

# Identification: Target Trial / Normative Commitment

労働経済学

川田恵介

## Table of contents

1	識別	2
1.1	Recap: Estimation . . . . .	2
1.2	Estimand . . . . .	2
1.3	Estimand . . . . .	3
1.4	まとめ . . . . .	3
2	格差の識別	3
2.1	規範命題 . . . . .	4
2.2	$Y, D =$ 格差を定義 . . . . .	4
2.3	$X =$ 格差を定義 . . . . .	4
2.4	例: 取り締まり (Rose 2023) . . . . .	4
2.5	まとめ . . . . .	5
2.6	まとめ . . . . .	5
2.7	注意: 適切な変数選択 . . . . .	5
2.8	注意: 不適切な変数選択 . . . . .	5
3	因果効果の識別	5
3.1	動機 . . . . .	6
3.2	課題 . . . . .	6
3.3	因果についての枠組み . . . . .	6
3.4	実験アプローチ . . . . .	6
3.5	理想的な実験: Controlled Experiment . . . . .	7
3.6	代替案: Randomized Controlled Trial . . . . .	7
3.7	実行可能な実験 . . . . .	7
3.8	Example: Resume Experiment . . . . .	7
3.9	Example. Resume Experiment . . . . .	7
3.10	Example . . . . .	8

3.11	自然実験 . . . . .	8
3.12	$X$ の選択への含意 . . . . .	8
3.13	例: STEM の効果 . . . . .	9
3.14	例: Fukai, Kawata, and Nakabayashi (2023) . . . . .	9
3.15	警句 . . . . .	9
3.16	まとめ . . . . .	9
4	推定への含意 . . . . .	10
4.1	RCT データへのコントロール . . . . .	10
4.2	例: ResumeName . . . . .	10
4.3	定式化依存への誤解 . . . . .	11
	Reference . . . . .	11

## 1 識別

- 労働経済学研究のデザイン: 何をどのように推定するのか?
  - ここまでの議論はどのように  $E[Y|d', X] - E[Y|d, X]$  を推定するのか (Estimation)
  - 何を推定するのかは (より?) 重要

### 1.1 Recap: Estimation

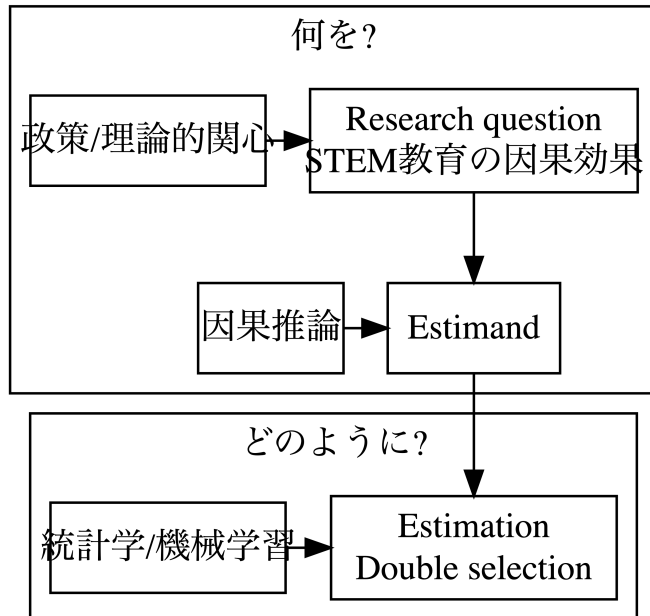
- 前回までは、 $E[Y|D = d, X] - E[Y|D = d', X]$  の特徴を推定する方法を議論
  - 暫定的に  $E[Y|D = d, X] - E[Y|D = d', X] = \tau$  を仮定
  - OLS を使用  $Y \sim D + X$ 
    - \* Double selection によって、 $X$  の重要な部分集合  $Z$  を特定し、 $Y \sim D + Z$  を推定
- Linear model が前提であっても、そもそも何を  $X$  として使用するか?

### 1.2 Estimand

- 「何を」を Research question の設定と Estimand の定義に分割
  - Research question = 研究者が頑張って決める
    - \* 経済学理論的動機、政策的動機、などを活用
  - Estimand = 母集団上で定義される推定対象: ここまでは  $E[Y|d', X] - E[Y|d, X]$ 
    - \* 無限大の事例数の元で、どのように Research question に答えられるか?

