The Best Batsmen and Bowlers in One-Day Cricket

by

David Beaudoin
B.Sc., Laval University, 2001

A PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the Department
of
Statistics and Actuarial Science

SIMON FRASER UNIVERSITY
April 2003

APPROVAL

Name: David Beaudoin

Degree: Master of Science

Title of project: The Best Batsmen and Bowlers in One-Day Cricket

Examining Committee: Dr. Richard Lockhart
Chair

____

Dr. Tim B. Swartz
Senior Supervisor
Simon Fraser University

____

Dr. Paramjit S. Gill
Co-Senior Supervisor
Okanagan University College

____

Dr. Larry K. Weldon
Simon Fraser University

____

Dr. Jinko Graham
External Examiner
Simon Fraser University

Date Approved: ______________________________

Abstract

This project proposes a new measure for assessing the performance of batsmen and bowlers in one-day cricket. Current measures of performance all suffer from serious drawbacks, thus expressing the need for a more sensible measure. The ratio of runs scored to resources consumed (where resources are defined according to the Duckworth-Lewis method of resetting targets) provides a statistic which is developed and studied in this project.

The approach used in defining the new measure is explained in detail, as well as a description of the Duckworth-Lewis method. The statistic is then calculated for a subset of batsmen and bowlers. In addition, a standard error is provided to help determine real differences in performance. Finally, comparisons are made with traditional measures of performance by looking at players' rankings.

Acknowledgements

There are no words to express how grateful I am to my senior supervisor, but most of all my good friend, Dr. Tim Swartz. He was always available and he believed in me even more than I believed in myself. I owe him immense gratitude for having me spend a month with his family in Australia. I am not about to forget about that trip! Thank you so much to his wife Kim and his daughters Philippa and Fiona for making my experience out West more enjoyable.

I also wish to thank my co-supervisor Paramjit Gill for his advice and suggestions. For introducing Tim and me to the game of cricket and for his hospitality, thanks to Basil de Silva back in Melbourne at RMIT. The whole staff and faculty at SFU guided me very well throughout these two years, especially Dr. Robin Insley, Dr. Richard Lockhart, Dr. Randy Sitter, Sylvia Holmes and Ian Bercovitz.

I am blessed with a great family; I would not have made it that far without my parents and my brother who have always supported me. I think of you very often even though we live so far away from each other.

For sharing some moments of craziness with me over the years, I wish to acknowledge my good friends Patrick "The Mammoth" Malenfant, Francois "F" Verret, Jean-Marc Daigle, Jacquie "Wacky" Gregory and Dawn Bodurtha. To everyone in K9501, I appreciate your help, but I would also like to thank you for making me laugh during my time in BC.

Contents

Abstract . . . iii
Acknowledgements . . . iv
List of Tables . . . vi
List of Figures . . . viii
1 Introduction . . . 1
2 One-Day Cricket . . . 4
2.1 The Game . . . 4
2.2 Common Statistics . . . 8
3 New Statistics . . . 12
3.1 The Approach . . . 12
3.2 Batting . . . 20
3.3 Bowling . . . 26
4 Discussion . . . 30
Appendices
A Tables . . . 32
B Figures . . . 47
Bibliography . . . 61

List of Tables

A.1 The Duckworth-Lewis resource table where the percentage of resources remaining is given as a function of the number of wickets lost and the number of overs available. 32

A.2 Batting statistics for 12 of the best active batsmen in ODI cricket listed in order of their runs per match (RM). The batting average (BA) and the strike rate (SR) are also provided along with the standard error for RM and the number of matches used in the calculations. 33

A.3 Components of the variance of the runs per match statistic for 12 of the best active batsmen in ODI cricket. 34

A.4 P-values of pairwise multiple comparisons for 12 of the best active batsmen in ODI cricket with values less than 0.05 in bold face. 35

A.5 Rankings based on common statistics (batting average (BA) and strike rate (SR)) and rankings based on runs per match (RM) for 12 of the best active batsmen in ODI cricket. 36

A.6 Sample correlation coefficients for the batting average (BA), strike rate (SR) and runs per match (RM) statistic based on 12 of the best active batsmen in ODI cricket. 37

A.7 First and second innings statistics (batting average (BA), strike rate (SR) and runs per match (RM)) for 12 of the best active batsmen in ODI cricket. The number of matches used in the calculations are also provided. 38

A.8 Correlation coefficients between the first and second innings batting statistics for 12 of the best active batsmen in ODI cricket. 39

A.9 Bowling statistics for 12 of the best active bowlers in ODI cricket listed in order of their runs per match (RM). The bowling average (BA), the economy rate (ER) and the strike rate (SR) are also provided along with the standard error for RM and the number of appearances used in the calculations. 40
A.10 Components of the variance of the runs per match statistic for 12 of the best active bowlers in ODI cricket. 41
A.11 P-values of pairwise multiple comparisons for 12 of the best active bowlers in ODI cricket with values less than 0.05 in bold face. 42
A.12 Rankings based on common statistics (bowling average (BA), economy rate (ER) and strike rate (SR)) and rankings based on runs per match (RM) for 12 of the best active bowlers in ODI cricket. 43
A.13 Sample correlation coefficients for the bowling average (BA), economy rate (ER), strike rate (SR) and runs per match (RM) statistic based on 12 of the best active bowlers in ODI cricket. 44
A.14 First and second innings statistics (bowling average (BA), economy rate (ER), strike rate (SR) and runs per match (RM)) for 12 of the best active bowlers in ODI cricket. The number of appearances these statistics are based on are also provided. 45
A.15 Correlation coefficients between the first and second innings bowling statistics for 12 of the best active bowlers in ODI cricket. 46

List of Figures

B.1 Average runs scored according to D/L method for overs available and wickets lost; $w = 0$ at top, $w = 9$ at bottom. Source: Duckworth, F.C. and Lewis, A. J. (2002), "Review of the application of the Duckworth/Lewis method of target resetting in one-day cricket." *Mathematics and Computers in Sport*, G. Cohen and T. Langtry (editors), Bond University, Queensland, Australia, **6**, 127-140. 47

B.2 Multiple comparisons without taking into account the multiple comparisons problem (level of significance = 0.05) for 12 of the best active batsmen in ODI cricket. 48

B.3 Multiple comparisons using Bonferroni's approach (level of significance = 0.00076) for 12 of the best active batsmen in ODI cricket. 49

B.4 Runs per match (RM) statistic for 12 of the best active batsmen in ODI cricket presented in increasing order. Standard errors multiplied by the 97.5-th percentile of a standard normal are attached to the RM values. 50

B.5 First versus second innings batting average for 12 of the best active batsmen in ODI cricket. 51

B.6 First versus second innings strike rate for 12 of the best active batsmen in ODI cricket. 52

B.7 First versus second innings runs per match for 12 of the best active batsmen in ODI cricket. 53

B.8 Multiple comparisons without taking into account the multiple comparisons problem (level of significance = 0.05) for 12 of the best active bowlers in ODI cricket. 54
B.9 Multiple comparisons using Bonferroni's approach (level of significance = 0.00076) for 12 of the best active bowlers in ODI cricket. 55
B.10 Runs per match statistic for 12 of the best active bowlers in ODI cricket presented in increasing order. Standard errors multiplied by the 97.5-th percentile of a standard normal are attached to the RM values. 56
B.11 First versus second innings bowling average for 12 of the best active bowlers in ODI cricket. 57
B.12 First versus second innings economy rate for 12 of the best active bowlers in ODI cricket. 58
B.13 First versus second innings strike rate for 12 of the best active bowlers in ODI cricket. 59
B.14 First versus second innings runs per match for 12 of the best active bowlers in ODI cricket. 60

Chapter 1

Introduction

The game of cricket has been a popular sport around the world for a long time. In fact, a series of Codes of Law has governed this sport for over 250 years. Just like baseball (which could be called cricket's cousin because of their similarities), cricket's main feature is the strategy involved and that is the main reason why cricket fans love the game so much. From a statistician's point of view, cricket is a great game since the finite number of outcomes and balls makes it conducive for modelling. Also, a great amount of data has been collected on cricket matches.

There exist two varieties of cricket: test matches and one-day matches. The former last five days, so fans attending the first few days of the game may not be happy about going home not knowing which team would end up winning. Moreover, many test matches end in draws which is unsatisfying to many fans. That is why one-day matches were introduced 25-30 years ago. Its popularity has been going up, although purists of the game still prefer test matches.

The International Cricket Council (ICC) is the governing body of all cricket matches at the international level. It is responsible for the organisation of major tournaments around the world. As of now, there are 10 full-member countries of the ICC: Australia, Bangladesh, England, India, New Zealand, Pakistan, South Africa, Sri Lanka, West Indies and Zimbabwe.

The most important tournament organised by the ICC is the World Cup, which is held every four years. Few Canadians know about the last World Cup which was held in South Africa in February 2003. In fact, did you know that Canada's national team played in this major event for only the second time in its history? "You tend to find with the Canadians that they are completely surprised they've actually got a cricket team, and even more surprised they have qualified for the World Cup" said squad member Ian Billcliff. Canada's first World Cup appearance occurred in 1979 where they lost all three games. They did better this year, earning their first ever World Cup win over Bangladesh before losing their remaining five games. We note that only one player on the Canadian team was a native born player, the rest of them being expatriates from the West Indies and Asia.

Here are some interesting facts about cricket in Canada. The first international cricket match played in the modern world involved the USA versus Canada. It was held in the 1840s and was watched by 10,000 spectators at Bloomingdale Park in New York. The fixture is thought to be the oldest international sporting event in the modern world, predating even today's Olympic Games by nearly 50 years.

Since then, Canada has been the host of numerous international cricket games. Indeed, until the late 1920s it was commonplace for the West Indies, England and Australia to tour Canada. Toronto held a number of matches between India and Pakistan in 1996 because, according to the *Montreal Gazette*, "The angle being that Canada was pretty much the only place on Earth these two death's-door rivals could play without causing a riot." Unfortunately, it was also in Toronto that Pakistan batsman Inzamam decided to jump into the crowd and attack a fan who had called him a potato! Let's also mention that the best cricketer to ever play the game, Sir Donald Bradman, enjoyed his experience in Canada, describing Stanley Park in Vancouver as his "favorite cricket ground".

The issue that we address in this project is the assessment of the performance of batsmen and bowlers in one-day cricket. It is not an easy task to come up with sensible statistics in one-day cricket. Today's traditional measures of performance all suffer from serious drawbacks, thus calling for new ideas. We believe that the measure proposed in this project and the companion paper (Beaudoin and Swartz, 2003) takes into account what is really important in winning a match. To do so, we use the Duckworth-Lewis (D/L) method, a statistical approach which has been adopted by the ICC for all of its competitions (in the context of resetting targets in interrupted matches).

In chapter 2, we discuss the game of cricket, more precisely one-day cricket. The basic rules of the game are reviewed, as well as the most common statistics used to assess batting and bowling performance. As we shall see, none of them reflects correctly the players' performances on the field. In chapter 3, we define a new measure of batting and bowling prowess in one-day cricket. The approach used in defining this measure is carefully explained, along with a description of the D/L method. Then, the new measure is applied to the context of batting. The measure is calculated for a subset of batsmen based on data collected from one-day international (ODI) cricket matches. Rankings are produced and comparisons are made with traditional measures. We then carry out similar tasks with respect to bowling. We conclude with a discussion in chapter 4.

Chapter 2

One-Day Cricket

2.1 The Game

We begin by describing one-day cricket in some detail, as this is necessary to understand the traditional and the proposed measures of performance. In cricket, there are two teams (called "sides") with eleven players each. In one-day matches, each side bats only once. So, the first team bats its "innings" (notice that an innings in cricket corresponds to a half-inning in baseball), followed by the second team batting its innings. A coin flip prior to the game decides which team has the option to bat first or second. Whoever scores the most runs wins the game.

The game is played on a large oval-shaped field with a rope marking the outer edge of the cricket ground. In the middle is a rectangular area called the "pitch" with one wooden wicket at each end. In fact, each wicket is located outside the pitch, but not very far away from it. The pitch is about 66 feet long and 10 feet wide. A wicket is about two feet tall and is formed of three vertical pieces ("stumps") and two horizontal ones ("bails"). When the wicket is hit by the ball, one or both of the crosspieces will fall off. This is one of the most common ways to get a batsman out.

When an innings starts, the batting side sends up its first two players in their batting order. This is called the first "partnership". These two players bat until one of them makes out, in which case the third player in the order replaces the batsman who just made out. The second partnership starts and continues until one of the two batsmen makes out. Then, the fourth man in the order comes in the game, and so on.

