By the Numbers

Volume 33, Number 1 The Newsletter of the SABR Statistical Analysis Committee March, 2024

Review

Academic Research, 2022-2023

Charlie Pavitt

Charlie reviews several recent studies from the academic literature.

In this issue

The previous issue of this publication was August, 2022 (Volume 31, Number 1).

Richard J. Paulsen (2022), It's a party in the MLB: An analysis of shirking between games in major league baseball, Journal of Sport Management, Vol. 36 No. 4, pp. 319-329.

This is Richard Paulsen's third attempt at demonstrating that players with long-term contracts shirk. This time, he used 2015-2017 game level data from DailyBaseballData.com, with a sample of 1009 non-pitchers and 117,427 games.

Consistently with other shirking studies, batter wOBA dropped with more years left on contracts, which is difficult to explain away by aging effects. But, in this case, the performance decrease was limited to the sample of day games immediately following night games.

Unlike the overall shirking effect, this specific finding can easily be accounted for by player aging rather than shirking as such; perhaps older players need more time to bound back after a previous game. However, the author found the effect extenuated for games played in "party cities",[1] particularly for players who were under the age of 30, childless, and unmarried. Some of this is not easily explained by player aging.

An interesting sidebar was that long rest (day/day, night/night, or longer) resulted in worse performance, which Paulsen speculated as possibly due to opposing pitchers being advantaged by the extra hours more than the batters.

[1] defined as Miami, Houston, New York, San Diego, Los Angeles, Chicago, and Philadelphia, all listed in the top ten nightlife cities according to https://www.skyscanner.com/tips-and-inspiration/10-best-nightlife-cities-united-states; and replicated with a list from https://travel.usnews.com/rankings/best-bar-club-scenes-in-the-usa/ which replaced Philadelphia and Houston with San Francisco and Boston.

Meghan A. Thornton-Lugo, Matthew W. McCarter, Jonathan R. Clark, William Luse, Steven J. Hyde, Zahra Heydarifard, and Lulu S. R. Huang (2023), Makeup calls in organizations: An application of justice to the study of bad calls, Journal of Applied Psychology, Vol. 108 No. 3, pp. 374-402.

These authors claimed to have evidence that home plate umpires have been making makeup calls after mistakes, based on 2008-2014 data from BaseballSavant plus contextual info from Retrosheet. Their results suggest that, after missing a ball, umps were 23 percent more likely to call a ball in the next five pitches than otherwise to the same batter, and 10 percent more likely for other batters from the same team. (The authors didn't consider that an extra strike could lead to more balls for strategic reasons – like tempting the batter to chase a bad pitch on an 0-2 count.)

After missing a strike, umps were 15 percent more likely, across the board, to call a strike in the next five pitches to the same batter. However, this was conditional on leverage; the higher the leverage, the less likely the extra strike. There was no analogous leverage impact on missed balls.

The authors' major findings are in line with those reported by Tobias Moskowitz and Jon Wertheim in their book *Scorecasting*.

Vivek Bhattacharya and Greg Howard (2022), Rational inattention in the infield, American Economic Journal: Microeconomics, Vol. 14 No. 4, pp. 348-393.

These authors were interested in determining the best pitch type mix (fastballs versus non-fastballs) for different game situations. They used a model that, if I understand it correctly, implies that the fastball is the best pitch, and so pitching anything other than a fastball is a losing strategy. The authors knew better than this, and the point of the paper was really to uncover the situations in which it is rational for pitchers to use pitches other than fastballs.

The study used pitch-level data from 2008 to 2017, from various sources including MLB, Retrosheet, FanGraphs, Baseball Reference, and Bill Petti. Outcomes were based on team wins, pitcher FIP, and opposition wOBA.

The most important finding is not exactly news: as the count increased in strikes, fastballs led to worse outcomes for the pitcher's team. Pitchers did recognize this, and so threw fewer fastballs given more strikes.

Extrapolating, the authors found that -- generalizing across all game situations and assuming no strategic changes from batters -- teams would have won one additional game if they had used 20 percent more fastballs in no-strike counts and 10 percent fewer fastballs in two-strike counts.

