# Essays on Behavioural Finance and Sports Economics

by

Husain Aboualhasan

***A thesis submitted for the degree of PhD in Finance***

**Essex Business School**

**University of Essex**

**January 2021**

## Abstract

We examined two fields within the behavioural finance framework as well as sports economics. We explored sports clubs' wages and other team characteristics to explain managerial behaviour with regard to players' playing time. Additionally, we conjectured that playing at home could affect team performance, and subsequently derived a team success model based on the home venue.

The sunk cost concept is central to behavioural economists' arguments, but evidence has mainly stemmed from laboratory experiments. Empirical evidence pertaining to sunk costs and their role in the decision-making processes is rather sparse. We used a novel dataset of the English Premier League (EPL), which has not been previously utilised, particularly in the context of sunk costs. We examined the frequent manager turnover within teams, where the self-justification explanation suggested that sunk costs may not apply in this case due to managerial change. Our findings provide evidence for a sunk cost effect within a football context where higher transfer fees could predict more playing time for players within different positions. The effect was robust to a number of specifications. Additionally, we provide evidence for managers' bias towards granting free transferred players more playing time as a way to show their skills in securing bargain deals for their clubs. Lastly, we estimated a regression discontinuity (RD) model to test the effect of expensive players, and the results support the sunk cost hypothesis.

Racial discrimination is a topic of considerable debate in the labour economics literature. We employed a novel market test approach that complements the established wage equation estimate approach in a sporting context while controlling for player performance metrics. We estimated the effect of nationality (of both players and managers) on player playing time. The findings suggest that British managers demonstrated bias towards their home players by

allocating them extra playing time. This extra playing time could not be explained merely by the players' performance level, their transfer fees, or playing position. We also re-examined the models of Szymanski (2000) and Pedace (2008) with our recent dataset and found no evidence for discrimination against non-white players or for foreign nationalities.

Home field (HF) advantage in competitive sports has garnered much empirical attention. Our panel regression results provide unequivocal evidence of a significantly strong and robust home field effect. We used bootstrap techniques to explore the HF effect in predicting team success and failure in the PL. Our results suggest that teams who outperformed the expected bootstrap results in their early home (as opposed to away) games would be secure in the Premier League. This was still the case even if they performed poorly in away games over this period. This indicates a strong implication that home results dominate away results.

## Acknowledgements

I am really grateful to my supervisors, Prof. Jerry Coakley and Prof. Andrew Wood. First, they have provided me with a precious opportunity to study the economics of the English Premier League. This is something I wished for a long time ago, and it could not have materialised without their acceptance and interest in taking the journey with me. Second, their early guidance on how to build a reliable dataset along with consequent hypotheses cemented my previous knowledge during the master’s degree phase. Third, their continuous feedback and immense support was a cornerstone to my academic achievements, without which I could not have reached this level. It was a privilege and an honour being supervised by such experienced professors. I would also like to express my deepest gratitude to them, as well as Essex Business School, which helped me achieve a great goal in my life journey.

# Contents

## 1 Introduction

Economic behaviour is often modelled as rationale, where agents are expected to allocate resources based on maximum utility and ensuring that future benefits should exceed future costs (e.g. Vroom, 1964). However, the rationality of decisions has been challenged by psychology literature related to decision-making processes. Staw (1976) provided evidence on the motives behind people’s tendencies towards rationalising the negative outcomes (or feedback) that were consequences of their initial decisions.

Many studies have based their findings on experimental setups, where surveys or laboratory experiments were conducted. These setups have revealed certain investment-related decision-making processes and behaviours. Research may show to what extent agents are risk averse or how much they can bear to lose before they quit. Apart from pure financial investment frameworks, the economics of professional team sports have attracted academic interest since the mid-1950s. Both theoretical and econometric aspects of the discipline have contributed to the literature, as many economists consider professional sports to be a legitimate area for scientific research. These studies cover broad economic aspects and implications of the sports market structure (e.g. free agency in players’ labour market, nature of production, utility maximisation). There were early contributions to data science in football in particular, such as forecasting models for football match results.

In a seminal paper, Staw and Hoang (1995) departed from the conventional experimental setups to explain agents’ decision-making behaviour, which is related to escalation of commitments, in a real-world setup instead. A sports field was chosen to test for such established behavioural tendencies. They provided evidence of sunk cost effects associated with draft order in the National Basketball Association (NBA). A draft order ranks players based on their skills and quality of play, and the evidence suggests that a player who is in the top rankings is likely to get excessive playing time that is not justified by his performance. A team manager’s perception of those highly ranked players, along with their high salaries, may lead to a sunk cost effect where excessive playing time is based on costs spent on the player rather than his on-field performance.

The shift in salaries and compensation patterns stimulated a new wave of professional sports research. Theoretical contributions to sports stars’ salary determinations were proposed by some economists (e.g. Lazear & Rosen, 1981). Data on players’ salaries in European professional football are scarce due to confidentiality and disclosure restrictions, so the scope for empirical evidence is rather thin. Nevertheless, a number of studies have attempted to determine factors affecting players’ salaries in Europe (e.g. Lucifora & Simmons, 2003; Lehmann & Schulze, 2008; Frick & Deutscher, 2009).

Racial discrimination is a topic of considerable debate in the labour economics literature. The focus of the related empirical discrimination literature is on the earnings gaps (i.e. salary differences) between specific groups, and discrimination through unequal pay is estimated through constructing earning functions (e.g. Szymanski, 2000). Szymanski argued that productivity measurements in sport economics studies were much more transparent than in other labour economic settings. With a relative scarcity of individual productivity measurements (at least before the beginning of the new millennia), an alternative was to relate wages (or salary) data with teams’ productivity measures, which were mainly measured by their position in the league rank. Szymanski (2000) and Pedace (2008) found that clubs’ wage payrolls could explain nearly 90% of the variation in team productivity. Thereafter, wage equation estimates were empirically employed as the cornerstone of employer taste-based discrimination, where some personal player characteristics such as race and nationality should not possess explanatory power in a non-discriminatory environment.

Another area of competitive sports literature that has garnered much empirical attention is the home field (HF) advantage. This advantage pertains to a higher probability of winning games at a team’s home ground as opposed to those at away grounds. The literature discusses several factors that may drive such an advantage (e.g. fans’ encouragement, players’ familiarity with stadium facilities, travel fatigue of opponents). The HF advantage is prevalent in many sports, such as basketball, football, rugby, American football, hockey, and baseball. The findings from these different sports are inconclusive; however, a positive association has been found between home advantage and the psychological state of athletes prior to sports events (e.g. Bray et al., 2000; Terry et al., 1998).

With regard to home advantage historic patterns, even though the home win percentage in English football decreased from around 50% in the 70s and 80s to around 43% later in the new millennia, it remained well above the away win ratio when it reached its highest percentages of around 30% later in the decade (Dobson & Goddard, 2009). Allen and Jones (2014) analysed premier league results in the seasons 1992/93–2011/12 and showed that winning home games, on average, secured 60.7% of total points earned by teams.

A thorough documentation and interpretation of HF advantage has been mainly directed towards the psychological aspects, while empirical research with respect to the effect on field performance is relatively limited. One explanation for this is the complicated nature of precisely quantifying factors that are closely related to match outcomes, such as individual players’ skills and qualities, teams’ principles and standards, managerial and coaching qualities, and match significance within the league fixtures. Despite that, several studies have attempted to empirically investigate HF advantage by means of (but not confined to) field performance style, match-based production functions, assessing scoring rate variations of the home and away teams during a match, HF outcomes and betting market, and distance between clubs’ grounds (fields).

This thesis revisits established behavioural biases, but from different perspectives. Two pertain to football managers and how they are affected by players transfer fees as well as some personal characteristics (nationality and ethnicity/common culture or language), and one pertains to players (or the team) and how they are affected by playing at home. First, it complements the sunk cost literature with a fresh football dataset that comprises different structures and financial features from the established NBA studies and some other sports. Second, it examines racial discrimination in a sporting context that explores players' playing time rather than the conventional wage equations approach. Finally, it contributes to the well-documented home field advantage literature by incorporating a new empirical approach.

### 1.1 Contributions

In Chapter 2, we demonstrate our use of a novel dataset of the PL that has not previously been utilised, particularly in the context of sunk costs. We provide evidence for a sunk cost effect within a football context where higher transfer fees could predict more playing time for players in different positions. Sunk cost effects were observed with incremental increases in transfer fees as well as a distinctive effect of expensive players in particular. The effect was robust to a number of specifications, and most importantly, the effect persisted even after a managerial change. This contradicts a rational decision-making standpoint where assigning a new manager should remove any sunk costs effects demonstrated by his predecessor (e.g. Wicklund & Brehm, 1976). In addition, to rule out the possibility that free transferred players could bias the results (as they do not incur a large sunk cost element), we tested the sunk cost hypotheses while removing those players from the sample. The effect remained robust, though with slightly lower magnitude.

We also show evidence for managers' bias towards granting free transferred players more playing time as a way to show their skills in securing bargain deals for their clubs (e.g.

Schindler, 1998). Additionally, we estimated the RD model to test the effect of expensive players. The results converged to support the sunk cost hypothesis. In other words, football managers tended to engage expensive players, even if they did not perform to the expected standard. This may help them to feel like they did not waste the fees spent on those players (e.g. Arkes & Blumer, 1985).

The second contribution of this thesis (Chapter 3) is its examination of the nationality effect on player playing time in the English Premier League (EPL). The empirical evidence establishes a unique discrimination effect as it accounts for the effective use of players (i.e. their playing time relative to their performance). We employed a novel market test approach that complements the established wage equation estimate approach while controlling for player performance metrics. The analysis was conducted using a unique dataset of player performance metrics and their transfer fees. The findings suggest that British managers were biased towards their home players. We found evidence that British managers allocated extra playing time to British players of about 100 minutes per season. This extra playing time could not be explained merely by the players' performance levels, their transfer fees, or their playing positions, therefore providing evidence of discrimination. Our findings also indicate that even if manager nationality was ignored, British players still got more playing time than non-British players after controlling for performances. Specifically, they got more playing time than European, South American, and African players by 102, 170, and 196 minutes per season, respectively. We also re-examined the models of Szymanski (2000) and Pedace (2008) with our recent dataset and found no evidence for discrimination either against non-white players or for foreign nationalities. These findings suggest a shift of the discrimination effect towards the individual player level from the club level.

The final contribution of this thesis (Chapter 4) is its unique empirical approach to test for HF advantage in the Premier League (PL). Our findings provide unequivocal evidence of a significantly strong and robust HF effect of 0.46 additional points per home game. It held with almost identical effects for top six teams, relegation teams, and the remaining teams in the league. After establishing the HF effect, we used bootstrap techniques to explore the HF in predicting team success and failure in the PL. More specifically, our bootstrap approach to modelling match results follows that of Bell et al. (2013). The findings suggest that teams who outperformed the expected bootstrap results in their early home (relative to away) games would be secure in the PL. This was still the case even if they performed poorly in away games over this period. This suggests a strong implication that home results dominate away results. Moreover, a great home performance, even if accompanied by poor away performance, could lead to a top four ranking. An average team performance at home that was associated with poor performance away from home increased the probability of relegation to the Championship by about 40%.

Lastly, an average performance in home games that was accompanied by great away performance did not suffice to secure a top four ranking. This is additional evidence in favour of the HF effect in the context of predicting promotion in elite European football competitions. The implication of our results is that teams can evaluate their probability for promotion to the top four to six rankings or for relegation on the basis of the home results of only the first 20 games each season.

# 2. **Sunk Costs and Premier League Transfer Fees**

Sunk costs are a subject of profound debate among neoclassical and behavioural economists. They show agents' tendencies towards spending more resources on a project once a large fixed cost has been incurred. The sunk cost concept is central to behavioural economists' arguments, but evidence has mainly stemmed from laboratory experiments. Staw and Hoang (1995) introduced a practical field study in a sporting discipline, and estimated the impact of draft position on playing time in the National Basketball Association (NBA). Rational decision making suggests that the most productive players would have more playing time irrespective of their cost to the team. In reality, sports pundits would argue that managers (agents) can be biased towards granting more playing time to high-cost players. One way of justifying this behaviour is by means of sunk cost or commitment effect.

Empirical evidence pertaining to sunk costs and their role in the decision-making processes is rather sparse. Staw and Hoang (1995) and Camerer and Weber (1999) provided the main contributions to sunk cost studies with a real-world setup. Staw and Hoang (1995) provided evidence of sunk cost effects associated with draft order in the NBA, showing that highly drafted players would be granted more playing time and retained longer in their teams, even after controlling for performance. They rejected three explanations that were based on rational decision making. The first was the salary cap imposed by the NBA. They claimed that this could be mitigated by teams by the possibility of releasing their underperforming players. Second was the effect of a popular player, where managers might feel under pressure from fans to play them, but Staw and Hoang claimed that fans' behaviour towards popular players could be fickle. Third, they tested the patience of teams with their highly drafted players hoping that they would perform as expected. They found that draft position could not predict performance in the longer run (in the 4$^{th}$ and 5$^{th}$ years) and thus it should not justify teams' patience.

Camerer and Weber (1999) re-examined Staw and Hoang's findings with different specifications, such as unbundling the performance index created by Staw and Hoang and introducing predicted performance, controlling for previous season performance, using a proxy for popular players, and testing for the first round pick effect. Their results generally supported those of Staw and Hoang but produced a lower magnitude of the sunk cost effect. Thus, their alternative specifications provided some rational explanations for the sunk cost effect.

Leeds et al. (2015) used a more sophisticated method, a regression discontinuity RD, to disentangle the first-round draft or lottery pick in the NBA from those drafted later. The sunk cost effect found in the previous couple of studies had been diminished (or removed) in their model specification. This is due to the following: First, their RD model challenged the possibility of a mistaken conclusion that may have been a result of a global linear specification used in the previous literature. Second, they used different player performance measurements (wins produced per 48 mins). Lastly, a more precise measure of playing time was used by taking into account injuries and suspensions. Keefer (2017) used the RD method to test the sunk cost for the American National Football League (NFL). The results contrasted with Leeds et al. (2015) and provided evidence for a sunk cost element where a 10% increase in the salary cap value of NFL players resulted in an additional 2.7 game starts.

Therefore, PL transfers provide the basis for an interesting study of sunk costs. We chose English Football as a fresh field for applying the sunk cost theory. We contribute to the literature with the following: first, we used a novel dataset of the PL that has not been previously utilised, particularly in the context of sunk costs. Previous literature mainly relied on broad player performance measures, such as goals scored and goal assists. From 2006 onwards, the EPL started to provide more detailed player performance metrics that covered major aspects of playing, including the areas of attack and defence, in addition to team play measures and player discipline, for a total of 37 performance metrics. We explored these measures, through correlation matrices and principal component analysis (PCA), to reflect a reliable player performance and conduct the hypotheses testing. Football (soccer) presents a different setup than the NBA and some other sports (e.g. NFL). Mainly, there are no draft picks in football, and instead players are traded through two transfer windows during a season. The NBA imposes constraints and limitations to the draft picks in both financial and technical aspects (e.g. salary caps, top teams and lower teams picking rules/lottery). Football clubs can freely choose players through their agents with no salary and transfer caps, but many operate under financial constraints. This implies a possible endogenous commitment effect. Second, we explored the frequent manager turnover within teams. The self-justification explanation suggests that sunk costs may not apply in this case due to managerial change. This provides an opportunity for testing that hypothesis. The managerial change effect has not been estimated in previous studies, and thus it is introduced as a robustness factor for the analysis of sunk costs. Third, we introduced a reverse sunk cost effect where free transferred players may be granted more playing time, showing a bargain deal bias of managers. We also tested whether transfer fees in the PL had similar effects to those in the NBA and NFL using the regression discontinuity method (RD).

We provide evidence for a sunk cost effect within a football context where higher transfer fees could predict more playing time for players in different positions. The effect was robust to a number of specifications. It was associated with incremental increases in transfer fees as well as a distinctive effect of expensive players in particular. More importantly, the effect persisted even after a managerial change. This contradicts a rational decision-making standpoint where assigning a new manager should remove any sunk cost effect demonstrated by his predecessor (e.g. Wicklund & Brehm, 1976). To rule out the possibility of free transferred players biasing the results (as they do not incur a large sunk cost element), we tested the sunk cost hypothesis while removing those players from the sample. The effect remained robust, though with slightly lower magnitude. We also showed evidence for managers' bias towards granting free transferred players more playing time as a way to show their skills in securing bargain deals for their clubs (e.g. Schindler, 1998). Lastly, we estimated the RD model to test the effect of expensive players. The results converged to support the sunk cost hypothesis.

### 2.1 Literature and Hypotheses

The economics of professional team sports has attracted academic interest since the mid-1950s, with most literature originating in the US. Both theoretical and econometric aspects of the discipline have been explored in the literature, as many economists consider professional sports to be a legitimate area for scientific research. Articles by Rottenberg (1956) and Neale (1964) are regarded as fundamental to sports research, while Sloane (1971) influenced football economics research in particular. These studies covered broad economic aspects and implications of the sports market structure (e.g. free agency in players' labour market, nature of production, utility maximisation). There were also early contributions to data science in football, such as forecasting models for football match results (e.g. Moroney, 1956; Reep et al., 1971) and game theory and football games (e.g. Chiappori et al., 2002).

Abolition of the maximum wage in English football in 1961 and the introduction of the Bosman rule through the European Court of Justice in 1995 (i.e. free agency contracts) provided the basis for a fundamental shift in players' salaries. Additionally, since introducing TV broadcasting rights for English football matches in the late 1980s, footballers' salaries and transfer fees have been escalating. The annual fee for exclusive TV rights went from £191 million in 1992 to as much as £5.1 billion in 2016 (see Table 11).

The shift in salaries and compensation patterns stimulated a new wave of professional sports research. Theoretical contributions to sports stars' salaries determinations were proposed by economists (e.g. Lazear & Rosen, 1981; Rosen, 1981). Data on players' salaries in European professional football are scarce due to confidentiality and disclosure restrictions, so the scope for empirical evidence is rather thin. A number of studies attempted to determine factors affecting players' salaries in Europe (e.g. Frick & Deutscher, 2009; Lehmann & Schulze, 2008; Lucifora & Simmons, 2003). Collectively, some performance measures were found to be significant in determining salaries, such as goal scoring and assists, matches played, and experience. Lucifora and Simmons (2003) found that goal scoring rate and assistance were associated with highly convex salaries.

### 2.1.1 Literature review

A rational economic decision suggests that future benefits should exceed future costs (Vroom, 1964). It is expected from a rational economic perspective that allocating resources would follow that principle. The traditional models of Edwards (1954) and Vroom (1964) provide a basis for these economic-based decisions. The rationality of decisions has been challenged by the psychology literature that is related to decision-making processes, and individuals' tendencies to rectify past failures or losses were scrutinised. Staw (1976) provided evidence on the motives behind people's tendencies towards rationalising the negative outcomes (or feedback) that were consequences of their initial decisions. Rationalising a poor decision is manifested in an individual's willingness to increase commitment in a failing course of action, labelled as escalation of commitment (EoC). Such behaviour increases after an initial decision receives negative feedback and more resources are likely to be invested in an attempt to recover any losses incurred, in addition to self-justification motives. Staw and Ross (1987) considered those predicaments along with the uncertainty of the consequences of persistence or withdrawal from the course of action as characteristics of an escalation situation.

The progress of Staw's initial studies reveals the complex nature of EoC processes. Self-justification is a predominant theory attempting to explain the EoC phenomenon (see Wicklund & Brehm, 1976). Staw's research (e.g. Staw & Fox, 1977; Staw & Ross, 1978) provide support for self-justification as a factor influencing EoC, and that justification of previous decisions is a function of personal responsibility. Schaubroeck and Williams (1993) found that factors affecting escalation of commitment were stronger under personally responsible conditions. However, the appeal of self-justification was reduced following a prolonged course of action. This led to the use of external justification in addition to the established self or internal justification (Staw & Fox, 1979). That distinction considers exogenous causes of a setback, such as job loss or stiff public resistance to policies. They found that EoC was influenced by exogenous factors, which are external demonstrations of competence.

Another source of escalating commitment is norms of consistency. Staw and Ross (1980) showed that being consistent in a course of action with ultimate success leads to escalating commitment. Their political context survey demonstrates that being consistent was generally a perception of the public, or the cultural norms surrounding decision makers.

Whyte (1986) and Bowen (1987) were opponents of the self-justification explanation of escalating commitments. Whyte (1986) argued that the prospect theory of Kahneman and Tversky (1979) provided a better explanation for decisions involving escalating commitments. The theory states that there is a reference point from which a decision maker will be risk averse in the domain of gains and risk taking in the domain of losses. Hence, the decision will be directly linked to that reference point rather than rational factors. Both prospect theory and self-justification seem to explain escalation when it is operationalised under a negatively framed condition (e.g. half empty bottle) with high personal responsibility. Nevertheless, the findings of Davis and Bobko (1986) that escalation would occur in some positively framed conditions with no effect of personal responsibility gives greater support to self-justification than prospect theory.

It was suggested by Northcraft and Neale (1986) that explicit consideration of opportunity cost alters the framing of situations, leading to more rational decisions. They illustrated that confining the choices between certain loss and low probability of no loss or small gain was a misleading setup. Though this illustration is not intended to support self-justification, it points to possible weaknesses of prospect theory.

Bowen (1987) proposed a decision dilemma theory and suggested that factors such as economic considerations, curiosity, need for greater efforts to reach success, and learning curve during the course of action are likely to influence decision makers, rather than their need for self-justification. Bowen's view was that participants in the previous escalation research had not clearly received negative feedback with respect to their initial resource allocation decisions. However, some studies explicitly showed the negative feedback experienced by decision makers (e.g. Conlon & Parks, 1987).

Brockner (1992) reviewed evidence supporting self-justification theory. He concluded that escalation behaviour was well explained by self-justification. The escalation studies used different strategies (i.e. differentiated control variables and manipulations) to explain the behaviour. This led Brockner to believe that the evidence converged to validate the self-justification explanation, and he viewed other theoretical perspectives as complementary to the self-justification perspective, adding more insight to the explanation of the behaviour. Staw and Ross (1987) showed that the trade-off between economic factors and social factors depended on the stage of the decision-making process. Based on their argument, the initial stages of a course of action are likely to account for economic considerations, whereas later stages are likely to involve behavioural characteristics within the decision process.

Arkes and Blumer (1985) introduced another cognitive aspect to explain the behaviour: the sunk cost effect (SCE). They defined SCE as ‘a greater tendency to continue an endeavour once a non-refundable investment in money, effort or time has been made’. Because the decision maker does not wish to appear to be wasting resources, this becomes maladaptive behaviour. Arkes conducted a well-known experiment concerning theatre season tickets. Individuals who bought the tickets without a discount, which was provided randomly to others, were significantly more inclined to attend the theatre more often than those who bought the ticket at a discount. Arkes concluded that a sort of mental appreciation of sunk cost was experienced by those who bought tickets at full price. Several other experiments were conducted using different contexts (e.g. aeroplane radar, ski trips), which all revealed behaviour indicating a need to avoid waste. Arkes (1996) maintained that SCE was mainly explained by the this need, and the motive for justifying sunk costs was a continuous aspect during operations (Ayton & Arkes, 1998).

The above studies attempt to explain the phenomenon of escalating commitments. Arkes and Blumer described it as a sunk cost effect. Decision makers’ tendencies to escalate commitments are likely to be grounded in sunk costs, whether it be money or time and effort. Hence, we may use both terminologies interchangeably. Research experiments are preceded by a sunk cost or loss scenario, followed by behavioural reasoning attempts such as personal responsibility, framing (i.e. processing of information), or other cognitive aspects that are being manipulated within these experiments. Thus, theoretical perspectives attempting to explain the phenomenon are embedded with sunk costs while different strategies are employed. The norm ‘too much is expended to quit’ is likely to be an antecedent to behavioural reasoning. Hence, sunk cost is strongly associated with escalation of commitment (EoC), either as a main driver leading to escalation, or as an antecedent to needs for self-justification or need not to waste. Arkes and Ayton (1999) were of the view that sunk cost works with EoC as one entity; this can be understood from their statement, 'Justification plays another role in the analysis of the sunk cost effect'[1]. Staw and Hoang (1995) considered sunk costs to be central to the escalation question and confronting the assumptions of rational economic decision making.

Staw and Hoang (1995) incorporated real organisational structure in escalation studies, incorporating NBA draft numbers and playing time. A main weakness they addressed was that previous escalation literature had been laboratory-based, which does not easily allow generalisation to natural fields. Thus, their study's main purpose was to fill that particular gap rather than focusing on theoretical processes underlying the behavioural effect.

They found significant sunk cost effects on managers' decisions pertaining to playing time, trading the players, and their continuation in the NBA. High-cost (top draft) players gained longer playing time and were kept longer in the team even after controlling for their performance on court. Yet, Staw and Hoang pointed out that seeking a single theoretical model as a causal explanation for the behaviour was far from reality. A number of psychological processes (e.g. justification, wastefulness, commitment) could have simultaneously contributed to the effect.

Camerer and Weber (1999) re-examined the results of Staw and Hoang (1995). They tested commitment to first round picks and included player trading cards as a proxy for popularity. Substitute players were a salient variable included in their test. In addition, they controlled for performance in previous and current seasons. They also incorporated predicted performance measures as a replacement for observed measures. Their results were generally consistent with Staw and Hoang; they found a positive impact of draft positions on playing time, specifically earlier in the players' career (i.e. second and third year). The first round pick gained significantly higher playing minutes (up to 475 minutes during a season) than the second

[1] p. 597

round pick. However, their later career years showed significantly lower playing time for players than Staw and Hoang's results, which is partly attributable to their added specifications.

Leeds et al. (2015) found contrasting results to the previous NBA studies. They found no evidence of sunk cost (i.e. draft position) effect on playing time. They used a more sophisticated model specification, regression discontinuity (RD), at the cut-off between the two draft rounds, which distinguished their results from Staw and Hoang's. First round players' salaries are based on draft number, whereas second round players' salaries are negotiated with their agents. This raises a possible endogenous effect of commitment, which made the distinction of using an RD method.

Keefer (2017) employed a similar method to Leeds et al. (2015) using playing time in the NFL; however, his results showed an effect of sunk costs on playing time decision. The sample of the study had no salary cap on first round picks (i.e. all salaries were determined through free negotiations). Accordingly, Keefer (2017) suggested that financial commitment may increase when the cost is endogenously determined. This supports the personal responsibility effect when associated with sunk costs (e.g. Bazerman et al., 1982).

### 2.1.2 Hypotheses

Transfer fees are a sunk cost; the higher the fees, the more significant the sunk cost. Classical economics suggests that team managers should play their most productive players (e.g. Leeds et al., 2015) and hence not be influenced by the size of a player's transfer fee. This gives rise to the following sunk cost hypotheses:

H1: Playing time is independent of a player's transfer fee after controlling for a player's performance.

H1A: Playing time is positively related to a player's transfer fee after controlling for a player's performance.

The rationale for the sunk cost hypothesis is that managers do not want to feel they are wasting a large transfer fee spent on a player. Their personal responsibility in signing a player along with negative feedback (e.g. if a player is underperforming) may induce self-justification. They may also be under pressure from fans and pundits to play high transfer fee players (or stars), which could illustrate the external justification or exogenous causes of Staw and Fox (1979). Thaler (2015, p. 284) stated, 'If a team is paying a high draft pick a lot of money, it feels under pressure to put him in the game, regardless of how well he is playing.'[2] Thus, they may be biased towards granting playing time based on sunk costs (transfer fee) incurred rather than on player performance. The implication of the null hypothesis (H1) is that player transfer fees have no effect on their playing time, as the manager is expected to base his decisions purely on player performance metrics. However, if sunk cost affects the manager's decisions, then the transfer fee will be a significant factor in determining a player's playing time. This is the implication of the alternative hypothesis (H1A).

Modifications and alternatives to the sunk cost hypothesis have been proposed. The self-justification explanation (e.g. Wicklund & Brehm, 1976) suggests that the sunk cost effect is removed or at least mitigated (i.e. a negative relation) when a new manager takes over from the manager who completed the transfer. If the sunk cost is considered a mental account, then when the owner of the mental account leaves, the account is closed, and the sunk cost effect is removed. This is consistent with personal responsibility, as mental accounting applies to

[2] Draft pick is a term related to the National Basketball Association (NBA). It has a different structure than the football transfer window in which players are ranked based on their qualities. Teams can trade those draft picks following sets of financial caps and lottery pick specifications by the NBA.

whoever bears the costs. We explore the frequent turnover of managers within teams to test this conjecture.

H2: Playing time is independent of a player transfer fee under a manager who was not in place when the transfer was completed, after controlling for player performance.

H2A: Playing time is positively related to a player transfer fee even under a manager who was not in place when the transfer was completed, after controlling for player performance.

The implication of H2 hypothesis is that the sunk cost effect (transfer fee) is mitigated under a manager different from the one under whom the transfer was completed.

There are players who run out of contract or may be highly established in their clubs (i.e. club legends) who progressed from the club youth system (e.g. Steven Gerrard in Liverpool, Ryan Giggs in Manchester United, Harry Kane in Tottenham). In both cases, the player does not add a sunk cost to his new team. If a player’s contract expires and his club (team) does not wish to renew it, then he will be able to move to another team without a transfer fee. Similarly, a player who progresses from a club youth system (or academy) will sign a contract that is free of a transfer fee for the first team of the club. Free transfers do not necessarily indicate that players are mediocre. For instance, Academy players signing for the same club indicate that they are within the required team standards. They might get more playing time than other established players once their performance is superior. This contrasts with NBA drafted players, where players drafted on the first round get more playing time than those from the second round, and the lottery-pick players play more than all of them (Leeds et al., 2015). NBA draft positions reflect player qualities while this is not necessarily the case with football while considering a free transfer condition. This allows us to test the conjecture that a free transfer is a bargain deal for the club, where established or academy players may provide a high performance level for a zero sunk cost.

The average playing time for free-of-cost players (who form about 12.4% of the data sample) was 1720 minutes, compared to 1794 minutes for others. The test of mean difference shows that this was statistically insignificant. Nevertheless, free transferred (or free-of-cost) strikers and defenders played on average fewer minutes than their non-free peers (1350 minutes compared to 1552 minutes and 1815 minutes compared to 2015 minutes, respectively), whereas the averages for free and non-free midfielders were very close (1721 minutes compared to 1730 minutes). Additionally, free players had on average a lower injury rate than non-free players (2.57 games compared to 3.27 games), and this was statistically significant. This shows that free-cost players played regularly, and a test for reverse sunk cost hypothesis is viable.

