

(教科書では) 基本的にパネルデータを扱っている。ここでのパネルデータは時系列と各項目(地域)。この期間中に何か介入があり、介入・非介入の各グループに同質なサンプルがない場合の効果を検証したい。

### **集計による効果検証 (①②)とその欠点 → DIDはその問題点を解決**

#### ① 介入を受けた地域と受けなかった地域で売上を比較

→問題:「地域と介入が完全に相関する、という可能性」問題。(地域ABCDでAのみ介入。地域Aのみに固有の効果があるとバイアスが発生。所得の格差、地理的な影響)

#### ② 前後比較

→問題:介入がなくても売上上昇(効果)があったかもしれない。(セレクションバイアスがサンプル間だけでなく、同一サンプルのタイミング間においても発生する可能性。イギリス全土での所得の向上による健康状態の向上。ロンドン全体におけるコレラの蔓延度合い)

**DID**・・・介入が行われた地域における介入の前後のデータに加えて、他の地域の介入の前後のデータを利用して因果推論する手法。

「地域における前後比較」をし、その後「その前後比較の結果」を地域間で比較する。

「興味のある介入」が「特定の共変量(特に地理情報)」と相関するケースだと、管理運営が地理的な区画ごとに行われる場合によくみられる(実店舗、マス広告、行政(法律や政策の変更))

**平行トレンド仮定** (common trend assumption)・・・『非介入グループのデータの変化』と『介入グループが介入していなかった場合の変化』が一致する、という仮定。この仮定が成り立たないと OVB が発生する。

「介入までの期間がある程度長いデータ」を分析している場合は、  
「介入までのトレンドが似ているか」を確認。しかし、介入までの期間が複数必要になる&明確な傾向が得られない。

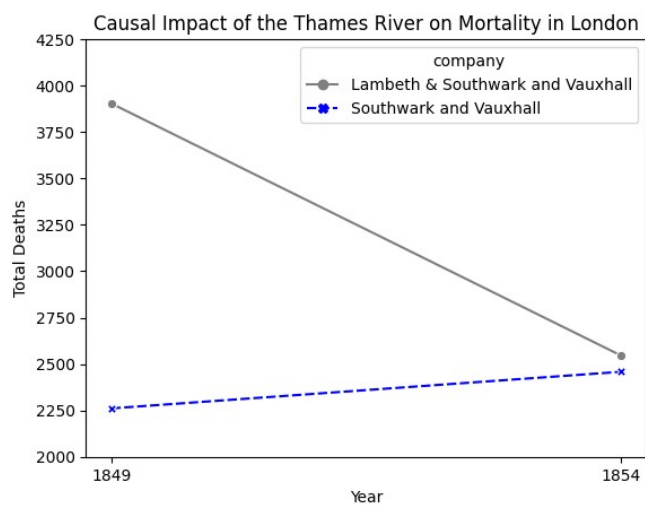
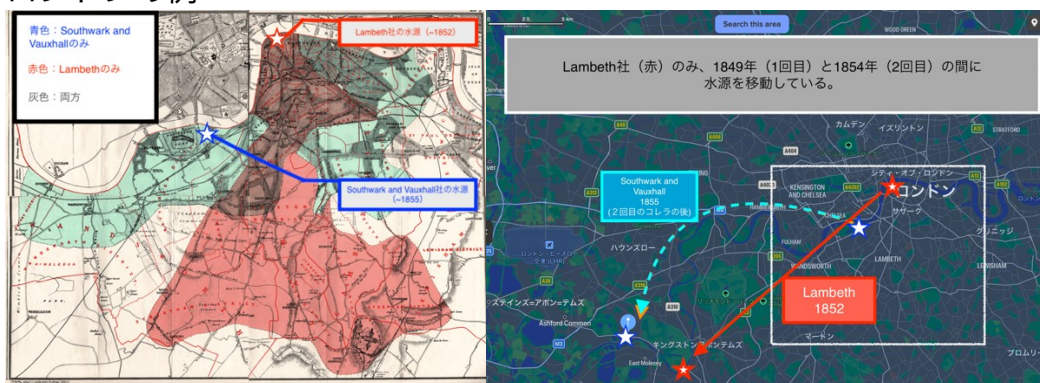
## トレンドが同一でない場合

