ANALYZING NFL DATA IS CHALLENGING, BUT PLAYER TRACKING DATA IS HERE TO HELP

Michael Lopez
November 2022

Overview

1. Brief look back

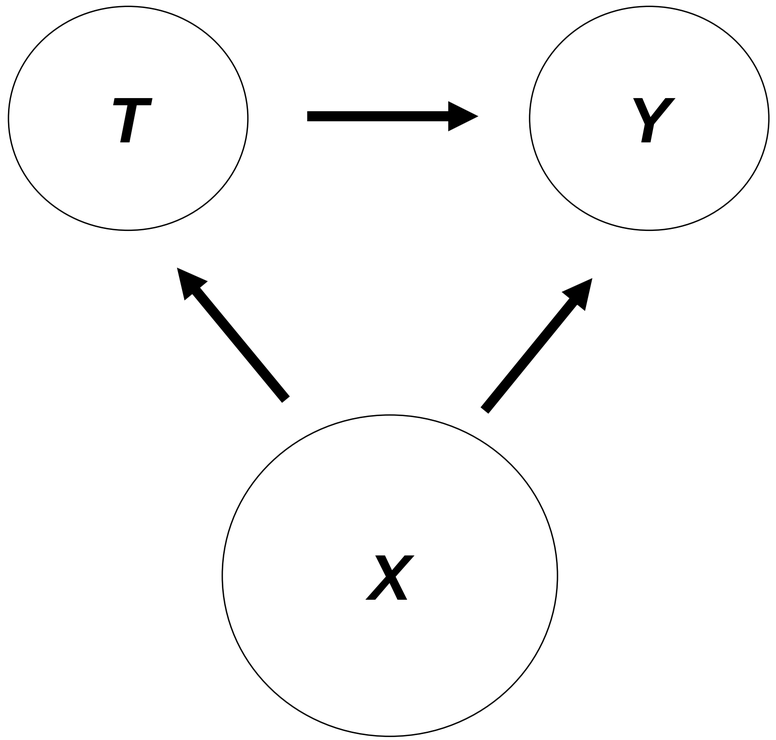
2. Limitations of historical analysis

3. How player tracking can help (with an example)

4. How you can get involved (Big Data Bowl)

5. Analytics at the NFL league office

Sports data is observational



We can’t estimate effect of *T* on *Y* without *X*

Previous learnings

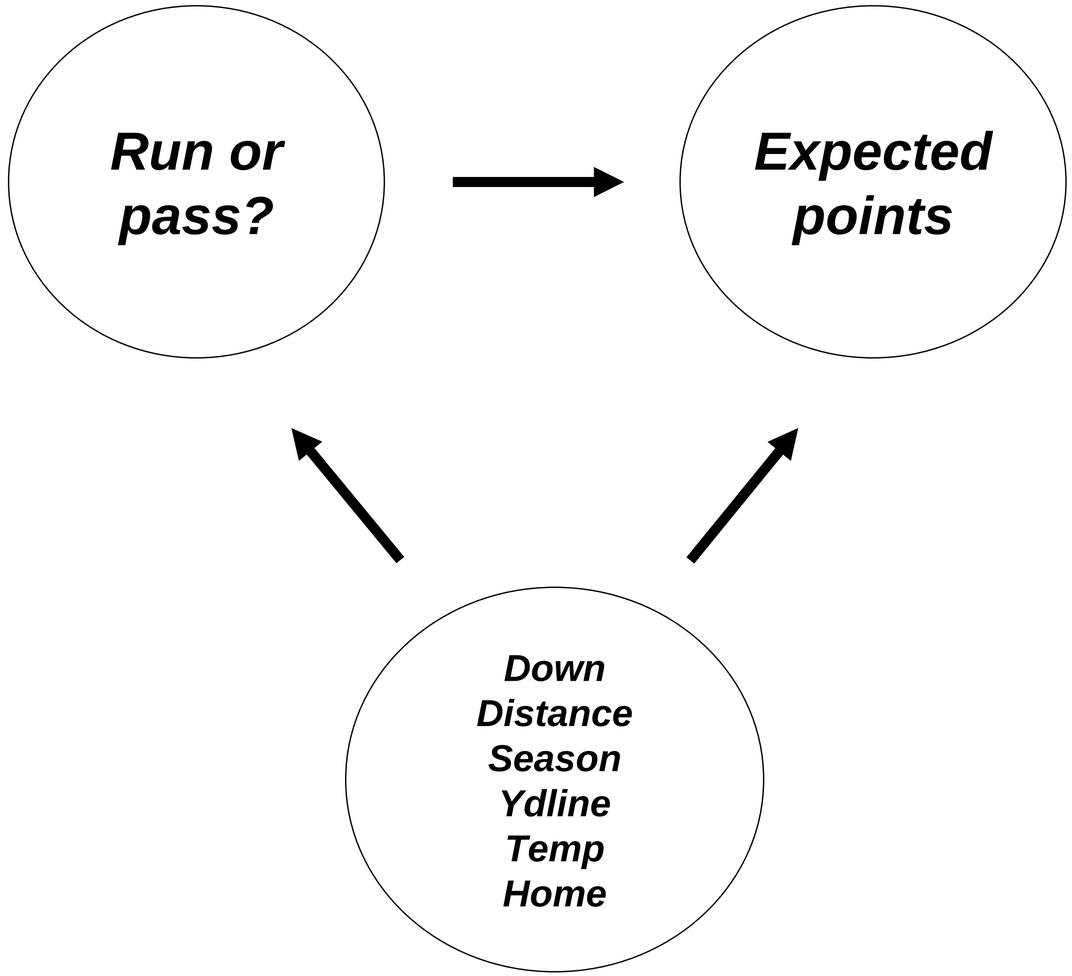
Kovash/Levitt (2009) on run/pass

Professionals Do Not Play Minimax: Evidence from Major League Baseball and the National Football League

Kenneth Kovash, Steven D. Levitt

NBER Working Paper No. 15347
Issued in September 2009
NBER Program(s): Law and Economics, Public Economics

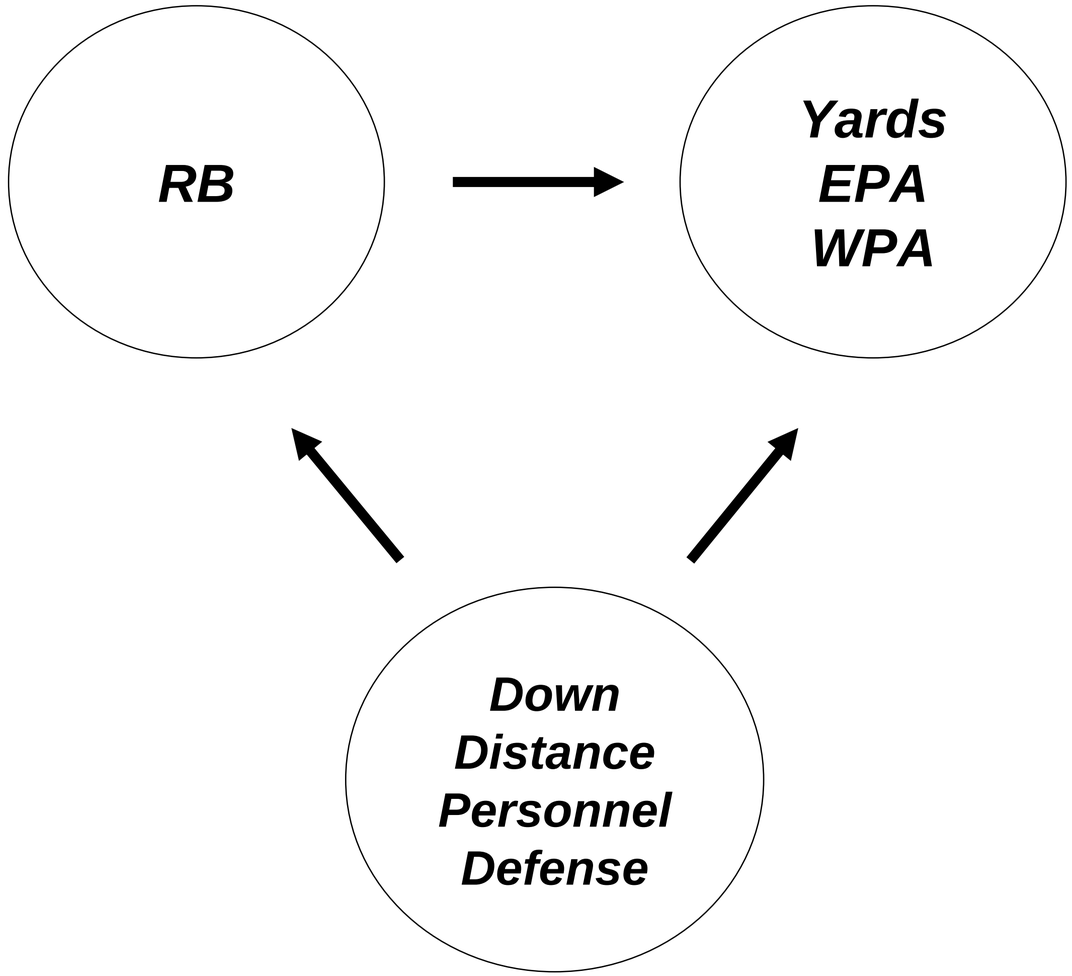
Game theory makes strong predictions about how individuals should behave in two player, zero sum games. When players follow a mixed strategy, equilibrium payoffs should be equalized across actions, and choices should be serially uncorrelated. Laboratory experiments have generated large and systematic deviations from the minimax predictions. Data gleaned from real-world settings have been more consistent with minimax, but these latter studies have often been based on small samples with low power to reject. In this paper, we explore minimax play in two high stakes, real world settings that are data rich: choice of pitch type in Major League Baseball and whether to run or pass in the National Football League. We observe more than three million pitches in baseball and 125,000 play choices for football. We find systematic deviations from minimax play in both data sets. Pitchers appear to throw too many fastballs; football teams pass less than they should. In both sports, there is negative serial correlation in play calling. Back of the envelope calculations suggest that correcting these decision making errors could be worth as many as two additional victories a year to a Major League Baseball franchise, and more than a half win per season for a professional football team.



