CADINHA&CO
INSTITUTIONAL SERVICES

Umpiring?
Victor A. Canto

Umpire selection, Type I and Type II errors

Most umpires go through an apprenticeship in the minor leagues, and only a few of those make it to the big show. At each stage of the process, the selection or appointment committee must make an assessment as to the quality of umpiring that each of the potential candidates may make at the next level. Given a lack of foresight, the committee must decide with an incomplete information set. Invariable errors will be made in the selection process. The committee must rely on various indicators such as past performance to make their decision, particularly the statistics needed to reach a conclusion regarding an umpire's prospective performance. Another being the impact on the quality of umpiring that the two type of errors generated by a decision-making process under uncertainty.

- One error is to falsely reject the true null hypothesis that an umpire will be a good umpire. The Type I error would falsely conclude that an umpire will not be a good umpire when in fact he or she will be a good one. The Type I error, commonly known as a "false positive", rejects many qualified candidates.

- At the other end of the spectrum we have the Type II error, which leads to what is known as a "false negative". The Type II error would falsely conclude that some candidates will make good umpires when in fact they will not. The Type II error fails to reject some unqualified candidates and thus lowers the quality of the umpires in the big show.

The net effect of the Type II of errors is to increase the pool of potential umpires while simultaneously lowering the overall quality of the umpiring crew. If our objective is to improve the quality of umpiring in the big show, the Type I error would be preferred to the Type II error. Efforts and resources should be directed at minimizing Type II errors.

The perils of making outcome-based inferences

Consider the possibility of someone using the following examples as evidence of umpire bias. A game when Nolan Ryan pitched a no hitter, struck out a double-digit number of batters and his team won the game, or the fact that the Red Sox won over 100 games beat two 100 game winners in the playoffs and the won the world series. Obviously, these outcomes are consistent with bias, but that is not the only explanation or possibility. As we say in the business correlation does not necessarily imply causality.

A baseball aficionado would argue that Nolan Ryan was a future Hall of Famer whose fastball was unhittable when he was in the Zone. Regarding the Red Sox, the aficionado would argue that Sox are an amazing team with incredible talent at every position, that they just happen to be the best team in baseball. A baseball *cognoscenti* would conclude that the outcome of these two examples was talent based not the result of umpire bias. In contrast a non-fan who does not have the knowledge nor the access to the sabermetric information, may be tempted to make an outcome-based decision and conclude that there is umpire bias when in fact there is none.

The outcome-based conclusion fails to reject the false null hypothesis, i.e. it produces a false negative or Type II error. Form a dynamic perspective, this Type II error will have a devastating effect on the integrity of the game. If umpires know that they are being judged based on outcomes, they will have every incentive to ensure that teams play 500 ball - that player statistics do not deviates too much from the average league statistics. That way they will never be accused of bias and their jobs will be safe. The umpires' behavior and rulings thwart the players incentives and efforts to excel, they will not try or do their best. The average quality of the game, the players skill and efforts will unambiguously decline. Under this scenario, over the games' outcome will be no different than a random one. That is the danger of making output-based inferences.

Information based decision-making

Baseball keeps statistics on everything. Things such as launch angle, exit velocity and umpires' calls. The information collected on the umpires could be quite useful evaluating their performance, identifying and promoting the better umpires. The baseball statistics may help us reduce the likelihood of Type I and Type II errors: Identify the good as well as the biased umpires. Reduce likelihood of firing a good umpire or promoting subpar umpires.

 The evaluation process may begin by examining the umpire strike zones to determine whether there is a systematic bias or not. That is and ask whether the strike zones are the same for both teams? If they are, this reduces the likelihood of systematic bias calling of balls and strikes. With the advent of instant replay, one can gather additional information such as the frequency of challenges to the decision made by the different umpires. How often they are these decisions overturned? Is there a systematic pattern in the overturned calls? Collectively the information helps identify which umpires adhere to the rule book, whether there are systematic biases in their decisions and how close are the umpire calls to the theoretical strike zones? The latter issue reduces the likelihood of inadvertently introducing a systematic bias favoring good hitting or good pitching baseball teams.

Do the umpires views or preference matter?

The amount of balls and strikes called by an umpire is a necessary, but not a sufficient condition to establish bias. The percentage of balls and strikes called by an umpire also depends on the quality or accuracy of the pitcher. Luckily, baseball collects data that allows one to compute the accuracy of calls made by an umpire in relation to the theoretical strike zone. These statistics are available by players, teams' seasons etc. And that allows us to test multitude of hypothesis regarding umpires' potential biases. The statistics allows us to identify the umpires with the characteristic we desire.

The data allows us to focus on another issue of interest, whether an umpire preference affect his or her decisions. One way to frame the issue is to ask a simple question regarding two umpires with identical accuracy records, does it matter that one may have preference for a team or a

player? If the results do not show any evidence of differential bias, the answer is a rotund no. It does not matter what the umpire preferences are, what matters is that the umpire calls balls and strikes based on the rule book. Once the evaluation criteria are specified, say the percentage accuracy in calling balls and strikes, is all one needs to judge the home plate umpires. Given the strike zone, one can identify which umpires have general systematic biases, i.e. a different strike zone, biases against different teams and or biases against different players. The league has the required information to evaluate and weed out bad umpires. The umpires who have an unacceptable percentage of bad calls should not be promoted to the next level. The process weeds out bad umpires and or those whose biases lead them to deviate from the norm. Those who remain are the ones are the better callers of balls and strikes irrespective of their preferences for specific ballplayers and or teams.

