The Baseball Hall of Fame is Not the Kiss of Death

Gary Smith
Department of Economics, Pomona College

Abstract

Abel and Kruger (2005) reported that the median life expectancy of Major League Baseball players after election to the Baseball Hall of Fame is five years shorter than that of players of the same age who are not elected to the Hall of Fame. This conclusion is surprising because there is no compelling explanation for such a dramatic reduction in life expectancy. However, the data used in that study are incorrect. A robust test applied to correct data shows that there is no statistically persuasive difference in the life expectancy of players elected to the Hall of Fame and their peers.

running head: Hall of Fame Kiss of Death

key words: baseball, longevity

Contact: Gary Smith
Department of Economics
425 North College Avenue
Claremont, CA 91711
e-mail: gsmith@pomona.edu
phone: 909-607-3135
fax: 909-621-8576

The Baseball Hall of Fame is Not the Kiss of Death

Abel and Kruger (2005) report that, "Baseball fame may have a hitherto unrecognized price" in that Major League Baseball (MLB) players who are elected to the National Baseball Hall of Fame die, on average, five years younger than age-matched MLB players who are not elected to the Hall of Fame. This is not only a hitherto unrecognized phenomenon, but one that contradicts a considerable body of research suggesting that positive self-esteem enhances physical health (Antonucci, Peggs, & Marquez, 1989; McGee & Williams, 2000; and Trzesniewsky, et al, 2006). Indeed, in another paper, Abel and Kruger (2009) argue that MLB players whose first names begin with the letter "D" die younger, on average, than do players with other first initials because of the low self-esteem associated with having a name that begins with the letter D.

Baseball players who are elected to the Hall of Fame are likely to have relatively high self-esteem that has been accumulated and reinforced by a lifetime of successes and accomplishments. If something seemingly as trivial as having D for a first initial can affect longevity negatively, then surely something as real and tangible as being the best at one's chosen profession should affect longevity positively.

In addition, it seems implausible that election to the Hall of Fame would be a life changing event. Election to the Hall of Fame is not a surprise comparable, say, to winning a lottery. Election to the Hall of Fame is based on lifetime achievement and, while there have been controversial borderline cases, election or nonelection is clear cut for most players. If election to the Hall of Fame is a formality for many players, yet reduces life expectancy by 5 years, the reduction in life expectancy for the borderline cases must be very large.

Finally, remember that we are not comparing Hall of Famers to couch potatoes, but rather

Hall of Famers to other MLB players who, by virtue of playing in the major leagues, have demonstrated that they are an elite group compared to every baseball player who never made it to the major leagues. It seems implausible that any differences in self esteem between MLB players who are elected to the Hall of Fame and those who are not is large enough to explain a five-year difference in life expectancy.

Problems With Abel and Kruger's Analysis

Abel and Kruger take their data from Sean Lahman's Baseball Archive (2010), a well-respected and widely used compilation of statistics for every baseball player who has ever played in the major leagues. Abel and Kruger restrict their analysis to Hall of Fame players who were alive at the time of their election. For controls, they matched each Hall of Fame player to Major League players who were also alive in the election year and had the same birth year as the Hall of Fame player. They then recorded the year of death for each Hall of Fame player and control. If no year of death was listed, Abel and Kruger assumed that the player is still alive.

Unfortunately, the Sean Lahman archive has incomplete data for some players, especially for relatively obscure persons who played in the early years of Major League baseball. In particular, the Lahman archive does not give a death date for more than half the players in the data base. In most cases, this is because the player is still alive, but in many cases no death date is listed because the death date is unknown, and there is no way to tell from the Lahman database which is the case. Thus, when Abel and Kruger write that 64% of the Hall of Famers were deceased compared to 47% of the controls, what they really mean is that 64% of the Hall of Famers have listed death dates, compared to 47% of the controls.

