

Using Quasi-Experimental Methodologies to Analyze Pitcher Performance Post-Tommy John Surgery



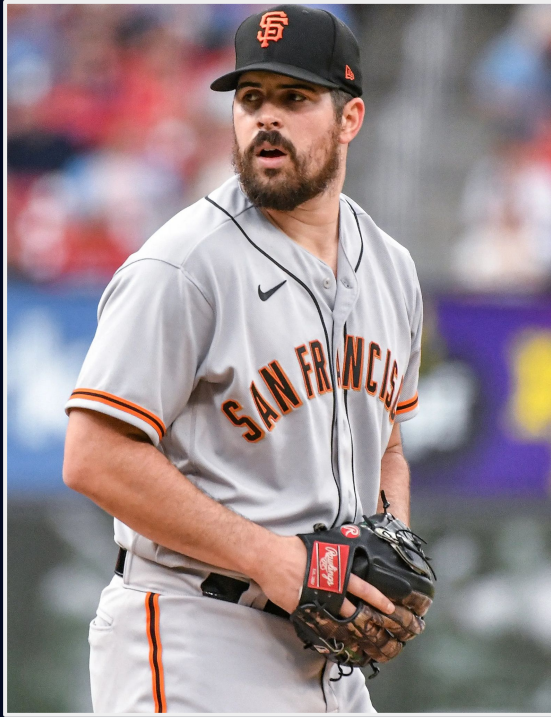
Atul Venkatesh
Dartmouth Class of 2027



01.

Setting the Scene





“I don’t think [Tommy John Surgery] stalled my career, I think it helped my career ’cause I got to work on some things, like mechanically on things throughout my delivery that got me to where I am now”.

Carlos Rodón

A closer look

	ERA	WHIP	FIP	SO9
2017	4.15	1.370	4.69	9.9
2018	4.18	1.260	4.95	6.7

Tommy John Surgery Recovery

A closer look

	ERA	WHIP	FIP	SO9
2017	4.15	1.370	4.69	9.9
2018	4.18	1.260	4.95	6.7
Tommy John Surgery Recovery				
2021	2.37	0.957	2.65	12.6
2022	2.88	1.028	2.25	12.0

Was Tommy John Surgery the reason why Carlos Rodón improved so drastically? How would he have performed if he never needed surgery?

02.

Methods



A biomechanical perspective

The unique and powerful characteristics of a pitching motion puts an incredible amount of stress on the elbow and shoulder

- The goal of **Tommy John Surgery** is to reconstruct the ulnar collateral ligament (UCL) and restore elbow/joint stability
- Recovery **retrains movement patterns** that can correct inefficient biomechanical habits
- Recovery post-Tommy John Surgery involves **understanding and recorrecting potential inadequacies** with the rest of the body

Why Quasi-experimental methodologies?

How do we estimate the effect of Tommy John Surgery?

Why random assignment doesn't work:

- UCL injuries are **uncontrollable**
- Surgery occurs as a response to injury severity not due to experimental design

Instead, we opt for a Quasi-experimental design:

- **Causal effect estimation** is observed by constructing a *counterfactual comparison group* to approximate what would have happened absent treatment
- **Core assumptions** that allow for comparison between treated and counterfactual groups

Pipeline

Compile a dataset
of *treated* players

*Players who
underwent Tommy
John Surgery

Obtain
pre-treatment and
post-treatment
periods for each
player

*Last three full
seasons before
surgery and first
three seasons after
surgery

Create a
customized *donor*
pool for each player

*Players who were
fully healthy during
the entirety of the
pre-treatment and
post-treatment
periods

Apply
quasi-experimental
methodology to the
donor pool

*Synthetic Control
Method or Difference
in Differences

Compute the
Average Treatment
Effect

*Look at the change in
Fielding Independent
Pitching (FIP) and
Strikeouts Per 9 (SO9)

Method 1: Difference-in-Differences (DiD)

Is there a significant change in a pitcher's performance after undergoing Tommy John Surgery when compared to the population of healthy pitchers in that same timeframe?

$ATE = (\text{Post-Pre treatment for TJ pitchers}) - (\text{Post-Pre treatment for donor pool})$

*Statistical significance determined through **placebo testing***

Method 2: Synthetic Control Method (SCM)

How does a pitcher's performance post-Tommy John surgery differ from a *synthetic control* built out of a weighted combination of healthy players that matches the pitcher's performance pre-Tommy John surgery?

