EJW
ECON JOURNAL WATCH
Scholarly Comments on
Academic Economics

ECON JOURNAL WATCH 10(1)
January 2013: 70-86

The Peer Effect of Jose Canseco: A Reply to J. C. Bradbury

Eric D. Gould[1] and Todd R. Kaplan[2]

LINK TO ABSTRACT

Our paper in *Labour Economics* (Gould and Kaplan 2011) examined the general issue of how workers affect the productivity of co-workers, and in particular whether unethical practices that boost performance are transmitted between workers. To do this, we investigated the steroid epidemic in Major League Baseball and Jose Canseco's claims that he taught his teammates how to acquire and use steroids and human growth hormone. Our results support Canseco's claim by presenting a striking pattern whereby a player's performance indeed tends to increase after being a teammate of Canseco. Our results are based upon standard measures of performance, and we found very little systematic evidence across outcome measures that any comparable player increased the performance of his peers in a similar way.

J. C. Bradbury (2013) has authored an extensive critique of our paper; the present paper is our reply. Bradbury makes no claims that our results are non-replicable or that they involve a programming error. Rather, he presents a series of criticisms about our choices regarding specification, sample, and other matters. His critique does not provide any convincing reason to reject our conclusions.

We disagree with all of Bradbury's major criticisms—which can largely be summarized as his belief that we should have imposed several very restrictive assumptions on our specification and estimation, and that we used too much data and did not adequately censor our sample on an outcome variable. It is highly unusual to criticize a paper for not deleting years of data and for not censoring the sample on an outcome measure. Moreover, Bradbury mounts criticisms that reveal

1. Hebrew University, Jerusalem 91905, Israel.
2. University of Haifa, Haifa 31905, Israel, and University of Exeter, Exeter EX4 4ST, UK.

a severe misunderstanding of basic econometrics. We will detail our response to Bradbury's numerous comments below, but it is important to note that Bradbury does not present any convincing argument for how the putative flaws in our analysis would lead to a spurious result. Lacking any explanation for why our results might be biased in support of the hypothesis that Canseco boosted peer performance, Bradbury argues that "even if the estimates are taken at face value, they do not support the conclusion that Canseco left a visible trail of steroid users in his wake" (2013, 42).

That statement is perhaps Bradbury's most serious one, since his other statements are not directly about what we did, but rather what he thinks we should have done differently. Bradbury's argument for why we have a "distorted interpretation of [our] own regression estimates" (2013, 41) is based on his summary of our findings in his Table 1. That table shows that we never found Canseco to have a significant positive effect on a player's performance while they were on the same team, and the "after-Canseco coefficient is positive and significant in 12 of the 27 specifications (44 percent)" (Bradbury 2013, 45). Twelve out of 27 is presumably too few.

Before we address the first point about the non-significant results for being "with Canseco," we want to point out the dishonesty in discussing the 12 significant "after-Canseco" coefficients as if all 27 specifications should be treated equally. Five of the 27 "after-Canseco" coefficients presented by Bradbury as not significant are testing for whether Canseco affected other players in terms of their steals, fielding percentage, and fielding errors. Here is what we wrote about these specific coefficients:

> The first three columns show that Canseco had no discernible effect on steals, fielding percentage, and fielding errors. Neither of these outcomes is considered particularly important for power hitting, nor are they typically thought of as being affected by physical strength. So, the lack of any effect for these outcomes strengthens the interpretation of the results in Table 4 that Canseco had a significantly positive effect on the hitting power of his former teammates by affecting their physical strength. (Gould and Kaplan 2011, 342-343)

One could argue with our interpretation that these results are essentially placebo tests which *should* be insignificant, but a fair critique of our paper would at least attempt to do so. Instead, Bradbury presents these insignificant results as if they obviously support the idea that Canseco had no effect on his peers, while accusing us of distorting our findings.

Regarding the other estimates for the "after Canseco" effect, we invite all readers to look at our paper and see that these coefficients are highly significant for the type of outcome that is relevant for each type of player. For power hitters, the "after Canseco" effect is positive and significant for home runs, RBIs, strikeouts, bases on balls, and at-bats, while the coefficients for batting average, slugging percentage, and intentional walks are positive but insignificant (2011, 343-344). For position players, the "after Canseco" effect is positive and significant for batting average and on-base percentage (343 n. 17). Even for pitchers, there is a positive and significant "after Canseco" effect on innings pitched (344).[3] Assuming that managers tend to play players more when they are playing better, at-bats and innings pitched are perhaps the best overall measures of a player's performance level. Given all these findings, we stand by our interpretation that Canseco had a significant, positive impact on his teammates after playing with him.

