Synthetic Control (Sam's Version)

ILRST/INFO/STSCI 3900: Causal Inference

12 Nov 2024

Learning goals for today

At the end of class, you will be able to:

1. Explain the intuition behind synthetic control
2. Understand how synthetic control relates to other causal inference methods

Logistics

- This week, read Ch 10 of The Causal Inference Mixtape
- Problem Set 5 peer reviews due Nov 15
- Task 3 and 4 Check-in (assigned Nov 5, due Nov 17)
- In class project check-ins next week
- Problem Set 6 (assigned Nov 14, due Nov 21)

What is the effect of personal events on google searches?

What is the effect of personal events on google searches?

- Who was the last celebrity you googled?
- What do you usually google about celebrities?

NFL Top 100

Before the start of each season, all current NFL players vote on the top players

(1) Mahomes

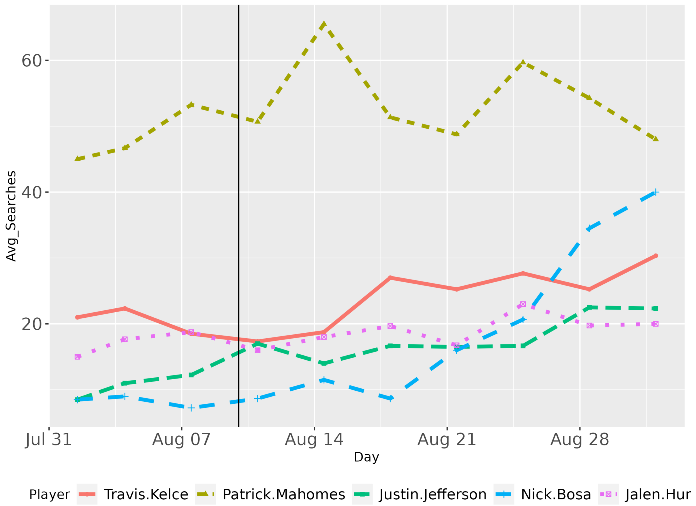
(2) Jefferson

(3) Hurts

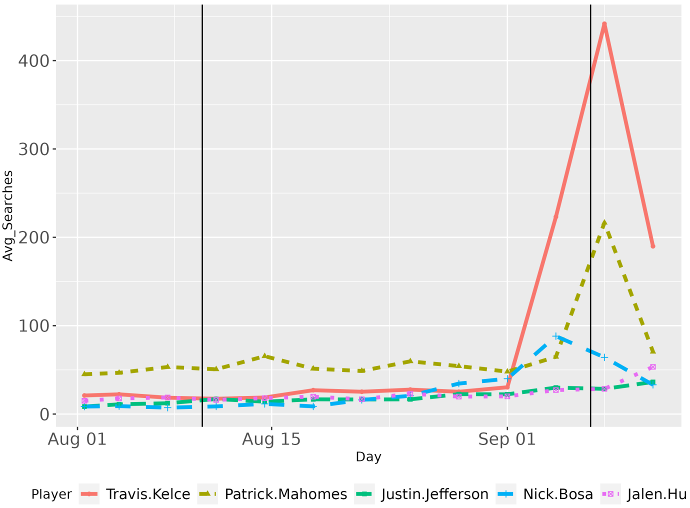
(4) Bosa

(5) Kelce

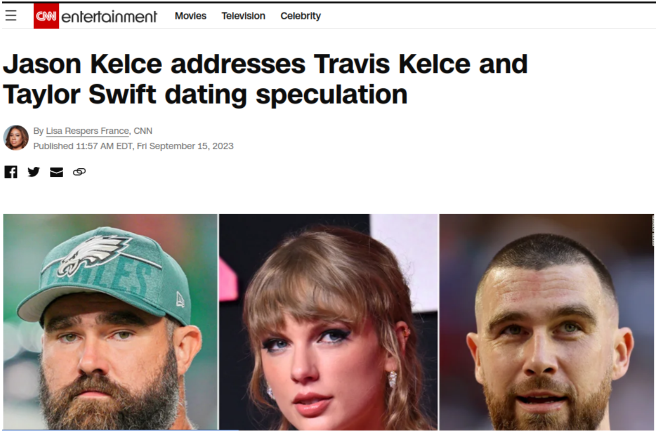
Google searches for NFL players



Google searches for NFL players



Google searches for NFL players

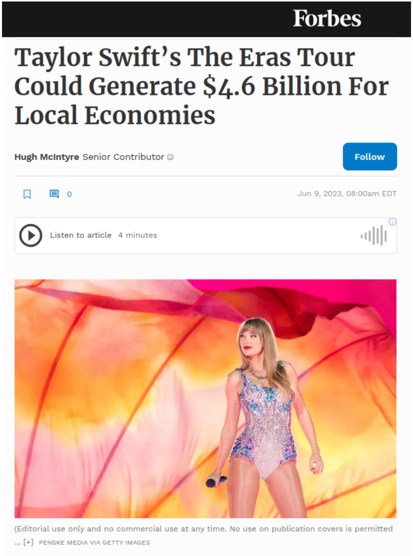


CNN entertainment Movies Television Celebrity

Jason Kelce addresses Travis Kelce and Taylor Swift dating speculation

By Lisa Respers France, CNN
