Shivakumar, Ram

Working Paper

What Technology Says About Decision-Making : Evidence from Cricket's Decision Review System (DRS)

What Technology Says About Decision-Making: Evidence from Cricket's *Decision Review System* (DRS)

Ram Shivakumar

This version: August 12, 2015

Correspondence: Ram Shivakumar, Booth School of Business, University of Chicago, 5807 S. Woodlawn Ave, Chicago, IL 60637: Email: ram.shivakumar@chicagobooth.edu

Acknowledgements: I thank Sanjiv Akhand, Ray Ball, Saunak Basu, Andy Brod, Pratik Chakraborty, Pradeep Chintagunta, Amit Gupta, Anup Gupta, Muhammad Islam, Ivan Joseph, Shyam Kothari, Richard Leftwich, N.R. Prabhala, Chris Ruhm, Rajagopal Shivakumar and two anonymous referees for comments and suggestions that have improved this paper. All errors of commission and omission are mine alone.

Electronic copy available at: http://ssrn.com/abstract=2700405

What Technology Says About Decision-Making:
Evidence from Cricket's *Decision Review System* (DRS)

***Abstract*.** We assemble and analyze a novel data set that provides evidence on the quality of decision-making by on-field umpires and teams in the sport of cricket. The decision review system (DRS), introduced in November 2009, gives teams a limited right to challenge an on-field umpire's decision. Of the 1201 on-field umpire decisions that were challenged in test matches by 9 playing teams between 2009 and 2014, 310 (25.81%) were reversed by a 3rd umpire. Controlling for several match and team characteristics, logistic regressions show that decisions that are ruled "out" (as opposed to "not out") by the on-field umpire are between 75% and 182% *more* likely to be reversed when the challenges involve LBWs. When challenges involve catches, decisions that are ruled "out" are *equally* likely to be reversed as decisions that are ruled "not out". For both LBW and caught challenges, home and away teams are *equally* likely to win a 3rd umpire reversal. Interacting a time variable with the main regressors, we are unable to reject the null hypothesis that the coefficients on the interaction terms are stable through time.

Keywords: Cricket, 3rd umpire, challenge, caught, DRS (decision review system), hawk-eye, Hosmer-Lemeshow test, hot-spot, interaction terms, LBW, logistic regression, Nagelkerke R^2, odds ratio, on-field umpire, p value, pitch-map, reversed, stump microphones, upheld

Electronic copy available at: http://ssrn.com/abstract=2700405

What Technology Says About Decision-Making: Evidence from Cricket's *Decision Review System* (DRS)

Introduction

Anyone watching a game of sport today on TV can, with the benefit of slow motion replays and other technological aids, immediately tell whether an on-field umpire's decision is incorrect. To prevent umpiring errors from influencing the outcome of games, the administrators of sport have provided teams with a limited right to challenge an on-field umpire's decision. Modern technology, i.e., high-definition cameras, ball-flight simulators, computer graphics, microphones, etc., is then used to reassess contentious plays and deliver final verdicts. For instance, in tennis, a player's challenge to an umpire's decision is resolved by the use of hawk-eye- a computer system that visually re-creates the path of the actual tennis ball to determine whether the ball is likely to have landed inside or outside the tennis court. In American football, a coach's challenge to an on-field umpire's call is resolved by having the on-field umpires review the play (on a television on the sidelines) in slow-motion with the aid of high-speed cameras operating from multiple angles.

In November 2009, the International Cricket Council (ICC), which administers the game of cricket, introduced a new set of rules to adjudicate decisions that is known as the umpiring decision review system (DRS).[i] The DRS provides teams with a maximum of two unsuccessful opportunities per inning[ii] to challenge the decision of an on-field umpire. Responding to a challenge by one of the playing teams, a 3rd umpire, with the aid of multiple technologies that are described below, can uphold or reverse the decision of an on-field umpire. In the event that the 3rd umpire does not find evidence that contradicts the on-field umpire's call, the on-field umpire's decision is upheld and the team (that challenged the on-field umpire's decision) loses one opportunity to challenge. In the event that the 3rd umpire finds evidence that contradicts the on-field umpire's call, the on-field umpire's decision is reversed and the team (that challenged the decision) retains its right to challenge an on-field umpire's decision. If a team is on the wrong end of both of its challenges, it cannot challenge any more on-field umpire decisions during that inning.

The DRS and the technologies used to support it have provoked many passionate debates among cricket players (current and former), cricket writers and fans.[iii] Proponents of the DRS view it as an improvement over the status quo.[iv] Some critics of the DRS, such as Indian cricket's governing body, the BCCI, have argued that the technologies used to support the DRS are flawed and are likely to distort the way the game is played.[v] Some other critics of the DRS have argued that the DRS will ultimately lead to the loss of authority for, and trust in, on-field umpires.[vi]

Despite the very large number of opinion columns in newspapers, magazines, websites and blogs on the DRS, what is surprising that there has been no empirical assessment of the DRS and its influence on decision-making by umpires and teams. Cricket is a particularly good laboratory to study the link between the use of modern technologies and decision-making for several reasons. First, on-field umpires apply a great deal of subjective judgment in adjudicating plays. For instance, one of the most contentious forms of dismissal in cricket is the leg-before-wicket (LBW) decision where the umpire is called upon to rule whether the ball has hit a batsman's pad and whether the ball would have gone on to hit the wicket.[vii] On average, 4 batsman are declared out LBW in every test match[viii] though, as anyone who has watched a test match can attest to, the number of appeals (by the fielding team) for an LBW is generally several multiples of 4. The DRS provides teams with an opportunity to receive restorative justice though, given the limited rights to challenge on-field umpire decisions, teams must use their challenges prudently.

Second, an umpire's decision to incorrectly give a batsman "out" or to incorrectly give a batsman "not out" can significantly alter the course of a match.[ix] The DRS was introduced to protect teams from decisions that are obviously wrong.

Third, the on-field umpires and players have to make their decisions rapidly and are prohibited by ICC rules from consulting with others outside the playing field. For instance, an on-field umpire usually responds to an appeal by the fielding team within a matter of seconds. Furthermore, teams challenging the on-field umpire's decision must do so within 30 seconds of the on-field umpire making his decision. Since test matches are played for up to 5 days in a row (30 hours of play), on-field umpires must display a great deal of sustained concentration.

The aim of this paper is to conduct an empirical assessment of the DRS. At the outset, it is important to emphasize that we do not question the efficacy of the technologies that the 3rd umpire employs to render his decision. We take the functions, attributes and capabilities of the technologies used by the 3rd umpire to make his decision as a given. We provide answers to two questions that will be of interest to cricket (and sport) aficionados as well as academics. First, what does modern technology say about the accuracy of the decisions made by the on-field umpires? If on-field umpires are shown to make "few" errors, then ontological authority has been legitimized. If on-field umpires are shown to make "many" errors, an obvious implication is that on-field umpires have to be selected more carefully and that they must receive more training and guidance. Importantly, an assessment of the types of errors that on-field umpire errors make can help the ICC revise the guidelines and training systems required to reduce umpiring errors. Second, how is the 3rd umpire's decision (to reverse or uphold the on-field umpire's decision) related to the other variables that also exert an influence on the outcome? To be sure, the details surrounding each contentious play matter. These details include match information and team characteristics such as the identity of the batsman and the bowler, the nature of the challenge, the on-field umpire's identity, the identity of the player/team that is challenging the decision, etc.

