THE BEAUTIFUL GAME?
AN ECONOMETRIC STUDY OF AUDIENCES, GAMBLING AND EFFICIENCY IN ENGLISH FOOTBALL

ECONOMICS PhD THESIS

1997

TIM KUYPERS
UNIVERSITY COLLEGE LONDON

For the 96

Sheffield 15th April 1989

I would like to thank
Donald Verry, Nick Rau and Simon Emerson for their useful comments
Mark Clifton for sharing his gambling knowledge
and Jennifer for her love and support.

Football is the world's most watched sport. This thesis investigates three related aspects of football in England using econometric techniques.

An investigation of the reasons why people watch football, both live and televised matches is undertaken. Particular attention is paid to outcome uncertainty, both match and seasonal. Two equations are developed to explain match attendance and BSkyB television audiences for the 1993/94 English Premier League season. In the match attendance equation capacity constraints are accounted for by use of the Tobit model. It is found that quality factors, outcome uncertainty and supporter loyalty are all important determinants of football attendances but that televising a match on BSkyB does not significantly affect audiences.

The second study focuses on the efficiency of the fixed odds betting market for football in England. It is the efficiency of how market participants utilise available information that is tested. A model of bookmaker behaviour is presented in which the bookmaker maximises their expected share of the total amount bet. It is found that an expected profit maximising bookmaker could set market inefficient odds. Several empirical tests using the ordered probit model with data on prices, publicly available information and experts' predictions are carried out. Evidence of market inefficiency is identified offering profitable betting opportunities.

Productive efficiency of football clubs is the focus of the third study. It investigates how efficiently clubs utilise their inputs to produce playing success. Unlike most previous sports productive efficiency studies, true inputs (i.e. playing ability proxied by wages) and not intermediate outputs (e.g. goals scored) are used in the efficiency estimations. Two techniques, econometrics and Data Envelopment Analysis are used, allowing a useful comparison of their relative benefits. Efficient clubs are identified and the features which make them efficient discussed.

TABLE OF CONTENTS

1. INTRODUCTION

"To say that these men paid their shillings to watch twenty-two hirelings kick a ball is merely to say that a violin is wood and catgut, that Hamlet is so much paper and ink. For a shilling the Bruddersford United AFC offered you Conflict and Art..."

-J.B. Priestley "The Good Companions" 1929

Football is the world's most watched sport. Despite certain commentators and economists doom and gloom[1], post Hillsborough, English football is resurgent. League attendance has risen from 16.5 million in 1985/86 to 22 million in 1995/96[2], stadia are improving rapidly and football is receiving record revenue from television and sponsorship.

This introductory chapter places the following three chapters in context. Section 1.1 discusses sports economics of which these chapters form a part. Section 1.2 looks at the structure of English football and its current importance and Section 1.3 summarises the three chapters.

1.1. Sports Economics

The birth of sports economics has been traced back to a paper on the baseball labour market written in 1956[3]. Since this date a large number of studies have been produced focusing on many different aspect of sport. The areas investigated include: the demand for sports; the labour market including racial discrimination; cartels and competition policy; and productive efficiency[4].

As with the first sport economics paper, the majority of studies produced since have focused on American sports, possibly as a result of the statistically rich nature of top US sports (e.g. American football and baseball). Indeed access to a large robust data set is one of the main advantages of empirical study in sports economics. Not only is the data widely available but it is often of a high accuracy, something that cannot be said for many areas of empirical study.

Sport around the world is big business, making sports economic studies interesting in their own right. However, as with the three main chapters in this thesis, there is often a wider applicability of the work. In particular studies of racial discrimination[5], market efficiency[6] and even crime[7] have used sports data but drawn conclusions for the wider economy.

[1] For example see Szymanski and Smith (1993).
[2] Source Rollin (1996).
[3] Rottenberg (1956).
[4] For a good review see Cairns et al. (1980).
[5] For a survey see Kahn (1991).
[6] For example Figlewski (1979).
[7] For example see McCormick and Tollison (1990).

Goff and Tollison (1990) identify a recent growth in sports economics literature directly testing hypothesis using econometric techniques. Whilst making the rather arbitrary distinction between "sports as economics" and "the economics of sport" they coin the phrase "sportometrics" to describe papers which test economic hypothesis using sports as their laboratory. The following three chapters all make strong use of econometric techniques and as such can be seen as part of this growing literature rather glibly labelled "Sportometrics".

1.2. Football In England

In the 1994/1995 season English league clubs had a turnover of around £470 million[8]. Whilst being a significant industry in itself, football's success also impacts on several other industries, most notably sports leisurewear manufacturing and broadcasting. Indeed in 1994/1995 only 42% of the Premier League clubs' turnover was made up of gate receipts[9].

The acquisition by satellite television of football's television rights has had a major impact on the broadcasting industry in Britain[10]. Films and sport have been identified as the key drivers in subscription television growth[11]. In Britain no sport is watched more than football and BSkyB (British Sky Broadcasting) used its live Premier League football rights as a key marketing device in developing its customer base. Currently many clubs are investigating the possible exploitation of pay per view television, which will be made much easier by the increased channel availability resulting from the imminent digitalisation of television. France and Italy already run pay-per-view systems where viewers may purchase "armchair season tickets" for their favourite club. In the longer term many football clubs envisage setting up their own local cable television station. Newcastle United already transmits its games live to clubs, pubs and cinemas in its area.

Football is a high profile industry and in 1996 the Premier League and its deal with BSkyB for television rights attracted the attention of the competition authorities. The Office of Fair Trading (OFT) is taking the Premier League to the Restrictive Trade Practices Court in an attempt to have the television deal struck down. They claim that the Premier League is acting as a cartel in the marketing of its television rights.

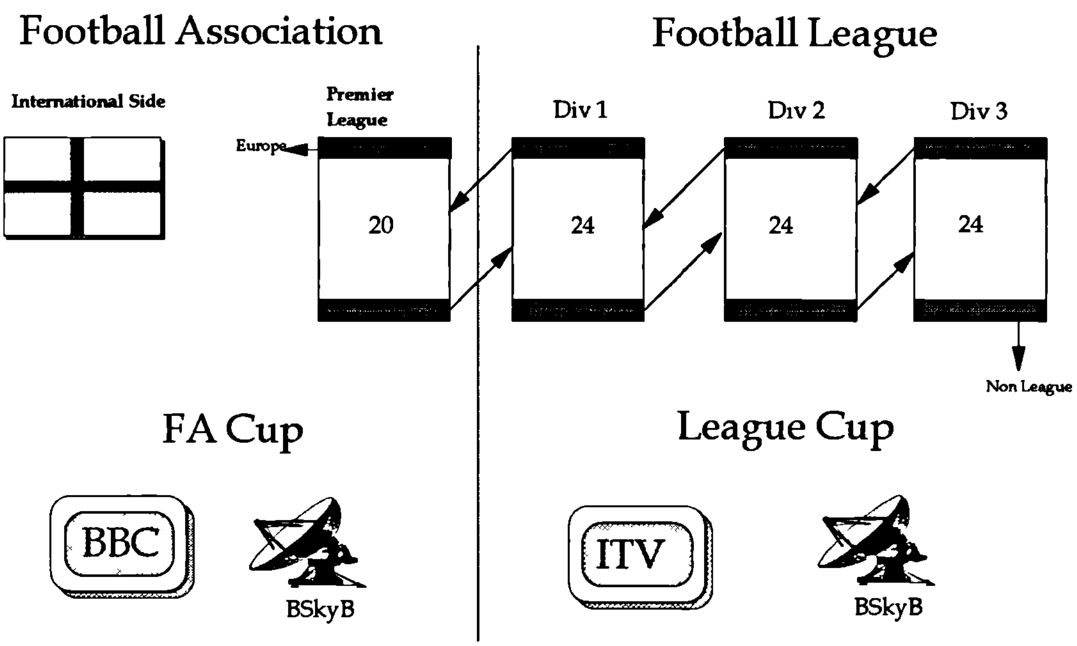
Football in Europe is currently in a state of flux. Both England and Scotland have recently reorganised their league structure and there is increasing talk of a European super league, in which the major clubs in each country would play each other regularly in front of a Europe wide television audience. In what many see as the first step towards the super league, the European Cup has changed from a purely two legged knockout competition into a league and knockout system grandly called the 'Champions League'.

Figure 1.1 below sets out the structure of English professional football. There are two governing bodies, the Football Association and the Football League. The Football Association is in charge of the international side that competes on behalf of England in the

[8] Source Deloitte & Touche (1996).
[9] Source Deloitte & Touche (1996).
[10] See for example Horne & Clarke (1994) pp. 220-224.
[11] See for example Hall et al. (1994).

World Cup and European Championships and is made up of English players from any club side. The Football Association is also the umbrella organisation of the Premier League, a league containing 20 clubs who play each other twice a season[12]. Clubs finishing at the top are allowed to enter into European competition in the subsequent season. The league winners enter the Champions League and the runners up and usually third places enter the UEFA[13] Cup. The bottom clubs are relegated into the First Division of the Football League. There are three Football Leagues with relegation and promotion operating in each[14]. The bottom of the Third Division is replaced by the winner of the Vauxhall Conference, the most prestigious amateur league (subject to minimum ground requirements). The Football Association runs a knockout cup competition (the FA Cup) that is entered by all clubs including non-league clubs. The winner of this competition enters next seasons European Cup Winners Cup. The League Cup[15] is run by the Football League for all League and Premier League clubs with the winner entering the UEFA Cup.

Figure 1.1: Structure of English Football 1996/1997



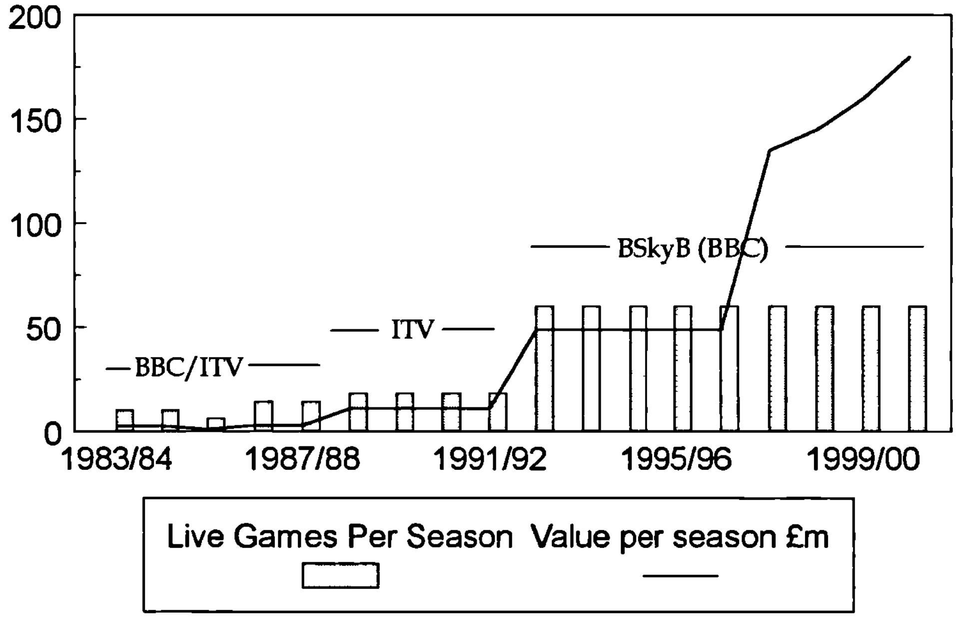
Currently BSkyB and the BBC have a joint deal with the Football Association. BSkyB is the domestic satellite channel transmitting several different channels, three of which are subscription sports channels. BSkyB has rights to show 60 live Premier League games a season plus live England internationals. The BBC can show highlights of both Premier League and England games and both BSkyB and BBC have live broadcast rights for the FA Cup. The FA Cup Final is one of eight protected events the Department of National

[12] 22 prior to the 1995/96 season.
[13] Union of European Football Associations.
[14] Prior to 1992/1993 the four leagues were known as Divisions 1,2,3 and 4.
[15] Currently sponsored by Coca-Cola.

Heritage has banned from being pay-per view events[16]. BSkyB and the BBC paid the Football Association £304 million for domestic football television rights for five years[17].

It is interesting to look at the rise in value of television rights for the Premier League as an indication of the growing importance of football. Prior to BSkyB's entry into the market there was a duopsony in which BBC and ITV were the only purchasers. When BSkyB entered the market in 1992/93 there was a corresponding increase in price. Figure 1.2 shows the number of live games broadcast each season along with the annual fee for television rights between 1983/84 and 2000/01. From 1983/84 to 1987/88 BBC and ITV shared coverage and paid on average £2.4m a season. In the five following seasons ITV had an exclusive contract paying an increased fee of £11 million pounds a season. With the entry of BSkyB into the market there was a large increase in both the price paid per season (£49m) and the number of live games shown (60). The value of rights has again increased with the signing of a new contract between BSkyB and the Premier League. The new deal gives BSkyB rights until 2000/01 when the annual price will be £180m reflecting the importance of football rights for a viable subscription television service.

Figure 1.2: Value of Premier League TV Rights



1.3. Outline and Summary

All three chapters focus on English football, countering some of the current American bias in the sports economics. Each of the chapters looks at an aspect of football not investigated before and each chapter has econometrics at its heart. The following sections summarise the chapters.

[16] The others are the Scottish FA Cup, the FIFA World Cup, home cricket test matches, the Olympics, the Derby, the Grand National and finals weekend at Wimbledon.

[17] Source BSkyB (1994).

1.3.1. An Econometric Study of Why People Watch English Football

For the first time the reasons why people attend football matches and/or watch live televised matches are modelled and contrasted. Two independent equations are developed to explain match attendance and BSkyB live television audiences for the 1993/1994 English Premier League season.

This chapter investigates factors which affect football audiences and can be influenced by the structure and regulations of the league. A particular focus is placed on the effect of outcome uncertainty on audiences. Other factors investigated include quality, supporter loyalty, the effect of television and the day on which the game is played. Longer term influences on football audiences such as: price; income; changes in the supply and medium of delivery (e.g. the move from terrestrial to satellite television); the availability and price of alternative leisure activities; quality of stadia; and the change in the level of hooliganism are not looked at. Individual match data is used and these longer term influences are assumed constant over the season.

There have been several approaches to measuring demand for sport and in particular demand for football. All these papers are reviewed with particular focus on studies using outcome uncertainty measures. Three types of outcome uncertainty have been identified in the literature: match outcome uncertainty, seasonal outcome uncertainty and the absence of long run domination. As this study focuses on short term effects, only the first two types of outcome uncertainty are investigated.

The measure of television audiences used is the proportion of potential viewers (i.e. Sky Sports subscribers) watching a particular football match. The measure of match attendance is simply the reported match attendance. In both equations the audience is assumed determined by the expected quality of the game, supporter loyalty, outcome uncertainty and the opportunity cost of watching football.

In measuring the determinants of attendance, several variables were often utilised. Four different aspects of quality are investigated:

- player quality as measured by the number of internationals in a team;
- expected excitement proxied by recent goal scoring performance;
- games of special interest; and
- expected success proxied by recent performance.

The opportunity cost of watching football was looked at by including variables reflecting the day on which the game is played, if it is televised and the weather. Loyalty was proxied by an average of previous years attendances and two measures of outcome uncertainty were developed. The measure of match outcome uncertainty uses bookmakers odds in a refinement of a variable used by Peel and Thomas (1988). The measure of seasonal outcome uncertainty is original and an improvement on measures developed previously by Jennett (1984) and Cairns (1987). Many of the variables were calculated for the home team, for the away team and for the match (i.e. sum of home and away values).

In previous sports demand studies little attention has generally been paid to the econometric techniques used; most studies employed ordinary least squares estimation. This study utilises both models of groupwise autocorrelation and heteroscedasticity and a tobit model (with and without heteroscedasticity) to take account of the censored nature of the data.

1.3.2. Information and Efficiency: An Empirical Study of A Fixed Odds Betting Market

There has been a large volume of literature testing the efficient market hypothesis in both financial and more recently betting markets. This chapter looks at a particular form of betting largely ignored, betting on football fixed odds. Fixed odds betting has similar characteristics to options markets and in particular markets for European call options. A fixed odd market is where the bookmaker offers odds prior to an event (in football's case three days before) and these odds do not move in response to betting patterns or additional information.

Fama's (1970) definition of efficiency is adopted where an efficient market is a market whose prices fully reflect all available information. Previous studies of efficiency in gambling markets give mixed results, some identify inefficiency but few are able to turn the identified inefficiencies into exploitable opportunities. The chapter summarises these studies, the majority of which focus on horse racing.

A model of bookmakers' odds setting decisions is developed. The model looks at bookmaking as an information market. It assumes that the bookmaker is an expected profit maximiser, maximising his expected share of the handle (total amount bet). Several rules of punters' betting behaviour are assessed. The punter's decision as to which outcome to bet on is modelled but not their decision to enter the market. The model shows that if punters display a bias in their expectations (e.g. supporters believe their team has a higher chance of winning than is really the case) then the expected profit maximising bookmaker would set inefficient odds, i.e. odds that do not reflect fully all available information.

Several empirical tests of efficiency are carried out. The data used is from the four English professional leagues in the 1993/94 and 1994/95 seasons (3382 matches). Initially, implied probability calculated from the odds is compared with the outcome probabilities (i.e. the number of times the event actually occurred) for about 20 odds groupings.

Several semi-strong tests of efficiency were carried out using the ordered probit model, the first time the ordered probit has been used in a study of this type. Initially equations were estimated with the outcome of matches being explained by the odds for that match. Further equations were estimated using additional publicly available information (e.g. goals scored in the last three matches). Betting rules were developed based on the results of the estimations.

Another test of efficiency was carried out using "experts" predictions. If the "experts" (newspaper tipsters) predictions offered more information than was contained in the odds, this would indicate inefficiency.

1.3.3. Productive Efficiency in English Football

While productive efficiency has been investigated in many sectors including sports, there has been no rigorous investigation of efficiency in English football. The fourth chapter assesses the productive efficiency of English football clubs using both econometric and non-parametric techniques. Data Envelopment Analysis (DEA), the non-parametric technique chosen, has not been used in sports economics before. The study provides a useful comparison of the relative benefits of DEA and econometric techniques in assessing efficiency.

Much of the previous work in sports productive efficiency studies has employed "intermediate outputs" and not true inputs. Intermediate outputs are outputs of the production process but not the final result (e.g. goals scored in a football match). This study uses wages as a proxy for playing talent and thus avoids the use of intermediate outputs.

Two data sets were used one containing 47 clubs from 1974-1989 and another with 39 clubs from 1974-1994. Two econometric methods were used; the fixed effects model which estimates a deterministic frontier; and a stochastic frontier estimation. The output measure in these estimations was league performance and the inputs were ability per squad member (proxied by relative wages), number of professionals in the squad, number of players used in the league and management tenure (i.e. how long the current manager has been in the job). The results were encouraging with all variables significant and of the expected sign. Efficiency of individual clubs was then calculated and reported in detail.

Two DEA specifications were estimated for both constant and variable returns to scale. The first specification used identical variables to the econometric equations thus facilitating comparison. The second specification made use of the feature of DEA allowing multiple inputs <u>and</u> outputs. In addition to league performance, outputs included performance in the FA Cup an League Cup.

1.4. References

BSkyB (1994). *Sky Media Pack 1994.* BSkyB Public Relations Department.

Cairns, J.A., Jennett, N. & Sloane, P.J. (1986). The Economics of Professional Team Sports: A Survey Of Theory and Evidence. *Journal Of Economic Studies.* Vol. 12, pp. 3-80.

Cairns, J.A. (1987). Evaluating Changes In League Structure: The Organisation of the Scottish Football League. *Applied Economics.* No. 19(2), pp. 259-275.

Deloitte & Touche (1996). *Annual Review of Football Finance.* Manchester.

Figlewski, S. (1979). Subjective Information and Market Efficiency in a Betting Market. *Journal of Political Economy.* Vol. 87, No.1. pp. 75-88.

Fleisher, A. (1990). Crime or Punishment? Enforcement of the NCAA Football Cartel in *Sportometrics.* (eds. Goff, B. & Tollison, R.). Texas A&M University Press, USA.

Goff, B.L. & Tollison, R.D. (1990). Sport as Economics in *Sportometrics.* (eds. Goff, B. & Tollison, R.). Texas A&M University Press, USA.

Hall,T., Beilby, M. & Duckworth, R. (1994). European Pay TV - BSkyB and Beyond. *S.G.Warburg Media Research.* October.

Horrie, C. & Clarke, S. (1994). *Fuzzy Monsters: Fear and Loathing at the BBC.* Mandarin Books, London.

Jennett, N. (1984). Attendances, Uncertainty of Outcome and Policy in the Scottish Football League. *Scottish Journal of Political Economy.* Vol. 31 No. 2, pp. 176-198.

Kahn, L.M. (1991). Discrimination in Professional Sports: A Survey of the Literature. *Industrial and Labor Relations Review.* Vol. 44, No. 3.

McCormick, R.E. & Tollison, R.D. (1990). Crime on the Court in *Sportometrics.* (eds. Goff, B. & Tollison, R.). Texas A&M University Press, USA.

Peel, D. & Thomas, D. (1988). Outcome Uncertainty and the Demand For Football: An Analysis of Match Attendances in The English Football League. *Scottish Journal of Political Economy.* No. 35(3), pp. 242-249.

Rollin, J. (ed.) (1996). *Rothmans Football Yearbook 1994-95,* Headline Book Publishing.

Rottenberg, S. (1956). The Baseball Players Labour Market. *Journal of Political Economy.* Vol. 64, No. 3, pp.242-258.

Szymanski, S. & Smith, R. (1993). The English Football Industry: Profit Performance and Industrial Structure. *Unpublished Mimeo, London Business School.*

2. WHY PEOPLE WATCH ENGLISH FOOTBALL: AN ECONOMETRIC STUDY OF MATCH ATTENDANCE AND TELEVISION AUDIENCES

2.1. League Structure, Regulations And Outcome Uncertainty

The football industry is in a state of flux. Recently both the English and Scottish Leagues have been re-organised and there is much talk about the formation of a European "Super League". Thus the relationship between audiences and the structure and regulations of the league are particularly interesting. This chapter investigates the factors which affect football audiences and are influenced by the structure, rules and regulations of the league as determined by football's governing authorities.

The major factors influencing football audiences investigated in this chapter include: the quality of football; supporter loyalty; whether the game is televised; the day the game is played; and outcome uncertainty. Several studies have investigated similar influences on sports audiences[18]. This chapter is intended as an extension and refinement of this literature. Longer term influences on audiences for football such as: price; income; changes in the supply and delivery medium of football; the availability and price of alternative entertainment; quality of stadia; and change in the level of hooliganism are not investigated here.

Clearly the football authorities can directly determine the day the game is played and whether the game is televised. However, they can also influence the quality of the game by affecting the availability of top quality players through transfer regulations and their attitude to foreign players. Longer term controls on quality that are not investigated in this chapter include altering the laws of football (e.g. the recent introduction of a revised offside law and harsher punishment for tackles from behind), changing the number of the games within the season (e.g. from 1995/96 the Premier League was reduced to twenty clubs) and instituting training programmes for youth players (e.g. the FA School of Excellence at Lilleshall). It is less clear that the authorities can influence fan loyalty. The introduction of supporters' clubs and the encouragement of product discrimination between teams can help but it is likely to be largely historically determined[19]. Of all these influences on demand, the one that has been investigated most extensively in the literature is outcome uncertainty.

It is argued in this chapter that the football supporter, both the match attendee and the television viewer, values uncertainty of outcome. Football is more attractive and therefore they are more likely to watch if outcome uncertainty exists. Three distinctive forms of outcome uncertainty have been identified in the literature: the absence of long run domination; seasonal uncertainty; and match uncertainty.

---

[18] See for e.g. Peel & Thomas (1988) and Cairns (1987).

[19] See Dobson & Goddard (1993).

Long run domination by one or two clubs in a league is likely to have a negative impact on attendances and television audiences. At the beginning of every season if it is clear that only one or two clubs could win the championship[20], interest in the season is likely to be diminished. Supporters of teams may become disillusioned if they believe their team never has a chance of winning the league. For example, in Scotland, Glasgow Rangers have won the Scottish Premier Titles in every season since 1988/89, leading to the belief amongst most supporters that their club will not win the title.

Seasonal outcome uncertainty is uncertainty surrounding the championship and relegation. Two types of seasonal uncertainty were distinguished by Cairns et al. (1986), one when the team you support is involved (i.e. they are in contention for the championship) and two where the ultimate uncertainty surrounding the championship is valued independently of your team's involvement. This distinction does not help to test the effect of outcome uncertainty on attendances and especially not on television audiences. A championship contender will attract television viewers from its own supporters, supporters from the opposing team and from so called 'neutrals'. This is also true in attendance at games, a championship contender will attract fans from both sides plus neutrals. It is impossible to distinguish in the data between fans watching or attending because their team is in the championship race and those supporters from other clubs or neutrals who value the championship race independently of their team's involvement.

Match outcome uncertainty will vary with each game and is uncertainty surrounding the result of a single match. This chapter introduces new measures of both match uncertainty and seasonal uncertainty but no attempt is made to look at the absence of long run domination.

Outcome uncertainty has implications for the structure and organisation of the league. Throughout English football's history there has been cross subsidisation between clubs. For example, between 1960 and 1985 the away club in league fixtures kept 20% of gate receipts. This, in effect, transferred revenue from clubs with large home gates to clubs with small home gates. Today, cross subsidisation comes via television revenue distribution. In addition to a match fee, each club receives revenue from the television deal independent of the number of times they appear on television. One defence of cross subsidisation can be made on outcome uncertainty grounds. dell'Osso and Szymanski (1991) show that there is a high correlation between expenditure, both on wages and transfers, and success. Thus, reducing the concentration of wealth in the league reduces the concentration of playing ability and therefore increases outcome uncertainty. If outcome uncertainty is an important determinate of revenue, some form of cross subsidy may increase total league profits.

Outcome uncertainty can be influenced by a number of other factors, the most important of which are transfer regulations and relegation and promotion arrangements. For example, a maximum number of non-home grown players per club or the introduction of a draft

[20] Throughout this paper the terms champions is used to describe the winners of the league and championship to describe the league competition which determines the champion.

system, as currently used in many North American sports, would all restrict the concentration of playing talent. A draft system allows the team that comes bottom of the league to have first pick of the new talent entering the league in the subsequent season. A recent example of how relegation and promotion rules affect outcome uncertainty is the introduction of the playoff system in the English lower leagues. Instead of the top three clubs being promoted, the top two clubs and the winner of a knockout competition between the third to sixth placed teams were promoted. This increased the number of clubs in contention for promotion and increased seasonal outcome uncertainty.

In the same way that outcome uncertainty is important for football clubs, it has important implications for television companies. Subscription television channels receive their income from three main sources, subscription, sponsorship and advertising all of which are affected by outcome uncertainty. Advertising and sponsorship revenues are a function of both the size and wealth of the audience. The more attractive the package that is shown (e.g. the more outcome uncertain the schedule) then the more subscribers the channel will have. If outcome uncertainty is an important determinant of television audiences then outcome uncertainty is an important determinant of advertising and sponsorship revenue.

2.2. Previous Sports Economics Studies

There is an extensive sports economics literature[21]; some of the more interesting and relevant papers are discussed below. The majority of the work has been carried out in the United States but there are several studies in the UK, most of which focus on cricket and football. The discussion below concentrates on those papers that look at English football and also those papers that try and assess the effect of outcome uncertainty on attendances.

Dobson & Goddard (1993) identify what they call medium term and long term influences on demand for football in England. Medium term factors are defined as loyalty, admission prices and the home team's style. Competition from clubs within the area, historical playing success and socio-economic factors are seen as longer term factors. A two stage analysis is used to determine the effects of these medium and long term factors. They find evidence of both a significant price effect and a strong loyalty component. Interestingly they also find clubs with an early admission to the league maintain an advantage in terms of their base level of support.

An earlier study by Bird (1982) attempted to explain what was then considered the long term decline in Football League attendances. He used real admission prices and travelling costs, weather, the level of hooliganism and the number of goals scored to explain the decline in attendances. He found that football was an inferior good but found no support for the hypothesis that changes in weather, hooliganism and goals explained the decline in attendances.

[21] For a good review see Cairns et al. (1986 .

A paper by Szymanski and Smith (1993) extended demand estimation into a system of equations which describe the behaviour of a utility maximising owner subject to demand and production constraints. However, the implied demand function in this model is very simple and their approach does not identify the factors upon which this chapter concentrates.

As well as the studies concentrating on English football there has been extensive debate in the literature about outcome uncertainty and its affect on the demand for sports. Table 2.1 sets out the most important papers along with the exact form of outcome uncertainty tested, the variable(s) constructed, the data used and the conclusions reached.

Cairns et al. (1986) noted that discussion of outcome uncertainty had been limited, unmethodical and confused with inadequate attention being paid to empirical specifications. Since 1986 there have been several other studies including one by Cairns himself[22] but many of Cairns' criticisms remain valid. All previous studies have looked at outcome uncertainty and its effects on match attendance, none have investigated the effect on television audiences.

Three studies specifically focus on match uncertainty. Hart et al. (1975) and Peel and Thomas (1988) look at English football while Jones and Ferguson (1988) concentrate on the National Hockey League (based in the United States and Canada). Hart et al. study four English league clubs for two seasons in the late 1960s. They estimate an equation for each club separately and proxy outcome uncertainty by the logged difference in league positions. Their results show mixed support for the outcome uncertainty hypothesis. Jones and Ferguson in their study of the 1977/78 National Hockey League season set up a dummy variable equal to one if the match involves two teams from the top three or two teams from the bottom three league positions. This dummy is intended to reflect close games. Jones and Ferguson find no support for the outcome uncertainty hypothesis. Both these proxies for outcome uncertainty are rather crude. They do not take into account many important factors such as home advantage, injuries or star player absence all of which can influence the expected outcome of the game without being reflected in league position.

An improvement on these measures was the use by Peel & Thomas (1988) of betting odds to reflect outcome uncertainty. If the odds were calculated efficiently then all available information prior to the game should have been taken into account. However, Peel & Thomas used the probability of a home victory as the proxy for outcome uncertainty. This is incorrect because at both high and low probabilities of a home win, outcome uncertainty is low. Thus by using home probability they are really testing how supporters value expected success of the home team.

More attention has been focused on testing seasonal outcome uncertainty and there is a greater amount of evidence to support this hypothesis than there is in defence of the match outcome uncertainty hypothesis. Noll (1974) found weak support for a seasonal outcome

---

[22] See Cairns (1987).

uncertainty effect on attendances in his study of ice hockey and baseball. In the baseball study Noll used a dummy variable which indicated whether or not the championship race was close in a particular year. However, this dummy could not reflect how close this race was as it did not take account of either the number of teams in contention nor the length of time that the race was close.

Borland (1987) in his study of Australian rules football experimented with four different measures of seasonal uncertainty, but found support for only one and this was only in one specification of the model. The four proxies he used were: the difference in games won between the first and last team; the sum of the coefficients of variation in games won; the average number of games behind the leader; and the number of teams in contention.

Several studies look at the championship significance of single games. Jennett (1984) calculates the reciprocal of the number of matches required to win the championship prior to each match. He finds this measure significant but utilises information not available to the spectator when he makes his decision to watch a match or not, namely the number of points required to win the championship. Cairns (1987) also finds that his championship significance variable is significant. He improves on Jennett by assuming that people compare teams' performance with each other and not against an unknowable championship winning total. However in doing this, Jennett relies on arbitrary assumptions regarding the anticipated success rates of clubs. Whitney (1988) develops a variable of the average expected probability of winning the championship. Whitney includes both this variable and its square in his estimation but finds only weak support for the outcome uncertainty hypothesis.

Absence of log run domination has only been tested explicitly once. Borland (1987) found that there was no relation between long run domination and lower attendances.

Table 2.1: Outcome Uncertainty In The Literature

Authors	Testing	Variable	Data	Result
Noll (1974)	Seasonal	- Whether team in contention for playoff -Whether championship race close	Ice Hockey Baseball	Weak support Weak support
Hart et al. (1975)	Match	Log difference in league positions	4 English Football Clubs 1969/70-1970/1971	Mixed
Jennett (1984)	Seasonal	- Championship/relegation significance of each game	Scottish League Football 1975-1981	Strong support
Borland (1987)	Seasonal	-Diff. in games won between first and last -Sum of coefficients of variation of game won -Average no of games behind the leader -Number of teams in contention	Victorian Football League (Australian Rules) 1950-1986	Weak support
Cairns (1987)	Seasonal	-Dummy of contention in championship	4 Scottish football clubs 1969/70-1979/80	Strong support
Jones & Ferguson (1988)	Match	-Dummy for top of the table and bottom of the table matches	NHL Season 1977/78	No support
Whitney (1988)	Seasonal	Average expected probability of winning	Baseball 1970-1984	Weak support
Peel & Thomas (1988)	Match	-Betting odds (probability of home win)	1981/82 English football league matches	Prob home win significant

2.3. Model Assumptions

Two independent equations are estimated, one for television audiences and one for match attendance. In order to isolate the factors affecting football audiences that can be influenced by the football authorities, individual match data for one season was looked at. Looking at one season, in this case 1993/1994, allows a number of longer term influences to be assumed constant. This is similar to the approach taken by Peel & Thomas (1988).

Live Premier League football is only available on the Sky Sports channel. Sky Sports is obtained in the UK either via DTH[23] satellite systems or via cable and is an encrypted subscription channel. Satellite and cable companies offer several different subscription packages with varying combinations of the 30 or so different channels currently available.

When looking at the television audiences, the decision modelled is whether the viewer with access to live football watches football or chooses some other activity. The decision to buy satellite or cable television and subscribe to Sky Sports is not investigated.

Although Sky Sports is a subscription channel, it is not a pay-per-view system, i.e. individuals are not charged per programme watched. The minimum subscription contract for satellite is one calendar year and for cable at least one month. A Sky Sports' viewer could subscribe in order to mainly watch one sport or perhaps a number of sports. Thus, although there is undoubtedly a link between the availability and quality of football on Sky Sports and the number of subscribers, it would be impossible to characterise this link in any useful way for the purpose of this study.

In order to isolate the required effects the basic model is described in terms of the proportion of Sky subscribers who watch each live football match.

Proportion of Sky Sports Subscribers watching each football match
= F(expected quality, loyalty, outcome uncertainty,
opportunity cost of watching football)

For the Sky Sports subscriber the marginal cost of watching a live match is zero. It is suggested that the individual decides whether to watch a game based on the attractiveness of that game to her and the opportunity cost. It is assumed that the availability, quality and cost of other activities both leisure and work (pecuniary and non pecuniary) remain constant throughout the season. The opportunity cost of watching football is narrowly defined and the proxy variables used are described in more detail in the next section. Individuals will have different preferences for the various attributes identified but the more attractive the game the bigger will be the aggregate television audience.

It is possible that as the subscriber base grows, people less committed to watching sports (and in particular football) will become subscribers. In order to test this, the regressions

[23] Direct To Home satellite systems. There are also a small number of MATV (Mast Antenna Television) in the UK.

were repeated using audience as the dependent variable (not percentage of subscribers) and adding a time dummy to take account of the growth in subscribers. The results of these regressions were almost identical to those reported below.

The dependent variable is calculated using data from BARB (British Audience Research Board) which produces the official ratings for the UK television industry. The BARB figures are based on a weekly diary of over 4000 households whose televisions and videos are electronically monitored. The panel members represent all regions and are based on a multi-stage stratified and unclustered sample. The population characteristics are determined by an annual survey of 44,000 households. BARB produces figures for the number of individuals who have access to Sky Sports and the number of individuals who watch each live match (average over every minute of the programme).

Figure 2.1 below shows the proportion of subscribers who watched each game over the 1993/94 season. This proportion varies widely ranging from 21% (Manchester United versus Blackburn 26/12/1993) to 6% (Norwich versus Arsenal 13/2/1994). There were 5.9 million subscribers[24] at the beginning of the season rising to 7.5 million by the end of the season.