Published 11:57 AM EDT, Fri September 15, 2023

Google searches for NFL players



Forbes

Taylor Swift's The Eras Tour Could Generate $4.6 Billion For Local Economies

Hugh McIntyre Senior Contributor

Follow

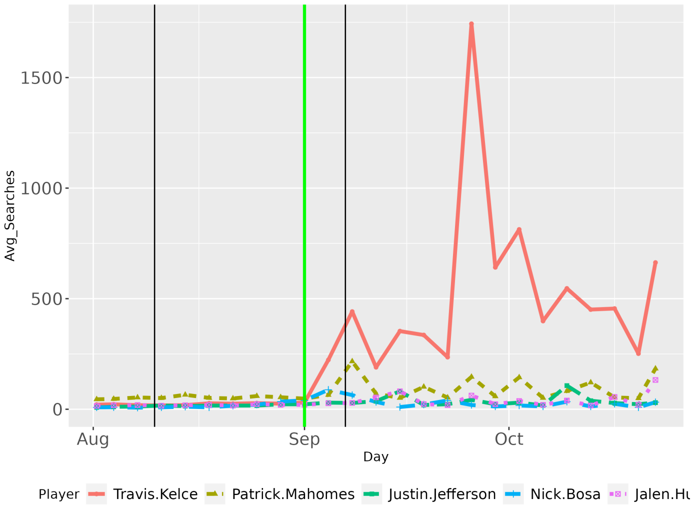
0

Jun 9, 2023, 08:00am EDT

Listen to article 4 minutes

(Editorial use only and no commercial use at any time. No use on publication covers is permitted ... [+] PENSKE MEDIA VIA GETTY IMAGES

Google searches for NFL players



Google searches for NFL players

What is the causal effect of dating Taylor Swift on google searches?

Google searches for NFL players

What is the causal effect of dating Taylor Swift on google searches?

- Causal effect may vary over time

Google searches for NFL players

What is the causal effect of dating Taylor Swift on google searches?

- Causal effect may vary over time
- Causal effect at time t

$$\tau_{t,Kelce} = Y_{t,Kelce}^{\text{Swift}} - Y_{t,Kelce}^{\text{NoSwift}}$$

- For notation, let T_0 denote the time that the treatment occurs
- We observe $Y_{t,Kelce}^{\text{S}}$ for $t > T_0$ and $Y_{t,Kelce}^{\text{NS}}$ for $t < T_0$, but not at the same time!

Google searches for NFL players

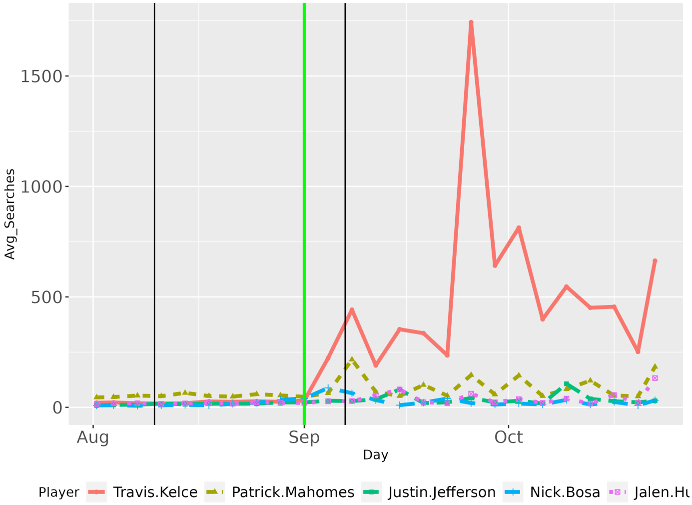
What is the causal effect of dating Taylor Swift on google searches?

