New England Symposium on Statistics in Sports

NESSIS

PROGRAM

September 29, 2007

Harvard University
Science Center, Lecture Hall C
1 Oxford Street
Cambridge, Massachusetts 02138

2007 New England Symposium on Statistics in Sports

September 29, 2007

Symposium Co-Organizers:

Mark E. Glickman, Department of Health Policy and Management, Boston University School of Public Health

Scott R. Evans, Department of Biostatistics, Harvard University School of Public Health

Sponsors:

- Boston Chapter of the American Statistical Association
- Section on Statistics in Sports of the American Statistical Association
- Section on Operations Research in Sp*OR*ts of the Institute for OR and the Management Sciences (INFORMS)
- Harvard University Department of Statistics

Acknowledgments: We wish to thank everyone who helped to make this event possible. Special thanks go in particular to Tom Lane for creating and maintaining the NESSIS web site and sending out BCASA announcements, Jim Cochran for helping to contact potential speakers and serving as the liaison to the Section on Operations Research in Sp*OR*ts of INFORMS, Ben Alamar for contacting potential speakers and for creating a special NESSIS issue of the Journal of Quantitative Analysis in Sports, and Mike Zarren for arranging a pre-conference get-together at the Fours and contacting potential speakers. We would also like to thank Betsey Cogswell, Ellen Weene, Dale Rinkel, Peter Bohlin, Carl Morris, Iram Farooq, and Rui Wang for their parts in helping with the conference.

2007 New England Symposium on Statistics in Sports

September 29, 2007

Schedule

All talks take place in the Harvard University Science Center, Lecture Hall C. The morning and afternoon breaks will be in the open area outside the lecture hall.

9:00am – 9:15am:	Welcome
9:15am – 9:45am:	Alan Schwarz, New York Times *"Can't we all just get along: Overcoming impasses between sports insiders and outsiders"*
9:45am – 10:15am:	Justin Wolfers, The Wharton School *"Racial Discrimination among NBA referees"*
10:15am – 10:45am:	A. James O'Malley, Harvard Medical School *"The Tennis Formula: How it can be used in professional tennis"*
10:45am – 11:00am:	Break
11:00am – 11:30am:	Dan T. Rosenbaum, Cleveland Cavaliers and UNC at Greensboro *"The pot calling the kettle black: Are NBA statistical models more irrational than 'irrational' decision-makers?"*
11:30am – 12:00pm:	Shane T. Jensen, The Wharton School *"Statistical models for the evaluation of fielding in baseball"*
12:00pm – 12:30pm:	Peter Vint, U.S. Olympics Committee *"Real-world challenges in the pursuit of Olympic excellence"*
12:30pm – 1:45pm:	Lunch break
1:45pm – 2:15pm:	Cyrus R. Mehta, Cytel Inc., and Arun M. Mehta, iClub Inc. *"iClub: A digital feedback device for improving your golf swing"*
2:15pm – 2:45pm:	Ben Baumer, New York Mets and CUNY *"Why on-base percentage is a better indicator than batting average"*
2:45pm – 3:15pm:	Jan Vecer, Columbia University *"On probabilistic excitement of sports games"*
3:15pm – 4:30pm:	Poster Session
4:30pm – 5:30pm:	Panel Discussion – *"Incorporating statistics into sports."* Moderator: Sean Grande – sportscaster, "voice" of the Boston Celtics Panelist: Aaron Schatz – Football Outsiders Panelist: Dean Oliver – Denver Nuggets Panelist: Carl Morris – Harvard University

Oral Presentation Abstracts

WHY ON-BASE PERCENTAGE IS A BETTER INDICATOR THAN BATTING AVERAGE

Ben Baumer†

New York Mets and City University of New York

† E-mail: *bbaumer@nymets.com*

Batting Average (AVG) and On-Base Percentage (OBP) are two of the most commonly cited statistics in baseball. While many researchers have observed that these statistics differ in the extent to which they reflect a hitter's true ability, we will use a simple algebraic decomposition to explain much of that discrepancy, and interpret the remainder within the theoretical framework of Defense Independent Pitching Stats (DIPS). Specifically, we will prove that OBP depends more heavily upon a particularly unpredictable variable, hits per balls in play, than does batting average. This result, interpreted in the context of DIPS, will explain why for both batters and pitchers, on-base percentage is a better indicator of future performance than batting average.

