The Sports Labor Market – Part 3

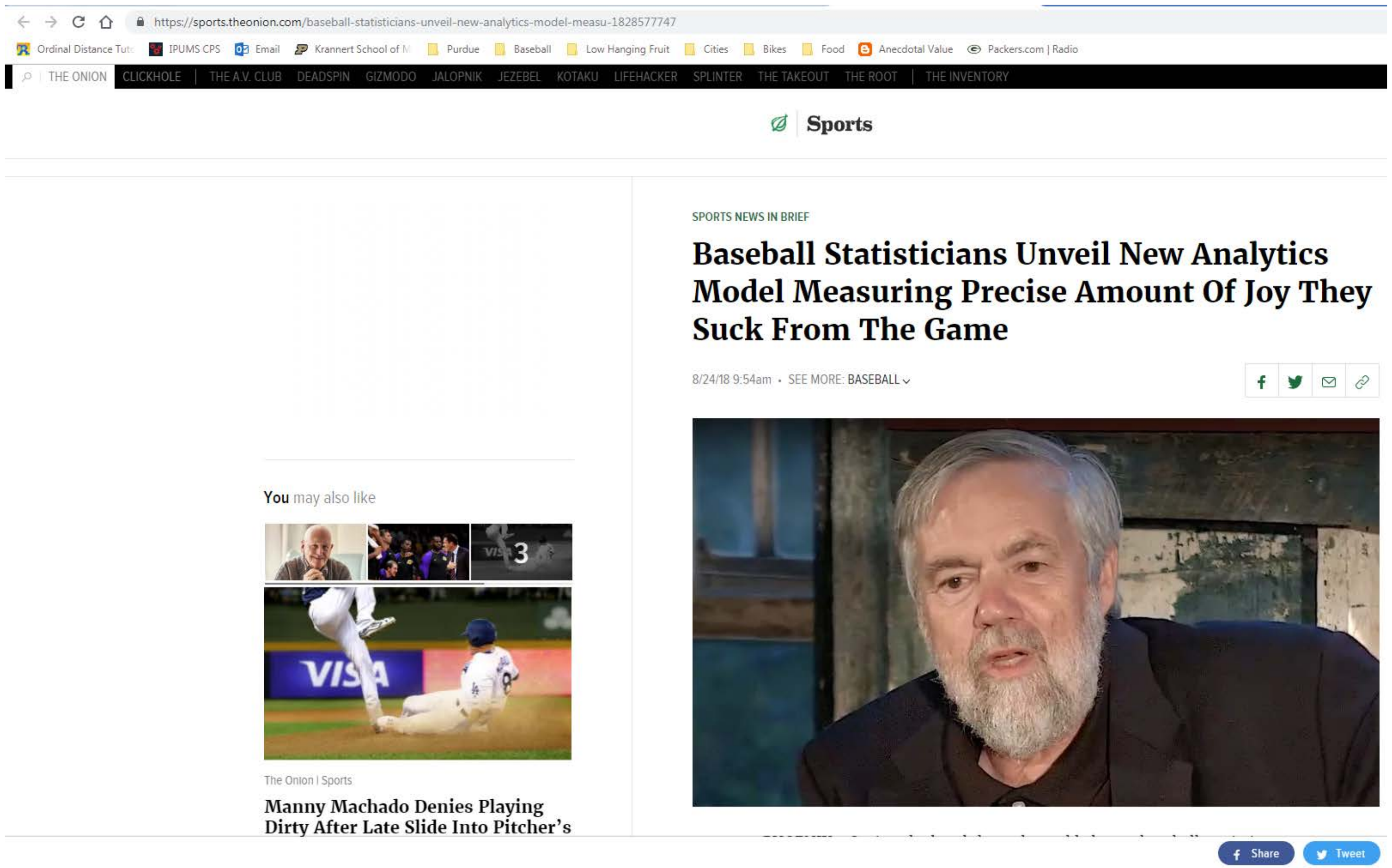
ECONOMICS OF SPORTS (ECON 325)

BEN VAN KAMMEN, PHD

Introduction

In this lecture we replicate, in an introductory way, and evaluate the methods used to measure player and coaching skill in sports.

Source.



Coaches are "technology"; players are "inputs"

One of the great things about studying sports is that in sports, "We count everything."

- The problem is not, as with many other occupations, how to observe productivity.
- With so many performance measures, it's hard to pick the signal out of the noise.

It's specific to the sport in question, but the basic idea is to identify each variable's ("stat column's") effect on the likelihood of winning games.

- David Berri is the symbol of this method.
- How did we arrive at slugging percentage (later OPS) and strikeout-walk ratio as the eminent measures of baseball performance?

