

Do Steroids Increase Home Runs?

by Ryan Young¹

Working Paper May 2009

Introduction

Was the sudden outburst of home run records in the 1990s and 2000s aided by steroids? Many players think so. Jose Canseco, who admitted to having used steroids for his entire major league career, absolutely believes they padded his statistics. He even mused that he would not have even made it to the big leagues without them. "You feel the strength, and the stamina," he told *60 Minutes*.²

Former Boston Red Sox pitcher Curt Schilling describes steroid use as "cheating," adding that, "if you get caught using steroids, you should have everything you've done in this game wiped out for any period of time that you used it."³

In the academic world, Tufts University physicist Roger Tobin estimates that, while steroids have little effect for non-power hitters, they could increase a top player's home runs by 50 to 100 percent.⁴

The consensus is that steroids increase home runs. Economist Art De Vany disagrees. He denies that performance-enhancing drugs actually enhance performance, at least as far as home runs are concerned.

This is a difficult case to make. Decades-old records set by names like Ruth, Maris, and Aaron all fell within a few short years. And most of the record-breakers also happened to be steroid users.

The case is not black-and-white. Nobody knows how many players used steroids. Some pitchers also used steroids, such as Roger Clemens. If a juiced pitcher faces a juiced batter, who has the edge?

Uncertainties aside, few people argue that the coexistence of steroids and record-setting home run hitting is a coincidence. Is De Vany on to something, or is he simply being contrarian?

¹ Ryan Young is the Warren T. Brookes Journalism Fellow at the Competitive Enterprise Institute in Washington, D.C.

² <http://www.cbsnews.com/stories/2005/08/05/60minutes/main761932.shtml>

³ http://www.boston.com/sports/baseball/redsox/extras/extra_bases/2007/05/curt_on_bonds_1.html

⁴ Roger G. Tobin, "On the Potential of a Chemical Bonds: Possible Effects of Steroids on Home Run Production in Baseball," *American Journal of Physics*, vol. 76, 15-20 (January 2008).
<http://www.tufts.edu/~rtobin/Tobin%20AJP%20Jan%2008.pdf>

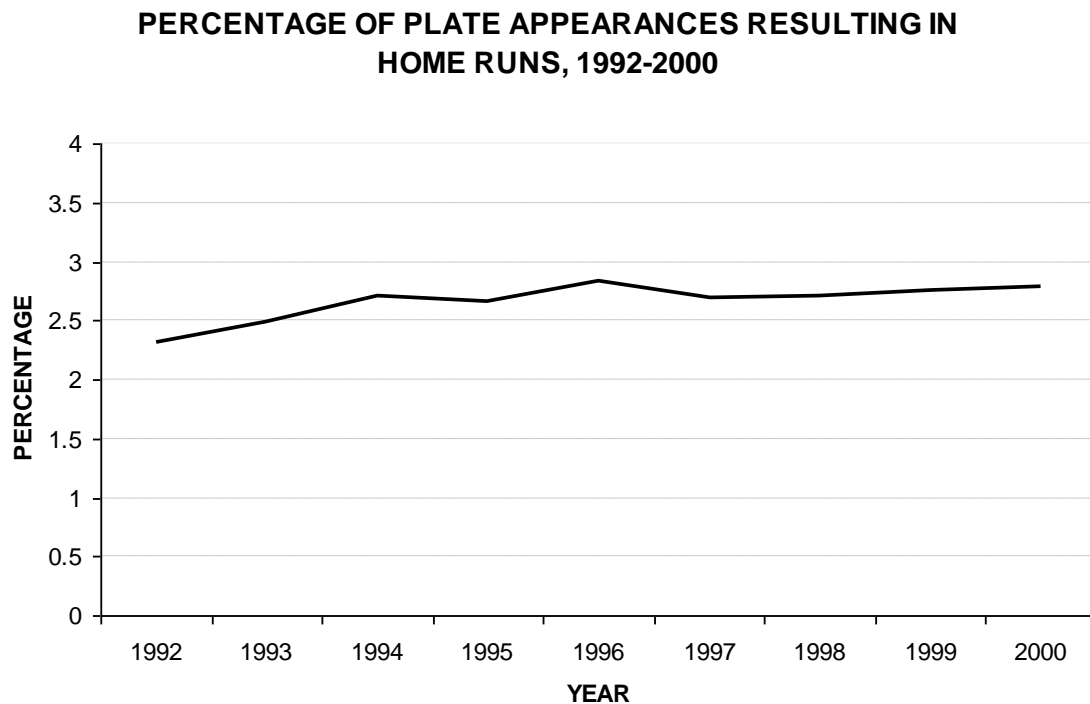
He offers several reasons why the age-old battle between pitcher and batter tilted the batters' way during the steroid era. The baseball is made differently than it used to be. Stadiums are smaller. So is the strike zone. Today's players are bigger and stronger, and not just because of steroids.

De Vany also believes that simple chance played a role in all the broken records. He makes a good case; it is still likely that steroids played a causal role. It is also likely that that role has been exaggerated.

Home Runs Have Been Increasing for a Long Time

Exhibit A for why people believe steroids have an effect on home runs is shown below in Figure 1: home runs went up during the steroid era. The change looks small in the chart. But there are around 50,000 plate appearances in an average season. Even a small change in the percentage of those plate appearances resulting in home runs can make a big difference.

