Rain rules for limited overs cricket and probabilities of victory

Ian Preston
Department of Economics, University College London, UK
Jonathan Thomas
Bank of England, UK

Summary. The paper discusses the properties of a rule for adjusting scores in limited overs cricket matches so as to preserve probabilities of victory across rain interruptions. Such a rule is argued to be attractive on grounds of fairness, intelligibility and tactical neutrality. Comparison with other rules also offers a useful way of assessing the way in which application of such rules will affect the fortunes of teams in rain-affected games. Simulations based on an estimated parametrisation of dismissal hazards and on numerical dynamic programming methods are used to compare a probability preserving rule with the now widely used Duckworth-Lewis method.
Keywords: Cricket, rain rules, dynamic programming

1 Introduction

The game of cricket suffers from a requirement for clement weather of a type rarely guaranteed by the climates in which the game is played. Interruptions due to poor weather have proven a particular difficulty for design of rules for the limited overs version of the game, since such interruptions affect the fortunes of the two teams very differently depending upon the point in the game at which they occur and, if prolonged and ignored, would destroy the contest and therefore much of the interest of the game.

Many rules have been suggested for adjusting the terms of the game in such a way as to maintain fairness between the two teams in cases of weather affected matches. However none have proven wholly satisfactory. For example, the "average run rate" rule which scales down second innings targets proportionately is widely felt to favour teams batting second while the "most productive overs" rule which reduces second innings targets only by runs scored in lowest scoring overs is equally widely seen as favouring teams batting first. The many methods, their strengths and their weaknesses, are summarised by Duckworth and Lewis (1998, 1999), whose own rule has won attention and increasingly widespread adoption.

We begin with a description of the game. Limited overs cricket is played between two teams of 11 players. Each team has one "innings" in which it "bats" while the other team "bowls". At any point in the game, the team batting has two players at the centre of the pitch, known as the "wicket", facing balls delivered to them by the bowling side. Balls are delivered in groups of

[1] *Address for correspondence*: I. Preston, Department of Economics, University College London, Gower St, London WC1E 6BT, UK.
E-mail: i.preston@ucl.ac.uk

six known as “overs”. The batting team scores “runs” by attemping to strike the ball and run between ends of the wicket while the bowling side recovers and returns the ball. In doing this the batsmen expose themselves to the risk of being dismissed (“getting out”, “losing a wicket”) in a number of ways. All major forms of dismissal are made more likely by a high run scoring rate and batting teams therefore need to balance the advantage of fast scoring against the risk of losing wickets.

If there is no intervention by the weather then the innings ends after a set number of overs have been bowled or all but one member of the batting team have been dismissed, whichever happens sooner. The total runs accumulated by the team batting first is then the target to be achieved by the team batting second if they are to win the match. However, the game can only be played in clement conditions. Breaks in play due to bad weather raise a problem if occurring after the game has started and lasting sufficiently long to require curtailment of total playing time. Some fair way needs to be found of reducing the target if the overs available for the second innings are reduced and of inflating the target if those for the first innings are. The manner in which this is done needs to recognise that the harm done to a team’s interests will differ depending upon the point in the innings at which the break occurs.

The Duckworth-Lewis rule is based on the notion of correcting for loss of “run-scoring resources.” In this it can be seen as a development of the idea underlying simpler rules based on run rates. We are sympathetic to the general notion of run-scoring resources but there is more than one way to value the several dimensions to resources against each other. Duckworth and Lewis appear implicitly to interpret the value of run-scoring resources as “average total score” given match circumstances. We explore an alternative approach in which they are valued according to their contribution to probability of victory.

This approach has several advantages. For one thing a rain rule which adjusts runs so as to preserve teams’ advantage judged according to their own objectives appears to us to capture well the notion of fairness. Secondly, while application of the rule would require complex calculations, the underlying concept is easily understood. Thirdly, rules such as this which recognise the team’s own objectives are tactically neutral in the sense that they will give teams no incentive to vary strategy in anticipation of rain. If, in all possible future circumstances, probabilities of victory after rain rule adjustment remain the same as before the interruption then teams expecting rain will have no reason to adjust behaviour in anticipation of application of the rule. The main disadvantage of the rule is probably the difficulty, explained below, in applying it consistently to cases where play has to be abandoned.

The paper is organised as follows. Section 2 outlines theories of optimum batting strategy in the context of which we address the properties of rain rules in Section 3. These properties are compared to those of the Duckworth Lewis rule in Section 4 and numerical simulations are used to expand on these comparisons in Section 5. Section 6 concludes.

2 Optimum Batting Strategy

Clarke (1988, 1998) has used dynamic programming to establish optimal run scoring rates in first and second innings of examples of limited overs games. Preston and Thomas (2000) extend this by deriving some general analytical results for problems of this type as well as estimating a model of actual strategies. We exploit their results to prove some general properties of fair rain rules.

We begin with a general formulation of the problem facing the batting team in either innings. There are T_1 balls in the first innings and T_2 in the second. In an uninterrupted game $T_1 = T_2$. Runs accumulated after the tth ball of the ith innings are denoted $R_{i,t}$. The objective of the team batting second is to maximise the probability that it accumulates the target set by the first team, $R_{2,t} > R_{1,T_1}$ for some $t \leq T_2$. Denote the value of that probability after t balls of the innings have expired, $R_{2,t}$ runs have been accumulated and k wickets have fallen by $F_{2,t}(R_{2,t}, k; R_{1,T_1}, T_2)$. (Arguments after the semi-colon do not vary within innings in uninterrupted games and will be suppressed where convenient in what follows). The objective of the team batting first is to maximise the probability that it sets a target which the second team fails to reach. Denote the value of that probability after t balls of the innings have expired, $R_{1,t}$ runs have been accumulated and k wickets have fallen by

$$F_{1,t}(R_{1,t}, k; T_1, T_2) = 1 - \mathcal{E}_{R_{1,T}} F_{2,0}(0, 0; R_{1,T_1}, T_2) \tag{1}$$

where the $\mathcal{E}_{R_{1,T}}$ operator denotes expectation over the distribution of R_{1,T_1} given chosen strategy.

We suppose that the control variable is *expected* run rate $r_{i,t}$. The effect of run scoring rates on the probability of dismissal are captured in a dismissal hazard which depends also on the number of wickets fallen, $h_{i,t} = \eta_i(r_{i,t}, k)$. We assume $\eta(.)$ to be positive, increasing and convex in its first argument. In the simplest case, there would be no uncertainty in run scoring other than that associated with the possibility of dismissal and actual runs achieved would equal $r_{i,t}$. More generally, we would wish to allow for the runs scored to depend uncertainly on intended run rate even if not dismissed so that actual runs achieved would be a random variable with expectation $r_{i,t}$.

