

# 留学効果の検証

中尾 走      Ran NAKAO

広島大学大学院教育学研究科・日本学術振興会特別研究員（DC2）

ran-nakao@hiroshima-u.ac.jp

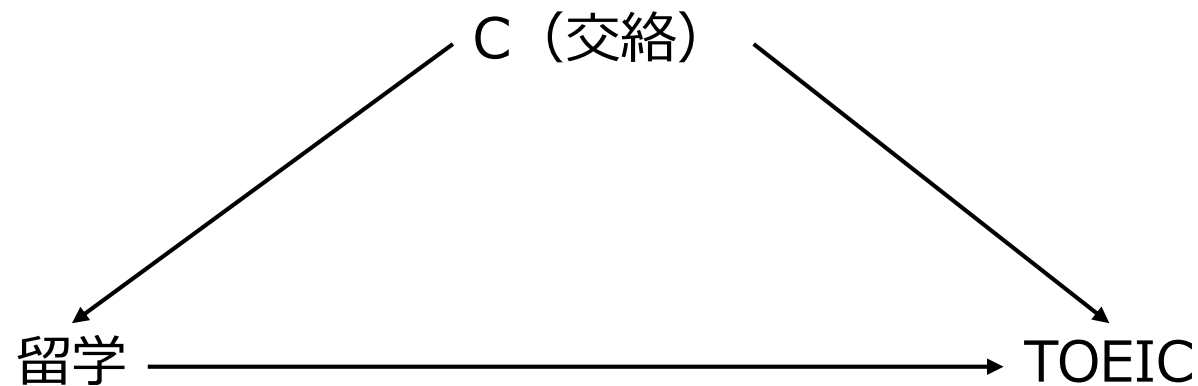
樊 怡舟      Yizhou FAN

広島大学大学院教育学研究科

fanyizhou921106@yahoo.co.jp

# 研究の目的

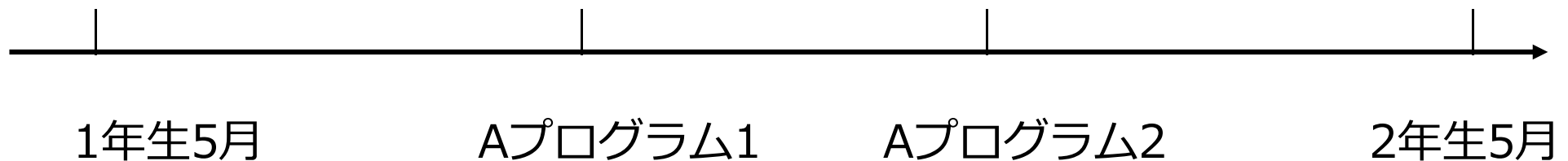
- 短期留学プログラムの効果（以下，留学効果）を明らかにすること
- 留学効果を明らかにする上で対処すべきこと
- 交絡変数（留学効果？それ以外の効果？）
- 分析手法に付随するいくつかの仮定
  - ① 平行トレンド仮定
  - ② SUTVAの仮定（第一仮定）



- 本研究では，このような問題に対処した上で，留学効果を明らかにしたい

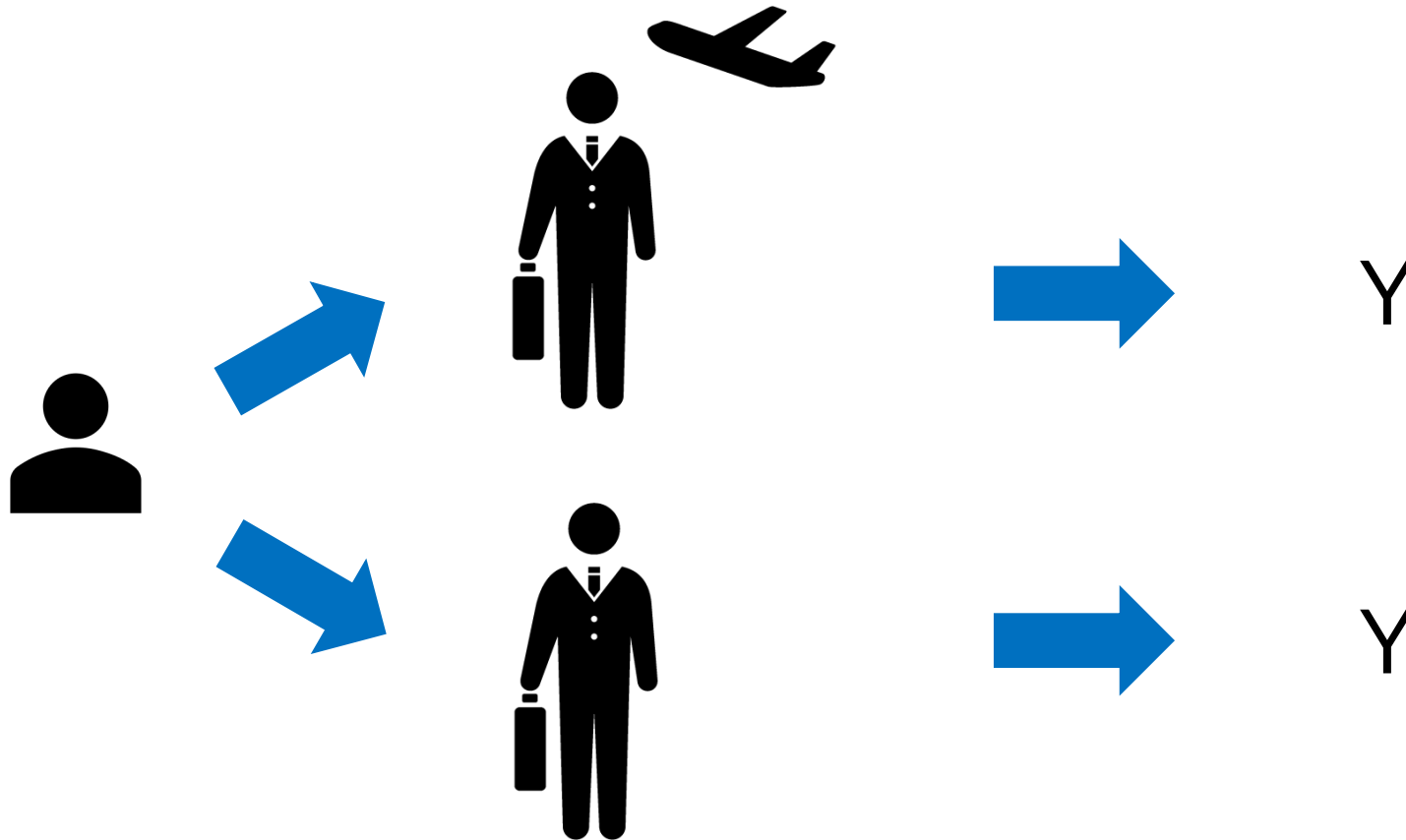
# データと変数

- データ
  - X大学で実施されている短期留学プログラム（以下, Aプログラム）を用いる
  - 2019年度のみ（理由は後述）
  - 1年生のみ（1年生以外でもAプログラムのアドバンス版に参加可能だが, ①質的に異なると考えた, ②時点を合わせるのが大変）
- 
- 変数
  - Aプログラムの参加の有無：1 = 参加, 0 = 不参加
  - TOEICの点数：連続量（1年生5月と2年生5月の二時点）
  - Aプログラム1, 2は前期・後期を示す