A theory of basketball

Even though he's holding a baseball in the picture, basketball is the sport he writes about the most.

In a 1999 paper, Berri proposes structuring empirical analysis around "a theory of basketball."

- He also has a website with analysis and links to data.
- And 2 excellent books: Stumbling on Wins (with Martin Schmidt) and The Wages of Wins (with Schmidt and Stacey Brook).

His theory of basketball revolves around how teams (allow opponents to) acquire and utilize possessions.

Professor David Berri. From wagesofwins.com.

The “wins of possessions?”

Following Berri: there are 3 ways to acquire possession.

- The other team scores,
- You take it away from them,
- You rebound a missed shot by the other team.

Once you have possession, your own scoring (which ultimately is what wins games) depends on ball movement (assists-turnovers ratio), shooting, and as my favorite color man, Jon McGlocklin, likes to say, “reloads” (offensive rebounds).

Here’s how

Using team data by season (from 1994-1998), here are his regression estimates.

PPS=Points Per Shot

DTO=Takeaways

Table 3. Estimated Coefficient for Equation (5) (Dependent Variable is PCT[a])

Independent variables	Coefficient	S.E.	*t*-Statistic
PPS	2.043	0.158	12.896
FT	0.887	0.194	4.573
FTA	0.013	0.002	5.426
RBO	0.029	0.005	6.340
ASTO	0.212	0.035	6.027
DTO	0.020	0.004	4.657
RBD	0.020	0.004	5.106
DPTS	(0.020)	0.001	(17.562)
R^2	0.963	Mean dependent variable	0.491
Adjusted R^2	0.945	S.D. of the dependent variable	0.169
S.E. of regression	0.039	Sum of squared residuals	0.120
Observations: 114			

The “wins of the box score”

Since some of the familiar stats from the box score enter Berri’s model as, say, the denominator in a ratio, he does the calculus to arrive at marginal effects.

I.e., If I get one more offensive rebound, holding other things constant, my team gets about 1/20 of a win.

Table 5. Marginal Values and Average Elasticities

	Marginal value	Elasticity
Player statistics		
Offensive rebound	0.058	1.506
Three point field goal made	0.052	0.574
Turnovers	(0.042)	(1.333)
Opponent’s turnover (steal)	0.037	1.188
Two point field goal made	0.026	1.648
Defensive rebound	0.026	1.483
Missed field goal	(0.025)	(2.222)
Made free throw	0.021	0.847
Assist	0.014	0.616
Missed free throw	(0.012)	(0.173)
Personal foul	(0.007)	(0.333)
Tempo statistics		
Field goal attempt	(0.023)	(3.805)
Free Throw Attempted	(0.009)	(0.488)
Defensive statistics		
Opponent’s three point field goals made	(0.026)	(0.286)
Opponent’s two point field goal made	(0.013)	(0.818)
Opponent’s missed field goal (blocked shot)	0.013	1.104
Opponent’s assist	(0.012)	(0.565)
Opponent’s free throw missed	0.009	0.126
Opponent’s free throw made	(0.003)	(0.127)

Just fit the model

Once you know the effect of each stat column on wins, you just plug in a player's season stats to estimate his production.

Berri performs a couple additional adjustments to this figure to account for:

- Differences in teams' pace of play, i.e., both teams will accumulate more counting stats if they each use 125 possessions per game than if they use 95,
- And position differences; if you're going to compare players of different positions, you want to see how they did relative to a substitute player of the same position rather than comparing a guard to a center.

Examples from 1997-98

Table 9. Top Ten Regular Season Wins Producers

Player	Team	Minutes	Wins Rank	Wins per minute	Wins	IBM rank	IBM
Dennis Rodman	Bulls	2856	1	0.0073	20.79	6	88.31
Karl Malone	Jazz	3030	2	0.0062	18.83	1	99.69
Jayson Williams	Nets	2343	3	0.0080	18.79	20	75.82
David Robinson	Spurs	2457	4	0.0071	17.50	3	96.66
Tim Duncan	Spurs	3204	5	0.0054	17.45	2	98.70
Michael Jordan	Bulls	3181	6	0.0052	16.44	8	85.58
Charles Barkley	Rockets	2243	7	0.0072	16.22	17	77.77
Gary Payton	Super Sonics	3145	8	0.0050	15.75	9	84.50
Charles Outlaw	Magic	2953	9	0.0052	15.37	15	80.29
Jason Kidd	Suns	3118	10	0.0048	14.88	13	81.38

From Berri (1999), p. 421. Dennis Rodman's extreme rebounding prowess gave him the most value of any player in the league that year.