Let $r^*_{i,t}$ denote the optimal intended rate and $h^*_{i,t} = \eta_i(r^*_{i,t}, k)$. By Bellman's Principle of Optimality (see Bertsekas 1995, pp.16ff)

$$F_{i,t}(R_{i,t}, k) = (1 - h^*_{i,t+1})\mathcal{E}_{R_{i,t+1}} F_{i,t+1}(R_{i,t+1}, k) + h^*_{i,t+1} F_{i,t+1}(R_{i,t}, k+1) \tag{2}$$

where the $\mathcal{E}_{R_{i,t+1}}$ operator denotes expectation over the distribution of $R_{i,t+1}$ given no dismissal and expected run rate $r^*_{i,t+1}$.

Assuming for heuristic purposes that we can regard scoring rate as a continuous variable and that first order conditions characterise the optimum, then

$$\begin{aligned} 0 &= (1-h^*_{i,t+1})\mathcal{E}_{R_{i,t+1}}\frac{\partial F_{i,t+1}(R_{i,t+1},k)}{\partial R_{i,t+1}} \\ &\quad -\frac{\partial h^*_{i,t+1}}{\partial r^*_{i,t+1}}[\mathcal{E}_{R_{i,t+1}}F_{i,t+1}(R_{i,t+1},k)-F_{i,t+1}(R_{i,t},k+1)] \end{aligned}$$

The first term is the expected marginal gain from run accumulation whereas the second is the expected marginal loss from the increased likelihood of dismissal. Using (2) we can rewrite this as

$$\begin{aligned} 0 &= (1-h^*_{i,t+1})\mathcal{E}_{R_{i,t+1}}\frac{\partial F_{i,t+1}(R_{i,t+1},k)}{\partial R_{i,t+1}} \\ &\quad -\frac{\partial \ln h^*_{i,t+1}}{\partial r^*_{i,t+1}}[\mathcal{E}_{R_{i,t+1}}F_{i,t+1}(R_{i,t+1},k)-F_{i,t}(R_{i,t},k)]. \end{aligned}$$

Optimising strategies can be shown to differ considerably between innings and to depend on the nature of the uncertainty in runs achieved. Preston and Thomas (2000) discuss several cases. Suppose $k = 9$, so that all but one wicket have fallen. In the typical second innings problem with no run scoring uncertainty, then the optimum strategy is to pursue a constant run rate equal to the required rate. For $k < 9$, there are good reasons to expect a higher run rate than this to be optimal. On the other hand in the first innings the optimum strategy is typically an increasing run rate. As the innings progresses the potential overs to be lost by early dismissal are declining and high scoring rates become more attractive. If there is uncertainty then simulations suggest that the run rate should increase modestly over the course of the second innings even for $k = 9$.

3 Probability preserving rain rules

Suppose a rain interruption requires play to be shortened by δ balls. If this happens during the second innings, the batting team has to resume with $T_i' = T_i - \delta$ balls available in the innings. If it happens during the first innings it is usual to reduce the number of balls available to both teams by $\delta/2$. In either case there is a need to adjust scores if one team is not to be seriously disadvantaged by the interruption. It is not regarded as acceptable to revise wickets lost and usual practice is to revise the run target. It is clearly equivalent and expositionally more convenient to represent any adjustment as revising the total of runs accumulated to R'.

Define a probability preserving (PP) rain rule as one which adjust run totals so as to maintain the probability of each team winning. Such a rule falls into the wider class of tactically neutral rain rules which are such that the probability of either team winning after an interruption is a monotonically increasing function of the probability before. It is only for rules within this class that the anticipation of rain will give batting teams no incentive to change strategy.

3.1 Second innings adjustments

Take firstly the case of the second innings. A PP rule decides R' by

$$\begin{aligned} F_{2,t}(R_{2,t}, k; R_{1,T_1}, T_2) &= F_{2,t}(R', k; R_{1,T_1}, T_2 - \delta) \\ &= F_{2,t+\delta}(R', k; R_{1,T_1}, T_2). \end{aligned} \quad (3)$$

That is to say R' is chosen so that the team has the same probability of winning after the interruption as it had before.

It is illuminating to consider adjustment in the case of the shortest possible interruption, $\delta = 1$, since this yields particularly readily interpretable expressions. If batsmen behave optimally according to the results described above, then it can be shown (see Appendix A.1), that

$$R' - R_{2,t} \simeq r^*_{2,t+1} - \frac{h^*_{2,t+1}(1 - h^*_{2,t+1})}{\partial h^*_{2,t+1} / \partial r^*_{2,t+1}}. \quad (4)$$

The appropriate rule therefore accords the team its optimal expected rate $r^*_{2,t+1}$ less an adjustment which is itself a function of $r^*_{2,t+1}$. Let us call this the adjusted optimal rate. The need for the adjustment arises to compensate for the certainty of the batting team surviving the interruption without loss of wickets.

3.2 First innings adjustments

The case of the first innings is complicated slightly by the possibility that total balls available may be modified in both innings. Suppose balls available in each innings are reduced by $\delta/2$. A PP rule decides R' by

$$\begin{aligned} F_{1,t}(R_{1,t}, k; T_1, T_2) &= F_{1,t}(R', k; T_1 - \delta/2, T_2 - \delta/2) \\ &= F_{1,t+\delta/2}(R', k; T_1, T_2 - \delta/2) \end{aligned} \quad (5)$$

Again R' is chosen so that the team has the same probability of winning after adjustment to the overs remaining for each team.

Suppose that the interruption last for two balls, $\delta = 2$, the shortest interruption leading to reduction in both innings. Then by similar reasoning (outlined fully in Appendix A.2), it can be shown that

$$R' - R_{1,t} \simeq r^*_{1,t+1} - \frac{h^*_{1,t+1}(1 - h^*_{1,t+1})}{\partial h^*_{1,t+1} / \partial r^*_{1,t+1}} - \mathcal{E}_{R_{1,T}}[r^*_{2,1} - \frac{h^*_{2,1}(1 - h^*_{2,1})}{\partial h^*_{2,1} / \partial r^*_{2,1}}] \quad (6)$$

Hence the runs accorded will be close to the difference between the current adjusted optimal rate and the expected adjusted optimal rate at the commencement of the second innings.

3.3 Premature termination of match

All discussion so far has assumed that the interruption allows the match to resume with overs still to be played. In such a case a PP rule is implementable with knowledge only of probability level sets in (R, t)-space. However there is a discontinuity in application of the rule as the interruption reaches the scheduled end of the match since at this point a winner has to be decided without the possibility of further play.