Figure 2.1: Proportion Of Sky Sports Subscribers Watching Live Football

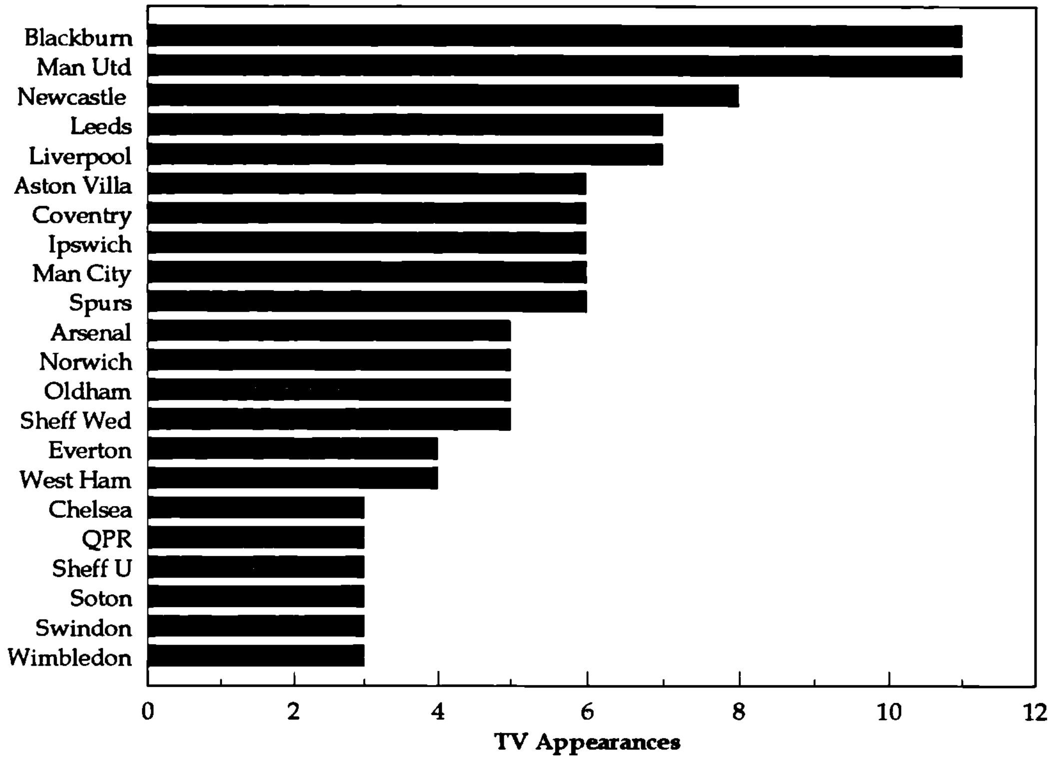
Proportion

Figure 2.2 below illustrates the number of television appearances per club. The clubs with most appearances were the teams in contention for the championship, Blackburn and Manchester United. The majority of clubs had 5, 6 or 7 appearances; however there were a number of clubs with only three appearances. An equal distribution of games would see

[24] Or more formally 5.9 million people in subscriber households.

each club playing 5 or 6 games. For 10 of the 60 televised games the stadiums were at capacity.

Figure 2.2: Television Appearances 1993/94



One potential problem with the dependent variable is that the BARB survey only includes households and does not include public houses or clubs, many of whom show satellite television. This could potentially bias the results; however the Sunday licensing laws were not relaxed until after the 1993/94 season and thus approximately half the televised games were shown outside licensing hours.

A similar approach is taken to the match attendance investigation with the dependent variable being the official attendance at the ground. As well as the assumptions made about alternative time use in the television equation there are assumptions that prices and income are constant over the season.

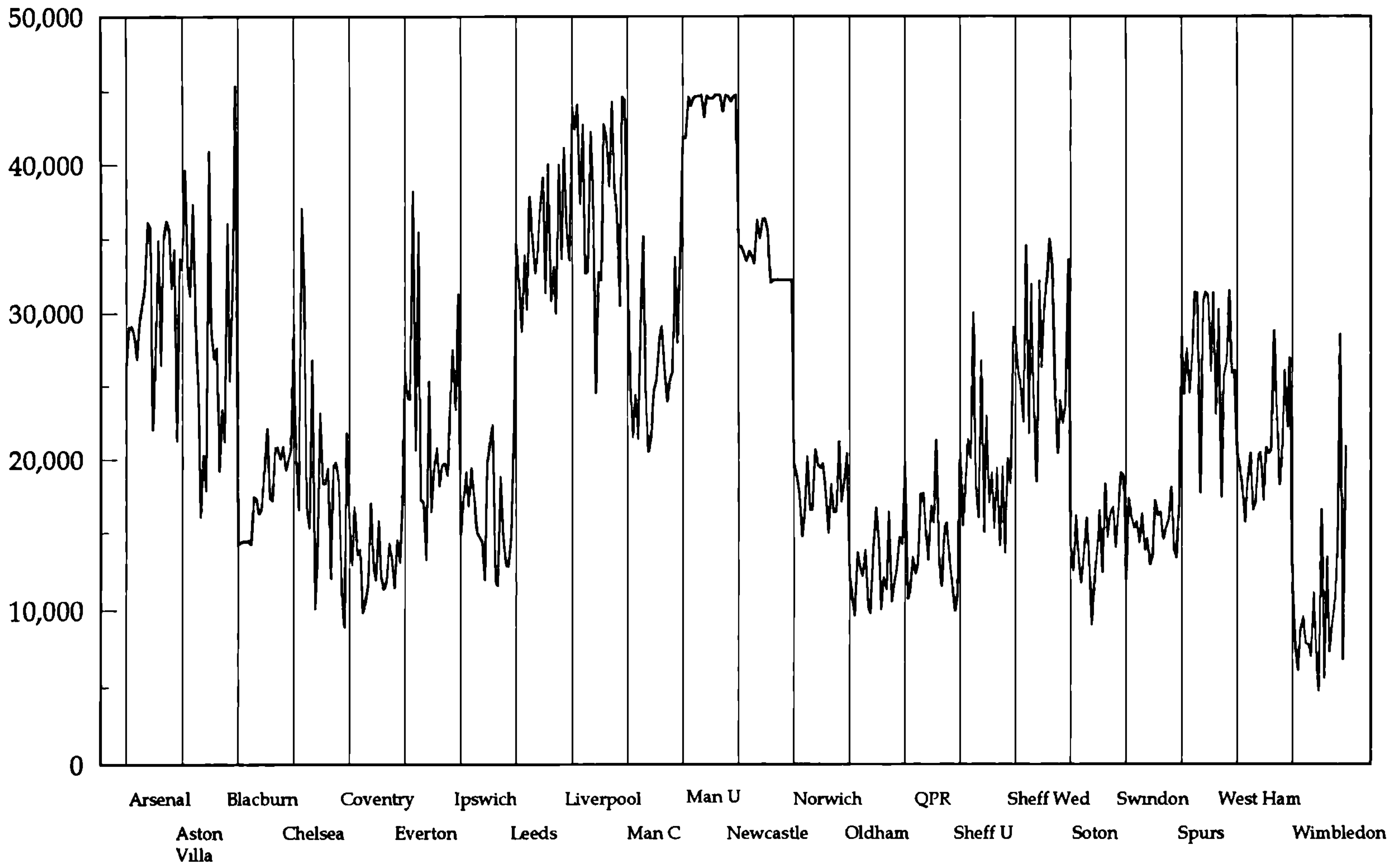
The football season is just under nine months long and runs from late August to mid May. Over this period it would seem plausible to assume that income, quality of stadia, changes in the level of hooliganism and the supply and medium of delivery of live football are constant. The assumption that the price of attendance is constant over the period is important and thus needs more consideration. This assumption, also adopted by Peel & Thomas (1988), does not seem too restrictive as football clubs do set their prices at the beginning of the season and these remain constant all season. There are however often price differences between different parts of the ground and there are certainly regional price differences. It is intended that the regional price differences be picked up by the loyalty variable (or base level of support variable). There is little direct competition between football clubs because of differentiation of products due to supporter loyalty, spatial distribution of teams and the fact that fixtures are arranged to avoid direct clashes.

Thus the only factors varying in the individual's decision to attend a match or not are the attractiveness of that match to her and the opportunity cost of attending. In the aggregate the function estimated is:

Attendance at game = F (expected quality, loyalty, outcome uncertainty, opportunity cost of attending match)

Figure 2.3 shows the dependent variable i.e. Premier League attendance in the 1993/94 season. Each section of the graph shows one clubs home games (21 a season). As can be seen there is considerable difference between clubs in both the level of their attendance and the variation over the season. Manchester United's attendance is fairly constant with an average in excess of 44,000. This is in sharp contrast to Wimbledon who have an average attendance below 10,000 and have much more variation over the season. The graph indicates however, that most clubs do have a considerable amount of variation in attendance and Manchester United is fairly untypical. Some of the variation can be put down to capacity changes due to redevelopment work. For example Newcastle's fall in attendance mid-season was caused by reduced capacity due to building work.

Figure 2.3: Premier League Football Attendance 1993/1994



2.4. Variables And Data

For many of the following variables values are calculated for the home team, the away team and the match (sum of home and away). As discussed in more detail in the results section this is done to allow isolation of home club and away club effects where appropriate.

2.4.1. Expected Quality

This category of variables attempts to measure the expected quality of each fixture independently of outcome uncertainty. It is the expected quality when the individual makes the decision to watch the game that is important and not the actual quality of the game. The higher the expected quality of the fixture the higher would be the expected attendance. Four differing quality attributes are identified. These attributes are described below along with the variables used to proxy them.

2.4.1.1. Player Quality

One aspect of the quality of a fixture is the skill level of the players. The variable used to proxy this was the number of players in a game who had represented their country (i.e. internationals) within the last three seasons. The higher the number of internationals, the higher would be the expected quality and therefore the higher the crowd. Thus the expected sign on this variable is positive. The source of this information was Rollin (1994).

2.4.1.2. Excitement

A major part of the excitement generated at a football match revolves around goals scored. A proxy of expected excitement is the expected number of goals scored. The measures constructed from Rollin (1994) to proxy expected goals were goals scored for and against in the previous three games. The higher the expected number of goals scored the higher the crowd and thus the expected sign of this variable is positive.

2.4.1.3. Games of Special Interest

Some games have an inherent quality because of historical factors or regional rivalry. Supporters will often attend a match regardless of other quality factors if it is a game of special interest. A dummy variable was set to 1 when the game is of particular interest, for example, when it is a local derby[25] or a fixture with a special history. The expected sign is positive. This category was subjectively but narrowly defined, totalling 5% of all fixtures. The games included in this category were Merseyside, Manchester, Yorkshire, North London and East Anglian derbies plus Manchester United versus Liverpool, Everton and Leeds.

[25] A derby is a match between two local sides, after the Earl Of Derby who also game his name to the classic horse race at Epsom.

2.4.1.4. *Success*

One aspect of quality for supporters is success. Identification with a winning side is part of the enjoyment gained from watching football. Thus expected success of your team is an aspect of the expected quality. Points in the last three home games was used as a proxy for expected success. The higher the recent success, i.e. the higher the recent points then the higher would be the crowd, thus the expected sign of this variable is positive. The source of this data was Rollin (1994).

2.4.2. Loyalty.

An important determinant of audience is the historical or base level of support a team enjoys. Football clubs engender extreme loyalty and some supporters will watch a game involving their club regardless of the opposition, competition or uncertainty of outcome. Supporters develop and value their relationship with their club. Dobson and Goddard (1993) find that this base level of support is important and is dependent to a large extent on history and tradition. They found that the earlier a club joined the league the larger their base support. They suggest that older clubs have built up loyalties which have been passed on through the generations. Unlike Dobson and Goddard no attempt is made to analyse why this base level of support is different between clubs only to measure it. The proxy variable used to capture fan loyalty is average home attendance over the previous three seasons. The higher the loyalty variable (i.e. average attendance in the last three years) the higher would be the expected current attendance, thus the expected sign is positive. This variable can also be interpreted as a pre-sample average dependent variable.

2.4.3. Opportunity Cost Of Watching Football

2.4.3.1. *Weather*

Weather could affect the attractiveness of both attendance at a game and viewing of a live television game. The two weather factors concentrated upon were temperature and rainfall. In the television equation, if the weather was good the opportunity cost of staying in and watching football would be increased given the added benefit of alternative outdoor activities. The opposite is true for the attendance equation, when the weather is bad the benefit from an indoors activity is increased.

Average daily temperature and daily rainfall in millimetres were obtained from the Meteorological Office for the eleven different weather stations that were closest to the twenty-two Premier League grounds[26]. An additional dummy variable was created which was set to 1 if there was any rainfall and set to 0 otherwise. In the television equation the weather station used was the London weather centre.

[26]The leven weather stations were: Liverpool; London; Manchester; Newcastle; Norwich; Preston; Sheffield; Southampton; Birmingham; Leeds; and Bristol.

2.4.3.2. *Televised Game*

A dummy variable was created to test whether a game being televised affects match attendance. For some supporters a televised match creates a cheaper substitute for match attendance. The dummy was set to 1 when the game was being televised and 0 otherwise. This variable is only appropriate for the match attendance equation and the expected sign is negative.

2.4.3.3. *Day Of Game*

Different days will present different leisure activity opportunities. For example, a game on Saturday would be easier to attend (a lower opportunity cost) for most people than a midweek game because of time constraints possibly due to work or other commitments. In the attendance equation a dummy was set up to distinguish between games being played at the weekend or a bank holiday and games being played mid-week. The dummy is set equal to 1 for midweek games and 0 for weekends or bank holidays and thus the expected sign is negative.

2.4.4. Outcome Uncertainty

It is postulated that outcome uncertainty is an attractive characteristic for supporters, thus it is thought that supporters will be more likely to attend a game if outcome uncertainty is high.

2.4.4.1. *Match Uncertainty*

Betting fixed odds were used as a measure of match outcome uncertainty. Table 2.2 gives the odds for the Premier League matches on the 11 September 1993[27]. The first, third and fifth columns give the odds for a home win, draw and away win. The odds indicate the payout you would receive if you bet and won. For example, if you bet £1 on Liverpool to beat Blackburn, which they did, then you would win £0.91 and have your £1 stake returned[28].

Chapter 3 contains a study of efficiency in the football fixed odds betting market. Despite its finding that the odds quoted by bookmakers were not always efficient, fixed odds represent an improvement on previous measures of outcome uncertainty.

The odds set are constrained by the underlying probabilities they represent and as such the measure used to proxy outcome uncertainty was the maximum odds minus the minimum odds (column six in the table). The lowest maximum-minimum figure gives the match with the highest outcome uncertainty, i.e. the outcomes are equally likely.

---

[27] Source Ladbrokes.

[28] For non televised games in the 1993/94 season the minimum bet was a treble.

On this day the most outcome uncertain match was Sheffield United versus Spurs with the odds of a home and away win being equal. The least outcome uncertain match was Arsenal versus Ipswich, with Arsenal expected to win. The more outcome uncertain a game (i.e. the smaller this measure) then the bigger the expected crowd, thus the expected sign of this variable is negative.

As already mentioned, Peel & Thomas (1988) mistakenly used odds of a home win to measure outcome uncertainty and found this to be significant. As a result of their significant finding, odds of a home win will be included as an alternative to the maximum - minimum measure described above. This will not be testing for outcome uncertainty but the expected success of the home team (as discussed above). The odds of a home win as a measure of expected success is more sophisticated than the points variable described above, as it will take into account the quality of the opposition both in the upcoming game and the previous games (i.e. will weight previous results) and will also allow for any other factors such as absence of star players.

Table 2.2: Ladbrokes Fixed Odds 11/9/93

Home Win Odds	Home Team	Draw Odds	Away Team	Away Win Odds	Max-Min
8/13	Arsenal	12/5	Ipswich	4/1	3.38
8/11	Aston Villa	9/4	Coventry	10/3	2.61
9/4	Chelsea	9/4	Man Utd	Ev	1.25
10/11	Liverpool	9/5	Blackburn	10/3	2.42
6/5	Man City	12/5	QPR	7/4	1.20
Ev.	Newcastle	9/4	Sheff Wed	9/4	1.25
4/5	Norwich	12/5	Wimbledon	11/4	1.95
5/4	Oldham	9/4	Everton	7/4	1.00
6/4	Sheff Utd	11/5	Spurs	6/4	0.70
5/4	Southampton	9/4	Leeds	7/4	1.00
5/6	West Ham	5/2	Swindon	5/2	1.67

2.4.4.2. *Seasonal Uncertainty*

Unlike Jennett (1984) and Cairns (1987) the measure of championship importance used does not require subjective judgements to be made nor use information unavailable to the spectator prior to the end of the season. There are three suggested measures of championship significance, all of which are a function of games left to be played and the number of points behind the leader. An analogous measure of relegation significance is calculated with the number of points behind the leader being replaced by the number of points ahead of the relegation zone (in the 93/94 season the bottom three places). It is to be investigated whether spectators will be attracted to games of relegation importance.

The basic idea is that the fewer games left and the fewer points behind the leader you are, the more championship significant is the game. The three measures differ only in the relative importance they attach to games left and points behind.

The three measures are

(a) $(PB/GL)^{3/4} \times G$

(b) $PB \times G$

(c) $(PB/3GL) \times (Gl^2 + 2GL)$

where PB= Points behind the leader and GL=games left.

If the team is the leader or is in the relegation zone, points behind (ahead) is assumed to be 1. When the team cannot mathematically win the championship the measure is 0. Only after 20 games have been played (out of a season of 42) does the measure become positive. The points behind the leader are adjusted for any difference in games played between the team and the leader. The championship importance of both home and away teams and the match (i.e. sum of home and away) is calculated.

Table 2.3 shows these measures for a number of different games left and points behind scenarios. A team gets 3 points for a win, 1 for a draw and 0 for a loss. Equation a puts more emphasis on the number of games left and less on the number of points behind than does equation c, equation b represents the middle ground. For example, the championship significance for a team one point behind with 20 games to be played is 42 times less than if there is only one game to go with equation a, with equation b it is 20 times and with equation c 7 times. With 20 games to be played the championship significance of a team 1 point behind is 21 times more than a team 60 points behind in equation a, 60 with equation b and 63 with equation c.

The more championship important (relegation important) a game, the smaller will be each of the three measures outlined but the bigger will be the expected crowd, thus the expected sign of both the championship and relegation importance variables is negative.

Table 2.3: Seasonal Outcome Uncertainty Measures

Equation a: $(PB/GL)^{3/4} * GL^2$

		Games Left			
		20	**10**	**5**	**1**
	60	912	0	0	0
Points	**30**	542	228	0	0
Behind	**10**	238	100	42	0
	3	96	41	17	2
	1	42	18	7	1

Equation b: $GL*PB$

		Games Left			
		20	**10**	**5**	**1**
	60	1200	0	0	0
Points	**30**	600	300	0	0
Behind	**10**	200	100	50	0
	3	60	30	15	3
	1	20	10	5	1

Equation c: $(PB/3GL) * GL^2+2GL$

		Games Left			
		20	**10**	**5**	**1**
	60	440	0	0	0
Points	**30**	220	120	0	0
Behind	**10**	73	40	23	0
	3	22	12	7	3
	1	7	4	2	1

2.5. Estimation And Results

There are several new features in this study. In particular, television audiences for live football have not been investigated before and thus the determinants of match attendance and television viewing have not been compared and contrasted. Three new variables have been introduced: the number of internationals as a measure of quality; a new measure of match uncertainty and a new measure of seasonal uncertainty. In addition, little attention has been previously paid to choice of econometric techniques employed. Most studies have employed ordinary least squares, which does not address the problem of match attendance capacity constraints. Observed match attendance data is censored and therefore this study utilises the Tobit model.

2.5.1. Match Attendance Equation

2.5.1.1. Ordinary Least Squares (OLS)

Table 2.4 gives the results of the OLS equations using in effect a pooled cross-section data set (i.e. 22 clubs each with 21 home games). Columns a and b give the results for the regression without the pre-sample average dependent variable (i.e. the loyalty variable which is average attendance over the last three years). The results indicate a positive autocorrelation problem with a Durbin Watson statistic of 0.702 and an estimated p for the AR(1) process of 0.649. This would indicate, along with knowledge of the data structure that there was some home club specific effects that were not being picked up. The inclusion of the pre-sample average dependent variable would be expected to remove at least some of this unobserved heterogeneity. Looking at the results of this equation (columns c and d) it can be seen that the problem has been reduced but not removed. The estimated p has fallen to 0.390 but the Durbin Watson statistic still indicates that positive autocorrelation exists. The length of the pre-sample average (i.e. three years) was chosen arbitrarily. Several different lengths of average pre-sample dependent variables were tried, but they made little difference to the autocorrelation or coefficient estimates.

The coefficient estimates in column c, while biased are still of interest. All coefficients are of the expected sign except for the television dummy and the relegation importance variable. All the estimates are significant, except for odds of a home win odds, rain dummy and television dummy.

In all the match attendance regressions it was found that the maximum - minimum odds coefficient which was intended to measure match outcome uncertainty was not always significant and sign was inconsistent. As an result the odds on a home win was included instead as a measure of likely success. This measure worked better than the maximum minus minimum odds variable and thus only the results for the home win odds are presented. The most appropriate variable to measure the effect of rain was found to be the rain dummy and not the rain in millimetres. The temperature variable did not perform at all well in any regression and on reflection is unlikely to be an important determinant of attendances and is excluded from the results presented.

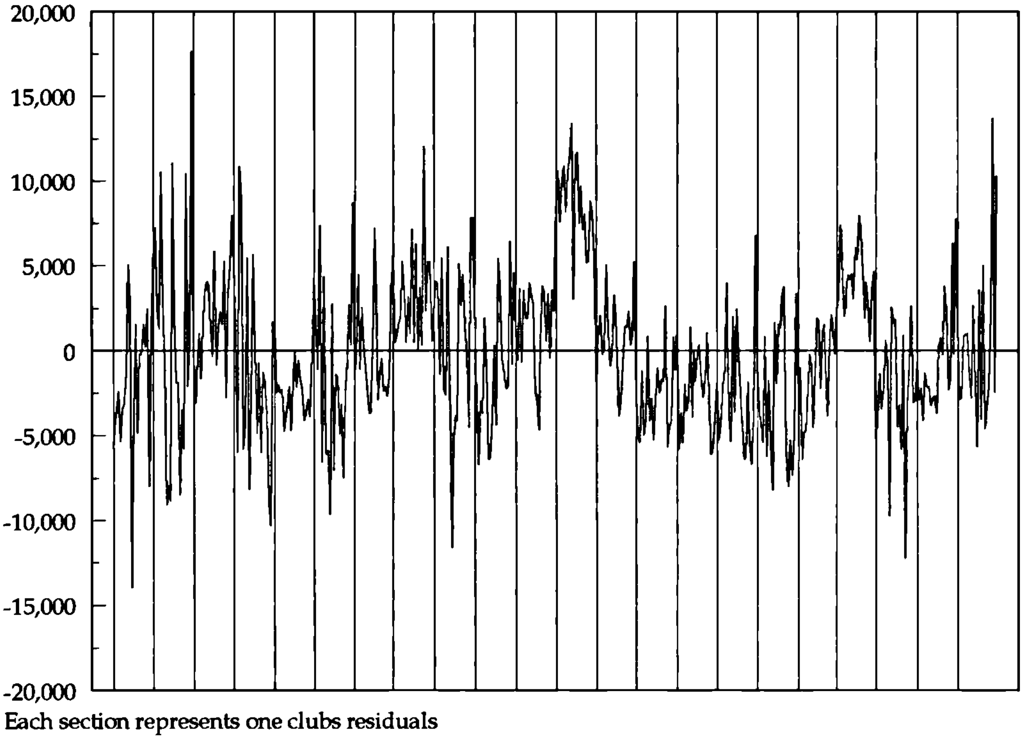
Table 2.4: OLS Estimation Of The Match Attendance Equation

Dependent Variable = Match Attendance Season 1993/1994 (Source Rollin 1994)
462 Observations

	No pre-sample average		With pre-sample average	
	a	b	c	d
	Coefficient	t-stat	Coefficient	t-stat
Number Of Internationals (Away)	692.63	3.11	583.70	5.89
Goals (scored for sum)	1107.10	6.57	170.97	2.19
Games Of Interest	6441.40	2.86	3285.70	3.27
Points	1466.20	6.50	219.21	2.10
Relegation Importance (Eqn B)	9.69	2.42	5.25	2.93
Championship Importance (Eqn B)	-2.37	-0.93	-2.73	-2.41
Odds of Home Win	879.06	1.02	-546.89	-1.48
Rain (Y/N)	3750.10	3.81	-396.43	-0.88
Television	-355.16	-0.24	130.72	0.20
Day Of Game	99.79	0.09	-1253.20	-2.51
Average Attendance Home	**	**	0.93	42.77
Adjusted R squared	-0.0914		0.7837	
Durbin Watson Stat	0.702 (3.30)	dL=1.571; dU=1.779	1.220 (2.78)	dL=1.561; dU=1.791
p AR(1)	0.649		0.390	

The presence of a residual problem is confirmed by looking at Figure 2.4 which plots the residuals from the second equation including the pre-sample average dependent variable. Each section represents one club's home games. The teams are currently arranged in alphabetical order and reordering the teams would affect the type and/or presence of autocorrelation. Thus the problem of cross sectional correlation of the residuals is more likely to be as a result of groupwise heteroscedasticity and/or groupwise autocorrelation rather than simple autocorrelation.

Figure 2.4 : OLS Residuals



2.5.1.2. *Groupwise Heteroscedasticity and Autocorrelation.*

The chosen approach to the residual problem was to use models of groupwise heteroscedasticity, correlation and autocorrelation. An alternative approach using first differences was also tried but these results were unsuccessful, most likely due to the low variation over time in several independent variables. It is worth noting that using groupwise heteroscedastic and autocorrelated models will not necessarily remedy the residual problem if it is caused by correlation between fixed club effects and the independent variables. The results described below however, support the approach taken.

As was done in the OLS estimation, stacking the data into a pooled regression model we get:

$$y = X\beta + \varepsilon \quad \text{....(2.1)}$$

The groupwise heteroscedastic model implies:

$$V = E[\varepsilon\varepsilon'] = \begin{bmatrix} \sigma_1^2 I & 0 & \dots & 0 \\ 0 & \sigma_2^2 I & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & \sigma_n^2 I \end{bmatrix} \quad \text{.........(2.2)}$$

The method of estimation used is two step feasible generalised least squares[29]. Two statistics are computed to test homoscedasticity:

$$\text{Wald} = (T/2)\sum_i \left[s^2 / s_{ii} - 1\right]^2 \quad \text{...(2.3)}$$

$$\text{LM} = (T/2)\sum_i \left[s_{ii} / s^2 - 1\right]^2 \quad \text{.......(2.4)}$$

where s^2 is the pooled OLS residual variance[30]. Both these statistics have a limiting chi-squared distribution with N-1 degrees of freedom under the hypothesis of homoscedasticity.

Extension of this model to allow correlation of the disturbances across clubs was also tried. That is:

$$E[\varepsilon_i \varepsilon_j'] = \sigma_{ij} I \quad \text{.......(2.5)}$$

Assuming that observations are uncorrelated across time, you obtain:

$$V = \begin{bmatrix} \sigma_{11} I & \sigma_{12} I & \dots & \sigma_{1n} I \\ \sigma_{21} I & \sigma_{21} I & \dots & \sigma_{2n} I \\ & \vdots & & \\ \sigma_{n1} I & \sigma_{n2} I & \dots & \sigma_{nn} I \end{bmatrix} \quad \text{..........(2.6)}$$

Again the appropriate method of estimation is feasible generalised least squares. To test the appropriateness of groupwise heteroscedasticity versus groupwise heteroscedasticity with cross group correlation, the following statistics were computed:

$$\text{LM} = T\sum\sum_{j,i} \left[s_{ij}^2 / s_{ii}s_{jj}\right]^2 \quad \text{....(2.7)}$$

[29] See Greene (1990) Chapter 16 for more detail.
[30] See Greene (1992) Chapter 28 for more detail.

$$LR = T\left(\sum_i s_{ii} - \ln|S|\right) \quad \text{...........(2.8)}$$

A further extension of the model is also possible, relaxing the assumption of non-autocorrelation. Suppose that for each i,

$$E[\varepsilon_i \varepsilon_i'] = \sigma_i^2 \Omega_i \quad \text{...... ....(2.9)}$$

with the earlier assumption of groupwise heteroscedasticity we get:

$$V = E[\varepsilon\varepsilon'] = \begin{bmatrix} \sigma_1^2\Omega_1 & 0 & \dots & 0 \\ 0 & \sigma_2^2\Omega_2 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & \sigma_n^2\Omega_n \end{bmatrix} \quad \text{.......(2.10)}$$

In order to estimate a model with both groupwise heteroscedasticity and groupwise autocorrelation, several steps are required. First the residuals from the ordinary pooled OLS regression are used to estimate p. Using this estimated p the data is transformed using the Cochrane-Orcutt transformation. A second OLS equation is then computed using the data which has been transformed to remove autocorrelation. The residual sums of squares and cross products are then used to estimate $S=s_{ij}$, which is used in the final GLS estimation.

The significance of the autocorrelation coefficients can be tested by using:

$$(T-1)r_i^2 / \left(1-r_i^2\right) \approx \chi^2[1] \quad \text{.......(2.11)}$$

Table 2.5 gives the results for the three different regression models described above. The tests for homoscedasticity in model 1 indicate strongly that groupwise heteroscedasticity is present (i.e. Wald statistic of 348 and a LR statistic of 126 with a critical value of 32.67). The extension to allow correlated disturbances across clubs is also strongly supported by the LM test in model 1 and the LR test in model 2 (541 and 1258 against a critical value of 267.17). However the final extension to allow groupwise autocorrelation is not supported as the correlation coefficient of 0.35 is not significant with a test statistic of 2.89 being below the critical value of 3.84.

The test statistics support model 2 with groupwise heteroscedasticity and cross group correlation. In this equation all the estimated coefficients are the expected sign apart from the television dummy and the relegation significance variable. All the coefficients are significant except for goals scored for, the rain dummy and television dummy.

Table 2.5: Groupwise Heteroscedasticity And Autocorrelation

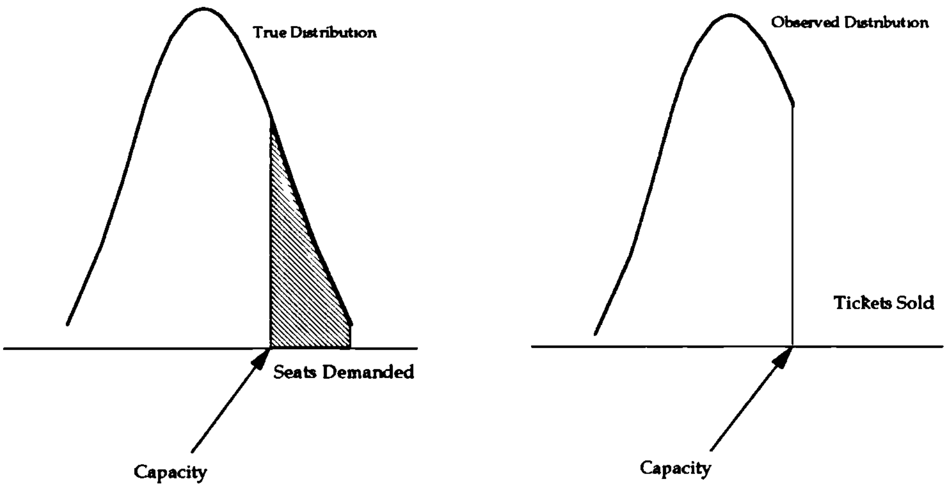
Dependent Variable = Match Attendance Season 1993/1994 (Source Rollin 1994)
462 Observations

	Groupwise Heteroscedasticity		Groupwise Het and Correlated		Groupwise Het. and Correlated with Autocorrelation	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Number Of Internationals (Away)	464.34	6.02	465.93	6.26	414.95	6.59
Goals scored (match)	68.12	1.06	49.57	0.80	90.72	1.59
Games Of Interest	3005.40	3.93	2925.50	3.94	2316.50	3.75
Points	275.02	3.40	273.59	3.45	123.71	1.47
Relegation Importance (Eqn B)	4.57	3.60	4.37	3.60	2.58	1.93
Championship Importance (Eqn B)	-2.11	-2.36	-2.00	-2.35	-1.16	-1.23
Odds of Home Win	-408.63	-1.36	-499.64	-1.72	30.06	0.11
Rain (Y/N)	-286.52	-0.80	-309.06	-0.89	-463.09	-1.61
Television	74.36	0.15	138.69	0.29	-274.37	-0.70
Day Of Game	-999.50	-2.52	-978.38	-2.57	-371.01	-1.22
Average Attendance Home	0.95	53.82	0.96	56.54	0.97	47.62
Test For Homoscedasticity	w=348.29 lr=125.51	**	**	**	**	**
Test Against Correlation	lm=540.84	**	LR=1257.83	**	LR=7562	**
Autocorrelation Coefficient	**	**	**	**	0.36	**
Test for Autocorrelation	**	**	**	**	2.89 (3.84)	**

2.5.1.3. *Tobit Estimation.*

One aspect of the data that has not been considered so far, is that in over ten per cent of matches in the 1993/94 season the ground was at capacity. Thus the observed distribution of attendances is a censored distribution. Figure 2.5 below illustrates this point.

Figure 2.5: Partially Censored Distribution



The appropriate regression model for this type of distribution model is the Tobit model which is based on the classical regression model.

$$y_i^* = \beta' x_i + \varepsilon_i \quad \text{.......(2.12)}$$

However, in this case y* is not directly observed if the ground is at capacity, i.e.

$$y_i = c \qquad if \quad y_i^* \geq c$$
$$y_i = y_i^* \qquad if \quad y_i^* < c \quad \text{.......(2.13)}$$

where c=capacity, y*= latent or index variable and y=observed variable.

Amemiya (1973) showed that the most appropriate method of estimation of this model is maximum likelihood estimation (MLE) as this produces all the normal desirable properties associated with MLE. The log likelihood for this upper censored regression is:

$$\ln L = \Sigma_{y_i < c} - \frac{1}{2}\left[\ln(2\pi) + \ln \sigma^2 + \frac{(y_i - \beta' x_i)^2}{\sigma^2}\right] + \Sigma_{y_i \;\, c} \ln\left[1 - \Phi\left(\frac{c - \beta' x_i}{\sigma}\right)\right]$$

.....(2.14)

Within the Premier League there were several different capacities so that the censoring point varies over the sample. This affects the log likelihood only in that the appropriate capacity should be used in calculating the second term. Figure 2.6 below shows the number of games played at each capacity. The variation in capacities is due in part to the fact that during the season several stadia underwent improvement work in response to the Taylor Report[31], causing their capacities to vary over the season.