Among other interesting findings:

- Higher leverage situations appear to accentuate the relationship between strikes and ideal fastball proportions;
- Closer-to-optimum fastball proportions were used by teams that shifted more, a result the authors attribute to those teams being more analytically savvy;
- Closer-to-optimum decisions were also more frequent from pitchers younger than 30, for which the authors blame cognitive decline by the over-30s;
- Temperatures below 90 degrees Fahrenheit lead to closer-to-optimum fastball usage. The authors posit that pitchers put less effort into thinking when the weather is hot.

Christopher W. Callahan, Nathaniel J. Dominy, Jeremy M. DeSilva, and Justin S. Mankin (2023), Global warming, home runs, and the future of America's pastime, Bulletin of the American Meteorological Association, Vol 104 No. 5, pp. e1006-e1016.

This paper got a lot of press; in fact, I learned about it from a mention in, in all things, The Skimm, which is an online news compiler directed at young women.

The authors studied the relationship of game-day temperature to home runs from 1962-2019 (100,000+ games), and batted balls from 2015-2019 (220,000+ games). They found an increase of 1 degree Celsius (1.8 degrees Fahrenheit) led to:

- A 1.96 percent increase in HR in open-air parks (2.4 percent day, 1.7 percent night);
- An smaller increase in HR in domed stadiums (0.3 percent day, 0.2 percent night);
- An increase in HR/FB of 0.16 percent.

The percentage increase in HR is equivalent to an increase of about 0.041 to 0.044 actual HR/G.

Also, the higher observed temperatures of 2010-2019 caused an increase of 58 extra HR per season, the authors estimate. That's the total for all of MLB, and is the equivalent of 95 HR per degree Celsius.

Finally, extrapolating to future temperatures predicted by climate studies: models postulating effective mitigation of climate change lead to about 130 extra HR per year, but models that postulate high CO_2 emissions suggest an increase of 192 HR by 2050, rising to an increase of 467 HR by 2100.

Muyuan Li, Greg Plithides, and Max Plithides (2023), Quantifying the effect of offseason contract extensions on short-term performance, Baseball Research Journal, Vol. 52 No. 1, pp. 72-78.

I'll keep this one short, as I assume everyone that reads this has access to it. The authors examined the impact of pre-free agency contract extensions by comparing players' performance after the extension to their performance the year before the extension. Their sample included all 182 pre-free agency off-season contract extensions from January 2000 through June 2020.

They discovered that performance was about 1¼ WAR worse the year after the extension was signed, compared to the year before -- but no worse than performance two years before, and no worse than the average of the three years before. This seems to have surprised them, but it didn't surprise me. My guess is that teams are motivated to sign players pre-free agency right after particularly good seasons from the player, which are likely to be random improvements over the player's usual level. The year after, the player would tend to return to his normal production as represented by his statistics for the previous and subsequent seasons.

Brandon J. Erickson, Paul Buchheit, Joseph Rauch, Rob Segedin, Michael G. Ciccotti, and Steven B. Cohen (2023), Hidden pitches in major league baseball, Orthopaedic Journal of Sports Medicine, Vol. 11 No. 4, Article 23259671231162864.

Pitch counts characteristically include only those thrown during actual game action, but, with the help of the Phillies organization, these authors studied pitches thrown by 137 pitchers in the Phillies system (MLB and MiLB) outside of the game itself, in warmups either before entering or between innings. Averaged over those 137 pitcher-seasons, the average total of pitches thrown during games was 771.0 per season, but the "hidden pitch" (authors' term) count averaged 678.3. Thus, the usual pitch count represents only 55 percent of the total pitches thrown on game days. An implication is that our customary pitch count estimates are way too low.

Sixty-six of the 137 pitchers spent time on the injured list during the season, and, contrary to the authors' initial hypothesis, they found no differences in either in-game, hidden, or total pitches between the injured and not-injured groups.

Unfortunately, the count of 66 included all injuries, of which only 30 were shoulder or elbow problems, so the hypothesis was not cleanly tested. As the authors note, we need a larger study, but, in addition, I suggest we need to pay attention to the injury types rather than lumping all of them together.

Brittany Bond and Ethan Poskanzer (in press), Striking out swinging: Specialist success following forced task inferiority, Organization Science.