One natural extension of the rule to such cases would be to award victory in such a prematurely terminated match to the team most likely to win at the time of interruption. That is to say the team batting would be declared to have won if $F_{i,t}(R_{i,t}, k) > 0.5$. However it should be noted that such a rule requires more information since it needs knowledge not only of the shape of probability level sets but also of the value of the probability. The shape of the probability level sets could in principle be independent, or close to independent, of the relative abilities of the two teams even when one would not expect this to be true of the actual value of the probabilities.

Note also that such an extension to the rule would not be strictly probability preserving since the team ahead at the point of interruption would be declared the winner with certainty. As a consequence the rule would no longer be tactically neutral if there were any possibility of the match being prematurely terminated. As a referee has pointed out, for example, second innings strategy under anticipation of an imminent downpour terminating the match would become a matter of achieving targets over-by-over. Tactical neutrality could be achieved by an alternative extension randomising the declared winner in accordance with the probabilites of victory before the interruption but such indeterminacy might not be a popular feature of a rule.

4 Comparison with Duckworth and Lewis

We start with a description of the Duckworth-Lewis (1998, 1999) (henceforth denoted DL) rule. The DL rule is not a PP rule and is not intended to be, as the authors make clear: "The D/L method maintains the margin of advantage. It does not maintain the probability of winning or losing (Duckworth and Lewis 1999, p.27)." In our opinion, each rule has things to be said in its favour. Comparing adjustments made by the DL and PP rules may nonetheless provide worthwhile insight into how application of the former will or will not alter probabilities of victory at different points in the match.

The DL rule can be represented as adjusting $R_{i,t}$ in line with proportion of "run scoring resources" lost as a consequence of the interruption. To explain this we need to outline the notion of "run scoring resources." In the interpretation offered by Duckworth and Lewis (1998, p.222) these are to be measured by the "average total score" in the remaining overs available given the number of wickets lost and in practice these are calculated from first innings scores. Accordingly let $Z_{1,t}(k; T_1)$ denote the expected score in the remaining $T_1 - t$

balls of the first innings if k wickets have fallen.

The evolution of $Z_{1,t}(k)$ across the course of the innings is crucial to the application of the DL rule and it is therefore worth saying something about this. (In general, behaviour depends on runs already scored so that the expressions below should be thought of as involving averages across the distribution of $R_{1,t}$. We neglect this complication in the ensuing exposition.) Clearly,

$$Z_{1,t}(k) = (1 - h_{1,t+1})[Z_{1,t+1}(k) + r_{1,t+1}] + h_{1,t+1}Z_{1,t+1}(k+1) \tag{7}$$

and therefore, by rearrangement,

$$Z_{1,t}(k) - Z_{1,t+1}(k) = (1 - h_{1,t+1})r_{1,t+1} + h_{1,t+1}[Z_{1,t+1}(k+1) - Z_{1,t+1}(k)].$$

The rate of decline of $Z_{1,t}(k)$ therefore depends on the typical run rate at that point in the first innings and on the likelihood and cost of dismissal.

Let us suppose that it is not too inaccurate to think of teams in the first innings seeking to maximise expected final score. Clarke (1988), for example, is happy to simply assume this as the objective. Then

$$(1 - h_{1,t+1}) + \frac{\partial h_{1,t+1}}{\partial r_{1,t+1}}[Z_{1,t+1}(k+1) + r_{1,t+1} - Z_{1,t+1}(k)] = 0$$

and

$$Z_{1,t}(k) - Z_{1,t+1}(k) = r_{1,t+1} - \frac{h_{1,t+1}(1 - h_{1,t+1})}{\partial h_{1,t+1}/\partial r_{1,t+1}}, \tag{8}$$

so that this decline is in fact exactly the adjusted optimal rate after t balls of the first innings and k wickets lost. This suggests that it may be possible to draw out close links between the nature of DL and PP rules.

4.1 Second innings

The DL formula for adjustment of second innings targets can be represented by

$$R' = R_{2,t} + \frac{R_{1,T_1}}{Z_{1,0}(0)}[Z_{1,t}(k) - Z_{1,t+\delta}(k)]. \tag{9}$$

If $R_{1,T_1} = Z_{1,0}(0)$, so that the target set is equal to the average score from an uninterrupted first innings, then the rule restores to the team affected the decline in resources $Z_{1,t}(k)$ across such an interruption for a first innings team with k wickets lost. Otherwise they are accorded such a reward appropriately scaled.

We can note a number of differences from a PP rule. Firstly, suppose the interruption is for one ball and take the case in which $R_{1,T_1} = Z_{1,0}(0)$. We note from the results above, (8) and (4), that the DL rule attributes something close to the adjusted optimal rate for a typical first innings team whereas a PP rule would attribute the adjusted optimal rate for the team currently batting in the second innings. Since we know that optimal strategies differ in the two innings the two rules can not make the same adjustments.

Secondly, if R_{1,T_1} and $Z_{1,0}(0)$ are not equal then the DL rule scales the adjustment by their ratio. The reason is that, at any point in the innings, the DL rule aims to revise the initial target according to proportion of innings resources lost. The rule takes no account of runs already scored up to the point of the interruption and, as a consequence, a team could in fact find it had already achieved its revised target on resumption ie $R' > R_{1,T_1}$ even if $\delta < T_2 - t$. DL recognise this explicitly: "if sufficient overs are lost it may not even be necessary for them to resume their innings. This is not an inconsistency. It happens with every other method of target revision that has been used (Duckworth and Lewis 1999, p.27)." This property alone makes immediately plain that the rule can not preserve probabilities of winning.

However, it would plainly be possible to modify the rule to avoid this property by having it revise the *remaining* target according to the proportion of *remaining* resources left.

$$R' = R_{2,t} + \frac{[R_{1,T_1} - R_{2,t}]}{Z_{1,t}(k)}[Z_{1,t}(k) - Z_{1,t+\delta}(k)]. \tag{10}$$

The option of providing targets by such a rule has been incorporated into the latest version of the DL computer program.

4.2 First innings

The DL rule for first innings adjustments differs depending upon which team suffers most from the interruption. We concentrate comments on the case holding "in the great majority of instances (Duckworth and Lewis 1998, p.223)" - that in which the team batting first loses the most resources. In this case the rule stipulates

$$R' = R_{1,t} + Z_{1,t}(k) - Z_{1,t+\delta/2}(k) - Z_{1,0}(0) + Z_{1,\delta/2}(0).$$

The runs accorded are the difference between the resources lost by the batting team $Z_{1,t}(k) - Z_{1,t+\delta/2}(k)$ and those that would be lost by a typical first innings team suffering a similar interruption at the beginning of the innings $Z_{1,\delta/2}(0) - Z_{1,0}(0)$.