Players born in the 1800s without listed death dates are almost certainly deceased with

unknown death dates; we can only speculate about players with more recent birth dates and no listed death date. Deceased players with unknown death dates are most likely relatively obscure players from the early years of baseball. Because of their celebrity status, every Hall of Fame player is either known to be alive or has a known death date.

What is certain is that every deceased player with an unknown death date died at an age that is less than his calculated age if we assume that he is still alive. Thus, Abel and Kruger's treatment of deceased players with unknown death dates as living players skews upward the calculated longevity of players who are not in the Hall of Fame. Their conclusion that players who are not in the Hall of Fame live five years longer, on average, than do players who are in the Hall of Fame is at least partly, and perhaps entirely, explained by the incorrect assumption that all players in the Lahman data base without listed death years are still alive.

Instead of calculating the age at death for each player, Abel and Kruger calculate the number of years they survived after the election year. This distinction is unimportant if we are comparing a single Hall of Fame player with other players who were born in the same year and were alive in the election year. However, it does make a difference for the aggregate calculations made by Abel and Kruger because life expectancies at birth and election-year ages are not constant. Specifically, the number of years that a player survives after an election year depends on when he was born (because life expectancies have increased over time) and on his election-year age (because life expectancy declines with age).

Abel and Kruger analyze their data with a Cox proportional hazards model because this model can handle cases where some of the subjects survive past the conclusion of the study. However, they lump together all players, ignoring the confounding effects of birth year and age.

Thus they do not distinguish between Elmer Flick, who was born in 1876 and was 87 years old when he was elected to the Hall of Fame in 1963, and Cal Ripken, Jr., who was born in 1960 and was 47 years old when he was elected to the Hall of Fame in 2007.

Their statistical procedure assumes that each player, Hall of Famer or control, has a survival probability that does not depend on the player's birth year or age. However, survivor probabilities are not the same for a 40-year old born in 1900, a 40-year old born in 1960, and an 80-year-old born in 1900. Thus, this conclusion by Abel and Kruger (2005, p. 961) is not meaningful: "The median length of post-induction survival for the Hall of Famers was 18 years (95% CI = 15.0-21.0) versus 23 years (CI = 22.1-23.9) for matched controls." Median post-induction survival length is conditional on a player's birth year and election-year age.

The confounding effects of birth year and age can be controlled for by analyzing disaggregated data that compare each Hall of Fame player who was alive at the time of his election with other Major League baseball players who share the same birth year and were alive during his election year.

Specifically, for each Hall of Fame player i, D_i is equal to the difference between his age at death A_i and the average age at death AAD_i of his peers who shared his birth year: $D_i = A_i - AAD_i$. (With disaggregated data, this difference D_i obviously does not depend on whether we calculate the difference between the ages at death or the difference between the number of years lived past the election year.)

Morrison and Smith (2005) show that grouping decedents by birth year provides a valid test of the null hypothesis that mortality rates are the same for different groups of decedents. Specifically, if a Hall of Fame player and his peers have the same mortality rates, then the

expected value of D_i over any horizon is zero. For example, if the mortality rates are the same for a Hall of Fame player who was born in 1900 and for his peers with the same birth year, the expected value of A_i and AAD_i both increase as we expand the horizon from, say, 25 to 50 years; however, for each horizon, the expected value A_i is equal to the expected value of AAD_i and the expected value of D_i is equal to zero.

The null hypothesis is that mortality rates are not affected by being elected to the Hall of Fame, so that the expected value of each paired difference is zero: $E[D_i] = 0$. The nonparametric Wilcoxon (1945) signed-rank test for paired differences tests the null hypothesis that the median of the paired differences is equal to zero. With more than 25 observations, the probability distribution of the Wilcoxon test statistic is well approximated by the normal distribution, which can then be used to calculate two-sided p-values.