STATISTICAL MODELS FOR THE EVALUATION OF FIELDING IN BASEBALL

Shane T. Jensen†

Department of Statistics, The Wharton School, University of Pennsylvania

† E-mail: *stjensen@wharton.upenn.edu*

The use of statistical modeling in baseball has received substantial attention recently in both the media and academic community. We focus on a relatively under-explored topic: the use of sophisticated statistical models for the analysis of fielding. We combine high-resolution ball-in-play data with spatial modeling in order to evaluate the performance of individual fielders. We also present a hierarchical Bayesian model for the sharing of information between players at each position. We also discuss our procedure in the context of past procedures for fielding evaluation. This is joint work with Kenny Shirley and Abraham Wyner in the Department of Statistics at The Wharton School.

iCLUB: A DIGITAL FEEDBACK DEVICE FOR IMPROVING YOUR GOLF SWING

Cyrus R. Mehta†[1], Arun M. Mehta[2]

[1] *Cytel Inc.;* [2] *iClub Inc.*

† E-mail: *mehta@cytel.com*

New motion-capture technologies that dont rely on work-intensive multi-camera setup are revolutionizing the game of golf by creating new opportunities to analyze the golf swing. At the recent "Better Golf Through Technology" symposium hosted by MIT, a hundred golf professionals, technology engineers and scientists gathered to sample the latest technological innovations available that can be used to improve their or their students game. The Body Motion System by iClub is a strap-on vest that easily and accurately digitizes, stores, and plays back a golfers torso movement during their swing, including measurements such as X-Factor, Shoulder Rotation Speed, Spine Tilt, and Hip Pivot (see www.iclub.net). 50 subjects used the vest to gather data before and after receiving training utilizing one of two similar swing-technique drills. The hypothesis tested was that the two drills would be equivalent in their effects on performance. A brief demonstration of the technology involved will be shown, as well as a demonstration of the training drills used.This will be followed by a statistical analysis of the data. New innovations, such as the Body Motion System, have made it possible to statistically test hypotheses in the golf industry that were previously un-testable because the data could not be easily and accurately gathered.

THE TENNIS FORMULA: HOW IT CAN BE USED IN PROFESSIONAL TENNIS

A. James O'Malley†

Department of Health Care Policy, Harvard Medical School

† E-mail: *omalley@hcp.med.harvard.edu*

In this talk I will derive expressions for the probability that a player wins a game of tennis, the "tennis formula," and also the probability of them winning the whole match. Under independence and homogeneity assumptions, the tennis formula reduces to an easy to evaluate expression. We study the important properties of the tennis formula and how tennis players can use the formula to improve their odds of beating a particular opponent. We then use the tennis formula to predict the probability of important events such as: winning a game when down a break point, winning a set when losing or ahead, winning a tiebreaker. By comparing the actual winning probabilities to the values predicted by the tennis formula, we evaluate which players perform the best under pressure. Finally, we use the tennis formula to assess which players might benefit the most and which players will benefit the least from proposed rule changes in tennis.

THE POT CALLING THE KETTLE BLACK: ARE NBA STATISTICAL MODELS MORE IRRATIONAL THAN "IRRATIONAL" DECISION-MAKERS?

Dan T. Rosenbaum[†1], David Lewin[2]

[1] *Cleveland Cavaliers and University of North Carolina at Greensboro;* [2] *Macalester College*

[†] E-mail: *rosenbaum@uncg.edu*

Statistical models of player productivity, especially those in academics, have often made a strong claims about the "rationality" of NBA decision-makers. Yet these statistical models have rarely been subjected to any rigorous examination of their ability to forecast the future. We examine how well several player productivity metrics, including (a) John Hollinger's Player Efficiency Rating, (b) Wages of Wins Wins Produced, and (c) the NBA Efficiency metric do in predicting future team wins and future player productivity as measured by plus-minus statistics. In addition to a comprehensive examination of the player productivity metrics used by NBA statistical analysts, this paper is the first academic presentation of plus-minus statistics. Our findings suggest that models that assume very simplistic NBA decision-making often outperform more sophisticated models.

CAN'T WE ALL JUST GET ALONG: OVERCOMING IMPASSES BETWEEN SPORTS INSIDERS AND OUTSIDERS

Alan Schwarz[†]

New York Times

[†] E-mail: *alan@alanschwarz.com*

What sports statistics mean and how they should be used has long been the subject of vehement debate between two groups far apart: team management and media directly involved in the games, and outsiders (often academics) who explore new ideas and strategies through the numbers. For most of the 20th century there has been little middle ground or understanding between the two sides, who often get along about as well as the Hatfields and the McCoys.

