ECONOMIC PAPERS

ECONOMIC PAPERS, VOL. 30, NO. 4, DECEMBER, 2011, 455–465

An “Economic” Ranking of Batters in Test Cricket*

Nicholas Rohde[1]

There is much debate within the cricket community over the relative greatness of various batters. Attempts to guide this debate using statistical techniques have thus far been unsatisfying due to difficulties in determining appropriate trade-offs between certain performance criteria. By applying the concepts of opportunity cost and supernormal profit to batting performance we are able to produce a cardinal ranking system that uses non-arbitrary weightings to rank players. The proposed method is used to score past and current players and we find that the Australian batsman Sir Donald Bradman is the highest performer with India’s Sachin Tendulkar a close second. We also note that there is little public awareness of the greatest women batters and rank England’s Rachael Heyhoe-Flint and Australia’s Betty Wilson in the first two positions.

Keywords: opportunity cost, ranking, sports statistics, supernormal profit.

1. Introduction

The game of cricket is one of the world’s most popular sports and test cricket is considered an important format. Due to its popularity the sport attracts an extensive public dialogue, much of which focuses on comparing the merits of past and current players. Indeed it is commonplace to the point of cliché for cricketers, supporters and institutions to name their own selections for short-lists such as the “Australian team of the century,” the “100 greatest players ever” or the ubiquitous “all time XI.” Though such rankings are generally light hearted and undertaken for recreational purposes the outcomes are not necessarily trivial. Important issues such as team selection, player earnings and sponsorship are all linked to perceived performance and on rare occasion honours such as knighthoods are awarded for outstanding achievement. Furthermore, leading cricketers can become iconic cultural figures (Nicholson, 2008) and possession of a world class player can be a point of national pride (Purandare, 2005). Given the importance placed upon such perceptions it is desirable to be able to objectively identify and measure batting prowess.

Judgements on cricketing greatness are usually justified with appeals to statistics. There is, however, no clear cut methodology for determining the superiority of one player’s career performance over another, and as a result, rankings are often constructed in an *ad hoc* fashion. A good example is the system used by the International Cricket Council (ICC) which gives an indication of time discounted player performance, allowing one to identify the leading players at a single moment in

*The author would like to thank Prasada Rao, Andrew Hodge, Sriram Shankar, Andrew Trappett, Tom Nguyen and Athula Naranpanawa as well as an anonymous referee. Any errors are my own.

[1]Department of Accounting, Finance and Economics, Griffith University

JEL classification: L83

Correspondence: Nicholas Rohde, Department of Accounting, Finance and Economics, Griffith University, Gold Coast, QLD, Australia. Email: n.rohde@griffith.edu.au

doi: 10.1111/j.1759-3441.2011.00138.x

time. The statistical method employed by the ICC is admirably thorough in certain respects while open to criticism on other fronts. Perhaps, the most serious objection that could be made to the ICC method is that it claims that ''no subjective assessment is made'' while simultaneously explaining that various weightings are given for more recent performances, successes against more competitive opposition, finishing on the winning side and other such factors.[2] While the weightings used may or may not be reasonable it is unlikely that they are all completely non-arbitrary, and hence some subjective assessments as to the relative trade-offs between these and other criteria must be made.[3] Though one might claim that the inexplicit weighting system used in the ICC rankings must be appropriate as the resultant rankings are consistent with popular beliefs, this argument has a degree of circularity. For example, consider that (i) the rankings themselves are instrumental in shaping perceptions of player ability; and (ii) as the exact formulae are not justified or readily available; it may be that the various weightings were chosen to provide acceptable rankings. In these cases the ICC method may be following rather than leading public perception.

The objective of this paper is to provide a complete ranking of test batters while avoiding, as far as possible, the use of these subjective or arbitrary judgements. While we are unable to entirely eliminate all subjectivity, we find that a simple application of some economic concepts produces a neat and novel solution to this problem. The method employed involves estimating the opportunity cost of a batter's place in a team and ranks players with respect to the excess of their career aggregate over their total opportunity cost. It is in this sense that the rankings produced are ''economic.''

