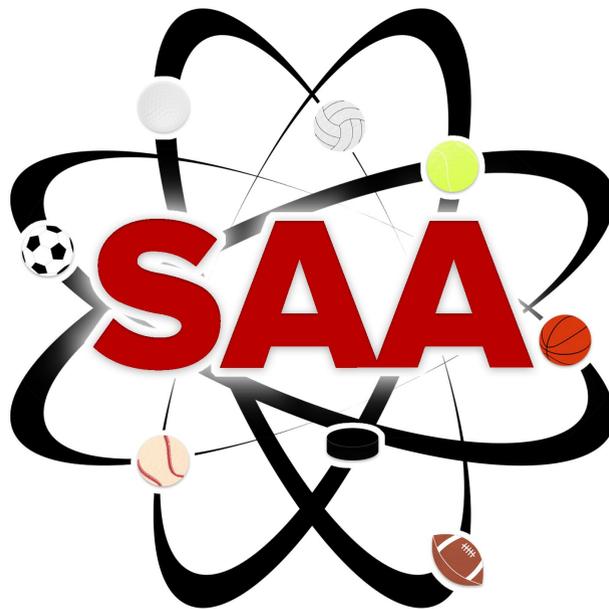


1st Annual OHIO STATE SPORTS ANALYTICS CONFERENCE 2020

Brought to you by the Ohio State [Sports Analytics Association](#) with assistance from the [Sports and Society Initiative](#).



Friday, November 13

8:00 am– 6:00 pm

Conference Rooms:

Discussion Panels: [Room 1](#) [Room 2](#)

Student Research Presentations: [Room 1](#) [Room 2](#)

WELCOME

The Sports Analytics Association aims to foster student interest in the application of analytical methods to sports, to widen the scope of sports analysis, and to provide opportunities for students to develop a broad range of analytical tools. This conference is an extension of those goals. We will provide an outlet for university students of all levels to exhibit their own original research, learn marketable data skills, and learn from industry leaders. We will continue to make this forum available to students for years to come. This is just the beginning.

Follow us on Twitter [@OSU_SAA](#) #OSUSAAC and visit our [website](#) for future events.

Organizers and Moderators

This event is the result of the hard work of faculty and students volunteering their time and energy. We appreciate the support of the Sports and Society Initiative, Department of Management Sciences at the Fisher College of Business, and the Economics Department at the Ohio State University.



John Draper: Faculty Co-Advisor



Ryan Ruddy: Faculty Co-Advisor



Trenton Butz: Treasurer of SAA



Noah Ritchie: Vice President of SAA



Hunter Green: Secretary of SAA



AJ Turner: Marketing Director of SAA



Alex Harding: SAA Member



Alex Williams: Former President of SAA



Nayan Patel: President of SAA



Bailey Weinstein: SAA Member



THE OHIO STATE UNIVERSITY

Sports and Society

— INITIATIVE —

This event has been made possible by a co-sponsorship from the [Sports and Society Initiative](#). SSI has helped to develop sports analytics at Ohio State in many ways including student research competitions, research grants, and helping to develop a sports analytics course currently offered in the Economics department.

SSI is a cross-discipline, collaborative endeavor that spans the university community and beyond, with the goal of deepening understanding of society through its intersection with sports. To accomplish this goal, it works to build knowledge and opportunities for students, gather and creating data to advance impactful research across an array of disciplines, and engage the community through meaningful conversations.

Keep apprised of SSI's upcoming events through their [website](#) or by following [@SportsSociety](#) on twitter. For more information please [email SSI](#).

Schedule

8:00 AM - 8:10 AM

Welcome Hosted in [room 1](#)

8:10 AM -10:00 AM

Student Research Presentations:

Presentation Room 1

Baseball Presentations:

Projecting from the KBO to MLB	Ben Howell (University of Texas)
Inefficiencies in Veteran MLB Contracts	Sam Stankivicz (The Ohio State University)
Controlling Launch angle to Limit Damage	Noah Thurm (Georgetown University)
Saving Free Agency: Analyzing the Modern Major League Baseball Labor Market	Matthew Lehman (University of Chicago)
Was the 2020 Cincinnati Reds Offense Unlucky	Max Couruso (The Ohio State University)
WAR Ball: MLB War Optimization with Budgetary Constraints	Brett Stowell (Xavier University)
Applicable Methods in Measuring the Baseball Ability of a Left-Handed Catcher	Red Li (The Ohio State University)

Hockey and Soccer Presentations:

Using Shot Tracking and Transitional Play Data to Quantify the Systemic Differences Between Men's and Women's Hockey	Nayan Patel (The Ohio State University)
Salary Cap Allocation By Position: A Function of Year and Success	Jacob Eckert (The Ohio State University)
Machine Learning Approach to Predict NHL Scoring	Brad Behan (Michigan State University)
Poisson Modeling and Predicting English Premier League Goal Scoring	Quang Nguyen (Loyola University Chicago)

Schedule

Presentation Room 2

Football Presentations:

<u>The Quarterback and the Situation</u>	Isaac Spear (University of Pennsylvania)
<u>MAYFIELD: Machine Learning Algorithm for Yearly Forecasting Indicators and Estimation of Long-run Player Development</u>	Alex Williams, Ben Clarke, and Seth Brugler (The Ohio State University)
<u>Value of an Interception in the NFL</u>	Hunter Green (The Ohio State University)
<u>The Impact of Temperature on the NFL</u>	Patrick Smith (The Ohio State University)
<u>Finding the Next Defensive Player</u>	Becca Skolnick, Tamara Dzolic, & Rediet Habtegebriel (Denison University)
<u>Fantasy Football Consistency Analysis</u>	Aidan Corey (The Ohio State University)

Basketball Presentations:

<u>Linear Regression for Basketball Team Scores Analysis</u>	Hong_Jui Shen & WenYi Shi (Denison University)
<u>The Impact of Conference Tournament Momentum on NCAA Tournament Success</u>	Cole Wagner (The Ohio State University)
<u>The Effect of the "Hot Hand" in NBA Games</u>	Chenchen Zhou, Yingyan Ma, & Yuliang Li (The Ohio State University)
<u>The Impact of Foreign Born Players in China Basketball Association</u>	Cecily Yeung & Meilin Hou (The Ohio State University)
<u>Performance Analysis of Batsman against Spin Bowling and Fast Bowling in Cricket</u>	Swarup Ranjan Behera (Indian Institute of Technology Guwahati, India)

10:00 AM - 12:00 PM

Baseball Analytics with R Workshop hosted by Jim Albert

Hosted in [Room 1](#). Over two sessions, [Jim Albert](#) will cover the use of Lehman data set and statcast data. His code for the sessions will be available at his [github site](#).

