DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH

THE DEMAND FOR TEST MATCH – AND ONE DAY INTERNATIONAL CRICKET IN SOUTH AFRICA – AN ECONOMIC ANALYSIS

by

Melissa van Rensburg

Assignment presented in partial fulfillment of the requirements for the degree of Bachelors of Commerce with Honours at the University of Stellenbosch.

Supervisor: Mr. F.K. Siebrits
March 2009

TABLE OF CONTENTS

List of tables

Page

List of figures

1. INTRODUCTION

“It's a game that reveals itself in layers, it's a game that demands commitment and not fleeting indulgence”
- Bal, editor of Cricinfo (2005)

The game that is cricket is remarkably different from any other sport in the world. One of the most distinguishing features is the lengthy duration thereof – a test match can last up to five days. There is evidence to suggest that the popularity of cricket – test match cricket in particular – has decreased over the years, and is continuing to do so. This is especially the case in South Africa (SA), where attendances at particularly test matches are modest, to say the least. A cursory glance at ticket prices suggests that there is a much higher demand for one-day internationals (ODI’s), than for test matches. The aim of this study is to quantify the demand for cricket in SA in terms of physical attendance at the different venues. Attendance functions were estimated for test matches and ODI’s. An inspection of the data revealed a clear drop in attendance at test matches post-1999.

This paper determined which factors influence attendance at both test matches and ODI’S, and whether the factors differ significantly between these two formats of the game. An attempt is also made to ascertain what effects the introduction of a new shorter version of the game, pro-20 cricket, have had on the demand for test matches and ODI’s.

Literature on the economics of sports has grown over the years starting in the 1950’s with seminal contributions by Rottenberg (1956) and Neale (1964). Since then, extensive research has been done to determine the demand for sports such as baseball, horseracing, soccer, and rugby, to name but a few. Cricket has, however, been neglected. A mere four articles on the demand for cricket have been published to date, none of which considers the South African case. This is not a surprise, since the process of gathering sufficient data was not an easy one.

The rest of the paper will proceed as follows. The subsequent section presents a literature overview, while section 3 gives a concise background. Section 4 specifies the model and data sources and – manipulations. Section 5 describes the estimation procedure that is followed for test matches and ODI's respectively, whereas section 6 presents descriptive statistics. This is followed by a discussion on the empirical results in section 7, whilst the subsequent section delivers diagnostic tests. The paper is concluded in section 9.

2. LITERATURE OVERVIEW

The application of economic analysis in various different areas has increased tremendously over the years. Examples include the economics of crime, politics, war, and even the marriage market. Gary Becker, who in his own words "uses the economic approach to analyze social issues that range beyond those usually considered by economists" (Becker 1993), won the Nobel Prize for Economics in 1992. This is an indication that economists have successfully ventured into new waters.

More than 50 years ago, the first professional article on the economics of sport, '*The baseball players' labor market*' (Rottenberg 1956), was published. Neale's (1964) article, '*The peculiar economics of professional sports'* was also very significant. Since then, diverse enquiries were made into different areas of sport. Examples include the measurement of the economic impact of major sporting events, the necessity of competitive balance between sports teams, and the economic rationale for subsidizing sports facilities, to name but a few.

There have also been numerous studies to determine the demand for professional sports. Most of the early studies focussed on traditional American sports, such as baseball (see e.g. Eisenberg & Siegfried 1980), basketball (see e.g. Burdekin & Idson 1991), and American football (see e.g. Peel and Thomas 1992). Attendance functions were later also estimated for various other sports, such as soccer (see e.g. Peel and Thomas 1988), horseracing (see e.g. Narayan & Smyth 2003), rugby (see e.g. Giles *et al.* 2000, Garland *et al.* 2004), and Australian Rules football (see e.g. Borland 1987). The major finding from the early work is the important effect the uncertainty of outcome of a match has on attendance (see e.g. Gratton & Taylor 2000 and Rosentraub *et al.* 2004). "*Other things being equal, games with evenly matched teams draw the biggest audiences".* (Rosentraub *et al.* 2004: 32, italics in original)

The economic aspects of cricket have not received as much attention as those sports mentioned above. However, the literature is growing rapidly. Schofield (1988) estimated a production function for cricket in the United Kingdom by determining the relationship between inputs (e.g. batting and bowling averages) and output (team performance). Bairam *et al.* (1990) followed on Schofield's (1988) research by estimating production functions for the Australian and New Zealand cricket teams. Furthermore, Duckworth & Lewis (1998) proposed a method for resetting the target in rain-interrupted ODI's. This method has been employed and is still in use.

Although the literature on various economic aspects of cricket have increased, enquiries into the demand for cricket remains under-researched. To date, there are only four published articles on the subject. Schofield (1983) estimated an attendance function for the John Player League, a one-day domestic English competition. Team performance and locational factors were found to contribute significantly to attendance, while most economic factors were insignificant. Hynds and Smith (1994) estimated the demand for test match cricket in England over the period 1984 to 1992. The authors point to the important effect the long duration of a cricket match has. They argue that "higher earnings may induce individuals to substitute other leisure activities for attendance at cricket matches" (Hynds & Smith 1994). Significant variables in their regression include the opposing side, venue played at, the weather, and uncertainty of series outcome. Bhattacharya and Smyth (2003) studied the determinants of attendance at test matches in Australia in the period 1911 to 1984. They report a drop in attendance since the 1930's. The aim of their paper was to determine the possible reasons for this. In accordance with previous studies, the coefficients on income and ticket prices were insignificant. Significant variables include opposition, uncertainty of series outcome, a dummy variable for rain-interrupted matches, and the presence of star players. The uncertainty of match outcome was insignificant, possibly because it is already captured in other variables, such as the opposing side. Finally, Morley and Thomas (2007) estimated an attendance function for domestic one-day cricket in England, emphasizing the importance of "habitual attendance" (Morley & Thomas 2007). They reported similar results.

3. BACKGROUND

"It has a history and heritage and that needs to be protected."

- Shaun Pollock on test match cricket (Polly worried about test cricket 2008)

The future of test match cricket, the purest form of the game, is uncertain. Attendances at test matches continued its downward trend when ODI's were introduced. Now the development of an even shorter form of the game, pro-20, is threatening the continuation of test match cricket. Sambit Bal (2005), editor of Cricinfo, argues that the form of the game should not be modified in order to attract more spectators, but rather that the standards of test match cricket must be maintained at a high level. This will guarantee support for test match cricket. Bal (2005) referred to the exciting Ashes series[1] in 2005, which held huge crowds captivated for days. Obviously, not all people enjoy watching cricket, and this should be accepted. As Bal (2005) puts it, "It is futile for cricket to try to appeal to all".

However, pro-20 does not necessarily pose a threat to the existence of test match cricket. It is possible that the shorter format has increased the scoring rates of batsmen, and hence, increased entertainment at test matches. This raises the question of whether test matches and pro-20 games are substitutes or whether they are complementary. The same can be asked of test matches and ODI's.

Attendances at test matches are not everywhere as appalling as it is in SA. It is not uncommon to see large crowds in test matches in countries such as Australia and England. Thus, the problem is not with test match cricket *per se*; a good contest can still succeed in attracting big crowds. Therefore, at first glance, it would appear that there simply is not such a big demand for test match cricket in SA. A recent survey by BMI on the popularity of the various sports in the country paints a different picture. According to this survey, cricket is the second most popular sport in the country, with its popularity growing with 10.4% in 2007 (Turner 2007). However, the data does not support this finding. The last time that a day of a test match in SA had a capacity

[1] Test match series played between Australia and England. The series originated in 1882 in "affectionate remembrance of English cricket" (Williamson 2008), after they were beaten by Australia.

crowd was in 2005 versus England, with a large proportion of the crowd consisting of English supporters (Turner 2007). It could be that, even though cricket is the second most popular sport in the country, people prefer to watch the matches on television rather than go to the venue. The percentage of households that own televisions in SA have increased strikingly over the years. This percentage stood at 56.5% in 2002, and increased to 67% in 2007 (Statistics South Africa 2007: 6). Thus, this may be a possible reason for the decline in attendance at test matches.