Alan Schwarz, a staff reporter for The New York Times, has been arguably the most prominent mainstream sports journalist trying to translate sophisticated statistical work for an often reluctant mainstream audience. Besides writing the *Keeping Score* statistics column on Sundays since 2004, Schwarz has also written stories that created national news, from coverage of academic studies on mental problems suffered by former NFL football players to a paper that claimed to find an own-race bias in the calling of fouls by NBA refereeing

crews. Schwarz will discuss the challenges he faces attempting to foster dialog and understanding between two sides, sports insiders and outsiders, who often share a glaring enmity and distrust.

ON PROBABILISTIC EXCITEMENT OF SPORTS GAMES

Jan Vecer†

Department of Statistics, Columbia University

† E-mail: *vecer@stat.columbia.edu*

In this talk we introduce a quantitative measure of the excitement of sports games. This measure can be thought of as the variability of the expectancy of winning as a game progresses. We illustrate the concept of excitement at soccer games for which the theoretical win expectancy can be well approximated from a Poisson model of scoring. We show that in the Poisson model, higher scoring rates lead to increased expected excitement. Given a particular strength of a team, the most exciting games are expected with opponents who are slightly stronger. We apply this theory to the FIFA World Cup 2006 games, where the winning expectancy was independently estimated by betting markets. Thus it was possible to compute the expected and the realized excitement of each given game from the trading data. Joint work with Tomoyuki Ichiba and Mladen Laudanovic.

REAL-WORLD CHALLENGES IN THE PURSUIT OF OLYMPIC EXCELLENCE

Peter F. Vint†

United States Olympic Committee

† E-mail: *peter.vint@usoc.org*

At the elite level of Olympic competition, differences between medal and non-medal winning performances are often on the order of fractions of a percent. I will describe several areas in which sport scientists and performance specialists at the United States Olympic Committee are currently working and will share some of the real-world challenges we face that could be of interest to the quantitative statistics community. Specific examples will include:

Identifying and prioritizing deterministic factors of performance: Quantifying the importance of factors that define the outcome of an athletic performance is a multifactorial problem. We need to identify analytical techniques which can help us understand which factors are most important to performance, most sensitive to training, and most relevant to differentiating between various levels of performance.

Modeling multi-event, multi-element performance: Outcomes in sports like team gymnastics, decathlon, and modern pentathlon are based on cumulative scores from different events. Given the known weighting of specific events and competencies of our athletes, we are interested in modeling optimal training, competition, and team selection strategies that could help maximize overall scores.

Data mining from notational analysis: Notational or match analysis is a technique by which specific information from a competition is recorded and reviewed for trends or measures like work intensity or work-rest ratios. The nature of these data make it difficult to extract meaningful information that can be used by our coaches and athletes. We need to identify tools and techniques that can more readily extract trends or situational probabilities from these data.

RACIAL DISCRIMINATION AMONG NBA REFEREES

Justin Wolfers[†1], Joe Price[2]

[1,2] *The Wharton School, University of Pennsylvania*

[†] E-mail: *jwolfers@wharton.upenn.edu*

The NBA provides an intriguing place to test for taste-based discrimination: referees and players are involved in repeated interactions in a high-pressure setting with referees making the type of split-second decisions that might allow implicit racial biases to become evident. Moreover, the referees receive constant monitoring, and feedback on their performance. (Commissioner Stern has claimed that NBA referees are the most ranked, rated, reviewed, statistically analyzed and mentored group of employees of any company in any place in the world.) The essentially arbitrary assignment of refereeing crews to basketball games, and the number of repeated interactions allow us to convincingly test for own-race preferences. We find thateven conditioning on player and referee fixed effects (and specific game fixed effects)that more personal fouls are awarded against players when they are officiated by an opposite-race officiating crew than when officiated by an own-race refereeing crew. These biases are sufficiently large that we find appreciable differences in whether predominantly black teams are more likely to win or lose, according to the racial composition of the refereeing crew.

Poster Presentation Abstracts

IMPROVED PARK FACTOR ESTIMATES FOR BETTER MLB PLAYER EVALUATIONS

Rohit Acharya[†1], Haibo Lu[1], Carl Morris[2]

[1]*Harvard Sports Analysis Collective, Harvard University;* [2]*Department of Statistics, Harvard University*

[†] E-mail: *racharya@fas.harvard.edu*

Park factors (PF), applied to runs scored, often (e.g. ESPN) are calculated for the 30 MLB parks as follows (RS, RA = runs scored for and against a team and G = games played):

$$\text{PF} = ((\text{homeRS} + \text{homeRA})/(\text{homeG})) \;/\; ((\text{roadRS} + \text{roadRA})/(\text{roadG}))$$

This PF formula stems from a reasonable assumption, but when applied directly its failure to account for unbalanced schedules produces bias, and its estimates are overly variable. E.g. Dodger Stadium's PF values jumped from being a significant pitcher's park in 2004 to a hitter's park in 2006. Kauffman Stadium's (KC) PF was 3rd best and 2nd best for hitters in 2006 and 2002, but 24th in 2004; and Cincinnati's PFs follow a pattern similar to KC's. We aim to reduce both bias and variability via better statistical estimation methods and with augmented data. We will include information about parks that have made structural changes, and will relate results to park measurements, e.g. foul area, fence distances and heights, elevation, humidity, turf, day/night play, and wind.