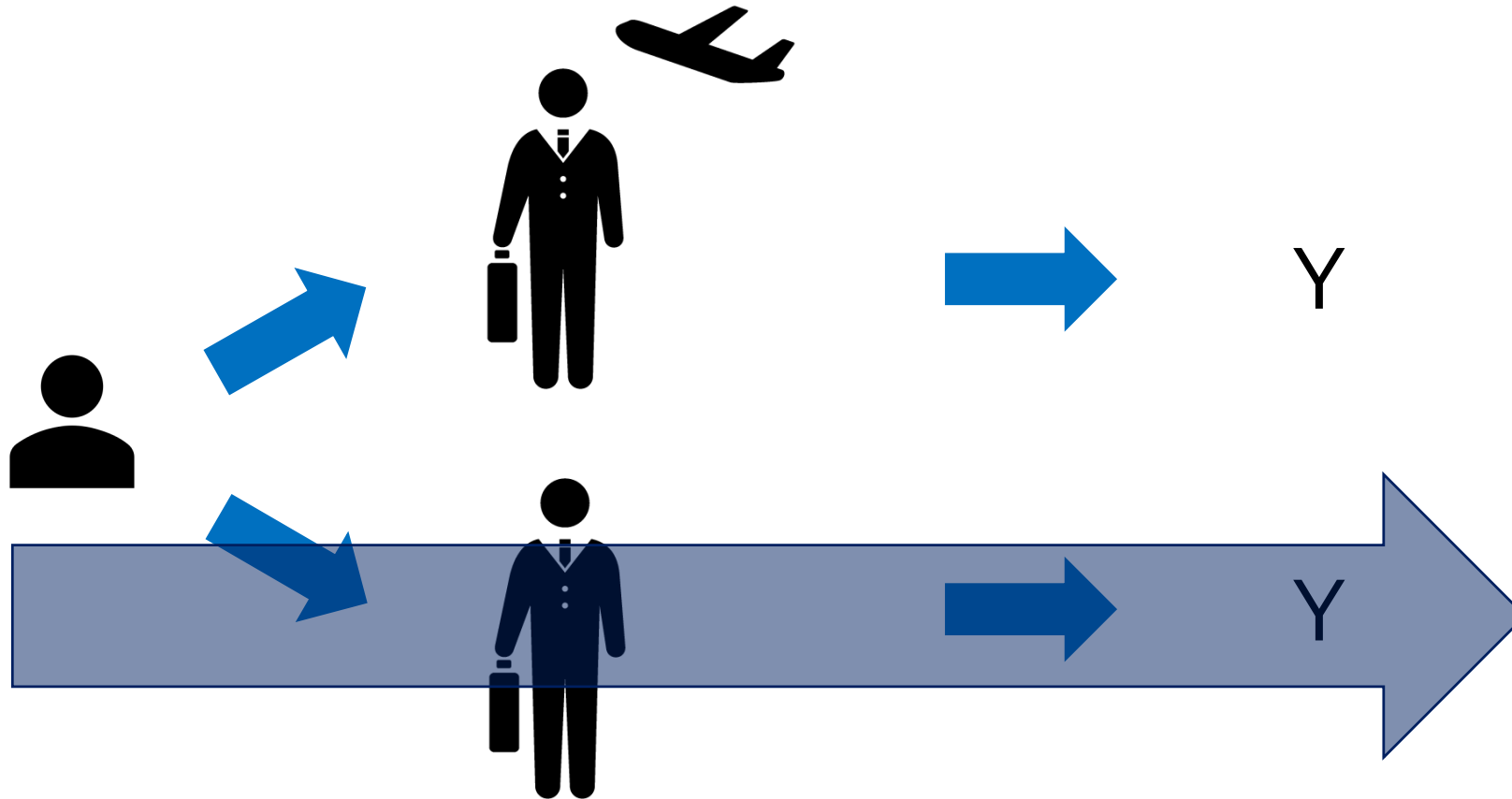
# 因果推論の根本問題

- 留学の効果



# 因果推論の根本問題

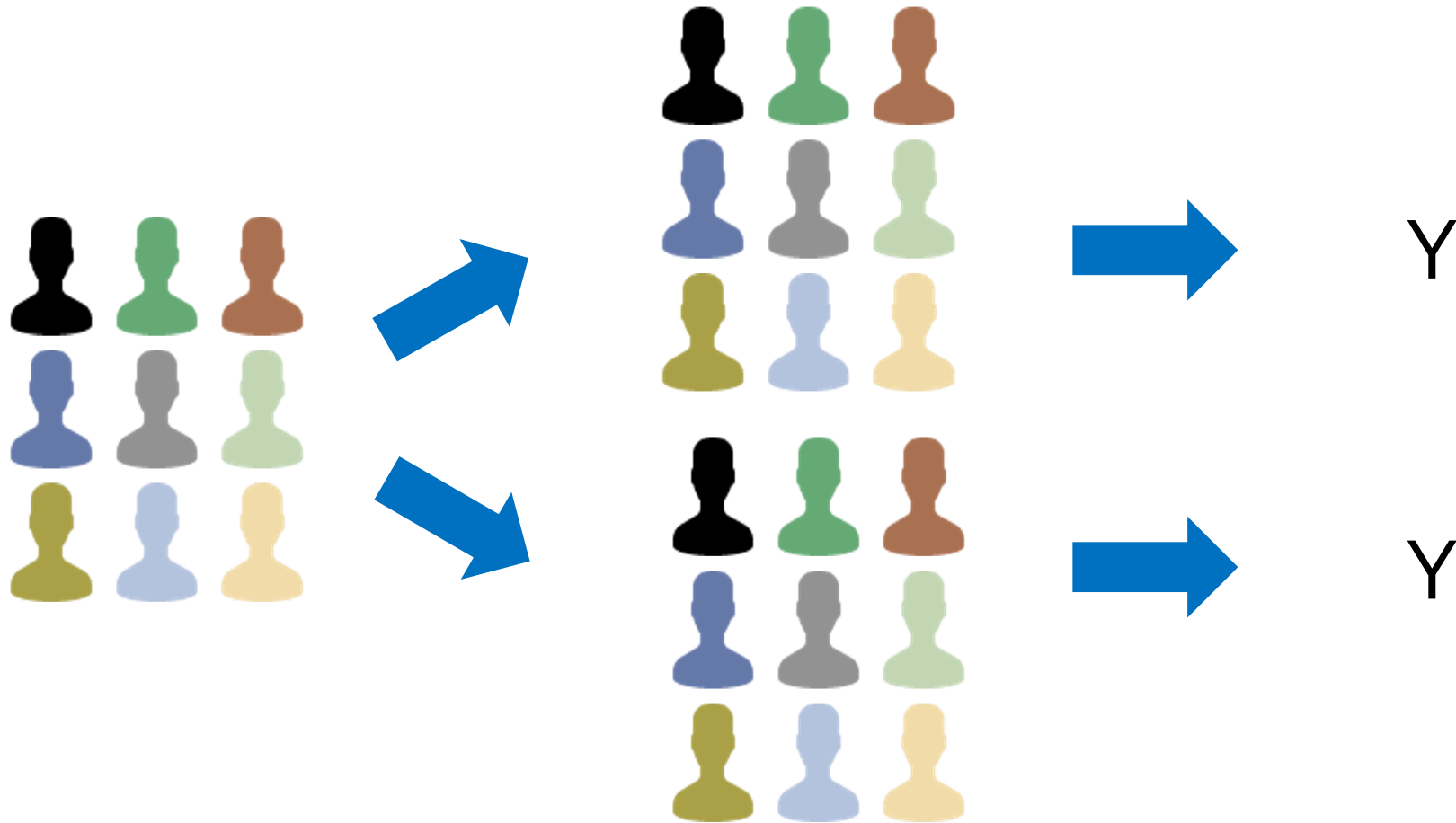
- 留学の効果



観測されない

# 因果推論の根本問題

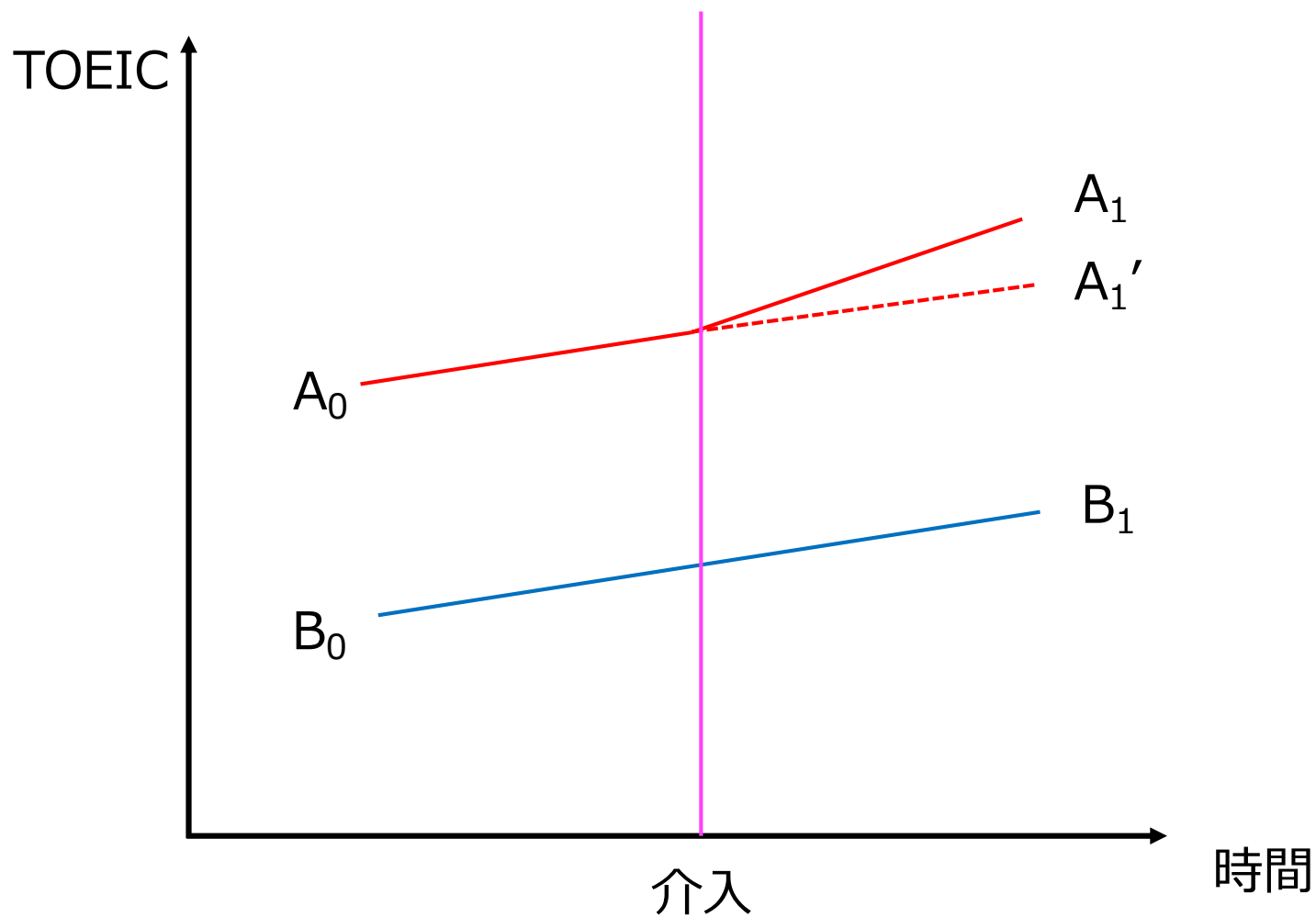
- 集団への因果効果なら同定できる



- 因果推論とは、基本的には推定できないものをデータから推定しようとしている

