2020

An Efficiency Analysis of Football Clubs in the English Premier League

Aakash Thapa
The College of Wooster, athapa20@wooster.edu

An Efficiency Analysis of Clubs in the English Premier League

Aakash Thapa

Economics Department, College of Wooster

Senior Independent Study

2019-2020

Advised by Dr. Sophie Wu

Abstract

This research analyses the productive efficiency of football clubs in the English Premier League by using a cost frontier model. Result suggests that among the more successful clubs that often have the same objective as winning the league, efficiency is varied and is likely due to key differences in personnel like team mangers.

Keywords: Efficiency Analysis, Cost Frontier Model, Team Managers

Table of Contents

List of Tables

1. Introduction

Association football[1] is a global sport and is only growing in popularity. This growth in popularity has aided plenty of professional teams in gaining revenue and increasing their brand values through worldwide broadcasting. Most of the beneficiaries of this positive change have been European football clubs. Clubs like Real Madrid, Manchester United, Barcelona, Manchester City, Bayern Munich and others are some of the most valuable franchises in the world with all clubs estimated to be valued over $2 billion (Badenhausen 2019).

With the growth in value and revenue more of these European football teams have started to spend considerable amounts to acquire the most talented players in pursuit of sporting success i.e. win as many titles and tournaments as possible. For instance, the value of the players employed by the English football club Manchester City is about $1.16 billion and all the top five teams with the most valuable players exceed $1 billion in squad valuation (Poli, Besson & Ravenel 2019). At the same time the values of players have also spiked with Brazilian forward Neymar being the most expensive player of all time sold at the price of $263 million[2](Gaines 2017).

Despite regulations like Financial Fair Play (FFP) where, clubs are restricted to exceed their spending to prevent unsustainable debt, instituted by UEFA (Union of European Football Associations) the governing body of Europe, owners of big football clubs continue to invest large sums of money to gain sporting success. Hence, plenty of questions arise. Are these teams able to succeed on the field? Does sporting success coincide with corporate success? Are these

[1] Known as Soccer in the US; Will be referred to as football for the rest of this paper.
[2]Players are treated like assets and are traded among the clubs. The most capable players are often the most expensive ones

teams able to sustain themselves without additional funds? How do the smaller teams without the capital and the resources compare?

Thus, in this paper I have investigated football teams by considering both on-field success and revenue turnovers to understand if football clubs are efficient businesses in their pursuit of sporting success. More precisely this research is focused on examining if clubs become inefficient production units in pursuit of sporting success. At the same time, the research also attempts to compare clubs within a league and see if clubs with less success are less inefficient. For this study, the football teams from the English Premier League are chosen as this league consists of some of the most successful and the most valuable teams in the world. The league is considered one of the most competitive football leagues in the world where players of the highest quality perform every season.

The rest of the paper is organized as follows: **Chapter 2** involves literature review where studies conducted on efficiency analysis of football teams are discussed. **Chapter 3** describes the theoretical framework of the study with discussion of the model and the variables used. **Chapter 4** discusses data and results and **Chapter 5** concludes the study.

2. Literature Review

Important note: The first four papers reviews are of studies conducted for the English Premier League (EPL). It is important to know the structure of the league. The league is comprised of 20 clubs. Every new season three clubs are relegated into the second division and three new teams from the second division move into the first division. The teams that finish in the top six positions are guaranteed to play in the two European tournaments: Champions League and Europa League. All the clubs in the league also get to participate in two separate domestic knockout tournaments: EFL Cup and FA Cup.

Dawson, P., Dobson, S., & Gerrard, B. (2000). Stochastic Frontiers and the Temporal Structure of Managerial Efficiency in English Soccer. *Journal of Sports Economics*, *1*(4), 341–362. https://doi.org/10.1177/152700250000100402

The researchers are interested in investigating performance of the managers in corporate organizations, but problems exist in gathering data. Hence, they chose to examine the efficiency of managers of soccer teams. They cite previous research that has been done in understanding the efficiencies of sports teams. They also mention that very few studies have been dedicated towards the managers performance because of the difficulty in measurement. The researchers deal with this problem by measuring the contribution of the managers separately from the contribution of the players in team performance by using a frontier production function rather than estimating a standard empirical production function.

A frontier production function allows the researchers to estimate the production frontier where level of production is optimal and estimate the different points at which the unit of

observations i.e. managers lie. By comparing the distance between the frontier (optimal level) and the levels where the managers are performing the researchers estimate an efficiency score. The researchers employ a stochastic production frontier which allows them to have two error terms, a two-sided term that captures statistical noise and a positive error term that accounts for managerial inefficiency. This is appealing to the researchers as team production is not only reliant upon managers decision like team selection but also on factors that are not under control of the manager like bad weather.

