

**BOSTON COLLEGE**  
Department of Economics

**EC 370.01: Sports Econometrics (Spring 2016)**  
Campion 302: M W (4:30 – 5:45)

Christopher Maxwell  
[maxwellc@bc.edu](mailto:maxwellc@bc.edu)  
<http://www.cmaxxsports.com>

Maloney Hall, 337  
Hrs: TBA  
x2-8058 (no voicemail, svp)

**Course Description:** This is an advanced stats/econometrics course; it is not a sports history or trivia class. We'll be developing various statistical tools of analysis and then applying those tools to a wide variety of sport-related topics, perhaps including:

- forecasting team performance,
- the efficiency of wagering markets,
- the drivers of home field advantage in sports,
- the business and economics of professional team sports,
- measuring and valuing parity in sports leagues,
- the importance of population in driving competitive imbalance,
- the efficacy of leagues' competitive balance initiatives,
- peer effects in team performance,
- the relationship between performance and player compensation,
- understanding the what drives ticket prices,
- valuing draft picks,
- ... and so forth.

We could easily work with other data, but there's so much publicly available sports-related data available... so why not? ... and besides, it's so much fun!

We will also weave some sports economics into the course. Most of that material will focus on the notion of *competitive balance* and the *Uncertainty of Outcome Hypothesis*, which many would say is the most important concept in sports economics..

And again, this is not a sports history or trivia class.

**Prerequisites:** Intermediate Microeconomics (EC201 or EC203) and Econometrics (EC228 and/or EC327). Students are expected to know how to run simple econometric models (OLS: SLR and MLR) using Stata and to be comfortable interpreting regression results.

This course will make extensive use of both Excel and Stata:

- You should have worked with Stata in your Econometrics course. At the start of the semester, we will review how to access and run Stata through BC's apps server.

To avoid traffic jams with Citrix and the apps server, you may want to purchase a six-month Stata IC license for \$75 (sorry, but small Stata will not suffice for EC 370). For details, go to: <http://www.stata.com/order/new/edu/gradplans/student-pricing/>.

- This course also makes extensive use of Excel. You should not take this course if you do not have strong Excel skills. To brush up on your Excel skills, you might look at the materials

## EC 370: Sports Econometrics

assembled by the ITS department:

[http://www.bc.edu/offices/its/tandc/training/course\\_materials.html](http://www.bc.edu/offices/its/tandc/training/course_materials.html) .

Unfortunately, the features offered by Excel differ somewhat across platforms and over time. For this course, you may need to install the *Analysis ToolPak* and the *SolverAdd-in* if your Excel does not already offer those features. Let me know if you need help with these installations. One option: You can always run Excel on BC's apps server, which has these capabilities and is faster than you might imagine.

**Analytic tools/methods:** While the list may change, at the moment I anticipate that we'll be focusing on the following statistical tools of analysis, working with both Stata and Excel:

- SLR and MLR estimation and inference (review)
- Assessing importance/meaningfulness of estimates: elasticities and beta regressions
- Non-linear least squares (and the *residual*)
- Binary dependent variables – I: Linear (and truncated linear) probability models (LPMs)
- Functional forms: fixed effects, percentile dummies, polynomials, and splines (linear and cubic)
- Binary dependent variables - II Maximum likelihood estimation (MLE), and logit, probit and even arc-tangent models
- More about limited dependent variables: Ordered logit and probit, and perhaps censored and truncated regressions
- Runs, streaks and testing independence: chi squared tests, binomial tests, regression analysis and *runs* tests

**Applications:** We'll illustrate the analytic methods with multiple applications, working with sports-related data... often working in both Excel and Stata. The applications that I have in mind fall into the following general categories:

- The business of professional sports
- Strategy and valuing game-states
- Forecasting game/match outcomes (e.g. runs, wins, points, putts, etc.)
- Umpire/referee/judging bias (and home field advantage)
- Assessing player and team performance
- Efficiency of wagering markets
- Pay and performance (for players, coaches and teams)
- Momentum effects (streaks and runs)

**Texts:**

- **Required:** Tobias Moskowitz and Jon Wertheim, **Scorecasting: The Hidden Influences Behind How Sports Are Played and Games Are Won**, Three Rivers Press (paperback), 2012.
- **Recommended, but not required:** Rodney Fort, **Sports Economics**, Prentice Hall.