Let us now address the insignificant findings for the "with Canseco" coefficients. Bradbury (2013, 46) writes: "The fact that the 'with Canseco' indicator is never positive and significant in 27 estimates reported in the Gould and Kaplan paper is strong evidence against the Canseco effect." Bradbury then accuses us of downplaying the lack of any findings for this coefficient by our not reporting it for many specifications. But, we stated very clearly that this coefficient is not significant, and we discussed why this might be the case:

> However, Table 4 again reveals no significant impact of playing with Canseco at the same time. The reason why playing with Canseco has a much smaller effect than playing "after Canseco" may be due to the idea, mentioned above, that players who learn about steroids from Canseco do not take steroids during the whole time they are playing "with Canseco," but do use them during the entire time that they are former teammates with him. Alternatively, it may take some time for Canseco's positive effect to be realized, or this pattern may be due to the fact that players who play with him spend more of their time as former teammates of Canseco than being current teammates of him. For example, power hitters who played at least one season with Canseco in our sample spent 15% of their seasons on a team with Canseco and 39% of their seasons being former teammates with him. Also, the smaller effect of playing with Canseco may be due to the idea that Canseco took away scarce team resources such as playing time, attention from coaches and trainers, etc. (Gould and Kaplan 2011, 342)

3. In unreported results, the coefficient on "after Canseco" for at-bats for all non-pitchers in one sample is 16.24 with a standard error of 9.36.

In footnote 3 (2011, 339), we discussed why it might take time for a player to learn to use steroids effectively. According to Canseco, its effectiveness depends on the proper mixing and cycling of growth hormone with various types of steroids, as well as a proper diet, a rigorous weight lifting routine, and abstinence from recreational drugs.

Given these explanations, it seems reasonable to interpret the significant "after Canseco" effect as strong evidence of a peer effect despite the insignificant "with Canseco" effect. But, once again, Bradbury does not make any attempt to counter the logic of these arguments. Rather, he implicitly denies one of them by writing that the "lack of performance improvement while playing with Canseco, especially because many players in this cohort played with Canseco for several years—some of whom admitted using steroids during this time—contradicts the steroid spillover hypothesis" (Bradbury 2013, 46). As we describe in the quotation from our paper above, most people did not play with Canseco for several years. Over sixty percent of the people in our sample in Table 3 played only one year with Canseco, and another 25 percent played only two years with him. So, if one were really open to the idea that Canseco did affect the performance of his teammates, should the hypothesis really be dismissed because the effect shows up a few years after their initial exposure to Canseco rather than having an immediate impact? Canseco never claimed to teach players how to use steroids on the first day that they became teammates.

Moreover, we mentioned what happens if no distinction is made in the specification between "with" and "after Canseco". We wrote: "If, however, we do not differentiate between current and former players by using one variable which indicates whether the player either plays currently or in the past with Canseco, the coefficient for home runs is 1.40, and is still highly significant with a standard error of 0.52" (Gould and Kaplan 2011, 342). Bradbury dismisses this specific point by saying (correctly) that: "As the other estimates show, however, this is being driven by after-effects" (Bradbury 2013, 45). But, wait a second—Bradbury claims that the "estimates presented in this paper do not provide evidence of a post-Canseco positive spillover" (46). So, how could the joint "with and after" effect be driven by an "after" effect that doesn't exist?

To further his claim that we over-interpret our results, Bradbury argues that the coefficients for Canseco's peer effect are not very different, and perhaps weaker, than the estimated peer effects from other power hitters that we examine:

> The slugging average is never statistically significant in any estimate for Canseco; however, it is positive and statistically significant in estimates of teammate spillovers for five other comparable hitters reported in Gould and Kaplan (2011): Rafael Palmeiro (Table 6), Ryne Sandberg (Table 6), Matt Williams (Table 9), Chili Davis (Table 9), and Dante Bichette (Table 9). Based on slugging average, several other players are more strongly associated with teammate improvement spillovers than Jose Canseco. (Bradbury 2013, 48)