So how do two players bat? Let's call them B1 and B2. B1 stands at the left end of the pitch, whereas B2 stands at the right end. On each play, only one of the two players is the "striker", that is only one of them is actually batting. Which of the two players is the striker depends on a few factors, as shall be explained later. Let's suppose B1 is the striker for now. Among the eleven players from the fielding side, one of them is going to be the bowler (not through the whole innings though) and another one is going to be the wicketkeeper. The bowler has to stand on the opposite end of the pitch from the striker. Thus, in this example, the bowler will deliver the ball from the pitch's right end. So, on the left end of the pitch will be B1 standing just outside of the line which demarcates the pitch, along with the wicket located in-between B1 and the wicketkeeper. One of the wicketkeeper's jobs is to catch the ball if it has not been hit by the striker (B1).

Once the ball is hit by the striker, the two batsmen B1 and B2 "change places" while the fieldsmen are chasing the ball and trying to throw it back to one of the wickets. Notice that B1 and B2 can change places more than once if they have enough time to do so. Every time they do, this scores a run for their side and for the striker.

However, the batsmen must be careful when changing places because they might make out. Indeed, if a wicket is broken (that is if the horizontal pieces are knocked off by the ball) while the batsman is "out of his ground", he is out! A player is said to be out of his ground if he and his bat are inside the pitch. Batsmen take their bats with them when they run and this has an important implication: in order to be "safe", they do not have to physically cross the outer line which demarcates the pitch. The bat is considered as being part of the batsman's body in this case, so all he has to do is touch safe territory with the tip of his bat.

There are more ways for the fielding side to get batsmen out. Just like in baseball, if the ball is hit and then caught by a fielder before it touches the ground, the striker is out. He is also out if the bowler manages to deliver the ball such that the striker can not prevent it from knocking the wicket over. If the ball hits the striker's leg and an umpire rules it would have hit the wicket if the leg had not been there, the striker is out by "lbw" (leg before wicket). The reason is that you can not use your body to defend your wicket. There are more ways for batsmen to make out, but those are the most common ones. An important piece of terminology: when a batsman makes out, it is expressed as "a wicket has been taken/lost". Every time a player is out, the batting side loses a wicket and this plays an important role towards ending an innings.

There are two special cases when scoring runs. If the striker hits the ball over the boundary (the rope marking the outer edge of the cricket ground), then six runs are automatically scored. If the ball goes across the boundary after touching the ground at least one time, the batting side is awarded four runs. They are known as "sixes" and "fours" respectively, and also "boundaries".

So when does an innings end? Basically, there are two ways for innings to end: when 10 wickets are lost or when 50 overs have been played (we explain overs in the following paragraph). To be more precise, let's mention that there exists a third way which can end an innings prematurely: it occurs when the team batting second surpasses the run total of the team batting first. The reason why the innings is over once 10 wickets have been lost is because the batting side is left with only one player who has not made out. Thus, there are no remaining batsmen to start a new partnership.

Now, let's explain the concept of "overs". An over simply consists of six balls being bowled by the fielding side. As we said earlier, an innings can not go more than 50 overs, in other words 300 balls. Sometimes more than 300 balls may be delivered in one innings because of "no balls" and "wides". The former occurs when the bowler performs an illegal action while delivering the ball; the latter is called by an umpire if he feels the ball was unreachable to the striker. In both cases, the batting side gets one run and the ball is not counted as part of the over, thus explaining why more than 300 balls can be bowled in an innings.

What determines whether B1 or B2 is the striker on any given ball? Let's suppose the innings starts with B1 standing at the left end of the pitch, B2 standing at the right end, and B1 is the striker. Now, the bowler is going to deliver from the right end for one full over (6 balls). So, if an even number of runs is scored on the first ball the bowler will face B1 again. Otherwise, B2 will now be the one standing on the left end (he ends up being the striker in this case). When an over is completed, a new bowler comes in (bowling two consecutive overs is prohibited, just as substitutions in the middle of an over are prohibited) and he has to bowl from the other end of the pitch. Therefore, if an even number of runs (including 0 runs) has just been scored, then the same batsman is going to be the striker, unless an over has been completed. One more exception occurs if the striker makes out, in which case he is replaced by the next batsman in the batting order. Let's mention one more rule regarding bowling: in one-day matches, a bowler is not allowed to bowl more than 10 overs.

One rule that is strange to baseball fans is that batsmen in cricket do not have to run when the ball is hit! If they think that they do not have enough time to change places, they just stand still. This rule might seem bizarre, but recall that in one-day matches a limited number of balls is bowled. So, in order to score runs you eventually need to run and take some risks!

A cricket game can be won by runs or by wickets. Cricket terminology states that the magnitude of victory is measured in terms of runs if the team batting first won the match, otherwise it is measured in wickets. For example, if the team batting first scored 250 runs, followed by their opponents scoring 220 runs, then we say team 1 has beaten team 2 by 30 runs. On the other hand, had team 2 surpassed team 1's run total with seven wickets lost, then the correct expression is "team 2 won by three wickets". It expresses the fact that team 2 had three more wickets with which to keep batting, but they did not need them.

2.2 Common Statistics

Various statistics are reported concerning players and teams in any major sport. Indeed, sports fans like to rate players/teams, as well as point out good/bad performances accomplished over a certain period of time (hot/cold streaks). In some cases, the field of statistics plays a major role in sports. For example, the BCS (Bowl Championship Series) formula ranks major US college football teams by taking into account many factors like the number of wins and losses, the strength of the schedule, the margin of victory, whether the games were played on the road or at home, etc. Ultimately, the BCS formula determines which two teams meet in order to crown the national champion. Obviously, for a university to have its football team playing in the championship game is a really big deal because they receive publicity and millions of dollars in revenue.

Now, going back to statistics concerning players, we note that all major North American sports report statistics which are widely accepted among fans. For instance, everyone agrees that the number of points per game for a basketball player is a statistic that makes sense and is a good measure of how good a player is. The same comment can be made about the goals against average (usually called GAA) for a hockey goaltender, the number of points scored (goals + assists) for a hockey player, the batting average for a baseball player, the earned runs average (called ERA) for a baseball pitcher, the rushing yards for a football running back, and so forth. In one-day cricket, which is a major sport outside of America, there does not seem to be such widely accepted player statistics. All of them suffer from serious drawbacks in the sense that a weak player could end up looking like a great player based on these measures, and vice-versa.

Unlike other sports, it is not an easy task to define a good measure of performance in one-day cricket. We want to derive a statistic that takes into account what is really important in winning a match. So, let's ask ourselves the following question: what exactly is important in winning a match? Basically, a batsman should score a high number of runs relative to the resources used. On the other hand, a bowler should allow a small number of runs relative to the resources used. There is no problem determining the number of runs scored/allowed by a cricketer, but what about the "resources used"? A depletion of resources consists of a loss of overs and/or wickets. Therefore, a good measure of a cricketer's performance must account for three things: runs, overs and wickets.

As shall be presented here, every currently reported cricket statistic is missing an important piece of information. With respect to batting, there are two statistics that are widely reported, each of which purports excellence via large values. The first statistic is the *batting average*. It is defined as the total number of runs scored by a batsman divided by the number of innings in which he was dismissed (i.e. the number of wickets lost). The problem with this statistic is that overs are not taken into account. In his explanation of cricket on the Cricinfo website (www.cricinfo.org), David Mar states that "A batting average above 30 is very good, 40 excellent, and 50 is legendary." A pathological case can be seen where a batsman scores a total of 100 runs during 100 overs, but is dismissed only once (which is not that difficult to achieve if you are batting in a very conservative way). Such a batsman would have an incredibly high batting average of 100.0 yet would be a detriment to his side as he scores only one run per over! At this rate, his side would score 50 runs, which is extremely low compared to the average of 233.1 (with standard error 3.0) in ODI matches from mid-1997 through early 2002 (Duckworth and Lewis, 2002).

The second batting statistic which is widely reported is called the *strike rate*. It is calculated as the number of runs scored by a batsman per 100 balls. It reflects the number of overs, but it does not account for wickets lost. It is generally agreed that only very good batsmen have a strike rate above 80. Now, consider a player batting according to the pattern of scoring a six on the first ball and then being dismissed on the second ball. Such a batsman would have an incredibly high strike rate of 300.0 yet would hurt his team as he uses up wickets so quickly. At this rate, his side would lose all of its 10 wickets after scoring only 60 runs!

There exist three widely reported statistics with respect to bowling. Each of them accounts for only two of the three important pieces of information (runs, overs, wickets) which are necessary to measure excellence in a sensible way. The first statistic is called the *bowling average*. It is similar to the batting average in the sense that it measures the runs per wickets lost. However, instead of counting the number of runs scored we are now interested in runs allowed. Also notice that small values are deemed to be good. The problem with this statistic can be seen through the following example: if both bowlers A and B have allowed 100 runs and taken 1 wicket, but bowler A has been involved in only 10 overs compared to 100 overs for bowler B, which of the two would you choose on your team? Based on their high bowling average of 100.0, they both seem to be weak bowlers (an average of 40 is considered to be bad). In reality, many teams would love to have bowler B on their team because he allows only one run per over (i.e. 50 runs over a full game)!

The *economy rate* is another popular statistic. It is defined as the total number of runs allowed by the bowler divided by the number of overs bowled. Even though it is interesting to know how many runs, on average, a bowler allows per over, there is no mention of the number of wickets that have been taken by him. So, two different bowlers could have the same economy rate with one of them having obtained twice as many wickets as the other one.

Finally, the bowler's *strike rate* is defined as the total number of balls bowled divided by the number of wickets taken. Again, this sounds like a sensible measure of excellence, but it is missing a very important piece of information: the number of runs allowed. Clearly, no one is going to complain about a bowler who is rarely obtaining wickets if he is tough to score runs against.

Several attempts have been made in the past to find more sensible measures in cricket (either test or one-day). For example, Kimber and Hansford (1993) proposed an alternative to the batting average by considering interevent times (where each dismissal is viewed as an event). They considered the number of runs scored by a batsman between successive times out. They pointed out that "Such processes commonly arise, though not usually with discrete interevent times, in repairable systems reliability."

It was argued in this section that it is not an easy task to come up with a batting/bowling statistic which is going to reflect a cricketer's performance on the field. The problem arises from the difficulty of measuring the resources that have been used by a player. All of the currently reported statistics suffer from serious drawbacks. We propose in the following chapter a new statistic which is more sensible because it takes into account what is really important in helping your side's cause towards winning a match. Also, this statistic (which we will call the "runs per match" statistic) is easily interpretable and is applicable to both batsmen and bowlers.

Chapter 3

New Statistics

3.1 The Approach

In the previous chapter, we expressed the need for a new measure of batting and bowling prowess in one-day cricket. Clearly, batsmen do well if they score runs at a high rate and bowlers do well if they allow runs at a low rate. The problem is to determine the relevant units when calculating the rate. We have mentioned earlier that the units should be a combination of the number of overs completed and the number of wickets taken during a cricketer's tenure as either a batsman or a bowler. A solution to this problem comes from the Duckworth-Lewis (D/L) method. It was originally devised to improve "fairness" in interrupted one-day cricket matches. However, it can be used in the context of defining a new measure of excellence because it expresses resources as a function of both overs and wickets, which is exactly what we require. Thus, we will briefly review the D/L approach before moving on to the definition of the brand new batting/bowling statistic.

Most cricket matches are played outside, thus causing some of them to be interrupted by external factors like rain. Other reasons like crowd trouble can sometimes lead to a delay in a match. When this situation arises, it is necessary to adjust the teams' scores because the match must terminate in one day. For example, if the team batting second was deprived of 12 overs (one over is deducted for every 3 minutes lost), they are then going to bat for only 38 overs. Obviously, this team has to be compensated for the loss of overs. That is, the target they must achieve in order to win the match has to be lowered. Since 1997, the Duckworth-Lewis method has been used throughout senior levels in these situations. Its main goal is to reset the target in games which have been shortened for some reason.

The method is based on the following model, which calculates *Z(u,w)*, the average number of runs to be scored when u overs are left to be played ($0 \leq u \leq 50$) and w wickets have been lost ($0 \leq w \leq 9$):

$$Z(u,w) = Z_0 \cdot F(w) \left[1 - exp\left(\frac{-b_0 u}{F(w)} \right) \right]$$

Thus, three terms need to be estimated: the positive constants *Zo* and *bo*, and also the function *F(w)*. The latter is a positive decreasing function. It is interpreted as the average runs scored when w wickets have been lost divided by the average runs scored when no wickets have been lost, both of them seen as if there was an infinite number of overs remaining. For instance, *F(0)* must be 1. This function was estimated based on Duckworth and Lewis' knowledge of cricket. Then, estimates of *Zo* and *bo* were obtained by fitting the model to some data. Figure B.1 plots the average runs scored versus the number of overs remaining for w = 0, 1, ..., 9. This graph was obtained from Duckworth and Lewis (1998).

Now, the quantity that is of more interest is the proportion of resources remaining, *P(u,w)*, when u overs are left to be played and w wickets have been lost. It can be easily obtained once we have *Z(u,w)* for $0 \leq u \leq 50$ and $0 \leq w \leq 9$. Indeed,

$$P(u,w) = \frac{Z(u,w)}{Z(50,0)}$$

Table A.1 presents values of *P(u,w)* for some u's and w's, allowing us to determine the fraction of resources left from these positions in an innings.