2. Review of Academic Literature

A few alternatives to the ICC rankings have come from the economic literature. The recent and prominent work of Borooah and Mangan (2010) provides a good example by ranking all players according to adjusted batting averages that emphasise various aspects of performance. The first adjustment the authors make is to consider a player's scores relative to his team rather than in absolute terms, while the second is to model player consistency. Although the first adjustment is very appealing, the second claim that batting consistency for a given average is a desirable trait is more debatable. To account for consistency, the authors use Sen's welfare index (Sen, 1976) which (in this context) is equal to the batter's average multiplied by one less the Gini coefficient of inequality between the scores. This rewards a batter for even scoring such that twin scores of fifty are worth more than scores of 100 and 0. It can be seen, however, that the former case is not necessarily superior to the latter case and under certain circumstances may even be inferior.[4]

Several other studies exist along similar lines, including the work of Barr and Kantor (2004) who produce rankings of batters in one day cricket using a technique that balances output and the speed of scoring using a well designed parametric index. Additional academic work comes from Brown (2009) who uses Z transformations to allow for meaningful comparisons of batting averages across eras.

Although academic literature on cricket rankings is quite sparse, econometric rankings of performance in other sports are much more readily available. The most commonly employed techniques are those borrowed from efficiency analysis such as Data Envelopment Analysis which have been applied to various sports, most notably baseball (see Ueda and Amatatsu, 2009; Koop, 2004;

[2]See http://www.reliancemobileiccrankings.com/about.php.

[3]For instance it is not clear that an individual's performance in a victory should be more highly regarded than an equal performance in defeat as in the ICC rankings. As the difference in outcomes is entirely dependent on the actions of other players, having successful teammates would in the ICC case *add* to the measured quality of a player's innings, while standard reasoning would suggest that an equivalent innings is more impressive if it is greater relative to those that surround it.

[4]The relative merits of consistent vs. inconsistent performance can be analysed using a thought experiment. Consider a case where a weaker team plays a stronger team. If both teams play as expected the stronger team will win with a high degree of certainty. However, if one or both teams play inconsistently there is a possibility that the variation in performances may be enough to offset the differences in strengths, allowing the weaker team a chance to win. For this reason a weaker team can be advantaged by inconsistent play.

Dawson and Dobson, 2002; Hofler and Payne, 2006; Richardson, 2002; Carmichael *et al.*, 2001; Schofield, 1988). These studies typically define a set of players which are most efficient at converting inputs into outputs and work by measuring the proximity of each player to an efficient equivalent. A difficulty with all the above approaches however is that they circumvent, rather than solve the issue as to how different player attributes should be combined and weighted. As a result, the rankings that are produced are either unable to differentiate between the leading players where most of the public interest lies (such as the efficiency analysis) or use indices that are sensitive to weighting parameters that are difficult to justify.

3. Developing a Ranking System

As it is a batter's intention to score as prolifically as possible without being dismissed, it is common to assess a player's performance using his or her batting average, which is simply the ratio of aggregate runs to dismissals. While such a metric is undoubtedly useful, there are other player characteristics that are often thought to be of importance that are not included in such a measure. Followers of cricket will often cite factors such as the quality of the opposition, the location of the matches played, the quality of the playing area, variations in luck, equipment and many other influences as reasons not to trust a batting average fully as a complete measure of performance. Despite these concerns, such factors may generally be dealt with simply by (i) comparing player averages to historical means as a method of controlling for variations through time; and (ii) by employing the cricketer's maxim that cross-sectional variations in luck, conditions and opposition tend to even themselves out over a long career.

An issue of significant and neglected importance that is not so easily addressed in a statistical analysis is player longevity. The judgement that both a player's average and longevity are important is implicit in the commonly held view that Sir Donald Bradman is statistically the greatest batsman of all time due to his career average of 99.94. Though Bradman's career average is extraordinarily high (more than double that of many highly regarded batsmen) it is second in men's cricket to the little known West Indian Andy Ganteaume. Ganteaume scored 112 in his only test innings and thus has a claim to be the greatest ever if average is to be the sole criterion. Justification for the belief that Bradman's career is superior to Ganteaume's comes from Bradman sustaining his level of performance while accumulating 6996 runs[5] while Ganteaume sustained his level for only 112.[6]

To create a thorough measure of batting achievement it is thus important to recognise that a batter's career prowess must be measured across two desirable criteria: the intensity of performance (as measurable by the player's average) and for how long the performance was sustained (measureable by player's aggregate contribution). Furthermore, it is necessary to determine how these criteria are to be weighted in order to provide a complete cardinal ranking. Should a batter's average be considered more important than his or her aggregate output in determining greatness, or vice versa?

For the purposes of examining this issue we define the vector of batting results $x_i = \left(x_{i1}, x_{i2}, \ldots, x_{ip}\right)$, where x_{ij} denotes the number of runs scored by player i in her jth innings. We consider the careers of n players, each of whom play p_i innings in total and is dismissed k_i times. This allows the average runs per dismissal of the ith batter to be written as $A_i = R_i/k_i$, where $R_i = \sum_{j=1}^{p_i} x_{ij}$.