- Causal effect may vary over time
- Causal effect at time t

$$\tau_{t,Kelce} = Y_{t,Kelce}^{\text{Swift}} - Y_{t,Kelce}^{\text{NoSwift}}$$

- For notation, let T_0 denote the time that the treatment occurs
- We observe $Y_{t,Kelce}^{\text{S}}$ for $t > T_0$ and $Y_{t,Kelce}^{\text{NS}}$ for $t < T_0$, but not at the same time!
- Blank space in our data

Google searches for NFL players



- Kelce and Mahomes play for the same team
- Kelce and Jefferson play similar positions
- Kelce and Bosa both went to college in Ohio

Synthetic Control

Synthetic Control

- Google searches for NFL players are affected by many things that change over time
- Trend prior in pre-season may not be a good trend for during season
- Estimating the effect far away from the treatment seems iffy

Synthetic Control

- Google searches for NFL players are affected by many things that change over time
- Trend prior in pre-season may not be a good trend for during season
- Estimating the effect far away from the treatment seems iffy
- Kelce doesn't quite match any individual player exactly, but is similar to other players in different ways

Synthetic Control

Synthetic Control

- We don't observe $Y^{\mathrm{NS}}_{t,Kelce}$ after T_0
- We do observe $Y^{\mathrm{NS}}_{t,Mahomes}$, $Y^{\mathrm{NS}}_{t,Hurts}$, etc.

Synthetic Control

- We don't observe $Y_{t,Kelce}^{\mathrm{NS}}$ after T_0
- We do observe $Y_{t,Mahomes}^{\mathrm{NS}}$, $Y_{t,Hurts}^{\mathrm{NS}}$, etc.
- Create a "synthetic" version of of Kelce by weighting other players

$$Y_{t,Kelce}^{\mathrm{NS}} \approx w_1 Y_{t,Mahomes}^{\mathrm{NS}} + w_2 Y_{t,Hurts}^{\mathrm{NS}} + w_3 Y_{t,Bosa}^{\mathrm{NS}} + w_4 Y_{t,Jefferson}^{\mathrm{NS}}$$

where $w_j \geq 0$ and $\sum w_j = 1$

Synthetic Control

- We don't observe $Y^{\mathrm{NS}}_{t,Kelce}$ after T_0
- We do observe $Y^{\mathrm{NS}}_{t,Mahomes}$, $Y^{\mathrm{NS}}_{t,Hurts}$, etc.
- Create a "synthetic" version of of Kelce by weighting other players

  $$Y^{\mathrm{NS}}_{t,Kelce} \approx w_1 Y^{\mathrm{NS}}_{t,Mahomes} + w_2 Y^{\mathrm{NS}}_{t,Hurts} + w_3 Y^{\mathrm{NS}}_{t,Bosa} + w_4 Y^{\mathrm{NS}}_{t,Jefferson}$$

  where $w_j \geq 0$ and $\sum w_j = 1$
- So perhaps, Synthetic Kelce is
  - 50% Patrick Mahomes
  - 25% Justin Jefferson
  - 25% Nick Bosa
  - 0% Jalen Hurts

Synthetic Control

- Estimate counterfactual Travis Kelce $Y^{NS}_{t,Kelce}$ by using Synthetic Kelce

$$Y^{NS}_{t,Synthetic} = .5 \times Y_{t,Mahomes} + .25 \times Y_{t,Bosa} + .25 \times Y_{t,Jefferson}$$

Synthetic Control

- Estimate counterfactual Travis Kelce $Y^{\mathrm{NS}}_{t,Kelce}$ by using Synthetic Kelce

$$Y^{\mathrm{NS}}_{t,Synthetic} = .5 \times Y_{t,Mahomes} + .25 \times Y_{t,Bosa} + .25 \times Y_{t,Jefferson}$$