Let's go back to the problem of interest, that is defining a sensible measure of performance in ODI cricket. The Duckworth-Lewis model provides the relevant units regarding the depletion of resources. Therefore, we propose a new statistic, the *runs per match* for a cricketer, defined as

$$\text{RM} = 100 \cdot \left(\frac{\text{total number of runs}}{\text{total resources used}} \right)$$

Note that the totals are taken over all of the cricketer's appearances.

The statistic RM is appealing as it is applicable to both batsmen and bowlers. Also, its definition probably provides the best way to measure a player's performance. Unlike the actual reported statistics, a weak batsman (bowler) can not have a high (low) value of RM. However, the main obstacle in using the measure is the difficulty of calculation. The calculation is not difficult in a mathematical or computational sense. Rather, the difficulty arises from the form in which one-day cricket data is recorded. Extracting the necessary data from the Cricinfo website requires an enormous effort.

One more appealing feature of the RM statistic is its interpretability. The RM statistic can be thought of as the number of runs obtained had the cricketer initiated batting/bowling for his side and was allowed to continue until all of the overs or wickets were used. Recall from section 2.2 that an average of 233.1 runs were scored (with standard error 3.0) in ODI matches from mid-1997 through early 2002. We can then use this piece of information as a standard for determining the excellence of batsmen and bowlers. Clearly, a good batsman has an RM value much larger than 233, and a good bowler has an RM value well below 233.

Not often do we see (if ever) standard errors reported on sport statistics. They are of interest to us since they allow us to make comparisons between various players. Moreover, standard errors turn out to be useful tools when it comes to determining whether a given player is significantly better than another player.

Consider a cricketer (either a batsman or bowler) for whom x_i runs and r_i resources are recorded in the i-th appearance, $i = 1, \ldots, n$. Now, let

$$\mathrm{T} = \frac{\mathrm{RM}}{100} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} r_i} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i}{\frac{1}{n}\sum_{i=1}^{n} r_i}$$

From the multivariate central limit theorem, we have the approximate result

$$\begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} x_i \\ \frac{1}{n}\sum_{i=1}^{n} r_i \end{pmatrix} \sim \mathrm{Normal}_2 \left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \frac{1}{n} \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right]$$

Using a Taylor series expansion and an application of the Delta theorem, we get:

$$\begin{aligned} \frac{\frac{1}{n}\sum_{i=1}^{n} x_i}{\frac{1}{n}\sum_{i=1}^{n} r_i} &\approx \frac{\mu_1}{\mu_2} + \left(\frac{\partial \frac{\mathrm{t}_1}{\mathrm{t}_2}}{\partial \mathrm{t}_1} \frac{\partial \frac{\mathrm{t}_1}{\mathrm{t}_2}}{\partial \mathrm{t}_2} \right) \Bigg|_{\substack{\mathrm{t}_1=\mu_1 \\ \mathrm{t}_2=\mu_2}} \begin{pmatrix} \mathrm{t}_1 - \mu_1 \\ \mathrm{t}_2 - \mu_2 \end{pmatrix} \\ &= \frac{\mu_1}{\mu_2} + \left(\frac{1}{\mathrm{t}_2} \frac{-\mathrm{t}_1}{\mathrm{t}_2{}^2} \right) \Bigg|_{\substack{\mathrm{t}_1=\mu_1 \\ \mathrm{t}_2=\mu_2}} \begin{pmatrix} \mathrm{t}_1 - \mu_1 \\ \mathrm{t}_2 - \mu_2 \end{pmatrix} \\ &= \frac{\mu_1}{\mu_2} + \left(\frac{1}{\mu_2} \frac{-\mu_1}{\mu_2^2} \right) \begin{pmatrix} \mathrm{t}_1 - \mu_1 \\ \mathrm{t}_2 - \mu_2 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
\Rightarrow \mathrm{Var}(\mathrm{T}) &= \mathrm{Var}\left(\frac{\frac{1}{n}\sum_{i=1}^{n} x_i}{\frac{1}{n}\sum_{i=1}^{n} r_i}\right) \\
&\approx \left(\frac{1}{\mu_2}\, \frac{-\mu_1}{\mu_2^2}\right) \mathrm{Var}\begin{pmatrix} \mathrm{t}_1 - \mu_1 \\ \mathrm{t}_2 - \mu_2 \end{pmatrix} \begin{pmatrix} \frac{1}{\mu_2} \\ \frac{-\mu_1}{\mu_2^2} \end{pmatrix} \\
&= \left(\frac{1}{\mu_2}\, \frac{-\mu_1}{\mu_2^2}\right) \mathrm{Var}\begin{pmatrix} \mathrm{t}_1 \\ \mathrm{t}_2 \end{pmatrix} \begin{pmatrix} \frac{1}{\mu_2} \\ \frac{-\mu_1}{\mu_2^2} \end{pmatrix} \\
&= \frac{1}{n}\left(\frac{1}{\mu_2}\, \frac{-\mu_1}{\mu_2^2}\right) \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \begin{pmatrix} \frac{1}{\mu_2} \\ \frac{-\mu_1}{\mu_2^2} \end{pmatrix}
\end{aligned}$$

In order to get $\widehat{\mathrm{Var}}(\mathrm{T})$, we estimate each of those quantities with consistent estimators:

- $\hat{\mu_1} = \frac{1}{n}\sum_{i=1}^{n} x_i$ because

$$\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}[x_i] = \frac{1}{n}\sum_{i=1}^{n} \mu_1 = \frac{1}{n}\cdot n \cdot \mu_1 = \mu_1$$

- $\hat{\mu_2} = \frac{1}{n}\sum_{i=1}^{n} r_i$ because

$$\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} r_i\right] = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}[r_i] = \frac{1}{n}\sum_{i=1}^{n} \mu_2 = \frac{1}{n}\cdot n \cdot \mu_2 = \mu_2$$

- $\hat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ because

$$\begin{aligned}
\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right] &= \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[x_i^2\right] - \mathrm{E}\left[\bar{x}^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{\mathrm{E}^2\left[x_i\right] + \mathrm{Var}\left(x_i\right)\} - \{\mathrm{E}^2\left[\bar{x}\right] + \mathrm{Var}\left(\bar{x}\right)\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\{\mu_1^2 + \sigma_1^2\} - \{\mu_1^2 + \frac{\sigma_1^2}{n}\} \\
&= \mu_1^2 + \sigma_1^2 - \mu_1^2 - \frac{\sigma_1^2}{n} \\
&= \frac{n-1}{n}\sigma_1^2
\end{aligned}$$

  Now, $\frac{n-1}{n}\sigma_1^2 \overset{n\to\infty}{\longrightarrow} \sigma_1^2$ so $\hat{\sigma}_1^2$ is a consistent estimator of σ_1^2.

- $\hat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2$ because

$$\begin{aligned}
\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2\right] &= \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^{n} r_i^2 - n\bar{r}^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[r_i^2\right] - \mathrm{E}\left[\bar{r}^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{\mathrm{E}^2\left[r_i\right] + \mathrm{Var}\left(r_i\right)\} - \{\mathrm{E}^2\left[\bar{r}\right] + \mathrm{Var}\left(\bar{r}\right)\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\{\mu_2^2 + \sigma_2^2\} - \{\mu_2^2 + \frac{\sigma_2^2}{n}\} \\
&= \mu_2^2 + \sigma_2^2 - \mu_2^2 - \frac{\sigma_2^2}{n} \\
&= \frac{n-1}{n}\sigma_2^2
\end{aligned}$$

  Now, $\frac{n-1}{n}\sigma_2^2 \overset{n\to\infty}{\longrightarrow} \sigma_2^2$ so $\hat{\sigma}_2^2$ is a consistent estimator of σ_2^2.

- $\hat{\sigma}_{12} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(r_i - \bar{r})$ because

$$\begin{aligned}
\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(r_i - \bar{r})\right] &= \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^{n} x_i r_i - \bar{x}\sum_{i=1}^{n} r_i - \bar{r}\sum_{i=1}^{n} x_i + n\bar{x}\bar{r}\right] \\
&= \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^{n} x_i r_i - n\bar{x}\bar{r} - n\bar{x}\bar{r} + n\bar{x}\bar{r}\right] \\
&= \frac{1}{n}\mathrm{E}\left[\sum_{i=1}^{n} x_i r_i - n\bar{x}\bar{r}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{\mathrm{E}\left[x_i r_i\right]\} - \mathrm{E}\left[\bar{x}\bar{r}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\{\mathrm{Cov}\left(x_i, r_i\right) + \mathrm{E}\left[x_i\right]\mathrm{E}\left[r_i\right]\} \\
&\quad - \left(\mathrm{Cov}\left(\bar{x}, \bar{r}\right) + \mathrm{E}\left[\bar{x}\right]\mathrm{E}\left[\bar{r}\right]\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\{\sigma_{12} + \mu_1\mu_2\} - \left(\frac{\sigma_{12}}{n} + \mu_1\mu_2\right) \\
&= \sigma_{12} + \mu_1\mu_2 - \frac{\sigma_{12}}{n} - \mu_1\mu_2 \\
&= \frac{n-1}{n}\sigma_{12}
\end{aligned}$$

where

$$\begin{aligned}
\mathrm{Cov}\left(\bar{x}, \bar{r}\right) &= \mathrm{Cov}\left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{j=1}^{n} r_j}{n}\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}\left(x_i, r_j\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Cov}\left(x_i, r_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\{\sigma_{12}\} \\
&= \frac{n\sigma_{12}}{n^2} \\
&= \frac{\sigma_{12}}{n}
\end{aligned}$$

and where we used the fact that $\text{Cov}(x_i, r_j) = 0$ if $i \neq j$ (i.e. runs scored in the i-th appearance are independent of resources used in the j-th appearance).

Now, $\frac{n-1}{n}\sigma_{12} \overset{n\to\infty}{\longrightarrow} \sigma_{12}$ so $\hat{\sigma}_{12}$ is a consistent estimator of σ_{12}.

Finally, $T = \frac{RM}{100}$ so

$$\widehat{\text{Var}}\,(\text{RM}) = \frac{10000}{n}\left(\frac{1}{\bar{r}}\,\frac{-\bar{x}}{\bar{r}^2}\right)\begin{pmatrix} \frac{n-1}{n}S_x^2 & \frac{n-1}{n}S_{xr} \\ \frac{n-1}{n}S_{xr} & \frac{n-1}{n}S_r^2 \end{pmatrix}\begin{pmatrix} \frac{1}{\bar{r}} \\ \frac{-\bar{x}}{\bar{r}^2} \end{pmatrix}$$

where S_x^2 is the sample variance of the runs scored, S_r^2 is the sample variance of the resources used and S_{xr} is the sample covariance between these two variables.

Before turning to the analysis of ODI data, we offer two cautionary remarks. The first applies generally to all batting and bowling statistics and is based on the recognition that a team's strategy in the second innings depends largely on the number of runs scored in the first innings. If the team batting first scores a large number of runs, then batsmen in the second innings tend to be more aggressive. Also, the fielding side in the second innings is more willing to allow singles and low scoring strokes in the hope of preventing 4's, 5's and 6's. For this reason, we also recommend the calculation of separate statistics for both the first and second innings. When substantial differences arise, it may provide some added insight. To our knowledge, this has not been previously done with any of the traditional statistics.

The second remark primarily concerns upper right hand entries in the D/L resource table. In many instances, as one goes down these columns, the number of resources does not change. The implication of this for the RM statistic is that a cricketer can accumulate runs in the numerator with no resources being added in the denominator. Although this is clearly not sensible, in practice, there are very few matches where the upper right hand section of the table is relevant. In other words, not often do we see a team losing many wickets (about 7 or more) within the first few overs played in the match (about 20 or less).

3.2 Batting

In the context of batting, the runs per match measure translates to the total number of runs scored by a batsman divided by the total number of resources used by the batsman taken over all matches. Although the numerator is easily obtained, the calculation of the denominator is problematic.

Ideally, one would calculate the denominator of the RM statistic from each of the batter's "by ball" contributions. Recall that 2 batsmen are up at a time. So, even if the batter we are interested in was on the field for 10 overs (60 balls), he might have faced only 25 balls. The ideal way to come up with the resources used by a batsman is to take the difference in resources remaining prior to and after every ball bowled to him. This is what we meant by a "by ball" contribution. Let's go over the concept once again. Prior to the delivery of every ball, the status of the game dictates the current number of resources remaining according to the D/L table. After the ball is bowled to the cricketer of interest, the number of resources change according to the reduction in overs remaining and whether he makes out. The difference in resources provides the batter's "by ball" contribution to the denominator. Note that interpolation is typically used in D/L applications where a single ball corresponds to 1/6 of an over. Now, we must add up the batsman's contributions over all balls he faced and over all games in which he made an appearance! This is too huge of a task to accomplish because the data on the Cricinfo website is presented in the format of commentary. Therefore, a different approach was taken in order to approximate the denominator.

