Synthetic Difference in Differences (SDiD)

Prof. Mike Aguilar

https:/www.linkedin.com/in/mike-aguilar-econ

Outline

1. Motivation - Case Study
2. SDiD Intuition
3. Synthetic DiD as a Weighted 2-way FE Specification
4. Details
   - Double Bias Adjustment
   - Weights
   - Advantages/Disadvantages

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences
- Similar to Synthetic Control
  - Uses unit and time weights to construct a synthetic comparison

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences
- Similar to Synthetic Control
  - Uses unit and time weights to construct a synthetic comparison
  - Balances pre-treatment outcomes to improve credibility

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences
- Similar to Synthetic Control
  - Uses unit and time weights to construct a synthetic comparison
  - Balances pre-treatment outcomes to improve credibility
- Synthetic DiD
  - Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences
- Similar to Synthetic Control
  - Uses unit and time weights to construct a synthetic comparison
  - Balances pre-treatment outcomes to improve credibility
- Synthetic DiD
  - Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes
  - Combines weighting and differencing to reduce bias

Outline

1. Motivation - Case Study
2. SDiD Intuition
3. Synthetic DiD as a Weighted 2-way FE Specification
4. Details
   - Double Bias Adjustment
   - Weights
   - Advantages/Disadvantages

Case Study - Cigarette Taxes



Image created with Bing Image Creator

Case Study - Background

- In 1988 California approved Proposition 99 (Prop.99)

Case Study - Background

- In 1988 California approved Proposition 99 (Prop.99)
- Prop.99 was a large-scale tobacco control program that increased cigarette excise tax by 25 cents per pack

Case Study - Background

- In 1988 California approved Proposition 99 (Prop.99)
- Prop.99 was a large-scale tobacco control program that increased cigarette excise tax by 25 cents per pack
- The (partial) intent of the program was to curb consumer demand for cigarettes

Case Study - Background

- In 1988 California approved Proposition 99 (Prop.99)
- Prop.99 was a large-scale tobacco control program that increased cigarette excise tax by 25 cents per pack
- The (partial) intent of the program was to curb consumer demand for cigarettes
- The program went into effect Jan1989

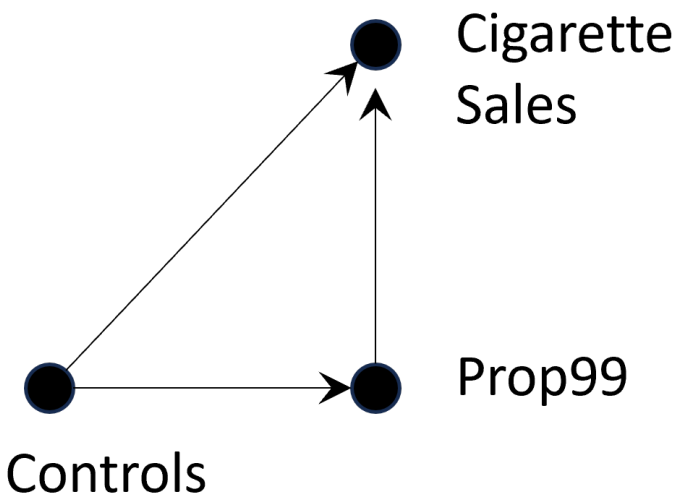
Case Study - Background

- In 1988 California approved Proposition 99 (Prop.99)
- Prop.99 was a large-scale tobacco control program that increased cigarette excise tax by 25 cents per pack
- The (partial) intent of the program was to curb consumer demand for cigarettes
- The program went into effect Jan1989
- The following is based upon the work of Abadie, Diamond, and Hainmueller (2010)

Causal Path

Q: Can you draw the DAG and define the treatment effect?

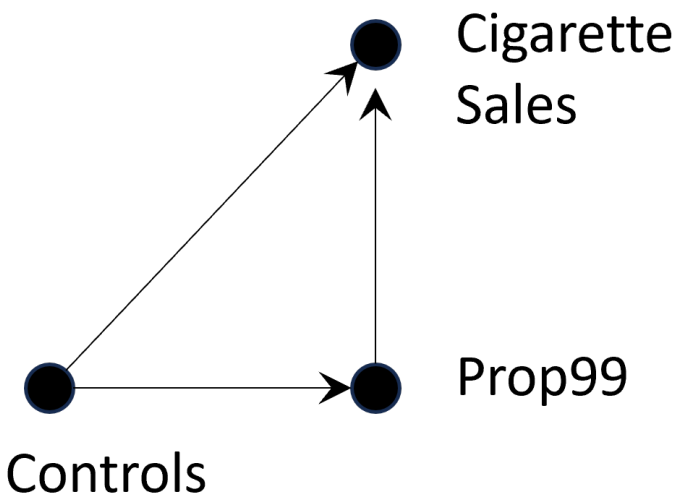
A:



Causal Path

Q: Can you draw the DAG and define the treatment effect?

A:

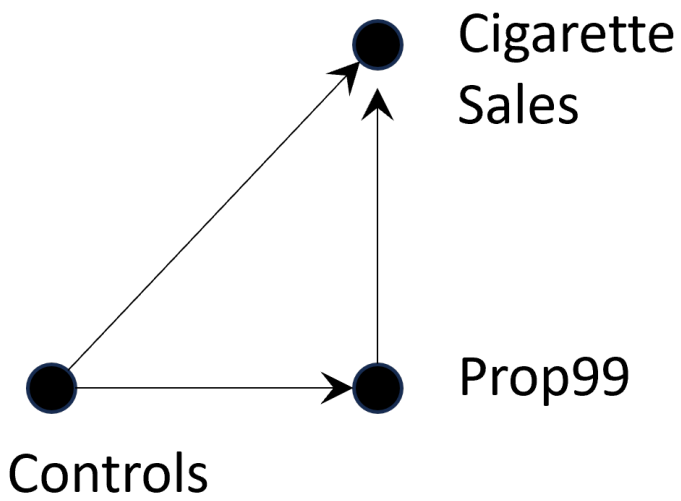


- Outcome:

Causal Path

Q: Can you draw the DAG and define the treatment effect?

A:

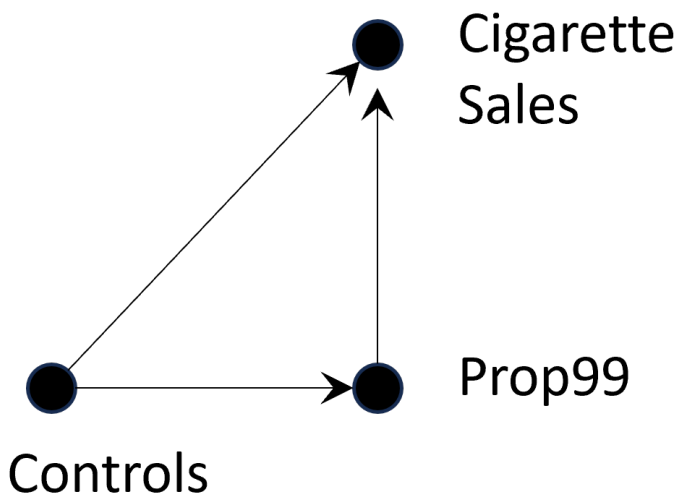


- Outcome: Cigarette sales

Causal Path

Q: Can you draw the DAG and define the treatment effect?

A:

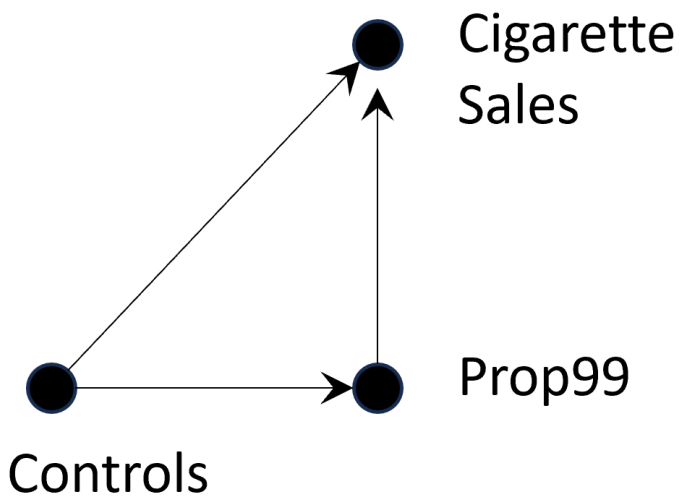


- Outcome: Cigarette sales
- Treatment:

Causal Path

Q: Can you draw the DAG and define the treatment effect?

A:

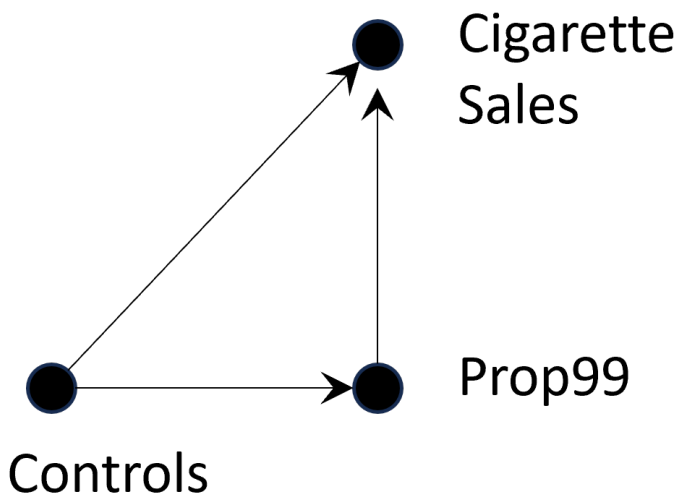


- Outcome: Cigarette sales
- Treatment: Prop.99

Causal Path

Q: Can you draw the DAG and define the treatment effect?

A:

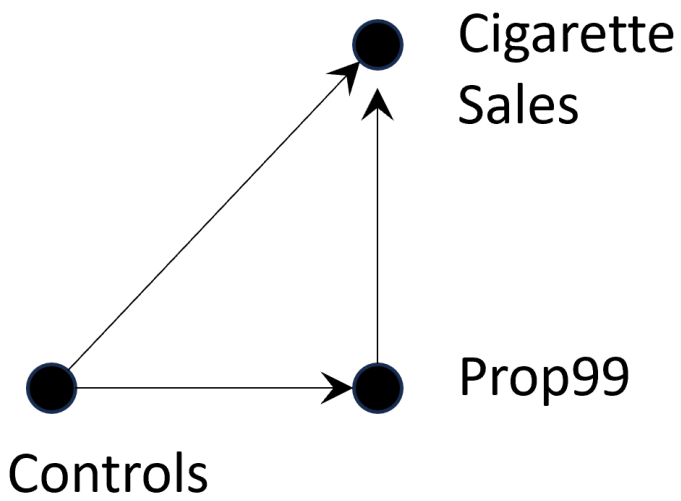


- Outcome: Cigarette sales
- Treatment: Prop.99
- Treatment Effect:

Causal Path

Q: Can you draw the DAG and define the treatment effect?

A:



- Outcome: Cigarette sales
- Treatment: Prop.99
- Treatment Effect: Impact of Prop.99 on cigarette sales

Dataset

- California
- Annual data
- 1970-2000
- Cigarette sales are available for the state on an aggregate basis each year (e.g. per capita cigarette sales in 1980 were 120 packs)
- We also have controls on an aggregate basis such as GDP per capita, demographics, retail price of cigarettes, and beer consumption

Dataset

- California
- Annual data
- 1970-2000
- Cigarette sales are available for the state on an aggregate basis each year (e.g. per capita cigarette sales in 1980 were 120 packs)
- We also have controls on an aggregate basis such as GDP per capita, demographics, retail price of cigarettes, and beer consumption
- Cigarette sales data for all other states
- We can assume that California is the only state with cigarette taxes at that time.