We assemble and analyze a novel data set that encompasses the entire population of DRS challenges and 3rd umpire decisions in test matches played between 2009 and 2014. Of the 1201 on-field umpiring decisions that were challenged by the playing teams, 310 decisions (25.81%) were reversed by the 3rd umpire- a *conditional* on-field umpire error rate that is higher than the *unconditional* error rate (approximately 8%) attributed to umpires by the ICC. Controlling for the influence of several market and team characteristics, logistic regressions are used to assess the influence of several variables on the likelihood of the 3rd umpire reversing the on-field umpire's decision.

The logistic regressions show that "out" decisions are between 75% and 182% *more* likely to be reversed (to "not out") as "not out" decisions (are to "out") when the challenges involve LBWs whereas "out" decisions are *equally* likely to be reversed (to "not out") as "not out" decisions (to "out") when the challenges involve catches, holding other variables constant. Simply stated, with caught challenges, on-field umpires are equally likely to make errors in either direction whereas with LBW challenges, on-field umpiring errors are consistently biased against batsman. This result provides credence, though not proof, for the contention that with the advent of the DRS, on-field umpires appear to be less willing to give the benefit of the doubt to the batsman, a tradition, though not a rule, in cricket for more than a century.[x] The logistic regressions also show that for both LBW and caught challenges, one cannot reject the null hypothesis that home and away teams are *equally* likely to win a 3rd umpire reversal. In other words, there is no evidence of bias (in favor of the either team) in the way that 3rd umpires adjudicate decisions.

To test for the evolution of decision-making by umpires through time, we interact a time variable with the two main independent variables- the on-field umpire decision ("out" or "not out") and the challenge team characteristic (home or away). The interacting terms are meant to capture the influence of technological change on 3rd umpire reversals through its influence on learning by on-field umpires and the playing teams. Because none of the coefficients on the interaction terms are statistically significant, we are unable to conclude that there has been a significant change in the influence of the interaction terms on the proclivity for 3rd umpire reversals.

Two recent papers examine decision-making by umpires in cricket. To test for the presence of bias in umpiring, Sancheti et al (2015) study LBW decisions in 1000 test matches played between 1986 and 2012. They find that when both on-field umpires are home umpires, the home team receives 16% fewer adverse decisions than visiting teams though this home-bias falls to 0.7% when both umpires are neutral umpires. They conclude that the source of the home-team bias (in the early years of their sample) was outright favoritism rather than social pressure from home crowds. They also observe that the number of LBWs (given against home and away teams by the on-field umpires) have fallen after the introduction of the DRS.

Borooah (2015) challenges the idea that the Hawk-Eye technology (used to determine LBW decisions) is always right. Referring to Paul Hawkins's (the founder of Hawk-Eye) claim that Hawk-Eye's prediction of the ball-path is accurate up to 0.2 inches and that the overall prediction (when the ball crosses the plane of the stumps) is accurate up to 0.8 inches,[xi] Borooah rightly notes that the failure to provide confidence intervals around the mean error can lead the players, umpires and the viewing public to draw misleading conclusions. Taking Hawkins's data-estimate as the true estimate of the mean error rate and assuming a normal distribution of error rates, Borooah constructs a confidence interval for prediction paths. If standard errors are incorporated, he demonstrates via example that many marginal calls that have gone against batsman should be given “not out” because the odds of the ball missing the stumps are not-trivial. Given that there are two ways of improving decision-making in cricket- better technology (through higher speed cameras, for instance) or better umpires (through better training)- he argues, using cost-benefit analysis, that an efficient solution calls for greater training for umpires.

It is worth noting, for the benefit of the reader not familiar with cricket and the DRS, that LBW and caught challenges have distinct standards for adjudication (by the 3rd umpire). With catches, the 3rd umpire is (usually) expected to find conclusive evidence as to whether the on-field umpire’s decision was correct or incorrect. With LBW challenges, Hawk-eye produces a margin of error around its predicted ball path so that the 3rd umpire is sometimes only able to conclude that the on-field umpire’s decision cannot be shown to be wrong. In such cases, the 3rd umpire chooses the “umpire’s call”.[xii]

We mobilize our thoughts in four sections. The next section discusses the mechanics of the DRS and the technological aids that the 3rd umpire uses to make his decision. The following section discusses the dataset that we have assembled and presents the descriptive statistical facts. In the penultimate section, we estimate logistic regressions that summarize the influence of match and team-specific variables on the 3rd umpire’s decision. A final section offers concluding remarks.

The DRS and Technological Aids

Figure 1 illustrates the mechanics of the DRS with the aid of a decision tree. At the 1st decision node, the on-field umpire responds to a query (an "appeal" in cricket jargon) from the fielding team by declaring the batsman "out" or "not out." At the 2nd decision node, the teams responds to the on-field umpire's decision by challenging it or by not challenging it. If no challenge occurs, the on-field umpire's decision is final and the game proceeds.[xiii] If one of the teams challenges the on-field umpire's decision, we proceed to the 3rd decision node where the 3rd umpire, a qualified umpire who is sitting in a room outside the playing field that is equipped with television monitors, gadgets and other support systems (see below), re-examines the contentious play. If a review of the contentious play is unable to conclusively reject the on-field umpire's decision, the on-field umpire's decision is upheld. If a review of the play provides conclusive evidence that the on-field umpire's decision is incorrect, the on-field umpire's decision is reversed. In the event of a reversal, the player who was declared out (not out) by the on-field umpire gets to continue (not continue).

The DRS is generally supported by 3 distinct technological aids: (1) Hawk-eye (2) Hot-spot and (3) Snickometer.[xiv] Hawkeye, a ball-tracking system which is also used in tennis and soccer, is essential in helping the 3rd umpire adjudicate LBW challenges. It relies on several high-definition cameras that record each play from multiple angles from the moment that the ball leaves the bowler's hand till. The cameras record the ball's speed, path and bounce and upload the data frame-by-frame to a computer software system. This software system uses the data to predict the future path of the ball beyond the point where it made contact with the batsman, hence determining whether the ball would have gone on to hit the stumps. The still photograph on the top left-hand side of Figure 2 is an illustration of what the 3rd umpire sees when he watches a slow-motion replay of the contentious decision. [xv]

Hot-spot is an infra-red imaging system that helps determine whether the ball has made contact with the batsman's bat (or glove) or if the ball has made contact with the batsman's pad. It is a vital technological aid in adjudicating both LBW and caught decisions.[xvi] Hot-spot is implemented with the use of infra-red cameras that are positioned at opposite ends of the cricket ground. The cameras measure the heat-friction generated by contact between the ball and the bat (or the pad) and generate a series of black-and-white frames that zero in on the point of contact. The still photograph on the top right-hand side of Figure 2 is an example of what the 3rd umpire sees when he watches a slow-motion replay of the contentious decision.[xvii]

The snickometer, a technology used to assess challenges involving catches, provides audiovisual evidence on whether the ball might have made contact with the batsman's bat. In the event that the ball has made contact with the batsman's bat, an elevated soundwave will appear on the screen at the moment when the ball passes the bat. The still photograph at the bottom of Figure 2 is an example of what the 3rd umpire sees when he reviews a catch challenge.[xviii]

Data and Descriptive Statistics

The data used in this study is assembled from the text commentary of test matches (provided in real-time) from ESPN Cricinfo's website (www.cricinfo.com). We start by cataloging every DRS challenge in test matches from the introduction of the system in November 2009 through December 2014.[xix] In addition, we catalog other information such as the index number and dates of the test match, the identity of the home and away teams, the city (specifically, the cricket ground) in which the match was played, the identity of the team that initiated the challenge, the identity of the batsman and bowler when the on-field umpiring decision was challenged, the batsman's rank (whether top 7 or bottom 4 as listed on the team sheet), the bowler's style (fast or spin), the type of challenge (LBW or caught), the identity of the on-field umpire, the stage of the game when the challenge was initiated (over #), the decision that the on-field umpire made and, importantly, the decision of the 3rd umpire.