Accurate PF measures are key to adequately assessing batter and pitcher performances. Fans, the sports media, and post-season award voters should be aware of PFs when considering player evaluations, to properly discount players in advantageous surroundings. From a strategic viewpoint, general managers especially need accurate PF values to evaluate properly free agents and trade proposals.

ISOLATING THE EFFECT OF INDIVIDUAL LINEMEN ON THE PASSING GAME IN THE NATIONAL FOOTBALL LEAGUE

Benjamin C. Alamar[†]

Menlo College

[†] E-mail: *quantsports@gmail.com*

Protecting the quarterback is an integral part of the passing game in the National Football league, yet the relationship between the abilities of an individual lineman and the effectiveness of a passing game remains unexplored. One of the principal reasons for this lack of study is the absence of publicly available data that is needed in order to track the performance of a specific lineman. In order to create the relevant data, several NFL games were charted. The performance of each lineman was recorded on every pass play, as well as the amount of undisturbed time the quarterback was given (time in the pocket) to make a throw. These data were used in a series of regressions to determine how likely a lineman was to successfully hold his block in relation to the time it took for the quarterback to throw the ball, for each lineman in the sample. These data were also used to estimate the correlation between time in the pocket and completion rate. The results of these regressions were then used to simulate the effects that different linemen have on the passing game. The analysis shows the position on the line with the highest variance in performance to be center which indicates that improving at the center position would have the biggest impact on a teams passing game.

A COMPOSITE-POISSON MODEL FOR GOAL SCORING

Phil Everson[†1], Paul Goldsmith-Pinkham[2]

[1,2] *Swarthmore College*

[†] E-mail: *everson@swarthmore.edu*

Goal scoring in sports such as Hockey and Soccer is often modeled as a Poisson process. We work with a Poisson model where the mean goals scored by the home team is the sum of parameters for the home team's offense, the road team's defense, and for a home advantage. The mean goals for the road team is the sum of parameters for the road team's offense and for the home team's defense. The best teams have a large offensive parameter value and a small defensive parameter value. A level-2 model connects the offensive and defensive parameters for the k teams. Parameter inference is made by imagining that goals can be classified as being strictly due to offense, to (lack of) defense, or to home advantage. Though not a realistic description, such a breakdown is consistent with our model assumptions, and we can work out the conditional distributions and generate random partitions to facilitate inference about the team parameters. We use the conditional Binomial distribution, given the Poisson totals and the current parameter values, to partition each observed goal total at each iteration in an MCMC algorithm.

SKILL IMPORTANCE IN BYU WOMEN'S VOLLEYBALL: A BAYESIAN APPROACH

Lindsay Florence[†1], Gilbert W. Fellingham[2]

[1,2]*Department of Statistics, Brigham Young University*

[†] E-mail: *gwf@byu.edu*

The BYU womens volleyball team recorded all skills (pass, serve-receive, set, etc.), rated each skill, and recorded rally outcomes (point for BYU, rally continues, point for opposition) for the entire 2006 home volleyball season. Only sequences of events occurring on BYUs side of the net were considered. Events followed one of these general patterns: serve–outcome, serve-receive–set–attack–outcome, or block–dig–set–attack–outcome. These sequences of events were assumed to be Markov chains where the quality of each contact depended only on the quality of the previous contact but not on contacts further removed in the sequence. We represented these sequences in an extensive matrix of transition probabilities where the elements of the matrix were the probabilities of moving from one state to another. Counts of events in each row of the transition matrix were assumed to be multinomial. A Dirichlet prior was formulated for each row, so posterior estimates of the transition probabilities were then available using Gibbs sampling. The different paths in the transition matrix were then followed through the possible sequences of events at each step of the MCMC process to compute the posterior probability density that a perfect serve-reception resulted in a point, a perfect set resulted in a point, etc. These posterior probability densities are used to address questions about relative skill importance in BYU womens volleyball.

