

Discrete Choice Model of Recruit's Selection Decision in College Football

Presented by
Mark Witte

Which school will they choose?

- 19,815 High School recruits (2002-2012)
 - Choosing between 5.7 schools (on average)
 - Prospective schools identified by Scouts.com
- What factors are most important?
 - Coach? School? Location?
- Different factors for different recruits?
 - 5 star recruits may have different career expectations than 3 star recruits

Motivation

- Langlett (2003), Caro (2012), Lloyd (2011)
 - Examines recruit classes
- Huffman and Cooper (2012)
 - Survey of 73 current college athletes
 - Informed our choice of RHS variables
- Dumond, Lynch and Platania (2008)
 - 3,395 total prospects and 13,394 potential schools
 - 4-5 star recruits
 - Informed our methodology

The Data

- All publicly available
- Exclusively online
- Data scraping (via SAS, try R in future)
 1. Specific program to sequentially pull from HTML
 - > Ex. player information, coach information
 2. Specific program to search
 - > Ex. distance, same state, family connections

Methodology

- Theory

$$E \sum_{t=1}^Y \beta^{t-1} U \left(J_W WIN_Z^t, J_B COACH_Z, J_P PLAY_{Z,k}^t, J_A AMEN_Z^t, \right. \\ \left. J_D DIST_{Z,j}, J_M MEDIA_Z^t, J_F AFFINITY_Z \right) \\ + E \sum_{t=Y}^{\infty} B^{t-1} U(J_C GRADPr_Z, J_P PRO_{j,k})$$

- Empirical model

- Conditional logit (McFadden, 1974)
 - Individual and choice characteristics
- Odds ratio calculated by exponentiation
- Time-consistent

Data cohorts (part 1/3)

- Win info
 - Recent and historical success at prospective school
- Coach info
 - First-time coach? NFL experience? Alma mater? Historical success and tenure
- Playing time
 - Last year's recruiting class, last year and current year's recruiting at same position

Data cohorts (part 2/3)

- Amenities
 - Football program expenses, total enrollment, annual rainfall, avg. low temp. in January, proximity to beach, restaurants/bar per capita
- Distance
 - Driving distance, same state? same hometown?
- Media
 - TV broadcasts, Age/Size/Location of stadium, BCS school, conference realignment, NCAA infractions
 - Sorry, we did not do “dome” indicator

Data cohorts (part 3/3)

- Affinity
 - Family connection to program? Official visit info, pipeline from H.S. to college?
- Graduation
 - Incoming class SAT, 4-year graduation rate
- Professional career
 - Former players in NFL

Model Fit

	3 star	4 star	5 star	All
Prospects	9,088	7,321	3,406	19,815
Prospect-School Obs.	44,240	46,717	22,427	113,204
Adjusted Estrella	.9182	.8473	.7566	
Percentage correctly predicted by models	72.46	59.54	55.28	64.73
Percentage correctly predicted by random guess	24.82	18.52	16.91	21.14
Percentage correctly predicted by closest school	29.34	27.11	34.91	29.47

Results (part 1/3, all recruits)

- SCHOLARSHIP! (700-1500x more likely)
 - Baseline probability = “walk-on”
 - Less than 1% of 3,4 or 5 star recruits “walk-on”
- Official visits (6-16x)
 - especially during the season
 - especially in large groups
- Attend Football camp (1.6-2.3x)
- Family previous at school (2.0-3.5x)
- Recruits like to stay in-state/close to home
- Former players drafted into NFL (no effect)

Results (part 2/3, recruit specific)

- 4-5 star players dislike cold
 - Every degree increases signing probability by 1%
- 4-5 star players like good facilities and media exposure
 - On-campus stadiums, larger stadiums, institutional spending, more TV broadcasts
- 3-4 star players like good academics
 - Every % of improved graduation rate increases signing probability by 181-261%

Results (part 3/3, new to lit)

- 4 star recruits value H.S. pipeline
 - Each recruit from previous year (4% more likely)
- NCAA sanctions – Scholarship reductions
 - 4 star recruits less likely (-5%)
 - 3 star recruits more likely (5%)

Conclusion

- Work is never done
 - Application to recruiting staffs (no longer Rivals)
 - So many lessons learned
 - Easy to renew data and redo results
- Attendance stipends/medical coverage
 - Richer schools may dominate recruiting