$ATE = \text{Pitcher post-treatment} - \text{Synthetic control post-treatment}$

Statistical significance determined through **placebo testing**

Synthetic Control - An example



Matching Carlos Rodón's Strikeouts Per 9

Synthetic Control - An example



Matching Carlos Rodón's strikeouts per 9

Synthetic Control - An example



Matching Carlos Rodón's Strikeouts Per 9



Michael Lorenzen
76.14%



José Quintana
8.36%

Craig Kimbrel
0.15%

Kenley Jansen
15.35%

A close pre-surgery fit

During the pre-surgery period, *Synthetic Rodón's* strikeouts per 9 should match his **actual** strikeouts per 9

*Synthetic Rodón approximates Actual Rodón's performance as if he **never** had surgery.*

Pre-surgery season	Synthetic Rodón	Actual Rodón
2016	9.31	9.2
2017	9.69	9.9
2018	6.86	6.7

Important numbers

49

Treated
players

13,596

Healthy pitching
seasons

9-105

Donor pool
size range

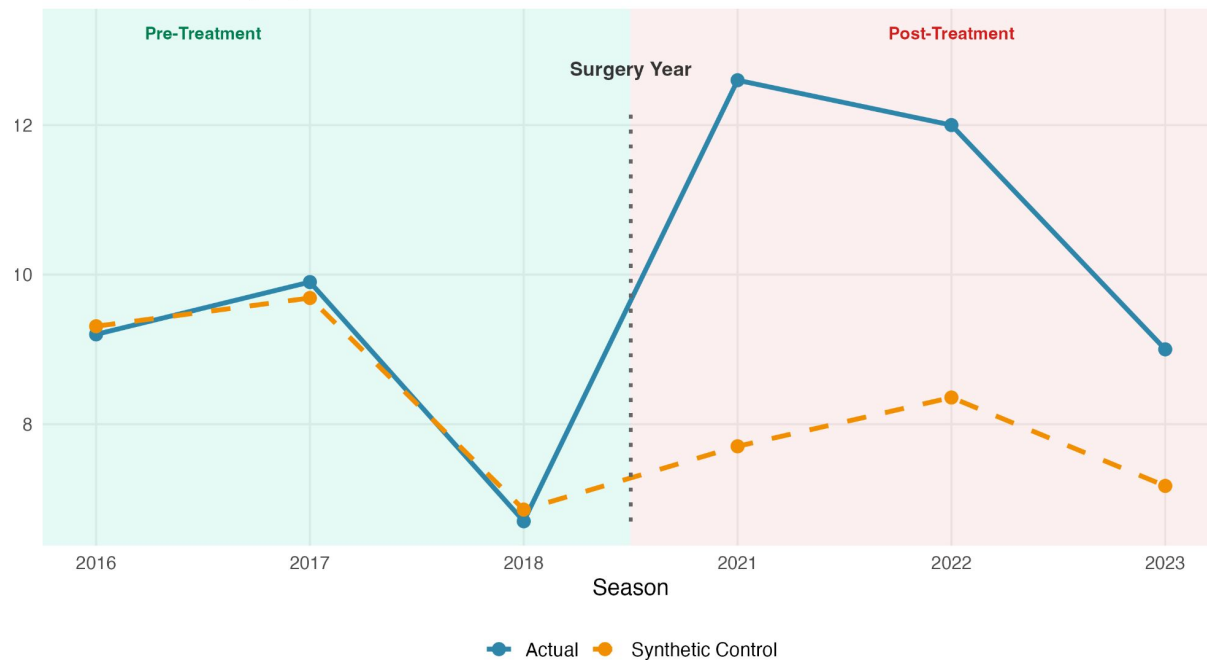
03.