HERRLIN-MASSEY MODEL: A BAYESIAN ANALYSIS OF FANTASY BASEBALL

Craig Massey[†1], Dan Herrlin[2]

[1,2]*San Diego State University*

[†] E-mail: *dherrlin@gmail.com*

Fantasy sports generated three billion dollars in revenue last year alone (FSTA, 2006). With the great interest that has been sparked by fantasy sports, and in particular fantasy baseball, many player indices have become available. The goal of these indices is to rank overall player value from highest to lowest. While these indices rank player performance for the immediate future (ie. upcoming season), they are static and incoming data cannot be used to update the index.

The following research develops a Bayesian fantasy player index for the 2006 Major League Baseball Season. The basis of any Bayesian analysis is to use prior events to update and predict future outcomes. This shift from past performance to future production is exactly what fantasy mangers attempt to predict when they select their players. As a result, this index incorporates multiple statistical categories and random nuances of the game (age, injuries, peak seasons, etc.) from previous seasons to predict player performances for the 2006 season.

The attractive aspect of using a Bayesian framework is that in creating our 2006 fantasy index, we obtained an adaptive model that can be used where previous indices could not. Our model can be updated at any point in the season, where additional data is available. Thus, the Herrlin-Massey Model can be used throughout the season to evaluate trades etc. which is not possible with a static index.

PROBABILITY AND OPTIMIZATION MODELS FOR RACING

Victor S. Y. Lo†

Fidelity Investments and University of British Columbia

† E-mail: *Victor.lo@fmr.com*

This presentation will review quantitative models for multi-entry racing competitions (e.g. horse races, dog races, car races, boat races). First, various probability distributions (exponential, normal, gamma) for modeling running times will be discussed and tested on actual race data. The implications of these distributional assumptions for estimating higher order ranking probabilities (e.g., predicting those finishing second and third) will be discussed. These higher order probabilities are critical to many types of sports where second and third positions are also important and are traditionally hard to estimate in practice because of the involvement of high-dimensional integrations. Approximations to the higher order ranking probabilities will then be proposed and analyzed mathematically and numerically so that they can be used in real practice. Last but not least, nonlinear stochastic programming models will be employed to optimize wages in actual data using higher order ranking probabilities along with their distributional assumptions. Examples on horse race data will be used to illustrate the techniques.

RATING THE RATINGS: A COMPARISON OF METHODS FOR RANKING COLLEGE ICE HOCKEY TEAMS

Robin Lock[1], Julia Palmateer†[2]

[1,2] *St. Lawrence University*

† E-mail: *jwpalm03@stlawu.edu*

A number of different methods can be used to rank teams that play unbalanced schedules as is often the case in NCAA sports where teams are organized into leagues based on geography and institutional commitment to the sport. These rankings may play an important role in comparing teams, for example in the selection and seeding of teams for national championship tournaments. Passionate debates occur when comparing the merits of a team with a great record from a relatively weak conference to a mid-level team in a power conference. In this paper we compare the accuracy and variability of several methods currently being used to rate NCAA Division I ice hockey teams. These methods include raw winning percentage, a rating percentage index (RPI) taking into account strength of opponents, a Bradley-Terry based model (KRACH Kens Rating of American College Hockey) and multiplicative model based on Poisson scoring rates (CHODR College Hockey Offensive/Defensive Rating). Our analysis is based on Monte Carlo simulations of many seasons where we can know and manipulate the relative strengths of the teams and then check the performance and stability of the various rating methods.

NOT ALL .300 SEASONS ARE CREATED EQUAL A PERMUTATION METHOD TO MEASURE CLUTCH PERFORMANCE

Shaun Lysen†

The Wharton School, University of Pennsylvania

† E-mail: *slysen@wharton.upenn.edu*

The 2005 American League MVP race was one of the more controversial races in recent history. Alex Rodriguez of the New York Yankees was named the American League MVP despite the belief by many that David Ortiz of the Boston Red Sox should have been. Specifically, the critics noted that Rodriguez often got his hits at less important times, e.g. when no one is on base, or when his team is already winning by a large margin. Ortiz, on the other hand, got his hits at the most important opportunities, which would in some sense, make him more "valuable" to his team. We examine this by first casting the batting average in a loss function framework. Specifically, the batting average is simply one minus the empirical risk under 0/1 loss. Then, we propose a more sensible loss function based on the expected runs that would score given the current state of the game. This gives us at bat weights so that we can weight important at bats more heavily, e.g. bases loaded with 2 outs. After that, we propose a general permutation idea that can be used for any baseball metric where we have different weights. For example, if we permute which weighted at bats that a batter got his hits, our statistic such as batting average will change. We can generate the full permutation distribution and find the corresponding p-value. This p-value can be viewed as a measure of clutch. Finally, we apply this permutation method to historical data.