Figure 2.6 : Capacity Distribution

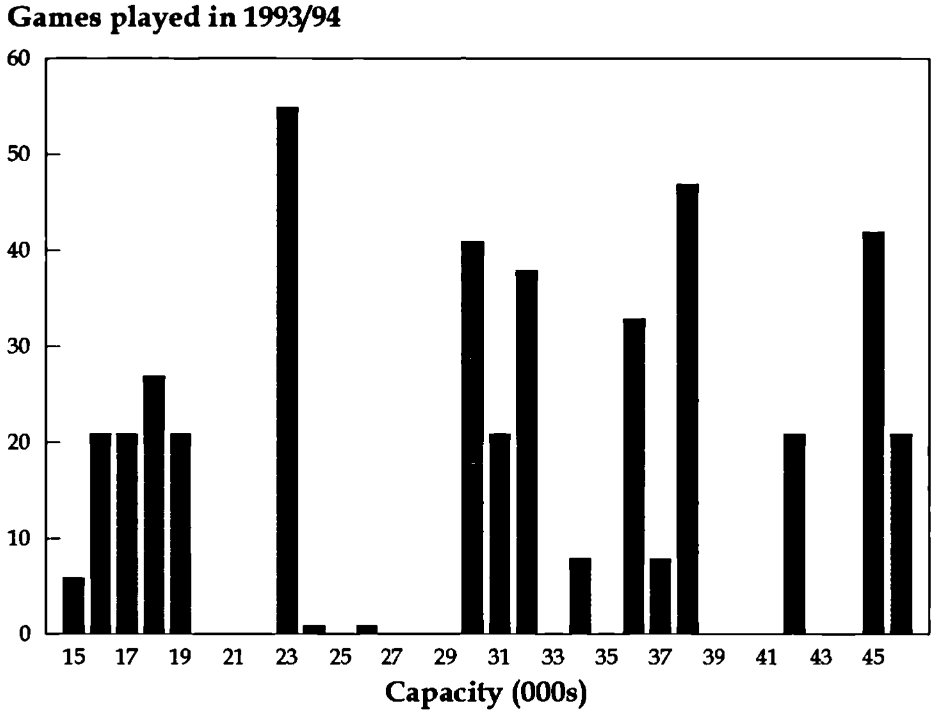


Table 2.6 shows the Tobit regression for match attendance along with tests for heteroscedasticity and normality. As with the groupwise heteroscedastic equation all the coefficients are the expected sign apart from the television dummy and relegation significance variable. All estimates are significant except for the television and rain dummies.

The tests for heteroscedasticity and normality are conditional moment tests as set out by Pagan and Vella (1989)[32]. The heteroscedasticity test is based on the conditional moment condition:

$$y^* = \beta' x + \varepsilon$$

$$E\left[E\left(z\varepsilon^2 \middle| y\right) - \sigma^2\right] = 0 \quad \text{......(2.15)}$$

In this test, z includes all the coefficients contained in the equation. Given the results of the previous regressions, heteroscedasticity would be expected to be present. Indeed, the test indicates strongly that heteroscedasticity exists.

[31] Report produced by Lord Justice Taylor which recommended all seater stadia in response to the Hillsborough disaster of 1989.

[32] Also see Greene (1993) pp. 701-706.

The test for normality is based on the conditional moment restrictions:

$$E\left[E\left(\varepsilon^3 | y\right)\right] = 0 \text{; and}$$

$$E\left[E\left(\varepsilon^4 | y\right) - 3\sigma^4\right] = 0 \quad \text{......(2.16)}$$

The test rejects the null hypothesis of normality. It would be possible to assume an alternative distribution but it is not clear that this would improve the estimates and may even make them worse.

Table 2.6: Tobit Estimation

Dependent Variable = Match Attendance Season 1993/1994 (Source Rollin 1994)
462 Observations

	Coefficient	Standard Error
Number Of Internationals (Away)	583.87	99.09
Goals (scored for sum)	170.97	78.17
Games Of Interest	3285.70	1005.00
Points	219.21	104.50
Relegation Importance (Eqn B)	5.25	1.79
Championship Importance (Eqn B)	-2.73	1.13
Odds of Home Win	-546.89	385.70
Rain (Y/N)	-396.43	448.90
Television	130.72	659.70
Day Of Game	-1253.20	499.3
Average Attendance Home	0.93	0.02
Heteroscedasticity	LM=36.42	
Normality	LM=98.38	

In order to combat the heteroscedasticity, a further model is estimated.

$$e_i = y_i - \beta x_{i,}$$

$$\theta_i = \sigma e^{\gamma' z_i}, \quad \text{...(2.17)}$$

Then the log likelihood becomes:

$$\ln L = \Sigma_{y_i < c} - \frac{1}{2}\left[\ln(2\pi) + \ln \theta_i^2 + \frac{(y_i - \beta' x_i)^2}{\theta_i^2}\right] + \Sigma_{y_i\ c} \ln\left[1 - \Phi\left(\frac{c - \beta' x_i}{\theta_i}\right)\right] \quad \text{....(2.18)}$$

The results of this model are given in Table 2.7. Again all the variables in the equation are included in the z vector

Table 2.7: Tobit Model With Heteroscedasticity

Dependent Variable = Match Attendance Season 1993/1994 (Source Rollin 1994)
462 Observations

	Coefficients	Standard Error
Number Of Internationals (Away)	652.84	127.20
Goals (scored for sum)	221.00	95.06
Games Of Interest	3287.00	1047.00
Points	201.79	129.50
Relegation Importance (Eqn B)	5.60	2.20
Championship Importance (Eqn B)	-3.20	1.14
Odds of Home Win	-525.54	382.60
Rain (Y/N)	-398.57	548.70
television	130.68	768.50
Day Of Game	-1253.20	667.00
Average Attendance Home	0.93	0.03
Heteroscedasticity Terms:		
Number Of Internationals (Away)	0.04	0.02
Goals (scored for sum)	0.03	0.02
Games Of Interest	-0.12	0.21
Points	-0.03	0.02
Relegation Importance (Eqn B)	-0.27 E-03	0.37 E-03
Championship Importance (Eqn B)	-0.11 E-03	0.26 E-03
Odds of Home Win	-0.30	0.13
Rain (Y/N)	-0.02	0.10
Television	-0.07	0.11
Day Of Game	0.17	0.01
Average Attendance Home	0.93 E-05	0.61 E-05

All the coefficients in the heteroscedastic and the homoscedastic models are of the same sign. The size of the coefficients and their standard errors have not changed drastically for most of the coefficients except for odds of a home win. The coefficient on odds of a home win increased from -942 with homoscedasticity to -525 with heteroscedasticity whilst also becoming insignificant.

When looking at the results of the Tobit model, care must be taken in interpreting the marginal results.

The marginal effects for the latent variable, y* are:

$$\frac{\delta E[y_i^* \; x_i]}{\delta x_i} = \beta \qquad ......(2.19)$$

Looking at the marginal effects of the latent variable will (given a capacity) overestimate the effects on attendance of any particular variable. The latent variable marginal effects can be appropriate if a club is planning ground developments and thus the capacity could be changed.

The marginal effects for y given the censoring are:

$$\frac{\delta E[y_i | x_i]}{\delta x_i} = \text{Prob [non limit]}* \beta \text{.....(2.20)}$$

In this case, with an upper censor that is variable, the marginal effects were computed as follows:

$$\frac{\delta E[y_i | x_i]}{\delta x_i} = \left[\Sigma_s \left(\Phi\left(\frac{c_s - \beta' x_i}{\sigma} \right) * \left(\frac{MPC_s}{MP} \right) \right) \right] * \beta \text{......(2.21)}$$

where MPC_s= matches played with capacity C
MP= Total matches played=462
S= number of different capacities.

The marginal effects given the censoring are approximately 70% of the latent variable marginal effects. For example the marginal effect of international in the model with heteroscedasticity (Table 2.7) is 0.7*652.84 =456.99

Table 2.8 compares the two favoured regressions with the OLS estimates. Column b gives the estimated coefficients for the groupwise heteroscedasticity model allowing for cross group correlation and column c gives the coefficients from the heteroscedastic Tobit model.

Table 2.8: Summary Table

Dependent Variable = Match Attendance Season 1993/1994 (Source Rollin 1994)
462 Observations

	a OLS	b Groupwise Het With Correlation	c Tobit With Het.
Number Of Internationals (Away)	583.70	465.93	652.84
Goals scored (match)	170.97	49.57	221.00
Games Of Interest	3285.70	2925.50	3287.00
Points	219.21	273.59	201.79
Relegation Importance (Eqn. b)	5.25	4.37	5.60
Championship Importance (Eqn. b)	-2.73	-2.00	-3.20
Odds of Home Win	-546.89	-499.64	-525.54
Rain (Y/N)	-396.43	-309.06	-398.57
Television	130.72	138.69	130.68
Day Of Game	-1253.20	-978.38	-1253.20
Average Attendance Home	0.93	0.96	0.93

In all of the three regression the sign on the estimated coefficients are identical. All the signs are as expected except the relegation importance variable and the television dummy variable.

In all the regressions the rain variable and the television variable were not significant. In addition in the heteroscedastic Tobit and OLS estimates the odds of a home win is not significant and in the groupwise heteroscedastic equation match goals scored for is not significant. All other coefficients were significant.

The results confirm that quality is an important determinant of attendance. Both the estimated coefficients for the game of interest variable and the number of international players are significant and positive. Importantly it is the away number of internationals that is significant. This makes intuitive sense as the number of internationals in a team is reasonably constant over a season and the quality of the home team is probably picked up in the loyalty variable which was significant and positive (home team average attendance over the last three years).

The points variable is in effect measuring recent form: the more successful the team recently the more people will want to come and support their team. As expected this is positive and significant. The odds of a home win variable is also significant in the groupwise heteroscedastic model. The failure of this variable (although it is the correct sign) in the other two regressions is possibly due to the partial correlation of the points and odds of home win variables as they are measuring expected success.

The opportunity cost variables did not perform as well. The rain variable was insignificant but the correct sign in all equations, suggesting that the weather is not an important determinant of attendance. Given that nearly all Premier League grounds are completely under cover this result is perhaps not too surprising.

More surprisingly in all the equations the television variable was the incorrect sign and insignificant. This suggests that televising a game live on Sky Sports does not have, as is widely argued, a negative effect on attendances. This is not to say that television does not have negative affects on attendance but just that the current Sky televising arrangement has no effect. Several factors could explain this phenomenon. Although Sky Sports had over 7million subscribers at the end of the 93/94 season, this is still a very small percentage of the viewing public. Thus the televising of a game on Sky is likely to have a smaller impact than televising a game on terrestrial television. In addition, Sky will pick the most attractive games for broadcast, for example consider the number of appearances by the championship contenders (see Figure 2.2). The effect of televising games is thought to be diminished the higher the attractiveness of the game. In high quality games, often television will not be a substitute but a complement for attendance, people will attend the game and video the game to watch later.

The day of the game variable was significant and negative indicating that if a game is played mid-week then this will have a negative impact on attendances. This makes sense as you would expect people to find it easier to attend matches at the weekend. This however could be used to argue that televising games does have an indirect effect on attendances. Sky shows almost half of its games live on a Monday evening, thus ensuring a negative effect on attendances via the day of game variable.

The outcome uncertainty variables' performance was mixed. The match uncertainty variable (maximum-minimum odds) was not significant and the sign varied. As an experiment the odds of a home win replaced match uncertainty and was the expected sign but not always significant. The smaller the odds (the bigger the chance of a home victory) the bigger the attendance, suggesting that supporters value success more than match uncertainty of outcome.

The championship importance variable was significant and negative as expected in all equations. This indicates that the more important a game in deciding the championship the bigger the audience. The relegation importance variable was also significant but the sign indicated that the more relegation important a game the smaller the crowd. The significance of the relegation variable could be due to it acting as a success and quality proxy. That is the more relegation significant the game the less successful the teams involved are likely to be and the lower the quality of the game. For example even if a team had several international players but was involved in a relegation important match, this would indicate that the players and team were not doing very well and as such the expected quality of the game would be low.

2.5.2. Television Equation.

Table 2.9 contains the television regression. Initially ordinary least squares was used and there was no reason to reject this method as the tests for autocorrelation and heteroscedasticity show. A log linear model was tested but rejected.

Two quality variables were significant. The number of internationals is significant and the expected sign, suggesting the more international players on the pitch (i.e. the higher the quality) the bigger the television audience. As with most of variables in the television equation it is the match variables (i.e. sum of home and away) that are significant. This is plausible because the equation is estimating nation-wide television audiences and the team specific factors (i.e. home or away) will be less important. Similar reasoning explains why the game of interest variable is insignificant. The majority of the games of interest were derbies and as such the extra interest in these games will often only concern the supporters of the two clubs. Generally the passion of local derbies can override the quality of the game and are not necessarily of extra interest to non supporters of the two clubs involved. Of all the goal variables, only the goals scored against the away team was significant. This is surprising but the sign suggests that the greater the number of goal scored against the away team recently the bigger the television audience. This variable could be picking up the effects of teams that have very negative (i.e. very defensive) away strategies such as Arsenal or Ipswich. That is the away team has a defensive strategy aimed at conceding few goals but making the game unattractive.

Again the loyalty variable was significant and the expected (average attendance over the last three years). Team dummy variables were added to test if they could improve the equation. The Newcastle dummy was significant and the sign was positive. This is entirely plausible given that Newcastle have been on an incredibly quick ascent through the divisions creating exceptional interest. After narrowly missing relegation to the Second Division in 1991/92, they won the First Division in 1992/93 and came third in the Premiership in 1993/94. Currently Newcastle's ground is sold out every week and as mentioned earlier all their games are shown live at local cinemas and clubs. It is a clear that average attendance over the last three seasons would not adequately reflect the current interest in Newcastle.

Table 2.9 : Television Audience Regression

Dependent Variable= Proportion of Sky Sports subscribers watching live football matches 1993/94.
60 Observations Ordinary Least Squares.

	Coefficient	t -ratio
Match championship significance (equation A)	-0.3130 E-04	-2.15
Match relegation significance (equation A)	0.4029 E-04	2.79
Match Number of Internationals	0.2978 E-02	2.95
Match Average Attendance over last three years	0.1382 E-05	6.19
Odds min - max	0.1859 E-2	0.76
Rain Y/N	-0.3901 E-02	-0.59
Dummy for Newcastle United	0.3294 E-01	3.45
Game of Interest	-0.3687 E-02	-0.35
Goals Scored Away Against in last three matches.	0.4588 E-02	3.193
Adjusted R^2	0.61	
Est Autocorrelation AR(1)	-0.099	
Heteroscedasticity (Goldfeld Quand Test)	1.59 F_{95} (17,20)=2.172	
DW Statistic	2.19 +ve; 1.81 -ve (dl=1.21;du=1.64)	

As with the match attendance equation the rain variable was insignificant but the expected sign.

Again the outcome uncertainty variables performance was mixed. Match uncertainty, as proxied by betting odds was not significant and indeed was the opposite sign to that which was suggested in the earlier discussion. Again, home odds were added to the regression, but were found to be insignificant, probably due to home success only being important to home supporters. The championship variable was significant and of the expected sign (i.e. negative) indicating the more championship important the game the bigger the television audience. The relegation importance variable was significant but again indicated the more relegation significant a game the smaller the audience. As with the attendance equation, it is possible that the relegation importance variable is a proxy for success and quality. An alternative explanation is that not a large enough number of relegation battles were televised.

2.6. Conclusions

The importance of seasonal uncertainty of outcome was confirmed using a measure that utilises only information available prior to the match and does not require arbitrary assumptions about future success. A refinement of Peel & Thomas' (1988) measurement of match uncertainty is presented and it is found that the probability of success of the home team is more important than match uncertainty. The introduction of internationals as a measure of quality was successful and the use of the Tobit model to account for capacity constraints was justified.

Similar factors are important determinants of both match attendance and television audience. Championship importance, number of internationals, loyalty and relegation importance are all significant in both equations. The main difference between the equations is that the match attendance equation includes home variables while the television audience equation includes match variables.

In the match attendance equation the day of the game was important as were games of interest and home success rates (i.e. recent points). This illustrates the internal contradiction of a league system. It is in every club's interest to succeed and increase its gates receipts and prize money but it is also true that outcome uncertainty requires a strong league without domination by any one club. Measures that ensure lack of concentration of playing talent, such as cross subsidisation or transfer regulations, can thus be defended in a league profit maximising framework. The significance of the day of game variable would indicate that the league should try and schedule as many weekend fixtures as possible.

The results of the television equation have impact for television schedulers and contract negotiators. In contract negotiations, television companies should maximise their freedom to show any game, allowing flexibility in scheduling as the season progresses. Some contracts enforce a minimum number of times each club must appear. It is in the interests of the television companies to show games that have an impact on the championship race, include clubs with a large base support and large number of internationals. It will only be possible to detect championship significant games as the season progresses. The insignificance of the game of interest dummy shows that local derbies, despite generating a large local interest are not necessarily the best games to televise. Another interesting result for the television companies is the fact that televising a game had no significant impact on match attendance. In previous television contracts there was always a clause to reimburse the home team in televised games for loss of gate revenue.

Seasonal outcome uncertainty is shown to be an important determinant of attendance and this suggests that cross subsidisation between clubs or transfer restrictions can be defended with a health of league argument. However, the historical reasons for club support including local team rivalries should not be underestimated nor undermined by any rule changes. Indeed, in planning the European super league it will be important to assess whether the absence of these historical factors and local rivalries will impinge on its growth.

2.7. References

Amemiya, T., (1973). Regression Analysis When the Dependent Variable is Truncated Normal. *Econometrica.* No. 41, pp. 928-934.

Arnold, T. & Benveniste, I. (1988). Cross Subsidisation and Competition Policy in English Professional Football. *Journal Of Industrial Affairs.* Volume 15, pp. 2-14.

Bird P.J.W.N. (1982). The Demand for League Football. *Applied Economics.* Vol. 14, No. 6, pp. 637-649.

Borland, J. (1987). The Demand For Australian Rules Football. *The Economic Record.* September pp. 221-230.

British Audience Research Board (1991). *Television Audience Measurement Reference Manual.* December.

British Audience Research Board. *Astra Satellite Panel Weekly TV Audience Report.* Various Issues August 1993- May 1994.

BSkyB (1994). *Sky Media Pack 1994.* BSkyB Public Relations Department.

Burkitt, B. & Cameron, S. (1992). Impact of League Restructuring on Team Sport Attendances: The Case of Rugby League. *Applied Economics.* No. 24, pp. 265-271.

Cairns, J.A., Jennett, N. & Sloane, P.J. (1986). The Economics of Professional Team Sports: A Survey Of Theory and Evidence. *Journal Of Economic Studies,* Vol. 12, pp. 3-80.

Cairns, J.A. (1987). Evaluating Changes In League Structure: The Organisation of the Scottish Football League. *Applied Economics.* No. 19(2), pp. 259-275.

Cairns, J.A. (1990). The Demand For Professional Team Sports. *British Review of Economic Issues.* October, pp. 1-20.

dell'Osso, F. & Szymanski, S. (1991). Who are the Champions? (An Analysis Of Football And Architecture). *Business Strategy Review.* Summer, pp. 113-130.

Demmert, H.G. (1973). *The Economics of Professional Team Sports.* Lexington Books, D.C. Heath and Co., Lexington Massachusetts.

Dobson, S.M. & Goddard, J.A. (1993). The Demand For Professional League Football in England and Wales. *Research Paper Series, School Of Accounting, Banking and Economics, Bangor, Prifysgol Cymru.* RP 93/37.

Football Trust (1994). *Digest Of Football Statistics 1992-1993.* London.

Greene, W.H. (1992). *LIMDEP Version 6.0. User's manual & Reference Guide.* Econometric Software, Inc. New York.

Greene, W.H. (1993). *Econometric Analysis.* Second Edition. Macmillan Publishing, United States.

Hall,T., Beilby, M. & Duckworth, R. (1994). European Pay TV - BSkyB and Beyond. *S.G.Warburg Media Research.* October

Hart, R.A., Hutton, J. & Sharot, T. (1975). A Statistical Analysis of Association Football Attendances. *Applied Statistics.* Vol. 24, No. 1, pp. 17-27.

Horrie, C. & Clarke, S. (1994). *Fuzzy Monsters: Fear and Loathing at the BBC.* Mandarin Books, London.

Hynds, M. & Smith, I. (1994). The Demand For Test Match Cricket. *Applied Economic Letters.* No. 1, pp. 103-106.

Jennett, N. (1984). Attendances, Uncertainty of Outcome and Policy in the Scottish Football League. *Scottish Journal of Political Economy.* Vol. 31 No. 2, pp. 176-198.

Jones, J.C.H. & Ferguson, D.G. (1988). Location and Survival In The National Hockey League. *The Journal Of Industrial Economics.* Volume XXXVI, No. 4, pp. 443-457.

Jones, J.C.H. The Economics of the National Hockey League. *Canadian Journal Of Economics.* Vol. 2, No. 1, pp. 1-20.

Maddala, G.S. (1983). *Limited Dependent And Qualitative Variables In Econometrics.* Econometric Society Monographs. Cambridge University Press.

McDonald, J. & Moffitt, R. (1980). The Uses Of Tobit Analysis. *The Review Of Economics And Statistics.* No. 62, pp. 757-758.

Neale, W.C. (1964). The Peculiar Economics Of Professional Sports. *Quarterly Journal Of Economics.* February, No.1 , pp.1-14.

Noll, R.G. (1974). *Government and The Sports Business.* Washington DC, Brookings Institution.

Pagan,A. & Vella, F. (1989). Diagnostic Tests for Models based on Individual data: A Survey. *Journal of Applied Econometrics.* No. 4, Supplement, pp. S29-S59.

Peel, D. & Thomas, D. (1988). Outcome Uncertainty and the Demand For Football: An Analysis of Match Attendances In The English Football League. *Scottish Journal of Political Economy.* 35(3), pp. 242-249.

Rollin, J. (ed.) (1994). *Rothmans Football Yearbook 1994-95,* Headline Book Publishing.

Schofield, J.A., (1982). The Development of First Class Cricket In England: An Economic Analysis. *The Journal Of Industrial Economics.* June, Volume XXX, No. 4, pp. 337-360.

Szymanski, S. & Smith, R. (1993). The English Football Industry: Profit Performance and Industrial Structure. *Unpublished Mimeo, London Business School.*

Walker,B. (1986). The Demand for Professional League Football and the Success Of Football League Teams: Some City Size Effects. *Urban Studies.* 23, pp. 209-219.

Whitney, J.D. (1988). Winning Games Versus Winning Championships: The Economics Of Fan Interest and Team Performance. *Economic Inquiry.* Vol.XXVI, October, pp. 703-724.

3. INFORMATION AND EFFICIENCY: AN EMPIRICAL STUDY OF A FIXED ODDS BETTING MARKET

3.1. Introduction

There has been extensive debate in the literature about the efficiency of information markets, both financial markets and betting markets. This chapter investigates a particular form of betting, the fixed odds betting market.

Several studies have rejected the efficient market hypothesis for financial markets[33]. An explanation advanced for the inefficient market finding is that agents employ information in an inefficient manner. The extreme of this is Keynes' view that financial markets are dominated by unstable psychological factors or "animal spirits".

Betting markets have similar characteristics to financial markets but have advantages for empirical investigation of efficiency. Betting markets give detailed price and outcome information with a neatly specified time between purchase of the bet and the outcome of the event.

Fixed-odds betting systems have only been modelled once before: the main focus of betting markets studies has been on pari-mutuel and starting price systems. The various betting systems and previous studies are discussed in detail in the next section. The uniqueness of fixed odds betting is that odds are set several days in advance and do not move in response to betting before the event.

Peel and Pope (1990) claim that fixed odds betting markets are similar in character to options markets. More specifically a European call option is similar in character to a fixed odds bet. A European call option[34] gives its owner the right to buy stock at a specified "exercise price" on a given day. The buyer pays a fixed price: if on the specified day the price of the share is below this price the option is worthless (in effect she has lost the bet); however if the share is worth more than the exercise price then the owner will in effect have gained the difference between the actual price and the exercise price (she has won the bet). In betting terms, the option buyer has placed a fixed odds bet that the share price will increase above a particular point (the exercise price).

The particular fixed odds market that this paper looks at is English football which has a turnover in the UK of approximately £300m and is growing. The efficiency of fixed odds betting markets is used as an assumption in various papers investigating the demand for football[35] where odds have been used as measure of outcome uncertainty. This chapter tests the validity of this assumption.

[33] See for example Figlewski & Wachtel (1981).

[34] For more details see Brealey & Myers (1991).

[35] See for example Chapter 2 and Peel & Thomas (1988).

3.2. Information And Efficiency

The definition of efficient markets adopted is set out by Fama in his much referenced Journal of Finance article[36]. Fama examines the stock market as an information market and not a service industry. Thus he defines an efficient market as a market whose prices fully reflect all available information. He distinguishes three forms of tests dependent on the information set utilised:

- *Weak* form tests using past prices only;
- *Semi-strong* tests using all publicly available information; and
- *Strong* tests using all information including any information over which certain groups have a monopoly.

In betting markets, weak form efficiency implies that no abnormal returns, either to the bookmaker or to the punter, can be achieved using just price information. Abnormal returns are defined as returns different from the bookmakers' take. For example, in the Tote pari-mutuel system the bookmakers' take is set by regulation at around 20%. If bookmakers were obtaining returns greater than 20% or punters were able to make returns better than -20% then this would imply that abnormal returns were being made.

Semi-strong efficiency implies that the incorporation of publicly available information should not improve on the accuracy of outcome predictions based on odds. Thus there should not be abnormal returns to a betting strategy utilising odds and publicly available information for either the bookmaker or the punters.

Strong efficiency implies that no group in society, as a result of private information, can make abnormal returns.

3.3. Gambling Market Studies

Three major betting systems have been investigated in the literature: pari-mutuel; spread betting; and quoted odds. In pari-mutuel betting systems, such as the Tote, winning punters share the betting pool (i.e. the total amount bet) less a fixed percentage for the bookmaker. The effective odds are determined by how much money is bet on each possible outcome.

Spread betting is most popular in the United States but certain forms are also available from specialist bookmakers in Britain. In US spread betting, the bookmaker announces a points difference between two teams in a match (the bookmakers' "line"). The public is offered even money (less the bookmakers take) on a bet that the actual spread is above (or below) the bookmakers line.

[36] Fama (1970).

In quoted odds systems, the bookmaker offers odds on each outcome (e.g. each horse in a race). The odds fluctuate up to the event in response to the amount of money that has been wagered on each outcome. This fluctuation is not determined by a fixed formulae as in the pari-mutuel systems but is a subjective decision made by the bookmaker. The punter can bet by either accepting the odds as quoted at the time she makes the bet or can take the starting price. The starting prices are the odds at which a sizeable bet could have been made on the course just before the beginning of the race[37]. A special form of quoted odds systems is fixed odds. This is where the bookmakers quotes odds on particular outcomes prior to an event but the odds do not fluctuate in response to betting patterns. It is on this type of betting that this chapter focuses.

3.3.1. Pari-Mutuel Systems

The majority of studies of the efficiency of pari-mutuel systems have used data from American horse racing. Snyder (1978) finds through weak tests that although punters display strong and stable biases, these biases are not large enough to earn positive profits. Snyder compared the rates of return for different levels of odds and found that positive rates of return were achieved on horses with odds less than 5 to 1. He explains the bias by suggesting that punters bet a smaller proportion on lower odd horses than their chances of winning justifies. This is due to a general tendency for punters to prefer low probability-high prize combinations over high probability-low prize combinations. Snyder's results are consistent with the results of a study by Ali (1977) who calculates the subjective and objective probabilities in harness races in the US and finds that smaller (or shorter) odd horses yield positive rates of return.

In a more recent study of thoroughbred racing Asch et al. (1984 & 1986) develop betting strategies in an attempt to obtain positive rates of return. They estimate a logit model with horses' win/loss record as the dependent variable and both the forecast starting price and pari-mutuel odds at different times before the race starts as independent variables. The forecast starting prices are the odds printed in the daily racing programme which are the handicapper's best estimates of the winning probabilities for each horse. They then use this model along with several betting rules to try and obtain a positive return They obtain returns better than the track take (-18.5%) but are unable to make positive returns. Their results do however indicate that inefficiencies exist.

Hausch et al. (1981) look at more "exotic" bets, that is place (first or second) and show (first, second or third) pari-mutuel betting systems. They investigate the difference between subjective probabilities and outcome probabilities. The inefficiency they identify by comparing odds with outcome probabilities is used to develop a betting system based simply on starting prices. They claim not only to achieve above average returns, as do Asch et al., but go further and claim positive returns. Their results indicate there are significant inefficiencies in the place and show pools pari-mutuel system.

[37] See Dowie (1976) p 140.

Figlewski (1979) tests whether subjective information (in this case newspaper tipsters) is efficiently incorporated into the pari-mutuel horse betting market. He tests to see whether tipsters' information can be combined with the track odds to produce significantly more accurate results forecast than the track odds alone. Figlewski finds that punters at the track appear to discount the tipsters information fully, but those betting through New York's off track betting system do not, suggesting some inefficiency in the off track market.

3.3.2. Quoted Odds Systems

Research into quoted odds systems has predominantly focused upon British horse racing. Dowie (1976) investigates the efficiency of odds from the 1973 flat season. He focuses on the differences between forecast starting prices (forecast by newspapers on the morning of the race) and starting prices. It is assumed that any superior or inside information that individuals or groups posses will be reflected in the actual starting prices. Thus, if inside information plays a significant role, then the correlation between starting prices and outcomes should be significantly higher than the correlation between forecast starting prices and outcomes. Although this is what Fama would describe as a strong test of efficiency, Dowie used the term "equity test" as the test determines whether one group in society can achieve abnormal returns. Dowie finds that the forecast starting price and starting prices are equally correlated with outcomes, suggesting that insider information does not play a significant role. Dowie's findings imply that the British flat horse race betting market is strongly efficient.

Crafts (1985) refutes Dowie's method and results. He states that Dowie's test is not enough: the difference between forecast starting prices and starting prices might be due to additional public information being reflected and thus the market could still be efficient. The real test of inefficiency is whether the starting prices are more correlated with outcomes than the forecast prices <u>and</u> there are opportunities for profitable arbitrage. Using data for the four month period from September 1982 (16,800 runners) and despite using a stronger test than Dowie, Crafts does find evidence of inefficiency. He finds that a punter can reduce his expected losses by placing a bet just before the off in certain odds categories, thus violating weak efficiency. Also, by concentrating on horses whose odds moved substantially before the race and by looking at descriptions in the racing press of large bets, Crafts concludes that there are exploitable opportunities for people with insider information. This conclusion violates the strong efficiency or Dowie's "equity" assumptions.

Johnson & Bruce (1992) investigate what they term as "Gluck's Second Law", a particular form of inefficiency in the quoted odds market. Gluck's Second Law states that the best time to bet a favourite is in the last race. Punters are believed to bet more on outsiders as the day progresses and thus there may be advantageous odds on favourites in the last races due to heavy betting on outsiders. Gluck's Second Law thus attempts to exploit a specific form of punter bias. However, by studying actual returns at different points during the day they find no evidence of an exploitable opportunity.

In addition to horse racing studies of the quoted odds market, Cain et al. (1990) have investigated the British greyhound market. They looked at the percentage of wins and returns to placing a unit bet on all dogs in particular odds categories. They found that for the majority of odds there were negative returns but did find evidence of the "favourite long shot bias". The favourite long shot bias is believed to occur because punters prefer low probability high return bets to high probability low return bets. This shows up in their study with the actual percentage of wins up to odds of 6/4 being higher than the probabilities implied by the odds The evidence of favourite long shot bias is formalised by using a weighted least squares and a logistic model. The results of the regression analysis which indicate inefficiency were then utilised in an attempt to achieve abnormal returns with certain betting strategies. They were unable to achieve any positive returns.

Gabriel & Marsden (1990) compare the pari-mutuel system with the quoted odds system. They study the returns to starting price bets at bookmakers (quoted odds) and the returns to Tote bets (pari-mutuel). They find that the returns to the pari-mutuel bets are consistently higher, even though both betting systems are of similar risk and the payoffs widely reported. This suggests that the British racetrack market does not meet the conditions for semi-strong efficiency.

There has only been one study of fixed odds betting, focusing on English football. In this paper, Pope and Peel (1990), concentrate on the prices offered by different bookmaking firms in the UK in the 1981/82 season. They divide match outcomes into seven different odds categories. For each category they compare the number of times the predicted match outcomes were correct with the number implied by the odds They find that for most odds grouping the odds imply a higher probability that the actual outcomes suggesting positive bookmaker margins and the absence of systematic profit opportunities. As a more rigorous check they regress odds on outcomes and do find a bias in draw odds but are unable to translate this into a profitable betting strategy.

3.3.3. Spread Betting Systems

Most studies of spread betting have focused on the United States. Spread bets are accepted on most US sports but the majority of bets are taken on American football.

Zuber et al. (1985) try two tests of market efficiency in the spread betting market for American football. The first uses OLS to test whether forward prices (i.e. the bookmakers' line) are the best unbiased predictor of outcomes (i.e. the actual points spread). In the second test, they develop a predictive model of the points spread using variables reflecting previous playing performance. Using this model they try to see if they can achieve a return better than the bookmakers' take (in this case -5%). They obtain inconclusive results for the first test but do find exploitable opportunities using their predictive model. However, the sample they use is very small and they admit themselves that the results may therefore be misleading.

Sauer et al. (1988) criticise Zuber's methods and conclusions. Sauer re-estimates Zuber's equations over the whole season and does not split the season into sixteen weekly portions as Zuber did. Sauer's results offer support for the efficiency of the market, in particular they reject the hypothesis that the bookmakers line and the actual points score are unrelated, something that Zuber was unable to do. In addition, Sauer shows that the method Zuber used for exploiting inefficiency in the betting market, when extended out of the sample on which it was based, incurs substantial losses.

Gandar, Zuber, O'Brien and Russo (1988) extend the analysis of Zuber et al. (1985) by presenting two types of "rationality tests" of the point spread betting market. Statistical tests of rationality look at statistical properties of markets such as price correlations, while economic tests attempt to detect unexploited profitable opportunities. The test used by Zuber (1985), i.e. the test of whether the bookmakers' line is an unbiased predictor of actual outcomes is classed as a statistical test. Their results do not contradict the hypothesis that the bookmakers' line is a rational expectation of the games outcome. A further statistical test of whether the difference between the bookmakers' line and the actual points spread is correlated with publicly available information is carried out. Market rationality would imply that there was no correlation. Again the results cannot reject market rationality.

Their second type of tests, namely economic tests, attempt to find profitable betting rules. The betting rules used fall into two categories:

- *ad hoc or mechanical rules*: e.g. bet on the favourite team or bet on the team with the highest average winning margin; and
- *behavioural rules:* based on some specific idea of irrational public betting behaviour.

The mechanical rules are not profitable but they do identify three profitable behavioural rules. The three rules they identify are:

- bet on the team that becomes less favoured over the course of the weeks betting (i.e. from when the line is opened to just prior to the kick off of the game). In effect they are betting against the public believing them to bet irrationally.

- bet against the public as the above rule, but only for games in weeks following winning weeks for the public. Winning weeks were defined as when at least 50% of line changes in the week moved the betting line closer to the eventual game outcome. The idea being that losing concentrates punters' minds and their information processing will be more irrational after a winning week.

- bet on the underdog against a favoured team who in the previous week beat the predicted spread by at least 10 points. This attempts to exploit punters' overreaction to a large victory in the previous weeks results,

Thus Gandar et al. find that statistical tests cannot reject rationality but that economic tests do.

Lacey (1990) also carries out what Gandar et al. identify as economic tests on the NFL betting markets. Lacey tests 15 trading rules based on publicly available information over the 1984-1986 seasons. He does find a small number of profitable trading rules but concludes that the market is still in effect efficient.

Even & Noble (1992) refute the methodology of both Zuber et al. (1985) and Sauer et al. (1988). They state the requirement that the market forecast be an unbiased predictor of the actual outcome is neither necessary nor sufficient for market efficiency, unless the mean and median of the forecast errors are equal. However, ignoring the bookmaker's take, a 0.5 probability that the market forecast is higher than the actual outcome is both necessary and sufficient for market efficiency.

Dobra et al. (1990) look at spread betting but use basketball data. They develop a model of the bookmaker's line setting decision such that the bookmaker maximises his share of the handle (total amount of money wagered). They identify three different testable propositions:

If both the bookmakers and the public have "rational expectations" then:
(i) the bookmaker will balance the book and set the line where the probability of being above or below the line is 0.5

If the public is biased, (e.g. more likely to bet on a favoured team), then:
(ii) the bookmaker can balance the books, i.e. set the line such that the number of people wagering above or below the line are equal, in this case the probability of being below or above the line is not 0.5; or
(iii) the bookmaker can take a wagering position against the public by setting the line where the probability of being above or below it is 0.5.