Appropriate modifications are made to their empirical model to account for time varying effects on manager efficiency (expect managers to be better overtime) and to sperate the effects of players on team performance to accurately estimate manger efficiency. Hence, mangers contribution to teams winning percentage is estimated and managers' efficiency as well. The researchers conclude that managerial efficiency is declining due to the intense pressure for success and high rate of managerial turnover. Since their research is only based on one dimension of the managers i.e. their ability to maximize team performance given the plyers available, further research could be done in their ability of finding players at a bargain.

Haas, D. J. (2003). Productive Efficiency of English Football Teams--A Data Envelopment Analysis Approach. *Managerial and Decision Economics*, *24*(5), 403–410. https://doi-org.wooster.idm.oclc.org/http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291099-1468/issues

This research is concerned with the productive efficiency of English Premier league teams. The researchers use Data Envelopment Analysis (DEA) to estimate a production frontier

and believe that this is appropriate for non-marketed (hard to value or price) inputs and outputs. They used clubs' total wages and salaries as team inputs with salary of the head coach as a separate input variable. The researchers mention that the value and size of the playing squad are implicitly reflected by these variables 'total wages and salaries.' These teams come from different parts of the country with varying populations and different demands for football. This difference affects team revenues hence, population of the club's hometown is also included as a non-discretionary input variable.

The outputs include points awarded during the 2000/2001 season and the total revenues o the season. The researchers mention that only using points awarded as the output would be misleading as some of these also play in international tournaments. Hence, they add revenue as an additional variable as that reflects the teams output nationally and internationally. Total revenues also include revenues from other domestic tournament which all teams participate in and having success in these tournaments can increase club revenue significantly. Therefore, the researchers believe that total revenue is a comprehensive variable that reflects the team output.

The efficiency scores of a team is relative to all the other teams all these teams either lie on the efficient frontier or below it. By employing an input-oriented DEA model which can process non-discretionary variables efficiency scores are acquired. Assuming constant returns to scale global technical efficiency (TE) scores are acquired and assuming input-oriented variable returns to scale local pure technical efficiency (PTE) scores are acquired. Further decomposing TE into PTE and scale efficiency allows insight into the sources of inefficiency i.e. inefficiency in transforming inputs into outputs or inefficiency in small scale operation or both.

Based on the empirical results the researchers conclude that for the inefficient operation is the main source of overall inefficiency.

Barros, C. P., & Garcia-del-Barrio, P. (2008). Efficiency Measurement of the English Football Premier League with a Random Frontier Model. *Economic Modelling*, *25*(5), 994–1002. https://doi.org/http://www.sciencedirect.com/science/journal/02649993

This research is concerned with the efficiency measurement of football teams of English Premier League by estimating cost functions. Unlike the other studies this study estimates the cost frontier of the clubs rather than a production frontier. Consequently, an efficient iso-quant is estimated form the sample and the clubs in the sample have efficiency scores as ratios of the minimum cost and actual cost of operating these clubs. Hence, the closer the score is to 1 the more efficient the team is.

The study discusses two theories as theoretical references for the research i.e. the strategic-group theory and the resource-based theory. Based on strategic-group theory the differences in efficiency between the clubs are explained by the differences in the structural characteristics of these clubs. Hence clubs with different structures and assets implement different strategies and have different results. Resource-based theory explains the differences in efficiency are explained by the differences in resources and capabilities of the clubs upon which the clubs base their strategies.

As inputs they use prices of workers and the capital and as outputs, they use sales, points awarded to the team and their attendance. They estimate their results through Cobb-Douglas non-random frontier model, Cobb-Douglas random frontier model and a translog random frontier

model for comparison. The empirical results suggest that heterogeneity should be considered a major issue in English Premier league and heterogenous models are better in capturing the cost structure of the English Premier League. They also mention that larger data sets could be used to further validate their results.

Guzman, I., & Morrow, S. (2007). Measuring Efficiency and Productivity in Professional Football Teams: Evidence from the English Premier League. *Central European Journal of Operations Research*, *15*(4), 309–328. https://doi.org/https://link.springer.com/journal/volumesAndIssues/10100

This study examines the productive efficiency of football clubs in the English Premier League. They use Data Envelopment Analysis(DEA) model to gain efficiency scores. This is done by implementing a Canonical Correlation Analysis(CCA) to ensure cohesion of the input and the output variables. They assume that all the clubs have access to common technology, so the productivity of the clubs is based on the management of their resources. As a result, constant returns to scale is assumed to asses club efficiency. They conduct an additional study to establish the productivity of the clubs in the sample by using Malmquist productivity index. The researchers mention various studies that have implemented DEA to measure efficiency in and outside the field of sports economics.

For output variables points won in a season and the corresponding total revenue are selected. Points won is used as proxy for sporting success. Revenue from non-football activities are excluded. Hence, the variable is total revenue sourced by gate revenues, media broadcasting, merchandising or other activity related to football. The researchers argue that due to the clubs'

participation in other domestic and continental tournaments total revenue is the most appropriate variable that reflects the financial success of the clubs involved.