For the record, our Table 6 presents results for the peer effect of a few selected players on skilled position players only, while Table 9 summarizes the results for a larger set of 27 players (that were similar to Canseco in terms of career achievements) on power hitters and position players, separately. All of the positive coefficients mentioned in the quote above appear in Table 9 (including those attributed to Table 6). Therefore, we invite all readers to look at Table 9 (Gould and Kaplan 2011, 347) and decide whether the results support our conclusions. Specifically, Canseco is the only player in Table 9 to exhibit a positive peer effect for all seven outcome measures, with four of them significant. As we write: "With the possible exception of Williams, there is no other player that has a systematically large and significant positive effect across several outcomes" (345).

Bradbury specifically mentions positive coefficients on slugging percentage by Palmeiro, Sandberg, Williams, Davis, and Bichette. It is true, as we mentioned in the paper, that Williams had three positive and significant coefficients. However, it is entirely misleading to suggest that the others display any evidence in favor of a positive peer effect. Palmeiro has a positive and significant coefficient for slugging percentage on position players, but also a negative coefficient on slugging percentage for power hitters. In total, Palmeiro shows five coefficients that are negative and significant, and two that are positive and significant. Sandberg displays only one positive and significant coefficient out of seven. Not surprisingly, Bradbury mentions only the one, not the other six. Davis had only two positive and significant coefficients out of seven, while Bichette had only one—but he had two that were negative and significant.

Given all this, we believe that a fair reading of Table 9 is that the results for Jose Canseco look unusually positive and significant, relative to the other 26 players of a similar caliber during the same era. In sum, there is no reason to take seriously Bradbury's claim that our interpretation of the results is distorted.

We now turn to Bradbury's comments about what we should have done differently. His main criticism concerns the specification of Canseco's treatment effect on his teammates. As discussed above, we tested for Canseco's influence on the performance of other players by separating a player's exposure to Canseco into two distinct periods—the period spent with Canseco on the same team and the period after the last year spent with Canseco (Gould and Kaplan 2011, 341):

<u>Our specification:</u>
$Performance_{it} = \gamma_0 + \gamma_1(with\ Canseco_{it}) + \gamma_2(after\ last\ year\ with\ Canseco_{it}) + u_{it}$

Bradbury calls this a "flawed empirical approach" (2013, 41), and he "corrects a major defect" with the following (55):

<u>Bradbury's specification:</u>
$Performance_{it} = \beta_0 + \beta_1(with\ Canseco_{it}) + \beta_2(after\ first\ year\ with\ Canseco_{it}) + u_{it}$

where "with Canseco" is similar to the variable we use for the time a player spends with Canseco, and "after first year with Canseco" equals one for every season after the first season playing with Canseco. It is hard to imagine that our specification is somehow stacking the deck in favor of finding large after-Canseco effects on his teammates, and Bradbury does not even try to suggest that it does. Bradbury is adamant, however, that his specification is right and ours is wrong.

To understand the differences between the two specifications, let's examine the parameterization of the Canseco effect for a player that joins Canseco's team in his second season and then plays a total of four consecutive years with Canseco:

	Canseco Effect	
Season	Bradbury	Gould and Kaplan
1 - without Canseco	0	0
2 - with Canseco	β_1	γ_1
3 - with Canseco	$\beta_1+\beta_2$	γ_1
4 - with Canseco	$\beta_1+\beta_2$	γ_1
5 - with Canseco	$\beta_1+\beta_2$	γ_1
6 - without Canseco	β_2	γ_2
7 - without Canseco	β_2	γ_2
8 - without Canseco	β_2	γ_2

As depicted above, Bradbury's specification is describing three distinct periods, but he breaks from convention by using dummy variables for categories that are not mutually exclusive (i.e., the player described above has a 1 in seasons 3 to 5 for both dummy variables: "with Canseco" and "after first year with Canseco"). Thus, he makes the unnecessary and rather strong assumption that the effect of the first year with Canseco plus the effect of "after last year with Canseco" is equal to the effect of playing with Canseco in all years except for the first year. Or, put differently, Bradbury assumes that β_1 is the change in performance between the season before playing with Canseco to the first year with Canseco, while the player's performance changes from the last year with Canseco to the first year without him by the same exact magnitude, but opposite direction (i.e., $-\beta_1$).