12:00 PM - 12:50 PM

Lunch Break

Schedule

12:50 PM - 1:00 PM

Panel Introduction [Room 1](#)

1:00 PM - 1:45 PM

Hockey - [Room 1](#)

[Dani Chu](#), [Alison Lukan](#), [Zac Urback](#), and [Sam Ventura](#) have all made waves bringing new and creative ideas to the sport of hockey as well as pushing hockey analytics to the mainstream. Our guests will talk about the current state of analysis within the sport, where the future of hockey analytics lies, and other hot topics within the sport. Nayan Patel will be moderating the panel.

Ohio State Athletics- [Room 2](#)

[Dan Delucia](#), [Brad Goldberg](#), [Nick Domicone](#), [Riley Ross](#), and [Kyle Davis](#) all have important positions within The Ohio State athletics programs and The Ohio State sports analytics programs. From strength training and injury prevention to baseball and basketball, these five utilize analytics to improve performance in their respective athletic fields. The panel will be discussing the different ways they apply analytics here at Ohio state and on the college as a whole, as well as the use going forward. This panel is moderated by Bailey Weinstein.

2:00 PM -2:45 PM

Sports Business - [Room 1](#)

[Ryan Chenault](#), [Shelly Cayette](#), [Joseph Moeller](#), and [Joe Odoguardi](#) all have very big rolls in the sports world Ohio. These four each face different challenges each day on the business aspect of sports. The panel will discuss what it is like when analytics and business collide in the world of sports, and how they work towards making a sports team into a successful business. This panel will be moderated by Hunter Green.

Private Companies/Consulting- [Room 2](#)

[Greg Strizek](#), [Justin Kubatko](#), and [Meghan Chayka](#) have all found a niche in sports analytics where they do not work directly for a team. The three guests will discuss entrepreneurship, data collection, and consultant work in wide variety of sports including NASCAR, basketball, and hockey. The panel will be moderated by Ryan Ruddy.

Schedule

3:00 PM - 3:45 PM

Basketball - Room 1

[Neil Johnson](#), [Justin Kubatko](#), [Seth Partnow](#), and [Andre Snellings](#) have all made strides within the basketball industry by bring attention to and using analytics. Our guests will discuss the growing landscape of basketball analytics and their different career paths that led them to where they are today. Alex Harding will moderate this panel.

Baseball- Room 2

[Jim Albert](#), [Walter King](#), [Katie Krall](#), and [Keith Woolner](#) bring a variety of perspectives to America's pastime from the front office to academic research. The panel will discuss how they ply their analytical skills in furthering team development and formation as well as many of the current unanswered questions and the future of the game. Dr. John Draper will moderate this panel.

4:00 PM -4:45 PM

Football - Room 1

What is the past, present, and future of football analytics? Join us for a discussion with seasoned football analytics professionals [Josh Hermsmeyer](#), [Sara Bailey](#), [Eric Eager](#), and [Zach Feldman](#) to find out! These speakers bring varied and unique perspectives for their vision of the football analytics industry. Alex Williams will moderate this panel.

5:00 PM -5:45 PM

Analytics in Media - Room 1

Somewhere in between freelance writing and directing sports analytics operations at ESPN, [Alison Lukan](#), [Sara Ziegler](#), [Josh Hermsmeyer](#), [Brian MacDonald](#), [Neil Johnson](#) and [Seth Partnow](#) have all found a home in sports media. Our guests will discuss effective methods of presenting data to the general public and share how they've progressed in their careers to this point. Trenton Butz will moderate this panel.

5:45 PM -6:00 PM

Student Research Award Announcements and Closing Remarks - Room 1

Guests



Jim Albert recently retired as Distinguished University Professor at Bowling Green State University. His interests include Bayesian modeling, statistical computing, and applications of statistical thinking in sports. He has written or co-authored books on Bayesian modeling, the use of the statistical language R, and applying statistical thinking in baseball.



Sarah Bailey is the Manager of Analytics and Data for the Los Angeles Rams. Prior to the Rams Bailey earned her Masters degree in Statistics from Simon Fraser University where she did research in baseball, hockey, and basketball, completing a thesis in baseball. Prior to attending Simon Fraser she interned with the San Diego Chargers as a digital media analyst, and earned her undergraduate degree in mathematics from University of the Pacific.



Shelly Cayette Weston is the Senior Vice President of Global Partnerships for the Cleveland Cavaliers where she started with the organization in August of 2012. Shelly oversees the Corporate Partnership Department including 40+ Team Members that span across the Cleveland Cavaliers (NBA), Cleveland Monsters (AHL), Canton Charge (G League), and Cavs Legion (Esports). Her responsibilities include leading key personnel to drive brand campaigns and platforms for clients, cultivate new business partnerships for the organization, as well as retain and grow current partner brand strategies and revenue. Previously, Shelly comes from New Orleans, Louisiana where she held positions in Marketing, Sponsorship, and Community Investment for the former New Orleans Hornets (currently New Orleans Pelicans) and Harrah's Casino for 8 years prior to moving to Cleveland, Ohio and has been recognized by the Sports Business Journal as a 2019 Game Changer and 2019 Crain's Business Cleveland 40 Under 40 honoree.