Table 1 provides a summary of the dataset. It shows that there were 1201 challenges by the playing teams during the 126 test matches in which the DRS was employed. Among the many statistical facts that will be of interest to cricket aficionados and others, two stand out. First, 310 of the 1201 challenges were reversed by the 3rd umpire- a reversal rate of 25.81%. In simple language, more than 1 in every 4 decisions made by the on-field umpire and challenged by one of the playing teams is reversed. The ICC claims that umpires get more than 91% of their decisions right.[xx] However, they are not explicit in telling us what their sample includes and excludes. Presumably their sample does not count the many on-field decisions ("out" and "not out") that are patently obvious and do not require the on-field umpire to render a judgment. The ICC also does not disclose how it separates the many frivolous appeals that fielding teams make (in optimism rather than conviction) from the more serious appeals (that are not challenged after due contemplation). Second, teams (as a group) challenge approximately 10% *more* decisions in away games than in home games. This may be because: (a) teams have greater familiarity with playing conditions at home and less familiarity with playing conditions away from home and/or (b) home teams still benefit (relative to visiting teams) from on-field umpire bias (despite the presence of neutral umpires) so teams have to challenge more on-field umpire decisions in away games to offset the advantage that accrues to home teams.[xxi]

Table 2 is a summary matrix of all on-field umpire decisions that were challenged by the playing teams. The vertical axis displays the decision of the on-field umpire- "out" or "not out"- while the horizontal axis identifies the decision of the 3rd umpire- "out" or "not out". The diagonal terms are those decisions that the on-field umpire got correct while the off-diagonal terms are those decisions that the on-field umpire got wrong. The top row of the matrix measures the challenges made by the batting team and the bottom row of the matrix measures the challenges made by the fielding team. Notice that fielding teams account for approximately 57% of all challenges though batting teams win a higher proportion of their challenges.[xxii]

Table 3 provides a breakdown of DRS decisions by type of challenge: "LBW" and "Caught." Notice that LBW challenges constitute more than 75% of all challenges though 3rd umpire reversals are less frequent for LBW challenges than caught challenges. Readers will also notice the greater dispersion in reversal rates across countries for caught decisions as opposed to LBW decisions.[xxiii]

Tables 4A and 4B provides a more granular look at the LBW and Caught challenges. They identify the combination of "out" and "not out" decisions taken by the on-field and 3rd umpires respectively. From Table 4A, it is easy to show that the odds of an LBW being reversed from "out" to "not out" is equal to the probability of an “out” being reversed to “not out” divided by the probability of an “out” being upheld as “out” or (119/420)/(301/420) = .2833/.7166 = .3953. The odds of an LBW being reversed from “not out” to “out” is equal to the probability of a “not out” being reversed to “out” divided by the probability of a “not out” being upheld as “not out” or (91/492)/(401/492) = .1849/.8150 = .2260. Hence, the odds ratio of an LBW challenge that is given “out” by the on-field umpire being reversed (relative to an LBW challenge that is given “not out” being reversed) is equal to 1.7514. That is, LBW challenges that are given “out” are about 75% *more* likely to be reversed (than LBW challenges that are given “not out”). The logistic regressions in the next section provides further evidence on the robustness of this result.

From Table 4B, it is possible to show that the odds of a Caught being reversed from “out” to “not out” is equal to the probability of an “out” being reversed to “not out” divided by the probability of an “out” being upheld as “out” or (40/90)/(50/90) = ..4444/.5555 = .80. The odds of a Caught being reversed from “not out” to “out” is equal to the probability of a “not out” being reversed to “out” divided by the probability of a “not out” being upheld as “not out” or (60/199)/(139/199) = .3015/.6984 = .4317. Hence, the odds ratio of a Caught challenge that is given “out” by the on-field umpire being reversed (relative to a caught challenge that is given “not out” being reversed) is equal to .80/.4317 = 1.8531. That is, Caught challenges that are given “out” are about 85% *more* likely to be reversed (than caught challenges that are given “not out”). In fact, the logistic regressions in the next section show this effect is not statistically significant.

Table 5 separates lbw challenges and decisions (made by the 3rd umpire) initiated by the home team from those initiated by away teams and Table 6 separates caught challenges and decisions (made by the 3rd umpire) initiated by the home team from those challenges initiated by away teams.

Logistic Regressions

We estimate logistic regressions of the form:

$$\text{Log}\ [q_i / (1 - q_i)] = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \beta_{3M} X_{M,i} X_{3,i} + \ldots.\ \varepsilon_i$$

Where $q_i = p(Y_i = 1)$ is the probability of a reversal by the 3rd umpire (Y_i is the ith observation of the dependent variable), the X's are the predictor variables and the β's are the estimated regression coefficients. Some of the predictor variables are products of other predictor variables- that is, they are interacting terms. The interacting terms are included alongside the main terms so that we can separate the direct effect of the main terms (on the dependent variable) from their indirect effect through one main term's interaction with another main term.[xxiv]

Table 7 is a list of the variables used in the logistic regressions. The dependent variable, *3rd umpire decision* (3RDUD), is a binary variable that takes on a score of 1 when the 3rd umpire *reverses* the on-field umpire's decision and a score of 0 when the 3rd umpire *upholds* the on-field umpire's decision. The 1st independent variable, *on-field umpire decision* (OFUD), takes on a value of 1 when the on-field umpire's decision is "out" and a value of 0 when the on-field umpire's decision is "not out." The odds-ratio for OFUD provides an estimate of the likelihood of a 3rd umpire reversal when the on-field umpire gives a decision of "out" as opposed to a decision of "not out", holding other variables fixed. The 2nd independent variable, *challenge team characteristic* (CTC) takes on a value of 1 when the team challenging the on-field umpire's decision is the home team and a value of 0 when the team challenging the on-field umpire's decision is the away team. The odds-ratio for CTC will tell us whether home teams are more (or less) likely to win reversals than away teams, holding other variables constant. The 3rd independent variable, *TIME*, takes on an integer value between 1 and 5 depending on the year in which the challenge occurred- a value of 1 is assigned for all challenges that occurred in 2010, a value of 2 is assigned for all challenges that occurred in 2010 and so on.[xxv] TIME is included as an independent variable so as to capture the influence of the evolution of the technologies (used to implement the DRS) on reversal rates. The 4th independent variable, *OFUD*TIME*, is calculated as the simple observation-by-observation product of OFUD and TIME. It is included as a separate variable from TIME and OFUD so as to capture the influence of the interaction between technological advances and changes in on-field umpire decision-making on the proclivity for a 3rd umpire reversal. The 5th independent variable, *CTC*TIME*, is the observation-by-observation product of CTC and TIME. It is included as a separate variable (from TIME and CTC) so as to measure the influence of the interaction between technological advances and changes in team decision-making on 3rd umpire reversals. Since we separate the pool of LBW challenges from Caught challenges, we (separately) estimate and report coefficients for OFUD, CTC, TIME as well as the interacting terms for each type of challenge.