Figure 1.



Source: data from <http://www.baseball-reference.com>.

Percentage of plate appearances resulting in home runs is the most accurate available measure for detecting trends in major league home run hitting. Tallying up the number of

home runs hit per year would be misleading because there are more teams than before (30 to 16), and the schedule contains more games than it used to (162 to 154). The numbers would curiously jump every time an expansion team begins play.

Home runs per game is a better measure. It is not sensitive to the number of teams or the length of the schedule; it is still inaccurate because every game is different. Low-scoring games have fewer plate appearances than high-scoring games. Some games go into extra innings. Others end prematurely due to rain.

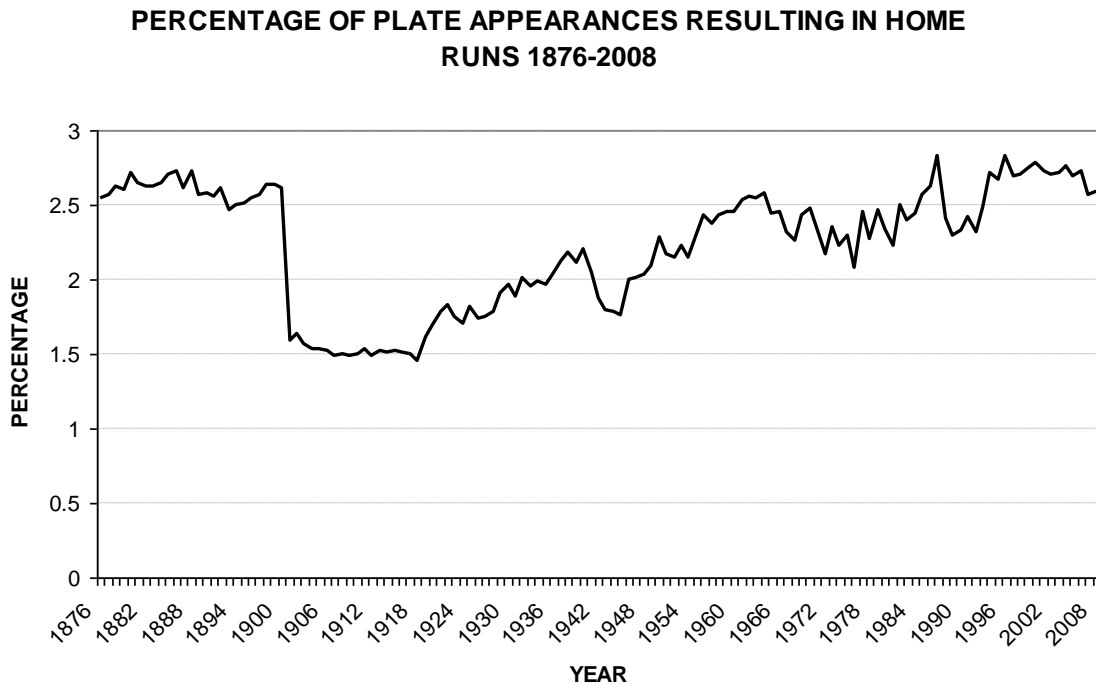
The best approach is to add up every opportunity hitters have to hit a home run, and see what they made of those opportunities. Plate appearances are used instead of at-bats because at-bats do not count all opportunities to hit a home run. If a player is walked or hit by a pitch, then a plate appearance does not count as an at bat. Plate appearances include every single opportunity a player has to hit a home run.

From the beginning of major league play in 1876, the average percentage of plate appearances resulting in home runs is 2.24%. In 1996, 2.83% of plate appearances resulted in home runs. This was an all-time high, and a sharp increase from just a few years before. Home runs declined somewhat after the league cracked down on steroids in the wake of the Mitchell report in 2007; they remain historically high.

The trend is clear. Home runs went up during the steroid era, and went down after. From that view, it is certainly plausible that steroids aided the increase – and their (presumed) absence caused the decrease.

Taking a longer view of baseball history reveals something more. Home runs have been increasing for a long time. In fact, the steroid era's home run numbers appear to be roughly in line with a 90-year trend.

Figure 2.



Source: data from baseball-reference.com.

The battle between pitcher and hitter has had many systoles and diastoles over the years. Steroid-unrelated reasons can explain the various twists and turns. They probably also explain most – but not all – of what happened to home runs during the steroid era.

The Baseball

The first such twist was the steep drop at the beginning of the 20th century, known as the deadball era. This first swing of the pendulum went toward the pitcher, and away from the hitter.

Today, the average lifespan of a ball is 6 pitches. Most games will go through more than 40 balls. This is necessary to keep conditions consistent. But professional baseball had humble beginnings. To save money, a single ball would often have to last for an entire game.

That single ball would become misshapen and discolored by the later innings. Not only did no-longer-spherical balls move differently coming out of the pitcher's hand, but dirt and grass stains made it harder to see. Battered balls don't carry as far as new ones. Hitting for power was extremely difficult during the deadball era.

A rule change meant foul balls began counting as strikes, making hitters far more cautious about fouling off pitches that they didn't like. Large new ballparks – much larger than today's – were built to accommodate baseball's growing popularity.