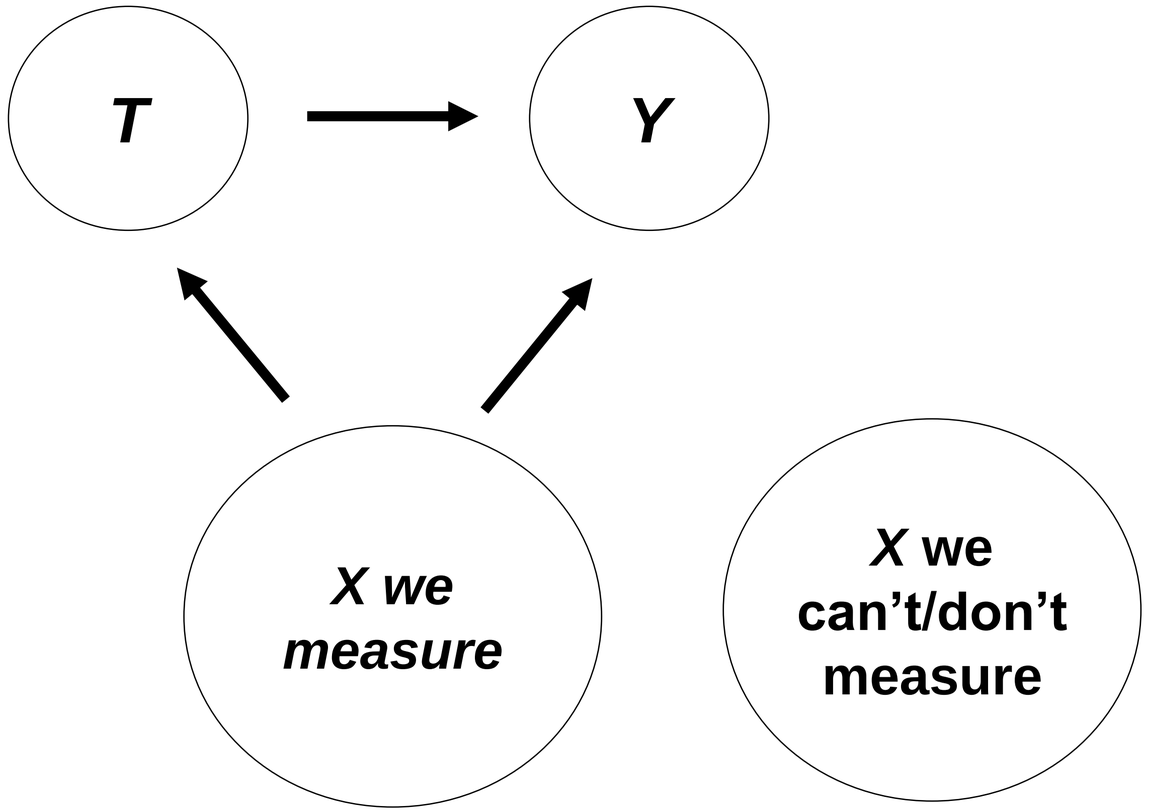
Perception (current) of running back value

THE RINGER NFL SHOW NFL NFL DRAFT

Do Running Backs Matter?

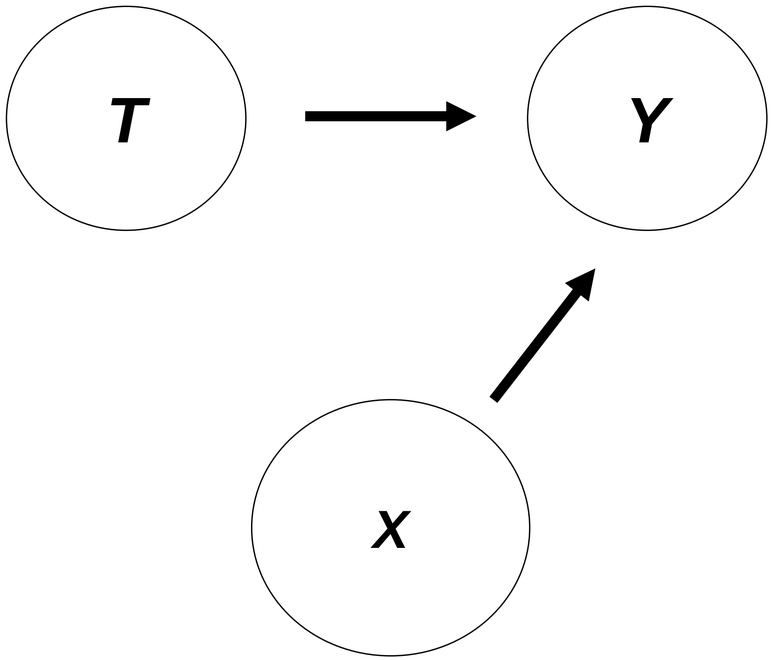


Observational data

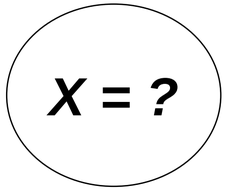


What may I be missing?

Randomized trial



Confounding variables in a randomized trial



Confounding variables in football

Down, Distance, Yardline, Time, Score, Hash mark, Head coach, Offensive/Defensive Coordinator, Formation, Personnel, Man/Zone, Blitz, Pass pattern, Motion, Player health, Temperature, Wind, Field surface, Dome, Defensive line game, Power/zone scheme, Timeouts (offense/defense), Two-minute warning, challenges remaining, audible

Football’s not that easy

Confounding variables in football

Additive?

Down, Distance, Yardline, Time, Score, Hash mark, Head coach, Offensive/Defensive Coordinator, Formation, Personnel, Man/Zone, Blitz, Pass pattern, Motion, Player health, Temperature, Wind, Field surface, Dome, Defensive line game, Power/zone scheme, Timeouts (offense/defense), Two-minute warning, challenges remaining, audible

Confounding variables in football

Additive?

Down, Distance, Yardline, Time, Score, Hash mark, Head coach, Offensive/Defensive Coordinator, Formation, Personnel, Man/Zone, Blitz, Pass pattern, Motion, Player health, Temperature, Wind, Field surface, Dome, Defensive line game, Power/zone scheme, Timeouts (offense/defense), Two-minute warning, challenges remaining, audible

Score*Time, Down*Yardline, etc

Interactions

Confounding variables in football

Down, Distance, Yardline, Time, Score, Hash mark, Head coach, Offensive/Defensive Coordinator, Formation, Personnel, Man/Zone, Blitz, Pass pattern, Motion, Player health, Temperature, Wind, Field surface, Dome, Defensive line game, Power/zone scheme, Timeouts (offense/defense), Two-minute warning, challenges remaining, audible

Score*Time, Down*Yardline, etc

Positivity: requires that there be both unexposed and exposed at each combination of the confounders

Background

Confounding variables in football

Down, Distance, Yardline, Time, Score, Hash mark, Head coach, Offensive/Defensive Coordinator, Formation, Personnel, Man/Zone, Blitz, Pass pattern, Motion, Player health, Temperature, Wind, Field surface, Dome, Defensive line game, Power/zone scheme, Timeouts (offense/defense), Two-minute warning, challenges remaining, audible