THE PASSING PREMIUM PUZZLE RECONSIDERED

Duane W. Rockerbie†

Department of Economics, University of Lethbridge, Alberta

† E-mail: *rockerbie@uleth.ca*

The passing premium puzzle, suggested by Alamar (2006), queries why NFL coaches choose a fairly even number of passing and running plays on offense when historical data suggests that the expected return to passing is significantly greater than running. This paper explores this puzzle in greater depth by analyzing both league-wide and team-specific return (yardage) distributions for passing plays and running plays that are more detailed than used by Alamar (2006). The analysis also makes use of risk-averse regression functions used in economics for the 2006 NFL season. The results suggest that the passing premium is much smaller than thought for the entire league, and that there may exist a running premium for some teams. The paper argues that this is the natural result of specialization by coaches according to comparative advantage.

A MARKOV MODEL FOR BASKETBALL

Kenny Shirley†

Department of Statistics, The Wharton School, University of Pennsylvania

† E-mail: *kshirley@wharton.upenn.edu*

I develop a discrete-time Markov chain model for basketball in which the states are defined in terms of which team has possession of the ball, and how that team gains possession. The simplest model has $2 \times 4 = 8$ states, one for each combination of team (2 teams) and possession type (4 types: Inbound pass, Defensive Rebound, Offensive Rebound, and Steal). The state space can be expanded to model free throws and three-point shots, so that the exact number of points scored by each team is known for each transition; such models have approximately 16–20 states. Simulations from 2003–2004 NBA data indicate that transition probabilities can be well-estimated using existing data, like shooting percentages, rebounding percentages, and turnover counts, although these estimates depend on certain assumptions. To estimate transition probabilities without such assumptions, transition counts must be made by watching games; a relatively small sample of such data is analyzed. With a large sample of such data, the potential of the model is virtually unlimited. The most obvious use is to estimate in-game win probabilities. Results from the Markov model are compared to a previously developed Brownian motion model for basketball. The Markov model can also be used to analyze (1) the relationship between the number of possessions in a game and each teams chances of winning, (2) the relative impact of defensive rebounds, offensive rebounds, and steals on scoring, and (3) the importance of certain recently-developed basketball statistics, such as deflections, which represent transitions in the Markov model.

SOURCES OF THE HOME COURT ADVANTAGE IN NBA BASKETBALL

Dylan Small†

Department of Statistics, The Wharton School, University of Pennsylvania

† E-mail: *dsmall@wharton.upenn.edu*

The home team in the NBA wins about 60% of the time. Several potential sources of this home court advantage are rest (more teams play away without a day off), the home crowd and players being able to sleep in their own homes. In this work, we investigate the sources of the home court advantage in the NBA. We fit a mixed effects model and examine whether there is evidence of variation in teams home court advantages and the factors that are associated with it, in particular the role of the home attendance on a teams home court advantage. In addition, we examine the role of rest in the home court advantage and how a teams road disadvantage depends on how long it has been on the road.

DOES POLE POSITION REALLY MATTER: MIXED MODEL RELATIONSHIP OF MOTOGP QUALIFYING PRACTICE TIMES AND FINAL RACE TIMES

Leanne Streja†

Department of Biostatistics, UCLA

† E-mail: *lstreja@ucla.edu*

In professional motorcycle racing MOTOGP, the location of where the rider starts the race has been thought to be very important in determining the race outcome. Riders compete in practice for the pole position or best location with respect to the starting line. This position is granted to the rider who has the fastest lap time achieved at the qualifying practice the day before the actual race. The location of all riders with respect to this pole position and the starting line is known as the race grid. However, it is of question whether this grid location is truly associated with final race time, and thus is qualifying practice time really a good predictor of final race time. In this analysis, we use a mixed model to assess the relationship of qualifying practice best lap times, which determine grid location, with the final race time for MOTOGP 2007 races. We examine the relationship of the variables, using several possible covariance structures. Additionally, we evaluate the impact of other possible random and fixed effects: racer, tire type, team, country of origin for the rider, bike, ground conditions, weather conditions, length of track, and number of laps.