Results



Analyzing Rodón's performance - SO9

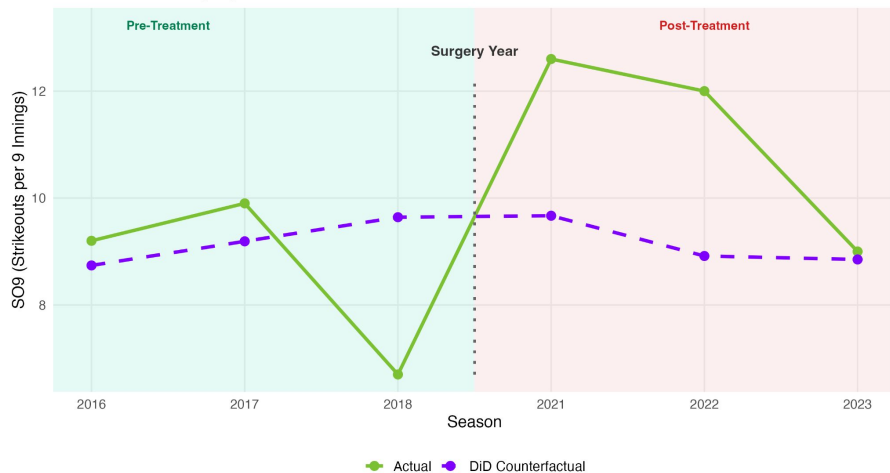
Comparing Rodón's **strikeouts per 9** to his *synthetic control*



Placebo testing for Rodón

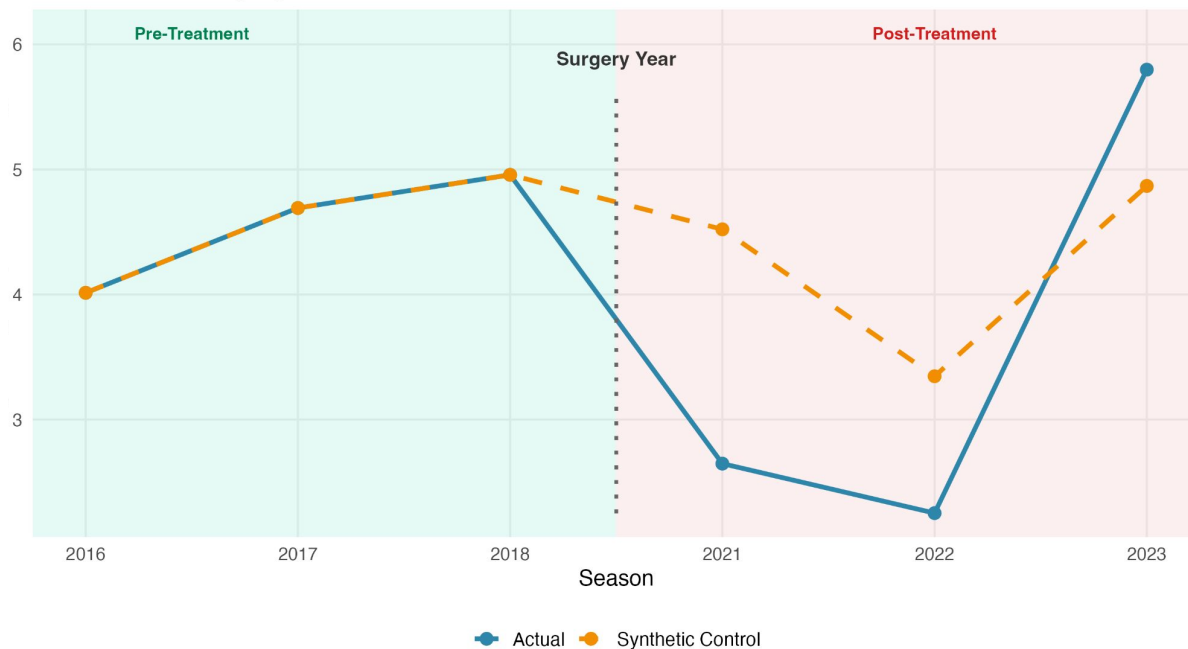
Across all 51 of Rodón's donors, **no pitcher** observed a greater change in strikeouts per 9 than Rodón himself