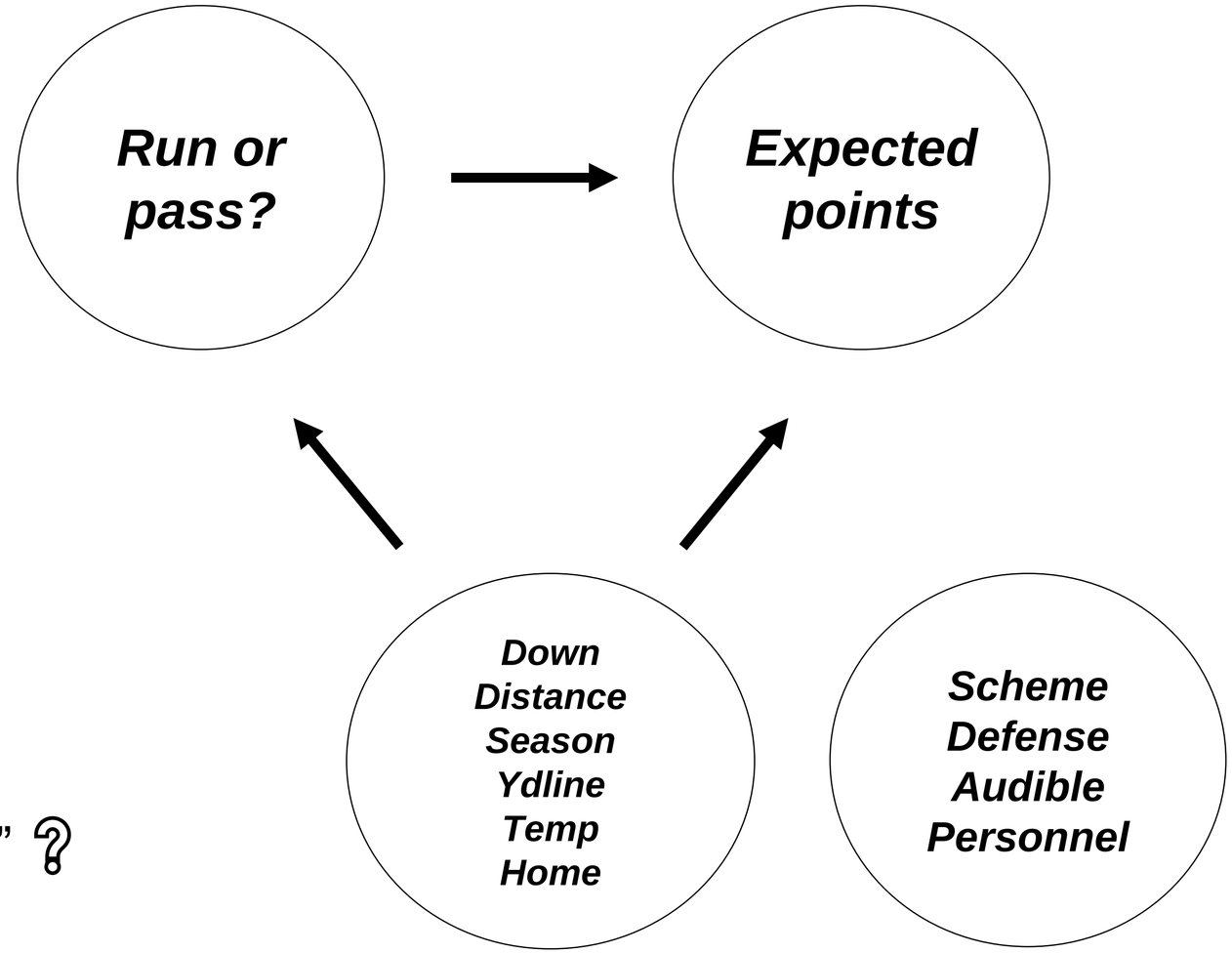
Score*Time, Down*Yardline, etc

Positivity: “positive” chance of receiving each treatment/therapy/intervention

Previous learnings -- limitations

Kovash/Levitt (2009) on run/pass

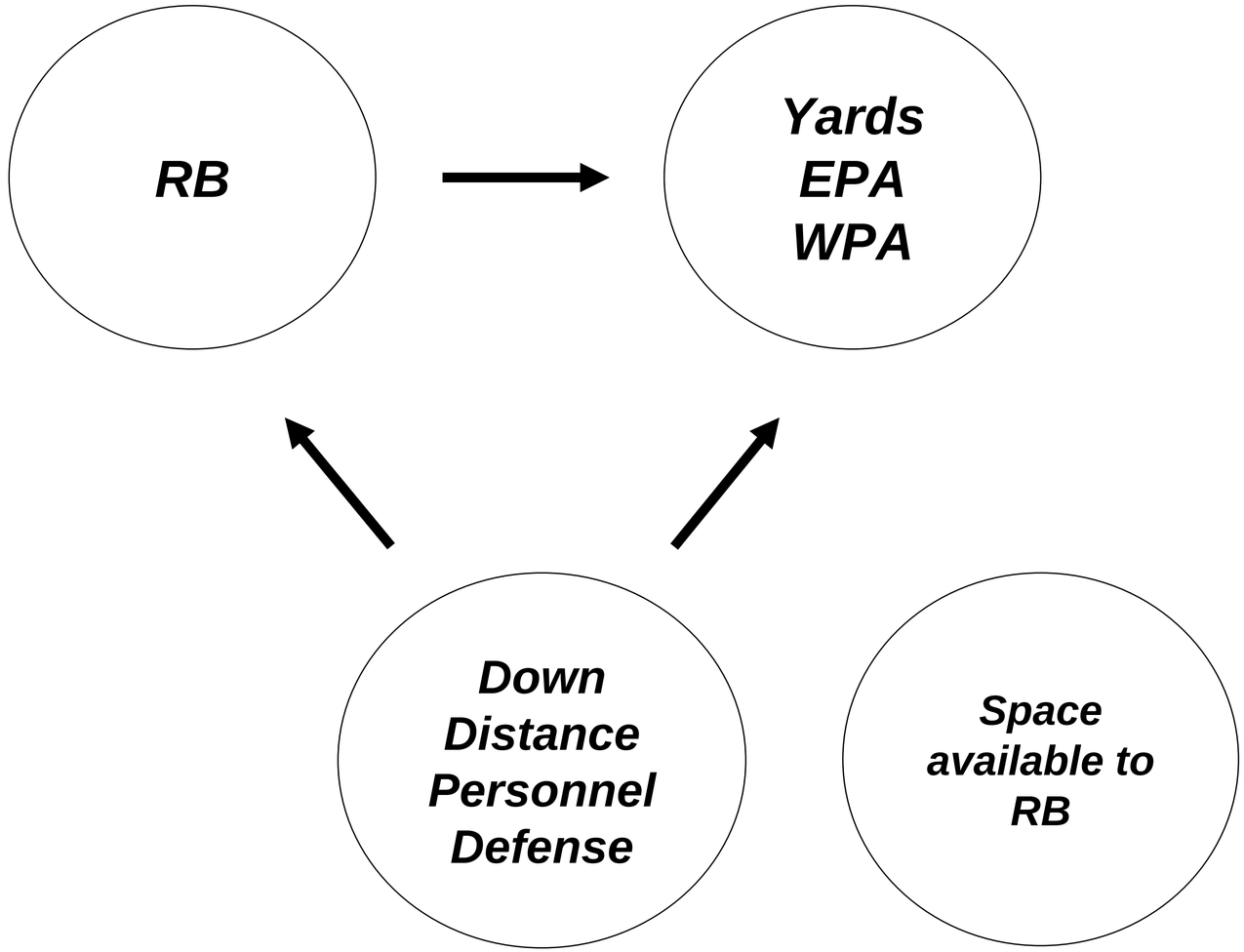
1. Passing linked to higher EPA ☑
2. Passing linked to higher EPA given down, distance, season, yard line, etc ☑
3. Among teams who that ran the ball, they should've passed more often ❓
4. "Football teams pass less than they should" ❓



Previous learnings -- limitations

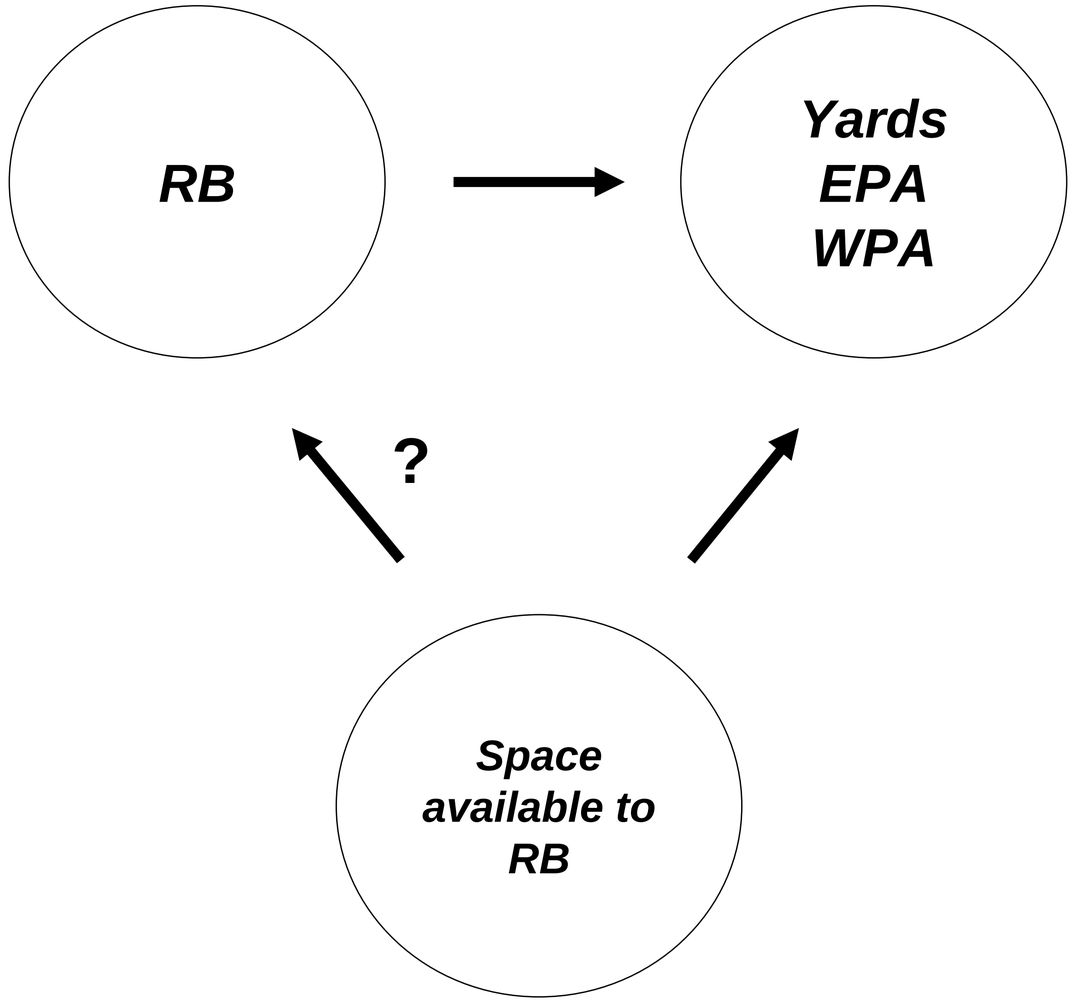
Perception (current) of running back value

1. Certain running backs have higher YPC ✅
2. Performances under various down, distance, personnel, defense combinations ✅
3. Running backs don't matter ❓