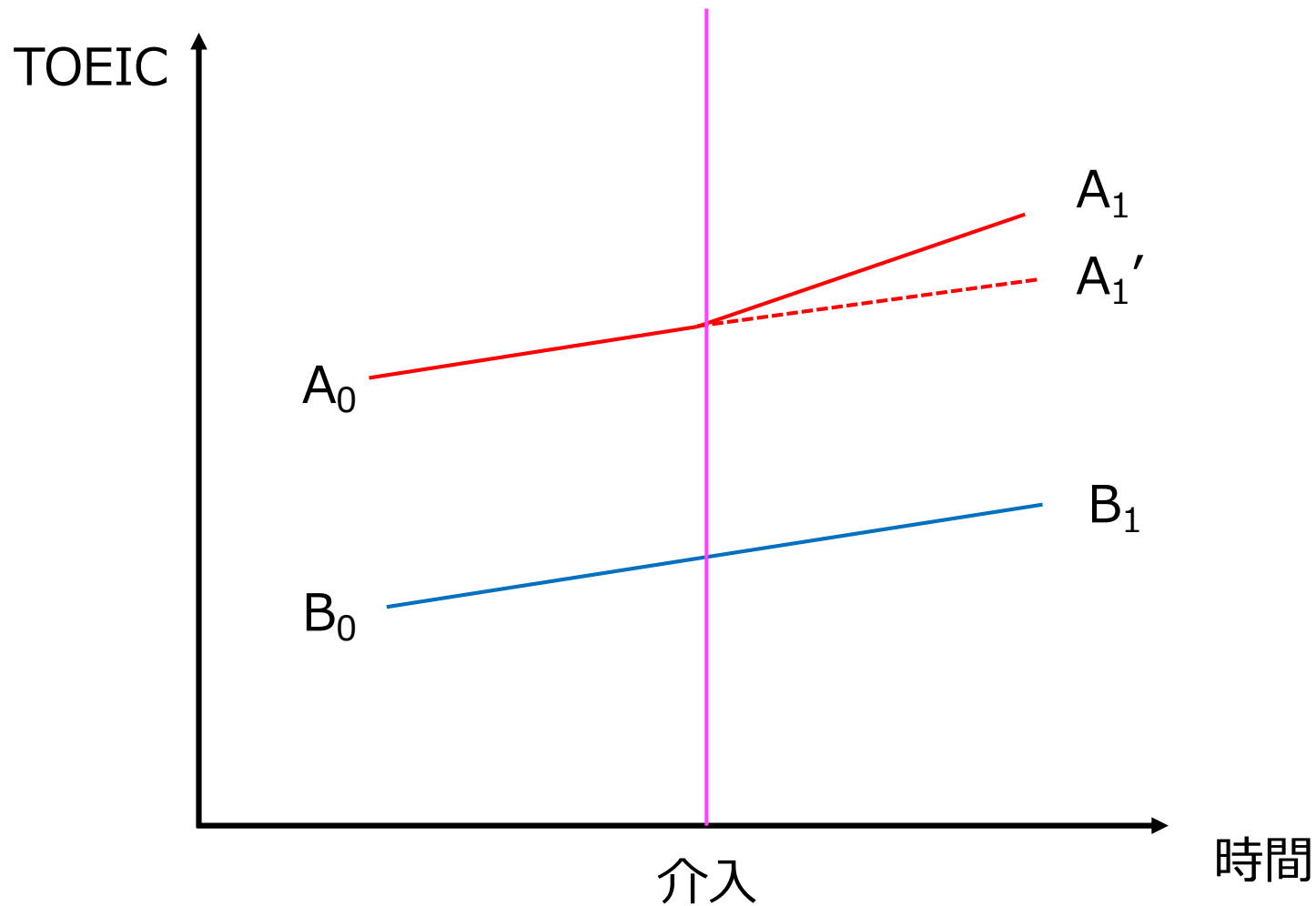
# 方法

- 差の差分分析 (Difference-in-difference : DID)
- 赤 : 留学参加, 青 : 不参加
- $A_1$ と $A_1'$ を比較すれば留学効果分かる



# 方法

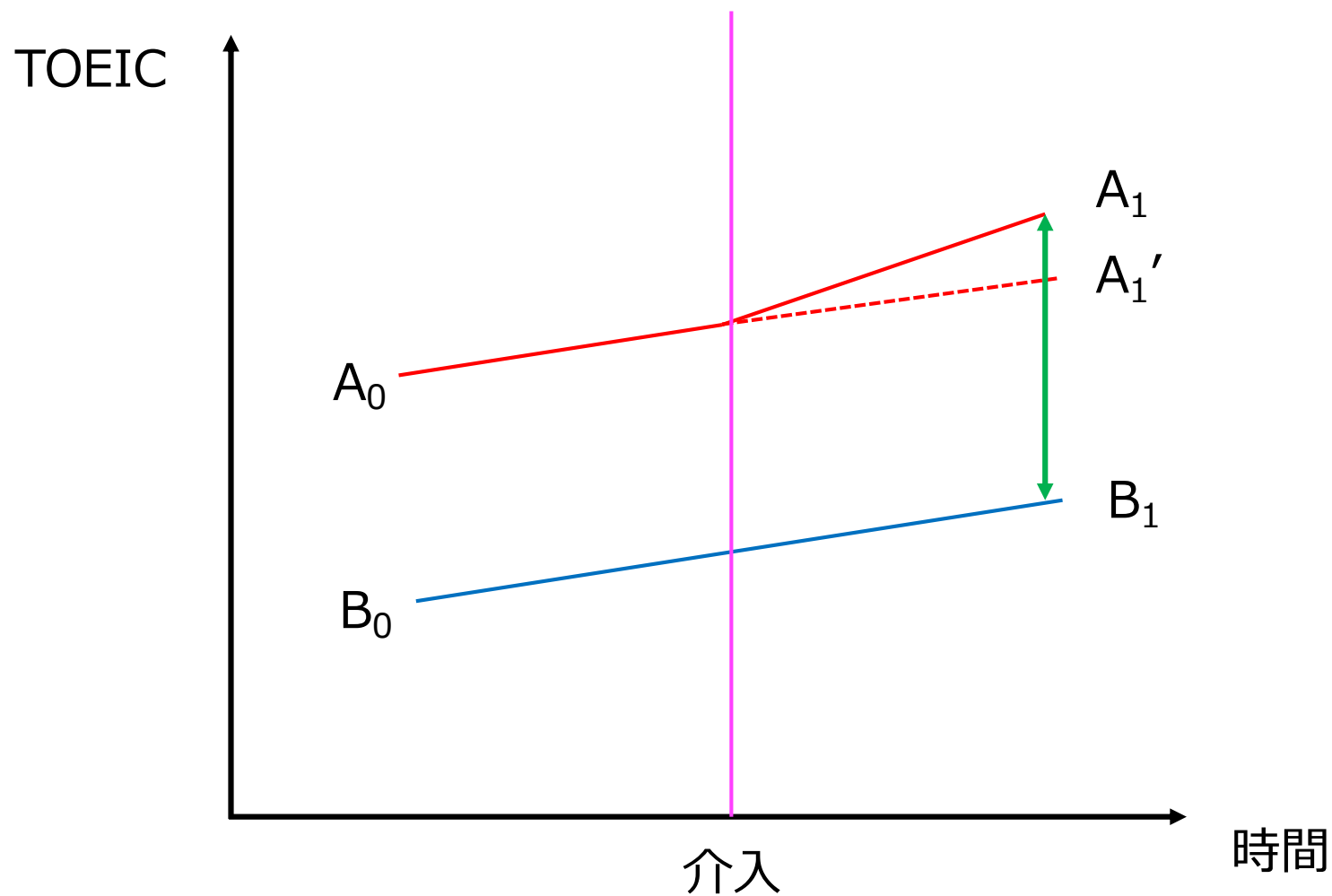
- しかし,  $A_1'$ が観察されない
- では対照群の $B_1$ と比較すれば良いか？それとも介入前の $A_0$ と比較すれば良いか？