The psychological value of a bargain has been discussed in consumer psychology literature. The smart shopper hypothesis is an attempt to explain the value of a bargain, and proposes that ego-expressive goals may be satisfied by getting a discount or a bargain deal. Schindler (1998) suggested that the consequence of getting a bargain could lead to a noneconomic component whereby consumers (managers in our case) would attribute the success of a bargain deal to their own skills. This implies managers who have a successful bargain will communicate their deal through the players' playing time. This is based on the assumption that managers select their team players. There might be instances where club owners or other technical staff select players; however, there is no reliable way to address that issue. Despite that, the general perception is that managers are usually approached before sealing a player deal.

A rational economic decision suggests that player performance should determine his playing time rather than his bargain deal. This is a reverse sunk cost scenario. Thus, we test the following hypothesis:

H3: Playing time is independent of a free transfer player after controlling for a player's performance.

H3A: Playing time is positively related to a free transfer player after controlling for a player's performance.

## 2.2 Empirical Analysis

### 2.2.1 Data and sample

The sample included all Premier League (PL) players who played a minimum of nine games (appearances) in any season for the period spanning 2006/07 to 2016/17. That period is used for strikers, while we confined the observations of midfielders and defenders to the period spanning 2012/13 to 2016/17. We opted to collect a longer sample period for strikers due to the fact that this position was held by a smaller number of players. There were more players in non-striker positions; thus, five seasons provided sufficient data to conduct the analysis. Each season spent by a player at a club was considered an observation. The maximum number of minutes a player can play in a single season is 3420 (90 minutes per game over a total of 38 games). The season runs from approximately mid-August to mid-May, during which a winter transfer window opens in January (mid-season window), in addition to the summer window (pre-season window).[3]

A minimum of nine games (appearances) amounted to approximately 24% of the season's games. The average player's appearance was 24, with a standard deviation of 8. Players who made a minimum of nine appearances had an average of 440 minutes of playing time. This is sufficient time to play frequently, at least as a substitute player. Players who

[3] Pre-season window spans usually from beginning of June to end of August, while mid-season window spans from 1st January to 31st January (https://www.fifatms.com/itms/worldwide-transfer-windows-calendar/).

appeared less than the average (i.e. less than 24 appearances) had more injuries than those who played more than the average. Those had an average of roughly five games injured, with a standard deviation of 6.4 and a maximum of 29 games. On the other hand, players who made more than the average number of appearances had 1.54 average games injured with a standard deviation of 2.6 and a maximum of 13 games. This suggests that injury had a considerable effect on players who played a below average number of games. Thus, we consider the minimum of nine appearances as a threshold by which a player becomes established within a team squad (see Tables 7 and 8 for PT averages). The sample excludes observations of players who had injuries for a full season and goalkeepers.[4]

We used a number of data sources (websites) to obtain the dataset, including Transfermarkt.co.uk for playing time metrics, Premierleague.com for players' performance metrics, and Soccerbase.com for managerial turnover records, players' transfer fees, and signing dates. The provided performance metrics are essentially derived from Opta sports, which is the leading sports data provider for major sports entities.

### 2.2.2 Econometric model

We ran the following panel regression models to test our hypotheses:

$$PT_{it} = \alpha + \beta_1\, TF_{it} + \beta_2\, TF_{it}\, D_{1it} + \gamma Perf_{it} + \boldsymbol{\delta}\, Controls_{it} + \boldsymbol{\eta_i} + \boldsymbol{\lambda} C_{\boldsymbol{i}} + \boldsymbol{\theta}\mathrm{V}_{\boldsymbol{t}} + \psi FT + \boldsymbol{\varepsilon_{it}} \qquad \textbf{(1)}$$

$$PT_{it} = \alpha + \beta_1\, TFhigh_{it} + \beta_2\, TFhigh_{it}\, D_{1it} + \gamma Perf_{it} + \boldsymbol{\delta}\, Controls_{it} + \boldsymbol{\eta_i} + \boldsymbol{\lambda} C_{\boldsymbol{i}} + \boldsymbol{\theta}\mathrm{V}_{\boldsymbol{t}} + \psi FT + \boldsymbol{\varepsilon_{it}} \qquad \textbf{(2)}$$

where $PT_{it}$ is the dependent variable representing playing time in minutes and in a separate panel—within the same regression model—as a percentage of total possible minutes for player *i* in season *t*. The percentage of *PT* is calculated by dividing the minutes played by the total number of theoretically possible playing minutes. Those possible minutes are calculated by

[4] Goalkeeper position is occupied by only one player. It is rare to substitute him unless an injury occurs.

subtracting suspension, injury, on-loan, and winter window game minutes from the season total minutes of 3420. We excluded added time (or injury time), which is usually added to each game's total minutes (90 minutes), from the total possible minutes.

The $TF_{it}$ variable in model (1) represents the transfer fee and captures the potential sunk cost effect. The transfer fee is adjusted for inflation for each player $i$ in all seasons (see Tables 4, 5, and 6 for TF spend averages). Discounting transfer fee is not applied, as its effect could be problematic due to the following: First, a player's value could increase or decrease during his contract. Second, managers tend to refer to the original transfer fee when evaluating their players.[5] This is consistent with the agent reference point of Kahneman and Tversky (1979). The expectation is that the $TF_{it}$ coefficient $\beta_1$ will be positive according to hypothesis H1A. However, there is at least one scenario where the sunk cost effect may be mitigated, and this is captured by $TF_{it}$ interacted with a dummy. The $TF_{it}D_{1it}$ variable reflects the transfer fee conditional on managerial change in the transferred player's team where $D_{1it}$ takes the value of 1 for the period during which a new manager is in charge. This is coded for each player once a team's manager is changed after his signing date in that season. The expectation is that the $\beta_2$ coefficient will be negative according to hypothesis H2, as the sunk cost effect should not apply in this case. If $\beta_2$ is insignificant, then the implication is that a positive $\beta_1$ coefficient (if any) is not influenced and hence the sunk cost is still positively related to playing time even after assigning a new manager to the team.

The $TFhigh_{it}$ variable in model (2) is a dummy that represents the transfer fee in the top quartile. It takes a value of 1 if the transfer fee of a player is within the top quartile values and 0 otherwise. This captures the potential sunk cost effect of expensive players. We used the dummy to test any distinctive effect of expensive players, which is analogous to the first and

[5] Sir Alex Ferguson manifested this kind of thinking for managers in his books Autobiography and Leading.

second draft picks of other sports fields (e.g. NBA, NFL). The draft ranks the players based on their qualities and subsequently their related financial commitments, and these escalate with higher draft positions. The first round contains top players, while the second contains players with lower quality and lower financial commitment. Expensive players in football often have better qualities than the less expensive players (see Table 15 for comparison statistics). After controlling for their performance, classical economics predicts that there should be no effect of sunk costs (transfer fees) incurred. The expectation of $\beta_1$ and $\beta_2$ coefficients are the same as model (1) coefficients. The difference between models (1) and (2) lies in the econometric treatment of the transfer fee, where (1) takes into account each increment in TF in the cross-section but (2) takes into account the effect of the top quartile TF. The rest of the variables apply the same way for models (1) and (2) as follows.

The $Perf_{it}$ variable captures a player's performance, and the $\gamma$ coefficient is the effect on playing time (*PT*). A player's performance is measured by several metrics, and we ran a separate regression for each position (strikers, defenders, and midfielders). Goals, assists, and passes were the main performance metrics applied for all positions, excluding assists from defenders. The other metrics—crosses, tackles, interceptions, and duels-lost—differed slightly. Performance metrics were selected based on their correlation coefficients, along with principal component analysis (PCA) (see Tables 9 and 10). Based on that, the selected variables in each position explained more than 67% of the variation in all performance metrics. We chose not to construct a performance index based on the PCA to reflect the distinctive variables effect. In addition to the main performance metrics mentioned above, tackles and crosses were assigned to strikers, tackles, and interceptions for midfielders and defenders, respectively, and duels-lost for both midfielders and defenders.

A main difference between our model and the NBA-related models (e.g. Staw & Hoang, 1995; Leeds et al., 2015) is that NBA players' performance is captured, to a considerable extent, by their experience. The NBA draft structure guarantees certain contract durations for players, which allows for a more balanced data panel. This is not the case in a football context due to the higher frequency of trading players. Football clubs may receive offers for their players that cannot be ignored due to their profit potential, whereas the financial caps imposed in the NBA and NFL prevent the clubs from that potential. The NBA salary cap could be exceeded in some cases, but with a luxury tax to be paid. The consequent limitation to our model is the lack of proper lagging performance specifications, especially if a player moves out of the EPL. Thus, we confine the analysis to current season performance metrics.

The *Controls* is a vector that includes the following variables: The first is the number of games for which a player was injured, suspended and/or on loan. The second is a dummy for winter window transfers (WW) to account for the lower possible playing time for players being transferred in that window, as they did not participate in the first half of the season with their new team. Hence, those observations are included in the sample after the winter transfer takes place. (See Table 16 for definitions of performance metrics and control variables.)

Lastly, $\eta$ is the fixed effect for players. The *C* variable is a dummy to capture club effects. The *V* variable is a dummy to capture the seasonal trend of *TF*. The *FT* variable is a dummy representing free transferred players to capture the effect of players who went out of contract or academy players.

### 2.3 Regression Results

We estimated fixed effects (FE) panel regressions for models (1) and (2). The *TF* is the transfer fee value in £million and captures the potential sunk cost effect. The *TFhigh* is a dummy that represents the top quartile TF and captures the potential sunk cost effect of expensive players. This takes a value of 1 for any player whose transfer fee lies in the top quartile. For strikers and midfielders, the top quartile TF value starts from £10m, while it starts from £6m for defenders. The £10m value represents a psychological level, as it is the first two digits TF in millions. It is in sync with the top quartile starting values. The lower starting value for defenders may be attributed to the football industry valuation of this position along with goalkeepers. Their performance measures are less clear than those of strikers and midfielders, where goals and assists are the major contributions to the team's success.

We estimated a second panel regression (Panel B) within each model but without including free transferred and club academy system players. This is a robustness test to rule out any potential influence on the results that may be caused by zero sunk cost players, and is based on the argument that free transfers (either out of contract or academy players) lack a sunk cost element and thus are not relevant to the EoC subject and its estimation model. In a rational world, it is expected that only playing performance determines playing time. Hence, it was expected that the regressions would yield significant coefficients for performance and control variables and insignificant coefficients for all other variables.

We included the win ratio of clubs as a subsample to represent distinctive club characteristics. Higher paid players often possess stronger qualities. It was therefore expected that clubs with higher spending would yield higher win ratios (Figure 3). This allowed us to test whether winning clubs appreciated sunk costs differently. The average win rate was 37%. We conducted a chi-square test for homogeneity in playing time distribution among two groups of teams (i.e. winning vs. non-winning), where winning teams had above average wins and non-winning teams had below average wins. We defined playing time as above average when a player played more than 50.3% (the median of playing time) of his possible playing time, and below average when it was less than the median. The null hypothesis was not rejected. This indicates that the two groups of teams had a similar playing time distribution.

### 2.3.1 Striker results

The striker regression results are presented in Table 1. The dependent variable of these results is playing time in terms of percentage of minutes, whereas the results in terms of played minutes are presented separately in Table 12. The results of models (1) and (2) indicate that the transfer fee *TF* for strikers was positively related to their playing time. The results are statistically significant for both minutes and percentage of minutes terms. These results support the H1A sunk costs hypothesis. The economic impact of higher fees on playing time *PT* was strong. Model (2) indicates that top TF quartile players play on average an extra 14.8% of their possible playing time per season. In terms of minutes, (2) indicates 425 more minutes played per season. This is roughly equivalent to five more games per season on average. The results of Staw and Hoang (1995) indicated that a second round drafted player played on average 552 fewer minutes in his second year in the NBA (in NBA, the first-round players are at a higher skill level than the second-round players and the higher the draft position—within each round—the lower the player quality). The NBA season consists of 82 games with a total of 3936 minutes, which is 15.1% more than the total minutes of the EPL football season. Despite the differences in the two sporting fields, our results were fairly close to those of Staw and Hoang (1995) once we considered top quartile players in football to be similar to first round drafted players in the NBA in terms of performance quality. The performance metrics of players in the top TF quartile suggest that they possessed stronger qualities (e.g. their goals mean was 0.447 per 90 minutes relative to 0.307 for lower quartiles, and the assist mean was 0.192 relative to 0.141).

These results are supported by the more general model (1) results, which account for each increment in the TF. They indicate that an additional £1m in *TF* for a player resulted in 0.68% extra playing time per season (or 18 minutes). Staw and Hoang's (1995) results showed that each increment in draft position in the NBA led to 23 fewer playing minutes in the second season. This is fairly close to (1) results. We may express the results in the following way: a player who cost £10 million will have on average, all else equal, 6.8% more playing time than a player who cost only £1 million. This is roughly equivalent to 2.5 more games per season on average. The results of (1) and (2) converge to provide evidence of the considerable economic impact of the higher transfer fees on playing time. Within (2), the transfer fee showed the highest economic impact of all variables. This is a striking result considering that player performance should be the ultimate time predictor from a rational standpoint.

The results were robust to managerial changes in teams. A new manager can be expected to alter his new team players. We ran the regression first with only a managerial change variable[6] (i.e. no interaction with the transfer fee). The variable was significant at the 5% level with a high negative sign of 0.042. This indicates less playing time, on average, when a new manager was assigned. However, the interaction variable *TFD1* was not significant. This suggests that a new manager's behaviour can alter with increasing transfer fee and hence may still hold the mental account of his predecessor where the feeling of not wanting to waste a large fee on a player persists. This is consistent with H2A, and playing time was still positively related to transfer fees even after a new manager was assigned. The self-explanation was surpassed with a new manager bias towards keeping an old mental account.

The results were also robust to the removal of free transferred and academy players from the estimation model. These players did not add a sunk cost element to their teams (TF=0). The expectation of the sunk cost hypothesis is that managers choose to play high TF players ahead of low TF players, including where TF might equal zero. However, panel B results show that the TF effect remained positive even after removing players with a zero sunk cost from the estimation model, where they accounted for 16.6% of the sample. The magnitude of the TF

[6] Not reported in the tables.

coefficient was, however, slightly reduced to 13% from 14.8% in (2) and reduced to 0.59% from 0.68% in (1).

The *FREE-TF* variable was significant and positively related to playing time for teams with a less than average win ratio (the average win ratio for teams was 37.6%). We tagged these teams as non-winning teams. The economic impact of the variable was strong, with at least 19% extra playing time per season. These teams seemed to highly appreciate players with a bargain deal either through free transfer or through the club academy system. The results in columns 2 and 5 of panel A support the bargain deal hypothesis H3A. Managers of non-winning teams may grant those players more playing time as a way to show their own skills of signing zero sunk cost players with good qualities.

Within the *Perf* variable, Models (1) and (2) show that a striker's *PASS* was the most statistically significant factor in predicting his playing time. It had a negative coefficient of 0.39 in percentage terms and (2) confirmed it with a relatively similar magnitude of 0.34. A striker's scoring performance *GOAL* was another main performance predictor of his playing time. The results indicate that goals were positively related to playing time, which is quite intuitive. The coefficient was significant at the 5% level in percentage terms and its economic impact was strong, where the percentage of extra minutes played was 10.6% in (1) and 10.3% in (2). This is a higher economic impact than TF within (1) but not within (2). This suggests the crucial importance of a top-paid player to the managers, sometimes to an irrational extent.

From the rest of the performance metrics, *TACKLE* shows a significant statistical and economic impact. It indicates a negative relationship with playing time, where playing time of a player decreased by 5.6% in percentage terms for each increment in his tackles. The negative coefficients in passes and tackles may be explained by the different playing characteristics among strikers. Expensive players cost more than £10m and the rest cost less, but they formed

74% of the sample. The latter players played less on average (1461 minutes compared to 1798 minutes); however, they had 10% more tackles than expensive ones (1.09 compared to 0.99). On the other hand, they had fewer passes than expensive players (27 compared to 30) but with much higher standard deviation and skewness (the maximum number of passes for expensive players was just 60, while the less expensive players had a maximum of 174). Panel (B) shows a higher negative *TACKLE* coefficient of 7.1%. This could imply that the higher the fees of a striker, the more he deviated from defensive roles. The results indicate that *ASSIST* was an insignificant factor. This is attributed to its high standard error relative to goals.

The TF results were driven by both below-and above-average winning teams. Most of these subsamples' coefficients were significant at least at the 5% level, except for panel B in model (2) non-winning teams where it was insignificant, and for panel A in model (1) where it was significant only at the 10% level. This provides restrained evidence that winning teams might be more influenced by higher fees.

Season effect coefficients were negative and highly significant from 2013 onwards (see Table 12). There has been a higher spending trend for players since 2013, which mainly relates to new TV broadcasting rights deals (see Tables 6 and 11 for trends of TF spending and TV deals). That higher spending may induce fiercer competition among players to gain more playing time, and thus the possibility of reaching that aim may be reduced.

### 2.3.2 Midfielders results

Midfielders' results are presented in Table 2. The dependent variable of these results is playing time in terms of percentage of minutes, and the results in terms of played minutes are presented separately in Table 13. The results of models (1) and (2) collectively indicate that transfer fee *TF* for midfielders, in most cases, was not related to their playing time. This does not support the alternative H1A sunk costs hypothesis.

Panel B estimation results show that top quartile TF was highly significant for non-winning teams with a strong coefficient of 31.8%. Managers of these teams seemed to appreciate their midfielder sunk costs differently. However, this effect was totally mitigated by the managerial change variable TFD1 with a strong negative coefficient of 43.9%. This suggests that a managerial turnover was sufficient to remove the sunk cost effect for midfielders. Additionally, the TF was significant in column (1) Panel A, but only at the 10% level, and this effect was also mitigated by TFD1. These results support the H2A hypothesis.

The TFD1 variable was highly significant within the top quartile results, despite the fact that the TF variable was insignificant. This supports the evidence that a new manager will close his predecessor's mental account, thus removing the sunk cost effect. This gives rise to the self-justification phenomenon of former managers. These results were more mixed and weaker than strikers' results. This may be attributed to having more midfielders in the team squad than strikers and simultaneously having a variety of playing roles (e.g. defensive, attacking, left or right wings, wide or deep position). This creates more options for team managers in the sense that the probability of player rotation could be higher.

The *FREE-TF* variable was significant and positively related to playing time in model (1). The economic impact of the variable was strong, with 13% extra playing time per season. Model (2) results were also significant but only at the 10% level. This converges with striker results to support H3A hypothesis pertaining to bargain deal bias.

The performance metrics *GOAL*, *DUEL-LOST,* and *TACKLE* were shown to be significant playing time predictors. The results indicate that the *GOAL* variable was positively related to midfielders' playing time. Most of its coefficients were statistically and economically significant within all panels. Each extra goal scored indicated roughly 30% extra time per season in model (1) and less magnitude of roughly 21% in model (2). Each *DUEL-LOST* implied 5.18% less playing time per season. The negative coefficient is intuitive, as losing the ball more often in duels can hurt teams. *TACKLE* was significant at the 10% level in percentage terms but highly significant in minutes terms. Its coefficient indicates roughly 3% less playing time per season. It is more puzzling than the duels-lost. It may be due to the unsuccessful tackles (they usually account for a quarter to a third of the total tackles) or the role of the midfielder, which may or may not involve defensive tasks. The *ASSIST* variable was significant and negatively related to playing time only in panel B, but it was not supported by either panel A or the results in minutes terms, where it was insignificant.

The season dummy coefficients were negative and highly significant from 2013 onwards (see Table 13). This is consistent with striker results.

### 2.3.3 Defenders results

The results are presented in Table 3. The dependent variable of these results is playing time in terms of percentage of minutes, and the results in terms of played minutes are presented separately in Table 14. The results indicate that the transfer fee for defenders was positively related to playing time minutes. They were highly significant in all model panels and support the H1A sunk costs hypothesis. The economic impact of higher fees on playing time *PT* was strong within models (1) and (2). The results in model (1) indicate that each incremental increase in *TF* for a player resulted in 1.63% extra playing time per season, while in model (2) it resulted in 22.8% extra playing time. These were driven by the winning teams, and their coefficient in (2) was 47%. This is an immense magnitude and suggests that playing time for defenders in winning teams was mainly driven by their transfer fees. This is supported by their insignificant intercept, which was an exception within all panel results reported. Defenders played more on average than midfielders (average minutes of 1998 compared to 1706) with a lower standard deviation, and this is consistent with the common practice of football managers, with defenders usually less often rotated and substituted during games than midfielders. This may explain why the defenders' results were much clearer than those of midfielders.

There is some evidence that this sunk cost effect could be partially mitigated following a managerial change. The results of model (2), which are reported in minutes terms, show a significant negative *TFD1* variable with a coefficient of 303 minutes. This reduces about 20% of the effect magnitude and is consistent with H2 hypothesis.

The *PASS* and *DUEL-LOST* are significant performance metrics to predict playing time for winning teams with coefficients of 0.76% and -5.06%, respectively. The negative *DUEL-LOST* coefficient is intuitive, as losing the ball more often in duels can hurt teams. Non-winning teams' players have no significant variable to predict their play time except for the intercept with a magnitude of 70.4% in (1) and 87.9% in (2). This implies that their squads are less vulnerable to changes within the defender position. This is supported by their insignificant *ON-LOAN* variable (Table 14). The *TACKLES* variable for defenders was replaced by the *INTERCEPTION* variable due to their high correlation coefficient of 0.40; however, it was not significant.

Panel B results indicate that the *GOAL* variable was significant and positively related to playing time in winning teams. Its economic impact was enormous, as an extra goal scored indicated 69% extra playing time per season. This result mitigates the effect of TF within the winning teams panel; however, the effect remained robust for the overall sample.

The *FREE-TF* variable was insignificant. This may be attributed to the higher percentage of free transferred players relative to other positions (accounting for 25.7% of the sample) and the lower cost for a defender in general. This may make the bargain bias less appealing. The result supports the H3 hypothesis, where a free transfer is not related to playing time. On the other hand, this may explain the *GOAL* results in Panel B, where the removal of free transferred players emphasised the importance of scoring to the winning teams. The season dummy coefficients were negative and highly significant from 2013 onwards (see Table 14).

**2.4 Regression Discontinuity Effect (RD) for Strikers**

The RD effect is presented in Figure 1. We employed RD to test for the expensive strikers' effect. This is an analogous treatment effect to the draft effect of NBA conducted by Leeds et al. (2015) and NFL draft effect by Keefer (2017). The first-round draft pick in the NBA involves significantly higher financial commitments than the second round. In our sample, the intercept gap between the green-fitted line (players worth £10m or more) and the orange-fitted line (players worth less than £10m) represents the regression discontinuity (RD) effect of TF being equal to £10m or more. Thus, we consider the £10m TF as a cut-off line for being an expensive value. The average performance metrics of players in that category are more distinctive than those who are below that category value (see Table 15). However, using a narrower bandwidth around the cut-off line reduces the variation in players' performance. For instance, using a bandwidth of £3m around the cut-off line shows a reduction in the GOAL mean difference from 0.14 (when using the total observations above the cut-off line) to only 0.03. Additionally, the top TF quartile starts from the £10m value. This is an inflation-adjusted value.

The result implies that the £10m TF had a positive effect on playing time while controlling for performance. In other words, strikers who were just above the cut-off line of £10m TF were likely to get more playing time than strikers who were just below that same line. We use the RD method, which relies on the nearest-matching technique and randomises the playing time of roughly similar quality players around the cut-off line (i.e. players below and above the cut-off line). The test result shows that strikers worth £10m or slightly above got 10.7% more playing time than strikers who were worth less than £10m, despite the similarities in players’ productivity. The result was highly significant. This supports Keefer’s (2017) findings, which showed that a 10% increase in the salary cap value of NFL players resulted in an additional 2.7 games starting when using the RD regression.

On the other hand, this generally contradicts the findings of Leeds et al. (2015) concerning NBA players. They employed RD and considered it a main contribution relative to Staw and Hoang’s (1995) seminal paper. Their results showed that there was no draft pick effect on NBA playing time. Hence, the employed RD removed the positive effect of draft order found by Staw and Hoang (1995) and Weber and Camerer (1999). Keefer (2017) indicated that the different results may be explained by the different contract negotiation schemes between the NBA and the NFL. NBA players who are drafted in the first round cannot freely negotiate their contracts’ compensation, while NFL players can do it through their agents with their potential teams. This induces an endogenous financial commitment, and thus the escalation of commitment is relevant when the decision maker bears the sunk cost. This is supportive of the personal responsibility-related findings of Bazerman et al. (1982)., which showed that the decision maker of hiring an employee was more inclined to promote their chosen employee than others, despite similarities in their productivity.

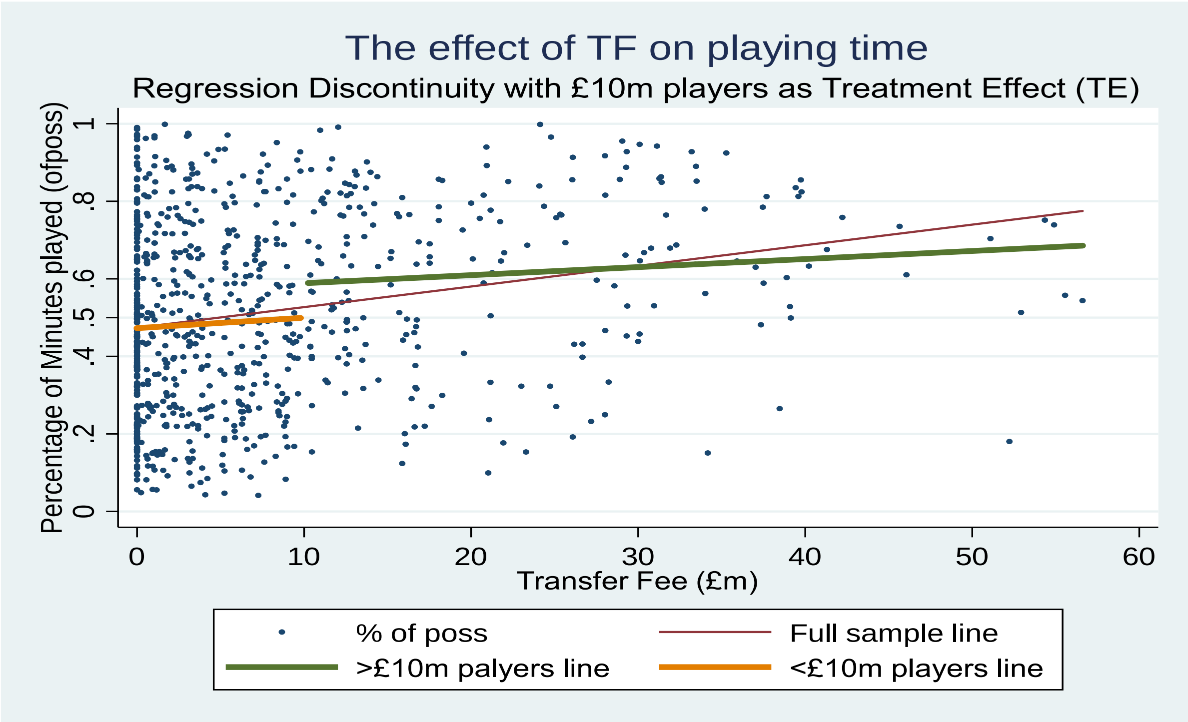
### 2.5 Conclusion

We introduce new evidence that supports the sunk cost hypothesis. We used a novel dataset from the English Premier League to test the effect of transfer fees on playing time. The rationale was that managers should play their most effective players regardless of their cost. However, when a player’s transfer fee escalates, it may induce the manager’s need to not feel like money was wasted on him, in addition to pressure from fans to play the team stars (i.e. high cost players). Hence, those players may get more playing time even after controlling for their performance. Moreover, we explored the frequent manager turnover within clubs to test whether the sunk cost effect could be mitigated. The rationale of this conjecture was that a self-justification explanation of the escalation of commitment or sunk cost effect was removed when a new manager replaced the one who was involved in signing the players (i.e. who bore the sunk cost account).

Our empirical findings support the sunk cost hypothesis. The transfer fee was shown to be positively related to playing time. This was especially true for strikers and defenders. Expensive strikers got on average 14.8% more playing time per season, while expensive defenders got on average 22.8% more. These results were robust to a managerial change in teams and when zero sunk cost players were removed from the estimation model. Managerial change could in some cases mitigate the sunk cost effect. This change reduced expensive defenders' playing time by roughly 20% and by 9.4% for expensive midfielders, even with the absence of a sunk cost effect. The RD estimation model results showed that the £10m transfer fee worked as a burden to teams where players may have had more playing time due to that value privilege. Lastly, we showed the bargain deal bias of managers pertaining to free transferred players.

### 2.6 Tables and Figures

#### 2.6.1 Figure 1




*Notes*: The orange line represents players who cost less than £10m while the green line represents players who cost more than £10m. The gap after the end of the orange line followed by the beginning of the green line on top of it represents the extra percentage of playing time that high-cost players got due to the perception that they were star players.

