ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2015)
178, *Part* 3, *pp.* 741–755

Home bias in officiating: evidence from international cricket

Abhinav Sacheti,

Nottingham University Business School, UK

Ian Gregory-Smith

University of Sheffield, UK

and David Paton

Nottingham University Business School, UK

[Received September 2013. Final revision July 2014]

Summary. We use data on leg before wicket decisions from 1000 test cricket matches to quantify the systematic bias by officials (umpires) to favour home teams. We exploit recent changes in the regulation of test cricket as a series of natural experiments to help to identify whether social pressure from crowds has a causal effect on home bias. Using negative binomial regressions, we find that home umpires favour home teams and that this effect is more pronounced in the later stages of matches.

Keywords: Bias; Cricket; Favouritism; Negative binomial

1. Introduction

In this paper, we explore the effect of social pressure that is exerted by crowds on the decision making of officials (umpires) in sporting contests by using data from test cricket matches. Previous work (e.g. Garicano *et al.* (2005) and Buraimo *et al.* (2010)) has identified a tendency for sporting officials to favour home teams. Test cricket is a particularly interesting laboratory for studying this issue for several reasons. First, cricket umpires can apply a high level of subjective judgement in decisions that can be critical for match outcomes. Second, regulatory changes to the appointment and functioning of umpires over the last two decades provide a series of natural experiments that can help to distinguish the effect of social pressure from other factors such as outright preference from home umpires. Third, test cricket matches are typically played for up to 5 days. The fact that crowds tend to be larger in the first few days of a match provides another potential source of identification of the effect of social pressure on decision making.

We focus on one particularly controversial area of decision making by umpires—that of deciding whether or not a batsman is dismissed ('out') 'leg before wicket' (LBW). Using data on 1000 test matches from 1986 until 2012, we estimate negative binomial regression models to identify the extent and determinants of bias towards the home team. We find strong evidence

Address for correspondence: David Paton, Nottingham University Business School, Jubilee Campus, Nottingham, NG8 1BB, UK.
E-mail: David.Paton@nottingham.ac.uk

 0964–1998/15/178741

that umpires have historically made decisions to the advantage of home teams and that the source of this advantage is more likely to be favouritism rather than social pressure from home crowds. We also find that the introduction of neutral umpires has virtually eliminated this bias.

In Section 2, we discuss previous work on the influence of social pressure on decision making and describe the relevant institutional features of test cricket, drawing out some specific hypotheses. In Section 3 we describe our data. We present our analysis, including our empirical methodology, in Section 4 and make some concluding comments in Section 5.

2. Social pressure in test cricket

Test match cricket involves two national teams, where the fielding team is trying to 'bowl' the batting team 'out'. Batsmen can be dismissed in several ways, but the relevant dismissal here is LBW. If a bowler from the fielding team delivers a ball which strikes a batsman on the leg, the fielding team can appeal for an LBW decision to one of the two on-field umpires, requiring him to judge whether or not the batsman should be given out (see Marylebone Cricket Club (2010) for the current rules underpinning the LBW decision). The umpire must take into account several specific regulations in making his decision but (at least until recent years) has complete discretion over the decision which, once made, is considered final. The LBW dismissal is very common in cricket, with more than four such dismissals in every match, on average.

Given the level of discretion that is afforded to umpires and also the intense time pressure on them (umpires must make their decision within a few seconds), there is clear potential for LBW decisions to favour one team in a systematic way. Indeed, Ringrose (2006) suggested that the difficulty of making LBW decisions lends them to conspiracy theories and claims of bias. Chedzoy (1997) statistically examined the effect of umpiring decisions on the performances of test cricket batsmen and teams, showing that they can alter the match outcome.

A striking feature of test cricket is that, until 1994, both on-field umpires had the same nationality as the home team. In 1994, cricket's governing body, the International Cricket Council (ICC), ruled that one neutral umpire should stand in every test match. Later, in 2002, the ICC ruled that henceforward both umpires would be neutral, chosen from the Elite Panel of umpires. Both of these changes were implemented in these years. From this point onwards, the best performing umpires at lower levels of cricket have been able to receive promotion to the Elite Panel, whereas poorly performing Elite Panel umpires can be demoted: for example, in 2004 three of the original eight Elite Panel umpires did not have their contracts renewed. This shift from fully home to fully neutral officials in test cricket is rare among major team sports (see Andreff and Szymanski (2006) for a review of major team sports). In the early years of our sample, test umpires were appointed by each country's board. This hiring process has gradually been centralized towards the ICC, with umpires on the Elite Panel now reportedly paid an annual salary of around £200000 plus expenses (Hoult, 2013). Bryson *et al.* (2011) observed that a change in payment system affects referees' behaviour and performance in professional football, raising the possibility that the change in terms and conditions for umpires following the introduction of the Elite Panel may have affected umpires' behaviour and performance in test cricket.

Contemporaneously with these regulatory developments, test umpiring has also undergone notable technological changes. From the early 2000s, television broadcasters of test matches have used computer-based technology that enables a retrospective judgement to be made on whether an umpire's LBW decision was correct. Recently, this technology has been included in the actual decision-making process: in 2008, the ICC began trialling a technological referral system called the 'decision review system' (DRS) and then officially introduced this system in late 2009. The system works by allowing both the batting and the bowling teams to challenge umpires' decisions. Teams are limited to no more than two unsuccessful challenges per innings but may make any number of successful challenges. At the time of writing, the DRS scheme can be used in any particular match if both teams agree.

Umpires are tasked with making decisions on a pair of teams, of which the home side typically attracts larger live crowds. Certainly, umpires gain utility from making correct decisions. Aside from the inherent pleasure of making a correct decision and the improvement to their professional reputations, the more decisions umpires make correctly, the more likely they are to be retained by the ICC. Nevertheless, umpires may be constrained by social pressure to favour the home team, particularly where the 'correct' decision is ambiguous, as frequently happens with LBW decisions.