Outline

1. Motivation - Case Study
2. SDiD Intuition
3. Synthetic DiD as a Weighted 2-way FE Specification
4. Details
   - Double Bias Adjustment
   - Weights
   - Advantages/Disadvantages

The Setting

Recall the hierarchy of causal settings

1. Randomized Experiments
2. Natural Experiments
3. Observational Studies

The Setting

Recall the hierarchy of causal settings

1. Randomized Experiments
2. Natural Experiments
3. Observational Studies

Q: Why is our case a Natural Experiment?

The Setting

Recall the hierarchy of causal settings

1. Randomized Experiments
2. Natural Experiments
3. Observational Studies

Q: Why is our case a Natural Experiment?
A: An exogenous policy event (Prop.99) allocated treatment among the States (e.g. the businesses and consumers didn't choose to be "treated" with the taxes).

As usual, we can debate whether this is a natural experiment or observational study. Regardless, the method doesn't change; only the confidence we have in the results.

Combining Prior Approaches

- DiD was useful in a natural experiment setting. It created a counterfactual for the treated group (i.e. potential outcome if untreated) from many control units via two-way FE (e.g. State and Year).

Combining Prior Approaches

- DiD was useful in a natural experiment setting. It created a counterfactual for the treated group (i.e. potential outcome if untreated) from many control units via two-way FE (e.g. State and Year).
- Synthetic Control was useful when we had only 1 treated unit. It created a pseudo treated unit as if it was untreated by weighting the control units before treatment (e.g. Synthetic CA without Prop.99)

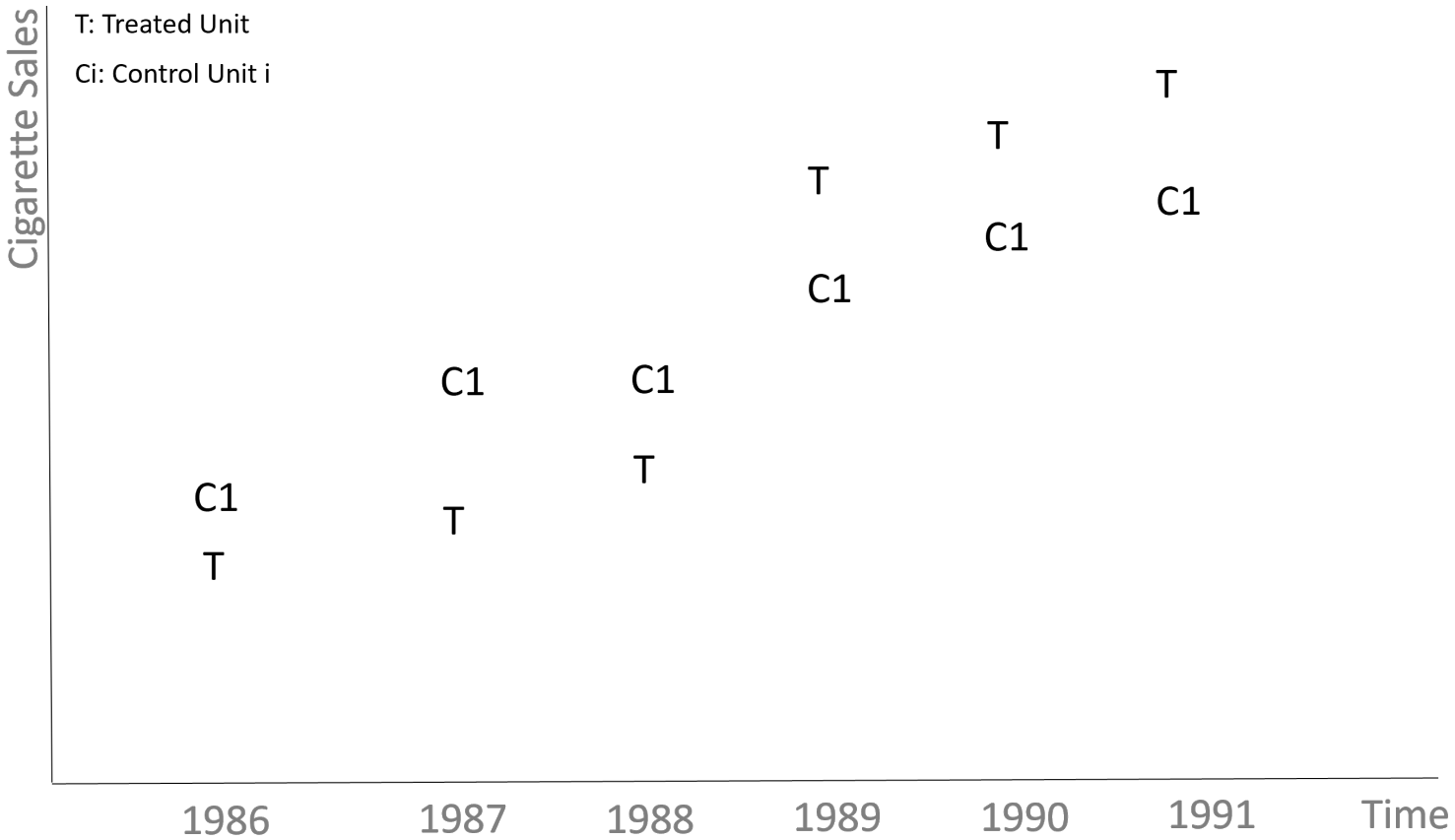
Combining Prior Approaches

- DiD was useful in a natural experiment setting. It created a counterfactual for the treated group (i.e. potential outcome if untreated) from many control units via two-way FE (e.g. State and Year).
- Synthetic Control was useful when we had only 1 treated unit. It created a pseudo treated unit as if it was untreated by weighting the control units before treatment (e.g. Synthetic CA without Prop.99)
- Synthetic DiD combines these approaches by creating a pseudo treated group via weighting the control units and weighting each pre-treatment time period.

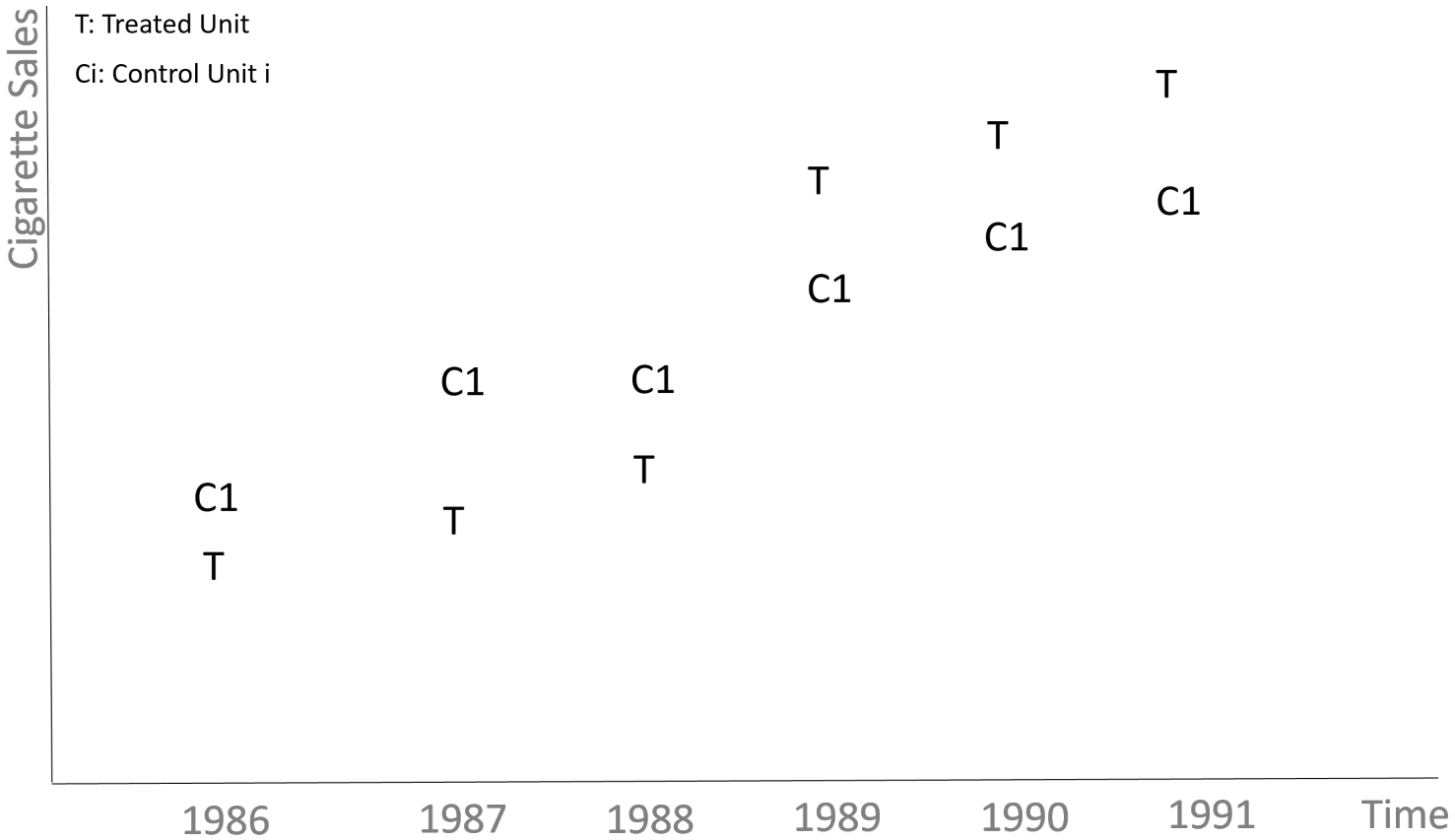
Combining Prior Approaches

- DiD was useful in a natural experiment setting. It created a counterfactual for the treated group (i.e. potential outcome if untreated) from many control units via two-way FE (e.g. State and Year).
- Synthetic Control was useful when we had only 1 treated unit. It created a pseudo treated unit as if it was untreated by weighting the control units before treatment (e.g. Synthetic CA without Prop.99)
- Synthetic DiD combines these approaches by creating a pseudo treated group via weighting the control units and weighting each pre-treatment time period.
- The following are crude visual illustrations of the different approaches.