**2.6.2 Figure 2(a) A scatterplot of minutes played and their relevant TF for the period 2006/07–2016/17 - Strikers**

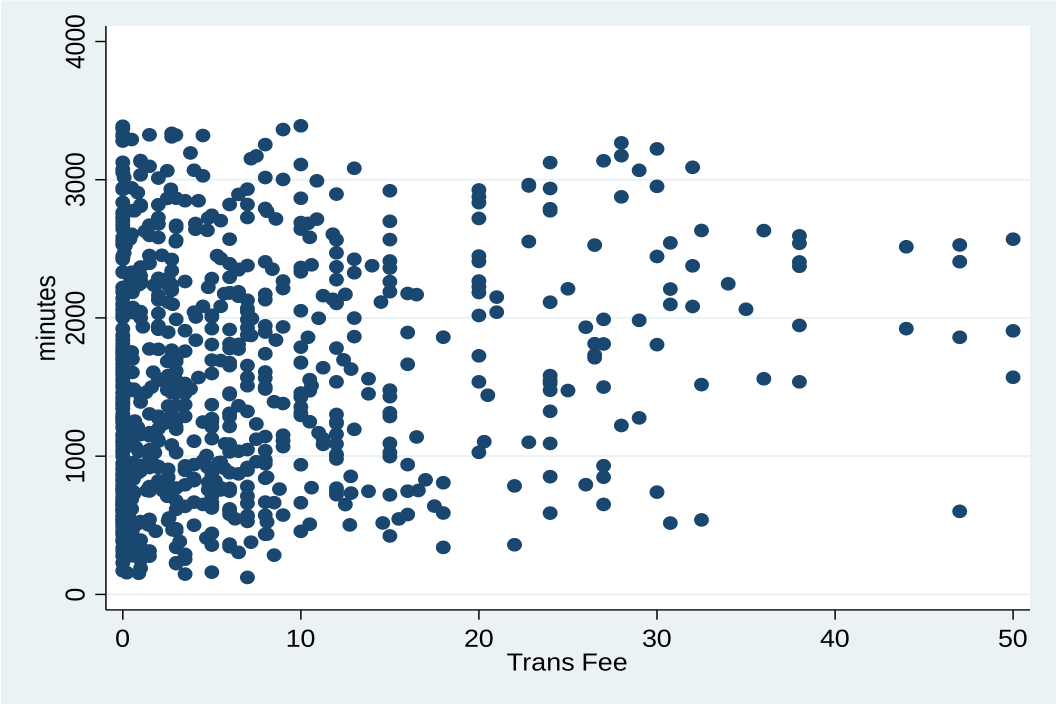
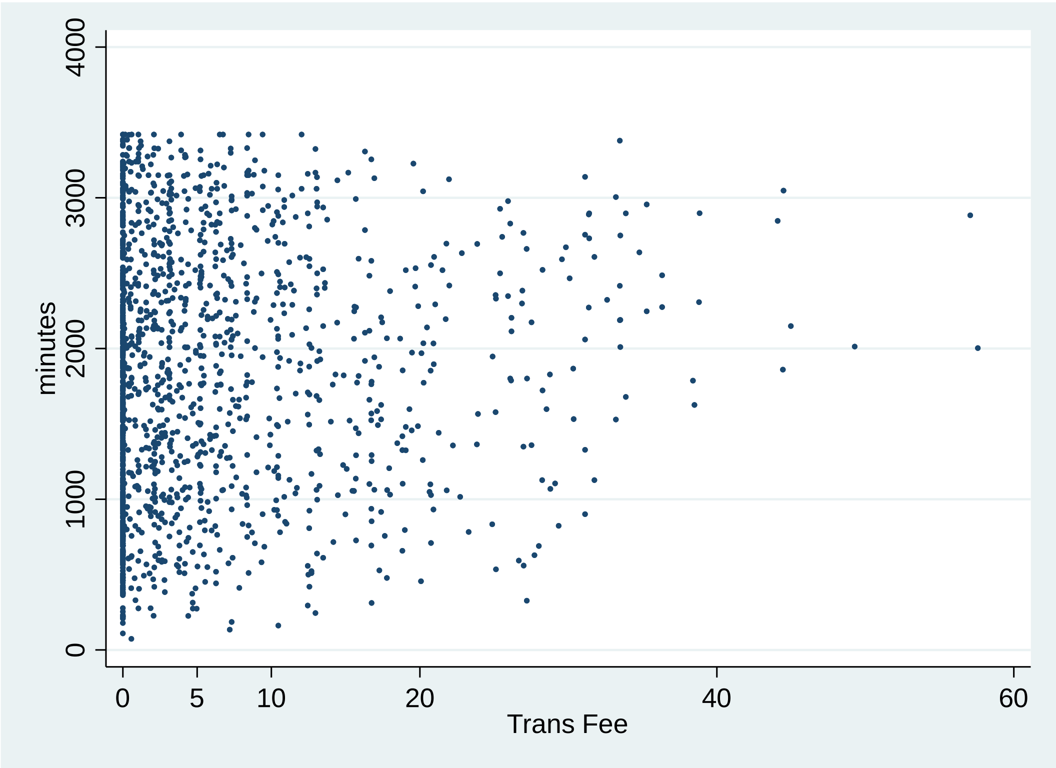
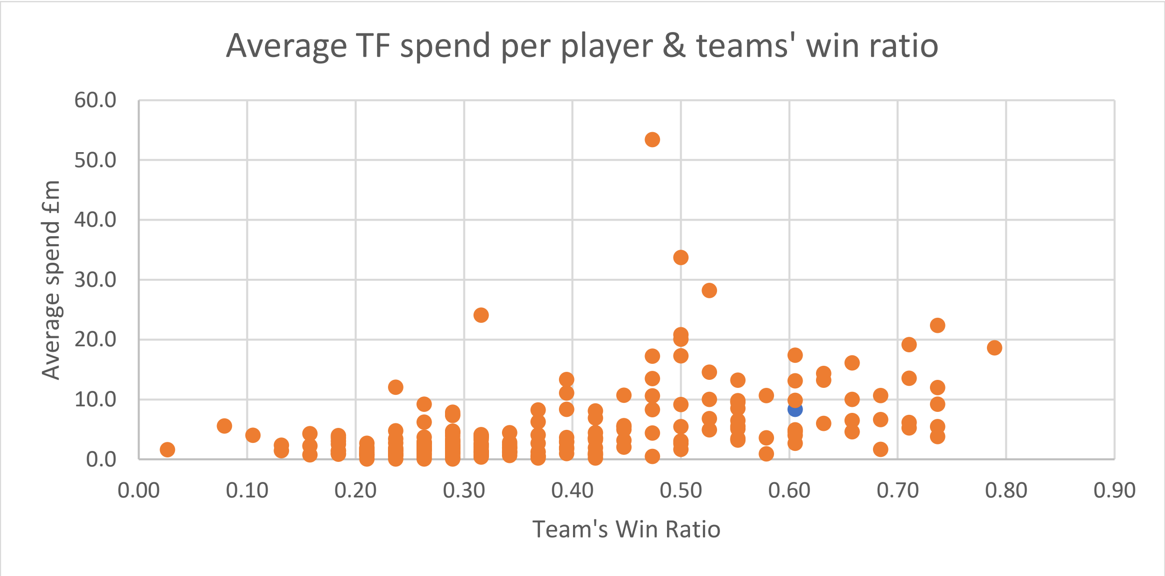



**Figure 2(b) A scatterplot of minutes played and their relevant TF for the period 2012/13–2016/17 - Midfielders and Defenders**



### 2.6.3 Figure 3




*Notes*: The y-axis represents the average transfer fee spent by clubs per player. The x-axis represents the win ratio for teams, calculated by dividing the number of won games by the total number of played games per season.

## 2.6.4 Table 1 Strikers' Results - 2006/07 to 2016/17

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| FE Model | (1) Full Sample | (2) WIN<37% | (3) WIN>37% | (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
| *Panel A: Zero TF player included* | | | | | | |
| TF | 0.0068*** | 0.0088* | 0.0059** | 0.1485*** | 0.1866** | 0.1664*** |
| | (0.0020) | (0.0046) | (0.0030) | (0.0402) | (0.0892) | (0.0600) |
| TFD1 | -0.0022 | -0.0056 | -0.0017 | -0.0288 | -0.0453 | -0.0182 |
| | (0.0015) | (0.0048) | (0.0017) | (0.0367) | (0.0787) | (0.0467) |
| GOAL | 0.1059** | 0.1390* | 0.1226* | 0.1031** | 0.1308 | 0.1304* |
| | (0.0477) | (0.0819) | (0.0690) | (0.0476) | (0.0816) | (0.0683) |
| ASSIST | 0.0635 | 0.2095 | -0.0141 | 0.0594 | 0.1976 | -0.0113 |
| | (0.0833) | (0.1372) | (0.1207) | (0.0831) | (0.1375) | (0.1195) |
| PASS | -0.0039*** | -0.0056 | -0.0070*** | -0.0034*** | -0.0045 | -0.0061** |
| | (0.0013) | (0.0036) | (0.0027) | (0.0013) | (0.0037) | (0.0027) |
| CROSS | -0.0125 | -0.0015 | -0.0012 | -0.0131 | -0.0040 | -0.0017 |
| | (0.0114) | (0.0227) | (0.0170) | (0.0114) | (0.0224) | (0.0168) |
| TACKLE | -0.0566** | -0.0781** | -0.0533 | -0.0572** | -0.0761** | -0.0559 |
| | (0.0243) | (0.0366) | (0.0406) | (0.0242) | (0.0366) | (0.0402) |
| FREE-TF | 0.0162 | 0.2167** | 0.2313 | 0.0161 | 0.1934** | 0.1838 |
| | (0.0594) | (0.0939) | (0.2448) | (0.0590) | (0.0920) | (0.2392) |
| CONSTANT | 0.5398*** | 0.8638*** | 0.7826*** | 0.5408*** | 0.8479*** | 0.7219*** |
| | (0.1009) | (0.1520) | (0.1469) | (0.1002) | (0.1525) | (0.1488) |
| N | 796 | 487 | 309 | 796 | 487 | 309 |
| $R^2$ | 0.2497 | 0.4147 | 0.2771 | 0.2528 | 0.4160 | 0.2927 |
| *Panel B: Zero TF player excluded* | | | | | | |
| TF | 0.0059** | 0.0138** | 0.0075** | 0.1301*** | 0.1749 | 0.2189*** |
| | (0.0025) | (0.0064) | (0.0036) | (0.0471) | (0.1096) | (0.0732) |
| TFD1 | -0.0020 | -0.0065 | -0.0016 | -0.0329 | -0.0535 | -0.0186 |
| | (0.0016) | (0.0052) | (0.0018) | (0.0380) | (0.0837) | (0.0476) |
| GOAL | 0.1062** | 0.1591* | 0.1208 | 0.1051** | 0.1542 | 0.1335* |
| | (0.0534) | (0.0946) | (0.0750) | (0.0533) | (0.0951) | (0.0737) |
| ASSIST | 0.0835 | 0.2261 | 0.0682 | 0.0814 | 0.2063 | 0.0616 |
| | (0.0926) | (0.1622) | (0.1330) | (0.0924) | (0.1631) | (0.1309) |
| PASS | -0.0034** | -0.0038 | -0.0068** | -0.0031** | -0.0033 | -0.0061** |
| | (0.0014) | (0.0042) | (0.0028) | (0.0014) | (0.0043) | (0.0028) |
| CROSS | -0.0151 | -0.0022 | -0.0059 | -0.0161 | -0.0042 | -0.0070 |
| | (0.0127) | (0.0262) | (0.0186) | (0.0127) | (0.0265) | (0.0182) |
| TACKLE | -0.0714** | -0.0991** | -0.0490 | -0.0718** | -0.0997** | -0.0443 |
| | (0.0279) | (0.0422) | (0.0484) | (0.0279) | (0.0425) | (0.0477) |
| CONSTANT | 0.5731*** | 0.7532*** | 0.8308*** | 0.5860*** | 0.7819*** | 0.7557*** |
| | (0.1136) | (0.1959) | (0.1605) | (0.1122) | (0.1965) | (0.1611) |
| N | 664 | 406 | 258 | 664 | 406 | 258 |
| $R^2$ | 0.2650 | 0.4532 | 0.3029 | 0.2676 | 0.4456 | 0.3272 |

* p<0.10, ** p<0.05, *** p<0.01. Fixed effect (FE) panel regressions are estimated. Dependent variable is percentage of played minutes. Panel A presents the results using the full sample data. Panel B presents the results while excluding free transferred and academy players from the sample (i.e. 0 TF players). However, we kept on-loan players in the Panel B sample, as some of them did have loan fees reported. Columns (2), (3), (5), and (6) are win ratio subsamples for both models. They account for different clubs' characteristics. The mean club win ratio was 37.6%. The average spend per player was considerably higher for clubs with a win ratio of more than 37% (see Figure 3 in appendix). Regression including season dummies and played minutes coefficients are presented in Table 12.

## 2.6.5 Table 2 Midfielder Results - 2012/13 to 2016/17

| FE Model | Model 1 (1) Full Sample | (2) WIN<37% | (3) WIN>37% | Model 2 (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
|---|---|---|---|---|---|---|
| *Panel A: Zero TF player included* | | | | | | |
| TF | 0.0051* | 0.0004 | 0.0038 | 0.0530 | 0.0388 | -0.0869 |
| | (0.0031) | (0.0084) | (0.0066) | (0.0558) | (0.1136) | (0.1304) |
| TFD1 | -0.0026* | -0.0129* | -0.0030* | -0.0940*** | -0.2987** | -0.0980*** |
| | (0.0015) | (0.0072) | (0.0016) | (0.0341) | (0.1220) | (0.0364) |
| GOAL | 0.3044*** | 0.5266*** | 0.2242* | 0.3005*** | 0.5105*** | 0.2177* |
| | (0.0941) | (0.1644) | (0.1247) | (0.0939) | (0.1629) | (0.1229) |
| ASSIST | -0.0991 | -0.0177 | -0.1769 | -0.1095 | -0.0232 | -0.2032* |
| | (0.0821) | (0.1404) | (0.1181) | (0.0821) | (0.1394) | (0.1168) |
| PASS | 0.0008 | -0.0006 | 0.0045* | 0.0004 | -0.0011 | 0.0039* |
| | (0.0016) | (0.0026) | (0.0023) | (0.0016) | (0.0026) | (0.0023) |
| TACKLE | -0.0296* | -0.0344 | -0.0234 | -0.0314* | -0.0431* | -0.0257 |
| | (0.0164) | (0.0250) | (0.0250) | (0.0164) | (0.0253) | (0.0247) |
| DUEL-LOST | -0.0518*** | -0.0636*** | -0.0501*** | -0.0519*** | -0.0637*** | -0.0473*** |
| | (0.0091) | (0.0152) | (0.0155) | (0.0090) | (0.0152) | (0.0151) |
| FREE-TF | 0.1299** | 0.0643 | 0.1238 | 0.1150* | 0.0765 | -0.0254 |
| | (0.0658) | (0.0885) | (0.1963) | (0.0651) | (0.0831) | (0.1821) |
| CONSTANT | 0.6181* | 1.1651*** | 0.7016*** | 0.7168** | 1.2040*** | 0.8808*** |
| | (0.3292) | (0.1859) | (0.2482) | (0.3275) | (0.1857) | (0.2564) |
| N | 811 | 481 | 330 | 811 | 481 | 330 |
| $R^2$ | 0.2817 | 0.2826 | 0.3006 | 0.2860 | 0.2923 | 0.3197 |
| *Panel B: Zero TF player excluded* | | | | | | |
| TF | 0.0062 | 0.0225** | 0.0161* | 0.0531 | 0.3178** | -0.0588 |
| | (0.0046) | (0.0113) | (0.0095) | (0.0723) | (0.1438) | (0.1502) |
| TFD1 | -0.0023 | -0.0196** | -0.0026 | -0.0940*** | -0.4389*** | -0.0952** |
| | (0.0015) | (0.0076) | (0.0016) | (0.0340) | (0.1232) | (0.0376) |
| GOAL | 0.2171** | 0.4979** | 0.2204 | 0.2162** | 0.4397** | 0.2368* |
| | (0.1022) | (0.1920) | (0.1338) | (0.1015) | (0.1863) | (0.1323) |
| ASSIST | -0.1847** | 0.1440 | -0.2827** | -0.2002** | 0.1429 | -0.3149** |
| | (0.0932) | (0.1787) | (0.1308) | (0.0930) | (0.1739) | (0.1305) |
| PASS | 0.0011 | 0.0001 | 0.0040 | 0.0008 | -0.0013 | 0.0039 |
| | (0.0018) | (0.0029) | (0.0027) | (0.0018) | (0.0029) | (0.0027) |
| TACKLE | -0.0420** | -0.0317 | -0.0189 | -0.0450** | -0.0496 | -0.0262 |
| | (0.0191) | (0.0311) | (0.0281) | (0.0190) | (0.0312) | (0.0281) |
| DUEL-LOST | -0.0350*** | -0.0550*** | -0.0451** | -0.0348*** | -0.0555*** | -0.0387** |
| | (0.0102) | (0.0184) | (0.0175) | (0.0101) | (0.0180) | (0.0172) |
| CONSTANT | 1.1343*** | 0.7889*** | 0.5925** | 1.1611*** | 0.8716*** | 0.9067*** |
| | (0.1734) | (0.2642) | (0.2751) | (0.1730) | (0.2558) | (0.2785) |
| N | 616 | 349 | 267 | 616 | 349 | 267 |
| $R^2$ | 0.2775 | 0.3475 | 0.3281 | 0.2857 | 0.3748 | 0.3364 |

* p<0.10, ** p<0.05, *** p<0.01. Fixed effect (FE) panel regressions are estimated. Dependent variable is percentage of played minutes. Panel A presents the results using the full sample data. Panel B presents the results while excluding free transferred and academy players from the sample (i.e. 0 TF players). However, we kept on-loan players in the Panel B sample, as some of them did have loan fees reported. Columns (2), (3), (5), and (6) are win ratio subsamples for both models. They account for different clubs' characteristics. The mean club win ratio was 37.6%. The average spend per player was considerably higher for clubs with a win ratio of more than 37% (see Figure 3 in appendix). Regression including season dummies and played minutes coefficients are presented in Table 13.

## 2.6.6 Table 3 Defenders Results - 2012/13 to 2016/17

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| FE Model | (1) Full Sample | (2) WIN<37% | (3) WIN>37% | (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
| *Panel A: Zero TF player included* | | | | | | |
| TF | 0.0163*** | 0.0225 | 0.0217** | 0.2284*** | -0.1645 | 0.4707** |
| | (0.0046) | (0.0309) | (0.0095) | (0.0694) | (0.2587) | (0.1890) |
| TFD1 | -0.0004 | 0.0103 | -0.0027 | -0.0096 | 0.1606 | -0.0660 |
| | (0.0025) | (0.0137) | (0.0029) | (0.0366) | (0.1612) | (0.0436) |
| GOAL | -0.0953 | 0.2939 | 0.0049 | -0.0698 | 0.2090 | 0.0189 |
| | (0.2073) | (0.3872) | (0.2874) | (0.2076) | (0.3833) | (0.2846) |
| PASS | 0.0038** | 0.0012 | 0.0078*** | 0.0039** | 0.0009 | 0.0076*** |
| | (0.0019) | (0.0035) | (0.0027) | (0.0019) | (0.0036) | (0.0026) |
| INTERCEPTION | -0.0076 | -0.0281 | 0.0096 | -0.0083 | -0.0300 | 0.0089 |
| | (0.0229) | (0.0342) | (0.0398) | (0.0229) | (0.0343) | (0.0398) |
| DUEL-LOST | -0.0211 | -0.0159 | -0.0449** | -0.0235 | -0.0178 | -0.0506** |
| | (0.0151) | (0.0259) | (0.0210) | (0.0152) | (0.0257) | (0.0208) |
| FREE-TF | 0.1200 | 0.2535 | 0.0363 | 0.1199 | 0.2765 | 0.2175 |
| | (0.1046) | (0.1951) | (0.2177) | (0.1064) | (0.2024) | (0.2486) |
| CONSTANT | 0.2406 | 0.7041** | 0.3005 | 0.2833 | 0.8791*** | 0.1493 |
| | (0.1902) | (0.2918) | (0.2762) | (0.1879) | (0.2709) | (0.3041) |
| N | 653 | 384 | 269 | 653 | 384 | 269 |
| $R^2$ | 0.2030 | 0.2264 | 0.2334 | 0.1990 | 0.2234 | 0.2471 |
| *Panel B: Zero TF player excluded* | | | | | | |
| TF | 0.0236*** | 0.0019 | 0.0194 | 0.3421*** | 0.0399 | 0.4171 |
| | (0.0075) | (0.0368) | (0.0187) | (0.1110) | (0.3253) | (0.2756) |
| TFD1 | -0.0003 | 0.0090 | -0.0019 | -0.0233 | 0.0946 | -0.0572 |
| | (0.0026) | (0.0157) | (0.0030) | (0.0365) | (0.1798) | (0.0416) |
| GOAL | 0.4042 | 0.4013 | 0.6954** | 0.4116 | 0.4089 | 0.6942** |
| | (0.2535) | (0.5092) | (0.3244) | (0.2536) | (0.5063) | (0.3197) |
| PASS | 0.0050** | 0.0029 | 0.0083*** | 0.0047** | 0.0029 | 0.0082*** |
| | (0.0021) | (0.0046) | (0.0026) | (0.0021) | (0.0046) | (0.0026) |
| INTERCEPTION | -0.0063 | -0.0375 | -0.0023 | -0.0094 | -0.0409 | -0.0015 |
| | (0.0266) | (0.0452) | (0.0409) | (0.0268) | (0.0454) | (0.0411) |
| DUEL-LOST | -0.0228 | -0.0339 | -0.0183 | -0.0265 | -0.0349 | -0.0247 |
| | (0.0174) | (0.0339) | (0.0225) | (0.0174) | (0.0336) | (0.0216) |
| CONSTANT | 0.4580** | 0.7546*** | 0.4534** | 0.3330* | 0.7527*** | 0.3514* |
| | (0.1876) | (0.2320) | (0.1818) | (0.1995) | (0.2240) | (0.2053) |
| N | 485 | 258 | 227 | 485 | 258 | 227 |
| $R^2$ | 0.1536 | 0.2028 | 0.1936 | 0.1523 | 0.2035 | 0.2095 |

* p<0.10, ** p<0.05, *** p<0.01. Fixed effect (FE) panel regressions are estimated. Dependent variable is percentage of played minutes. Panel A presents the results using the full sample data. Panel B presents the results while excluding free transferred and academy players from the sample (i.e. 0 TF players). However, we kept on-loan players in the Panel B sample, as some of them did have loan fees reported. Columns (2), (3), (5), and (6) are win ratio subsamples for both models. They account for different clubs' characteristics. The mean club win ratio was 37.6%. The average spend per player was considerably higher for clubs with a win ratio of more than 37% (see Figure 3 in appendix). Regression including season dummies and played minutes coefficients are presented in Table 14.

**2.6.7 Table 4 Average TF (£m) spend per player sorted by seasons – All teams**

*Panel A: Average TF per player (Nominal Value)*

| Season | Average | SD | Min | Max |
|---|---|---|---|---|
| 2006 | 3.0 | 4.1 | 0.3 | 18.6 |
| 2007 | 3.3 | 2.5 | 0.7 | 11.9 |
| 2008 | 3.8 | 2.4 | 1.2 | 11.1 |
| 2009 | 3.7 | 3.5 | 0.2 | 13.3 |
| 2010 | 4.3 | 5.5 | 0.0 | 19.2 |
| 2011 | 3.3 | 3.3 | 0.5 | 13.1 |
| 2012 | 3.7 | 3.7 | 0.0 | 13.5 |
| 2013 | 5.7 | 6.1 | 0.0 | 22.3 |
| 2014 | 5.6 | 5.9 | 0.5 | 20.8 |
| 2015 | 6.3 | 6.0 | 1.4 | 28.2 |
| 2016 | 11.9 | 12.9 | 1.6 | 53.4 |

*Panel B: Average TF per player (Inflation Adjusted)*

| | | | | |
|---|---|---|---|---|
| 2006 | 3.87 | 5.34 | 0.33 | 24.24 |
| 2007 | 4.20 | 3.15 | 0.84 | 15.31 |
| 2008 | 4.61 | 2.98 | 1.41 | 13.50 |
| 2009 | 4.42 | 4.23 | 0.22 | 16.04 |
| 2010 | 4.89 | 6.35 | 0.00 | 22.38 |
| 2011 | 3.66 | 3.63 | 0.58 | 14.56 |
| 2012 | 3.98 | 4.03 | 0.05 | 14.65 |
| 2013 | 6.04 | 6.50 | 0.00 | 23.64 |
| 2014 | 5.86 | 6.15 | 0.47 | 21.74 |
| 2015 | 6.58 | 6.30 | 1.47 | 29.52 |
| 2016 | 9.55 | 8.65 | 1.61 | 34.98 |

**2.6.8 Table 5 Average TF (£m) spend per player sorted by seasons and positions – All teams**

*Panel A: Average TF for Strikers*

| Obs | Season | Average | SD | Min | Max |
|---|---|---|---|---|---|
| 74 | 2006 | 5.00 | 7.18 | 0.00 | 39.09 |
| 74 | 2007 | 6.79 | 8.03 | 0.00 | 38.47 |
| 73 | 2008 | 7.39 | 9.32 | 0.00 | 39.60 |
| 73 | 2009 | 7.86 | 10.85 | 0.00 | 56.60 |
| 72 | 2010 | 7.77 | 10.59 | 0.00 | 54.90 |
| 71 | 2011 | 8.34 | 12.46 | 0.00 | 55.56 |
| 66 | 2012 | 8.21 | 11.49 | 0.00 | 54.35 |
| 68 | 2013 | 7.98 | 10.16 | 0.00 | 52.93 |
| 76 | 2014 | 8.15 | 9.28 | 0.00 | 39.74 |
| 83 | 2015 | 9.08 | 10.60 | 0.00 | 46.06 |
| 66 | 2016 | 11.50 | 11.78 | 0.00 | 45.65 |

*Panel B: Average TF for Midfielders*

| | | | | | |
|---|---|---|---|---|---|
| 158 | 2012 | 5.30 | 7.51 | 0 | 34.78 |
| 164 | 2013 | 6.11 | 8.75 | 0 | 44.99 |
| 165 | 2014 | 6.87 | 10.08 | 0 | 62.44 |
| 165 | 2015 | 7.66 | 9.89 | 0 | 57.58 |
| 159 | 2016 | 10.12 | 12.68 | 0 | 92.65 |

*Panel C: Average TF for Defenders*

| | | | | | |
|---|---|---|---|---|---|
| 135 | 2012 | 3.96 | 5.64 | 0 | 32.61 |
| 136 | 2013 | 3.98 | 5.54 | 0 | 31.76 |
| 140 | 2014 | 4.31 | 6.03 | 0 | 33.47 |
| 127 | 2015 | 5.15 | 6.39 | 0 | 33.50 |
| 116 | 2016 | 6.47 | 8.63 | 0 | 49.28 |

***Notes**: All values are adjusted for inflation.

**2.6.9 Table 6 Average TF (£m) spend per player sorted by clubs**

| Team | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Chelsea | 8.3 | 4.9 | 6.4 | 4.6 | 19.2 | 6.5 | 13.5 | 10.6 | 16.1 | 6.6 | 24.0 |
| Man UTD | 18.6 | 11.9 | 6.2 | 3.8 | 5.2 | 13.1 | 9.2 | 22.3 | 20.8 | 28.2 | 33.7 |
| Man City | 0.5 | 4.5 | 11.1 | 13.3 | 17.2 | 9.8 | 5.4 | 17.4 | 13.5 | 14.3 | 20.0 |
| Liverpool | 2.6 | 4.9 | 5.0 | 10.0 | 8.3 | 5.6 | 8.2 | 6.9 | 10.6 | 10.6 | 8.1 |
| Arsenal | 0.9 | 2.6 | 6.0 | 10.0 | 4.6 | 5.4 | 9.5 | 8.5 | 13.2 | 3.6 | 14.5 |
| Spurs | 4.0 | 5.4 | 7.5 | 4.1 | 3.2 | 0.8 | 6.8 | 13.2 | 3.4 | 9.1 | 17.3 |
| Stoke | - | - | 1.6 | 3.6 | 1.8 | 4.4 | 1.8 | 0.8 | 0.6 | 3.0 | 6.3 |
| Sunderland | - | 2.9 | 1.9 | 4.8 | 4.7 | 1.9 | 3.9 | 1.3 | 2.0 | 3.6 | 2.7 |
| Newcastle | 4.3 | 2.3 | 2.2 | - | 0.9 | 2.4 | 1.6 | 0.0 | 2.8 | 9.2 | - |
| Norwich | - | - | - | - | - | 1.4 | 0.3 | 2.7 | - | 1.8 | - |
| Crystal Palace | - | - | - | - | - | - | - | 0.6 | 1.2 | 1.9 | 7.8 |
| Aston Villa | 2.2 | 3.4 | 3.5 | 3.1 | 10.7 | 3.3 | 2.5 | 2.8 | 0.8 | 3.7 | - |
| Everton | 5.5 | 3.6 | 3.0 | 5.0 | 0.2 | 1.8 | 1.9 | 3.5 | 5.5 | 4.0 | 4.5 |
| Southampton | - | - | - | - | - | - | 2.0 | 12.0 | 8.3 | 4.4 | 7.6 |
| Westham | 0.6 | 3.9 | 2.8 | 0.2 | 0.6 | - | 1.4 | 2.5 | 2.8 | 3.8 | 4.4 |
| QPR | - | - | - | - | - | 0.8 | 1.6 | - | 4.0 | - | - |
| Hull City | - | - | 1.2 | 0.8 | - | - | - | 2.2 | 2.4 | - | 1.6 |
| West Brom | - | - | 1.9 | - | 0.0 | 1.1 | 1.3 | 1.2 | 1.3 | 4.4 | 6.2 |
| Leicester | - | - | - | - | - | - | - | - | 1.5 | 7.3 | 9.8 |
| Swansea | - | - | - | - | - | 0.5 | 1.2 | 1.8 | 0.8 | 1.9 | 2.5 |
| Burnley | - | - | - | 0.6 | - | - | - | - | 0.5 | - | 3.9 |
| Watford | 0.3 | - | - | - | - | - | - | - | - | 1.4 | 3.7 |
| Bournemouth | - | - | - | - | - | - | - | - | - | 3.0 | 2.4 |
| Blackburn | 1.2 | 1.0 | 1.7 | 1.5 | 1.2 | 1.2 | - | - | - | - | - |
| Reading | 0.7 | 0.7 | - | - | - | - | 0.0 | - | - | - | - |
| Portsmouth | 0.7 | 2.6 | 3.2 | 1.0 | - | - | - | - | - | - | - |
| Mid'brough | 3.0 | 4.1 | 2.6 | - | - | - | - | - | - | - | 1.9 |
| Fulham | 0.8 | 1.5 | 1.9 | 0.5 | 1.0 | 1.0 | 0.5 | 1.4 | - | - | - |
| Wigan | 1.7 | 1.4 | 2.8 | 1.2 | 0.7 | 0.7 | 0.5 | - | - | - | - |
| Sheffield UTD | 0.8 | - | - | - | - | - | - | - | - | - | - |
| Charlton | 1.6 | - | - | - | - | - | - | - | - | - | - |
| Bolton | 1.1 | 1.1 | 3.4 | 1.2 | 0.4 | 1.2 | - | - | - | - | - |
| Birmingham | - | 2.0 | - | 2.6 | 2.4 | - | - | - | - | - | - |
| Derby | - | 0.9 | - | - | - | - | - | - | - | - | - |
| Wolves | - | - | - | 1.6 | 0.0 | 3.1 | - | - | - | - | - |
| Cardiff | - | - | - | - | - | - | - | 2.3 | - | - | - |

*The average is calculated by dividing the total value spent by a club over the number of transferred players to the club in each season, while we excluded players from the youth squad and who were on loan from the calculation.