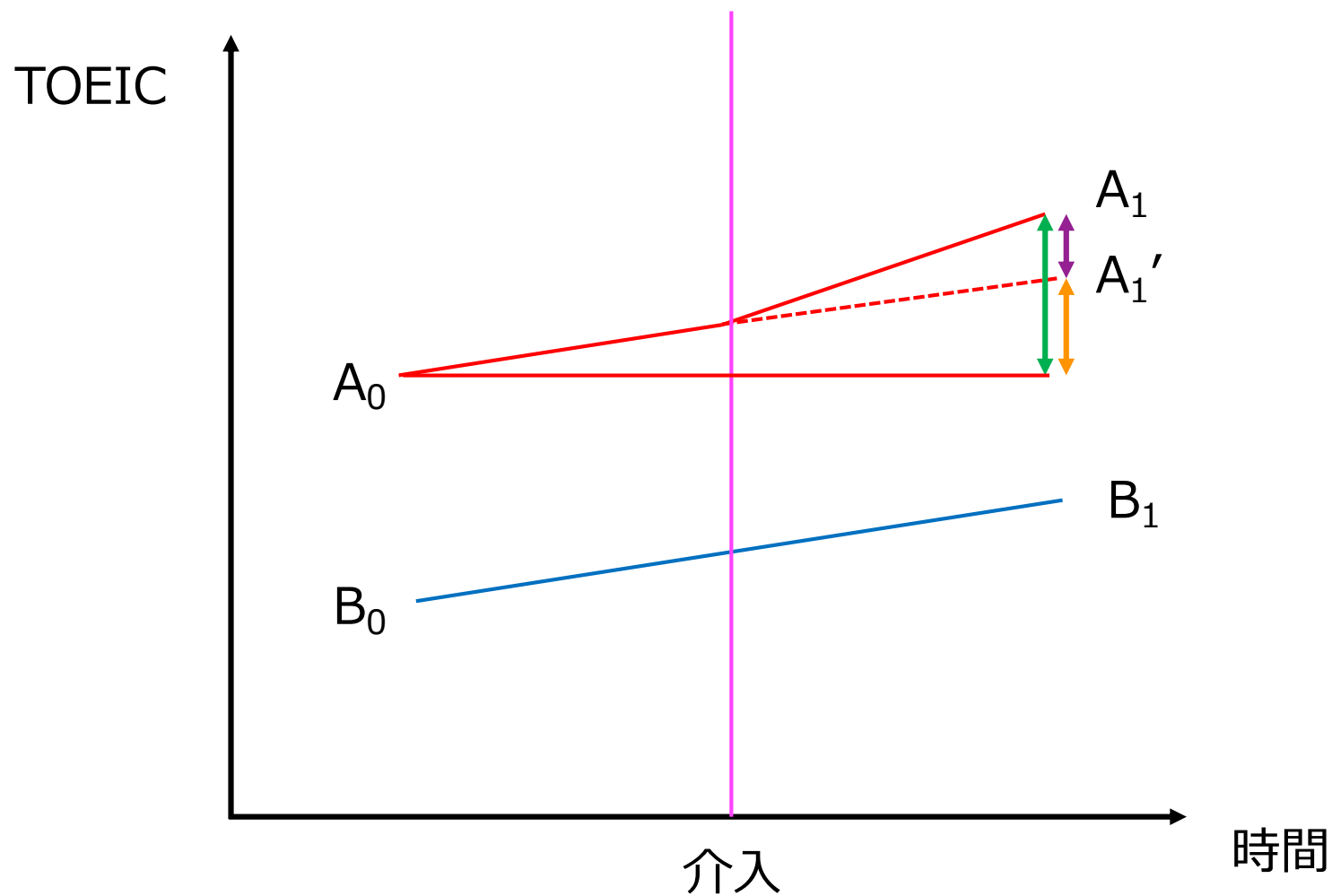
# 方法

- 対照群の $B_1$ との比較では, 効果がかかなり大きく出てしまう  
(緑)



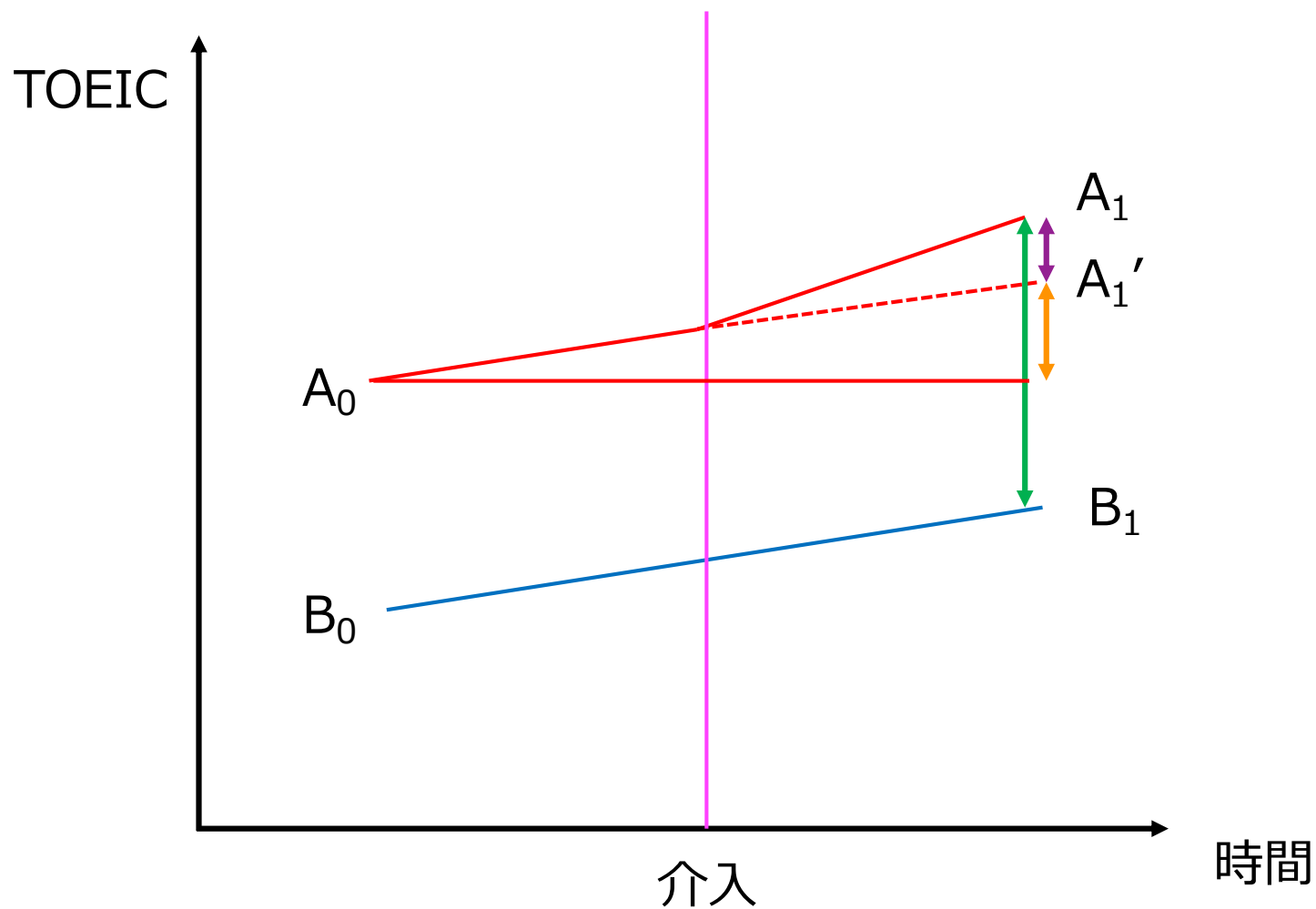
# 方法

- 過去のA ( $A_0$ )との比較では, 時間トレンドと識別ができない
- 緑は, 時間トレンド (オレンジ) + 介入効果 (紫)