In practice, we approximate the denominator of the RM statistic using the following procedure. For each partnership in which a batsman is involved, we approximate the number of resources used by the batsman and this forms a "by partnership" contribution to the denominator. As a first step, we obtain the number of resources remaining at the beginning of the partnership according to the Duckworth-Lewis table. We then obtain the number of resources remaining one ball prior to the end of the partnership (i.e when one of the batsmen makes out or the match terminates). The difference in resources between the two states describes the resources lost during the partnership up to but not including the final ball. Now, the two batsmen may not be equally responsible for the resources used. So, we count the number of balls faced by each of the two players during the partnership (up to but not including the final ball). Then, we compute the fraction of balls faced by the batsman in question, which we multiply by the resources used. This provides a reasonable approximation to the resources used by the batsman. Finally, to this quantity we add the number of resources lost by the batsman on the final ball. We note that the batsman is not responsible for resources lost when his partner makes out. On the other hand, if he makes out, then he is fully responsible for the loss of resources.

We now consider the analysis of batting data obtained from the Cricinfo website. The data consists of all ODI matches from October 1998 through September 2002. We were unable to go further back in time as the commentary log, essential for data extraction, is unavailable. For our analysis, we considered batsmen listed in the top 50 according to lifetime batting average. From that list, we chose 12 active batsmen who have batted in a minimum of 93 ODI matches throughout their career.

In Table A.2, we present the batting average (BA), the strike rate (SR) and the runs per match statistic (RM) for these 12 batsmen based on the 1998-2002 data. Each player's ranking among the group of twelve is also provided in brackets for all 3 statistics. Also shown in the table are the standard errors for the runs per match statistic and the number of games played by each batsman during the time period. Notice that the batsmen are listed according to their calculated runs per match. The quantities required to calculate the variance of the runs per match statistic are presented in Table A.3. It may appear strange that four of the 12 batsmen have RM values below the average of 233.1 runs but one must keep in mind that only a small subset of batsmen do most of the batting. One more thing to point out is that Klusener (South Africa) and Lara (West Indies) have large standard errors associated with RM; this highlights their tendency to have dominant performances together with notably lesser performances.

We are interested in all pairwise comparisons to determine differences between the players according to the RM statistic. In other words, we assume that the runs per match statistic for the i-th cricketer

$$\mathrm{RM}_i \sim \mathrm{Normal}\left(\mu_i, \hat{\sigma}_i^2\right)$$

where the cricketers are independent and $\hat{\sigma}_i^2$ is assumed to be known (we use the estimates of variance in Table A.2). We are interested in testing

$$\begin{aligned} \mathrm{H_{0ij}} :& \mu_i = \mu_j \\ \mathrm{H_{1ij}} :& \mu_i \neq \mu_j \end{aligned}$$

for $i = 1, \ldots, 12$ and $j = 1, \ldots, 12$ $(i \neq j)$.

The test statistics can then be easily derived as follows:

$$\begin{aligned} \mathrm{RM}_i - \mathrm{RM}_j &\sim \mathrm{Normal}\left(\mu_i - \mu_j, \hat{\sigma}_i^2 + \hat{\sigma}_j^2\right) \\ \Rightarrow \frac{(\mathrm{RM}_i - \mathrm{RM}_j) - (\mu_i - \mu_j)}{\sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}} &\sim \mathrm{Normal}\,(0, 1) \end{aligned}$$

Under H_{0ij},

$$\mathrm{Z_{obs}} = \frac{\mathrm{RM}_i - \mathrm{RM}_j}{\sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}} \sim \mathrm{Normal}\,(0, 1)$$

$$\Rightarrow \mathrm{p - value} = 2 * \mathrm{P}\left(\mathrm{N}\,(0, 1) > |\mathrm{Z_{obs}}|\right)$$

We then considered two different levels of significance:

- Without taking into account the multiple comparisons problem

  $\Rightarrow$ level of significance is 0.05

- Bonferroni

  We make $m = \binom{12}{2} = 66$ comparisons

  $\Rightarrow$ level of significance is $0.05/66 = 0.00076$

Table A.4 presents the p-values obtained from all pairwise comparisons for the 12 batsmen. In bold face characters are those which turn out to be below 0.05. We summarize the comparative performances by a line plot, indicating each non-significant difference by a rule in Figures B.2 and B.3 (using 0.05 and 0.00076 as the levels of significance respectively). There does not seem to be clear-cut groups which can be formed. Bonferroni's method only declares SR Tendulkar (India) to be significantly different from Y Youhana (Pakistan). All of the other comparisons are non-significant.

Figure B.4 plots the runs per match statistic for the 12 batsmen in increasing order. Confidence limits which are calculated without taking into account the multiple comparisons problem are attached to the RM values. From this graph, it seems like there is a gap between the best 2 batsmen and the remaining 10. However, the standard errors are too large to pick up a significant difference between Tendulkar (India) + Klusener (South Africa) and the rest of the group.

We have mentioned earlier how the batting average and the strike rate suffer from serious drawbacks. Recall that they do not account for overs and wickets respectively. However, we could argue that some sort of combination of these 2 statistics would provide a sensible measure of performance (which is the idea behind the RM statistic). Thus, for each batsman we have summed their ranking based on BA and their ranking based on SR to get a "top 12 based on common statistics." We expect the latter to be fairly similar to the "top 12 based on the RM statistic." Both are presented in Table A.5 and we can see immediately that they are very comparable. This indicates that not only does the runs per match statistic has a sound theoretical basis, but it does not deviate greatly from common perceptions.

In Table A.6, we present the correlation coefficients between each pair of the three statistics taken over the 12 batsmen. All of the correlations turned out to be positive, which is what one would expect. Indeed, a good batsman should have high values of BA, SR and RM, whereas weak players should have low values of all three statistics. We observe that strike rate correlates highly (0.95) with runs per match. Given that the RM statistic is the most sensible measure because of its sound theoretical basis, this suggests that strike rate might be the preferred batting statistic among the currently reported ones. This seems to contradict prevailing wisdom that tends to prefer batting average as a measure of batting proficiency. Quoting David Mar in his explanation of cricket posted on the Cricinfo website, "Over a single player's career, the two most important statistics are ... batting average and bowling average."

Table A.7 shows the batting average, the strike rate and the RM statistic calculated by splitting the first and second innings appearances. The batsmen's rankings are also provided in brackets. Also shown in the table are the number of matches used in the calculations. One would expect the first and second innings statistics to be positively correlated since a good player should do well no matter the innings. Figures B.5, B.6 and B.7 plot the 12 batsmen's first innings versus second innings BA, SR and RM respectively.

At first glance, there does not seem to be much correlation in these plots, except for the strike rate graph. Table A.8 confirms this observation: the correlation is very small between the first and second innings RM (0.07) and it even turns out to be negative in the case of the batting average (-0.26)! One of the most extreme cases occurs for Klusener whose batting average is 35 in the first innings (ninth among the 12 batsmen) as opposed to 59 in the second innings (best among the group of 12 batters). Perhaps one explanation for this lack of correlation between innings is that we do not have a sufficiently large sample of data (i.e. 12 batsmen). Another explanation may have to do with one's team. If a batsman plays for a strong (weak) team, then he would more often need a small (large) number of runs in the second innings. As for SR, we obtained a fairly high correlation (0.69). Its graph shows an interesting feature: 11 of the 12 points lie below the $y = x$ line. Thus, the strike rate in the first innings is greater than the strike rate in the second innings for almost all of the batsmen in the analysis, suggesting a possibly less aggressive approach in the second innings. To see this, recall the definition of the strike rate, which is the number of runs scored per 100 balls. So, an aggressive player who is not afraid to make out will have a strike rate whose value is large.

3.3 Bowling

In the context of bowling, the runs per match measure translates to the total number of runs allowed by a bowler divided by the total number of resources used while bowling taken over all matches. The calculation of the denominator was problematic in the batting case because there were two batters up at a time. In the bowling case, its calculation is more straightforward since a bowler must bowl for a full over (no substitutions allowed during the course of an over, except in the case of injuries).

When calculating RM for bowlers, we consider every bowling appearance that a bowler makes and note that he usually makes several appearances in a match. However, recall from the rules of ODI cricket that a bowler can not bowl more than ten overs. The contribution to the numerator for each appearance is readily available from commentary logs. As we said earlier, the calculation of the denominator is also straightforward. Prior to each appearance, the status of the game dictates the current number of resources remaining according to the D/L table. After the appearance, the number of resources change according to the reduction in overs remaining (usually one over, but may be less in the case of a bowler getting hurt or the innings ending) and wickets being possibly taken during the bowler's appearance. The difference in resources between the two states forms a "by appearance" contribution to the denominator of the RM statistic. We therefore observe that although the calculation of the runs per match for bowlers is fairly easy, data extraction is time consuming due to the multiple appearances by a bowler during a match.

We now consider the analysis of bowling data obtained from the Cricinfo website. The data consists of all ODI matches from October 1998 through January 2003. Again, we were unable to go further back in time as the commentary log, essential for data extraction, is unavailable. For our analysis, we considered bowlers listed in the top 50 according to lifetime bowling average. From that list, we chose 12 active bowlers who have bowled in a minimum of 56 ODI matches.

In Table A.9, we present the bowling average (BA), the economy rate (ER), the strike rate (SR) and the runs per match statistic (RM) for these 12 bowlers based on the 1998-2003 data. Just as we did for batsmen, we also provide in brackets each player's ranking among the group of twelve (for all four statistics). Also shown in the table are the standard errors associated with the RM statistic and the number of appearances by each bowler during the 1998-2003 period. Notice that the bowlers are listed according to their calculated runs per match. The quantities required to calculate the variance of the runs per match statistic are presented in Table A.10.

It seems like the best pair of bowlers is more dominant than the best pair of batsmen. Indeed, batsmen Tendulkar and Klusener have a runs per match statistic 49.6 and 43.2 greater than the ODI average of 233.1. Now, the RM value of the best bowlers (Muralitharan and McGrath) is 77.9 and 69.6 smaller than the average! We should point out that all 12 bowlers have RM values below the average. Also, Khan's RM standard error is a bit large, suggesting a lack of consistency in his performances (he mixes very good performances with pretty bad ones). Finally, we observe that the standard errors of the RM statistic are smaller for bowlers than for batsmen.

Like batsmen, we are interested in making pairwise comparisons among bowlers. The same approach was used and we present the p-values obtained from all pairwise comparisons for the 12 bowlers in Table A.11. Once again the p-values which are below 0.05 are shown in bold. Figures B.8 and B.9 present line plots for the two methods considered in this project; that is, without accounting for the multiple comparisons problem and Bonferroni, respectively. We rely more on the first method because Bonferroni's approach is too conservative when making many comparisons (66 in this case). There does not seem to be clear-cut groups which can be formed. Figure B.10 is a useful tool in forming groups of bowlers as it plots the runs per match statistic for the 12 bowlers in increasing order. Confidence limits which are calculated without taking into account the multiple comparisons problem are attached to the RM values. From this graph it seems like we could form three groups of bowlers. Indeed, from Figure B.8 we see that M Muralitharan (Sri Lanka) and GD McGrath (Australia) could be in group 1, then S Akhtar (Pakistan), SM Pollock (South Africa) and AA Donald (South Africa) could be in group 2 with the remaining seven bowlers in group 3. However, the standard errors are too large to pick up these differences.

We came up with a "top 12 bowler ranking based on common statistics" in a similar way to the batting case. That is, we summed each bowler's ranking based on BA, ER and SR in order to obtain the top 12. It is presented in Table A.12, along with the "top 12 based on the RM statistic." We observe that the two rankings are similar, just as they were for batsmen.

In Table A.13, we present the correlation coefficients between each pair of the four statistics taken over the 12 bowlers. For the same reasons as those given in the batting section, we expect the correlations to be all positive. Surprisingly, one of them turned out to be negative (-0.16 for the correlation between ER and SR). The RM statistic correlates highly with BA and ER, the highest being 0.93 with the batting average. Given that the runs per match statistic has a sound theoretical basis yet is more difficult to calculate, this suggests that bowling average might be the preferred bowling statistic among the currently reported ones. Unlike the batting case, this does agree with the prevailing wisdom (which tends to prefer bowling average as a measure of bowling proficiency).

Table A.14 presents the bowling average, the economy rate, the strike rate and the RM statistic calculated by splitting the first and second innings appearances. The rankings are also presented in brackets. Also shown in the table are the number of appearances used in the calculations. Once again we expect the first and second innings values to be positively correlated for all four statistics. Figures B.11, B.12, B.13 and B.14 plot the 12 bowlers' first innings versus second innings BA, ER, SR and RM respectively. The same phenomenon as we saw in the batting section occurs here: there does not seem to be much correlation in these plots, except in one of them (the economy rate graph). The actual correlation coefficients are shown in Table A.15. We observe that all four of them are positive. However, the bowling average and the strike rate coefficients are pretty small (0.15 and 0.07 respectively). RM's correlation is 0.53, whereas ER's turns out to be the largest at 0.76. We keep in mind that outliers can greatly impact correlations.