For input variables staff costs, director's remuneration and general operating costs are selected. Staff costs include salaries of players, managers and coaching staff and other staff involved in the football business. Director's remuneration is used as a proxy for the human capital responsible for guiding the club based on the club's long-term objectives. Operating costs include items like lease and rental charges, fixed asset depreciation, repairs and renewals, training ground costs and professional fees. They mention that general expenses and director's remuneration represent the investment by the club to attain both financial and sporting success and this is also an extension on previous studies done on this topic.

The sample size consists of data of six seasons from 1997-1998 to 2002-2003. Using Canonical Correlation Analysis, the strength of the model in explaining the data is tested for all the six seasons. Similarly, the strength of all the variables both the inputs and the outputs are tested for all six seasons and they find that Director's remuneration is not significant for three seasons (half the sample size). Hence, their final model did not include Director's remuneration as a variable.

An additional study is conducted to measure the change in productivity of the clubs within the six seasons included in the sample. This is done by estimating the Malmquist Total Factor Productivity index (TFP) for 11 clubs included in the sample. They deconstruct the index into two components: catching up and technological change. Catching up shows the relative change in technical efficiency of the clubs within periods and technological change reflects to the

technical change of the sector/league within periods (t, t+1). This allows the researchers to distinguish between the performance of the clubs and the performance of the league in its entirety. The researchers conclude with the application of CCA, DEA models are suitable to measure the efficiency of professional football clubs.

Important note: The following review is of a study conducted on soccer clubs of Ligue 1 (French football league). The league consists of 20 teams. Each new season two teams that finish in the 19th and the 20th positions are relegated into the second division and the team that finishes in 18th position play a relegation play-off with the third ranked team form the second division to get a chance to stay in the first division. The first four teams are guaranteed to play in two continental tournaments: Champions League and Europa League. All the teams in the league also participate in two domestic knockout tournaments: Coup de la Ligue and Coupe de France.

Barros, C. P., Peypoch, N., & Tainsky, S. (2014). Cost Efficiency of French Soccer League Teams. *Applied Economics*, *46*(7–9), 781–789. https://doi.org/http://www.tandfonline.com/loi/raec20

This research is concerned with the cost efficiency of the French soccer league, Ligue 1. By using a cost frontier that represents the minimum cost required to produce an output given input prices the researchers specify an empirical model that allows the deconstruction of the error term into two components: statistical noise(v) that is assumed to follow a normal distribution centered around zero and inefficiency(u) that is assumed to follow a half-normal distribution. Hence, $u = 0$ would indicate that a club is producing at the lowest possible cost and a $u > 0$ would indicate

that the club is producing above the minimum cost frontier. And the positive value of u would show the actual level of inefficiency.

The researchers mention that previous studies conducted using stochastic frontier models have assumed that all teams use the same technology or have the same technological capabilities and can lead to an overestimation of inefficiency scores if the assumption is wrong. Hence, the researchers cluster the sample based on similar characteristics and create distinct clusters to address the issue of heterogeneity. As a result, the sample is divided into two clusters: Group 1 and Group 2. Group 1 consists of top teams that are internationally known that have had success in European tournaments with high operating costs. Group 2 consists of smaller regional clubs with less notoriety and much lower operating costs.

As outputs points gained in a season is chosen. As inputs, prices of labor and capital are chosen. The proxy for labor prices is the total wages paid to the players. The proxy for price of capital premises is the measure of amortization and reconstruction expenditure divided by net assets and liabilities and the proxy for capital funding is the measure of dividing debs by the total assets. The data is collected from nine seasons from 2002 to 2011. The researchers conclude that future research should consider the presence of heterogeneity.

3. Theoretical Framework

3.1Model:

In examining the economic efficiency of a firm one can test if the firm is able to maximize the output with the available resources i.e. productive efficiency and if the firm is using the optimal proportions of the resources used to satisfy its customers' wants, given their prices and the production technology i.e. allocative efficiency.

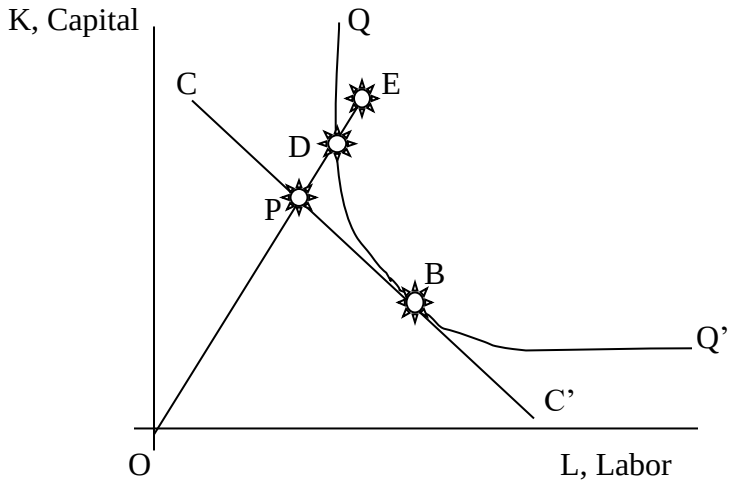


Figure 1.1

The above diagram shows both the efficient isocost and the isoquant curves of a firm. If the firm uses the quantities of output represented by E to produce quantity QQ' then it is not productively efficient, and its inefficiency is represented by DE. If the firm is using the quantities of output represented by D, then it is not able to fully minimize its cost of production to produce QQ'. This means that the firm gains productive efficiency but does not attain allocative efficiency. And its allocative inefficiency is represented by DP. To attain both productive and allocative efficiency in producing QQ' quantity the firm needs to reduce the proportion of its