## 2.6.10 Table 7 Playing time metrics averages-sorted by clubs

| | | Possible minutes played | | | Minutes played | | | Appearances | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Obs | Mean | Median | SD | Mean | Median | SD | Mean | Median | SD |
| *Panel A: Total Clubs Averages* | | | | | | | | | | |
| | 829 | 0.503 | 0.493 | 0.251 | 1501 | 1392 | 847 | 23.8 | 25.0 | 8.8 |
| *Panel B: Individual Clubs averages* | | | | | | | | | | |
| Arsenal | 50 | 0.530 | 0.541 | 0.250 | 1531 | 1416 | 841 | 24.1 | 25.0 | 8.7 |
| Man City | 48 | 0.453 | 0.465 | 0.234 | 1345 | 1157 | 782 | 22.2 | 23.0 | 9.2 |
| Stoke | 44 | 0.495 | 0.481 | 0.244 | 1570 | 1359 | 845 | 24.9 | 26.5 | 8.8 |
| Westham | 40 | 0.475 | 0.491 | 0.215 | 1274 | 1274 | 606 | 20.9 | 22.0 | 7.9 |
| Everton | 40 | 0.442 | 0.375 | 0.267 | 1272 | 1079 | 852 | 20.6 | 18.5 | 9.7 |
| Man Utd | 39 | 0.584 | 0.622 | 0.238 | 1814 | 1860 | 742 | 26.7 | 28.0 | 6.3 |
| Liverpool | 39 | 0.525 | 0.495 | 0.255 | 1472 | 1241 | 805 | 23.3 | 24.0 | 8.1 |
| Newcastle | 38 | 0.474 | 0.450 | 0.228 | 1414 | 1332 | 737 | 22.4 | 24.0 | 8.8 |
| Chelsea | 38 | 0.491 | 0.521 | 0.280 | 1534 | 1556 | 912 | 24.8 | 27.5 | 8.7 |
| Spurs | 37 | 0.473 | 0.484 | 0.277 | 1481 | 1478 | 931 | 23.7 | 27.0 | 10.3 |
| Sunderland | 33 | 0.516 | 0.544 | 0.260 | 1568 | 1763 | 906 | 23.1 | 26.0 | 9.4 |
| West Brom | 32 | 0.445 | 0.383 | 0.231 | 1355 | 1166 | 719 | 24.3 | 24.0 | 7.6 |
| Aston Villa | 31 | 0.589 | 0.688 | 0.272 | 1767 | 1812 | 928 | 25.2 | 28.0 | 9.0 |
| Fulham | 28 | 0.587 | 0.550 | 0.241 | 1714 | 1533 | 904 | 24.8 | 27.0 | 9.5 |
| Blackburn | 27 | 0.450 | 0.432 | 0.219 | 1372 | 1446 | 689 | 22.6 | 23.0 | 7.1 |
| Bolton | 20 | 0.609 | 0.669 | 0.290 | 2045 | 2204 | 985 | 29.4 | 30.5 | 6.7 |
| Crystal Palace | 20 | 0.431 | 0.444 | 0.245 | 1185 | 984 | 871 | 19.9 | 19.0 | 8.8 |
| Wigan | 20 | 0.566 | 0.511 | 0.208 | 1699 | 1655 | 854 | 25.8 | 26.5 | 8.2 |
| Portsmouth | 19 | 0.429 | 0.381 | 0.224 | 1345 | 1048 | 805 | 22.5 | 19.0 | 7.5 |
| Southampton | 19 | 0.536 | 0.549 | 0.272 | 1573 | 1538 | 1026 | 25.1 | 30.0 | 10.7 |
| Swansea | 19 | 0.471 | 0.496 | 0.262 | 1387 | 1148 | 899 | 23.1 | 20.0 | 8.6 |
| Hull City | 15 | 0.457 | 0.426 | 0.172 | 1176 | 1195 | 371 | 20.3 | 20.0 | 5.3 |
| QPR | 14 | 0.477 | 0.497 | 0.238 | 1255 | 1148 | 739 | 19.8 | 19.5 | 8.8 |
| Middlesbrough | 13 | 0.524 | 0.609 | 0.267 | 1576 | 2084 | 926 | 23.4 | 29.0 | 9.3 |
| Leicester | 12 | 0.510 | 0.508 | 0.262 | 1664 | 1637 | 854 | 28.3 | 29.5 | 7.9 |
| Norwich | 12 | 0.470 | 0.485 | 0.201 | 1470 | 1536 | 690 | 26.5 | 29.0 | 8.2 |
| Watford | 12 | 0.600 | 0.595 | 0.263 | 1772 | 1564 | 992 | 24.3 | 19.0 | 10.1 |
| Birmingham | 12 | 0.502 | 0.472 | 0.223 | 1516 | 1498 | 822 | 24.3 | 27.5 | 9.2 |
| Bournemouth | 11 | 0.495 | 0.562 | 0.235 | 1090 | 899 | 699 | 20.0 | 15.0 | 8.2 |
| Burnley | 11 | 0.577 | 0.601 | 0.249 | 1777 | 1874 | 853 | 28.1 | 30.0 | 7.7 |
| Reading | 11 | 0.480 | 0.436 | 0.224 | 1614 | 1492 | 789 | 27.2 | 29.0 | 7.7 |
| Wolves | 9 | 0.523 | 0.399 | 0.200 | 1709 | 1365 | 703 | 26.9 | 28.0 | 6.2 |
| Derby | 4 | 0.508 | 0.506 | 0.183 | 1453 | 1161 | 668 | 22.0 | 21.0 | 5.1 |
| Sheffield Utd | 4 | 0.515 | 0.527 | 0.214 | 1396 | 1109 | 617 | 22.5 | 23.5 | 5.5 |
| Cardiff | 3 | 0.465 | 0.365 | 0.252 | 1592 | 1250 | 861 | 24.7 | 22.0 | 9.2 |
| Charlton | 3 | 0.531 | 0.433 | 0.221 | 1817 | 1480 | 757 | 29.0 | 30.0 | 2.9 |
| Blackpool | 2 | 0.544 | 0.544 | 0.224 | 1530 | 1530 | 1097 | 20.0 | 20.0 | 11.0 |

**Notes:* Possible minutes played are the percentage of the actual minutes played relative to the total available minutes for a player, taking into account his injuries, suspensions, and other participations such as with the national team or on loan to other clubs.

### 2.6.11 Table 8 Playing time (PT) metrics averages-sorted by positions and TF

| Variable | Obs | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| *Panel A: PT average for strikers* | | | | | |
| Minutes | 796 | 1552.15 | 825.45 | 123.00 | 3390.00 |
| Possible minutes | 796 | 0.52 | 0.25 | 0.04 | 1.00 |
| Appearance | 796 | 24.48 | 8.26 | 9.00 | 38.00 |
| *Panel B: PT average when TF is less than £10 m* | | | | | |
| Minutes | 582 | 1461.68 | 825.52 | 123.00 | 3387.00 |
| Possible minutes | 582 | 0.48 | 0.24 | 0.04 | 1.00 |
| Appearance | 582 | 23.78 | 8.31 | 9.00 | 38.00 |
| *Panel C: PT average when TF is more than £10 m* | | | | | |
| Minutes | 214 | 1798.18 | 775.28 | 340.00 | 3390.00 |
| Possible minutes | 214 | 0.61 | 0.22 | 0.10 | 1.00 |
| Appearance | 214 | 26.38 | 7.84 | 9.00 | 38.00 |
| *Panel D: PT average for Midfielders and Defenders* | | | | | |
| Minutes | 1,462 | 1849.55 | 846.69 | 74.00 | 3420.00 |
| Possible minutes | 1,462 | 0.61 | 0.25 | 0.03 | 1.00 |
| Appearance | 1,462 | 24.68 | 8.43 | 9.00 | 38.00 |
| *Panel E: PT average when TF is less than £5 m* | | | | | |
| Minutes | 896 | 1,793.70 | 873.21 | 74.00 | 3,420.00 |
| Possible minutes | 896 | 0.59 | 0.26 | 0.03 | 1.00 |
| Appearance | 896 | 24.04 | 8.67 | 9.00 | 38.00 |
| *Panel F: PT average when TF is more than £5 m* | | | | | |
| Minutes | 566 | 1937.97 | 795.69 | 135.00 | 3420.00 |
| Possible minutes | 566 | 0.66 | 0.23 | 0.04 | 1.00 |
| Appearance | 566 | 25.70 | 7.92 | 9.00 | 38.00 |

**Notes:* Panels B & C relate to strikers, while Panels E & F relate to midfielders and defenders. Possible minutes played are the percentage of the actual minutes played relative to the total available minutes for a player, taking into account his injuries, suspensions, or other participations, such as with national team or on loan to other clubs.

## 2.6.12 Table 9 Correlation matrix and PCA for strikers' performance metrics

### Correlation Matrix

| | pass | cross | tackle | block | interc~n | cleara~e |
|---|---|---|---|---|---|---|
| pass | 1.0000 | | | | | |
| cross | 0.1113 | 1.0000 | | | | |
| tackle | 0.1514 | 0.2386 | 1.0000 | | | |
| block | 0.1298 | 0.1055 | -0.0259 | 1.0000 | | |
| interception | 0.1874 | 0.3370 | 0.4817 | 0.0518 | 1.0000 | |
| clearance | 0.1184 | -0.1464 | 0.0777 | -0.0787 | 0.1214 | 1.0000 |

### Principal components' analysis (PCA)

```
. pca pass cross tackle block interception clearance, mineigen(1)

Principal components/correlation                  Number of obs    =        796
                                                  Number of comp.  =          3
                                                  Trace            =          6
    Rotation: (unrotated = principal)             Rho              =     0.6768
```

| Component | Eigenvalue | Difference | Proportion | Cumulative |
|---|---|---|---|---|
| Comp1 | 1.82319 | .638995 | 0.3039 | 0.3039 |
| Comp2 | 1.18419 | .130585 | 0.1974 | 0.5012 |
| Comp3 | 1.05361 | .274703 | 0.1756 | 0.6768 |
| Comp4 | .778903 | .10592 | 0.1298 | 0.8066 |
| Comp5 | .672983 | .185855 | 0.1122 | 0.9188 |
| Comp6 | .487128 | . | 0.0812 | 1.0000 |

Principal components (eigenvectors)

| Variable | Comp1 | Comp2 | Comp3 | Unexplained |
|---|---|---|---|---|
| pass | 0.3318 | 0.1034 | 0.6114 | .3928 |
| cross | 0.4489 | -0.4242 | -0.2035 | .3759 |
| tackle | 0.5490 | 0.1580 | -0.2530 | .3535 |
| block | 0.1213 | -0.4813 | 0.6474 | .2573 |
| interception | 0.6024 | 0.0933 | -0.1316 | .3098 |
| clearance | 0.0970 | 0.7376 | 0.2905 | .2498 |

## 2.6.13 Table 10 Correlation matrix and PCA for midfielders' and defenders' performance metrics

### Correlation Matrix

| | Pass | Cross | Tackle | Block | Interc~n | Cleara~e | Dualwon | Duallost | Recovery |
|---|---|---|---|---|---|---|---|---|---|
| Pass | 1.0000 | | | | | | | | |
| Cross | -0.0631 | 1.0000 | | | | | | | |
| Tackle | 0.2261 | -0.1038 | 1.0000 | | | | | | |
| Block | 0.1541 | 0.3959 | -0.0380 | 1.0000 | | | | | |
| Interception | 0.0602 | -0.3690 | 0.4006 | -0.3407 | 1.0000 | | | | |
| Clearance | -0.3106 | -0.4817 | -0.1370 | -0.4883 | 0.3544 | 1.0000 | | | |
| Dualwon | 0.0487 | -0.0523 | 0.5328 | 0.1150 | 0.1475 | -0.0248 | 1.0000 | | |
| Duallost | 0.1379 | 0.3138 | 0.2795 | 0.4639 | -0.2227 | -0.5431 | 0.5487 | 1.0000 | |
| Recovery | 0.4329 | 0.0154 | 0.3992 | 0.1475 | 0.2449 | -0.3339 | 0.3193 | 0.3204 | 1.0000 |

### Principal components analysis (PCA)

```
. pca Duallost Recovery Interception Pass Cross Tackle Block Clearance Dualwon, mineigen(1)

Principal components/correlation                 Number of obs    =      1,462
                                                 Number of comp.  =          3
                                                 Trace            =          9
    Rotation: (unrotated = principal)            Rho              =     0.6974
```

| Component | Eigenvalue | Difference | Proportion | Cumulative |
|---|---|---|---|---|
| Comp1 | 2.87447 | .647009 | 0.3194 | 0.3194 |
| Comp2 | 2.22746 | 1.05264 | 0.2475 | 0.5669 |
| Comp3 | 1.17482 | .461829 | 0.1305 | 0.6974 |
| Comp4 | .712994 | .154107 | 0.0792 | 0.7766 |
| Comp5 | .558887 | .0602169 | 0.0621 | 0.8387 |
| Comp6 | .49867 | .0798612 | 0.0554 | 0.8941 |
| Comp7 | .418809 | .0783925 | 0.0465 | 0.9407 |
| Comp8 | .340416 | .146952 | 0.0378 | 0.9785 |
| Comp9 | .193464 | . | 0.0215 | 1.0000 |

Principal components (eigenvectors)

| Variable | Comp1 | Comp2 | Comp3 | Unexplained |
|---|---|---|---|---|
| Duallost | 0.4853 | 0.0079 | 0.3171 | .2048 |
| Recovery | 0.3307 | 0.3307 | -0.3171 | .3238 |
| Interception | -0.1491 | 0.5104 | -0.0693 | .3501 |
| Pass | 0.2410 | 0.1840 | -0.6654 | .2375 |
| Cross | 0.2849 | -0.3696 | 0.0868 | .4535 |
| Tackle | 0.2298 | 0.4818 | 0.1333 | .3102 |
| Block | 0.3861 | -0.2531 | 0.0068 | .4289 |
| Clearance | -0.4619 | 0.1921 | 0.2250 | .245 |
| Dualwon | 0.2801 | 0.3556 | 0.5248 | .1694 |

## 2.6.14 Table 11 Premier League TV rights deals: Historic data

| Period | Years | Broadcaster | Games Per Year | Total Games in deal | Cost for Whole Deal | Cost Per Game |
|---|---|---|---|---|---|---|
| 1992–97 | 5 | Sky | 60 | 300 | £191m | £0.64m |
| 1997–01 | 4 | Sky | 60 | 240 | £670m | £2.79m |
| 2001–04 | 3 | Sky | 110 | 330 | £1200m | £3.64m |
| 2004–07 | 3 | Sky | 138 | 414 | £1024m | £2.47m |
| 2007–10 | 3 | Sky & Setanta | 138 | 414 | £1706m | £4.12m |
| 2010–13 | 3 | Sky & ESPN | 138 | 414 | £1773m | £4.28m |
| 2013–16 | 3 | Sky & BT | 154 | 462 | £3008m | £6.53m |
| 2016–19 | 3 | Sky & BT | 168 | 504 | £5100m | £10.12m |

Source: https://www.totalsportek.com/money/premier-league-tv-rights-deals-history-1992-2019/

## 2.6.15 Table 12 Strikers' results including season dummy (played minutes as a dependent variable)

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| FE Model | (1) Full Sample | (2) WIN<37% | (3) WIN>37% | (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
| *Panel A: Zero TF player included* | | | | | | |
| TF | 18.20*** | 30.89** | 19.74** | 424.78*** | 585.79** | 566.02*** |
| | (6.48) | (14.76) | (9.92) | (130.36) | (289.09) | (198.55) |
| TFD1 | -5.71 | -15.84 | -4.69 | -72.03 | -135.10 | -41.27 |
| | (4.93) | (15.74) | (5.67) | (117.16) | (255.62) | (150.99) |
| GOAL | 290.16* | 322.08 | 300.38 | 281.72* | 290.52 | 324.65 |
| | (151.70) | (269.67) | (225.84) | (151.05) | (269.43) | (223.02) |
| ASSIST | 34.15 | 466.49 | -130.45 | 25.73 | 442.24 | -118.96 |
| | (264.85) | (457.35) | (388.43) | (264.03) | (458.68) | (383.56) |
| PASS | -12.54*** | -21.02* | -21.92** | -11.28*** | -17.87 | -19.07** |
| | (4.20) | (11.68) | (8.51) | (4.21) | (12.00) | (8.44) |
| CROSS | -37.36 | 22.35 | 3.09 | -38.34 | 12.60 | 2.11 |
| | (36.39) | (73.59) | (54.45) | (36.32) | (72.99) | (53.71) |
| TACKLE | -172.02** | -221.39* | -215.45 | -173.29** | -215.91* | -225.70* |
| | (77.10) | (117.89) | (130.47) | (76.90) | (118.01) | (128.77) |
| FREE-TF | 90.80 | 713.99** | 828.03 | 97.73 | 620.81** | 664.96 |
| | (188.62) | (302.81) | (781.73) | (186.96) | (296.71) | (762.86) |
| SUSPENSION | -58.12* | -53.93 | -33.40 | -56.82* | -57.68 | -32.65 |
| | (32.51) | (51.86) | (46.67) | (32.34) | (52.14) | (46.14) |
| INJURY | -48.56*** | -53.62*** | -51.89*** | -50.06*** | -54.99*** | -53.81*** |
| | (6.35) | (10.55) | (10.04) | (6.38) | (10.54) | (9.91) |
| ON-LOAN | -50.61*** | -53.23*** | -82.90** | -51.65*** | -55.15*** | -86.69** |
| | (11.68) | (14.64) | (38.40) | (11.66) | (14.64) | (37.70) |
| WW | -1106.68*** | -1111.14*** | -1134.56*** | -1127.51*** | -1146.01*** | -1168.52*** |
| | (165.66) | (242.14) | (338.24) | (165.72) | (242.28) | (329.77) |
| 2007.season | -70.82 | -504.16** | -81.09 | -89.78 | -492.51** | -148.74 |
| | (126.77) | (243.74) | (201.44) | (127.19) | (245.17) | (200.13) |
| 2008.season | 14.34 | -118.38 | -17.56 | -6.26 | -105.82 | -79.27 |
| | (137.82) | (237.62) | (225.11) | (137.90) | (238.10) | (222.06) |
| 2009.season | -37.12 | -73.09 | -51.34 | -72.90 | -83.62 | -130.07 |
| | (144.03) | (262.02) | (209.32) | (144.41) | (264.55) | (206.82) |
| 2010.season | -243.44* | -591.81** | -34.78 | -263.51* | -626.56** | -75.76 |
| | (141.90) | (244.07) | (226.85) | (141.81) | (244.68) | (224.27) |
| 2011.season | -218.79 | -651.60** | -154.80 | -242.69* | -674.22** | -225.94 |
| | (145.97) | (260.74) | (230.02) | (146.13) | (262.95) | (228.15) |
| 2012.season | -183.74 | -786.56*** | 33.58 | -222.85 | -821.28*** | -52.45 |
| | (157.13) | (275.78) | (235.60) | (157.35) | (277.43) | (234.86) |
| 2013.season | -424.19** | -1164.52*** | -268.31 | -474.79*** | -1196.71*** | -363.22 |
| | (168.01) | (309.62) | (246.33) | (168.55) | (313.71) | (243.88) |
| 2014.season | -498.34*** | -1114.34*** | -461.34* | -552.57*** | -1145.49*** | -566.22** |
| | (168.20) | (308.59) | (254.24) | (169.10) | (311.30) | (254.03) |
| 2015.season | -683.26*** | -1456.73*** | -412.46 | -735.84*** | -1500.26*** | -502.39* |
| | (178.47) | (318.49) | (287.04) | (179.54) | (321.30) | (286.32) |
| 2016.season | -665.05*** | -1478.35*** | -337.51 | -723.13*** | -1530.32*** | -458.33 |
| | (192.04) | (336.65) | (292.86) | (191.15) | (340.96) | (286.52) |
| CONSTANT | 1920.85*** | 2995.18*** | 2732.42*** | 1914.52*** | 2988.87*** | 2535.52*** |
| | (319.91) | (495.31) | (471.20) | (317.72) | (496.23) | (475.01) |
| N | 796 | 487 | 309 | 796 | 487 | 309 |
| $R^2$ | 0.36 | 0.48 | 0.41 | 0.37 | 0.48 | 0.43 |

* p<0.10, ** p<0.05, *** p<0.01

## Table 12 Strikers' results including season dummy (played minutes as a dependent variable)

| FE Model | Model 1 (1) Full Sample | (2) WIN<37% | (3) WIN>37% | Model 2 (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
|---|---|---|---|---|---|---|
| *Panel B: Zero TF player excluded* | | | | | | |
| TF | 16.62** | 43.52** | 22.72* | 394.28** | 485.04 | 665.80*** |
| | (7.95) | (20.75) | (11.89) | (152.86) | (353.73) | (238.50) |
| TFD1 | -5.68 | -17.50 | -4.52 | -79.86 | -129.38 | -41.11 |
| | (5.11) | (16.93) | (5.85) | (122.11) | (272.19) | (154.38) |
| GOAL | 303.28* | 426.68 | 272.06 | 298.92* | 425.42 | 310.40 |
| | (171.25) | (313.49) | (246.24) | (170.76) | (315.95) | (242.52) |
| ASSIST | 146.57 | 627.54 | 271.26 | 144.14 | 539.70 | 257.53 |
| | (296.52) | (541.07) | (427.21) | (295.59) | (543.71) | (420.96) |
| PASS | -11.27** | -19.22 | -22.07** | -10.32** | -18.12 | -20.12** |
| | (4.46) | (13.50) | (8.99) | (4.46) | (13.85) | (8.80) |
| CROSS | -49.15 | 16.60 | -24.48 | -52.22 | 11.28 | -28.22 |
| | (40.68) | (85.49) | (59.20) | (40.59) | (86.43) | (57.90) |
| TACKLE | -196.15** | -261.38* | -236.38 | -196.71** | -266.14* | -219.64 |
| | (89.55) | (136.32) | (157.66) | (89.29) | (137.40) | (155.35) |
| FREE-TF | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | (.) | (.) | (.) | (.) | (.) | (.) |
| SUSPENSION | -35.97 | -48.09 | -18.85 | -35.76 | -56.87 | -18.31 |
| | (35.34) | (56.36) | (48.15) | (35.12) | (56.72) | (47.42) |
| INJURY | -47.48*** | -62.49*** | -52.27*** | -48.96*** | -60.96*** | -51.86*** |
| | (7.20) | (11.85) | (11.59) | (7.23) | (11.80) | (11.37) |
| ON-LOAN | -47.25*** | -55.28*** | -92.97** | -48.06*** | -57.12*** | -95.39** |
| | (12.64) | (15.69) | (39.13) | (12.62) | (15.74) | (38.31) |
| WW | -1069.25*** | -1236.48*** | -1125.89*** | -1072.79*** | -1251.55*** | -1168.04*** |
| | (201.10) | (289.47) | (342.77) | (200.40) | (291.33) | (332.72) |
| 2007.season | -106.22 | -357.05 | -195.56 | -120.81 | -330.08 | -269.14 |
| | (144.68) | (318.77) | (213.54) | (145.08) | (324.16) | (211.30) |
| 2008.season | 19.82 | 164.06 | -133.72 | -2.67 | 168.71 | -187.76 |
| | (161.07) | (334.93) | (243.23) | (161.14) | (338.87) | (239.38) |
| 2009.season | -59.97 | 196.53 | -202.84 | -97.62 | 183.69 | -272.24 |
| | (168.47) | (361.54) | (228.40) | (169.09) | (369.30) | (224.60) |
| 2010.season | -212.38 | -340.48 | -136.88 | -230.93 | -368.26 | -146.40 |
| | (168.53) | (353.23) | (243.90) | (168.43) | (358.22) | (239.62) |
| 2011.season | -213.27 | -437.64 | -183.02 | -240.70 | -461.46 | -241.49 |
| | (170.79) | (364.13) | (252.05) | (170.96) | (370.90) | (248.23) |
| 2012.season | -156.17 | -615.71 | -42.60 | -195.26 | -652.94* | -92.18 |
| | (181.46) | (376.09) | (263.47) | (181.43) | (381.40) | (258.31) |
| 2013.season | -429.69** | -1072.77*** | -301.76 | -486.72** | -1123.48*** | -372.51 |
| | (195.06) | (401.71) | (273.42) | (195.61) | (410.29) | (267.51) |
| 2014.season | -559.27*** | -1136.31*** | -587.02** | -620.38*** | -1185.96*** | -691.19** |
| | (192.47) | (404.14) | (273.45) | (193.52) | (410.27) | (270.69) |
| 2015.season | -671.78*** | -1320.93*** | -555.46* | -734.35*** | -1383.01*** | -645.96** |
| | (202.45) | (413.87) | (305.93) | (204.04) | (420.81) | (303.87) |
| 2016.season | -716.12*** | -1461.78*** | -497.50 | -783.78*** | -1516.03*** | -606.03* |
| | (216.53) | (429.02) | (317.65) | (215.52) | (436.91) | (307.04) |
| CONSTANT | 1979.17*** | 2889.77*** | 3008.98*** | 2007.93*** | 3004.70*** | 2781.55*** |
| | (362.96) | (635.40) | (516.20) | (358.42) | (638.19) | (519.04) |
| N | 664 | 406 | 258 | 664 | 406 | 258 |
| $R^2$ | 0.36 | 0.51 | 0.44 | 0.36 | 0.50 | 0.46 |

* p<0.10, ** p<0.05, *** p<0.01

### 2.6.16 Table 13 Midfielders' results including season dummy (played minutes as a dependent variable)

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| FE Model | (1) Full Sample | (2) WIN<37% | (3) WIN>37% | (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
| *Panel A: Zero TF player included* | | | | | | |
| TF | 10.87 | -3.54 | 16.57 | 120.71 | 158.61 | -151.29 |
| | (9.41) | (26.54) | (20.73) | (172.18) | (374.45) | (401.59) |
| TFD1 | -8.79* | -40.31* | -10.44** | -290.68*** | -811.86** | -334.35*** |
| | (4.62) | (23.18) | (4.87) | (104.79) | (404.49) | (111.04) |
| GOAL | 936.03*** | 1377.25*** | 971.33** | 921.39*** | 1310.02** | 963.12** |
| | (289.83) | (514.30) | (386.83) | (288.76) | (512.18) | (380.80) |
| ASSIST | -212.28 | 293.89 | -459.59 | -243.89 | 258.33 | -556.96 |
| | (252.57) | (444.97) | (366.60) | (252.44) | (444.45) | (362.90) |
| PASS | 1.06 | -7.39 | 15.88** | 0.08 | -8.56 | 14.01* |
| | (4.87) | (8.19) | (7.34) | (4.86) | (8.20) | (7.24) |
| TACKLE | -112.98** | -162.51** | -66.92 | -118.84** | -183.06** | -72.04 |
| | (50.77) | (79.30) | (77.44) | (50.61) | (80.31) | (76.40) |
| DUEL-LOST | -152.65*** | -181.01*** | -159.30*** | -152.47*** | -182.74*** | -150.83*** |
| | (28.17) | (48.46) | (47.35) | (28.01) | (48.48) | (46.35) |
| FREE-TF | 416.18** | 206.63 | 455.83 | 383.80* | 265.80 | -4.64 |
| | (202.38) | (278.19) | (599.28) | (200.09) | (262.05) | (555.41) |
| SUSPENSION | -6.21 | -34.36 | 33.08 | -5.17 | -34.94 | 32.36 |
| | (24.28) | (37.26) | (32.67) | (24.16) | (37.12) | (32.18) |
| INJURY | -55.65*** | -52.23*** | -65.81*** | -55.43*** | -52.33*** | -65.69*** |
| | (5.94) | (9.35) | (9.10) | (5.90) | (9.37) | (8.90) |
| ON-LOAN | -69.53*** | -77.13*** | -48.46** | -68.77*** | -71.17*** | -44.60** |
| | (15.45) | (24.60) | (21.95) | (15.40) | (25.03) | (21.35) |
| WW | -1175.08*** | -1204.92*** | -1046.81*** | -1180.62*** | -1233.89*** | -992.68*** |
| | (152.50) | (209.08) | (325.80) | (152.56) | (214.87) | (321.87) |
| 2013.season | -209.01*** | -169.57 | -157.19 | -474.79*** | -1196.71*** | -363.22 |
| | (75.60) | (110.10) | (116.86) | (168.55) | (313.71) | (243.88) |
| 2014.season | -285.83*** | -147.29 | -203.21 | -552.57*** | -1145.49*** | -566.22** |
| | (82.29) | (127.12) | (128.21) | (169.10) | (311.30) | (254.03) |
| 2015.season | -365.21*** | -352.11*** | -206.79 | -735.84*** | -1500.26*** | -502.39* |
| | (85.30) | (123.12) | (140.10) | (179.54) | (321.30) | (286.32) |
| 2016.season | -474.28*** | -461.07*** | -388.80*** | -723.13*** | -1530.32*** | -458.33 |
| | (95.44) | (140.56) | (145.60) | (191.15) | (340.96) | (286.52) |
| CONSTANT | 1919.13* | 4041.53*** | 2287.65*** | 1914.52*** | 2988.87*** | 2535.52*** |
| | (1017.32) | (589.44) | (755.27) | (317.72) | (496.23) | (475.01) |
| N | 811 | 481 | 330 | 796 | 487 | 309 |
| $R^2$ | 0.46 | 0.49 | 0.50 | 0.37 | 0.48 | 0.43 |

* p<0.10, ** p<0.05, *** p<0.01

## Table 13 Midfielders' results including season dummy (played minutes as a dependent variable)

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| FE Model | (1) Full Sample | (2) WIN<37% | (3) WIN>37% | (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
| *Panel B: Zero TF player excluded* | | | | | | |
| TF | 16.76 | 70.36* | 45.04 | 227.49 | 1152.63** | -98.75 |
| | (14.02) | (35.64) | (29.47) | (221.43) | (468.85) | (460.35) |
| TFD1 | -7.47 | -61.69** | -9.07* | -288.85*** | -1406.77*** | -322.73*** |
| | (4.58) | (24.52) | (4.99) | (104.30) | (421.19) | (114.28) |
| GOAL | 756.45** | 1413.16** | 1027.09** | 749.82** | 1182.89** | 1090.84*** |
| | (313.96) | (600.64) | (416.56) | (311.61) | (585.37) | (409.07) |
| ASSIST | -455.12 | 715.93 | -718.07* | -503.05* | 668.55 | -848.21** |
| | (286.23) | (560.42) | (406.32) | (285.22) | (546.10) | (404.58) |
| PASS | 4.18 | -3.05 | 14.08* | 3.39 | -6.42 | 13.56 |
| | (5.61) | (9.41) | (8.43) | (5.58) | (9.21) | (8.36) |
| TACKLE | -124.62** | -110.02 | -46.72 | -132.69** | -163.22 | -61.47 |
| | (59.17) | (98.64) | (86.81) | (58.81) | (98.82) | (86.58) |
| DUEL-LOST | -113.93*** | -177.61*** | -133.36** | -113.67*** | -173.42*** | -117.88** |
| | (31.23) | (58.26) | (53.37) | (30.98) | (57.13) | (52.26) |
| FREE-TF | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | (.) | (.) | (.) | (.) | (.) | (.) |
| SUSPENSION | 20.22 | 49.57 | 39.37 | 20.96 | 46.54 | 41.34 |
| | (28.45) | (55.21) | (34.73) | (28.29) | (54.12) | (34.30) |
| INJURY | -53.13*** | -52.65*** | -65.23*** | -53.14*** | -51.64*** | -65.16*** |
| | (6.60) | (11.04) | (9.52) | (6.54) | (10.92) | (9.40) |
| ON-LOAN | -46.33** | -38.10 | -60.77** | -45.07** | -15.08 | -57.45** |
| | (19.60) | (34.60) | (24.79) | (19.49) | (35.41) | (24.56) |
| WW | -1114.27*** | -1041.02*** | -1133.76*** | -1127.69*** | -1174.97*** | -1008.93*** |
| | (163.14) | (226.04) | (337.28) | (162.11) | (225.45) | (332.31) |
| 2013.season | -257.94*** | -167.71 | -189.76 | -247.21*** | -223.37* | -182.28 |
| | (86.14) | (132.05) | (128.87) | (85.50) | (127.81) | (128.03) |
| 2014.season | -393.94*** | -205.80 | -261.36* | -392.43*** | -301.13** | -278.35* |
| | (93.51) | (149.30) | (142.66) | (92.75) | (145.08) | (142.49) |
| 2015.season | -433.55*** | -387.14*** | -266.78* | -431.68*** | -428.08*** | -289.04* |
| | (96.08) | (136.75) | (155.68) | (95.33) | (135.96) | (154.04) |
| 2016.season | -567.78*** | -502.25*** | -490.05*** | -547.19*** | -585.34*** | -464.37*** |
| | (107.87) | (167.79) | (157.18) | (105.12) | (161.80) | (154.87) |
| CONSTANT | 3635.94*** | 2656.11*** | 1965.93** | 3723.88*** | 2820.84*** | 2836.01*** |
| | (532.83) | (835.76) | (838.08) | (531.14) | (808.58) | (847.43) |
| N | 616 | 349 | 267 | 616 | 349 | 267 |
| $R^2$ | 0.47 | 0.53 | 0.53 | 0.47 | 0.55 | 0.54 |