Fenway Park's predecessor, the Huntington Avenue Grounds, opened in 1901 with its centerfield wall 530 feet away from home plate. It was lengthened to 635 feet (!) in 1908. Home runs were few and far between at parks like these. All these factors added up to advantage the pitcher.

Babe Ruth put an end to the deadball era. He hit a league-record 29 home runs in 1919. Fans loved it, and began flocking to games. Other players copied his approach, and new players were hired based on how much power they brought to the plate. Home runs steadily increased.

The league denies to this day that it helped the Babe by ordering harder, sturdier balls that would carry farther; the rumors persist anyway. Ball manufacturers independently confirmed that technological progress and mechanization allowed them to make more consistent and more tightly-wound balls.⁵

The Japanese professional leagues use a ball that is even more tightly wound than the modern major league ball. Japan's home run king, Sadaharu Oh, hit 868 career home runs. Ball composition matters.

Whether through conspiracy or merely the march of progress, the ball changed. Home runs went up and up, all without the aid of steroids.

Strategy Has Changed

Baseball began as a game of small-ball. Instead of concentrating on fence-clearing blasts, teams moved runners up a base at a time. The emphasis was on getting base hits and playing solid defense. The rise of the home run changed the way teams approached offense.

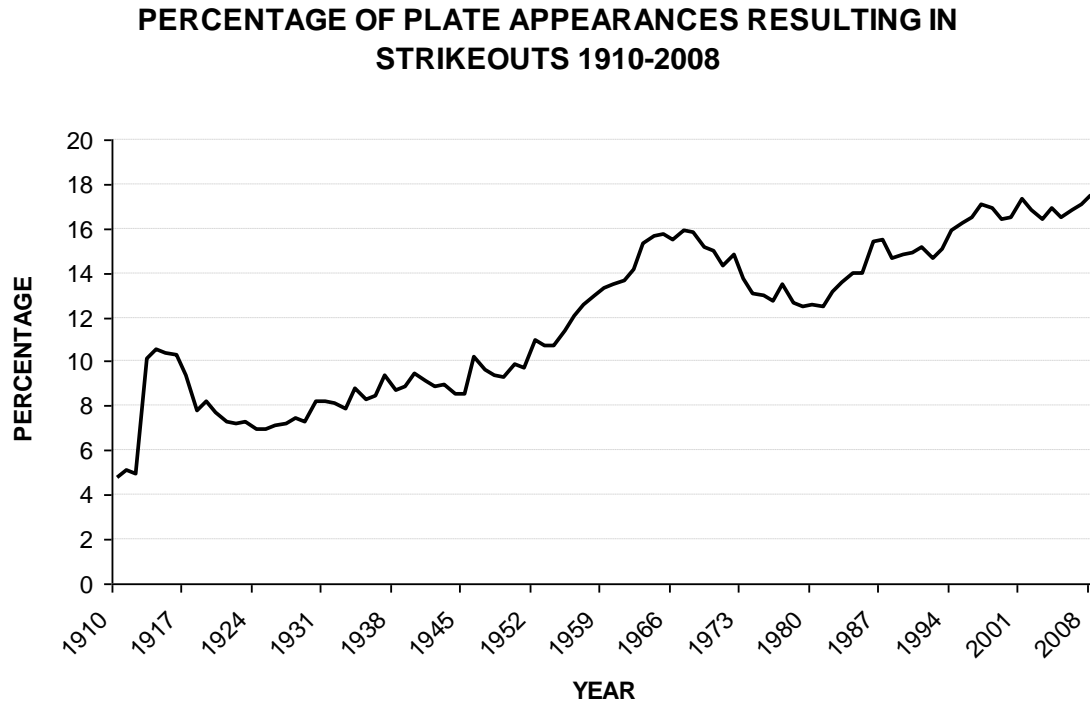
Home runs are feats of remarkable precision. The batter must hit the ball about half an inch below center. Regulation baseballs are only 2 7/8 to 3 inches in diameter. Pitchers routinely throw at over 90 miles per hour, and rarely on a straight line. The batter must meet this with a bat speed of about 70 miles per hour.

It also helps to swing with a slight uppercut to give the ball a steeper trajectory coming off the bat. With neither ball nor bat moving straight – and both moving extremely fast – the odds of hitting the ball just right are extremely low.

⁵ George Vecsey, *Baseball: A History of America's Favorite Game*, (New York: Modern Library, 2006), 70.

The trade-off to more home runs is more strikeouts. This makes some intuitive sense. The harder a player swings, the less time the bat spends in the strike zone. Less time in the strike zone means a lower chance of contact, and more swings-and-misses. The rise of the home run was also the rise of the strikeout. Baseball was becoming an all-or-nothing game well before the steroid era.

Figure 3.



Source: data from baseball-reference.com.

The continuing increase in strikeouts is evidence of a change in strategy. Hitters are hitting more home runs because they are trying to. If steroids have an effect, it is smaller than the effect of this new strategy.

Other than the long-run increasing trend, two things stand out in Figure 3 – the spikes at the end of the dead-ball era and in the 1960s. The dead-ball bump is probably due to players emulating Babe Ruth’s swing-for-the-fences approach without necessarily emulating his success. As teams began selecting for better power hitters, strikeout rates reverted to trend.

