The Dilemma of Choosing Talent:
Michael Jordans are Hard to Find

by

Peter A. Groothuis
Associate Professor
3093 Raley Hall
Appalachian State University
Boone, NC 28608
(828) 262-6077

J. Richard Hill
Professor
310 Sloan Hall
Central Michigan University
Mt. Pleasant, MI 48859
(517) 774-3706

and

Timothy Perri
Professor
3092 Raley Hall
Appalachian State University
Boone, NC 28608
(828) 262-2251

Fall 2005

Economics has a long history of situations where agents have *ex post* regrets from decisions made under uncertainty. In the now classic case of the winner's curse, agents who have differing beliefs about an amenity value will find the winner of an auction will be the bidder who overvalued the amenity. Capen, Clap and Cambell (1971) provide one of the fist references to the winners curse looking at competitive bidding for oil leases, while Cassing and Douglas (1980) provides an example of the winner's curse in baseball free agency. More recently Lazear (2004) identifies the Peter Principle as a situation where individuals who are promoted may have been lucky in a stochastic sense, and may be promoted above their performance level.

Nowhere is this problem of *ex post* regrets more pronounced then in the pursuit of talent. In sports, teams are in pursuit of the next Michael Jordan; in movie production studios in pursuit of the next Titanic; and in the music industry the next Beatles. Yet player after player, movie after movie, and singer after singer fall short and fail to meet expectations. In the pursuit of superstars there are many false positives. We identify this problem as the dilemma of choosing talent. In section one; we develop the model of the dilemma of choosing talent when the distribution of talent is known to be from the upper portion of an ability distribution. In section two we apply this theory to an upper tail of a normal distribution. In section three, we test the theory using a panel study of player's in the NBA during the 1990s. We conclude with a discussion the dilemma of choosing talent and how it relates to the economics of superstars.

Section 1: The Model

To formally model the problem of choosing talent, consider what happens to the probability of finding high quality talent when the lower bound for high quality increases. Assume:

- x = talent, $x_L \le x \le x_H$
- x ~ continuously with a p.d.f of $f(x)$ & a c.d.f of $F(x)$
- x^* is the minimum level for high quality talent
- A potential employer observes a binary signal which is either *favorable* or *unfavorable*
- $P = \text{prob}(x > x^*|\text{favorable})$

Thus, from Bayes theorem we have:

$$P = \frac{\text{prob}(\text{favorable} \mid x > x^*)\text{prob}(x > x^*)}{\text{prob}(\text{favorable} \mid x > x^*)\text{prob}(x > x^*) + \text{prob}(\text{favorable} \mid x < x^*)\text{prob}(x < x^*)} \quad . \qquad (1)$$

Note $\text{prob}(x > x^*) = 1\text{-}F(x^*)$ and $\text{prob}(x<x^*) = F(x^*)$.

Now suppose the probability of a favorable signals increases linearly in x: $\text{prob}(\text{favorable}|x) = x/x_H$. This means those with $x = x_H$ have a probability of one of receiving a favorable signal; others have a smaller probability of a favorable signal.

Now $\text{prob}(\text{favorable}|x > x^*) = \int_{x^*}^{x_H} \frac{x}{x_H} f(x)dx \,/[1 - F(x^*)]$, and $\text{prob}(\text{favorable}|x < x^*) =$ $\int_{x_L}^{x^*} \frac{x}{x_H} f(x)dx \,/\, F(x^*)$. We then can simplify eq.(1):

$$P = \int_{x^*}^{x_H} xf(x)dx \Big/ \int_{x_L}^{x_H} xf(x)dx \,. \qquad (1')$$

The denominator of (1') is the population mean of x, $\overline{X}$. Clearly $\partial P/\partial x^*$ is negative: the higher the level of talent desired ($dx^* > 0$), the smaller the probability one with a favorable signal exceeds the cut off for high talent (x^*). Also $\partial P/\partial \overline{X}$ is negative: the more talented the population, on average, the smaller the probability one with a favorable signal exceeds the cut off for high talent.

For further insight, suppose x ~ uniformly on [$\overline{X}$-Δ, $\overline{X}$+Δ]. We have:

$$P = \frac{(X_H)^2 - (x^*)^2}{4\Delta\overline{X}} \qquad (1'')$$

Now $\partial P/\partial \Delta < 0$, so a larger variance of x (which is positively related to Δ) implies a smaller probability one with a favorable signal exceeds the cut off for high talent.