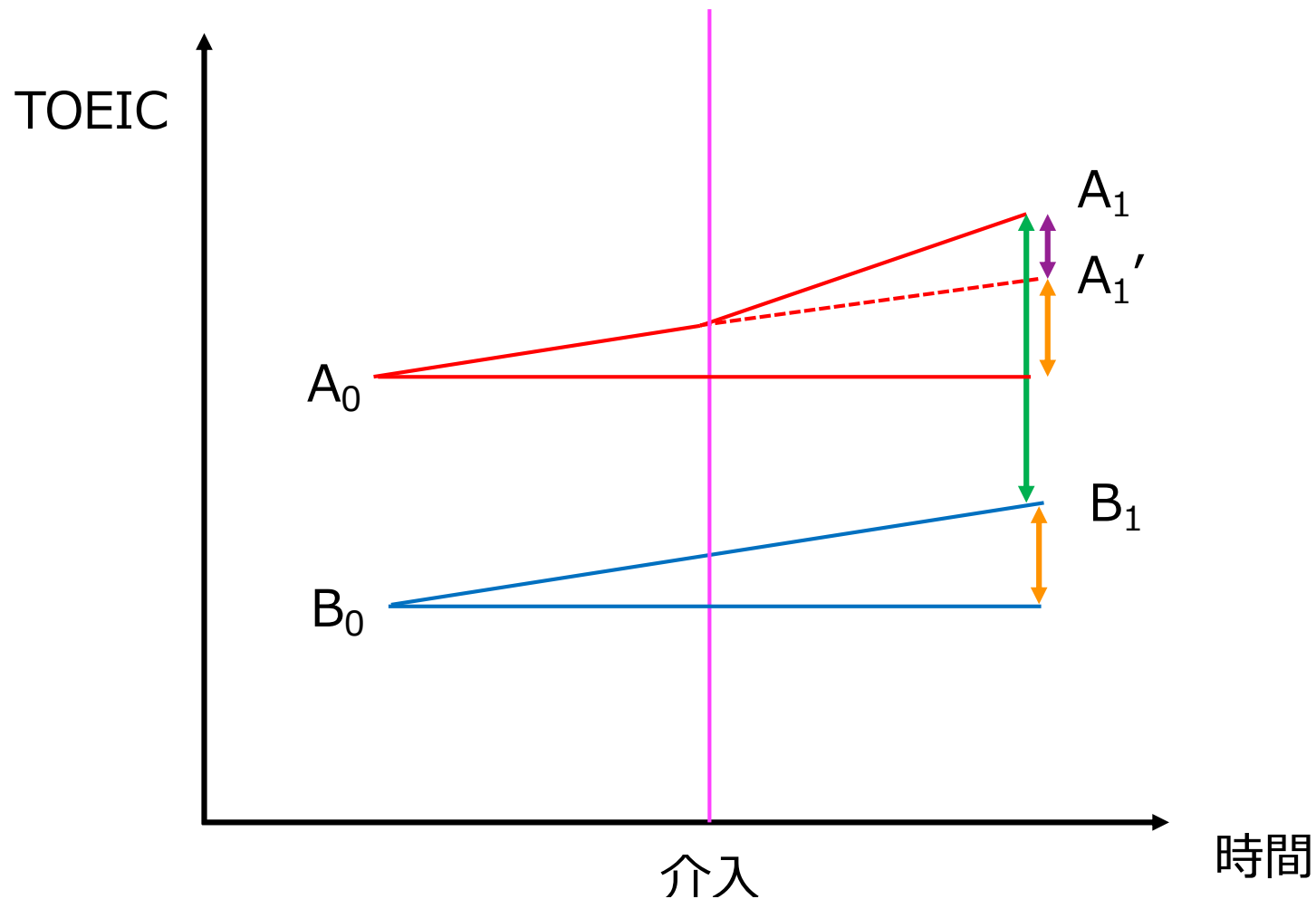
# 方法

- DIDの発想は、緑からオレンジを引けば、紫（介入効果）が推定できる
- そのために必要な仮定が平行トレンド仮定



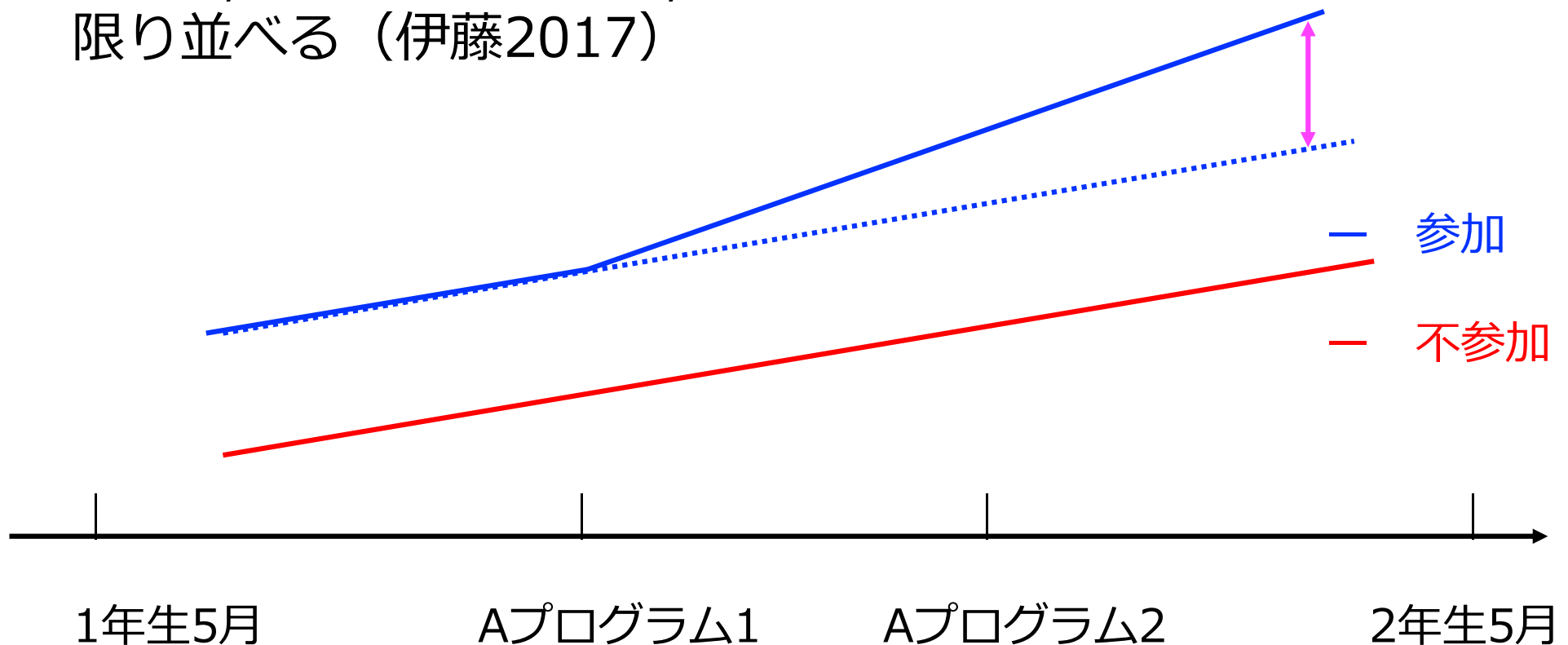
# 方法

- 処置群の反実仮想 ( $A'$ ) と対照群が介入前後で平行であること：平行トレンド仮定
- 平行であれば、同じ時間トレンド（オレンジ）を持つので、時間トレンド分を削除できる



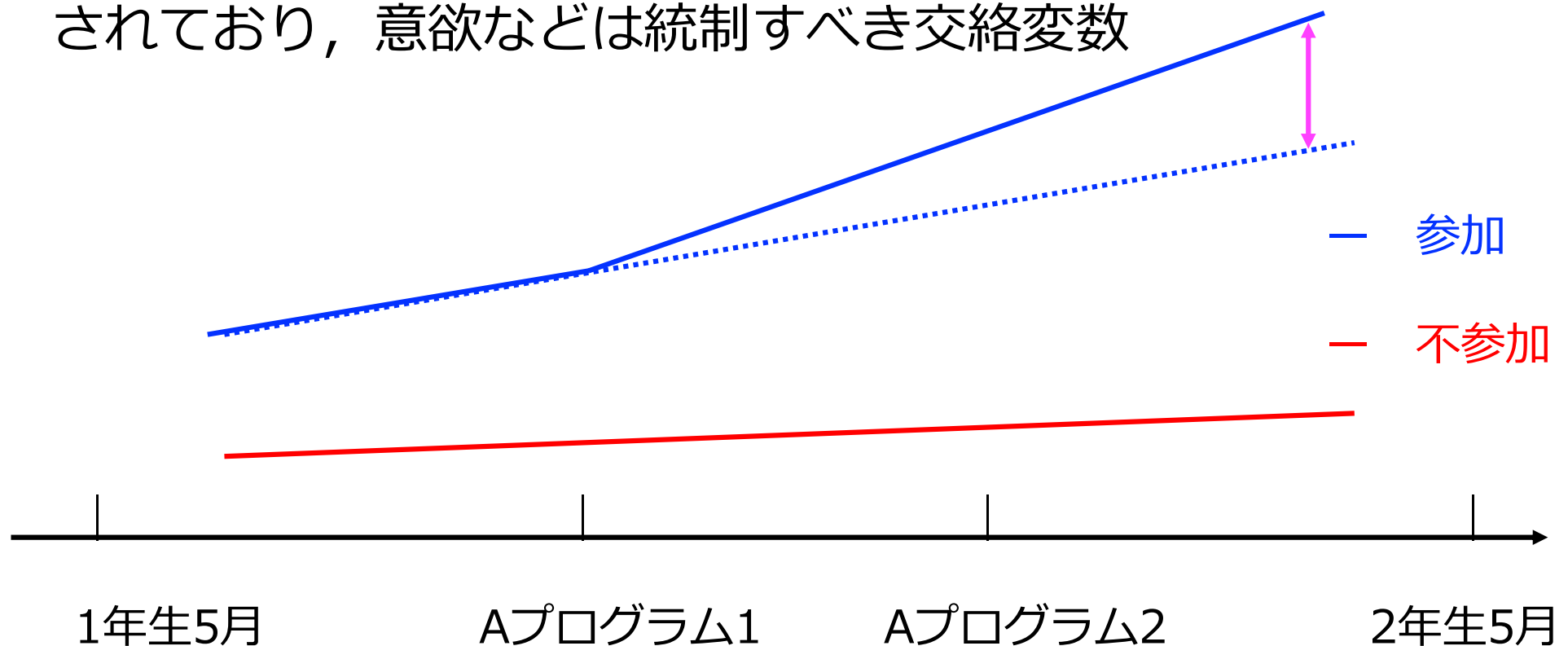
# 方法

- ただし，平行トレンド仮定はデータから証明が出来ないため，傍証として他人を納得させることでしか，仮定が満たされているという主張はできない
- 平行トレンド仮定が成り立ちそうな対照群を選択する（安井2020）や介入前のトレンドが平行であることを示す（伊藤2017；笹渕2019）など，仮定が成立している証拠をできる限り並べる（伊藤2017）