inputs by DE and reduce its cost of production by DP. This is only possible at point B which represents the cost minimizing combination of the inputs to produce QQ'. Point B represents the cost-effective combination of inputs to produce QQ'. Hence, by engaging in cost minimization a firm can attain both Productive and Allocative efficiency (Coelli, Rao & O'Donnell 2005).

Therefore, this research will study the cost efficiency of football teams and examine if the clubs are able to minimize their costs in attaining a certain level of output.

Two methods have been widely used in studying the efficiency of firms and industries: DEA and frontier analysis. For this research I have chosen a frontier analysis to examine efficiency. One of the main attractiveness of frontier analysis over DEA is that it allows one to account for the impact of statistical noise hence reducing the chances of getting less accurate results (Coelli, Rao & O'Donnell 2005).

Hence, here a cost frontier is used to characterize the production process of football clubs. This cost frontier represents the minimum expenditure required to produce any output given input prices. This means that the clubs either lie on or above the frontier (Barros & Garcia-del-Barrio 2008; Barros, Peypoch, & Tainsky 2014). Thus, the general cost frontier is defined as:

$$C = C * (x, y, T) \qquad eq(1)$$

where, *C* is operational cost, *x* is input prices, *y* is output, and *T* represents the state of technology and the unit of analysis is a football club. For panel data the model is specified as:

$$C_{it} = C^{*}\left(x_{it}, y_{jt}, T_{t}\right) \cdot e_{it} \quad eq(2)$$

where,

C_{it} = Operational cost of the i^{th} club in the t^{th} period

x_{it} = input prices of the i^{th} club in the t^{th} period

y_{it} = output of the i^{th} club in the t^{th} period

T_t = State of technology for the t^{th} period

e_{it} = Positive error term indicating inefficiency of the i^{th} club in the t^{th} period

Here the key variable is the error term which is assumed to positive because at the ideal cost frontier clubs would maximize their output and produce without any inefficiency. However, this value can be misleading if heterogeneity is not considered. Football clubs are of different sizes, have different levels of resources and strategies and consequently have different objectives. This means that the clubs could be implementing different technology in their production (Barros & Garcia-del-Barrio 2008). Considering this the sample data is likely to have different clusters that represent clubs of different characteristics. This would cause variation in efficiency scores that does not necessarily represent the differences in the ability among the clubs to optimally maximize their output. Also, at the same time this research is also concerned with analyzing the efficiencies of clubs that are not successful in terms of winning trophies it is more appropriate to have two different clusters. Hence, for this research the sample data will be stratified into two different clusters.

Before we identify the different clusters a thorough understanding of the league and its structure is necessary. This research will include football clubs from the English Premier League which consists of 20 clubs. The team that finishes in the 1^{st} position with the most points wins the title. The teams that finish in the top 6 positions including the winning team qualifies for one of two continental tournaments: Champions League and Europa League. Due to their prestige, qualifying for continental tournaments is a major incentive for teams to gain more points and

move up in rankings. At the same time getting a chance to play in such tournaments also helps clubs gain more revenue through broadcasting, ticket sales and tournament merit payouts. The teams that finish in the last three positions 18^{th}, 19^{th} and 20^{th} are relegated into the second division. Also, all the teams in the league also participate in two domestic knockout tournaments and additionally the teams that have qualified for continental tournament in the previous season also participate in one other tournament.

Clusters based on the most common club objectives are:

Cluster1:

Clubs that aim to win the league are often the most popular clubs that are internationally well known. Since, the competition to win the league is extremely high teams need to maximize their wins and minimize their losses. As a result, these clubs often employ highly talented players and managers. Along with this their infrastructures such as training facilities and stadiums often are of the highest quality. Hence, clubs that aim for the league title usually have the highest level of expenditure and are often expected to get the most amounts of points. Since, only one team can win the title these teams that fail to win the title are at least expected to qualify for a continental tournament. These teams almost always finish in the top 6 positions in the league.

Cluster2:

The rest of the teams in the league will be included in cluster 2. These teams are not as successful or as famous as the clubs in cluster 1. These clubs have somewhat talented players but

often these players are bought by more successful teams like the ones in cluster 1. These clubs are also defined by their modest infrastructure and spending capabilities.

These clusters will be further validated with empirics and the efficiency analysis for each of these clusters will be conducted separately. The implications of the analysis in the context of the hypothesis of this research will be discussed in the next chapter.