Natural Experiment Data

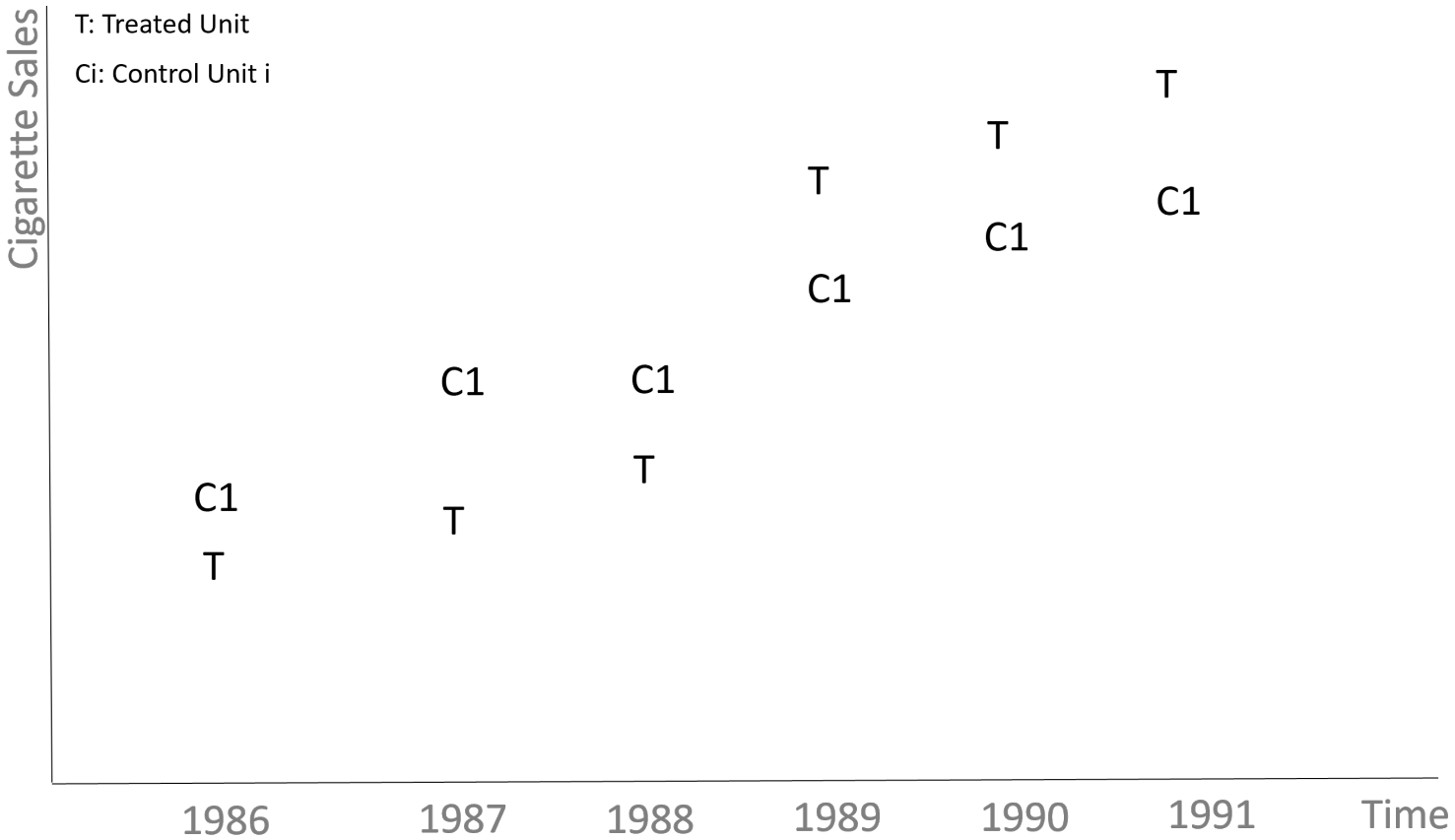


Natural Experiment Data



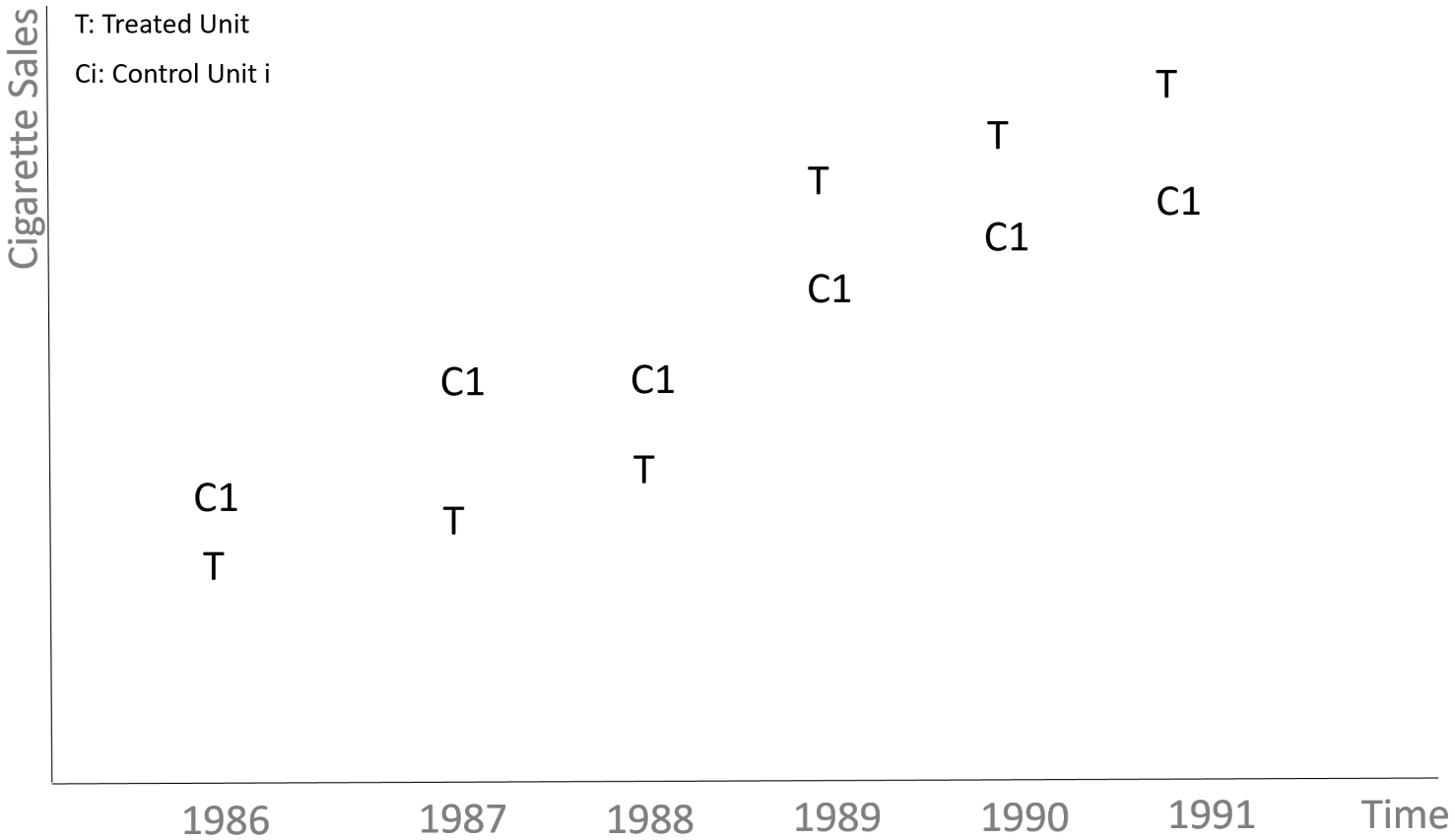
Notice there appears to be a jump in the Treated unit near 1989 (i.e. CA might be impacted by Prop.99).

Natural Experiment Data



Notice there appears to be a jump in the Treated unit near 1989 (i.e. CA might be impacted by Prop.99). Q: Does this match your apriori?

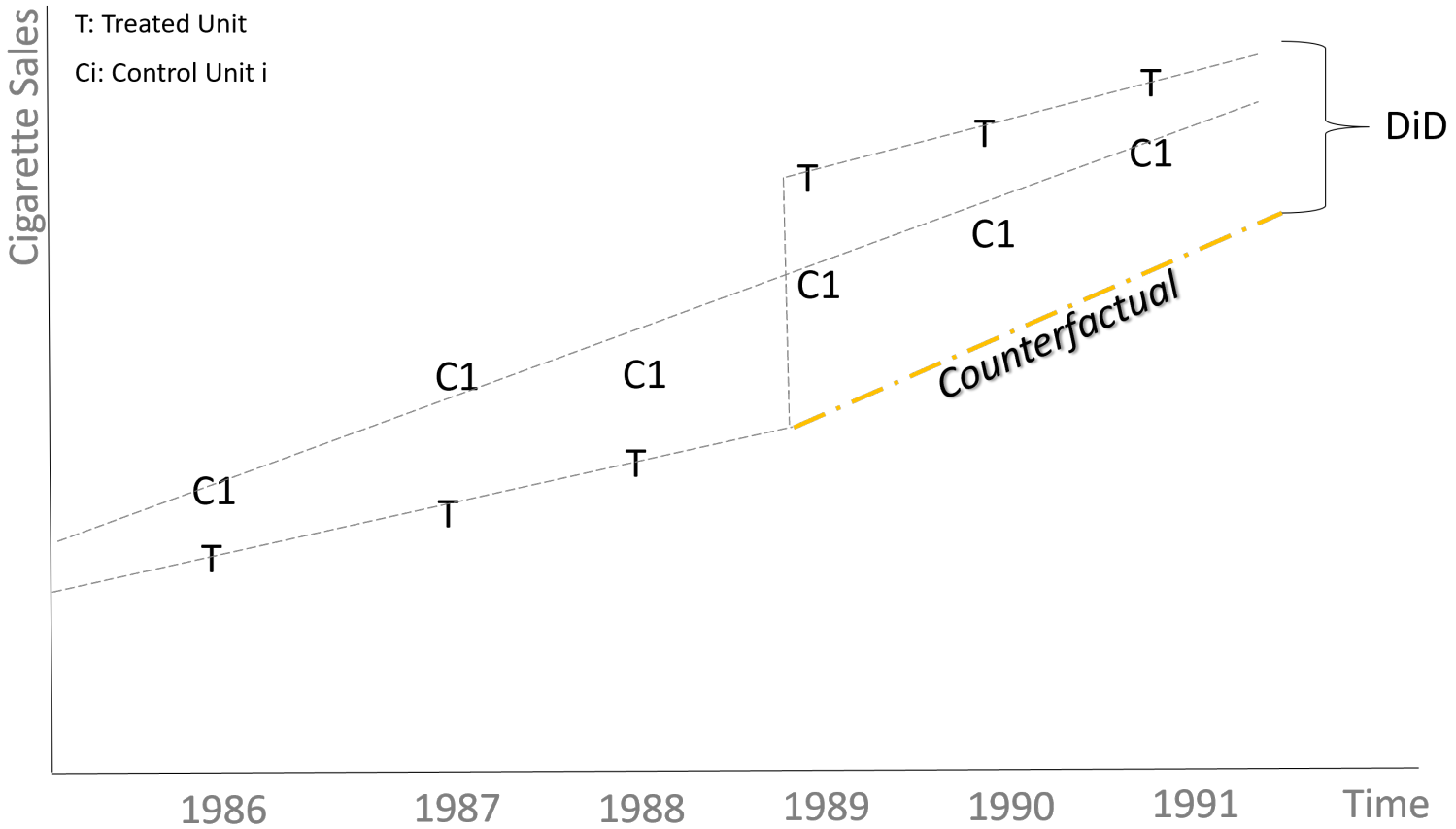
Natural Experiment Data



Notice there appears to be a jump in the Treated unit near 1989 (i.e. CA might be impacted by Prop.99). Q: Does this match your apriori? A: No. We would expect a negative impact. (This is for illustrative purposes only.)