* p<0.10, ** p<0.05, *** p<0.01

## 2.6.17 Table 14 Defenders' results including season dummy (percentage of PT as a dependent variable)

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| FE Model | (1) Full Sample | (2) WIN<37% | (3) WIN>37% | (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
| *Panel A: Zero TF player included* | | | | | | |
| TF | 51.85*** | -27.30 | 70.43** | 690.43*** | -1251.88 | 1529.07** |
| | (14.79) | (99.34) | (31.13) | (221.12) | (834.29) | (609.63) |
| TFD1 | -4.57 | 25.55 | -13.99 | -90.87 | 342.55 | -303.91** |
| | (8.01) | (45.22) | (9.30) | (116.48) | (536.91) | (137.85) |
| GOAL | 74.15 | 1080.19 | 534.51 | 147.98 | 875.56 | 625.02 |
| | (661.61) | (1242.41) | (902.21) | (663.45) | (1218.60) | (889.36) |
| PASS | 12.59** | 3.12 | 24.92*** | 12.54** | 1.90 | 24.45*** |
| | (5.96) | (11.45) | (8.27) | (5.98) | (11.41) | (8.14) |
| INTERCEPTION | -27.10 | -140.08 | 62.33 | -30.22 | -134.67 | 59.00 |
| | (72.99) | (110.16) | (129.15) | (73.28) | (109.91) | (128.79) |
| DUEL-LOST | -67.22 | -23.92 | -168.54** | -74.74 | -32.14 | -189.07*** |
| | (48.12) | (84.23) | (66.51) | (48.38) | (82.74) | (65.89) |
| FREE-TF | 250.53 | 478.55 | 99.83 | 243.47 | 699.67 | 674.65 |
| | (341.17) | (670.77) | (690.81) | (349.66) | (679.92) | (789.31) |
| SUSPENSION | 13.90 | 36.95 | 23.75 | 6.31 | 28.63 | 16.42 |
| | (44.37) | (71.68) | (72.88) | (44.44) | (71.64) | (71.26) |
| INJURY | -66.77*** | -79.22*** | -75.34*** | -67.02*** | -80.23*** | -77.90*** |
| | (7.32) | (11.70) | (11.22) | (7.36) | (11.57) | (11.08) |
| ON-LOAN | -5.99 | -55.38 | 132.73* | -4.11 | -56.75 | 138.48* |
| | (35.88) | (43.61) | (76.77) | (35.99) | (43.31) | (75.63) |
| WW | -1452.22*** | -1442.07*** | -1733.12** | -1502.66*** | -1403.42*** | -1773.43*** |
| | (230.01) | (359.09) | (684.58) | (231.11) | (361.94) | (665.94) |
| 2013.season | -153.03* | -193.61 | 23.51 | -150.72* | -167.50 | 46.93 |
| | (89.22) | (134.67) | (141.11) | (89.48) | (128.50) | (139.82) |
| 2014.season | -263.85*** | -437.08*** | -153.29 | -252.17** | -409.78*** | -117.45 |
| | (99.27) | (148.39) | (169.07) | (99.50) | (147.02) | (165.74) |
| 2015.season | -277.01*** | -362.56** | -227.84 | -269.06** | -325.77** | -211.64 |
| | (105.51) | (157.80) | (183.68) | (106.16) | (154.62) | (179.99) |
| 2016.season | -361.17*** | -316.05* | -371.58* | -330.99*** | -248.30 | -297.11 |
| | (115.39) | (167.76) | (191.33) | (115.69) | (165.07) | (186.24) |
| CONSTANT | 710.72 | 2686.90*** | 965.26 | 896.68 | 3015.00*** | 474.98 |
| | (612.31) | (945.70) | (861.41) | (602.88) | (873.37) | (940.96) |
| N | 653 | 384 | 269 | 653 | 384 | 269 |
| $R^2$ | 0.39 | 0.47 | 0.44 | 0.39 | 0.48 | 0.46 |

** p<0.10, ** p<0.05, *** p<0.01*

**Table 14 Defenders' results including season dummy (percentage of PT as a dependent variable)**

| FE Model | Model 1 (1) Full Sample | (2) WIN<37% | (3) WIN>37% | Model 2 (4) Top Quartile | (5) WIN<37% | (6) WIN>37% |
|---|---|---|---|---|---|---|
| *Panel B: Zero TF player excluded* | | | | | | |
| TF | 73.03*** | -139.57 | 44.20 | 987.63*** | -1125.73 | 1154.14 |
| | (23.65) | (120.01) | (57.67) | (355.60) | (1073.16) | (841.63) |
| TFD1 | -5.14 | 28.67 | -13.60 | -137.90 | 321.20 | -278.86** |
| | (8.16) | (51.47) | (9.51) | (115.85) | (607.78) | (130.28) |
| GOAL | 1788.73** | 1536.39 | 2627.74** | 1831.78** | 1613.89 | 2706.34*** |
| | (799.65) | (1642.34) | (1001.99) | (800.76) | (1638.47) | (983.86) |
| PASS | 16.57** | 9.51 | 26.87*** | 15.41** | 9.71 | 26.46*** |
| | (6.60) | (14.98) | (7.98) | (6.59) | (15.02) | (7.83) |
| INTERCEPTION | -30.91 | -203.53 | 75.29 | -37.80 | -185.85 | 75.40 |
| | (84.70) | (148.03) | (130.01) | (85.30) | (150.18) | (130.28) |
| DUEL-LOST | -71.80 | -48.77 | -105.49 | -84.39 | -60.73 | -120.89* |
| | (55.22) | (111.97) | (70.96) | (55.36) | (110.79) | (67.77) |
| SUSPENSION | 31.74 | 5.38 | 93.46 | 29.06 | -3.65 | 82.07 |
| | (53.31) | (84.40) | (87.81) | (53.23) | (85.54) | (85.63) |
| INJURY | -67.81*** | -82.84*** | -70.67*** | -68.59*** | -81.46*** | -72.56*** |
| | (8.16) | (15.17) | (10.97) | (8.17) | (15.14) | (10.73) |
| ON-LOAN | -38.63 | -51.26 | 0.00 | -43.46 | -51.27 | 0.00 |
| | (40.94) | (46.56) | (.) | (41.19) | (46.65) | (.) |
| WW | -1515.25*** | -1561.78*** | -1888.59*** | -1547.47*** | -1500.83*** | -1894.72*** |
| | (245.37) | (399.95) | (648.57) | (246.93) | (404.02) | (639.95) |
| 2013.season | -144.55 | -183.38 | -36.73 | -143.34 | -159.91 | -15.11 |
| | (100.25) | (185.30) | (135.30) | (100.26) | (174.73) | (133.44) |
| 2014.season | -241.73** | -384.13* | -143.11 | -227.92** | -366.04* | -118.57 |
| | (112.47) | (195.66) | (166.09) | (112.44) | (193.71) | (164.47) |
| 2015.season | -252.21** | -385.93* | -193.90 | -234.21* | -384.67* | -182.61 |
| | (120.16) | (204.96) | (180.67) | (119.79) | (200.99) | (178.24) |
| 2016.season | -338.92** | -263.95 | -384.18** | -288.68** | -244.93 | -321.60* |
| | (135.50) | (228.71) | (188.11) | (133.96) | (223.42) | (180.48) |
| CONSTANT | 1515.86** | 2865.96*** | 1711.30*** | 1192.93* | 2769.97*** | 1461.21** |
| | (592.13) | (744.22) | (624.72) | (632.12) | (719.87) | (681.99) |
| N | 485 | 258 | 227 | 485 | 258 | 227 |
| $R^2$ | 0.41 | 0.51 | 0.43 | 0.40 | 0.51 | 0.45 |

** p<0.10, ** p<0.05, *** p<0.01*

2.6.18 Table 15 The difference in some performance metrics between expensive players (worth £10 or more) and others

| Variable | Obs. | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| *Panel A: Performance of players worth more than £10m* | | | | | |
| Goal | 214 | 0.4467 | 0.2255 | 0 | 1.2209 |
| Assist | | 0.1916 | 0.1289 | 0 | 0.6274 |
| Shoot_accuracy | | 0.3971 | 0.0874 | 0.1764 | 1.0806 |
| Tackle | | 0.9817 | 0.5138 | 0.1113 | 3 |
| Pass | | 30.7 | 9.04 | 9.2 | 60.3 |
| *Panel B: Performance of players worth less than £10m* | | | | | |
| Goal | 582 | 0.3067 | 0.2116 | 0 | 1.15 |
| Assist | | 0.1408 | 0.1199 | 0 | 0.6225 |
| Shoot_accuracy | | 0.3686 | 0.1018 | 0 | 1 |
| Tackle | | 1.09 | 0.6046 | 0 | 3.17 |
| Pass | | 27.5 | 9.4 | 0 | 174 |

*Notes*: Metrics are presented as per 90 minutes of a game. For instance, a top (expensive) player needs, on average, almost two games to score one goal.

### 2.6.19 Table 16 Definitions of performance metrics for the *Perf*$_{it}$ variable and *Control* variables

| Performance Metric | Definition |
|---|---|
| Goal per Minute (GPM) | Number of goals per 90 minutes. |
| Assist per Minute (APM) | Number of assists per 90 minutes. |
| Cross per Minute (CPM) | Number of crosses per 90 minutes.<br>A medium-long range ball passing towards the opponent's area. |
| Pass per Minute (PPM) | Number of passes per 90 minutes.<br>All ball passes. |
| Tackle per Minute (TPM) | Number of tackles per 90 minutes.<br>to dispossess an opponent of the ball. |
| Interception per Minute (IPM) | Number of interceptions per 90 minutes.<br>When the ball is intended for a player of the same team but caught by a player of the opposing team. |
| Duel-lost per Minute (DLM) | Number of duel-losts per 90 minutes.<br>When the ball is lost to the opponent. |

| Control Variables | Definition |
|---|---|
| WW | Winter window transfers. |
| Win | Win ratio of teams; calculated by dividing the won games over 38 during the season. |
| Suspension | Number of matches that a player is legally suspended. |
| On loan | Number of games not played while being on loan (either part or full season) to another team. |
| Manager change | A managerial change that takes place after a player transfer date. |
| Free transfer | A player transferred to another team with zero transfer fee. |
| Injury | Number of matches that a player could not play due to an injury. |

**Notes*: The relevant data are taken from Premierleague.com, the official website that represents the PL authority. We opted to adjust the variables per 90 minutes rather than per match appearance. This gives a proper reflection of the playing metrics and players' productivity per full game minute.

# 3. Do British Premier League managers exhibit home bias?

Racial discrimination is a topic of considerable debate in the labour economics literature. The focus of the related empirical discrimination literature is on the earnings gaps between specific groups. Those groups have to share similar levels of productivity or talent to enable one to conclude their pay gap as a discrimination factor. Discrimination through unequal pay is estimated through constructing earning functions (Szymanski, 2000). After controlling for productivity or performance, there should be no residual race or ethnicity difference in earnings in a non-discriminatory setup. However, if there are some unobserved productivity measurements that are correlated with a specific group, then the empirical estimates of discrimination might be biased (Heckman, 1998). Szymanski (2000) argued that this critique may be resolved by an alternative market test approach. In sport economics studies, productivity measurements are much more transparent than in other labour economic settings.

Studies concerning European professional sports are sparse relative to those on North American sports. This is primarily due to the availability of relevant data (e.g. individual player salaries, wide range of productivity-related data). A number of studies (e.g. Pedace 2006; Szymanski 2000; Wilson & Ying 2003) provided evidence for some discrimination within the European football profession with regard to race, ethnicity, and nationality. Wage equation estimates are the cornerstone of an employer's taste-based discrimination. With a relative scarcity of productivity measurements, the alternative way is to relate wage (or salary) data with teams' productivity measures, which are mainly measured by their position in the league rank. Both Szymanski and Pedace found that clubs' wage payroll could explain nearly 90% of the variation in team productivity. Hence, if there is discrimination in the wage structure, then some personal players' characteristics, such as race and nationality, may have a significant impact on team performance. In other words, teams may achieve a given productivity or success level with a lower wage payroll.

Pedace (2008) found evidence that players from South America received preferential treatment in terms of salary. Teams with a larger number of South American players tend to perform poorly, and this implies that these players are overpaid. However, this result is justified by the increased stadium attendance when a team plays with a South American player on the pitch. The author considers it a rational reaction from club owners to hire South American players if they seek more gate revenue. The findings of Szymanski (2000) indicated that teams that hired more black players tended to perform better. This implies that there was racial discrimination in the sense that black players were being underpaid. This was based on the evidence that wages reflected team productivity (performance), and hence performance measures were not expected to be significant. However, if it turned out to be significant, then the implication is that a group of players were underpaid if the team performed systematically above average with them, and would be overpaid if the team performed poorly with them. In these studies, team performance was used as a dependent variable, and hence discrimination was determined through team (or club) level measurement. This is an intuitive approach, as the club wage bill is almost the only factor that can be matched with productivity measurement at the club level, provided that it explains about 90% of its variability.

This paper contributes to the literature by examining the nationality effect on player playing time in the English Premier League (EPL) from 2012–2016. The empirical evidence establishes a unique discrimination effect as it accounts for the effective use of players (i.e. their playing time relative to their performance). We employ a novel market test approach that complements the established wage equation estimate approach while controlling for player performance metrics. Recently, the explanatory power of wages for team success has been considerably reduced to about half. We conducted a regression of team position on wages for

33 EPL clubs during the period 2009–2016 and showed that wages explained about 46% of the variation in team performance. This result challenges both the financial efficiency of clubs and their discrimination levels. While escalating transfer fees for players (which are mainly due to the exponential increases in TV broadcasting revenues) are beyond the full control of clubs, the best use of their players should be at their discretion. We draw on this conjecture to estimate the effect of nationality (both of players and managers) on player playing time.

The second contribution is that the analysis was conducted using a unique dataset of player performance metrics and their transfer fees. These have not been utilised, to the best of our knowledge, in the discrimination literature to date. The data for player performance were first published in the 2006 season. We have hand-collected the data from the relevant sporting websites, which supply them in a sporting context, thereby constructing a novel sporting dataset for empirical research.

Our third contribution relates to the behaviour of British football managers. The findings suggest that British managers are biased towards their home players. We found evidence that British managers allocated extra playing time to British players of about 100 minutes per season. This extra playing time cannot be explained merely by the players' performance levels and their transfer fees and playing positions. Accordingly, discrimination in this context was evident. On the other hand, we found that if the manager nationality was ignored, British players still got more playing time than non-British players. Specifically, they got more playing time than European, South American, and African players by 102, 170, and 196 minutes per season, respectively.

We also re-examine the models of Szymanski (2000) and Pedace (2008) with our recent dataset and find no evidence for discrimination, either against non-white players or foreign nationalities. These findings suggest a shift in the discrimination effect towards the individual player level from the club level.

### 3.1 Literature Review

Racial discrimination in labour markets has been well documented. Becker (1971) found that owners (firms) who were non-discriminatory against their workers had a competitive advantage over discriminating firms. If firms systematically generate higher profits while hiring more of a particular group of workers than average, then there is evidence of discrimination (Szymanski, 2000). This implies that this particular group may earn less while possessing equivalent talent or productivity levels to their higher earning peers. Less discriminating employers seek to hire those who are earning a lower market wage rate, and hence discrimination would be competed away (Arrow, 1973). Szymanski indicated that employers who discriminated the most would earn lower profits than their less discriminating peers, but they would compensate that profit gap with higher psychic profits. Becker (1971) related wage discrimination to three sources: customer preferences, owners' prejudice, and co-workers' preferences.

Gwartney and Haworth (1974) researched racial discrimination in a sports setting, namely Major League Baseball (MLB). They showed that attendance (an index for customer preference) could be significantly affected by the team's racial composition. Burdekin et al. (2005) found a similar attendance preference pattern in the National Basketball Association (NBA). Scully (1974) found some evidence for attendance preference to watch non-black players in the MLB. In Europe, Carnibella et al. (1996) found that during the 1970s and 1980s, discrimination persisted among fans through racist chants in stadiums, but some of this has been moderated in recent years.

Kahn and Sherer (1988) found evidence of owners' discrimination in the sense that NBA black players outperformed their white peers while earning similar salary levels. Hamilton (1997) found that white players in the NBA earned an 18% premium in the upper quantile of the salary distribution, while Holmes (2011) found that black player salaries in the MLS were discounted by 25% in the lower quartile. Other studies also uncovered evidence of earnings premiums of more than 10% for white players in the NBA (e.g. Koch & Vander Hill 1988; Wallace 1988; Brown, Spiro, & Keenan 1991). However, Bodvarsson and Bradstow (1999) provided evidence of diminished wage discrimination in the NBA during the 1990s. Sommers and Quinton (1982) found no evidence of a wage differential based on teams' racial composition.

In a co-worker or teammate environment, the story of MLB player Jackie Robinson back in 1947 has been documented in a film named "42," which refers to the player's shirt number. He suffered discrimination and humiliation from his teammates and opponents. The story is told as a demonstration and indictment of the racial discrimination phenomenon.

Wage discrimination is conventionally identified through an earnings function where it could be influenced by personal characteristics matched with productivity (Szymanski, 2000). Heckman (1998) raised the problem of possible omitted-variable bias associated with that approach where discrimination may falsely appear due to unobserved characteristics that may correlate with one group's productivity.

Szymanski (2000) proposed a market test approach utilising a sport setting that aimed to overcome the omitted-variable bias. He built on the assumption that English football was a competitive labour market, where player performance was reasonably measurable, and wages would tend to reliably explain their productivity and hence team performance. His sample panel regression of teams' performance (represented by team position in the league rank) on their wages produced an $R^2$ of almost 90%. This supports the competitive market hypothesis.

In addition, European sports clubs are generally assumed to be utility maximisers (i.e. primarily seeking team success subject to a break-even constraint) rather than profit maximisers, which prevail in US sports (Sloane, 2006). In North America, sports are bounded by financial and competition constraints such as salary caps, players drafts, selling players for cash, and the absence of international competitions. The freedom of such constraints in Europe should lead to more efficient sporting decisions, even though Sloane (2015) indicated that evidence was limited in that respect. Another aspect of competition is stadium attendance (spectators) and sponsorship. It is plausible to assume that spectators will seek entertainment through watching talented players, and hence clubs will compete to acquire them. Club density is high in the UK, where fifty clubs might exist within a hundred miles (e.g. in Manchester) (Szymanski, 2000). It surpasses those of North America, resulting in a higher competition for talent. Szymanski and Zimbalist (2006) regarded American baseball as an effective monopoly under a closed league system, while football (or soccer) is more competitive under an open league system where promotions and relegations are allowed. This explains the difference in profitability.

If a particular group (e.g. specific nationality or race) of players appears to be positively related to team success, this is an indication of wage discrimination against them. Team success should not be affected by players' personal characteristics in a discrimination-free environment. Szymanski (2000) found evidence of wage discrimination against black players in English football for the 1978–93 period. Teams recruiting a higher proportion of black players tended to perform better, and their ability surpassed their white peers in certain aspects of the game, such as career longevity and participation in their national teams. That discrimination effect was higher during the 1986–93 period, which is mostly attributed to the increasing immigration to the UK in general and the increased number of black players in particular. His results also showed that teams who discriminate—in the sense that they do not hire black players—would end up paying a 5% wage premium just to maintain their position in the league table. Szymanski concluded that clubs hiring more black players tended to perform better, and that black players were paid less for their talent than white players with equal ability.

In parallel, Pedace (2008) found evidence of wage discrimination in a nationality context in English football. However, this came in a form of preferential treatment in which players from South America received higher salaries, but their teams did not perform better. The author attributes this behaviour as a rational owner's reaction to the evidence of increased stadium attendance that is associated with the appearance of a South American player on the team. It is perceived as a revenue-seeking approach.

Wilson and Ying (2003) used a specification similar to Szymanski (2000) to determine discrimination. They used both performance and attendance regressions. Their dataset covered the top five European leagues (England's Premiership, France's Le Championnat, Germany's Bundesliga, Italy's Serie A, and Spain's Primera Division) during the 1997–2000 period, which were post-Bosman seasons. Their performance regression results suggest that the more players from South America and Eastern European countries were used, the better teams performed. However, teams were not hiring an adequate composition of such nationalities. Attendance regression results do not indicate fans' discrimination against non-domestic players. Therefore, Wilson and Ying concluded that management and/or ownership preferences would lead to lower hiring of non-domestic talent (i.e. race discrimination). These results were criticised by Pedace (2008) as the lack of salary controls in their specification leads to misinterpretation of the positive coefficients of nationalities. Moreover, the exclusion of team fixed effect may have biased the estimates due to its probable correlation with time-varying team nationality.

After the Bosman rule in 1995, clubs were exposed to a significant change towards a more internationalised player transfer market. Subject to their budget constraints, clubs would compete and seek to acquire the best players. Despite that, clubs in different league positions could experience relative concentration of a particular nationality player group (see Table 5). Such concentration or hiring bias is not a simple reflection of an open transfer market. McGovern (2002) rationalised that management behaviour by emphasising their social network roles. For instance, some managers or coaches in the Premier League praised the merits of Northern European players (especially Scandinavians) and, on the other hand, the unique styles of play of South American players. The author challenged the literature of globally integrated markets (e.g. Dicken, 1998; Hirst & Thompson, 1996) by testing the market for professional footballers. He argued that players' market trends evolved over regional rather than global lines. While the decision to hire players has its economic base, it is socially embedded with recruitment networks and personal relationships. When clubs hire players, they seek to resemble their own local concept of a professional player. Their scouts or foreign sources draw on talents who can match the local standards in terms of climate, culture, language, and style of football.

### 3.2 Research Hypotheses

In the sport economics literature, wage functions are widely used to test for discrimination. Within English football, Szymanski (2000) found that black players earned less than their white peers. Pedace (2008) found that players from South America earned more than those from other continents, provided that their talents were matched with their peers. They used team performance as the dependent variable. Until these studies were published, there was a lack of sufficient player performance measures, and hence, a wage function seemed the most convenient and appropriate specification. This approach is reinforced by the evidence that supports clubs' efficiency, where wages explain about 90% of team performance variation and hence their estimation is based on the crucial assumption that the market for players is efficient. In recent years, even though the explanatory power of wages has remained reasonably high, it falls to about half of its 90% level. Starting in the 2006 season, more sophisticated player performance measures were published. We explore these data and establish a new perspective for addressing discrimination that complements the existing literature. Building on the work of Szymanski (2000) and Pedace (2008), we use player playing time as a dependent variable and test for discrimination using a player performance function where individual player productivity metrics are employed.

To test for a nationality effect, we estimate models at the player and club levels. First, we test for a discrimination effect in player playing time while controlling for his individual performance. We devised two types of hypotheses for this test. One uses a common nationality variable shared between managers and players as an explanatory variable, while the other constructs five nationality subgroups, using a dummy for each player relating him to his specific subgroup. These are two distinct hypotheses, but both are concerned with nationality rather than race, where the second one disregards the fact that a manager may or may not share nationality with his players. Second, we re-examine the work of Szymanski (2000) and Pedace

(2008) using a recent EPL dataset. This can capture discrimination effects on team performance and hence involves a club-level hypothesis.

### 3.2.1 Player level hypotheses

H1: Sharing a common cultural background (nationality) between managers and players is independent of the players' playing time after controlling for performance.

H1A: Managers sharing a common cultural background (nationality) with players increase their playing time after controlling for performance.

The second hypothesis is as follows:

H2: Managers do not discriminate in playing time against any specific nationality subgroup after controlling for player performance.

H2A: Managers discriminate in playing time against specific nationality subgroups, even after controlling for player performance.

### 3.2.2 Club level hypotheses

H3: Managers do not discriminate against particular player subgroups after controlling for club wages and number of squad players.

H3A: Managers discriminate against particular player subgroups after controlling for club wages and number of squad players.

### 3.3 Data and Methodology

The sample includes all PL players who made a minimum of one appearance in any season for the period spanning 2012/13 to 2016/17. We opted to choose the most recent five seasons of available data as they sum to form sufficient (1809) observations. It also takes into account the new Sky TV contract with the PL that began in the 2013/14 season, which is higher in value and associated with higher spending on players (see Table 4 on p. 42). Each season spent by a player at a club is considered an observation. The maximum number of minutes a player can play in a single season is 3420 (90 minutes per game over a total of 38 games).

We used a number of data sources (websites) to build the dataset, including Transfermarkt.co.uk for player playing time metrics and personal profiles (through which we observe their pictures to determine skin colours), Premierleague.com for players' performance metrics, Soccerbase.com for managerial records, and players' transfer fees and signing dates. The performance metrics provided are derived from Opta Sports, which is the leading sports data provider for major sports entities. Club turnover and wage bill data are obtained from the David Conn blog in the Guardian newspaper. These are derived from the published annual reports at Companies House.

To examine discrimination in the PL, we tested for the nationality effect on player playing time. This was estimated based on the player-level hypotheses H1 and H2. We departed from the widely used earning function to a unique market test where (player) playing time was employed as the dependent variable. The rational decision of employing a player in the best interests of a team should be based on his talent and performance. However, other factors with respect to the personal player characteristics may play a role. McGovern (2002) argued that the decision to hire a player in English football was embedded within the social networks of a club's managers and scouts. They tended to praise players who resembled the local culture professional standards even if foreign players with similar quality levels were available. Therefore, we predicted a nationality effect that influences player playing time even after controlling for performance.

Descriptive statistics for playing time of different player nationalities are reported in Table 1.

[ insert Table 1 here]

These show that British players played more than non-British players, with an average of 1827 minutes per season compared to 1749 minutes per season, respectively. The difference of 78 minutes is close to a full game, and hence British players played on average almost one more game per season than non-British players. Additionally, British players got more playing time even under different managerial scenarios. They got 1834 minutes of playing time compared to 1700 minutes for non-British players while playing under a British manager, and 1815 minutes compared to 1787 minutes for non-British players while playing under a non-British manager.

To examine the representation of different nationalities among clubs, we divided the teams into four categories, and this is presented in Table 2.

[insert Table 2 here]

These are based on five league position levels (i.e. first five teams or top teams followed by second five teams, third five teams, and bottom five teams). This reveals player recruitment patterns among different club levels. Table 2 shows that the top five teams tended to recruit a majority of European players. A team in that level had a mean of eight European players in its squad, and this decreased in the lower league levels, reaching a mean of four European players for the teams in the lower half of the league table. By contrast, teams in the last three levels had a mean of nine British players compared to six players for the top five teams. These statistics are generally consistent with the findings of Pedace (2008), who showed that one additional player from the British Isles decreased team performance by one position on average, while one additional player from Europe increased team performance by one position on average. Top teams also recruited more South American players than the rest (a mean of 2.7 compared to 1.5), while they recruited fewer African players (a mean of 1.3 compared to 2.2-2.5).

The representation of British managers according to clubs' turnover and league positions is presented in Table 3.

[insert Table 3 here]

This shows the patterns of recruiting British managers among different club levels. It reveals that clubs who earned the most (a turnover above the median of £121 million) tended to recruit fewer British managers. There are only 233 observations in that category that include a British manager compared to 714 observations in clubs who earned less than the median. The prevalence of British players within the smaller teams may be attributed to the British managers of those teams who are expected to have a preference for their home players over the uncertainty of hiring players from abroad. That preference may be more pronounced when clubs' resources are limited compared to top teams, and their overseas networks for players are consequently limited.

For a manifestation of different players' characteristics with respect to their nationalities, we present player performance descriptive statistics in Table 4.

[insert Table 4 here]

They show that British players' performance records were mediocre relative to players from other nationalities. The goals average for British players was 0.121, which is below the goal metric mean (for the full sample) of 0.135, while for non-British players the goals average was 0.146.[7] British players were also less expensive.

In a player-level analysis, the nationality effect is estimated by means of a random-effect (RE) panel regression model. This is to allow for the inclusion of time-invariant variables such as player nationality and playing position. We also include club and season dummies in the estimation to capture specific club characteristics and time trends to control for the heterogeneity of these unobserved factors. As Pedace (2008) indicated, including club and year dummies (or fixed effects) addresses the possible correlation between them and ethnic player appearances.

In a club-level analysis, we estimated the ethnicity and nationality effects by means of fixed-effects models. This was to allow for a within-estimation among different clubs, whereas nationality and ethnicity are time-varying covariates, as they encompass the number of relevant players in the team squad.