4. DATA AND MODEL SPECIFICATION

Attendance functions were estimated separately for ODI's and test matches played in SA versus SA for the period 1995 to 2007. The estimation methods for the two forms of the game are discussed in detail in the subsequent section. The dependent variable in the case of ODI's is the amount of people that actually attended the game, i.e. it includes season-ticket holders that attended the game. Unfortunately, data on member-attendance could not be gathered. Hence, paid attendance and member-attendance could not be separated. This is a shame, since as Schofield (1983) notes; data on paid tickets are of particular interest, as it conveys important information regarding the popularity of the game. The dependant variable in the test match regression is average daily attendance divided by each ground's capacity.[2] The rationale for the use of capacities will be elaborated on in section 5. Actual daily attendance would have been ideal, since some days of a test match are usually better attended than others are. However, the data was not available. Attendance figures were gathered from the *Wisden Cricketers Almanack* (Cricinfo 2008). Unfortunately, for some matches the attendance figures were not available. Some of these gaps were filled from figures from Cricket South Africa. A discussion on the independent variables that were considered in the regression model follows in the subsequent section. Some variables that were included in previous similar studies, but not included in this model, are also discussed, together with reasons for their exclusion.

[2] Average attendance was calculated by dividing the total attendance by the number of days the test match lasted. In cases where the last day did not yield sufficient play (at least half the day), it was ignored.

4.1 ECONOMIC AND DEMOGRAPHIC FACTORS

4.1.1 INCOME

For every city a match was played in, the real per capita gross domestic product for the province the city is situated in was included in the model as a measure of income. The income of each city would have been ideal, but the data was not available. Most previous studies on the demand for sport have judged the income variable to be insignificant. The gross domestic product per region (GDPR) for the provinces in constant 2000 prices was collected from Statistics South Africa (2006). The expected sign on the coefficient of income is ambiguous, as it depends on whether a cricket match is considered a normal or an inferior good. If a cricket match is an inferior good, an increase in income will lead to a fall in the demand. According to Hynds & Smith (1994), the long duration of a cricket match brings about a shift away from cricket towards less time-consuming leisure activities. This is one explanation for the introduction of shorter versions of the game as an alternative to test matches.

4.1.2 TICKET PRICES AND PRICES OF COMPLEMENTARY GOODS

Naturally, the price of a product should be included in the estimation of the demand for that product. A robust finding from previous published papers on the demand for cricket is that the price variable is not statistically significant. This is also true for other sports: see e.g. Haupert *et al.* (1992: 76 – 77) (baseball), Welki and Zlatoper (1999: 291) (American football), and Burdekin and Idson (1991: 184) (basketball). Unfortunately, data on ticket prices for cricket matches in SA could not be gathered. Nevertheless, qualitative information on ticket prices for Bloemfontein and Port Elizabeth was obtained[3]. A close inspection of this information revealed that, for the above-mentioned two regions, the real price of tickets has remained relatively stable. According to Bhattacharya and Smyth (2003: 85), the insignificance of the price variable might be due to the lack of variation in ticket prices. Hence, the exclusion is not necessarily a dilemma.

[3] The author would like to thank Helene Pieterse, administrative manager of the Free State Cricket Union, and Leigh Deyzel, commercial manager of the Chevrolet Warriors, for the provision of information on ticket prices.

The prices of complementary goods, such as parking, refreshments, and merchandise, should ideally also be included in the model. However, it was not possible to gather data on these expenditures. Consequently, it was ignored, which is in accordance with previous studies. It should, however, be kept in mind when analysing the results. The prices of complementary goods are rather expensive at South African cricket grounds. Therefore, it could be a deterrent to attendance. Nevertheless, since it is not known whether the prices of these complementary goods have increased in real terms over the years, no conclusion can be reached.

4.1.3 POPULATION

The urban population of the province the match was played in was included in the model to account for market size (Bhattacharya & Smyth 2003: 80). Unfortunately, the population for every city, which would have been ideal, was not available for every year. The exclusion of the non-urban population is essential since in some provinces, e.g. the Eastern Cape, there are a high percentage of people living in rural areas. These people are not likely to attend a cricket match. The expected sign of the coefficient is positive. The data was acquired from mid-year population estimates from SSA (1996 – 2007).

4.2 MATCH SPECIFIC FACTORS

4.2.1 UNCERTAINTY OF OUTCOME

It is well established in the literature of the economics of sport that more evenly matched games will be better attended, *ceteris paribus* (see e.g. Borland 1987 and Haupert *et al.* 1992). This is what differentiates the sport industry from other industries (Gratton & Taylor 2000: 193). Gratton & Taylor (2000: 194) observe that 3 criteria of uncertainty of outcome plays a role in determining the demand for a game, namely uncertainty of the outcome of the match, uncertainty of the outcome of the season/series, and the uncertainty of the level of domination of a team in the long-run. It is the first two that are important in the context of this analysis.

For test matches, two measures were used to determine match uncertainty. Method 1 applies an *ex post* evaluation, while the second utilises an *ex ante* estimate of the prospective closeness of a contest. A discussion on both methods follows.

METHOD 1:

For test matches, the uncertainty of the outcome of a match is determined by several factors. A match was judged "certain" if:

- the team batting first scored less than 150 or more than 500, or
- a team has a first innings lead of more than 150, or
- the team batting last is chasing less than 150 or more than 350.

With "certain", it is meant that the probability of only one team being able to win is very high. Even though a draw may still be likely, most people find this outcome less fascinating. These criteria were only used as a guideline; e.g. if a match was clearly uncertain as revealed by an inspection of the scorecard, but the criteria indicated the opposite, the match was judged uncertain. A dummy variable that took the value of 1 if a match was uncertain and 0 otherwise was included in the model. In the regression output, the variable is called 'match uncertainty'.

This proxy for match uncertainty is comparable to the method employed by Bhattacharya and Smyth (2003: 80 – 81). They used two proxies, one for whether an early finish is expected, and one for whether a team won by an innings or more. Both variables turned out to be statistically insignificant. Hynds and Smith (1994: 103 – 104) also included a dummy variable for whether the outcome of a day of a test match is certain. They do not elaborate on how outcome uncertainty was determined, and their variable also turned out to be statistically insignificant. The insignificance of these proxies for match uncertainty prompted the search for a superior alternative.

METHOD 2:

Another measure of the uncertainty of the outcome of a match is the absolute value of the difference between the ratings of the two teams. The ratings that were used are a variation of the Elo ratings system that is used to rank chess players (Wikipedia 2008). The basic idea of the system is that when two teams play in a test series, their current ratings are used to determine the probable outcome of the series. If a team performs better than predicted their rating will improve, if they perform worse, their rating will deteriorate, and if the predicted result is correct, the rating will remain unchanged. This produces relatively stable ratings for teams over time (Test cricket rating service 2008). Unfortunately, these ratings were not available for ODI's.

An alternative measure of match uncertainty was employed by Morley and Thomas (2007: 2090 – 2091). They estimated the probability of a home team win by betting odds. This variable was statistically significant and had a negative coefficient, which is supportive of the match uncertainty hypothesis. However, the validity of the use of betting odds as a proxy for match uncertainty is questionable. Buraimo *et al.* (2005: 645) explains that there is good reason to believe that betting odds are not unbiased estimators of true winning probabilities. There exist theoretical and empirical evidence in support of this view (see e.g. Levitt 2004 for the former and Avery & Chevalier 1999 for the latter). Because of the above-mentioned problems with this proxy for match uncertainty, it was not considered in this model.

The duration of a one-day international means that supporters cannot know in advance whether a match outcome will be uncertain. Due to the unavailability of ratings for ODI's, the absolute run rate difference was used as a proxy for competitive balance.[4] Thus, even though it will not always be the case, one would expect that a larger run rate difference would indicate less competitive balance between two teams.

[4] The run rate difference was calculated as follows: for each match in the data period the average runs scored per over of the two teams were subtracted from each other.

A dummy variable for the uncertainty of the outcome of a series was included in the model. This variable was assigned a value of 1 if the outcome of the series was uncertain at the beginning of the game and 0 otherwise.

4.2.2 OPPOSITION

The data revealed that matches against some teams are better attended, *ceteris paribus*. This is already accounted for by the variable for match uncertainty to a certain extent, but there may also be other factors at play. Table 1 shows the average attendance against the different oppositions for ODI's and test matches. It is evident from the table that attendance is noticeably higher at ODI's than at test matches. In fact, the average attendance for all oppositions at ODI's is nearly twice the average attendance at test matches (16 596 versus 8 889). The highest average attendance at ODI's is versus Australia, the West Indies, and England. Note the high standard deviation of attendance versus New Zealand. This is due to the fact that 2 out of the 3 ODI's in the sample were played at the De Beers Oval in Kimberley, which has a much lower capacity. The lowest attendance is versus Bangladesh. This is to be expected, since there is not a good competitive balance between SA and Bangladesh. For test matches, the highest average daily attendance is versus England, India, and the West Indies, while the lowest is versus Bangladesh and Zimbabwe. Unfortunately, no attendance figures for test matches versus New Zealand could be acquired. A dummy variable was included for every country SA has played against in the sample period, except West Indies who is the reference category.