PAIRED COMPARISON MODELS FOR RANKING NATIONAL SOCCER TEAMS

Andrew W. Swift†1, Shawn E. Hallinan[2]

[1]*Department of Mathematics, University of Nebraska at Omaha,* [2]*PHARMetrics*

† E-mail: *aswift@mail.unomaha.edu*

National soccer teams are currently ranked by soccers governing body, FIFA. Although the system is thorough, taking into account many different factors, many of the weights used in the systems calculations are somewhat arbitrary. It is investigated here how the use of a statistical model might better compare the teams for ranking purposes. By treating each game played as a pair-wise comparison experiment and by using the Bradley-Terry model as a starting point some suitable models are presented. A key component of the final model introduced here its ability to differentiate between types of matches.

A GENERAL FRAMEWORK FOR COMPARING AND IMPROVING THE JUDGMENTAL FORECASTS OF SPORTS PROFESSIONALS – APPLICATIONS OF AN EXPANDED LENS MODEL EQUATION

Ryan M. Taylor†

Department of Public Administration and Policy, Rockefeller College, University of Albany, SUNY

† E-mail: *ryantaylor79@gmail.com*

This paper discusses the development of an expanded Lens Model Equation (LME) to decompose correlation coefficients of judgmental forecasts and task systems. Such a framework facilitates the comparison of forecasting skill across decision makers in professional sports as well as identification of deficiencies in individual forecasters. The expanded LME can increase efficiency in resource allocation and improve quality of judgments and decision making. Utility of such a model in wide-ranging task contexts are demonstrated in two areas: (1) decomposing team wins in relatively-uncertain conditions and (2) improving forecasts of player value for contract negotiations using time-series data. While variants of this framework can be readily adapted for use by professionals in any sports context, this paper applies the model to data from Major League Baseball and the National Basketball Association.

A PATH ANALYTIC EXAMINATION OF HOME-FIELD ADVANTAGE IN THE NATIONAL FOOTBALL LEAGUE: THE MEDIATING ROLE OF PENALTY DIFFERENTIAL

Daniel E. Tuccitto†

Department of Applied Physiology and Kinesiology, University of Florida

† E-mail: *dtuccitto@hhp.ufl.edu*

Previous research has indicated that there exists a home-field advantage in the National Football League (NFL).Although several explanations for this advantage have been offered (e.g., travel, crowd support, officiating, park familiarity), few have been examined empirically. In addition, few, if any,studies on home-field advantage have employed robust multivariate statistical techniques (e.g., structural equation modeling) to explore the complexity of the phenomenon. Therefore, using data from 256 NFL games played during the 2005 and 2006 regular seasons, I tested a series of structural equation models estimating the influence ofvenue (home/away) on victory (win/loss), as well as the role of penalty differential as amediating variablein the venue/victory relationship. Furthermore, I conducted multi-group invariance analyses to determine if the mediating role of penalty differential is influenced by (a) whether or not a given game is between divisional opponents, and (b) whether or not a game was closely contested. In addition to further supportingthe existence of an NFL home-field advantage, results revealed significant mediation for penalty differential, and provided preliminary evidence thatthe mediating role of penalty differential is unaffected by the divisional or closely contested nature of an NFL game. As research on the NFL home-field advantage remains in its nascent stages, further research examining increasingly complex models of the phenomenon, is clearly necessary.

IN SEARCH OF THE ADVANTAGE TO BATTING LAST IN BASEBALL

Theodore L. Turocy†

Department of Economics, Texas A&M University

† E-mail: *turocy@econmail.tamu.edu*

It is taken as an article of faith that there is a strategic advantage to batting last in baseball. Yet, data corroborating that position is hard to come by. For example, it is an empirical regularity that the first-batting team wins a majority of long extra-inning games; this result, however, may be the result of selection bias. To complement this, this paper uses a Markov game approach at the play level to modeling a baseball game. It is easy to compute optimal strategies in this setting, despite the large number of possible states. Using this model,

it is shown that "one-run" strategies such as the sacrifice bunt and stolen base cannot justify a significant advantage to batting last, and including a defensive strategy such as the intentional walk results in an advantage to batting first. In all cases, though, the advantage is trivial, amounting to one or two extra wins per thousand games. It is argued that in order to generate a significant advantage, strategy would have to have a significant effect at the pitch, rather than the play, level. Since the defense makes a significant decision – the type and location of pitch – on every pitch, it is plausible that there is a significant strategic advantage to batting first. Finally, it is hypothesized that the persistence of conventional wisdom that batting last is an advantage can be explained by a framing bias which incorrectly equates the opportunity to score points with strategic leverage.