Player achievement can be ranked from these variables using an index of the form $f(A, R, \beta) \to \Re$, where f is the index, A is the batter's average, R is career output and β is a parameter that weights between A and R. Ideally when determining such an index, the functional form chosen for f would be uncontroversial and similarly the choice of β would unambiguously capture the true trade-off between the two criteria. While it is difficult to immediately specify f in such a

[5]Bradman achieved this total despite his career being interrupted by the Second World War. Bruce Chapman estimates that if it were not for the war Bradman would have a career aggregate of 9873 runs at an average of 100.74. See http://blogs.mbs.edu/fishing-in-the-bay/wp-content/uploads/2007/08/bradman1.pdf.

[6]It is of interest to observe that despite this remarkable achievement, Ganteaume's innings was not regarded as successful as he scored so slowly as to allow the game to be forced into a draw.

fashion, it is instructive to determine some of the properties that it should satisfy to be consistent with the fundamental aims of batsmanship. For the functional f we suggest two characteristics an index should satisfy for ranking leading batters.

Property (1) Average monotonicity:[7] The index f should be monotonically increasing in career average i.e. $\partial f(A, R, \beta)/\partial A > 0$ for $R \geq 0$, $k > 0$.

Property (2) Runs monotonicity: The index f should be monotonically increasing in aggregate career runs i.e. $\partial f(A, R, \beta)/\partial R > 0$ for $R \geq 0$, $k > 0$.

Both properties seem reasonable for ranking highly achieving batters as it is their objective to score as many runs as possible at the highest sustainable average. Using this idea, Properties (1) and (2) are sufficient to establish a dominance criterion for pairwise comparisons. This is: $x_i \succ x_j$ if $R_i \geq R_j$ and $A_i > A_j$ or $R_i > R_j$ and $A_i \geq A_j$ where ''$\succ$'' denotes the rank ordering. The result holds as $f(A_i, R_i, \beta)$ will necessarily exceed $f\left(A_j, R_j, \beta\right)$ provided that the properties are met. Thus, if a batter outscores a competitor in career aggregate at an equal or better average he is judged as exhibiting a greater overall performance. Similarly if two players have the same career aggregate the batter with the higher average will be judged as superior.

By applying this dominance criterion to all binary player relations we can determine the un-dominated set for male and female batters. For men this is [Ganteaume, (112, 112), Bradman (99.94, 6996), Hammond (58.45, 7249), Sobers (57.78, 8032), Tendulkar (56.54, 14532)] and for women we get the set [Seneviratne, (148, 148), Drumm, (144.33, 433), Broadbent (109.25, 437) Annetts (81.9, 819), Bakewell (59.88, 1078), Hockley (52.04, 1301), Brittin (49.61, 1935)]. The first term within each set of brackets gives the batter's average and the second gives their aggregate career output. As these players possess the leading averages for batters with a certain output this gives a good summary of the leading players. The primary advantage of such a system is that it was determined without specifying f and thus the results produced may be argued to be very general in nature. There are, however, some significant drawbacks. Most notably, the inability of the method to rank between these and other players is unsatisfying, and furthermore it does not give the relative differences between the players. The system also makes no allowances for changes in the game through time, such that players coming from eras favourable for batsmanship will be unduly advantaged. This last point is of particular importance as we will see later that a position in the un-dominated set does not necessarily guarantee a high ranking when the era that a batter plays in is accounted for.

To address the drawbacks of this approach we use the cardinal ranking equation f with an explicit functional form and we employ parameter β to arbitrate between A and R. A primary difficulty of this method is determining exactly how the trade-off should occur between output and average without being chosen arbitrarily such that another system of trade-offs with a corresponding alternative set of rankings could not be selected with equal validity. Indeed, if a non-arbitrary system of trade-offs is unable to be found, f could legitimately be designed such that the effect of Property (1) is sufficiently strong relative to Property (2) to allow the index to order purely in terms of batting average, making Andy Ganteaume the leading male player. Conversely if the effect of Property (2) is chosen to be strong relative to Property (1) the ranking will coincide purely with career aggregate, ranking prolific players with low averages (such as the well respected but historically unexceptional former English captain Michael Atherton) ahead of batters with shorter careers but high averages such as Bradman.