4.2.3 VENUE

Table 2 lists the various stadia that were played at in the sample period. It also shows which venues are generally used for test matches. Buffalo –, Goodyear –, and Sedgars Park were also used *occasionally*, but this is the exception rather than the rule. An inspection of the data revealed that some venues were consistently better attended than others were. In some cases, this may be due to supply constraints. According to Hynds & Smith (1994: 104), "Demand variation by venue will reflect both the size of the catchment population, the attractiveness of the stadium, and local interest in live international cricket”. For ODI’s there exist clear supply constraints; attendance is frequently near capacity at most venues. This is not the case for test matches.

Table 1: Attendance by opposition for test matches and ODI’s

	Tests			***ODI's***		
	Average	*Standard deviation*	*n*	*Average*	*Standard deviation*	*n*
Australia	9200	1880	6	20226	5834	15
Bangladesh	1942	1389	2	4230	2408	4
England	11284	4720	10	19035	5065	11
India	10132	6361	5	16467	2792	4
Kenya	-	-	0	5856	2133	2
New Zealand	-	-	0	17404	12297	3
Pakistan	6578	2104	5	14833	5006	13
Sri Lanka	8557	2493	6	13504	8302	9
West Indies	9566	2191	9	19525	4276	13
Zimbabwe	3539	1084	2	15141	7628	9
All	8889	4049	45	16348	6969	84

The capacities of the venues remained unchanged throughout the sample period. Even though capacities were increased for the 2003 World Cup, it does not affect this study. The reason for this is twofold. Firstly, the change in capacity was only temporary. Secondly, no World Cup matches were included in the regression.

As can be seen from table 3, Newlands in Cape Town is consistently the best-attended ground for test matches when accounting for capacity. Table 4 shows the average attendances at ODI’s for the different stadia. Both these tables also show the ‘average attendance capacity ratio’ (aacr). A dummy variable was included for every venue, except Centurion in Pretoria, which was the reference category.

Table 3 shows that, for test matches, there are big variations between the ratios of attendance to capacity for the different stadia. Newlands has the highest aacr with 0.51, and Wanderers the lowest with 0.33. The aacr ought to perform better as dependent variable in the regression for test matches, since it gives a better indication of the true demand for test match cricket in a particular region. The importance of using the aacr as the dependent variable is evident from the fact that if average daily attendance is used, Wanderers it at the top, but when capacity is taken into consideration, it is at the bottom. At first glance, it would appear from table 4 that the ratios of attendance to capacity for ODI's also differ significantly between the stadia. However, the low ratios are all at venues that are not used all that often. These venues usually get matches versus weaker oppositions, which would explain the lower aacr. Considering this, there is not enough variation in the aacr to use it as dependent variable in the case of ODI's. This issue is further explored in section 5.

Table 2: Different stadia played at

Town	**Stadium name**	**Test**
Benoni	Willowmoore Park	No
Bloemfontein	Goodyear Park	Yes
Cape Town	Newlands	Yes
Durban	Kingsmead	Yes
East London	Buffalo Park	No
Johannesburg	Wanderers	Yes
Kimberley	De Beers Diamond Oval	No
Paarl	Boland Park	No
Port Elizabeth	St. George's	Yes
Potchefstroom	Sedgars Park	No
Pretoria	Centurion	Yes

Table 3: Attendance by venue for test matches

	Capacity	*Average*	*Standard deviation*	*n*	*aacr*
Centurion	18000	7602	1980	10	0.42
Kingsmead	23500	9157	3213	7	0.39
Newlands	22000	11155	4459	11	0.51
St. George's	18000	7027	3683	6	0.39
Wanderers	33000	10954	3800	8	0.33

Source: Own calculations

Note: aacr = average attendance capacity ratio

4.2.4 WEATHER

If rain is predicted for the day a match is played on, less people will attend, *ceteris paribus*. Since data on weather predictions could not be gathered, a dummy variable was included for whether a game was rain interrupted or not (1 if it was, 0 otherwise). The author's discretion was used to decide when a match was sufficiently rain-interrupted to assign a 1.

Table 4: Attendance by venue for ODI's

	Capacity	*Average*	*Standard deviation*	*n*	*aacr*
Centurion	18000	15143	1634	8	0.84
Kingsmead	23500	19264	4167	13	0.82
Newlands	22000	18904	3122	13	0.86
St. George's	18000	14307	3741	11	0.79
Wanderers	33000	25150	6975	14	0.76
OUTsurance Oval (previously Goodyear Park)	16000	12117	3758	8	0.76
Sedgars Park	8000	4766	3674	2	0.60
Buffalo Park	16000	13130	2692	8	0.82
De Beers Diamond Oval	8500	5203	1895	4	0.61
Willowmoore Park	7000	4643	2140	2	0.66

Source: Own calculations

4.2.5 PRESENCE OF STAR PLAYERS

Two of the previous studies on the demand for cricket have included the presence of star players as an independent variable. Some of the studies on the demand for other sports have also included this variable. Bhattacharya and Smyth (2003) defined Sir Donald Bradman as such a player, and included a dummy variable to indicate whether he played in a match or not. His presence turned out to have a significant positive effect on attendance. Schofield (1983) also included a dummy variable for the presence of star players that also turned out to be significantly positive, but less convincingly so. For this paper, it was decided that the choice of which players to classify as star players would be too subjective. It is doubtful whether there is a player in the sample period that would have boosted attendance radically merely by pitching up.

4.2.6 OTHER

Dummy variables that indicate what month the match was played in were included in the model, with January as the reference category. The other months that test matches were played in are February, March, November, and December. One would expect that test matches played in the December-holidays, i.e. in December and January, will be better attended, *ceteris paribus*. Cricket has been broadcasted on television for the entire sample period; hence, it is not necessary to control for this fact.

5. ESTIMATION PROCEDURE

5.1 TEST MATCHES

As mentioned previously, attendance at test matches is rarely close to the stadium's capacity. Because of considerable differences between the different capacities of the stadia, using the average daily attendance as the dependent variable would lead to delusory results. Instead, the average daily attendance was divided by the stadium's capacity to give a better indication of the demand for test cricket in a particular region.

This complicates the estimation procedure somewhat, since ordinary least squares cannot simply be applied with proportions as the dependent variable.

For every observation, the proportion of the stadium that is full was calculated as follows:

$$\hat{P}_i = \frac{n_i}{N_i} \qquad (1.1)$$

where $\hat{P}_i$ = estimate of the true population proportion

n_i = average daily attendance

N_i = capacity of each stadium

The estimated proportions were then used to estimate the logit for every observation:

$$\hat{L}_i = \ln\left(\frac{\hat{P}_i}{1-\hat{P}_i}\right) = \hat{\beta}_1 + \hat{\beta}_2 X_{1i} + \ldots + \hat{\beta}_k X_{ki} \qquad (1.2)$$

where $\hat{L}_i$ = estimated logit for each observation i

$\hat{\beta}_k$ = k^{th} estimated parameter

X_{ki} = k^{th} regressor of the i^{th} observation

Since $\hat{P}_i = \frac{n_i}{N_i}$ (1.1), the estimated logit can also be rewritten as follows:

$$\hat{L}_i = \ln\left[\frac{\frac{n_i}{N_i}}{1-\frac{n_i}{N_i}}\right]$$

$$= \ln\left[\frac{\frac{n_i}{N_i}}{\frac{N_i - n_i}{N_i}}\right]$$

$$= \ln\left[\frac{n_i}{N_i - n_i}\right] \qquad (1.3)$$

To ensure that $\hat{P}_i$ does not equal 0 or 1, ½ is added above and below the line of 1.3. Thus:

$$\hat{L}_{i\_adjusted} = \ln\left[\frac{n_i + \frac{1}{2}}{N_i - n_i + \frac{1}{2}}\right]$$

$$=\ln\left[\frac{\dfrac{n_i+\frac{1}{2}}{N_i}}{\dfrac{N_i-n_i+\frac{1}{2}}{N_i}}\right]$$

$$=\ln\left[\frac{\hat{P}_i+\frac{1}{2N_i}}{1-\hat{P}_i+\frac{1}{2N_i}}\right] \qquad (1.4)$$

Ordinary least squares cannot now simply be blindly applied to 1.4 to obtain the regression estimates, as it can be shown that the disturbance term will not be homoscedastic, as is required by the assumptions of the classical normal linear regression model (Gujarati 2003: 598 – 600). If we proceed without taking into account heteroscedasticity, the estimators will be inefficient even though it will still be unbiased (Gujarati 2003: 599). There are two ways to overcome this problem: firstly, OLS can be used with robust standard errors, and secondly, the weighted least squares method can be applied. Both these methods were employed and compared, in order to determine the most appropriate for this specific model.