Suppose $\overline{X} = 6$ & $\Delta = 5$. A firm that desired an above-average worker ($x^* = 6$) would, choosing at random, obtain one with a 50% probability. Using (1"), the signal would correctly identify such an individual 71% of the time. If the firm desired one with $x > 10$, choosing at random, it would obtain such an individual 10% of the time. Using the signal, it would obtain such an individual 17.5% of the time.

Section 2: Choosing Talent and the Normal Distribution

To further explore the dilemma of choosing talent consider the problem that arises when individuals are chosen from the upper tail of the normal distribution. For instance, in basketball there are about 5000 athletes who play division 1 college ball. Of these players, about 30 are drafted into the NBA each year. In any given year out of the total of 300 players in the NBA players only 24 make the all-star game. Of all these players who play in the All Star Game, only a few rise to the caliber of Michael Jordan or a Magic Johnson.

To illustrate the dilemma of choosing talent, suppose talent is normally distributed and both NBA players and superstars come from the upper tail of the normal distribution. Suppose there are four groups of players: Bench Warmers whose talent is 1 to 2 standard deviations above the mean; Role Players whose talent is 2-3 standard deviations above the mean; Stars whose talent is 3-4 standard deviations above the mean; the rare superstars whose talent is 4-5 standard deviations above the mean; and the very rare phenomenon players whose talent is greater than 5 standard deviations from the mean. Now suppose agents can evaluate performance with 90% accuracy with only 10% being judged better than their actual performance. In Table 1, we report the probability of a correct decision.

Standard Deviation	Marginal Density	Population 4 Million	Correct Decision (.9)	False Positives (.1)	Percent Correct
Cut 0-1	.3413	1,365,379	1,228,841		
Bench 1-2	.1359	543,620	489,259	136,538	.782
Role Player2-3	.0214	85,600	77,040	54,362	.586
Star 3-4	.0013	5,273	4,746	8560	.357
Superstar 4-5	.00003	126	113	527	.177
Phenom >5	.0000003	1.1	1.0	13	.076

In column one, we report the probability of any particular player's ability. In this column note that half the distribution of talent is above the mean, with 34% being of the quality that should be cut, while about 14% are of Bench warmer quality, 2% of role player quality, and less than 1 percent are of either star, superstar, or phenomenon quality. To illustrate what these probabilities mean, consider of population of 4 million basketball players. This number was chosen because the probability of a player being of phenomenon quality is one in four million. Column three illustrates how many of each quality is expected to occur in a population of 4 million. Note in a population of four million we expect about 5000 of star quality, about 125 of superstar quality, and only 1 player of phenomenon quality.

Column four provides the number of correct decisions made while column five reports the number of false positives. For instance ninety percent of the players of superstar quality will be predicted of superstar quality which is 113 players, while ten percent of the players of star quality will be predicted to be of super star quality which is 527 players. The last column reports the probability of a correct decision. This column show the difficulty of choosing talent; when picking bench warmers the odds are in your favor-- 73% of the time you make a correct decision. When picking a star, however, you are only about 35% correct; a superstar you are only 18% correct; and with a phenomenon player you are only correct about 8% of the time.

Section 3: Empirical Results

To empirically test the model of the dilemma of choosing talent we focus on NBA data of performance during the 1990s[1]. We use a measure of player performance called the efficiency formula to develop a distribution of talent. As reported by NBA.com, this index is calculated per game as: (points + rebounds + assists + steals + blocks) – ((field goals attempted – field goals made) + (free throws attempted - free throws made) + turnovers). This measure provides a measure of quality based upon performance in all aspects of the games. In table 1, we report the mean, median, standard deviation and highest level of the efficiency rating. We find in all cases the mean is higher than the median suggesting a skewed right distribution of talent. We also find the highest value is always over three standard deviations from the mean. In figure one, we plot a

[1] See Groothuis, Hill and Perri (forthcoming) for a more detailed description of the data set.

distribution of efficiency ratios for one year. The distribution is skewed right with only a few players in the top tail of the distribution.

In table 2, we focus on the players whose efficiency rating is two standard deviations from the mean. We find from 12 to 22 players a year have efficiency ratings over two standard deviations from the mean in any given year. During the 90's, we find that only two players who were in this elite category were undrafted, Ben Wallace in 2001-2002 season and Brad Miller in the 2003 and 2004 season. Many were on the list a multiple of times; some as many as 9 years. During this time, we find many of the number one picks and lottery picks are in the elite category. Some number one picks, however, never show up on the list. Still others only make the list one time in their career.