Six covariates are included in all the regressions. The variable, *batsman characteristic* (BATC), takes on a value of 1 when the batsman involved in the challenge is a top-order batsman (listed in the team sheet as a top 7 batsman) and a value of 0 when the batsman involved in the challenge is a lower-order batsman (listed in the team sheet as a bottom 4 batsman). The variable, *bowler characteristic* (BOWC) takes on a value of 1 when the bowler involved in the challenge is a fast bowler and a value of 0 when the bowler involved in the challenge is a spin bowler. The variable, *umpire id* (UID), is a dummy variable that takes on a value of 1 when the on-field umpire involved in the challenge is, for instance, Aleem Dar, and 0 otherwise. The model contains 17 dummy variables, one for each umpire, with K. Dharamasena as the reference umpire.[xxvi] The variable, *challenge team id (CTID)*, is a dummy variable that takes on a value of

1 when the challenging country is a given country (say England) and 0 otherwise. Since we have 9 countries in the sample, the model contains 8 dummy variables with Australia as the reference country. The variable, *challenge team rank* (CTR), takes on an integer value between 1 and 9 (reflecting the ICC's rank-ordering of teams in the test match format at the time that the test match was played). The variable, *Location* (LOC), is a dummy variable that takes on a value of 1 if the match is held in a given country (say England) and 0 otherwise. Since test matches were held in 8 countries during the sample period,[xxvii] the model contains 7 dummy variables with Australia as the reference country.

Tables 8 and 9 report the results from the logistic regressions. In each table, column 1 reports the estimated β coefficient for the predictor variables and columns 2 and 3 report the associated standard errors and Wald test statistics respectively. Column 5 reports the estimated odds ratios [Exp (β)] for each variable and column 4 lists the corresponding p-value. Columns 6 and 7 report the lower and upper bounds for the estimated odds ratios at the 95% confidence limit. Each regression also reports the Nagelkerke R^2, a pseudo coefficient of determination that takes on a value between 0 and 1, as well as the estimates from a Hosmer-Lemeshow test- a chi-squared goodness-of-fit test.

Table 8 reports the results from a pair of logistic regressions- the first for LBW challenges and the second for Caught challenges. The dependent variable is the likelihood of a 3rd umpire reversal. From inspection (of the p values) on the pair of regressions, notice that OFUD|LBW is the only variable that is statistically significant. The odds-ratio of 1.752 tells us that "out" decisions are 75.2% *more* likely to be reversed from "out" to "not out" than vice-versa. Simply stated, when it comes to LBWs, on-field umpiring errors have been biased against batsman.

To understand the import of this result, it is worth noting that for more than a century, one of the unwritten rules of cricket has been that on-field umpires give the benefit-of-doubt to batsman. That is, if umpires believe that there is even a small probability that batsmen (involved in a contentious play) are "not out", they are to be given "not out."[xxviii] Some cricket players and journalists have observed that the advent of the DRS has upset this tradition because on-field umpires appear to be willing to give batsman "out" more frequently than earlier.[xxix]

As Peter Willey (the former English umpire) notes[xxx]:

> "I do find myself thinking differently about decisions, especially, LBWs. And I shouldn't think like that. I have been umpiring for 15 years and for the last 13 certain decisions were definitely not out- now due to television I think they are close… when the technology is used fully a lot more people are given out. Reassessing my decision-making means I am giving decisions based on Hawk-Eye graphics on the television."

Findlay (2011) bluntly summarizes the impact of the DRS on on-field umpiring decisions:

> "The DRS has clearly shown that umpires have been too conservative in considering lbw appeals in the past, and that batsmen have been getting away with murder for years! The higher incidence of lbw dismissals for both spin and quick bowlers is a result not only of the direct intervention of video replays, but also, in all probability, of a realisation by umpires that they are safer in giving out what they originally would have considered to be marginal decisions only a year or two ago."

The fact that it is 75% more likely for "out" decisions to be reversed (as opposed to "not out" decisions) suggests that the pendulum may have gone too far in the other direction- on-field umpires appear to give the benefit of the doubt to the bowler.

While there is some evidence (in cricket and other sports) that home teams receive the marginal calls in their favor,[xxxi] notice (from inspection of the p-values in Table 8) that the variable, CTC, is statistically insignificant in both sets of regressions. Consequently, one cannot reject the null hypothesis that home and away challenges are *equally* to win (or lose) reversals from the 3rd umpire. In simple language, we find no evidence that home teams have any advantage over visiting teams in winning 3rd umpire reversals.

Table 9 reports the results from a pair of regressions when time and time-interaction effects are included as explanatory variables. Notice (by inspection of p values) that the only variable that is statistically significant is OFUD|LBW. With the inclusion of time and time-interaction effects, an "out" decision is now 182% *more* likely to be reversed to "not out" (than "not out" decisions are reversed to "out"). What is also noteworthy is that neither TIME nor the interaction variables, OFUD*TIME and CTC*TIME, are statistically significant. That is, we are unable to find evidence that technological change through its interaction with changes in the way umpires and teams make decisions, has had any influence in altering 3rd umpire reversal rates.

Conclusion

Technology is exerting a significant influence on the way sport is adjudicated and played. However, with the exception of anecdotal evidence, precious little is known about the quality of decision-making by on-field umpires. We take a first step towards redressing this imbalance by assembling and analyzing a novel data set that provides evidence on the quality of decision-making by on-field umpires in test-match cricket. Our analysis leads to 4 conclusions. First, more than 1 in every 4 decisions made by the on-field umpire are incorrect. Second, for decisions involving LBWs, on-field umpire decisions of "out" are between 75% and 182% *more* likely to be reversed by the 3rd umpire to "not out" than on-field umpire decisions of "not out". Third, home and away teams are *equally* likely to have their challenges successfully reversed by the 3rd umpire. And fourth, testing for time interaction effects, we are unable to reject the null hypothesis that the coefficients are stable through time.