Indeed, the limited existing literature provides some support for this hypothesis. Looking at test matches played between 1877 and 1980, Sumner and Mobley (1981) noted that home teams suffered significantly fewer adverse LBW decisions than away teams in Australia, India and Pakistan. Crowe and Middeldorp (1996) compared LBW rates for the first six batsmen in a batting line-up in test matches played in Australia from 1977 to 1994. Using logistic regression models, they found that three of seven opposition teams were given out LBW more frequently than the Australian team, but it was unclear whether this was because of umpiring bias. They suggested that differentials in LBW rates between the Australian team and its opponents could be explained by differences in playing styles.

More recently, Ringrose (2006) examined test matches played between 1978 and 2004 for evidence of favouritism towards home teams, finding that location and batting team were significant: home teams suffered fewer adverse LBW decisions than away teams. However, he did not find significant evidence that the introduction of neutral umpires reduced the home bias. We build on this result in two ways. First, in Ringrose (2006) data were available for only the very early years of neutral umpires. We now have the luxury of being able to observe a considerably longer period during which only neutral umpires have been used. Second, we adopt an alternative estimation strategy to that of Ringrose (2006), who measured LBW dismissals as a rate relative to the total number of dismissals. Such an approach may understate the true level of home team bias as the fewer LBW decisions awarded by a biased umpire will also reduce the total number of dismissals. Instead we adopt a count data approach which avoids this problem. This is because, with each ball that is bowled in an innings, there is the possibility of a dismissal caused by LBW. Using the total number of dismissals in an innings as the exposure variable is problematic because of the endogeneity between total dismissals and LBW dismissals. If there is indeed home team bias such that home teams are given out LBW less frequently than away teams then, other things being equal, the value of total dismissals will be lower for home teams (on average) than for away teams. An alternative exposure variable would be the number of LBW appeals, but these data are not available.

Drawing on Akerlof (1997), Dawson and Dobson (2010) explained that utility for decision makers can be context specific. In the case of test cricket, utility gained by umpires may differ depending on whether or not they are umpiring a match involving their own country's team. We propose that home umpires may

(a) prefer the team that shares their own nationality, either consciously or unconsciously and
(b) be more susceptible to pressure from home crowds.

This leads to an interesting question which has not been explored in the existing empirical literature. What is the origin of the observed preference by home umpires for home teams? Is it inherent favouritism towards their own nation or is it the home crowd exerting pressure on the home umpire? To distinguish between these types of bias, we exploit the fact that test matches are played over a period of up to 5 days and that, consequently, the size of the crowd (and by proxy the extent of social pressure on umpires) varies with the stage of the match. Given that crowds tend to be very much bigger in the early stages of a test match (Hynds and Smith, 1994) social pressure from crowds should be greater in the early innings of a match relative to the later innings.

For this reason, in our empirical work, we estimate home team bias in LBW decisions both by the type of umpires (home, neutral or a mixture) and by the stage of the match (the first or second innings compared with the third or fourth innings). This allows us to identify the origin of the home bias and to distinguish between favouritism and social pressure.

3. Data

We use data from all test matches played between January 1986 and July 2012, excluding two matches that were abandoned without a ball bowled; 16 matches played at neutral venues and one match played between Australia and a 'World XI'. This leaves 1000 matches and 3601 innings, with each innings treated as one observation. Of these 1000 test matches, 206 were played with two home umpires, 348 with one home umpire and one neutral umpire and 446 with two neutral umpires. The data were collected by the authors in their entirety from the ESPNcricinfo Web site `http://www.cricinfo.com`. They and the programs that were used to analyse them can be obtained from

`http://wileyonlinelibrary.com/journal/rss-datasets`

136 umpires took to the field in the sample period, with the combined experience of the umpires in each test match ranging between two and 201 tests. The average combined umpire experience per match in the sample is around 53 test matches.

In Table 1, we report the number of LBW decisions per innings for home and away teams, both for all batsmen and for just the top seven batsmen. The mean rate of LBW decisions is lower for home teams, both for all batsmen and for the top seven batsmen in the team. Standard t-tests suggest that these differences are statistically significant at conventional levels.

In Table 1, we provide the breakdown of LBW decisions by umpire neutrality, presence of the DRS and host nation. When both umpires are from the same country as the home team, the home team received 16% fewer adverse LBW decisions per innings than the away team. With the introduction of one neutral umpire, this advantage decreases to 9.9% and with two neutral umpires the bias in favour of home teams is virtually eliminated (0.7%). One confounding factor could be that home teams receive fewer adverse LBW decisions because of superiority in familiar conditions. If so, it is possible that the introduction of two neutral umpires may have led to a bias towards away teams should neutral umpires feel pressure to favour away teams. We control for this in the formal analysis in Section 4.

We also report, in Table 1, the rate of LBW decisions with and without the DRS (with the latter sample restricted to matches with neutral umpires). Contrary to received wisdom at the time of writing, the mean LBW values have *fallen* for both home and away teams with the introduction of the DRS. In terms of different stages of the match, the second innings produces the most LBW decisions per innings, whereas the fourth innings produces the least. The latter finding is unsurprising as fourth innings of test matches are, on average, shorter than the others. Although there is evidence of bias towards home teams in each innings, there is no clear trend in the bias between earlier and later stages of matches.