Examples from 2010-11, team estimates, source: wagesofwins.com

Team	*Actual Wins*	*Summation of* *Wins Produced*	*Difference in* *Absolute Terms*		*Team*	*Actual Wins*	*Summation of* *Wins Produced*	*Difference in* *Absolute Terms*
Atlanta	44	38.6	5.4		Milwaukee	35	38.7	3.7
Boston	56	55.1	0.9		Minnesota	17	23.2	6.2
Charlotte	34	30.3	3.7		New Jersey	24	24.7	0.7
Chicago	62	60.4	1.6		New Orleans	46	43.6	2.4
Cleveland	19	16.8	2.2		New York	42	43	1
Dallas	57	52	5		Oklahoma	55	51.5	3.5
Denver	50	53.4	3.4		Orlando	52	55.5	3.5
Detroit	30	31.5	1.5		Philadelphia	41	45.2	4.2
Golden State	36	34.8	1.2		Phoenix	40	39	1
Houston	43	46.9	3.9		Portland	48	45	3
Indiana	37	38	1		Sacramento	24	26.8	2.8
LA Clippers	32	32.7	0.7		San Antonio	61	56.1	4.9
LA Lakers	57	57.3	0.3		Toronto	22	24.2	2.2
Memphis	46	47.3	1.3		Utah	39	36.2	2.8
Miami	58	60.8	2.8		Washington	23	21.4	1.6
						Average Difference	2.61	

A “theory of football”

In a 2007 paper (and Chapter 9 of Wages of Wins), Berri estimates the analogous effects of counting stats on wins in the NFL:

- Getting possessions: opponent kickoffs, opponent punts, takeaways, failed FG attempts and 4th down conversions.
- Use of possessions: return yards, rushing and passing yards per play, net penalty yards, propensity to score TDs rather than FGs.

The team stats used to estimate the “production function” are difficult to attribute to individual players in football, because of the interaction among them in executing a play.

- As opposed to baseball, where you know who threw the pitch, who (didn’t) hit it, and who fielded it well (badly),
- And basketball, where you know who made (missed) the shot and who turned the ball over to the opponent.

QBs and RBs in football may be the exceptions, though.

Berri's results for NFL

Using team data spanning 1998-2005.

These are the marginal effects, ceteris paribus, on points of things that QBs and RBs have some control over in a football game.

Once you know how many of these things each player does, you can plug in his stats and sum up the point values of the good (gain more yards per play) and bad (turn the ball over), to make comparisons.

TABLE 14.4
Marginal value of various quarterback and running back statistics.

Variable	Impact on point differential of a one unit increase
Yards	0.080
Plays	–0.214
Interceptions	–2.745
Fumbles Lost	–2.899

Top performers in the data

The top quarterbacks: 2000-2005, ranked by *QB Score* per play and the NFL's quarterback rating, minimum 224 pass attempts per season.

Quarterback	Rank *QB Score* per play	Rank *QB Rating*	Year	Yards	Plays	Turnovers	*QB Score* per play	*QB Score*	*QB Rating*
Peyton Manning	1	1	2005	3,711	503	12	1,842	3.66	104.1
Ben Roethlisberger	2	3	2005	2,325	322	10	1,059	3.29	98.6
Matt Hasselbeck	3	4	2005	3,429	509	11	1,572	3.09	98.2
Trent Green	4	8	2005	3,892	574	14	1,750	3.05	90.1
Tom Brady	5	6	2005	4,011	583	17	1,752	3.01	92.3
Peyton Manning	1	1	2004	4,494	535	11	2,559	4.78	121.1
Daunte Culpepper	2	2	2004	4,885	682	15	2,389	3.50	110.9
Donovan McNabb	3	4	2004	3,903	542	14	1,857	3.43	104.7
Trent Green	4	7	2004	4,449	613	21	1,980	3.23	95.2
Brett Favre	5	10	2004	4,023	568	18	1,779	3.13	92.4
Trent Green	1	4	2003	3,992	569	13	1,895	3.33	92.6
Peyton Manning	2	2	2003	4,186	612	11	2,020	3.30	99.0
Steve McNair	3	1	2003	3,245	457	13	1,484	3.25	100.4
Jake Plummer	4	5	2003	2,314	353	9	985	2.79	91.2
Daunte Culpepper	5	3	2003	3,705	564	17	1,503	2.66	96.4
Trent Green	1	4	2002	3,774	527	13	1,803	3.42	92.6
Chad Pennington	2	1	2002	3,034	450	7	1,474	3.28	104.2
Rich Gannon	3	2	2002	4,631	704	13	2,129	3.02	97.3
Kerry Collins	4	14	2002	3,918	613	15	1,629	2.66	85.4
Brad Johnson	5	3	2002	2,958	485	8	1,263	2.60	92.9
Kurt Warner	1	1	2001	4,657	612	26	2,041	3.33	101.4
Steve McNair	2	5	2001	3,513	543	15	1,434	2.64	90.2
Brett Favre	3	4	2001	3,826	570	21	1,486	2.61	94.1
Jeff Garcia	4	3	2001	3,678	602	15	1,422	2.36	94.8
Peyton Manning	5	8	2001	4,056	611	26	1,443	2.36	84.1
Kurt Warner	1	3	2000	3,331	385	19	1,606	4.17	98.3
Jeff Garcia	2	5	2000	4,537	657	11	2,236	3.40	97.6
Brian Griese	3	1	2000	2,651	382	7	1,295	3.39	102.9
Trent Green	4	2	2000	1,987	284	7	925	3.26	101.8
Peyton Manning	5	6	2000	4,398	628	17	2,004	3.19	94.7