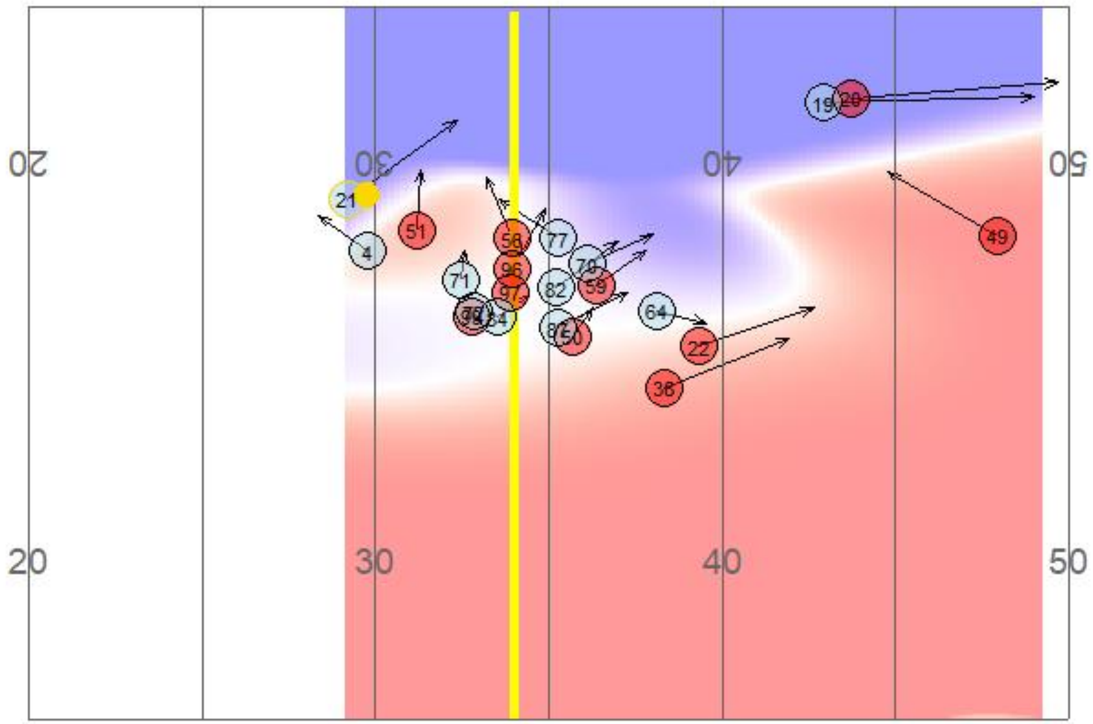
Previous learnings -- limitations

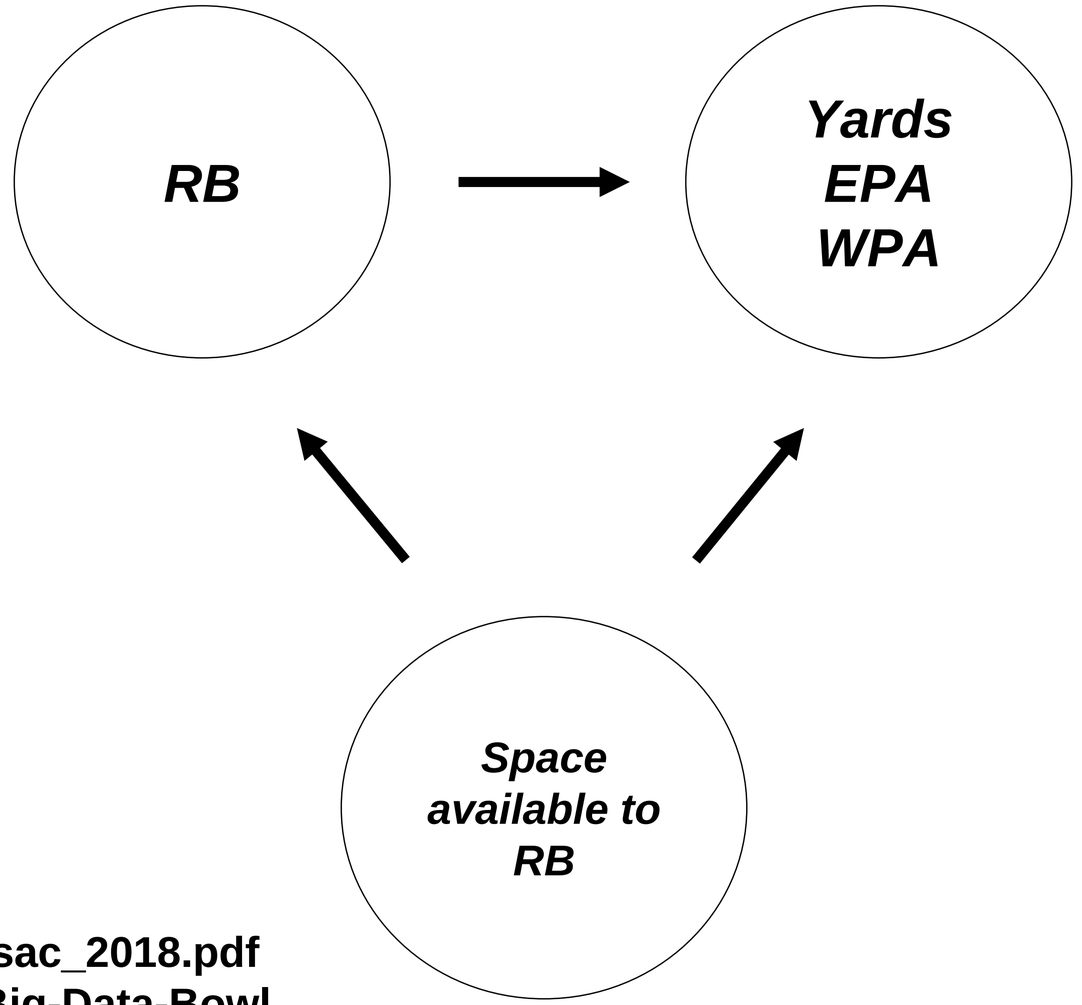
Perception (current) of running back value



Previous learnings -- limitations

Perception (current) of running back value





See http://www.lukebornn.com/papers/fernandez_ssac_2018.pdf
See Chu et al entry into Big Data Bowl ops.nfl.com/Big-Data-Bowl

Why does this stuff matter?

Why does this stuff matter? An example

THE LANCET
Volume 362, Issue 9382, 9 August 2003, Pages 419-427

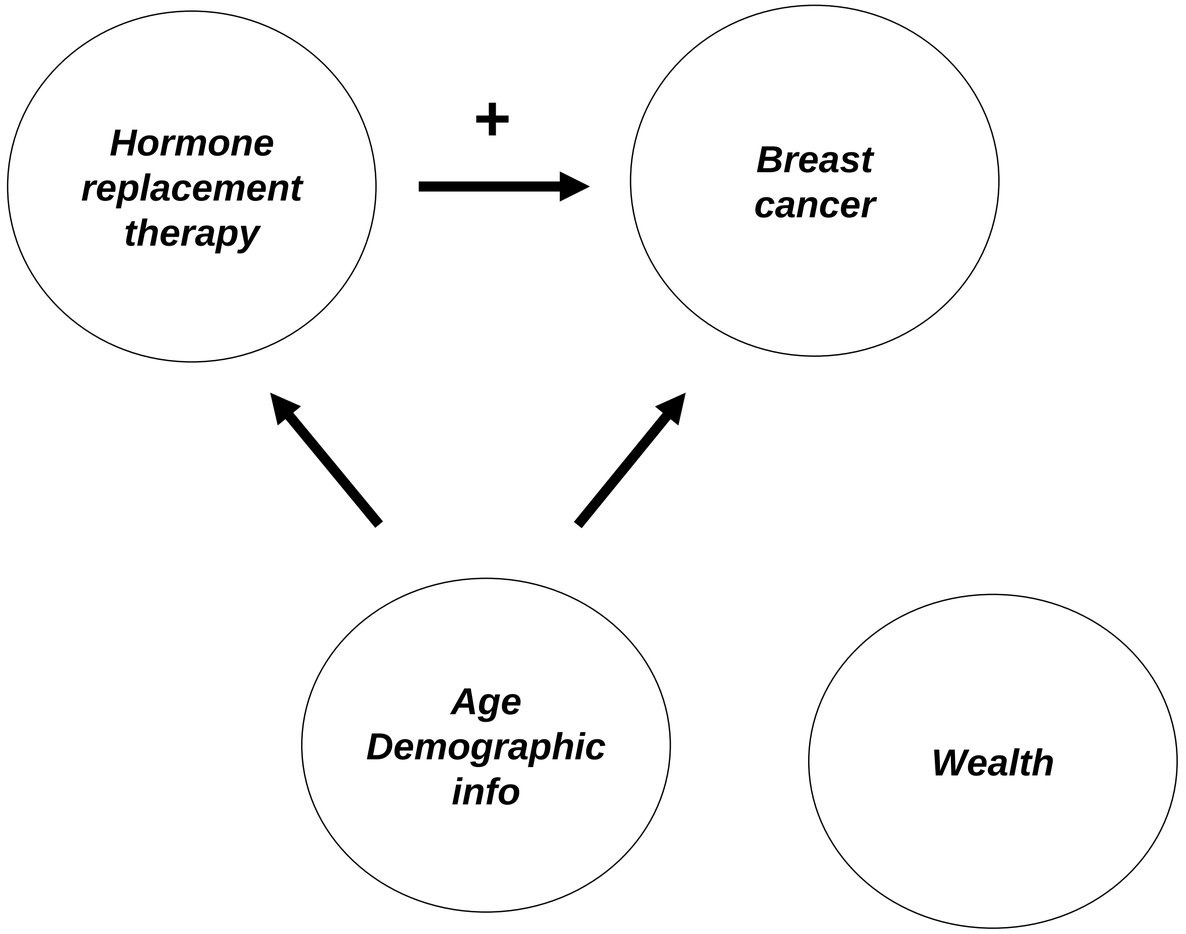
Articles

Breast cancer and hormone-replacement therapy in the Million Women Study

Million Women Study Collaborators
Show more

Interpretation

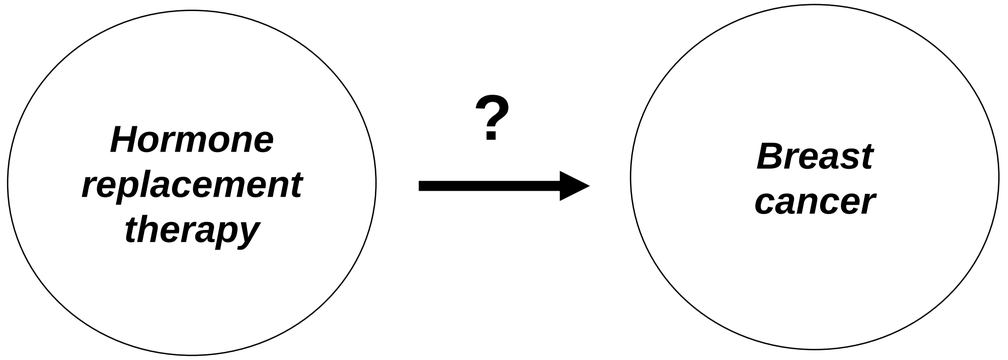
Current use of HRT is associated with an increased risk of incident and fatal breast cancer; the effect is substantially greater for oestrogen-progestagen combinations than for other types of HRT.



https://www.sciencedirect.com/science/article/pii/S0140673603140652 (h/t Julian Wolfson)

Big study finds no rise in death risk among women who took hormone therapy

A new twist to menopausal hormone replacement dilemma



More than a decade after a massive study of hormone replacement therapy was halted due to higher rates of breast cancer, heart attack and stroke among women assigned to the drugs to treat menopausal symptoms, a new follow-up study has found those women had no higher risk of death as of 2014 than participants who took a placebo.

The latest findings are published today in JAMA by investigators in the Women's Health Initiative, the same multi-institutional research program that conducted the original trial begun in 1993 and stopped in 2002 because of those adverse effects. The new study focuses on "all-cause mortality," or deaths from any cause, up to 18 years after the start of therapy and 10 to 12 years after the therapy was stopped. It offers reassurance to postmenopausal women who took hormone replacement therapy during the trial that they have not increased their risk of dying.

https://www.fredhutch.org/en/news/center-news/2017/09/death-risk-menopausal-hormone-replacement.html

Why does this stuff matter? A football example

Go for It: Sean McVay's Achilles' Heel Is His Fourth-Down Decision-Making

The Rams head coach is known as an X's and O's wunderkind, but the one situation he's most conservative in could prove crucial going up against Bill Belichick and the Patriots

By Riley McAtee | Jan 23, 2019, 6:10am EST

NFL

Charts: Just How Wimpy Are NFL Coaches On Fourth Down Calls?