Looking at different venues for Test matches, the most LBW decisions per innings are given in Pakistan, followed by Bangladesh, India, Sri Lanka and the West Indies. The least LBW

Table 1. LBW decisions per innings for home and away teams†

	Results for home team	*N*	*Results for away team*	*N*	*Difference*	*Results for all teams*	*N*
All innings	1.37 (0.03)	1771	1.48 (0.03)	1830	0.11	1.42 (0.02)	3601
All innings, top 7	1.02 (0.02)	1771	1.09 (0.02)	1830	0.07	1.06 (0.02)	3601
Both umpires home	1.32 (0.07)	360	1.57 (0.07)	360	0.25	1.45 (0.05)	720
One neutral umpire	1.28 (0.05)	620	1.42 (0.05)	633	0.14	1.36 (0.04)	1253
Both umpires neutral	1.46 (0.05)	791	1.47 (0.04)	837	0.01	1.46 (0.03)	1628
No DRS	1.47 (0.05)	654	1.48 (0.05)	697	0.01	1.48 (0.03)	1351
With DRS	1.41 (0.12)	134	1.41 (0.11)	143	0.00	1.41 (0.08)	277
First innings	1.33 (0.05)	506	1.43 (0.05)	494	0.10	1.38 (0.04)	1000
Second innings	1.53 (0.07)	490	1.57 (0.05)	498	0.04	1.56 (0.04)	988
Third innings	1.41 (0.06)	455	1.62 (0.06)	503	0.21	1.52 (0.04)	958
Fourth innings	1.13 (0.07)	317	1.18 (0.07)	338	0.05	1.15 (0.05)	655
Australia	0.95 (0.06)	260	1.48 (0.07)	281	0.53	1.23 (0.05)	541
England	1.43 (0.07)	299	1.42 (0.07)	300	−0.01	1.42 (0.05)	599
South Africa	1.23 (0.09)	170	1.15 (0.08)	194	−0.08	1.18 (0.06)	364
New Zealand	1.45 (0.10)	189	1.19 (0.08)	182	−0.26	1.32 (0.07)	371
West Indies	1.77 (0.11)	207	1.27 (0.08)	217	−0.50	1.51 (0.07)	424
India	1.30 (0.10)	178	1.76 (0.09)	195	0.46	1.54 (0.07)	373
Pakistan	1.54 (0.11)	134	2.14 (0.13)	143	0.60	1.85 (0.09)	277
Sri Lanka	1.15 (0.09)	169	1.90 (0.11)	174	0.75	1.53 (0.07)	343
Zimbabwe	1.60 (0.13)	89	1.03 (0.12)	87	−0.47	1.32 (0.09)	176
Bangladesh	1.74 (0.15)	72	1.34 (0.17)	61	−0.40	1.56 (0.11)	133

†Standard errors are given in parentheses. *N* is the number of innings in each sample. The sample of games with no DRS is restricted to matches with both umpires neutral. 'Difference' is the away team average minus the home team average.

decisions per innings are given in South Africa, followed by Australia. As Ringrose (2006) mentioned, the slower nature of pitches in Asian countries tends to mean that the ball bounces less high and makes LBW decisions more likely, in contrast with quicker and bouncier pitches in Australia and South Africa. The evidence of home team bias varies quite considerably across countries. In the West Indies, New Zealand, Zimbabwe and Bangladesh, the mean number of LBW decisions is notably higher for home teams, whereas in Australia, India, Pakistan and Sri Lanka the reverse is true. In England and South Africa the mean values are very similar. We shall control for this in the subsequent analysis.

In Table 2, we summarize 'caught-behind' decisions in our full sample of matches. Such decisions (meaning dismissals caused by the batsman edging the ball to the wicketkeeper behind the stumps) frequently involve an element of judgement by the umpire and, as such, may also be subject to home team preference. However, in contrast with LBW decisions, some caught-behind decisions do not require adjudication by the umpire and we cannot distinguish in the data which decisions require a formal decision. For this reason, differences in LBW decisions provide a cleaner measure of home team advantage. Nonetheless, we observe a clear pattern in the caught-behind decision data, and one which is consistent with the effect of neutral umpires on LBW decisions. With two home umpires, away teams suffer 5.8% more caught-behind decisions than home teams. When the one neutral umpire policy is introduced, this advantage to home teams declines to 3.8%. When two neutral umpires are present, the home team's advantage declines yet further to 1.5%.

To summarize, the descriptive statistics are consistent with home teams being favoured by umpires and with this bias being reduced by the presence of neutral umpires. There is also

Table 2. Caught-behind decisions per innings for home and away teams†

	Results for home team	*N*	*Results for away team*	*N*	*Difference*	*Results for all teams*	*N*
All innings	1.54 (0.03)	1771	1.59 (0.03)	1830	0.05	1.57 (0.02)	3601
Both umpires home	1.47 (0.07)	360	1.56 (0.07)	360	0.09	1.45 (0.05)	720
One neutral umpire	1.55 (0.05)	620	1.60 (0.05)	633	0.05	1.36 (0.04)	1253
Both umpires neutral	1.57 (0.04)	791	1.59 (0.04)	837	0.02	1.58 (0.03)	1628

†Standard errors are given in parentheses. *N* is the number of innings in each sample. 'Difference' is the away team average minus the home team average.

a priori evidence that the bias towards home teams varies considerably by the country in which the test match is being played. However, these differences may be explained by the relative strengths of particular countries and so we now go on to use negative binomial regression analysis to see whether differences that are suggested in the descriptive statistics are robust to the inclusion of our control variables.