1. 仮定が満たされていないと考えられる他の地域を取り除く。  
Synthetic Control、近隣の地域のみを選ぶ
2. 共変量としてトレンドの乖離を説明するような変数をモデルに加える  
e.g. 水源の変化しなかったいくつかの地域で施設が新設→衛生改善  
「施設が改善」という情報が共変量としてモデルに入れば OK

## 介入の他に別の介入や施作が行われる(Activity Bias)

複数の効果が混じってしまうために、DID,CasualImpactでの検証ができない。(e.g. 商品の入れ替え、Web ページの刷新、広告の同時実施)

## ロンドンの例



### DIDにおける標準誤差

- ・時系列データでのデータを利用。自己相関の発生。

e.g. 店舗の売上は自己相関が高い。

- ・誤差項の値は同一店舗で似通った値になる。

→クラスター標準誤差を利用してパラメータの標準誤差を算出。

× サンプルごとに誤差を観測している

○ 指定した観測対象ごと（店舗ごと）に誤差を観測している（と考える）

e.g. ある店舗の売上データ

× 毎月の誤差を考える

○ その人物の全ての観測された期間での誤差を考える

（教科書には出てこないが、）ロバスト標準誤差というものもある。

---

### CasualImpact

①②の欠点を補える。ただし、複数の期間におけるデータが必要。DIDのアイデアは「介入が行われたサンプルの”介入が行われなかった場合”の結果を非介入グループの結果で補う」→「介入が行われたサンプルの”介入が行われなかった場合”の結果」を予測したい。

様々な種類の変数\_Xを利用してYを予測できるモデルを介入前の期間のみで作成

e.g. ある店舗で価格を変えた効果を推定する

Y ... 売り上げ

X ... （価格を変えた）商品名の検索数、

複数種類の検索回数、

web アクセス回数

### CasualImpactでは平行トレンド仮定は大事

どのデータをそもそも使うのか、は分析者の仮定が必要になる。

○ 「売り上げ\_Y」と強い相関がありつつ、「介入\_Z（価格変更）の影響」を受けない「変数\_X」を探すことが大事。

× 「売り上げ\_Y」だけでなく、「介入\_Zの影響」を受ける「変数\_X」

× 「売り上げ\_Y」から影響を受ける「変数\_X」

(∵  $Z \rightarrow Y \rightarrow X$  と順路で介入の影響を受けてしまう)