These results have important implications for way cricket is played, adjudicated by on-field and 3rd umpires and administered by the ICC. First, since teams have a limited right to challenge on-field umpire decisions and since technology is getting better at exposing umpiring errors, the pressure on on-field umpires to reduce errors, and on the ICC to amend the DRS, will likely increase. Cricket is relatively unique among team sports (like American football, soccer and baseball) that have instituted a review of on-field umpiring decisions in three significant ways: (a) cricket uses technology more intensively than other sport to resolve differences of opinion over umpiring decisions. This is attributable, in part, to the fact that test-match cricket is, unlike other games of sport, played over a five-day period with the ultimate result (win, lose or draw) often depending on a small number of decisions by the on-field umpires (b) the players on the cricket field, rather than coaches who have access to greater information, are the ones who must choose in real-time whether to challenge an on-field umpire's decision. In major league baseball

(MLB), the managers of the respective teams, the only ones with the authority to challenge umpiring decisions, will have likely received a green light from their assistants before they invoke their authority to challenge. (c) In cricket, on-field umpiring errors are generally corrected only when the playing teams use their limited right to challenge decisions. Simply stated, the DRS forces teams to play a game within a game. The English Premier League (EPL) has begun to use goal-line technology, a system that instantaneously alerts the referee and match officials by transmitting a signal to the referee's watch. In theory, there is no limit to the number of times the referee can correct his own decision.[xxxii]

Second, because LBW challenges constitute 75% of all challenges and since a quarter of LBW challenges are reversed, the ICC is likely to focus its attention on helping on-field umpires reduce their error rates even. As the margin of error associated with the Hawk-eye's predicted ball path gets smaller, umpiring errors will likely increase placing greater pressure on the ICC to amend the rules that 3rd umpires use to assess on-field umpire decisions.

Third, on one important dimension, the DRS can be said to be working quite well: there is no evidence of bias against home or away teams in the way that the 3rd umpires make their decisions. Given the paucity of umpires on the elite panel (umpires deemed to be the best), it has been suggested by some writers that at least one home umpire should be allowed to umpire in test matches.[xxxiii] However, given the limited number of challenges that each team can call upon, and most importantly, our limited knowledge of the quality of on-field umpiring decisions when on-field umpire decisions are not challenged, we recommend that cricket administrators proceed with caution.

The analysis conducted here ought to be extended in several ways. First, relatively little is known about how the quality of decision-making varies across teams and across umpires, whether this variation is significant and the nature of the policies and systems that might be adopted to reduce such variance. Second, relatively little is known about the quality of on-field umpiring when teams do not challenge on-field umpire decisions. The analysis of these questions will provide a more comprehensive picture of the quality of decision-making by on-field umpires and teams.

References

Borooah, V., (Forthcoming 2015). Upstairs and Downstairs: The Imperfections of Cricket's Decision Review System. *Journal of Sports Economics*.

Brettig, D., (November 11, 2013). Clarke claims that DRS distorts the game. *ESPN Cricinfo*.

Brettig, D., (November 20, 2013). Snicko is Hot Spot's Insurance Policy. *ESPN Cricinfo*.

Chappell, I., (July 14, 2013). Leave the DRS to the umpire. *ESPN Cricinfo*.

Chedzoy, B., (1997). The Effect of Umpiring Errors in Cricket. *Statistician*, 46, 529-540.

Collins, H., (2010). The Philosophy of Umpiring and the Introduction of Decision-Aid Technology. *Journal of the Philosophy of Sport.* 37, 135-146.

Collins, H. and Evans, R., (2012). Sports Decision-Aids and the "CSI-effect": Why Cricket uses Hawk-Eye Well and Tennis uses it Badly. *Journal of the Philosophy of Sport,* 21, 904-921.

Coverdale, B., (August 1, 2013). DRS breaking spirits of players. *ESPN Cricinfo*.

Coverdale, B., (August 7, 2013). Watson believes DRS upstages benefit of doubt. *ESPN Cricinfo*.

Coverdale, B., (December 14, 2013). England baffled by Root's DRS dismissal. *ESPN Cricinfo*.

Cozier, T., (December 6, 2009). Umpire Review System Not Working. *ESPN Cricinfo*.

Crowe, M., (March 24, 2014). Why the BCCI won't be swayed by Richardson's claims. *ESPN Cricinfo*.

Dobell,G., (February 14, 2012). DRS has affected the game more than we thought it would. *ESPN Cricinfo.*

Dobell, G., (August 6, 2013). DRS "hasn't worked well"- Flower. *ESPN Cricinfo*.

Dobell, G. and J. Kimber. (August 12, 2013). Rogers, the master of the DRS. *ESPN Cricinfo*.

Findlay, R., (April 5, 2011). The DRS effect on LBW decisions. *ESPN Cricinfo.*

Gilchrist, A., (July 17, 2013). The DRS is not the solution, it is the problem. *ESPN Cricinfo*.

Hughes, A., (December 13, 2014). The DRS's in-built safety mechanisms. *ESPN Cricinfo*.

International Cricket Council. (2009). *Umpire decision review system- A guide.* Dubai, UAE, 1-49.

Jayaraman. S., (March 3, 2014). Umpiring is about excellence, not neutrality. *ESPN Cricinfo*.

Moonda, F., (March 9, 2012). Kallis unconvinced by ball-tracking technology. *ESPN Cricinfo*.

Moskowitz, T. and J. Wetheim. (2012). *Scorecasting: The Hidden Influences Behind How Sports are Played and Games are Won*. Three Rivers Press.

Nicholas, M., (January 2, 2015). Stop embarrassing yourself over the DRS, India. *ESPN Cricinfo*.

Ozer-Balli, H., and B. Sorensen. (August 2013). Interaction Effects in Econometrics. *Empirical Economics*, 583-603.

Pringle, D. (November 8, 2010). The Ashes 2010: Graeme Swann goes to great lengths to thank bowlers. *Daily Telegraph*.

Rajesh, S., (2006). A Threat or an Advantage, *ESPN Cricinfo*.

Sancheti, A., D. Patton and I.G. Smith, (forthcoming 2015). Home bias in officiating: Evidence from international cricket. *Journal of the Royal Statistical Society*.

Selvey, M., (February 4, 2012). DRS is a friend to nobody- for the good of the game it needs a rethink. *The Guardian.*

Smyth, R., (February 4, 2012). Why all the controversy over DRS during Pak-Eng Series. *The Guardian*.

Steen, R., (December 2011). Going upstairs: The decision review system- velvet revolution or thin edge of an ethical wedge. *Sport in Society*, 14, 1428-1440.