4. Analysis

As our aim is to examine how LBW decision making in test matches is affected by the introduction of neutral umpires, we use the number of LBW decisions in an innings as the dependent variable. This is a form of count data, so we first consider using a Poisson regression model for our empirical analysis rather than an ordinary least squares model. This is because we cannot reasonably assume that the dependent variable and the error terms are normally distributed as the count of LBW decisions is a whole number bounded between 0 and 10. In our sample, the variance of the dependent variable is slightly higher than its mean, suggesting that the data vary more than might be expected in a Poisson distribution. Where the data are overdispersed in this way, Koop (2008) suggests that the most common alternative is the use of a negative binomial regression model, which allows the mean to differ from the variance. In the Poisson model, overdispersion can lead to very small standard errors, causing false positive results. Although formal tests do not suggest that the overdispersion in our data is statistically significant, we proceed with a negative binomial model on the grounds of generality although, in fact, our conclusions are unaltered by the use of a Poisson model.

The overdispersion in the binominal model is captured by a quadratic variance function which includes the unknown parameter α:

$$V(y|\mu,\alpha)=\mu(1+\alpha\mu). \tag{1}$$

From Cameron and Trivedi (2005), pages 675–676, the negative binominal maximizes the log-likelihood of the probability mass function

$$\Pr(Y=y|\mu,\alpha)=\frac{\Gamma(\alpha^{-1}+y)}{\Gamma(\alpha^{-1})\,\Gamma(y+1)}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu}\right)^{\alpha^{-1}}\left(\frac{\mu}{\mu+\alpha^{-1}}\right)^{y} \tag{2}$$

where μ is parameterized as $\exp(X'\beta)$, as in the standard parameterization of the Poisson model, and α is the variance parameter of the gamma distribution, which is a constant. The treatment of the variance parameter as being gamma distributed is convenient as it allows the analytical expression that is given in equation (2). This derives from the usual 'mixture interpretation' that is given to the overdispersion, meaning that it bases the overdispersion as arising from unobserved gamma heterogeneity in a Poisson process. As $\alpha \to 0$, the model reduces to the Poisson model. Whereas the overdispersion parameter α enters equation (2) and hence affects the probability distribution, the expectation of the conditional mean is the same as it is in the Poisson model:

$$E[y|\mathbf{X}] = \exp(\mathbf{X}'\boldsymbol{\beta}). \tag{3}$$

In our model, y is the LBW count, $\mathbf{X}$ is the vector of explanatory variables and $\boldsymbol{\beta}$ are the corresponding coefficients. The interpretation of the regression coefficients in a negative binomial model remains the same as for a Poisson regression. The marginal effect of the coefficients on the number of LBW decisions is given by

$$\frac{\mathrm{d}E[Y_i]}{\mathrm{d}X_i} = \beta \exp(\beta X_i). \tag{4}$$

As this is a non-linear model, the value of the marginal effects will vary with different values of the covariates. When we calculate the marginal effects we report the 'average marginal effect' as in Cameron and Trivedi (2010), page 576. This is the average of the marginal effects calculated at the sample values. Since we do not have a specific value of the covariates in mind when we are calculating the marginal effect, we prefer the average marginal effect.

To determine whether bias towards home teams occurs in test cricket, we are interested in how the explanatory variable 'Home batting' impacts the LBW count. Home batting is a dummy variable taking the value 1 if the batting side is playing at home. We estimate equation (3) for the whole sample and then separately for three subsamples: test matches with two home umpires, test matches with one home and one neutral umpire and test matches with two neutral umpires. Our hypothesis is that a significantly negative estimate of the coefficient on Home batting would provide evidence of systematic bias towards home teams. To the extent that home umpires are more likely than neutral umpires to favour home teams, the Home batting coefficient should be highest in magnitude (more negative) when there are two home umpires and lowest (less negative) when there are two neutral umpires.

We consider the top seven batsmen in an innings separately, as social pressure on umpires may differ when making decisions against the specialist batsmen (including all-rounders) compared with the less skilled lower order batsmen. In the full sample estimates, we also include period dummy variables based on the neutral umpire rule for the period. These period dummy variables control for changes in the character of test match pitches over time and the effect of the introduction of HawkEye ball tracking technology to television broadcasts from 2001. Given the possibility of within-match correlation (as up to four innings can be played in one match, with both teams batting up to two times), we cluster the standard errors by match. We control for the fact that the number of LBW decisions in each innings will be affected by various factors. Some innings are curtailed prematurely because of the end of the match or because of lack of time, leading to the batting captain's voluntarily 'declaring' the innings closed. As noted above, using the total number of dismissals to control for this is inappropriate as the total number of wickets falling will be directly correlated with any home bias. Hence, we include the logarithm of the number of overs bowled in each innings ('ln Overs') as our exposure variable. This does not entirely solve this issue. If wickets are less likely to fall when the home side is batting, the innings is more likely to be extended and the number of overs may be greater.

As a check of robustness, we control for the length of the innings with the ‘Over dummies’ variables.

We also include the combined umpire experience in the match (‘Umpire Experience’) and a dummy variable indicating whether the DRS was in use during the match (‘DRS’). In some matches, the DRS has not included ball tracking technology that is necessary for LBW referrals. These matches are treated as not having the DRS in place for the purposes of this study.

There may be systematic differences in the number of LBW decisions across different innings of a match and so we include dummy variables for three of the four innings (‘Innings2 dummy’, ‘Innings3 dummy’ and ‘Innings4 dummy’). Similarly, as discussed, LBW decisions may be relatively more common in certain countries owing to systematic differences in pitches. For example, pitches in Australia and South Africa are traditionally bouncier than in England, making it more likely that a ball will bounce over the stumps and so less likely for an LBW decision to be given against the batsman. These impacts will be even more apparent by pitch: for example, venues within countries have different characteristics. For this reason, one of our checks of robustness below involves including dummy variables for the venue at which the game is being played (‘Lord’s’, ‘Melbourne Cricket Ground’ etc.).

Finally, to control for the relative strengths of the teams, we also include two sets of team dummy variables: one for the batting side and one for the bowling side in each innings. We provide a full description of the variables in Table 3.