**Canvas:** All handouts, exercises, exercise answers, data, etc. will eventually be posted to the course's Canvas site.

## EC 370: Sports Econometrics

**Accommodations:** If you are a student with a documented disability seeking reasonable accommodations in this course, please contact Kathy Duggan (x2-8093; [dugganka@bc.edu](mailto:dugganka@bc.edu)) at the Connors Family Learning Center regarding learning disabilities and ADHD, or Paulette Durrett, (x2-3470; [paulette.durrett@bc.edu](mailto:paulette.durrett@bc.edu)) in the Disability Services Office regarding all other types of disabilities, including temporary disabilities. Advance notice and appropriate documentation are required for accommodations.

**Academic Integrity:** You will be held to Boston College's standards of academic integrity. If you have any questions as to what that means, please go to <http://www.bc.edu/offices/stserv/academic/integrity.html>.

**Pass/Fail:** It's perfectly OK to take this course Pass/Fail; however it is not fair to your fellow students to shirk on team assignments. I expect equal contributions by all team members, but history has proved that students taking the course Pass/Fail have at times unfortunately failed to pull their weight. And that's just not fair to the other students. If you are taking the course Pass/Fail, please let me know at the start of the semester.

**Laptops and wireless devices:** If you are in the classroom... turn them off; shut them down. If that doesn't work for you, then this course is not for you.

**Course Structure:** There are four components to the course; they are (%'s of course grade are in parentheses):

1. **Mid-Term Exam** (35%)
2. **Six-or-so Exercises** (40%)
3. **Research Paper and Presentation** (20%)
4. **Tuesday Topics/Participation** (5%)

### 1. **Mid-Term Exam** (35% of total grade):

There is one exam in this course... a Mid-Term exam towards the end of the semester covering the empirical methods and applications developed in this course. Exam grades are curved. The exam date is yet to be determined, but I'm thinking about Wednesday, April 27<sup>th</sup> (in the next to last week of classes).

Note: Only in extraordinarily compelling situations will I even consider the possibility of a "make up" exam. It is your responsibility to plan your schedule accordingly.

### 2. **Exercises** (40% of total grade):

I anticipate having six-or-seven-or-so exercises (of equal value) over the course of the semester. Final grades on exercises are curved.

These will typically be team assignments (usually with two students per team) lasting about two weeks. I will assign the teams, which will change from exercise to exercise. The set of exercises is not yet set, but here's what I have in mind at the moment:

- a. *The Sports League Challenge* (modeled around the NBA circa 2013; the economics of professional team sports; population driving competitive imbalance; the efficacy of

## EC 370: Sports Econometrics

various competitive balance initiatives; scored using NASCAR scoring rules). Here's the proposed schedule:

- 0\_SamePops: M 1/25 – F 1/29 (pre-season; all teams at pop=4.7M; does not count towards NASCAR scoring)
  - 1\_NBAPops: M 2/1 – W 2/10 (introduce three pop tiers: 2.2, 4.4 and 7.9)
  - 2\_RevShare: M 2/15 – W 2/24
  - 3\_SalCap: M 3/14 – W 3/23
  - 4\_Draft: W 3/30 – F 4/8
  - 5\_NBA: W 4/13 – F 4/22
- b. *Moneyball* revisited: *OBP* v. *SLG* (review of SLR and MLR analysis; measuring explanatory power; statistical significance v. importance/meaningfulness (economic significance); elasticities; *beta* regressions)
  - c. Valuing Draft Picks using Trade Data (NFL) (Weibull and exponential distributions; non-linear estimation; the importance (and at times, arbitrariness) of the residual)
  - d. *March Madness* and the *Ratings Performance Index* (RPI) (ratings models; model calibration; concordance; Kendall *tau* correlations; non-linear estimation)
  - e. Valuing Game States: Run production in MLB (*quick slant*; fixed effects; 11+ million observations)
  - f. Strategy and Field Position (NFL) (valuing game states; non-linear functional forms; tests of independence (chi squared, binomial, OLS regression, and Wald-Wolfowitz *runs* tests))
  - g. Wagering Market Efficiency (US and European football) (MLE estimation; linear probability models; logit and probit models; ordered logit and ordered probit models; testing market efficiency)

In many cases, there are faster and slower ways to complete the exercises. Let me know if progress is painfully slow, and I'll be happy to make suggestions to help speed things up. No late work accepted.

### 3. Research Paper and Presentation (20% of total grade):

The research paper (and its presentation) is an empirical project, which will kick off with team assignments after Spring Break. I will assign teams, which will likely have three team members. Students' grades will reflect both their individual performance as well as the quality of the final team product.