They distinguish between i or iii and ii by seeing if the bookmakers line is an unbiased predictor of the actual points spread. If it is not, then they accept proposition ii. Propositions i and iii are distinguished by looking at the returns to bookmakers (i.e. the bookmakers' profitability); if this is bigger than their take (the profit they would make if the market was efficient) then this would support proposition iii. Their results give strong support for efficient utilisation of information by the market (i.e. proposition i).

Table 3.1 summarises the literature, illustrating the debate surrounding efficiency of betting markets. The majority of work on pari-mutuel systems and quoted odds systems find some form of inefficiency although claims about the availability of abnormal returns vary greatly. The most controversial area is spread betting markets. There are results in favour of both the efficiency and inefficiency of spread betting markets.

Table 3.1: Betting Market Efficiency In The Literature

Study	Sport	Test	Inefficiency	Positive Returns
Pari-Mutuel				
Snyder (1978)	US horse	Weak	✓	✗
Ali (1977)	US harness	Weak	✓	✓
Asch et al (1984 & 1986)	US thoroughbred	Weak	✓	✗
Haush et al (1981)	? exotic bets	Weak	✓	✓
Figlewski (1979)	US horse	Semi-strong	✓	?
Quoted Odds				
Dowie (1976)	UK horse	Strong	✗	✗
Crafts (1985)	UK horse	Strong	✓	✓
Johnson & Bruce (1992)	UK horse	Weak (?)	✗	✗
Cain et al. (1990)	UK greyhound	Semi-strong (?)	✓	✗
Gabriel & Marsden (1990)	UK horse PM vs Quoted	Semi -strong	✓	?
Pope & Peel (1990)	English football	Weak	✓	✗
Spread Betting Systems				
Zuber et al. (1985)	US football	Semi-strong	✓	✓
Sauer et al. (1988)	US football	Semi -strong	✗	✗
Gandar et al. (1988)	US football	Semi- strong	✓	✓
Dobra et al. (1990)	US basketball	Weak	✗	✗

3.4. Bookmaking In The UK

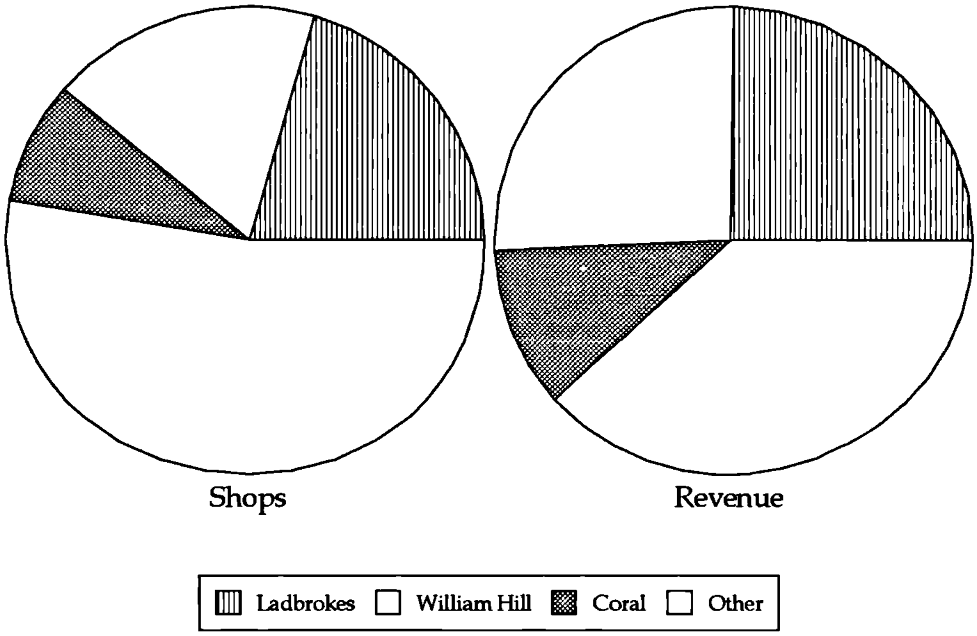
3.4.1. Structure of the Industry

The UK bookmaking industry is regulated by the Betting, Gaming and Lotteries Act 1963. The regulations are said to be based on the philosophy of "providing facilities to meet unstimulated demand and the prevention of crime"[38].

There are two types of bookmakers in the UK, on course and off course bookmakers. On course bookmakers are licensed by local magistrates, pay no betting duty and operate only within race courses and dog track grounds. Football grounds do have betting facilities but these are treated as off course bookmakers.

In 1994 there were 9,760 betting office licences held and around 4,700 bookmakers' permits[39]. Of these 9,760 shops approximately 43% were owned and operated by the "big three" namely Ladbrokes (1870), Coral (772) and William Hill (1594) (see figure 3.1). The big three also dominate in terms of revenue taking 62% of the £4196 million spent in 1994.

Figure 3.1: Bookmaking in the UK 1994



Off track bookmakers must have a permit and each of their premises must be licensed, again by the local magistrates courts. Off track bookmakers levy 10% 'tax' on customers' bets.[40] The customer can either pay 10% of her stake and not pay tax on her winnings or pay no tax on the stake and pay 10% of any winnings. The 10% of revenue is used by the bookmaker to pay betting duty and the horse race betting levy. The current betting duty tax rate is 7.75% of the total bet (including the tax element) and is paid to the Inland Revenue. The horse race betting levy incorporates two separate schemes, the company scheme for the big three and an alternative for the remaining bookmakers. The company scheme requires payment of 1.37% of turnover on all horse racing bets. The alternative scheme allows £262,000 worth of horse race betting before bookmakers must pay 2.2% of all bets. Despite only having to pay the levy on horse racing bets, bookmakers still charge 10% "tax" on all bets. The levy is paid to the Horse Race Levy Board who distribute the money to racecourses for facilities and for use as prize money. The amount of the levy each year is set by joint agreement between the Levy Board and the Bookmakers Committee (made up of representatives from all sizes of bookmakers). If the two parties fail to agree the Home Secretary sets the level.

In addition to the taxation regulations, the act places restriction on many of the off course bookmakers activities. Through the act, off course bookmakers are regulated in the hours

[38] Commission of the European Communities (1991).
[39] Source Home Office (1993).

they open, the type and medium of advertising (including shop front displays) and in the other uses to which a licensed bookmakers' premises can be put (e.g. gaming machines, alcohol and tobacco sales). Recently however these restrictions have become less stringent, for example bookmakers can now open on Sundays and are allowed to use text in their window displays.

As smaller bookmakers have closed, the big three have increasingly dominated the market and this process is continuing. It has been argued that price competition between the bookmakers, especially within football betting is not prevalent. The bookmakers themselves would no doubt dispute this. Competition is apparent in terms of facilities, both location and within shop services. However, local competition is diminished through the act as to obtain a new bookmakers permit you have to prove to the local magistrate that there is unserved demand in the area. The argument that you could provide bookmaking services more efficiently than a local incumbent is not acceptable.

3.4.2. The Football Betting Market

There are varying methods of betting on football in Britain. There are several specialist firms[41] that offer spread betting and bookmakers often quote odds on bigger games (especially televised games). However this chapter concentrates on the biggest form of football betting in Britain, via fixed odds coupons. These coupons offer betting odds on match results, the time and scorer of the first and last goal plus half-time scores.

In fixed odds betting the bookmakers produce a coupon three to four days before the football programme. The general coupon will contain odds for results and correct scores for each English and Scottish league match (or alternatively cup games if it is a cup weekend). The results bet offer odds for a home win, away win and draw. The correct scores odds are quoted in relation to the results odds regardless on which team you are betting. For example for any team whose odds to win the game are 4/5, the odds for a 3-0 win will always be 12/1. Thus free scoring teams and lower scoring teams are given the same odds. For the bigger games (e.g. live televised matches) often a single coupon will be produced which offers results odds, correct score odds and odds on the first/last player to score.

There are varying combinations of bets that are allowed. Single bets (i.e. bet on the result of a single game) are only usually allowed on televised matches. Trebles (simultaneously bet on the result of three matches) are allowed on approximately 30 matches usually including the Premier League, the First Division and Scottish Premier Division plus a handful of lower division games. Five fold bets (simultaneous bets on the results of five matches) are allowed on any combination of English and Scottish matches. The odds on the same match for different categories of bet are identical. Singles and upwards are accepted on correct scores.

Once the odds have been set it is extremely rare that they will change before the kick off of the matches. This is where fixed odds betting is most different from both pari-mutuel

---

[40] Reduced to 9% in 1996.

[41] For example Sporting Index.

systems and quoted odds systems, which reflect and react to the amount of money bet on each outcome right up to the start of the event.

3.5. Model Of Bookmakers' Odds Setting Decision

3.5.1. Introduction

The bookmaking business can be looked at as both a service and an information market. Bookmakers are providing a service, i.e. the opportunity and facilities to place a bet, but also creating an information market similar to markets in stocks and shares. The price for each aspect of the business can be separated. The price for providing the facilities to bet is the bookmakers' take while the prices in the information market are the relative odds.

In pari-mutuel betting systems the bookmaker's take is a fixed percentage of the handle or total amount bet. In fixed price odds betting this is not true but the bookmakers' take can be estimated from the "over-roundness" of the book.

The bookmaker's theoretical gross margin (over-roundness), assuming a balanced book and ignoring betting duty can be easily calculated. This is best illustrated by a simple example. In all matches the home team is the team whose ground the game is played on and is signified by coming first in the fixture. The odds on a Liverpool v Sheffield Wednesday match are 5/6 home win, 13/5 away win and 12/5 draw. A home win in this case is a Liverpool victory. For each outcome the percentage (amount bet required to win £100) is calculated as follows:

home win	100/ (1+5/6)	=	54.5
away win	100/ (1+ 13/5)	=	27.8
draw	100/ (1+ 12/5)	=	29.4

The over-roundness of the book is then the sum of the percentages less 100:
over-roundness = (54.5 + 27.8 + 29.4) - 100 = 11.7.

Thus if the book is balanced, that is, the bookmaker takes stakes on the three outcomes in the proportion 54.5: 27.8: 29.4 then the bookmaker will keep 11.7% of the stakes whatever the outcome of the match. The bookmakers' return is then 11.7/111.7=10.5% of the total stake (known as the handle)[42].

In football fixed odds betting the over-roundness of the book is remarkably constant at around 11.5% for all the major bookmakers. The average over-roundness in the sample of 3382 games is 11.47% with a standard deviation of only 0.34.

The second part of the price can be considered the relative odds or more properly the implied probabilities derived from the odds. Assuming that the over-roundness of the book is fixed and is 11% the formula below translates odds into implied probabilities. The

[42] For discussion of margins in different forms of gambling see The Royal Commission On Gambling (1978), Annex b.

probabilities calculated are the 'percentage' calculations above but adjusted so as to sum to one. The formula is:

$$implied\ probability = \frac{1}{1.11(1+odds)} \qquad \text{......(3.1)}$$

The prices investigated in this chapter are the prices in the information market. No attempt is made to explain or investigate how the 11% over-roundness is set. It is assumed that it is a result of the competitive and regulatory processes. Efficiency of the football betting market is looked at in terms of the information market only.

Adopting Fama's terminology, the football betting market can be assumed to be weakly efficient if abnormal returns can not be made using price information only (i.e. odds). Abnormal returns are defined as returns better than the bookmakers take, thus in this case returns better than -11%. The market can be assumed to be semi-strongly efficient if prices reflect all publicly available information and thus no abnormal returns can be made using publicly available information. The market can be assumed to be strongly efficient if the implied probabilities reflect all publicly available information and no one group in society can use its private information to achieve abnormal returns.

3.5.2. The Model

In the model it is assumed that bookmakers have no private information but are able to evaluate publicly available information as well as any other individual or organisation. In football it is unlikely that there is a great deal of insider information: games are played in front of large audiences and reported widely in the media. The sort of private information that could be available would be information on an injured player that a club is keeping secret. It is unlikely that this sort of information would vest with the bookmakers and not the media. Another form of private information could be the knowledge that certain players are going to throw matches. Despite the on-going high profile case[43] there is no evidence that this is widespread. The existence of this form of information does not however invalidate the assumption about the bookmakers information set.

There are three decision points in the model. Firstly the bookmaker decides which odds to quote. Secondly the punters decide on which outcome to bet and thirdly nature decides the outcome of the game. The bookmaker's decision process is modelled and thus will incorporate reaction functions for the punters' decision on which outcome to bet.

There are three distinct outcomes on which a punter can bet i.e. home win, away win and draw. These outcomes are denoted by use of the subscripts 1, 2 and 3.

---

[43] The Crown versus Messrs. Grobelaar, Fashanu and Segers.

Let H be the handle.

Let h_1 ,h_2 ,h_3 be the amount bet on each outcome. Thus the share of H on each outcome is defined as:

$$s_1 = \frac{h_1}{H} \qquad s_2 = \frac{h_2}{H} \qquad s_3 = \frac{h_3}{H} \qquad \text{.....(3.2)}$$

Let the bookmakers subjective probability that the result of the event be 1, 2 or 3 be represented by b_1, b_2 *and* b_3. By definition, $b_1 + b_2 + b_3 = 1$.

Let the bookmaker's posted odds be denoted by o_1, o_2 *and* o_3

As we are assuming that the overoundness of the book is 11% then:

$$\frac{1}{1+o_1} + \frac{1}{1+o_2} + \frac{1}{1+o_3} = 1.11. \qquad \text{.....(3.3)}$$

Let d be the implied probabilities from the odds i.e.

$$d_1 = \frac{1}{1.11(1+o_1)},\ d_2 = \frac{1}{1.11(1+o_2)}\ \text{and}\ d_3 = \frac{1}{1.11(1+o_3)}. \qquad \text{...(3.4)}$$

By definition $d_1 + d_2 + d_3 = 1$

Three different models of the punters' decision are investigated. The decision to enter the market has already been made. That is, the punters accept the bookmaker's price for offering his service (i.e. the over-roundness of the book). The model is only concerned with how punters spread their bets over the three outcomes. Thus the punters' reaction functions are only used to determine the share of the handle (i.e. s_i) which is bet on each outcome.

The punters' decision rules, set out below, are unlikely to hold for all punters at once. Indeed it could be argued that all three rules will be used by different punters simultaneously. Reaction function b in which punters bet according to the odds is a fairly unsophisticated view of their behaviour. There are likely to be marginal punters who are better informed. However if the informed punters form the minority in terms of the share of the handle, then the applicability of the results will still hold.

Throughout the model we are assuming that bookmakers know the appropriate punters' reaction function. Bookmakers are assumed to be risk neutral and as such are expected profit maximisers. The relaxation of this assumption and the introduction of tax are discussed later.

Reaction Function a: Punters bet a fixed share.

It is assumed that punters will always bet a fixed share on a particular outcome, thus the shares $s_1, s_2, and\, s_3$ are exogenous i.e. independent of the odds the bookmaker sets. This assumption is not as unrealistic as it might first appear. For example, if punters are attached to a particular team they may bet on this team to win whatever the posted odds. Thus if a well supported team is playing a less well supported team then the bookmaker will know what the shares of the handle will be on each outcome.

This phenomenon can often be seen when comparing the quoted odds of an England victory within England and abroad. For example, the odds on England winning the European championship would likely be much lower in England than in the rest of Europe due to punters backing England out of emotional attachment regardless of the odds. This is a particular type of market inefficiency due to punters' inefficient use of information.

The bookmaker's expected profit is:

$$E(\Pi) = H - b_1\left[h_1 o_1 + h_1\right] - b_2\left[h_2 o_2 + h_2\right] - b_3\left[h_3 o_3 + h_3\right] \text{.......(3.5)}$$

That is the bookmaker's expected profit is the handle less his subjective probability of each outcome multiplied by the payout for each outcome. Winning punters receive the odds multiplied by their stake ($h_i o_i$) and their stake returned (h_i).

Note: $h_i = Hs_i \quad and \quad o_i = \frac{1}{1.11d_i} - 1$, substituting into 3.5 gives:

$$E(\Pi) = H \quad b_1\left[Hs_1\left(\frac{1}{1.11d_1} - 1\right) + Hs_1\right] - b_2\left[Hs_2\left(\frac{1}{1.11d_2} - 1\right) + Hs_2\right] - b_3\left[Hs_3\left(\frac{1}{1.11d_3} - 1\right) + Hs_3\right]$$

....(3.6)

Note that $d_1 + d_2 + d_3 = 1$ and therefore $d_3 = 1 - d_1 - d_2$. Substituting into 3.6 and re-arranging you get:

$$E(\Pi) = H - \frac{b_1 Hs}{1.11d} - \frac{b_2 Hs_2}{1.11d_2} - \frac{b_3 Hs_3}{1.11(1 - d_1 - d_2)} \text{................(3.7)}$$

The bookmaker wishes to maximise his expected profit by setting odds given the punters' reaction function. For ease of calculation equation 3.7 contains implied probabilities but not odds, the bookmakers decision variable. Thus profit will be maximised by setting d (implied probability) which implicitly sets odds as each d implies a unique odd.

Thus differentiating with respect to $d_1 \; and\, d_2$

$$\frac{\delta E(\Pi)}{\delta d_1} = \frac{b_1 H s_1}{1.11 d_1^2} - \frac{b_3 H s_3}{1.11(1 - d_1 - d_2)^2} = 0 \qquad \text{.........(3.8)}$$

$$\frac{\delta E(\Pi)}{\delta d_2} = \frac{b_2 H s_2}{1.11 d_2^2} - \frac{b_3 H s_3}{1.11(1 - d_1 - d_2)^2} - 0 \qquad \text{.........(3.9)}$$

from 3.8 we get $d_2 = 1 - d_1 - \frac{\sqrt{b_3 s_3} d_1}{\sqrt{b_1 s_1}}$ ....(3.10)

substituting 3.10 into 3.9 you obtain

$$d_1 = \frac{\sqrt{b_1 s_1}}{\sqrt{b_1 s_1} + \sqrt{b_2 s_2} + \sqrt{b_3 s_3}} \qquad \text{........(3.11)}$$

similarly

$$d_2 = \frac{\sqrt{b_2 s_2}}{\sqrt{b_1 s_1} + \sqrt{b_2 s_2} + \sqrt{b_3 s_3}} \qquad \text{..........(3.12)}$$

For semi-strong efficiency it would be expected that $d_i = b_i$. That is, the implied probabilities from odds are equal to the bookmakers subjective probability. Bookmakers are assumed to possess all publicly available information and are able to process it as well as any other individual or organisation. This condition implies that their odds are the best predictor of outcomes and that no abnormal profits can be made.

Using equations 3.11 and 3.12 it can be shown that the market can be efficient, but it is dependent on the shares bet by the punters:

For all i if $s_i = b_i$ then $d_i = b_i$. That is, if the shares bet equal the subjective probabilities, then the implied probabilities equal the bookmakers subjective probability and odds are set an efficient level;

for a given i, if $s_i > b_i$ then $d_i > b_i$. That is, if the share bet on an outcome is greater than the subjective probability, then the implied probability is greater than bookmakers subjective probability and odds are set at less than their efficient level; or

for a given i, if $s_i < b_i$ then $d_i < b_i$. That is, if the share bet on an outcome is less than the subjective probability, then the implied probability is less than bookmakers subjective probability and odds are set greater than their efficient level.

Reaction Function b: Punters are guided by the odds.

It is assumed that punters believe that the odds are a good indication of how likely each outcome is. This is a fairly unsophisticated assumption. The extreme form of this assumption is used, namely that the handle is split in proportion to the implied probabilities. That is $d_i = s_i \; for \; i = 1,2,3$.

Using equation 3.7 from above:

$$E(\Pi) = H - \frac{b_1 H s_1}{1.11 d_1} - \frac{b_2 H s_2}{1.11 d_2} - \frac{b_3 H s_3}{1.11(1 - d_1 - d_2)} \quad \text{.....................(3.7)}$$

Substituting in $d_i = s_i \; for \; i = 1,2,3$ into 3.7

You obtain:

$$E(\Pi) = H - \frac{b_1 H}{1.11} - \frac{b_2 H}{1.11} - \frac{b_3 H}{1.11} \quad \text{.................................(3.13)}$$

As $b_1 + b_2 + b_3 = 1$ then:

$E(\Pi) = 0.0991H$

Thus the level of expected profit is independent of the implied probabilities d_i. That is, whatever the odds the bookmaker sets, his expected profit is still the same.

Reaction Function c: Comparison of subjective probability and odds.

A third, more sophisticated view of the punter is that she compares her subjective probability of an outcome with the odds. She will bet on outcomes which she believes are more likely than the odds suggest.

Thus, the share bet is a function of the implied probability from the bookmakers odds and the distribution of punters subjective probabilities over the three outcomes labelled P. Thus, $s_i = f(d_i, P)$.

Thus the bookmakers expected profit function becomes:

$$E(\Pi) = H - \frac{b_1 H s_1(d_1, P)}{1.11 d_1} - \frac{b_2 H s_2(d_2, P)}{1.11 d_2} - \frac{(1 - b_1 - b_2) H s_3(d_1, d_2, P)}{1.11(1 - d_1 - d_2)} \quad \text{.....(3.14)}$$

Differentiating with respect to $d_1 \,\&\, d_2$ the bookmakers decision variables:

$$\frac{\delta E(\Pi)}{\delta d_1}=\frac{b_1 H s_1(d_1,P)}{1.11 d_1^2}-\frac{\delta s_1(d_1,P)}{\delta d_1}\bullet\frac{b_1 H}{1.11 d_1}-\frac{(1-b_1-b_2)H s_3(d_1,d_2,P)}{1.11(1-d_1-d_2)^2}-\frac{\delta s_3(d_1,d_2,P)}{\delta d_1}\bullet\frac{b_3 H}{1.11(1-d_1-d_2)}=0 \text{.....(3.15)}$$

$$\frac{\delta E(\Pi)}{\delta d_2}=\frac{b_2 H s_2(d_2,P)}{1.11 d_2^2}-\frac{\delta s_2(d_2,P)}{\delta d_2}\bullet\frac{b_2 H}{1.11 d_2}-\frac{(1-b_1-b_2)H s_3(d_1,d_2,P)}{1.11(1-d_1-d_2)^2}-\frac{\delta s_3(d_1,d_2,P)}{\delta d_2}\bullet\frac{b_3 H}{1.11(1-d_1-d_2)}=0 \text{...(3.16)}$$

In this model in order for the market to be efficient then $d_i = b_i$ i.e. implied probability from the odds = the bookmakers' subjective probability which is assumed the best possible subjective probability. Substituting this into equations 3.15. and 3.16 and simplifying and rearranging:

$$s_1(d_1,P)=b_1\left[\frac{\delta s_1(d_1,P)}{\delta d_1}+\frac{s_3(d_1,d_2,P)}{(1-b_1-b_2)}+\frac{\delta s_3(d_1,d_2,P)}{\delta d_1}\right] \text{.....(3.17)}$$

$$s_2(d_2,P)=b_2\left[\frac{\delta s_2(d_2,P)}{\delta d_2}+\frac{s_3(d_1,d_2,P)}{(1-b_1-b_2)}+\frac{\delta s_3(d_1,d_2,P)}{\delta d_2}\right] \text{.....(3.18)}$$

Thus in order for the market to be efficient then the function that determines the shares evaluated at $d_1 = b_1$ *and* $d_2 = b_2$ must satisfy equations 3.17 and 3.18 above. This is by no means necessarily the case and thus a situation can be envisaged in which the expected profit maximising implied probabilities (in effect odds) are not equal to their subjective probability.

The simple numerical example below illustrates a case where assuming punter reaction function c can result in an inefficient outcome.

There is a single match Liverpool versus Manchester United. There are ten punters each betting 1 unit and following the same betting rule:

bet on outcome i such that:

$$\begin{array}{ll} i=\arg\max(p_i-d_i) & p_i \neq d_i \ \forall i \\ i=\arg\max(p_i) & p_i = d_i \ \forall i \end{array} \text{..........(3.19)}$$

That is bet on the outcome which maximises the difference between their subjective probability (p_i) and the probability implied by the odds (d_i). If for each outcome the subjective probability equals the probability implied by the odds, bet on the most likely event.

Let us assume that there are two types of punters, 6 Manchester United fans and 4 neutrals. The Man U fans believe that Man U have a better chance of winning the game than the bookmakers and have subjective probabilities $p_{1mu} = 0.4, p_{2mu} = 0.2 \, and \, p_{3mu} = 0.4$. The neutrals are of the same opinion as the bookmaker i.e.

$$b_1 = p_{1n} = 0.5, b_2 = p_{2n} = 0.2 \, and \, b_3\_p_{3n} = 0.4 .$$

Now assume that the bookmaker sets the market efficient level of odds $(d_i = b_i)$ i.e. $d_1 = 0.5, d_2 = 0.2 \, and \, d_3 = 0.3$. Using equation 3.19, all the Manchester U fans would bet on an away win, i.e. i=3 and all the neutral fans on a home win i=1. Thus the shares of the bet are $s_1 = 0.4, s_2 = 0 \, and \, s_3 = 0.6$.

Thus equation 3.14 becomes:

$$E(\Pi) = 10 - \frac{0.5*10*0.4}{1.11*0.5} - \frac{0.2*10*0}{1.11*0.2} - \frac{0.3*10*0.6}{1.11*0.3} = 0.99$$

However the bookmaker could set odds to take account of the Manchester U supporters' bias. For example he could set odds such that $d_1 = 0.41 \, d_2 = 0.2 \, and \, d_3 = 0.39$. With these odds and using equation 3.19, the punters would bet in an identical way to when the odds were efficient. Thus equation 3.14 becomes:

$$E(\Pi) = 10 - \frac{0.5*10*0.4}{1.11*0.41} - \frac{0.2*10*0}{1.11*0.2} - \frac{0.3*10*0.6}{1.11*0.39} = 1.45$$

Thus using punter reaction function c , the bookmaker can increase his expected profit by setting market inefficient odds.

3.5.3. Related Issues

In the model above it is assumed that the bookmaker is an expected profit maximiser implying risk neutrality. In the example above the bookmaker increases his expected profit by setting non market efficient odds. However his actual profit will depend on the outcome of the match. There is another risk minimising strategy that the bookmaker could adopt. If he always sets the odds such that the underlying probability equals his subjective probability (i.e. $d_i = b_i$) he would be guaranteed a return, whatever the outcome of the event of 0.0991H.

The bookmaker's decision can be represented in a payoff matrix as follows:

Nature Bookmaker	Home Win	Away Win	Draw
"Risk Minimising Strategy" $d_t = b_t$	0.0991H	0.0991H	0.991H
Expected Profit Maximising	$H - \frac{Hs}{1.11d_1}$	$H - \frac{Hs_2}{1.11d_2}$	$H - \frac{Hs_3}{1.11d_3}$

If for example $s_1 > d_1, s_2 = d_2$ *and* $s_3 < d_3$, then the expected profit maximising strategy will give a greater return to the bookmaker than the risk minimising strategy only if the outcome of the game is a draw. An away win would leave the bookmaker's position unchanged and a home win would make him worse off.

Two further issues also need discussing, namely taxation and multiple bet restrictions. "Taxation" is levied at 10% on all football bets. Tax will not affect the share of money bet on different outcomes, however it might affect a punter's decision to enter the market. In the three behavioural rules described earlier the variables that affect a punters decision on which outcome to bet are the odds and the punter's own subjective probabilities. The level of taxation does not affect either of these variables and can be seen as an addition to the charge for the bookmakers service (over-roundness). However, taxation will affect the level of returns required before the punter becomes profitable.

The second issue is multiple bets. As noted in Section 3.4 bookmakers often place restrictions on the number of games that make up a bet, for example often a minimum treble is required. As with taxation multiple bet restriction increase the over-roundness of the bet. For example on a triple bet if each of the matches has an over-roundness of 11% then the over-roundness on the triple will be $(1.11)^3$-1=0.368= 36.8%. The restriction on triple bets will not affect how the punter actually places her bet but might affect the decision to enter the market.

3.6. Empirical Tests

3.6.1. Weak Test Of Efficiency

A weak test of efficiency is to compare the subjective probability implied from odds with outcome probability. In an efficient market these would not be systematically different as this would allow profitable betting opportunities. This type of test has discovered long shot bias (i.e. more people bet on low probability high return combinations than on high probability low return bets) in pari-mutuel systems where lower odds categories have yielded positive returns[44].

[44] See for example Snyder 1978).

The sample used is 3382 matches from the 1993/1994 and 1994/1995 in the four divisions of the English football league. Odds are grouped into 24 categories and the implied probability calculated using equation 3.1 is compared with the outcome probability.

For example, there were 193 bets with odds of evens in the sample. Of these 193 bets 85 were correct (that is if the odds on a home win was evens and the result of the game was a home win then the bet was correct). Thus the outcome probability is 85/193=0.44. This is then compared with the implied probability from the odds i.e. a probability of 1/[1.11(1+1)]=0.45.

The ability to make profits is tested by calculating the return on £1 (plus tax) placed on each bet in a given odds category. In effect this is testing a strategy of placing a bet on every outcome with a particular level of odds. The returns are then compared with the expected return to these strategies under the assumption of odds efficiency. The expected returns given an efficient market are calculated below:

Using the notation of section 3.5

$$\text{odds } o_1, o_2, o_3 \ \textit{where} \ \sum_{i=1}^{3} \frac{1}{1+o_i} = 1.115 \text{ ....(3.20)}$$

that is the assumed over roundness of the book is 11.5% which as shown earlier is the case in practice.

Given the assumption that that the true probability of event i = the implied probability (from odds) of event i = $\frac{1}{1.115(1+o_i)}$

Then the expected payout on a £1 bet on event i =expected probability*stake*odds

$$= \frac{1}{1.115(1+o_i)} * 1 * (1+o_i) = \frac{1}{1.115} = 0.8969$$

Thus the expected return without tax is the payout minus stake divided by the stake i.e. (0.8969-1)/1= -10.31%

The expected returns tax paid is again the payout minus stake divided by the stake i.e. (0.8969-1.1)/1.1=-18.46%

Table 3.2 below sets out the implied probabilities, the odds from which these probabilities are derived along with the outcome probability and the return to a £1 bet tax paid. The odds are taken from the fixed odds coupons of a leading bookmaker (Ladbrokes) and the results taken from Rollin (1994 and 1995).

Table 3.2: Implied Versus Outcome Probability

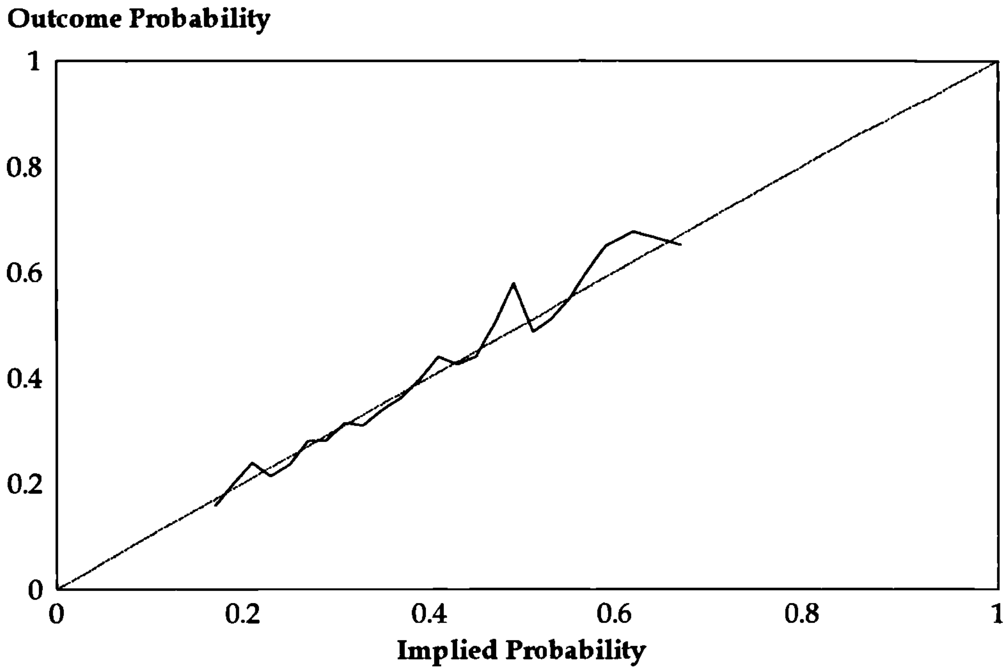
(a) Implied Probability	(b) Odds	(c) Outcome Probabilities	(d) Returns (%)
0.68	0.335	0.651	-20.99
0.63	0.441	0.678	-11.20
0.59	0.527	0.649	-9.84
0.57	0.581	0.601	-13.60
0.55	0.638	0.547	-18.51
0.53	0.700	0.511	-20.99
0.51	0.766	0.488	-21.71
0.49	0.839	0.580	-3.13
0.47	0.917	0.502	-12.47
0.45	1.002	0.440	-19 84
0.43	1.095	0.425	-18.98
0.41	1.197	0.440	-12.06
0.39	1.310	0.397	-16.67
0.37	1.435	0.362	-19.90
0.35	1.574	0.338	-20.89
0.33	1.730	0.310	-23.09
0.31	1.906	0.315	-16.91
0.29	2.107	0.282	-20.41
0.27	2.337	0.280	-14.98
0.25	2.604	0.236	-22.71
0.23	2.917	0.215	-23.60
0.21	3.290	0.240	-6.47
0.19	3.742	0.201	-13.48
0.17	4.299	0.158	-23.70

Notes:
(1) Implied probability as per equation 3.1
(2) Columns a and b are mid points of categories
(3) Column d is return to a £1 bet tax paid on each outcome that is in the implied probability category.

For an efficient betting market implied probability (column a) would equal the outcome probability (column c) and the returns (column d) would equal -18%.

Figure 3.2 below plots the outcome probability versus implied probability from the odds (columns c and a in the table). The 45 degree line represents where the outcome and implied probability are equal.

Figure 3.2: Implied versus Outcome Probability



As can be seen the implied probability seems a good match with the outcome probability except around a probability of 0.5. In order to do a more formal test of whether the outcome probability equals implied probability a simple OLS regression was carried out:

The estimated equation was : implied probability= b*outcome probability.

Table 3.3: Implied Versus Outcome Probability Regression

coefficient	0.97
t stat	74.74
R squared adjusted	0.92
95% lower limit	0.9463
95% upper limit	1.0002

The hypothesis that b=1 (i.e. implied probability =actual probability) could not be rejected at the 95% confidence level. Thus the regression analysis backs up Figure 3.2, there is no systematic bias.

The returns column in Table 3.2 shows that although there are returns that are better than the -18% there are no positive returns. The higher returns directly correspond to the odds categories where the actual probability is higher than the implied probability. Thus in order for there to be a systematic bias in the returns there would have to be a systematic difference between the implied and outcome probabilities, a hypothesis rejected by the above analysis. Over all odds categories the return was -18.11 %. Although in this sample the punter could achieve a return as high as -3%, there is no proof of a systematic bias.

Pope and Peel (1990) found that there was a systematic bias in draw odds. Table 3.4 below sets out the average return to home bets, away bets and draws along with the highest and lowest returns to a particular level of odds in each outcome. This is calculated by placing a £1 on each bet that falls into the category from the sample.

Table 3.4: Betting Returns

	Home	Away	Draw
Return to all odds categories			
Return %	-16.5	-22.4	-15.4
Return to a particular odds category:			
Highest	-2.5	+4	-4.5
Lowest	-47.6	-44	-23.4

The results above give no support to Pope and Peel's assertion that there is systematic bias in the odds for draws. As a further check the regression analysis carried out above was repeated for home win, away win and draw odds separately. Table 3.5 shows that all three regressions could not reject the hypothesis that b=1, i.e. implied equals actual probability.