The 1960s spike is more complicated. Several rule changes were put in place to benefit pitchers, and they worked a little too well.

The pitching mound was raised. This increased the ball's vertical movement as it traveled from the pitcher's hand to the plate, making it harder for hitters to track. The strike zone was enlarged, forcing hitters to swing at pitches they'd rather not.

The changes culminated in 1968, the Year of the Pitcher. Bob Gibson had a record 1.12 ERA. Denny McClain won 31 games. Carl Yastrzemski's .301 batting average was the lowest ever to lead the American League.

Things quickly swung back the batters' way. Four expansion teams began play in 1969. Their inexperienced pitching staffs padded their opponents' statistics. The pitching mound was lowered from 15 inches to 10 inches. Home runs immediately went back up, and strikeouts immediately went down. They never returned to their pre-1960s levels, but they did revert to trend.

The all-or-nothing approach was now so ingrained in hitters that strikeout rates continued their slow climb even as the strike zone shrank again, allowing hitters to be more selective. The trajectory has continued more or less uninterrupted to the present day. If steroids had an effect, it is smaller than changes in strategy and rules.

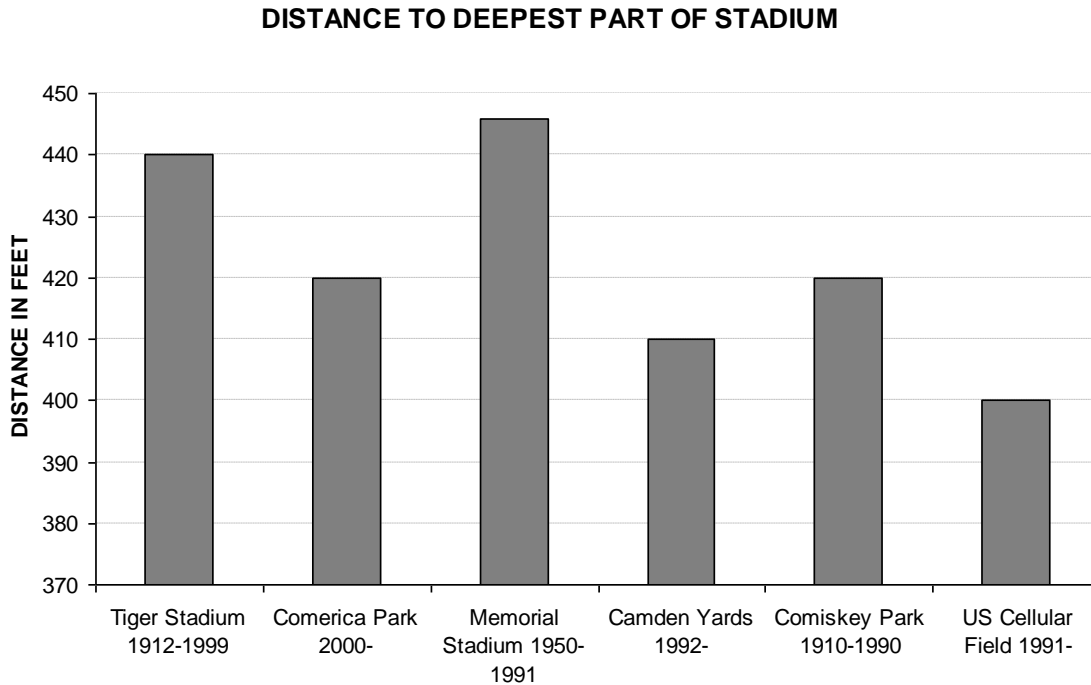
Today's Ballparks Are Smaller

There was a stadium-building boom underway by the dawn of the steroid era. 20 of baseball's 30 teams have built new stadiums since 1990. They typically have shorter fences than older stadiums, making home runs easier to hit. This makes some economic sense; fans love home runs, and will pay good money to see top sluggers practice their craft.

New, smaller ballparks gave batters an even stronger incentive to use an all-or-nothing approach. Shorter fences mean that what used to be a warning track fly was now a home run. The odds of slugging success became slightly higher.

Home runs plateaued near their historic highs throughout the steroid era until 2007, when the league took a stricter stance on steroid use. Though down slightly from the steroid era, home runs remain historically high, even though steroids are believed to have mostly disappeared from major league clubhouses. That the average stadium is smaller than before has something to do with it.

Figure 4.



Source: data from <http://www.andrewclem.com/Baseball/Dimensions.html>.

Figure 4 compares the sizes of old and new stadiums for three teams: The Detroit Tigers, Baltimore Orioles, and Chicago White Sox. In all three cases, the new stadium has smaller dimensions than the old. This pattern is typical.

There are some exceptions. Citi Field's deepest fence is 415 feet, compared to 410 for its predecessor, Shea Stadium. But Citi's dimensions taper off quickly, and are several feet smaller than Shea's everywhere except for deep right-center.

New stadiums are built to favor hitters. As more teams move into smaller parks, players should be expected to hit more home runs, steroids or not.

Players Are Bigger and Stronger, Naturally

Today's athletes are bigger, stronger, and fitter than ever before – even the ones who don't use performance-enhancing drugs. There are two reasons for this: conditioning and nutrition have improved over time, and players are drawn from a more numerous population.