3.2. Variables:

As football clubs want their respective teams to move up in the table, teams are always seeking to maximize the points they can get. In every game the winning team gets 3 points and the losing teams gets 0 points. If a game ends in a draw each team is rewarded with 1 point. Avoiding losses by either winning or drawing can help the team maximize their points and consequently move up in rankings. Thus, attaining higher points strongly correlates with attaining higher rankings. At the same time points rewarded to a team also shows their relative performance compared to other teams. Since, points gained in a season represents the team's performance and its competitiveness against other teams', it is chosen as the output to be observed in this research. This variable is commonly used as the output of the football clubs in previous studies e.g. Hass 2003; Guzman & Morrow 2007; Barros & Garcia-del-Barrio 2008; Barros, Peypoch & Tainsky 2014. This variable is the primary sporting output of the football team and is also positively correlated with the club's revenue potential and fan interest (Hass 2003).

The second variable chosen as output is the club's revenue that represents the commercial output of the football teams. As the clubs are also involved in business activities through ticket

sales, merchandising, sponsorships, broadcasting etc., only including points awarded for the league is not enough to examine the performance of football clubs and could lead to misleading efficiency estimates. At the same time all clubs participate in two different domestic tournaments and on top of this some clubs also participate in continental tournaments every season. Success in these tournaments are not reflected by points awarded (Haas 2003). Hence, total revenue attained in a season is also included as an output.

Since, the analysis will be conducted in a cost minimizing framework the prices for labor and capital are included as input variables (Barros & Garcia-del-Barrio 2008; Barros, Peypoch & Tainsky 2014). These are the usual variables considered in estimating a cost frontier (Barros & Garcia-del-Barrio 2008). Proxy for price of labor is measured by dividing the total wages paid to the players by the number of players (Barros & Garcia-del-Barrio 2008; Barros, Peypoch & Tainsky 2014) along with salaries paid to team managers and head coaches (Haas 2003). Capital is divided into capital premises and capital investment. Proxies for prices of capital premises is measured by dividing amortization costs by total assets (Barros & Barrio 2008) and capital investment is measured by dividing total debt by total asset (Barros & Peypoch & Tainsky 2014).

Thus, the empirical specification of the model is as follows:

$$lnCost_{it} = \beta_0 + \beta_1 Trend + \beta_2 lnPL_{it} + \beta_3 lnPK1_{it} + \beta_4 lnPK2_{it} + \beta_5 lnPoints_{it} + \beta_6 lnRev_{it} + (v_{it} + u_{it}) \qquad eq(3)$$

where,

Trend = time dummy for each season to account for temporal changes,

PL_{it} = price of labor of i^{th} club in the t^{th} season

$PK1_{it}$ = price of capital-premises of i^{th} club in the t^{th} season

$PK2_{it}$ = price of capital-investment of i^{th} club in the t^{th} season

$Points_{it}$ = points awarded to the i^{th} club in the t^{th} season

$Rev._{it}$ = Total revenue collected of i^{th} club in the t^{th} season

e_{it} = Positive error term indicating inefficiency of the i^{th} club in the t^{th} period

The specification is of random effects model where heterogeneity is assumed to be constant within time periods. This is an appropriate assumption as the same club can have different characteristics in different seasons. For instance, a single team can change players and coaches between seasons included in the sample time period. This is likely to affect the team quality and the performance and bring changes in both input and output of the football club.

4. Data and Results

4.1 Sample

All the data on the sample was collected from individual financial statements of the clubs for each year[3]. The time periods for the sample was 7 years starting from the season 2009/2010 and ending with 2015/2016. The first cluster included data for 6 clubs so there were 42 observations (6 clubs over 7 years) in cluster 1. All the clubs in cluster 1 were able to remain in the Premier League and were not relegated. Hence, the panel data in cluster 1 is completely balanced.

Table 1: Descriptive Statistics of data in cluster 1

Variable	Observations	Mean	Std.Dev.	Min	Max
cost	42	244327.9	64107.07	132748	376615.3
points	42	71.92857	9.988757	50	89
rev	42	211561.7	59419.7	113833	350207.1
pl	42	376.0772	156.7742	201.043	835.763
pk1	42	.1421429	.0546918	.052	.295
pk2	42	.645881	.6273364	.112	3.924
trend	42	4	2.024243	1	7

The values are expressed in 2010-pound sterling.

Similarly, cluster 2 included data for 10 clubs for the same number of time periods. So, there were 70 observations (10 teams over 7 years) in cluster 2. But some of the clubs during the sample time period were relegated and were not performing in the Premier League. Since the nature of competition is different in the lower leagues data for the years where clubs were not in the Premier League were not included in the sample. This resulted into 10 less observations. Hence, the panel data in cluster 2 had 60 observations and was not balanced.

[3] Financial statements were accessed from a directory of public records; companieshouse.gov.uk

4.2 Preliminary Statistics

After implementing the random effects model discussed in the previous chapter results for both the clusters were attained.