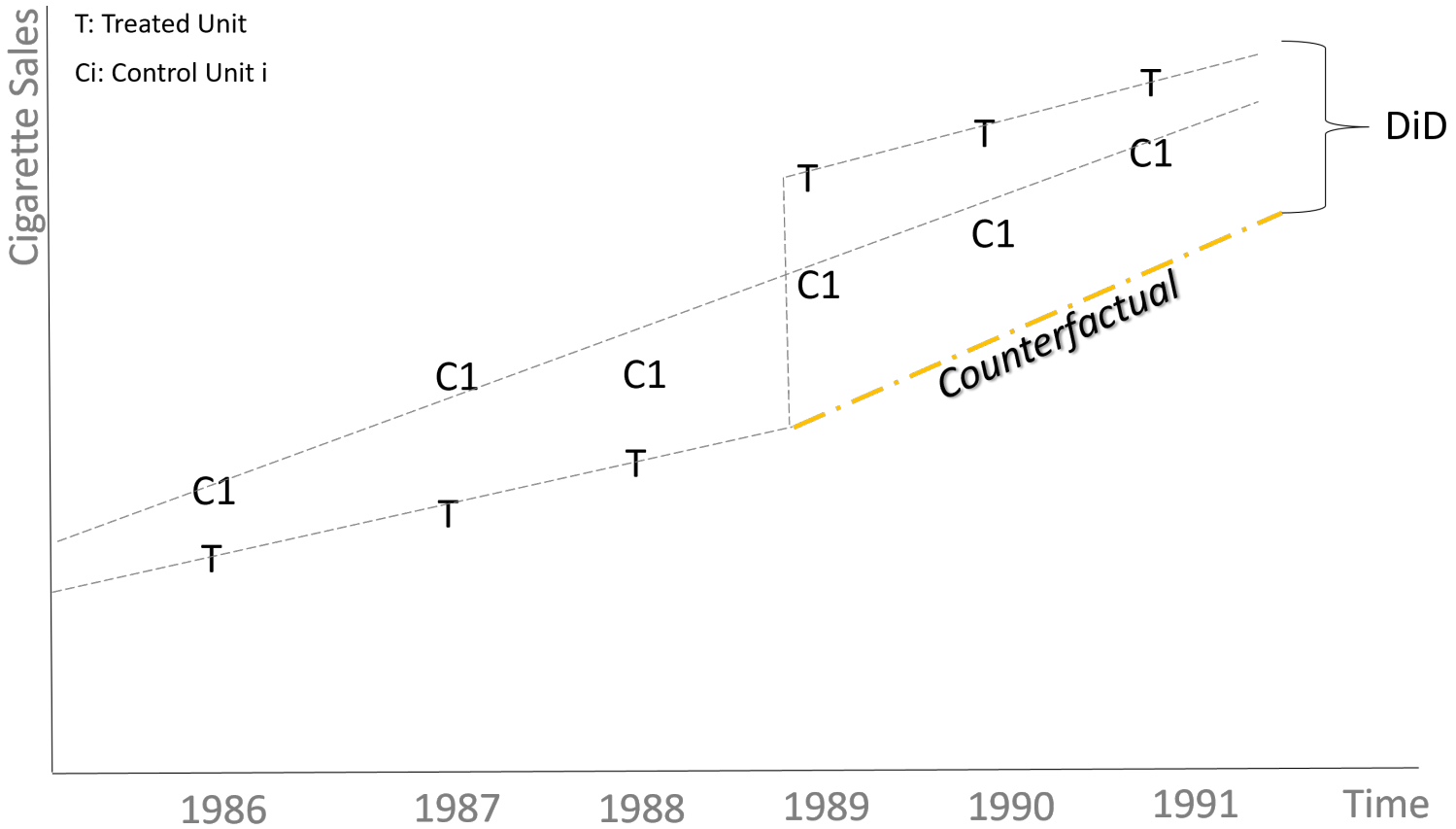
2×2 DiD

Illustrated over multiple periods, so not technically 2×2



2×2 DiD

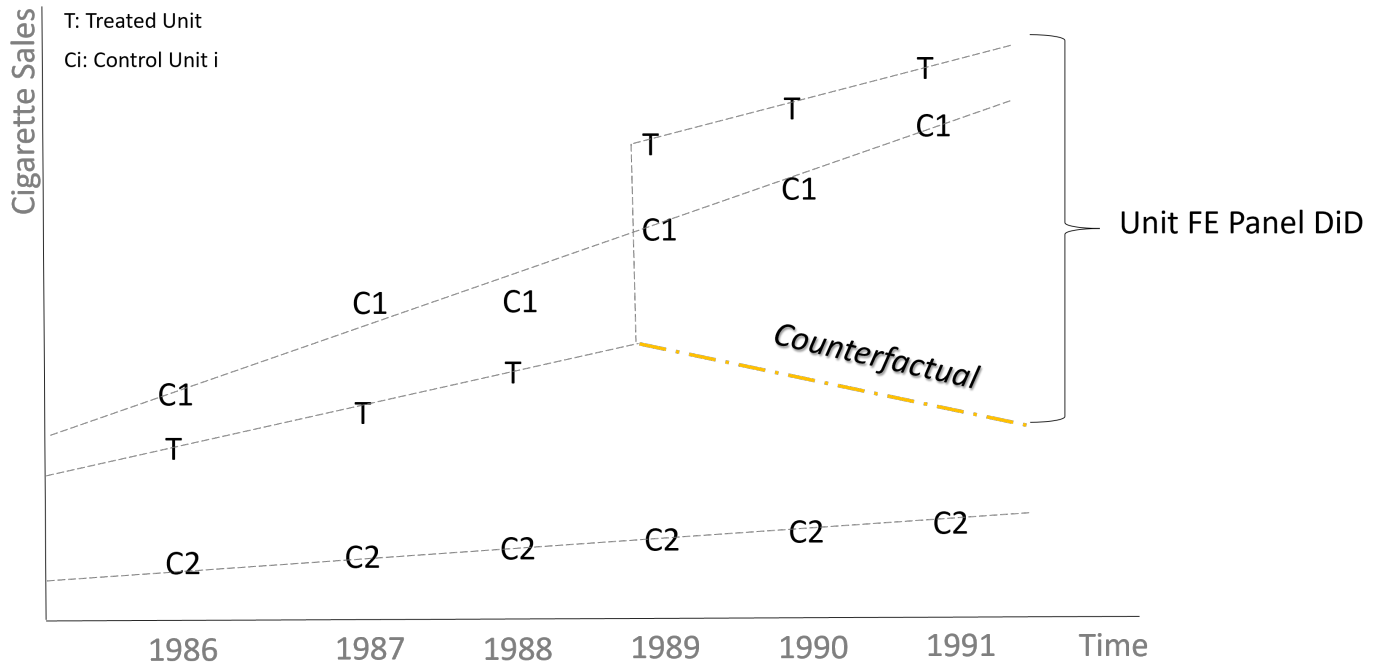
Illustrated over multiple periods, so not technically 2×2



The DiD TE estimate compares the observed cigarette sales for the treated group with the counterfactual cigarette sales, which are built from the control unit.

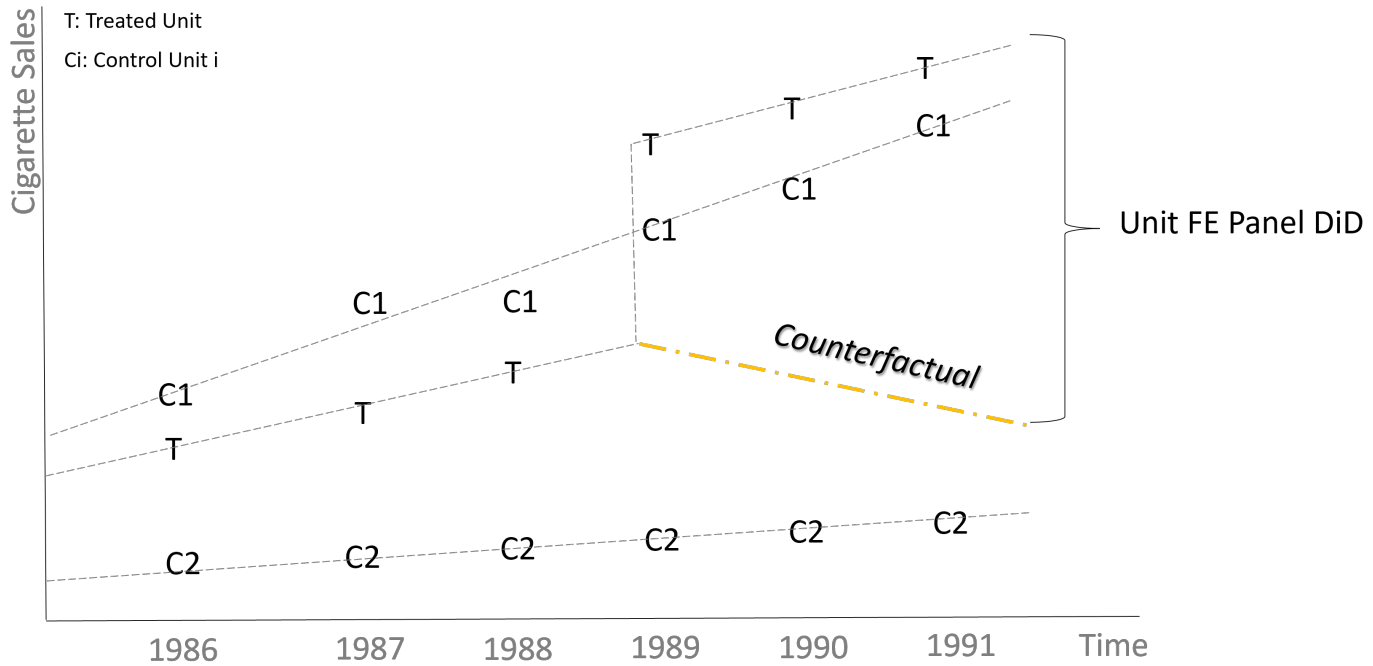
Unit FE Panel DiD

Depicted with 1 Treat / 2 Controls for illustrative purpose only



Unit FE Panel DiD

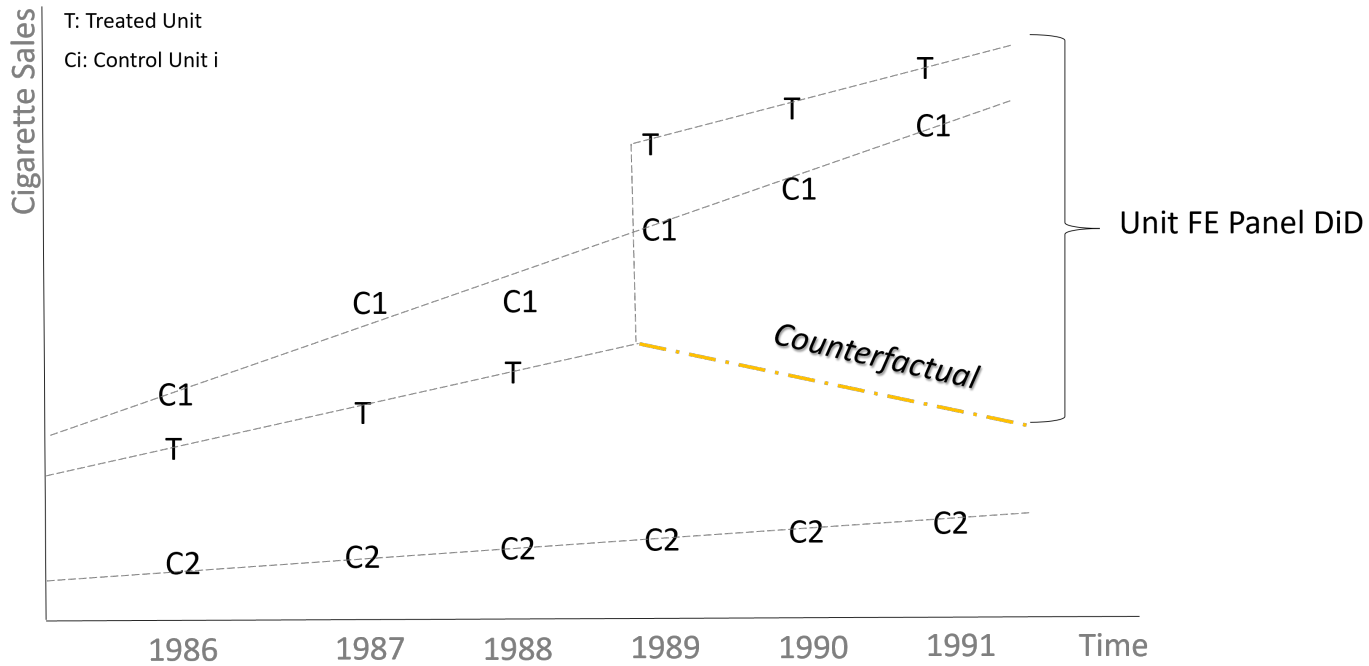
Depicted with 1 Treat / 2 Controls for illustrative purpose only



Q: Why is the FE estimate different than 2×2?