It is interesting to note that the highest correlation occurred for the statistic which does not take wickets into account, both in the batting and bowling cases. A possible explanation is the fact that an additional wicket changes the value of a statistic much more than an additional run or over.

Chapter 4

Discussion

We have argued that the traditional measures of batting/bowling performance in one-day cricket all suffer from serious drawbacks. The main problem is the calculation of the number of resources used by a cricketer in each game. This project has proposed a new measure which we called the *runs per match* statistic. This new statistic uses the widely accepted Duckworth-Lewis method and it does not have any major flaws, unlike all of its predecessors. Moreover, the RM statistic is readily interpretable and has a symmetry that is applicable to both batting and bowling. The measure also yields a standard error. Recall that we calculated the number of resources used by summing over the "by partnership" contributions for batsmen and the "by appearance" contributions for bowlers.

We analyzed data on 12 of the best batsmen and bowlers in ODI cricket. Ranking the players according to the RM statistic and the more common measures led to similar results. This proves to be an indication that the statistic introduced in this project does not deviate greatly from commonly held perceptions. The most highly correlated statistics with the runs per match statistic were strike rate for batters and economy rate for bowlers. By looking at all measures, it seemed like Tendulkar and Klusener are the best batsmen, whereas Muralitharan and McGrath are the best bowlers (among their respective groups of 12 cricketers which we considered in our analysis).

At this time, the main obstacle in using the RM statistic is the difficulty of calculation. It arises from the form in which one-day cricket data is recorded. The calculation of this measure for a given player over numerous games requires a huge effort in data extraction. We suggest that ball by ball results on every match should be collected in matrix form, where rows correspond to balls and columns represent each of the possible outcomes. This would not only open up the possibility of the widespread use of the RM measure but permit deeper analyses of many aspects of ODI cricket. We remark that ball by ball results are currently recorded in a convenient form for all matches in the Australian Open tennis championship.

Appendix A

Tables

Table A.1: The Duckworth-Lewis resource table where the percentage of resources remaining is given as a function of the number of wickets lost and the number of overs available.

	Wickets Lost									
Overs Left	0	1	2	3	4	5	6	7	8	9
50	100.0	92.4	83.8	73.8	62.4	49.5	37.6	26.5	16.4	7.6
49	99.2	91.8	83.3	73.5	62.2	49.4	37.6	26.5	16.4	7.6
48	98.3	91.1	82.7	73.1	62.0	49.3	37.6	26.5	16.4	7.6
47	97.4	90.3	82.2	72.7	61.8	49.2	37.6	26.5	16.4	7.6
46	96.5	89.6	81.6	72.3	61.5	49.1	37.5	26.5	16.4	7.6
45	95.5	88.8	81.0	71.9	61.3	49.0	37.5	26.4	16.4	7.6
40	90.3	84.5	77.6	69.4	59.8	48.3	37.3	26.4	16.4	7.6
35	84.2	79.3	73.4	66.3	57.7	47.2	36.9	26.3	16.4	7.6
30	77.1	73.1	68.2	62.3	54.9	45.7	36.2	26.2	16.4	7.6
25	68.7	65.6	61.8	57.1	51.2	43.4	35.1	25.9	16.4	7.6
20	58.9	56.7	54.0	50.6	46.1	40.0	33.2	25.2	16.3	7.6
15	47.5	46.1	44.4	42.1	39.1	35.0	30.0	23.7	16.0	7.6
10	34.1	33.4	32.5	31.4	29.8	27.5	24.6	20.6	14.9	7.5
5	18.4	18.2	17.9	17.6	17.1	16.4	15.5	14.0	11.5	7.0
4	14.9	14.8	14.6	14.4	14.1	13.6	13.0	11.9	10.2	6.6
3	11.4	11.3	11.2	11.1	10.9	10.6	10.2	9.6	8.5	6.0
2	7.7	7.7	7.6	7.6	7.5	7.4	7.2	6.9	6.3	4.9
1	3.9	3.9	3.9	3.9	3.9	3.8	3.8	3.7	3.5	3.1
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table A.2: Batting statistics for 12 of the best active batsmen in ODI cricket listed in order of their runs per match (RM). The batting average (BA) and the strike rate (SR) are also provided along with the standard error for RM and the number of matches used in the calculations.

Batsman		BA	SR	RM	Standard error(RM)	Number of games
SR Tendulkar	IND	51.2 (1)	89.0 (2)	282.7 (1)	14.2	84
L Klusener	SA	45.0 (4)	90.1 (1)	276.3 (2)	21.0	68
RT Ponting	AUS	40.6 (6)	79.6 (3)	251.4 (3)	13.9	74
SC Ganguly	IND	40.7 (5)	76.9 (5)	248.4 (4)	11.5	111
BC Lara	WI	37.2 (10)	79.0 (4)	240.0 (5)	21.2	57
MG Bevan	AUS	49.1 (2)	74.3 (6)	239.7 (6)	13.8	64
RP Arnold	SL	39.6 (7)	72.8 (7)	237.0 (7)	13.4	69
JH Kallis	SA	46.5 (3)	68.9 (10)	235.8 (8)	09.9	86
NV Knight	ENG	37.6 (9)	69.3 (9)	231.0 (9)	15.7	48
G Kirsten	SA	37.0 (11)	69.5 (8)	230.5 (10)	12.9	71
MS Atapattu	SL	36.5 (12)	65.8 (12)	224.2 (11)	11.3	89
Y Youhana	PAK	38.1 (8)	68.3 (11)	216.3 (12)	12.9	83

Table A.3: Components of the variance of the runs per match statistic for 12 of the best active batsmen in ODI cricket.

Batsman		n	$\bar{x}$	$\bar{r}$	S_x^2	S_r^2	S_{xr}
SR Tendulkar	IND	84	45.1	15.9	2009.7	98.2	416.9
L Klusener	SA	68	26.4	9.6	504.5	23.2	72.9
RT Ponting	AUS	74	35.6	14.2	978.6	49.4	199.1
SC Ganguly	IND	111	38.9	15.7	1576.3	88.6	353.6
BC Lara	WI	57	31.4	13.1	957.5	49.5	166.4
MG Bevan	AUS	64	32.2	13.4	706.3	50.1	160.8
RP Arnold	SL	69	28.7	12.1	510.5	39.0	115.3
JH Kallis	SA	86	39.0	16.5	913.5	70.1	226.6
NV Knight	ENG	48	33.6	14.6	877.1	55.8	198.8
G Kirsten	SA	71	34.9	15.2	1037.1	55.9	229.6
MS Atapattu	SL	89	32.8	14.6	793.9	50.7	178.5
Y Youhana	PAK	83	33.0	15.3	935.5	56.1	202.1

Table A.4: P-values of pairwise multiple comparisons for 12 of the best active batsmen in ODI cricket with values less than 0.05 in bold face.

	Klus.	Ponting	Ganguly	Lara	Bevan	Arnold	Kallis	Knight	Kirsten	Atap.	Youha.
Tendulkar	0.8036	0.1155	0.0616	0.0947	**0.0300**	**0.0193**	**0.0070**	**0.0147**	**0.0066**	**0.0013**	**0.0005**
Klusener		0.3215	0.2441	0.2236	0.1446	0.1141	0.0815	0.0836	0.0629	**0.0289**	**0.0149**
Ponting			0.8696	0.6541	0.5495	0.4554	0.3627	0.3298	0.2697	0.1291	0.0640
Ganguly				0.7282	0.6268	0.5180	0.4093	0.3705	0.3000	0.1343	0.0634
Lara					0.9886	0.9036	0.8578	0.7307	0.6994	0.5092	0.3387
Bevan						0.8895	0.8218	0.6768	0.6258	0.3855	0.2158
Arnold							0.9444	0.7694	0.7246	0.4642	0.2649
Kallis								0.7929	0.7412	0.4394	0.2301
Knight									0.9803	0.7262	0.4709
Kirsten										0.7148	0.4383
Atapattu											0.6472

Table A.5: Rankings based on common statistics (batting average (BA) and strike rate (SR)) and rankings based on runs per match (RM) for 12 of the best active batsmen in ODI cricket.

Batsman		Ranking based on BA and SR	Ranking based on RM
SR Tendulkar	India	1	1
L Klusener	South Africa	2	2
RT Ponting	Australia	4	3
SC Ganguly	India	5	4
BC Lara	West Indies	7.5	5
MG Bevan	Australia	3	6
RP Arnold	Sri Lanka	7.5	7
JH Kallis	South Africa	6	8
NV Knight	England	9	9
G Kirsten	South Africa	10.5	10
MS Atapattu	Sri Lanka	12	11
Y Youhana	Pakistan	10.5	12

Table A.6: Sample correlation coefficients for the batting average (BA), strike rate (SR) and runs per match (RM) statistic based on 12 of the best active batsmen in ODI cricket.

	BA	SR	RM
BA		0.56	0.67
SR			0.95

Table A.7: First and second innings statistics (batting average (BA), strike rate (SR) and runs per match (RM)) for 12 of the best active batsmen in ODI cricket. The number of matches used in the calculations are also provided.

Batsman		BA 1st inn	BA 2nd inn	SR 1st inn	SR 2nd inn	RM 1st inn	RM 2nd inn	# games 1st inn	# games 2nd inni
SR Tendulkar	IND	63.5 (1)	41.2 (7)	91.4 (1)	86.2 (2)	298.3 (1)	265.4 (2)	37	47
L Klusener	SA	34.6 (9)	58.9 (1)	84.1 (4)	95.5 (1)	237.7 (8)	317.3 (1)	36	32
RT Ponting	AUS	40.8 (7)	40.3 (9)	84.2 (3)	75.2 (4)	268.6 (2)	235.6 (8)	33	41
SC Ganguly	IND	42.7 (5)	39.1 (10)	78.2 (5)	75.7 (3)	249.9 (4)	247.0 (3)	50	61
BC Lara	WI	41.6 (6)	32.9 (11)	84.6 (2)	72.9 (5)	261.0 (3)	217.9 (11)	25	32
MG Bevan	AUS	48.0 (3)	50.4 (3)	76.9 (6)	71.4 (7)	243.1 (7)	235.9 (7)	32	32
RP Arnold	SL	32.2 (11)	55.4 (2)	73.0 (8)	72.6 (6)	231.3 (10)	244.5 (4)	41	28
JH Kallis	SA	49.9 (2)	43.4 (4)	71.2 (10)	66.6 (10)	249.6 (5)	222.5 (10)	41	45
NV Knight	ENG	35.4 (8)	40.9 (8)	71.4 (9)	66.6 (9)	235.5 (9)	225.3 (9)	26	22
G Kirsten	SA	31.5 (12)	41.8 (6)	70.2 (11)	69.1 (8)	220.0 (11)	237.8 (5)	31	40
MS Atapattu	SL	32.3 (10)	42.0 (5)	67.2 (12)	64.5 (11)	212.1 (12)	237.5 (6)	47	42
Y Youhana	PAK	45.2 (4)	26.2 (12)	75.1 (7)	54.1 (12)	244.4 (6)	162.6 (12)	52	31

Table A.8: Correlation coefficients between the first and second innings batting statistics for 12 of the best active batsmen in ODI cricket.

Batting statistic	Correlation between first and second innings statistics
Batting average	-0.26
Strike rate	0.69
Runs per match	0.07

Table A.9: Bowling statistics for 12 of the best active bowlers in ODI cricket listed in order of their runs per match (RM). The bowling average (BA), the economy rate (ER) and the strike rate (SR) are also provided along with the standard error for RM and the number of appearances used in the calculations.

Bowler		BA	ER	SR	RM	Standard error(RM)	Number of appearances
M Muralitharan	Sri Lanka	18.2 (1)	3.35 (1)	32.5 (5)	155.2 (1)	7.8	785
GD McGrath	Australia	19.9 (2)	3.79 (3)	31.5 (4)	163.5 (2)	10.3	657
S Akhtar	Pakistan	20.6 (3)	4.47 (10)	27.6 (1)	185.2 (3)	13.3	473
SM Pollock	South Africa	23.0 (5)	3.72 (2)	37.1 (10)	186.6 (4)	9.0	910
AA Donald	South Africa	21.1 (4)	4.29 (5)	29.6 (2)	188.6 (5)	13.6	410
W Akram	Pakistan	24.9 (8)	4.02 (4)	37.1 (11)	204.7 (6)	11.2	708
D Gough	England	25.4 (9)	4.37 (8)	34.9 (8)	207.0 (7)	12.1	513
A Razzaq	Pakistan	24.7 (7)	4.36 (7)	34.0 (7)	216.9 (8)	12.3	713
W Younis	Pakistan	24.7 (6)	4.87 (12)	30.4 (3)	217.5 (9)	13.5	606
SK Warne	Australia	29.0 (12)	4.35 (6)	40.1 (12)	219.6 (10)	12.2	657
S Mushtaq	Pakistan	27.3 (11)	4.45 (9)	36.9 (9)	226.3 (11)	13.2	579
Z Khan	India	26.6 (10)	4.81 (11)	33.1 (6)	228.3 (12)	15.2	449

Table A.10: Components of the variance of the runs per match statistic for 12 of the best active bowlers in ODI cricket.