Comparing this with (6) and noting (8) reveals differences and similarities. For a short interruption, the PP rule makes an adjustment based on the difference between the current adjusted optimal scoring rate and that expected to apply to the lost balls at the beginning of the subsequent innings. The DL rule adjustment, on the other hand, reflects the difference between the same current adjusted optimal scoring rate and that applying at the beginning of the current innings.

Secondly, the averages used in the DL rule are not conditioned on the current run total $R_{1,t}$. This is probably of little significance as regards the optimal rate now and of none as regards that at the beginning of a typical first innings.

However it will matter to the optimal rate expected to apply at the beginning of the second innings and therefore to the PP rule. The consequences of this will depend on the state of play at the point of interruption. If the team batting first is doing well then the opposing team would need to aim at a higher run rate than otherwise and to the extent that the DL rule does not take account of this it appears that it should favour teams already performing well.

5 Simulated examples

Dynamic programming simulations allow us to investigate the magnitude and direction of differences between rain rule adjustments implied by Duckworth-Lewis and probability preserving rules. The simulations are similar in character to those in Preston and Thomas (2000). We consider a forty over match and, for manageability, we adopt a specification with the over as the unit of decision. The hazard is assumed to be the same for all teams and to take the multiplicative form $\eta_i(r_{i,t}, k) = \beta r_{i,t}^{\alpha} \gamma(k)$ where $\gamma(k)$ is a step function taking the value 0.5 in the range $k < 6$, 1 in the range $5 < k < 8$ and 1.5 for $7 < k$. To allow for uncertainty in run scoring there is a probability p that the team fails to score in any over even if not dismissed. Thus each over has the possible outcomes that a wicket is lost, no wicket is lost and no runs scored and that no wicket is lost and the intended run rate scored. We have found conclusions to be insensitive to the fineness of grid for run scoring rates and the simulations used here allow choice only over integer values.

Choice of α, β and p is clearly crucial. Preston and Thomas (2000) estimate a partnership duration model using 133 matches from the 1996 Axa Equity and Law Sunday League without imposing optimal behaviour. Their estimates suggest a value for α of about 1.1. We have also explored choice of parameter values by fitting simulated distributions of first innings scores under optimising behaviour to the distribution observed in the same data. To be precise, we took a grid of 1430 combinations of values for α, β and p with $1.1 \leq \alpha \leq 1.9$, $-6.5 \leq \ln \beta \leq -2.25$ and $0 \leq p \leq 0.5$. For each combination we solved for the optimum strategy, simulated 10,000 games and took the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles of the distribution of first innings scores. A simple grid search, minimising the equally-weighted quadratic distance from the corresponding observed percentiles, gives values $\alpha = 1.6$, $\ln \beta = -4.5$ and $p = 0.2$. The closeness of fit is shown in Table 1. (More details are available on request from the authors). We adopt these values for simulation.

Table 1: Actual and fitted percentiles for first innings totals

Percentile	*Actual value*	*Fitted value*
5th	127	122
10th	139	140
25th	167	170
50th	201	202
75th	234	233
90th	265	262
95th	276	277

The optimum batting strategies calculated for such parameter values are similar in properties to those discussed in Preston and Thomas (2000). These allow one to calculate victory probability given circumstances at any point in the two innings. Figure 1, for example, shows curves describing second innings victory probability as a function of outstanding run target, assuming three wickets lost, at four different points in the innings. These curves clearly contain all information required for application of a probability preserving rule.

To compare with the DL rule we need to calculate expected run totals given wickets lost and balls remaining at all feasible points in the first innings. We do this by taking means across 100,000 simulated matches with teams applying the known optimising strategies. The estimated DL curves are shown in Figure 2 and clearly have a sensible shape, similar to those estimated directly by Duckworth and Lewis (1998, p.222) from actual data. Note that we take the actual simulated averages without smoothing by fitting a parametric form as they do. The gaps in the illustrated curves occur where the circumstances involved occurred insufficiently often (by which we mean less than 100 times) for accurate estimation through simulations.

These curves can then be used to calculate Duckworth Lewis adjustments for comparison with those required by a probability preserving rule. We have done this for many first and second innings situations and Figures 3 and 4 illustrate typical examples. In each case we consider an interruption occurring after 15 overs with three wickets lost and in the second innings case we assume the initial target was the average set by first innings teams. We then take three situations. In one the current scores are such that the two teams are equally likely to win while, in the other two, one or other team is well ahead so that the better placed team has two to one odds of victory.

Take for example Figure 3. The uppermost set of three lines correspond to a case where 133 runs remain to be scored in 25 overs with 7 wickets in hand. A team pursuing the optimum batting strategy from this point onward would achieve the target with probability 0.33 given the parameter values used here. The three lines plot the adjustments called for under the three rules as overs left at resumption get fewer. The bottom set of lines illustrate adjustments for a case where only 105 runs needed initially to be scored and the probability of victory was correspondingly higher at 0.67.

For both first and second innings the pictures indicate compatible conclusions. In either innings the DL rule is more generous relative to a PP rule the better the batting team is performing. In each case illustrated batting teams already performing well in the innings have their probability of victory enhanced if rain intervenes and the DL rule is applied. In the first innings the DL rule benefits the bowling side only if the batting team is doing well and in the second innings only if the batting team is considerably behind.

The modified rule (10) applicable in the second innings is also illustrated. Interestingly the modification brings the DL rule closer to a PP rule for batting teams performing well but takes it further away for teams behind par who are treated noticeably more generously in comparison with a genuine PP rule.

An interesting question to ask is how probabilities of victory after application of the DL rule relate to probabilities before. Figure 5 plots these probabilities against each other for interruptions of the same length at the same point in the second innings, as runs required and wickets lost are varied. There are three distinct lines corresponding to the three different possibilities considered for number of wickets lost. Victory probabilities after application of the rule are not a simple function of probabilities before - for a given probability of victory before rain the probability at resumption is higher under the DL rule the fewer wickets are lost. This has interesting implications for the tactical nonneutrality of the rule. Since the rule does not preserve the ranking of situations by victory probability, teams anticipating rain and seeking to maximise probability of winning after anticipated application of the DL rule might behave differently from how they would if no rain were expected. Anticipation of rain can lead to changes of strategy as the team batting switches towards maximising the probability after the anticipated application of the rule. In these simulations, for example, we can see that a team responding in this way to the DL rule would adopt a more cautious batting style than it would under a tactically neutral rule.

