

パネルデータ

川田恵介

1 Panel Data の活用

1.1 パネルデータ

- 同一事例を追跡調査したデータ: 事例 i について、複数時点 t の $\{Y_{it}, D_{it}, X_{it}\}$ が観察可能
 - ▶ 動学効果の推定や新しい識別方法などが活用可能に!!!
- 要約/推定方法が、急速進歩 (Baker et al., 2022; De Chaisemartin & d'Haultfoeuille, 2022; Roth et al., 2023)
 - ▶ 実証結果に深刻な影響 (Baker et al., 2022)

1.2 例

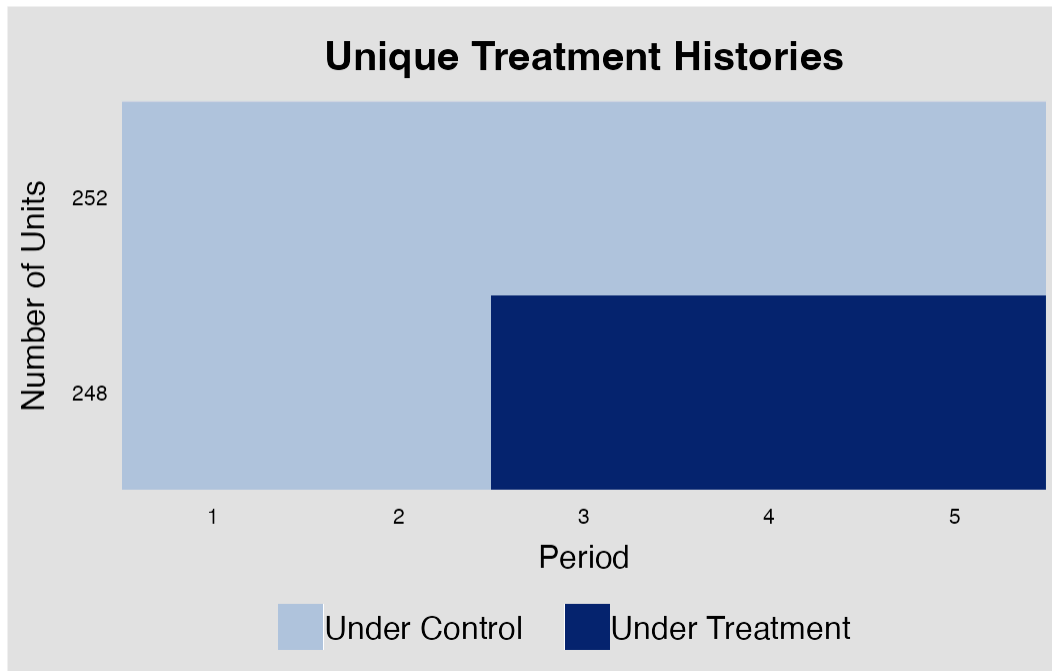
```
# A tibble: 9 × 5
  ID Period TreatGroup    Y    D
<int> <int>      <int> <dbl> <dbl>
1     1     1         1 -1.71     0
2     1     2         1 -8.94     0
3     1     3         1 -10.5     1
4     2     1         1 -6.26     0
5     2     2         1  4.24     0
6     2     3         1 -1.94     1
7     3     1         1 11.6     0
8     3     2         1  5.59     0
9     3     3         1  5.88     1
```

