



AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®

732 North Washington Street, Alexandria, VA 22314 • (703) 684-1221 • Toll Free: (888) 231-3473 • www.amstat.org • www.twitter.com/AmstatNews

Why Did Home Runs Surge in Baseball? Statistics Provides Twist on Hot Topic

Vancouver, Canada (July 12, 2018): Around the middle of the 2015 season, something odd started happening in Major League Baseball (MLB): Home runs surged. They surged again in 2016, from the previous year's 4,909 to 5,610, and then again in 2017 to an all-time high of 6,105.

What was going on? For a stats-mad sport, the mystery was irresistible. There was the theory of the "Juiced Ball." Some subtle, possibly unintentional change in the manufacturing process had given balls just enough extra bounce to change history. Then there was the batter approach theory, which speculated that just a little bit more of an uppercut swing—perhaps in part due to defensive shifts—was giving the ball extra lift. Maybe batters were just cranking it as hard as they could and going for home runs given this shift to stronger defensive tactics?

And then there was a massive investigation requested by the MLB commissioner, who asked 10 scientists to find out what was going on. They tested a lot of balls and concluded it was a case of reduced drag combined with the launch angle of the ball coming off the bat.

But Jason Wilson, a statistician at Biola University in Southern California, has a different explanation. The poorer the pitch, the easier it is to whack a home run—and the quality of pitching between 2015 and 2017 had gotten worse if you broke a pitch down into measurable components and then measured pitching quality over time. Wilson called this measure "Quality of Pitch" (QOP).

The idea for measuring pitch quality began in 2010, with Jarvis Greiner, one of Wilson's students. Greiner combined an interest in statistics with being a film major and a pitcher on the college baseball team. "He had the idea that we could quantify the quality of a curve ball," says Wilson, "and for his class project, he videotaped curve balls against tape measures. The data turned out to be great, and we ended up publishing it as an academic paper. Then his father, Wayne Greiner, who works for a sports distribution company and is absolutely passionate about baseball stats, asked, 'Could this be scaled up to analyze all kinds of pitches in the MLB?' Thanks to the introduction of cameras in stadiums in 2008, we had access to tons of PITCHf/x data, and—yes—our original model did generalize quite nicely."

With Greiner senior, Wilson refined the QOP statistic. At its simplest, QOP describes how difficult a pitch would be to hit on a scale of zero to 10. "The first thing we did [was] break a pitch down into six components," says Wilson. "The first component is rise on the pitch. If there's any rise, that's a tell that it's probably a curve ball, and that counts against the quality of the pitch.

"Then there's the distance until the ball starts to break and go down. The farther out, the better. Third is the total vertical break; again, the more break, the better. Fourth is the horizontal break, and the more break horizontally, the better. We also incorporate velocity, so the faster the pitch, the better. And the

final component is location, the strike zone. The corner's the best spot, the middle is bad, and if you are far outside the strike zone, well that's obviously bad, too. We combine all these into a single number, which is the QOP value."

Wilson and Greiner then began to model what happened on the field between 2016 and 2017. From the six components of the QOP, vertical break was the most important predictive variable—and it had dropped sharply. What that meant in practice was that after looking at more than 700,000 pitches per season, they found the balls were being pitched more directly than previously at the batter. They were higher in the zone; there was less variation in where they crossed.

Wilson is quick to add that with more than 700 pitchers per season, a single factor cannot explain the entire surge. But the drop in vertical break makes sense if you think about it as a way of combating the batter's upward swing—pitching higher up would make it harder to pull off a home run.

Of course, Wilson's analysis shows that if this was indeed a pitching strategy, it didn't work. QOP says Wilson can explain between two to four percent of the change in the home run number (113 to 226 home runs) based on pitching, which turns out to be 23 percent to 46 percent of the home run increase between 2016 and 2017.

The big news for 2018? Home runs are down—and if you look at the data through Wilson's model, the quality of the pitching is up.

Talk details:

The Home Run Spike of MLB 2017: Drop in Quality of Pitch (QOP) Is a Missing Factor

Tuesday, July 31, 2018

3:05–3:50 p.m.

<http://ww2.amstat.org/meetings/jsm/2018/onlineprogram/AbstractDetails.cfm?abstractid=332642>

For details, contact Jason Wilson.

Email: jason.wilson@biola.edu

Webpage: www.biola.edu/directory/people/jason-wilson

Phone: (951) 743-2172

About JSM 2018

JSM 2018 is the largest gathering of statisticians and data scientists in the world, taking place July 28–August 2, 2018, in Vancouver. Occurring annually since 1974, JSM is a joint effort of the American Statistical Association, International Biometric Society (ENAR and WNAR), Institute of Mathematical Statistics, Statistical Society of Canada, International Chinese Statistical Association, International Indian Statistical Association, Korean International Statistical Society, International Society for Bayesian Analysis, Royal Statistical Society and International Statistical Institute. JSM activities include oral presentations, panel sessions, poster presentations, professional development courses, an exhibit hall, a career service, society and section business meetings, committee meetings, social activities and networking opportunities.

About the American Statistical Association

The ASA is the world's largest community of statisticians and the oldest continuously operating professional science society in the United States. Its members serve in industry, government and academia in more than 90 countries, advancing research and promoting sound statistical practice to inform public policy and improve human welfare. For additional information, please visit the ASA website at www.amstat.org.