Difference in Differences comparison



Rodón's FIP

Comparing Rodón's **Fielding Independent Pitching** to his *synthetic control*



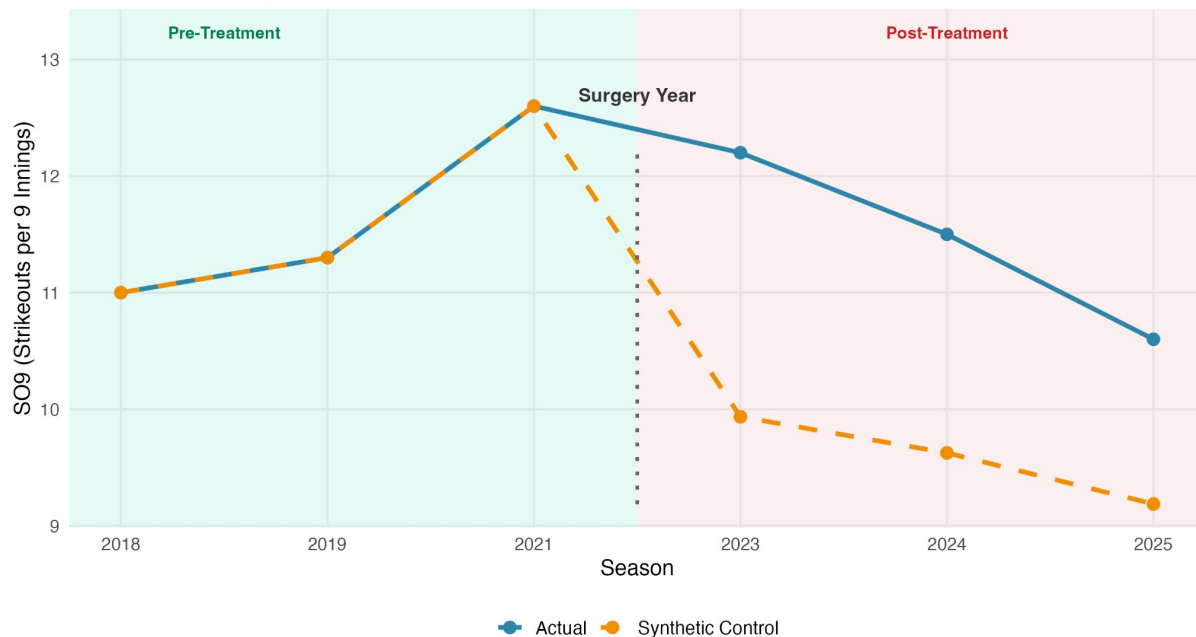
Findings for Rodón

- Tommy John was indeed a **contributing factor** to Rodón's drastic improvement post-surgery
- Absent needing surgery, Rodón would have still improved as a pitcher, but **not nearly** at the observed rate