We report the estimates for the whole sample and then separately for the first two innings of a match and for the final two innings of a match. As argued above, crowds are generally much larger for the first and second innings of a test match. Hence, if the evidence of home team bias is relatively stronger for the first two innings, this would be consistent with crowd pressure being the source of the bias.

Table 3. Variables and definitions

Variable	*Definition*
LBW	Number of LBW decisions in the innings
Top 7 LBW	Number of LBW decisions against the first 7 batsmen in the innings
Home batting	1 if the batting team was playing at home; 0 otherwise
Away batting	1 if the batting team was playing away from home; 0 otherwise
Umpire Experience	Combined number of matches officiated by the two umpires
DRS	1 if the match had the DRS in place; 0 otherwise
Time	Number of the match in the sample
Innings1 dummy	1 if first match innings; 0 otherwise
Innings2 dummy	1 if second match innings; 0 otherwise
Innings3 dummy	1 if third match innings; 0 otherwise
Innings4 dummy	1 if fourth match innings; 0 otherwise
Runs	Number of runs scored in the innings
Overs	Number of overs bowled in the innings
ln Overs	Logarithm of the number of overs bowled in the innings
Over dummy variables	Dummy variables for innings length in terms of number of overs bowled
Country (host) dummy variables	Dummy variables for the country where the match is played
Ground dummy variables	Dummy variables for the venue of the match
Umpire period dummy variables	Period dummy variables for two home umpires, one neutral and one home umpire and two neutral umpires

We report the negative binomial regression estimates by using the full sample in Table 4. For all four models, we report a likelihood ratio test for overdispersion in the data. Use of Poisson regression (which is not reported here but is available on request) leads to very similar results to those obtained with the negative binomial model.

We first estimate the model on all matches. Consistent with home bias, the coefficient on the key variable of interest, Home batting, is negative and statistically significant, suggesting that over the whole sample home batsmen are given out less often than away batsmen. The marginal effect is -0.128. At the mean value of 1.37 LBW decisions per innings, this represents an approximately 10% decrease in the number of LBW decisions per innings. The DRS does not appear to have a significant independent effect on the number of LBW decisions, but we do observe marginally fewer LBW decisions as umpire experience increases. We also observe relatively more LBW decisions in the second, third and fourth innings of tests compared with the first innings.

Although we control for differences in average team abilities, some of the home advantage might be attributed simply to the home team being better able to avoid LBW decisions, rather than poor decision making by umpires. Put another way, teams may simply play better (and lose fewer wickets) at home than away. Therefore our main identification strategy is to estimate the model separately for test matches with two home umpires, test matches with one home and one neutral umpire and test matches with two neutral umpires. Home advantage due to other factors such as greater familiarity of the pitch should, on average, be constant across the three groups of tests. Any advantage to home teams in LBW decisions that are made by umpires between the groups can be attributed to home bias. These results are reported in columns (2)–(4) of Table 4.

There is a very clear pattern. The magnitude of home advantage is biggest when there are two home umpires and smallest when there are two neutral umpires. The marginal effect when there are two home umpires is -0.284, implying a decrease of approximately 21 percentage points in the number of LBW decisions per innings given against home teams. This is roughly equivalent to one extra LBW decision in favour of the home team in every innings, which is certainly enough to have a major influence on the outcome of the match. The effect is halved (-0.134) when there is one neutral umpire and reduced again when both umpires are neutral, to the point of statistical insignificance. To check that columns (2), (3) and (4) are not nested within column (1), we constructed a set of interactions between the covariates in column (1) and the dummy variables for one home umpire and two home umpires respectively, and then performed a joint test of significance on both sets of interaction terms. This is akin to a Chow test. In both cases, the set of interactions was significant (one home umpire, $\chi^2(30) = 61.42$, Prob $> \chi^2 = 0.0006$; two home umpires, $\chi^2(33) = 48.06$, Prob $> \chi^2 = 0.0438$), suggesting that it is appropriate to split column (1) into columns (2)–(4).

So we find strong evidence that the introduction of neutral umpires reduced bias in favour of home teams. Next, we seek to distinguish between the types of bias: does it arise from social pressure or inherent favouritism towards the home team? To do this, we make use of our second source of identification, namely the variation in crowd attendance over the four innings in the test match. To explore the source of this bias, we next estimate the model separately for the first two and the final two innings. If the team's best batsman is incorrectly given out on the first day of the match, this will affect the team's first-innings total and will impact the outcome. However, the influence of decisions early on is somewhat indirect as there are potentially three more innings left in the match for the team to recover from a flawed decision, lessening the effect that a biased umpire can have on the match's outcome. By contrast, in the final stages of the match (i.e. the final two innings), a similarly flawed decision against the team's best batsman will have a direct effect on the match (to give one example, a biased decision to give the team's best