According to *Baseball Digest's* Peter Schmuck the 1927 Yankees had an average height and weight of 5'11" and 176 pounds. They were nicknamed "The Murderer's Row."

They did not look the part compared to the average 2001 Yankee, who stood 6'2" and weighed 204 pounds.⁶

As nutrition, medical care, and living standards improved over the 20th century, the average person became larger and larger. It makes sense then, for players to increase in size along with the general population.

Modern players also train year-round. It was a rarity for players to so much as lift a weight in the old days. Offseason training regimens were almost non-existent, since low salaries required most players to take second jobs in the off-season.

Not only are people today larger, there are more of them. Population exploded in the 20th century. Teams have a bigger pool of bigger people from which to choose their players, and not just because of steroids. The National and American leagues merged in 1901. That year there were 16 teams with 25-man rosters, for a total of 400 players. U.S. population at the time was 78 million, so these were the top 1-in-195,000 players. This is shown in Table 1 below.

Table 1.

YEAR	U.S. POPULATION	MLB PLAYERS	NUMBER OF PEOPLE PER MLB PLAYER
1901	78 million	400	195,000
1951	154 million	400	385,000
2001	285 million	750	380,000

Sources: U.S. Census data, baseball-reference.com

50 years later there were the same number of teams and players, but nearly double the population; 1951's major leaguers were the lucky 1-in-385,000.

This is actually understating the increase. Baseball was hampered by a color barrier until Jackie Robinson made his debut in 1947, so its potential recruiting population in 1901 was actually quite a bit smaller than 78 million.

By 2001, baseball had expanded to 30 teams. Population had grown almost in lockstep, so the ratio of players-to-population was slightly down. This does not mean that teams had become less selective.

⁶ Peter Schmuck, "Does Size Really Matter? Today's Major leaguers Are Bigger and Stronger than those of Earlier Eras," *Baseball Digest*, July 1, 2001. http://www.accessmylibrary.com/coms2/summary_0286-738872_ITM

The U.S. population is no longer the only one that matters. Rosters now feature players from all over Latin America. Some of Japan's top players – themselves 324 of the best players from among Japan's 127 million people – are now playing in the United States. Baseball is growing in Korea and China, and those countries are beginning to export talent.

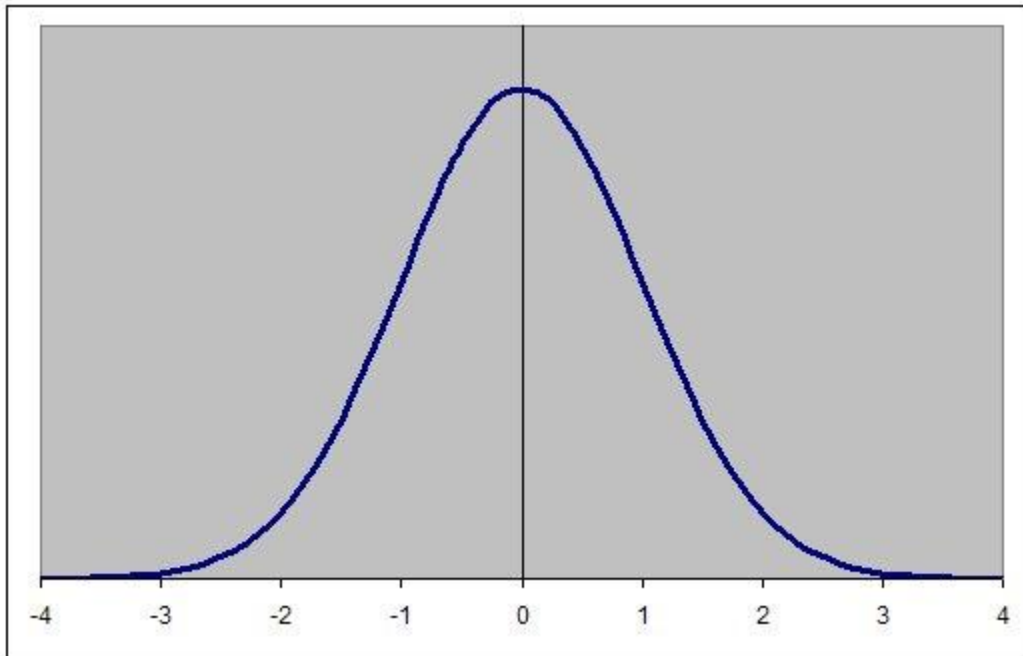
Since teams today have hundreds of millions more potential players, they can be choosier than ever. And the larger a population, the more people there are who will be several standard deviations above-average in terms of size and strength. Steroids are not the only reason why today's players are bigger, stronger, and hitting more home runs.

Still, people just seem to *know* that steroids are what explain inflated home run totals. Why would players use them if they didn't help? Steroid skeptics like De Vany have one more argument: according to the laws of probability, records will almost always fall, given enough time and enough opportunities. The achievements of McGwire, Bonds, and Sosa are well within the bounds of chance.

Extreme Accomplishment vs. Normal Distribution

Most people are familiar with what a Gaussian, or normal, distribution looks like. It is the same thing as the famous Bell Curve, shown below in Figure 5.

Figure 5. Example of a Bell Curve.