The top running backs: 2000-2005, ranked by *RB Score* per play and rushing yards.

Quarterback	Rank *RB Score*	Rank Rushing Yards	Year	Rushing Yards	Receiving Yards	Plays	Funbles lost	*RB Score*
Tiki Barber	1	2	2005	1,860	530	411	1	1,127
Larry Johnson	2	3	2005	1,750	343	369	4	866
Shaun Alexander	3	1	2005	1,880	78	385	1	773
Warrick Dunn	4	8	2005	1,416	220	309	1	679
LaDainian Tomlinson	5	6	2005	1,462	370	390	1	632
Tiki Barber	1	5	2004	1,518	578	374	2	914
Edgerrin James	2	4	2004	1,548	483	385	2	816
Brian Westbrook	3	28	2004	812	703	250	1	735
Curtis Martin	4	1	2004	1,697	245	412	0	706
Shaun Alexander	5	2	2004	1,696	170	376	3	648
LaDainian Tomlinson	1	3	2003	1,645	725	413	0	1,131
Priest Holmes	2	9	2003	1,420	690	394	1	898
Clinton Portis	3	5	2003	1,591	314	328	1	891
Ahman Green	4	2	2003	1,883	367	405	5	885
Jamal Lewis	5	1	2003	2,066	205	413	6	852
Priest Holmes	1	3	2002	1,615	672	383	1	1,108
Charlie Garner	2	22	2002	962	941	273	0	1,084
Clinton Portis	3	4	2002	1,508	364	306	3	864
LaDainian Tomlinson	4	2	2002	1,683	489	451	1	789
Ricky Williams	5	1	2002	1,853	363	430	5	776
Marshall Faulk	1	5	2001	1,382	765	343	3	1,028
Priest Holmes	2	1	2001	1,555	614	389	3	912
Ahman Green	3	4	2001	1,387	594	366	4	763
Tiki Barber	4	19	2001	865	577	238	1	698
Garrison Hearst	5	10	2001	1,206	347	293	1	644
Marshall Faulk	1	8	2000	1,359	830 3	34	0	1,187
Robert Smith	2	2	2000	1,521	348	331	1	846
Edgerrin James	3	1	2000	1,709	594	450	5	803
Tiki Barber	4	22	2000	1,006	719	283	3	786
Ricky Watters	5	11	2000	1,242	613	341	2	772

Porter & Scully's paper (1982)

Conceive of the manager as the club's technology in turning inputs (player skills) into output (Wins). In baseball the skills are Hitting and Pitching, and P&S model them as a Cobb-Douglas production function, where Managers can multiply production with their skill.