WHEN TO WALK? A STATISTICAL APPROACH TO INTENTIONAL WALKS

Michael B. Valdes†

Harvard Sports Analysis Collective, Harvard University

† E-mail: *mvaldes@fas.harvard.edu*

Do intentional walks actually put a team in a better position to minimize runs allowed? A measure that is relatively in-game specific can be obtained by relying upon the Net-Expected Runs Value table, which displays the expected amount of runs a team will score from a particular situation until the end of the inning. Utilizing readily obtainable statistics and the aforementioned table, it is possible to determine the expected value of a particular players at-bat in a given in-game situation (i.e. Runner on 2nd, 2 Outs).

First, take the difference in expected runs between the new situation that would exist for each possible outcome in that at-bat (i.e. Hitter hits a single, sacrifice fly, double play, etc.) and the existing situation. Second, multiply that difference by the probability of that outcome happening for that particular player. The sum of all run differentials weighted by the respective probabilities is the expected value of that players at-bat in that situation. By comparing this expected value with the value obtained when taking the difference between the situation, post-intentional walk, and the initial situation, it is possible to determine whether it is optimal for the team to walk the batter or pitch to him.

The results obtained when performing these calculations for major league hitters have demonstrated the potential over reliance and misuse of the intentional walk. The data implies that walks should be reserved for few players and primarily in 2 out situations, contrary to the prevailing conventional wisdom that walks are useful when setting up double plays.

A SIMPLE AND FLEXIBLE RATING METHOD FOR PREDICTING SUCCESS IN THE NCAA BASKETBALL TOURNAMENT

Brady T. West†

University of Michigan–Ann Arbor

† E-mail: *bwest@umich.edu*

This poster will first briefly review potential rating tools and methods for predicting success in the NCAA basketball tournament, including those methods (such as the Ratings Percentage Index, or RPI) that receive a great deal of weight in the selection and seeding of teams for the tournament. A simple and flexible rating method based on ordinal logistic regression and expectation (the OLRE method) that is designed to predict success for those teams selected to participate in the NCAA tournament will then be introduced. A method of simulating tournament outcomes based on the parametric Bradley-Terry model for paired comparisons will also be introduced, and the ability of the computationally simple OLRE method to predict success in the tournament will be compared to that of the simulation method, using actual NCAA tournament data from 2006 and 2007. The OLRE method can incorporate several different predictors of success in the NCAA tournament when calculating a rating, and will be shown to have comparable or even better predictive power than the model-based simulation approach. Limitations of the OLRE method, mostly related to the collection of historical predictors of success, will also be discussed, along with additional potential applications of the method.

GENERALIZED PAIRED-COMPARISON MODELS FOR COMPARING TEAM STRENGTHS

Jayson D. Wilbur†[1], Gregory J. Matthews[2], Andrew W. Swift[3]

[1]*Department of Mathematical Sciences, Worcester Polytechnic Institute;* [2]*Department of Statistics, University of Connecticut;* [3]*Department of Mathematics, University of Nebraska at Omaha*

† E-mail: *jwilbur@wpi.edu*

In this talk we consider two generalizations of paired-comparison models for sports data. In the first, we investigate the effect of data transformations on the model for NFL point-spread data of Glickman and Stern (1998). Robustifying transformations are employed to reduce the influence of "blowouts" on predictions and evaluated against the point spread. In the second, we consider the two-dimensional Bradley-Terry model of Causer and Husson (2005) for application to NFL and MLB data, identify some inadequacies and propose a Bayesian alternative.

ESTIMATING SITUATIONAL EFFECTS ON OPS

Phil Yates†

Department of Mathematics and Statistics, California State Polytechnic University, Pomona

† E-mail: *philipayates@gmail.com*

"What is the offensive value of Player A?" Of all the metrics that sabermetricians have developed to try to answer that question, OPS (on base percentage plus slugging percentage) has been one of the first for the mainstream media to slowly embrace as an alternative to batting average. Looking at baseball statistics for the Chicago Cubs on ESPN.com's baseball page, for example, one sees that the batting statistics are sorted by OPS as the default sort. Part of the appeal of OPS is that it is simpler to calculate than a more accepted metric in sabermetric circles like WARP (Wins Above Replacement Player); however, the raw value of OPS does have limitations such as not taking into consideration ballpark effects or the differences between the two leagues.

What if the question asked was "What is the offensive value of Player A in Situation B versus Situation C?" In 1994, Jim Albert used the Gibbs sampler to estimate what was the effect different in-game situations had on batting average. Examples of two such situations are a player's breakdown statistics home and away and hitting ahead and behind the count. By employing the Gibbs sampler multiple times on each situation, one can compute the situational effect on a player's OPS. The data will look at the hitting performance of major league regulars during the 2006 season that qualified for the batting title.