6 Practical application

6.1 Feasibility

The information needed to calculate the PP adjustments comprises a tabulation of the victory probabilities $F_{1,t}(R_{1,t}, k; T_1, T_2)$ and $F_{2,t}(R_{2,t}, k; R_{1,T_1}, T_2)$. This is more extensive than the information required for current alternatives and could not, for example, be condensed into tables as brief as those provided by Duckworth and Lewis (1999). Nonetheless the calculations could be programmed to be done quickly and simply on a computer capable of storing the required information and the limit on feasibility is therefore the accessibility to umpires of the required computer resources. (With advances in modern telecommunications this is likely to become decreasingly prohibitive). If it were possible to approximate the probability functions accurately with a sufficiently parsimonious specification then computation would be rendered even simpler. That remains a topic for further research.

6.2 Examples

To facilitate comparison with other methods and demonstrate feasibility we provide a couple of examples. These both come from the first round of the 1992 World Cup and are both used as examples in Duckworth and Lewis (1998). Since these were 50 over matches we repeated the above dynamic programming and simulation exercises using the same parameter values to produce estimates of the required functions for 50 over matches.

The first example is the match between South Africa and Pakistan at Brisbane on 8 March 1992. South Africa set a target of 211 runs in a full 50 overs. Pakistan were 74 for the loss of 2 wickets when rain intervened with 29 overs remaining. The match resumed with 14 overs lost.and the most productive overs method in use for this tournament set a revised target of 193 which Pakistan failed to achieve. Duckworth and Lewis (1998) report that their method would have set a lower target of 164, in line with the observation that the most productive overs method tends to penalise teams batting second. The methods of this paper provide an alternative set of estimates for the DL curves using a different set of data and a different estimation method and therefore could give an alternative DL target but in fact give exactly the same target of 164. The PP target calculated to keep the probability of victory constant across the interruption is very close at 161.

The second example is the match four days later between South Africa and England at Melbourne. South Africa again batted first, without interruption, and set a target of 236 runs. England had reached 63 without loss in 12 overs before rain robbed the match of 9 overs. The revised target of 226 was reached by England. Duckworth and Lewis (1998) recommend a lower target of 207. The alternative DL target using our estimated parameters is quite close at 203. The PP target is 207.

7 Conclusion

Rain rules based on the idea of keeping probabilities of victory constant across interruptions offer an intelligible and fair alternative to other rules. Furthermore even if the case for their use is not regarded as persuasive it is still interesting to compare them to alternative rules to judge how the application of those other rules is likely to change the balance of advantage within a game. For example, simulations demonstrate that under the widely used Duckworth-Lewis rule, probabilities of victory after an interruption are not in a one-to-one relationship with probabilities before, teams already doing well at the point of interruption are favoured and anticipation of rain may lead teams towards cautious batting.