\* 因果推論、などを活用

### 1.3 Estimand



### 1.4 まとめ

- 推定にまつわるモデル化 (変数選択を含む)、Estimand の定義にまつわるモデル化は、分けて議論する必要がある (Manski 1995; Lundberg, Johnson, and Stewart 2021)

非推奨!!!

- “Y に影響を与えそうな変数” を全て  $X$  として用いる
  - 事例数が無限あったとしても、望ましくない帰結をもたらす

## 2 格差の識別

- 多くの実証研究で用いられている支配的な枠組みを、(川田は) 知らない
  - 因果推論とは、質的に異なる難しさ (規範命題) がある
  - 格差は因果的定義に馴染みにくいという立場 (Rose 2023; Jackson and VanderWeele 2018) に従う

## 2.1 規範命題

- 格差 (差別、不平等) := 本来的に許容すべきではない差
  - 「である」(事実命題) ではなく、「べき」(規範命題)
- 因果効果か否かは、事実命題であり、例えば理想的な実験を用いて区別可能
- 客観的事実から、規範を導くことには、慎重であるべき (できない?)
  - [ヒュームの法則](#)

## 2.2 $Y, D =$ 格差を定義

- $Y, D =$  研究関心となる格差/差別/不平等を定義する変数
- $Y =$  大学進学、 $D =$  出身地
  - 地域間教育格差
- $Y =$  [おいりの消費](#)、 $D =$  出身地
  - 地域間格差?/文化的多様性?

## 2.3 $X =$ 格差を定義

- $X =$  研究関心となる格差/差別/不平等を定義する変数

## 2.4 例: 取り締まり (Rose 2023)

- 研究課題に応じて、格差の識別は異なる
- 例: スピード違反の取り締まりに Race は影響しているのか?
  - $X =$  超過速度: 同じ速度で捕まった人について、量刑が同じであれば、“現場”において差別はない
- 例: Race 間で格差はあるのか?
  - 「犯罪を犯さざるを得ない状況」の偏りも格差に含めるのであれば、超過速度はコントロールすべきではない

## 2.5 まとめ

### 格差研究における $X$ の選択

- 価値判断へのコミット
  - 母集団が直接観察可能 (無限大の事例数をもつデータがある) 場合に、社会的に望ましくない差はどのように定義するか
  - ある  $X$  内で差がなければ、研究課題となっている格差/差別/不平等はないと**定義する**

## 2.6 まとめ

- 格差の定義: 議論の余地は大きい
  - 議論のためにも、定義を明示する必要がある
- 対話のための実践的な枠組みは存在するのか?
  - 規範理論 (厚生経済学、政治/法哲学) の応用?

## 2.7 注意: 適切な変数選択

- Double selection: **推定** に用いる  $Z(\subset X)$  を、統計的基準で選ぶ
- Research question の変更:  $X$  が推定できない (観察できないも含む) ので、研究課題を変更する
  - 格差の定義に必要な  $X$  も変化する

## 2.8 注意: 不適切な変数選択

- 識別に用いる  $X$  を、統計的基準で選ぶことは困難
- 例: Research question = 警察官の差別
- 「Double selection を用いて  $X$  を選べる」のであれば、以下を受け入れる必要がある
  - 事例数が少ないので、 $X$  から超過速度を除外する
    - \* 規範的判断を「マイノリティに対して、“下駄”をはかさない警察官は差別している」に変更する???

## 3 因果効果の識別

- 課題:  $E[Y|d, X] - E[Y|d', X]$  は因果効果と**解釈**できるのか?

- $D$  の実際の決まり方と理想の実験との乖離、に依存する

### 3.1 動機

- STEM - 非 STEM での平均収入差は容易に推定できるが、、、
  - 両親の教育経験や生まれ年など背景属性にも違いがある
  - 平均賃金差 = 学歴の効果 + 背景属性の違い
- 学歴によらず”稼げる人” はそもそも大学に行かない
  - 背景属性の違い  $< 0$
- そもそも数理的思考が得意な人が STEM に行く
  - 背景属性の違い  $> 0$

### 3.2 課題

- $X$  の選択 および データ選択をどうするか？
  - 観察できる変数の中から、どの  $X$  の選択すべきか
  - どのような変数をデータから観察したいのか
- 解決: 「因果 (“真の値”)」を定義する
  - 新たな概念的枠組みを導入
  - 母集団と同じ戦法