We don't see why anyone would place such strong, and rather strange, restrictions on the parameters. Bradbury does not even acknowledge that the restrictions exist, let alone justify them. In fact, if β_1 is negative, Bradbury's specification implies that there will be an increase in performance after leaving Canseco's team in season 6. According to Bradbury's Table 3 (2013, 58), almost all of his estimates are negative, which is consistent with our results showing a significant boost in performance when the player is no longer on the same team as Canseco.

In contrast to Bradbury, our specification divides a player's career into two distinct periods: *during* and *after* the player's exposure to Canseco as a teammate. We follow the convention of using dummy variables for categories that are mutually exclusive, so no restrictions are made regarding how the effect may or may not change between periods. In addition to the discussion described above about why the "with Canseco" effect might differ from the "after Canseco" effect, we justified the distinction between the two periods by writing: "the distinction between playing 'with Canseco' and playing 'after Canseco' is important since even if a player did learn about steroids from Canseco, we do not know when he learned about it during his time with Canseco, but we can be sure that he already acquired the knowledge after playing with Canseco" (Gould and Kaplan 2011, 341-342). We never claimed that this is the only way to examine this issue, but nothing Bradbury writes creates any doubts in our minds that ours is a sensible and conventional specification.

Let us now examine the two specifications for a player that plays only one year with Canseco before moving on:

	Canseco Effect	
Season	Bradbury	Gould and Kaplan
1 - with Canseco	β_1	γ_1
2 - without Canseco	β_2	γ_2
3 - without Canseco	β_2	γ_2
4 - without Canseco	β_2	γ_2
5 - without Canseco	β_2	γ_2

For this type of player, the two specifications are identical—they both separate the player's career into two distinct periods with independent effects for each period. Since approximately 60 percent of the players who ever played with Canseco actually played only one season with him, Bradbury is unwittingly proposing the same specification for a majority of the players. However, as we've shown, Bradbury's specification makes additional, unnecessary restrictions on how Canseco may have affected players that played more than one year with him. One can only guess why Bradbury believes so fervently in these restrictions.

Bradbury also takes a strong stand against our treatment of players who play with Canseco in multiple stints (Gould and Kaplan 2011, 341 n. 11), calling our strategy "convoluted":

> The "after-Canseco" interpretation is even more convoluted in the authors' method for players who play on the same team as Canseco but on non-consecutive occasions, involving their playing together on different teams. Players are coded as "with Canseco" for the first year playing with Canseco for every year all the way through to the last year the player is on a team with Canseco. That is, it is not until the first season after the final year of playing with Canseco that the player is coded as "after Canseco." (Bradbury 2013, 44)

Let's compare the parameterizations of the Canseco effect under both specifications for a player who plays with Canseco in non-consecutive seasons (and in a pattern that happens to coincide with Roger Clemens' career during this time period):

	Canseco Effect	
Season	Bradbury	Gould and Kaplan
1994 - without Canseco	0	0
1995 - with Canseco	β_1	γ_1
1996 - with Canseco	$\beta_1+\beta_2$	γ_1
1997 - without Canseco	β_2	γ_1
1998 - with Canseco	$\beta_1+\beta_2$	γ_1
1999 - without Canseco	β_2	γ_2
2000 - without Canseco	β_2	γ_2

Starting from 1994 as a base year, Bradbury specifies that the subsequent changes in Clemens' performance in the following years go as follows: first there is a change equal to β_1 in his first year with Canseco, then a change of β_2 in the next year, followed by a change of exactly β_1 but in the opposite direction (viz., $-\beta_1$), then going back in the other direction by β_1, and then reversing course with another change of exactly $-\beta_1$. Bradbury offers no justification for why a player's performance would jump around in such a stringent manner.

But, he calls our strategy "convoluted." In our specification, γ_1 captures the effect of the period when the player was directly exposed to Canseco and potentially learned how to use steroids from him, while the "after exposure" effect is represented by the parameter γ_2, which is not restricted to be the same or different from γ_1. We will not argue that this is the only or even the best possible way to handle this non-trivial issue, but Bradbury fails to convince us that the "convoluted" approach is ours.

Bradbury continues with a long series of comments that amount to throwing everything but the kitchen sink at us. At best, some could be considered worthwhile robustness checks. We'll proceed largely in the order that Bradbury presents them.