1.3 推定対象: 動学効果

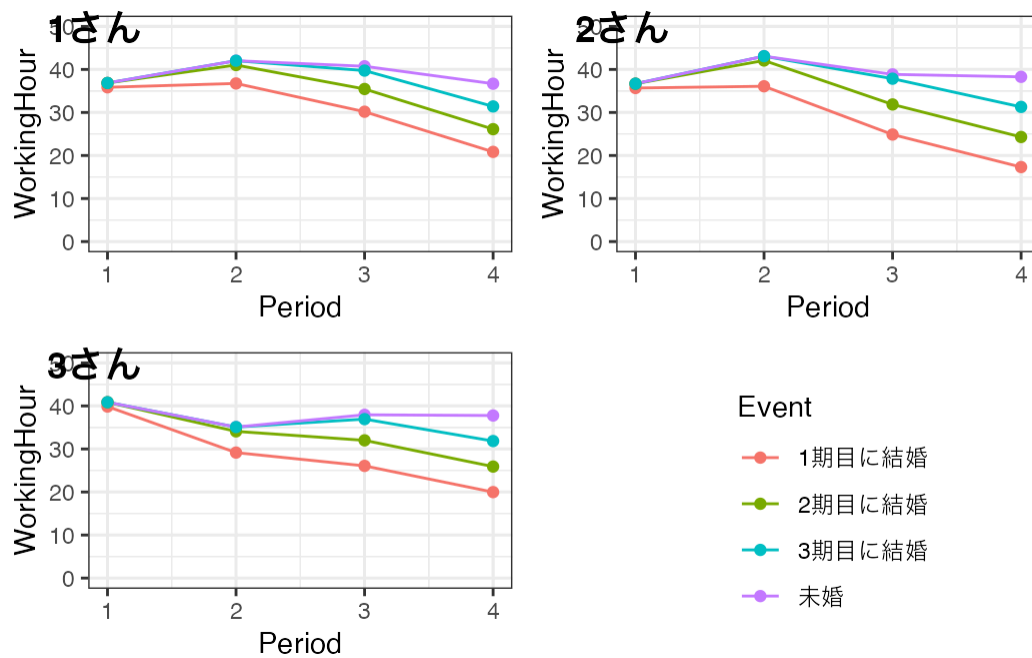
- 多くの介入で、効果は介入からの期間に応じて変化しうる
- 介入が発生した場合、しなかった場合の差
 - ▶ いつ介入したか (j) + 介入からどのくらい経ったか (l) に応じて、大量に定義できる
- 例: 2024 年における労働時間についての因果効果
 - ▶ 2023 年に結婚 VS 未婚のまま ($j = 2023/l = 1$)
 - ▶ 2018 年に結婚 VS 未婚のまま ($j = 2018/l = 1$)

- ▶ 2018 年に結婚後 5 年 VS 未婚のまま ($j = 2018/l = 5$)

1.4 panel view



1.5 例



2 因果効果の要約

- 1 期間の平均因果効果と比べて、因果効果の要約方法がより大量に存在する
 - ▶ どのような要約方法を採用するか、明確に定める必要がある

2.1 平均動学効果

- 介入から l 期後の平均効果を集計
- $E[\text{個別効果} \mid l]$

2.2 コホート別平均動学効果

- 介入した時点 j ごとに、介入発生から l 期後の平均効果を集計
- $E[\text{個別効果} \mid l, j]$
 - ▶ 例: 結婚の効果は、時代によって異なる
- 平均動学効果の推定においても、重要な役割を果たす

3 識別

3.1 識別の仮定

- これまでと同様に、Positivity/No interference
- 緩和した Conditional independence: Parallel trends
- No anticipation