Based on 1999 to 2018 data from Retrosheet, the authors uncovered evidence that when pitchers batted and made out, they were slightly more likely to get the next half inning's leadoff hitter out, as compared with that individual pitcher's baseline performance. This effect decreased with subsequent hitters and was gone completely by the fourth batter of the inning.

The effect was also greatest with a tie score. It weakened with increasing differences in score, and disappeared completely with a margin of 4+ runs. Also, pitchers were also more likely to throw strikes and walk fewer leadoff hitters after having made an out at the plate. That effect worked out to 0.018 runs.

Looking at the data in the other direction, there was no correlation between pitching performance so far that game and the results of the pitcher's own subsequent plate appearance.

When interviewed on the topic, several MLB pitchers reported that making out at the plate motivated them to pitch more aggressively.

Ryan S. Brill, Sameer K. Deshpande, and Abraham J. Wyner (2023), A Bayesian analysis of the time through the order penalty in baseball, Journal of Quantitative Analysis in Sports, Vol. 19 No. 4, pp. 245-262.

Is the "times-through-the-order penalty" a sudden drop in effectiveness at the beginning of each time through, as some research has suggested, or instead a steady degradation of pitcher effectiveness as the game progresses? This study, based on 2012 to 2019 Retrosheet data, supports the "steady degradation" hypothesis.

The authors controlled for batter and pitcher quality (via wOBA), handedness, home team, and away team. Cognizant of the selection bias that most TTOP studies have suffered from – that looking at pitchers who went deep in the game biases the sample towards pitchers who were luckier than normal that outing – they compared the pitcher's results to what would be the expected wOBA for a particular pitcher/batter matchup, rather than the pitcher's actual performance that day. They did this for all PA in the sample, whether in the control group (first three times through the order) or treatment group (4^{th}+ time through).

The results? The study found a linear – not sudden – degradation of starter performance across the game.

As for the magnitude of the degradation, the fourth-time-through-the-order effect was about the same as home-field advantage, and again about the same size as the platoon effect. As expected, it was smaller than the impact of batter and pitcher quality.

Charlie Pavitt, chazzq@udel.edu ♦

The Competitive Advantage of the Pitcher's Park

Bill James

Bill James finds that teams who play in pitcher's parks have had better records, historically, than teams who play in hitter's parks. He presents the data showing the effect, and then offers a suggestion for why this may be happening.

This research addresses the question of whether and to what extent teams which play in a pitcher's park have a competitive advantage in putting together championship teams.

The idea that there seems to be a competitive advantage to playing in pitcher's parks first occurred to me about 1971, when the Baltimore Orioles were the strongest organization in baseball. The Orioles played in a pitcher's park. The then-recent publication of the 1969 Macmillian Baseball Encyclopedia had introduced some organized information about park effects to the public, so I was aware that the Yankees throughout their storied history had generally played in parks which favored the pitcher. On the other side, I had grown up as a compulsive fan of the Kansas City A's, who (a) played in a hitter's park, and (b) lost with even more regularity than the Yankees won. In 1968 the A's moved to Oakland, a pitcher's park, and edged over .500 at once, then moved on to championship teams. I thought I saw a pattern here. Winning teams tend to play in pitcher's parks.

Of course I was aware of the Brooklyn Dodgers and other teams which have done well in hitter's parks. And of course the pattern was mixed. I know that I wrote that I believed that this was true more than 40 years ago. I have probably done some research related to this idea sometime in the last 40 years, but if so I don't remember it, and I'm sure others have researched the idea as well, but if so I haven't seen it or studied it. So this is all news to me.

Method

I started with a file of 329,988 game scores from 1921 to 2018, a file I have used many times.

First, I figured the park factor for each team each season. There is, of course, a "raw" park factor, which contrasts the runs scored per game at home with the runs scored per game on the road, and an "adjusted" park factor which adjusts the raw park factor for the overall impact of the park on the team's players' statistics. For this study I used the Raw Park Effect, since we are not studying anything related to player statistics here.