Table 3.5: OLS Outcome Probability and Implied Probability

	Home	Away	Draw
coefficient	0.97	0.93	1.02
t stat	49.41	19.02	43.50
R squared adjusted	0.873	0.414	0.799
95% lower limit	0.934	0.822	0.971
95% upper limit	1.015	1.028	1.076

3.6.2. Tests of Semi-Strong Efficiency

Assuming an efficient market, it would not be possible for a punter to exploit publicly available information to achieve systematic abnormal returns. This section attempts to do just that, using ordered probit models and publicly available information, it tests the semi-strong efficiency of football betting markets by attempting to achieve abnormal returns.

Abnormal returns would at best be profitable returns, but are defined as better than the expected return of -18% (tax paid) reflecting the bookmakers take as set out in section 3.6.1 above.

The dependent variable in these estimations is the result of the match, a home win=0, a draw=1 and an away win=2.

The ordered probit was chosen to take account of the ordinal nature of the dependent variable. For example, if a very strong home team was playing a weak team then the most likely outcome would be a home win, then a draw then an away win. Reversing the teams strengths you would most likely see an away win, then a draw then a home win. Thus there is a natural ordering of the outcomes which is best addressed by using the ordered probit model.

Throughout this section, as well as ordered probit models, ordered probit models adjusting for heteroscedasticity and ordered logit models were also estimated. However there was no evidence to reject the ordered probit. Neither the logit nor the ordered probit adjusted for heteroscedasticity performed better than the ordered probit, nor did they give significantly different results.

The basic model using ordered probit is:

$$y^* = \beta' x + \varepsilon \qquad \text{......(3.21)}$$

where y* is the probability of the outcome and is unobserved. The x vector is made up of the explanatory variables. What is observed is the outcome of the game:

$$\text{home win} \quad y = 0 \quad \textit{if } y^* \leq 0, \text{...........(3.22)}$$

$$\text{draw} \quad y = 1 \quad \textit{if } 0 \langle y^* \leq \mu, \text{...........(3.23)}$$

$$\text{away win} \quad y = 2 \quad \textit{if } \mu. \leq y^* \text{............(3.24)}$$

μ is unknown and is estimated along with β.

The predicted probabilities are

$$\text{Prob}(y = 0) = 1 - \Phi(\beta' x) \text{.....................(3.25)}$$

$$\text{Prob}(y = 1) = \Phi(\mu - \beta' x) - \Phi(-\beta' x) \text{....(3.26)}$$

$$\text{Prob}(y = 2) = 1 - \Phi(\mu - \beta' x) \text{.................(3.27)}$$

In the ordered probit model information given by the estimated coefficients is limited. The marginal effects of the coefficients on the predicted probability of each outcome offer more insight. In this model the marginal effects are calculated as follows:

$$\frac{\delta \text{Prob}(y = 0)}{\delta x} = -\phi(\beta' x)\beta, \text{..................................(3.28)}$$

$$\frac{\delta \text{Prob}(y=1)}{\delta x} = (\phi(-\beta' x) - \phi(\mu - \beta\, x))\beta, \qquad (3.29)$$

$$\frac{\delta \text{Prob}(y=2)}{\delta x} = \phi(\mu - \beta' x)\beta. \qquad (3.30)$$

The data set used is for two English League seasons, 1993/94 and 1994/95. All four divisions are used giving 1733 match observations for 1993/94 and 1649 match observations for 1994/95. Four different approaches were tried and these are described below.

3.6.2.1. *Model with Odds Only*

The first model estimates the ordered probit model with only the posted fixed odds as explanatory variables, i.e. the odds of a home win, a draw and an away win. Table 3.6 below gives the results of the estimation using both seasons data.

Table 3.6: Ordered Probit-Odds Only

Dependent variable = outcome of 3382 matches in the 93/94 and 94/95 seasons

Variable	Coefficient	t-stat
Constant	-0.316	-1.416
odds of home win	0.351	5.655
odds of draw	0.111	1.192
odds of away win	-0.089	-3.240
mu	0.761	34.441
log likelihood	-3499.931	

Marginal Effects

on the probability of a home win	
constant	0.126
odds of a home win	-0.139
odds of a draw	-0.044
odds of an away win	0.035
on the probability of a draw	
constant	-0.024
odds of a home win	0.027
odds of a draw	0.008
odds of an away win	-0.007
on the probability of an away win	
constant	-0.101
odds of a home win	0.112
odds of a draw	0.036
odds of an away win	-0.029

The results of the estimation are mixed. Both the home odds and the away odds variable are significant but the draw odds variable is not. A Likelihood ratio test of the null hypothesis that all the coefficients on the odds variables are zero was calculated as:

$$LR = -2\left[\ln \hat{L}_r - \ln \hat{L}\right] \text{.......................(3.31)}$$

Here the restricted equation is where all the coefficients except the constant are 0 and the unrestricted equation is contained in the Table 3.6. The test rejects the hypothesis that all the coefficients are 0 thus supporting the model[45].

As the probabilities must sum to 1, the sum of the marginal effects for each variable should to be 0. This is indeed the case, for example the sum of the marginal effects of draw odds are: -0.044+0.008+0.036=0

You would expect that as the odds of a home win increase then the probability of a home win would decrease and the probability of an away win and a draw would increase. Thus you would expect the coefficient on the marginal effects on the home probability of home win to be negative and the marginal effects on an away win and on a draw to be positive. Extending this analysis to all outcomes the expected sign for the marginal effects are

Table 3.7: Expected Signs of Marginal Effects

	Probability of...		
	home win	draw	away win
home win odds	-ve	+ve	+ve
draw odds	+ve	-ve	+ve
away win odds	+ve	+ve	-ve

The signs on the marginal effects of home win odds are as expected for all the probabilities. However the sign of marginal effects of the draw odds is as expected for the away win probability but opposite to the expected sign for both home and draw probabilities. The away win odds marginal effect is as expected for both away win probability and home win probability but not for the draw probability.

In summary, the home odds and away odds are significant and the marginal effects are as expected. However the draw odds variable does not perform as expected. This could be due to the fact that draws are particularly hard to predict or it could perhaps be the indication of a systematic bias in the draw odds variable. To test this out the results of the model were used to enact a betting strategy and assess its profitability.

[45] Thus LR=-2[-3594.700--3499.931]=189.54 is Chi- squared with 3 degrees of freedom.

For each observation the predicted probability using the estimated coefficients and equations 3.25 - 3.27 were calculated. In addition the implied probability from the odds was calculated using[46]:

$$\text{Prob outcome } i = \frac{\dfrac{1}{1+odds_i}}{\sum_{1}^{3} \dfrac{1}{1+odds_i}} \quad \text{..........(3.32)}$$

A betting rule was then developed:

place £1 on the outcome of a particular match if:

$$\text{Probability Ratio} = \frac{\text{Predicted probability using model}}{\text{Implied probability from odds}} > X \quad \text{.....(3.33)}$$

The results are presented for different values of X. This betting rule could mean that a bet was placed on two outcomes in a particular game. An alternative rule could have been used to just place a bet on the maximum probability ratio in a given match if the probability ratio satisfied equation 3.33. This is an alternate but not necessary a better betting rule.

Table 3.8: Betting Results Using Odds Only Equation

	X			
	1.1	1.2	1.3	1.4
No bets (Max 3382)	392	26	14	11
% Bets Correct	30	54	43	27
Return No Tax (%)	-13	43	53	27
Return Tax Paid (%)	-21	30	39	16

Model estimated using both seasons data (3382) and the betting rule applied to both seasons data.

Table 3.8 indicates that abnormal returns can be made. Indeed, using the betting strategy described above and X's of 1.2 or above positive returns can be made. The returns rise up to X =1.3 and then drop when X=1.4 whilst remaining positive. This drop could be due to the smaller number of bets chosen as X increases, causing the reported returns to be more sample dependent. The positive returns imply that the betting market for football is not efficient and systematic abnormal returns can be made. It supports the existence of a systematic bias in the odds. However the strategy above could not have been practically implemented. The model was estimated using data to which the betting strategy was then applied.

In order to test the robustness of this result under more realistic conditions the data set was split into seasons. The model was estimated using 93/94 data only and then the betting

[46] This is a refinement of equation 3.1.

strategy using these coefficients was applied to the 94/95 season. The coefficients estimated were similar to the full sample estimations (see Table 3.6) for home and away odds and mu although the coefficient on draw odds was approximately three times larger and significant. The results using these coefficients are contained in the Table 3.9 below. As can be seen the returns are not as good and indeed are non-positive. However the returns are abnormal using the definition described earlier (i.e. better than -18%) and thus still support the inefficiency finding.

Table 3.9: Betting Returns Using Odds Only Equation 94/95

	X			
	1.1	1.2	1.3	1.4
No bets (Max 1649)	315	70	39	29
% Bets Correct	35	26	23	24
Return No Tax (%)	-10	-9	-5	8
Return Tax Paid (%)	-18	-17	-13	-2

Model estimated using 93/94 season data (1733 observations) and the betting rule applied to the 94/95 season (1649 observations).

3.6.2.2. *Publicly Available Information: Odds and Performance Data*

In order for prices to be efficient (in this case odds) then they must take account of all publicly available information. This test uses publicly available performance data to test if they improve on the model of match outcomes using just the odds as independent variables. If any of the performance variables was significant and thus did improve the odds as an indicator of match outcome then this would imply semi strong inefficiency, prices did not fully reflect all available information.

The publicly available information variables are all derived from Rollin (1994 & 1995). "Difference in teams'" just refers to the value of the variable for the home team minus the value of the variable for the away team. The variables used were:

- difference in teams' average points per game over the season;
- difference in teams' cumulative points over the season;
- difference in teams' league position;
- difference in teams' average points over the last three games;
- difference in teams' average goal difference (goals scored minus goals conceded);
- difference in teams' cumulative goal difference;
- difference in teams' goal difference in the last three games;
- difference in teams' average weighted points in last three games. Points gained in away fixtures are weighted more than points gained in home fixtures as these are more difficult to obtain; and
- difference in points between the home teams cumulative points won in home matches and the away teams cumulative points won in its away matches.

As the variables described above increase they would be expected to increase the probability of a home win and decrease the probability of an away win. As the variable approaches 0 the probability of a draw would be expected to be greatest. The only exception is the difference in league position, as this increases the probability of an away win should increase and the probability of a home win decrease.

All the above variables and combinations of variables were tried in addition to the odds as explanatory variables. None of the variables were found to be significant. Two examples are given below:

Table 3.10: Ordered Probit: Odds And Performance Variables 1

Variable	Coefficient	t-stat
constant	-0.338	-1.491
odds of home win	0.330	4.659
odds of draw	0.118	1.257
odds of away win	-0.082	-2.726
difference in home team's home points & away team's away points	0.017	0.549
diff. In cum goal difference	-0.002	-0.873
diff in goal difference last 3 games	-0.002	-0.334
mu	0.761	34.424
log likelihood -3499.359		

Dependent variable = outcome of 3382 matches in the 93/94 and 94/95 seasons

The likelihood ratio test cannot reject the hypothesis that the coefficients on the performance variables are 0. Here the restricted equation is the one described in Table 3.6 thus LR=-2[-3499.931--3499.359]=1.144.

Table 3.11: Ordered Probit: Odds And Performance Variables 2

Variable	Coefficient	t-stat
Constant	-0.314	-1.407
odds of home win	0.358	5.680
odds of draw	0.110	1.182
odds of away win	-0.091	-3.284
difference in average points last 3 matches	0.012	0.554
mu	0.761	34.441
log likelihood -3499.771		

Dependent variable = outcome of 3382 matches in the 93/94 and 94/95 seasons

The likelihood ratio test cannot reject the hypothesis that the coefficient on the performance variable is 0. Again the restricted equation is the one described in Table 3.6, thus LR=-2[-3499.931--3499.771]=0.32. Thus using publicly available performance data does not add to the explanatory power of the odds. The hypothesis of efficient markets cannot be rejected using this test.

3.6.2.3. *Performance Data Only.*

The test in 3.6.2.1 found that there were exploitable opportunities using just the odds model. Thus as an additional test of whether abnormal returns can be made was carried out using just performance data described above to develop models of match outcome. Two different models seem to have similar explanatory power to the odds only models and these were used to apply betting strategies.

Table 3.12 Ordered Probit: Performance Data Only

Variable	Coefficient	t-stat
Constant	0.119	4.419
difference in home team's home points & away team's away points	-0.057	-1.974
diff. in cum goal difference	-0.010	-7.564
diff in goal difference last 3 games	-0.009	-1.775
mu	0.754	34.455
log likelihood -3522.493		

Dependent variable = outcome of 3382 matches in the 93/94 and 94/95 seasons

Marginal Effects

on the probability of a home win

constant	-0.047
difference in home team's home points & away team's away points	0.022
diff. In cum goal difference	0.004
diff in goal difference last 3 games	0.004

on the probability of a draw

constant	0.009
difference in home team's home points & away team's away points	-0.004
diff. In cum goal difference	-0.001
diff in goal difference last 3 games	-0.001

on the probability of an away win

constant	0.038
difference in home team's home points & away team's away points	-0.018
diff. In cum goal difference	-0.003
diff in goal difference last 3 games	-0.003

Using the betting strategy described in 3.6.2.1 and the coefficients in Table 3.12 above, the returns detailed below are achieved:

Table 3.13: Betting Returns Performance Data Only

	X			
	1.1	1.2	1.3	1.4
No bets (Max 3382)	1812	594	236	70
% Bets Correct	30	25	23	19
Return No Tax (%)	-13	-13	-9	-11
Return Tax Paid (%)	-20	-21	-17	-19

Model estimated using both seasons data (3382) and the betting rule applied to both seasons data.

Table 3.13 indicates that no abnormal returns can be made using this model again supporting the hypothesis of efficiency.

3.6.2.4. *Performance Data and Selected Odds*

A fourth method using performance data and selected odds variables was developed. The results are described below.

Table 3.14: Ordered Probit: Performance Data And Selected Odds

Variable	Coefficient	t-stat
Constant	-0.231	-3.469
Odds of an away win	0.133	5.417
diff in league position	0.057	17.564
mu	0.804	34.596
log likelihood	-3363.494	

Dependent variable = outcome of 3382 matches in the 93/94 and 94/95 seasons

Marginal effects...

on the probability of a home win	
constant	0.092
away win odds	-0.053
diff in league position	-0.023

on the probability of a draw	
constant	-0.019
away win odds	-0.011
diff in league position	-0.005

on the probability of an away win	
constant	-0.072
away win odds	0.041
diff in league position	0.018

Using the method described in 3.6.2.1 and the coefficients in Table 3.14 above the returns detailed below are achieved:

Table 3.15: Betting Returns Using Performance Data & Odds Model

	X			
	1.1	1.2	1.3	1.4
No bets (Max 3382)	3533	1681	1019	607
% Bets Correct	44	53	53	50
Return No Tax (%)	16	31	36	44
Return Tax Paid (%)	6	19	24	31

Model estimated using both seasons data (3382) and the betting rule applied to both seasons data.

Table 3.15 indicates that there are abnormal returns can be made. Indeed, using X's of 1.1 or above positive returns can be made. This implies that the betting market for football is not efficient.

Again, however the strategy above could not have been implemented because the model was estimated using data to which the betting strategy was then applied. In order to test the robustness of this result under more realistic conditions the data set was split into seasons. The model was estimated using 93/94 data and then the betting rule using these coefficients applied to the 94/95 season. The coefficient estimates and significance were all very similar to those estimated with the full sample described in Table 3.14. The results using these coefficients are contained in the Table 3.16 below. As can be seen the returns are not as good but are still positive and thus abnormal, supporting the inefficiency of the betting markets.

Table 3.16 : Betting Returns Using Performance & Odds 94/95

	X			
	1.1	1.2	1.3	1.4
No bets (Max 1649)	1638	723	421	267
% Bets Correct	44	50	49	44
Return No Tax (%)	18	36	44	45
Return Tax Paid (%)	7	23	31	32

Model estimated using 93/94 season data (1733 observations) and the betting rule applied to the 94/95 season (1649 observations).

The betting strategies described above are practical in that the models are tested on data not included in the estimation sample. There is one difficulty in implementing the strategy due to the restriction on single bets. The strategy assumes that single bets are allowed on all games and this is not the case. This has two implications: firstly in order to cover all possible outcomes, the weekly number of bet is increased; and secondly the over-roundness of the bet is increased, reducing profitability.[47]

[47] Discussed further in Chapter 5.

3.6.3. Are there any experts?

A further test of efficiency is to see if so-called experts, for example newspaper tipsters, achieve a higher than average return. These experts' predictions are public knowledge, it is their job to publicise their forecasts. They have an incentive to get as many forecasts right as possible because that is how they keep their job. However McCloskey's jibe of "if you're so smart why ain't you rich?"[48] does seem to apply here - that is, why if they are so good do they not make their living from privately placing bets using their knowledge.

The test below attempt to discover if the "experts" predictions carry any more information over and above the information contained in the odds. Thus these tests are further tests of semi-strong efficiency looking at specific features of publicly available information.

Two data sets were used, weekly betting recommendations from The Times and from the Racing Post. Over the two seasons 1993/94 and 1994/95, The Times recommends 597 bets and The Racing Post 704.

Two separate tests were carried on both the data sets. The first uses ordered probit models to determine whether the inclusion of dummy variables indicating the recommended bet, increased the fit compared to a model which just included odds. The second test not only takes account of the success of the predictions but also the profitability of the bets. It places a £1 bet on each of the recommended bets and looks at the return that would have been made.

3.6.3.1. *Ordered Probit*

In order to test whether the experts prediction do add any information to the odds, two separate models were estimated. The first model uses just odds as the explanatory variable while the second uses both odds and the dummy variables indicating the recommended bet.

Table 3.17 below gives the results for the model using just odds for the Racing Post and The Times.

Table 3.17: Odds Only Ordered Probit

	Racing Post		The Times	
Variable	Coefficient	t-stat	Coefficient	t-stat
Constant	0.940	0.131	-0.458	-0.563
odds of home win	0.203	1.173	0.271	1.676
odds of draw	0.068	0.231	0.286	0.845
odds of away win	-0.133	-2.086	-0.175	-2.441
mu	0.810	16.193	0.699	13.414
log likelihood	-726.663		-608.299	

Racing Post: Dependent variable = outcome of 704 matches for which the Racing Post had a recommended bet
Times: Dependent variable = outcome of 597 matches for which The Times had a recommended bet

[48] See McCloskey (1992).

Despite the smaller samples, the results of the regression are very similar to the results of the odds regression on the full sample described in 3.6.2.1.

To capture the predictions of the "experts" two dummy variables were included.

Table 3.18: Prediction Dummy Variables

	Dummy Variable Value	
Recommendation	Home prediction	Draw prediction
Home win	1	0
Draw	0	1
Away Win	0	0

The results of the regressions for both the Racing Post and The Times are described below.

Table 3.19: Ordered Probit including Prediction Dummies

	Racing Post		The Times	
Variable	Coefficient	t-stat	Coefficient	t-stat
Constant	0.245	0.336	-0.291	-0.353
odds of home win	0.123	0.669	0.196	1.118
odds of draw	0.063	0.212	0.270	0.800
odds of away win	-0.122	-1.900	-0.149	-1.941
home prediction	-0.163	-1.430	-0.217	-1.393
draw prediction	0.132	0.320	-0.049	-0.387
mu	0.813	16.178	0.700	13.422
log likelihood	-725.562		-607.257	

Racing Post Dependent variable = outcome of 704 games for which the Racing Post had a recommended bet
Times Dependent variable = outcome of 597 matches for which The Times had a recommended bet

When estimating a probit model it is inappropriate to calculate the marginal effects for dummy variables. Instead the predicted probability using the average value of all other variables and all potential dummy values can be calculated. This is contained in the Table 3.20 below.

Table 3.20: Predicted Probabilities

	Racing Post Prediction..			The Times Prediction..		
Probability of	home win	draw	away win	home win	draw	away win
home win	0.490	0.374	0.426	0.510	0.443	0.423
draw	0.294	0.314	0.308	0.256	0.268	0.271
away win	0.216	0.312	0.266	0.234	0.289	0.306

Both the prediction dummies are insignificant in both the Racing Post and the Times' equations. However the predicted probabilities calculated above do show some expected effects. Table 3.20 shows that the probability of a home win is highest at 0.490 when a home win is predicted. For the Racing Post when a draw is predicted the probability of a draw is just highest (i.e. 0.314 versus 0.308 and 0.294). However for the Times the probability of draw is highest when an away win is predicted For the Racing Post the probability of an away win is highest at 0.314 when a draw is predicted while for the Times the probability of an away win is highest when an away win is predicted.

These results do not give great support to the skill of the "experts". However, a more formal test of whether the recommendations add more information to the model can be carried out using the Likelihood ratio test described earlier. Here the restricted models are the models with only the odds in (see Table 3.17) i.e. in effect the coefficient on the prediction variables are zero. The unrestricted models are the models calculated above using both odds variables and the prediction dummies. In both equations the LR statistics is chi-squared with 5 degrees of freedom.

The Racing Post LR= -2[-726.663- -725.562]=2.202

The Times LR=-2[-608.299- -607.257]=2.084

For both "experts" the hypothesis that the prediction dummies coefficients are zero cannot be rejected. Thus the conclusion that the prediction do not give more information than contained in the odds is confirmed. This suggests that the odds are semi-strong efficient when looking at this particular type of public information.

3.6.3.2. *Profitability*

The preceding analysis only addresses one part of the effectiveness of the predictions as it does not take account of the odds. For example predictors may get the majority of bets wrong but go for long odds which would provide a positive return. The return to the predictions is calculated below in Table 3.21. The returns to betting £1 tax paid on each of the forecasters' predicted outcomes is shown. The expected return if the forecasters added no information would be -18% (see 3.6.1.1). The calculated overall return for the Post and The Times at -17% and -16% is not much different from this expected outcome.

Table 3.21: Racing Post Season 1993/94 - 1994/95

By Prediction

	Home	Away	Draw	Total
Bets	359	336	9	704
% Correct	55	32	33	44
Return % (post tax)	-15	-20	-2	-17

By Division

	Prem	1st	2nd	3rd	Total
Bets	170	178	196	160	704
% Correct	49	38	46	42	44
Return % (post tax)	-7	-26	-16	-18	-17

Table 3.22: The Times Season 1993/94 - 1994/95

By Prediction

	Home	Away	Draw	Total
Bets	261	168	168	597
% Correct	57	40	25	43
Return % (post tax)	-11	-9	-33	-16

By Division

	Prem	1st	2nd	3rd	Total
Bets	111	152	160	174	597
% Correct	49	41	45	41	43
Return % (post tax)	-12	-18	-12	-21	-16

The Racing Post steered away from predicting draws (only 1% of games predicted). The majority of the bets suggested were in the home category (51%) and they were predicted with an accuracy of 55%. This gave a return of -15% indicating that the bets that were correct had low odds. The best performance was in the Premier League with a return of -7% possibly due to the greater amount of information in the public domain on the Premiership (e.g. more coverage on television and in newspapers). In no category of bets did the Racing Post achieve a positive rate of return.

The Times did not steer away from predicting draws (28% of games predicted) but did very badly in predicting them, getting only a quarter right and obtaining a return of -33%. Again the performance on home matches was the best, with the predictions being correct 57% of the time, but again the -11% return reflects the low odds on these bets. The best return in any category was slightly lower than the Racing Post at -9%, again indicating that a there were no positive returns.

Although in some categories there were returns better than the stated -18%, there is no evidence of systematic bias being exploited by the "experts".

3.7. Conclusions.

The theoretical framework used identified that it was possible for an expected profit maximising bookmaker to set odds at an market inefficient level. Market inefficient odds imply there would be exploitable opportunities for punters. The empirical tests were designed to try and identify any such exploitable opportunities.

The weak efficiency test, using just prices (i.e. odds), could not find any evidence of inefficiency. A comparison of implied probability from odds and outcome probability was made. There was no systematic difference and therefore no opportunity to make abnormal returns. Thus unlike in pari-mutuel horse betting, no evidence of favoured long shot bias was found whereby positive returns were achieved for certain odds categories.

Several efficiency tests aimed at identifying profitable betting rules were carried out. The data employed were prices and publicly available information The unique use of the ordered probit model was found to be successful. Four different approaches were used.

The first approach modelled outcomes with just odds as explanatory variables, it was found that the draw odds did not perform as expected. Thus the results of this model were used in a betting strategy. The strategy achieved abnormal returns (i.e. better than -18%), with positive returns when applied to the estimation sample but negative returns when tested out of sample. These abnormal returns indicate inefficiency. A plausible explanation for this inefficiency is the reluctance of punters to bet on draws. This reluctance is illustrated in the lack of draw bets recommended by the Racing Posts' columnist. Taking advantage of this distortion in punters betting habits, an expected profit maximising bookmaker would set odds at an market inefficient level.

The second and third test were less successful in identifying inefficiency. The second test added publicly available performance data to odds, but this did not improve the explanatory power of the model, suggesting that the odds fully reflected all public information. In addition the third test, utilising performance data only as explanatory variables was unable to identify a betting rule giving abnormal returns.

The fourth approach was more successful using a combination of publicly available data and selected odds as explanatory variables. Two useful models were identified and used as the basis of a betting strategy. One of these models, using away odds and difference in league position, achieved positive returns when combined with the betting strategy estimated both within and out of sample. Thus both inefficiency was identified and a positive return was achieved. An additional test of semi-strong efficiency indicated that "experts", in this case newspaper columnists, gave no further information than was available in the odds.

Thus whilst several tests did imply efficiency, using ordered probit as the main empirical tool and odds and publicly available information as explanatory variables, inefficiency in the football fixed odds betting market has been identified. A betting rule was found that exploited this inefficiency giving positive returns. Thus unlike the majority of the literature both inefficiency and profitable betting opportunities were identified in the football fixed odds market.

3.8. References

Ali, M.M. (1977). Probability and Utility Estimates for Racetrack Bettors. *Journal of Political Economy*. Vol. 85, no. 4, pp. 803-815.

Ali, M.M. (1979). Some Evidence of the Efficiency of a Speculative Market. *Econometrica*. Vol. 47, No. 2, pp. 387-392.

Asch, P., Malkiel, B.G. & Quandt R.E. (1984). Market Efficiency in Racetrack Betting. *Journal of Business*. Vol. 57, no. 2, pp. 157-175.

Asch, P., Malkiel, B.G. & Quandt R.E. (1986). Market Efficiency in Racetrack Betting: Further Evidence and a Correction. *Journal of Business*. Vol. 59, no. 1, pp. 157-160.

Brealey, R.A. & Myers S.C. (1991). *Principles of Corporate Finance. Fourth Edition*. McGraw-Hill, Inc. New York.

Cain, M., Law, D. and Peel, D.A. (1990). Greyhound Racing: Further Empirical Evidence On Market Efficiency in a Wagering Market. *University College of Wales Working Paper*.

Commission of the European Communities (1991). *Gambling in the Single Market - A Study of the Current Legal and Market Situations*.

Cornish, D.B. (1977). Gambling: A Review Of The Literature And Its Implications For Policy And Research. *Home Office Research Study No. 42*. London: HMSO.

Crafts, N.F.R. (1985). Some Evidence of Insider Knowledge in Horse Race Betting in Britain. *Economica*. No. 52, pp. 295-304.

Dobra, J.L., Cargill, T.F. & Meyer, R.A. (1990). Efficient Markets for Wagers: The Case of Professional Basketball Wagering in *Sportometrics* (ed. Goff B and Tollison R.D.) Texas A&M University Press. pp. 215-249.

Dolbear, F.T. (1993). Is Racetrack Betting on Exactas Efficient? *Economica*. No. 60, pp.105-111.

Dowie, J.A. (1976). On the Efficiency and Equity of Betting Markets. *Economica*. No. 43, pp. 139-150.

Even, W. E. & Noble, N.R. (1992). Testing the efficiency in gambling markets. *Applied Economics*. No. 24, pp. 85-88.

Fama, E.F., (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal Of Finance*. No. 33, pp. 383-423.

Figlewski, S. (1979). Subjective Information and Market Efficiency in a Betting Market. *Journal of Political Economy*. Vol. 87, No.1. pp. 75-88.

Figlewski, S. & Wachtel, P. (1981). The Formation of Inflationary Expectations. *The Review of Economics and Statistics*. Vol LXIII, No1. pp. 1-10.

Gabriel, P.E. & Marsden, J.R. (1990). An Examination Of Market Efficiency in British Racetrack Betting. *Journal of Political Economy*. Vol. 98, No. 4, pp. 874-885.

Gandar.J.M, Zuber, R., O'Brien, T. & Russo, B. (1988). Testing Rationality in the Point Spread Betting Market. *Journal of Finance*. Vol. XLIII, No. 4, pp. 995-1008.

Hausch, D.B. & Ziemba, W.T. (1985). Transactions Costs, Extent Of Inefficiencies, Entries And Multiple Wagers In A Racetrack Betting Model. *Management Science*. Vol. 31, No.4, April, pp. 381-394.

Hausch, D.B. , Ziemba, W.T. & Rubinstein, M. (1981). Efficiency of the Market for Racetrack betting. *Management Science*. Vol. 27, No.12, pp.1435-1453.

Home Office Research and Statistics Department (1993). Betting Licensing. *Home Office Statistical Bulletin*. Issue 26/94.

Johnson, J.E.V. & Bruce, A.C. (1992). Betting Patterns in Early and Late Races: An Exploitable Opportunity? *Discussion Papers in Accounting & Management Science, University Of Southampton.*

Lacey, N.J. (1990). An estimation of market efficiency in the NFL point spread betting market. *Applied Economics*. No. 22, pp. 117-129.

McCloskey, D. (1992) If You're So Smart, Why Ain't You Rich? in Auspitz, J.L. et al. (eds). Praxiologies and the Philosophy of Economics. *The International Annual of Practical Philosophy and Methodology*. Vol 1. Transaction, New Brunswick and London. pp 93-111.

Peel, D. & Thomas, D. (1988). Outcome Uncertainty and the Demand For Football: An Analysis of Match Attendances In The English Football League. *Scottish Journal of Political Economy*. 35(3), pp. 242-249.

Pope, P.F. & Peel, D.A. (1990). Information, Prices and Efficiency in a Fixed-Odds Betting Market. *Economica*, 56, pp. 323-41.

Quandt, R.E. & Asch, P. (1987). Efficiency and Profitability in Exotic Bets. *Economica*. No. 54, pp. 289-298.

The Racing Post (Various Issues 1993-1995). London.

Rollin, J. (ed.) (1994). *Rothman's Football Yearbook 1994/1995.* Headline Publishing, London, England.

Rollin, J. (ed.) (1995). *Rothman's Football Yearbook 1995/1996.* Headline Publishing, London, England.

Rosett, R.N. (1965) Gambling and Rationality. *Journal of Political Economy.* Vol. 73 No.6, pp. 595-607.

Royal Commission on Gambling (1978). *Final Report.* Cmnd. 7200. London: Her Majesty's Stationery Office.

Sauer, R.D., Brajer, V., Ferris, S.P. & Marr, M.W. (1988). Hold Your Bets: Another Look at the Efficiency of the Gambling Market for National Football League Games. *Journal of Political Economy.* Vol. 96, No. 1, pp. 206-213.

Snyder, W.W. (1978). Horse Racing: Testing The Efficient Markets Model. *The Journal of Finance.* Vol. XXXIII, No. 4, September, pp. 1109-1118.

Thaler, R.H. & Ziemba, W.T. (1988). Pari-mutuel Betting Markets: Racetracks and Lotteries. *Journal of Economic Perspectives.* Vol. 2, No. 2 Spring, pp. 161-174.

The Times (various issues 1993-1995). Times Newspapers Limited, London.

Zuber, R.A., Gandar, J.M., & Bowers, B.D. (1985). Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games. *Journal of Political Economy.* Vol. 93, no. 4, pp.800-805.

4. PRODUCTIVE EFFICIENCY IN ENGLISH FOOTBALL

4.1. Introduction

Productive efficiency has been investigated in many sectors including sports. However, as with much sport economics there has been a distinct American bias in the literature. Very few studies have been carried out in England and there has been no rigorous investigation of productive efficiency in football.

This chapter investigates productive efficiency in English football using both econometrics and data envelopment analysis, a technique not employed in sport efficiency studies before. As discussed further in Section 4.3, many sports productive efficiency studies employ what I have termed intermediate outputs and not the true inputs. Intermediate outputs are outputs which are a result of the production process (i.e. playing the sport) but not the final result. For example, in football intermediate outputs could include goals scored, goals conceded and players booked. This study will address this by using wages as a proxy for the true inputs into the production process, namely managerial and playing talent.

As with much sports economics literature, the results of this study has impacts outside the sports sector. In the current UK regulatory environment comparative efficiency studies have become increasingly important. For example, the Office of Water Services' (Ofwat) price setting decision for water companies were informed by comparative efficiency studies using econometric analysis[49]. This study compares the results of both methods using a well defined data set, something the regulators are unable to do.

Section 4.2 below discusses Farrell's definition of efficiency which is employed here. Section 4.3 looks at some of the work that has been carried out in estimating sports production functions and assessing efficiency. Section 4.4 describes the model and data. Sections 4.5 and 4.6 present the results of the econometric and data envelopment analysis. The final section, compares the methodologies and draws out some conclusions.

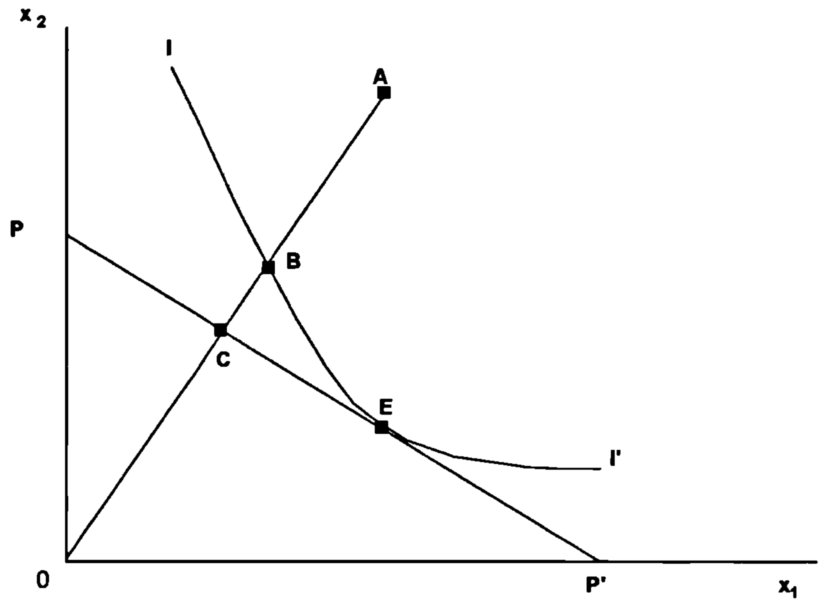
4.2. Efficiency

Farrell (1957) defined efficiency as the ratio of potential to actual performance. He decomposed overall efficiency into two mutiplicative components, technical efficiency and allocative efficiency. Consider Figure 4.1 where there are two inputs x_1 and x_2 into a production process. The production frontier can be represented by the isoquant II' producing a particular level of output, say y_1. The organisation at point A also produces y_1 and thus has technical efficiency equal to the ratio of potential to actual input consumption which is 0B/0A. Assuming that PP' is the isocost line defined by the ratio of prices then allocative efficiency can be defined as 0C/0B, the ratio of the minimum cost of producing y_1 to actual cost of producing y_1. Overall efficiency =technical efficiency * allocative efficiency = 0B/0A * 0C/0B = 0C/0A. Thus, boundary production alone is not sufficient for full

[49] See Ofwat (1994).

efficiency, it is also required that costs are minimised. Full efficiency (allocative and technical) is achieved at point E.