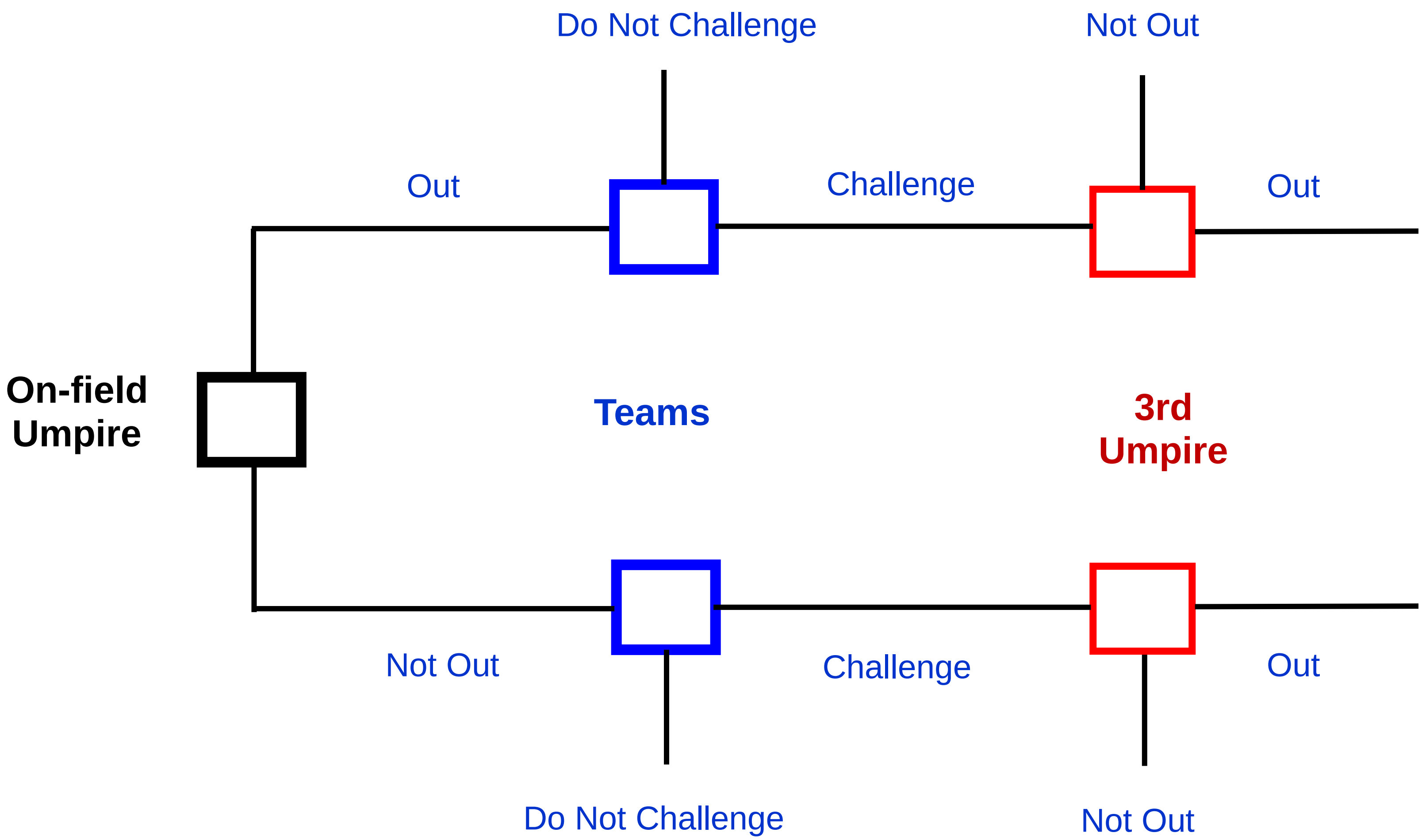
Steen, R., (March 20, 2013). Were the good old days really better? *ESPN Cricinfo*

Stewart, A., (2013). Ashes 2013: bring back 'home' umpires. *BBC,* London.

Thawfreeq, S., (2008). Review System is a Blessing for Bowlers- Bayliss. *ESPN Cricinfo.*

Willey, P., (2009). Why Umpiring Reviews are Flawed. *ESPN Cricinfo.*

Figure 1



Do Not Challenge
Not Out
Out
Challenge
Out
On-field
Umpire
Teams
3rd
Umpire
Not Out
Challenge
Out
Do Not Challenge
Not Out

Figure 2
Technologies used to support the DRS

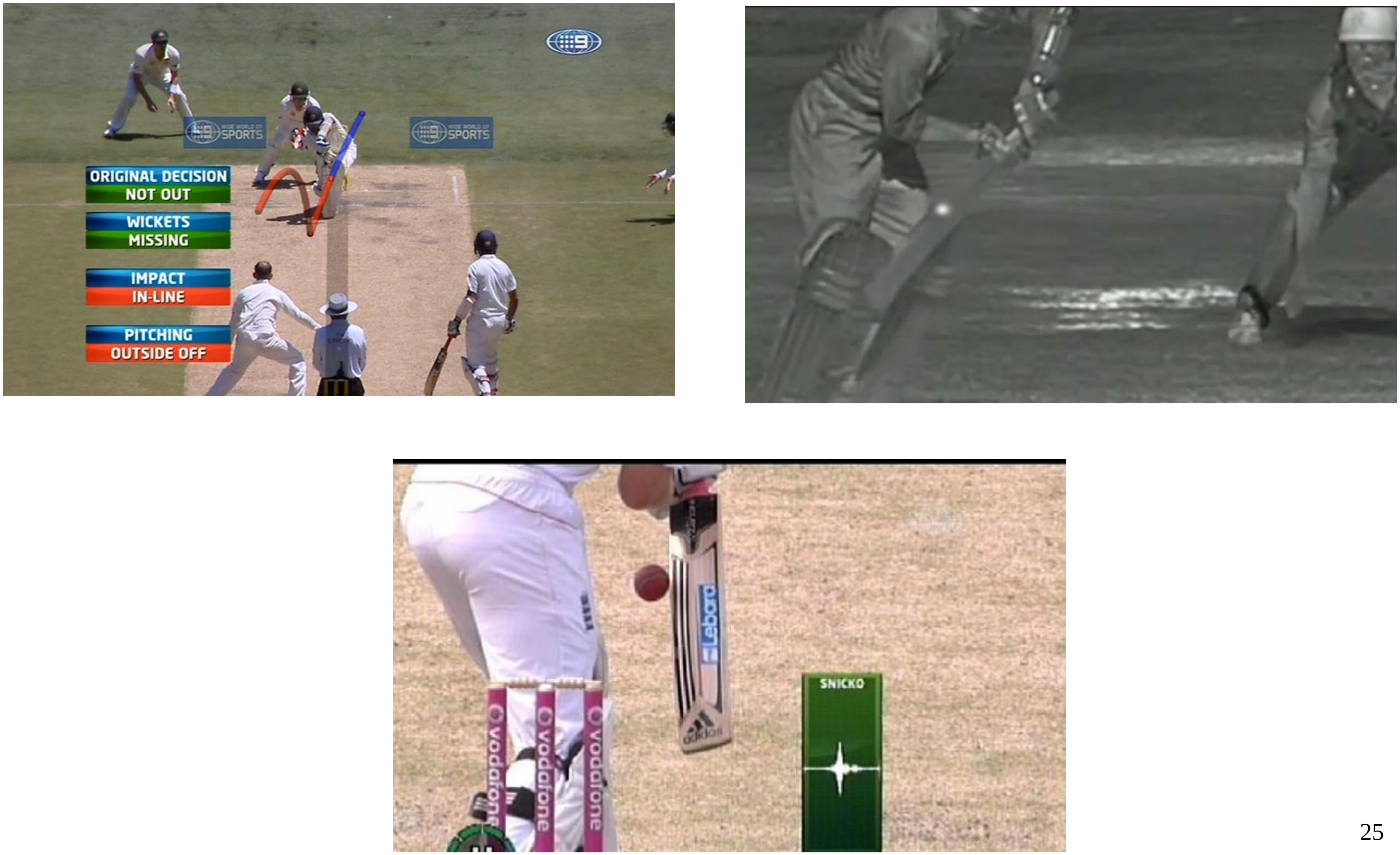


Table 1- DRS Use by Teams

Team	Matches	Challenges	Reversals	Reversal Rate	Home Challenges	Reversals	Away Challenges	Reversals
Australia	45	197	58	29.44	92	30	105	28
England	45	187	55	29.41	92	30	95	25
New Zealand	30	131	34	25.95	63	18	68	16
South Africa	34	154	42	27.27	82	22	72	20
Pakistan	30	194	42	21.65	90	13	104	29
Sri Lanka	26	148	35	23.65	67	14	81	21
West Indies	30	127	28	22.05	70	12	57	16
Zimbabwe	6	35	11	31.43	0	0	35	11
Bangladesh	6	28	5	17.86	11	1	17	4
TOTAL	**126**	**1201**	**310**	**25.81%**	**567**	**140**	**634**	**170**