- Post-treatment at time t, use difference between observed Kelce and Synthetic Kelce as estimate of the causal effect

$$\hat{\tau}_t = Y_{t,Kelce} - Y^{\mathrm{NS}}_{t,Synthetic}$$

Synthetic Control

- Estimate counterfactual Travis Kelce $Y^{\text{NS}}_{t,Kelce}$ by using Synthetic Kelce

$$Y^{\text{NS}}_{t,Synthetic} = .5 \times Y_{t,Mahomes} + .25 \times Y_{t,Bosa} + .25 \times Y_{t,Jefferson}$$

- Post-treatment at time t, use difference between observed Kelce and Synthetic Kelce as estimate of the causal effect

$$\hat{\tau}_t = Y_{t,Kelce} - Y^{\text{NS}}_{t,Synthetic}$$

- Straightforward approach boils down to picking "good" weights

Picking Weights

Picking Weights

- We want "Synthetic Kelce" to predict $Y^{NS}_{t,Kelce}$
- We observe $Y_{t,Kelce} = Y^{NS}_{t,Kelce}$ before treatment when $t < T_0$

Picking Weights

- We want "Synthetic Kelce" to predict $Y^{NS}_{t,Kelce}$
- We observe $Y_{t,Kelce} = Y^{NS}_{t,Kelce}$ before treatment when $t < T_0$
- Select weights to minimize

$$\sum_{t<T_0} \left(Y_{t,Kelce} - \underbrace{w_1 Y_{t,M} + w_2 Y_{t,H} + w_3 Y_{t,B} + w_4 Y_{t,J}}_{Y_{t,Synthetic}} \right)^2$$

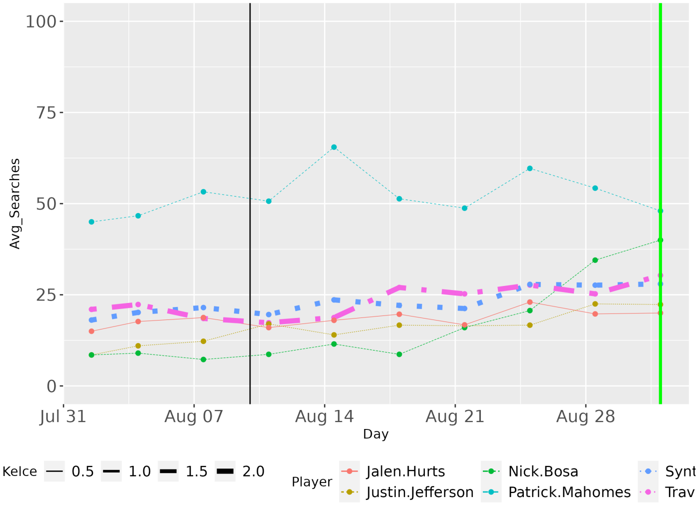
Picking Weights

- We want "Synthetic Kelce" to predict $Y^{NS}_{t,Kelce}$
- We observe $Y_{t,Kelce} = Y^{NS}_{t,Kelce}$ before treatment when $t < T_0$
- Select weights to minimize

$$\sum_{t<T_0} \left(Y_{t,Kelce} - \underbrace{w_1 Y_{t,M} + w_2 Y_{t,H} + w_3 Y_{t,B} + w_4 Y_{t,J}}_{Y_{t,Synthetic}} \right)^2$$

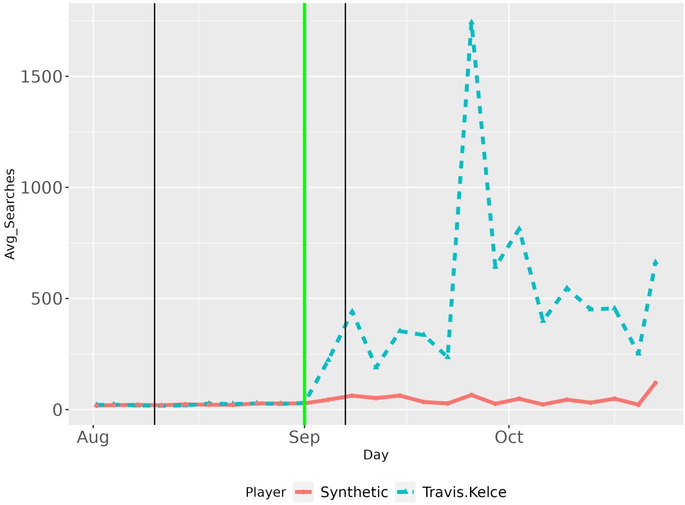
- Can also be selected to minimize discrepancy between other pre-treatment covariates (preview of discussion)