NFL

Message to NFL Coaches: Stop Being Stupid on Fourth Downs

JANUARY 17, 2019 Ahmed Cheema

The Dumb Punt Epidemic

Oliver Connolly Follow

Sep 28, 2015 · 4 min read

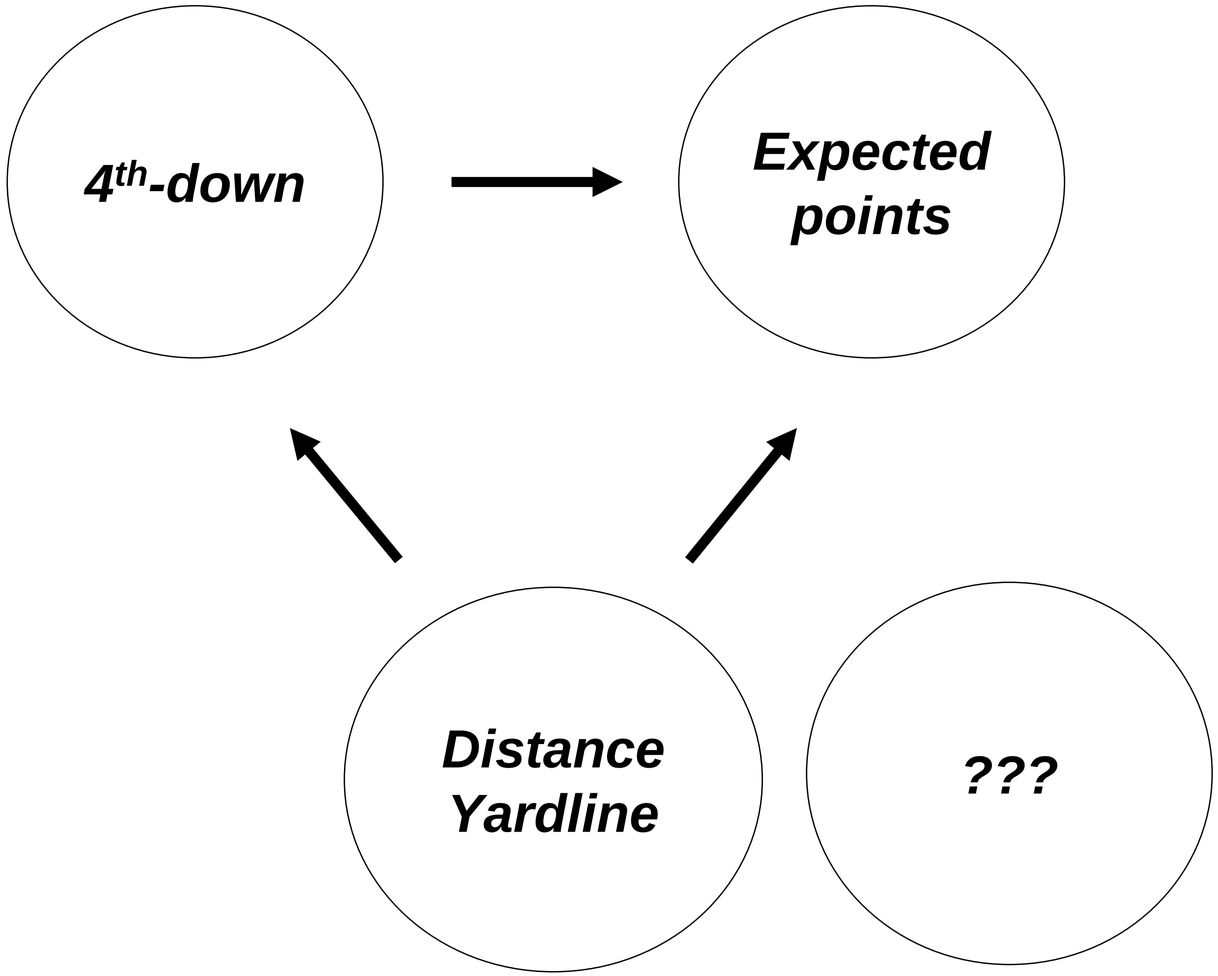
Previous learnings

Romer (2006) on 4th downs

David Romer

University of California, Berkeley and National Bureau of Economic Research

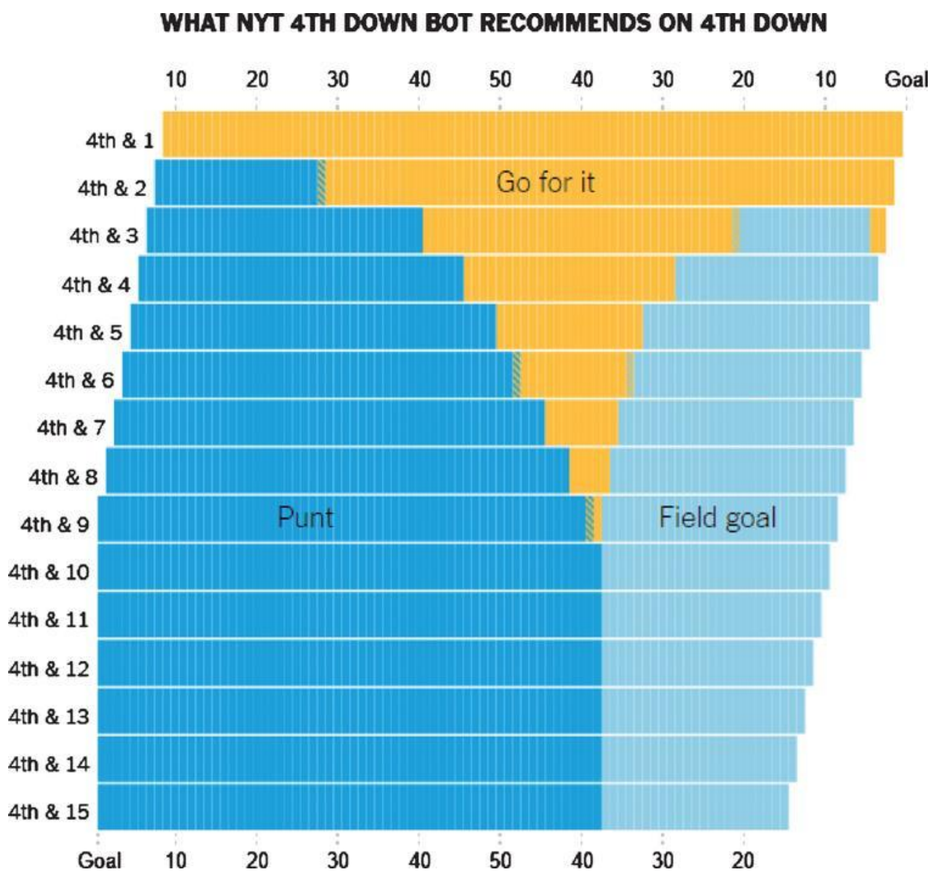
This paper examines a single, narrow decision—the choice on fourth down in the National Football League between kicking and trying for a first down—as a case study of the standard view that competition in the goods, capital, and labor markets leads firms to make maximizing choices. Play-by-play data and dynamic programming are used to estimate the average payoffs to kicking and trying for a first down under different circumstances. Examination of actual decisions shows systematic, clear-cut, and overwhelmingly statistically significant departures from the decisions that would maximize teams' chances of winning. Possible reasons for the departures are considered.

