

@11_tjr



OU MASTER PLAN
2027
生きがいを育む社会の創造

ピッチコールの公平性と観客の存在 : MLBの投球データ分析

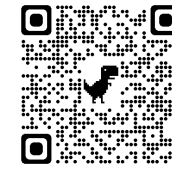
Reio Tanji

@Sports Analyst Meetup #13



大阪大学「ワニ博士」

自己紹介



@11_tjr



- 丹治 伶峰 (たんじ れいお): #spoana LTは今回が2度目
 - 0から始まる大学で経済学の研究をしています
 - 博士後期課程学生
 - 労働経済学・行動経済学分野の実証研究が専門、スポーツのデータも頻繁に利用
 - 好きなスポーツ: 野球など
 - 右投左打、大学時代は指名打者キャッチャーでした
 - T-岡田選手のファン、試合はセリーグから女子硬式まで何でも観ます
 - Twitterやら何やらでデータ分析など(たまに)公開しています
 - 今日の報告が面白かったらフォローしてね
 - スポーツ業界への就職も視野にFA宣言中です
 - ノンテnder間近なだけかもしれない

よろしくお願いします！



Abstract



@11_tjr

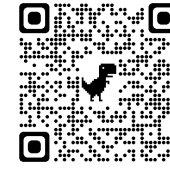


- **Research Question**

「Major League Baseballの投球データを用いて、球審の判定に影響を及ぼすバイアスの仕組みを明らかにしよう」

- **Home Advantage (ホームアドバンテージ)**: 特定地域・スタジアムをホームチームとする球団の攻撃中に、ストライクゾーンが狭くなる
- ホームアドバンテージが起こる仕組み: 球審にはそうするインセンティブがない
- 球審自身が自らの意思でそうしているわけではなく、スタジアムに詰めかけた**観客の存在 (Social Pressure)**がバイアスの原因に？
- Covid-19の感染防止措置が引き起こした観客動員数の激減を利用して、このメカニズムを実証分析

Literature



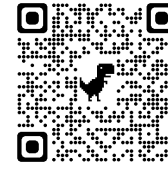
@11_tjr



- 研究対象としてのスポーツ
 - プレーやその評価がある程度定量化されており、そこからの乖離：バイアスの効果を推定しやすい
 - 観客に端を発するバイアスは、一般社会でも同様に表れる可能性がある
- 書籍『オタクの行動経済学者、スポーツの裏側を読み解く』
'Scorecasting: The Hidden Influences Behind How Sports Are Played and Games Are Won,' Moskowitz, Wertheim: 2011
- スポーツデータを扱ったバイアスの研究も様々
 - Parsons et al. (2011): 球審の人種バイアス
 - Higgs and Stavness (2022): 4大スポーツのホームチームの勝率
 - Sandberg (2018): 馬術、審査員による国籍バイアス



Behavioral Biases



@11_tjr



- ホームアドバンテージに限らず、審判の判定にはライブで観戦する観客の存在が影響することが知られている
 - 観客動員数の少ない試合では、ホームアドバンテージの影響が小さくなる (Moskowitz, Wertheim, 2011)
 - 国籍バイアスの影響は、動員数の少ない試合の方が大きい (Parsons et al., 2011)
- メカニズム: さまざまなモデルで解釈可能
 - 自軍のフランチャイズ球場の特性をよく知るホームチームの選手のパフォーマンスが上がる
 - 観客の存在が Social Pressure として球審に影響を与える
 - モニタリング効果: 観客が審判の判定精度を「監視」し、より公平な判定を促す
 - テレビ中継やトラッキング機器を用いたフィードバックの存在