$$W = MP^{\alpha}H^{1-\alpha}$$

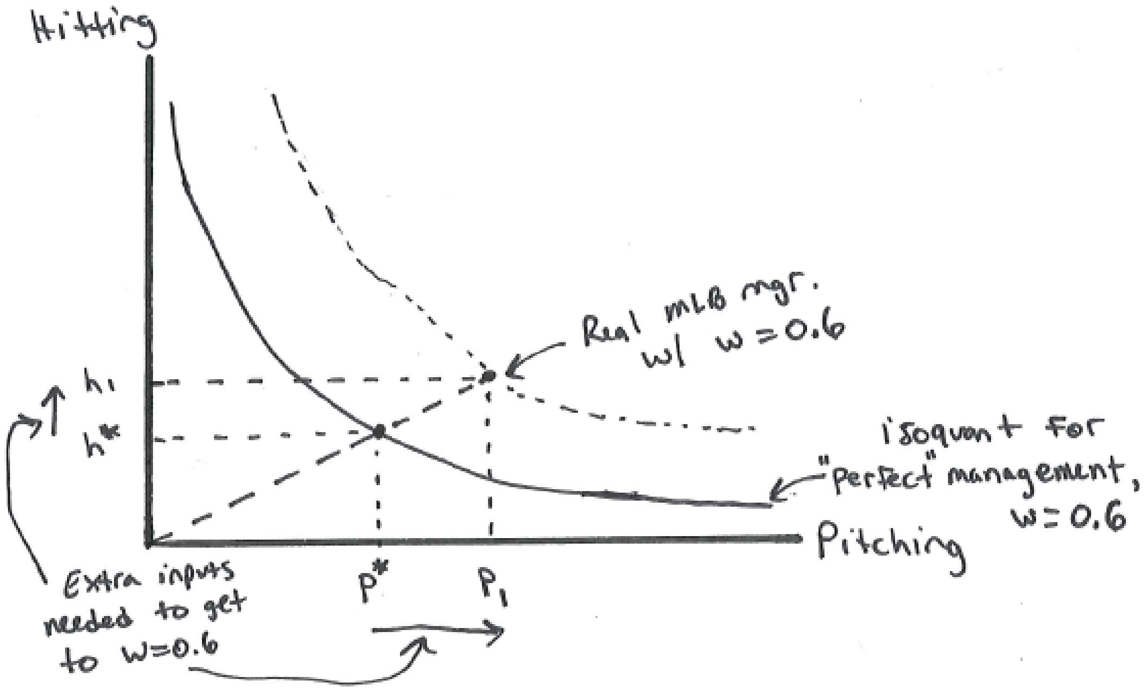
- A more skilled manager can produce the same number of wins with poorer players.
- His isoquants, lines showing combinations of hitting/pitching that yield the same level of wins, are closer to the origin—and closer to a *theoretical* optimum management.

Managerial (in)efficiency

Using Pythagorean Theorem, you can express how many more inputs a real MLB manager would need to achieve the same win % as the "ideal" manager:

$$E \equiv \frac{[(p^*)^2 + (h^*)^2]^{\frac{1}{2}}}{[p_1^2 + h_1^2]^{\frac{1}{2}}}; 0 \leq E \leq 1.$$

- This is the measure of the real manager's efficiency (1 is optimum).



Porter & Scully, results

Take the form of average inefficiency estimates:

- By Manager and
- By Club.

The estimates range from about 0.7 to 0.99. With the mean at about 0.85.

Earl Weaver is the best manager in the sample. Using the valuation methods from Scully's earlier paper on player MRP, he was worth about $675,000 /year (in $1969) to the Orioles.

- This was only a little less than Scully's estimate of Sandy Koufax's MRP.
- Elite managers are quite valuable.

Better teams have better managers.

Expansion teams' managerial efficiency improves over time.

Based on my own replication

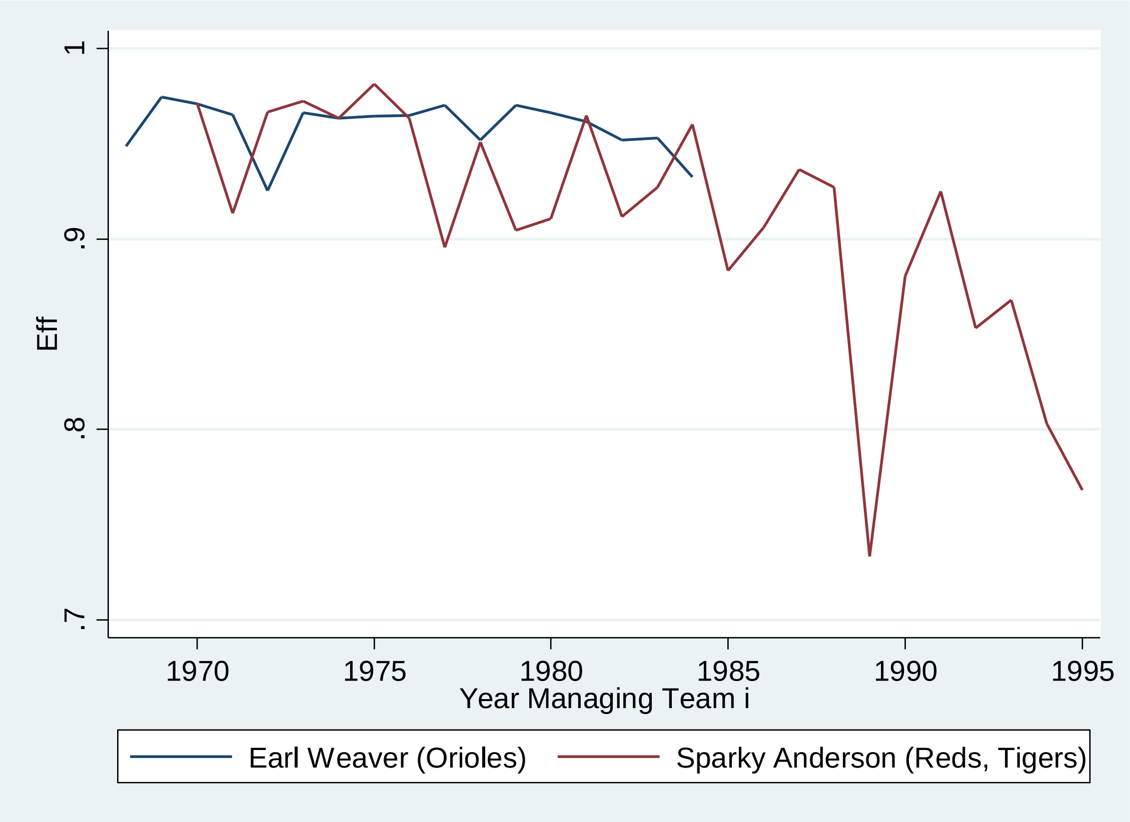
Weights on hitting, pitching are 0.75 and 0.25, respectively.