We run the following RE model to test hypothesis H1:

$$PT_{it} = \alpha + \beta_{1,n}\, NE_{it} + + \eta MC + ¥\, Perf_{it} + \psi\, Controls_{it} + C_i + \boldsymbol{\theta}_t V_t + \boldsymbol{\varepsilon_{it}} \qquad \textbf{(1)}$$

where $PT_{it}$ is the dependent variable that represents playing time for each player $_i$ in season $_t$. This is measured by the total minutes played per season. The *NE* variable is a vector that captures the nationality effect, by which a manager might allocate extra playing time to a player who shares his nationality or continent. This vector accounts for four possible scenarios (four dummies) and takes a value of 1 or 0 otherwise in each. We ran four versions of model (1) to

[7] We conducted an equality of means test and it showed that the British players' goals average was statistically different (less) than non-British players' goals average and significant at the 1% level. The maximum of the goal metric records is 1.02. The highest and consistent (played at least 3 seasons with above average playing time) records were achieved by players like Aguero in Manchester City, Lukaku in Everton, Walcott in Arsenal, and Harry Kane in Tottenham.

test for each scenario separately. Each one has a distinct acronym for the *NE* effect: (i) BB is for a British player and a British manager; (ii) BN is for a British manager and a non-British player; (ii) NB is for a non-British manager and a British player; and (iv) NN is for a non-British manager and a non-British player. There are 947 sample observations (out of 1809) where the team manager is British. The *MC* variable stands for managerial change. This is an interaction dummy that takes a value of 1, or 0 otherwise, if a new manager is appointed to the team after the start of a season (i.e. after August), and simultaneously there is a change in the manager nationality (i.e. either from British to non-British or vice versa). This captures the effect of a manager change on player playing time subject to a change in the manager's nationality.

The *perf* variable is a vector that captures player performance by accounting for the following performance metrics: goals, passes, crosses, and interceptions per player per season (see Table 13 for their definition). The number of each metric per season is divided by the number of games a player had during that season, and then multiplied by 90 to reflect a full game time. These metrics are common and available for players in all positions. We dropped both the assist and tackle metrics due to their high correlation with cross and interception metrics, respectively (see Table 5 for performance correlation matrix). The inclusion of defensive measures such as interceptions relaxes the argument of Pedace (2008) that most sports performance measures used in the literature (e.g. Lucifora & Simmons, 2003) are biased towards the offensive side of player roles. *Controls* is a vector of three control variables: injury, suspension, and on-loan players (i.e. when a player in team A is on loan to team B for only part of the season). We used the number of games to account for a player being unavailable due to any of those circumstances. *C* is club fixed effects, and *V* are season dummies. The *C* variable is essential even though the dataset is based on player observations. This is to allow for different team qualities that may affect player performance (e.g. midfielders who do not sufficiently provide their strikers with proper assistance).

There are other factors that may influence player playing time and are indirectly related to player performance. One is the player transfer fee. Although it does not reveal specific information with respect to player performance metrics, it works as a proxy for the general quality of a player. Higher transfer fees are expected to be associated with higher-level talent and better performance. Thus, managers might be inclined, especially in the early stage of player employment, to allocate him extra playing to allow for that talent to be revealed or to adapt to the English football atmosphere if it is a fresh foreign talent. The other is the player position. Different teams might have similar proportions of players employed in each playing position. However, it is expected that playing position would have a large impact on the retention of players in their teams. Goddard and Wilson (2009) indicated that those retained players had more appearances than others, and this is a natural outcome. In other words, player retention is closely related to player playing time.

To address possible omitted variable bias in model (1), we include these two variables (transfer fee and playing position) in the specification as follows:

$$PT_{it} = \alpha + \beta_{1,n}\, NE_{it} + ¥\, Perf_{it} + øControls_{it} + \eta MC + \theta TF_{it} + \psi PP_{it} + C_i + \boldsymbol{\theta}_t V_t + \boldsymbol{\varepsilon_{it}} \quad \textbf{(2)}$$

where *TF* is the player transfer fee measured in £million and *PP* is a dummy that indicates the player position (taking a value of 1 or 0 for each position). There are three main player outfield positions: striker (or forward), midfielder, and defender. In the regressions, we drop the midfield position to use as a basis for the other positions.

We ran the following RE model to test the H2 hypothesis:

$$PT_{it} = \alpha + \beta_{1,n}\ Nationalty_{it} + ¥\ Perf_{it} + \eta MC + \ C_i + \boldsymbol{\theta}_t V_t + \boldsymbol{\varepsilon_{it}} \qquad \textbf{(3)}$$

The variables $PT_{it}$, $Perf_{it}$, $C_i$, and $V_t$ are the same as those in models (1) and (2). The rest are different as follows. The *Nationality* variable consists of five subgroups: Africa, Europe, Britain, South America, and Asia. Each one is a dummy that takes a value of 1 if the player nationality belongs to that subgroup, and 0 otherwise. *MC* is a dummy that takes a value of 1 if a new manager is appointed to the team after the start of the season. This is different from the *MC* variable in model (1) in terms of the manager nationality consideration, which is disregarded in this case according to the hypotheses H1 and H2 differences.

Consistent with the modification in model (1), we then tested the hypothesis by including player transfer fees and playing position in specification (3), and the model is as follows:

$$PT_{it} = \alpha + \beta_{1,n}\ Nationalty_{it} + ¥\ Perf_{it} + \eta MC + \ \theta TF_{it} + \ \psi PP_{it} + \ C_i + \boldsymbol{\theta}_t V_t + \boldsymbol{\varepsilon_{it}} \quad \textbf{(4)}$$

where the *TF* and *PP* variables are identical to specification (2).

We run the following fixed effect model to test the H3 hypothesis:

$$P_{ct} = \alpha + \beta_1\ nonWhite_{ct} + \beta_2\ logWage_{ct} + \beta_3\ NPlayer_{ct} + \boldsymbol{\theta}_t V_t + \boldsymbol{\varepsilon_{ct}} \quad \textbf{(5)}$$

This demonstrates a re-examination of Szymanski's (2000) model, which uses a fixed effect approach. We included season dummies to capture time trends and used an EPL dataset, while other league divisions were not considered. Following Szymanski's specification, the dependent variable was the log of position calculated as ln[ P/ (21-P) ]. The *logWage* variable is the log of wages for club *i* in season *t*. Wage is represented by the aggregate wage bill for clubs, whereas wage payroll for individual players is confidential. The *NPlayer* variable is the number of squad players (or squad size) that clubs use per season. The *non-white* variable is the number of non-white players each club uses per season. A minimum of a single appearance was required for a player in both variables to be included in the sample. It is not clear how Szymanski (2000) selected black players. In this sample, we chose a non-white player based on the following criteria[8]: (i) Being African and/or from an African origin (i.e. even if a player held a different nationality than his origin). A salient example of this is an English player with Jamaican origins. (ii) Having clearly black-coloured skin (i.e. dark black). (iii) Originating from a black or non-white ethnicity. For instance, British players with a black ethnicity may have a skin colour that is not clearly black but is simultaneously non-white. (iv) South Americans and Asians are not considered black unless their skin is dark black. The $V_i$ represents season dummies.

We re-estimate model (3) using four nationality groups, rather than black players' subgroups, to test for nationality discrimination.

$$P_{it} = \alpha + \beta_1 \, logWage_{it} + \beta_2 \, NPlayer_{it} + \beta_3 British\ Isles_{it} +$$

$$\beta_4 Europe_{it} + \beta_5 Africa_{it} + \beta_6 \, S.America_{it} + \lambda MC + \boldsymbol{\theta}_t V_t + \boldsymbol{\varepsilon_{it}} \qquad \textbf{(6)}$$

Each nationality variable represents the number of players of that nationality that clubs use for their squad per season, relative to the average per season. *British Isles* represents British players from England, Wales, Scotland, Ireland, and Northern Ireland). The *Europe* variable represents players from Europe, and the same goes for *Africa* and *South America*.

[8] We used relevant data provided on player profiles from data sources that we used for the purpose of player groupings. However, some of the groupings are based on our discretion, as they were not explicitly mentioned on the profiles.

## 3.4 Empirical Analysis

### 3.4.1 Player level regression

The regression results for models (1) and (2) are presented in Tables 6 and 7, respectively.

[insert Tables 6 and 7 here]

The dependent variable is playing time represented by total played minutes per season. Model (1) results show that the BB coefficient was significantly positive at the 5% level. The economic impact of belonging to a particular nationality was considerable. A British player playing under a British manager got on average an extra 107 minutes playing time per season, which is equivalent to more than a full game's minutes. This result suggests that player playing time is positively associated with belonging to a particular nationality. Specifically, a British player who played under a British manager was likely to get extra playing time relative to non-British players. The BN coefficient was significantly negative at the 10% level. The economic impact was also considerable with a 95-minute coefficient, equivalent to just over a full game, including extra (or injury) time. The negative coefficient means that British managers allocated less playing time to non-British players in their teams. These results support the taste-based discrimination hypothesis H1A. The coefficients of NB and NN in columns 3 and 4 were not significant, and this implies that non-British managers were not discriminating in player playing time.

All the performance metric coefficients were significant at the 1% level, except for the crosses, which were not. Goals were the most significant playing time predictor, with a high coefficient of more than 600 minutes. This was followed by interceptions with a more than 120-minute coefficient and, lastly, passes with an above 7-minute coefficient (the economic impact of passes was also meaningful, as the average number of passes per game was naturally much higher than other performance metrics). Goals were expected to have an immense impact on playing time, as it is a clear-cut indication of a player's contribution to team success. Interceptions, while correlated with other defensive performance measures such as tackles, proxy for the contribution to pressing and defending aspects of teams, while passes proxy for the contribution to teamwork in general. These coefficient signs are intuitive and cover most of the playing attributes in the outfield. The MC coefficients were insignificant in all four scenarios. While there is some evidence in the literature that managerial changes may affect team performance (Pedace, 2008), it seems not to have affected player playing time. Injury and on-loan coefficients were, as expected, highly significant with a negative effect, as they made players unavailable, typically for several weeks or months. The suspension coefficient was interestingly significantly positive, suggesting that suspended players may possess talent that secured them playing time that outweighed their suspension period.

When we added player transfer fees and playing position to the general specification in model (1), the coefficients changed in significance and magnitude. We first ran the regression including only transfer fees, and this did not change the general pattern in (1), except for the NN coefficient, which became significant at the 10% level (this result is not reported in Table 7). In addition, the TF added variable was highly significant with a 23-minute coefficient, suggesting that higher transfer fees predicted more playing time. We then ran the full regression specification (2), and consequently the BB coefficient significance was reduced to the 10% level with a slight decrease in magnitude to 100 minutes. The NB coefficient turned out to be insignificant. The goals' coefficient dramatically increased to more than 1000 minutes, while crosses and interceptions were insignificant and passes remained highly significant with similar magnitude. This emphasises the player position role in determining the importance of each performance metric. The position coefficients were highly significant with a high economic impact, where the striker coefficient was negative with 285 minutes magnitude, and the defender coefficient was 407 minutes. These are interpreted with respect to the midfield position, and they suggest that strikers were more likely to be rotated, while defenders were expected to play the most. These are consistent with Goddard and Wilson's (2009) findings in the sense that playing positions had a considerable impact on playing time and revealed a high predictive power towards it. To wrap up, we conclude that British managers exhibited home bias as they allocated extra playing time to British players, even after controlling for performance. The result was robust to the inclusion of player transfer fees and playing position in the specification in (1).

The results were driven by teams that were in the top half of the league table (top 10). We focused on the BB variable for this purpose, as the other scenario variables were not significant in model (2). The results of the subsample regressions are presented in Table 8.

[insert Table 8 here]

Specifically, teams ranked in the 5th to 10th places in the league table (see Table 2) had more influence on the results.[9] The BB coefficient was highly significant at the 5% level, with a stronger economic impact of 174 minutes of playing time. This coefficient was not significant within the lower ranked teams. The implication is that British managers tended to allocate extra playing time to British players in the top ranks of the league table.

Clubs can freely trade players in the transfer market that takes place twice a year (between May and August in summer and in January in winter). Owners and/or team managers are able to recruit their required level of professional players subject to their budget constraints and UK immigration barriers.[10] We tested that conjecture using model (3), and the results are presented in Table 9.

---

[9] These teams included Liverpool, Westham, West Bromwich, Swansea and Stoke who were managed by a British manager for most of the 2012–2016 seasons and simultaneously achieved their top 10 ranks more than once.
[10] In order for foreign players to get a work permit to play in the UK there are some pre-requisites which have to be filled, such as playing for the national team and the position of that team in the FIFA rankings.

[insert Table 9 here]

The dependent variable of model (3) is playing time represented by total played minutes per season. The African coefficient was significant at the 1% level in most cases when matched against the European and British coefficients. These included the regressions when Europe, Britain, and Africa variables were dropped, as per the estimation specification. When Europe and Britain variables were dropped, then the Africa coefficients showed a negative sign, while if the Africa variable was dropped, the Europe and Britain coefficients showed a positive sign. This indicates less playing time for African players. The economic impact was large, with an Africa coefficient of 200 and 253 minutes less than Europe and Britain, respectively. On the other hand, Europe and Britain coefficients were 215 and 267 minutes, respectively, against the Africa variable. The results suggest that African players got less playing time per season compared to their British and European peers, and hence there is evidence of taste-based discrimination against African players. The results support the discrimination hypothesis H2A and are also generally consistent with the findings of Szymanski (2000). He showed evidence of wage discrimination against black players, of which Africans form a large part. In addition, part of the result was consistent with those pertaining to hypothesis H1A, where British players got extra playing time per season compared to non-British players.

Nonetheless, the modifications in model (4) have a considerable impact on the estimates. These results are presented in Table 10.

[insert Table 10 here]

The Britain coefficients were highly significant at the 5% and 1% levels against the South American and Africa coefficients, respectively, and significant at the 10% level against the Europe coefficient. The economic impact was large, with a magnitude of 170, 197, and 103 minutes against South America, Africa, and Europe coefficients, respectively. The results support the discrimination hypothesis H2A and were robust to the inclusion of player transfer fees and playing position. We conclude that the findings in models (3) and (4) combined to yield strong evidence for discrimination in playing time in favour of British players. By contrast, the discrimination against African players in model (3) reduces in extent in model (4), as the majority of the other nationality coefficients are also significant in a negative direction against the British coefficient. This implies that not only African players are being discriminated against, but that they indirectly suffer from greater playing time than British players. The descriptive statistics in Table 4 provide support to the evidence that British players are favoured in their playing time when quality level is held constant.

McGovern (2002) provided evidence that clubs used their social networks to hire players who matched their preferred local standards. Team scouts were likely to hire players they knew individually or through their outsourced scouting links. In other words, the availability of equal foreign talent might not alter the scout’s decision to hire their preferred locals. The evidence pertains to hiring players but not playing them. Nevertheless, if the assumption that those hired players played most of the time is true, then one may infer that the behaviour could, to some extent, be rationalised.

### 3.4.2 Club level regression

The regression results of models (5) and (6) are presented in Tables 11 and 12, respectively.

[insert Table 11 here]

The dependent variable in model (5) is the team rank in the league table (or its position), represented by the log of position equation suggested by Szymanski (2000). The non-white coefficient is significantly positive, but only at the 10% level. Its economic impact was not large compared to the other main predictors (club wages and size of team squad). We recall that the value of the dependent variable decreased with higher league table position.

Accordingly, variables with positive coefficients had negative effects on team performance and vice versa. Unlike Szymanski's result, this means that a higher usage of non-white players was not associated with better team performance. This suggests that teams recruiting an above average number of non-white players tended to systematically underperform in the league. If the player market was efficient (Szymanski, 2000; Pedace, 2008), then we may also imply that non-white players could earn higher salaries relative to other players with equal ability. However, considering the coefficient magnitude in addition to its large standard error, it is unlikely that team performance would be worse off by recruiting one additional non-white player.

Simultaneously, there is only weak evidence that non-white players were favoured in their salaries. They may belong to different nationalities and play on various team levels. The average squad size consisted of 27 players, of whom 9 players on average were non-white. There was no statistical difference either in the means of non-white players between top teams and lower tier teams, or in the means of league position between teams who recruited above average numbers of non-white players and the ones who recruited below average.[11] Accordingly, to disentangle a discrimination effect for non-white players is not a simple task.

The wage variable *logWAGE* was highly significantly negative at the 1% level, as expected. Its economic impact was huge, with a coefficient of -1.60 relative to -0.53 in Szymanski (2000). This supports the notion that the player market is competitive. The squad size variable *N_Player* was also highly significant in the expected positive direction. Szymanski indicated that it worked as a proxy for a high number of injuries and also for a lack of players' chemistry within teams.

The dependent variable in model (6) is identical to model (5).

[11] The tests for two means differences are available upon request.

[insert Table 12 here]

The wage variable *logWAGE* was significantly negative, as expected, at the 5% level with a coefficient of 1.40. This was the only significant predictor in the specification. In addition to the fixed effects model that we used, we also ran the regressions using OLS and random effects approaches, and it turned out that none of the nationality variables were significant. The $R^2$ is quite reasonable with 36.6%, but including the manager change and nationality variables did not add much explanatory power to model (5). The results suggest that there was no evidence for wage discrimination that was associated with the nationality composition of teams' squads. This is contradictory to the findings of Pedace (2008). His fixed effect estimation with respect to the Premier League did not produce significant salaries and manager change coefficients, but it did produce significant coefficients for the squad size (aggregate number of team players) and, more importantly, the South American players. These findings are interpreted as evidence for salary discrimination in favour of that subgroup, but this effect is rationalised by fans' appeal to watch Latino players.

### 3.5 Conclusion

This paper introduces a new discriminatory variable associated with player playing time in the English Premier League, which we call the nationality effect. The basic intuitive idea is that managers favour players who share a common cultural background with them, be it nationality, language, or even a particular playing style. Additionally, clubs' sources (e.g. managers, agents, scouts) who are responsible for searching and recruiting players may be embedded in social networks, and hence seek players who resemble the local standards, even if equal or even better foreign talent is available in the market. Therefore, we conjectured that players who share a common cultural background with their manager would be allocated extra playing time per season relative to those who do not, and to an extent that cannot be attributed to mere performance. We also conjectured that British players in particular would be allocated extra playing time, considering the embedded social networks of clubs' managerial power and the resemblance of a British player to the preferred local professional football standards.

The nationality effect was supported by our empirical findings. British players playing under a British manager got an extra 100 minutes of playing time per season. Additionally, even if the manager's nationality was not considered in the estimation, British players got more playing time than European, South American, and African players by 102, 170, and 196 minutes per season, respectively. Our tests showed that the nationality effect prevailed even when we included player transfer fees and playing positions, which are indirectly related to playing time.

Lastly, our empirical findings showed no evidence for wage discrimination either against non-white players or players from different nationalities. These findings are inconsistent with those of Szymanski (2000) and Pedace (2008), and they suggest a shift of the discrimination effect towards the individual player level instead of the established club level analysis. In other words, there is evidence for discrimination in player playing time rather than in player salary. However, managers' decisions could be embedded within social networks (e.g. scouts, player agents) that enforce the resemblance of home standards for professional footballers (McGovern, 2002).

### 3.6 Tables

#### 3.6.1 Table 1 Player playing time according to nationality differences

| | *N* | PT | SD | Min | Max |
|---|---|---|---|---|---|
| *Panel A* | | | | | |
| *Player nationality* | | | | | |
| British | 834 | 1827 | 866 | 110 | 3420 |
| Non-British | 975 | 1749 | 845 | 74 | 3420 |
| *Panel B* | | | | | |
| *NE scenario* | | | | | |
| BB | 520 | 1834 | 865 | 110 | 3420 |
| BN | 427 | 1700 | 846 | 74 | 3420 |
| NB | 314 | 1815 | 867 | 190 | 3420 |
| NN | 548 | 1787 | 843 | 157 | 3420 |

*Notes*: Panel A presents the player playing time average (PT) among British and non-British players, and this is in total minutes played per season. SD is the standard deviation and min and max represent the minimum and maximum, while N is the number of observations. Panel B presents the same averages as Panel A but with the consideration of the manager's nationality under whom the player is playing. NE stands for nationality effect, and the scenarios are as follows: i) BB when both the player and the manager are British, ii) BN when the manager is British and the player is not, iii) NB when the manager is non-British and the player is British, and iv) NN when both the player and the manager are non-British.

**3.6.2 Table 2 Means of nationality groups in teams among four position categories**

| Nationality of player | | Britain | | Europe | | South America | | Africa | | USA & Asia | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Team position | Obs (N) | Mean | Median | Mean | Median | Mean | Median | Mean | Median | Mean | Median |
| **1st Five** | **457** | 6 | 6 | 8 | 9 | 2.7 | 3 | 1.3 | 2 | .13 | 0 |
| **2nd Five** | **455** | 8.3 | 8 | 6 | 6 | 1.5 | 1 | 1.7 | 2 | .49 | 0 |
| **3rd Five** | **450** | 9.2 | 10 | 4.3 | 3 | 1 | 1 | 2.5 | 2 | .69 | 1 |
| **4th Five** | **447** | 9.6 | 10 | 4.3 | 4 | 1.5 | 2 | 2.2 | 2 | .38 | 0 |

*Notes*: 1st Five and the following tranches represent five positions in the league table sorted into an ascending order. Each tranche shows the average and the median number of players who belong to a particular nationality. For instance, a top team that is ranked in the top five of the league table recruits on average 6 British players, whereas a potentially relegated team that is in the bottom of the league table recruits about 10 British players.

### 3.6.3 Table 3 British managers within different clubs' turnover and positions

| | Clubs' Turnover (T.O) (£million) | | | |
|---|---|---|---|---|
| | >121 *(More than)* | | <121 *(Less than)* | |
| British manager Obsv.(N) | 233 | | 714 | |
| | Position in the league table (1–10 and 11–20) | | | |
| | *if T.O >121* | | *if T.O <121* | |
| | 1st half | 2nd half | 1st half | 2nd half |
| | 163 | 70 | 159 | 555 |

*Notes*: There are 947 observations in the dataset (out of 1809) where the team manager was British. We first split teams according to their clubs' turnover and the figure £121 million is their median. In the second part, the observations were sorted based on team positions in the league table subject to their clubs' turnover. For instance, teams with a turnover of more than £121m and ranked in the top 10 position included 163 observations where the manager of the team was British. Teams with a turnover of less than £121m and ranked in the top 10 position included 159 observations where the manager of the team was British.

**3.6.4 Table 4 Performance, playing time, and transfer fees comparison between different nationality players**

| **Full sample** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Nationality* | Britain | | Europe | | Africa | | S. America | |
| *Performance Metric* | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Goals | .121 | .163 | .141 | .177 | .166 | .191 | .158 | .218 |
| Assists | .106 | .117 | .118 | .142 | .120 | .116 | .133 | .142 |
| Passes | 39.2 | 13.1 | 43.3 | 14.3 | 36.2 | 15.1 | 43.3 | 12.4 |
| Crosses | 2.82 | 2.33 | 2.2 | 2.2 | 1.83 | 1.98 | 2.08 | 2.14 |
| Interceptions | 1.42 | .753 | 1.50 | .93 | 1.22 | .887 | 1.62 | .928 |
| PT (mins) | 1827 | 865 | 1812 | 847 | 1600 | 849 | 1715 | 831 |
| TF (£m) | 3.82 | 5.8 | 9.12 | 10.4 | 5.6 | 6.8 | 11.4 | 11.3 |
| TF Median (£m) | 1.65 | | 6 | | 3.5 | | 8 | |
| Obsv (N) | 834 | | 561 | | 196 | | 171 | |

*Notes*: The comparison is presented for the full sample and the top 10 ranked teams. This is to reflect performance among different clubs' characteristics. The 5 performance metrics are calculated using the total number of each metric per player per season divided by the number of games a player had, then multiplied by 90 to reflect a full game measurement. PT stands for playing time and is measured by the total minutes played per season. TF stands for transfer fees, and these are averages for the African and European players. SD stands for standard deviation.

**3.6.5 Table 5 Correlation matrix for player performance metrics**

| | Goals | Assists | Passes | Crosses | Tackles | Interceptions |
|---|---|---|---|---|---|---|
| Goals | 1 | | | | | |
| Assists | 0.320*** | 1 | | | | |
| Passes | -0.214*** | 0.0337 | 1 | | | |
| Crosses | 0.00707 | 0.435*** | -0.0391 | 1 | | |
| Tackles | -0.387*** | -0.137*** | 0.385*** | 0.0179 | 1 | |
| Interceptions | -0.508*** | -0.345*** | 0.341*** | -0.179*** | 0.559*** | 1 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

### 3.6.6 Table 6 Playing time according to nationality effect

| | (1)<br>BB | (2)<br>BN | (3)<br>NB | (4)<br>NN |
|---|---|---|---|---|
| BB | **107.29**** | **-** | **-** | **-** |
| | **(54.56)** | | | |
| BN | **-** | **-94.98*** | **-** | **-** |
| | | **(55.90)** | | |
| NB | **-** | **-** | **42.42** | **-** |
| | | | **(59.50)** | |
| NN | - | - | - | **-39.39** |
| | | | | **(60.65)** |
| Goals | 668.65*** | 657.62*** | 651.03*** | 653.65*** |
| | (135.83) | (135.72) | (135.89) | (135.97) |
| Passes | 7.54*** | 7.32*** | 7.57*** | 7.66*** |
| | (2.02) | (2.02) | (2.02) | (2.02) |
| Crosses | -4.75 | -4.26 | -3.61 | -3.78 |
| | (10.09) | (10.08) | (10.10) | (10.09) |
| Interceptions | 123.18*** | 122.63*** | 120.19*** | 119.97*** |
| | (29.66) | (29.63) | (29.66) | (29.65) |
| MC | -110.54 | -47.61 | -95.63 | -80.16 |
| | (86.48) | (47.09) | (89.09) | (90.14) |
| Suspension | 35.15* | 34.61* | 33.25* | 33.56* |
| | (18.69) | (18.69) | (18.68) | (18.69) |
| Injury | -54.24*** | -54.45*** | -54.48*** | -54.44*** |
| | (3.65) | (3.65) | (3.65) | (3.65) |
| On-Loan | -61.56*** | -60.79*** | -61.05*** | -61.23*** |
| | (10.05) | (10.06) | (10.06) | (10.07) |
| _cons | 1342.70*** | 1359.66*** | 1335.99*** | 1366.38*** |
| | (161.14) | (161.47) | (162.02) | (164.30) |
| N | 1809 | 1809 | 1809 | 1809 |
| $R^2$ | 0.199 | 0.199 | 0.198 | 0.198 |

*Notes*: This table reports the nationality effect on player playing time, which is represented by total playing minutes per season. We tested the nationality effect through four scenarios in columns 1 to 4, where (i) BB stands for British player and a British manager, (ii) BN stands for British manager and a Non-British player, (iii) NB stands for non-British manager and a British player, and (iv) NN stands for non-British manager and a non-British player. Each one of these is a dummy that takes a value of 1 or 0 considering the pair nationalities. Goals, passes, crosses, and interceptions are individual player performance metrics. These were calculated using the aggregate number of each metric per season divided by the total played minutes per season, and the resulting number was multiplied by 90 to reflect a full game minutes measurement. MC stands for managerial change, and this is an interaction dummy where its first part takes a value of 1 if a new manager was appointed to the team after the start of a season (i.e. after August) and 0 otherwise, and its second part takes a value of 1 if there was a change in the manager nationality (i.e. either from British to non-British or vice versa). Suspension, injury, and on-loan are control variables and are measured by the total number of games a player could not make per season due to these circumstances. Clubs and seasons fixed-effects coefficients are not reported. Standard errors are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

**3.6.7 Table 7 Playing time according to nationality effect (transfer fee and player position variables included)**

| | (1)<br>BB | (2)<br>BN | (3)<br>NB | (4)<br>NN |
|---|---|---|---|---|
| **BB** | **99.68***<br>**(52.56)** | **-** | **-** | **-** |
| **BN** | **-** | **-69.09**<br>**(53.93)** | **-** | **-** |
| **NB** | **-** | **-** | **77.03**<br>**(57.91)** | **-** |
| **NN** | **-** | **-** | **-** | **-95.88**<br>**(59.00)** |
| Goals | 1058.12***<br>(145.26) | 1054.38***<br>(145.40) | 1045.53***<br>(145.47) | 1050.77***<br>(145.34) |
| Passes | 8.73***<br>(2.11) | 8.58***<br>(2.12) | 8.77***<br>(2.11) | 9.01***<br>(2.11) |
| Crosses | -7.67<br>(10.22) | -7.28<br>(10.23) | -6.44<br>(10.22) | -6.58<br>(10.21) |
| Interceptions | 36.70<br>(31.24) | 35.86<br>(31.26) | 34.47<br>(31.21) | 34.47<br>(31.18) |
| MC | -87.92<br>(84.92) | -58.62<br>(84.67) | -98.96<br>(87.67) | -64.69<br>(88.66) |
| TF | 22.62***<br>(3.15) | 22.41***<br>(3.15) | 22.88***<br>(3.18) | 23.16***<br>(3.19) |
| Striker | -284.72***<br>(82.31) | -288.80***<br>(82.34) | -285.16***<br>(82.36) | -282.31***<br>(82.38) |
| Defender | 406.69***<br>(59.25) | 406.16***<br>(59.37) | 410.97***<br>(59.20) | 413.82***<br>(59.18) |
| Suspended | 36.13**<br>(18.32) | 35.13*<br>(18.32) | 34.52*<br>(18.32) | 35.15*<br>(18.32) |
| Injured | -54.44***<br>(3.56) | -54.63***<br>(3.56) | -54.68***<br>(3.56) | -54.64***<br>(3.56) |
| On Loan | -57.03***<br>(9.81) | -56.74***<br>(9.81) | -56.61***<br>(9.81) | -57.06***<br>(9.82) |
| _cons | 1073.32***<br>(175.13) | 1098.48***<br>(175.28) | 1053.15***<br>(176.26) | 1114.16***<br>(176.10) |
| N | 1809 | 1809 | 1809 | 1809 |
| $R^2$ | 0.21 | 0.21 | 0.212 | 0.212 |

*Notes*: The table presents the regression results of model (2). This is a modification to model (1) where two variables are added: player transfer fee (TF) and playing position (striker and defender). The TF is calculated as the value of the transfer in £million. The playing positions are dummy variables that take a value of 1 for each respective position to the player. The notes on the rest of variables in Table 6 apply here. Standard errors are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

### 3.6.8 Table 8 Playing time according to nationality effect and team ranks

| | (1) 1st 10 teams | (2) 2nd 10 teams |
|---|---|---|
| **BB** | 174.07** | 52.15 |
| | (83.11) | (65.96) |
| Goals | 1051.74*** | 1080.27*** |
| | (184.08) | (253.42) |
| Passes | 10.77*** | 6.19* |
| | (2.77) | (3.34) |
| Crosses | -5.31 | -11.41 |
| | (14.57) | (13.71) |
| Interceptions | 19.44 | 74.54 |
| | (44.04) | (45.83) |
| MC | -110.71 | -62.70 |
| | (182.39) | (126.58) |
| TF | 23.20*** | 19.43** |
| | (3.57) | (7.87) |
| Striker | -316.16*** | -243.76** |
| | (118.20) | (109.07) |
| Defender | 470.10*** | 339.63*** |
| | (87.01) | (75.63) |
| Suspended | 34.90 | 52.05* |
| | (23.95) | (29.33) |
| Injured | -55.50*** | -53.80*** |
| | (4.81) | (5.47) |
| On-Loan | -53.82*** | -61.74*** |
| | (15.43) | (13.43) |
| _cons | 939.37*** | 1524.95*** |
| | (221.91) | (353.78) |
| N | 912 | 897 |
| $R^2$ | 0.234 | 0.234 |