Synthetic Control



Synthetic Kelce $= .14 \times$ Mahomes $+ .20 \times$ Bosa $+ .66 \times$ Hurts

Synthetic Control



$$\text{Synthetic Kelce} = .14 \times \text{Mahomes} + .20 \times \text{Bosa} + .66 \times \text{Hurts}$$

Synthetic Control

Pros:

- Counterfactual prediction is easy to understand and explain

Synthetic Control

Pros:

- Counterfactual prediction is easy to understand and explain
- Works well when there are not many units and a single good match may be difficult to find

Synthetic Control

Pros:

- Counterfactual prediction is easy to understand and explain
- Works well when there are not many units and a single good match may be difficult to find
- Allows for extrapolation away from treatment time

Synthetic Control

Pros:

- Counterfactual prediction is easy to understand and explain
- Works well when there are not many units and a single good match may be difficult to find
- Allows for extrapolation away from treatment time

Cons:

- Requires lots of pre-treatment data to pick good weights

Synthetic Control

Pros:

- Counterfactual prediction is easy to understand and explain
- Works well when there are not many units and a single good match may be difficult to find
- Allows for extrapolation away from treatment time

Cons:

- Requires lots of pre-treatment data to pick good weights

Synthetic Control

Synthetic control "is arguably the most important innovation in the policy evaluation literature in the last 15 years" (Athey and Imbens 2017)

Synthetic Control

Synthetic control "is arguably the most important innovation in the policy evaluation literature in the last 15 years" (Athey and Imbens 2017)

Examples:

- What is the effect of political instability on the economy in Basque country in the 1960-70s?
  (Abadie and Gardeazabal 2003)
- What is the effect of a cigarette tax on smoking in California?
  (Abadie, Diamond, Hainmueller 2010)

Synthetic control and Matching

Synthetic control and Matching

In some ways, synthetic control can be seen as a specific form of matching

- Predict unobserved potential outcome using observed outcome of “similar” units
- Can choose “matches” (i.e., weights) to match untreated outcomes (of eventually treated unit)

Synthetic control and Matching

In some ways, synthetic control can be seen as a specific form of matching

- Predict unobserved potential outcome using observed outcome of "similar" units
- Can choose "matches" (i.e., weights) to match untreated outcomes (of eventually treated unit)
- Synthetic control differs in how weights are chosen
- Data across time (longitudinal) so we also observed untreated outcomes of (eventually) treated unit

Synthetic control and Matching

In some ways, synthetic control can be seen as a specific form of matching

- Predict unobserved potential outcome using observed outcome of “similar” units
- Can choose “matches” (i.e., weights) to match untreated outcomes (of eventually treated unit)
- Synthetic control differs in how weights are chosen
- Data across time (longitudinal) so we also observed untreated outcomes of (eventually) treated unit
- Can directly match to minimize pre-treatment fit

Synthetic control and Difference and Difference

Synthetic control and Difference and Difference

- Both have observations pre and post treatment
- Diff-in-Diff requires parallel trends assumption

Synthetic control and Difference and Difference

- Both have observations pre and post treatment
- Diff-in-Diff requires parallel trends assumption
- In synthetic control, we have a similar assumption, but parallel trends holds for synthetic unit
- Generally, Diff-in-Diff has fixed set of comparison units using prior knowledge (i.e., NJ vs PA)

Synthetic control and Difference and Difference

- Both have observations pre and post treatment
- Diff-in-Diff requires parallel trends assumption
- In synthetic control, we have a similar assumption, but parallel trends holds for synthetic unit
- Generally, Diff-in-Diff has fixed set of comparison units using prior knowledge (i.e., NJ vs PA)
- Synthetic control, we can start with a large "donor pool" and select weights using data

Learning goals for today

At the end of class, you will be able to:

1. Explain the intuition behind synthetic control
2. Understand how synthetic control relates to other causal inference methods