A solution to this problem of finding an appropriate weighting system can be found by recognising that persons playing test cricket face competition from their own compatriots for positions within a team. Consider batsman z with very low non-zero batting average A_z who embarks upon his $(k+1)th$ innings and maintains his average, scoring $x_{z,k+1} = A_z$. According to the dominance criterion, the player has *improved* his position as he has sustained the intensity of his performance over a greater quantity of output. This result is intuitive when considering high performing players, as sustained success at high level A is preferable to transitory success at the same average. However when performance is low (A_z) the converse to Property (2) is more appealing as continued batting failure is not preferable to brief failure. This point becomes especially evident when

[7]It is required that a batter is dismissed at least once such that the batting is well defined.

one considers that the inclusion of player z may come at the expense of another player's exclusion. Thus if the benchmark for batting performance is specified in terms of batting average β, then a player's sustained performance below this level can reasonably be argued to be more damaging to the team than a brief performance at the same low level. Simultaneously a continued performance above this level is still beneficial, indicating that player longevity is a desirable trait only when the player is outscoring their potential replacement.

Using this idea we may amend Property (2) to get:

Property (2*) Benchmark non-monotonicity: the index f should be increasing in career runs when a player averages above a certain benchmark i.e. $\partial f(A,R,\beta)/\partial R > 0$ for $A > \beta$ and should decrease with career runs when the player averages below the value i.e. $\partial f(A,R,\beta)/\partial R < 0$ for $0 < A < \beta$ where parameter β represents the performance benchmark.

A simple measurement index that satisfies both Property (1) and Property (2*) is $f(A,R,\beta) = R - \beta(R/A)$. It is trivial to show that $\partial f(A,R,\beta)/\partial A = \beta(R/A^2) > 0$, $\partial f(A,R,\beta)/\partial R = 1 - (\beta/A) > 0$ for $A > \beta$ and $\partial f(A,R,\beta)/\partial R = 1 - (\beta/A) < 0$ for $A < \beta$, and thus the index satisfies Properties (1) and (2*).

By recognising that $R/A = k$ the index may be rewritten as

$$f(R,k,\beta) = Ak - \beta k \tag{1}$$

and is simply equal to the aggregate difference between a player's career output and the career output of a ''benchmark'' player across the same number of innings.

Here, we may draw a parallel between the ranking of batters and the economic notions of ''opportunity cost'' and ''supernormal profit.'' As an economic concept, opportunity cost refers to the cost incurred by foregoing the second best alternative when making a mutually exclusive choice from a given range of options. An example in a traditionally economic context could be that funds used for an investment project (project 1) cannot also be used for a second project (project 2), so the decision to undertake project 1 implicitly leads to the forfeiture of project 2. This forfeiture is the opportunity cost of project 1.[8] It is clear that this analytical tool is appropriate for analysing batters in cricket, as the opportunity cost of selecting batter i in a team is the quality of the next best replacement player. That is, if the expected batting average of a replacement player is captured by parameter β, this may be interpreted as the opportunity cost of selecting player i. The concept of ''supernormal profit'' is also useful, which is defined as the excess payoff a decision yields over its opportunity cost. In the previous example, a supernormal profit would occur if the return on investment project 1 was greater than the opportunity cost of not selecting project 2. In the context of ranking batters, if batter i is able to score at an average that exceeds β then they are providing excess runs to the team that would not occur in the player's absence. Thus, $A_i - \beta$ provides a measure of the excess value a player brings to a team relative to his or her potential replacement which may be interpreted as a supernormal profit.

It can be seen that Equation (1) measures performance according to this standard. The advantage of this approach is that the appropriate trade-off between intensity and longevity of performance is captured purely by β, which has an intuitive cricketing interpretation and as a result it is possible to justify its choice based upon observing the game. As the parameter is defined as the batting average that is expected of a batter to be of ongoing value to his team, it may be estimated with a degree of objectivity from the data. Thus when an index of the proposed form is used, the relevant trade-off between output and average need not be selected on an *ad hoc* basis, making the methodology justifiable where other methods are not. Furthermore, the index has a unit of measurement of ''runs'' which allows for convenient cricketing interpretation.

4. Estimating the Opportunity Costs

An implication of this model is that the appropriate compromise between the intensity and longevity of a player's performance depends crucially upon what is expected of him as a batsman, which

[8] This assumes that project 2 represents the next best alternative to project 1.

in turn depends upon the ability of the team that he plays for. It is easy to verify, however, that this awkwardness is not merely a symptom of the model but is reflective of the underlying problem. A batter with the ability to average A_i is of little use to a team where the opportunity cost is greater than A_i, while the same player would be of considerable value in a team with lower expectations. As however we wish to rank all batters on their absolute rather than relative merits there is a need to determine a universal benchmark for batting performance rather than specific values for players from different teams. Fortunately, the topic of a universal benchmark for batting performance is a well established concept (in men's cricket) and it is widely regarded that an average of forty runs per dismissal is reasonably appropriate for judging passable batting success (Fraser, 2008; Kidd, 2008; Brenkley, 2008; Lynch, 2009) regardless of the country a player represents.

