結果の比較にはセレクションバイアスが生じる。また、参加希望者をランダムに処置群と対照群に分けることもできないという制約を設けた。

このような設定において、被験者に対して実験参加のインセンティブを付与する介入を行い、被験者プールを変化させることで追加的な情報を得ることを考えた。具体的には、金銭移転と情報開示の 2 種類を比較した。金銭移転は実験参加自体に金銭的な報酬あるいは支払いが生じる介入であり、情報開示は実験の効果に関する 1 回目の結果を開示することで 2 回目の参加を促進または抑制する介入である。ランダム化ができなければ基本的にはバイアスのない推定はできないが、被験者のコストパラメータが指数分布に従っており、かつ被験者の観察できない属性にその分布が依存しない場合には、金銭介入ありの実験で得られた追加情報によって、母集団全体の処置効果がバイアスなく推定できることを示した (厳密にいうと不偏性ではなく一致性である)。一方、同様の介入である情報開示ではそのような結果が得られないことが示された。ここで得られた理論的結果は、ハザード率が一定という指数分布の特性に大きく依存する。

図 1 は単純な平均比較 (est.base; 赤)、金銭移転 (est.m; 青)、情報開示 (est.info; 黄) それぞれの方法についてシミュレーションによって推定量のばらつき方を可視化した。真の処置効果は 0.2 に設定されている。平均比較や情報開示では処置効果を過大に推定しており、セレクションバイアスの影響を受けている。一方、金銭移転については平均的に真の値 0.2 周りに推定値が散らばるが、予測精度はそう高くないことが判明した。

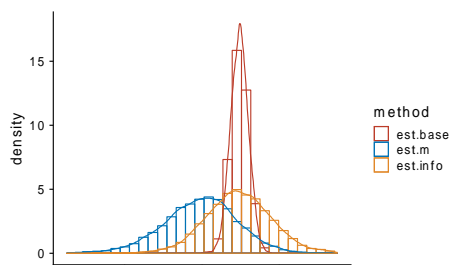


図 1: 推定量の分布

さらに、処置効果が個人間で異なる場合、処置効果のスケール効果が生じる。本モデルでは、これは最初の実証実験に参加するインセンティブを持つような属性を持つグループとそうでないグループとの間で処置効果

果が異なることに起因する。そのように拡張された設定で、RCT や操作変数法などの既存の計量手法と本研究で提示する方法を比較し、それぞれの統計量が推定する estimand の違いを数学的に記述した。

上記の結果をもとに、実験費用と推定精度のトレードオフについての費用便益分析としてシミュレーション分析を行った (図 2)。その結果、推定精度を費用の関数として見たときに凸関数の形状になる、つまり追加的な費用に伴う便益は逡減していくことが判明した。モデルの仮定が満たされる範囲では、大規模に実験を実施するときあまり変わらない予測精度を中規模実験において得られることを示している。

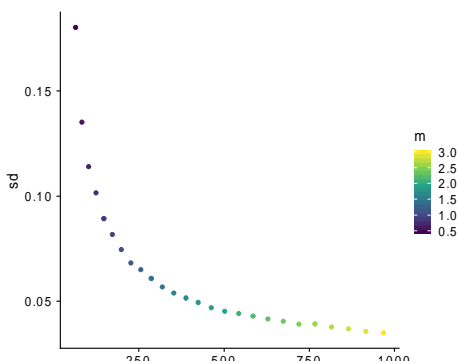


図 2: コストトレードオフ

上記の理論的分析を現実に応用する際の問題を類型化した。実験者の金銭移転の操作により属性の異なるプールを作ることが理論分析における基本的なアイデアであるが、その操作自体が様々な認知バイアスや行動バイアスを通じて処置効果に影響する可能性がある。特に処置効果が個人の努力量に左右される場合にはその問題が顕著となることが予想される (例えば Gneezy and List (2016) で扱われている問題など)。

またオンライン実験のように被験者の異質性が大きく事前情報の精度が低い状況では、理論モデルが仮定する分布のサポート条件を満たさない可能性が高い。

本研究期間を通じて、ランダムに処置を割り当てるのが難しい状況においても処置効果に関する一定の情報を得る方法を提案することができたが、一方で様々な限界があることも明らかになった。今後の研究において、既存の実験設計や効果検証方法との組み合わせや行動バイアスについて頑健な方法などの発展が期待される。

<引用文献

Banerjee, A. V., Chassang, S., Montero, S., & Snowberg, E. (2020). A theory of experimenters:

Robustness, randomization, and balance. *American Economic Review*, 110(4), 1206-1230.

Davis, J. M., Guryan, J., Hallberg, K., & Ludwig, J. (2017). The economics of scale-up. *National Bureau of Economic Research*.

Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics*, 4, 3895-3962.

Gneezy, U., & List, J. A. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74(5), 1365-1384.

Kowalski, A. E. (2022). How to examine external validity within an experiment. *Journal of Economics & Management Strategy*. *Forthcoming*.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 Wataru Tamura
2. 発表標題 Bayesian Persuasion with Quadratic Preferences
3. 学会等名 2022 Conference on Mechanism and Institution Design (国際学会)
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------