*Notes*: The results in the table are based on two subsamples: (i) the top 10 positions ranked teams, and (ii) the lower 10 ranked teams (the second half of the table). Only the first scenario of model (1) is presented, as the rest were insignificant in the subsequent model (2). Column 1 represents the first subsample, and column 2 represents the second subsample. The definitions of all variables presented in Tables 6 and 7 apply here. Clubs and seasons fixed-effects coefficients are not reported. Standard errors are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

### 3.6.9 Table 9 Nationality effect on playing time using player nationality subgroups

| | (1) European | (2) Britain | (3) S.America | (4) America.Asia | (5) Africa |
|---|---|---|---|---|---|
| Europe | - | -48.51 | 68.26 | 123.84 | 214.94*** |
| | | (58.56) | (87.28) | (153.21) | (83.26) |
| Britain | 56.83 | - | 120.89 | 175.62 | 266.99*** |
| | (58.52) | | (86.92) | (150.57) | (79.77) |
| S. America | -50.47 | -103.13 | - | 69.26 | 160.28 |
| | (87.62) | (87.32) | | (166.43) | (105.29) |
| America-Asia | -69.93 | -123.52 | -5.86 | - | 139.99 |
| | (155.47) | (152.91) | (168.24) | | (164.11) |
| Africa | -199.90** | -253.20*** | -135.87 | -80.98 | - |
| | (83.53) | (80.08) | (105.21) | (162.24) | |
| Goals | 670.06*** | 670.06*** | 670.71*** | 668.83*** | 669.12*** |
| | (135.69) | (135.71) | (135.65) | (135.68) | (135.64) |
| Passes | 7.41*** | 7.41*** | 7.39*** | 7.38*** | 7.38*** |
| | (2.01) | (2.01) | (2.01) | (2.01) | (2.01) |
| Crosses | -5.74 | -5.74 | -5.77 | -5.96 | -5.89 |
| | (10.08) | (10.08) | (10.08) | (10.08) | (10.08) |
| Interceptions | 119.03*** | 118.92*** | 119.37*** | 119.12*** | 118.63*** |
| | (29.71) | (29.71) | (29.70) | (29.70) | (29.70) |
| MC | -48.51 | -48.66 | -47.80 | -48.27 | -48.25 |
| | (47.06) | (47.06) | (47.05) | (47.06) | (47.05) |
| Suspended | 38.02** | 37.99** | 38.12** | 38.06** | 38.26** |
| | (18.72) | (18.72) | (18.72) | (18.71) | (18.71) |
| Injured | -54.69*** | -54.69*** | -54.68*** | -54.70*** | -54.71*** |
| | (3.64) | (3.64) | (3.64) | (3.64) | (3.64) |
| OnLoan | -61.29*** | -61.28*** | -61.29*** | -61.35*** | -61.29*** |
| | (10.04) | (10.04) | (10.04) | (10.04) | (10.04) |
| _cons | 1360.46*** | 1412.20*** | 1295.25*** | 1241.60*** | 1151.34*** |
| | (163.16) | (164.40) | (179.03) | (215.47) | (173.18) |
| N | 1809 | 1809 | 1809 | 1809 | 1809 |
| $R^2$ | 0.198 | 0.198 | 0.198 | 0.198 | 0.198 |

*Notes:* The dependant variable is the total played minutes for each player per season. The first five variables are the nationality subgroups where Europe stands for players from Europe, Britain for players from Great Britain (England, Wales, Scotland), Ireland and North Ireland, S. America for players from South or Latin America, and America-Asia for players from North America and Asia. These are dummy variables where each player takes a value of 1 or 0, according to his relevant nationality subgroup. In columns 1 to 5 (i.e. in each regression), one nationality subgroup dummy is withdrawn. This is to interpret the other nationality coefficients relative to it. The definitions of all other performance and control variables are identical to the ones in Tables 6, 7, and 8. The MC is a dummy that takes a value of 1 if a new manager was appointed to the team after the start of the season. Clubs and seasons fixed-effects coefficients are not reported. Standard errors are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

### 3.6.10 Table 10 Nationality effect on playing time using player nationality subgroups (TF and PP included)

| | (1)<br>Europe | (2)<br>Britain | (3)<br>S. America | (4)<br>America-Asia | (5)<br>Africa |
|---|---|---|---|---|---|
| Europe | - | -96.49* | 70.15 | 59.81 | 97.51 |
| | | (55.89) | (82.10) | (144.31) | (79.16) |
| Britain | 102.60* | - | 169.93** | 158.77 | 196.77*** |
| | (55.87) | | (82.82) | (141.61) | (76.10) |
| S. America | -58.01 | -157.28* | - | -1.01 | 36.58 |
| | (82.39) | (83.15) | | (156.84) | (99.66) |
| America-Asia | -21.45 | -121.65 | 45.84 | - | 72.55 |
| | (146.37) | (143.70) | (158.53) | | (154.68) |
| Africa | -86.37 | -186.18** | -19.21 | -30.10 | - |
| | (79.33) | (76.29) | (99.51) | (152.80) | |
| Goals | 1041.20*** | 1042.73*** | 1040.89*** | 1040.60*** | 1039.60*** |
| | (145.33) | (145.33) | (145.30) | (145.33) | (145.34) |
| Passes | 8.65*** | 8.65*** | 8.64*** | 8.62*** | 8.64*** |
| | (2.11) | (2.11) | (2.11) | (2.11) | (2.11) |
| Crosses | -7.64 | -7.66 | -7.65 | -7.83 | -7.72 |
| | (10.23) | (10.23) | (10.23) | (10.23) | (10.23) |
| Interceptions | 41.63 | 41.36 | 41.98 | 41.70 | 41.61 |
| | (31.35) | (31.36) | (31.35) | (31.35) | (31.35) |
| MC | -32.05 | -32.34 | -31.78 | -31.94 | -31.73 |
| | (46.33) | (46.33) | (46.33) | (46.32) | (46.33) |
| TF | 23.51*** | 23.46*** | 23.53*** | 23.48*** | 23.48*** |
| | (3.20) | (3.20) | (3.20) | (3.20) | (3.20) |
| Striker | -266.47*** | -267.85*** | -265.97*** | -267.25*** | -265.42*** |
| | (82.62) | (82.61) | (82.61) | (82.57) | (82.62) |
| Defender | 399.22*** | 399.25*** | 398.87*** | 398.46*** | 398.23*** |
| | (59.40) | (59.42) | (59.39) | (59.41) | (59.40) |
| Suspended | 38.15** | 38.15** | 38.21** | 38.18** | 38.33** |
| | (18.38) | (18.38) | (18.38) | (18.37) | (18.37) |
| Injured | -54.70*** | -54.69*** | -54.69*** | -54.70*** | -54.71*** |
| | (3.56) | (3.56) | (3.56) | (3.56) | (3.56) |
| On Loan | -56.78*** | -56.77*** | -56.78*** | -56.83*** | -56.79*** |
| | (9.81) | (9.81) | (9.81) | (9.81) | (9.81) |
| _cons | 1360.46*** | 1412.20*** | 1295.25*** | 1241.60*** | 1151.34*** |
| | (163.16) | (164.40) | (179.03) | (215.47) | (173.18) |
| N | 1809 | 1809 | 1809 | 1809 | 1809 |
| $R^2$ | 0.198 | 0.198 | 0.198 | 0.198 | 0.198 |

*Notes:* This is a modification to the model in Table 9 where three variables are added: TF, which stands for player transfer fees; and striker and defender, which are dummies for player playing position. Clubs and seasons fixed-effects coefficients are not reported. Standard errors are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

**3.6.11 Table 11 The effect of non-white players on team performance: Seasons 2009–2016**

| | (1)<br>Log Position<br>b/se |
|---|---|
| NON-WHITE | 0.0608* |
| | (0.0335) |
| logWAGE | -1.5979*** |
| | (0.5565) |
| N_PLAYER | 0.1571*** |
| | (0.0292) |
| _cons | 0.1241 |
| | (0.16) |
| N | 157 |
| $R^2$ | 0.3581 |

*Notes*: The regression is conducted by means of a fixed effect model. The dependant variable is the team performance represented by the log of its position on the league rank. NON-WHITE is the number of non-white players in the team squad. LogWAGE is the log of wages for club $_i$ in season $_t$. N_PLAYER is the total number of players in the team squad. Standard errors are reported in parentheses. Season dummies are not reported. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

**3.6.12 Table 12 The effect of nationality groups on team performance: Seasons 2009–2016**

| | (1)<br>Log Position<br>b/se |
|---|---|
| logWAGE | -1.402** |
| | (0.60) |
| N_PLAYERS | -0.288 |
| | (0.35) |
| MC | 0.031 |
| | (0.15) |
| BRITISH | 0.444 |
| | (0.35) |
| EUR | 0.482 |
| | (0.35) |
| SAMERICAN | 0.529 |
| | (0.35) |
| AFRICAN | 0.497 |
| | (0.35) |
| USASIA | 0.454 |
| | (0.36) |
| _cons | 0.476 |
| | (2.49) |
| N | 157 |
| R2 | 0.366 |

*Notes*: The regression is conducted by means of a fixed effect model. The dependant variable is the team performance represented by the log of its position on the league rank. LogWAGE is the log of wages for club $_i$ in season $_t$. N_PLAYER is the total number of players in the team squad. MC stands for managerial change and takes a value of 1 if the manager was new to team $_i$ in season $_t$ and 0 otherwise. BRITISH is the number of British players in a team squad who appeared in at least one league match. EUR is the number of European players in a team squad who appeared in at least one league match. SAMERICAN is the number of South American players in a team squad who appeared in at least one league match. AFRICAN is the number of African players in a team squad who appeared in at least one league match. USASIA is the number of American and Asian players in a team squad who appeared in at least one league match. Standard errors are reported in parentheses. Season dummies are not reported. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

### 3.6.13 Table 13 Definitions of performance metrics for the *Perf*$_{it}$ variable and *Control* variables

| Performance Metric | Definition |
|---|---|
| GOALS | Number of goals scored per 90 minutes. |
| CROSSES | Number of crosses performed per 90 minutes. It is a medium-long range ball passing towards the opponent's area. |
| PASSES | Number of passes performed per 90 minutes. All ball passes. |
| INTERCEPTIONS | Number of interceptions performed per 90 minutes. When the ball is intended for a player of the same team but caught by a player of the opposing team. |
| **Control Variables** | **Definition** |
| Suspension | Number of games that a player is legally suspended. |
| On loan | Number of games not played while being on loan to another team. |
| Injury | Number of matches that a player could not play due to an injury. |
| MC | A managerial change that takes place after the start of the season. |

*The relevant data are taken from Premierleague.com, the official website that represents the PL authority. We opted to adjust the variables per 90 minutes rather than per match appearance. This gives a proper reflection of the playing metrics and players' productivity per full game minute.

# 4. Can home field advantage predict team success?

Home field (HF) advantage in competitive sports has garnered much empirical attention. This advantage pertains to a higher probability of winning games at a team's home ground as opposed to those at away grounds (e.g. Dobson & Goddard, 2009). Despite the empirical difficulty in identifying with precision the sources of home field advantage, five main factors are discussed in the literature that may drive that advantage. These comprise fans (or crowd encouragement), players' familiarity with stadium facilities, travel fatigue of opponents, rules and probable biased referees' decisions under pressures favouring the home team, and lastly, the territoriality aspect (Carron et al., 2005; Courneya & Carron, 1992; Nevill & Holder, 1999; Pollard, 2006a).

Human physical conditioning naturally plays a major role in sports, with its determinants such as training programmes, nutrition, and genetics. However, athletes, coaches, and referees are also likely to be influenced by the psychological factors derived from HF advantage. The five main factors provide a basis for explaining the behavioural states of athletes while playing at home. For instance, the evidence on HF advantage did not show a strong long-term trend in MLB and NFL, and it was consistently lower in baseball (MLB) than in other sports. In English football, besides NBA and NHL (hockey), the trend seems to have declined in recent decades (Dobson & Goddard, 2009). One of the explanations for such differences among sport disciplines relates to the crowd size effect, where some sports facilities have lower capacity for crowds, especially indoor ones such as ice-hockey or basketball. The crowd effect, whether its influence is on players' encouragement or referee decisions, can be related to crowd density (e.g. Agnew & Carron, 1994).

This paper contributes to the HF literature in two respects. The effect continues as the rate of home wins regularly exceeds away wins. However, is this effect statistically and economically significant? The paper's first contribution is that it employs a unique sample of all PL games played during the 2012/13 to 2016/17 seasons to test for the continued relevance of the HF advantage. The total number of games played for each season is 380, and this yields a total of 1900 matches. However, our full sample contains 3,800 matches, as a home match for team A simultaneously constitutes an away match for team B. Our panel regression results provide unequivocal evidence of a significantly strong and robust home field effect. The HF effect over the full sample period is a resounding 0.46 additional points per home game, and it is significant at the 1% level. It holds with almost identical effects for both top six teams and remaining teams in the league.

The paper's second contribution is that it uses bootstrap techniques to explore the HF effect in predicting team success and failure in the PL. More specifically, our bootstrap approach to modelling match results follows that of Bell et al. (2013). This bootstrap approach is applied to the first half of the season's games (20) to predict the success and failure of teams at the end of the season. The implication of our results is that teams can evaluate their probability for promotion to the Top four to six rankings, or for relegation on the basis of the home results of the first 20 games each season.

Our empirical findings are novel in the context of the Premier League. They suggest that teams who outperform the bootstrap expected results in their early home (relative to away) games will be secure in the PL. This is still the case even if they perform poorly in away games over this period. This is a strong implication that home results dominate away results. Moreover, a great home performance, even if accompanied by poor away performance, can lead to a top four ranking. An average team performance at home that is associated with poor performance away from home increases the probability of relegation to the Championship by about 40%. Lastly, an average performance in home games that is accompanied by great away performance may not suffice to secure a top four ranking. This is additional evidence in favour of the home field effect in the context of predicting promotion in elite European football competitions.

## 4.1 Literature Review

It is established in the literature that the home field effect favours home teams and creates difficulties for away teams in professional sports. Evidence shows that winning home games is significant to the extent that it cannot be explained by mere chance. In Courneya and Carron's (1992) review of the literature, they identified that consistently winning over half of the games played at home represented a home advantage. Pollard (1986) defined home advantage as a higher percentage of points gained while playing on the home field. Bray (1999) defined home advantage as a difference of greater than 5% between home and away winning percentages. This is particularly related to sport competitions where the schedule for home and away games is balanced, i.e. each team plays an identical number of games in its home field and in away field within a season.

Studies related to home advantage were mainly established in the United States, covering many professional sports such as basketball, baseball, ice hockey, and American football. Historical evidence suggests a clear indication of an HF advantage despite variations in its magnitude between different sports, and some declining patterns were also identified (Dobson & Goddard, 2009). Courneya and Carron (1992) showed home win percentages for different sports as follows: baseball 53.5%, American football 57.3%, ice hockey 61.1%, basketball 64.4%, and football (soccer in American terms) 69%. Jamieson's (2010) extended analysis of 260,000 sports games showed 60.4% win percentage if games were played on the home field.

Social psychology provides explanations for the causes of HF advantage (e.g. Edwards & Archambault, 1989; Schwartz & Barsky, 1977). These fall within five main categories: (i) familiarity with stadium facilities, such as physical features of the playing ground (pitch) and stands; (ii) travel fatigue that may be experienced by the visiting team and its consequential effect on match preparation routines; (iii) rules factors through some privileges given to the hosting team, such as the last bat in baseball; and (iv) the crowds' psychological effect on hosting and visiting teams, where support is extended to home team players and intimidation towards visiting teams. Crowd pressure also may have an influence on referees' decisions where hosting teams could gain, for instance, soft fouls or penalties (at least before the introduction of the VAR). Finally, (v) territoriality, where certain geographic area norms arise and are reflected in home games (Catalonia norms, for instance). Generally, there is evidence for minimal effect with regard to travel fatigue (e.g. Ramchandani & Wilson, 2010), the familiarity aspect (e.g. Pollard, 2002), and to a lesser extent, rule factors (e.g. Courneya & Carron, 1990). Crowd effects have received more attention, but results vary among different sports, with no clear association established between crowd support and home advantage. Indoor sports such as basketball and ice-hockey, which display closer interrelation among players and the crowds, have shown clearer evidence for the positive effect of crowd support. This is due to the fact that indoor facilities are more uniform (e.g. size and quality of the playing surface, hall temperature, and atmosphere) than outdoor facilities.

British football also grasped attention (e.g. Clarke and Noman, 1995; Nevill et al., 1996; Pollard, 1986; Pollard & Pollard, 2005). Pollard and Pollard showed that since the league formation in 1888, there was only a small change in home advantage as it varied among the four divisions, from 63% to 65.5% over the seasons 1970–1981. They suggested that familiarity with home field conditions was more influential than crowd support and travel fatigue experienced by opponent teams. Their finding contrasts with that of Nevill et al. (1996), who suggested that home advantage varied significantly among eight major divisions of the English and Scottish football leagues. Average attendance, unlike some other sports disciplines, in each division was influential in those home advantage differences. Nevill et al. (2002) implied that crowd pressure provoked more aggressive behaviour from away team players as well as influencing referees' decisions. Downward and Jones' (2007) findings support the crowd size effect, in particular on referees' decisions.

With regard to home advantage historic patterns, even though the home win percentage in English football decreased from around 50% in the 70s and 80s to around 43% later in the new millennia, it remained well above the away win ratio when it reached its highest percentages of around 30% later in the decade (Dobson & Goddard, 2009). That declining trend might be attributable to the introduction of the Premier League beginning in 1992, where clubs gained more financial power, through which grabbing talents was more viable and hence competition increased. Allen and Jones (2014) analysed PL results in the seasons 1992/93–2011/12 and showed that winning home games, on average, secured 60.7% of total points earned by teams.

A thorough documentation and interpretation of HF advantage has been mainly directed towards the above-mentioned psychological aspects, while empirical research with respect to the effect on field performance is relatively limited. One explanation for this is the complicated nature of precisely quantifying factors that are closely related to match outcomes, such as individual players' skills and quality, teams' principles and standards, managerial and coaching qualities, and match significance within the league fixtures. Therefore, another strand of research has focused on the manifestation of home advantage within match outcomes. For instance, one can assume—ceteris paribus—that home teams may adapt a more offensive tactical approach for the match, while away teams may follow a more defensive approach. Moreover, a frustration hypothesis suggests that away team players may be more inclined to make fouls or violate rules (e.g. Volkamer, 1971). The findings of Edwards and Archambault (1989) suggest that the HF effect manifests itself in a non-outcome (non-result) aspect as well, where home teams tend to outperform away (visiting) teams.

Carmichael and Thomas (2005) extended that strand of literature through a more explicit empirical approach, where they examined match-play data from the EPL in the 1997/98 season. They estimated match-based production functions for home and away team performances. Their results indicated that adapting an attacking play style at home could produce more goals and consequently better results, while adapting a defensive style at away games was more important for achieving better results.

Team tactical preparation and starting line-up may vary based on whether a match is played at home or away. Carmichael et al. (2000) employed player performance statistics to test for match-based production functions of the FA cup 1997/98 season. They used goal differential as a proxy to measure team performance and incorporated a home field advantage variable, which had a significant positive effect on match results for the observed team. Crowder et al. (2002) focused on the probabilities of home win, draw, or away win, as their outcomes constituted the betting market, to estimate English teams' attacking and defending capabilities through a refined Poisson model. Clarke and Norman (1995) used English football match results to produce a home advantage effect in addition to a team rating. They investigated the reasons behind some home advantage differences among teams between the 1981/82 to 1990/91 seasons. They showed that the distance between club grounds (fields) was linearly related to home advantage. Dixon and Robinson (1998) used home advantage as a part of their soccer match model parameters to eventually assess scoring rate variations of the home and away teams during a match. Barnett and Hilditch (1993) assessed the home advantage of a few English teams during the 1980s and 1990s seasons by investigating the impact of adopting artificial grass on home teams' match performance and outcome.

## 4.2 Research Hypotheses

In the home field advantage literature, a positive association is found between playing at home and a better behavioural state of athletes prior to games. Most studies discuss psychological factors that may drive such a state. The findings differ across a wide range of sports with respect to the longevity of the HF advantage and the intensity of its effect, as well as its probable drivers. The HF advantage effect seems to have declined over the years and is at its highest levels in the early years of a particular sport (Dobson & Goddard, 2009). Nevertheless, the effect has remained relatively robust.

Some studies have focused on the advantage effect by country (e.g. Pollard, 2006) while some have focused on two-leg knockout ties played in European football competitions (e.g. Eugster et al., 2011). Bray et al. (2003) provided evidence in English football that higher division teams had greater home advantage than lower division teams. Familiarity with stadium facilities, including using artificial grass, size of attendance and effect on referees, travel to the home stadium, and territoriality are all psychological factors that have been studied within English football.

Statistics from the Sky Sports Football Yearbook show the mean proportion of home wins in the EPL across the 1970–2009 seasons was near the 50% threshold in the 1970s and 1980s, and declined to a low of 43% in 2009. Away wins, on the other hand, increased from around the 20% level in the 1970s and 1980s to reach around 30% from the beginning of the new millennium.

Another strand of the literature provides evidence on predicting match results and team success. Goddard and Asimakopoulos (2003, 2004) used forecasting models (e.g. Poisson models and ordered probit models) to predict match results and goals scored while focusing on the fixed-odds betting markets. Dobson and Goddard (2003) studied the persistence in sequences of consecutive football match results by means of Monte Carlo analysis, while taking into account the venue (home or away) at which the game was played. Other studies have used artificial neural networks (ANNs) to predict team success. For instance, Arabzad et al. (2014) found that football match results could be predicted through a machine learning approach. Their model predicted five out of the six Iranian football teams promoted to the Asian Football Confederation Champions League (ACL). Lebovic and Sigelman (2001) studied how information processes may affect team rankings in American college football. They showed that rankings exhibited strong inertial tendencies that changed only with the flow of new information about games played in prior weeks. Carmichael and Thomas (2005) provided evidence that a team's tactical approach (or within match performance) varied depending on whether a match was played at home or away. Thus, this could affect how the team performs at home, and their findings support those of Edwards and Archambault (1989), which indicated a tendency for home teams to outperform their visiting (away) teams on a number of non-outcome measures.

#### 4.2.1 Home field effect hypotheses

We combined the above two strands of literature to test for a new home field effect. First, we estimated a model following the approach of Bell et al. (2013) that accounted for several factors that may affect a team match performance. We added an HF variable to their original model, which captured the effect on team performance of playing at home represented by the average earned points. Hence, our first hypothesis is that playing at home can significantly improve team performance and hence earn higher points compared to away games.

H1: Playing at home has a positive effect on team performance after controlling for factors such as player absences, fatigue, and club financial resources.

H1A: Playing at home is independent of team performance after controlling for factors such as player absences, fatigue, and club financial resources.

Second, we estimated the expected points for teams by means of a bootstrap approach, and this provided a benchmark for our analysis. The bootstrap allows for reliable probabilities of match outcomes, and its prediction lies on outcome averages rather than a simple prediction of discrete match outcomes. We then compared the actual points won during the season to these expected benchmark points. Comparing these two can help evaluate whether the team performance at home games was superior to those of away games in predicting team success. To assess the predictive power of home games for the final team rankings in the league, we used bootstrap confidence intervals to generate success and failure signals. The second hypothesis is as follows:

H2: The information from home game results is superior to that from away games in predicting team success and survival in the Premier League.

H2A: Home and away game results have the same predictive power for team success and survival in the Premier League.

## 4.3 Data and Methodology

### 4.3.1 Data

The sample includes all Premier League games (matches) played during the 2012/13 to 2016/17 seasons. The total number of games played for each season was 380, and this yields a total of 1900 matches. However, our sample contains 3,800 observations, as we analysed each game from a team perspective (i.e. one game can generate a loss for one team and a win for another, or a draw for each). The relevant dataset used is the league table rankings at the end of each season, which totals 100 observations (20 team rankings in the PL in each of the five seasons). Table rankings are the outcomes of the 380 league games played.

We used a number of data sources (mainly websites) to build the dataset. These include Transfermarkt.co.uk for all game-related information, match fixtures and results, player injuries, absences and suspensions, extra games played in other competitions, and the net transfer spent by each club in each season. We use Soccerbase.com for teams' managerial records. We obtained club wage bill data from The David Conn Blog in the Guardian Newspaper, which were derived from the published annual reports at Companies House.

To examine the HF advantage in the PL, we tested for its effect on team performance. Match points won were used as a proxy for performance where a win was 3 points, a draw was 1 point, and a loss was 0 points. We followed the approach of Bell et al. (2013) to estimate team performance, but extended their model by including an HF variable. The HF advantage (e.g. Dobson & Goddard, 2009) shows that the percentage of home wins is higher than away wins. The minimum spread between those two percentages is 13%; this is the difference between the home win percentage 43% in 2009 and the corresponding 30% away win percentage in the same year. Thus, we predict that merely playing at home produces better team performance even after controlling for relevant team characteristics such as players' absences and fatigue as well as team financial resources.

The descriptive statistics for home and away team performance are reported in Table 1.

[Insert Table 1 here]

We categorised teams into three groups based on performance: top six, mid-table, and relegation zone teams. The latter consists of four teams, and the ten remaining teams occupied the mid-table slot. We conducted equality of means tests for points earned by teams in these three groups. These show that team performance was better on average when a team played at home than when it played away. Better performance persisted across the sample period and across all team groups. The overall average (i.e. from 2012 to 2016 seasons and for all teams) for team performance was 1.57 points earned at home with a standard deviation of 0.06, compared to 1.14 points earned in away games with a standard deviation of 0.04. The difference of 0.43 between the two averages was statistically significant at the 1% level with a *t*-statistic of 10.9. This confirms that teams earned more points at home than at away venues. The small standard deviations from average points indicate a consistent pattern. These results reflect the first 20 games played (10 home and 10 away games), which represents the first half of the league (38 games are played in total).

The detailed performance for the first 10 home games over all seasons was as follows: The top six teams averaged 2.2 points, mid-table 1.42 points, and relegation zone teams 1.1 points. The correspondence performance for the first 10 away games was as follows: top teams averaged 1.77 points, mid-table 0.99 points, and relegation zone teams 0.66 points. Table 1 shows that the equality of mean tests for all team levels and across all seasons was statistically significant. Hence, it made no difference whether a team was a European competition candidate or in danger of relegation to the lower tier league (i.e. championship league) in exploring the psychological advantage stemming from playing at home.

### 4.3.2 Panel regressions

Several characteristics influenced actual team performance in terms of points. Bell et al. (2013) devised a model that separated a team's manager impact from the effects of other team characteristics. This is estimated through a fixed effect model where the manager of a team is assumed to be an unobserved effect. This allows one to capture managerial skills relative to team resources and conditions. We followed their approach but took a different starting point. We expected that playing at home was associated with a higher probability of better performance and, consequently, a higher win rate. Its effect on team points is crucial. This is a prerequisite for the assessment of its predictive power for team success, which is discussed at a later stage of this model.

The home field effect was estimated by means of a fixed effects panel regression model. The dependent variable was team performance, represented by the team points measure[12]

[12] Points are more direct and clear performance measure than goal difference. Nevertheless, adding goal difference to the estimation is beyond the scope of this chapter.

(where the number of points scored for a win was 3, 1 for a draw and 0 for a loss) for team *i* in fixture *t*.

$$y_{i,t} = \eta_i + \sum_{j=1}^{k} \beta j \; x_{j,i,t} + u_{i,t} \qquad (1)$$

where $\eta_i$ is the team fixed effect, $\boldsymbol{x}$ is a vector of team characteristics, $u_{i,t}$ are zero mean i.i.d. disturbances, $i$=1,…,$N$; $t$=1,…,$T$ so that there are $N$ teams with a total of $T$ matches during the 2012/13–2016/17 seasons (total of 5 seasons). The number of matches in our sample ranged from a minimum of 38, where a team played only one season in the Premier League during our sample period, to a maximum of 190, where a team played the full five sample seasons in the PL. There were 1900 matches over the five-season period that involved a total of 28 teams playing in the PL (20 teams only in each season).

We estimated model (1) using variables that were expected to capture team performance. These include (i) the log of the total wage bill for team players over each season in £m. There is evidence in the literature (e.g. Pedace, 2008) suggesting that player wages explain almost 90% of the variability in team performance. This indicates that clubs hire the best available players subject to the constraints of their financial resources. We opted not to include the net transfer spend variable used by Bell et al. (2013) (which measures the extent to which the club is currently able to hire new players) due to its complicated cumulative effect. Simultaneously, it was insignificant in the model estimation, both in statistical and economic terms. The other variables are (ii) the total number of injured players in match *t*, (iii) the total number of suspended players in match *t*, and (iv) the total number of unavailable players in match *t* for reasons other than suspension or injury (e.g. playing with their national team or the reserve team of their own club), (v) the total number of games that the team played in competitions other than the PL (e.g. EFL or Carabao Cup, Champions or Europa Leagues), and finally, (vi) match *t* venue, which is a dummy variable that takes a value of 1 if the team *i* played at home, and 0 for its away games.

We expected that playing at home and the player wage bill would positively affect team performance. Conversely, we expected player absences to negatively affect it. The impact of non-premier league games on team performance is somewhat tricky to predict. This is because playing more games in other competitions (especially the European Champions League) indicates a higher quality team, while on the other hand, playing too many games may lead to a degree of exhaustion for (some) team players that may adversely affect performance levels. Additionally, the same above-mentioned characteristics of the opposing team for match $t$ are also included in the model as explanatory variables, and their coefficients were expected to have the opposite sign. Each game accordingly is included twice in the dataset (i.e. one regression each for home and away teams/games).

### 4.3.3 Bootstrap approach

Bell et al. (2013) used the bootstrap to analyse the impact of the manager effect. They extracted the manager fixed effect from their original model in order to generate bootstrap replications for team performance that were unaffected by manager skills or interventions. Then they compared the actual points achieved by each manager for their team with the bootstrap generated or expected points. We employed a similar bootstrap procedure to isolate the impact of the HF effect to test for the reliability of the HF effect over the full course of the season. More specifically, we tested whether the results of the first 20 games of the season could help to predict the final league position outcomes in terms of top four or relegation.

We proceeded with estimating model (1). This adds the HF effect to the Bell et al. (2013) model in the estimation. Then, we sampled with replacement from the dependent variable constructed in (1), where for each $j$=1,…,1000, bootstrap replications were run to generate the average number of points scored and their confidence interval for team $i$ at each replication $j$. We opted for 1,000 replications as experimentation with 10,000 replications for a sample of games that produced qualitatively similar results. The resulting sampling distribution represents the expected match points for each team *i* that could be collected relative to team characteristics and unobserved effects based on our model specification.