Though this figure is likely to be reasonable given its widespread acceptance, we seek a more formal treatment and estimate it explicitly from the data. While it is not entirely self evident how this should be determined, we proceed by calculating the expected value of a completed innings for an average player selectable purely on batting merit across all but the least successful test playing teams. This requires the assumption that these teams have access to reserve players that may achieve average success at test cricket if given the opportunity. While this is not likely to be true of all teams at all times, it is felt that the convenience and parsimony of this assumption outweighs its imperfections.

To estimate opportunity costs we take data on every test match played up until the end of 2010. The information is freely available from the well known ESPNCricinfo website (http://www.cricinfo.com) which also provides a number of facilities available for sorting and managing the data. Data is restricted to batters playing in the first four batting positions. Although cricket teams typically have batters from positions 1–6, batters in positions five and six are excluded as they are often chosen partially for their skills in other areas such as captaincy, bowling, wicket-keeping and fielding.[9] As a result of these compensatory strengths their batting performances are likely to be lower than what one would expect for a player selected on batting ability alone. We also exclude batters who have not played more than a solitary innings in these positions each year as it is not uncommon for a bowler or other player not chosen for batting prowess to be promoted to one of these higher positions for tactical reasons. Lastly, as the leading batters emerge overwhelmingly from the more established teams we feel it is incorrect to allow batting performances from players in the very weakest teams to be included in establishing the opportunity costs. Thus, we filter the data by excluding the poorest performing international teams by requiring a win/loss ratio greater than or equal to 1:5. This excludes performances by Zimbabwe, Bangladesh and the ICC World XI in men's matches and Pakistan and the Netherlands in women's cricket. Though other types of cut-offs and filters for inclusion are also reasonable, experience with the data suggests that these rules are parsimonious and effective at filtering out performances that are not representative of a full time batter.

Although we have argued that cross-sectional variations in performance need not be considered due to the averaging out of these influences over many matches, the same cannot be said for variations through time. It is recognised that batting conditions evolve over time and thus performances must be standardised accordingly for comparisons of players from different eras to be valid. To account for this phenomenon we partition x into t sequential annual subgroups where x_i^r denotes the rth the subset of innings played by batter i. As the subgroups are annual (e.g. $r = 1$ denotes the year 1877) they are mutually exclusive (as an innings in one year cannot also occur in another) and collectively exhaustive.[10] Once the partitions are formed we can give the aggregate output of player i as $R_i = \sum_{r=1}^{t} R_i^r = \sum_{r=1}^{t} \sum_{j=1}^{p} x_{ij}^r$ and the batter will be dismissed k_i^r times in year r leading to an average for the rth year as $A_i^r = R_i^r / k_i^r$. The idea behind the series partitions is to allow us to estimate r values of β, one for each year that test cricket has been played. This allows

[9]If batters in positions 1–6 were chosen the resultant opportunity cost would be based upon a mixture of both batting performance and performances in other facets of the game. It is the author's view that basing the opportunity costs on the scores of the first four batters is the most parsimonious method for identifying the value of a ''pure'' batter.

[10]To our knowledge there have been no innings that have spanned New Year's Eve and New Year's Day.

the index to adjust through time, valuing batting performances in leaner periods against a lower standard than performances in more plentiful years.

The yearly benchmark value for batting performance β_r can be estimated as

$$\hat{\beta}_r = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p} x_{ij}^r}{\sum_{i=1}^{n} k_i^r} \tag{2}$$

giving over one hundred intertemporal estimates. We now view the batter's performance as the sum of many yearly performances, each of which is compared relative to its corresponding annual benchmark.

A difficulty with this approach is that an endogenous feedback effect occurs when determining the performance of the ith batter based upon yearly benchmarks, when the benchmark itself depends partially on the performance of player i. A batter experiencing a productive year scoring a high value for $R_i^r = \sum_{j=1}^{p} x_i^r$ will push the numerator of Equation (2) upward, raising the estimated benchmark for performance for that year and simultaneously undercutting the value of his or her performance. While this effect may be present for all batters, it is stronger in years when there is relatively little cricket played as a single batter has a greater influence over the cross-sectional average score in this instance. The effect introduces a bias into the statistic that disadvantages leading players from earlier eras where test cricket was less frequently played. To eliminate this bias we allow the benchmark to vary for each batter, where the estimate for player i in year r depends on all qualified performances in that year *except* those by player i.