Table 4. Negative binomial model of number of LBW decisions per innings†

Variable	*(1) All innings*	*(2) Innings with 2 home umpires*	*(3) Innings with 1 neutral umpire*	*(4) Innings with 2 neutral umpires*
Home batting	−0.091	−0.196	−0.099	−0.057
	(0.030)‡	(0.069)‡	(0.049)§	(0.044)
DRS	−0.008			0.005
	(0.078)			(0.077)
Umpire Experience	−0.013	−0.002	−0.003	−0.001
	(0.0006)§	(0.003)	(0.001)§	(0.0006)
ln Overs	0.256	0.378	0.287	0.208
	(0.028)‡	(0.064)‡	(0.048)‡	(0.040)‡
Innings2 dummy	0.126	0.139	0.156	0.092
	(0.036)‡	(0.084)§§	(0.058)‡	(0.054)§§
Innings3 dummy	0.138	0.182	0.224	0.054
	(0.036)‡	(0.078)§	(0.06)‡	(0.052)
Innings4 dummy	0.022	0.130	0.111	−0.08
	(0.053)	(0.125)	(0.085)	(0.075)
‘One neutral umpire’ period dummy	−0.018			
	(0.050)			
‘Two neutral umpires’ period dummy	0.136			
	(0.056)§			
Bangladesh (host)	0.407		0.866	0.437
	(0.133)‡		(0.39)§	(0.155)‡
England (host)	0.245	0.208	0.258	0.272
	(0.07)‡	(0.135)	(0.144)§§	(0.105)‡
India (host)	0.331	0.330	0.459	0.260
	(0.082)‡	(0.184)§§	(0.174)‡	(0.107)§
New Zealand (host)	0.120	0.198	−0.135	0.287
	(0.087)	(0.174)	(0.155)	(0.136)§
Pakistan (host)	0.386	0.425	0.386	0.342
	(0.087)‡	(0.175)§	(0.149)‡	(0.136)§
South Africa (host)	0.165	−0.231	0.147	0.192
	(0.089)§§	(0.447)	(0.143)	(0.127)
Sri Lanka (host)	0.257	−0.130	0.282	0.366
	(0.091)‡	(0.212)	(0.155)§§	(0.132)‡
West Indies (host)	0.309	0.449	0.112	0.364
	(0.086)‡	(0.152)‡	(0.174)	(0.124)‡
Zimbabwe (host)	0.060	−0.99	−0.066	0.530
	(0.119)	(0.547)§§	(0.171)	(0.181)‡
Constant	−0.892	−1.449	−1.008	−0.583
	(0.156)‡	(0.357)	(0.275)‡	(0.231)

(*continued*)

batsman out when the team is chasing a target to win the match in the final innings will very directly impact the final outcome). In these cases, the effect of the umpire’s decision on match outcome is more direct and we argue that any home favouritism from umpires is likely to be strongest during these stages of the match for this reason. These estimates are reported in the first two rows of Table 5. Coefficients on the control variables are suppressed here for brevity.

In both cases, we see the same pattern as for the whole sample. The coefficient on Home batting is negative and largest when there are two home umpires and smallest when there are two neutral umpires. However, for the first two innings, the coefficient on Home batting is never significant and is much smaller than for the final two innings. Given that crowds tend to be much larger in early stages of test matches, it is difficult to attribute the bias towards home teams that

Table 4 (*continued*)

Variable	*(1) All innings*	*(2) Innings with 2 home umpires*	*(3) Innings with 1 neutral umpire*	*(4) Innings with 2 neutral umpires*
Batting team effects included	Yes	Yes	Yes	Yes
Bowling team effects included	Yes	Yes	Yes	Yes
Number of innings	3601	720	1257	1624
Log-pseudolikelihood	−5392.1	−1073.3	−1817.9	−2440.7
Likelihood ratio test	0.121	0.374	0.496	0.472
Home marginal effect	−0.128‡	−0.284‡	−0.134§	−0.083

†Robust standard errors are given in parentheses, clustered by match. There is no estimate for Bangladesh (host) in the estimates with two home umpires because no test matches were played with two home umpires in Bangladesh. ‘Home marginal effect’ is calculated as the average marginal effect as used by Cameron and Trivedi (2010), page 576. Each marginal effect is an average of the marginal effects calculated at the sample values. ‘Batting team effects’ are a control for the batting side in each innings, and ‘bowling team effects’ similarly control for the bowling side in each innings. These are included to control for batting- and bowling-team-specific effects. The ‘likelihood ratio test’ is the p-value from the likelihood ratio χ^2-test that the dispersion parameter α is equal to 0. The values do not reject the null hypothesis of no overdispersion but we nevertheless proceed with the negative binomial over the Poisson model as a matter of caution. The results by using Poisson estimation are qualitatively identical to those above and are available on request.
‡Significant at the 1% level.
§Significant at the 5% level.
§§Significant at the 10% level.

is displayed by home umpires as a response to crowd pressure. Instead it appears that home umpires give more LBW decisions in favour of the home team in the later stages of the match, where the decision is more likely to sway the outcome of the game.

5. Robustness checks

We next consider several alternative specifications to explore how robust these results are. These are as follows.

(a) Inclusion of country-specific time trends to allow for changes in LBW rates in different countries over time: an example of this might be pitches in a particular country becoming generally more (or less) bouncy over time and, as a result, LBW decisions becoming less (or more) common.
(b) Inclusion, also, of time trends that are specific to each batting and bowling side: this controls for systematic changes in abilities of particular teams over time.
(c) Restricting LBW decisions against the top seven batsmen only: generally, decisions relating to higher order batsmen are more critical (and, hence, potentially more susceptible to social pressure) than for lower order (or ‘tail end’) batsmen (see Crowe and Middeldorp (1996)).
(d) Inclusion of match fixed effects: in this specification, we include fixed effects for every match. This is a much more restrictive specification which allows for the possibility that match-specific factors (e.g. atmospheric conditions) may affect the number of LBW decisions. In this case, the estimates of the coefficient on Home batting are driven solely by within-match variation.
(e) Inclusion of over dummy variables: these are included as an alternative approach to controlling for the effect of innings length on LBW decisions, particularly shorter innings.