Figure 4.1: Technical And Allocative Efficiency



This chapter focuses on estimating technical efficiency, using both econometric frontier techniques and data envelopment analysis (DEA). Section 4.4 below discusses the model in more detail but it is useful to make some preliminary observations on the type of efficiency that will be looked at. Emphasis will be placed on playing and not financial success. The inefficiency on which the model will try to focus is organisational inefficiency and is not related to the skills of the team manager. The measure of wages used includes managers' wages and thus team managerial skill is treated in the same way as playing ability. In effect the assessment will be how efficient a club is in turning given inputs (player and team management ability) into outputs (playing success). Under this formulation it is possible for unsuccessful clubs (in terms of league position and cup performance) such as Huddersfield Town to be more efficient than traditionally successful clubs such as Liverpool.

4.3. Sports Productive Efficiency Studies

The following discussion divides the literature in three. Section 4.3.1 looks at studies which estimate production functions using econometrics. These functions are "average" production functions as there will be observations which both outperform and underperform the production function. Section 4.3.2 focuses on production frontiers estimated by either econometrics or linear programming techniques. The production frontier represents best practice and as such all observations will be on or below the frontier. The third section, 4.3.3, discusses efficiency studies which do not fall easily into either category including studies of marginal revenue product.

4.3.1. Production Functions

Schofield (1988) looks at production functions in cricket using data from the English County Championship and the John Player League from 1981-1983. His model comprises 5 equations:

Success = S(captaincy, fielding, batting, bowling, weather);
Captaincy = F_1 (Z);
Fielding = F_2 (Z, captaincy, weather);
Batting = F_3 (Z, captaincy, weather); and
Bowling = F_4 (Z, captaincy, fielding, weather).

Z includes such factors as: inherent player ability; form; experience; player availability; scouting and coaching skills; club management skills; quality of training facilities; and characteristics of playing facilities. In Schofield's framework I would classify Z plus captaincy as the true inputs and fielding, bowling and batting as intermediate outputs.

Whilst detailing an impressive model of true inputs, Schofield, because of measurement problems, concentrates only on batting and bowling performance. He ignores captaincy and fielding and the "intangible" variables in Z. Using OLS, both a linear and a multiplicative (i.e. Cobb Douglas) production function are estimated for both competitions. The data used is pooled cross section over three seasons normalised by sample season means. Team success is defined as the percentage of games won and the explanatory variables are again intermediate outputs (runs per game, runs per over, wickets per game and runs per wickets). The results are used to discuss the relative impact of bowling and batting on success. One interesting idea discussed is simultaneity, that is, not only does performance influence success but that success influences performance. For example, being top of the league may influence incentives and/or strategy. Schofield unsuccessfully tries to accommodate this simultaneity by using prior years performance as an instrumental variable.

Bairam et al. (1990) build on Schofield's work and provide international comparisons for his results using Australian and New Zealand data. Bairam uses a Box-Cox general transformation function which is not as restrictive as the Cobb Douglas functional form assumed by Schofield. The production function estimated by maximum likelihood estimation is:

$$(S^{\lambda}-1)/\lambda = A+\alpha_i[(B_i^{\lambda}-1)/\lambda]+\beta_j[(W_j^{\lambda}-1)/\lambda] \text{ ....(4.1)}$$

$$\text{where } \infty \; < \lambda < \infty^{+}, \alpha > 0, \beta < 0 .$$

Bairam shows that as long as $\lambda \leq 1$ then the function satisfies all the requirements of a neo-classical production function. If $\lambda = 1$ the function is identical to the Schofield's linear model while if $\lambda = 0$ it becomes Schofield's log linear estimate. For New Zealand it is found that λ is close to one supporting a linear function however the results for Australia were not as clear, depending on the variables included in the estimation.

Bairam, like Schofield, uses intermediate outputs in the estimation. In equation 4.1 above B= (runs per completed player innings, runs scored per over) and W=(runs per wicket taken, balls bowled per wicket, runs scored by opposition per over). S is again success and is defined as the percentage of maximum points possible.

Bairam, in line with Schofield, finds that attacking batting (as proxied by runs scored per over) is vital for maximising the probability of winning. This indicates that winning tactics in England, New Zealand and Australia need not differ significantly.

Carmichael and Thomas (1995) focus on productivity of English rugby league in the 1990/91 season. They model success (the percentage of games won) both as a function of intermediate outputs and playing and organisational characteristics. The intermediate outputs modelled are tries for and against and goals scored for and against. The player and organisational characteristics included are fitness of players (uses height and weight), experience (uses age), inherent ability (number of professionals, internationals and overseas players), team organisation (number of appearances per squad member) and coaching ability (coaching experience in years). They estimate a linear and multiplicative form (i.e. Cobb- Douglas) of their relationships using three models:

success= f(intermediate outputs);
success=f(player and organisational characteristics); and.........................(4.2)
intermediate outcomes=f(player and organisational characteristics).

Efficiency of each club was compared by the difference in the true and predicted output in terms of percentage of games won. A team with a positive (negative) value is interpreted as being of above average (below average efficiency). The use of player characteristics and organisational factors was only considered partially successful by the authors as not all the variables modelled were found to be statistically significant.

4.3.2. Frontier Analysis

Porter and Scully (1990) investigate major league baseball between 1961 and 1980. They attempt to investigate managers' efficiency in trying to maximise the win rate given a level of player skills. They model win percentage as a function of team hitting performance (total bases divided by the number of times bat) and team pitching performance (strike out to walk ratio). They calculate frontier unit isoquants for each year in the sample using a parametric linear programming technique. The technique they use requires that all observations of inputs per unit of output lie on or above the isoquant and minimise the sum of the squared deviations from the observations to the isoquant along rays through the origin. A particular manager's efficiency is calculated by the distance from the frontier isoquant. Porter and Scully are in effect measuring how effective managers are at turning intermediate outputs in to wins, not how efficient they are at producing these intermediate outputs.

A study by Zak et al. (1990) also estimates a production frontier but unlike Porter and Scully they use OLS. For a given vector of inputs the frontier production function is denoted by F(x). Observed output differs from the frontier by a factor u i.e. Y=F(x).u

If u is restricted to between 0 and 1 it can be used as a measure of production efficiency. Following Afriat (1972), when using a Cobb-Douglas form for the production function, it is assumed that v=-ln u, has a gamma distribution with parameter λ and thus the mean of the efficiency term u is equal to 2^{λ}. When λ<1 most observations are fairly efficient, when λ=1 a uniform distribution of efficiency results and when λ>1 then most observations are relatively inefficient. Thus, as is done, a teams average efficiency across games can be calculated by estimating λ which is shown under OLS to equal the variance of the regression. The estimated efficiency (2^{λ}) is then compared across teams.

The sport modelled is basketball using individual match data for the 1976-77 NBA season. In the Cobb Douglas (i.e. F(x)) production function intermediate output variables are used, for example ratio of home and away field goal percentages, ratio of personal fouls. The level of success modelled was the ratio of the final scores. Their results however do not adequately distinguish between teams. All teams appear efficient and the efficiency ranking is identical to the win/loss ratio ranking.

4.3.3. Other Studies

Kahn (1993) looks at managerial quality, team and individual success in major league baseball from 1969-1987. He argues that the market for managers is competitive; thus their salaries should approach their marginal revenue product, i.e. he uses salary to determine managerial quality. However he only has salary data for one year (1987) and therefore uses predicted salary using a regression with the 1987 data:

log managers salary= f (years experience, lifetime winning %, national league dummy) (4.3)

The estimated managerial quality variable is utilised in models of team performance and individual performance. He models team performance in two ways. In the first model the teams' win percentage is a linear function of managers quality, runs scored/runs allowed, win percentage in the previous year and a national league dummy. It could be argued that this equation underestimates managerial influence as scoring ratio is in some way determined by manager's quality. Therefore in addition Kahn estimates a team's win percentage as a linear function of: managerial quality; slugging percentage; stolen bases; earned run averages; fielding percentage; batting average; batters struck out by the teams pitchers; batters walked by the team pitchers and a dummy for strike shortened seasons.

In effect the first equation looks at how skilful managers are at turning runs ratio into wins, while in addition the second equation assesses the ability of the manager to turn all the intermediate output variables into runs scored/runs conceded. Kahn is in effect explaining win percentages with a combination of inputs (managerial quality) and intermediate outputs.

Kahn also tests whether players are able to get more from their ability whilst playing under high quality managers. Again, having only one year's salary data he regresses ability variables on salary and labels this function as long run quality. He estimates separate equations for both pitchers and nonpitchers. The example below is for non pitchers and all the variables are career averages:

Wages (long run quality) = F(slugging average, batting average, walks per bat,
stolen bases, fielding average)..............(4.4)

Kahn then calculates short run quality using the coefficients derived from 4.4 and the seasonal averages for the five variables. To assess whether good managers make a difference to players he then estimates:

short run quality-long run quality= f(managerial quality, plus dummies for players over 35 with more than 10 years experience, with less than three
years experience, infielders, catchers).......................(4.5)

In both formulations the managerial quality variable was positive. Thus Kahn concludes that managerial quality has a positive, usually significant effect on team winning percentage, controlling for team offensive and defensive inputs, as well as on player performance relative to player ability.

In effect because of the inclusion of intermediate outputs in the equations Kahn is looking at only a limited part of the managerial function; in particular no attempt is made to assess the managers ability in building a winning squad (i.e. employing the most appropriate players). What is being picked up is a managers ability to combine the players effectively (team performance equation) and motivate players (player ability equation).

Although there has been no rigorous treatment of productive efficiency in football two studies have addressed the issue in some form. In the first paper dell'Osso & Szymanski (1991) look at 12 leading clubs in the English football league between 1970 and 1989. They compiled a points table based on league positions and find that by far the most successful club is Liverpool. Using OLS they regress expenditure on players' wages and transfers against success (average league position). There are two major outliers, Liverpool and Notts Forest. Without these two clubs in the data the equation yields an R^2=0.81 but with them an R^2=0.4. The paper then discusses how these clubs could achieve this competitive advantage focusing on four different capabilities formal (legal), reputation, technology and architecture. Architecture is defined as the structure of contracts both formal and informal within an organisation. They conclude that Liverpool's competitive advantage is due to its superior architecture while Notts Forest's is due to the exceptional skills of their manager Brian Clough. Whilst interesting the paper uses a very small sample and the R^2 test identifying Liverpool and Notts Forest as efficient is not as rigorous as it could be.

In a later paper Szymanski and Smith (1993) set out a model of English football defining simultaneously a revenue function, a production function and a cost function. To estimate these they use OLS on pooled data for 48 clubs over 15 years allowing club specific intercepts. The production function estimated was:

quality (position in league)= f(difference in wages from average, quality last period).....(4.6)

They utilise the 48 club specific intercepts to indicate the position the club would have had in the league had it spent the average amount on wages. As with the earlier paper Liverpool are found to be outliers.

Following Scully (1974) a literature has developed assessing the marginal revenue product of sportsmen, mainly baseball players[50]. This literature is related but not of direct import to the current study and thus will only be discussed briefly. The literature is based on the (fairly unrealistic[51]) assumption that fans only attend a game to see the home team win. The approach identifies two equations. The first "performance" equation models team winning percentage as a function of intermediate outputs. In baseball for example, these intermediate outputs could include team slugging average and team strikeout to walk ratio. The second "revenue" equation models revenue as a function of team winning percentage and a number of market characteristics, for example size of local population, quality of stadia and so on. The second revenue equation can then be used to estimate the value in terms of revenue of a one point increase in percentage win. Individual players performance records are then used along with the results of the performance equation to estimate what effects their playing performance had on their teams winning percentage. Using this figure along with the revenue per increase in percentage win, a marginal revenue for each player can then be estimated.

4.3.4. Conclusions

As with most of the sports economics literature there has been an emphasis on American sports in productive efficiency studies. The majority of studies concentrate on the statistically rich baseball, basketball or American football. However there have been some studies in the UK of cricket and rugby union both discussed above. To date there have been no rigorous productive efficiency studies on English Football although Szymanski and Smith (1993) do estimate a simple production function.

All these studies utilise OLS with the exception of Porter and Scully (1990) who use a linear programming technique. In all papers, emphasis is placed on what I have termed intermediate outputs. Inputs should include factors such as player ability, coaching skills, training facilities, optimal team selection, tactics and so on. Although some of these factors are hard to proxy there seems to have been no attempt with the exception of Carmichael and Thomas (1995) to model them.

Whilst utilising true inputs Szymanski and Smith (1993) only model limited inputs and outputs i.e. league position and wages. As section 4.4 below shows, this chapter uses a

[50] See for example Sommers & Quinton (1982)
[51] See Chapter 2.

broader range of inputs and outputs. In addition the use of a stochastic production frontier and DEA as empirical techniques provide a more rigorous treatment. Szymanski and Smith do not focus on inefficiency of particular clubs nor explore potential reasons for it.

4.4. The Model and Data

As discussed in more detail below the central focus of the model is playing success and the major input into this process is playing and team management ability. The model incorporates a measure of team managerial skill and as such the inefficiency the model attempts to identify are factors which prevent the full utilisation of team management and playing talent. It is not the skill of the team manager that is under scrutiny.

The management studies literature identifies several areas where internal organisation or the structure of the club may lead to inefficient use of human resources[52]. dell'Osso & Szymanski (1991) use the general term "architecture" to describe these factors. Some of the more important features are discussed below:

Organisational Structure-put simply, who is responsible for what. For example does the board or the team manager deal with transfers? Does the team manager have other responsibilities that detract from his team management duties (e.g. press and public relations)?

Power-this is related to but more specific than the first point. The question is really within an organisation how is power distributed and enforced. For example you may have two clubs whose chairman are ultimately responsible for transfers. However in one club a chairman could be in place ruling by fear and retaining power by restricting information dissemination (e.g. Robert Maxwell the media tycoon when in charge at Derby County). Alternatively a chairman could just exercise a veto on the team managers decision, in most cases deferring to the team managers' greater playing knowledge.

Culture-these are the norms and beliefs which are held in the organisational. They determine what type of characters will succeed and be rewarded within an organisation. For example different clubs will have different disciplinary procedures. Many clubs might not have supported Eric Cantona after his attack on a fan in the way Manchester United did. Actions like this send signals for the level of acceptable behaviour amongst other players and the type of characters that will succeed in an organisation.

Strategy- Do clubs develop and maintain strategies for dealing with their changing environment or is the management style more reactive. Recent environmental changes include the formation of the Premier League and the results of the Bosman ruling which significantly altered the transfer market.

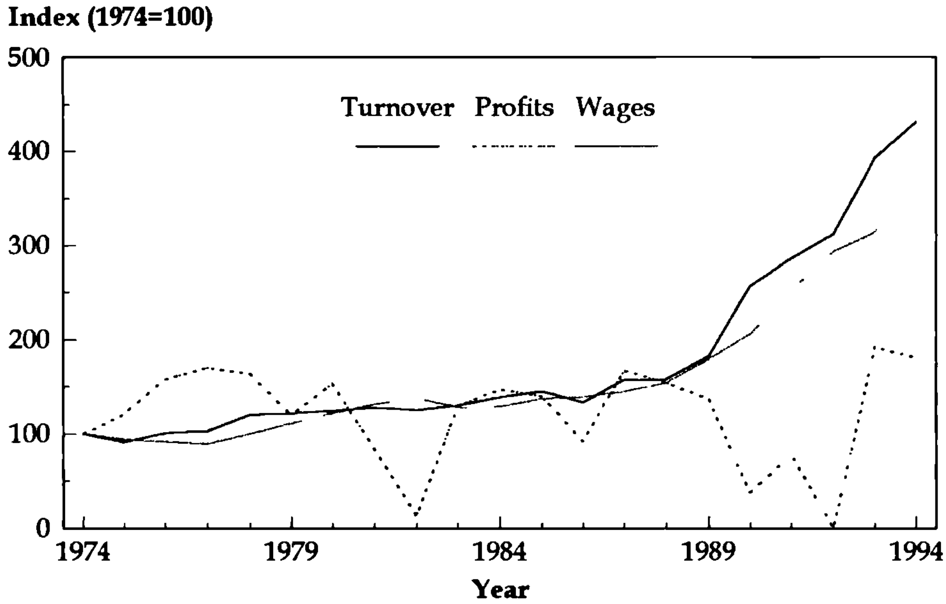
The following sections discuss in more detail the inputs and outputs of the production process modelled and the data set on which the estimations are based.

[52] See for example Handy (1985) or Wilson (1995).

4.4.1. Outputs

In line with other sports studies[53] the objective function of a football club is assumed to be playing success whilst remaining solvent. This assumption seems to fit with anecdotal evidence quite well. In 1994 whilst announcing the sacking of their manager, David Moores the Liverpool chairman stated that the only objective of the club was to win trophies and be a source of pride to their supporters. However, this could not be said for all clubs. In the last three to four years some bigger clubs, most notably Manchester United, have been floated on the stock exchange and are thus subject to meeting shareholders demands requiring a positive rate of return. Many of the clubs in the sample incur consistent losses, indeed over the sample used there is an average loss and not a profit. The change in real average profit in the sample is shown in Figure 4.2 along with turnover and wages. As can be seen, over the sample period there is an increase in both turnover and wages but no increase in profitability. Despite these losses since 1974 only a handful of clubs have gone bankrupt, most of whom reform immediately under a new board. The ownership structure of the clubs is characterised by a few directors, often family members or businessmen successful in another industry investing in football almost as a hobby.

Figure 4.2: Football Club Financial Performance 1974-1994



All figures are real
Sample of 39 clubs

Given the clubs objective function it is sensible that the output of the production process is defined as playing success. In this chapter playing success is predominantly measured via league performance. The league structure in England is made up of 92 league clubs in four divisions[54]. League games comprise the majority of fixtures and winning the league is considered the main objective of clubs.

[53] See Sloane (1971), Szymanski & Smith (1993) and Carmichael & Thomas (1993).

[54] See description in Chapter 1.

Performance in the league is measured by:

League performance = (93-league position)/ league position.....(4.7)

League position is the position in the league at the end of the season with 1 being the winner of the Premiership and 92 being the bottom club in the Third division. Thus the logarithnic version has reflecting barriers i.e. winner of the Premiership would obtain a value of 92 (ln=4.521) and the bottom of the Third division would obtain 0.011 (ln=-4.521). This formulation reflects more accurately success than just comparing the relative league positions (e.g. 92 with 91).

In addition to the league competitions there are two domestic cup competitions, the FA Cup and the League Cup (in its present guise known as the Coca -Cola Cup). The FA Cup is a knockout competition with 7 rounds and a final, the League Cup 6 rounds then a final. The variable used to capture cup success is the round the club was eliminated from the competition squared. An extra round value is given to the winner of the competition. Thus the winners of the FA Cup would get a value of 81, the runners up 64 and so on. In contrast the winner of the League Cup gets a value of 64, reflecting the greater importance of the FA Cup.

4.4.2. Inputs

The major input into a football clubs production process is labour. Player ability and team managerial ability (i.e. the ability to acquire, coach and combine playing talent into a successful team) are essential to playing success. The production of playing success is not a capital intensive process. The major capital investment in football, expenditure on stadia, is unlikely to impact on playing success. It could be argued that more sophisticated training and medical facilities could increase playing success, however in reality football is a simple game and effective training can be undertaken with only a pitch, a ball and some goal posts.

Unlike in other studies where playing is proxied by intermediate outputs (e.g. Bairam et al. (1990)) this study uses direct monetary valuation of ability, namely wages. Szymanski and Smith (1993) argue that unlike in many labour markets effort and performance of players is "easily" observed and monitored and thus asymmetry of information problems do not influence the market for football players. Thus it could be argued that wages will indicate playing talent. However even in a competitive market without asymmetry of information or bargaining power, wages are likely to reflect other factors than just playing ability such as the value of the player in non playing terms, e.g. marketability. Despite these problems wages are used to measure playing talent and are considered a better method than the intermediate outputs of Bairam et al. (1990) or the characteristics approach (e.g. age, experience) used by Carmichael and Thomas (1995).

A second feature of the labour market for players, namely transfer fees, has not yet been discussed. When a player moves between clubs, his new club compensates his previous club for their loss[55]. Thus the amount of money spent on transfer fees could also be used to measure playing talent. However transfer fees will not always cover all players in a squad, i.e. many clubs include home grown players developed through the club's youth system. In addition it is likely that transfer fees are subject to outside influences such as bargaining power and attitude to risk to a greater extent than wages and are thus likely to be further removed from playing ability. Carmichael and Thomas (1993) present a Nash bargaining model of transfer fees in which transfer fees not only depend on the value of the player to the club but also on the bargaining strength of the two parties based on several factors including attitude to risk.

In the model wages are employed as a relative measure. That is the playing talent of a club is relative to that of its peers. More formally, where there are c=1..C clubs for t=1..T time periods then:

$$ability_{ct} = \frac{wages_{ct}}{\sum_{c-1}^{C} wages_{ct} / C} \qquad ....(4.8)$$

In some formulations of the model wages have been expressed as a wage per squad member figure, approximating the average talent of the squad. This is an approximation as the wage bill will include more than the professional players, but still provides a good proxy for ability.

$$ability\ per\ squad\ member_{ct} = \frac{wages_{ct} / professionals_{ct}}{[\sum_{c=1}^{C}(wages_{ct} / professionals_{ct})] / C} \qquad ....(4.9)$$

In addition to wages several other inputs to the production process are considered, some as direct inputs others as a decomposition of inefficiency. When ability per squad member is used an additional measure of the number of professional players a club has is included. It is thought that given a certain average squad quality, the bigger the squad, the better equipped it will be to deal with the inevitable injuries that occur during a season and thus the more successful it will be.

Another input investigated is the number of different players used in a league campaign. Generally a successful teams will use a smaller number of players than unsuccessful teams. One of the major reasons large number of players are used is injuries. The inclusion of this variable is to identify player injuries as a random event out of the control of the club thus removing this factor from any estimates of inefficiency.

However there are other reasons why a team may use a large number of players during the season, for example the period of uncertainty after a new manager is appointed, player

[55] For a good description of the workings of the transfer market see Carmichael & Thomas (1993).

inconsistency and even managerial inconsistency. However other inputs should take account of these factors. The inconsistency of both players and management should be taken account of via the ability input and the transition period should be identified by managerial tenure variable which is discussed below. Thus in estimation the interaction of the player variable with the ability and tenure inputs was investigated to check whether the player variable was acting as intended.

The number of home produced players was included in some specifications of the model. The idea behind this is that initially there might be a wage discount on home grown players when compared with purchased players. For example while Robbie Fowler was establishing himself in the Liverpool team his contract was re-negotiated several times over a couple of seasons, an event unlikely to happen for a purchased player.

A input measuring managerial tenure was also investigated. The variable was set equal to the number of years the manager had been in his job. The variable was included to pick up the cost of change when a manager is replaced. Building a football team takes time and initially the change is likely to affect the productivity of the team. This variable is not really an input but tries to separate out a particular form of inefficiency. Some clubs have very little patience with managers who are not immediately successful, changing managers frequently despite the manager needing time to realise his ability (as reflected in his wages).

The source of the wage information was club accounts lodged with Companies House[56] and was taken as playing staff wages including managers' wages[57]. All other data are taken from various editions of Rothmans Football Yearbooks[58].

4.4.3. Data Set

The data set consists two panels:

- 47 clubs from 1974- 1989; and
- 39 clubs from 1974-1994.

All clubs were full member of the league throughout the sample period (i.e. any clubs relegated to non-league are excluded). There are 92 clubs in the league and thus the samples consist of 51% and 34% of league clubs. The reason the panel does not contain the full 92 clubs is that not all clubs submit full accounts to Companies House every year and thus the information on wages was missing for many clubs.

Table 4.1 shows the average performance of clubs in the two samples compared to the population. League Position and FA Cup and League Cup round eliminated are calculated as if all 92 clubs were in the sample, while attendance is the outturn value for the full population.

[56] I would like to thank Stefan Szymanski for use of this data.
[57] Some noise introduced in that not all clubs split out playing staff wages in each year.
[58] See Rollin (Various).

Table 4.1 Sample Vs Population Averages

	League Position	FA Cup Round Eliminated	League Cup Round Eliminated	Average League Attendance (000s)
Population Average	46.0	3.2	2.4	10.4
1974-1989 (47 clubs)	42.2	3.4	2.5	12.7
1974-1994 (39 clubs)	38.4	3.5	2.7	13.9

As Table 4.1 indicates the two samples are slightly biased in favour of more successful clubs, with the thinner longer panel (39 clubs) being slightly more biased. This bias is not too concerning as a major focus of this chapter is relative efficiency. However, it may be more of a concern when investigating the underlying importance of inputs in the production process. As a test of how important this bias is the underlying estimations were carried out for both panels allowing a comparison to assess if the results were affected by the larger bias in the 39 club sample.

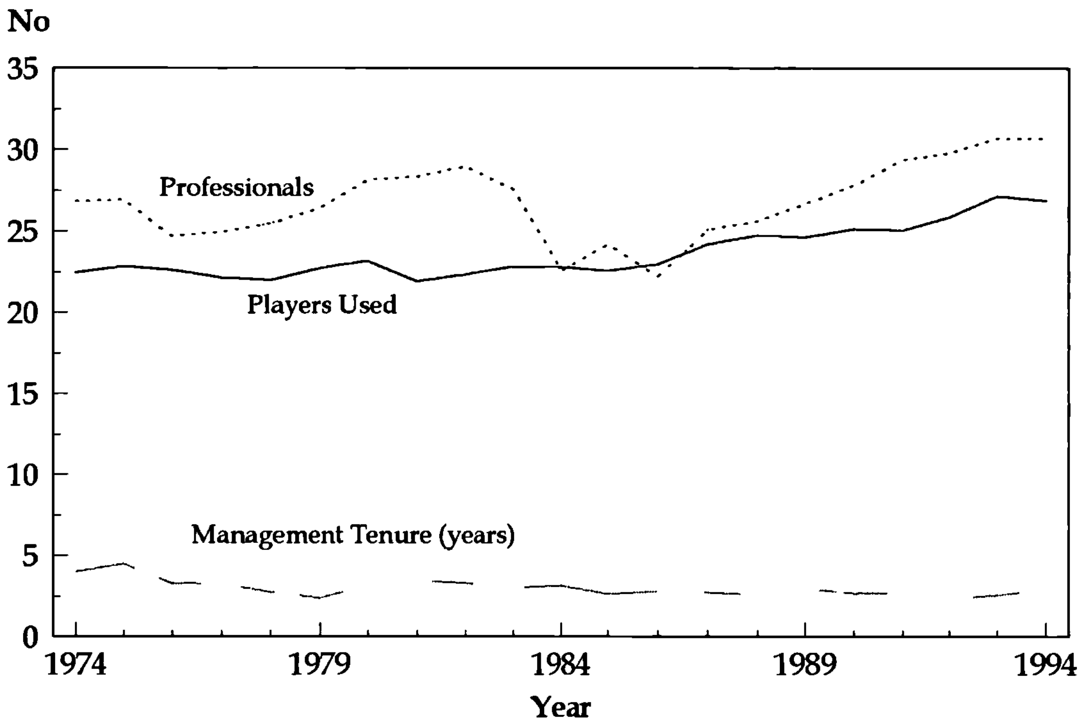
Appendix 1 and 2 give the clubs and sample means for the variables in both panels which are summarised in Table 4.2 below. Table 4.2 confirms Liverpool as the most successful side in the sample period. They obtain the highest score for all three competitions (i.e. league, FA Cup & League Cup) particularly dominant in their league performance. However it is unclear whether this will mean they are efficient given that they also top many of the input variables (wages and professionals). It could be important that Liverpool have the minimum players used average. A similar story is shown at the unsuccessful end of the sample, with Scunthorpe appearing in both the lowest output and lowest input columns.

Table 4.2: Ranges Of Sample Means 1974-1994

	League Performance	F.A. Cup	League Cup	Relative Wages	Professionals	Management Tenure	Home Grown Players	Players used
	Eqn 4.7	Round Eliminated	Round Eliminated	Eqn 4.9	No.	Years	No.	No.
Highest	Liverpool 61.70	Liverpool 5.90	Liverpool 4.81	Liverpool 2.12	Man U 37.58	Oldham 7.12	Arsenal 20.76	Birmingham 26.53
	Man Utd 26.74	Man Utd 5.67	Arsenal 4.33	Arsenal 2.07	Liverpool 33.90	West Ham 6.71	Spurs 20.57	Brentford 25.67
Average	6.49	3.52	2.68	1.00	26.80	3.01	12.05	23.66
Lowest	Mansfield 0.42	Southend 2.05	Bristol R 1.57	Rotherham 0.48	Bury 21.80	Hull Peterborough Plymouth 1.95	Cambridge 5.43	Arsenal 21.23
	Scunthorpe 0.17	Mansfield 2.04	Scunthorpe 1.48	Scunthorpe 0.46	Peterborough 21.76	WBA 1.62	Bury 5.14	Liverpool 19.76

As well as looking at the sample averages across clubs it is also interesting to look at the sample averages over time. Figure 4.3 below shows the number of professionals, players used and management tenure over the sample period. Management tenure is fairly stable over the period with a slightly lower value at the end of the sample than at the beginning. The number of professionals and players used are more volatile with both variables levelling off the end of the sample after an increasing trend since about 1986.

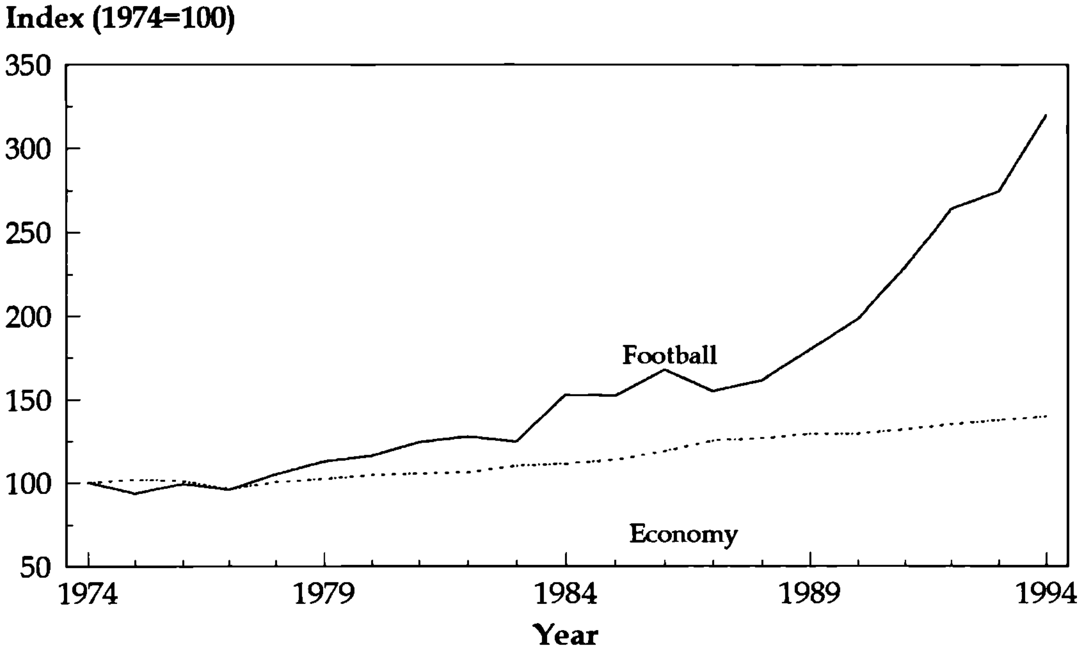
Figure 4.3: Players, Professionals & Management Tenure



Annual average from sample of 39 clubs

Figure 4.4 compares the trend in real wages in the economy as a whole and in football. It highlights the rapid increase in average wages in football well above the national average growth most notably after 1986. The increasing wages coincide with an increase in turnover as depicted in Figure 4.2, indicating that much of the value of football's recent post-Heysel popularity has been accumulated by the players and team managers.

Figure 4.4: Football and Economy Wide Real Average Wages



Economy real per employee annual wages CSO (various)
Football per player annual wages from sample of 39 clubs

4.5. Econometric Approach

4.5.1. Methodology

There have been two basic approaches in the econometric literature to estimating production functions. The first technique developed was deterministic production functions whereby all the stochastic element of the model is contained in the efficiency term.

For example using the linear production function below, where there are N firms indexed by i, each producing y_i output using x_i inputs. β are estimated coefficients and ε the error term.

$$y_i = \alpha + \beta' x_i + \varepsilon_i .....(4.10)$$

In the simplest case the equation can be estimated be OLS and as shown by Greene (1980) the constant term can be consistently estimated by simply shifting the least squares line upwards sufficiently such that the largest residual is zero. Thus the corrected constant term is $\alpha + \max_i(\varepsilon_i)$. The resulting efficiency measures $(\hat{u}_i)$ of the ith firm is then:

$$\hat{u}_i = \varepsilon_i - \max_i(\varepsilon_i) .....(4.11)$$

As a result of the treatment of the error term there is no room for random events that might affect productivity but are out of the control of the producer nor room for model misspecification. The stochastic production frontier technique, developed amongst others by Aigner, Lovell and Schmidt (1977) overcomes these problems by decomposing the error term as follows:

$$y_t = \alpha + \beta' x_t + \varepsilon_t$$
$$= \alpha + \beta' x_t + v_t - u_t \text{ ....(4.12)}$$

where $u_t \geq 0$ and is the efficiency term and v_t is the unrestricted error term.

The above discussion has been framed in terms of output. The same discussion can be had in terms of cost, i.e. there is a output maximusation frontier given a level of cost but also a cost minimisation frontier given a level of output. Distance from the output frontier will only give technical efficiency while distance from the least cost technique will take account of both allocative and technical efficiency (see Figure 4.1 above). This is because any sub-optimal technical or allocative input vectors will show up in the costs. However, this study focuses on technical efficiency and as such no cost frontiers are estimated.

Taking advantage of panel nature of the data set employed can improve the estimation.

Expanding the production function described in 4.12 above for use with a panel data set we get:

$$y_{it} = \alpha + \beta' x_{it} + v_{it} - u_{it} \text{ ....(4.13)}$$

Here there are N firms and T observations on each. If u_{it} and v_{it} are independent over time as well as over individual clubs then there is no benefit to having a panel data set. However this is unlikely. If we make the assumption that inefficiency is constant over time i.e. $u_{it} = u_i$ then the model becomes

$$y_{it} = \alpha + \beta' x_{it} + v_{it} - u_i \text{ ....(4.14)}$$

and panel data estimation techniques can be used. This assumption of constant inefficiency over time does not seem unreasonable given the type of institutional inefficiency that is being investigated.

Using a fixed effects model (FEM) the u_i's are treated as firm specific constants and the model may be estimated by ordinary least squares. The model becomes:

$$y_{it} = \alpha_i + \beta' x_{it} + v_{it} \text{ ....(4.15)}$$

Estimation produces a set of firm specific constants (α_is); these are then used to estimate the technical efficiencies. The estimated technical efficiency used in this study is calculated at the club means as follows:

$$\hat{TE}_i = \frac{y_i^*}{y_{iMAX}^*} = \frac{\alpha_i + \beta' \bar{x}_i}{\alpha_{MAX} + \beta' \bar{x}_i} \quad ...(4.16)$$

For the most efficient club (i.e. the club with the largest α_i) this measure will be 1. The formulation has similarities with the deterministic production functions described earlier. Here as with the deterministic function the most efficient club is used as a benchmark and all the α_is are assumed to be related to efficiency. Using this method means that the assumption that firm inefficiencies are uncorrelated with input levels and have a normal distribution (required in non-panel estimation techniques) can be dispensed with. However, problems with this model occur when there are time invariant firm attributes (e.g. capital stock) which affect efficiency. If these are included then the model cannot be estimated as above. If these factors are important but are omitted then they will reappear in the fixed effects indicating inefficiency. However in the specifications estimated there are no real time invariant effects and thus this model may be appropriate.

If the assumption of independence of inefficiencies and input levels can be maintained then a random effects model (REM) which directly estimates equation 4.14 might be more suitable. The advantage of REM is that it allows time-invariant firm specific attributes to enter the model unlike in the FEM.