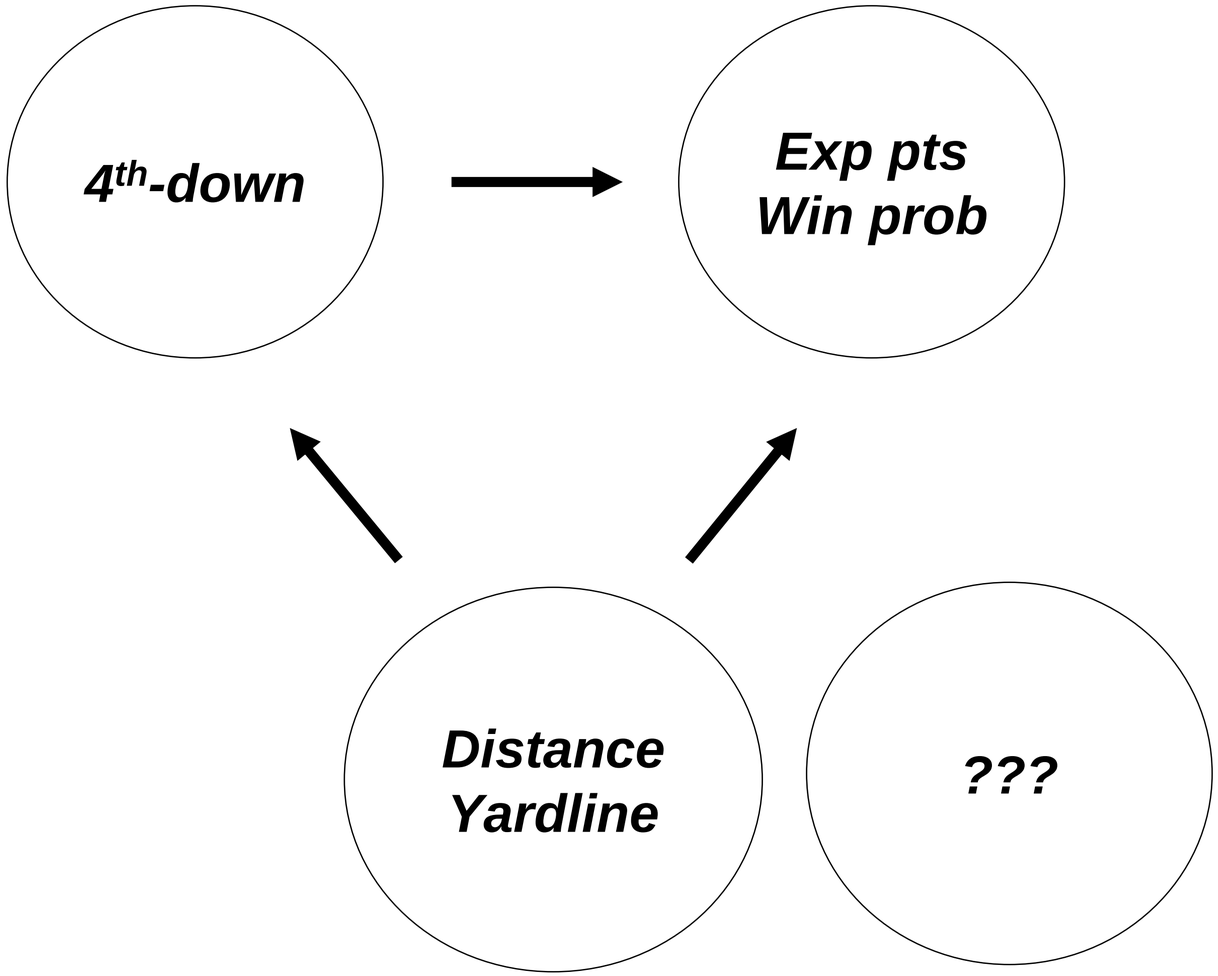


Additional work: Virgil Carter, Robert Machol, Brian Burke, Trey Causey

Previous learnings

4th down bot (2013) on 4th downs

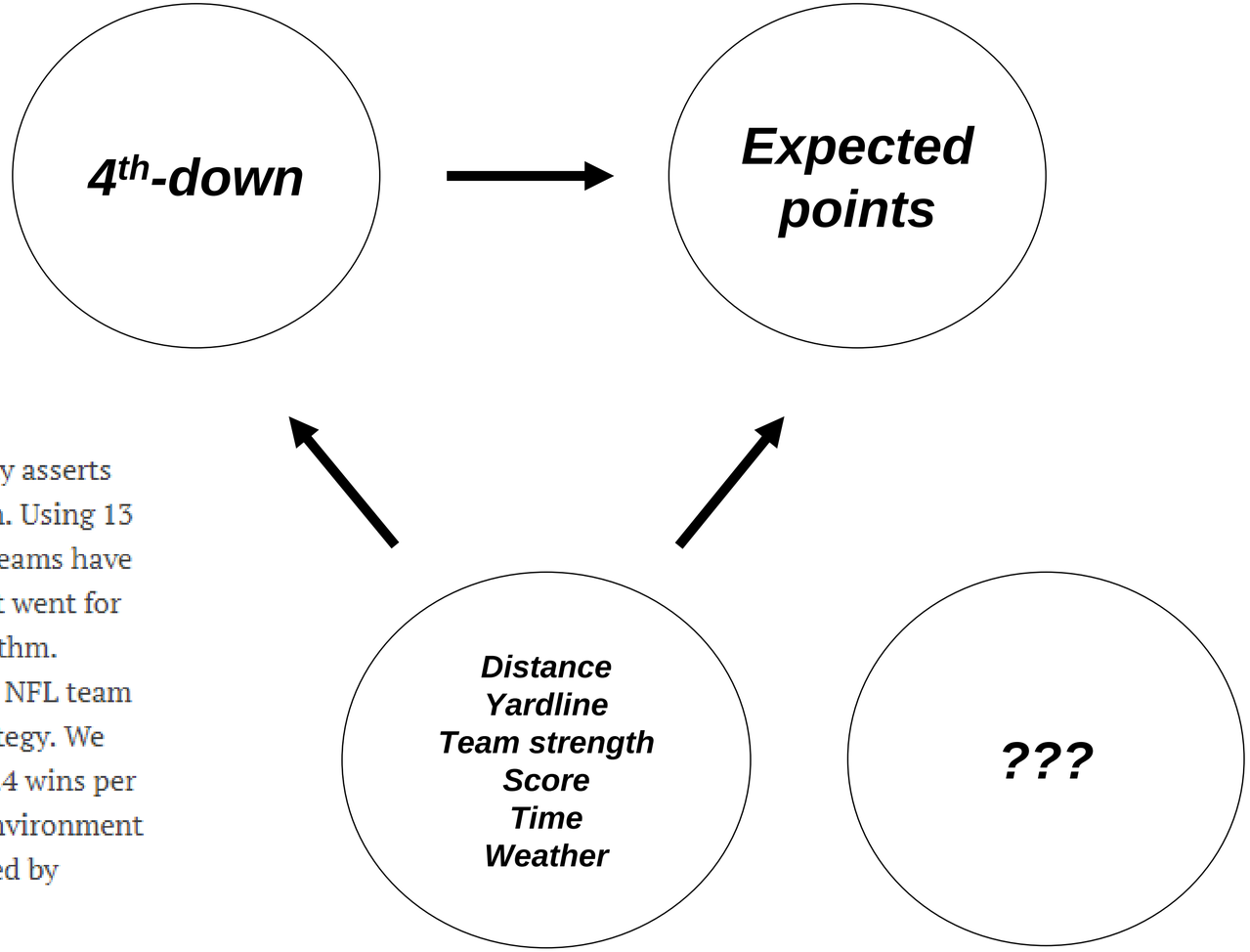




Additional work: Virgil Carter, Robert Machol, Brian Burke, Trey Causey

Previous learnings

Yam & Lopez (2017) on 4th downs



Abstract

In part driven by academic research, perception in the sports analytics community asserts that coaches in the National Football League are too conservative on fourth down. Using 13 years of data, we confirm this premise and quantify the unobserved benefit that teams have missed out on by not utilizing a better fourth down strategy. Formally, teams that went for it are paired to those who did not go for it via a nearest neighbor matching algorithm. Within the matched cohort, we estimate the additional number of wins that each NFL team would have added by implementing a basic but more aggressive fourth down strategy. We find that, on average, a better strategy would have been worth roughly an extra 0.4 wins per year for each team. Our results better inform decision-making in a high-stakes environment where standard statistical tools, while informative, have possibly been confounded by extraneous factors.

Limitations

Spot the difference

NGS data

On player tracking data

- RFID chips in shoulder pads

- Real time location, speed, direction (10+ frames/sec)
  - Players
  - Ball

- Beginning in 2018, teams receive on all teams in all games

- NFL play-by-play data: ~160 rows per game

- Next Gen Stats data: (at least) 300,000 rows per game

Limitations

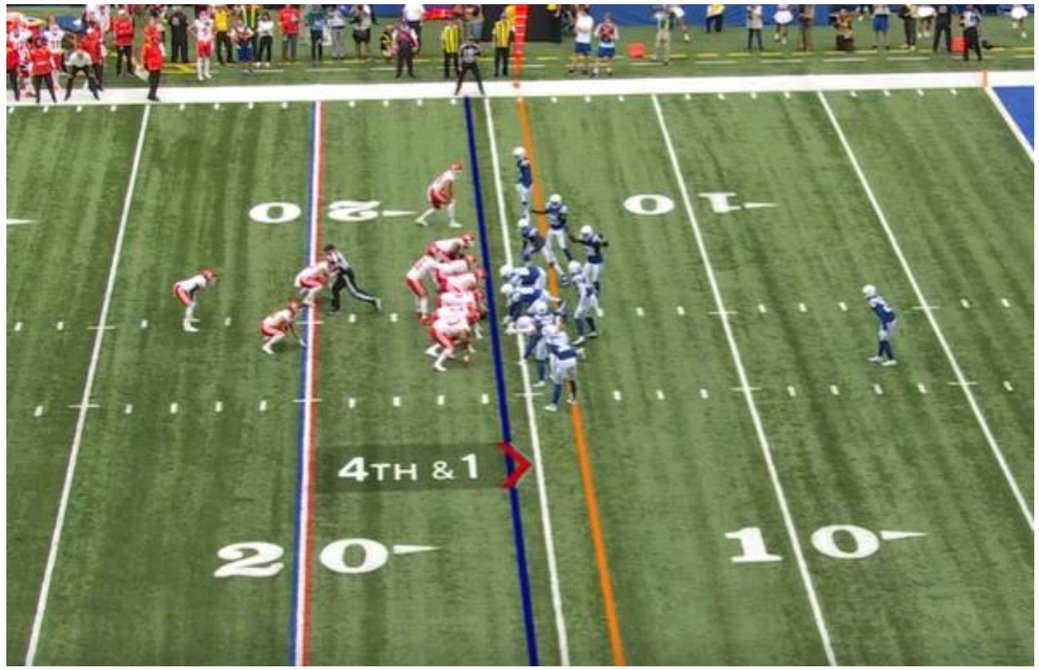
Spot the difference

4th-0.21 yards

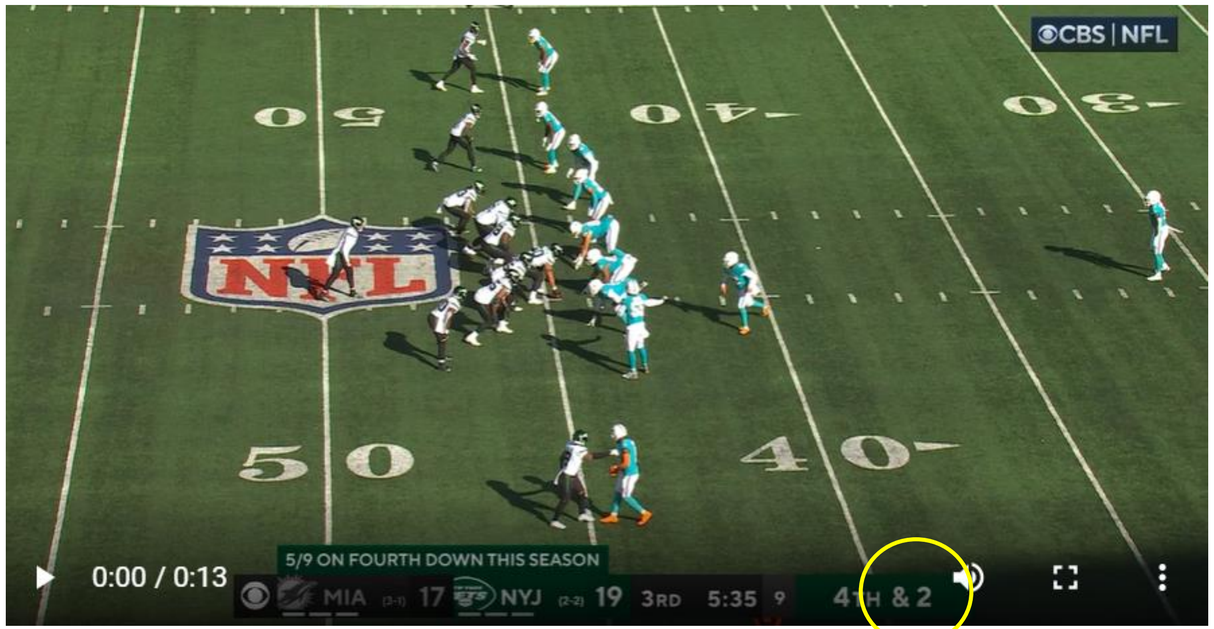
4th-1.58 yards

Limitations

NFL yards to go



NYJ 4 1 MIA-44 (5:30) (Shotgun) Z.Wilson pass incomplete deep left to C.Davis.