Unit FE Panel DiD

Depicted with 1 Treat / 2 Controls for illustrative purpose only

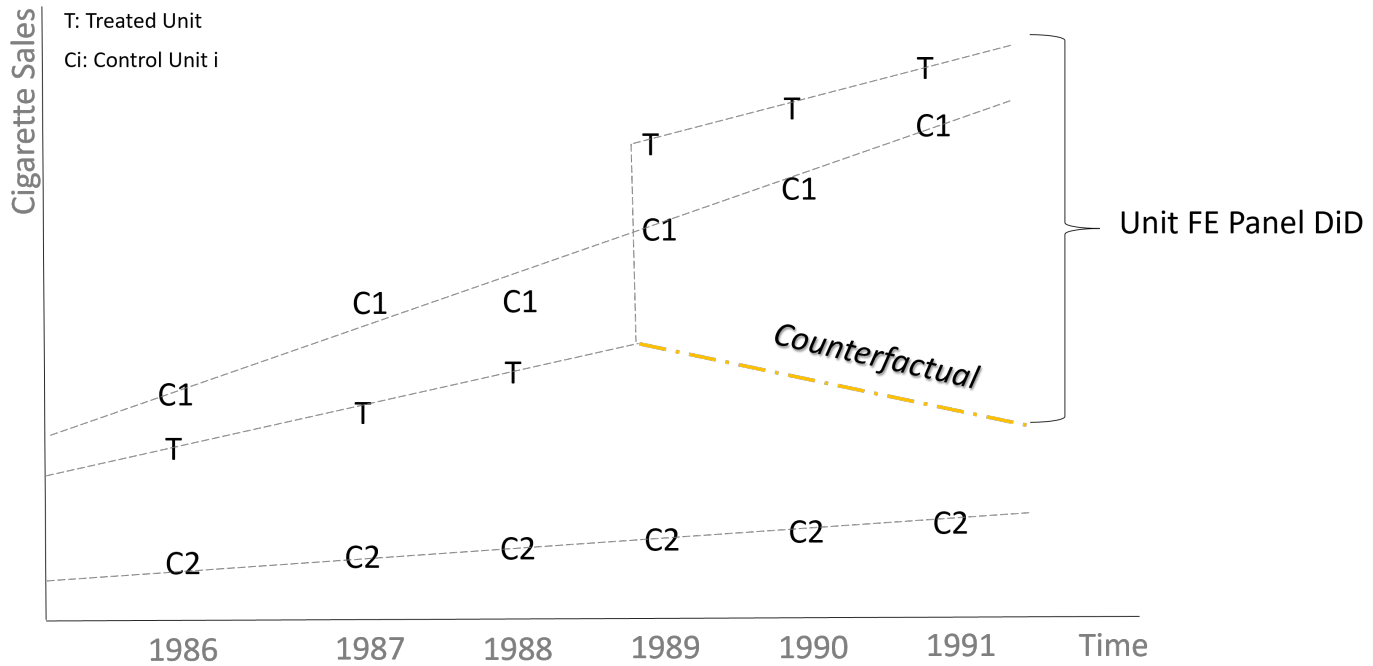


Q: Why is the FE estimate different than 2×2?

A: The presence of other controls (C2) alters the estimate.

Unit FE Panel DiD

Depicted with 1 Treat / 2 Controls for illustrative purpose only



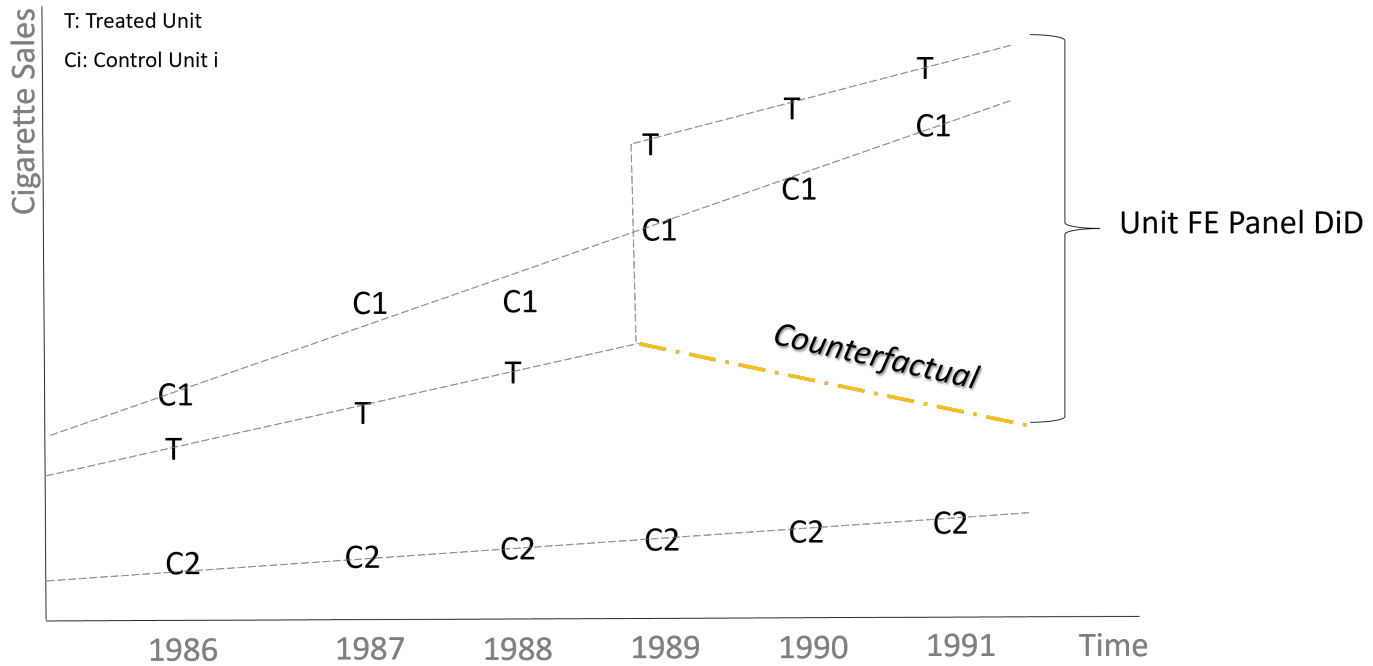
Q: Why is the FE estimate different than 2×2?

A: The presence of other controls (C2) alters the estimate.

Q: How "see" the unit fixed effects?

Unit FE Panel DiD

Depicted with 1 Treat / 2 Controls for illustrative purpose only



Q: Why is the FE estimate different than 2×2?

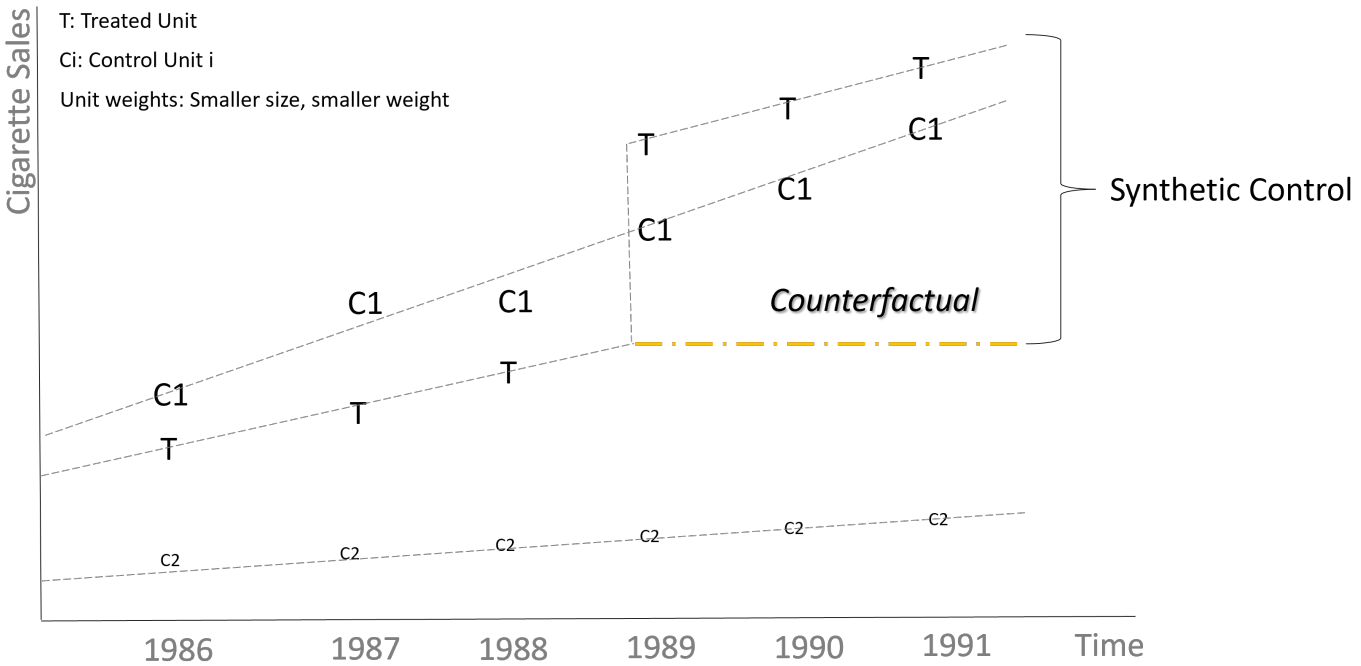
A: The presence of other controls (C2) alters the estimate.

Q: How "see" the unit fixed effects?

A: Each control unit (C1, C2) is permitted to have its own intercept (i.e. unit fixed effects)