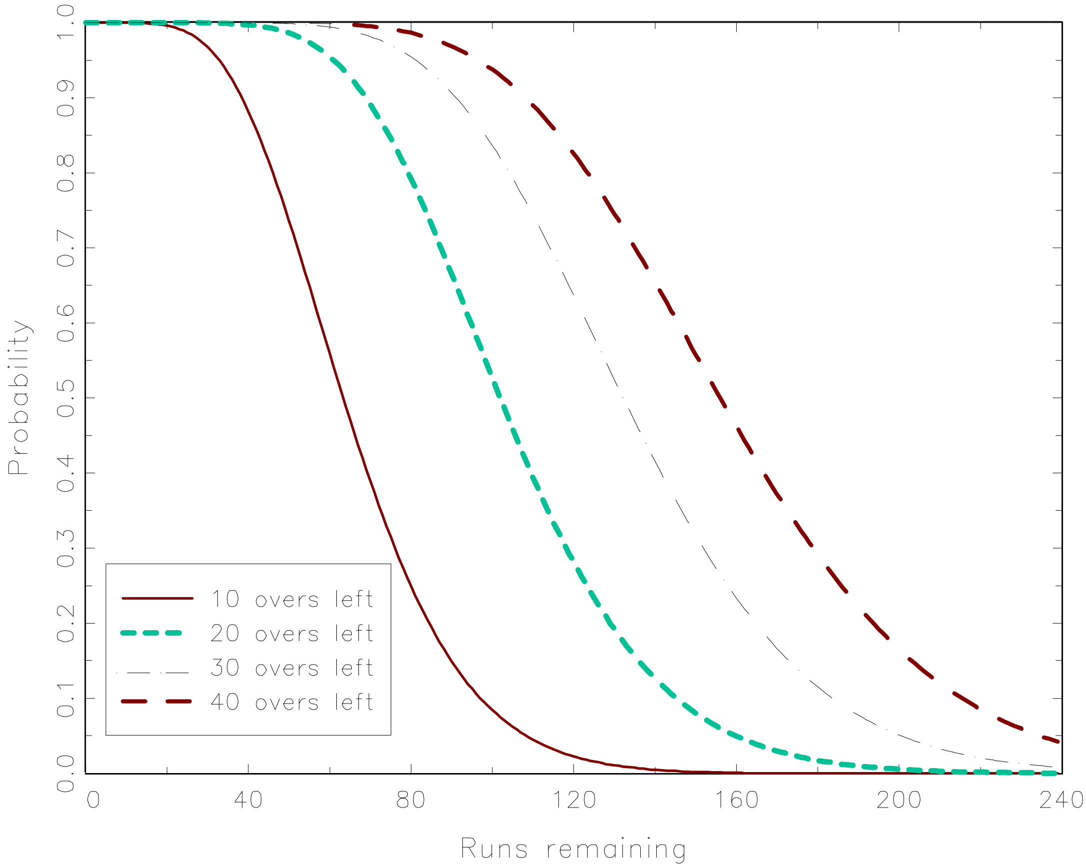

Figure 1 Second Innings Victory Probabilities; three wickets lost

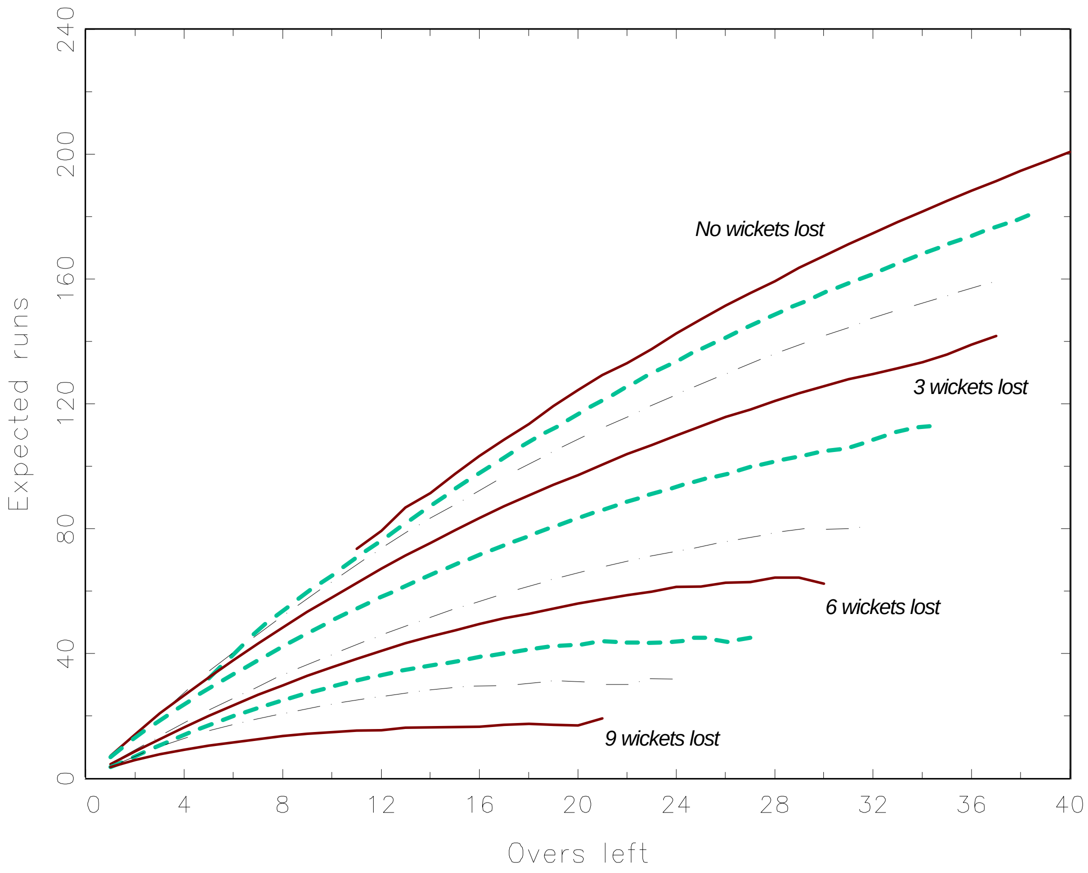

Figure 2 Simulated Duckworth Lewis Curves

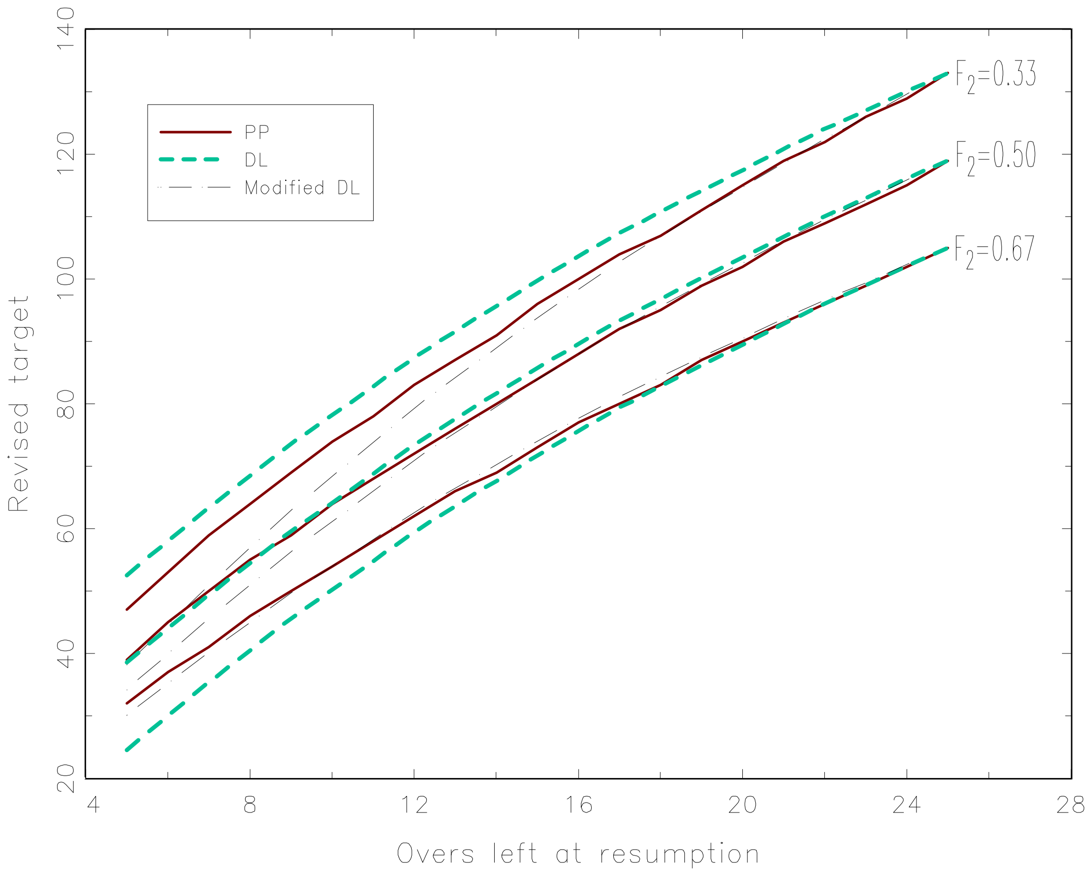

Figure 3 Second Innings Rain Rule Comparisons; interruption after fifteen overs with three wickets lost