Topics should showcase interesting sports econometric analysis. In this paper you will first review and replicate an existing **published** piece of sports econometric analysis (of your

## EC 370: Sports Econometrics

choosing),<sup>1</sup> and then improve that analysis in some way (by adding more data, changing the specification of the model, changing the estimation technique, and so forth).

- **Phase I : *Review & Replication***

Replicate both the summary statistics presented in the paper (to show that you have indeed replicated the construction of the dataset) as well as at least one set of regression results of interest. **Leave plenty of time for this phase. You'll find this far more challenging and time consuming than you could ever imagine.**

- **Phase II: *Improvement***

**Your turn!** Your improvement to the published analysis. This should be a lot of fun... but again, it will not go quickly or smoothly, so budget your time accordingly.

Your paper should discuss your data sources and how you built your dataset. In some cases you may be able to obtain data from the original authors, which obviously greatly simplifies the replication phase. You can do that if you want, but since building datasets is hard work, you won't get as much credit for your Phase I efforts as you would had you built the dataset yourself. In other words: Phase I credit will reflect in part the level of difficulty.<sup>2</sup> If you do construct your own dataset, please attach your .do (or Excel) file to your submission.

Papers should be concise and to the point; shorter is always better - please do not make them longer than necessary. In the past, papers have typically run about 12-18 pages in total. I will say more about the format of the deliverables when teams are assigned.

Alternative formats: Feel free to submit a video or PowerPoint presentation.

There are two milestone dates/deliverables:<sup>3</sup>

- **Wednesday, April 6<sup>th</sup>: Topic selection (in-class presentations)**

Please send me a one paragraph description of your topic, by 6 PM on 4/5 (so I have time to put them into a handout. We'll have in-class presentations (six-eight minutes or so per team) on the 6<sup>th</sup> .

- **Wednesday, May 4<sup>th</sup>: Papers due (in-class presentations?)**

Please send me a one paragraph description of your paper by 6 PM on 5/3 (so I have time to put them into a handout). If time permits, we'll have in-class presentations on the 4<sup>th</sup> (again, six-eight minutes or so per team).

***Empirical work is slow going. Be sure to leave yourself enough time to complete the assignment to your satisfaction.***

---

<sup>1</sup> If you have any questions about relevance, just ask. The published paper that you are replicating must be published in an academic journal such as the *Journal of Sports Economics* or the *Journal of Quantitative Analysis in Sports...* and **published** does actually mean **published** (so no unpublished senior theses, web blogs, or the like). If you have a particular paper in mind and are wondering whether it meets this criterion, just ask.

<sup>2</sup> If you want a sense of *degree of difficulty*, just ask.

<sup>3</sup> Hard copy form, please (except for videos, obviously).

## EC 370: Sports Econometrics

### 4. Tuesday Topics:

These will typically take place at the start of class every Tuesday (if we need more slots, we'll add some Thursday presentations). We'll devote the first 10 minutes or so of class time to a discussion of a current relevant issue. Given the class size, the discussion will be led by a team of three students (team assignments will be distributed once the class list is finalized). The team leading the discussion may want to prepare a brief set of *talking points* to guide and focus the discussion. Presentations should include some of **your own** empirical analysis of the topic. To provide a sense of how this might work, I'll do the first presentation. Presentations will be graded, and along with *participation*, count towards 5% of your course grade.

**Important:** You will be limited to 10 minutes, and at most five slides (not counting the title slide).

**Proposed Schedule of Topics:** The schedule and set of applications will likely evolve as we work through the semester, but here's a sense of the topic schedule:

#### A. *Introduction*

1. Introduction to the course
2. Kicking off the *Sports League Challenge* (SLC)<sup>4</sup>

#### B. *Econometrics and (other) Statistics*

3. Review of Simple Linear Regression (SLR) analysis (Excel & Stata)
  - a. MLB: The Pythagorean Theorem; NBA: shooting success and distance; NFL: field goals success and distance
4. Review of Multiple Linear Regression (MLR) analysis (Excel & Stata)
  - a. MLB: The Pythagorean Theorem; NFL: Ticket Prices; NBA referees and *own-race* bias; NFL: field goal success, altitude and distance
5. Non Linear regression analysis ... and the *residual* (Excel & Stata)
  - a. MLB: The Pythagorean Theorem; NFL: field goal success, altitude and distance
6. Binary Dependent Variables I: The *Linear Probability Model* (LPM) (Excel & Stata)
  - a. NBA: *The Myth of the Hot Hand*; NFL: Icing the Kicker; MLB: *Win Expectancy* and *Leverage*; PGA: *Putting Prowess*; NFL: *Deflategate*
7. Binary Dependent Variables II: Maximum Likelihood Estimation (MLE) (Excel & Stata)... logit, probit and even arc-tangent.
  - a. Repeat the examples in 6a.