Background

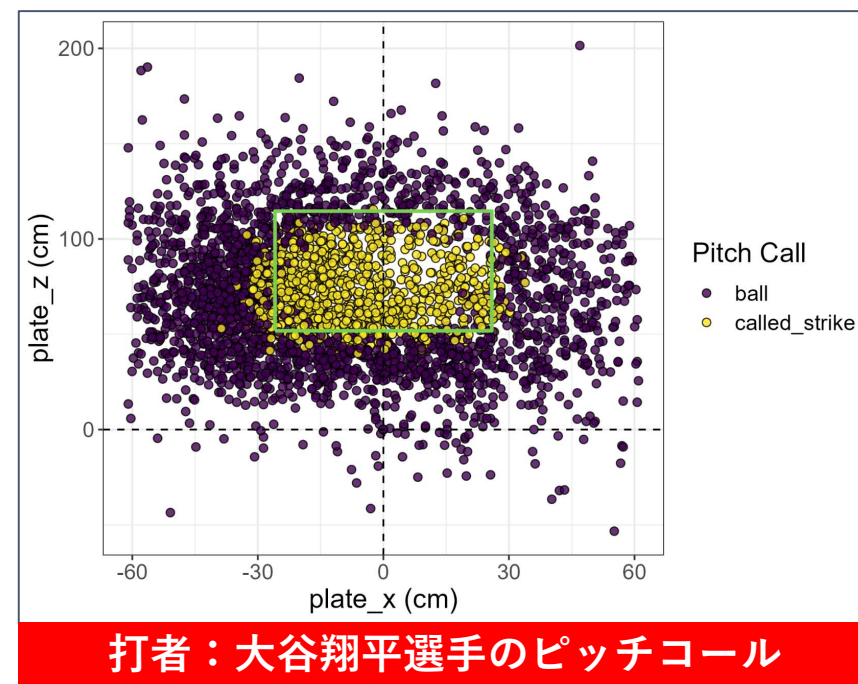


@11_tjr



OSAKA UNIVERSITY
MASTER PLAN
OU 2027
生きがいを育む社会の創造

- ピッチコール: ストライク・ボールの判定
 - 球審が目視で投球の通過位置を確認し、判定を行う
 - 一度下した判定は覆らないため、ゲームに系統的なバイアスの入り込む余地がある
 - 一方で、トラッキングデータに基づいた「あるべき判定」も定義可能
 - 全投球の6割前後が見送り投球: バイアスの影響は小さい
- 審判員は原則4人一組のクルーで各ポジション (PL, 1B, 2B, 3B) をローテーション
 - 判定精度はトラッキングシステムを用いてフィードバックされる