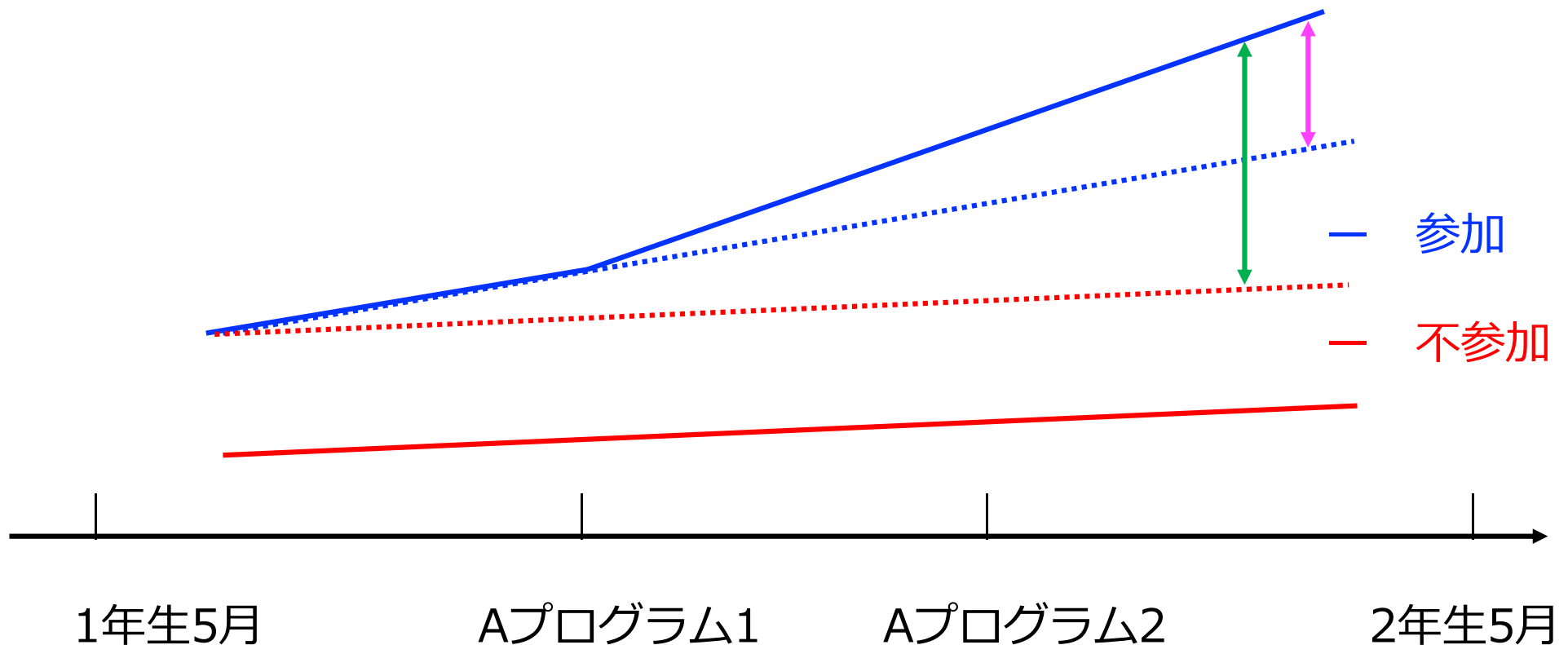
# 方法

- Aプログラムに参加する学生と参加していない学生の間で学習習慣，学習意欲によってTOEICの点数の伸びが異なるかもしれない
- 川田・西谷（2017）では，この要因については，統制されていない（そもそも統制が，非常に困難）
- 川田ほか（2019）で，参加決定による学習意欲の増加が考察されており，意欲などは統制すべき交絡変数



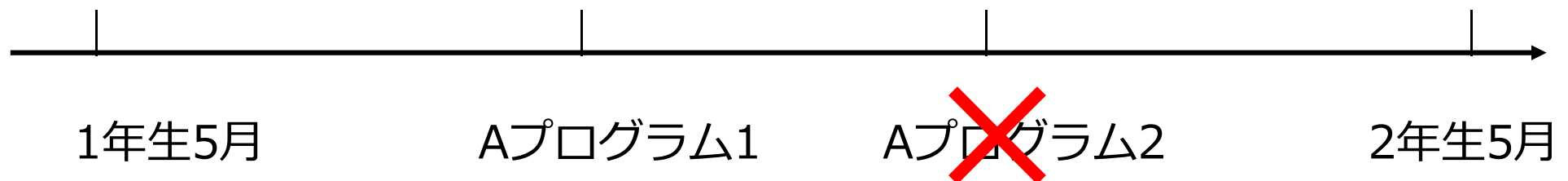
# 方法

- 例えば，不参加の人はそもそも学習意欲が低く，今回留学に参加した人よりも伸びが小さいことが予想される場合は，平行トレンド仮定が成り立たない
- その場合，留学効果が過大推定される（緑の矢印）
- そこで，平行トレンドが満たされそうな対照群を持ってこることが必要となる



# 方法

- 2019年度は、新型コロナウイルスの影響で後期のプログラムのみが中止
- つまり、Aプログラムに参加する予定だった人の中で参加、不参加が混在する自然実験の状況が発生した
- この二つの群は、学習習慣や学習意欲など観測困難な交絡変数や未観測要因の交絡変数が等しいことが確率的に期待される
- DIDは、他の介入によって平行トレンド仮定が満たされない（Craig et al. 2017）ことが指摘されるが、Aプログラム同様に他の留学プログラムも中止されている可能性が高く、他の介入によって平行トレンド仮定が崩れる可能性は小さいことが想定できる



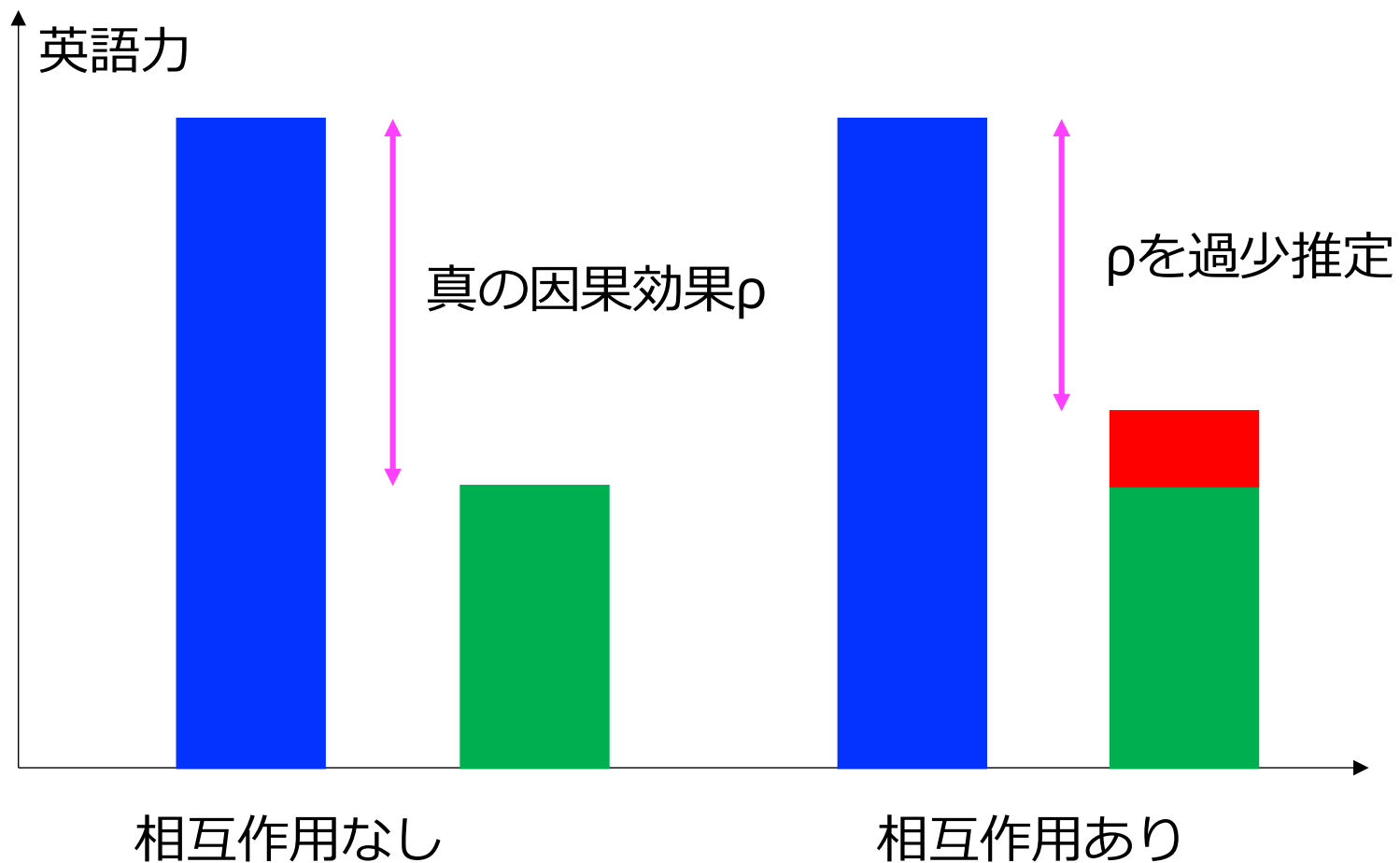


# 方法

- DIDを用いる際に必要な仮定 (Wing et al. 2018)
  - ① 平行トレンド仮定
  - ② SUTVAの仮定 (第一仮定)
- ②のSUTVAの第一仮定は満たされているだろうか？
- SUTVAの第一仮定：個体 $i$ の潜在的な結果 $\{Y_i(1), Y_i(0)\}$ は、他の個体の受ける処置に依存しない (no inference)
- 個人 $i$ が留学に行った場合、行かなかった $i$ は観察できない
- もし相互作用がなければ、 $i$ が留学に行ったときと、行かなかった場合の差分を取ればそれが因果効果となる
- けれども、もし仲の良い友達 $j$ が留学に行ったことに影響を受けて、 $i$ が留学に行かなくても英語力が上昇したとすればどうなるか・・・ (いわゆるピア効果)
- 今、反実仮定が想定できること世界に來たとする
- ピア効果がある場合、効果なしという結論が効果あり結論に変化することはあり得るため、検証すべき課題