In table 3, we look at only the top 5 players in efficiency ratings. We find in our 17 year panel only 19 players fill the 85 spots in this time period. Most were on the list a multiple of times. The lowest rank in the draft on this list was the 13th pick—two players, Karl Malone in 1985 and Kobe Bryant in 1996. Many of the top players were number one draft picks. Many number one picks, however, did not make the top 5 players in the NBA. In fact, many of the top picks did not make it to two deviations above the mean. There are many false positives.

In Table 4 the mean, standard deviation, minimum value, maximum value, and number of observations for efficiency are reported by draft number. The figures in this table reveal some interesting results. First the drop off in efficiency between the first pick in the draft and the second pick is statistically significant.[2] The decrease in mean efficiency is also statistically significant between the fifth and sixth picks. There is a general negative relationship between mean efficiency and draft number; exceptions to

[2] The value of the test statistic is 6.5239. This is greater than the critical value at the .005 level of significance given the degrees of freedom.

this trend occur when lower picked players overachieve (e.g. Both Karl Malone and Kobe Bryant were thirteenth picks in the draft). Overall the draft appears to represent either an efficient judge of talent or a self-fulfilling prophesy (teams may give number one picks more minutes and more opportunities to be a superstar).

To further test the dilemma of choosing talent, we use a random effects panel model to estimate player's efficiency ratings. A simple equation to represent the model is:

$$Eff_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it-1} + \varepsilon_{it} \qquad (2)$$

where i refers to the individual player, Eff_{ii} represents the efficiency of the player in year t, X_1 is a vector of time-invariant player characteristics, $X_{2(t-1)}$ is a vector of experience measures, and ε_t is vector of disturbances. The only time-variant player characteristics included in the model are experience and experience squared; no performance statistics are used since efficiency is computed from these stats. Time invariant personal characteristics used to explain efficiency are player height, weight, years of college and a dummy variable equal to one for white players.

Two options for estimating this model are the fixed effects approach and the random effects approach. In the fixed effects formulation of the model, differences across individuals are captured in differences in the constant term; thus any time-invariant personal characteristics are dropped from the regression. In this formulation of the model it is impossible to determine if differences exist between players in terms of efficiency due to draft number or other time-invariant variables. Therefore the fixed effects model will not be used.

In the random effects formulation, the differences between individuals are modeled as parametric shifts of the regression function. This technique of estimating panel data allows for estimates of all of the time-invariant personal characteristics as well as the experience statistics. Breusch and Pagan (1980) developed a Lagrange multiplier test (LM Test) for the appropriateness of the random effects model compared to the OLS format.[3] The Lagrange Multiplier test statistic is 9481.09, which greatly exceeds the 95 percent chi-squared with one degree of freedom, 3.84. Thus the simply OLS regression model with a single constant term is inappropriate.

In table 5 we report these results. In regression I, draft number, experience, experience squared, years of college and race are all statistically significant determinants of efficiency; height and weight are not. As expected, efficiency declines as draft number rises. Efficiency initially rises with experience then declines. Efficiency declines as years of college rises; this reflects the early entry of outstanding college or high school players. The negative coefficient for white players is interesting. *A priori* we would expect this coefficient to equal zero. The result suggests white players may be drafted higher than the future performance would indicate. Regression II is run minus the white variable. There is no change in sign or significance of the remaining variables.

The R-square of the models is around 16%-17% overall. It is somewhat higher in explaining variation in efficiency between players, approximately 22 %, and between years for the same players, 23%. In general the results suggest a great deal of unexplained variation in player efficiency from season to season.

Conclusions

The dilemma of choosing talent suggests that when talent is relatively rare more false positive signals exist then correct decisions. Using NBA data, we find there is much uncertainty in selecting talent. Our results show, if superstars are found they are usually identified early; however, more false positive exist than correct decisions with high draft picks. Our results suggest the dilemma of choosing talent is not so much a

[3] See Stata Release 6 , Reference SU-Z pp. 438-439 for details or Greene (2000) , pp. 572-573.

winners curse but more like a purchase of a lottery ticket. Most of the time you expect to lose, but if you are going to win, you know you must buy the ticket.