Table 2: Estimation Results for Cluster 1

lgcost	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
lgpoints	-.3147634	.1600158	-1.97	0.049	-.6283885	-.0011383
lgrev	.7264044	.1085908	6.69	0.000	.5135704	.9392385
lgpl	.3199573	.0658942	4.86	0.000	.190807	.4491076
lgpk1	.0315838	.0651619	0.48	0.628	-.0961312	.1592987
lgpk2	.0841258	.0311786	2.70	0.007	.0230168	.1452348
trend	.0041601	.012601	0.33	0.741	-.0205375	.0288576
_cons	3.070656	1.232951	2.49	0.013	.6541168	5.487195

As you can see the variable lgpk1 i.e. log of price of capital 1 and the variable trend are not significant at 10% level. Variable lgpoints i.e. log of points attained in a season is not significant at 1% level and the remaining variables are significant at 1% level. The signs of all the coefficients are as expected (positive) except for lgpoints. Here, attainment of higher revenue (lgrev) along with higher labor and capital are strongly correlated with higher costs.

Table 3: Estimation Results of Cluster 2

lgcost	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
lgpoints	-.5776565	.5050684	-1.14	0.253	-1.567572	.4122593
lgrev	-.1375943	.7015345	-0.20	0.845	-1.512577	1.237388
lgpl	.1638981	.2588734	0.63	0.527	-.3434844	.6712806
lgpk1	-.0511864	.2210844	-0.23	0.817	-.4845038	.382131
lgpk2	-.0240875	.210531	-0.11	0.909	-.4367206	.3885456
trend	-.1940574	.069902	-2.78	0.006	-.3310628	-.057052
_cons	6.802627	7.238402	0.94	0.347	-7.384381	20.98963

For cluster 2 none of the variables are significant at 10% level except trend which is significant at 1% level. This is likely because of the data set not being balanced. The purpose of having two clusters made up of different types of clubs was to investigate their performance as production firms and comparing their results to see if the more successful clubs in cluster 1 are more inefficient firms than the clubs in cluster 2. This is not possible as the estimation results of cluster 2 are not valid. However, we can still investigate the clubs within cluster 1 to see if clubs have varying degrees of efficiency and answer if chasing sporting success is correlated with high inefficiency.

The results presented above are from a random effects model. This means that the error term which indicates productive inefficiency cannot be explained by unobserved heterogeneity in the sample that exists due to the different characteristics of the clubs and only indicates a club's inability to maximize their output given the inputs implemented. If random effects model is not appropriate, then then fixed effects model should be implemented which means that the error

term can be explained by the heterogeneity of the clubs. This means that the error term is composed of the effects of unobserved heterogeneity and the productive inefficiency of the clubs. Hence, a Hausman Test is performed to see if random effects model is appropriate for the panel data in cluster 1.

Table 4: Hausman Test Statistics

	Coefficients ----			
	(b)	**(B)**	**(b-B)**	**sqrt(diag(V_b-V_B))**
	fe	**re**	**Difference**	**S.E.**
lgpoints	-.1629012	-.3147634	.1518621	.
lgrev	.2082266	.7264044	-.5181778	.053259
lgpl	.41601	.3199573	.0960527	.1170783
lgpk1	.1023674	.0315838	.0707836	.0233314
lgpk2	.0476912	.0841258	-.0364346	.
trend	.0451675	.0041601	.0410074	

.

b = consistent under Ho and Ha; obtained from xtreg

B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(6) = (b-B)'[(V_b-V_B)^(-1)](b-B)

= 98.98

Prob>chi2 = 0.0000

The low chi2 probability valve indicates that null hypothesis should be rejected. This means that differences in coefficients between the random effects and fixed effects are statistically significant and that fixed effects model is more appropriate for the panel data in cluster 1. Moreover, this implies that the varying error terms for different clubs is explained by both their inability to maximize their output and their differences in characteristics such as different objectives and resources available.

4.3 Efficiency Statistics

To see the productive inefficiencies of the clubs within the sample period the mean scores of error terms for each club was attained. The results are as follows:

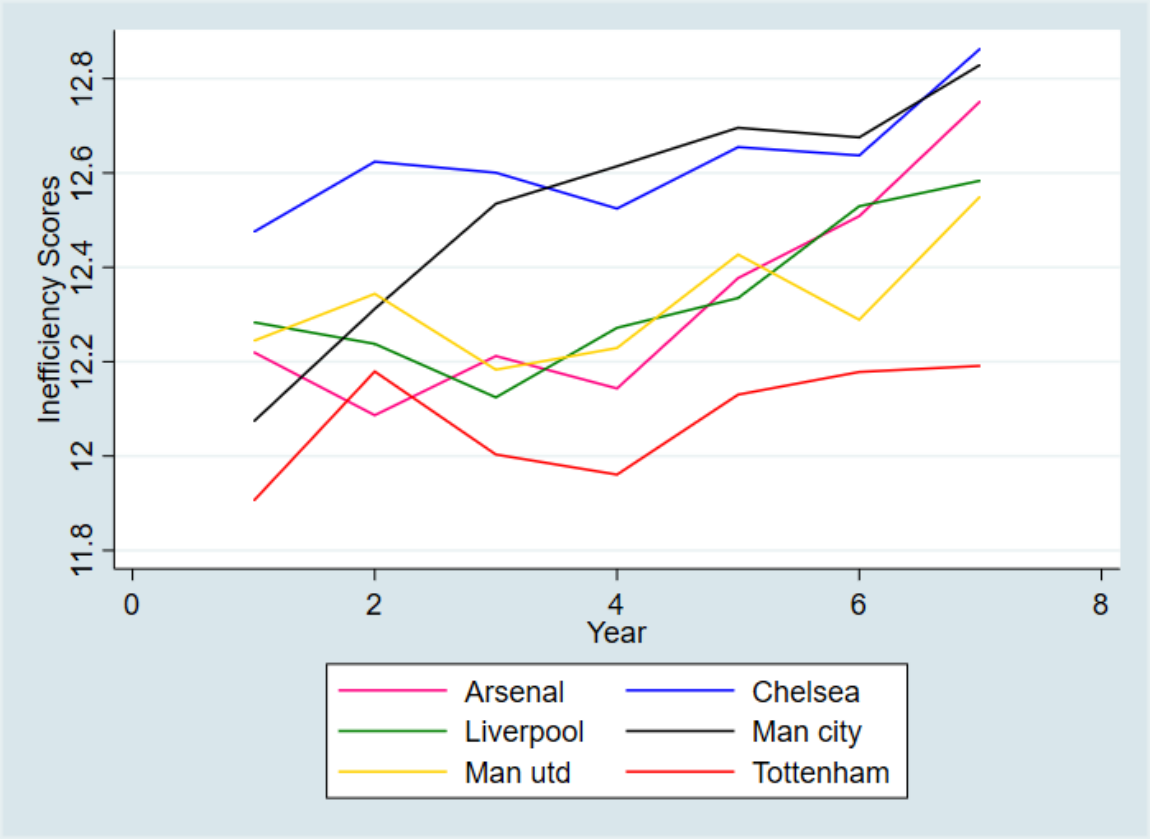
Table 5:Inefficiency Ranking

Inefficiency Rank	**Club**	**Inefficiency Score (μ_e)**	**Avg. Points**
1st	Chelsea	12.62559	73.57143
2nd	Man. City	12.5335	76.57143
3rd	Liverpool	12.33801	62.85714
4th	Arsenal	12.32856	73
5th	Man. United	12.3238	77.57143
6th	Tottenham	12.07827	68

Here the club with the lowest mean inefficiency score for the sample period is Tottenham and the club with the highest score is Chelsea and at the same time these two teams have average points of 68 and 73.67, respectively. Manchester united has the highest average points for the sample period but is the 2nd most efficient club. Liverpool has the lowest average points and is the third most inefficient club. So, far it is not clear that clubs attempt to gain more points and achieve greater accolades lead to higher inefficiency.

We can further examine the relationship between the clubs' inefficiencies and their sporting achievements by investigating the inefficiency score of each club for every year in the time period.

Graph 1: Inefficiency across the years



Here we see an increasing trend of inefficiency with succeeding years. By comparing the season achievements with the respective inefficiency score we can gain some insight between the relationship between the club's inefficiency and their sporting success.

Table 6:League winners

SEASONS	WINNERS
2009/10	Chelsea
2010/11	Manchester United
2011/12	Manchester City
2012/13	Manchester United
2013/14	Manchester City
2014/15	Chelsea
2015/16	Leicester City[4]

Within the sample time period only three clubs were able to win the league with Chelsea, Manchester United and Manchester City each winning in two different seasons. Looking at the Graph 1 we can see that Chelsea and Man. City are somewhat similar in terms of inefficiency. Only Manchester has lower inefficiency score as a club that was able to win the league. It seems that compared to Chelsea and Manchester City, Manchester United was more efficient in their pursuit of sporting success. This could be due to the Manchester United's legendary manager Sir Alex Ferguson who oversaw the football team the years Manchester United won the leagues. This highlights the importance of effective mangers that can get the best players that are available. We can also see that the years when the clubs won the league, inefficiency scores increased from the previous years. This could mean that the pursuit of sporting success like winning league does lead a club to be more inefficient than they used to be.

[4] Leicester City FC winning the league is something of a miracle and is a statistical anomaly.

5. Discussion and Conclusion

First and foremost, not all the objectives of this study were accomplished. The results for a large portion of the data collected i.e. Cluster 2 were invalid. This cluster included of teams that were not that successful compared to clubs in Cluster 1 and the exclusion of this data from the analysis prevents us from gaining insight about the efficiency of the less successful clubs. Hence, only cluster 1 with 42 observations could be analyzed.