For the first method, White Heteroskedasticity-consistent standard errors and covariance was used. For the weighted least squares, the following weight is applicable:

$$w_i = N_i \hat{P}_i (1-\hat{P}_i) \qquad (1.5)$$

These two methods will be compared in section 6.

5.2 ODI'S

Attendances at ODI's have remained relatively high throughout the sample period. There may have been a slight decrease, but no drastic change is apparent. Furthermore, attendances at all the grounds are consistently near capacity. Hence, it could be argued that, for ODI's, there exist supply constraints. This makes it difficult to capture the true demand for ODI's (see Buraimo *et* *al.* 2005: 643). Even when the data does not report a capacity attendance, it often would not have been possible to buy a ticket because of season-ticket holders that do not necessarily attend every game. Since there is not enough variation in attendances when weighted with each ground's capacity, the appropriate dependant variable for the ODI regression is actual attendance. Hence, the method of ordinary least squares would be adequate in this case. However, the reader is advised to interpret the results with care, owing to the possible problems that may arise due to the mentioned supply constraints.

6. DESCRIPTIVE STATISTICS

Figures 1 and 2 show the attendance at test matches and ODI's respectively over the sample period. A reduction in attendances at test matches is clearly visible, while attendances at ODI's have remained relatively stable.

Figure 1: Average daily attendance at test matches

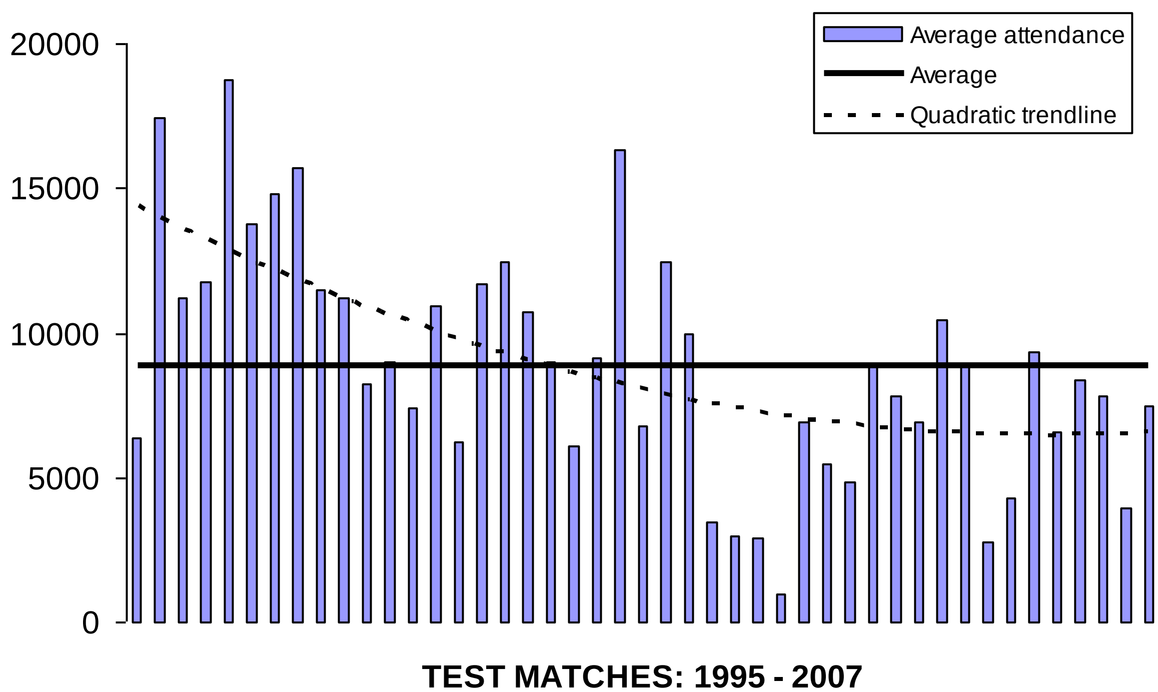
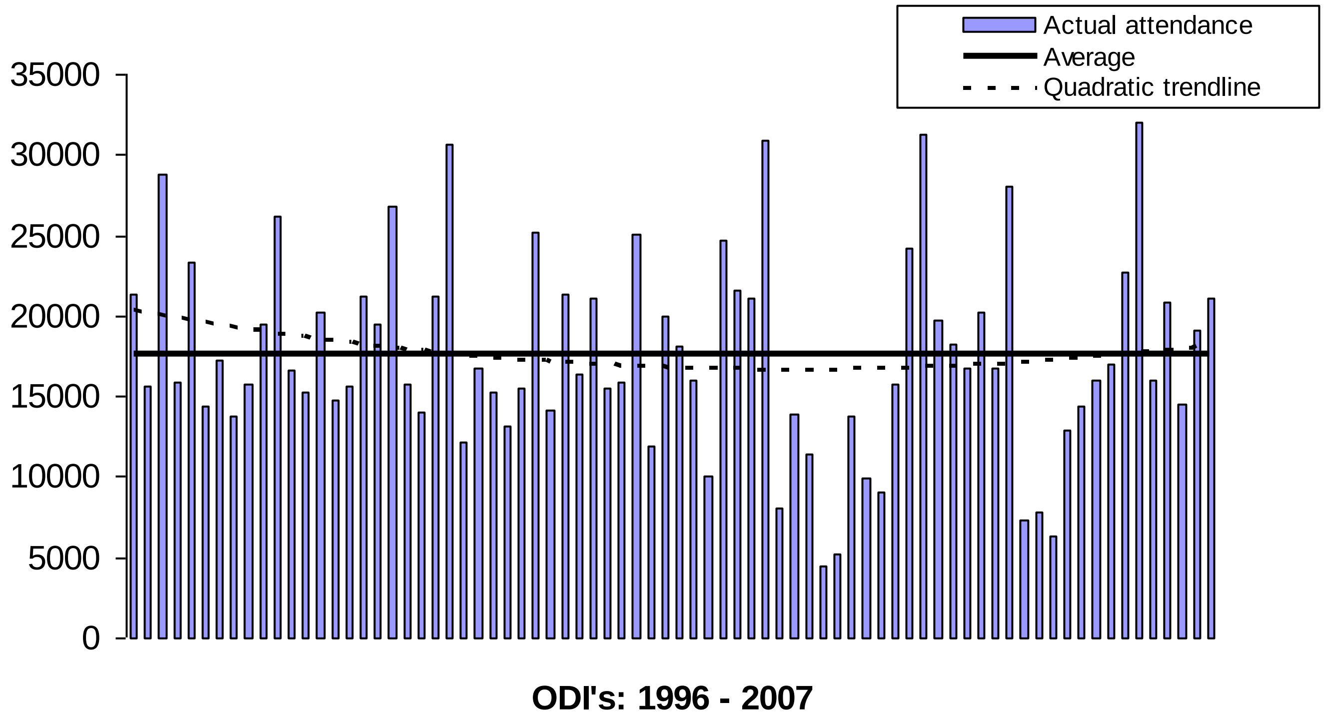


Figure 2: Actual attendance at ODI’s



7. EMPIRICAL RESULTS

7.1 TEST MATCHES

The method described in section 5 was applied to 45 test matches played in SA between 1995 and 2007. There is a good representation of all oppositions except for Bangladesh, New Zealand, and Zimbabwe. Even though the sample stretches over 12 years, the data was treated as cross-section since there are no regular periodic time intervals.

Both the weighted least squares and the robust standard errors measures for the correction of heteroscedasticity were employed. As mentioned in section 4, two methods were considered as a proxy for the uncertainty of the outcome of a match. The two estimation measures are compared for both methods of match uncertainty, the results of which are shown in table 5 and 6 respectively. Table 7 shows the regression results when both methods are included. This option has merit, since the two methods capture different dimensions of match uncertainty. This exercise of comparing regression outputs is also useful for another reason, namely that it provides information on the robustness of the variables to the different specifications of the model. Note that for three of the oppositions, Bangladesh, New Zealand, and Zimbabwe, and for three venues, Buffalo –, Goodyear – and Sedgars Park, there were not enough observations to render reliable estimates. Hence, they were excluded from the model.

A glance at tables 5 to 7 makes it clear that the two methods of correcting for heteroskedasticity do not differ much. The p-values and the size and sign of the coefficients are very similar in the two models. However, the robust standard errors method consistently yields the higher adjusted R^2. Yet, the Akaike information criterion (AIC) is lower in the case of the wls method. Ultimately, the robust standard errors method was chosen for two reasons. Firstly, it seems to fit the data better, as is clear from the higher adjusted R^2. Secondly, the interpretation of the coefficients is easier than in the wls scenario.