2 of the best managers of the P&S era.

The best managers are typically consistently good on the efficiency measure.

A diminishing return to experience as managers stay on longer.

- More apparent for Sparky Anderson, whereas Earl Weaver left “on top.”



Porter & Scully, conclusions

This is a milestone paper in Economics of Sports, but it makes several strong assumptions.

Among them is that the player performance (OBP and K:BB ratio) are taken as given and not influenced by managerial quality.

- Kahn (1993) relaxes this assumption and finds that players *do* play better when playing for better managers. Surely this would raise the MRP of an elite manager beyond P&S's estimates.
- One wonders whether managers are paid for this or whether players reap the rewards of playing for a good coach. Frick & Simmons (2008) suggest (in German soccer, at least) managers are paid <MRP.

Similar methods have been used on more recent samples and in different sports.

- As representatives of the work on the NBA, consider the estimates from Hofler & Payne (1997) and Lee & Berri (2008).

Each NBA team ranked by frontier wins 1992–1993 season

Team	Wins		
	Frontier[a]	Actual	Efficiency
Phoenix	78	62	79.3%
NY Knicks	64	60	93.5%
Houston	63	55	87.5%
Cleveland	62	54	86.8%
Chicago	60	57	94.2%
San Antonio	58	49	85.0%
Seattle	57	55	96.8%
Portland	57	51	90.9%
Boston	54	48	89.2%
Atlanta	50	43	84.8%
Utah	49	47	96.9%
Detroit	49	40	81.0%
Charlotte	48	44	91.0%
Orlando	45	41	91.7%
Indiana	45	41	91.7%
LA Clippers	45	41	91.0%
NJ Nets	44	43	97.5%
LA Lakers	43	39	90.4%
Miami	41	36	88.3%
Denver	37	36	98.5%
Golden State	35	34	96.4%
Washington	32	22	68.7%
Milwaukee	30	28	92.7%
Philadelphia	30	26	86.7%
Sacramento	27	25	92.6%
Minnesota	19	19	97.8%
Dallas	18	11	61.7%
Averages	46	41	89.0%

NBA (1992-93 season), from Hofler & Payne, p. 298.

Average efficiency levels time-invariant fixed effects (FE) vs. time invariant GLS

