

集計データの活用

労働経済学 2

川田恵介

Table of contents

1	集計データの活用	1
1.1	集計データの欠点	2
1.2	個表データの欠点	2
1.3	No interference	2
1.4	No interference の問題点	2
1.5	練習問題: RCT under constraints	3
1.6	No interference への対応	3
1.7	まとめ	3
2	Example. Shift-Share Analysis	3
2.1	関心	4
2.2	例	4
2.3	例	4
2.4	例	4
2.5	Shift-Share 分析のラフな動機	5
2.6	例: 人口と就業者数	5
2.7	分解 (Dunn Jr 1960)	5
2.8	分解 (Dunn Jr 1960)	5
2.9	分解 (Dunn Jr 1960)	6
2.10	例	6
2.11	まとめ	6
2.12	Referene	7

1 集計データの活用

- 今日の多くの労働経済学における実証研究は、比較的”小さい”単位のデータ(家計、企業等々)を用いている

- 個表を集計したデータ (地域、時系列) を活用した研究も以前重要

1.1 集計データの欠点

- (集計者が) 集計した情報しか活用できない
 - 例: 地域の平均賃金はわかるが、分散はわからない
- 分析単位が大きいため
 - 事例数が確保しにくい
 - 研究者による実験は、実質不可能

1.2 個表データの欠点

- Privacy の問題等が大きくなりがちで、限られた研究者しかデータにアクセスできない状況も多い
 - 再現可能性が低くなる
 - データの入手に時間がかかり、迅速な分析ができない
 - 個表データの入手が実質できない状況がある
- 分析単位の小さいため、Interference の問題が厄介

1.3 No interference

- 教科書的な因果推論においては、**No interference (SUTVA)** が仮定されている
 - ある経済主体への介入は、別の経済主体に影響を与えない
- 例: 労働経済学 2 を受講した効果は、各受講者の中で完結する
 - 「講義内容を友達に紹介する、などを行わない」と仮定する

1.4 No interference の問題点

- 「経済学理論」は、Interference の重要性をむしろ強調してきた
 - 「経済学 = 希少な資源の分配問題」: 誰かが資源を活用すると、他の主体は活用できない
 - 「ゲーム理論 = 戦略的環境に置かれている戦略的思考を行う主体の決定問題」: 他人の決定が、自身に影響をあた得る効果
 - 市場均衡/Peer effect/外部生等々

- 主体が小さいほどより深刻なる (かもしれない)

1.5 練習問題: RCT under constraints

- 日本全体からランダムに選ばれた高校生への介入が、「大阪大学経済学部に入学者の確率」に与える因果効果を推定したい
 - 非常に効果的な介入をランダムに行った結果、介入を受けた学生と受けなかった学生の間で非常に大きな差が見られた
 - 日本全体で同様の介入を行った場合、因果効果は消失する可能性が高い
 - * なぜか?

1.6 No interference への対応

- 近年、改めて問題視される (Wolf 2023; Sävje 2024)
- 少なくとも「関心のある母集団からランダムに選ばれた大量の被験者に対して、ランダムに介入を行う実験」を行えたとしても、多くの介入の効果はわからないことに注意
 - 介入のランダム化、外的妥当性、Positivity については大きな問題はない
 - Intereference が発生し、標準的な推定結果と真の因果効果はずれてしまう可能性がある

1.7 まとめ

- “Wolf (2023) shows how to combine credible micro and macro evidence to analyse policy-relevant counterfactual in macroeconomic models. We view this as an attractive area of future research.” (Arkhangelsky and Imbens 2024)
- 発展中の文脈であり、川田も勉強中
- とりあえず集計データ + 経済学のアイディアによる、すぐに活用できるアプローチを散文的に紹介
- メインは、サーチ理論の活用 (Şahin et al. 2014; Michailat and Saez 2021; Kawata and Sato 2021)