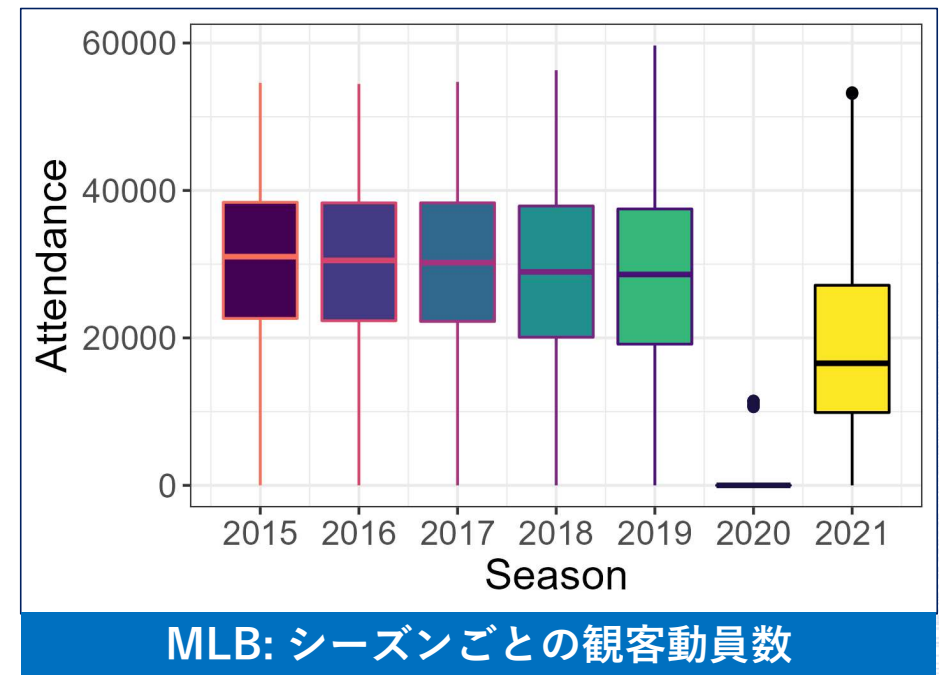
Behavioral Biases



@11_tjr



- Covid-19の感染拡大により、観客動員数が厳しく制限
 - 2020年のレギュラーシーズンは無観客、2021年も一部制限
 - それ以前のシーズンから予想することが難しい「外生的な」動員数の変動
 - 観客動員がバイアスの程度に及ぼす因果関係を識別できそう
- また、国境を跨ぐ移動・大規模な移動を減らすため、フランチャイズ球場で試合を行えないチームも存在
 - Blue Jaysの他、数チームが他球団の本拠地を間借りしたり、MiLB本拠地を利用



Data & Identification Strategy



@11_tjr



- 使用データ

- 一球データ: MLBAMがBaseball Savantで公開している投球のトラッキングデータ
- 試合データ: MLBAMが公開する試合球場、審判などのデータ
- 選手データ: Lahmanデータセット、Chadwick Baseball Bureau, Fangprahs, Baseball Reference, etc.

データの取得・成型にはRを利用 (baseballrパッケージがとても便利), $N = 2,338,649$

- 分析方法

- 打者がスイングしなかった投球について、ストライクを1,ボールを0とするダミー変数を作成、線形確率モデルでホーム/ビジターチームが攻撃時のストライク確率の変動を推定する
- 同じ位置を通過した投球に対する平均的なストライク確率、打者の体格、投球カウント、投手、打者、捕手、審判の特性などをコントロール変数として制御



Summary of Results



@11_tjr



MASTER PLAN
OU 2027
生きがいを育む社会の創造

Covid-19の感染防止対策以前

- ホームチームの攻撃時には、同じ場所を通過した投球に対するストライクコールの確率が0.4ppt上昇
 - 特に、ストライクゾーンの端(平均的なストライク確率が30-70%)を通過した投球については1.7ppt

ホームチームに有利な判定:ただし、そのメカニズムについては議論の余地がある

Covid-19 Era

- 2020年シーズンにおいては、その他の年度で観測されたホームアドバンテージが有意差を持たなかった
 - 点推定の値もゼロ近傍
- どちらのフランチャイズでもない:中立地で開催されたゲームについても、判定がいずれか一方に偏ることは確認されなかった

Effects of Spectators



@11_tjr



- より明示的な観客動員に関する変数：観客動員数
 - 観客動員数によるサブサンプル分析：動員数およそ5000人以下の試合ではホームアドバンテージが消失
 - 一方で、それ以上の観客が入った試合では、アドバンテージの効果量が大きく変動することはなさそう
- メカニズムに関する議論
 - 球場音声やスコアボードの仕様は変わらないはず
 - 観客のパネルや、スピーカーから流す「歓声」は？：NPBならより大きな影響が？
 - 無観客で声が通りやすい？：その効果がホーム・ビジターで異なるのか？

Concluding Remarks: What Comes Next?



@11_tjr



バイアス研究がもたらすものとは？

- MLBでは、判定の自動化に関する制度変更・議論が進む
 - 本研究で議論したようなバイアスは(システムに組み込まない限り)残らず取り除かれることになる
 - バイアスを取り除く：自動判定が導入された時に起こりうる変化のシミュレーション
 - 技術的な制約に基づく系統的なバイアスの発生
 - ピッチコールに対するジャッジレビュー制度
- 一方で、こうした傾向は競技の歴史が作り上げた「知恵」でもある
 - カウント・展開によって伸び縮みするストライクゾーン：ゲームコントロールとしての役割
 - 属人的特性に基づく差別は問題だが、ホームアドバンテージはどうか？
 - ピッチャーの打席：ストライクゾーンが広がる→二刀流選手の登場