Table 2- Decisions by on-field umpires and 3rd umpires

		3RD UMPIRE'S DECISION	
		OUT	NOT OUT
ON-FIELD UMPIRE'S DECISION	OUT	**351**	**159**
	NOT OUT	**151**	**540**

Table 3- Breakdown of decisions by type of challenge

Teams	LBW Challenges	Reversals	Reversal Rate	Caught Challenges	Reversals	Reversal Rate
Australia	140	37	26.43%	57	21	36.84%
England	134	32	23.88%	53	23	43.40%
New Zealand	106	27	25.47%	25	7	28.00%
South Africa	107	25	23.36%	47	17	36.17%
Pakistan	159	32	20.13%	35	10	28.57%
Sri Lanka	113	28	24.78%	35	7	20.00%
West Indies	105	19	18.10%	22	9	40.91%
Zimbabwe	28	6	21.43%	7	5	71.43%
Bangladesh	20	4	20.00%	8	1	12.50%
Total	**912**	**210**	**23.03%**	**289**	**100**	**34.60%**

Table 4- Further Breakdown of DRS Use by Type of Challenge

LBW

ON-FIELD UMPIRE'S DECISION	**3RD UMPIRE'S DECISION:** OUT	NOT OUT
OUT	**301**	**119**
NOT OUT	**91**	**401**

Caught

ON-FIELD UMPIRE'S DECISION	**3RD UMPIRE'S DECISION:** OUT	NOT OUT
OUT	**50**	**40**
NOT OUT	**60**	**139**

Table 5- LBW Challenges & Decisions, Home and Away

Team	Home Challenges	Reversals	Reversal Rate	Away Challenges	Reversals	Reversal Rate
Australia	62	19	0.3065	78	18	0.2308
England	67	19	0.2836	67	13	0.1940
New Zealand	51	16	0.3137	55	11	0.2000
South Africa	56	10	0.1786	51	15	0.2941
Pakistan	77	11	0.1429	82	21	0.2561
Sri Lanka	52	11	0.2115	61	17	0.2787
West Indies	58	9	0.1552	47	10	0.2128
Zimbabwe	0	0	-	28	6	0.2143
Bangladesh	9	0	0.0000	11	4	0.3636
TOTAL	**432**	**95**	**0.2199**	**480**	**115**	**0.2396**

Table 6- Caught Challenges and Decisions, Home and Away

Home Team	**Home challenges**	**Reversals**	**Reversal rate**	**Away challenges**	**Reversals**	**Reversal rate**
Australia	30	11	0.3667	27	10	0.3704
England	25	11	0.4400	28	12	0.4286
New Zealand	12	2	0.1667	13	5	0.3846
South Africa	26	12	0.4615	21	5	0.2381
Pakistan	13	2	0.1538	22	8	0.3636
Sri Lanka	15	3	0.2000	20	4	0.2000
West Indies	12	3	0.2500	10	6	0.6000
Zimbabwe	0	0	-	7	5	0.7143
Bangladesh	2	1	0.5000	6	0	0.0000
TOTAL	**135**	**45**	**0.3333**	**154**	**55**	**0.3571**

Table 7- Regression Variables

DEPENDENT VARIABLE	EXPLANATION
3rd Umpire Decision (3RDUD)	Likelihood that 3rd umpire reverses (1) or upholds (0) on-field umpire's decision
INDEPENDENT VARIABLES	
On-Field Umpire Decision (OFUD)	Whether on-field umpire gives batsman out (1) or not out (0)
Challenge Team Characteristic (CTC)	Whether the challenging team is home team (1) or away team (0)
Time (TIME)	Values from 1 through 5 for each of the years during which challenges occurred
Time*On-Field Umpire Decision (OFUD*TIME)	An interaction term that is the product of OFUD and TIME
Time*Challenge Team Characteristic (CTC*TIME)	An interaction term that is the product of CTC and TIME
CONTROL VARIABLES	
Umpire ID (UMP ID)	Dummy variables (1 or 0) for 18 umpires relative to K. Dharamasena, the reference umpire
Location (LOC)	Dummy variables for location (1 or 0) relative to Australia, the reference country
Challenge Team ID (CT ID)	Dummy variables (1 or 0) for 8 countries relative to Australia
Bowler Characteristic (BOWC)	Whether bowler is a fast (1) or spin (0) bowler
Batsman Characteristic (BATC)	Whether batsman is a top 7 batsman (1) or bottom 4 batsman (0)
Challenge Team Rank (CTR)	Challenging team's ICC rank (1 through 9) at the time the match was played

Table 8- Logistic Regressions- with controls but without time interaction effects

Variable	**β**	**S.E**	**Wald**	**Sig**	**Exp (β) Odds ratio**	**95% confidence Lower bound**	**95% confidence Upper bound**
OFUD|LBW	0.561	0.164	11.648	0.001	1.752	1.270	2.419
CTC|LBW	0.068	0.166	0.167	0.683	1.070	0.773	1.481
CONSTANT	0.787	0.520	2.289	0.130	2.198		
No. of observations	912						
Nagelkerke R^2	0.062						
Hosmer- Lemeshow	Chi-Square	df	Sig				
	1.769	8	0.987				
Variable	**β**	**S.E**	**Wald**	**Sig**	**Exp (β) Odds ratio**	**95% confidence Lower bound**	**95% confidence Upper bound**
OFUD|CAUGHT	0.749	0.303	6.127	0.013	2.115	1.169	3.828
CTC|CAUGHT	-0.070	0.290	0.058	0.810	0.932	0.528	1.646
CONSTANT	-1.401	0.910	2.370	0.124	0.246		
No. of observations	289						
Nagelkerke R^2	0.187						
Hosmer- Lemeshow	Chi-Square	df	Sig				
	7.580	8	0.476				
Controls: CT ID, UMP ID, BOWC, BATC, LOC, CTR							

Table 9- Logistic Regressions- with controls and with time interaction effects

Variable	**β**	**S.E**	**Wald**	**Sig**	**Exp (β) Odds ratio**	**95% confidence Lower bound**	**95% confidence Upper bound**
OFUD|LBW	1.038	0.347	8.961	0.003	2.824	1.431	5.571
CTC|LBW	-0.011	0.375	0.001	0.997	0.989	0.474	2.065
TIME*OFUD|LBW	-0.592	0.521	1.294	0.255	0.553	0.199	1.535
TIME*CTC|LBW	0.473	0.551	0.737	0.391	1.605	0.545	4.727
TIME	0.292	0.599	0.237	0.627	1.339	0.413	4.334
CONSTANT	0.601	0.540	1.239	0.266	1.825		
No. of observations (LBW)	912						
Nagelkerke R^2	0.079						
Hosmer- Lemeshow	Chi-Square	df	Sig				
	5.000	8	0.758				