The benchmark performance for the qth batsman in time period r is thus estimated as

$$\hat{\beta}_q^r = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p} x_{ij}^r}{\sum_{i=1}^{n} k_i^r} \quad \text{for } i \neq q \tag{3}$$

when the specified qualifications hold. Using this equation we determine unique (though only slightly different) sets of benchmarks for all batters in all years using separate data from men's and women's competitions.

5. Ranking Batters in Test Cricket

To determine the rankings we apply the given methodology to the full data set[11] using the final specification:

$$f(R_i^r, k_i^r, \hat{\beta}_i^r) = \sum_{r=1}^{t} \left(A_i^r k_i^r - \hat{\beta}_i^r k_i^r\right) \tag{4}$$

and present the results in Table 1. The first column gives the ordinal ranking based upon the player's aggregate score determined from Equation (4) and the specific value for f is given in column five. The batter's overall aggregate R_i is given in column three while their average A_i appears in column four. The country abbreviations in column six are AUS – Australia; ENG – England; IND – India; PA – Pakistan; SA – South Africa; SLA – Sri Lanka; WI – West Indies[12] and ZIM – Zimbabwe.

[11] Note that once the opportunity costs have been established the filters discussed in Section 4 are no longer employed. Batters are ranked according to their full career aggregates in Tables 1 and 2.

[12] The West Indian team is not technically a national team as it consists of players from a number of small countries in the Caribbean region. This is similar to the England team which sources players from England and Wales.

Table 1. *Ranking of Male Batters in Test Cricket 1877–2010*

Ranking	Player	Runs (R)	Average (A)	Score (f)	Country
1	Bradman	6996	99.94	4067.10	AUS
2	Tendulkar†	14532	56.54	4038.94	IND
3	Kallis†	11677	56.41	3116.88	SA
4	Lara	11953	52.88	2862.65	WI
5	Sobers	8032	57.78	2811.01	WI
6	Border	11174	50.56	2794.07	AUS
7	Ponting†	12363	53.51	2629.76	AUS
8	Gavaskar	10122	51.12	2602.51	IND
9	S Waugh	10927	51.06	2540.89	AUS
10	Dravid†	12027	52.75	2520.87	IND
11	Miandad	8832	52.57	2502.83	PAK
12	Barrington	6806	58.67	2403.32	ENG
13	Hammond	7249	58.45	2230.38	ENG
14	Richards	8540	50.23	2189.89	WI
15	Chappell	7110	53.86	2168.33	AUS
16	Jayawardene†	9527	53.82	2088.65	SLA
17	Sangakkara†	8244	57.25	2003.01	SLA
18	Hobbs	5410	56.94	1922.11	ENG
19	Hutton	6971	56.67	1896.52	ENG
20	Ul-Haq	8830	49.6	1762.35	PAK
21	Boycott	8114	47.72	1662.95	ENG
22	Sutcliffe	4555	60.73	1565.46	ENG
23	Weekes	4455	58.61	1554.75	WI
24	Yousuf†	7530	52.29	1545.89	PAK
25	Sehwag†	7670	54.01	1501.13	IND
26	Harvey	6149	48.41	1420.23	AUS
27	Lloyd	7515	46.67	1389.02	WI
28	Chanderpaul†	9063	48.98	1360.12	WI
29	Hayden	8625	50.73	1256.03	AUS
30	Greenidge	7558	44.72	1246.01	WI
31	Flower	4794	51.54	1245.79	ZIM
32	Gower	8231	44.25	1240.11	ENG
33	Kanhai	6227	47.53	1199.35	WI
34	Walcott	3798	56.68	1172.45	WI
35	Gooch	8900	42.58	1077.82	ENG
36	Cowdrey	7624	44.06	1067.44	ENG
37	Compton	5807	50.06	1055.51	ENG
38	Walters	5357	48.26	1037.27	AUS
39	May	4537	46.77	1033.37	ENG
40	Kirsten	7289	45.27	996.60	SA

Note: †Current player as of 31 December 2010.

The table gives the rankings for the top forty batters from 1870 to 2010. The index gives the unsurprising result that the Australian Sir Donald Bradman is ranked first, scoring more than 4067 runs than a replacement batter of the same era playing the same number of test innings. India's Sachin Tendulkar is in second place with a score of 4039. It should be noted that Tendulkar and other batters still playing cricket after 2010 are likely to end up with different final scores than given in the table, and that they may move up or down the rankings depending on their future performances. It is therefore possible that Tendulkar could overtake Bradman by the end of his career.