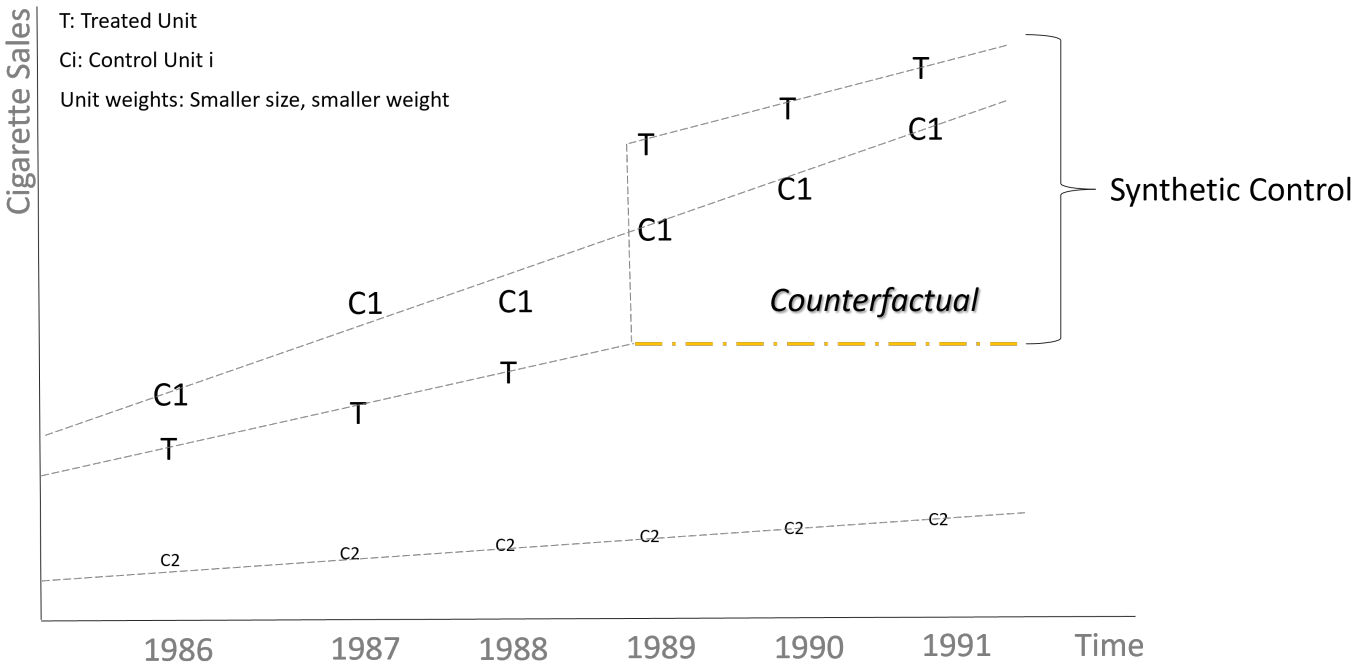
Synthetic Control

Depicted with 2 Controls for illustrative purpose only



Synthetic Control

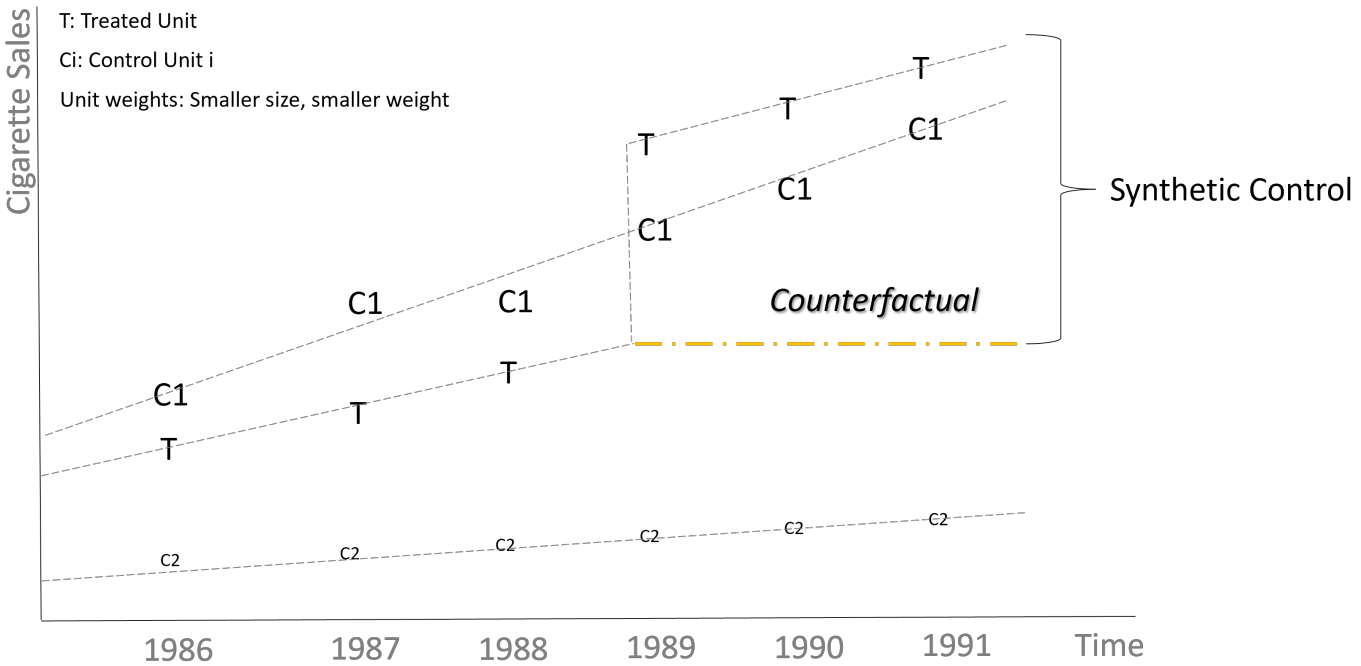
Depicted with 2 Controls for illustrative purpose only



Q: Why is the Synethic Control estimate attenuated relative to Unit FE Panel DiD?

Synthetic Control

Depicted with 2 Controls for illustrative purpose only

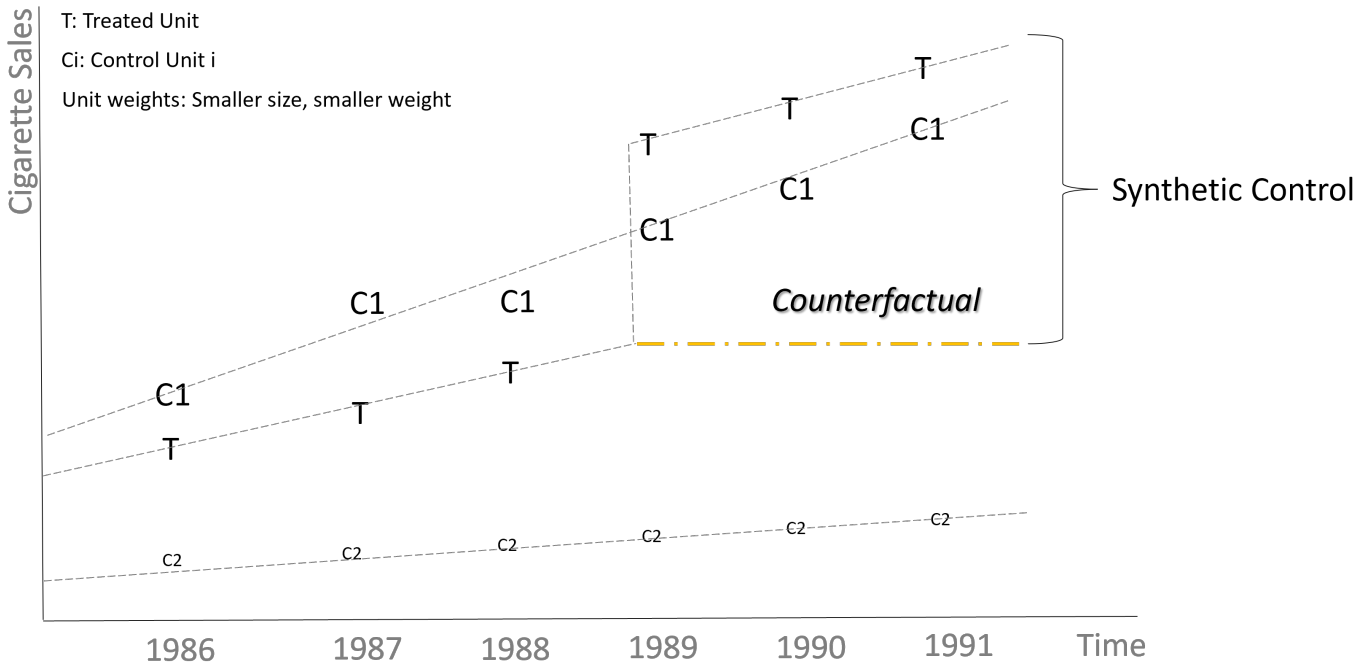


Q: Why is the Synethic Control estimate attenuated relative to Unit FE Panel DiD?

A: C2 is given less weight, because

Synthetic Control

Depicted with 2 Controls for illustrative purpose only

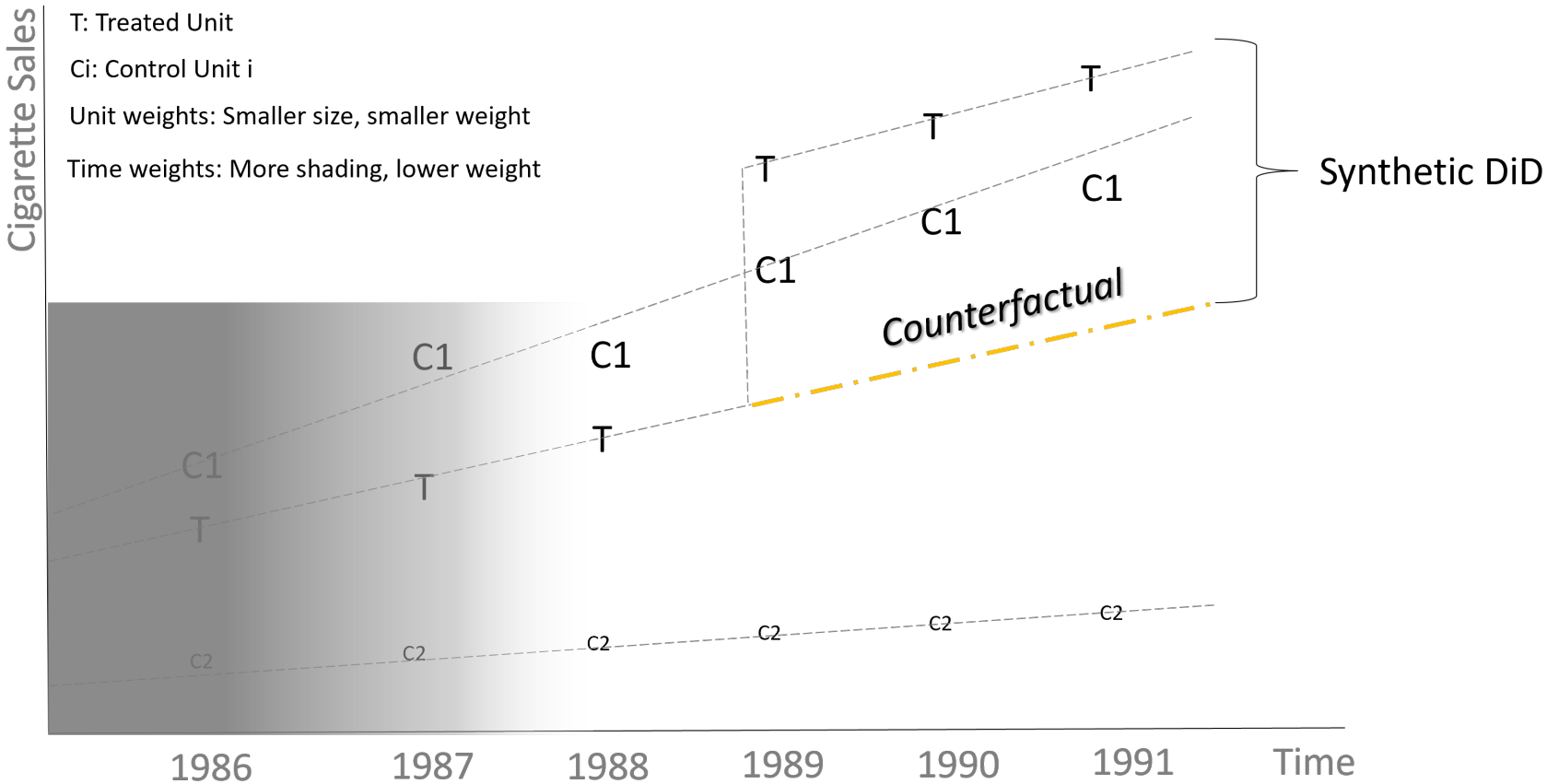


Q: Why is the Synethic Control estimate attenuated relative to Unit FE Panel DiD?

A: C2 is given less weight, because it appears to be a less informative control unit (e.g. New Hampshire's (C2) cigarette sales might not be as predictive of CA's sales as a neighboring state like Nevada (C1))