Second, I sorted the teams in the study into nine groups by Raw Park Effect (RPE):

Group 1	1.35 -	Extreme Hitter's Parks
Group 2	1.25 - 1.35	Very Good Hitter's Parks
Group 3	1.15 - 1.25	Better Hitter's Parks
Group 4	1.05 - 1.15	Hitter's Parks
Group 5	.96 - 1.05	Neutral Parks
Group 6	.88 - .96	Pitcher's Parks
Group 7	.80 - .88	Better Pitcher's Parks
Group 8	.72 - .80	Very Good Pitcher's Parks
Group 9	- .80	Extreme Pitcher's Parks

Third (and finally) I figured, for each group of teams, their winning percentage at home and their winning percentage on the road.

Data and Conclusions

The data is … what should I say? Proof? Nah; proof would be the wrong word. The data is 100% consistent with the theory that teams which play in pitcher's parks have a competitive advantage in building a team. Table 1 is a chart summarizing the data.

Table 1 – Team Winning Percentage by Raw Park Effect

Group	Raw Park Effect	# Teams		Wins	Losses	WPct	Home Field Adv.
1	Extreme Hitter's Parks 1.35 or higher	31	Home	1213	1071	.531	
			Road	910	1389	.396	
			Total	2123	2460	.463	.135
2	Very Good Hitter's Parks 1.25 to 1.35	57	Home	2186	1890	.536	
			Road	1796	2257	.443	
			Total	3982	4147	.490	.093
3	Better Hitter's Parks 1.15 to 1.25	160	Home	6331	5756	.524	
			Road	5392	6805	.442	
			Total	11723	12561	.483	.082
4	Good Hitter's Parks 1.05 to 1.15	451	Home	18148	15902	.533	
			Road	15594	18531	.457	
			Total	33712	34433	.495	.076
5	Neutral Parks .96 to 1.05	638	Home	26140	22400	.539	
			Road	22339	26185	.460	
			Total	48479	48585	.499	.078
6	Good Pitcher's Parks .88 to .96	520	Home	21520	17787	.547	
			Road	18137	21036	.463	
			Total	39657	38823	.505	.084
7	Better Pitcher's Parks .80 to .88	246	Home	10425	8480	.551	
			Road	8879	10003	.470	
			Total	19304	18483	.511	.081
8	Very Good Pitcher's Parks .72 to .80	63	Home	2627	2068	.560	
			Road	2253	2452	.479	
			Total	4880	4520	.519	.081
9	Extreme Pitcher's Parks Less than .72	10	Home	426	309	.580	
			Road	363	359	.503	
			Total	789	668	.542	.077

The teams which play in the very best hitter's parks have an aggregate winning percentage of .531 in their home parks, which is near the normal or average home-field winning percentage. In road games, however, they have an appalling .396 winning percentage. I feel like there should be an exclamation mark in there somewhere: they have a .396 winning percentage on the road! In the way this is usually interpreted, we could say that they have a much larger-than-normal home field advantage (.135). A look at the data, however, shows that this not really a home field "advantage". What it really is is a road team *dis*-advantage. These teams have a normal home-field advantage at home; they just absolutely stink on the road.

Of course, this is a group of only 31 teams, so you really can't draw any conclusion from that. Eleven of the 31 teams are Rockies teams of the last 30 years, and another six are Philadelphia Phillies assemblages from the Baker Bowl era, when the Phillies were hopeless. A small sample of teams heavily colored by some historic failures doesn't mean anything, write it off.

Except that that data is fully and completely consistent with the rest of the chart. All of the groups of teams, of course, have a higher winning percentage at home than on the road. For all four groups of teams which play in hitter's parks, however, the winning percentage on the road is further under .500 than the winning percentage at home is above .500, which means that all four groups have overall losing records, which get worse and worse as the park factor gets higher. On the other hand, all four groups of teams which play in pitcher's parks have overall winning percentages over .500. The overall winning percentages, by group, are:

Group 9	Extreme Pitcher's Parks	.542
Group 8	Very Good Pitcher's Parks	.519
Group 7	Better Pitcher's Parks	.511
Group 6	Pitcher's Parks	.505
Group 5	Neutral Parks	.499
Group 4	Hitter's Parks	.495
Group 3	Better Hitter's Parks	.483
Group 2	Very Good Hitter's Parks	.490
Group 1	Extreme Hitter's Parks	.463

Group 9 is just ten teams, so that doesn't mean anything, and yes, Group 2 data is a little bit out of line, but ... it's about as clear a pattern as you get with real-life data. The consistent pattern of the data is that teams that play in hitter-friendly parks have very low winning percentages on the road, which get steadily better as the environment shifts to favor the pitcher: .396, .443, .442, .457, .460, .463, .470, .479, .503. Teams which play in hitter's parks struggle in road games.