Table 1: NBA Efficiency: Means, Medians, and Standard Deviations: 1988-2003

Season	Mean	Median	Standard Deviation	Highest
1987-1988	11.79	10.45	6.82	35.04
1988-1989	10.05	8.68	7.19	36.9
1989-1990	9.96	8.02	7.23	34.6
1990-1991	10.32	9.06	6.89	33.5
1991-1992	9.91	8.38	6.98	32.6
1992-1993	9.94	8.49	6.66	34.4
1993-1994	9.43	8.35	6.49	34.0
1994-1995	9.50	8.14	6.41	32.4
1995-1996	9.33	8.00	6.43	32.0
1996-1997	8.93	7.21	6.42	30.2
1997-1998	8.81	7.59	6.14	29.2
1998-1999	8.05	7.12	5.94	28.8
1999-2000	8.94	7.93	6.03	33.8
2000-2001	8.88	7.29	6.20	31.0
2001-2002	8.98	7.88	6.09	31.2
2002-2003	8.78	7.46	6.19	32.1
2003-2004	8.60	7.22	5.97	33.1

Table 2: Superstar Seasons Based on Efficiency Ratings

Season	Draft Year and Draft Number of Players whose performance was two Standard deviations above the mean based on efficiency measure
1987-1988	84-3, 78-6, 84-5, 84-1, 83-14, 79-1, 85-13, 84-16, 80-3, 82-11
1988-1989	84-3, 79-1, 84-5, 84-1, 85-13, 83-14, 82-11, 85-1, 84-16, 87-7, 78-6, 85-7
1989-1990	84-3, 84-1, 85-1, 85-13, 84-5, 87-1, 79-1, 78-6, 84-16, 87-7, 85-7, 83-14, 82-11, 81-8
1990-1991	87-1, 84-3, 85-13, 84-5, 84-1, 85-1, 79-1, 86-7, 84-16, 87-7, 82-3, 85-7, 85-66, 86-1, 89-14, 78-6, 81-20, 83-14
1991-1992	87-1, 84-3, 85-13, 84-1, 85-1, 84-5, 86-1, 86-27, 89-1, 84-11, 83-14, 78-6, 87-5, 84-16, 85-7, 91-1, 82-3, 81-20, 89-14
1992-1993	84-1, 84-5, 84-3, 85-13, 87-1, 92-1, 86-1, 85-1, 91-1, 82-3, 90-1, 85-8, 92-2, 81-20, 89-14, 91-4
1993-1994	87-1, 92-1, 84-1, 85-13, 85-1, 84-5, 87-5, 89-17, 90-1, 84-16, 92-2, 84-11, 87-10, 93-1
1994-1995	87-1, 84-1, 92-1, 85-13, 84-5, 85-1, 87-5, 89-26, 84-16, 92-2, 84-3, 89-17, 93-1, 83-14, 90-1, 91-4, 89-16, 93-3
1995-1996	87-1, 84-1, 84-3, 85-13, 84-5, 92-1, 92-2, 94-3, 93-1, 89-17, 93-3, 85-1, 87-1, 91-1, 83-14, 84-16
1996-1997	85-13, 92-1, 84-5, 84-3, 94-3, 93-1, 88-53, 84-1, 85-1, 87-7, 93-8, 90-2, 87-5, 92-2, 92-24, 84-16, 92-6, 91-4
1997-1998	85-13, 92-1, 97-1, 87-1, 95-5, 93-1, 84-3, 92-6, 94-3, 84-1, 84-5, 85-1, 90-2, 88-19, 91-4, 86-24, 92-2
1998-1999	92-1, 85-13, 93-1, 97-1, 94-2, 92-2, 95-5, 84-5, 95-2, 84-1, 90-2, 94-3, 89-17, 87-1, 96-3, 92-6, 89-26, 91-4
1999-2000	92-1, 95-5, 93-1, 97-1, 85-13, 90-2, 92-2, 94-3, 91-4, 98-5, 96-3, 96-13, 87-1, 94-2, 95-21, 99-1
2000-2001	92-1, 93-1, 95-5, 97-1, 96-13, 97-9, 85-13, 95-2, 98-9, 98-5, 99-9, 90-2, 96-6, 99-1, 96-3, 94-2, 96-1, 96-5, 95-4, 98-10, 82-18, 99-2
2001-2002	97-1, 92-1, 95-5, 93-1, 98-9, 97-9, 99-1, 90-2, 98-10, 96-13, 85-13, 99-8, 99-9, 96-3, 96-Undrafted, 96-17, 94-2, 96-1, 96-6
2002-2003	95-5, 97-1, 92-1, 97-9, 96-13, 98-9, 93-1, 99-1, 99-9, 96-17, 98-10, 96-9, 94-2, 85-13, 01-3, 96-3, 99-2, 90-2
2003-2004	95-5, 97-1, 99-1, 92-1, 98-9, 97-9, 96-14, 99-undrafted, 96-13, 99-9, 02-35, 99-24, 96-17, 01-19, 96-5, 93-24, 02-1