The best method for estimating match uncertainty can also be judged by examining tables 5 to 7. In table 5 the first method, as discussed in section 4, is employed, while table 6 contains the second method, the absolute rating difference between the teams. The model that contains the second method slightly outperforms the model that uses the first, more subjective method. The adjusted R^2 is higher (0.73 versus 0.70), and the AIC is lower (1.14 versus 1.24). However, these marginal differences do not necessarily render the second method superior. As mentioned previously, there is merit in including both measures of match uncertainty, as it captures different aspects thereof. The pairwise-correlation between match uncertainty and absolute rating difference is not very large (-0.17), thus, including both variables will not be problematic. The model containing both methods (table 7), outperforms the two previous models. The adjusted R^2 is higher at 0.78 and the AIC is lower at 0.96. Furthermore, it is possible for two oppositions to be closely matched in terms of their rating differential, while the outcome of the match might be relatively certain. Consequently, it was decided to use the model that includes both measures of match uncertainty.

Table 5: Model 1 – match uncertainty method 1

	Robust standard errors	Weighted Least Squares
Constant	8.0517	7.6142
	0.0000	**0.0007**
Kingsmead	-4.6625	-4.3071
	0.0002	**0.0013**
Newlands	-2.4117	-2.4717
	0.0002	**0.0010**
St. George's	-6.5520	-6.2004
	0.0000	**0.0004**
Wanderers	-0.5936	-0.6499
	0.0014	**0.0095**
Australia	1.0962	1.1752
	0.0111	**0.0254**
England	0.0781	0.0257
	0.7466	0.8911
India	0.1335	0.1766
	0.5855	0.4519
Pakistan	-0.6292	-0.6427
	0.0085	**0.0309**
Sri Lanka	-0.2801	-0.1920
	0.2331	0.4584
Urban population	-0.0000003	-0.0000004
	0.0442	**0.0380**
Real GDPR per capita	-0.0001	-0.0001
	0.0090	**0.0631**
Math uncertainty	0.6181	0.6346
	0.0001	**0.0020**
Series uncertainty	0.6761	0.5664
	0.0117	**0.0305**
Rain	-0.5490	-0.5554
	0.0104	**0.0064**
February	-1.5558	-1.4948
	0.0064	**0.0259**
March	-1.3655	-1.4140
	0.0003	**0.0016**
November	-1.2105	-1.0527
	0.00001	**0.0021**
December	-0.7752	-0.6550
	0.0415	**0.0463**
n	40	40
R^2	0.8386	0.8254
Adjusted R^2	0.7003	0.6758
AIC	1.2441	1.0610
Durbin-Watson	1.9922	1.9928
Prob(F-statistic)	0.0001	0.0002

Table 6: Model 2 – match uncertainty method 2 (rating differentials)

	Robust standard errors	Weighted Least Squares
Constant	4.5736	4.2135
	0.0091	**0.0309**
Kingsmead	-2.6685	-2.1936
	0.0240	**0.0611**
Newlands	-0.7690	-1.2793
	0.2264	**0.0802**
St. George's	-3.4028	-2.9640
	0.0139	**0.0444**
Wanderers	-0.3757	-0.3106
	0.0991	0.1494
Australia	0.6479	0.5569
	0.1153	0.2167
England	0.5621	0.4300
	0.0109	**0.0446**
India	0.0508	-0.0126
	0.8528	0.9565
Pakistan	-1.0837	-1.0947
	0.0004	**0.0016**
Sri Lanka	0.2787	0.3048
	0.3924	**0.2759**
Urban population	0.00000001	-0.00000020
	0.9780	0.2913
Real GDPR per capita	-0.0001	-0.00005
	0.0824	0.3744
Rating difference	-0.0083	-0.0082
	0.0002	**0.0010**
Series uncertainty	0.6508	0.6142
	0.0147	**0.0175**
Rain	-0.3504	-0.4056
	0.0908	**0.0256**
February	-1.3642	-1.3981
	0.0254	**0.0304**
March	-0.7383	-0.7605
	0.0141	**0.0366**
November	-1.3637	-1.3414
	0.00001	**0.0002**
December	-0.7175	-0.7827
	0.0594	**0.0168**
n	40	40
R^2	0.8551	0.8366
Adjusted R^2	0.7308	0.6965
AIC	1.1366	0.9951
Durbin-Watson	2.4870	2.3611
Prob(F-statistic)	0.00003	0.0001

Table 7: Model 3 – both methods of match uncertainty

	Robust standard errors	Weighted Least Squares
Constant	5.9386	5.8662
	0.0008	**0.0036**
Kingsmead	-3.3703	-3.2179
	0.0022	**0.0077**
Newlands	-1.3227	-1.6211
	0.0327	**0.0203**
St. George's	-4.6636	-4.5946
	0.0007	**0.0041**
Wanderers	-0.5117	-0.5121
	0.0018	**0.0212**
Australia	0.8887	0.9826
	0.0418	**0.0340**
England	0.3736	0.2973
	0.0897	0.1325
India	0.0717	0.0376
	0.7658	0.8577
Pakistan	-0.9912	-1.0030
	0.0003	**0.0016**
Sri Lanka	0.1040	0.1493
	0.6826	0.5630
Urban population	-0.0000001	-0.0000002
	0.5956	0.2263
Real GDPR per capita	-0.0001	-0.0001
	0.0202	**0.0821**
Match uncertainty	0.3892	0.4224
	0.0093	**0.0253**
Rating difference	-0.0060	-0.0059
	0.0045	**0.0122**
Series uncertainty	0.6701	0.6511
	0.0049	**0.0067**
Rain	-0.5075	-0.5270
	0.0025	**0.0037**
February	-1.3585	-1.4832
	0.0166	**0.0130**
March	-1.0622	-1.1769
	0.0061	**0.0032**
November	-1.2833	-1.2614
	0.000001	**0.0002**
December	-0.7753	-0.7476
	0.0241	**0.0124**
n	40	40
R^2	0.8846	0.8735
Adjusted R^2	0.7750	0.7534
AIC	0.9587	0.7888
Durbin-Watson	2.4543	2.3954
Prob(F-statistic)	0.00001	0.00003

7.1.1 DISCUSSION OF VARIABLES

Table 8 reproduces the final regression model, i.e. the robust standard errors method was employed, and both measures of match uncertainty were included in the model. The variables will now be discussed following the classification in section 4.

7.1.1.1 Economic and demographic factors

a) Income

The income variable is negative and significant at the 1% level. This is in contrast with previous cricket demand studies which have mostly found the income variable to be statistically insignificant. The negative coefficient suggests South Africans regard test cricket as an inferior commodity, i.e. as people's income increase they attend less test matches. Therefore, SA's good economic performance since 2000 may be one explanation for the decreased attendance at test matches. People may have shifted away to less time-consuming activities, including shorter versions of cricket such as ODI's and pro-20.

The use of the logit model complicates the interpretation of the variables somewhat. Recall from equation 1.4 that

$\hat{L}_{i\_adjusted} = \ln \left[\dfrac{\hat{P}_i + \frac{1}{2N_i}}{1 - \hat{P}_i + \frac{1}{2N_i}} \right]$ (see section 5.1). Therefore, the approximated odds ratio is obtained by taking the antilog of the estimated logit. For each regressor, $(\exp(\beta_j) - 1)*100$, will give the percentage change in the odds for a unit increase in the j^{th} regressor (Gujarati 2003: 602). Thus, a one-unit increase in real GDPR per capita will decrease the odds of attending a test match by 0.0123%.

Table 8: Model 3

Constant	5.9386
	0.0008
Kingsmead	-3.3703
	0.0022
Newlands	-1.3227
	0.0327
St. George's	-4.6636
	0.0007
Wanderers	-0.5117
	0.0018
Australia	0.8887
	0.0418
England	0.3736
	0.0897
India	0.0717
	0.7658
Pakistan	-0.9912
	0.0003
Sri Lanka	0.1040
	0.6826
Urban population	-0.0000001
	0.5956
Real GDPR per capita	-0.0001
	0.0202
Match uncertainty	0.3892
	0.0093
Rating difference	-0.0060
	0.0045
Series uncertainty	0.6701
	0.0049
Rain	-0.5075
	0.0025
February	-1.3585
	0.0166
March	-1.0622
	0.0061
November	-1.2833
	0.000001
December	-0.7753
	0.0241
n	40
R^2	0.8846
Adjusted R^2	0.7750
AIC	0.9587
Durbin-Watson	2.4543
Prob(F-statistic)	0.00001

The coefficients can also be interpreted by computing the probabilities and change in probabilities of attending a test match. The important thing to note is that the impact of income on the probability will be different at *each* income level. If a linear probability model was used, the effect of income on the probability would have been constant (Gujarati 2003: 596).