In a bell curve, most events lie toward the peak in the middle, with a few outliers at either end. They are rare, though. Records are difficult to break in a Gaussian world.

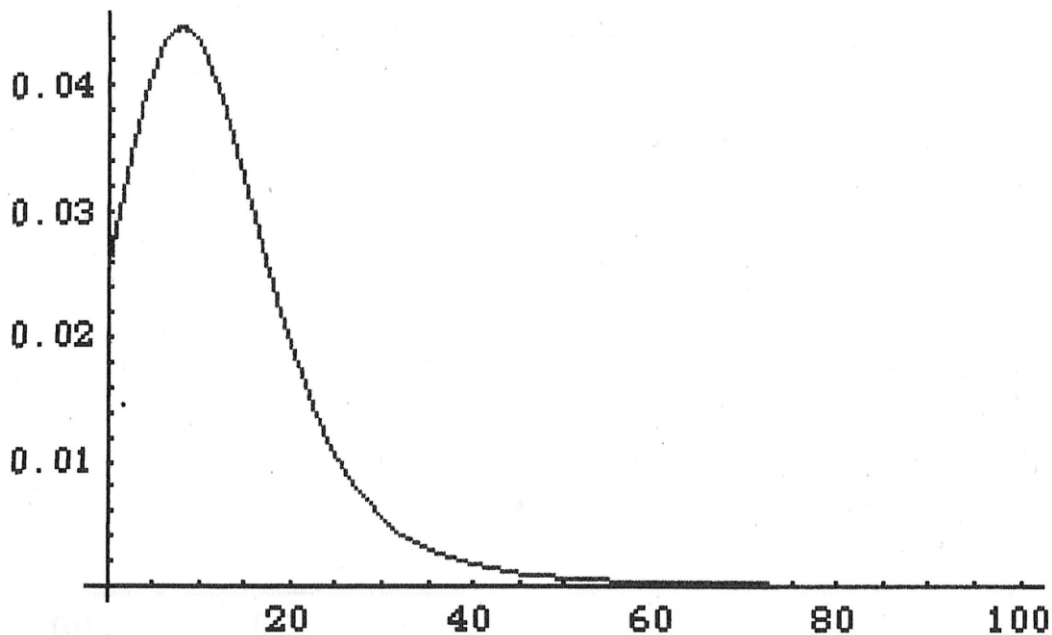
Human beings have an inherent cognitive bias to see the world as Gaussian. For whatever evolutionary reason, it is how our brains work. Extreme outliers like Barry Bonds instantly arouse suspicion, especially since he very likely used steroids. His statistics are so far away from the meat of the bell curve as to seem almost unreal.

De Vany argues that home run hitting cannot be explained by the Gaussian Bell Curve model. Babe Ruth and Barry Bonds are so many standard deviations above the mean that their very existence would be all but impossible in a Gaussian world. In such a place, “hitting 60 home runs... would not occur in millions of years [of play].”⁷

Home run outliers in baseball are far more common than the traditional bell curve model would predict. Therefore, the bell curve does not explain them. Actual major league data yields Figure 6 below:

Figure 6. (Figure 4 from De Vany)

Fig. 4.— The Law of Home Runs: Stable Probability Density Function of Home Runs



Source: Arthur De Vany, “Steroids, Home Runs, and the Law of Genius,” p. 35.

This curve is derived from the totals of how many home runs each major league player hits in each season with more than 200 at-bats. The average player hits 7.8 home runs per season, so the peak is toward the left of the graph, not the middle. After the initial crash

⁷ De Vany, pp. 8-9.

from the peak, the tail of the graph extends very far to the right – much further than a bell curve. The general pattern is very stable from season to season.

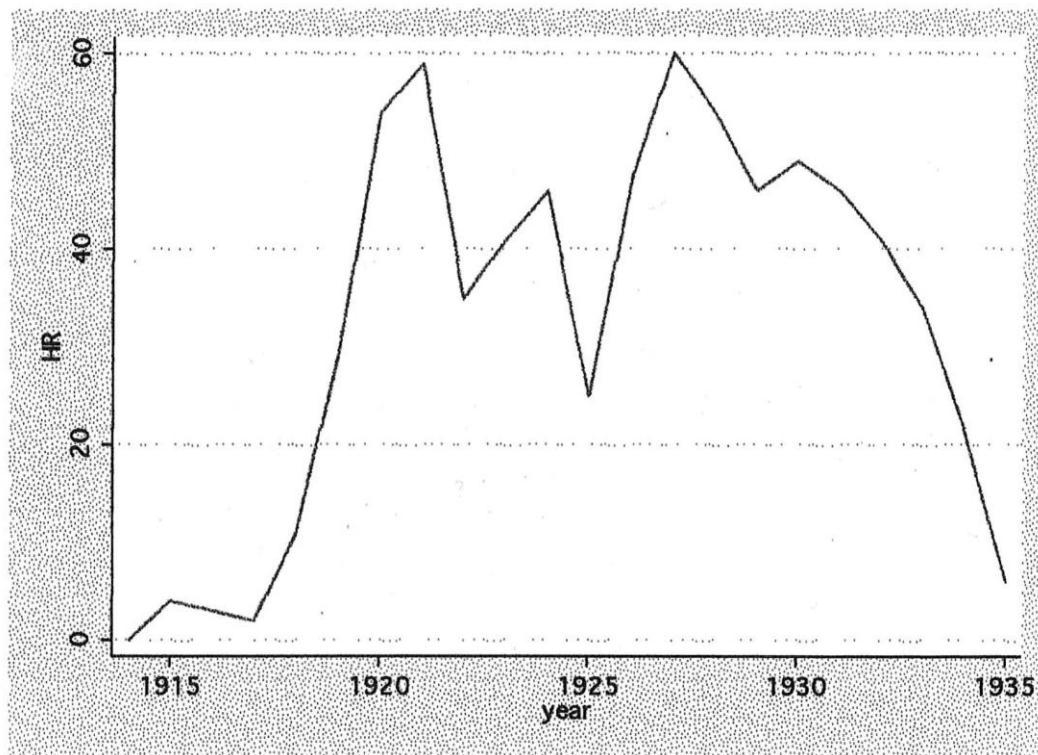
This shape reflects the fact that prolific home run hitting is extremely difficult, but also possible. Players who can hit twenty home runs are considered power hitters. Forty and above is purely elite territory. But, contrary to the Gaussian model's predictions, these elites keep showing up year after year. Maybe the feats of great individual hitters like Bonds, McGwire, and Sosa weren't necessarily aided by steroids.