A New Analysis of Abel and Kruger's Proposition

Because there is no way to determine whether players without a listed death date are alive or deceased, I restrict my analysis to players with listed death dates. As of 2010, there were 292 persons in the Baseball Hall of Fame, of whom 164 were major league players who were alive in the year they were elected to the Hall of Fame. Of these 164 Hall of Fame players who were alive at the time of their election, 62 are still alive and 102 are deceased. Each deceased Hall of Fame player was matched with all other deceased Major League players who were born in the same year as the Hall of Fame player and, like the Hall of Fame player, were alive when he was elected to the Hall of Fame.

Four of the 102 deceased Hall of Fame players died during their election year. Two of these players died before the election results were announced and two died afterward. It is not practical

to look at every major league player who died during an election year to determine if they died before or after the election results were announced. I consequently did two separate calculations, one assuming that anyone who died during an election year was alive when the election results were announced and one assuming anyone who died during an election year died before the election results were announced.

For the calculations including players who died during an election year, there are 102 deceased Hall of Fame players and the median value of D_i, the difference between the Hall of Famer's age at death and the average age at death of his peers is 0.99 years ($Z = 0.06$, two-sided $p = 0.952$). Excluding players who died during an election year, there are 98 deceased Hall of Fame players and the median value of D_i is 1.05 years ($Z = 0.15$, two-sided $p = 0.882$).

In each case, the median value of the difference between the age at death of the Hall of Fame player and the average age at death of the matched players is approximately one year. Contrary to the conclusion reached by Abel and Kruger, Hall of Fame players, on average, lived longer than their peers. However, the observed differences are not close to being statistically persuasive. The Z values are close to zero and the two-sided p-values are close to one.

Conclusion

Abel and Kruger (2005) report that, "Baseball fame may have a hitherto unrecognized price" in that election to the Baseball Hall of Fame reduces a player's life expectancy by 5 years. There is no compelling theoretical reason for such a dramatic decline. If anything, we might expect that Hall of Fame players have high self-esteem and are physically healthier than players of the same age who are not elected to the Hall of Fame.

The puzzling conclusion reached by Abel and Kruger can be attributed to their incorrect

assumption that all players who do not have known death dates are still alive. Because all deceased players with unknown death dates are not Hall of Fame players, this assumption biases upward the calculated average longevity of players who are not in the Hall of Fame. In addition, Abel and Kruger's statistical analysis does not take into the confounding effects of birth year and age on longevity.

The Wilcoxon signed-rank test for paired differences is a very robust test that is well-suited for handling confounding effects like birth year and age that are not easily parametrized. Here, the application of this test to deceased MLB players with known death dates shows that there is no statistically persuasive difference in the life expectancy of players elected to the Baseball Hall of Fame and their peers.

References

Abel, E. & Kruger, M. (2005). The longevity of baseball hall of famers compared to other players. *Death Studies*, 29, 959-963.

Abel, E. & Kruger, M. (2009). Athletes, doctors, and lawyers with first names beginning with "D" die sooner. *Death Studies*, 34, 71-81.

Antonucci, T., Peggs, J., & Marquez, J. (1989). The relationship between self-esteem and physical health in a family practice population. *Family Practice Research Journal*, 9, 65-72

McGee, R. & Williams, S. (2000). Does low self-esteem predict health compromising behaviours among adolescents? *Journal of Adolescence*, 23, 569-582.

Morrison, S., & Smith, G. (2005) Monogrammic determinism? *Psychosomatic Medicine*, 67, 820-824.

Sean Lahman's Baseball Archive (2010), retrieved January 26, 2010, from http:// www.baseball1.com.

Trzesniewsky, K., Donnellan, M., Moffitt, T., Robins, R., Poulton, R., & Caspi, A. (2006). Low self-esteem during adolescence predicts poor health, criminal behavior, and limited economic prospects during adulthood. *Developmental Psychology*, 42, 381-390.

Wilcoxon F. (1945). Individual comparisons by ranking methods. *Biometrics*, 1, 80-83.