Table 3: Top Five Players Based on Efficiency Ratings: 1988-2003 Seasons

Season	Player Name, Draft Year, and Draft Number
1987-1988	Michael Jordan: 84-3, Larry Bird: 78-6, Charles Barkley: 84-5, Hakeem Olajuwon: 84-1, Clyde Drexler: 83-14
1988-1989	Michael Jordan: 84-3, Magic Johnson: 79-1, Charles Barkley: 84-5, Hakeem Olajuwon: 84-1, Karl Malone: 85-13
1989-1990	Michael Jordan: 84-3, Hakeem Olajuwon:84-1, Patrick Ewing: 85-1, Karl Malone: 85-13, Charles Barkley: 84-5
1990-1991	David Robinson: 87-1, Michael Jordan: 84-3, Karl Malone: 85-13, Charles Barkley: 84-5, Hakeem Olajuwon: 84-1
1991-1992	David Robinson:87-1, Michael Jordan: 84-3, Karl Malone: 85-13, Hakeem Olajuwon: 84-1, Patrick Ewing: 85-1
1992-1993	Hakeem Olajuwon 84-1, Charles Barkley: 84-5, Michael Jordan: 84-3, Karl Malone: 85-13, David Robinson: 87-1
1993-1994	David Robinson: 87-1, Shaquille O'Neal: 92-1, Hakeem Olajuwon: 84-1, Karl Malone: 85-13, Patrick Ewing: 85-1
1994-1995	David Robinson:87-1, Hakeem Olajuwon: 84-1, Shaquille O'Neal: 92-1, Karl Malone: 85-13, Charles Barkley: 84-5
1995-1996	David Robinson: 87-1, Hakeem Olajuwon: 84-1, Michael Jordan 84-3, Karl Malone: 85-13, Charles Barkley: 84-5
1996-1997	Karl Malone: 85-13, Shaquille O'Neal: 92-1, Charles Barkley: 84-5, Michael Jordan: 84-3, Grant Hill: 94-3
1997-1998	Karl Malone: 85-13, Shaquille O'Neal: 92-1, Tim Duncan: 97-1, David Robinson: 87-1, Kevin Garnett: 95-5
1998-1999	Shaquille O'Neal: 92-1, Karl Malone: 85-13, Chris Webber: 93-1, Tim Duncan: 97-1, Jason Kidd: 94-2
1999-2000	Shaquille O'Neal: 92-1, Kevin Garnett: 95-5, Chris Webber: 93-1, Tim Duncan: 97-1, Karl Malone: 85-13
2000-2001	Shaquille O'Neal: 92-1, Chris Webber: 93-1, Kevin Garnett: 95-5, Tim Duncan: 97-1, Kobe Bryant: 96-13
2001-2002	Tim Duncan: 97-1, Shaquille O'Neal: 92-1, Kevin Garnett: 95-5, Chris Webber: 93-1, Dirk Nowitzski: 98-9
2002-2003	Kevin Garnett: 95-5, Tim Duncan: 97-1, Shaquille O'Neal: 92-1, Tacy McGrady: 97-9, Kobe Bryant: 96-13
2003-2004	Kevin Garnett: 95-5, Tim Duncan: 97-1, Elton Brand: 99-1, Shaquille O'Neal: 92-1, Dirk Nowitzski: 98-9