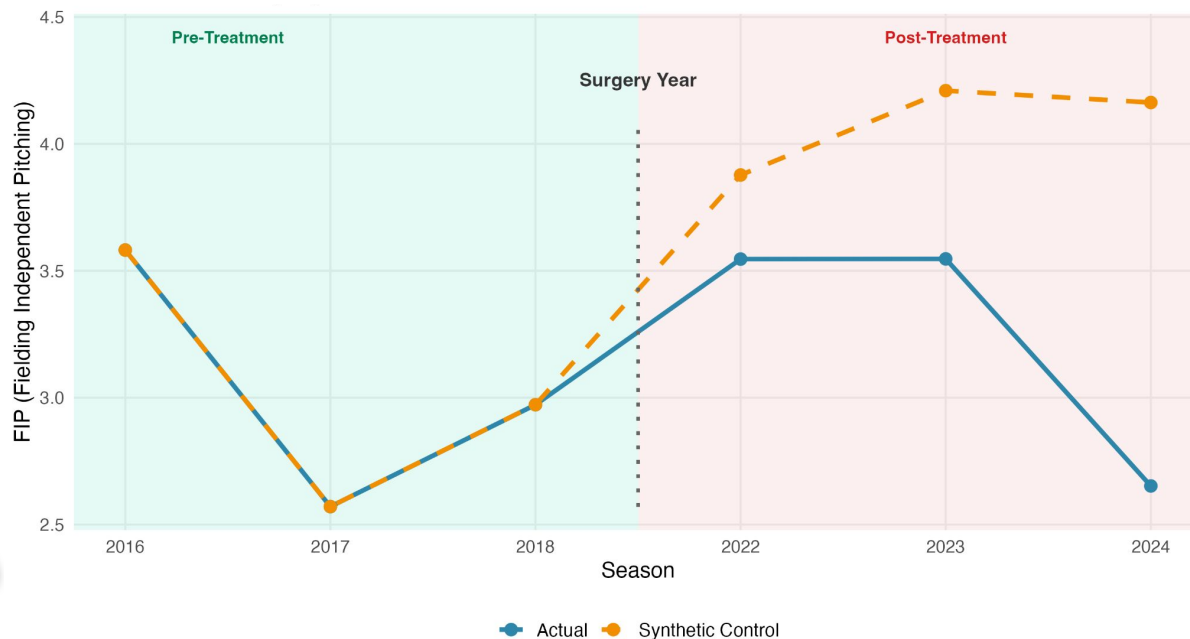
Other notable trends - Tyler Glasnow

Despite recording fewer strikeouts per 9 compared to pre-treatment, Glasnow **still outperformed** his synthetic control ($p = 0.1$)



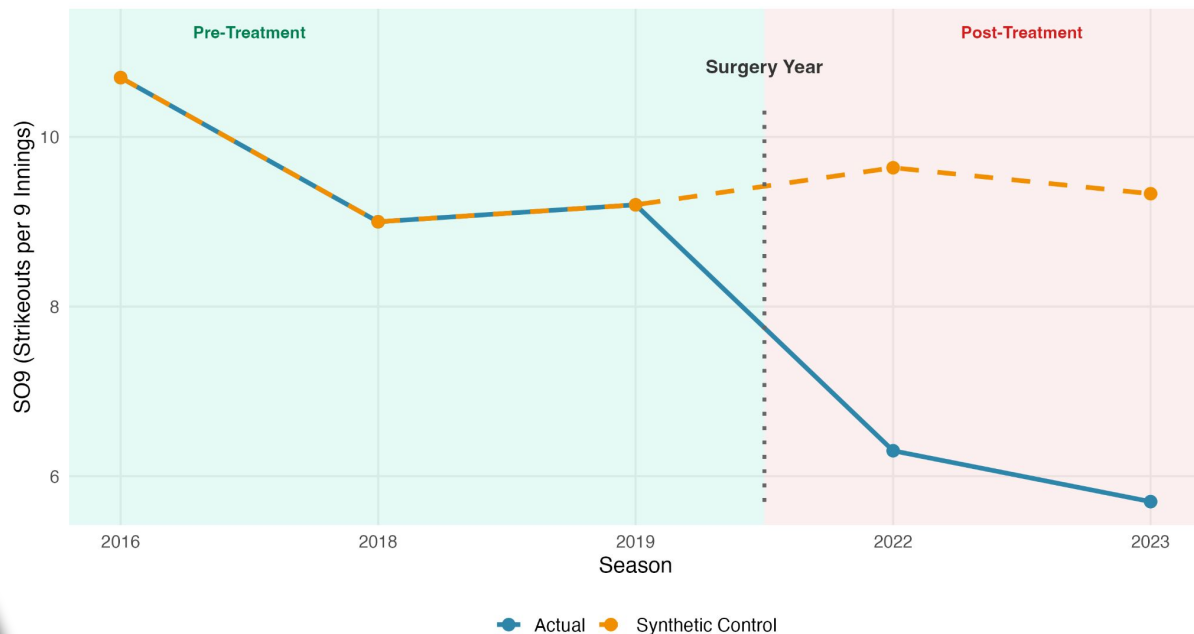
Other notable trends - David Robertson

Robertson's yearly FIP treatment effect was relatively constant until he had a fantastic 2024 in year 3 post-surgery ($p = 0.13$).