Bowler		n	$\bar{x}$	$\bar{r}$	S_x^2	S_r^2	S_{xr}
M Muralitharan	Sri Lanka	785	3.33	2.15	7.7	5.6	-0.24
GD McGrath	Australia	657	3.75	2.30	12.3	8.3	-0.68
S Akhtar	Pakistan	473	4.44	2.40	11.6	8.8	-1.69
SM Pollock	South Africa	910	3.69	1.98	9.5	5.0	-0.49
AA Donald	South Africa	410	4.28	2.27	10.7	6.7	-1.17
W Akram	Pakistan	708	4.00	1.96	9.7	5.6	-0.23
D Gough	England	513	4.36	2.11	9.7	5.3	-0.16
A Razzaq	Pakistan	713	4.34	2.00	10.1	6.1	-1.02
W Younis	Pakistan	606	4.85	2.23	11.9	8.2	-0.97
SK Warne	Australia	657	4.33	1.97	10.5	5.2	-0.68
S Mushtaq	Pakistan	579	4.39	1.94	9.8	5.3	-0.27
Z Khan	India	449	4.80	2.10	13.1	6.0	-0.31

Table A.11: P-values of pairwise multiple comparisons for 12 of the best active bowlers in ODI cricket with values less than 0.05 in bold face.

	McGr.	Akhtar	Pollock	Donald	Akram	Gough	Razzaq	Younis	Warne	Mush.	Khan
Muralitha.	0.5198	0.0521	**0.0084**	**0.0330**	**0.0003**	**0.0003**	**2e-05**	**6e-05**	**9e-06**	**4e-06**	**2e-05**
McGrath		0.1985	0.0918	0.1412	**0.0069**	**0.0062**	**0.0009**	**0.0014**	**0.0005**	**0.0002**	**0.0004**
Akhtar			0.9293	0.8563	0.2617	0.2237	0.0801	0.0871	0.0565	**0.0283**	**0.0322**
Pollock				0.9015	0.2083	0.1750	**0.0470**	0.0560	**0.0296**	**0.0130**	**0.0179**
Donald					0.3610	0.3108	0.1228	0.1302	0.0896	**0.0467**	0.0508
Akram						0.8880	0.4649	0.4642	0.3688	0.2135	0.2102
Gough							0.5678	0.5613	0.4635	0.2823	0.2714
Razzaq								0.9715	0.8740	0.6029	0.5572
Younis									0.9080	0.6430	0.5938
Warne										0.7124	0.6548
Mushtaq											0.9180

Table A.12: Rankings based on common statistics (bowling average (BA), economy rate (ER) and strike rate (SR)) and rankings based on runs per match (RM) for 12 of the best active bowlers in ODI cricket.

Bowler		Ranking based on BA, ER and SR	Ranking based on RM
M Muralitharan	Sri Lanka	1	1
GD McGrath	Australia	2	2
S Akhtar	Pakistan	4	3
SM Pollock	South Africa	5	4
AA Donald	South Africa	3	5
W Akram	Pakistan	8	6
D Gough	England	9	7
A Razzaq	Pakistan	6.5	8
W Younis	Pakistan	6.5	9
SK Warne	Australia	12	10
S Mushtaq	Pakistan	11	11
Z Khan	India	10	12

Table A.13: Sample correlation coefficients for the bowling average (BA), economy rate (ER), strike rate (SR) and runs per match (RM) statistic based on 12 of the best active bowlers in ODI cricket.

	BA	ER	SR	RM
BA		0.61	0.68	0.93
ER			-0.16	0.81
SR				0.40

Table A.14: First and second innings statistics (bowling average (BA), economy rate (ER), strike rate (SR) and runs per match (RM)) for 12 of the best active bowlers in ODI cricket. The number of appearances these statistics are based on are also provided.

Bowler		BA 1st inn	BA 2nd inn	ER 1st inn	ER 2nd inn	SR 1st inn	SR 2nd inn
M Muralitharan	Sri Lanka	23.3 (7)	14.6 (1)	3.73 (2)	3.00 (1)	37.5 (11)	29.1 (5)
GD McGrath	Australia	21.6 (3)	18.0 (3)	3.77 (3)	3.81 (3)	34.4 (9)	28.4 (4)
S Akhtar	Pakistan	22.3 (5)	19.2 (4)	4.42 (8)	4.52 (9)	30.3 (2)	25.5 (1)
SM Pollock	South Africa	20.6 (1)	27.2 (9)	3.71 (1)	3.72 (2)	33.3 (6)	43.9 (11)
AA Donald	South Africa	23.0 (6)	17.9 (2)	4.46 (9)	3.96 (5)	31.0 (3)	27.2 (2)
W Akram	Pakistan	21.7 (4)	28.5 (10)	4.11 (4)	3.95 (4)	31.7 (5)	43.3 (10)
D Gough	England	20.9 (2)	34.0 (12)	4.16 (5)	4.64 (11)	30.2 (1)	44.0 (12)
A Razzaq	Pakistan	24.5 (8)	24.9 (6)	4.34 (6)	4.39 (8)	33.9 (8)	34.1 (6)
W Younis	Pakistan	26.2 (9)	23.5 (5)	4.71 (11)	5.05 (12)	33.4 (7)	27.9 (3)
SK Warne	Australia	28.8 (12)	29.5 (11)	4.37 (7)	4.32 (7)	39.5 (12)	40.9 (9)
S Mushtaq	Pakistan	27.7 (11)	27.0 (7)	4.62 (10)	4.26 (6)	35.9 (10)	38.0 (8)
Z Khan	India	26.3 (10)	27.0 (8)	4.99 (12)	4.60 (10)	31.6 (4)	35.3 (7)

Bowler		RM 1st inn	RM 2nd inn	# of appearances 1st inn	# of appearances 2nd inn
M Muralitharan	Sri Lanka	184.7 (3)	131.8 (1)	371	414
GD McGrath	Australia	172.6 (1)	153.3 (2)	369	288
S Akhtar	Pakistan	195.7 (6)	176.8 (4)	223	250
SM Pollock	South Africa	182.9 (2)	191.9 (5)	527	383
AA Donald	South Africa	206.5 (7)	159.0 (3)	269	141
W Akram	Pakistan	188.0 (5)	222.0 (9)	324	384
D Gough	England	187.4 (4)	236.5 (12)	292	221
A Razzaq	Pakistan	219.2 (10)	215.0 (6)	328	385
W Younis	Pakistan	218.3 (9)	216.8 (8)	307	299
SK Warne	Australia	217.1 (8)	223.5 (10)	389	268
S Mushtaq	Pakistan	235.2 (12)	216.5 (7)	303	276
Z Khan	India	225.4 (11)	232.4 (11)	248	201

Table A.15: Correlation coefficients between the first and second innings bowling statistics for 12 of the best active bowlers in ODI cricket.

Bowling statistic	Correlation between 1st and 2nd innings statistics
Bowling average	0.15
Economy rate	0.76
Strike rate	0.07
Runs per match	0.53

Appendix B

Figures

Figure B.1: Average runs scored according to D/L method for overs available and wickets lost; $w = 0$ at top, $w = 9$ at bottom. Source: Duckworth, F.C. and Lewis, A. J. (2002), "Review of the application of the Duckworth/Lewis method of target resetting in one-day cricket." *Mathematics and Computers in Sport*, G. Cohen and T. Langtry (editors), Bond University, Queensland, Australia, **6**, 127-140.

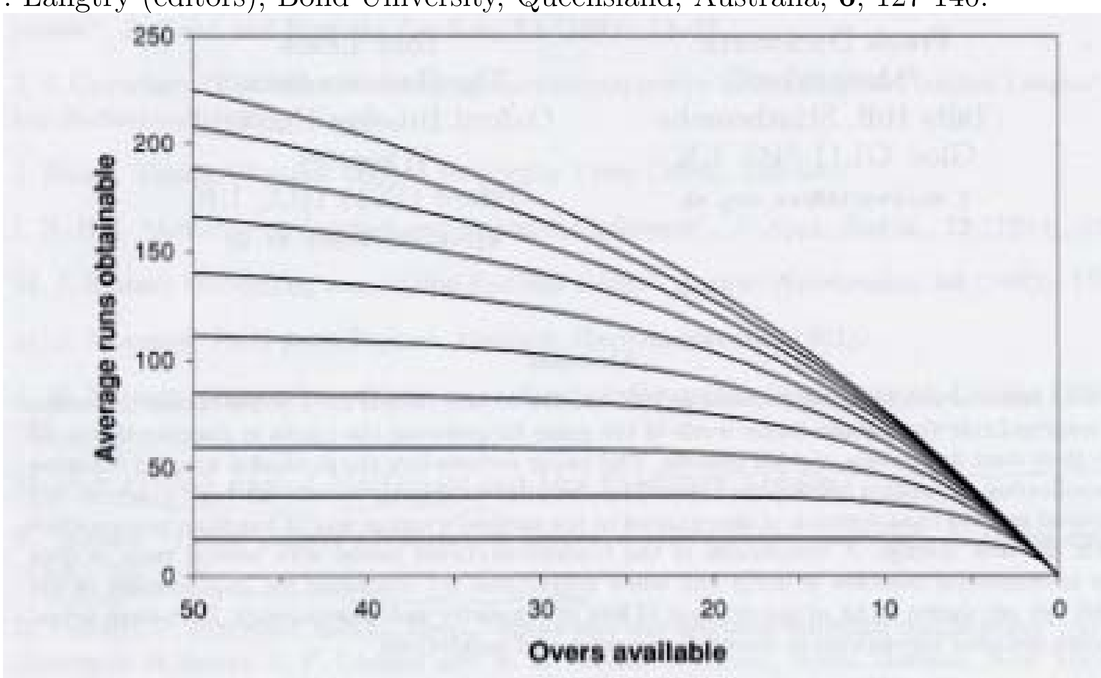


Figure B.2: Multiple comparisons without taking into account the multiple comparisons problem (level of significance = 0.05) for 12 of the best active batsmen in ODI cricket.

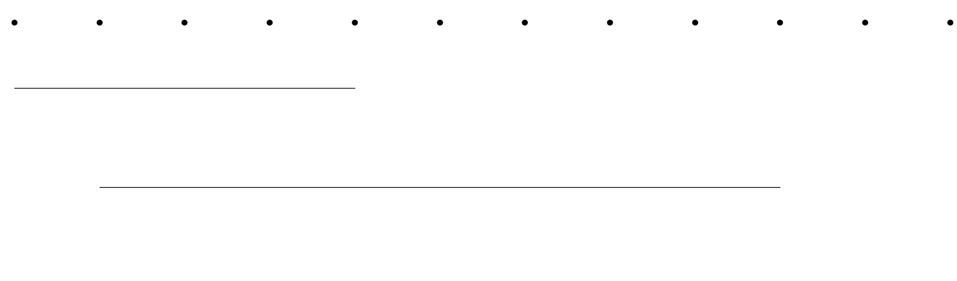


Figure B.3: Multiple comparisons using Bonferroni's approach (level of significance = 0.00076) for 12 of the best active batsmen in ODI cricket.

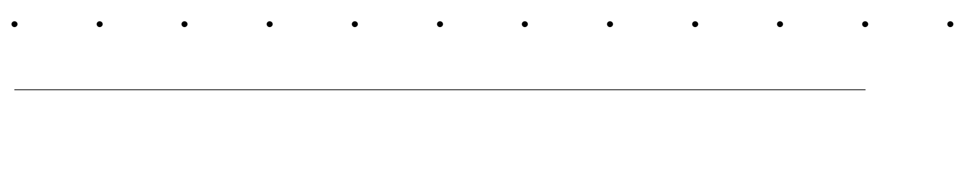


Figure B.4: Runs per match (RM) statistic for 12 of the best active batsmen in ODI cricket presented in increasing order. Standard errors multiplied by the 97.5-th percentile of a standard normal are attached to the RM values.