Figure 1

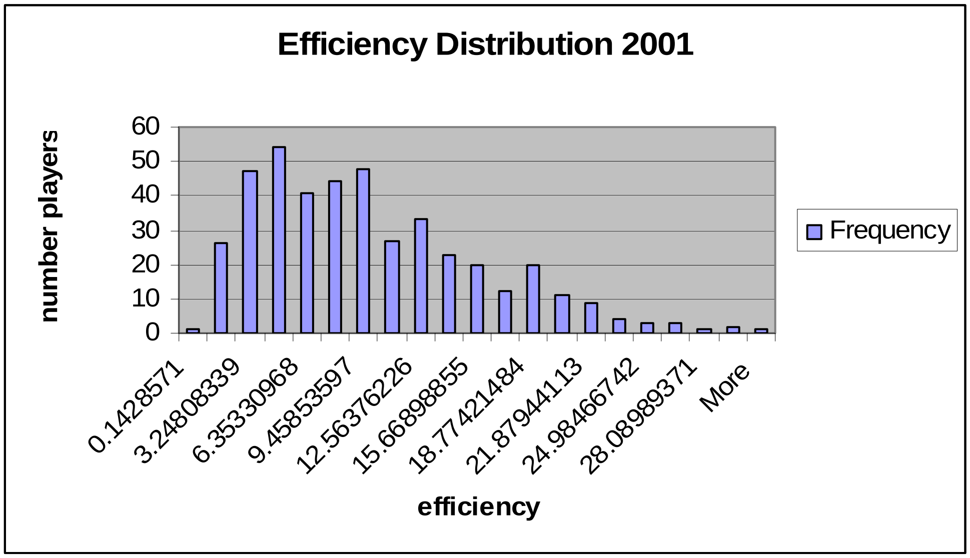
Efficiency Distribution 2001
number players
60
50
40
30
20
10
0
0.1428571
3.24808339
6.35330968
9.45853597
12.56376226
15.66898855
18.77421484
21.87944113
24.98466742
28.08989371
More
efficiency
Frequency

Table 4
Mean and Standard Deviation by Draft Number: (1987-2003)

Draft Number	Mean	Std. Dev.	Min	Max	Obs.(N/n)
1	19.47	7.84	1.67	34.41	213/26
2	15.07	5.55	1.89	26.33	184/22
3	15.75	6.68	.67	36.99	188/24
4	13.79	5.17	1.95	23.80	182/20
5	14.44	7.14	1.08	33.13	199/24
6	10.92	6.26	.75	34.01	157/23
7	12.59	6.00	1.09	29.2	177/25
8	11.83	5.89	-.52	26.1	177/24
9	12.30	6.54	.14	28.8	193/23
10	12.11	5.45	2.36	27	156/21
11	11.47	5.68	1.19	27.48	191/22
12	9.36	5.06	1.33	23.44	148/23
13	12.11	7.59	-.67	31.88	167/21
14	10.58	6.90	-1	28.87	142/22
15	8.86	4.60	-.4	20.06	119/18
16	9.54	6.21	-.25	27.40	146/22
17	9.46	6.05	.67	24.73	112/19
18	10.07	5.22	.43	21.67	139/21
19	8.70	5.83	-.33	22.05	116/20
20	9.17	5.55	0	24.51	117/23
21	8.14	5.19	.33	22.08	127/19
22	7.94	5.38	.33	19.89	99/21
23	8.86	4.80	.2	21.7	118/20
24	10.26	6.02	-2	22.87	128/19
25	7.18	5.53	-1	23.06	79/18
26	7.83	6.27	.2	24.45	76/16

Table 5
Random Effects GLS Efficiency Regression Results: (1987-2003)

Variable	I	II
Constant	13.399 (4.046)	16.290 (4.973)
Draft Number	-.105 (-21.393)	-.108 (-22.078)
Height	-.028 (-0.506)	-.071 (-1.29)
Weight	.007 (0.987)	.007 (1.042)
Experience	.956 (24.655)	.952 (24.533)
Experience Squared	-.102 (-33.605)	-.101 (-33.530)
Years of College	-.539 (-5.066)	-.489 (-4.586)
White	-1.321 (-4.709)	
R-Sq: Within Between Overall	.2327 .2268 .1745	.2331 .2184 .1637

Z-statistics are in parentheses below the coefficients.

Bibliography

Capen, E.C., R.V. Clapp and W.M. Campbell "Competitive Bidding in High-Risk Situations,' Journal of Petroleum Technology, June 1971, pp.641-653.

Cassing, James and Richard Douglas "Implications of the Auction Mechanism in Baseball Free Agent Draft" Southern Economics Journal, vol.47, July 1980, pp. 110-121.

Groothuis, Peter A., James Richard Hill, and Timothy J. Perri "Early Entry in the NBA Draft: The Influence of Unraveling, Human Capital and Option Value" Journal of Sports Economics, forthcoming.

Lazear, Edward, "The Peter Principle: A Theory of Decline" Journal of Political Economy, vol.112, no.1, 2004, pp s141-s163.

Pepall, Lynne M and Daniel J. Richards "Reach for the Stars: A Strategic Bidding Game" Economica, vol.68 2001, pp.489-504.

Rosen, Sherwin, "The Economics of Superstars" The American Economic Review, vol.71, no5, 1981, pp.845-858.