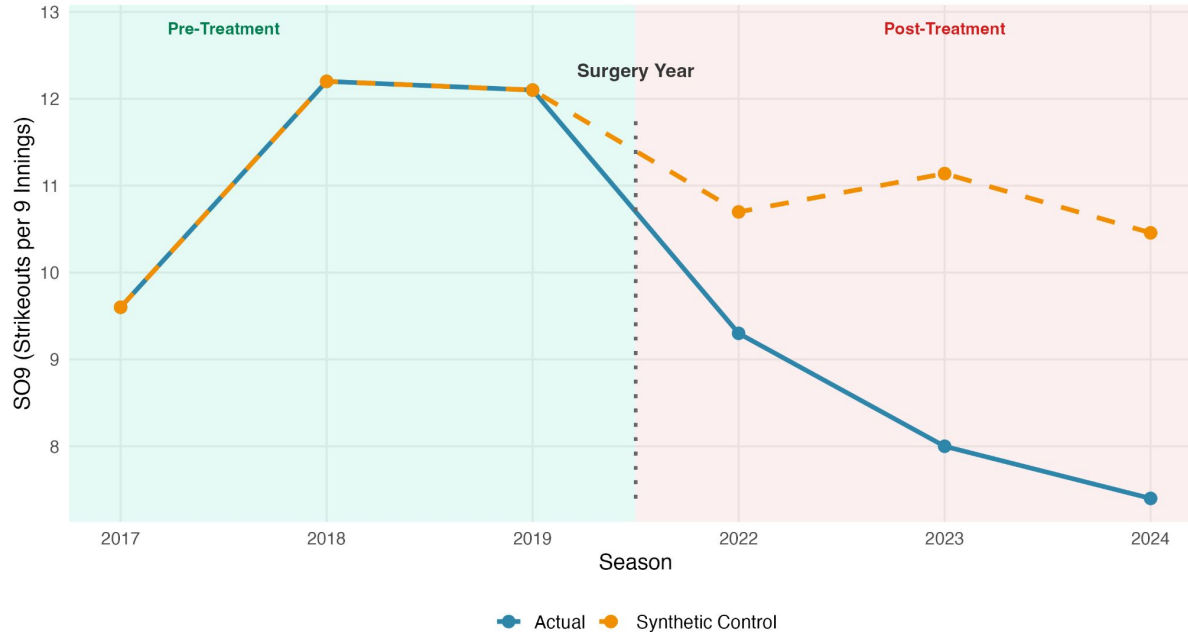
Other notable trends - Noah Syndergaard

Despite his performance in 2016, 2018, and 2019, Syndergaard hasn't been able to reach pre-surgery levels ($p = 0.064$).



Other notable trends - Justin Verlander

While Verlander's FIP, especially the year after surgery, was still comparable to pre-surgery levels, his strikeouts per 9 have taken a steep decline ($p = 0.064$).



04.

Final words



A three-year post-treatment effect

One year annual treatment effect (among all treated pitchers):

Negative SO9 indicates a lower SO9 than synthetic control. Positive FIP indicates a higher FIP than synthetic control

	Mean SO9	Median SO9	Mean FIP	Median FIP
Year 1	-0.879	-0.931	0.366	0.455

A three-year post-treatment effect

Holistic **two** year annual treatment effect (among all treated pitchers):

It gets worse before it gets better

	Mean SO9	Median SO9	Mean FIP	Median FIP
Year 1	-0.879	-0.931	0.366	0.455
Year 2	-0.988	-1.025	0.579	0.528

A three-year post-treatment effect

Holistic **three** year annual treatment effect (among all treated pitchers):

Back to pre-treatment levels

	Mean SO9	Median SO9	Mean FIP	Median FIP
Year 1	-0.879	-0.931	0.366	0.455
Year 2	-0.988	-1.025	0.579	0.528
Year 3	-0.149	-0.201	-0.006	0.020

Other considerations

- It is difficult to model pitcher performance from year to year
 - Pitching performance is inherently volatile
 - The sample size of healthy seasons are normally smaller than a hitter
- Confounding factors impede us from crediting 100% of a pitcher's performance post-treatment to Tommy John surgery
- Survivorship bias

However, by taking these limitations into account, analyzing stable statistics, and conducting placebo testing, we can extract insights that reveal more about a pitcher's recovery process

Acknowledgements

- This project was carried out as a part of Dartmouth Presidential Scholars program. Thank you to Dartmouth for the opportunity to conduct this project.
- Thank you to Professor Michael Herron for all of his advice and assistance throughout the project
- Thank you to SABR for giving me the opportunity to present my research.

This project was inspired by a presentation given by Professor Lee Kennedy-Shaffer at the 2025 Connecticut Sports Analytics Conference on using SCM/DiD to understand the effect of MLB's ban on infield shifts.

Thank you to Jon Roegele (@MLBPlayerAnalys) for the extensive Tommy John surgery dataset.