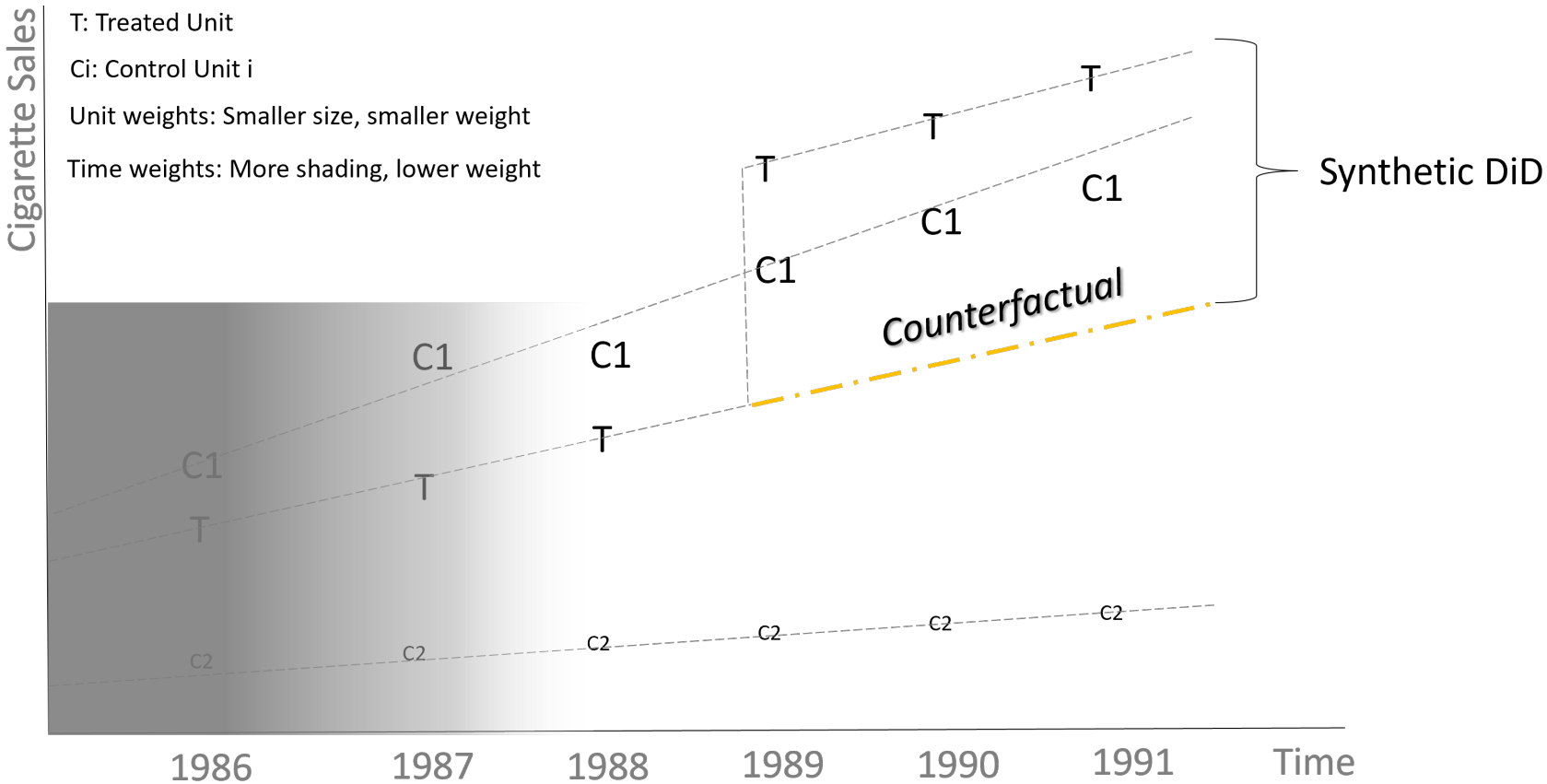
Synthetic DiD Imposes State and Time Weights

Depicted with 2 Controls for illustrative purpose only



Synthetic DiD Imposes State and Time Weights

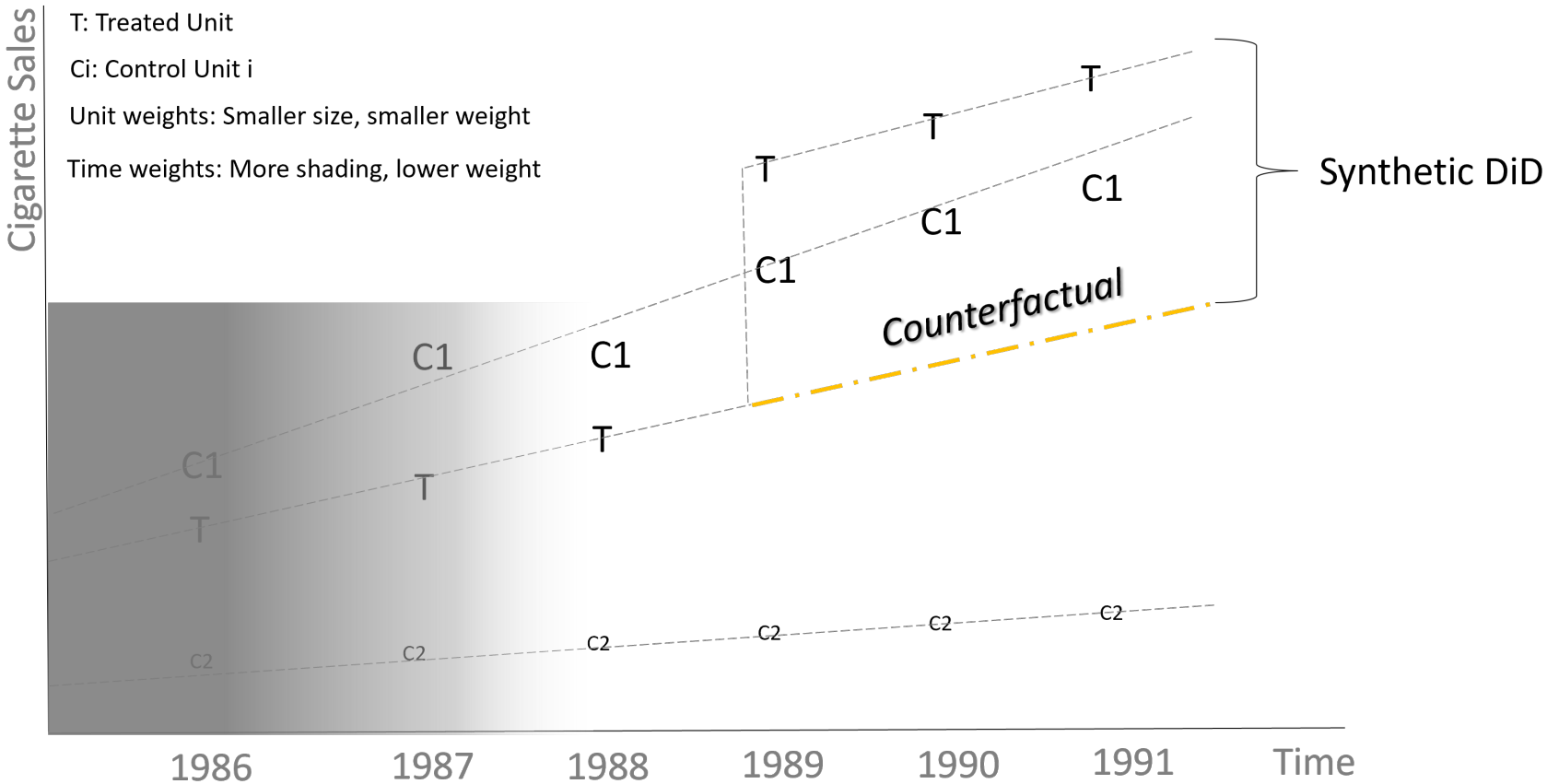
Depicted with 2 Controls for illustrative purpose only



Q: Why is 1988 given more weight than 1987, which is given more weight than 1986?

Synthetic DiD Imposes State and Time Weights

Depicted with 2 Controls for illustrative purpose only

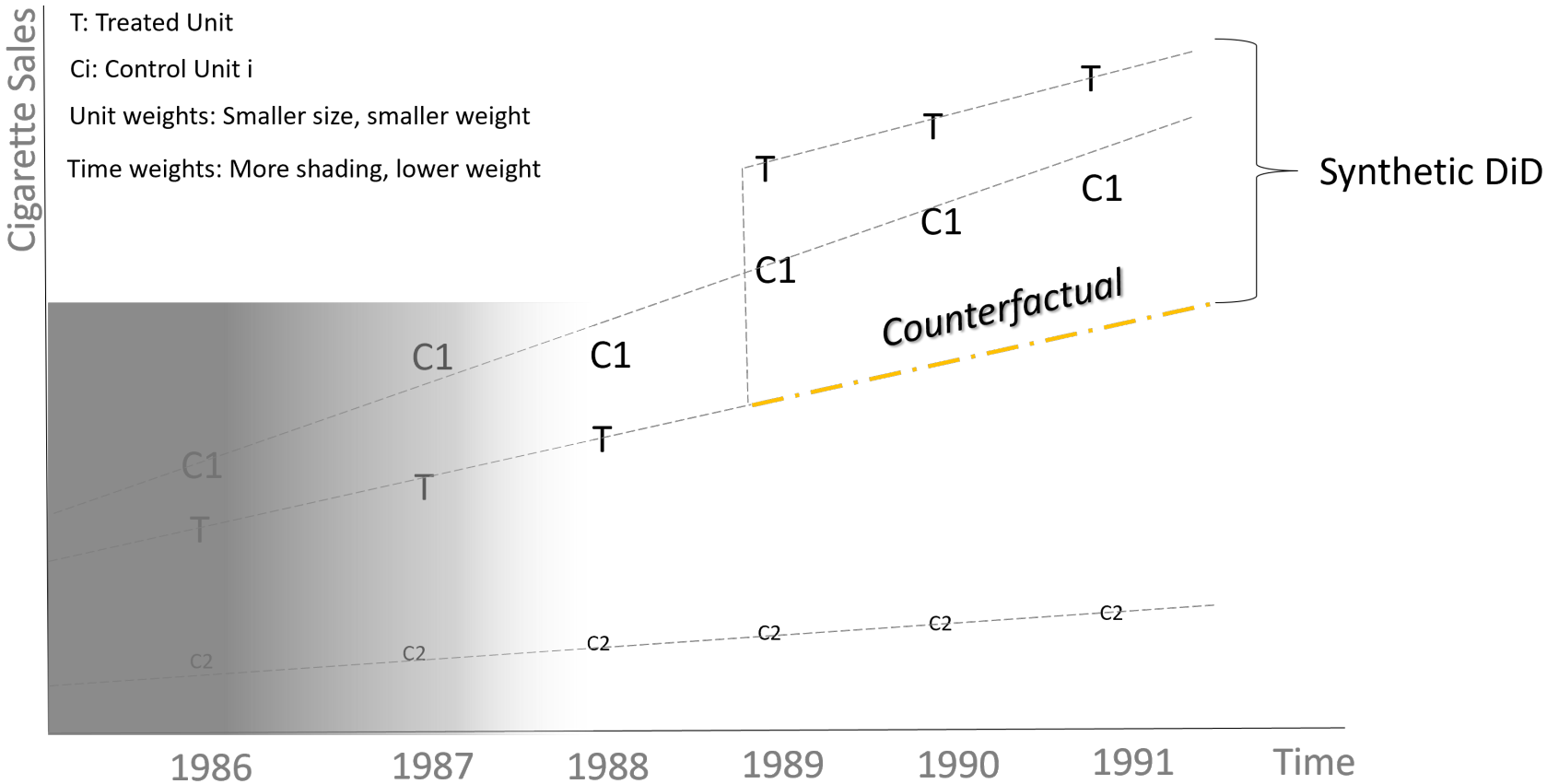


Q: Why is 1988 given more weight than 1987, which is given more weight than 1986?

A: Periods close to Prop99 might be more informative of cigarette sales after Prop99.

Synthetic DiD Imposes State and Time Weights

Depicted with 2 Controls for illustrative purpose only



Q: Why is 1988 given more weight than 1987, which is given more weight than 1986?

A: Periods close to Prop99 might be more informative of cigarette sales after Prop99. Note that weights need not be chosen in this manner.

Outline

1. Motivation - Case Study
2. SDiD Intuition
3. Synthetic DiD as a Weighted 2-way FE Specification
4. Details
   - Double Bias Adjustment
   - Weights
   - Advantages/Disadvantages

Two Way FE DiD

Define $D_{it} = 1$ to indicate treatment for unit i in period t and 0 otherwise.

Two Way FE DiD

Define $D_{it} = 1$ to indicate treatment for unit i in period t and 0 otherwise. Assuming no controls for simplicity we can write a 2-way FE panel DiD model as

$$Y_{it} = \alpha + \beta D_{it} + u_i + \tau_t + \eta_{it}$$

Two Way FE DiD

Define $D_{it} = 1$ to indicate treatment for unit i in period t and 0 otherwise. Assuming no controls for simplicity we can write a 2-way FE panel DiD model as

$$Y_{it} = \alpha + \beta D_{it} + u_i + \tau_t + \eta_{it}$$

Q: What does it mean to say this is a "two-way" model?

Two Way FE DiD

Define $D_{it} = 1$ to indicate treatment for unit i in period t and 0 otherwise. Assuming no controls for simplicity we can write a 2-way FE panel DiD model as

$$Y_{it} = \alpha + \beta D_{it} + u_i + \tau_t + \eta_{it}$$

Q: What does it mean to say this is a "two-way" model?
A: The composite error term has a unit fixed effect u_i and a time fixed effect τ_t.