4th-1.999 yards :: 4th-1

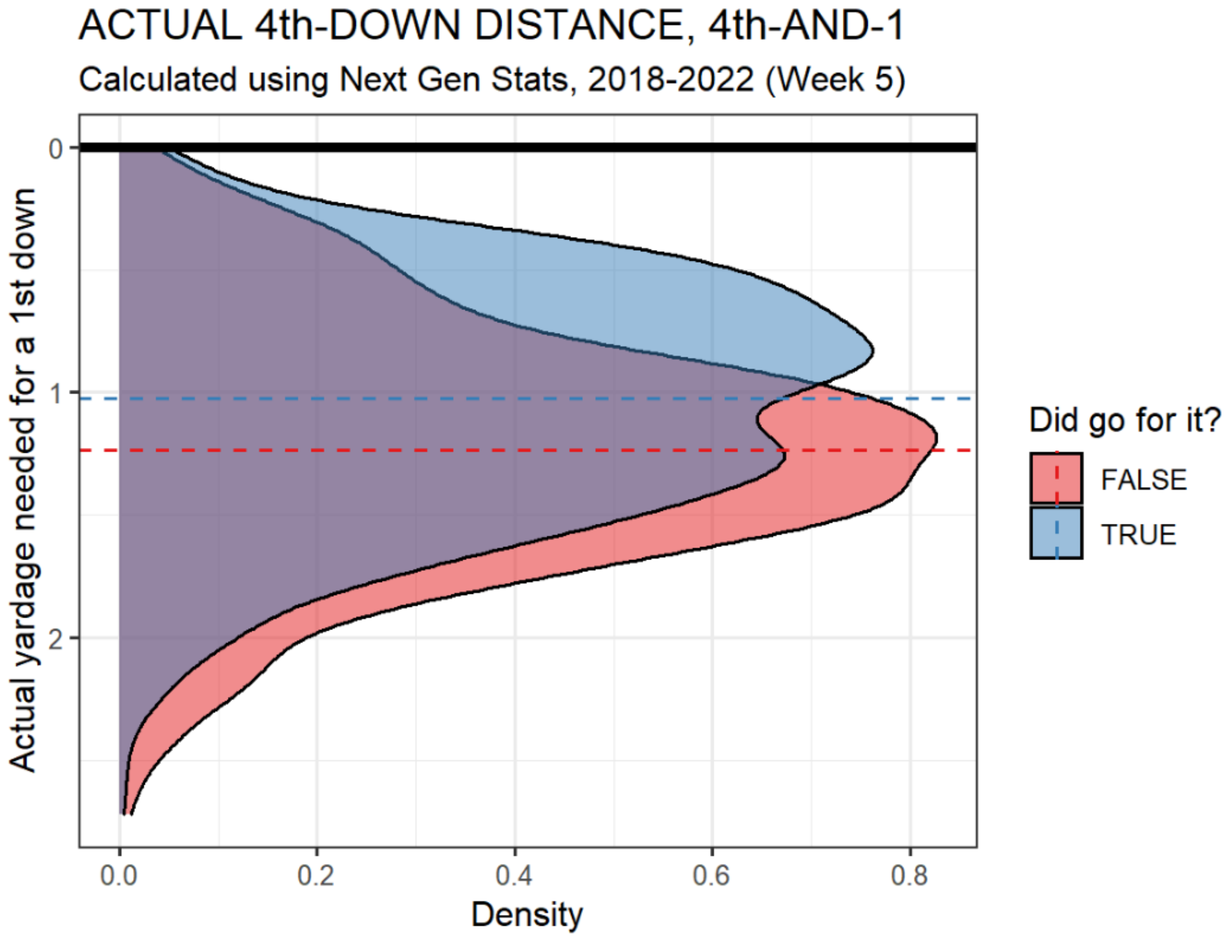
A missing piece

Teams that went for it had shorter distances to go

4th-down

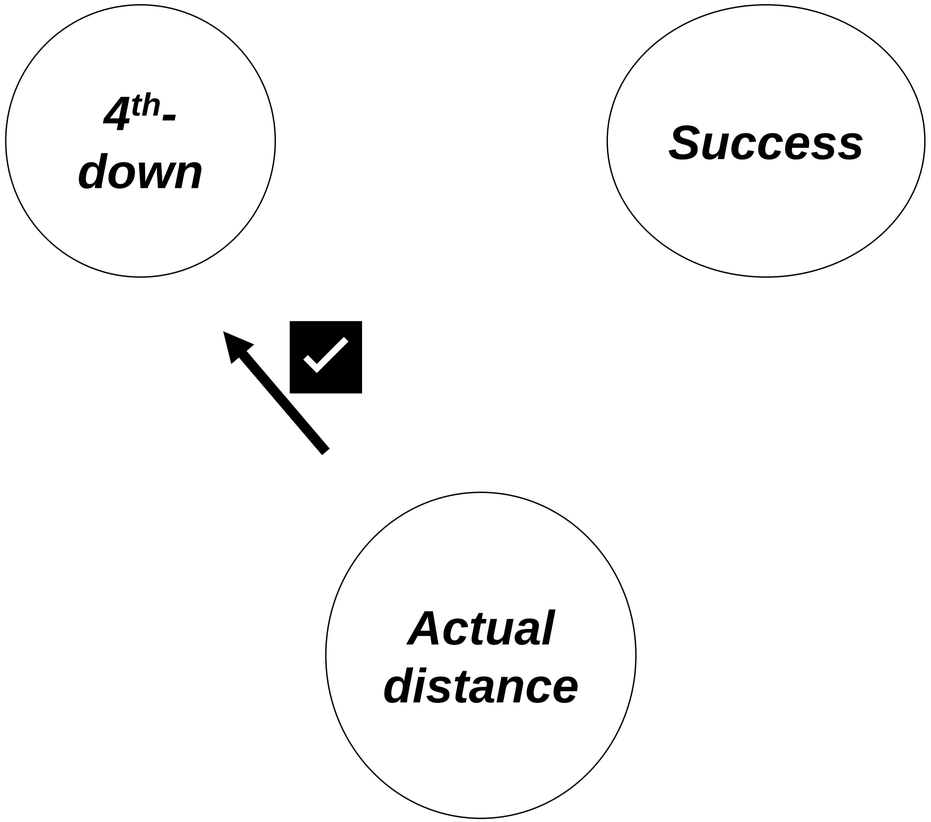
Success

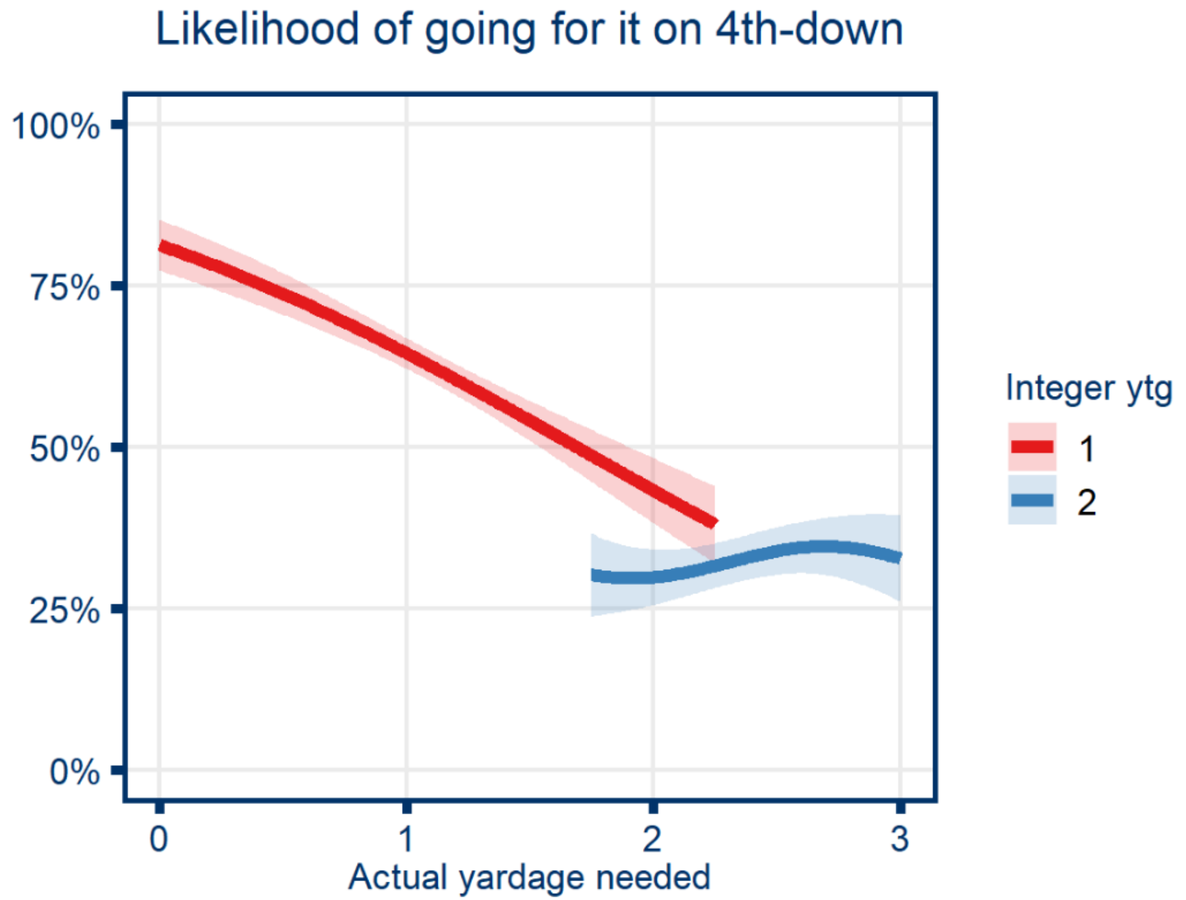
Actual distance



A missing piece

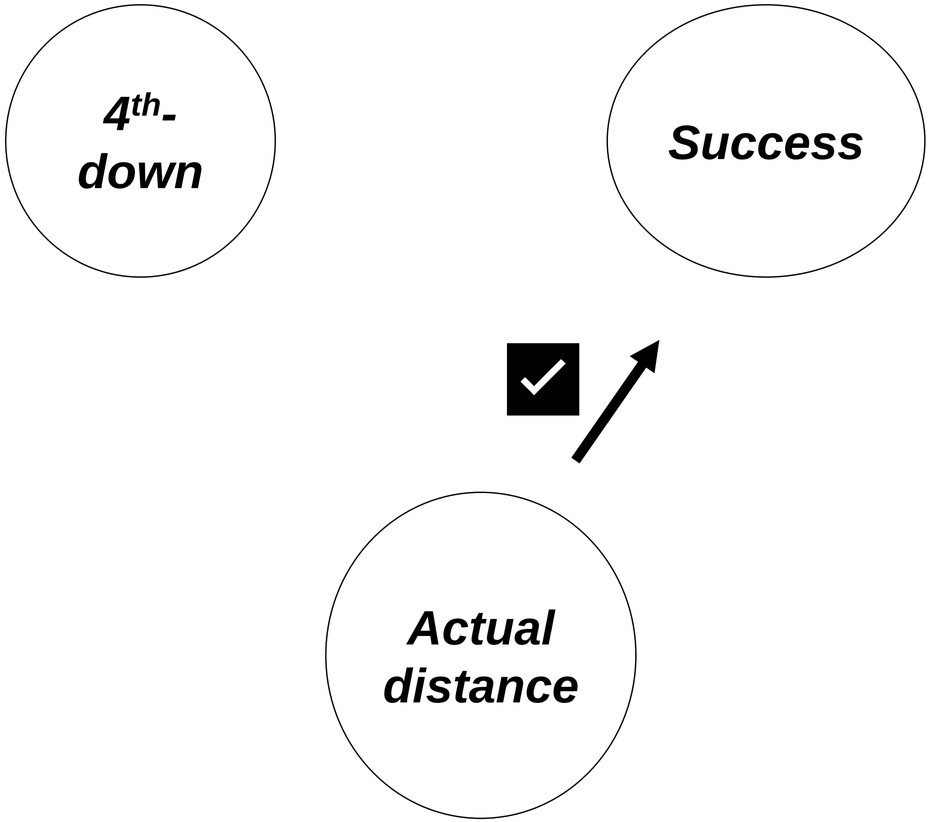
Teams go for it on 4th-and-inches more often