Guests



Meghan Chayka Meghan is an entrepreneur and the Co-Founder/CEO of Stathletes, a hockey data and analytics company providing insights to over 22 leagues worldwide. She was featured on the insert cover of the Hockey News' Top 100 of power and Influence, listed at 95 overall, Top 10 for Hockey Business Executives and awarded Top Young Entrepreneur of the Year for 2018/2019 by the Ontario Chamber of Commerce.



Ryan Chenault Ryan Chenault was named vice president of marketing for the Columbus Blue Jackets in June 2019. In this role, he is responsible for creating and implementing a comprehensive marketing strategy for the hockey club with emphasis on brand development, e-marketing, advertising, promotion, media planning/buying, creative services, market research and integration with social media. Prior to joining the Jackets, Chenault served as team president of the Ohio Machine of Major League Lacrosse for nearly three years, a tenure that was highlighted by an MLL championship in 2017. With the Machine, he was responsible for all business and lacrosse operations. He also established and chaired the Ohio Machine Lacrosse Foundation, the nonprofit extension of the organization. Chenault's previous experience also includes ten years with Northwestern where he led day-to-day management of the athletic department's marketing efforts for all 19 varsity programs. Chenault also served on the athletic department's senior administrative team and was a sports administrator for the women's lacrosse program. Prior to his time at Northwestern, Chenault also spent time with the Chicago Wolves of the American Hockey League and the athletic department at Xavier University. He holds a pair of undergraduate degrees from The Ohio State University and a master's degree from Xavier University. He currently resides in New Albany with his wife, Kristin, and two children, Andrew and Madeline.



Dani Chu is a Quantitative Analyst with the Seattle Kraken. He recently graduated with his Masters in Statistics from Simon Fraser University. At SFU, he was the co-president of the SFU Sports Analytics Club with Lucas Wu and Matthew Reyers. Along with Lucas, Matt and James Thomson, he was the winner of the College Division of the 2019 NFL Big Data Bowl and the 2018 Sacramento Kings Case Competition. Dani has also interned as a statistician at the NBA, Best Buy Canada and Fraser Health Authority.

Guests



Kyle Davis begins his fifth season as the Video & Operations Coordinator for the Ohio State men's basketball program in 2020-21. Davis is a two-time Ohio State graduate with a Bachelor's degree in sport industry (2013) and a Master's in sports management (2015). Davis' main duties include oversight of the program's video operations including opponent and self-scouting, breaking down recruit film, and statistical analysis combined with video & analytics. He also assists in the daily operations management, including on-campus recruiting activities, gameday and travel logistics, directing the creative content team, and organizing camps and clinics.



Dan DeLucia is in his second season as the Buckeyes' pitching coach after spending six as Ohio State's volunteer assistant coach. DeLucia has been a part of the Buckeye coaching staff that has guided to the team to three NCAA Regional appearances in the last five years as well as two Big Ten Tournament titles. He helped guide the youngest pitching staff in the nation to a postseason title run as the OSU arms shattered the school record for strikeouts in a season (583). DeLucia also took over the first base coaching duties during the team's title run in 2019.



Nick Domicone joined The Ohio State University in September of 2018. As the Director of Sports Science, he works to improve performance and reduce injury risk of all student athletes at OSU. In 2014, Nick received a Bachelor of Science in Biomedical Engineering with a specialization in Biomechanics from the Ohio State University. Immediately after graduation, Nick worked as an engineer for TS Tech Americas, Inc. where he developed and tested new-model seats for Honda vehicles. In his two years there, he was the project lead for seven automotive projects, which are all currently in production.

Guests



Eric Eager is the Executive Director of Research and Development at PFF, where he analyzes data for all 32 National Football League teams and as well as over 80 college football teams. Before joining PFF in 2018, he was a professor in the Department of Mathematics and Statistics at the University of Wisconsin – La Crosse, where he published over 20 papers in mathematical biology and the scholarship of teaching and learning.



Zach Feldman is a researcher for NFL Next Gen Stats. He attended Ohio State University during his undergraduate education, studying Economics. Feldman has previously held analytics positions with NumberFire, STATS LLC, Hoop-Ball.com, and the Columbus Destroyers Arena Football Team.



Brad Goldberg spent the 2020 season as the Buckeyes' director of pitching development and has been elevated to volunteer assistant coach for the Ohio State staff. Assisting pitching coach Dan DeLucia with the Buckeye staff in 2020, Ohio State led the nation with 12.7 strikeouts per nine innings. Goldberg, a third team All-Big Ten selection in 2013, went 6-1 with a 2.99 earned-run average in 15 starts for the Buckeyes as a senior before being selected in the 10th round of the 2013 Major League Baseball Draft. The Beachwood, Ohio, native spent six seasons in pro ball between the White Sox and Arizona Diamondbacks organizations, earning a call up to the big leagues in 2017, making 11 appearances for Chicago.

Guests



Josh Hermsmeyer is a football writer and analyst at FiveThirtyEight. He is also the author of the Air Yards Buy-Low Model, a weekly column on establishtherun.com which aims to predict the upcoming fantasy football success of wide receivers by gauging the players which have recently underperformed their opportunity.



Neil Johnson has worked with ESPN Analytics since 2016. The 2013 graduate from Ohio State (B.S. in Computer Science & Engineer) primarily works with basketball data and has specific interests in machine learning techniques and applying analytical analysis across all levels of competition.



Walter King is an Ohio State alum (B.S. Economics, '16) who works in Football R&D for the Las Vegas Raiders. He previously spent 3 1/2 seasons with the Los Angeles Angels working as an Analyst, Baseball Operations Assistant, and Intern.