The change in the probability of attending a test match was calculated as follows:

$\delta P/\delta X_j = \beta_j * P * (1 - P)$ (Gujarati 2003: 596). Figure 3 shows the change in the aacr for different levels of income, with the other regressors held constant at their averages. The probabilities were computed for a test match played in January at Centurion versus Pakistan that was not rain interrupted (henceforth reference match). Also, the match outcome was uncertain while the series outcome was not.

The figure underlines the importance of using the logit model since the probabilities are not confined to change linearly with X. As can be seen from figure 3, the negative impact that income has on attendance reaches a maximum at a real GDPR per capita of approximately R35 200. The change in probability at this income level is about -0.000031. Table 9 shows the aacr for different income levels.

b) Population

The coefficient on the urban population of the province the match was played in was negative, but statistically insignificant. A possible reason for why it is not significant might be the high correlation (0.82) between real GDPR per capita and the urban population. Therefore, provinces with bigger urban populations are also the provinces with a higher GDPR per capita. Consequently, the effects of population are already captured in the income variable. The negative coefficient is contrary to previous sport demand studies. A possible reason for the negative coefficient might be that provinces with larger populations have more choices in terms of leisure activities. These provinces also get more matches in a season, further contributing to the lower attendance.

Figure 3: Change in aacr at different income levels

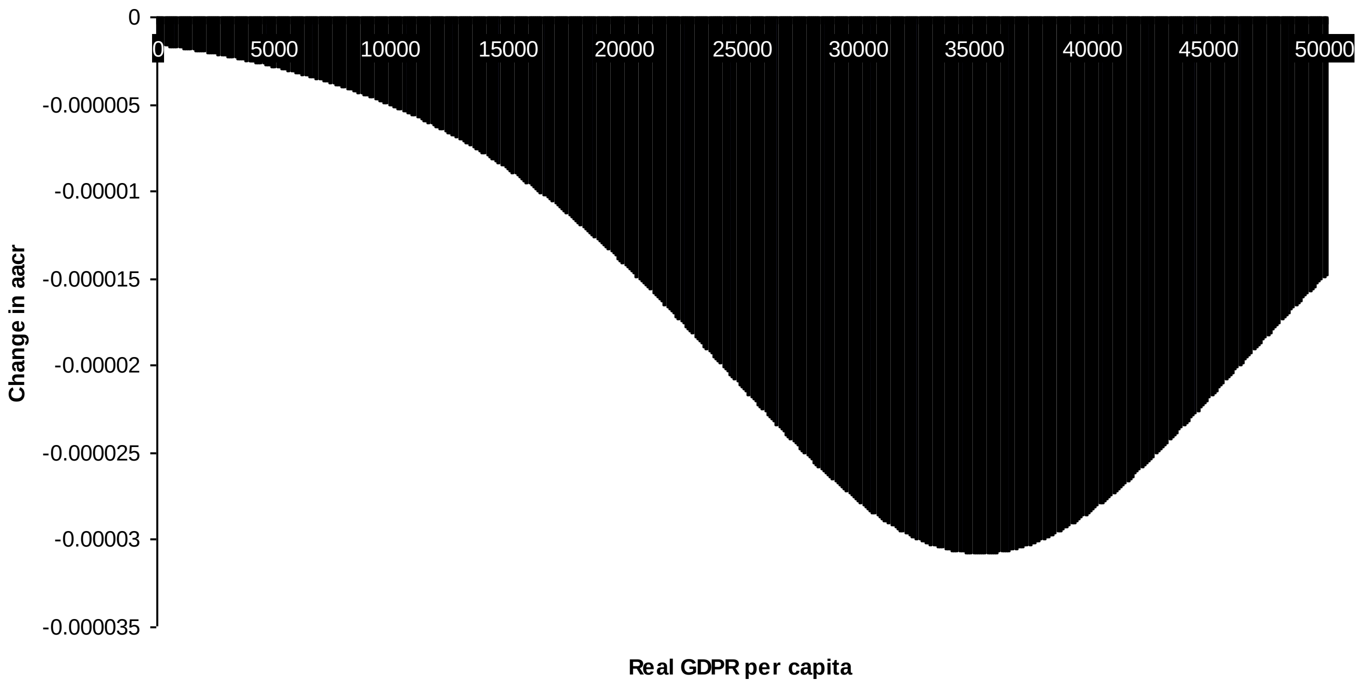


Table 9: Estimated aacr by real GDPR per capita

Real GDPR per capita	aacr
10000	0.9570
15000	0.9232
20000	0.8667
25000	0.7785
30000	0.6551
35000	0.5067
40000	0.3570
45000	0.2308

7.1.1.2 Match specific factors

a) Uncertainty of outcome

Firstly, consider method 1 of estimating match uncertainty. It is a dummy variable that was assigned 1 if the match was deemed uncertain and 0 otherwise. The variable is positive and significant at the 1% level, which is as expected. The probability of attending a match was calculated for both a certain and an uncertain match outcome, holding the other variables constant at their averages. Once again, the probability of attending was calculated for the reference match. For a certain series outcome, the probability of attending a match is 0.0913 higher if the match is uncertain. For an uncertain series outcome, the probability of attending a match is 0.0704 higher if the match is uncertain. It is clear that the increase in the probability of attending is lower when the series outcome is uncertain. The reason for this is that a share of the increase in attendance is already captured in the uncertainty of the series outcome.

Consider now method 2 of estimating match uncertainty. The coefficient on the absolute rating difference is negative and significant at the 1% level, as expected. Thus, the bigger the difference between the ratings of two teams, the less uncertain is the match outcome, and hence, the less people will attend the match, *ceteris paribus*. A one-unit increase in the absolute rating difference decreases the odds of attending a match by 0.602%. The estimated aacr was calculated for the reference match for different levels of rating differentials. This is shown in table 10.

Table 10: Estimated aacr by rating differentials

Rating difference	**aacr**
0	0.7773
25	0.7501
50	0.7207
75	0.6894
100	0.6562
125	0.6214
150	0.5852
175	0.5482
200	0.5106
225	0.4729
250	0.4355
275	0.3988
300	0.3632
325	0.3291
350	0.2966

The coefficient on series uncertainty is positive and significant at the 1% level. This is once again as expected. For the same reference match as before, the probability of attending a match with an uncertain outcome is 0.1297 higher when the outcome of the series is still unknown. If the reference match has a certain outcome, the probability of attending is 0.1506 higher when the series outcome is uncertain. Two things should be noted here. Firstly, the effect that series uncertainty has on attendance is stronger than the effect of match uncertainty. Thus, it would seem that people care more about if the series is “still alive”, than the excitement of a specific match. Since people buy tickets well in advance, if the outcome of a series is a one-way bet, less people will attend, *ceteris paribus*. The second thing to note is that, once again, the effect of an uncertain series outcome is lower when the match outcome is also uncertain. The reason is the same as before, i.e. that a share of the increase in attendance is already captured in the uncertainty of the match outcome.

b) Opposition

The dummies for Australia, England, Sri Lanka, and India have positive coefficients. The dummy variables for Australia and England are statistically significant (Australia at 5% and England at 10%). This means that Australia and England are on average better attended than the West Indies, which is the reference opposition. Attendance for matches played versus India and Sri Lanka are about the same as for versus the West Indies. Keep in mind that there are already being controlled for the differences between the teams’ ratings. Thus, the opposition variable measures whether a match versus a specific team is better attended for reasons other than competitive balance. This is why there are differences in the signs of the opposition variables in tables 6 to 8. The dummy variable for Pakistan has a negative coefficient, which is significant at the 1% level. Thus, matches versus Pakistan are on average less well attended than matches versus the West Indies are.

Table 11 was constructed for a rainless match played at Newlands in December, of which both the match – and series outcomes were already known. The population – and income variables were kept at their averages, while the rating differentials were varied by its average for every opposition. The result is an estimated aacr for every opposition. Australia has the highest aacr of 0.5135, while Pakistan has the lowest at 0.1865.

Table 11: Estimated aacr by opposition

Opposition	aacr
Australia	0.5135
England	0.3975
India	0.3578
Pakistan	0.1865
Sri Lanka	0.2694
West Indies	0.3284

c) Venue

All of the venues' coefficients are negative and statistically significant. All but Newlands is significant at the 1% level, while Newlands is significant at the 5% level. This means that, other things equal, Centurion is the best-attended venue. This contradicts the finding in section 4.2.3 (see table 3), where Newlands had the highest aacr, followed by Centurion. A possible reason might be that Newlands' higher attendance is already captured with other variables. Newlands always hosts the New Year's test (2 to 6 January), which is normally very well attended. Furthermore, if a strong opposition (say Australia), plays a 3-match series in the country, it is usually played at Newlands, Kingsmead, and Wanderers. Table 12 shows the estimated aacr for a rainless match played versus Pakistan in January. The income and population variables were held constant at their averages for each venue. Both the match – and series outcomes were certain.