Dependent variables

Our analysis examines many different measures of a player's performance to see if the results are consistent across different measures. Some of them are normalized by playing time (measured as at-bats) like batting average or slugging percentage, and some of them are raw output measures like the number of home runs or number of RBIs. For home runs, we also show what happens if the number of at-bats is included as a control variable. Bradbury insists that we should only use the measures which are normalized by at-bats. He writes: "One problem with using total home runs per season is that the data is positively skewed to an extreme degree, which raises the possibility of strong influence by outliers. This potential bias is never addressed in the paper" (Bradbury 2013, 47).

The reason why this "potential bias" is not addressed in the paper is because we have no idea what bias he is talking about. In any case, what is the evidence that home runs is skewed after controlling for experience and fixed effects for each individual and year? Furthermore, the outliers that he is referring to are probably the records that were smashed because of the use of steroids (by Barry Bonds, Mark McGwire, etc).

Bradbury then makes a serious point that "output is affected by performance opportunities as well as performance level" but misses the main point by saying that "changes in total output over time may be the product of managers choosing to play a player more in the future due to aging, health, or other reasons, outside of any influence steroids might have" (2013, 47). That statement misses the main point because we need to be concerned that steroids affected both the output and the opportunities. But, Bradbury recovers a bit by acknowledging that "Gould and Kaplan briefly address this potential bias by including at-bats as an independent variable in one specification", but that "using raw home run totals without controlling for opportunities is not the best solution. A home run rate normalized for opportunities is a far better dependent variable choice for measuring changes in performance over time, as it avoids the endogeneity issue and does not ignore an obviously relevant factor" (47).

Unfortunately, Bradbury mixes a bunch of stuff together (some of it true) and ends with a recommendation that, contrary to his claim, does not avoid the endogeneity issue. If steroids make players better, it is possible that Canseco increased both his home run production and his playing time, and therefore, both the numerator and denominator in a measure such as "home runs per at-bat." In fact, our results demonstrate just this, showing as well that the effect on home runs is still significant for power hitters after controlling for at-bats. Contrary to what Bradbury thinks, there is no clean solution to this issue, but our approach of looking at it in several directions strikes us as the most reasonable way to proceed.

Bradbury then criticizes our choice of output measures from the perspective of someone who is ensconced in the specialized world of discussing, creating, and analyzing every minor detail and statistic in baseball. Some of the alternative measures are worth checking, but the purpose of our study was not to take a stand on which precise measure is the best measure of power hitting and so forth. We used conventional measures of performance in order to speak to a more general audience, and to avoid any appearance that we were cherry picking the best results from all the various alternative measures of power hitting that baseball experts have proposed.

Furthermore, as we discussed in the paper, Canseco claimed that steroids could affect a player's performance in ways that are hard to measure, such as faster recovery from injuries (Gould and Kaplan 2011, 343 n. 16). Also, greater prowess for power hitting could show up in a non-power hitting statistic like batting average if, say, the defense is forced to play deeper in the outfield. In addition, opposing teams are likely to "pitch around" power hitters, pitch to them in a fashion intended to give up a single instead of an extra base hit, or even bring in a new pitcher to specifically handle a dangerous batter. In these instances, the total number of home runs will understate the power-hitting prowess of a player. Given the difficulty in identifying precisely what steroids should be affecting, it seems prudent to check the most conventional measures, as we did in our analysis.

Control variables

Bradbury believes "the control variables employed by Gould and Kaplan (2011) are also problematic, and thus call into question the specification choices and the significance of the estimates they produce" (Bradbury 2013, 48). He does not "believe the included variables are necessarily contributing to significant bias in the regression estimates" (48), but nonetheless, that does not stop him from complaining about them. He specifically calls using the slugging percentage of the division "an odd choice" and complains that "[c]areer wins by a manager provide little information regarding managerial ability" (49). We have no idea why he feels this way. The coefficient on our managerial ability variable has a t-statistic of 2.89 for home runs in Table 3.

Bradbury argues that we should have used "age" instead of "tenure", since he believes that tenure is correlated with ability ("Superior baseball players tend to enter the league at younger ages than inferior players" (49)). We don't see his point, since better players will also stay in the league longer and at older ages. Also, we controlled for ability with individual fixed effects. But, in principle, we don't have a problem of using age instead of tenure, and after making this change, we can report that the results are similar.