### 3.3 因果についての枠組み

- 複数の有力なアプローチ (潜在結果や構造的因果モデル (DAG で表現される)) がある
  - [Chap 2, 4-7 in CausalML](#)
    - \* 他にも “経済理論?” (Heckman and Pinto 2024)、頑健性 (Peters, Bühlmann, and Meinshausen 2016) などもある
  - 不毛 (?) な”学派” 論争も散見されるが、(反実仮想分析など) 共通点が多い (Goldberg 2019)。

### 3.4 実験アプローチ

- 理想的な (思考) 実験 (Target Trial) を実施するか/結果を近づける

- 本講義で採用
- 他の枠組みの参照点としても機能

### 3.5 理想的な実験: Controlled Experiment

- 背景属性が全く同じ被験者を 2 名以上用意して、一部にのみ介入  $D = 1$  を行う
  - 例: 食塩を入れると、水の沸騰温度は上がるのか?
- 労働経済学においては実現不可能
  - “全く同じ” 人間とは?
  - 例: 一卵性双子

### 3.6 代替案: Randomized Controlled Trial

- 無限の被験者が存在し、被験者間相互作用がない環境において、各被験者の  $D$  をランダムに決める
  - データから観察可能/不可能な背景属性と”無関係”に  $D$  は決定
  - $Y$  の分布に差があれば、 $D$  の違いによってもたらされたと解釈する”しかない”状況に持ち込める
- 労働経済学の研究関心と親和的

### 3.7 実行可能な実験

- 有限の被験者に対して、 $D$  をランダムに割り振る
  - 背景属性は”偶然”偏るが、その偏りは信頼区間で評価できる

### 3.8 Example: Resume Experiment

- 履歴書の名前は、採用確率に影響を与えるのか?
  - Race が”伝わる”ことの因果効果を推定
    - \* 労働市場における差別の影響へ含意

### 3.9 Example: Resume Experiment

- Bertrand and Mullainathan (2004)
- 求人”偽の”履歴書を送り、返信があるかどうかを測定

- $D$  = 履歴書の内容: 特に Caucasian/African-American 系の名前かどうか
- $Y$  = Callback があるかどうか
- $X$  = 企業の属性、他の履歴書属性
- AER package にデータが収録

### 3.10 Example

```
lm(Y ~ D,
    Data) |>
marginaleffects::avg_comparisons(
  vcov = "HC3", # Robust standard error
  variables = "D")
```

Term	Contrast	Estimate	Std. Error	z	Pr(> z )	S	2.5 %	97.5 %
D	mean(1) - mean(0)	-0.032	0.00779	-4.11	<0.001	14.6	-0.0473	-0.0168

Columns: term, contrast, estimate, std.error, statistic, p.value, s.value, conf.low, conf.high, pred  
Type: response

### 3.11 自然実験

- Target Trial を実際に実施することは、労働経済学においては難しい
  - 社会の中で”勝手に起きた”実験的状况 (自然実験) を活用する方法が発展
    - \* Target Trial との間にギャップがあるので、補正を試みる
- 社会の中で”勝手に起きた”実験的状况
  - Estimand = 理想の実験

### 3.12 $X$ の選択への含意

- Target Trial をどうすれば、模倣できるのか?
  - Target Trial において、 $D$  間で差異が生じない変数は  $X$  に含める
  - 差異が生じる変数は含めない
    - \* 含めると Target Trial の模倣ができなくなる



- ・ Bad Controll

### 3.13 例: STEM の効果

- $D$  = STEM 系学部出身かどうか /  $Y$  = 30 歳時点での所得 /  $X$  = 出身地
- Bad Control = STEM 系大学院に進学したかどうか
  - STEM 系学部への進学が Random に決まっていたとしても、因果効果が推定できなくなる (後述)

### 3.14 例: Fukai, Kawata, and Nakabayashi (2023)

- 研究課題: 「“科学的処置” の負の結果」が科学への信頼に与える影響
  - $D$  = コロナ腕の発生/非発生、 $Y$  = ワクチン/自然科学/承認プロセスへの信頼
- 識別の仮定:  $X = \{ \text{性別、年齢、病歴、ワクチンの種類} + \alpha \}$  が同じ回答者内では、コロナ腕はランダムに発生

### 3.15 警句

介入なくして因果なし

- 「理想的な実験を” 想像できない” のであれば、因果効果は定義できない」
  - 議論が続いている主張
- (思考) 実験における介入が  $D$  を明確に定義する
  - 例: STEM 進学とは?  $D = 1 \iff$  (大阪大学理学部、関大工学部..へ進学)
  - 例: Social category (“性別”、“Race”) とは?