---

<sup>4</sup> If we decide to run an abridged version of the SLC, this will likely be moved to later in the semester.

## EC 370: Sports Econometrics

8. Functional forms (percentile dummies, polynomials, splines (linear and cubic) and fixed effects; big v. small datasets) (Stata)
  - a. NBA: shooting success; NFL: field goals tries; PGA: *Putting Prowess*
9. Testing independence in play calling and success: Chi-squared tests, Binomial tests, OLS analysis and the *Wald Wolfowitz* runs test (Excel & Stata)
  - a. NBA: More Hot Hand; MLB: pitch selection; maybe PGA or tennis?

### C. *Selected Topics*<sup>5</sup>

10. More ratings models
  - a. Simple *retrodictive* models (NCAA football); ordered logit and probit models (European football); the *google PageRank* model (NCAA basketball)
11. Wagering market efficiency
  - a. Football (NCAA and NFL (lines and spreads)); Thoroughbred racing (pari-mutuel odds)
12. Competitive balance
  - a. Theory and evidence (MLB, NBA, NFL, NHL and European football/soccer) (at the season and game levels, and across seasons); testing the *Uncertainty of Outcome Hypothesis*
13. Peer effects in performance
  - a. Estimating production complementarities (NBA synergies on the court)

### Additional Resources

- Rodney Fort: <https://sites.google.com/site/rodswebpages/codes>
- John Vrooman: <http://www.vanderbilt.edu/econ/faculty/Vrooman/sports.htm>
- Journal of Sports Economics (JSE): <http://jse.sagepub.com/>
- Journal of Quantitative Analysis in Sports (JQAS): <http://www.degruyter.com/view/j/jqas>
- Multi-author blog: [www.thesportseconomist.com](http://www.thesportseconomist.com)
- Sports Business Daily: <http://www.sportsbusinessdaily.com/Daily.aspx> (expensive but informative; two week trial subscription; student rates (still expensive))
- Sports Business Journal: <http://www.sportsbusinessdaily.com/Journal.aspx> (I believe the library has acquired a subscription to this journal)
- SportsBiz: <http://thesportsbizblog.blogspot.com/>

---

<sup>5</sup> The topic list may very well change depending on time and interest.

## EC 370: Sports Econometrics

- Sports Law: <http://sports-law.blogspot.com/>
- National Sports Law Institute (Marquette): <http://law.marquette.edu/national-sports-law-institute/welcome>
- The “Wages of Wins” Journal: <http://dberri.wordpress.com/>
- and <http://www.cmaxxsports.com/misc/misc.html> (you’ll find useful web pages devoted to MLB, the NBA, the NCAA, the NFL, and European football/soccer... and more)

### and some books:

- **Fair Ball – A Fan’s Case for Baseball**, Bob Costas, Broadway Books, 2001.
- **Sports Economics**, Roger Blair, Cambridge University Press, 2011.
- **The Economics of Sports**, 4<sup>th</sup> ed., Michael Leeds and Peter von Allmen, Prentice Hall, 2010.
- **The Economic Theory of Professional Team Sports: An Analytical Treatment**, Stefan Kesenne, Edward Elgar Publishing, 2007.
- **Playbooks and Checkbooks: An Introduction to the Economics of Modern Sports**, Stefan Szymanski, Princeton U. Press, 2009.
- **The Oxford Handbook of Sports Economics: Volume 1: The Economics of Sports**, Leo H. Kahane and Stephen Shmanske (eds.), Oxford University Press, 2012.
- **The Oxford Handbook of Sports Economics: Volume 2: Economics Through Sports**, Stephen Shmanske and Leo H. Kahane (eds.), Oxford University Press, 2012.
- **Pay Dirt: The Economics of Professional Team Sports**, James Quirk and Rodney Fort, Princeton U. Press, 1997.
- **Sports, Jobs and Taxes: the Economic Impact of Sports Teams and Stadiums**, Roger Noll and Andrew Zimbalist, The Brookings Institution, 1997.
- **International Handbook on the Economics of Mega Sporting Events (International Library of Critical Writings in Economics series)**, Wolfgang Maennig and Andrew Zimbalist (eds.), Edward Elgar Publishing, 2012
- **The Game of Life: College Sports and Educational Values**, William Bowen and James Shulman, Princeton U. Press, 2002.
- **Reclaiming the Game: College Sports and Educational Values**, William Bowen and Sarah Levin, Princeton U. Press, 2005.