As it currently stands, the list is primarily dominated by English and West Indian batsman who occupy eleven and nine of the forty positions respectively. Eight of the ranked batters are Australian, including Bradman, while there are four from India, three from Pakistan, two from South

Africa and Sri Lanka and one from Zimbabwe. That no Bangladeshi batters appear in the ranking is expected, though there are also no batters from New Zealand, with their leading player Martin Crowe scoring 808.21 and appearing in position forty eight.

A surprising feature is the prevalence of modern players, which might lead one to conclude that the index unfairly favours these players. This result is not unexpected or indicative of bias however, as the volume of men's test cricket played has increased sharply in recent decades, leading to the observed trend in the results. Of the 1944 test matches played until the end of 2010, approximately 30 per cent were played after 1st Jan 1996. If player representation in the top forty occurs proportionately to the quantity of test cricket played, then one would expect a corresponding proportion of 30 per cent (twelve from forty batters) on the list to be from this era. Of the listed batsman, six have careers contained entirely within this period while nine others had careers that have overlapped the period. Given that a small majority of the tests played by these batters occurred after 1st January 1996 it seems reasonable to assign around two thirds of these players' appearances to the post-1996 era and one third to the pre-1996 era. This gives the expected twelve appearances such that batters from the most recent 30 per cent of tests occupy approximately 30 per cent of positions.

For women batters the results are given in Table 2. As data is much sparser for women cricketers many of the rankings from twenty-one to forty depend on a very limited number of performances (in some cases only two or three innings). For this reason we feel that there is insufficient information to rank these players with any degree of confidence, and hence we only provide rankings for the leading twenty batswomen.

The list is well populated with English (8) and Australian (7) batswomen, the remainder being from New Zealand (3) and India (2). Unlike for men's cricket, the current women cricketers are not so well represented in the corresponding rankings, with only three of the twenty batters still playing as of the end of 2010. Again this is not unexpected as there has not been a strong increase in the number of women's test matches played in the last two decades.

Appealingly, the rankings for both men's and women's cricket are roughly consistent with popular opinion on the greatest batters of all time, though we emphasise that the index was not designed with this objective in mind. Tendulkar for instance is generally thought to be second to

Table 2. *Ranking of Batters in Women's Test Cricket 1934–2010*

Ranking	Player	Runs (R)	Average (A)	Score (f)	Country
1	Heyhoe-Flint	1594	45.54	564.74	ENG
2	Wilson	862	57.46	514.86	AUS
3	Hockley	1301	52.04	503.55	NZ
4	Rolton	1002	55.66	500.57	AUS
5	Bakewell	1078	59.88	494.75	ENG
6	Annetts	819	81.90	480.40	AUS
7	Brittin	1935	49.61	472.70	ENG
8	Maclagan	1007	41.95	412.84	ENG
9	Taylor†	1030	41.20	398.58	ENG
10	Edwards†	1380	46.00	382.96	ENG
11	Haggett	762	58.61	328.21	AUS
12	Drumm	433	144.33	326.24	NZ
13	Hill	499	62.37	291.85	AUS
14	Raj†	572	52.00	284.89	IND
15	Hide	872	36.33	282.21	ENG
16	Agarwal	1110	50.45	257.70	IND
17	Broadbent	437	109.25	234.05	AUS
18	Clark	919	45.95	212.55	AUS
19	Snowball	613	40.86	211.40	ENG
20	Doull	779	43.27	210.53	NZ

Note: †Current player as of 31 December 2010.

Bradman as the greatest male batter, with Lara, Sobers and Ponting in close contention. Similarly England's Rachael Heyhoe-Flint is ranked first amongst women and is perhaps the world's most famous female cricketer, ahead of Betty Wilson who was known as ''the female Bradman'' during her career for Australia in the 1950s. Third in the women's rankings comes Debbie Hockley who scored prolifically for New Zealand in the 1980s and 1990s. It is interesting to note that in the women's rankings, neither Heyhoe-Flint, Wilson or Rolton were in the un-dominated set of players determined in Section 2, as former England Captain Jan Brittin outscored all three in career aggregate at a better average. Brittin's relegation to position seven in the rankings is attributable to playing in an era of relatively easy run-scoring, which has resulted in her performances being benchmarked against higher standards than other batters.

6. Conclusion

This paper has proposed a simple methodology for measuring career performances of batters in test cricket and has ranked the leading male and female players. The approach uses the concepts of opportunity cost and economic profit to produce non-arbitrary weightings for batting success. An approximate time variant opportunity cost for the selection of each batter is estimated over his or her career and players are ranked according to the extent to which they outscore this value. The rankings produced are broadly consistent with prior beliefs of player performance.