Furthermore, the Hausman test concluded that random effects model was not appropriate for the panel data in Cluster 1. This means that the inefficiency scores attained from the error terms in the previous chapter is impacted by unobserved heterogeneity. For the sample data analyzed the differences in traits among the clubs can explain the differences in inefficiency scores rather than their inability to produce efficiently. Hence, this research can not conclude that the clubs that are differentiated by success have different inefficiencies due to their inability to produce given the resources they implemented.

However, for the sample period we saw that among the clubs that were able to win the league only one club i.e. Manchester United was relatively less inefficient and the two other clubs Chelsea and Manchester City that won titles during the sample period. Given the similarity between these clubs in terms of spending capabilities, infrastructure, and quality of personnel it is unlikely that the difference in inefficiency scores between these clubs is due to unobserved heterogeneous traits. More importantly, Manchester United won the league titles under the legendary Manager Sir Alex Ferguson who was well known for getting the best out of the players available. This is an important distinction because it highlights the significance of key

personnel such as talented team managers in a club's overall production. Hence, for the sample analyzed some clubs were more efficient than other in achieving the same goal.

Since, the analysis was conducted on a very small sample of only 6 clubs this research cannot make any conclusions about most of the clubs in the English Premier League. The analysis conducted on the most successful clubs in the league indicate that clubs have varying degrees of inefficiency due to an inability to maximize their outputs given the inputs they implemented. The results indicate that for the most successful clubs to be less inefficient, finding effective team mangers that can get the best out of the players they have is crucial. However, larger data set is needed to further validate these results.

Future research on analyzing the relationship between team managers and productive efficiency is suggested to further investigate the significance of team mangers on a club's overall performance.

References

Badenhausen, K. The World's 50 most Valuable Sports Teams 2019. Retrieved from: https://www.forbes.com/sites/kurtbadenhausen/2019/07/22/the-worlds-50-most-valuable-sports-teams-2019/#fa47b8e283da

Barros, C. P., & Garcia-del-Barrio, P. (2008). Efficiency Measurement of the English Football Premier League with a Random Frontier Model. *Economic Modelling*, *25*(5), 994–1002. https://doi.org/http://www.sciencedirect.com/science/journal/02649993

Barros, C. P., Peypoch, N., & Tainsky, S. (2014). Cost Efficiency of French Soccer League Teams. *Applied Economics*, *46*(7–9), 781–789. https://doi.org/http://www.tandfonline.com/loi/raec20

Coelli, T. J., Rao, D. S. P., & O'Donnell, C. J. (2005). *An Introduction to Efficiency and Productivity Analysis*. Second edition. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0802741&site=ehost-live

Dawson, P., Dobson, S., & Gerrard, B. (2000). Stochastic Frontiers and the Temporal Structure of Managerial Efficiency in English Soccer. *Journal of Sports Economics*, *1*(4), 341–362. https://doi.org/10.1177/152700250000100402

Guzman, I., & Morrow, S. (2007). Measuring Efficiency and Productivity in Professional Football Teams: Evidence from the English Premier League. *Central European Journal of Operations Research*, *15*(4), 309–328. https://doi.org/https://link.springer.com/journal/volumesAndIssues/10100

Gaines. C. (2017). Neymar's move to PSG will cost the French soccer giant more than $500 million. Retrieved from: https://www.businessinsider.com/how-much-neymar-salary-transfer-fee-psg-barca-2017-8

Poli, R. Besson, R. & Ravenel, L. (2019). Squad transfer values: English clubs at the top. Retrieved from: https://football-observatory.com/IMG/sites/b5wp/2018/253/en/

Appendix

A1: Descriptive Statistics of Error term and Points

Club = Arsenal

Variable	Obs.	Mean	Std. Dev.	Min	Max
e	**7**	**12.32856**	**.235322**	**12.08655**	**12.75201**
Points	**7**	**73**	**3.696846**	**68**	**79**

Club = Chelsea

Variable	Obs.	Mean	Std. Dev.	Min	Max
e	**7**	**12.62559**	**.1233349**	**12.47477**	**12.86366**
points	**7**	**73.57143**	**13.30234**	**50**	**87**

Club = Liverpool

Variable	Obs.	Mean	Std. Dev.	Min	Max
e	**7**	**12.33801**	**.1633559**	**12.12436**	**12.58381**
points	**7**	**62.85714**	**10.00714**	**52**	**84**

Club = Man. city

Variable	Obs.	Mean	Std. Dev.	Min	Max
e	**7**	**12.5335**	**.2587481**	**12.07302**	**12.82953**
points	**7**	**76.57143**	**8.997354**	**66**	**89**

Club = Man united

Variable	Obs.	Mean	Std. Dev.	Min	Max
e	**7**	**12.3238**	**.1283976**	**12.18309**	**12.55065**
points	**7**	**77.57143**	**10.78359**	**64**	**89**

Club = Tottenham

Variable	Obs.	Mean	Std. Dev.	Min	Max
e	**7**	**12.07827**	**.1190818**	**11.90502**	**12.19115**
points	**7**	**68**	**3.605551**	**62**	**72**