# 方法

- 青：留学に行った時の英語力
- 緑：留学に行かなかった時の英語力
- 赤：ピア効果（仲の良い友達が行ったことに影響を受けた効果）



# 方法

- 実際に、学部ごとにSTARTプログラムの参加率は異なる（川田・西谷2017）
- 例えば、文学部であれば相互作用を受けやすいかもしれない
- さらに今回は履修科目のデータがある

	START非参加者	START参加者	総数
総合科学	124	3	127
文学	116	11	127
教育	470	27	497
法学	139	5	144
経済学	164	6	170
理学	214	2	216
医学	224	6	230
歯学	86	0	86
工学	475	4	479
生物生産	98	6	104
薬学	62	0	62
総数	2,172	70	2,242

表 2 : 学生の出身学部

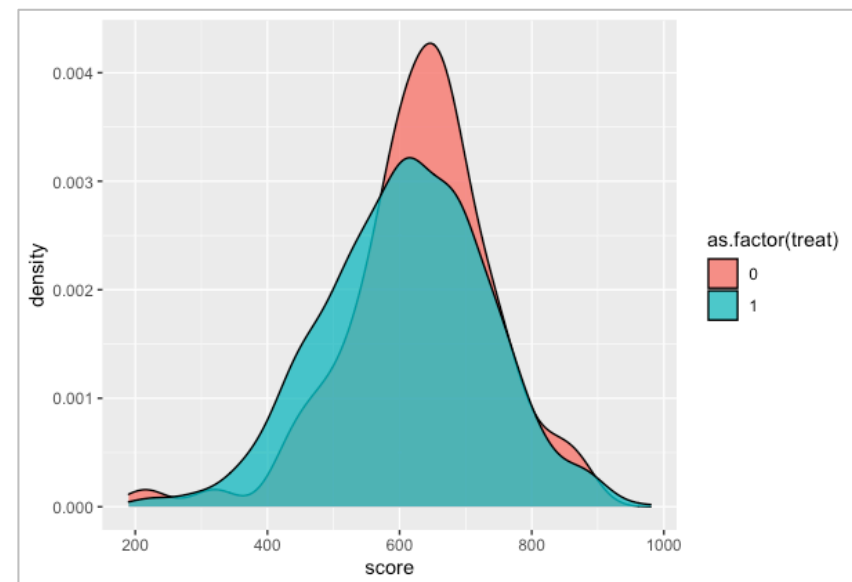
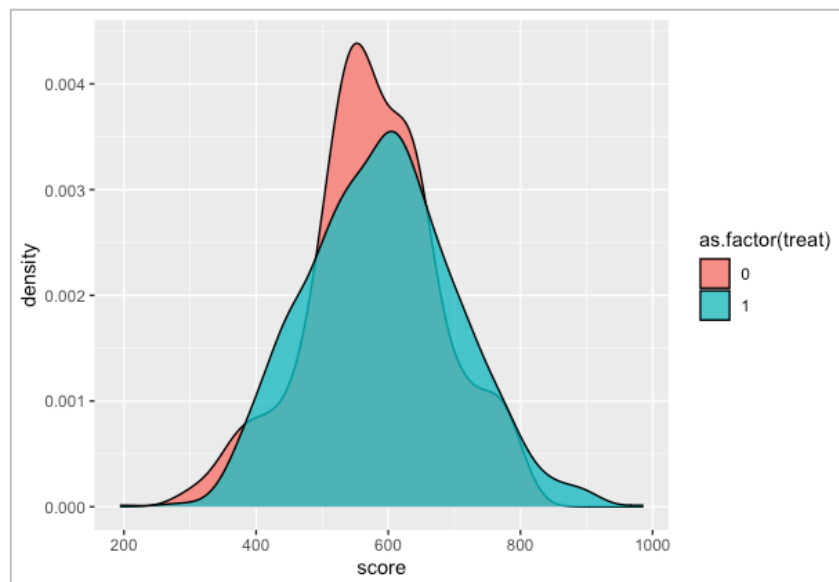
# 方法

- 基本的に、社会科学の観察データではSUTVAの第一仮定が満たされないことが多い (Imai 2019) が、SUTVAの第一仮定を満たさない可能性を考慮して、相互作用効果を統制した上で因果効果を推定する手法 (Clark 2017) を用いる
- 仲の良い友達：同じ学部内で履修科目の類似性を代理変数とする (0~100)
- 仲の良い友達であれば、相互作用による効果を与えている可能性があるかと仮定する
- その中で、最も仲の良い友達から影響を受けると仮定する
- また、学部内の相互作用はあるが、学部間を超えて広がらないことを仮定する
- つまり、学部内の履修科目の類似性のみを計算
- 類似性は、距離行列をバイナリーデータから計算 (全て一緒なら0, 全く違うなら100)

# 分析

- 記述的な分析
- 青：STARTプログラムに参加した学生（処置群：n=191）
- 赤：中止になった学生（対照群：n=107）

time	mean	sd	median	min	max
1	585.28	111.74	585	195	915
2	633.68	119.57	640	215	900



# 分析①

- DIDの分析結果
- 効果なし (null association)
- 自然実験を用いたDIDの結果によると留学効果なし (先行研究は, 意欲など交絡の効果?)

```
Call:
lm(formula = score ~ time * treat, data = did_2019)

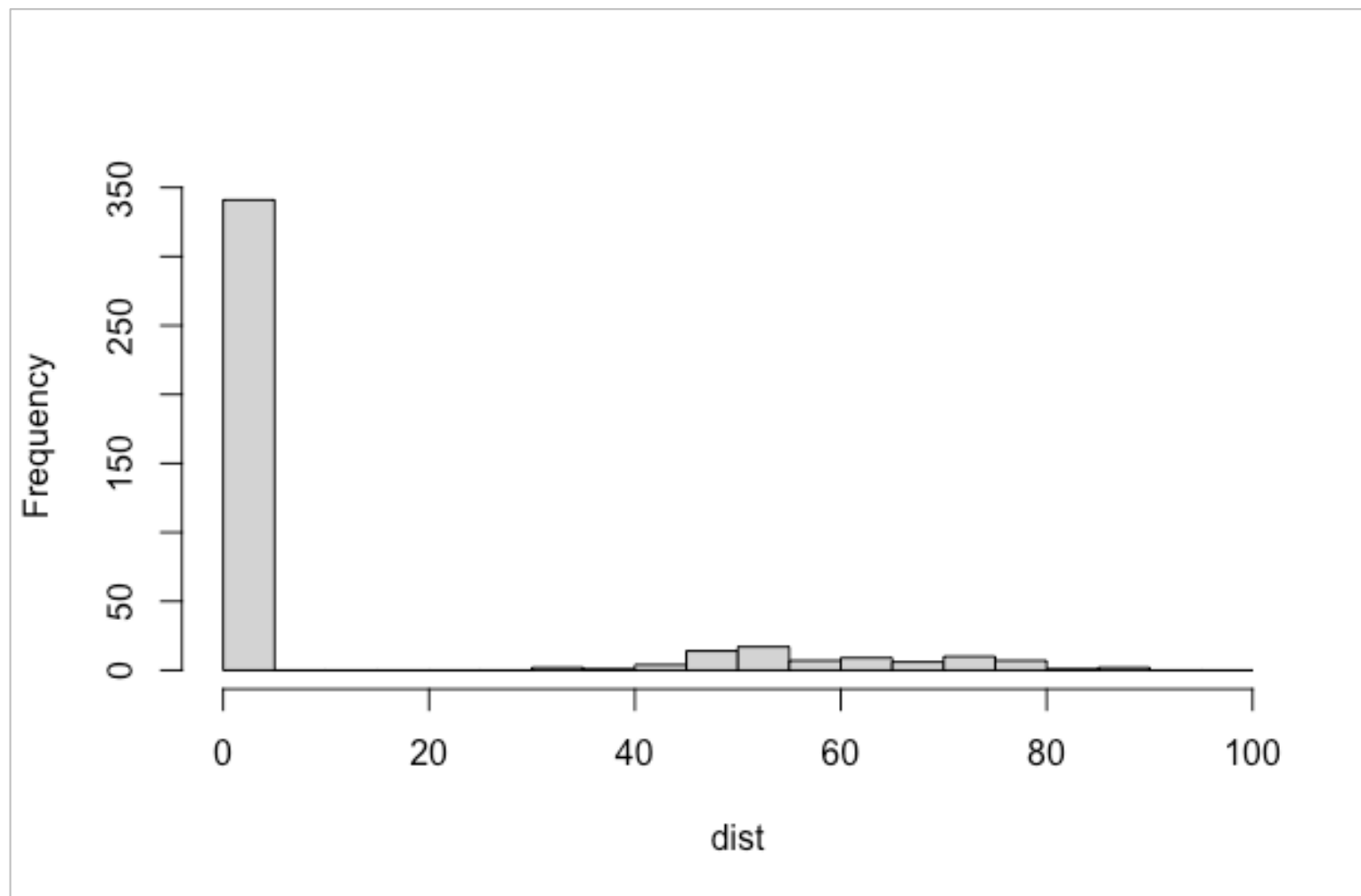
Residuals:
    Min       1Q   Median       3Q      Max
-416.62  -69.64    5.26   70.26  325.46

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   525.10      25.80  20.353 < 2e-16 ***
time           53.26      17.07   3.120  0.00191 **
treat          19.24      32.54   0.591  0.55470
time:treat     -8.06      21.33  -0.378  0.70562
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 115.5 on 511 degrees of freedom
(47 observations deleted due to missingness)
Multiple R-squared:  0.04319,    Adjusted R-squared:  0.03757
F-statistic: 7.688 on 3 and 511 DF,  p-value: 4.936e-05
```