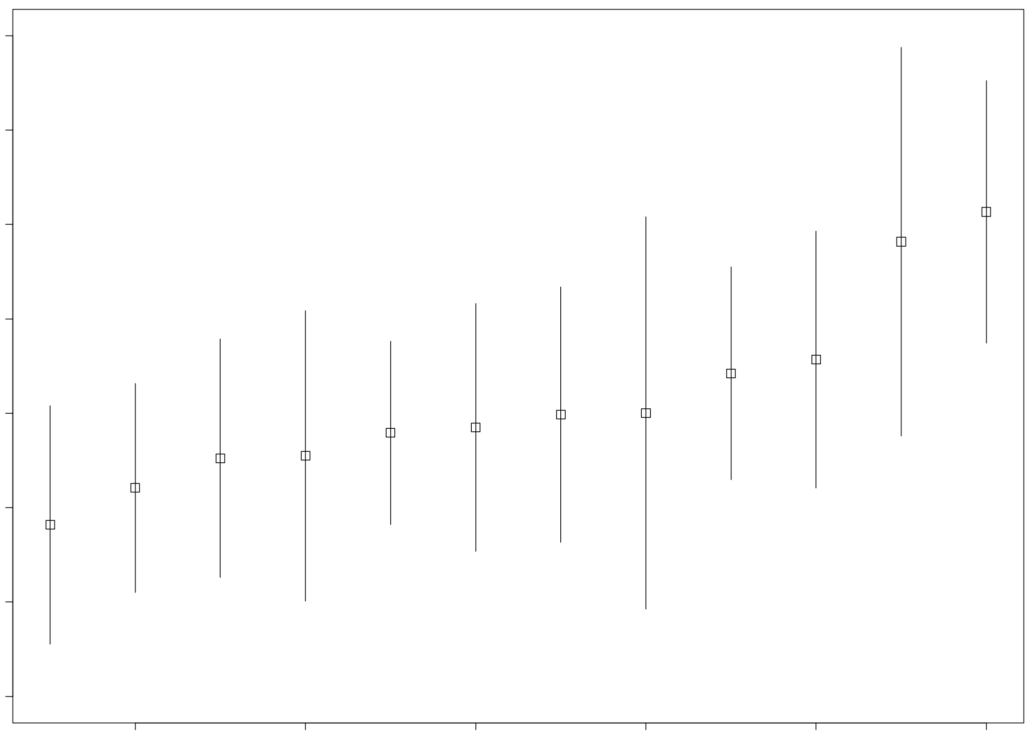


Figure B.5: First versus second innings batting average for 12 of the best active batsmen in ODI cricket.

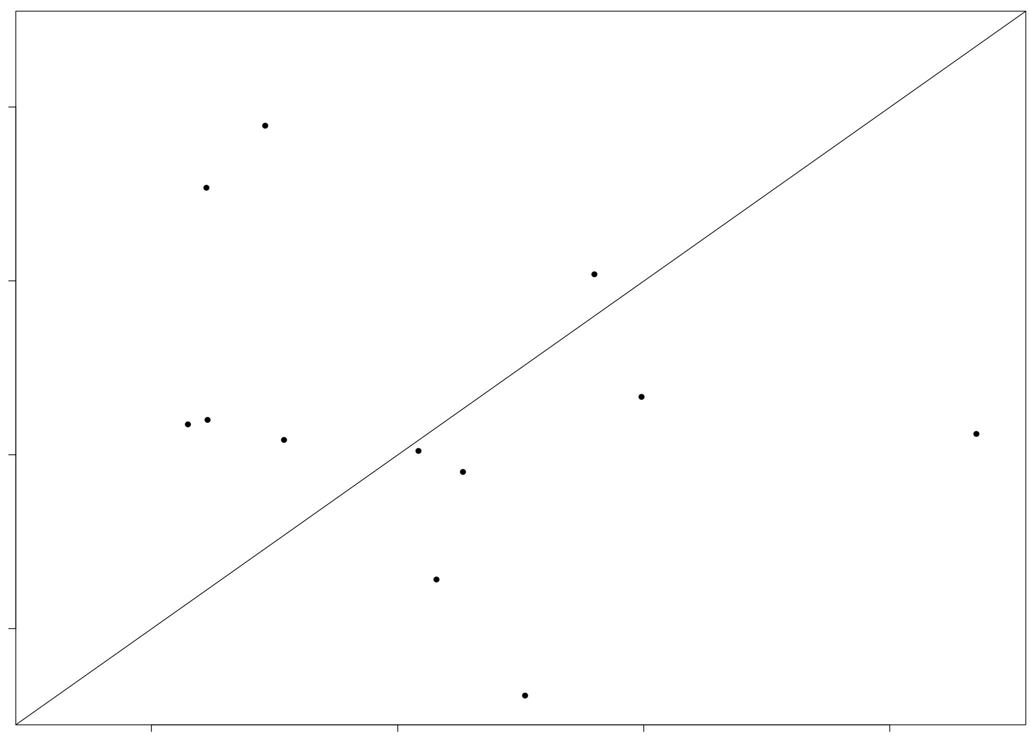


Figure B.6: First versus second innings strike rate for 12 of the best active batsmen in ODI cricket.

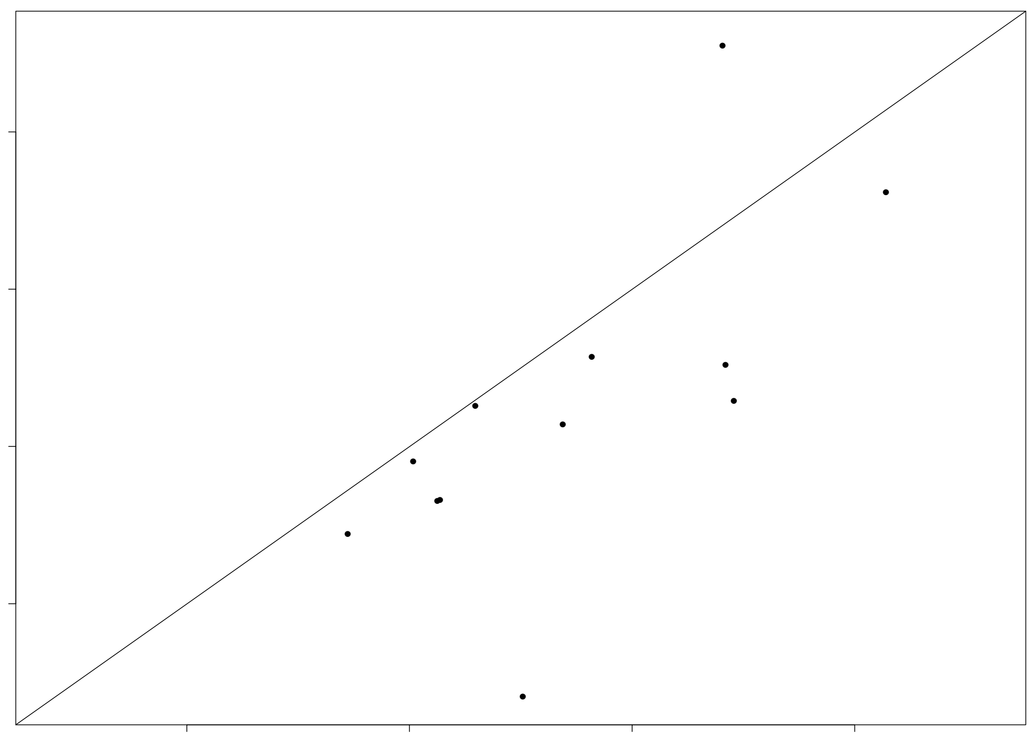


Figure B.7: First versus second innings runs per match for 12 of the best active batsmen in ODI cricket.

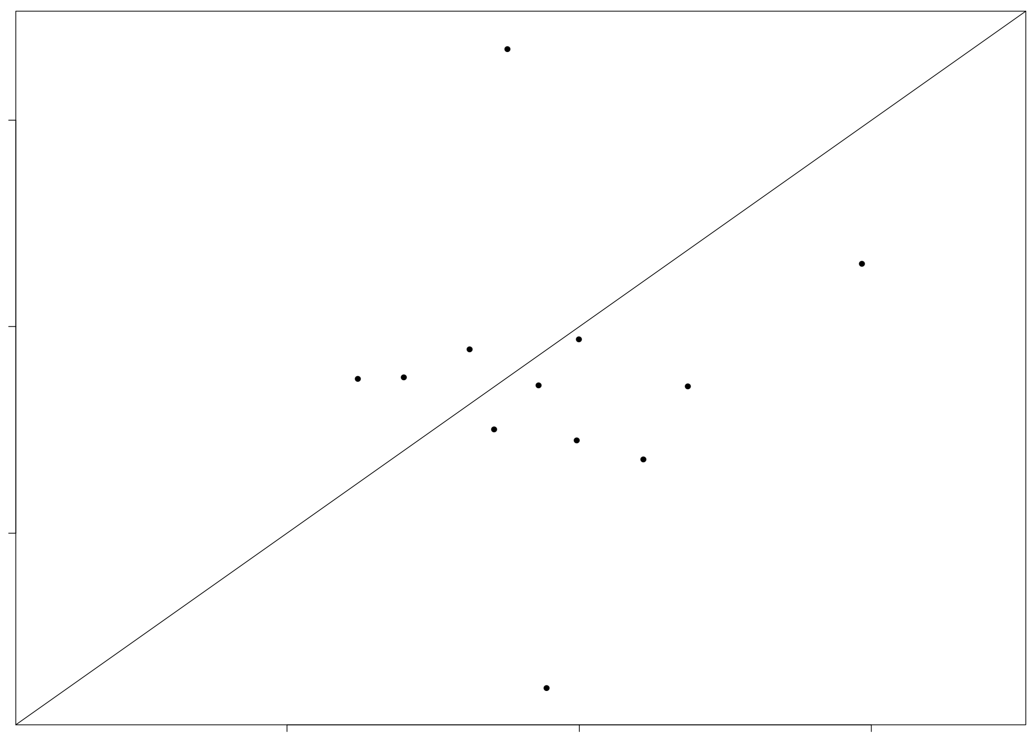


Figure B.8: Multiple comparisons without taking into account the multiple comparisons problem (level of significance = 0.05) for 12 of the best active bowlers in ODI cricket.

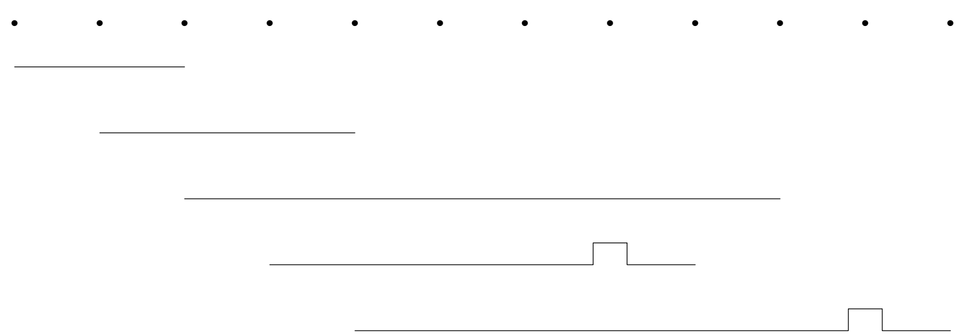


Figure B.9: Multiple comparisons using Bonferroni's approach (level of significance = 0.00076) for 12 of the best active bowlers in ODI cricket.

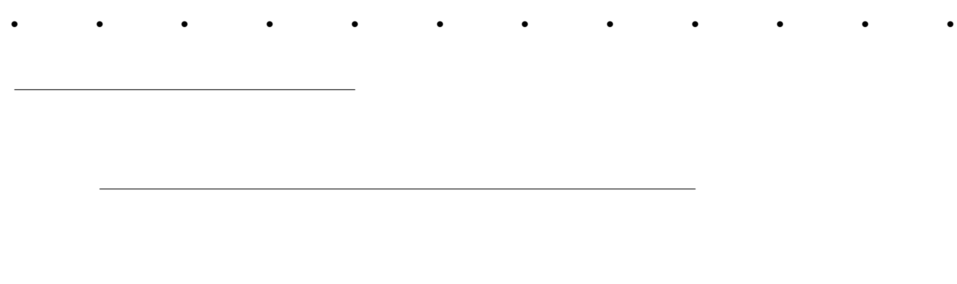


Figure B.10: Runs per match statistic for 12 of the best active bowlers in ODI cricket presented in increasing order. Standard errors multiplied by the 97.5-th percentile of a standard normal are attached to the RM values.