There are a number of caveats that apply to the analysis. Firstly, it is important to recognise that the technique used represents only one candidate method for ranking batters rather than the only admissible method. It may therefore be possible to construct alternative procedures that produce different rankings that are just as defensible. Secondly, an objection could be made to the ''assuming away'' of cross-sectional variations in luck, opposition, playing conditions and other factors by invoking a Law of Large Numbers argument. While this argument is frequently employed in cricket (especially with regard to umpiring decisions), failing to control for these factors is to some degree imperfect as it is not clear exactly how strong this ''averaging-out'' effect is. It is therefore uncertain how long a player's career has to be before it can be claimed that the ranking has not been unduly affected by these distorting factors.

Although it is unlikely that the averaging-out process will completely eliminate all biases from cross-sectional variations, it is felt that attempts to control for any residual effects are unlikely to improve upon the results presented. As there are a great many factors that could have some marginal effect upon the rankings, using statistical procedures to account for all possible extraneous variables is prohibitively difficult. Furthermore, methods of controlling for such factors usually require making additional assumptions, leading to significant costs in terms of reduced objectivity, transparency and parsimony. As a primary motivation of the paper is to produce rankings with these strengths, it seems unlikely that the benefits of this approach would outweigh the costs. Regardless however of whether these adjustments are made, there is always room for disagreement about the validity of these and other rankings. Given the enjoyment that cricket followers derive from such debate this is no bad thing.

REFERENCES

Barr, G. and Kantor, B. (2004), 'A Criterion for Comparing and Selecting Batsmen in Limited Overs Cricket', *Journal of the Operational Research Society*, **55**, 1266–74.

Borooah, V. and Mangan, J. (2010), ''The ''Bradman Class'': An Exploration of Some Issues in the Evaluation of Batsmen for Test Matches', 1877–2006', *Journal of Quantitative Analysis in Sports*, **6** (3), Article 14, doi: 10.2202/1559-0410.1201.

Brenkley, S. (2008), 'Batsmen Think They are Great but They are Average at Best', *The Independent*, 25 May. Available at: http://www.independent.co.uk/sport/cricket/batsmen-think-they-are-great-but-they-are-average-at-best-833945.html.

Brown, S. (2009), 'Comparing Batsmen Across Different Eras: The Ends of the Distribution Justifying the Means', *Economic Analysis and Policy*, **39** (3), 443–54.

Carmichael, F., Thomas, D. and Ward, R. (2001), 'Production and Efficiency in English Premiership Football', *Journal of Sports Economics* **2** (3), 228–43.

Dawson, P. and Dobson, S. (2002), 'Managerial Efficiency and Human Capital: An Application to English Association Football', *Managerial and Decision Economics*, **23**, 471–86.
Fraser, A. (2008), 'Bewildered Batsmen Lose Sight of Big Score', *The Independent*, 11 March. Available at: http://www.independent.co.uk/sport/cricket/bewildered-batsmen-lose-sight-of-big-score-794009.html.
Hofler, R. and Payne, J. (2006), 'Efficiency in the National Basketball Association: A Stochastic Frontier Approach With Panel Data', *Managerial and Decision Economics*, **27**, 279–85.
Kidd, P. (2008), 'Stephen Fleming Aims to be Cut Above Average', *The Sunday Times*, 24 March. Available at: http://www.timesonline.co.uk/tol/sport/cricket/article3607548.ece.
Koop, G. (2004), 'Modelling the Evolution of Distributions: An Application to Major League Baseball', *Journal of the Royal Statistical Society Series A, Royal Statistical Society*, **167**, 639–55.
Lynch, S. (2009), *Preface to Wisden on the Ashes*. A&C Black Publishers, London.
Nicholson, R. (2008), *Icon of World Sport – Don Bradman*. Roving Eye Publishing Group, Springwood, N.S.W.
Purandare, V. (2005), *Sachin Tendulkar – A Definitive Biography*. Roli Books, New Delhi.
Richardson, S. (2002), 'Efficiency Estimation Using the Stochastic Production Frontier Approach: Evidence From the National Rugby League', Massey University Applied and International Economics Discussion Paper No. 02-05, Department of Applied and International Economics, Massey University, Palmerston North, New Zealand.
Schofield, J. (1988), 'Production Functions in the Sports Industry: An Empirical Analysis of Professional Cricket', *Applied Economics*, **20**, 177–93.
Sen, A.K. (1976), 'Real National Income', *Review of Economic Studies*, **41**, 19–39.
Ueda, T. and Amatatsu, H. (2009), 'Determination of Bounds in DEA Assurance Region Method: Its Application to Evaluation of Baseball Players and Chemical Companies', *Journal of the Operations Research Society of Japan*, **52**, 459–67.