## 分析②

- 履修科目の類似性
- ほとんどの人が全く一緒の講義を取っている



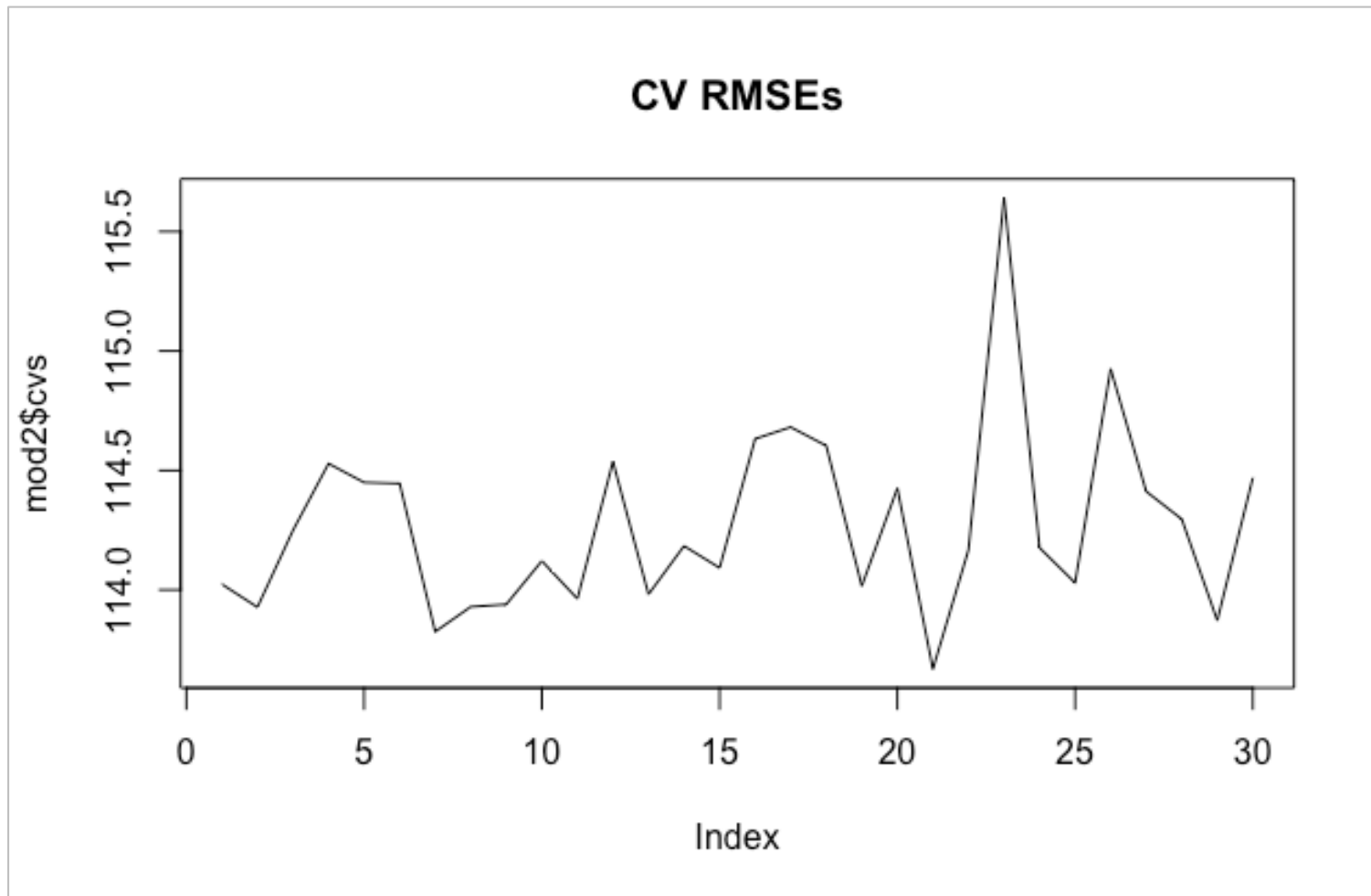
# 手法

- $y_{it} = \delta(t) + \alpha D(i, t) + \eta(i) + v(i, t) \quad \dots \textcircled{1}$
- $y_{it} = \delta(t) + \alpha D(i, t) + \beta R(i, t) + \eta(i) + v(i, t) \quad \dots \textcircled{2}$
- 相互作用がなければ,  $\alpha$ が因果効果
- けれども, 相互作用があれば, 因果効果 ( $\alpha$ ) と相互作用効果 ( $\beta$ ) に分けられる
- 相互作用も距離が近いほど大きく, 遠ければ小さい場合,  $\textcircled{2}$ 式は以下の通りとなる
- $y_{it} = \delta(t) + \alpha D(i, t) + \beta_1 R_1(i, t) + \beta_2 R_2(i, t) + \dots + \beta_k R_k(i, t) + \eta(i) + v(i, t)$
- 仮に距離が0から100の範囲で定義でき, 5分割ごとに相互作用が異なることが分かっていたら  $k=5$ で5個の相互作用の係数が推定できるはず
- 今は, 分からないので1~30分割し, 以下の $\textcircled{3}$ 式でRMSEが最小化される点を分割数とする (leave-one-out法)
- $CV(k) = \frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}(X_i(k), \hat{\theta}_i))^2 \quad \dots \textcircled{3}$



## 分析②

- 今回のデータだとデータ数が少なくて収束しない
- データ数を増やしてやってみたら収束は少し良くなるが、ピア効果はないと推定される



# まとめ

## • 結論と考察

- 自然実験的な状況を用いたDIDによれば、AプログラムによるTOEICの点数の伸びはない（留学に行かなくてもTOEICの勉強は出来る）
- 今回はあくまでも、TOEICの点数を従属変数に用いたが、その他の従属変数（BEVIを含む非認知的能力など）で測定すると、Aプログラムの効果はあるかもしれない

## • 今後の課題

- 平均効果はないが、効果の異質性（heterogeneity）はあるかも
- 外的妥当性（external validity）
- 仲の良さを他の変数で代替？（今回の変数のようなものではなく、重複が小さいもの）

## • 示唆

- 自然実験の状況が生まれる可能性があるので、データの定点観測は重要

【付記】本研究推進に際し、JSPS科研費JP18K18651・JP18K186512・JP19H00621・JP19H01690・JP20H01643・JP20J14673の資金提供と西谷元教授よりデータの提供を受けた。ここに記して感謝申し上げる次第である。

# 参考文献

- Clarke, D., 2017, " Estimating Difference-in-Differences in the Presence of Spillovers, " *Munich Personal RePEc Archive*.
- Craig, P., Katikireddi, S. V., Leyland, A, and Popham, F, 2017, "Natural Experiments: An Overview of Methods, Approaches, and Contributions to Public Health Intervention Research," *The Annual Review of Public Health*, 38: 39-56.
- 川田恵介・西谷元, 2017, 「短期留学プログラムが語学到達度に与える影響について：広島大学STARTプログラムの事例から」  
(<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxrZWlzdWtla2F3YXRhN3xneDoxOTQwNDhhNDNkZGIwZWNm>, 2021年3月20日最終確認日) .
- 川田恵介・西谷元・森田光宏, 2019, 「短期の留学でも語学力は向上するのか？」  
(<https://premium.toyokeizai.net/articles/-/20825>, 2021年3月20日最終確認日) .
- Imai Kosuke, 2019, "Design and Analysis of Two-Stage Randomized Experiments"  
(<https://imai.fas.harvard.edu/talk/files/MIT19.pdf>)
- Wing, C., Simon, Kell-Gomez, R. A., 2018, "Designing Difference in Difference Studies- Best Practices for Public Health Policy Research, " *The Annual Review of Public of Health*, 39 : 453-69.