2 Example. Shift-Share Analysis

- 古典的だが、今でも使われている手法
 - [wiki](#)
 - 特に政策レポートなどで重宝されている印象

- 後に紹介する Shift-Share IV についてにおいても重要

2.1 関心

- 全体の成長率 (G) とあるグループの成長率 (g) の比較: $g - G$
- 例

$$G = \frac{t\text{時点での人口} - \text{初期時点での人口}}{\text{初期時点での人口}}$$

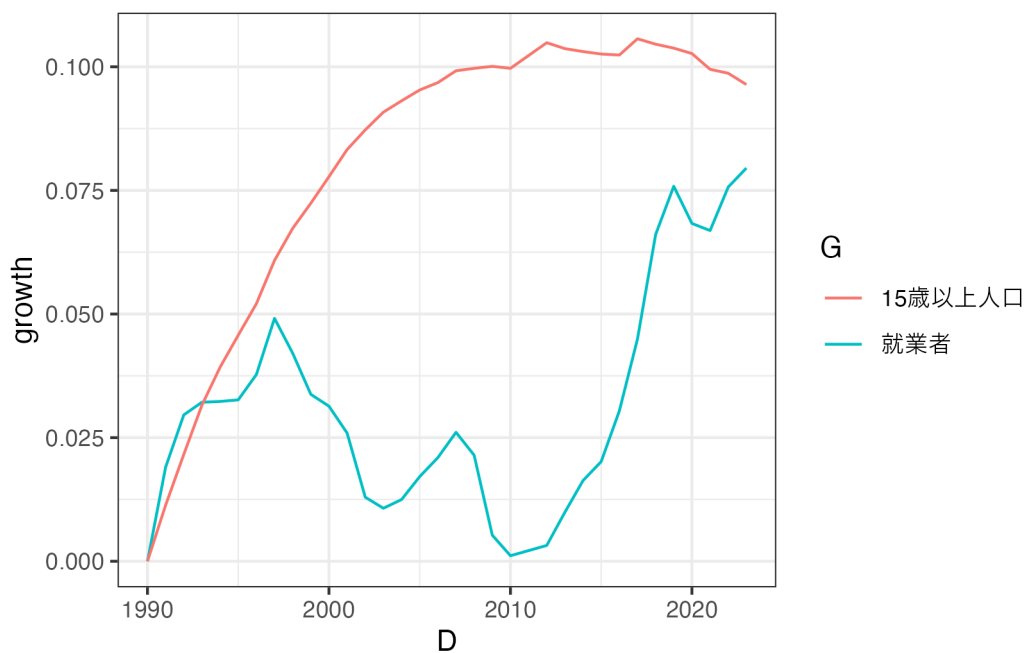
$$g = \frac{t\text{時点での就業者} - \text{初期時点での就業者}}{\text{初期時点での就業者}}$$

2.2 例

- 定義より

(15歳以上)人口 = 就業者 + 非労働力 + 失業者

2.3 例



2.4 例

- 2010 年以降、就業者数の顕著な上昇が見られる
 - 米国とは対照的 (Binder and Bound 2019; Abraham and Kearney 2020)

- なぜか?

2.5 Shift-Share 分析のラフな動機

- “外生的な変化” がもたらす影響を知りたい
 - 外生とは?
- 外生: “The term external variable appears to originate in Wright [1934]. Frisch [1933] wrote about autonomous relationships. Given the numerous conflicting definitions of “exogenous” and “endogenous” variables documented by Leamer [1985], the “internal-external” distinction is a useful one for focusing on what is determined in a model and what is specified outside of it.” (Heckman 2000)

2.6 例: 人口と就業者数

- 日本において、人口構造の変化は、かなり安定的に予測できる (と言われている)
 - 経済発展とともに少子高齢化に向かうのは、かなり頑強な fact
- 少なくとも短期的な経済・社会ショックの”外側”で決まっていると考えられる = “外生”
- 就業者数の変化は、人口構造の変化からどの程度”予測”できるか?

2.7 分解 (Dunn Jr 1960)

$$\begin{aligned}
 & \bullet \\
 & \text{就業者数の成長率} = \\
 & \left[\underbrace{\frac{\text{年齢}X\text{における}t\text{期の就業者} - X\text{における初期の就業者}}{\text{年齢}X\text{における初期の就業者}}}_{Shift} \right. \\
 & \quad \times \left. \underbrace{\frac{X\text{における初期の就業者}}{\text{初期の就業者}}}_{Share} \right] \text{の総和}
 \end{aligned}$$