Guests



Katie Krall was hired by the Cincinnati Reds as a Baseball Operations Analyst in January 2020. In her role she develops and integrates new tools and technology to improve Baseball Operations decision making processes, provides comprehensive scouting coverage, and statistical request support to the front office. Krall previously worked for a year and a half at Major League Baseball in the Commissioner's Office in New York City where she advised Clubs on 40-man roster management, MLB rules and compliance, major league administration, and salary arbitration. She was the first female chosen for MLB's inaugural Diversity Fellowship program, an initiative designed to promote women and people of color into front office executive roles. She graduated from Northwestern University with a B.A. in History and is currently pursuing her MBA from the University of Chicago's Booth School of Business.



Justin Kubatko is Chief Data Officer at StatMuse, the leader in natural language processing for sports. He is perhaps best known for creating the popular website Basketball-Reference.com. Prior to StatMuse, Justin served as Vice President at Sports Reference LLC. TIME magazine recognized Sports Reference as one of the "50 Best Websites" of 2010, and the company won an Alpha Award for "Best Analytics Innovation/Technology" at the 2013 MIT Sloan Sports Analytics Conference. Justin has also served as a statistical consultant for the Portland Trail Blazers and the NBA's Statistics & Analytics Technologies group. He has a bachelor's degree in mathematics from Grove City College and a master's degree in Applied Statistics from The Ohio State University.



Alison Lukan is a freelance data-driven storyteller covering the NHL. She also currently serves as the co-host of the Too Many Men podcast.

Guests



Brian MacDonald is currently the Director of Sports Analytics in the Stats & Information Group at ESPN. He was previously the Director of Hockey Analytics with the Florida Panthers Hockey Club, and an Associate Professor in the Department of Mathematical Sciences at West Point. He received a Bachelor of Science in Electrical Engineering from Lafayette College, Easton, PA, and a Master of Arts and a Ph.D. in Mathematics from Johns Hopkins University, Baltimore, MD.



Joseph Moeller is currently his tenth season with the Browns and his third as the team's director of business analytics. In his role, Moeller and the Browns business analytics department are responsible for collecting and structuring business data, evaluating the information obtained and incorporating findings to support multiple areas of the organization's day-to-day operations, including identifying opportunities to improve Browns fans' experience throughout the year. He initially joined the franchise's information technology department in 2011 before becoming a member of the business analytics team in 2014. Moeller is a graduate of Hiram College, where he received degrees in computer science and biology.



Joe Odoguardi joined the Ohio State University Department of Athletics in 2015. He currently serves as Executive Associate Athletic Director & CFO. Odoguardi is primarily responsible for the planning, monitoring and managing of all the athletic department's financial activities. He is also responsible for all of the department's core business functions including financial forecasting, budgeting, analysis, reporting, and business operations. Odoguardi began his career at the Greater Columbus Convention Center in 1982 as a Junior Accountant, and left that organization, as its Chief Operating Officer, in 1998, to come to Ohio State as the Chief Financial Officer of The Jerome Schottenstein Center. Odoguardi is originally from Youngstown Ohio. He graduated from Youngstown State University with a degree in accounting, and has an MBA from the University of Dayton. He lives in Dublin, Ohio with his wife Susan, and has four daughters, three are Ohio State graduates, and one that is currently attending, and is in her junior year.

Guests



Seth Partnow currently covers the NBA and the world of basketball analytics for The Athletic. In the past, he has worked as the Director of Basketball Research for the Milwaukee Bucks and he was the managing editor of the basketball analytics website Nylon Calculus. Seth earned a B.A. in Economics from Carleton College and a J.D. from the University of Minnesota Law School.



Riley Ross joined the staff as the Director of Wrestling Sports Science & Development prior to the 2018-19 season. He is responsible for integrating analytics, technology and strength and conditioning strategies, and tasked with developing Buckeye wrestlers to their full genetic potential. He owns unique background in performance, assessment and prescriptions, as well as analytics in the tactical and private sector.



Andre Snellings is a Senior Writer and on-air personality at ESPN, focusing on fantasy basketball, becoming a two-time FSWA Basketball Writer of the Year. During his previous career in neural engineering, he was a senior writer for Rotowire, and a contributor for Nylon Calculus, TYTSports, Fansided, and Sumit Hoops. He earned his BS in electrical engineering at Georgia Tech, before getting his Ph.D in biomedical engineering at the University of Michigan and becoming a research fellow at Duke University.

Guests



Greg Strizek is the Director of Research for Strategic Analytics, Inc. He leads a research effort supporting the NASCAR Sprint Cup program for a major automobile manufacturer. He has worked with Achievement Metrics to develop models predicting NFL player behavior and performance based upon personality traits identified through analysis of player speech. In addition to his sports analytics work, Strizek has worked for a number of government and corporate clients.



Zac Urback is a hockey analyst with the Columbus Blue Jackets. Previously he was director of player evaluation/assistant to the GM for the Mississauga Steelheads.



Sam Ventura is the Director of Hockey Research for the Pittsburgh Penguins and an affiliated faculty member at Carnegie Mellon University's Department of Statistics & Data Science. Prior to that, he was a professor of Statistics at CMU, where he also received his PhD (Statistics, 2015). He previously served as an assistant coach for Carnegie Mellon's ice hockey team and as faculty advisor to the CMU Sports Analytics Club. His academic research focuses on clustering, prediction, record linkage, synthetic data, infectious diseases, and sports (particularly hockey and football). He is an associate editor for the Journal of Quantitative Analysis in Sports. Sam has co-authored multiple R packages for open-source data collection and analysis, including `nhlscrapr`, `nflscrapR`, and `spew`, and he co-founded `war-on-ice.com`. He co-organizes the annual Carnegie Mellon Sports Analytics Conference.