### 3.16 まとめ

格差研究における  $X$  の選択

- 規範的判断へのコミット

因果推論における  $X$  の選択

- 理想的な実験結果へのコミット
- Double selection を  $X$  について使用すると、コミットが事例数に応じて変化することを受け入れる必要がある

- 非推奨

## 4 推定への含意

- RCT は、推定方法についても大きな恩恵をもたらす
  - $D$  は  $X$  と統計的に独立している
    - \*  $\Pr[D = d|X] = \Pr[D = d]$  が保証される
  - 推定モデルの定式化への依存度が低い (Aronow et al. 2021)

### 4.1 RCT データへのコントロール

- RCT データに対して、 $X$  をコントロールすべきか?
  - コントロールすることが望ましい (推定誤差を減らし、信頼区間を狭くすることができる)
- Lin (2013) :  $D$  との交差項も入れれば、定式化が誤っていたとしても、漸近的に推定結果が悪化することはない
  - 小規模サンプルでないのであれば、以下を推定すべき

$$Y \sim D + X + D \times X$$

- 有限標本における補正 (@ Chang, Middleton, and Aronow forthcoming)

### 4.2 例: ResumeName

```
lm(
  Y ~ D*(.), # Y,D 以外との交差項を導入
  Data # Y,D,X のみを格納
) |>
marginaleffects::avg_comparisons(
  variables = "D",
  vcov = "HC3"
)
```

Term	Contrast	Estimate	Std. Error	z	Pr(> z )	S	2.5 %	97.5 %
	D mean(1) - mean(0)	-0.0313	0.00775	-4.03	<0.001	14.2	-0.0464	-0.0161

Columns: term, contrast, estimate, std.error, statistic, p.value, s.value, conf.low, conf.high, pred

Type: response

### 4.3 定式化依存への誤解

- しばしば、Simple な Linear model で十分だと考えられてきた。
- 理想的な自然実験であれば十分
- 多くの応用では、 $D$  と  $X$  は関連しているので、定式化依存の問題を校了する必要がある
  - 一つの手段は Double Selection

### Reference

- Aronow, PM, James M Robins, Theo Saarinen, Fredrik Sävje, and Jasjeet Sekhon. 2021. “Non-parametric Identification Is Not Enough, but Randomized Controlled Trials Are.” *arXiv Preprint arXiv:2108.11342*.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination.” *American Economic Review* 94 (4): 991–1013.
- Chang, Haoge, Joel Middleton, and PM Aronow. forthcoming. “Exact Bias Correction for Linear Adjustment of Randomized Controlled Trials.” *Econometrica*, forthcoming.
- Fukai, Taiyo, Keisuke Kawata, and Masaki Nakabayashi. 2023. “Updated Beliefs and Shaken Confidence: Evidence from Vaccine Hesitancy Caused by Experiencing ‘COVID Arm’.” *BMC Infectious Diseases* 23 (1): 612.
- Goldberg, Lisa R. 2019. “The Book of Why: The New Science of Cause and Effect: By Judea Pearl and Dana Mackenzie, Basic Books (2018). ISBN: 978-0465097609.” Taylor & Francis.
- Heckman, James, and Rodrigo Pinto. 2024. “Econometric Causality: The Central Role of Thought Experiments.” *Journal of Econometrics*, 105719.
- Jackson, John W, and Tyler J VanderWeele. 2018. “Decomposition Analysis to Identify Intervention Targets for Reducing Disparities.” *Epidemiology (Cambridge, Mass.)* 29 (6): 825.
- Lin, Winston. 2013. “Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman’s Critique.” *Annals of Applied Statistics*, 295–318.
- Lundberg, Ian, Rebecca Johnson, and Brandon M Stewart. 2021. “What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory.” *American Sociological Review* 86 (3): 532–65.
- Manski, Charles F. 1995. *Identification Problems in the Social Sciences*. Harvard University Press.
- Peters, Jonas, Peter Bühlmann, and Nicolai Meinshausen. 2016. “Causal Inference by Using Invariant Prediction: Identification and Confidence Intervals.” *Journal of the Royal Statistical Society Series B: Statistical Methodology* 78 (5): 947–1012.
- Rose, Evan K. 2023. “A Constructivist Perspective on Empirical Discrimination Research.” *Journal of Economic Literature* 61 (3): 906–23.