Bradbury then embarks on an ill-fated attempt to criticize our use of league-year fixed effects:

> League-year fixed effects offer some controlling influence for the change in the offensive environment of each league, but they also reflect other factors not related to changing offensive environments altering player performance. For example, in two seasons during the timeframe of the study, a significant number of games were lost to a labor strike. In 1994, teams played an average of 48 fewer games and in 1995 teams played an average of 18 fewer games. When count totals are used as dependent variables, the lower numbers were a product of playing fewer games with an uncertain effect as to how the propensity for home runs and runs scoring were affected. In rate terms, home runs and runs per game were at historically high levels; however, the raw totals of players will indicate that home run performance declined. With totals as a dependent variable, an explicit control for the offensive performance in each season is needed to be sure that the deviation in performance is offensive-environment related rather than the result of an unrelated exogenous shock like the 1994-1995 strike. Therefore, it is important to control for changes in league performance explicitly, to ensure that such problems are not biasing the main coefficients of interest. (Bradbury 2013, 49-50)

No, this is way off. Using league-year fixed effects precisely controls for anything and everything that may have affected the mean in any year and league. The inclusion of these fixed effects means that all performance outcomes are measured relative to the mean in that given year and league. This is basic Econometrics 101.

Sample years

Bradbury continues in his confusion about what league-year fixed effects control for with a comment about our choice of sample years. Our sample includes years 1970 to 2009, despite Canseco's career spanning the years 1985 to 2001. Bradbury calls this "problematic" because performance measures are improving over time. He writes:

> [N]o justification is offered for extending the sample back this far into the past, and doing so only risks weighting the Canseco coefficients positively by including multiple observations from players many years before Canseco entered the league that adds numerous observations of "0" for the with- and after-effect variables from a low-offense era. (Bradbury 2013, 51)

Once again, Bradbury does not understand that league-year fixed effects control for changes in average performance over time. We did not justify going back to 1970 because we felt it was obvious that more years of data is better than less when one has controls for each year. Furthermore, not every player's career coincided with Canseco's, and it is important to include years before players were exposed to Canseco in order to estimate the "with" and "after" effects while controlling for individual fixed effects.

Bradbury displays a similar lack of understanding when he writes:

> Offense tended to be rising for most of Canseco's career; therefore, it is natural to expect anyone who played with Canseco to see his offense improve after leaving Canseco. The authors argue—possibly in anticipation of a critique like the present one—that if such effects are driving the results, then there would be similar effects observed from a sample of ten comparable power hitters, and they describe the estimates for the comparison cohort to be "strikingly different." (2013, 51)

No, we would not respond to this comment by comparing Canseco to other players. Our response once again is that we included league-year fixed effects to control for changes in average performance measures over time. Our examination of other players is unrelated to the issue of increasing average outcomes over time—rather, the purpose is to examine whether Canseco's effects are spurious by seeing if they are typical or unusual compared to his peers.

Regarding sample years, Bradbury also takes issue with our analysis that shows that the "after Canseco" effect disappears after 2003, which coincided with the advent of new policies conducted by Major League Baseball to test players for performance-enhancing drugs. We justified our choice of using an "after 2003" variable by citing the policy change (the institution of testing with penalties), and we noted that the results of the testing revealed a significant drop in steroid use from 2003 to 2004 (Gould and Kaplan 2011, 345 and 345 n. 18). But Bradbury doesn't buy this at all:

> I believe 2003 does not represent the pivotal season for a change in drug use. As a part of the 2002 collective bargaining agreement, the 2003 drug tests were implemented for survey purposes only—the tests were anonymous and there was no sanction for testing positive. In 2004, a player would have to fail at least two drug tests before being subject to suspensions or fines. Testing in 2003 and 2004 did not result in any suspensions, fines, or public announcements of positive tests for any major-league player: hardly a sign that Major League Baseball was identifying and punishing steroid users in a manner that would deter use. It was not until 2005 that Major League Baseball instituted a revised drug testing program that included a ten-day suspension without pay for a first positive test, and several major-league players received suspensions for positive tests, including former Canseco teammate Rafael Palmeiro. (Bradbury 2013, 56-57)

So, Bradbury is arguing that the measured drop in steroid use in 2004 is irrelevant because suspensions weren't given until 2005. It is hard for us to understand why a decrease in steroid use is not relevant for seeing whether Canseco affected his teammates by teaching them about steroids. But, in any case, Bradbury is saying that we included a "steroid-using" year, 2004, in our definition of the "post-steroid" era. If true, this should bias the results against finding a difference in the "after Canseco" effects between the two periods. Instead, we find a huge difference—the "after Canseco" effect completely disappears after 2003. Instead of acknowledging or discussing the powerful evidence that this finding represents in terms of bolstering our conclusions, Bradbury offers a tiresome tirade for why our results and conclusions would actually be stronger if we did it his way.