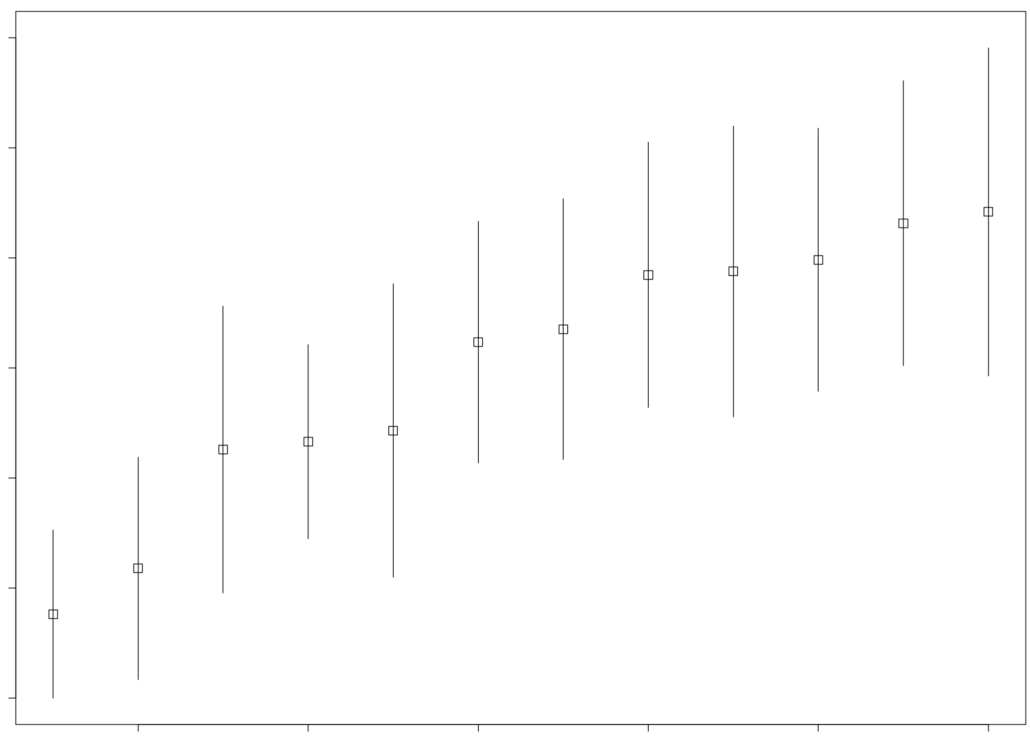


Figure B.11: First versus second innings bowling average for 12 of the best active bowlers in ODI cricket.

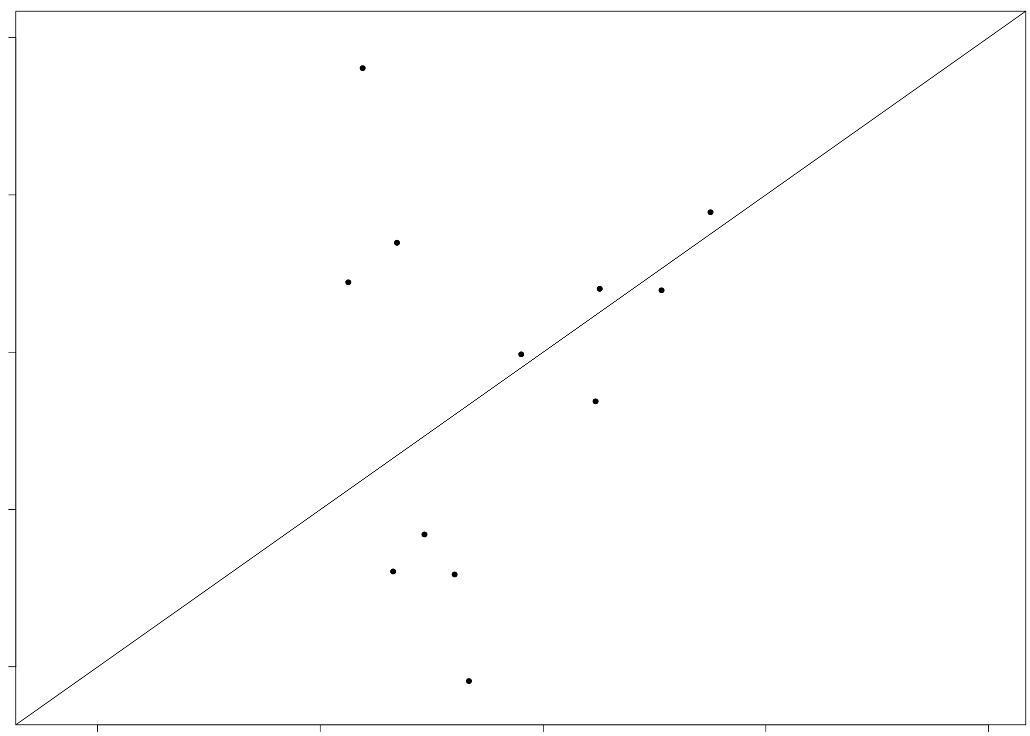


Figure B.12: First versus second innings economy rate for 12 of the best active bowlers in ODI cricket.

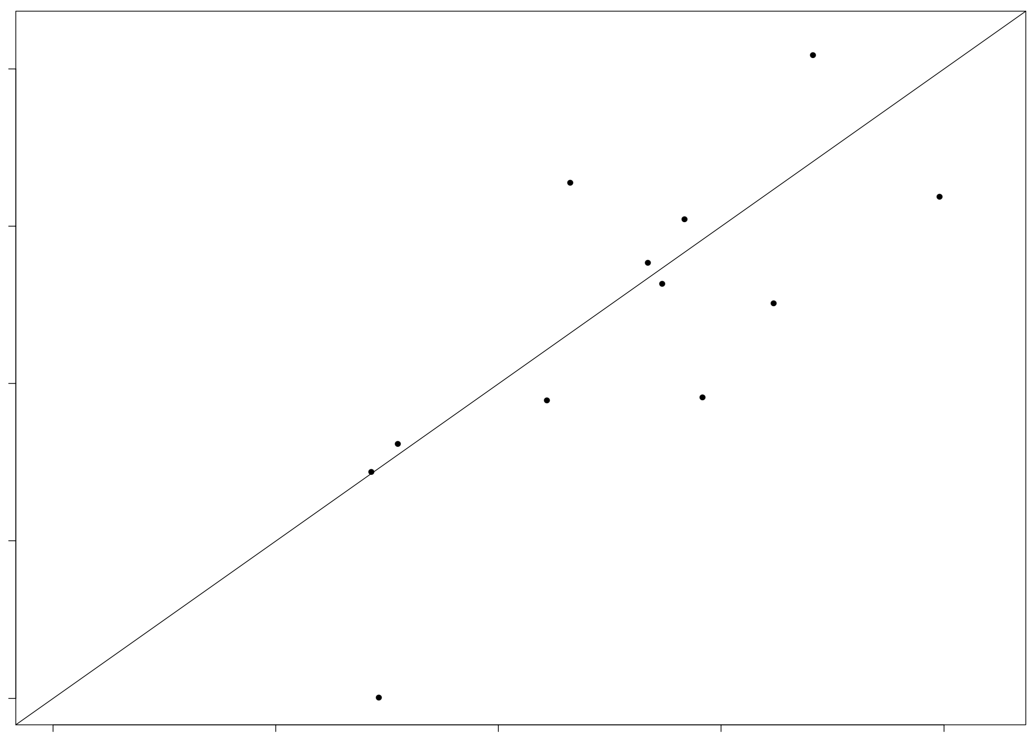


Figure B.13: First versus second innings strike rate for 12 of the best active bowlers in ODI cricket.

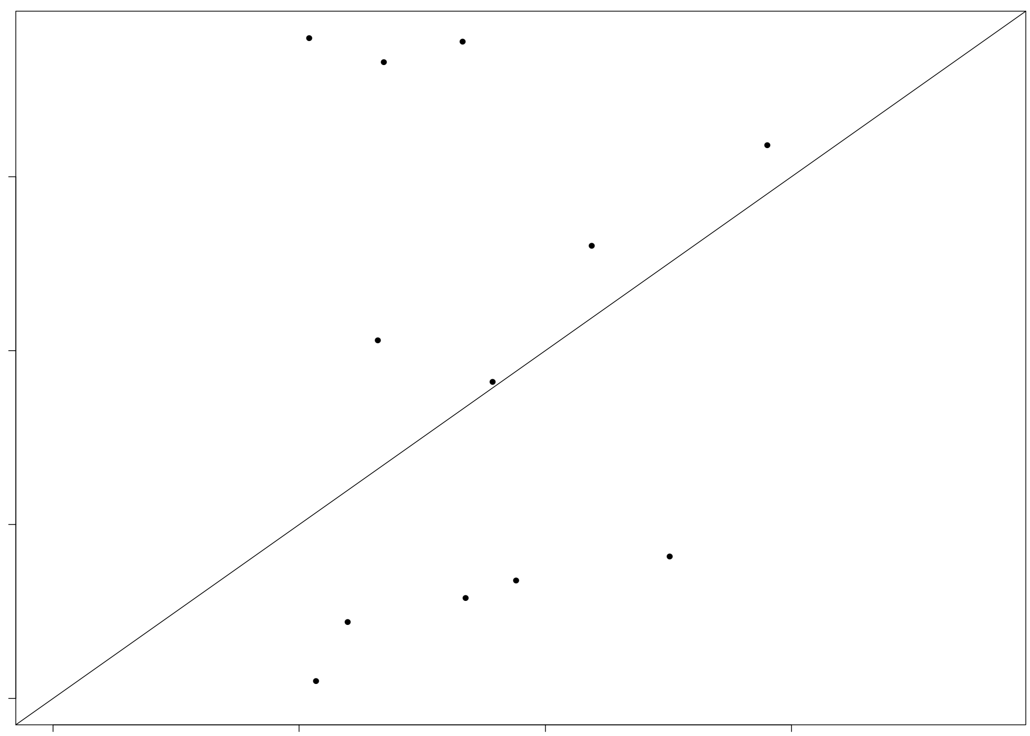
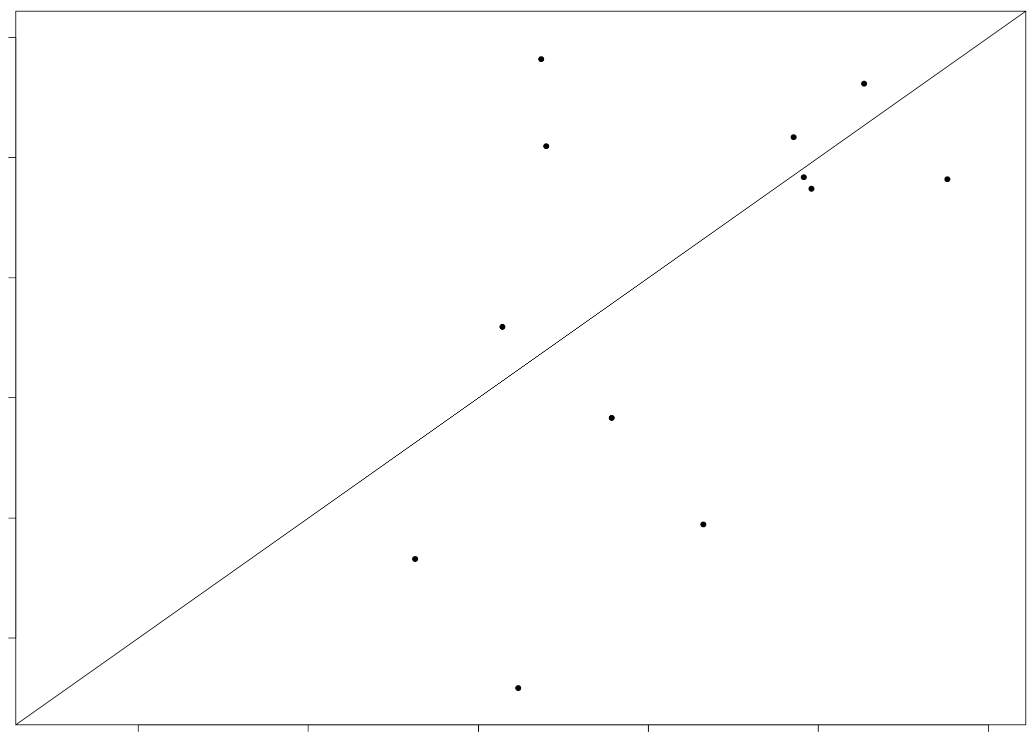


Figure B.14: First versus second innings runs per match for 12 of the best active bowlers in ODI cricket.



Bibliography

[1] Beaudoin, D. and Swartz T.B. (2003), "Performance Statistics in One-Day Cricket.", manuscript.

[2] Duckworth, F.C. and Lewis, A. J. (1998), "A fair method for resetting the target in one-day cricket matches." *Journal of the Operational Research Society*, **49**, 220-227.

[3] Duckworth, F.C. and Lewis, A. J. (2002), "Review of the application of the Duckworth/Lewis method of target resetting in one-day cricket." *Mathematics and Computers in Sport*, G. Cohen and T. Langtry (editors), Bond University, Queensland, Australia, **6**, 127-140.

[4] Kimber, A. C. and Hansford, A. R. (1993), "A Statistical Analysis of Batting in Cricket." *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, **156**, Part 3, 443-455.

[5] Knight, Ron, "Cricket Explained (for novices)", retrieved from *http://www-usa.cricket.org/link_to_database/ABOUT_CRICKET/EXPLANATION/CRICKET_EXPLAINED_FOR_NOVICES.html* on 01/06/2002.

[6] Mar, David, "An Explanation of Cricket - Rules and Game Play Described for Novices", retrieved from *http://www-usa.cricket.org/link_to_database/ABOUT_CRICKET/EXPLANATION/EXPLANATION_OF_CRICKET.html* on 01/06/2002.

[7] Tucker, Jeff, "Cricket Explained (An American Viewpoint)", retrieved from http://www-usa.cricket.org/link_to_database/ABOUT_CRICKET/EXPLANATION/CRICKET_EXPLAINED_AMERICAN.html on 01/06/2002.