Guests



Keith Woolner enters his 13th season in Baseball Analytics with Cleveland, having joined the Indians organization in 2007. In his role, he drives innovative research to improve the organization's in-game strategy, player forecasting and acquisition decisions through the use of technology, data management, machine learning and statistical analysis. Prior to joining the Indians, Keith was Director of Research & Development at Baseball Prospectus for 10 years. He has co-authored 10 books, including several editions of the Baseball Prospectus annual and the award-winning *Baseball Between The Numbers*. Keith is the inventor of VORP (Value Over Replacement Player), a well-known sabermetric statistic. He worked in the software industry for 15 years before joining the Indians, including stints at Oracle and SAS Institute. He holds dual Bachelor's degrees in Mathematics with Computer Science and Management Science from the Massachusetts Institute of Technology (M.I.T.), and a master's degree in Management Science & Engineering from Stanford University.



Sara Ziegler is the sports editor at FiveThirtyEight and the host of the *Hot Takedown* podcast. She worked as an editor at the *Omaha World-Herald* for 15 years and is the president of ACES: The Society for Editing. Sara earned a B.A. in journalism and political science from Iowa State University.

Research Presentation

Abstracts

Projecting from the KBO to MLB

Ben Howell (Ben.Howell@utexas.edu)

University of Texas at Austin

The popularity of the Korean Baseball Organization (KBO) sky-rocketed this year when ESPN broadcasted their games in place of the postponed Major League Baseball (MLB) season. The KBO Wizard is a resource created to manually track pitches from KBO pitchers to provide advanced statistics and insights into the KBO. The 26,000+ pitches collected and tagged inspired a project creating a projection system using these insights to predict MLB success for KBO players transitioning to MLB. The project investigates the differences between MLB and KBO styles of play and how that transition may affect players in terms of their plate discipline and batted ball data. Through linear regression analysis, this project worked to identify some important indicators of MLB success as measured by Wins Above Replacement (WAR) then adjusted those indicators for the KBO's different style of play. With those adjusted indicators, the project analyzed the collected data to produce an estimate of MLB success and potential. Drastically different playing environments have made comparing the KBO to MLB difficult, a problem exacerbated by a lack of publicly available pitch-by-pitch data for foreign leagues. The KBO Wizard and this project aim to address that lack of data and explore how to project player performance from one league to another.

Inefficiencies in Veteran MLB Contracts

Sam Stankivicz (stankivicz.1@osu.edu)

The Ohio State University

This research delves into a potential inefficiency for younger players in the MLB and clubs favoring veterans and past production over the deterioration a veteran player will inevitably experience. Using the Lahman database (version 7.1) built into R, along with calculations from FanGraphs to find wOBA and FIP, models were constructed to figure out a base correlation of performance and age to a player's salary. Then, lag and lead models were constructed to show how much better or worse a salary correlates when stacked up against a prior or future year's numbers. From both of these sets of models, age and contract status affect how much performance metrics affect a player's salary. It can be said that veterans are being paid more for similar production than younger players, even those in arbitration and early free agency. Also, a definite ex post mindset can be seen as salary can be better correlated with performances up to 3 years prior to that contract. Overall, this tends to show a large inefficiency in club operations where veterans are favored more over young, potentially marketable stars for equal levels of production.

Research Presentation

Abstracts

Controlling Launch Angle to Limit Damage

Noah Thurm (njt17@georgetown.edu)

Georgetown University

Successful pitchers limit damage by minimizing the quality of contact they allow, as pitchers tend to focus on some combination of deception, movement, and location to try and miss barrels. I propose that the most important pitcher-influenced variable to quality of contact is Launch Angle, and understanding and influencing it ought to be a priority for all pitchers. It is clear that Exit Velocity is the single most important predictor of a batter's success, but that relationship cannot be manipulated much, if at all, by any pitcher. Across baseball, batters' Exit Velocity distributions are much tighter than their Launch Angle distributions. This means pitchers are likely better able to directly influence Launch Angle than Exit Velocity, which is quite "sticky" around the mean for a given hitter. As batted ball trajectories move towards the positive and negative extremes of launch angle, outcomes suffer *at all exit velocities*. This suggests a new pitching approach centered around "Launch Angle Deflection," or the attempt to induce weak contact and get outs by "deflecting" batted balls to the nearest extreme (and therefore suboptimal) launch angle given a player's batted ball profile. A series of variable importance tests across pitch types returned **pitch height** as the most significant pitch-level variable to launch angle (~40% VI). The recommended strategy is to throw predominantly high pitches to batters with high median launch angles, and throw mostly low pitches to batters with low average launch angles, pushing each batter to their nearest unproductive extreme.

Saving Free Agency: Analyzing the Modern Major League Baseball Labor Market

Matthew Lehman (matthewlehman@uchicago.edu)

University of Chicago

Over the past several years, MLB free agency has faced amplified scrutiny from players, agents, and fans over its inability to fulfill its foundational purpose: improving the salaries and status of MLB athletes. Common criticisms include the slowing rate of free agent activity, the growing prevalence of non-competitive "tanking" teams, and the failure of players – especially MLB's "middle class" – to recoup their fair value on the market. I analyzed the progression and success of the 2009-2020 MLB labor system from the perspective of these present public concerns. Following established methodology, I constructed linear regressions to project expected free agent contract values. For players who signed contract extensions prior to testing free agency, I developed additional regressions to evaluate their actual contract results against their projected worth. I particularly focused on current criticisms, namely shifts from 2016-2020 and shifts among middle-class players. I find that over the past twelve seasons, contract extensions have become progressively more lucrative than free agent contracts. This has been especially true for middle-class free agents from 2016-2020. Both of these conclusions substantiate public criticisms about the diminishing returns of free agency. Interestingly, I also find that more upcoming free agents are entering the weakened free agent market than before, indicating further supply-side contract concerns in addition to the demand-side complications confirmed in this paper. While these results are crucial for individual players to consider, they also underscore imperative problems with – and provide evidence of potential solutions for – the troubling direction of MLB's current labor market structure.

Research Presentation

Abstracts

Was the 2020 Cincinnati Reds Offense 'unlucky'?