Cutoffs for inclusion in individual years

In our analysis, we restrict the sample to include seasons where a player had at least 50 turns at bat. This is a fairly minimal restriction, motivated by the intention to exclude years when a player was injured. Bradbury has major problems with this cutoff: "The cutoff choices of 50 plate appearances for hitters and ten games for pitchers do not make much sense" (2013, 51). He later argues that restricting the sample to those with at least 200 plate appearances is preferable. But, as we show in the paper, Canseco had a direct effect on at-bats. Therefore, as noted above and in the paper, the number of times at bat is a legitimate measure of performance. Again, it is very unusual to see a paper criticized for not censoring the sample enough on an outcome variable.

Arbitrary partitioning of the sample

Bradbury does not like that we separate hitters into "power hitters" and "position players." For reasons that are not clear, he suggests that this makes it hard to interpret our results. For those that do not want a separation, Table 3 in our paper shows similar results for a sample composed of all hitters—significant effects on home runs, strikeouts, RBIs, and batting average. Table 3 also tests for whether the coefficients are the same for both types of players, and this hypothesis is rejected for almost all outcomes. For someone who claims that his "first objection to this partitioning is that it is not necessary" (2013, 52), it is puzzling that Bradbury did not acknowledge or address this table in his critique.

Re-estimating Jose Canseco's effect on his peers

Bradbury estimates the effect of Canseco on his peers after making many changes to the specification, control variables, and sample. Not surprisingly, he claims to find results that do not support our conclusions.

For some reason, however, he deletes from the sample any player that switched teams during the year. For example, his sample does not include Mark McGwire in 1997 when he hit 58 home runs for Oakland and St. Louis, or Fred McGriff in 1993 and 2001 when he hit 37 and 31 home runs respectively. In fact, we count 1,397 observations missing in Bradbury's sample that are included in our sample in Table 3. It appears that the main reason for the missing observations is that the player switched teams during the year. Bradbury drops them because "it is unclear how spillover effects may impact performance when the season is split on multiple teams" (2013, 57). We don't understand this point, especially since most of these cases involve seasons where players did not play with Canseco on either team during the same season. In addition, deleting observations where players switch teams creates an odd situation whereby 44 players (see the Appendix to this article) in his data are coded with a 1 for the "after the first year with Canseco", but never have a year when they are coded with a 1 for "with Canseco". So, according to Bradbury's raw data, many players became a former teammate of Canseco without ever being a teammate of Canseco. Logic would seem to dictate that knowing whether the two played together would be a necessary condition for determining that they no longer play on the same team.

There are several other reasons for Bradbury's results being different from ours. As outlined in the estimating equations above, he estimates different parameters than we do, and provides no justification for the unnecessary parameter restrictions imposed on the way Canseco potentially affects his teammates as they play together and then move on to separate teams. He doesn't seem to realize that his negative and significant estimates for β_1 actually are consistent with our findings of a significant boost in performance when a player is no longer on the same team as Canseco after playing multiple seasons with him. But, interpreting the meaning of that coefficient is difficult since it represents the change in performance after playing with Canseco and the effect (in the other direction) of being together with Canseco for the first time. This kind of confusion is what happens when you break from convention and use dummy variables for overlapping categories.

Another important difference is that Bradbury imposes an AR(1) process of the error term on the estimation of the parameters (using the "xtregar" command in STATA). Bradbury is correct that we did not address the issue of serial correlation, but our results using OLS are still consistent under any general form of serial correlation. Bradbury's method assumes, without any justification, a very specific form of serial correlation—an AR(1) process—and imposes the AR(1) specification on the estimation procedure to obtain his estimates. His estimates are problematic if his specific form of serial correlation is incorrect (see Wooldridge 2002, 278). So, if the estimates are different under the two estimation procedures, this is a sign that the AR(1) assumption is incorrect. Therefore, a more conservative strategy would be to impose no assumptions on the form of serial correlation.