It is possible to estimate the REM in two ways, using generalised least squares or maximum likelihood estimation. The maximum likelihood estimator was used as this has efficiency advantages[59]. To actually estimate u_i, an assumption about its distribution is required. As with much of the literature throughout the following it is assumed that u_i has a half normal distribution. Following Battese and Coelli (1988) then for a random effects panel data estimate of 4.14 assuming the half normal distribution, then estimated efficiency:

$$E\left[u_i \middle| \varepsilon_{i1} \ldots \varepsilon_{iT_i}\right] = M_i^* + \sigma_{*i}\{\phi(M_i^*/\sigma_{*i})[\Phi(-M_i^*/\sigma_{*i})]^{-1}\} \quad ....(4.17)$$

where

$$\begin{aligned} M_i^* &= \gamma_i \mu + (1-\gamma_i)(-\bar{\varepsilon}_i) \\ \bar{\varepsilon}_i &= T_i^{-1} \textstyle\sum_{t=1}^{T_i} \varepsilon_{it} \\ \sigma_{*i}^2 &= \sigma_u^2 \gamma_i \\ \gamma_i &= 1/(1+\lambda T_i) \end{aligned} \quad ....(4.18)$$

[59] The log likelihood function is given in Appendix 3.

If the frontier production function is defined in terms of logarithm of production the technical efficiency of the ith firm becomes:

$$E\left[e^{u_i}\left|\varepsilon_{it}\ldots\ldots\varepsilon_{iT_i}\right.\right]=\left\{\frac{\Phi[M_i^*\ \sigma_{*_i}-\sigma_{*_i}]}{\Phi(M_i^*\ \sigma_{*_i})}\right\}\exp(-M_i^*+\tfrac{1}{2}\sigma_{*_i}^2)\ldots.(4.19)$$

There is a further refinement that can be made namely using a two-way effects model which allows relaxation of the assumption of fixed inefficiencies over time. This becomes computationally cumbersome with a reasonably sized panel and its value is questionable[60], thus is not carried out here.

4.5.2. Results

Two formulations of the production function were estimated, a semi-log function and a log-linear function. The home grown player variable was investigated but found to be insignificant in all specifications and thus has been dropped from the results described below. Therefore the two specifications investigated are:

semi-log
ln(performance)=F{ability, players used, professionals, management tenure}.....(4.20)

log linear
ln(performance)=F{ln(ability), ln (players used), ln(professionals), ln(management tenure)} ....(4.21)

Tables 4.3 and 4.4 give the results for the semi-log and log linear model for the OLS model, the fixed effects model (equation 4.15) and the stochastic frontier model assuming a half-normal distribution for u (equations 4.14 and 4.17). For each estimation method there are three specifications combining different combinations of variables from the right hand side of 4.20 and 4.21. In each equation the results for the preferred ability variable per squad member (equation 4.9) are shown.

Before each estimation the skewness of the OLS residuals were checked. Waldman[61] has shown that if the OLS residuals are negatively skewed then the maximum likelihood estimator is simply OLS for the slopes and both σ_v^2 and σ_u^2 equal 0. If the residuals were negatively skewed this would imply that model was not well specified. In each estimation the residuals were not negatively skewed.

Also included in the table are the adjusted R^2, two likelihood ratio statistics and a Hausman statistic. The LR statistic is computed as: $LR=-2(\ln\hat{L}_{res}-\ln\hat{L})$ and is chi-squared distributed with J degrees of freedom where J is the number of restrictions.

---

[60] See Greene (1993).
[61] See Waldman (1982).

The first LR statistic computed compares the fixed effects model to the OLS model. The OLS regressions is the restricted model, i.e. fixed effects equal 0. Thus the statistic is chi-squared distributed with 38 degrees of freedom. High values of this statistic favour the fixed effects model above the OLS model.

The second LR statistic compares the stochastic frontier model against the OLS. Again the OLS equation is the restricted equation with the restriction that $\lambda = 0$. As such the statistic is chi-squared with 1 degree of freedom. High values of the statistic favour the stochastic model versus OLS.

The Hausman statistic[62] compares the fixed effects model versus the random effects model (i.e. the stochastic frontier model estimated via GLS). As the Hausman statistics is based on the inconsistency of GLS it cannot be produced for maximum likelihood estimations. Thus the statistic shown is based on the generalised least squares estimators. However all the coefficient estimates are maximum likelihood as this is more efficient. The Hausman statistic is chi-squared distributed with degrees of freedom equal to the number of independent variables. A high value of the statistic argues in favour of the fixed effects model.

[62] See Hausman (1978)

Table 4.3: Semi-Log Production Functions

Dependent variable = Ln(league success) Sample 39 clubs 1974-1994 T-stats in brackets.

	OLS			Fixed			Stochastic		
Specification	1	2	3	1	2	3	1	2	3
ability per squad member	1 93	1.81	1.77	1.46	1.41	1.39	1.72	1.61	1.58
	(32.81)	(31.24)	(30.35)	(15.20)	(15.48)	(15.32)	(41.99)	(37.58)	(32.14)
no. of professionals	0.07	0.08	0.08	0.04	0.05	0.05	0.05	0.06	0.064
	(11.01)	(13.34)	(13.31)	(5.98)	(8.20)	(8.15)	(11.92)	(14.18)	(14.11)
no. players used in league		-0.08	-0.07		-0.07	-0.07		-0.08	-0.08
		(9.02)	(-8.46)		(-9.04)	(8.58)		(14.46)	(-12.82)
management tenure			0.05			0.05			0.05
			(4.04)			(4.36)			(4 45)
constant	-3.23	-1.61	-1.80				1.90	-0.36	-0.60
	(21.11)	(-6.96)	(-7.71)				(-13.43)	(2.32)	(-3.86)
$\lambda \quad \sigma_u / \sigma_v$							1 18	1.20	1.16
							(2 84)	(2.78)	(2.74)
σ_v							0 57	0.51	0.50
							(29.74)	(27.07)	(26.77)
Log Likelihood	-1040.06	-1001.11	-993.00	-905.30	-864.38	-854.49	-988.18	944 34	933.67
Adjusted R^2	0.70	0.73	0.74	0.78	0.80	0 80			
LR Fixed vs OLS				269 52	273.46	277.01			
LR Stochastic vs OLS							103.76	113.54	118.66
Hausman (Fixed vs Random)				22.28	19.34	18 66			

Table 4.4: Log Linear Production Functions

Dependent variable = Ln(league success) Sample 39 clubs 1974-1994. T-stats in brackets.

	OLS			Fixed			Stochastic		
Specification	1	2	3	1	2	3	1	2	3
log ability per squad member	2.13	2.00	1.96	1.77	1.69	1.66	2.00	1.88	1.85
	(37.85)	(36.27)	(35.60)	(18.99)	(18.71)	(18.71)	(34.14)	(31.23)	(31.49)
log no. of professionals	1.86	2.19	2.17	1.34	1.64	1.62	1.64	1.93	1.90
	(11.89)	(14.32)	(14.38)	(7.76)	(9.64)	(9.67)	(13.07)	(15.27)	(15.02)
log no. players used in league		-1.74	-1.61		-1.47	1.37		-1.62	-1.51
		(-9.30)	(-8.65)		(-8.25)	(7.73)		(-12.5)	(-11.41)
log management tenure			0.18			0.18			0.19
			(4.75)			(5.10)			(5.36)
constant	-5.20	-0.82	-1.31				-3.93	0.13	-0.30
	(-10.08)	(-1.21)	(-1.93)				(-9.50)	(0 23)	(-0.54)
$\lambda \quad \sigma_u / \sigma_v$							0 93	0 81	0.82
							(2.64)	(2 70)	(2.76)
σ_v							0.51	0.47	0.45
							(22.95)	(23.35)	(22.55)
Log Likelihood	-971.01	-929.73	-918.52	-855.66	-821.30	-807.80	-937.57	-897.68	-884.01
Adjusted R^2	0.75	0.77	0.78	0.81	0.82	0.82			
LR Fixed vs OLS				230.72	216.87	221.45			
LR Stochastic vs OLS							66.88	64.10	69.02
Hausman (Fixed vs Random)				11.99	16.16	15.13			

The results in Tables 4.3 and 4.4 are as expected. All the variables are significant in each formulation and are of the correct sign. As expected playing and team managerial ability is strongly correlated with success. The larger the squad of professional players the more successful the club as shown by the positive and significant finding on the number of professionals variable.

The player used variable is also significant in all specifications and of the correct sign. It was postulated that a more successful team would use fewer players throughout the season. The fact that the introduction of the playing variables does little to alter the ability and number of professionals variables is encouraging. This supports the argument that number of players used is picking the random effect of injuries on a team.

The management tenure variable is also significant and positive, supporting the postulate that there are advantages to consistency in management. However as argued above short management tenure is really a form of inefficiency that should be included in the efficiency estimates and thus specification 3 of the models will not be used in calculating the efficiency of each club.

The coefficients and significance of the variables do not change considerably with the three differing estimation methods (i.e. OLS, fixed effects and stochastic frontier). The Likelihood Ratio tests in all 3 specifications of the model, both the semi-log and log linear versions, support strongly both the fixed effects model and the stochastic effects model over the OLS. This is backed up in the stochastic case by the significance λ. The Hausman statistics support the fixed effects model over the random effects model but as discussed earlier this test is of only limited validity.

Both the log linear and semi-log versions of the model seem to work well with the log linear version giving a slight advantage on adjusted R^2.

Given the above results the most favoured regressions for calculating the efficiency of football clubs are the log linear fixed effects and stochastic frontier specification 2. This model does not include management tenure as it is thought this should be included in the inefficiency component of clubs.

As a check on the underlying relationship identified above the favoured specification was calculated for the 1974-1989 panel, a shorter time period but including an additional 8 clubs. The results are contained in the Table 4.5 below. Encouragingly, while there are differences in the estimated coefficients the underlying relationship seems to be supported.

Table 4.5 Log Linear Specification 2 With Different Sample Lengths

Dependent Varaible Ln (home success). T stats in brackets.

	1974-1994 39 Clubs			1974-1989 47 Clubs		
Specification	OLS	Fixed	Stochastic	OLS	Fixed	Stochastic
log ability per squad member	2.00	1.69	1.88	2.13	1.53	1.91
	(36.27)	(18.71)	(31.23)	(35.60)	(15.15)	(42.10)
log no. of professionals	2 19	1.64	1.93	2.11	1.21	1.74
	(14.32)	(9.64)	(15.27)	(12.98)	(6.50)	(13.49)
log no. players used in league	-1.74	-1.47	-1.62	-1.89	-1.46	-1.68
	(-9.30)	(-8.25)	(-12.5)	(-9.0)	(-7.36)	(-9.75)
constant	-0.82		0.13	-0.27		0.86
	(-1.21)		(0.23)	(-0.343)		(1.30)
λ σ_u / σ_v			0.81			1.22
			(2.7)			(2.38)
σ_v			0.47			0.49
			(23.35)			(26.26)
Log Likelihood	-929.73	-821.30	-897.68	-891.09	-758.98	-861.70
Adjusted R^2	0.77	0.82		0.77	0.82	
LR Fixed vs OLS		216.87			264.23	
LR Stochastic vs OLS			64.10			58.78
Hausman (fixed vs Random)		16.16			43.09	

Using equations 4.16 and 4.19 and the fixed and stochastic frontier estimates of the favoured regression (specification 2 log linear) the efficiency of each club was calculated. The results are contained in Tables 4.6 and 4.7 below.

Table 4.6: Efficiency

	Fixed Effects		Stochastic	
	Effic	Rank	Effic	Rank
Arsenal	0.426	8	0.702	18
Aston Villa	0.502	3	0.846	6
Barnsley	0.333	16	0.744	12
Birmingham City	0.304	19	0.605	21
Blackburn Rovers	0.378	14	0.773	10
Bolton W	0.227	29	0.490	30
Brentford	0.238	26	0.569	24
Bristol R	0.474	4	0.914	3
Burnley	0.143	38	0.318	38
Bury	0.206	31	0.506	28
Cambridge	0.297	22	0.704	17
Coventry C	0.408	10	0.749	11
Everton	0.469	5	0.781	7
Huddersfield	0.197	32	0.451	32
Hull C	0.325	18	0.727	13
Leeds	0.393	12	0.705	16
Leicester	0.372	15	0.724	14
Liverpool	1.000	1	0.959	1
Luton	0.330	17	0.641	20
Man Utd	0.447	7	0.716	15
Mansfield	0.179	36	0.445	33
Newcastle	0.295	23	0.551	25
Oldham	0.457	6	0.880	4
Peterborough	0.187	34	0.461	31
Plymouth	0.255	25	0.571	23
Preston	0.217	30	0.502	29
Reading	0.185	35	0.437	35
Rotherham U	0.287	24	0.682	19
Scunthorpe	0.090	39	0.238	39
Sheff U	0.380	13	0.776	8
Sheff W	0.303	20	0.593	22
Shrewsbury	0.413	9	0.878	5
Southampton	0.583	2	0.914	2
Southend	0.229	27	0.551	26
Swindon	0.193	33	0.430	36
Spurs	0.301	21	0.512	27
WBA	0.403	11	0.774	9
West Ham	0.227	28	0.417	37
Wrexham	0.177	37	0.438	34
Average	0.329		0.633	

Table 4.7 Efficiency By Quartile

	Fixed Effects	Stochastic Frontier
Q1 0.76-1.00	Liverpool	Aston Villa Blackburn, Bristol R, Everton, Liverpool, Oldham, Sheff U, Shrewsbury, Southampton, WBA
Q2 0.51-0.75	Aston Villa, Southampton	Arsenal, Barnsley, Birmingham Brentford, Bury, Cambridge, Coventry, Hull, Leeds, Leicester, Luton, Man Utd, Newcastle, Plymouth, Preston, Rotherham, Sheff W, Southend, Spurs
Q3 0.26-0.50	Arsenal, Barnsley, Birmingham, Blackburn, Bristol R, Cambridge, Coventry, Everton, Hull, Leeds, Leicester, Luton, Man Utd, Newcastle, Oldham, Plymouth, Rotherham, Sheff U, Sheff W, Shrewsbury, Spurs, WBA	Bolton, Burnley, Huddersfield, Mansfield, Peterborough, Reading, Swindon, West Ham, Wrexham
Q4 0.00-0.25	Bolton, Brentford, Burnley, Bury, Huddersfield, Mansfield, Peterborough, Preston, Reading, , Scunthorpe, Southend, Swindon, West Ham, Wrexham	Scunthorpe

In both formulations, despite its high level of inputs Liverpool is the most efficient club. This is in line with dell'Osso and Szymanski (1991) findings. In the fixed effects model Liverpool is dominant with the second most efficient club being just over half as efficient. The fixed effects model gives all clubs, with the exception of Liverpool, a smaller efficiency score than the stochastic model as reflected in the averages. It seems that Liverpool's fixed effect is strongly dominant and as a result of the deterministic nature of the model the efficiency of other clubs is underestimated. However, when comparing clubs' ranking in the two models, the results are similar, with a rank correlation coefficient[63] of 0.944.

Concentrating on the stochastic results the clubs in the top quartile include a mix of clubs consistently in the top division throughout the sample (e.g. Liverpool, Everton), clubs only recently in the top division (Blackburn) and also clubs consistently in the lower divisions (e.g. Shrewsbury). The more inefficient clubs are concentrated in the lower divisions although West Ham are in the third quartile.

[63] Rank Correlation Coefficient $= 1 - \frac{6\sum d^2}{n(n-1)}$, where d= difference in ranks.

4.6. Data Envelopment Analysis (DEA)

4.6.1. Methodology

Data Envelopment Analysis (DEA) is a linear programming method capturing Farrell's (1957) approach to efficiency measurement. DEA was first effectively formulated by Charnes, Cooper and Rhodes (1978, 1979) and since has been put to many uses for example in prisons (Ganley & Cubbin 1987), US agriculture (Lee and Somwaru 1993) and hospitals (Burgess and Wilson 1993).

There are several versions of DEA that can be estimated, with input minimisation or output maximisation and constant or variable returns to scale being the main options available. The choice of constant or variable returns to scale affects both the formulation of the model and the resulting production frontier. The choice of an output or input focus of the model does not affect the estimated production frontier but does affect the resultant efficiency measurement.

Initially the discussion below will focus on a constant returns to scale model and then extended this to a variable returns model. Throughout the empirical work in order to aid comparison with the econometric approach output maximisation only has been estimated for both constant and variable returns to scale.

In the simplest case efficiency can be defined as output divided by input. However, organisations often have multiple incommensurate inputs and outputs and thus efficiency can be better defined as the weighted sum of outputs divided by the weighted sum of inputs.

One way to set these weights is by use of DEA. If you have Z decision making units (DMUs)[64] each using the same set of inputs to produce the same outputs then efficiency can be measured as the maximum of the ratio of weighted outputs to weighted inputs subject to constraints reflecting the performance of other DMUs. This can be put more formally: if there are q outputs Y_i where i=1..q and m inputs X_k, k=1..m then DEA:

$$\underset{V_i W_k}{Max} \frac{\sum_{i=1}^{q} V_i Y_{ip}}{\sum_{k=1}^{m} W_k X_{kp}} \quad \text{................(4.22)}$$

subject to Z constraints $0 \leq \frac{\sum_{i=1}^{q} V_i Y_{ic}}{\sum_{k=1}^{m} W_k X_{kc}} \leq 1$ where c=1...p,...Z

[64] As coined by Charnes et al. (1978).

and V_i, W_k >0 for all i and k. These constraints ensure that the optimal weights do not imply an efficiency score greater than unity for any DMU.

The program is computed for each DMU giving Z sets of optimal weights. The formulation above is termed the fractional program but cannot be solved because it has intractable non-linear and non convex properties. Charnes et al. (1978) suggested the following transformation of which only the output maximisation is shown. The linear program for the pth DMU is:

$$\underset{V_i,W_k}{MAX} \sum_{i=1}^{q} V_i Y_{ip} \quad \text{..............(4.23)}$$

subject to:

$$\sum_{i=1}^{q} V_i Y_{ic} \leq \sum_{k=1}^{m} W_k X_{kc} \quad \text{, c=1,...p,...Z}$$

$$\sum_{k=1}^{m} W_k X_{kp} = 1$$

and V_i, W_k >0 for all i and k.

The above program constrains the weighted sum of inputs to be unity and then maximises the weighted sum of outputs at the pth DMU choosing appropriate values of V_i, and W_k. In a subsequent article Charnes et al. (1979) restricted the weights such that:

$$\begin{matrix} W_k > \varepsilon, k = 1,....,m \\ V_i > \varepsilon, i = 1,....,q \end{matrix} \quad \text{......................(4.24)}$$

Where ε is an infinitesimal constant of the order 10^{-6}. These additional constraints were added as the 1978 formulation allowed a unity efficiency score for some inefficient units.

There is a more tractable dual formulation to the above primal. The primal constraints are indexed on all Z DMUs while the dual constraints (set out below) are indexed on inputs and outputs and sum over DMUs. The number of inputs and outputs are never likely to exceed the number of DMUs.

The dual of the output maximisation problem (4.23) is the input minimisation problem set out below:

$$\underset{\lambda_c}{Min}\, h_p - \varepsilon\left(\sum_{k=i}^{m} S_k + \sum_{i=1}^{q} S_i\right) \ldots\ldots\ldots\ldots(4.25)$$

subject to

$$X_{kp} \cdot h_p - S_k = \sum_{c-1}^{z} X_{kc}\lambda_c,\ k = 1,\ldots m$$

$$Y_{ip} + S_i = \sum_{c-1}^{z} Y_{ic}\lambda_c,\ i = 1,\ldots q$$

$\lambda_c \geq 0$ c=1,...,p,...,Z (weights on branches)
$S_k \geq 0$ k=1,...,m (input slacks)
$S_i \geq 0$ i=1,...,t (output slacks)

Again ε is an infinitesimal constant. Note that the pth DMU is relatively efficient if and only if :

$$h_p^* = 1\ with\ S_k^* = S_i^* = 0, for\ all\ k\ and\ i \ldots\ldots.(4.26)$$

where * denotes the optimal values of the variables in the dual program.

Figure 4.5: Input Minimisation

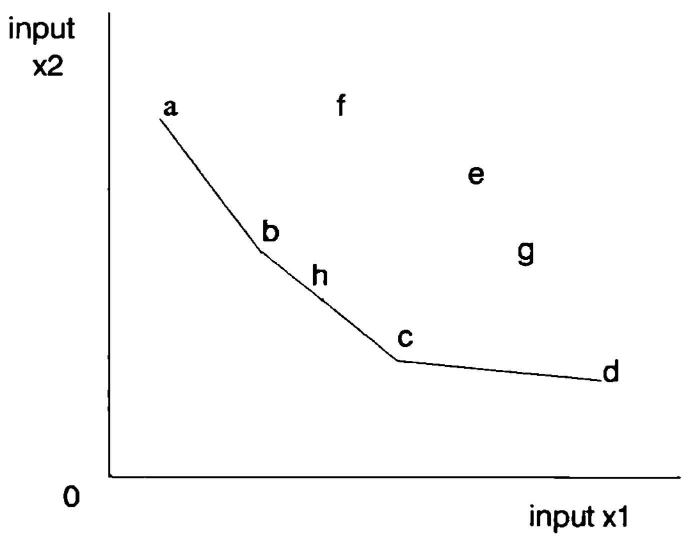


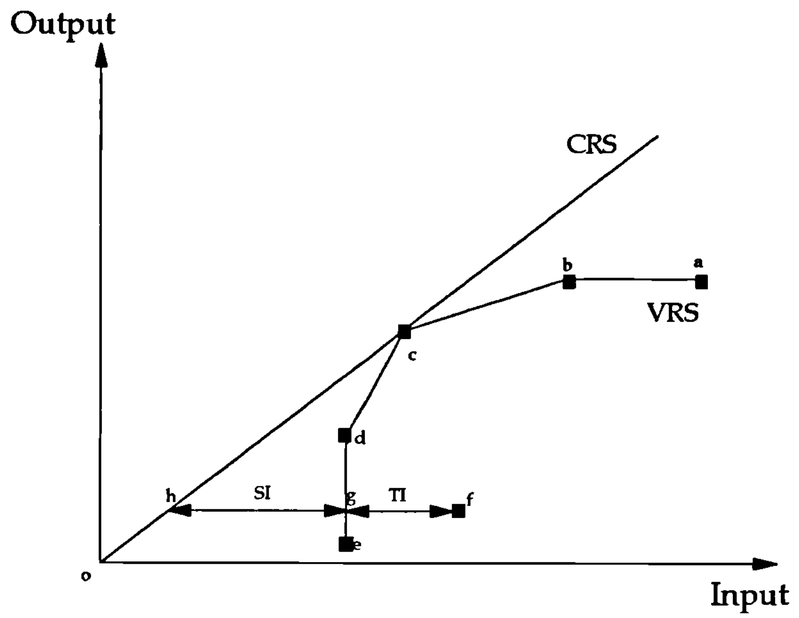
Figure 4.5 represents the input minimisation program described above. There are seven DMUs a-g producing a single output with two inputs x_1 and x_2. Four DMUs a, b, c, d define the frontier which implies they are best practice, i.e. no other DMU or linear combination of DMUs can be identified producing the same level of output for less of both or either inputs. These DMUs satisfy the conditions set out in equation 4.26 above, i.e. they have unity efficiency ratios and zero slacks. DMUs e, f, g are inefficient in that is it possible to find a DMU or linear combination of DMUs which while producing the same output use less inputs. For example looking at DMU e, a linear combination of b and c (denoted its peers) produce a similar output for fewer inputs. This inefficiency is reflected in e's efficiency ratio oh/oe.

The linear programme 4.25 above can easily be adjusted to allow variable returns to scale. The programme is identical, except for an additional constraint:

$$\sum_{c\ 1}^{Z} \lambda_c \leq 1 \text{ .....(4.27)}$$

This additional constraint ensures that the frontier is composed of multiple convex linear combinations of best practice where dominance is now more weakly defined to include regions of increasing and decreasing returns. This is illustrated in Figure 4.6. In this production process there is a single output and a single input. Points a-f are different companies with different productivities. If we assume constant returns to scale then the DEA frontier would be oc while under variable returns to scale it becomes abcde. Company c is the most productive company with companies a and b subject to decreasing returns to scale and companies d, e and f increasing returns. Under Constant Returns to Scale (CRS) all these companies except c would be considered inefficient because they are not on the frontier oc but under Variable Returns to Scale (VRS) a,b,d and e are also on the frontier and would thus be considered efficient. Using variable returns gives pure technological efficiency while constant returns looks at technological and scale efficiency. This can be illustrated by looking at company f. Under CRS the reference point is h, while with VRS it is g and thus inefficiency can be decomposed into fg which is pure technical inefficiency and gh which is scale inefficiency. Thus the number of inefficient DMUs under varying returns will be less than under constant returns.

Figure 4.6: Constant and Variable Returns To Scale



Rangan et al. (1988) developed a measure of scale inefficiency simply based on the ratio of the constant return and variable return to scale efficiency scores i.e.

scale efficiency = CRS efficiency / VRS efficiency ....(4.28)

Banker (1984) shows that the sum of the weights on branches in the constant return to scale version can be used as a local indicator of returns to scale. If the Banker indicator =1 then there are constant returns to scale, if Banker >1 then there are decreasing returns to scale and if Banker<1 then there are increasing returns to scale.

An alternative intermediate case, non increasing returns to scale estimation, can be carried out by changing the additional constraint under varying returns to:

$$1 = \sum_{c\ 1}^{z} \lambda_c \text{ ...(4.29)}$$

In dealing with a panel data DEA offers a number of approaches. In this chapter, two different approaches have been tried. Firstly a DEA calculation has been carried on a single set of sample means for each club. Secondly, a separate DEA calculation has been carried out for each year of the sample. Averaging the annual results and comparing this with the sample means estimation results in an almost identical ranking (rank correlation coefficient of 0.99) although the efficiency scores are lower with the annual approach[65]. The results for just the mean estimations are presented below.

[65] See Appendix 4.

4.6.2. Results

Two model specifications were estimated[66]. The first uses the same inputs and outputs as the favoured econometric specification and is calculated for direct comparison. That is the inputs are ability (proxied by wages), number of professionals and number of players used; output is league performance. The second specification utilises the feature of DEA that allows multiple inputs and outputs to be modelled. The inputs are identical to the first specification but performance in the FA Cup and League Cup are added to the outputs.

Tables 4.8 and 4.11 give the detailed results for the two specification. As well as efficiency and rank there is a list of the peer group and where appropriate the number of times a club is cited in a peer group. As noted above the peer group give the linear combinations of clubs to which the efficiency of the club under scrutiny is compared (see Figure 4.5). The number of times a club is cited can be used as an indication of how important the club is as an example of best practice although this doesn't take account of outliers. Also included in the constant returns to scale estimation is a column giving Bankers indication of returns to scale.

The constant returns to scale version of Specification 1 (Table 4.8) has one efficient club, the whole estimation being dominated by Liverpool which is cited 38 times. The second most efficient club is Manchester United which is less than half as efficient as Liverpool. The majority of clubs are in the bottom quartile of efficiency as Table 4.9 shows. Every club's Banker measure suggests local increasing returns to scale

[66] DEA estimates were calculated using a program developed by Professor John Cubbin.

Table 4.8: DEA Specification 1

Inputs = ability, no. players, professionals. Outputs= League performance

		Constant Returns to Scale				Variable Returns to Scale		
		Efficiency	Peer Group (times cited)	Rank	Banker	VRS	Peer Group (times cited)	Rank
1	Arsenal	0.397	18	4	0.916	0.474	18, 24	11
2	Aston Villa	0.359	18	5	0.703	0.406	18 29	13
3	Barnsley	0.086	18	20	0.288	0.267	18 28 29	19
4	Birmingham C	0.092	18	18	0.552	0.119	18 29	35
5	Blackburn R	0.169	18	14	0.472	0.441	10 18 24	12
6	Bolton W	0.061	18	25	0.401	0.113	10 18 29 32	37
7	Brentford	0.036	18	33	0.264	0.152	18 29	32
8	Bristol R	0.069	18	21	0.255	0.355	18 28 29	16
9	Burnley	0.078	18	24	0.387	0.137	18 29	34
10	Bury	0.029	18	37	0.255	1.000	(8)	1
11	Cambridge	0.060	18	26	0.236	0.532	18 29	9
12	Coventry C	0.156	18	10	0.642	0.184	18 29	25
13	Everton	0.407	18	3	0.855	0.582	18 24	8
14	Huddersfield	0.040	18	30	0.311	0.102	18 29	39
15	Hull C	0.064	18	22	0.288	0.197	18 29	24
16	Leeds	0.339	18	6	0.750	0.375	10 18 29	15
17	Leicester	0.129	18	16	0.486	0.181	18 29	28
18	Liverpool	1.000	(38)	1		1.000	(32)	1
19	Luton	0.123	18	17	0.547	0.159	18 29	31
20	Man Utd	0.486	18	2	0.892	0.503	18 29	10
21	Mansfield	0.027	18	38	0.255	0.378	10 18 29	14
22	Newcastle	0.152	18	12	0.613	0.184	18 29	26
23	Oldham	0.094	18	19	0.373	0.184	10 18 29 32	27
24	Peterborough	0.029	18	35	0.264	1.000	(5)	1
25	Plymouth	0.050	18	28	0.326	0.115	18 28 29	36
26	Preston	0.044	18	31	0.283	0.142	18 29	33
27	Reading	0.032	18	34	0.316	0.168	10 18 24	30
28	Rotherham U	0.047	18	29	0.226	1.000	(3)	1
29	Scunthorpe	0.013	18	39	0.217	1.000	(27)	1
30	Sheff U	0.132	18	15	0.396	0.226	18 29	20
31	Sheff W	0.172	18	13	0.557	0.220	18 29	21
32	Shrewsbury	0.068	18	23	0.250	1.000	(3)	1
33	Southampton	0.259	18	7	0.543	0.336	18 29	17
34	Southend	0.040	18	32	0.250	1.000	(0)	1
35	Swindon	0.051	18	27	0.373	0.104	10 18 29 32	38
36	Spurs	0.207	18	8	0.939	0.210	18 29	23
37	WBA	0.206	18	9	0.491	0.288	18 29	18
38	West Ham	0.125	18	11	0.778	0.174	18 24	29
39	Wrexham	0.035	18	36	0.274	0.217	10 18 29	22
	Average	0.153				0.390		

Table 4.9: Efficiency By Quartile (Specification 1)

	Constant Returns	Variable Returns
Efficient	Liverpool	Bury, Liverpool, Peterborough, Rotherham, Scunthorpe, Shrewsbury, Southend
Q1 0.76-1.00		
Q2 0.51-0.75		Cambridge, Everton, Man Utd
Q3 0.26-0.50	Arsenal, Aston Villa, Everton, Leeds, Man Utd, Southampton	Arsenal, Aston Villa, Barnsley, Blackburn, Leeds, Mansfield, Southampton, WBA
Q4 0.00-0.25	Barnsley, Birmingham, Blackburn, Bolton, Brentford, Bristol R, Burnley, Bury, Cambridge, Coventry, Huddersfield, Hull, Leicester, Luton, Mansfield, Newcastle, Oldham, Peterborough, Plymouth, Preston, Reading, Rotherham, Scunthorpe, Sheff U, Sheff W, Shrewsbury, Southend, Swindon, Spurs, WBA, West Ham, Wrexham	Birmingham, Bolton, Brentford, Bristol R, Burnley, Coventry, Huddersfield, Hull, Leicester, Luton, Newcastle, Oldham, Plymouth, Preston, Reading, Sheff U, Sheff W, Swindon, Spurs, West Ham, Wrexham

As expected there are more efficient clubs in the variable returns to scale estimation of specification 1 (Table 4.8). Six new clubs join Liverpool as efficient and interestingly not all the clubs are from the higher divisions (e.g. Shrewsbury). Liverpool is still dominant, being cited on 32 occasions in clubs' peer group. There is an extreme difference between the CRS and VRS models' ranking, indeed the rank correlation coefficient is 0. Some clubs experiencing large changes in efficiency are given in Table 4.10 below.

Table 4.10: Major Differences Between CRS and VRS

	CRS Rank	CRS Effic	VRS Rank	VRS Effic	Rangan Scale Inefficiency	Relative Wages
Bury	37	0.029	1	1	0.029	0.54
Peterborough	35	0.029	1	1	0.029	0.56
Rotherham	29	0.047	1	1	0.047	0.48
Scunthorpe	39	0.013	1	1	0.013	0.46
Shrewsbury	23	0.068	1	1	0.068	0.53
Southend	32	0.040	1	1	0.040	0.53

The changes under VRS are dramatic with several clubs moving from being the most inefficient clubs to the most efficient. The results imply that these clubs' inefficiency is purely due to scale inefficiency as shown by Rangan's scale inefficiency measure. The major differences in results between VRS and CRS is not unusual[67] but does seem rather implausible. This illustrates one of the problems with DEA, namely it gives the companies the benefit of the doubt. That is, it will give a company an efficiency score of one unless it can find a linear combination of two other companies which is more efficient. This can be a particular problem when DMU's are outliers. Thus more weight should be placed on evidence of inefficiency than on evidence of efficiency. All the clubs in Table 4.10 are outliers as evidenced by the expenditure on wages, thus it would seem sensible to treat the findings of efficiency for these clubs with caution.

Tables 4.11 and 4.12 contain the results for DEA specification 2 which adds F.A. Cup and League Cup performance to the outputs. Adding more variables to the DEA specification, as expected finds clubs more efficient as evidenced by the average efficiency in specification 2 of 0.772 compared with 0.153 under specification 1. DEA specification 2 finds 6 clubs efficient under constant returns to scale and 14 clubs efficient under variable returns. Liverpool while still efficient are less dominant under specification 2 than under specification 1 in terms of the number of times the club is cited as best practice. The Banker test finds that 20 clubs have increasing returns to scale, 12 have decreasing returns to scale and 1 club Bristol Rovers has constant returns to scale.

[67] See for example Ganley & Cubbin (1992)

Table 4.11: DEA Specification 2

Inputs= ability, no. of players, professionals. Outputs= League performance, FA Cup performance and league cup performance.