Max Caruso- (caruso.140@buckeyemail.osu.edu)

The Ohio State University

For my presentation I am looking at the 2020 Cincinnati Reds offense. It was a mystery throughout the whole season why the offense wasn't having success. They ranked above league average in XWOBA, Barrel%, XSLG, etc. They ranked amongst the bottom 3 teams in the league in BA, runs/game, BABIP, etc. In my presentation I think I found out the real reason the offense didn't have a lot of success. The Reds ranked amongst the worse in the league in hitting with RISP and 2 outs. This is an area where teams like the Dodgers and Padres did very well.

WAR BALL: MLB WAR Optimization with Budgetary Constraints

Brett Stowell (stowellbs@xavier.edu)

Xavier University

In this paper we create "All Star" teams based on ex post player WAR from 2018. We form eight person teams for the eight position players on the field and find the highest WAR team that could be formed from experienced players. We only consider players that have been in the league long enough to have reached the free agent time frame of their career. These are the players we consider to have their salary based on free market conditions. After establishing the highest WAR team that could be formed with these players, we perform constrained linear optimization to determine the highest WAR team that could be formed at different levels of spending. We create six different teams based on constrained budgets from \$20 million to \$123 million.

In our team formations, 23 different players are selected with 11 players selected for multiple teams and the remaining 12 selected for 1 team only. It would not actually be possible to form teams the way we have done so in the paper because many of the players are committed to multiple year contracts and are not freely available as we have assumed. However, teams could do similar analysis with players that they expect to be available using forecasts of future WAR and expected salary. Additionally, it is important to note that team decisions include a full 25 (or more) person roster along with minor league players that might be needed in the event of injury or poor performance. Finally, it is important to recognize that we have treated the optimization problem as a single period (season) game. In practice teams are likely to consider a longer horizon and may have single year or multiple years in which they would like to maximize wins while maintaining a certain level of wins for the other years. These non-focus years would become additional constraints in the model.

Research Presentation

Abstracts

Applicable Methods in Measuring the Baseball Ability of a Left-handed Catcher

Qingyi (Red) Li (li.9142@buckeyemail.osu.edu)

The Ohio State University

Ever since Benny Distefano step out of the diamond in 1989, there had been no other left-handed catcher on the stage of the Major League Baseball. According to a research published in 1996, there were only 27% players who were lefties whereas 63% were right-handed over the 110 years of major league baseball being played at the time¹. While the demand for left-handed pitcher is high, along with some demand for left-handed first-baseman and outfielders, it seems like there is no other position available for left-handed players in present-day Major League Baseball. In fact, there were 14 lefty catchers² in the history of the Major League Baseball, meaning that left-handed catchers are capable of being as competitive as right-handed catchers at professional level. Regardless of the numerous assumptions people make about why there is no more left-handed catchers in the higher-level baseball, there is no solid scientific evidence against left-handed catchers. The focus of this research is to determine an effective method of studying if left-handed catchers are really disadvantaged in the game of baseball. The main methods studied will include, but not exclusive to, data analytics and video analysis. The purpose of this research is to find out if there is an effective way in measuring the baseball ability of left-handed baseball catchers in comparison with that of a right-handed baseball catcher.

Using Shot Tracking and Transitional Play Data to Quantify the Systemic Differences Between Men's and Women's Hockey

Nayan Patel (patel.3000@osu.edu)

The Ohio State University

This project looks into the systemic differences between how men's and women's hockey is played on the tactical level. The game is identical except for the important "no body checking" rule present in the women's version of the game. I attempt to see how that rule changes how the game is played. I chose to compare men's DI and women's DI as it's the only level of play in the world where men and women play on the same level and have access to most of the same resources. By manually tracking micro-stats such as zone entries, zone exits, and neutral zone play, as well tracking shot attempt locations, I try to tease out how the "no checking" rule causes the game to be played differently on a systems level. To do this, I tracked every Big Ten men's hockey game and every WCHA women's hockey game from the last season, recording zone data in a spreadsheet and the shot data into Alyssa Longmuir's Hockey Plotter Shiny app to build heat maps.

Research Presentation

Abstracts

Salary Cap Allocation By Position: A Function of Year and Success

Jacob Eckart (eckert.138@buckeyemail.osu.edu)

The Ohio State University

Every year before the regular NHL season starts, GM's spend weeks crafting a roster they believe will help them eventually win the Stanley Cup. A major factor in creating these rosters is determining what percentage of the salary cap to allocate to each position. This research shows the relationship between salary cap allocation by position as a function of time and success. This project looks at how every team over the last nine seasons has spent their available salary cap by position as well as how successful teams (defined as teams earning on pace of 99 points in a season) differ from the average. Using linear regression analysis and covariance matrixes, this project analyzed the average spending per position as well as its relationship to success in the regular season. From the data collected, it became apparent that teams spending more money than average on forwards and defenders tended to earn more points in the regular season. A weak relationship between spending money on goalies, as well as a slight decrease in goalie salary cap allocation, shows that teams are not very good at valuing goalies. The data shows that the top teams have recognized this fact and have avoided spending big on goalies.

"Machine Learning Approach to Predict NHL Scoring"

Bradley Behan (behanbra@msu.edu)

Michigan State University

Hockey Analytics is utilized by teams in every league across the world. Statistical methods can be used to evaluate contracts, draft prospects, trades, and player/team evaluation. The area of hockey analytics that is persistently trying to evolve is player success prediction. In this report, we will be using linear regression to estimate player performance over the first 3 seasons of their NHL careers. This study will be limited to forwards who have played in the Canadian Hockey League (CHL), a conglomerate of 3 junior leagues across Canada and the United States.