Variable	**β**	**S.E**	**Wald**	**Sig**	**Exp (b) Odds ratio**	**95% confidence Lower bound**	**95% confidence Upper bound**
OFUD|CAUGHT	0.735	0.635	1.342	0.247	2.087	0.601	7.243
CTC|CAUGHT	-0.779	0.692	1.266	0.261	0.459	0.118	1.782
TIME*OFUD|CAUGHT	-0.269	1.049	0.066	0.798	0.764	0.098	5.970
TIME*CTC|CAUGHT	-0.972	0.91	1.141	0.285	0.378	0.064	0.251
TIME	-1.351	1.185	1.300	0.254	0.259	0.025	2.641
CONSTANT	-1.26	1.044	1.455	0.228	0.284		
No. of observations (CAUGHT)	289						
Nagelkerke R^2	0.224						
Hosmer- Lemeshow	Chi-Square	df	Sig				
	18.718	8	0.016				

Controls: CT ID, UMP ID, BOWC, BATC, LOC, CTR

[i] See the International Cricket Council (2009), *Umpire Decision Review System- A Guide*, Dubai, UAE

[ii] When the DRS system was introduced in 2009, teams were given a maximum of 2 unsuccessful challenges per innings. In 2013, the ICC modified this rule to allow teams an additional maximum of 2 (unsuccessful) challenges every 80 overs of an inning.

[iii] See, for instance, Brettig (2013), Chappell (2013), Cozier (2009), Gilchrist (2013), Moonda (2012) and Smyth (2012).

[iv] See Nicholas (2015).

[v] See Crowe (2014).

[vi] See Steen (2011).

[vii] For a batsman to be given out LBW, the laws of the game state that the ball: (1) did not pitch outside the leg stump (2) hit the batsman's pad first (before contact with the bat) (3) hit the batsman in line with the wicket and (4) would have gone on to hit the wicket. If the umpire doubts that any one of the 4 conditions identified above are not satisfied, he is obliged to give the batsman "not out."

[viii] See Sancheti et al (2015),

[ix] See Hedzoy (1997).

[x] See, for instance, Findlay (2011).

[xi] See Rajesh (2006).

[xii] The cricinfo text commentary does not identify 3rd umpire decisions that were adjudicated as "umpire's call".

[xiii] Note that our data set only catalogs on-field umpire decisions that are challenged by the playing teams.

[xiv] In many test matches in our sample, hawk-eye was the only technology used to aid the 3rd umpire.

[xv] Image retrieved from http://www.cricket.com.au/news/decision-review-system-up-for-debate-during-australia-india-test-series/2014-12-11.

[xvi] For LBW challenges, hot-spot ascertains whether the ball made contact with the batsman's pad before hitting anything else while for caught challenges, hot-spot ascertains whether the ball made contact with the bat before being caught by a fielder.

[xvii] Image retrieved from http://www.ebuyer.com/blog/2015/02/tech-in-cricket/

[xviii] Image retrieved from http://www.ebuyer.com/blog/2015/02/tech-in-cricket/

[xix] The data set begins with test # 1934 between New Zealand and Pakistan and runs through test # 2150 between South Africa and New Zealand. Test matches in which the DRS was not in use are excluded from the analysis.

[xx] See Steen (2011).

[xxi] The odds of a home challenge being reversed is equal to the probability that a home challenge is reversed divided by the probability that a home challenge is not reversed or (140/567)/(427/567) = .3278. The odds of an away challenge being reversed is equal to the probability that an away challenge is reversed divided by the probability that an away challenge is not reversed (170/634)/(464/634) = .3663. Hence, the *odds ratio* of a home challenge being reversed (relative to an away challenge being reversed) equals .3278/.3663 = .8948. While this suggests that a home-team challenge is 10% *less* likely to be reversed than an away-team challenge, analysis reveals that the odds-ratio is not statistically significant.

[xxii]The odds of an "out" decision being reversed is equal to the probability that an "out" is given "not out" divided by the probability that an "out" is upheld as "out" or (159/510)/(351/510) = .3117/.6882 = .4529. The odds that a "not out" is reversed is equal to the probability that a "not out" is given as "out" divided by the probability than a "not out" is upheld as "not out" or (151/691)/(540/691) = .2185/.7826 = .2791. Hence the odds ratio of an "out" being reversed to "not out" (relative to a "not out" being reversed to an "out") is .4529/.2791 = 1.6227. This tells us that, when LBW and Caught challenges are pooled and when the influence of other variables on the 3rd umpire's decision are ignored, an on-field umpire's decision of "out" is about 62% *more* likely to be reversed (to "not out") than a "not out" is likely to be reversed (to "out"). Because the criteria used by the 3rd umpire to adjudicate LBW and Caught challenges are distinct, the logistic regressions (in the following section) are estimated separately on LBW and Caught challenges.

[xxiii] The odds of an "lbw" challenge being reversed (from "out" to "not out" and from "not out" to "out") is equal to the probability that an "lbw" challenge is reversed divided by the probability that an "lbw" challenge is upheld or (210/912)/(702/912) = .2302/.7697 = .2990. The odds of a "caught" challenge being reversed (from "out" to "not out" and from "not out" to "out") is equal to the probability that a "caught" challenge is reversed divided by the probability than a "caught" challenge is upheld or (100/289)/(189/.289) = .3460/6539 = .5291. Hence, the odds ratio of an LBW reversal (relative to a caught reversal) equals .2990/.5291 = .5651. This tells us that, with *no other* variables influencing the 3rd umpire's decision, an "lbw" challenge is 44% *less* likely to be reversed than a "caught".

[xxiv] For an excellent discussion on interaction effects in regression models, see Ozer-Balli and Sorensen (2013).

[xxv] A value of 1 is assigned to the matches played at the end of 2009.

[xxvi] Since 6 (of the 24) umpires in the data set were involved in less than 12 challenges, we drop them from the model.

[xxvii] There were no test matches held in Zimbabwe in which the DRS was in use.

[xxviii] One piece of evidence that bowlers (and fielding teams) found it very difficult to win an LBW from the on-field umpire (in the pre-DRS era) is that spin bowlers won a total of 4 lbw decisions in the 1975, 1979 and 1983 world cups (when the matches were played in England). In the 2011 world cup (played in India), the proportion of LBWs given by the on-field umpire (and upheld by the 3rd umpire) rose to 16.28% of all decisions confirmed against the batsman. See Findlay (2011).

[xxix] See Coverdale (2013, Dobell (2012) and Steen (2011).

[xxx] See Willey (2009).

[xxxi] See Moskowtiz and Wertheim (2012).

[xxxii] In soccer, a goal is considered to be scored when the whole ball crosses the line between the goal posts and the cross bar. In most cases it is patently clear to everyone whether the ball has crossed the goal-line. However, in a small number of cases, it is difficult for the referee to tell whether a ball has crossed this line before it was saved by the goal keeper or cleared by a defender.

[xxxiii] See Stewart (2013).