Great Individuals

De Vany compares great sluggers to the Da Vincis and Michelangelos of human history. An exaggeration, yes. But in terms of how much Bonds, McGwire, and Sosa stood out from their peers, the comparison is justified. De Vany's point is that, as rare as they are, truly great hitters keep popping up every few generations. De Vany argues that Bonds and McGwire set their records because they really were the best, steroids or no.

The steady league-wide long-tail pattern hides the fact that individual home-run hitting actually varies wildly. Looking at individual player data instead of league-wide data makes that clear. Figure 7 below shows Babe Ruth's seasonal home run totals.

Figure 7. Babe Ruth's Seasonal Home Run Totals (Figure 2 from De Vany)



Source: Arthur De Vany, "Steroids, Home Runs, and the Law of Genius," p. 33.

Babe Ruth was arguably the greatest hitter of all time. His statistical lines are the sporting equivalent of Botticelli's *The Birth of Venus*. But even in his prime, Ruth's yearly totals varied wildly.

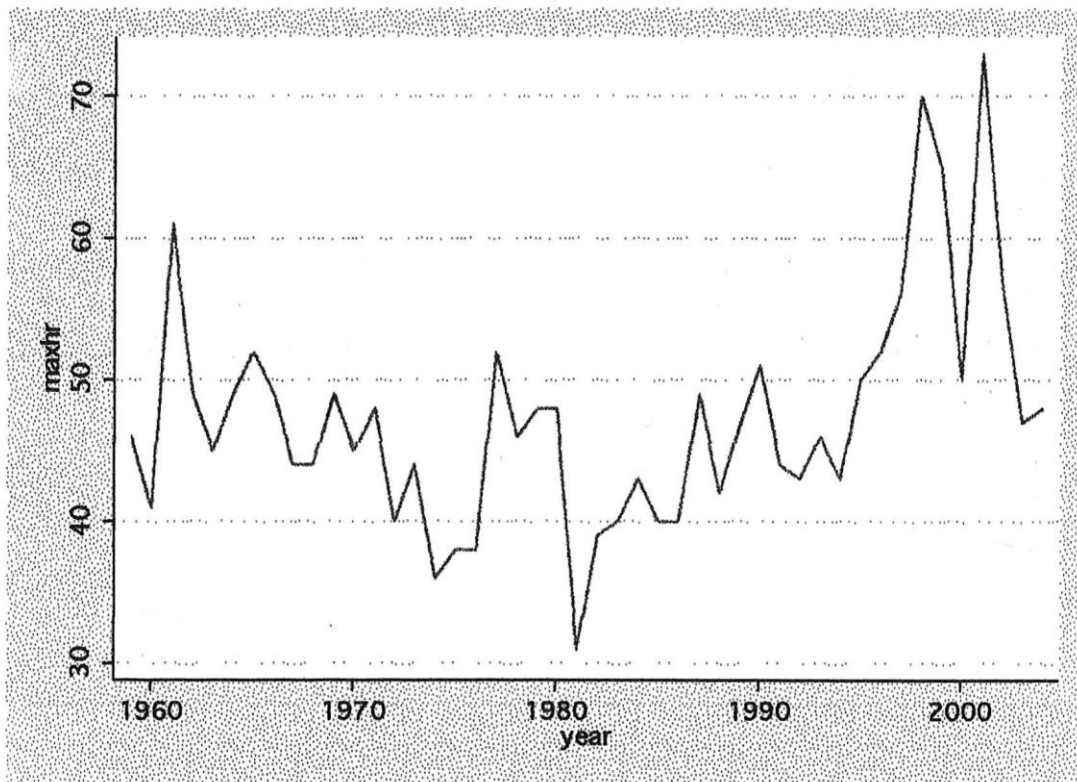
The 1921-22 dip from 59 home runs to 35 requires some explanation. It is lower mainly because Ruth played in 42 fewer games in 1922. He began the year serving a six-week suspension for participating in a barnstorming tour over the offseason, in violation of league rules for World Series participants (the rule was soon repealed).