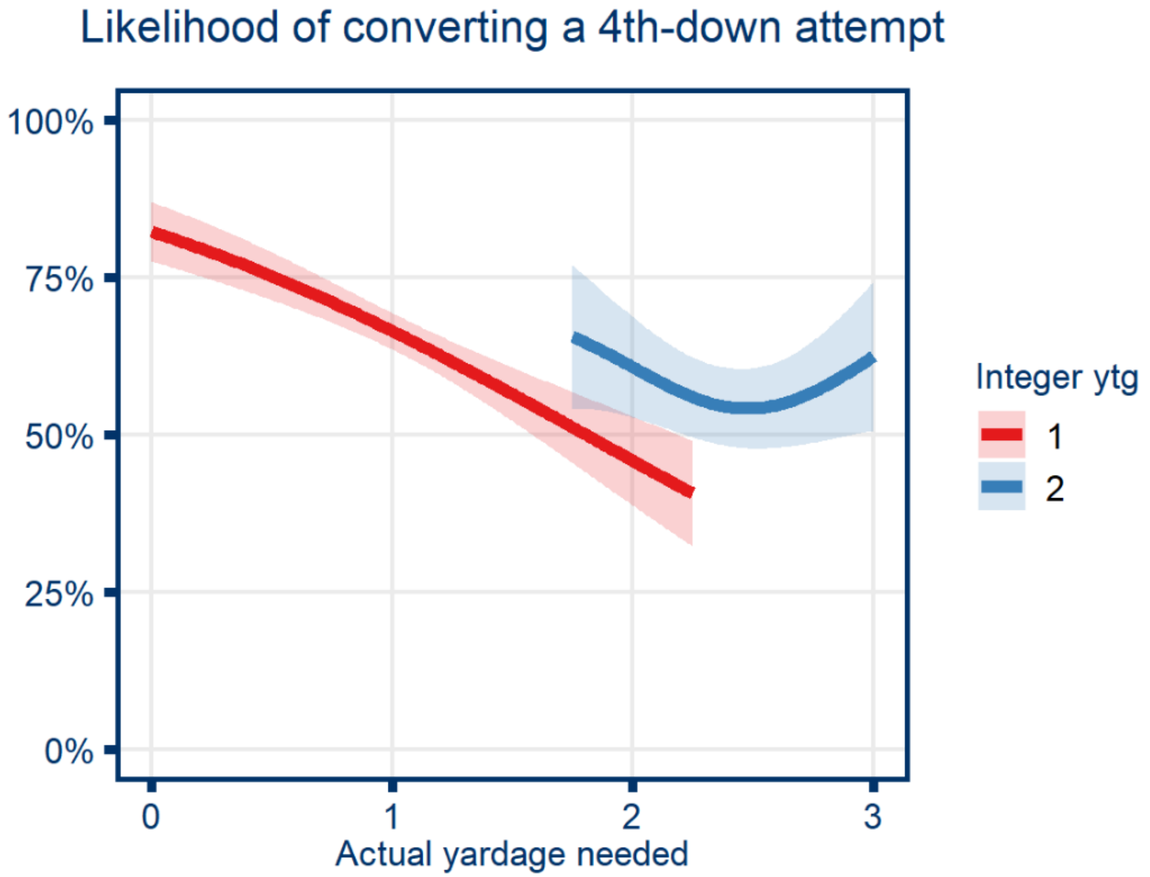




A missing piece

Teams convert on 4th-and-inches more often





Implications for fourth down studies

Romer (2006) on 4th downs

Assumes fixed conversion rate for each integer distance

Now we know: Likelihood of success among teams that went for it about 5 percentage points higher than for teams that did not go for it, given actual distance to go

Yam & Lopez (2017) on 4th downs

Matches on integer distance, 2005 through 2016

Now we know:

Matching on integer distance, 2017 & 18 seasons:

- 3.3% change in WPA per-play by going for it (CI: 1.9% to 4.7%)
- 0.27 wins per team per year

Matching on actual distance, 2017 & 18 seasons:

- 2.0% change in WPA per-play by going for it (CI: 0.0% to 3.9%)
- 0.14 wins per team per year

A missing piece

Recapping

- Sports analysis is tricky given that most data is observational
- In observational data, unmeasured confounding can bias an analysis
- Even when we thought we measured something, maybe we didn't measure it well
  - Roughly 40% attenuation in benefit to going for it*
  - *Several other caveats that are not part of the story (yet)
- Limitless possibilities – but substantive sport expertise (e.g, scouts and coaches) and data knowledge required

What’s next?

NFL’s Big Data Bowl

- Pipeline for club hires
- Innovation
- Support partners (NextGenStats, AWS)
- Engage core fans with analytics content

2023 Big Data Bowl Overview

Analytics competition: Assess the pass blocking or pass rushing performance of an individual player or team using Next Gen Stats data on dropbacks.

Mentoring program: Link participants from underrepresented groups with current team analytics staffers

Prize money: 100k (AWS)

Data: 2021 W1-W8 dropback passes, snap -> pass release/sack/scramble event, PFF data

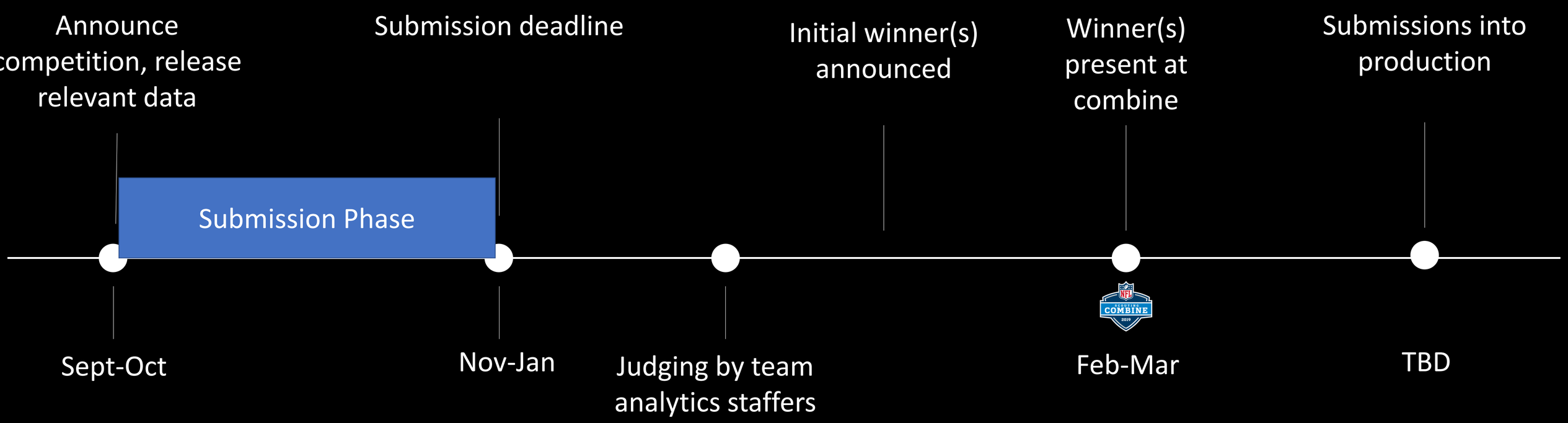
Topic tracks:

Undergraduates only (contest is open to *everyone,* but there will be a track for only undergraduates)

Create a metric to assess *individual* or *team* performance

Coaching focused

T I M E L I N E

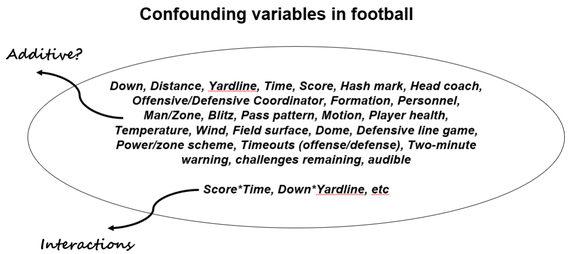

Announce competition, release relevant data
Submission deadline
Initial winner(s) announced
Winner(s) present at combine
Submissions into production
Submission Phase
Sept-Oct
Nov-Jan
Judging by team analytics staffers
Feb-Mar
TBD

Addendum: Over 50 participants from the first four Big Data Bowls went on to work in pro sports, including more than 30 (updated estimate, 34) by NFL teams or player tracking vendors

What’s next?

Limitations & final thoughts on player tracking data

- Not a panacea



- Football knowledge (or collaboration) needed. Also time.
- Some questions lead to dead ends, and that’s okay
- We don’t know much (yet)*
- We don’t even know what questions to ask (yet)*

*publicly

Football Data Science at the NFL

Mission statements

"To ensure that NFL football is the greatest competitive sport in the world"

- Mission Statement, NFL Football Operations

"[Use data] To ensure that NFL football is the greatest competitive sport in the world"

- Mission Statement, NFL Football Data & Analytics team

Football Data & Analytics – 2023 workstreams

League Priority	*What we are working on*
1 Future of the Game	**Research to improve the game** ➢Help spearhead the most efficient and timely review process in NFL history ➢Collection of new NFL stats, including expected yards, expected return yards, coverage scheme, and quarterback slides (Big Data Bowl & NGS collaborations) ➢Data driven Combine: player tracking data at 2021/2022 International Combines and 2022 Scouting Combine (Indianapolis) ➢Foundational DEI reporting on coaching and General Manager transitions ➢Rules Proposals for 2023 and beyond
2 Know the Business	**Integration of football data throughout the league** ➢With Player Health and Safety: identify impact and unintended consequences of potential changes ➢With Media, Player Health and Safety: Overtime 2022 rules change review presents sweeping perspective of club rules proposals ➢With Player Health and Safety: Collaborations to understand football impact of injuries ➢With Broadcast: How do perceived scheduling inequities impact competitiveness? ➢With Officiating: How can we use player tracking data to identify fouls?

Thank You!

@Statsbylopez

michael.lopez@nfl.com