Table 12: Estimated aacr by venue

Venue	aacr
Centurion	0.4689
Kingsmead	0.4035
Newlands	0.4290
St. George's	0.2950
Wanderers	0.3672

d) Weather

As expected, the rain variable is negative and significant at the 1% level. Thus, other things equal, if a match is rain interrupted, less people will attend. For a certain match – and series outcome, a rain interrupted match decreases the aacr by 0.126. If both the match – and series outcomes are uncertain, rain decreases the aacr by 0.0928. This once again shows the effect that the uncertainty of outcome has on attendance.

e) Month played in

All of the month dummies are negative and statistically significant. March and November are significant at the 1% level, while February and December are significant at the 5% level. It would seem as though the difference in attendance in the various months is due mainly to the occurrence of the school holidays. Table 13 shows the estimated aacr for the different months for the reference match, with certain match – and series outcomes. January has the highest aacr of 0.5844 while February has the lowest at 0.2655.

Table 13: Estimated aacr by month

Month	**aacr**
January	0.5844
February	0.2655
March	0.3271
November	0.2804
December	0.3931

7.2. EFFECT OF PRO-20 ON TEST MATCH CRICKET

In this section, the author will attempt to ascertain to what extent, if any, pro-20 has influenced attendance at test matches. Pro-20 was first introduced in SA in April 2004. To determine whether the introduction of the new, shorter version of the game diverted attendance away from test matches, a dummy variable was included in model 3 (see section 7.1) that was assigned the value 1 for every test match played after the introduction of pro-20. The results of this regression are shown in table 14.

There are some interesting points to note about table 14. Firstly, the coefficient on pro-20 is **negative** and significant at the 10% level. Thus, this would *suggest* that the introduction of pro-20 has had a harmful impact on attendance at test matches in the country. Secondly, there are variables that are no longer significant, namely the venues, real GDPR per capita, and the first method of match uncertainty. Thirdly, the adjusted R^2 of the model is higher than in model 3 (0.8041 versus 0.775). Additionally, the AIC is lower at 0.8188 compared to model 3's AIC of 0.9587.

Table 15 presents the same model as in the previous table, except that real GDPR per capita is dropped from the regression. The coefficient on pro-20 is still negative and is now significant at the 1% level. The problem now is to identify cause and effect. Three things happened since 2004; pro-20 was played for the first time, SA experienced high economic growth, and attendance at test matches dropped markedly. It is impossible to determine with a regression which factor is the driving force behind lower attendance. This might be why the coefficient on real GDPR per capita is insignificant in model 4. Nonetheless, model 5 outperforms model 4 in terms of adjusted R^2 (0.8138 versus 0.8041) and AIC (0.7695 versus 0.8188). The coefficients on match uncertainty and Wanderers are now significant at the 10% level. Using model 5, for a rainless match played against India at Centurion in January of which both the match – and series outcomes were unknown, the aacr was calculated with the other variables held constant at their averages. The outcome is that the introduction of pro-20 caused a 0.213 drop in the aacr.

Table 14: Model 4 – the effect of pro-20 on test match attendance

Constant	0.3494
	0.9244
Pro 20	-0.8450
	0.0658
Kingsmead	-0.0204
	0.9927
Newlands	0.1936
	0.8634
St. George's	-0.3918
	0.8918
Wanderers	-0.3379
	0.1015
Australia	0.8429
	0.0241
England	0.6271
	0.0443
India	0.0638
	0.7813
Pakistan	-1.0913
	0.0001
Sri Lanka	0.3424
	0.3747
Urban population	0.0000001
	0.6687
Real GDPR per capita	-0.00001
	0.9154
Match uncertainty	0.2755
	0.1189
Rating difference	-0.0094
	0.0042
Series uncertainty	0.7250
	0.0011
Rain	-0.4800
	0.0049
February	-1.3800
	0.0036
March	-0.7801
	0.0342
November	-1.4375
	0.0000
December	-0.9707
	0.0025
n	40
R^2	0.9046
Adjusted R^2	0.8041
AIC	0.8188
Durbin-Watson	2.4760
Prob(F-statistic)	0.00001

Table 15: Model 5 – model 4 without real GDPR per capita

Constant	0.0046
	0.9974
Pro 20	-0.8826
	0.0018
Kingsmead	0.1901
	0.8152
Newlands	0.2632
	0.7489
St. George's	-0.1282
	0.9124
Wanderers	-0.3291
	0.0762
Australia	0.8327
	0.0186
England	0.6388
	0.0172
India	0.0639
	0.7733
Pakistan	-1.1036
	0.00001
Sri Lanka	0.3601
	0.2347
Urban population	0.0000001
	0.6483
Match uncertainty	0.2683
	0.0855
Rating difference	-0.0096
	0.0007
Series uncertainty	0.7245
	0.0008
Rain	-0.4817
	0.0035
February	-1.3706
	0.0022
March	-0.7645
	0.0264
November	-1.4431
	0.00000002
December	-0.9767
	0.0016
n	40
R^2	0.9045
Adjusted R^2	0.8138
AIC	0.7695
Durbin-Watson	2.4636
Prob(F-statistic)	0.000002

7.3 ODI's

As mentioned in section 5.2, the fact that ODI-attendance is frequently near capacity makes it difficult to measure the actual demand for this format of the game. It also makes it impossible to weigh attendance with capacity, since there will then not be enough variation between the observations. Hence, the actual attendance of 75 ODI's between 1996 and 2007 played in SA versus SA was regressed with the use of the OLS method.

Two models were attempted – a linear – and a log-linear model. Because the dependent variables are not the same, the R^2 of the two models cannot be compared directly. Ramsey's regression error specification test was used to determine the correct functional form. The null-hypothesis of correct functional form was rejected comprehensively for the linear model (p-value = 0.0063), while it could not be rejected for the log-linear model (p-value = 0.5378). Thus, the log-linear model seems to be the most appropriate. This model is presented in table 16.

Due to space limitations, there will only be a brief discussion on the variables for the ODI – regression, with emphasis on the differences between the ODI – and test match regressions. Contrary to the test match regression, the coefficient of real GDPR per capita is positive, and is significant at 5%. Thus, it would *seem* that an ODI is regarded as a normal good as opposed to test matches being regarded inferior.

Since ratings were not available for ODI's, the absolute run rate difference was used as a proxy for match uncertainty. The coefficient on the absolute run rate difference is negative, as expected, and significant at 1%. The series uncertainty variable was also included, but was not statistically significant.

Contrary to the test match regression, only Wanderers and Kingsmead are significant for ODI's. However, since the dependent variable is not adjusted for the capacity of each stadium, this is of little consequence. The best-attended oppositions are Australia, Sri Lanka, Pakistan, and New Zealand.

Table 16: Model 6 – Log-linear model for ODI's

	Log-linear
Constant	8.5512
	< 0.0001
Buffalo Park	0.7441
	0.2174
De Beers Oval	-0.3721
	0.4545
Goodyear Park	0.3039
	0.4949
Kingsmead	1.0650
	0.0301
Newlands	0.2550
	0.3327
St. George's	0.9000
	0.1305
Wanderers	0.4455
	< 0.0001
Australia	0.3915
	0.0043
England	-0.0694
	0.3162
India	0.0140
	0.8636
New Zealand	0.3700
	0.0005
Pakistan	0.3746
	0.0075
Sri Lanka	0.3257
	0.0618
Zimbabwe	-0.3190
	0.0001
Rain	-0.1671
	0.0090
Absolute run rate difference	-0.0892
	0.0147
Series uncertainty	-0.0007597
	0.9908
Real GDPR per capita	0.0001
	0.0503
Urban population	-0.0000001
	0.2195
Pro-20	-0.4392
	0.0001
February	-0.0756
	0.2574

Table 16 continued:

March	-0.2595
	0.0186
April	-0.4689
	0.0019
October	-0.7708
	0.0004
November	-0.7191
	0.0038
December	-0.9117
	< 0.0001
n	75
R^2	0.8898
Adjusted R^2	0.8300
Durbin-Watson	1.9697
Prob(F-statistic)	< 0.0001

Pro-20 once again has a negative coefficient and is significant at 1%. Thus, this would imply that the introduction of pro-20 may have shifted attendance away not only from test matches, but from ODI's as well. Once again, these results have to be interpreted with care. However, the fact that pro-20 has a negative coefficient even though the coefficient on the income variable for the ODI regression is positive, does indicate that pro-20 might well have had a harmful effect on attendance at both test matches and ODI's.