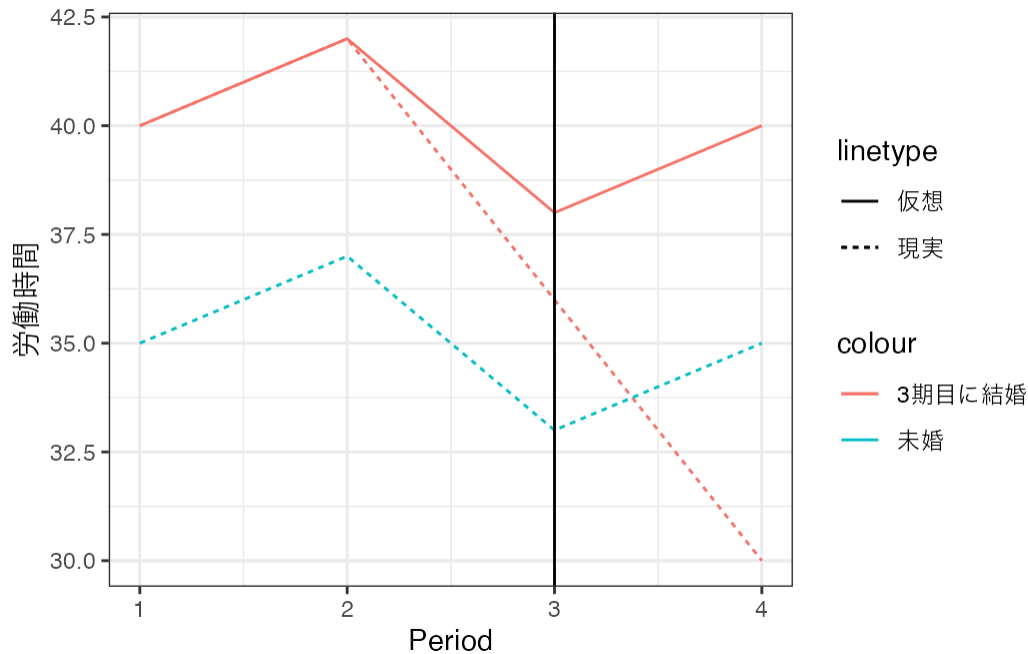
3.2 No anticipation

- 将来の介入は、過去に影響を与えない
 - ▶ 例: 2024 年に未婚から既婚に変化しても、2022 年の労働時間は変化しない

3.3 Parallel trends

- 介入が生じなければ、「介入が生じなかったグループ」と平均的に同じ変化をする。

3.4 例



3.5 Parallel trends の利点

- RCT/自然実験により保証できる
 - ▶ それ以外では？
- 実践的な利点は、介入発生前に複数期間あれば、部分的なテスト (Before-Before/PreTrend test)が可能
 - ▶ Event の発生状況に関わらず、 Y は同じように推移している
 - ▶ Roth (2022)

3.6 補論: コントロール変数の導入

- 条件付き Parallel trends: 同じ X 内で Parallel trends が成り立っている
 - ▶ 固定効果モデルでは、時間を通じて変化する変数しかできない
 - ▶ 推定方法を工夫する必要がある

4 推定: 2×2 Case

4.1 2×2 Case

- 2 期間 \times 2 グループデータ
 - ▶ Control Group: 介入が生じない

- ▶ Treatment Group: 2 期目に介入発生
- 確立された推定方法が存在
 - ▶ パネル分析法の良い出発点

4.2 Two Ways Fixed Effect Model

- 固定効果モデルを推定:

$$Y_{i,t} = \tau D_{i,t} + \underbrace{f_i}_{\text{個人固定効果}} + \underbrace{f_t}_{\text{時点固定効果}} + u_{i,t}$$

- ▶ OLS と同様の方法 (データへの適合度を最大化する) で推定可能
- ▶ 追加の仮定を導入することで、推定精度を高める方法(変量効果モデルの最尤法による推定)もある
- 識別の仮定のもとで、Average Treatment Effect on Treated について”信頼できる”信頼区間形成が可能

4.3 R Example

```
fixest::feols(
  Y ~ D + factor(Period),
  DataShort,
  panel.id = "ID"
)
```

```
OLS estimation, Dep. Var.: Y
Observations: 1,000
Standard-errors: Clustered (ID)
      Estimate Std. Error  t value Pr(>|t|)
(Intercept)  -0.655464    0.460420  -1.42362  0.15518
D              0.911598    0.896771   1.01653  0.30987
factor(Period)2 1.187798    0.799808   1.48510  0.13815
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 10.1  Adj. R2: 0.005762
```

5 Reference

Bibliography

Baker, A. C., Larcker, D. F., & Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates?. *Journal of Financial Economics*, 144(2), 370–395.

- De Chaisemartin, C., & d'Haultfoeuille, X. (2022). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey.
- Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3), 305–322.
- Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*.