Overall, Bradbury's point seems to be that it is possible to get different results from ours if you estimate different parameters which imply strong and non-intuitive restrictions on the way Canseco affects his teammates, impose strong assumptions about an AR(1) process of the error term on the estimation, drop a lot of observations for no good reason, and make other assorted changes to the sample and control variables. We cannot argue with that. The question is whether any of these changes is worth making, and we found Bradbury's attack on every single aspect of our analysis to be thoroughly unconvincing and unserious.

Given the stridency in Bradbury's critique, however, we suspect that Bradbury's main problem is not really with our analysis, but with our conclusions. Bradbury has a very public history of denying the efficacy of steroids and human growth hormone. He has claimed repeatedly that steroids had nothing to do with the longstanding home run records being smashed by Mark McGwire and Barry Bonds in 1998 and 2001 at the ages of 35 and 36 respectively. In his 2008 book *The Baseball Economist*, Bradbury stated unequivocally that league expansion, not steroid use, explains why power hitting has increased over the years: "In fact, their great achievements in home run hitting are exactly what we would expect given the current distribution of talent in the league. The statistics don't convict them" (Bradbury 2008, 94). In a *New York Times* article in 2007, Bradbury wrote: "the blame shouldn't be placed on pills, needles and balms. The true culprit is expansion" (Bradbury 2007).

Without having done any econometric analysis of the effects of steroids versus league expansion, Bradbury denies that steroids had any effect on home runs in baseball. This is a rather extreme position—one that is clearly at odds with our results, which came out a few years later. But, Bradbury went even further in his book by arguing that the whole topic should be avoided: "In fact, we know very little about the potential impacts of anabolic steroid use on performance, but that has done little to curtail the discussions about threats to the 'integrity' of the game" (Bradbury 2008, 108). Similarly he writes: "if you want to ruin any discussion of baseball, just bring up steroids" (2008, 108). Apparently, Bradbury doesn't even like that we raised the subject.

Appendix

Player IDs in Bradbury's data set that are listed as playing "after Canseco" but never "with Canseco":

aguilri01, aldresc01, baldwja01, bernato01, berroge01, bitkejo01, braggda01, brandma01, brewebi01, chrismc01, donnech01, downske01, embreal01, guthrma01, hillesh01, hollida01, jacobbr01, johnsjo07, johnsru01, johnto01, lewisri01, ludwier01, mahompa01, mantoje01, mathetj01, mitchke01, moyerja01, nenro01, penaal01, pennibr01, pulsibi01, randowi01, rodrifr02, rodrine01, russeje01, sierrru01, spraged02, stantmi02, sturtta01, suttodo01, tatumji01, trachst01, tynerja01, whitema01.

References

Bradbury, John Charles. 2007. What Really Ruined Baseball. *New York Times*, April 2. **Link**

Bradbury, John Charles. 2008. *The Baseball Economist: The Real Game Exposed*. New York: Plume.

Bradbury, John Charles. 2013. Did Jose Canseco Really Improve the Performance of His Teammates by Spreading Steroids? A Critique of Gould and Kaplan. *Econ Journal Watch* 10(1): 40-69. **Link**

Gould, Eric D., and Todd R. Kaplan. 2011. Learning Unethical Practices from a Co-Worker: The Peer Effect of Jose Canseco. *Labour Economics* 18(3): 338-348.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass.: MIT Press.

About the Authors

Eric D. Gould is a Professor of Economics at Hebrew University. He received his Ph.D. from the University of Chicago. His research examines various empirical issues, such as income inequality, marriage market behavior, the economics of crime and terrorism, education outcomes, and immigration. His email address is mseric@mscc.huji.ac.il.

Todd R. Kaplan is a Senior Lecturer in the Department of Economics at the University of Haifa and a Professor of Economics at the University of Exeter Business School. He earned his B.Sc. from Caltech and his Ph.D. from the University of Minnesota. His research interests include auctions, industrial organization, and experimental economics, with eclectic paper topics that include explaining gift-giving, modelling research contests, and testing understanding of uncertainty in weather forecasts. His email address is dr@toddkaplan.com.

Go to Archive of Comments section
Go to January 2013 issue