Team	Average wins	Efficiency FE	Efficiency GLS	Potential wins FE[a]	Potential wins GLS
San Antonio	58.667	1.000	0.986	58.667	59.472
Sacramento	58.333	0.826	0.929	70.647	62.806
Dallas	56.667	0.965	1.000	58.740	56.667
LA Lakers	54.667	0.841	0.931	64.971	58.748
Portland	49.667	0.708	0.874	70.161	56.825
Minnesota	49.333	0.842	0.958	58.619	51.492
Philadelphia	49.000	0.813	0.930	60.271	52.686
Utah	48.000	0.714	0.890	67.227	53.909
Charlotte (New Orleans)	45.667	0.815	0.933	56.060	48.922
Milwaukee	45.000	0.761	0.928	59.102	48.479
Detroit	44.000	0.865	0.953	50.861	46.160
Indiana	43.667	0.772	0.904	56.548	48.297
Phoenix	43.667	0.599	0.841	72.863	51.897
Boston	43.000	0.754	0.931	57.029	46.198
Orlando	43.000	0.678	0.881	63.394	48.785
Seattle	43.000	0.616	0.862	69.783	49.901
New Jersey	42.333	0.655	0.862	64.671	49.113
Houston	38.667	0.676	0.886	57.216	43.619
New York	38.333	0.733	0.886	52.268	43.280
Toronto	37.667	0.599	0.827	62.841	45.529
Miami	37.000	0.638	0.839	57.957	44.102
LA Clippers	32.333	0.563	0.826	57.441	39.142
Atlanta	31.000	0.584	0.832	53.046	37.247
Washington	31.000	0.679	0.878	45.655	35.288
Denver	28.000	0.508	0.788	55.107	35.541
Cleveland	25.333	0.452	0.753	56.010	33.642
Golden State	25.333	0.478	0.771	52.988	32.855
Vancouver (Memphis)	24.667	0.462	0.762	53.437	32.390
Chicago	22.000	0.447	0.735	49.217	29.929

Notes:
[a]Potential wins are calculated under the assumption that each team is 100% efficient and the remaining teams remain at the same efficiency level estimated.

NBA (1993-2003), from Lee & Berri, p. 64.

Okay here's one for the NFL, too

Table 1. *Frontier parameter estimates: offence dependent variable is natural log of total points scored on offence (all regressors are in natural logs)*

Variable	Estimate	Standard error	*t*-ratio
Intercept	−6.55	0.76	−8.59
Net yards gained, rushing	0.58	0.06	10.24
Net yards gained, passing	0.71	0.08	9.27
Third down efficiency (%)	0.22	0.10	2.18
Punt return yards	0.12	0.03	4.23
Successful field goals (%)	0.19	0.07	2.55

(left) Estimates of marginal products in NFL "production function" during 1989-1993, Hofler & Payne (1996).
(right) Efficiency measures compared to overall record in 1992. The best teams aren't necessarily the most efficient. From Hofler & Payne, p. 746.

Team	Deviation from best possible performance offence (%)[a]	(Rank)	Games won (Rank)	(Wins)
Atlanta	1.78	1	18	6
Buffalo	3.12	9	4	11
Chicago	3.02	6	21	5
Cincinnati	5.26	20	21	5
Cleveland	3.35	11	16	7
Dallas	3.41	12	2	13
Denver	4.00	16	15	8
Detroit	2.35	2	21	5
Green Bay	3.54	13	12	9
Houston	2.43	3	10	10
Indianapolis	9.21	26	12	9
Kansas City	2.54	5	10	10
L.A. Raiders	3.87	15	16	7
L.A. Rams	7.49	24	18	6
Miami	3.02	7	4	11
Minnesota	3.02	8	4	11
N.Y Giants	5.59	21	18	6
N.Y. Jets	10.52	27	25	4
New England	15.71	28	27	2
New Orleans	2.54	4	3	12
Philadelphia	3.14	10	4	11
Phoenix	8.93	25	25	4
Pittsburgh	4.95	19	4	11
San Diego	6.90	23	4	11
San Francisco	3.59	14	1	14
Seattle	6.68	22	27	2
Tampa Bay	4.53	18	21	5
Washington	4.20	17	12	9
Average	4.95			8

The difference between performance and talent

It can be tempting to look at these lists and conclude who is the most skilled player at each position.

If the measure was capturing something enduring about the player, though, you'd expect performance to be predictable based on the past.

- Or unpredictable if it is noisy, i.e., the effect of teammates, luck, or other things outside his control.

The difference between performance and talent

It turns out that all of our favorite NFL stats are (pretty much) the latter type. The ones with the strongest autocorrelation (last year's stat correlation with this year's same stat) are:

- QB completion % ($r = 0.54$) and
- QB rushing yards per attempt ($r = 0.6$).

Even though it seems to incorporate so much about QB efficiency, Berri's measure has only a ($r = 0.4$) modest autocorrelation coefficient.

- Except in very obvious cases, e.g., Peyton Manning, a lot of QBs' success seems to derive from combining them with a complementary supporting cast.

Running backs' stats have similarly low autocorrelation coefficients

Summary, caveats, extensions

Basketball and baseball player stats tend to be more persistent (predictive of future stats) than football player stats. Even harder for soccer.

- MLB pitchers' ERAs can be predicted (correlation $r = 0.877$, source: Bradbury, The Baseball Economist, p. 171) from their previous seasons' strikeout, walk, and home runs allowed rates.
- Hitters' SLG% (to justify their use in Scully's early paper and elsewhere) have an autocorrelation of 0.82.