Research Presentation

Abstracts

Poisson Modeling and Predicting English Premier League Goal Scoring

Quang Nguyen (nminhquang.99@gmail.com)

Loyola University Chicago

The first purpose of this research was to verify the consistency between goal scoring in the English Premier League and the Poisson process; specifically, the relationships between the number of goals and the Poisson distribution, the time between goals and the exponential distribution, and the time location of goals in a match and the continuous uniform distribution. It turned out to be that the Poisson process and the three probability distributions do perform a great job of describing Premier League goal scoring. In addition, Poisson regression was utilized to predict a Premier League season's results, using different sets of season data and with a large number of simulations being involved. Various metrics were examined and compared, including chances of being the champions, finishing in the top four and bottom three, and relegation points.

"The Quarterback and the Situation"

Isaac Spear (ibsppear@sas.upenn.edu)

University of Pennsylvania

The fundamental attribution error—or the tendency to underestimate the role of situational influences in behavior—has been called “as robust and reliable a phenomenon as any in the literature on person perception” (Quattrone, 1982, p. 376). Despite its prevalence in social psychology literature, there has been little evidence for the fundamental attribution error outside of the laboratory. To investigate whether people commit the fundamental attribution error in the real world, we examined popular assumptions about the factors that lead to quarterback success in football and used publicly available collegiate and professional performance data to assess the validity of these assumptions. In Study 1 (N = 70), participants rated individual ability as nearly twice as important for quarterback success as situational influences. In Study 2, we collected college performance data for quarterbacks drafted into the NFL from 2006-2016 (N = 131) and found that 27 individual ability factors combined to predict only 6.8% of the variance in quarterback success in the NFL, whereas two situational factors—passes dropped and quarterback protection—predicted 18.9%. Together, these findings provide evidence outside of the lab for the fundamental attribution error in quarterback evaluation.

Research Presentation

Abstracts

MAYFIELD: Machine Learning Algorithm for Yearly Forecasting Indicators and Estimation of Long-run Player Development

Alex Williams (williams.5889@osu.edu), **Ben Clarke** (clarke.256@osu.edu), **Seth Brugler** (seth.brugler@gmail.com)

The Ohio State University

Accurate statistical prediction of American football player development and performance is an important issue in the sports industry. We propose and implement a novel, fast, approximate k-nearest neighbor regression model utilizing locality-sensitive hashing in highly dimensional spaces for prediction of yearly National Football League player statistics. MAYFIELD accepts quantitative and qualitative input data, and can be calibrated according to a variety of parameters. Concurrently, we propose several new computational metrics for empirical player comparison and evaluation in American football, including a weighted inverse-distance similarity score, stadium and league factors, and NCAA-NFL statistical translations. We utilize a training set of comprehensive NFL statistics from 1970-2019, across all player positions and conduct cross-validation on the model with the subset of 2010-19 NFL statistics. Preliminary results indicate the model to significantly improve on current, publicly available predictive methods. Future training with advanced statistical datasets and integration with scouting-based methods could improve MAYFIELD's accuracy even further.

Value of an Interception in the NFL

Hunter Green (green.2103@buckeyemail.osu.edu)

The Ohio State University

Currently in the NFL when thinking of how to measure defensive back success the first thing that comes to mind is interceptions. At first glance this sounds like a good idea, but this stat doesn't really show much of a story, instead just counts the number of times the player picked the ball off. While this is valuable knowledge, this stat doesn't really show the impact behind the number.

This project focuses on that impact of these crucial plays by looking into what makes an interception more valuable. Through many linear regressions using data from the Big Data Bowl, it shows which stats go into giving certain intersections producing more expected points. It then gives each player a value adjusted interception rank, showing the flaws of ranking players just by interception count alone. Then it takes this to a team level showing how some teams get much more value from their interceptions than others, making their defense more impactful in games. The project is meant to highlight the impact behind different picks showing that not every pick has the same impact, and there is much more to the story than just an interception count when evaluating a defense or player's interception impact.

Research Presentation

Abstracts

The Impact of Temperature on the NFL

Patrick Smith (smith.12338@buckeyemail.osu.edu)

The Ohio State University

At the beginning of the NFL season in early September, the weather is more than ideal. As the season goes on, however, the weather starts to become less and less ideal. The temperature drops, the wind picks up, and snow starts falling. While several NFL teams avoid this weather change by playing in a covered dome, most teams do not have this luxury. For my project, I will be attempting to answer the question: Are NFL teams from cold places better in cold weather than teams from warmer areas? More specifically, I will be analyzing the impact cold weather has on offensive efficiency/production as well as the outcome of games for NFL teams, and if this impact varies depending on where the team is from. I will be using data from the nflfastR package in R and from the National Centers for Environmental Information for this investigation.

Finding the Next Defensive Player

Becca Skolnick (skolni_b1@denison.edu), **Tamara Dzolic** (dzolic_t1@denison.edu),
and **Rediet Habtegebriel** (habtegeg_r1@denison.edu)

Denison University

Predicting the efficacy and quality of a defensive player in the NFL from their college and combine performance is something the draft and scouts aim to do on a regular basis. In our study, we utilize the combine and statistics from each player's senior season of college, and the NFL 2019 season in an attempt to expose a relationship between an individual player's growth and the team's success. Success in our study is defined by the team's ranking in the top or bottom three teams of the NFL as of Super Bowl LIV. Using a principal component analysis coupled with a linear regression model, we attempt to find a relationship that links the explanatory variables: percentage of solo tackles, quarterback hits, 40 yard dash, 20m shuttle, bench press, broad and vertical jumps, to the success of the NFL team associated with each player. The ultimate goal of this study is to identify a relationship between our response and predicting variables so as to predict the ability of a defensive player to contribute to a team's success on gameday.

Research Presentation

Abstracts

Linear regression for basketball team scores analysis

Hong_Jui Shen (shen_h1@denison.edu) and **WenYi Shi**
(shi_w1@denison.edu)

Denison University

At present, there is a lot of prediction for team scores in sports. Therefore, we selected the data in the basketball reference website. Mainly focused on the overall game data of the Los Angeles Lakers, Milwaukee Bucks, and Houston Rockets from 2019 to 2020. And through the average performance of the entire team's players on the field from 2019 to 2020 to predict the next game's score. We will use Principle Component Analysis to make the composition of the main factors of a game. Through these factors, use linear regression to predict the team's next game score.