最大の障壁は「介入が行われる地域の Y」を上手に予測できる変数 X のデータを見つけること。

→ 「他の地域の Y」を上手に予測する必要性はない。

→ 分析者に利用できるデータの種類が多くなる。アイデアも大事。

例 1：「売り上げ」と強い相関がありつつ、「介入（価格変更）の影響」を受けない「変数\_X（検索数）」（Google Trend は役に立つ）

例 2：「売り上げ」と強い相関がありつつ、「介入（広告）の影響」を受けない変数として「別の国の検索数」

カリフォルニアのタバコ規制 Proposition99 （Abadie, et al.(2010)）

（ Proposition 《米》 《法》（住民投票による）州法改正案。）

1988 年から実施。1993 年まで同様の試みは他の州で実施されず。

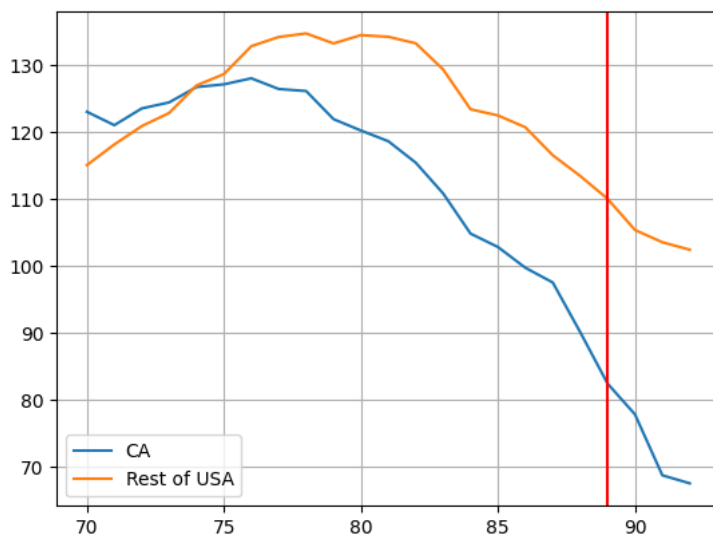
多岐にわたる活動が行われた

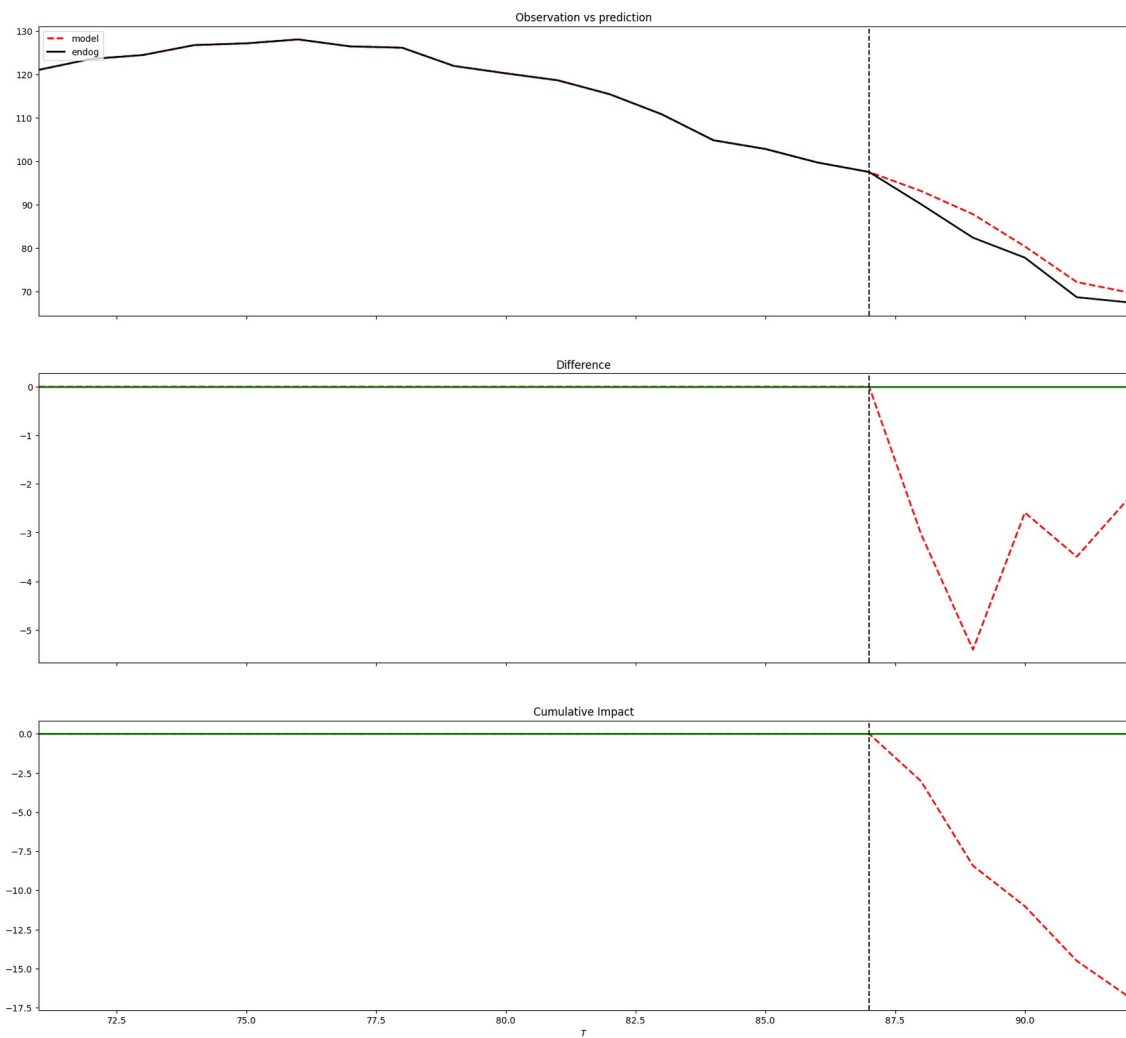
タバコ 1 箱に対して 25 セントの税金

→ 税収を「タバコの健康被害に対する教育」や「メディア上の禁煙キャンペーン」の予算に

室内での空気環境を改善するための条約を設立

1 億ドル/年





- 
- DID,CasualImpact は、実験が限定的にしか行われない環境で重宝される。