Table 5. Estimates of home bias in LBW decisions: alternative specifications†

	Results for all innings	*Results for innings with 2 home umpires*	*Results for innings with 1 neutral umpire*	*Results for innings with 2 neutral umpires*
Innings 1 and 2	−0.054	−0.121	−0.078	−0.031
	(0.035)	(0.081)	(0.059)	(0.052)
Innings 3 and 4	−0.116	−0.336	−0.102	−0.060
	(0.048)	(0.100)‡	(0.079)	(0.073)
Country-specific trends	−0.092	−0.195	−0.101	−0.056
	(0.030)‡	(0.070)‡	(0.048)§	(0.044)
Country, batting and bowling team trends	−0.098	−0.212	−0.095	−0.061
	(0.030)‡	(0.069)‡	(0.048)§	(0.044)
Top 7 batsmen only	−0.083	−0.163	0.102	−0.061
	(0.033)§	(0.078)§	(0.054)§§	(0.048)
Match fixed effects	−0.091	−0.192	−0.099	−0.063
	(0.029)‡	(0.066)‡	(0.051)§§	(0.044)
Over dummy variables	−0.076	−0.170	−0.078	−0.049
	(0.029)‡	(0.066)‡	(0.048)	(0.04)
LBW/100 overs	−0.21	−0.592	−0.136	−0.149
	(0.086)§§	(0.157)‡	(0.154)	(0.114)
LBW/wickets	−0.007	−0.023	−0.009	−0.002
	(0.005)	(0.012)	(0.008)	(0.008)
OLS	−0.127	−0.278	−0.148	−0.062
	(0.042)‡	(0.101)‡	(0.063)§	(0.062)
Ground dummy variables	−0.078	−0.168	−0.116	−0.017
	(0.031)§	(0.072)	(0.051)§§	(0.046)

†Robust standard errors are given in parentheses, clustered by match. All estimates are based on negative binomial regressions, with the exception of LBW/100 overs, LBW/wickets and OLS, all three of which are ordinary least squares regression estimates. All estimates include the control variables that are listed in Table 4. These coefficients have been suppressed for brevity. All estimates except those in the 'Match fixed effects' row include country fixed effects and team fixed effects for the batting and bowling sides. For reasons of multicollinearity, the 'Match fixed effects' row includes team fixed effects for the batting side only. In the 'Ground dummy variables' row, team batting and bowling effects are not included owing to multicollinearity with ground dummy variables.
‡Significant at the 1% level.
§Significant at the 5% level.
§§Significant at the 10% level.

We include dummy variables for five bands of overs: 1–25 overs; 26–50 overs; 51–100 overs; 101–150 overs; more than 150 overs.

(f) We estimate the model by using an alternative specification for the dependent variable, namely the number of LBW decisions per 100 overs bowled. In this case we use an ordinary least squares rather than negative binomial regression.
(g) To enable a more direct comparison with Ringrose (2006), we estimate the model by using the proportion of LBW decisions to wickets in the innings as the dependent variable.
(h) We estimate the baseline model by using an ordinary least squares regression.
(i) To control for the effect of venue on LBW decisions, we use ground dummy variables rather than country dummy variables. Owing to multicollinearity, we exclude team batting and bowling effects from this regression.

The results of the alternative specifications are generally consistent with the baseline model. In every case except (g), the coefficient on Home batting is significantly negative when there are two home umpires but insignificant (at least at the 10% level) when there are two neutral umpires, whereas the coefficient is larger in size when there is one neutral umpire than when there are two.

The result from (g), using the proportion of LBW decisions to wickets as the dependent variable, similarly to Ringrose (2006), differs from our other findings. In this case, the coefficient on Home batting is reduced in magnitude and is no longer significant. However, the presence of neutral umpires still reduces the magnitude of the coefficient even further. The insignificance of the coefficients when including the number of wickets can be explained by the fact that the number of LBW decisions is constrained by the number of wickets, so the number of wickets is likely to have a stronger influence on the number of LBW decisions than do other factors. The use of ln Overs as the exposure variable partially overcomes this difficulty, though decision making by a biased umpire is still likely to alter the number of overs in an innings.

We also explored the effect of crowd size on LBW decisions in a more indirect way by considering LBW decisions in the Boxing Day and New Year's test matches in Australia. These two test matches, which are traditionally held at Melbourne and Sydney respectively, are extremely high profile events in the Australian cricket calendar and tend to attract particularly large crowds. We found that the coefficient on Home batting is larger (meaning that it is more negative) for these two test matches compared with other test matches in Australia.

Finally, we consider the outcomes of DRS referrals by using data collected from on-line text commentary for the 71 matches in our sample in which the DRS system was in place. All these matches had two neutral umpires, so we cannot use these data to identify any effect of favouritism by home umpires. However, differences in the proportion of decisions going against home and away teams on referral may provide an indicator of any favouritism by neutral umpires towards home teams. Looking at all DRS LBW referrals, 222 out of 389 decisions (57.1%) went against the batting team. 112 out of 199 decisions (56.3%) went against away teams, and 110 out of 190 decisions (57.9%) went against home teams. The difference between home and away proportions is not statistically significant at the 5% level, which is consistent with our main finding that neutral umpires do not display bias.

In summary, we find strong evidence that the introduction of neutral umpires significantly affected the rate of LBW decisions going in favour of home teams. This finding is generally robust to the inclusion of country and team fixed effects and time trends over the sample period.

Further, our finding that home umpires favour home teams more in the final two innings than in the first two innings provides some insights into the source of the home bias in LBW decisions. Specifically our results are more consistent with home bias arising from favouritism (whether conscious or not) by home umpires for home teams than by crowd pressure.