The Impact of Conference Tournament Momentum on NCAA Tournament Success

Cole Wagner (Wagner.1383@osu.edu)

The Ohio State University

This work employs ordinary least squares (OLS) linear regression analysis to quantify the momentum built by conference-tournament winners as they head into the NCAA Men's Basketball Tournament, colloquially known as "March Madness." With a robust dataset comprising team factors and statistics from the previous nine (completed) college basketball seasons (2011-2019), the analytical model suggests that the average conference-tournament winner advances *no further* in the NCAA Tournament than non-conference tournament winners, all else equal. Following the examination of these results as well as discussion of prospective extensions and improvements this paper briefly overviews the model's forecasting capacity and proposes additional applications based on this potential.

Research Presentation

Abstracts

The Effect of the “Hot Hand” in NBA Games

Chenchen Zhou (zhou.1835@osu.edu), Yingyan Ma (ma.1790@osu.edu), and Yuliang Li (li.9013#@osu.edu)

The Ohio State University

Our group's project aims to uncover the mystery of “hot hand” in basketball games. The project retrieves the data from the past NBA seasons(2019-2020). We make collections on players' shooting performance when the previous shots by them were successfully made. We study if the second shots can be statistically proved to have a higher shooting percentage comparing to the players' season averages, and look for the correlation between the improvement in shooting percentages and numbers of offensive rounds taken before the second shooting attempts were made, if the improvement in shooting percentage does statistically exist. In our first attempt of sample data selection, we retrieved data from all the play-off games Miami Heat, Los Angeles Lakers, Denver Nuggets, and Boston Celtics played, expect the finals played between Heat and Lakers. The founding was not disappointing: the shots did not show obvious difference or improvement comparing to the expected shooting percentage that we calculated based on the players play-off average. In our second round of sample data selection, we will works on the regular season games of those four teams. With a relatively large sample size, we are excited to see if the results would differ, and if the differences can be explained by the different offensive and defensive styles between the play-off games and regular season games.

The Impact of Foreign Born Players in the China Basketball Association

Cecily Yeung (yeung.80@osu.edu) and Meilin Hou (hou.402@osu.edu)

The Ohio State University

Since the 1995-96 season of China Basketball Association (CBA), all teams are-permitted to have two registered foreign-born players on their teams. In our study, we hypothesis that teams win game has a relationship with foreign-born players through analyzing the relative important of numbers of games foreign players played in season, minutes played by the foreign players, the points they earned to team win percentage in past five years between 2015-2020 seasons. All the players whose nationality is not China are included as foreign players. Players from Hong Kong and Tai Wan are included as Chinese players.

Since the sample we chose basically satisfies the assumptions for multiple linear regression model, we run a standard multiple regression to estimate the cross-section correlation between the variables that describes the performance of foreign players and the winning percentage of the team. We used Gretl and Excel to analyze the data and establish the correlation. Our expectation is that the average game played by foreign players in percentage, the total minutes played by foreign players in parentage, and the total points earned by foreign players should all have positive relationship with the winning percentage.

The regression tells us that only the coefficient for the minutes played by foreign players in percentage has a positive significant coefficient. The the total minutes played by foreign players in parentage has a positive coefficient, but it is not statistically significant. The coefficient of total points earned by foreign players is significant, but it has a negative correlation with the winning percentage. Therefore, we concluded that there are significant correlations between the average game and points earned by foreign players with the win percentage.

Research Presentation

Abstracts

Performance Analysis of Batsman against Spin Bowling and Fast Bowling in Cricket

Swarup Ranjan Behera (b.swarup@iitg.ac.in)

Indian Institute of Technology Guwahati, India.

In today's world, as sports gets more competitive than ever, players and teams are looking for ways to get an edge over their rivals. The progressive trend of analyzing vast amounts of data has also emerged in cricket, as it brings a significant advantage against other teams in the championship. The traditional statistics, such as strike rate and average, are the raw counts of certain cricket events. While interesting, they fail to provide a complete picture of the match. Cricket text commentaries, on the other hand, are written narratives that give a detailed description of a minute-by-minute account of the game while it is unfolding. In this work, we analyze the performance of batsman against spin bowlers and fast bowlers using publicly available ball-by-ball text commentary data. We have extracted the bowling and batting features from the text commentary data and represented it in the form of a confrontation matrix. Every element in this matrix corresponds to how the batsman is confronted with the spin (or fast) bowlers. We have employed a dimensionality reduction method, namely correspondence analysis, to identify the performance of batsman and biplots to visualize the identified performance. We have analyzed the performance of more than 500 players by analyzing text commentaries over one million deliveries for 13 years (2006-2019).

Fantasy Football Consistency Analysis

Aidan Corey (corey.80@buckeyemail.osu.edu)

The Ohio State University

The following project analyzes the effects of player consistency on his team's win percentage. To answer this question, I first set up a chart of six players averaging 15 fantasy points a game, with each standard deviation increasing from one player to the next. I then simulated 120 matchups for each possible matchup of players and tallied the total wins, losses, and ties among each player. I then made a separate chart in which I retained the same average and deviation for each player but changed the direction in which each game is deviated from the average 15 points per game. My results displayed an oscillating function about the 50% Win % line, in which the ideal player has a low standard deviation with the majority of their games slightly above the average and minimal games slightly below the average. This oscillation of the graph also revealed that the more inconsistent a player becomes, the effects seen at higher consistency levels dwindle, and the inconsistent players' win % closes in on 50%. From a statistical perspective, this study has implications for fantasy football owners as it can lead to better decision making when choosing between players with similar scoring averages, whether it be for the beginning-of-the-year draft or for starting roles on a week-to-week basis.