The teams also improve *at home* as the offensive furnace cools off, but not by nearly as much and with much less consistency. In the home games, the winning percentages for the nine groups (in the same order as before) are .531, .536, .524,.533, .539, .547, .551, .560 and .580, the final figure based on just ten teams.

As the environment for the hitter gets better and better, the home field "advantage" grows larger and larger—not as much because the teams are playing better at home, but mostly because they get worse on the road.

An issue that should be addressed here is, "could this just be a Yankee thing?" A disproportionate share of the most successful teams in baseball history were Yankee teams that played in a pitcher's park. Could it be that this effect is just a manifestation of the excellence of the Yankee franchise, which isn't "caused" at all by playing in pitcher's parks?

I excluded all Yankee teams from the study, and re-ran the numbers. The effects noted were slightly *larger* in the non-Yankee sample than in the complete study.

Interpretation, or "Why, do you think?"

I have two reasons to offer you as to why this is true. The two explanations are cooperative, rather than competitive, meaning that they can both be true at the same time, and, in fact, I believe that both explanations ARE true. The explanations are based on:

- Confidence, and
- Perception.

Success in all things is, of course, tied to confidence. If you believe you can do it, that helps. I believe that it is easier for pitchers to build confidence when working in a pitcher's park. You throw a pitch in San Francisco, you get a routine fly ball to left. You throw the same pitch in Boston, it bangs off the wall for a double. A friendly park helps you to build confidence in your ability to pitch in the majors, but also, it helps you survive and gain experience. If you pitch in a pitcher's park it helps you to stay around and figure out what you are doing.

The argument will be made that if a pitcher's park helps build the confidence of a pitcher, wouldn't a hitter's park help build the confidence of a hitter? Yes, but the two activities are not remotely parallel in the relevant sense. Confidence is much more helpful to public speaking than it is to winning a 100-yard dash. Public speaking, if you think you can do it, you can figure out the rest of the problem. Sprinting, you can do it or you can't. Hitting a 95-mph fastball is much more of a you-can-do-it-or-you-can't type of thing. In my opinion, expressed with some trepidation, confidence is much more relevant to learning to pitch in the majors than it is to learning to hit.

The trepidation is this. In the infancy of sabermetrics I ran a journal called the Baseball Analyst, and we had a discussion about why so many home runs were hit in Atlanta, the park then being nicknamed The Launching Pad. The dimensions of the park were almost the same as the dimensions of Busch Stadium in St. Louis, but many, many more home runs were hit in Atlanta.

An engineer who was a member of our group studied the problem, and concluded that it was a psychological effect[1]. People thought they could hit home runs in Atlanta, so they did. This was, of course, a ridiculous explanation; actually it is hard to believe that just 40 years ago, we understood the problem so poorly that anyone would suggest that was the reason. Another gentleman who was part of that group broke the code with a later article. Dick O'Brien noticed that in the Texas League within a group of seasons all eight parks had nearly identical dimensions and fence heights, but very different elevations. Dick's study[2] showed that the number of home runs hit in each park tracked precisely with the elevation of the park. Now this is so commonly understood that it is hard to believe that there was ever a time when it wasn't. Fulton County Stadium was the highest park in the majors at that time, more than 1,000 feet above sea level. The first gentleman noted this but dismissed the significance of it.

In our society … I think this was more true 40 years ago than it is now, but it is still true. In our society we use psychology exactly the same way people centuries ago used witchcraft. We use it to explain whatever we don't understand. My attribution of this effect (in part) to *confidence* runs the risk of looking in retrospect like the attribution of home runs in Atlanta to *psychology*. But confidence is nonetheless a real thing, and I think it is true, so I will go ahead and say it.