8. DIAGNOSTIC TESTS

The assumptions that the models in the previous section are conditioned on were tested to confirm that the models are indeed valid. Tests for normality and mis-specification are shown in tables 17 and 18. It is clear that both the normality assumption and the assumption of a correctly specified model are valid for all of the models. Moreover, there is controlled for heteroscedasticity by the use of robust standard errors. Although it is not shown, the null-hypothesis of no serial autocorrelation could not be rejected by the Breusch-Godfrey test for any of the models.

Table 17: Tests for normality

Model	JB statistic	Prob (JB = 0)	Normality
1a	0.22	0.89	Not rejected
1b	0.91	0.64	Not rejected
2a	1.64	0.44	Not rejected
2b	2.25	0.33	Not rejected
3a	0.54	0.76	Not rejected
3b	0.51	0.77	Not rejected
4	0.9	0.64	Not rejected
5	0.95	0.62	Not rejected
6	3.92	0.14	Not rejected

Table 18: Tests for mis-specification

Model	F-stat	Prob(F-stat) = 0	Mis-specification
1a	0.04	0.97	Not rejected
1b	0.12	0.89	Not rejected
2a	0.37	0.70	Not rejected
2b	0.58	0.57	Not rejected
3a	0.24	0.79	Not rejected
3b	0.17	0.84	Not rejected
4	0.36	0.70	Not rejected
5	0.4	0.68	Not rejected
6	1.03	0.36	Not rejected

9. CONCLUSION

This paper was the first that considered the determinants of attendance at cricket matches in SA. Regressions were estimated for both test matches and ODI's. Some important conclusions can be drawn from this paper. Firstly, the effect that income has on attendance is different for test matches and ODI's. The analysis in this paper suggests that test matches are regarded an inferior good by the South African public, while ODI's are considered a normal good. Thus, the higher economic growth experienced in SA of late *may* have increased attendance at ODI's, but decreased attendance at test matches, *ceteris paribus*. Secondly, unlike some previous sport demand studies, the variables used to account for uncertainty of match outcome were statistically significant. Two methods were employed for test matches, the first of which is the author's subjective classification of the match as uncertain or otherwise, depending on various criteria. The second method used the absolute rating difference between teams. For ODI's the absolute run rate difference was used to measure the uncertainty of the outcome of a match. For both ODI's and test matches, a series uncertainty variable was included. This variable was significant in the case of test matches, but insignificant in the ODI model. Preliminary analysis suggests that the introduction of pro-20 may have deterred the demand for test matches and ODI's. However, without further research it cannot be stated categorically that pro-20 has shifted attendance away from longer forms of the game.

"May cricket continue to flourish and spread its wings. The world can only be richer for it."

- Sir Donald Bradman (1958: 239)

WORDS USED: 8 000

10. BIBLIOGRAPHY

AVERY, C., and CHEVALIER, J., 1999. Identifying investor sentiment from price paths – the case of football betting. *Journal of Business* 72(4): 493 – 521.

BAIRAM, E.I., HOWELLS, J.M., and TURNER, G.M., 1990. Production functions in cricket – the Australian and New Zealand experience. *Applied Economics* 22(7): 871 – 879.

BAL, S., 2005. *The revival of the beautiful game.* Cricinfo. [Online]. Accessed: 23 February 2008. Available: http://content-www.cricinfo.com/engvaus/content/story/219378.html

BECKER, G.S., 1993. Nobel lecture – the economic way of looking at behaviour. *Journal of Political Economy* 101(3): 385 – 409.

BHATTACHARYA, M., and SMYTH, R. The game is not the same – the demand for test match cricket in Australia. *Australian Economic Papers* 42(1): 77 – 90.

BORLAND, J., 1987. The demand for Australian rules football. *Economic Record* 63(182): 220 – 230.

BRADMAN, D., 1958. *The art of cricket.* London: Hodder and Stoughton.

BURAIMO, B., FORREST, D., and SIMMONS, R., 2005. Outcome uncertainty and the couch potato audience. *Scottish Journal of Political Economy* 52(4): 641 – 661.

BURDEKIN, R.C.K., and IDSON, T.L., 1991. Customer preferences, attendance and the racial structure of professional basketball teams. *Applied Economics* 23(1b): 179 – 186.

CRICINFO, 2008. Wisden Cricketers Almanack. [Online]. Accessed: various dates in June 2008. Available: http://content-rsa.cricinfo.com/wisdenalmanack.

DUCKWORTH, F.C., and LEWIS, A.J., 1998. A fair method of resetting the target in interrupted one-day cricket matches. *Journal of the Operational Research Society* 49(3): 220 – 227.

EISENBERG, J.D., and SIEGFRIED, J.J., 1980. The demand for minor league baseball. *Atlantic Economic Journal* 8(2): 59 – 69.

GARLAND, R., HAUGHEY, K., and MACPHERSON, T., 2004. Rugby fan attraction factors. *Marketing Bulletin* 15(3): 1 – 12.

GILES, D.E.A., JONES, J.C.H., and SCHOFIELD, J.A., 2000. Our fans in the north – the demand for British rugby league. *Applied Economics* 32(14): 1877 – 1887.

GRATTON, C., and TAYLOR, P., 2000. *Economics of sport and recreation.* New York: Spon Press.

GUJARATI, D.N., 2003. *Basic Econometrics*. 4th edition. New York: McGraw-Hill.

HAUPERT, M., KNOWLES, G., and SHERONY, K., 1992. The demand for major league baseball – a test of the uncertainty of outcome hypothesis. *The American Economist* 36(2): 72 – 80).

HYNDS, M., and SMITH, I., 1994. The demand for test match cricket. *Applied Economics Letters* 1(7): 103 – 106.

LEVITT, S.D., 2004. Why are gambling markets organised so differently from financial markets? *The Economic Journal* 114(495): 223 – 246.

MORLEY, B., and THOMAS, D., 2007. Attendance demand and core support – evidence from limited overs cricket. Applied Economics 39(16): 2085 – 2097.

NARAYAN, P.K., and SMYTH, R., 2003. Attendance and pricing at sporting events – empirical results from Granger causality tests for the Melbourne cup. *Applied Economics* 35(15): 1649 – 1657.

NEALE, W.C., 1964. The peculiar economics of professional sports – a contribution to the theory of the firm in sporting competition and in market competition. *The Quarterly Journal of Economics* 78(1): 1 – 14.

PEEL, D.A., and THOMAS, D.A., 1992. The demand for football – some evidence on outcome uncertainty. *Empirical Economics* 17(2): 323 – 331.

POLLY WORRIED ABOUT TEST CRICKET, 2008. *Cricket365.* [Online]. Accessed: 23 February 2008. Available: http://www.cricket365.com/story/0,,6575_3177216,00.html.

ROSENTRAUB, M.S., SANDY, R., and SLOANE, P.J., 2004. *The Economics of sport: an international perspective.* New York: Palgrave MacMillan.

RINGROSE, T.J., 2006. Neutral umpires and leg before wicket decisions in test cricket. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 169(4): 903 – 911.

ROTTENBERG, S., 1956. The baseball players' labor market. *The Journal of Political Economy* 64 (3): 242 – 258.

SCHOFIELD, J.A., 1983. The demand for cricket – the case of the John Player League. *Applied Economics* 15(3): 283 – 294.

SCHOFIELD, J.A., 1988. Production functions in the sports industry – an empirical analysis of professional cricket. *Applied Economics* 20(2): 177 – 193.

STATISTICS SOUTH AFRICA, 1996 – 2007. *Mid-year population estimates*.

STATISTICS SOUTH AFRICA, 2006. *Gross Domestic Product*.

STATISTICS SOUTH AFRICA, 2007. *General Household Survey*, July 2007.

TEST CRICKET RATING SERVICE, 2008. *The latest ratings for test cricket nations.* [Online]. Accessed: 29 October 2008. Available: http://www.geocities.com/colosseum/stands/8832/.

TURNER, A., 2007. *Cricket popularity questionable.* News24. [Online]. Accessed: 23 February 2008. Available: http://www.news24.com/News24/Columnists/Guest_Column/0,,2-1630-1647_2237381,00.html.

WELKI, A.M., and ZLATOPER, T.J., 1999. US professional football game-day attendance. *Atlantic Economic Journal* 27(3): 285 – 298.

WIKIPEDIA, 2008. *Elo rating system*. [Online]. Accessed: 29 October 2008. Available: http://en.wikipedia.org/wiki/Elo_rating_system.

WILLIAMSON, M., 2008. A short history of the Ashes. Cricinfo. [Online]. Accessed: 23 February 2008. Available: http://content-usa.cricinfo.com/ausveng/content/story/259985.html.