

# Do motivated baseball players have higher batting averages? Commentary on Pope and Simonsohn (2010)

Paul T. Scott and Phil Birnbaum\*

October 24, 2010

Pope and Simonsohn (2010) document that “[Major League] baseball players with a batting average just below .300 going into their last scheduled plate appearance of the season are almost *twice* as likely to get a base hit relative to players just above a .300 average.” We argue that this disparity is produced by sampling bias: near the end of the season, many players quit immediately after getting the hit that pushes them over .300.

An article in the *New York Times* surmised “hitters at .299 or .300 batted a whopping .463 in that final at-bat, demonstrating a motivation to succeed well beyond normal.” While Pope and Simonson do not claim to find evidence of increased performance, they do not rule out the possibility. They show that hitters in their sample are frequently replaced by pinch hitters; we document that pinch running and sitting out of games are also prominent. Once all these sources of bias are removed, the apparent performance disparity disappears. Moreover, tests with unbiased sampling criteria show nothing unusual about the performance of players on the cusp of .300.

## 1 Example

Here we illustrate the source of bias with a simple example.

Suppose every player hits with a .300 average independently across at-bats (ABs) in a season of at most twelve ABs. For simplicity, there is no walking—players either make a hit or an out in each AB. After each of his first eleven ABs, a player may decide to keep going or quit early. Players usually want as many ABs as possible, but each player who hits .300 or higher in at least ten ABs will receive a big bonus.

Each player will take the first ten ABs since he can’t get the bonus otherwise. After the tenth AB, players with three hits (a .300 average) will stop rather than risk losing the bonus. No other players stand to lose the bonus in the final two ABs, so the rest will continue through 12 ABs.

If we select final ABs in which players stand to cross the .300 mark, about 23.8% of the sample would consist of players who were 2-for-9 going into their 10th AB and then stopped (these batters had a hit for sure). The other 76.2% would consist of players who were 3-for-11 going into their 12th AB (these batters hit at a .300 rate). Even though

---

\*Thanks to Laura Blue, Stephen Chaudoin, and Marshall Reaves for comments and discussion.

players hit with probability .3 in each AB, we can expect to observe a .466 batting average in the sample.<sup>1</sup> The problem is that we sample only the ABs following 2-for-9 starts in which there was a hit, omitting the outs because those hitters continue to twelve ABs.

Thus, if a plate appearance's outcome can affect whether it is a player's last, then selecting final plate appearances will lead to a biased sample.

## 2 A biased sample

Using data from Retrosheet.org for the 1975-2008 regular seasons of Major League Baseball, we attempt to replicate Pope and Simonsohn's sample by selecting plate appearances (PAs) satisfying the following criteria by year:

- The PA is the batter's last.
- The batter has at least 200 ABs.
- The batter has a batting average below .2995 going into the PA.
- A hit in the PA would make the batter's average at least .2995.
- The date is September 25 or later.

We use the cutoff .2995 rather than .3000 because recorded batting averages are rounded to three decimal places.

There are 121 PAs in our sample, including 57 hits in 116 ABs for a batting average of .491, similar to Pope and Simonsohn's observations.

We consider three ways in which a batter's season can be stopped early to preserve a .300 average:

1. A pinch runner replaces him immediately (13 occurrences in our sample),
2. A pinch hitter replaces him the next time his position in the batting order is reached (13 occurrences),
3. He sits out his team's remaining games (34 occurrences, 24 distinct from 1 and 2).

Following 50 out of 121 PAs, the batter's season was apparently cut short. Removing these observations drops the sample's batting average from .491 to .299 (20 for 67), matching the batters' earlier performance exactly.

Thus, there is strong evidence that the high batting average is a result of sampling bias alone.

---


$${}^1 P(\text{3rd hit on 10th AB}) = \binom{9}{2} (.3)^2 (.7)^7 (.3) = .08005$$

$$P(\text{3 hits after 11th AB}) = \binom{11}{3} (.3)^3 (.7)^8 = .2568$$

$$P(\text{4th hit on 12th AB}) = \binom{11}{3} (.3)^3 (.7)^8 (.3) = .07705$$

$$P(\text{last AB in sample}) = P(\text{3rd hit on 10th AB}) + P(\text{3 hits after 11th AB})$$

$$P(\text{hit in last AB} \mid \text{last AB in sample}) = \frac{P(\text{3rd hit on 10th AB}) + P(\text{4th hit on 12th AB})}{P(\text{3rd hit on 10th AB}) + P(\text{3 hits after 11th AB})} = .4663$$

### 3 An unbiased sample

To implement Pope and Simonsohn's tests with an unbiased sample, we simply choose PAs based on criteria that are known before the PA's outcome.

We sample all PAs in a team's last two games of the season where the batter was hitting less than .300, but a hit in the current PA would push him over. These highly motivated players hit a combined .299 (155 for 517), almost exactly their season average.

Then, still looking only at teams' last two games, we sample PAs where batters were already at .300 or higher, but would drop below .300 if they made an out. Those players also should have been highly motivated, but their average was .297 (43 for 145), again closely matching their earlier performance.

While hitters in our sample do not increase their batting averages, it might be argued that they did hit better than expected. A batting average of .300 is substantially above average for the league (typically around .260). Therefore, it would be expected that players in our samples would hit less than .300, regressing to the mean somewhat. They regressed only slightly.

However, even for the larger sample of 517 AB, the standard error of observed batting average is about .020. Since the mean batting average is within two standard deviations of .300, the lack of regression to the mean is not significant evidence of increased performance, for players would not be expected to regress fully to the mean.

We find no statistically significant increase in performance for players around the .300 mark, and conclude that sampling bias fully accounts for Pope and Simonsohn's findings.

### References

- [1] Pope, D., & Simonsohn, M. (2010). Round Numbers as Goals: Evidence from Baseball, SAT Takers, and the Lab. *Psychological Science*, forthcoming.
- [2] Schwarz, A. (2010, October 2). Sniffing .300, Hitters Hunker Down on Last Chances. *The New York Times*. Retrieved from <http://www.nytimes.com/>