But he played a full season in 1923 and still hit only 41 home runs. Ruth actually had 6 more plate appearances than in 1921 (699 to 693), and hit 18 fewer home runs. His variation alone was more than twice the average player's entire seasonal output.

1925 was an even worse year for the Babe. He played in only 98 games out of 154 due to illness, and home runs were not his only statistic to suffer. Even accounting for these outliers, the yearly variation is immense.

Another way to show how much individual home run hitting varies is to look at each year's league leader. De Vany shows that in Figure 8 below.

Figure 8. Maximum Home Runs by Year (Figure 8 from De Vany)



Source: Arthur De Vany, "Steroids, Home Runs, and the Law of Genius," p. 39.

The totals are all over the map. A bell curve cannot explain such high variability. The outliers run in both directions, too. Gaudy totals like 70 or 73 home runs are countered by totals as low as 39 in 1982 and 36 in 1974.

The lowest outlier is 31 home runs, from the 1981 season. It was shortened by a strike, so it doesn't represent a full season. 1994 was also shortened by a strike. The great steroid-era climb began the very next season, so 1994's artificially low total slightly exaggerates the slope of the increase.

Whatever the mitigating factors, the sudden and sustained steroid-era increase looks suspicious. The totals of 70 and 73 are the two highest in baseball history, and they happened just three seasons apart. It seemed unprecedented; it actually wasn't.

How often should Records Fall?

Ned Williamson hit 27 home runs in 1884 to set the season record. It stood for 35 years, until Babe Ruth hit 29 in 1919. Ruth later pushed the record to 60. That mark stood for 34 years until Roger Maris hit 61 in 1961.

Maris' record also had longevity – 37 years. Then it fell twice in the same year. In 1998, Mark McGwire hit 70 and Sammy Sosa hit 66. Both hitters also put up Maris-busting numbers the following season. Barry Bonds hit 73 in 2001; his record has not been threatened since baseball cracked down on steroids.

Records do not often fall so quickly. It looks suspicious. De Vany counters that this outburst is one that can be explained by chance, even if it doesn't look that way.

It all happened before, with Babe Ruth. "Ned Williamson's record... was broken three years in succession," De Vany writes, adding that "In a 9 year period the [single-season] home run record was advanced 4 times."

De Vany would agree with one of the central tenets of Nassim Nicholas Taleb's thought – just because something hasn't happened before doesn't mean it won't ever happen. Just because nobody had hit 70 home runs before Mark McGwire doesn't make it an impossible feat without the help of steroids.

There is no pre-determined length of time that a record should stand. Given individual hitters' extreme variability, each year is a roll of the dice. McGwire and Bonds' records are within the realm of chance. Chemical help is not needed to set new records, according to De Vany.

Conclusion

De Vany makes a good argument for why steroid-era records are not necessarily the product of steroids. It very well could be simple chance; the luck of the draw; great individuals doing great things. Combine his analysis with the other mitigating factors in this paper, and the role of steroids does come into question.

Even so, steroids probably do have an effect on home runs. If they didn't, players probably would not use them. But their effects have almost certainly been exaggerated. League-wide steroid-era home run numbers are not radically different from the years immediately before and after. They fit in well with a trend that is nearly a century old.

Other factors have larger effects. Players try harder to hit home runs, as evidenced by their strikeout totals. Rule changes, such as a smaller strike zone, have tended to favor offense. New stadiums are almost always smaller than old ones, which induces more home runs.

The average human being is bigger and stronger today than in 1900 because of improved living standards. Steroids do not explain all of the size increase among professional athletes. There are also more people alive today, which allows for more exceptionally talented individuals to exist.

The laws of probability allow for records to be broken. A once-in-a-lifetime talent like Babe Ruth rapidly set a succession of records. There is no reason why another great individual, like a Bonds or a McGwire, couldn't come along and do it again.

If baseball becomes a demonstrably steroid-free game once again, then fans will have a chance to see it happen.

References

Jose Canseco, interview with *60 Minutes*,

<http://www.cbsnews.com/stories/2005/08/05/60minutes/main761932.shtml>

Arthur De Vany, "Home Runs, Steroids and the Law of Genius," working paper available at <http://www.arthurdevany.com/webstuff/images/DeVanyHomeRunMS.pdf>

Peter Schmuck, "Does Size Really Matter? Today's Major leaguers Are Bigger and Stronger than those of Earlier Eras," *Baseball Digest*, July 1, 2001.

http://www.accessmylibrary.com/coms2/summary_0286-738872_ITM

Steve Silva, "Curt on Bonds," *The Boston Globe*, May 7, 2007,

http://www.boston.com/sports/baseball/redsox/extras/extra_bases/2007/05/curt_on_bonds_1.html

Nassim Nicholas Taleb, *Fooled by Randomness*, (New York: Random House, 2004)

Roger G. Tobin, "On the Potential of a Chemical Bonds: Possible Effects of Steroids on Home Run Production in Baseball," *American Journal of Physics*, vol. 76, (January 2008), 15-20.

George Vecsey, *Baseball: A History of America's Favorite Game*, (New York: Modern Library, 2006)

Data from <http://www.baseball-reference.com>.