		Constant Returns to Scale				Variable Returns to Scale		
		Efficiency	Peer Group (times cited)	Rank	Banker	VRS	Peer Group (times cited)	Rank
1	Arsenal	0.876	18	13	0.916	0.949	5 13 18 24	18
2	Aston Villa	1.000	(3)	1		1.000	(2)	1
3	Barnsley	0.984	24 32 33	7	1.033	0.985	24 32 33	15
4	Birmingham C	0.845	20 33	15	0.926	0.855	18 32 33	21
5	Blackburn R	0.767	20 33	20	0.764	1.000	(1)	1
6	Bolton W	0.750	32 33	23	0.919	0.780	18 24 32	23
7	Brentford	0.490	24 32 33	38	0.972	0.490	18 24 32 33	38
8	Bristol R	0.752	18 32	22	1.000	0.753	11 18 32	27
9	Burnley	0.715	24 32 33	27	1.015	0.715	24 32 33	30
10	Bury	0.627	8 24 32 33	30	0.941	1.000	(2)	1
11	Cambridge	0.940	24 32 33	9	0.916	1.000	(1)	1
12	Coventry C	0.722	18 20 33	26	0.888	0.735	18 24 32	29
13	Everton	0.916	18	12	0.855	1.000	(2)	1
14	Huddersfield	0.583	24 32 33	33	1.056	0.583	24 32 33	37
15	Hull C	0.714	24 32 33	28	1.001	0.714	24 32 33	31
16	Leeds	0.588	2 18 24	32	1.020	0.590	2 18 24	35
17	Leicester	0.747	20 32 33	24	1.090	0.752	32 33	28
18	Liverpool	1.000	(11)	1		1.000	(17)	1
19	Luton	0.800	18 24 33	16	1.017	0.801	18 24 33	22
20	Man Utd	1.000	(7)	1		1.000	(2)	1
21	Mansfield	0.523	24 32 33	35	0.950	0.587	10 18 24 29 32	36
22	Newcastle	0.627	18 20 33	31	1.069	0.634	18 20 33	33
23	Oldham	0.920	24 32 33	10	0.959	0.938	18 24 32 33	19
24	Peterborough	1.000	(23)	1		1.000	(19)	1
25	Plymouth	0.630	24 32 33	29	1.020	0.630	24 32 33	34
26	Preston	0.488	24 32 33	39	1.003	0.488	24 32 33	39
27	Reading	0.499	24 32 33	37	0.886	0.675	10 18 24 32	32
28	Rotherham U	0.746	3 24 32 33	25	0.836	1.000	(0)	1
29	Scunthorpe	0.538	27 24 32	34	0.860	1.000	(1)	1
30	Sheff U	0.760	24 32 33	21	1.082	0.764	32 33	26
31	Sheff W	0.962	2 18 24	8	1.053	0.970	2 18 24	17
32	Shrewsbury	1.000	(21)	1		1.000	(18)	1
33	Southampton	1.000	(24)	1		1.000	(15)	1
34	Southend	0.509	24 32 33	36	0.884	1.000	(0)	1
35	Swindon	0.920	2 24	11	0.951	1.000	(0)	1
36	Spurs	0.775	18 20	18	0.969	0.777	18 20 33	25
37	WBA	0.777	18 32 33	17	0.957	0.778	18 24 32 33	24
38	West Ham	0.856	18 20	14	0.810	0.982	13 18 24	16
39	Wrexham	0.775	24 32 33	19	0.938	0.886	18 24 32	20
	Average	0.772				0.840		

Table 4.12: Efficiency By Quartile (Specification 2)

	Constant Returns	Variable Returns
Efficient	Aston Villa, Liverpool, Man U, Peterborough, Shrewsbury, Southampton	Aston Villa, Blackburn, Bury, Cambridge, Everton, Liverpool, Man Utd, Peterborough, Rotherham, Scunthorpe, Shrewsbury, Southampton, Southend, Swindon
Q1 0.76-1.00	Arsenal, Barnsley, Birmingham, Blackburn, , Cambridge, Everton, Luton, Oldham, Sheff U, Sheff W, Swindon, Spurs, WBA, West Ham	Arsenal, Barnsley, Birmingham, Bolton, Bristol R, Luton, Oldham, Sheff U, Sheff W, Spurs, WBA, West Ham, Wrexham
Q2 0.51-0.75	Burnley, Bury, Coventry, Huddersfield, Hull, Leeds, Leicester, Mansfield, Newcastle, Plymouth, Rotherham, Scunthorpe, Southend	Burnley, Coventry, Huddersfield, Hull, Leeds, Leicester, Mansfield, Newcastle, Plymouth, Reading
Q3 0.26-0.50	Brentford, Preston, Reading	Brentford, Preston
Q4 0.00-0.25		

The difference between the CRS and VRS version of specification 2 is not as marked as with specification 1 (see Table 4.13 below). Despite the difference in efficiency between specifications the ranking of clubs is similar for both CRS and VRS versions of the model.

Table 4.13: Rank Correlation Coefficients

CRS: Spec 1vs Spec2	0.51
VRS: Spec 1 vs Spec2	0.58
Spec 1: CRS vs VRS	0.00
Spec 2: CRS vs VRS	0.50

4.7. Conclusions

The econometric model estimated was successful in that all the variables were significant and of the expected sign except for home grown players. It was the first time that true inputs in to the football production process have been modelled and not intermediate outputs. Wages were used as a proxy for ability and were found to be strongly related to success. Success was also positively related to the size of the squad and management tenure and negatively related to the number of players used. This indicates that consistency, both in terms of management and the players used, is important for success. It is common for football clubs to change managers as often as every two years. This would be very rare in other industries where the benefits of avoiding change unless necessary is better understood.

The results of the two methods (fixed effect and stochastic frontier) were fairly similar in terms of ranking although the stochastic model gave higher efficiency estimates.

DEA has not been applied to sports economics and has the particular advantage of allowing both multiple inputs and outputs to be investigated. However DEA does not estimate a functional form for the production frontier and gives no opportunity for specification testing. Thus this method has nothing to say about the relevance of the variables used in the model. As expected the variable returns to scale estimation had a larger number of efficient clubs than the constant returns to scale estimation. Also adding variables to the DEA estimation increased the number of efficient clubs. The results highlighted a particular problem with DEA, namely the sensitivity of efficiency to the assumptions about returns to scale. This is particularly marked when there are several outlying observations.

It is useful to compare the DEA Specification 1 results with the econometric methods results which use identical variables. It would be most appropriate to compare the VRS model as no imposition of return to scale is made in the econometric model. However given the scepticism with which the results of the VRS model are being treated both CRS and VRS models are used for comparison. Table 4.14 gives some useful statistics.

Table 4.14: Econometric versus DEA

	Fixed	Stochastic	DEA 1 CRS	DEA 1 VRS
Average Efficiency	0.329	0.633	0.153	0.390
No. of Efficient Clubs	1	0	1	7
Correlation Coefficient				
Fixed vs..		0.94	0.81	0.20
Stochastic vs			0.63	0.27

The average efficiency for clubs is bigger in general under the econometric models although DEA finds more clubs fully efficient particularly under the VRS version. Both econometric methods and DEA with CRS all come out with similar rankings. However, despite this general similarity for certain clubs which method is chosen can have a big effect on ranking for example West Ham is ranked 37 under the stochastic model but 11 under the DEA CRS.

Despite all the uncertainties one club, Liverpool, was found to be efficient under all specifications. Anecdotal evidence suggest that Liverpool does well on the organisation features identified in section 4.4. The organisational structure of the company is well defined and consistent. In the 1960s Bill Shankly put in place a system of recording what happened at each Liverpool match including detailed information about the weather any injuries incurred. Thus there is a knowledge of football contained within the club once an individual team manager has left. A succession of team managers have been promoted from within the club and thus the objectives of club and manager are unlikely to clash. There is no interference on the playing side and the fabled "boot room" where all playing decisions are made is off limits to non playing staff. In addition the culture of Liverpool seems well defined, no one player is bigger than the club and the ultimate aim of the club is playing success.

Another, slightly less fashionable club that is deemed efficient is Southampton. They are ranked second in both econometric specifications and first under DEA specification 2. Again it seems that consistency of management is important. Between 1973 and 1985 (a large part of the sample) they had a single manager (Lawrie McMenemy) who was eventually replaced by an ex-player. Lawrie McMenemy is now Director of Football at Southampton, indicating the consistency of company beliefs, ideals and culture.

The sensitivity of efficiency estimates to specification and methodology with a well defined data set provides a warning for use of efficiency estimates in regulated industries. Regulated companies could drastically alter the view the regulator's view of its efficiency by successfully arguing for the addition of a particular variable or perhaps a slightly different estimation technique. For example if Scunthorpe was a regulated company it would have strong preferences for which of the models discussed above it believed to be true.

4.8. References

Afriat, S.N. (1972). Efficiency Estimation of Production Frontiers. *International Economic Review*. No. 13, October, pp. 568-98.

Aigner, D., Lovell K. & Schmidt, P. (1977). Formulation and Estimation of Stochastic Production Frontier Models. *Journal of Econometrics*. No. 6, pp. 21-37.

Ali, A.I. & Seiford, L.M. (1993). The Mathematical Programming Approach to Efficiency Analysis in *The Measurement of Productive Efficiency: Techniques and Application*. (eds. Fried, H.O., Knox Lovell, C.A. & Schmidt, S.S.). Oxford University Press, New York, Oxford. pp. 120-159.

Bairam, E.I., Howells, J.M. & Turner, G.M. (1990). Production Functions in Cricket: The Australian and New Zealand Experience. *Applied Economics*. No. 22(7), pp. 871-79.

Banker, R (1984). Estimating the Most Productive Scale Size Using Data Envelopment Analysis. *European Journal of Operational Research*. No. 17, pp.35-44,

Battese, G.E. & Coelli, T.J. (1988). Prediction of Firm-Level Technical Efficiencies With A Generalised Frontier Production Function and Panel Data. *Journal of Econometrics*. No 38, pp. 387-399.

Boussofiane, A, Dyson, R.G. & Thanassoulis, E. (1991). Applied Data Envelopment Analysis. *European Journal of Operational Research*. No. 52, pp.1-15.

Burgess, J.F. & Wilson P.W. (1993). Technical Efficiency in Veterans Administration Hospitals in *The Measurement of Productive Efficiency: Techniques and Application*. (eds. Fried, H.0., Knox Lovell, C.A. & Schmidt, S.S.). Oxford University Press, New York, Oxford. pp. 335-351.

Carmichael, F. & Thomas, D. (1993). Bargaining in the Transfer Market: Theory and Evidence. *Applied Economics*. No. 25, pp. 1467-1476.

Carmichael, F. & Thomas, D. (1995). Production and Efficiency in Team Sports: An Investigation of Rugby League Football. *Applied Economics*. No. 27, pp. 859-869.

Charnes, A., Cooper, W. & Rhodes, E. (1979). Short Communication: Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*. 3 p339.

Charnes , A., Cooper, W. & Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*. No. 2, pp. 429-444.

Clement, R.C. & McCormick, R.E. (1989). Coaching Team Production. *Economic Inquiry*. No. 27(2), pp. 287-304.

CSO(Central Statistical Office) (Various). *Annual Abstract of Statistics.* HMSO, London.

dell'Osso, F. & Szymanski, S. (1991). Who are the Champions? (An Analysis of Football and Architecture). *Business Strategy Review.* Summer, pp. 113-130.

Farrell, M.J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society.* A, 120:pp. 253-281.

Ganley, J.A. & Cubbin, J.S. (1987). Performance Indicators for Prisons. *Public Money.* No. 7, pp.57-59

Ganley, J.A. & Cubbin, J.S. (1992). *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis.* North Holland, Amsterdam.

Greene, W.H. (1980). Maximum Likelihood Estimation of Econometric Frontier Functions. *Journal of Econometrics.* No.13 (1), pp. 27-56.

Greene, W.H. (1993). The Econometric Approach to Efficiency Analysis in *The Measurement of Productive Efficiency: Techniques and Application.* (eds. Fried, H.O., Knox Lovell, C.A. & Schmidt, S.S). Oxford University Press, New York, Oxford. pp.68-119.

Greene, W.H. (1993a). *Econometric Analysis.* Second Edition. Macmillan Publishing, United States.

Handy, C.B. (1985). *Understanding Organisations.* Penguin Books, London. 3rd Edition.

Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica.* No. 46, pp. 1251-1271.

Kahn, L.M. (1993). Managerial Quality, Team Success and Individual Player Performance in Major League Baseball. *Industrial and Labor Relations Review.* No. 46 (3) ,pp. 531-47.

Lee, H. & Somwaru, A. (1993). Share Tenancy and Efficiency in U.S. Agriculture in *The Measurement of Productive Efficiency: Techniques and Application.* (eds. Fried, H.O., Knox Lovell, C.A. & Schmidt, S.S). Oxford University Press, New York, Oxford. pp.288-299.

Ofwat (1994). *Future Charges for Water and Sewerage Service.* Ofwat, Birmingham, England.

Porter, P.K. & Scully G.W (1990). Measuring Managerial Efficiency: The Case of Baseball in *Sportometrics* (eds. Goff, B. & Tollison, R.D.). Texas A&M University Press. pp. 93-102.

Rangan, N., Grabowski, R., Aly, H. & Pakura, C. (1988). The Technical Efficiency of US Banks. *Economic Letters.* No. 28, pp. 169-176.

Rollin, J. (ed.) (Various). *Rothmans Football Yearbook,* Headline Book Publishing. London.

Schofield, J.A. (1988). Production Functions in the Sports Industry: An Empirical Analysis of Professional Cricket. *Applied Economics*. No. 20(2), pp. 177-93.

Scully, G.W. (1974). Pay and Performance in Major League Baseball. *American Economic Review*. December, pp. 915-30.

Sloane, P.J. (1971). The Economics of Professional Football: The Football Club as a Utility Maximiser. *Scottish Journal of Political Economy*. Pp. 121-146.

Sommers, P.M. & Quinton, N. (1982). Pay and Performance in Major League Baseball: The Case of the First Family of Free Agents. *The Journal of Human Resources*. Summer, pp. 426-36.

Szymanski, S. & Smith, R. (1993). The English Football Industry: Profit Performance and Industrial Structure. *(Unpublished Mimeo), London Business School.*

Waldman, D. (1982). A Stationary Point for the Stochastic Frontier Likelihood. *Journal of Econometrics*. No. 18, pp.275-279.

Wilson, D.C. (1995). *Organisational Behaviour*. Prentice Hall, London. Second Edition.

Zak. T.A., Huang, C.J. & Siegfried, J.J. (1990). Production Efficiency: The Case of Professional Basketball in *Sportometrics* (eds. Goff, B. & Tollison, R.D.) Texas A&M University Press. pp. 103-117.

4.9. Appendix 1: Sample Means 1974-1989

	League Perform -ance	F A. Cup	League Cup	Relative Wages	Professionals	Players used	Management Tenure	Home Grown Players
	Eqn 4.7	Round Eliminated		Eqn 4.9	No.	No.	Years	No.
Arsenal	19.00	5.19	4.25	2.23	28 81	20.63	4.50	19 13
Aston Villa	13.20	4.19	4.19	1.52	29.00	23.25	3.25	15.36
Barnsley	1.49	3.13	2.31	0.66	24.63	23.25	2.25	11.69
Birmingham C	3.82	4.06	2.88	1.42	27.75	25.06	2.63	13.94
Blackburn R	1.81	3.81	2.25	0.88	22.56	21.69	2.31	8.00
Bolton W	1.66	3.19	2.19	1.01	24 38	21.63	2.31	13.56
Brentford	0.52	2.13	1.56	0 64	22.18	24 94	1.94	6.06
Bristol R	1.01	3.00	1.63	0.56	24 61	23.06	2.25	13.38
Burnley	2.31	3.31	2.50	1 01	26 81	23.13	2 19	15.69
Bury	0.45	2.38	2.25	0.61	21.13	22.38	2.43	5.50
Cambridge U	0.82	2.50	1.81	0.53	24.06	24.44	3 00	6.25
Carlisle	1.21	3.31	1.75	0.76	20.13	21.81	2.25	2.44
Chelsea	4.58	3.81	2.88	1.32	29.31	24.13	2.25	18.75
Colchester U	0.40	2.56	1.69	0.54	21.00	21.38	2.56	8.31
Coventry C	6.34	4.31	3.25	1 49	27.75	22.50	3.44	15.06
Everton	25.64	5.43	4.56	2.15	29.19	23.25	2.75	13.75
Gillingham	0.70	2.19	1.75	0.63	22.88	23.56	3.13	8.50
Huddersfield	0.79	2.56	1.94	0.76	25.13	23.63	3.31	10.19
Hull C	1.20	3.06	2.06	0.70	24.81	23.25	2.06	12.88
Leeds U	11.84	4.00	3.06	1.58	29.06	24.00	3.44	18.31
Leicester C	4.41	4.06	2.50	1.16	29 13	23.68	2.69	16.56
Liverpool	69.91	6.00	5.25	2.13	32.56	18.50	4 88	14.81
Luton T	4.55	4.25	3.44	1.28	26.75	22.88	3.75	14.56
Man Utd	19.49	5.44	3.63	1.94	37.25	22.38	3.00	20.63
Mansfield T	0.44	2.25	1.63	0.66	20 88	23.06	2.38	7.19
Newcastle U	5.10	4.06	2.88	1.30	31.50	25.19	4.25	16.62
Northampton	0.24	1.56	1.75	0.51	21.38	22.68	2.13	8.56
Oldham	1.73	3.38	2.19	0 77	23 12	22.81	6.25	11.56
Peterborough	0.37	2.75	2.13	0.61	21.38	22.93	2.00	6.06
Plymouth	0.96	2.81	1.88	0.79	23.63	22.38	2.06	10.63
Port Vale	0.43	2.38	1.38	0.65	21 75	23 81	2 75	8.69
Preston NE	0.86	2.31	1.94	0.71	24.63	22.94	1.69	11.25
Reading	0.58	2.50	1.75	0 73	21.19	22.69	3.56	5.50
Rochdale	0.08	1.69	1.31	0.35	22.06	24.69	1.63	6.93
Rotherham U	0.69	2.56	2.13	0.55	24.56	21.69	2.63	11.5
Scunthorpe	0.17	2.19	1.50	0.50	22.50	24.38	2 19	8.00
Sheff U	2.47	3.12	2.38	0.84	26.63	24.94	2.19	10.44
Sheff W	3.91	3.87	3.06	1.11	26.88	23.81	2.75	12.81
Shrewsbury	1.25	3.50	1.94	0.61	22.19	21.06	2 38	11 00
Southampton	9.42	4.68	3.50	1.21	28.75	23.19	5.50	16.75
Southend U	0.43	2.12	1.38	0 51	21.94	22.31	3 13	7 50
Swindon T	0.80	3.25	2.31	0 73	23.75	22.19	2 31	9 38
Spurs	12.00	4.87	3.81	2 11	31.75	24.06	4 31	20.38
Walsall	0.71	2.68	2.00	0 57	22.31	21.50	3 25	10.75
WBA	7.86	4.18	3.13	1 22	29.00	23.50	1 75	15.81
West Ham	6.77	4.81	3.75	1.75	27.75	21 38	8 13	16.50
Wrexham	0.71	3.00	1.94	0 71	21.00	22 25	3 81	11 43
Average	5.42	3.37	2.49	0 56	25.34	22 93	3 01	11 88

4.10. Appendix 2: Sample Means 1974-1994

	League Perfor -mance	F.A. Cup	Leagu e Cup	Relative Wages	Professionals	Players used	Management Tenure	Home Grown Players
	Eqn 4.7	Round Eliminated		Eqn 4.9	No.	No.	Years	No.
Arsenal	22.43	5.23	4.33	2.07	31.04	21.23	4.86	20.76
Aston Villa	15.58	4.33	4.24	1.49	29.76	23.43	2.95	14.81
Barnsley	1.52	3.38	2.24	0.61	25.29	23.81	2.24	11.05
Birmingham C	3.14	4.05	2.62	1.17	28.52	26.53	2.33	14.19
Blackburn R	4.91	3.86	2.43	1.00	23.90	22.71	2.48	7.86
Bolton W	1.50	3.43	2.19	0.85	25.00	22.38	2.62	12.29
Brentford	0 59	2.14	1.67	0.56	23.67	25.67	2.00	6.71
Bristol R	1.08	2.86	1.57	0.54	24.24	23.62	2.19	11.52
Burnley	1.85	3.24	2.24	0.82	27.19	23.81	2.10	13.43
Bury	0.46	2.14	2.05	0.54	21.80	23.10	2.29	5.14
Cambridge	0.88	2.90	1.95	0.50	24.24	24.57	2.81	5.43
Coventry C	6.17	4.19	3.29	1.36	28.67	23.71	3.14	14.48
Everton	21.50	5.14	4.38	2.05	29.00	23.05	2.71	12.52
Huddersfield	0.77	2.67	1.95	0.66	25.57	24.57	3.05	10.19
Hull C	1.13	2.81	2.10	0.61	25.67	24.28	1.95	13.00
Leeds	15.67	3.81	3.24	1.59	30.14	24.24	3.33	16.81
Leicester	3.86	3.86	2.47	1.03	29.95	24.71	2.57	15.90
Liverpool	61.70	5.90	4.81	2.12	33.90	19.76	4.52	13.48
Luton	4.16	3.90	3.09	1.16	28.33	23.48	3.33	16.00
Man Utd	26.74	5.67	4.09	1.89	37.58	22.43	3.48	20.05
Mansfield	0.42	2.04	1.76	0.54	22.10	23.95	2.62	7.57
Newcastle	5.76	3.90	2.90	1.30	32.33	25.48	3.57	15.81
Oldham	2.16	3.71	2.67	0.79	24.71	23.00	7.14	11.10
Peterborough	0.48	2.90	2.67	0.56	21.76	24.00	1.95	5.81
Plymouth	1.00	2.90	1.90	0.69	25.04	23.76	1.95	10.95
Preston	0.76	2.29	1.71	0.60	25.76	24.33	1.81	10.95
Reading	0.62	2.48	1.71	0.67	22 43	23.24	3.33	6.62
Rotherham U	0.66	2.57	2.05	0.48	24.67	22.76	2.52	10.62
Scunthorpe	0.17	2.14	1.48	0.46	22.43	24.47	2.19	7.43
Sheff U	3.22	3.43	2.52	0.84	27.67	25 38	2 62	9.43
Sheff W	5.92	4.14	3.62	1.18	24 43	23.90	2.52	13.00
Shrewsbury	1.05	3.19	1.95	0.53	23.05	22.76	2.24	10.14
Southampton	8.66	4.62	3.57	1.15	29 57	24.05	5.00	16.62
Southend	0.61	2.05	1.67	0.53	23.00	22.71	2.76	7.67
Swindon	1.17	3.00	2.81	0.79	24.62	22.67	2.10	9.00
Spurs	11.97	4.95	4.10	1.99	33.90	24.62	4.00	20.57
WBA	6.24	3.86	2.81	1.04	29.24	24.76	1.62	14.29
West Ham	6.01	4.86	3.71	1.65	28 52	22.28	6.71	15.62
Wrexham	0.59	2.67	1.81	0.58	22.42	23.57	3.62	11.05
Average	6.49	3.52	2 68		26 80	23.66	3.01	12 05

4.11. Appendix 3: Random Effects Model Likelihood Function

$$
\begin{aligned}
\ln L_i = & (-T_i / 2)\ln(2\pi) - \frac{1}{2}\ln 2 - (T_i / 2)\ln \sigma_v^2 + \frac{1}{2}\ln(1 + \lambda T_i) \\
& + \frac{1}{2}(1/\sigma_v^2)[\lambda / (1 + \lambda T_i)][\textstyle\sum_t (\varepsilon_{it} - \mu / \sigma_u)]^2 \\
& - (1/\sigma_v^2)[\textstyle\sum_t (\varepsilon_{it} - \mu)^2] \\
& + \ln \Phi\{-[\lambda/(1 + \lambda T_i)]^{\frac{1}{2}}[\textstyle\sum_t ((\varepsilon_{it} - \mu / \sigma_u) / \sigma_v) + (\mu / \sigma_u)T_i(1 - 1 / \lambda)]\} \\
& - \ln \Phi(\mu / \sigma_u)
\end{aligned}
$$

4.12. Appendix 4: Comparison Of DEA Estimation Techniques

Constant returns to scale: inputs= ability, professional, players used; outputs= league performance

	Sample Mean		Average of Annual Estimations	
	CRS	Rank	Efficiency	Rank
Arsenal	0.397	4	0.296	4
Aston Villa	0.359	5	0.228	5
Barnsley	0.086	20	0.060	20
Birmingham City	0.092	18	0.067	19
Blackburn Rovers	0.169	14	0.094	11
Bolton W	0.061	25	0.040	25
Brentford	0.036	33	0.024	33
Bristol R	0.069	21	0.050	22
Burnley	0.078	24	0.044	21
Bury	0.029	37	0.020	37
Cambridge	0.060	26	0.038	26
Coventry C	0.156	10	0.117	12
Everton	0.407	3	0.299	3
Huddersfield	0.040	30	0.027	31
Hull C	0.064	22	0.046	24
Leeds	0.339	6	0.207	6
Leicester	0.129	16	0.087	15
Liverpool	1.000	1	0.741	1
Luton	0.123	17	0.082	17
Man Utd	0.486	2	0.347	2
Mansfield	0.027	38	0.020	38
Newcastle	0.152	12	0.111	13
Oldham	0.094	19	0.066	18
Peterborough	0.029	35	0.021	36
Plymouth	0.050	28	0.035	28
Preston	0.044	31	0.027	30
Reading	0.032	34	0.022	35
Rotherham U	0.047	29	0.033	29
Scunthorpe	0.013	39	0.008	39
Sheff U	0.132	15	0.087	14
Sheff W	0.172	13	0.101	10
Shrewsbury	0.068	23	0.045	23
Southampton	0.259	7	0.158	7
Southend	0.040	32	0.025	32
Swindon	0.051	27	0.037	27
Spurs	0.207	8	0.150	8
WBA	0.206	9	0.136	9
West Ham	0.125	11	0.113	16
Wrexham	0.035	36	0.021	34

5. CONCLUSIONS

Each of the three main chapters provides new insight into a particular area of English football relying of econometric techniques for their analysis. Each chapter has implications both for the football industry and other related industries most notably television. It is these implications which are the subject of this following chapter.

5.1. Football Administrators and Sports Broadcasters

Chapter 2 is the first study to compare the reasons why people attend football with the reasons why people watch live television matches. A tobit model was used to take account of the censored nature of the attendance data, something largely ignored in the literature. Original measures of outcome uncertainty, both seasonal and match uncertainty and quality were introduced.

As discussed in Chapter 1 English football is experiencing a time of growth. Allied with the growth is the desire among some clubs for structural change. Ideas that have been floated include a Europe wide league and the introduction of some form of joint Anglo-Scottish club competition. The results of this Chapter 2 have implications for both league administrators and broadcasters in the current changing environment.

The results confirmed that quality was an important determinant of both match attendance and televisions audiences. One of the measures of quality used was the number of internationals in each team. It was found that the more internationals in teams the more attractive the fixture. The recent growth in Premier League income (see Chapter 1) has allowed Premier League clubs to compete in the world market for top players. This has meant that overseas internationals from successful countries such as Holland, Italy and Brazil now play in the Premier League. In addition top English players are no longer exported to the top European Leagues. The influx of foreign stars has not been welcomed by all sections of football. For example the Professional Footballers Association (PFA) has tried to block some foreign players' work permits. The results of Chapter 2 imply that in terms of attendance and television audiences the influx of internationals is good for the game. However this does not address all the critics concerns as they fear player imports will impede the development of talented English players.

It was found that the impact of games of interest[68] was localised as indicated by the game of interest variable being significant in the attendance equation but not the television equation. Thus television schedulers could increase audiences by broadcasting other games. The games that broadcasters should show are suggested by the significance of the quality and outcome uncertainty variables. These variables show that audiences wish to see quality players (i.e. internationals) playing in matches which have a bearing on the championship. It is very difficult to identify championship important games much in advance of them being played. Thus television companies should attempt to obtain a flexible contract allowing maximum freedom in selecting games as the season progresses.

[68] Narrowly defined to include derby matches and games with a history such as Liverpool versus Manchester United.

The importance of games of interest also has implications for administrators. It suggests that local and historical rivalries are important determinants of attendance. This should be a warning to advocates of major structural changes such as the introduction of a European super league. Another lesson for administrators can be found in the significance of the day of game variable in the attendance equation. It was found that weekend matches are much more attractive than mid-week games. This implies that the number of weekend fixtures should be maximised. There are several ways that this could be achieved including: lengthening the season and or introducing a mid-winter break to reduce the number of postponements; reducing the number of other competitions that take up weekend fixtures (e.g. League Cup); and changing the transmission date of live football from Monday nights. An alternative approach would be to introduce some form of pricing differential between mid-week and weekend fixtures. As noted in Chapter 2 the pricing structure in the Premier League is very inflexible, although since the study some teams have introduced a two tier pricing structure reflecting the attractiveness of their opponents.

The televising of a game did not impact on attendance according to the insignificant television variable in the attendance equation. This result is quite surprising as the standard belief is that televising a game will reduce attendances. It is possibly due to two factors: the small subscriber base of Sky Sports which is only available through cable or satellite; and the fact that Sky show many of the most attractive games, for which television becomes a complement and not substitute for many people. Thus this result should not lead football administrators to underestimate the effects of television on attendance. Indeed with the advent of digital television it is expected that subscription sports channels will be available to terrestrial households, greatly increasing the number of potential subscribers. In addition, pay-per-view football is being discussed in many quarters. With pay-per view, it is likely that the supply of live televised football will increase. This is currently the case in Italy where all top division games are televised on a pay per view basis. The combination of these factors could greatly increase the impact of televising games on attendances. However it is in the broadcasters interests to keep attendances high as this adds to the television spectacle and will affect television audiences. Indeed, in the US and Italy local blackouts often operate in attempt to maintain attendances for televised games.

The importance of outcome uncertainty for both attendances and television audiences was confirmed. This indicates the importance of strong league made up of teams with fairly equal playing ability. One method of ensuring a strong league with a fairly equal distribution of playing talent is to provide cross subsidisation between the bigger and smaller clubs within a league. This currently occurs via the television deals where each club receives a minimum payment regardless of how well they perform or how many times they are shown on television. However this form of cross subsidisation may be under threat. The Office of Fair Trading is taking the Premier League to the Restrictive Trade Practices Court for acting as a cartel in the selling of its television rights. The OFT seems to want each club to sell its home game television rights and not allow the Premier League to sell a rights package for all clubs as is currently the case. Individual licensing, while not making cross subsidisation impossible, does make it much more difficult as explicit revenue transfers between clubs are then required.

5.2. Gamblers and Bookmakers

The third chapter investigates the efficiency of fixed odds betting, the major form of betting on football in England. It provides both a model of bookmakers behaviour and a test of the efficient market hypothesis.

The model of bookmaker behaviour shows that market inefficient odds are possible when the bookmaker maximises expected profit. The model indicates that if there is a bias in punters' expectations, perhaps due to team loyalty, then the bookmaker can take advantage of this by setting non market efficient odds.

Empirical work on two seasons data was carried to test whether inefficient odds were set in practice. The unique use of the ordered probit model does identify inefficiency and exploitable betting opportunities using both odds and publicly available performance data as explanatory variables. A betting rule is developed which in the seasons tested provides a positive post tax return. This is a practical outcome of the study, a profitable betting rule. However, as well as the transactions costs of monitoring the system there are complications in implementing the solution. There has been a move by bookmakers in seasons subsequent to the study to further restrict the bets they allow. Currently only a minority of games can be bet on outside five-folds (i.e. you must bet of the results of five games simultaneously). This restriction affects the bookmakers take and complicates the implementation of the betting strategy.

For example taking Table 3.16 which looks at the results of a particular betting rule over season 94/95 and using a filter of 1.2 there are 723 bets which are recommended of which the rule predicts 50% correctly. In a season of 35 weeks this averages out at just over 20 matches a week forecasting correctly approximately 10 matches. If singles were allowed then this would amount to 20 bets, however to cover all possibilities with five-fold combinations this would require over 15,000[69] five fold bets. Thus given that most bookmakers have a minimum bet of around £0.25 the average weekly expenditure becomes very large. The return using five-folds is also diminished by the extra over-roundness on multiple bets as described in section 3.5.3.

An alternative option to overcome the restrictions would be to bet on half time/full time scores. The bookmaker allows bets predicting the score at half time and the score at full time. Thus there are 9 combination of bet as shown in Table 5.1. In order to cover one predition (e.g. a home win) three bets will be required (i.e. 1,4 and 7). Thus the average number of weekly bets becomes 60, still large but much less than the five fold combinations. The over-roundness of these types of bets is around 20%, greater than single bets but much less than five-folds.

[69] Calculated as $20!/(5!\,15!) = 15,504$

Table 5.1 HalfTime/Full Time Combinations

	Half Time	Full Time
1	Home win	home win
2	home win	away win
3	home win	draw
4	away win	home win
5	away win	away win
6	away win	draw
7	draw	home win
8	draw	away win
9	draw	draw

Another interesting outcome of the study is the warning it gives punters when evaluating so called experts' advice. A particular test of efficiency looks at predictions published in The Racing Post and The Times and find that these predictions contain no information over and above the odds. Indeed, following these tipsters advice provides negative returns.

This chapter was the first paper to rigorously test the efficiency of a fixed odds market and discovered the rare occurrence of both inefficiency and profitable betting opportunities. It provides evidence against the efficient market hypothesis and despite implementation difficulties gives a betting rule for punters to exploit.

5.3. Football Clubs and Utility Regulators

Chapter 4 is the first study of efficiency in English football. Emphasis was placed on true inputs and not intermediate outputs as has been the case in many studies. The chapter allows a useful comparison of efficiency techniques. This has implications for other industries, particularly regulated industries where efficiency estimates play an important role in price setting.

The efficiency investigated concentrates on organisation efficiency. The question addressed is how successful is a club in turning its available playing and team management talent into playing success (here mainly league position). It is often perceived that the management of football clubs is extremely inefficient and amateur. This criticism is probably overstated. In recent years football clubs have become big businesses and thus have been attracting better managers. It has been suggested that as many clubs are currently planning to float on the stock exchange there will be an increase in managerial ability as shareholders' discipline is encountered. This may be true but the floating of clubs does dilute the objective assumed in this study, namely playing success, requiring clubs to make positive returns.

The econometric model used gives an indication of the importance of certain factors for success. The results indicate that given an average ability in the squad, success is positively related to the size of squad. This makes sense, a bigger squad will provide better cover for injuries and loss of form. Also in a large squad there is competition for places and this increases the incentives for players to perform to the best of their ability. However the negative sign and significance of the number of players used variable indicates that there are benefits to be had from having a settled consistent side. The need for consistency is confirmed by the significance of the management tenure variable (i.e. how long a manager has been in a given job) indicating for a given level of managerial and playing ability there are success advantages to tenure. In a wider context, the importance of consistency is not surprising. Much management literature is concerned with the problems of organisations adapting to change. Indeed, the high turnover of team managers seen in football is not observed in other industries where consistency and stability are more highly valued.

The study identifies efficient football clubs, allowing other clubs to identify and possibly learn from their organisational structure and style. Liverpool is the most efficient club in all formulations of the model. Certain features of Liverpool's organisation can be identified as having contributed to its efficiency. Notably there are clear goals within the organisation, playing success is the aim, financial considerations are secondary. Stability and continuity of management styles are also important. Apart from West Ham, Liverpool has had the smallest number of managerial changes since its foundation, 14 managers in 104 years. In the sample period since 1974, Liverpool has had 5 managers, three of whom had worked for the club's training staff for many years and the remaining two being ex-players. Thus all were versed in the Liverpool style of management and were aware of the organisations goals and culture. The importance of stability is confirmed by the high ranking given to Southampton, a less fashionable club who also exhibited consistency in management.

Outside football this chapter has implications for utility regulators. Utility regulators, most notably Ofwat (Office of Water Services), use efficiency studies to inform their price setting decisions. Ofwat carried out its first Periodic Review of prices in 1995, just over five years after privatisation. At a Periodic Review Ofwat sets a 10 year price cap (RPI+K) for each of the water and sewerage companies. The price cap is set such that the companies can cover operating costs, capital depreciation and maintenance and achieve a reasonable return on their capital employed. Ofwat's view of appropriate operating and capital expenditure forecasts are subject to company specific efficiency targets. These targets set are informed by efficiency studies.

In 1995 while no formal model of capital expenditure was developed, Ofwat carried out econometric studies of companies' operating expenditure efficiency. The favoured technique was estimating deterministic production functions[70] using ordinary least squares. The results indicated that a 6% per annum efficiency target could be assumed for the most inefficient companies. In the actual price setting the econometric targets were reduced by half in recognition of the uncertainties surrounding the efficiency estimates due to data problems. The data employed were submitted to Ofwat by the companies and as such were subject to both the usual measurement problems and manipulation by companies.

The results of this Chapter 4 indicate that even with a well defined and accurate data set, efficiency estimates are extremely sensitive to the method employed (i.e. econometrics or DEA); the specification used; and in DEA's case returns to scale assumptions. Thus as was

[70] See 4.5.1

shown with football clubs, a regulated company could be deemed both efficient and inefficient in two different but equally plausible models.

Ofwat has reconfirmed its commitment to econometric efficiency estimates for the next Periodic Review in 1999. Whilst econometrics can indeed inform efficiency estimates, the sensitivity displayed in Chapter 4 suggests that less emphasis should be placed on econometrics alone, particularly if based on a single specification. A suggested approach would be to utilise a range of techniques and specifications. Thus consistent findings across specifications will be identified (e.g. the efficiency of Liverpool) as well as companies for which the specification matters (e.g. Scunthorpe). In this way the efficiency of companies particularly sensitive to specifications can be better understood, thereby avoiding inappropriate efficiency targets.