This may reveal something about the sports' production functions. Being talented at baseball or basketball will result in a good performance on a more reliable basis than in football.

I wonder whether the propensity of these statistics to regress to the mean (an above average year makes a decrease the following year more likely) can be explained by opponents.

- This would be especially true in the NFL because there are so many plays that can be run. Opponents see you succeeding with one style and learn to stop that style. Then, if you're smart, you find a new offense to beat them with.
- This back and forth creates the ups and downs from year to year, even if my team's talent is constant.

I still think in these terms

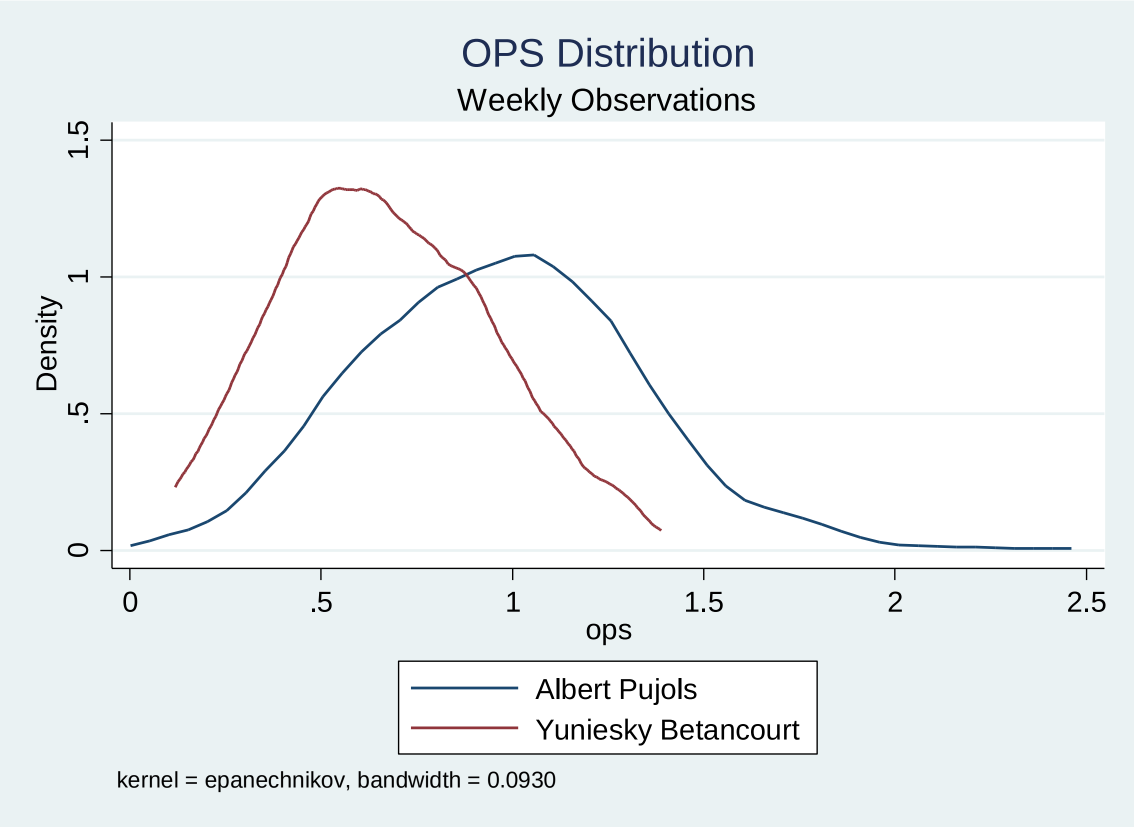
When a club signs a player, it's paying for draws from a distribution like these (right).

Future Hall-of-Famer, Albert Pujols has had plenty of weeks where he was worse than Yuniesky Betancourt offensively.

- Would be rare to see the average over 24 draws from the blue be worse than 24 from the red, though.

The mean is primarily what the club pays for.

- Maybe the variance, too.
- Not clear if more is better ("upside") or worse ("inconsistent").



Conclusion

The histograms on the previous slide are generatedex post. Clubs have to staff their rosters ex ante.

- Their task is to form accurate expectations of each player's *future* performance distribution, based on present observations.
- An arms race in terms of information.

This suggests 2 distinct paths to success in sports:

- Try to (legally) win the informational arms race to acquire more talent, and
- Coach players more efficiently, thereby *making* them more talented, shifting their performance distributions rightward. The ol' fashioned way!