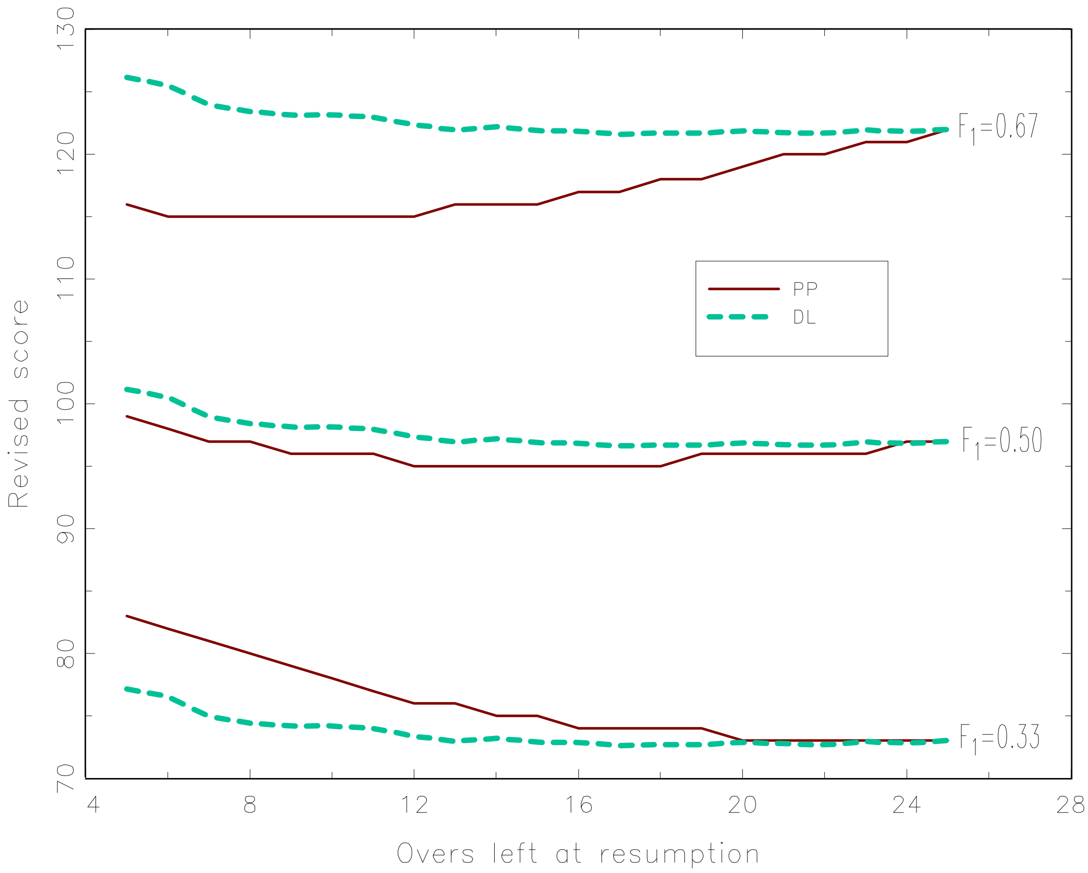

Figure 4 First Innings Rain Rule Comparisons; interruption after fifteen overs with three wickets lost

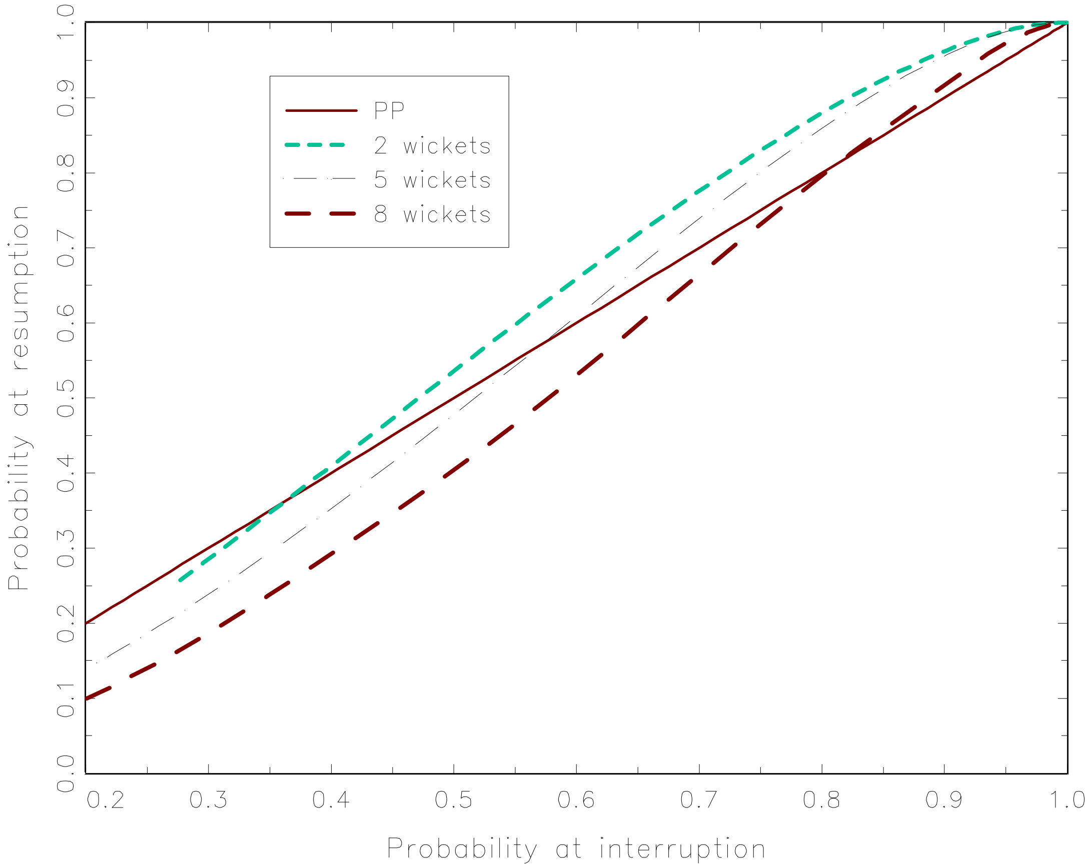

Figure 5 Victory Probabilities Before and After Interruptions; interruption of ten overs duration after fifteen overs of play

Acknowledgements

We have benefitted from discussions with or suggestions from Pelham Barton, Frank Duckworth, Jeremy Edwards, Charles Grant, Anthony Lewis, Hamish Low, Stephen Clarke, participants in seminars at UCL and the University of Birmingham and at meetings of the Royal Statistical Society and three anonymous referees. We are grateful to the Press Association and the England and Wales Cricket Board for access to data. All opinions remain our own.

Appendix A

A.1 Second innings adjustments

>From (3)

$$\begin{aligned} F_{2,t}(R_{2,t}, k; R_{1,T_1}, T_2) &= F_{2,t}(R', k; R_{1,T_1}, T_2 - \delta) \\ &= F_{2,t+\delta}(R', k; R_{1,T_1}, T_2). \end{aligned}$$

Using (2)

$$\begin{aligned} F_{2,t}(R_{2,t}, k) &= \mathcal{E}_{R_{2,t+1}} F_{2,t+1}(R_{2,t+1}, k) \\ &\quad + h^*_{2,t+1}[F_{2,t+1}(R_{2,t}, k+1) - \mathcal{E}_{R_{2,t+1}} F_{2,t+1}(R_{2,t+1}, k)] \\ &= \mathcal{E}_{R_{2,t+1}} F_{2,t+1}(R_{2,t+1}, k) \\ &\quad - \frac{h^*_{2,t+1}(1 - h^*_{2,t+1})}{\partial h^*_{2,t+1}/\partial r^*_{2,t+1}} \mathcal{E}_{R_{2,t+1}} \frac{\partial F_{2,t+1}(R_{2,t+1}, k)}{\partial R_{2,t+1}} \qquad (11) \end{aligned}$$