Unintended bias

Many baseball fans argue that consistency is the hallmark of a good umpire, and an inconsistent application of the strike one makes for a bad or deficient umpire. Baseball purists contend that strict adherence to the blue book is a must, that deviations from the rule book, even if consistently applied, will have predictable effects that systematically alters how the game will be played.

Since this bias is not aimed at benefiting or hurting a team, we have labelled any consistent deviation from the rule book as an unintended bias. Once the players know who is going to be behind home plate, if they know the umpire strike zone and how he likes to call a game, the teams will adjust their pitching and hitting strategy accordingly. Over time persistent variations in the strike zone will elicit a reaction on the part of the teams. Teams will be structured to take advantage of the "unofficially" modified strike zone and their drafting and players selectin process will change accordingly.

 One important implication of this insight is that the teams with better management will be able to draft the players that are more adequate to the new incentive structure, i.e. the umpiring and the "modified" strike zone. Yet the "unintended bias" will not necessarily alter the balance of power in the league. The teams with the better

management will on average have a better record, i.e. wins, than those with less competent management. In other words, A students will always be A students. Changing the incentive structure only changes the behavior in such a way that the smartest students will still be at the top of the curve. The only thing the changes in incentive structure does is to alter the quality and nature of the education.

Next, we consider the effect of alternative strike zones:

- **A smaller strike zone**: Once players become aware that umps have adopted a smaller strike zone, the hitters will be much more selective. The pitchers will have a smaller target. These changes increase the likelihood of the batter getting on base. The game offensive statistics will improve. Also, if more balls are put in play and there are more players on base, the strategy of the game changes. Will there be less bunts, more double plays? Will the defensive alignments change? Will there be more or less stolen bases? and will the value of a home run decline? The smaller strike zone will change the dynamics of the game.

- **A wider strike zone** increases the likelihood of a pitch being called a strike. Forcing the batters to widen their strike zone increases the numbers of strikeouts and quite probably reduce the on base percentage of the players thereby resulting in less runs scored. All of this points to a change in strategy. The teams will play for a single run as they do in softball. Bunting, advancing the runners will become a larger part of the game.

- **A random strike zone** will not overtly favor one group over another, yet it will have negative consequences for the game. If the players do not know what the strike zone will be like, it will be hard to prepare for it. The random strike zone diminishes the players hitting, pitching and fielding discipline. The net effect being a debasement of a player

skill levels. The random strike zone also alters the game strategy as a greater degree of randomness is introduced to a game that cherishes strategy, statistics and tradition.

It is the fans who ultimately attend the games. It is the commissioner and the rule making bodies that decide how to alter the game. If the fans like the changes made by the baseball rules committee, they will support the rule makers by voting with their pockets. The sport will gain viewers and increase in popularity. If the fans do not like the rule changes, the attendance will decline and eventually the rule makers will be forced to respond and change the rules. If the fans are happy there will be no need to change the rules of the game and the blue book will remain the same. It is not up to the umpires to change the game as they see fit. It is important that the game selects umpires that adhere to the rules otherwise the umpires would be usurping that function from the rulemaking committees.

Hence umpire selection is very important and it could have an impact on the integrity of the game and the direction to which the game evolves. One way to minimize the unintended bias is to select umpires that adhere to the rule book and to fire those who do not do so.

Deliberate bias

Special interest groups would love to have a strike zone tailored to their needs. For example, hitters would prefer an umpire with a small strike zone. Pitchers will prefer an umpire with a larger strike zone. While the strike zone yields systematic biases about offense and defense, *ex ante* there is no reason to expect that a strike zone will favor a particular team over another. But what if the umpires have a deliberate explicit bias against and or in favor of a team. In this case it is possible that when the team they favor is batting the umpire will widen the strike zone and shrink it when it is the other teams is batting turn. Doing so tilts the odds towards the favored team.

Most of us have experienced in life instances where good intentions have produced bad outcomes. During our recreational league days of coaching our kids, we've observed that the umps made it their mission to balance the different contests. They took into consideration the quality of the teams and their pitchers. They applied a

tougher standard and a smaller strike zone to the teams with the better pitcher in order to make the contest more competitive. It is not the umpires' job to solve skills and coaching inequalities. If the league wanted competitive teams, they should have modified the rules on player selection and talent distribution among the teams. But once the contest starts, the umpire obligation is to call the game according to the blue book rules.

One way to illustrate the pervasive effects of the equalization policies is that in the extreme the games would ideally end in a tie. There would be no winners or losers, and everyone would get a participation trophy. But if that were the case what would motivate the coaches to teach the game and work hard in practice to improve the players' skills? Back to the professional level, a policy that induces a competitive equality among the teams yields some interesting insights. If all teams play 500 ball, who wins and or loses may be truly random. On average the teams have no incentive to structure the best possible team. Add revenue sharing to the equation and the one way to maximize profits is to minimize costs. The effect of the bias that induces equality among teams is to destroy incentives to compete and to hire the best talent. That is not a desirable outcome. Deliberate bias is something that in principle should be rooted out from the get-go.

This baseball analogy suggests that a policy of selecting umpires that incorporate their biases into their decision making will have significant negative effects on the game. Umpires who attempt to equalize the outcome, not the opportunities, will destroy the game's incentive system with significant negative consequences to the outcome of the game. Umpires who use their bias systematically to favor specific interest groups are usurping the jobs of the commissioner and rules making body.