So how do our results compare with the growing literature on referee and umpire bias in professional sport? North American literature (e.g. Price and Wolfers (2010) on the National Basketball Association and Parsons *et al.* (2011) on Major League Baseball) has tended to focus on ethnicity bias, which is a type of bias that we do not consider in our study. European studies such as Dawson *et al.* (2007) and Buraimo *et al.* (2010) have looked more closely at biased decisions towards home teams and social pressure as proxied by crowd attendance. Our results contribute towards this second strand of the literature by indirectly assessing the effect of social pressure on decision making by using the stage of the match as a proxy for crowd attendance. Our finding that neutral umpiring reduces the extent of home bias is in contrast with that of Buraimo *et al.* (2012), who found that home team bias persists in Champions League football matches even with neutral referees. Buraimo *et al.* (2012), page 331, noted that the appointment of neutral match officials in the Champions League is a '... deliberate attempt to combat potential bias' but found that bias persists regardless. These contrasting results are most probably explained by crowd pressure acting in different ways on officials in cricket and football. For example, referees in football make decisions on more physical and perhaps more emotive events than do cricket umpires. Football crowds are often abusive towards referees and on occasion football referees have been physically assaulted. Cricket umpires are motionless and stand at a fixed distance of 22 yards from the 'on-strike' batsman, in the middle of the field and away from the crowd. In contrast, football referees are typically on the move and can be closer to the crowd. Distinguishing between these factors is left for future work.

6. Conclusions

Our main finding is that home bias has influenced umpires in test matches in the last two and a half decades but has receded with the introduction of neutral umpires. With two home umpires, home teams received significantly fewer adverse LBW decisions than did away teams. This advantage significantly declined with the introduction of the one-neutral-umpire policy and has declined further with the introduction of the two-neutral-umpires policy, which currently remains in place in test cricket. The advantage to home teams from home umpires was more pronounced in the third and fourth innings of test matches. This is consistent with home umpires displaying favouritism, as decisions made in those innings more strongly affect match outcomes, even as crowd pressure generally declines.

Recently, some influential commentators have suggested (see, for example, Stewart (2013)) a return to using home umpires on the grounds that this would increase the available pool of high quality umpires, and hence reduce errors. If such a proposal were to be followed, it may be that the presence of the DRS leads to a reduction in favouritism by home umpires than observed historically. It would be interesting to test this hypothesis at a later date, after more matches under the DRS have been played. However, the DRS still permits umpiring discretion and, as a result, the evidence from this paper suggests that any proposal to move away from neutral umpires should be treated with caution.

Acknowledgements

The authors thank participants at the 2012 Western Economics Association International conference; Bruce Morley, Paul Fenn and Chris O'Brien for many helpful suggestions and comments on this research and Usha Sacheti for kindly assisting with data collection. We also thank a referee and the Associate Editor for many helpful suggestions and comments.

References

Akerlof, G. A. (1997) Social distance and social decisions. *Econometrica*, **65**, 1005–1027.
Andreff, W. and Szymanski, S. (2006) *Handbook on the Economics of Sport*. Cheltenham: Elgar.
Bryson, A., Buraimo, B. and Simmons, R. (2011) Do salaries improve worker performance? *Lab. Econ.*, **18**, 424–433.
Buraimo, B., Forrest, D. and Simmons, R. (2010) The twelfth man?: refereeing bias in English and German soccer. *J. R. Statist. Soc.* A, **173**, 431–449.
Buraimo, B., Simmons, R. and Maciaszczyk, M. (2012) Favoritism and referee bias in European soccer: evidence from the Spanish league and the UEFA Champions League. *Contemp. Econ. Poly*, **30**, 329–343.
Cameron, A. C. and Trivedi, P. K. (2005) *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press.
Cameron, A. C. and Trivedi, P. K. (2010) *Microeconometrics using Stata*. Texas: StataCorp.
Chedzoy, O. B. (1997) The effect of umpiring errors in cricket. *Statistician*, **46**, 529–540.
Crowe, S. M. and Middeldorp, J. (1996) A comparison of leg before wicket rates between Australians and their visiting teams for test cricket series played in Australia, 1977–94. *Statistician*, **45**, 255–262.

Dawson, P. and Dobson, S. (2010) The influence of social pressure and nationality on individual decisions: evidence from the behaviour of referees. *J. Econ. Psychol.*, **31**, 181–191.
Dawson, P., Dobson, S., Goddard, J. and Wilson, J. (2007) Are football referees really biased and inconsistent?: evidence on the incidence of disciplinary sanction in the English Premier League. *J. R. Statist. Soc.* A, **170**, 231–250.
Garicano, L., Palacios-Huerta, I. and Prendergast, C. (2005) Favoritism under social pressure. *Rev. Econ. Statist.*, **87**, 208–216.
Hoult, N. (2013) Ashes 2013: England and Australia appeal over neutral umpires. *Daily Telegraph*. (Available from `www.telegraph.co.uk/sport/cricket/international/theashes/10176936/Ashes-2013-England-and-Australia-appeal-over-neutral-umpires.html`.)
Hynds, M. and Smith, I. (1994) The demand for test match cricket. *Appl. Econ. Lett.*, **1**, 103–106.
Koop, G. (2008) *Introduction to Econometrics*. Chichester: Wiley.
Marylebone Cricket Club (2010) *Laws of Cricket*. London: Marylebone Cricket Club. (Available from `www.lords.org/mcc/laws-of-cricket`.)
Parsons, C. A., Sulaeman, J., Yates, M. C. and Hamermesh, D. S. (2011) Strike three: discrimination, incentives and evaluation. *Am. Econ. Rev.*, **101**, 1410–1435.
Price, J. and Wolfers, J. (2010) Racial discrimination among NBA referees. *Q. J. Econ.*, **125**, 1859–1887.
Ringrose, T. J. (2006) Neutral umpires and leg before wicket decisions in test cricket. *J. R. Statist. Soc.* A, **169**, 903–911.
Stewart, A. (2013) Ashes 2013: bring back 'home' umpires. BBC, London. (Available from `www.bbc.co.uk/sport/0/cricket/23683136`.)
Sumner, J. and Mobley, M. (1981) Are cricket umpires biased? *New Scient.*, **91**, 29–31.