By averaging across Taylor expansions around $R' = R_{2,t+1}$

$$\begin{aligned} F_{2,t+1}(R', k) &\simeq \mathcal{E}_{R_{2,t+1}} F_{2,t+1}(R_{2,t+1}, k) \\ &\quad + \mathcal{E}_{R_{2,t+1}} \frac{\partial F_{2,t+1}(R_{2,t+1}, k)}{\partial R_{2,t+1}} [R' - R_{2,t+1}]. \qquad (12) \end{aligned}$$

Equating the two expressions we derive (4)

$$\begin{aligned} 0 &\simeq \mathcal{E}_{R_{2,t+1}} \frac{\partial F_{2,t+1}(R_{2,t+1}, k)}{\partial R_{2,t+1}} [R' - R_{2,t+1} + \frac{h^*_{2,t+1}(1 - h^*_{2,t+1})}{\partial h^*_{2,t+1}/\partial r^*_{2,t+1}}] \\ &\Rightarrow R' - R_{2,t} \simeq r^*_{2,t+1} - \frac{h^*_{2,t+1}(1 - h^*_{2,t+1})}{\partial h^*_{2,t+1}/\partial r^*_{2,t+1}} \end{aligned}$$

assuming that $\frac{\partial F_{2,t+1}(R_{2,t+1}, k)}{\partial R_{2,t+1}}$ is close to constant across the range of $R_{2,t+1}$ in question.

A.2 First innings adjustments

>From (5)

$$\begin{aligned} F_{1,t}(R_{1,t}, k; T_1, T_2) &= F_{1,t}(R', k; T_1 - \delta/2, T_2 - \delta/2) \\ &= F_{1,t+\delta/2}(R', k; T_1, T_2 - \delta/2). \end{aligned}$$

We expand both sides. By similar reasoning to that used in (11)

$$F_{1,t}(R_{1,t},k) = \mathcal{E}_{R_{1,t+1}} F_{1,t+1}(R_{1,t+1},k) - \frac{h^*_{1,t+1}(1-h^*_{1,t+1})}{\partial h^*_{1,t+1}/\partial r^*_{1,t+1}} \mathcal{E}_{R_{1,t+1}} \frac{\partial F_{1,t+1}(R_{1,t+1},k)}{\partial R_{1,t+1}}. \quad (13)$$

On the right hand side, we use the fact that

$$\begin{aligned} & F_{1,t+1}(R',k;T_1,T_2-1) - F_{1,t+1}(R',k;T_1,T_2) \\ = \; & \mathcal{E}_{R_{1,T}}[F_{2,0}(0,0;R_{1,T_1},T_2) - F_{2,1}(0,0;R_{1,T_1},T_2)] \\ = \; & \mathcal{E}_{R_{1,T}}[\mathcal{E}_{R_{2,1}} F_{2,1}(R_{2,1},0) - F_{2,1}(0,0)] \\ & + \mathcal{E}_{R_{1,T}} h^*_{2,1}[F_{2,1}(0,1) - \mathcal{E}_{R_{2,1}} F_{2,1}(R_{2,1},0)] \\ \simeq \; & \mathcal{E}_{R_{1,T}} \mathcal{E}_{R_{2,1}} \frac{\partial F_{2,1}(R_{2,1},0)}{\partial R_{2,1}} [r^*_{2,1} - \frac{h^*_{2,1}(1-h^*_{2,1})}{\partial h^*_{2,1}/\partial r^*_{2,1}}], \end{aligned} \quad (14)$$

where the arguments after the semicolon in the terms of the final two lines have been suppressed. The derivation of the final line from the previous one follows from Taylor expansion of the first expectation around $R_{2,1} = 0$ and use of the first order condition to substitute in the second. Note that the expectation operator $\mathcal{E}_{R_{1,T}}$ must here be read as taking expectations over the distribution of R_{1,T_1} conditional on $R_{1,t+1} = R'$ and we assume reasonably that we can ignore any difference made by the anticipated curtailing of the second innings duration.

Combining these and using an average of Taylor expansions of $F_{1,t+1}(R',k;T_1,T_2)$ around $R' = R_{1,t+1}$ as in derivation of (12), we derive

$$\begin{aligned} 0 \simeq \; & \mathcal{E}_{R_{1,t+1}} \frac{\partial F_{1,t+1}(R_{1,t+1},k)}{\partial R_{1,t+1}} [R' - R_{1,t+1} + \frac{h^*_{1,t+1}(1-h^*_{1,t+1})}{\partial h^*_{1,t+1}/\partial r^*_{2,t+1}}] \\ & + \mathcal{E}_{R_{1,T}} \mathcal{E}_{R_{2,1}} \frac{\partial F_{2,1}(R_{2,1},0)}{\partial R_{2,1}} [r^*_{2,1} - \frac{h^*_{2,1}(1-h^*_{2,1})}{\partial h^*_{2,1}/\partial r^*_{2,1}}]. \end{aligned}$$

Thus, to the extent that

$$\frac{\partial F_{1,t+1}(R_{1,t+1},k)}{\partial R_{1,t+1}} \simeq \mathcal{E}_{R_{1,T}} \frac{\partial F_{2,1}(R_{2,1},0)}{\partial R_{2,1}}$$

across the relevant ranges of $R_{1,t+1}$ and $R_{2,1}$, as seems reasonable given (1), we can derive (6)

$$R' - R_{1,t} \simeq r^*_{1,t+1} - \frac{h^*_{1,t+1}(1-h^*_{1,t+1})}{\partial h^*_{1,t+1}/\partial r^*_{1,t+1}} - \mathcal{E}_{R_{1,T}}[r^*_{2,1} - \frac{h^*_{2,1}(1-h^*_{2,1})}{\partial h^*_{2,1}/\partial r^*_{2,1}}].$$

References

Bertsekas, D. (1995) *Dynamic Programming and Optimal Control.* Athena Scientific: Belmont.

Clarke, S. R. (1988). Dynamic programming in one-day cricket - optimal scoring rates. *J. Opl. Res. Soc.* **39**, 331-337.
Clarke, S. R. (1998). Test Statistics. In *Statistics in Sport* (ed J. Bennett), pp. 83-103. Arnold: London.
Duckworth, F. C. and Lewis, A. J. (1998). A fair method of resetting the target in interrupted one-day cricket matches. *J. Opl. Res. Soc.* **49**, 220-227.
Duckworth, F. C. and Lewis, A. J. (1999). *Your comprehensive guide to the Duckworth/Lewis method for resetting targets in one-day cricket.* University of the West of England: Bristol.
Preston, I. P. and Thomas, J. M. (2000) Batting strategy in limited overs cricket. *J. R. Statist. Soc. D*, **49**, 95-106.