Two Way FE DiD

Define $D_{it} = 1$ to indicate treatment for unit i in period t and 0 otherwise. Assuming no controls for simplicity we can write a 2-way FE panel DiD model as

$$Y_{it} = \alpha + \beta D_{it} + u_i + \tau_t + \eta_{it}$$

Q: What does it mean to say this is a "two-way" model?
A: The composite error term has a unit fixed effect u_i and a time fixed effect τ_t.
The associated estimator is found via

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{DiD}) = \mathit{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2$$

Synthetic Control

Arkhangelsky et al show that S.C. can be written with a quadratic optimization

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{SC}) = \mathit{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - \tau_t)^2 \hat{w}_i^{SC}$$

Synthetic DiD

Arkhangelsky et al ALSO show that SDiD can be written with a quadratic optimization

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{SDiD}) = argmin \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2 \hat{w}_i^{SDiD} \hat{\lambda}_t^{SDiD}$$

Objective Function Similarities

Objective Function Similarities

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{DiD}) = \textit{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2$$

Objective Function Similarities

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{DiD}) = \textit{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2$$

$$(\hat{\alpha}, \hat{\tau}, \hat{\beta}^{S.C.}) = \textit{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - \tau_t)^2 \hat{w}_i^{S.C.}$$

Objective Function Similarities

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{DiD}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2$$

$$(\hat{\alpha}, \hat{\tau}, \hat{\beta}^{S.C.}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - \tau_t)^2 \hat{w}_i^{S.C.}$$

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{SDiD}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2 \hat{w}_i^{SDiD} \hat{\lambda}_t^{SDiD}$$

Objective Function Similarities

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{DiD}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2$$

$$(\hat{\alpha}, \hat{\tau}, \hat{\beta}^{S.C.}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - \tau_t)^2 \hat{w}_i^{S.C.}$$

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{SDiD}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2 \hat{w}_i^{SDiD} \hat{\lambda}_t^{SDiD}$$

Notice that there are two types of differences between the models

Objective Function Similarities

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{DiD}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2$$

$$(\hat{\alpha}, \hat{\tau}, \hat{\beta}^{S.C.}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - \tau_t)^2 \hat{w}_i^{S.C.}$$

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{SDiD}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2 \hat{w}_i^{SDiD} \hat{\lambda}_t^{SDiD}$$

Notice that there are two types of differences between the models

1. Some have fixed effects u_i and/or τ_t

Objective Function Similarities

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{DiD}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2$$

$$(\hat{\alpha}, \hat{\tau}, \hat{\beta}^{S.C.}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - \tau_t)^2 \hat{w}_i^{S.C.}$$

$$(\hat{\alpha}, \hat{u}, \hat{\tau}, \hat{\beta}^{SDiD}) = \underset{}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \alpha - \beta D_{it} - u_i - \tau_t)^2 \hat{w}_i^{SDiD} \hat{\lambda}_t^{SDiD}$$

Notice that there are two types of differences between the models

1. Some have fixed effects u_i and/or τ_t
2. Some have weights w_i and/or λ_t

The Role of FE and Weights

Q: What's the role of the fixed effects?

The Role of FE and Weights

Q: What's the role of the fixed effects?
A: Fixed Effects permit units/time periods to behave individually (via separate "intercepts").

Q: What's the role of the weights?

The Role of FE and Weights

Q: What's the role of the fixed effects?
A: Fixed Effects permit units/time periods to behave individually (via separate "intercepts").

Q: What's the role of the weights?
A: The weights alter the importance of each unit/period.

Comparing the Estimators

	DiD	S.C.	SDiD
u_i-Unit FE	✓		✓
τ_t-Time FE	✓	✓	✓
w_i-Unit Weights		✓	✓
λ_t-Time Weights			✓

Outline

1. Motivation - Case Study
2. SDiD Intuition
3. Synthetic DiD as a Weighted 2-way FE Specification
4. Details
   - Double Bias Adjustment
   - Weights
   - Advantages/Disadvantages

Outline

1. Motivation - Case Study
2. SDiD Intuition
3. Synthetic DiD as a Weighted 2-way FE Specification
4. Details
   - Double Bias Adjustment
   - Weights
   - Advantages/Disadvantages

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Q: Interpret the last two terms on the right hand side.

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Q: Interpret the last two terms on the right hand side.
A:

- $(Y_{1989}^{NC} - Y_{1988}^{NC})$ is a

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Q: Interpret the last two terms on the right hand side.
A:

- $(Y_{1989}^{NC} - Y_{1988}^{NC})$ is a bias correction for

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Q: Interpret the last two terms on the right hand side.
A:

- $(Y_{1989}^{NC} - Y_{1988}^{NC})$ is a bias correction for time trends

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Q: Interpret the last two terms on the right hand side.
A:

- $(Y_{1989}^{NC} - Y_{1988}^{NC})$ is a bias correction for time trends
- $(Y_{1988}^{CA} - Y_{1988}^{NC})$ is a

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Q: Interpret the last two terms on the right hand side.
A:

- $(Y_{1989}^{NC} - Y_{1988}^{NC})$ is a bias correction for time trends
- $(Y_{1988}^{CA} - Y_{1988}^{NC})$ is a bias correction for

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Q: Interpret the last two terms on the right hand side.
A:

- $(Y_{1989}^{NC} - Y_{1988}^{NC})$ is a bias correction for time trends
- $(Y_{1988}^{CA} - Y_{1988}^{NC})$ is a bias correction for state differences

Bias Correction in DiD

Consider a simple 2×2 DiD estimate with California and North Carolina across the Prop99 intervention of 1988 and 1989

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{CA}) - (Y_{1989}^{NC} - Y_{1988}^{NC})$$

Add and subtract Y_{1988}^{NC} and rearrange

$$\widehat{DiD} = (Y_{1989}^{CA} - Y_{1988}^{NC}) - (Y_{1989}^{NC} - Y_{1988}^{NC}) - (Y_{1988}^{CA} - Y_{1988}^{NC})$$

Q: Interpret the last two terms on the right hand side.
A:

- $(Y_{1989}^{NC} - Y_{1988}^{NC})$ is a bias correction for time trends
- $(Y_{1988}^{CA} - Y_{1988}^{NC})$ is a bias correction for state differences
- DiD is attempting to double bias adjust, which is difficult if model is misspecified and/or have small N/T.

Bias Correction in SDiD

SDiD attempts to improve upon this double bias correction with weightings and unit FE

Bias Correction in SDiD

SDiD attempts to improve upon this double bias correction with weightings and unit FE

The SDiD analogues to the three pieces of the DiD estimator are (with some abuse of notation)

$$\bar{Y}^{CA}_{\geq 1989} - \sum_{state}^{-(CA)} \sum_{t}^{<1989} \hat{w}^{SC}_{state} \hat{\lambda}^{SC}_{t} Y^{state}_{<1989}$$

$$\text{Diff. Between Before/After: } \sum_{state}^{-(CA)} \hat{w}^{SC}_{state} \left(Y^{state}_{\geq 1989} - \sum_{t}^{<1989} \hat{\lambda}^{SC}_{t} Y^{state}_{t} \right)$$

$$\text{Diff. Between Treat/Control: } \sum_{t}^{<1989} \hat{\lambda}^{SC}_{t} \left(Y^{CA}_{t} - \sum_{states}^{-(CA)} \hat{w}^{SC}_{state} Y^{state}_{t} \right)$$

Outline

1. Motivation - Case Study
2. SDiD Intuition
3. Synthetic DiD as a Weighted 2-way FE Specification
4. Details
   - Double Bias Adjustment
   - Weights
   - Advantages/Disadvantages

Optimal Weights

Akin to S.C., myriad weights are possible, but authors recommend the following

Optimal Weights

Akin to S.C., myriad weights are possible, but authors recommend the following

- Unit weights for the treated units(s) are equal weighted, sum to 1 and are non-negative

Optimal Weights

Akin to S.C., myriad weights are possible, but authors recommend the following

- Unit weights for the treated units(s) are equal weighted, sum to 1 and are non-negative
- Time weights for the post treatment period(s) are equal weighted, sum to 1 and are non-negative

Optimal Weights

Akin to S.C., myriad weights are possible, but authors recommend the following

- Unit weights for the treated units(s) are equal weighted, sum to 1 and are non-negative
- Time weights for the post treatment period(s) are equal weighted, sum to 1 and are non-negative
- Unit weights for the control group are set via

$$\hat{w}^{SC} = argmin \sum_{t}^{<1989} \left(\sum_{state}^{-(CA)} w_{state} Y_t^{state} - Y_t^{CA})^2 \right)$$

Optimal Weights

Akin to S.C., myriad weights are possible, but authors recommend the following

- Unit weights for the treated units(s) are equal weighted, sum to 1 and are non-negative
- Time weights for the post treatment period(s) are equal weighted, sum to 1 and are non-negative
- Unit weights for the control group are set via

$$\hat{w}^{SC} = argmin \sum_{t}^{<1989} \left(\sum_{state}^{-(CA)} w_{state} Y_t^{state} - Y_t^{CA})^2 \right)$$

- Time weights for the pre treatment periods are set via

$$\hat{\lambda}^{SC} = argmin \sum_{state}^{-(CA)} \left(\sum_{t}^{<1989} \lambda_t Y_t^{state} - Y_{1989}^{state} \right)^2$$

Outline

1. Motivation - Case Study
2. SDiD Intuition
3. Synthetic DiD as a Weighted 2-way FE Specification
4. Details
   - Double Bias Adjustment
   - Weights
   - Advantages/Disadvantages

Advantages and Disadvantages of SDiD

Relative to DiD and Synthetic Control

Advantages and Disadvantages of SDiD

Relative to DiD and Synthetic Control

Advantages

- Parallel trends required only after unit and time reweighting

Advantages and Disadvantages of SDiD

Relative to DiD and Synthetic Control

Advantages

- Parallel trends required only after unit and time reweighting
- Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes

Advantages and Disadvantages of SDiD

Relative to DiD and Synthetic Control

Advantages

- Parallel trends required only after unit and time reweighting
- Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes
- Handles multiple treated and control units naturally

Advantages and Disadvantages of SDiD

Relative to DiD and Synthetic Control

Advantages

- Parallel trends required only after unit and time reweighting
- Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes
- Handles multiple treated and control units naturally
- Reduces bias when treated and control units have different pre-treatment trends

Advantages and Disadvantages of SDiD

Relative to DiD and Synthetic Control

Advantages

- Parallel trends required only after unit and time reweighting
- Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes
- Handles multiple treated and control units naturally
- Reduces bias when treated and control units have different pre-treatment trends

Disadvantages

- Requires a sufficiently long pre-treatment period

Advantages and Disadvantages of SDiD

Relative to DiD and Synthetic Control

Advantages

- Parallel trends required only after unit and time reweighting
- Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes
- Handles multiple treated and control units naturally
- Reduces bias when treated and control units have different pre-treatment trends

Disadvantages

- Requires a sufficiently long pre-treatment period
- Involves nontrivial optimization and less transparent inference than DiD

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences
- Similar to Synthetic Control
  - Uses unit and time weights to construct a synthetic comparison

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences
- Similar to Synthetic Control
  - Uses unit and time weights to construct a synthetic comparison
  - Balances pre-treatment outcomes to improve credibility

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences
- Similar to Synthetic Control
  - Uses unit and time weights to construct a synthetic comparison
  - Balances pre-treatment outcomes to improve credibility
- Synthetic DiD
  - Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes

Synthetic DiD – Summary

- Most useful in policy or natural experiment settings with multiple treated units
- Similar to DiD
  - Identification relies on plausibly exogenous treatment timing
  - Requires panel data
  - Computes counterfactual via weighted differences-in-differences
- Similar to Synthetic Control
  - Uses unit and time weights to construct a synthetic comparison
  - Balances pre-treatment outcomes to improve credibility
- Synthetic DiD
  - Relaxes strict parallel trends by enforcing balance in pre-treatment outcomes
  - Combines weighting and differencing to reduce bias