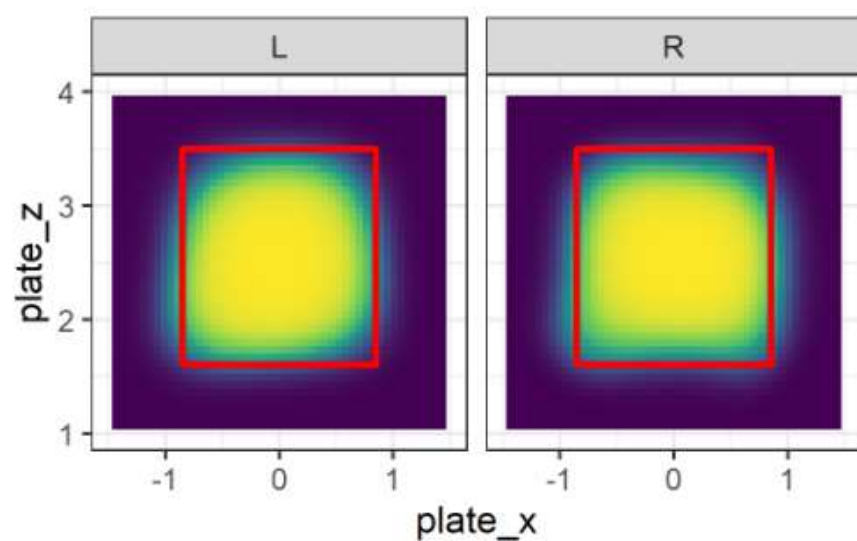
Appendix: 平均的なストライクコール確率



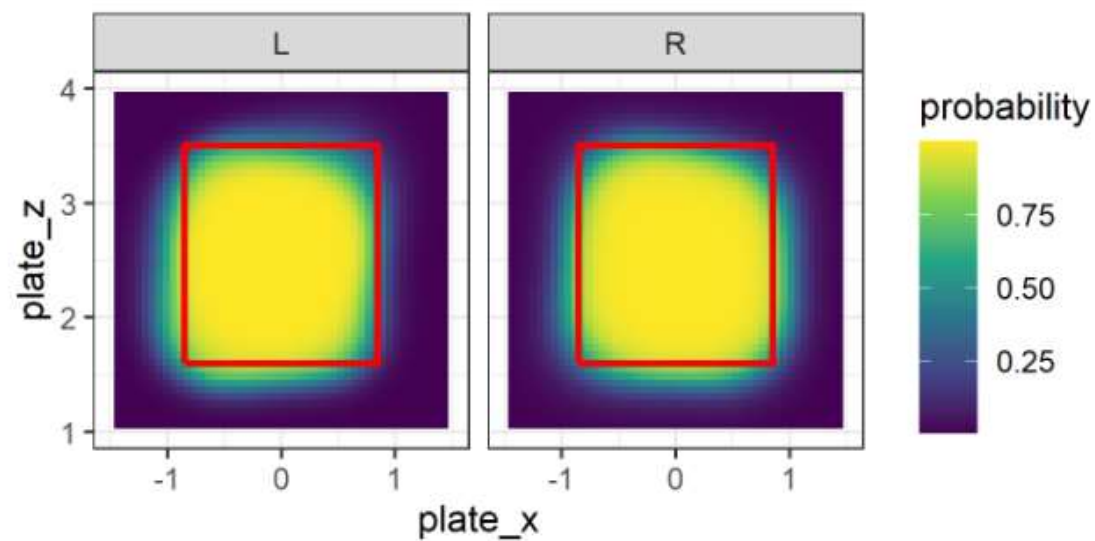
@11_tjr



OSAKA UNIVERSITY
MASTER PLAN
OU 2027
生きがいを育む社会の創造



カウント0-2



カウント3-0