The other reason that I believe this happens is that the hitting environment gives shape to a hitter's performance image. Historically, teams made near-zero effort to adjust their understanding of a player's performance for park effects. A good defensive middle infielder who hits .265 with single-digit home runs—that is, Dee Gordon, Jose Iglesius, Alcides Escobar, Juan Segura, Andrelton Simmons, Erick Aybar, Garry Templeton, Ozzie Guillen, Jose Lind, Robby Thompson, Alfredo Griffin, Ron Oester, Larry Bowa, Phil Garner, Frank White, Ivan DeJesus—such a player can be a valuable performer on a team in a "normal" offensive context. But if you put those same offensive numbers on a team in a high-offense context, you may have a player who is completely useless.

The Colorado Rockies shortstop from late 1996 to mid-2001 was Neifi Perez. Perez was an OK shortstop, perhaps a good shortstop, and he had offensive numbers which, without adjusting for context, looked excellent. From 1997 to 2001 he hit .291, .274, .280, .287, and .298. He hit as many as 39 doubles in a season. One season he led the league in triples with 11, and in a five-year stretch his triples were 10, 9, 11, 11 and 8, with about the same numbers of homers. Adjusted for context, he was useless. His career high in WAR in that era was 1.4. (He had one season later when he got it up to 2.1.) His career batting average in Colorado was .321. In all other parks, it was .245.

When I was with the Red Sox, one of our key executives came to us from the Rockies. He said that the other Rockies executives in that era thought that Perez was a treasure, an untouchable in trade talks, one of the best shortstops in baseball. That is the problem that hitter's parks cause.

[1] See http://tinyurl.com/HRpsychology --Ed.
[2] See http://tinyurl.com/obrienelevation --Ed.

Of course, the argument can again be offered that if this hurts you on one side of the game, offense, wouldn't it help you on the other side? Park illusions cause you to think that your hitters are better than they are, which is a bad thing, and park illusions cause you to think that your pitchers are worse than they are, which is also a bad thing. In essence, I believe that teams in pitcher's parks are usually trying to improve their hitting while allowing their pitchers to work and improve, while teams in hitter's parks get complacent with their offenses, while shuffling pitchers in and out.

Further research should be able to confirm:

- That teams in hitter's parks tend *not* to replace regular players with 1.0 to 2.0 WAR, or at least not are not as likely to do so as teams in pitcher's parks, and

- That teams in hitter's parks develop fewer successful pitchers than teams in pitcher's parks, but shuttle their pitchers in and out more quickly.

And further research could focus on the issue of to what extent this is still true, as modern executives may (or may not) have learned to combat the illusions created by park effects.

Bill James, biljames@aol.com ♦

Back issues

Back issues of "By the Numbers" are available at the SABR website, at http://sabr.org/research/statistical-analysis-research-committee-newsletters, and at editor Phil Birnbaum's website, www.philbirnbaum.com .

The SABR website also features back issues of "Baseball Analyst", the sabermetric publication produced by Bill James from 1981 to 1989. Those issues can be found at http://sabr.org/research/baseball-analyst-archives.

Submissions

Phil Birnbaum, Editor

Submissions to *By the Numbers* are, of course, encouraged. Articles should be concise (though not necessarily short), and pertain to statistical analysis of baseball. Letters to the Editor, original research, opinions, summaries of existing research, criticism, and reviews of other work are all welcome.

Articles should be submitted in electronic form, preferably by e-mail. I can read most word processor formats. If you send charts, please send them in word processor form rather than in spreadsheet. Unless you specify otherwise, I may send your work to others for comment (i.e., informal peer review).

I usually edit for spelling and grammar. If you can (and I understand it isn't always possible), try to format your article roughly the same way BTN does.

I will acknowledge all articles upon receipt, and will try, within a reasonable time, to let you know if your submission is accepted.

Send submissions to Phil Birnbaum, at 110phil@gmail.com .

"By the Numbers" notifications

SABR members who have joined the Statistical Analysis Committee will receive e-mail notification of new issues of BTN, as well as other news concerning this publication.

The easiest way to join the committee is to visit http://members.sabr.org, click on "my SABR," then "committees and regionals," then "add new" committee. Add the Statistical Analysis Committee, and you're done. You will be informed when new issues are available to download from the SABR website.