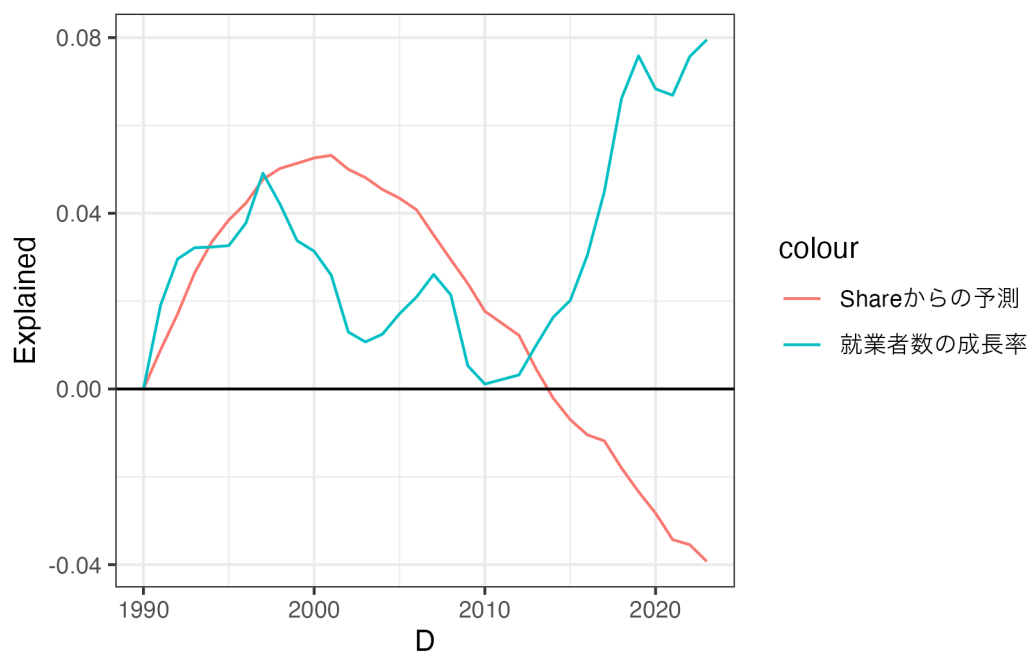
2.8 分解 (Dunn Jr 1960)

$$\begin{aligned}
 & \bullet \text{ 一般に} \\
 & \text{就業者の成長率} \\
 & = \text{Shareから予測される成長率} \\
 & \quad + \underbrace{\text{予測されない成長率}}_{\text{年齢}X\text{における就業者率の変化}}
 \end{aligned}$$

2.9 分解 (Dunn Jr 1960)

- $$\begin{aligned} & \text{Shareから予測される成長率} \\ &= \frac{\text{年齢}X\text{における}t\text{期の人口} - X\text{における初期の人口}}{\text{年齢}X\text{における初期の人口}} \\ & \quad \times \text{就業者における年齢}X\text{のシェア} \end{aligned}$$
- 「もし年齢ごとの就業率が変化しなければ」、人口構造の変化に応じて、就業者数はどのように変化したのか?

2.10 例



2.11 まとめ

- 人口構造の変化からの予測と実際の成長率の乖離は、2010年代で顕著
 - 年齢構造の変化のみを見ると、就業率は大きく低下しているはず
 - 実際は大きく上昇している
- 「もし年齢ごとの就業率が変化しなければ」を前提とした、会計的な分析であることに注意

2.12 Referene

- Abraham, Katharine G, and Melissa S Kearney. 2020. “Explaining the Decline in the US Employment-to-Population Ratio: A Review of the Evidence.” *Journal of Economic Literature* 58 (3): 585–643.
- Arkhangelsky, Dmitry, and Guido Imbens. 2024. “Causal Models for Longitudinal and Panel Data: A Survey.” *The Econometrics Journal* 27 (3): C1–61.
- Binder, Ariel J, and John Bound. 2019. “The Declining Labor Market Prospects of Less-Educated Men.” *Journal of Economic Perspectives* 33 (2): 163–90.
- Dunn Jr, Edgar S. 1960. “A Statistical and Analytical Technique for Regional Analysis.” *Papers in Regional Science* 6 (1): 97–112.
- Heckman, James J. 2000. “Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective.” *The Quarterly Journal of Economics* 115 (1): 45–97.
- Kawata, Keisuke, and Yasuhiro Sato. 2021. “A First Aid Kit to Assess Welfare Impacts.” *Economics Letters* 205: 109928.
- Michaillat, Pascal, and Emmanuel Saez. 2021. “Beveridgean Unemployment Gap.” *Journal of Public Economics Plus* 2: 100009.
- Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L Violante. 2014. “Mismatch Unemployment.” *American Economic Review* 104 (11): 3529–64.
- Sävje, Fredrik. 2024. “Causal Inference with Misspecified Exposure Mappings: Separating Definitions and Assumptions.” *Biometrika* 111 (1): 1–15.
- Wolf, Christian K. 2023. “The Missing Intercept: A Demand Equivalence Approach.” *American Economic Review* 113 (8): 2232–69.