The bootstrap estimates produced reliable confidence intervals (CI) of the sample distribution using data from the first 20 games of the season only. This allowed for success/failure predictions for the remaining games of the season. We then examined where the actual points achieved by each team fit relative to the 95% CI of the bootstrap distribution. This was examined for home and away games separately. We defined team performance as above average if its actual collected match points lay above the 95% CI boundary, as below if it was below the 5% lower boundary, and as standard if it lay between the upper and lower boundaries.

Based on these definitions, we sought to predict a team's promotion to the top four, survival in the Premier League, or relegation to the Championship League (next lower league division). Securing a place in the top four promotes English teams to the Europa Champions League, which is the most prestigious football tournament for European football clubs. The team in fifth place gains automatic promotion to the Europa League, which is a lower level but well respected European tournament, while the sixth-placed team qualifies for the knock-out competition for the Europa League. The bottom three teams in the league are relegated to the Championship League.

### 4.4 Regression and Bootstrap Results

#### 4.4.1 Regression results

The results from estimating model (1) are reported in Table 2.

[Insert Table 2 here]

Our focus in the results is on the association between team performance after controlling for other relevant factors that may influence the match result and the home field effect. The results show that the HF (venue) variable positively affected team performance with a coefficient of 0.46 and was highly significant at the 1% level. This implies that playing at home increased the expected points per match by 0.46 points on average. This huge economic impact strongly supports hypothesis H1 on the HF effect. The result was robust to teams other than the top six. While this is a novel finding in the context of the Premier League, it is in line with the supportive evidence for the HF advantage in other studies in different sports (Bray, 1999; Jamieson, 2010; Pollard, 1986).

The wage bill also positively affected team performance, as expected, with a coefficient of 0.32 that was significant at the 1% level. This is intuitive, as wages reflect the quality of players hired for the team and subsequently their expected contribution to team performance and success. We also used the wage bill in the estimation in a non-log form and obtained a coefficient of 0.0024 compared to the 0.007 of Bell et al. (2013). This indicates that, for each £100m increase in the home team wage bill (or fall of the same amount in the opposing team wage bill), there was an expected increase of 0.24 in match points. The lower effect of wage bills in our sample did not change when we removed the home variable from the estimation. This may be attributed to the escalating wage bills, especially in the last decade.

The absence of players (represented by injuries, suspensions, and other reasons) from the team squad on the match day was not significant for team performance, although the coefficients had the expected signs. However, injuries were significant at the 10% level in the estimation of a random effect panel regression. Player absences were a significant factor in the Bell et al. (2013) results. We interpreted our results as indicating that the magnitude of the HF effect outweighed it.

Injuries in the opposing team, on the other hand, were significant at the 1% level with a coefficient of 0.029 compared to 0.04 in Bell et al. (2013). Our results for the extra games variable (or non-Premier League) contrasted with those in Bell et al. (2013). Our estimation results show that home teams were negatively affected by those extra games with a significant coefficient of 0.015 at the 1% level, while the effect for opponents was statistically insignificant. Bell et al. (2013) found that opposing teams were positively influenced with a coefficient of 0.016, but that home teams were unaffected.

The variation in wage bills in our sample was higher than that reported by Bell et al. (2013). It ranged from £61m (Burnley) to £264m (Manchester City) for the 2016/17 season. This difference of £203m far exceeds that of £140m in the Bell et al. study (2013) sample for the 2008/09 season. This indicates a higher spending pattern on players, especially in the top six teams, in our sample. Therefore, we split the sample into two subsamples to account for the top six and the rest separately. The overall average wage bill was £104m with a median of £79.5m. The top six teams were well above the average. Three teams (Everton, Leicester, and Southampton) recorded higher than average wage bills in the 2016/17 season only, but we opted not to include them in the top six. This was to allow for the financial, team quality, and psychological differences between the two groups.

The regression results for the two subsamples are reported in Table 3(a).

[Insert Table 3(a) here]

The top six teams subsample (1) includes 988 observations (26% of the total sample), while the other teams subsample (2) accounts for the remaining 2812 observations. The results were similar for the two subsamples. First, the venue coefficient was virtually identical at 0.46 in the two subsamples and highly significant at the 1% level. This implies that the economic significance of the HF effect was not associated with a specific group of the total sample; rather, teams in both subsamples could exploit the HF advantage. Second, the opponent’s wage bill was significantly negative at the 1% level for both groups. However, it was stronger for the top six teams with a 0.75 point increase on average for a lower opponent’s wage bill as compared with a corresponding 0.59 point increase for the remaining subsample.

However, there were significant differences between the two groups. The wage bill was insignificant for the top six teams, whereas it was significant at the 5% level for the remaining teams with a coefficient of 0.28. The gap in wage bills was considerable between both groups (see Figure 1). The top six average wage bill was £196m with a standard deviation of £34m, compared to £72m with a standard deviation of £16.8m for the other subsample. While wages bills are perceived to play a foremost role in determining teams' competence, its coefficient results imply a reduced sensitivity to increments in wage bills within the top six teams. These teams may need to excel on factors beyond mere high wages (e.g. elite football technical director, coaching staff, composed players dressing room) to secure promotion to the Champions League or winning the Premier League title. On the other hand, the aim for mediocre teams may be confined to securing another season in the PL. This is, to a certain extent, less demanding than the top six level and thus they become less sensitive to high wages. The extra-game variable was significant at the 5% level for the top six teams but insignificant for the remaining teams. The top six team rankings promoted them to additional European games involving mid-week travel. The negative effect may thus result from player fatigue, which limits the options for team managers in their squad selections for PL games.

The other two differences relate to opponents' injuries and suspensions, which were significantly positive for the top six teams but insignificant for the remaining teams. This reflects the superior ability of the top six to cope with these due to the larger number of quality players in their squads.

As a further robustness check on the HF effect, we divided teams into three other categories. These were as follows: top four teams (which secured promotion to the Champions League), mid-table teams (from the 5$^{th}$ to the 17$^{th}$ league ranks, which secures a further season in the Premier League), and lastly relegation teams (from 18$^{th}$ to 20$^{th}$ league ranks, which move to the lower league division next season).

The regression results for these three subsamples are reported in Table 3(b).

[Insert Table 3 (b) here]

The home field effect prevailed across all subsamples and was highly significant at the 1% level. The economic impact remained strong and robust, with very close coefficients of 0.51, 0.44, and 0.49 for the three subsamples, respectively. These strongly support the results in Table 3(a) that even mediocre and relegation zone teams could benefit from the HF advantage. This is not very intuitive for lower-level teams, and hence the effect was novel in this context. Similar to the top four teams, the wage coefficient was insignificant for the relegation teams. This implies a different view as opposed to top teams. Team management might not have been on the right track within the relegation teams, which produces more manager sackings. In our sample, most of the sackings were associated with teams performing below average relative to our bootstrap expected results. Of 31 manager sackings in our sample, only three were associated with top teams (Chelsea in 2012, Tottenham in 2013, and Liverpool in 2012).

### 4.4.2 Bootstrap results

Table 4 shows the actual teams' performances relative to the bootstrap CIs.

[Insert Table 4 here]

Interestingly, in the overall bootstrap average (i.e. across all seasons), points from the first 20 games were calculated to be 1.37 points per game. This is identical to that of Bell et al. (2013), where it was the average expected points for the full season games during the period 2004/05 to 2008/09. The top six teams had an average score of 1.46, while the others had an average of 1.33. Teams with superior financial resources were always in the top four, with a few exceptions. These include Manchester United in 2013 when the Sir Alex Ferguson era had come to an end, and the phenomenal achievement of Leicester City in 2015 when they won the league title under Claudio Ranieri.

The average team points reported in Table 1 represent the actual points won. Comparing these to the bootstrap generated points (those predicted by team financial and player availability characteristics) enables one to better understand team performance. For home games of the top four teams, the difference between the actual and bootstrap points ranged from 0.46 for Arsenal in the 2012/13 season (4th place) to 1.5 for Manchester City in the 2013/14 season (winner of the league title), while the average difference was 0.88. Winning the league title was associated with an average points difference of more than 1.0 points in home games.

Leicester City was an exception in the 2015/16 season with just a 0.57 difference, despite the other top four teams enjoying a higher point difference that season. Leicester balanced their average home results with a very high away games point difference of 0.77 (the maximum point difference in the away games across the sample is 0.85). The point difference for relegated teams ranged from 0.31 for QPR in 2014/15 to -0.78 for QPR in 2012/13. The positive difference for QPR in 2014 is uncommon and far from the average difference for relegated teams of -0.25. This was due to their exceptional home results with average points of 1.8, which more often than not guarantees team survival in the Premiership. However, they had exceptionally bad results in their away games where not a single point had been collected.

The differences between the actual and bootstrap points for away games were wider. Some top four teams had a negative point difference, which shows that playing away from home negatively influenced team performance—this can be seen as an implication of the HF effect for teams outside the top four. The point difference ranged from 0.85 for Tottenham in 2013/14 to -0.51 for Manchester United in 2014/15. This is a range of 1.34 between the maximum and the minimum points relative to a difference of 1.04 in home games. The point difference criterion for title-winning teams is less clear in away games. Manchester City won the league in the 2013 season with a point difference of just -0.09. This may reflect the fact that teams outside the top four were more motivated (‘up for it’) in their home games against the top four opponents.

### 4.4.3 Comparisons relative to the bootstrap confidence interval (CI)

The next step in the analysis is comparing the actual collected points to the CI of the bootstrap distribution. This process provides a benchmark in predicting a team’s success and failure. Table 4 shows that the majority of teams across the sample fell either above or below that benchmark. This applied to both home and away games. The number of standard performance teams in home games ranged from 6 to 11, where 11 occurred only once, and there were no instances for 9 and 10. For away games, this ranged from 2 to 7. The number of instances of a bad (below average) performance in away games was, as expected, considerably higher than those in home games, while on the other hand, the number of good (above average) performances in home games was higher than in away games.

It is plausible to assume that higher wages can attract players with stronger qualities. Thus, predicting success for teams’ promotion to the top four or six is associated with higher spending of club financial resources. On the other hand, it is more challenging to predict survival or relegation teams, as differences in player wages among these teams are narrower and so are their playing qualities (see Figure 1). Clubs hire managers to boost their team standing through more efficient use of available club resources. With our bootstrap approach, we also aimed to assess if teams were on their right track in that regard. Analysing home games as opposed to away games results provided a basis for predicting team success or failure. Nevertheless, this process may also contribute to prediction patterns in sacking team managers.

### 4.4.4 Survival and relegation teams

We started with predicting survival and relegation teams, as these predictions were novel and beyond the notion of predicting success only for top teams. When comparing actual team results with the bootstrap generated CIs, five possible scenarios arose based on our team performance definition: (i) Below-Below where team performance was below the 5% bootstrap distribution benchmark in both home and away games; (ii) Standard-Below where team performance was standard at home and below average away; (iii) Above-Below where performance was above average at home and below average away; (iv) Standard-Standard where team performance was standard at both home and away games; and (v) Above-Standard where performance was above average at home and away.

Observations of teams in these categories are presented in Table 5.

[Insert Table 5 here]

Interestingly, there was no single incidence—within our performance definitions—of a team performing below average at home and above average away (i.e. Below-Above scenario). There were nevertheless only 15 cases out of 100 observations where a team collected more actual points in away games, but these did not suffice to move out of the CI boundaries.

There were seven teams in (i) representing 17 observations across the sample period. In 10 of these observations, teams managed to survive in the PL (Sunderland alone had four observations). We infer that below average performance at home increased relegation chances by around 41% on average (7 out of 17). The majority of teams that survived had some common factors. First, seven managerial changes took place during the season, and in six of these, the change was made before the 20$^{th}$ match week. Teams that did not change their manager were West Ham and Aston Villa in 2013/14, and Leicester in 2014/15. The remaining seven observations in (i) were relegated teams. These involved five managerial changes, of which three were made before the 20$^{th}$ match week. This indicates that early (before the 20$^{th}$ match week) managerial changes for teams that were underperforming could have been vital to improve team performance. Second, 8 out of the 10 survival teams achieved better results in the second half of the league. It is noteworthy that this particular achievement is beyond our prediction period used for the bootstrap (i.e. the first 20 games), but it provides an explanation for most teams that survived. The average ratio of the second half to the first half results in the league was 1.08, with a standard deviation of 0.39. This shows that teams on average performed 8% better in the second half (last 19 games). There were at least eight teams each season in our sample period that performed better in the second half of the league. However, the data show that teams who achieved a ratio higher than 1.47 (this is equivalent to the average ratio plus the standard deviation) were more likely to survive. The eight teams that managed to survive achieved higher than that required ratio, whereas only three relegated teams achieved that ratio. Those relegated teams, despite their better results in the second half, had very poor results in their away games (less than 0.6 points on average). The overall average point for away games was 1.15, while this fell to just 0.6 for relegated teams. The actual away games results for relegated teams who outperformed in the second half varied between 0.2 and 0.4. Hence, struggling teams need to significantly improve in the second half of the season in order to survive in the PL. One possible way to achieve this is to implement an early managerial change.

There were 18 teams under scenario (ii) with 21 observations. There was a 38% chance on average (8 out of 21) that teams in this scenario would be relegated.

Some 10 teams fall under scenario (iii) with 11 observations, and three teams fall under scenario (v) with four observations. Strikingly, all of these teams managed to survive in the PL. This suggests that above-average home performance, regardless of away performance, is sufficient for teams to survive in the PL. Interestingly, teams under those two scenarios kept their manager over the sample period with just two exceptions: Newcastle in 2014/15 and Leicester City in 2016/17. This indicates that club owners perceived good (above average) home performance as implying that a team is on the right track.

There are 12 observations (8 teams) under scenario (iv) where all teams survived well. This indicates that survival with a standard home performance was conditional on at least a standard away performance.

To sum up, teams need to achieve a good (above average) performance in their home games to secure survival in the Premiership. By contrast, a below average performance at either home or away games, or both combined, increases the relegation hazard by roughly 40%. In other words, for non-big 6 teams, it is possible to predict team success based on merely home field results. Away results count for less in this respect. This supports our hypothesis H2.

**4.4.5 Top six teams**

When comparing the top six teams' actual results with the bootstrap CIs, four possible scenarios arose based on team performance: (i) Above-Above where team performance was above the 95% boundary of the bootstrap distribution in both home and away games; (ii) Above-Below where performance was Above at home and Below away; (iii) Above-Standard where performance was Above at home and Standard away; and (iv) Standard-Above where team performance was Standard at home and Above at away. Observations of teams falling within these categories are presented in Table 6.

[insert Table 6 here]

There is no single incidence—within our performance definition—of a top six team with a bad (below average) home performance. In three seasons in our sample, all top six teams enjoyed good (above average) home performance, one season with five teams and another with four teams. In away games, above-average performance was less frequent, and there were two incidences of below-average performance.

A total of eight teams with 28 observations enjoyed good (above average) performance at home.[13] The majority of these teams (20 observations) secured a top four ranking as follows: 13 observations associated with an Above-Above scenario, 2 observations with Above-Below scenario, and 5 observations with Above-Standard scenario. There were only two observations under the Standard-Above scenario where teams did not manage to secure a top four place. This shows that qualifying for the elite European Tournament (Champions League) was associated with at least an above-average home performance. To qualify for the Europa League, above-average home performance was also required, but to a lesser extent.

To sum up, top teams need to achieve a good (above average) performance in their home games to secure promotion to the Champions League. By contrast, a standard performance at either home or away games, or both combined, did not suffice to secure a top four place. Nevertheless, it could lead to a promotion to the Europa League. In other words, away results count for less for the elite European Tournament (Champions League). This also supports our hypothesis H2.

All of these teams kept their manager during the sample period with just two exceptions: Chelsea in 2012/13 and Tottenham in 2013/14. Like mid- and lower-table teams, an above-average home performance signals a stable team that significantly reduces the chances of managerial sacking.

## 4.5 Conclusion

This paper introduced a novel home field effect in studying team performance in the English Premier League. It is well documented in the literature that, across different sporting

[13] These teams were; Manchester United, Manchester City, Chelsea, Liverpool, Arsenal, Tottenham, Everton and Leicester. The first six were perceived as the big six clubs in English football, mainly due to their financial power. Throughout English football history, Manchester United and Liverpool (in addition to Arsenal and Chelsea in the last two decades) have been specifically known for their dominance over football trophies and competitions.

disciplines, teams win more often when playing at home than when playing away. Several studies have focused on the psychological aspects (e.g. crowd effect, opponents travel fatigue, familiarity with stadium features, sporting rules favouring home team) that might be driving this phenomenon. Another strand of literature has attempted to test home advantage empirically, with a focus on within-match performance measures. These include, for instance, differences in teams' tactical approaches and playing style at home as opposed to away games, players' behaviour and aggressiveness while playing away from home, goal differentials, and how they reflect better overall performance for home teams. There is also a strand that focuses on forecasting match results through different modelling techniques, such as Monte Carlo simulations and artificial neural networks (ANNs).

We combined these strands of literature to test for a new HF advantage/effect and its predictive powers. We first used fixed effect regressions to test conjecture that playing at home could positively affect team performance even after controlling for other relevant team characteristics. Our results established a highly significant HF effect of 0.46 points per game over our full sample period. The result was robust for the top six teams and the lower level teams. Running a regression for five subsamples showed an almost identical effect where the coefficient preserved a narrow magnitude between 0.44 and 0.51. This is a novel effect in the context of Premier League football. Additionally, we found that wage coefficients were insignificant for the top and relegation teams. This indicates that the HF effect can dominate the wage effect once teams are fighting for promotions to elite tournaments or for survival in the Premiership.

Second, we employed a bootstrap approach to test whether home advantage could predict team success and failure in the Premier League. Early league games (i.e. first 20 games in the league fixtures) were used in the prediction process. The proposed predictive power of home over away games was supported by our empirical findings. Teams who outperformed in their early home games would stay in the Premier League even if they performed poorly in away games. A standard home performance combined with above average away performance may not suffice to secure a top four ranking. Moreover, a standard team performance at home that is associated with a below average performance away from home increased the probability of relegation to the Championship by about 40%.

## 4.6 Tables and Figures

### 4.6.1 Table 1 Descriptive statistics for team point averages

| Season | (1) Top 6 avg | (2) Mid-table avg | (3) Relegation avg | (4) Total avg |
|---|---|---|---|---|
| *Panel A: Home games* | | | | |
| 2012 | 2.10 | 1.55 | 0.98 | 1.54 |
| 2013 | 2.40 | 1.27 | 1.15 | 1.61 |
| 2014 | 2.17 | 1.43 | 1.20 | 1.60 |
| 2015 | 2.00 | 1.34 | 1.05 | 1.46 |
| 2016 | 2.32 | 1.49 | 1.13 | 1.64 |
| Total avg | 2.20 | 1.42 | 1.10 | 1.57 |
| *Panel B: Away games* | | | | |
| 2012 | 1.82 | 0.88 | 0.60 | 1.10 |
| 2013 | 1.70 | 1.02 | 0.80 | 1.17 |
| 2014 | 1.68 | 1.02 | 0.70 | 1.13 |
| 2015 | 1.70 | 1.26 | 0.60 | 1.19 |
| 2016 | 1.93 | 0.77 | 0.60 | 1.10 |
| Total avg | 1.77 | 0.99 | 0.66 | 1.14 |
| *Panel c: Equality of means tests* | | | | |
| | *t*-statistic | *t*-statistic | *t*-statistic | *t*-statistic |
| 2012 | 5.44*** | 7.24*** | 2.42** | 5.14*** |
| 2013 | 5.39*** | 6.16*** | 4.25*** | 4.48*** |
| 2014 | 5.40*** | 6.51*** | 4.31*** | 4.79*** |
| 2015 | 5.40*** | 6.11*** | 3.33*** | 3.48*** |
| 2016 | 5.88*** | 7.34*** | 3.70*** | 6.52*** |
| Total avg | 5.72*** | 10.23*** | 10.1*** | 10.9*** |

*Notes*: Average points are calculated from only 10 game results in panels A and B, while the rankings are based on the end of season results. The first column reports averages for teams placed in the top 6 rankings. The second reports mid-table teams' rankings from the 7$^{th}$ place to the 16$^{th}$ (total of 10 teams). The third reports the last 4 bottom teams in the table, out of which 3 teams will be relegated to the lower league (Championship League). The fourth is the total average of all teams. Panel C reports the *t*-statistics of the equality of mean difference tests for the corresponding pairs of games in each column, with their relevant significance levels. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

**4.6.2 Table 2 Regression results of model (1): Home field effect on team performance**

| | (1)<br>**Team Performance**<br>b/se |
|---|---|
| Match venue | 0.4608*** |
| | (0.039) |
| Log-wage | 0.3213*** |
| | (0.123) |
| Injured | -0.0114 |
| | (0.013) |
| N/A | 0.0007 |
| | (0.007) |
| Suspended | 0.0142 |
| | (0.065) |
| Extra-games | -0.0153*** |
| | (0.005) |
| Log-wage_opp | -0.6271*** |
| | (0.059) |
| Injured_opp | 0.0291*** |
| | (0.011) |
| N/A_opp | 0.0090 |
| | (0.007) |
| Suspended_opp | 0.0353 |
| | (0.064) |
| Extra-games_opp | -0.0059 |
| | (0.005) |
| _cons | 2.6148*** |
| | (0.536) |
| N | 3800 |
| $R^{2\ (\text{within})}$ | 0.1062 |
| $R^{2\ (\text{between})}$ | 0.4391 |

*Notes*: A fixed effect panel regression model is used. The dependent variable is team performance, represented by its match points. This takes a value of 3 for a win, 1 for a draw, and 0 for a loss. The following variables belong to the hosting team: Match venue is where the match is held. It is a dummy variable that takes a value of 1 for a home game and 0 for an away game. Log-wage is the wage bill of clubs in each season; its original values are in £million and reported in a log form. Injured is the number of injured players in a match day. N/A is the number of unavailable players in a match day for different reasons, such as playing for international or team of the reserves team within their club. Suspended is the number of suspended players in a match day due to either red/consecutive yellow cards or a misbehaviour suspension. Extra-games are the number of non-Premier League games that a team plays during a season. The rest of the variables (that end with _opp) are identical but relate to the opponent team. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

**4.6.3 Table 3(a) Regression results of model (1): Subsamples of top six teams and the rest of table ranking**

| | (1)<br>**Top six teams**<br>b/se | (2)<br>**All other teams**<br>b/se |
|---|---|---|
| Match venue | 0.4606*** | 0.4596*** |
| | (0.075) | (0.045) |
| Log-wage | 0.2849 | 0.2897** |
| | (0.330) | (0.134) |
| Injured | -0.0285 | -0.0063 |
| | (0.025) | (0.015) |
| N/A | -0.0085 | 0.0031 |
| | (0.014) | (0.009) |
| Suspended | -0.0287 | 0.0375 |
| | (0.125) | (0.076) |
| Extra-games | -0.0168** | -0.0119 |
| | (0.007) | (0.008) |
| Log-wage_opp | -0.7456*** | -0.5904*** |
| | (0.116) | (0.068) |
| Injured_opp | 0.0506** | 0.0198 |
| | (0.021) | (0.012) |
| N/A_opp | 0.0125 | 0.0077 |
| | (0.013) | (0.008) |
| Suspended_opp | 0.2954** | -0.0647 |
| | (0.122) | (0.075) |
| Extra-games_opp | -0.0020 | -0.0070 |
| | (0.009) | (0.005) |
| _cons | 3.7703** | 2.4295*** |
| | (1.628) | (0.561) |
| N | 988 | 2812 |
| $R^{2\ (within)}$ | 0.1285 | 0.1019 |
| $R^{2\ (between)}$ | 0.2816 | 0.0155 |

*Notes*: Identical variable definitions of Table 2.

**Table 3(b) Regression results of model (1): Subsamples of top four, mid-table, and relegation teams**

| | (1) **Top four teams** b/se | (2) **Mid-table teams** | (3) **Relegation teams** b/se |
|---|---|---|---|
| Match venue | 0.5079*** | 0.4382*** | 0.4900*** |
| | (0.083) | (0.049) | (0.091) |
| Log-wage | 0.4086 | 0.3752** | 0.4610 |
| | (0.421) | (0.152) | (1.061) |
| Injured | -0.0196 | -0.0099 | 0.0226 |
| | (0.029) | (0.017) | (0.033) |
| N/A | 0.0006 | 0.0067 | -0.0025 |
| | (0.017) | (0.009) | (0.022) |
| Suspended | 0.0573 | 0.0354 | -0.1341 |
| | (0.147) | (0.078) | (0.213) |
| Extra-games | -0.0125 | -0.0058 | -0.0549 |
| | (0.010) | (0.008) | (0.069) |
| Log-wage_opp | -0.6624*** | -0.6208*** | -0.5891*** |
| | (0.129) | (0.074) | (0.140) |
| Injured_opp | 0.0388* | 0.0269** | 0.0211 |
| | (0.023) | (0.014) | (0.026) |
| N/A_opp | 0.0044 | 0.0128 | -0.0101 |
| | (0.015) | (0.009) | (0.016) |
| Suspended_opp | 0.2291* | 0.0183 | -0.1310 |
| | (0.136) | (0.082) | (0.154) |
| Extra-games_opp | -0.0093 | -0.0063 | -0.0008 |
| | (0.010) | (0.006) | (0.011) |
| _cons | 2.8414 | 2.1669*** | 1.5980 |
| | (2.145) | (0.642) | (4.311) |
| N | 760 | 2470 | 570 |
| $R^{2\ (within)}$ | 0.1352 | 0.0988 | 0.1170 |
| $R^{2\ (between)}$ | 0.0068 | 0.4794 | 0.1243 |

*Notes*: Identical variable definitions of Table 2. Column (2) presents teams from the 5th rank in the league table up to the 17th rank. Column (3) presents teams from the 18th rank in the league table up to the 20th rank.

### 4.6.4 Table 4 Actual team points compared with the bootstrap C.I.s

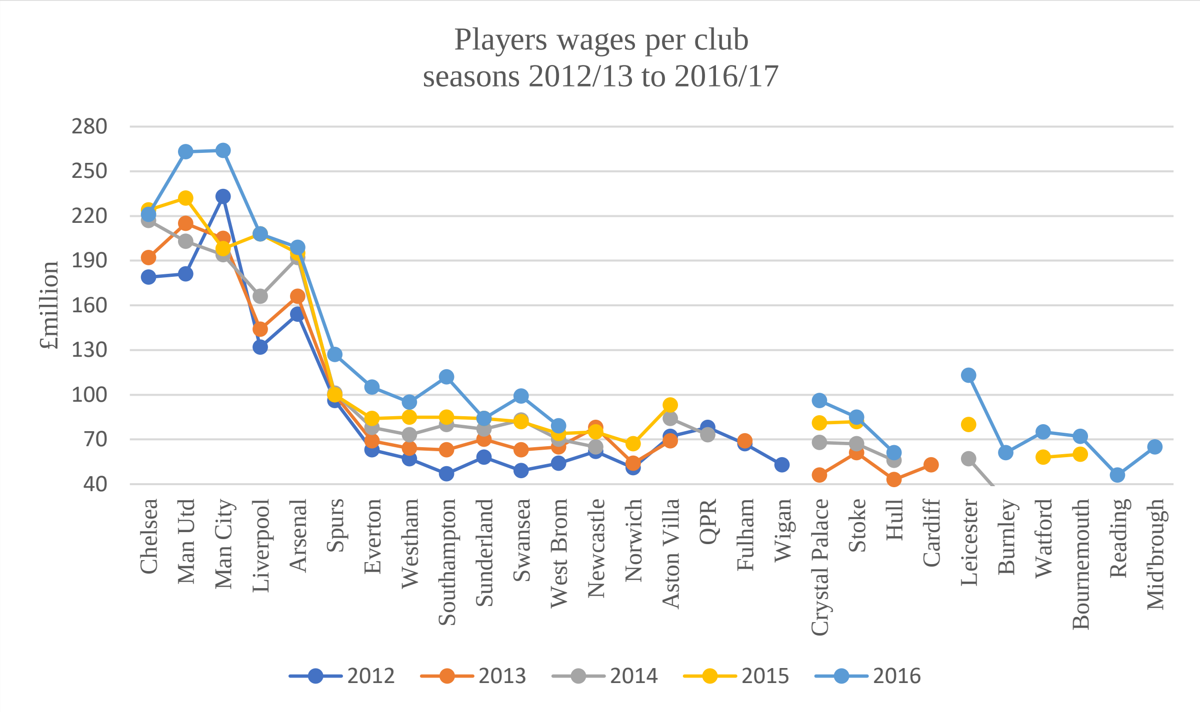
| | | (1) Home games | | (2) Away games | | (3) | (4) |
|---|---|---|---|---|---|---|---|
| Club | Season | Actual pts | Within C.I | Actual pts | Within C.I | Bootstrap pts | Table rank |
| Man Utd | 2012 | 2.7 | above | 2.2 | above | 1.64 | 1 |
| Man City | | 2.3 | above | 1.9 | above | 1.71 | 2 |
| Chelsea | | 1.9 | above | 2 | above | 1.37 | 3 |
| Arsenal | | 2 | above | 1.6 | - | 1.54 | 4 |
| Spurs | | 1.8 | above | 1.8 | above | 1.40 | 5 |
| Everton | | 1.9 | above | 1.4 | - | 1.39 | 6 |
| Liverpool | | 1.5 | - | 1.3 | - | 1.48 | 7 |
| West Brom | | 2.2 | above | 1.1 | below | 1.37 | 8 |
| Swansea | | 1.4 | - | 1.4 | - | 1.23 | 9 |
| Westham | | 1.5 | - | 0.8 | below | 1.38 | 10 |
| Norwich | | 1.7 | above | 0.8 | below | 1.27 | 11 |
| Fulham | | 1.4 | - | 0.7 | below | 1.53 | 12 |
| Stoke | | 2 | above | 0.9 | below | 1.41 | 13 |
| Southampton | | 1.2 | - | 0.6 | below | 1.33 | 14 |
| Aston Villa | | 1 | - | 0.8 | below | 1.31 | 15 |
| Newcastle | | 1.6 | above | 0.4 | below | 1.22 | 16 |
| Sunderland | | 1.2 | below | 1 | below | 1.39 | 17 |
| Wigan | | 0.9 | below | 0.9 | below | 1.26 | 18 |
| Reading | | 1.1 | below | 0.2 | below | 1.36 | 19 |
| QPR | | 0.7 | below | 0.3 | below | 1.49 | 20 |
| Man City | 2013 | 3 | above | 1.4 | - | 1.50 | 1 |
| Liverpool | | 2.7 | above | 1.2 | below | 1.71 | 2 |
| Chelsea | | 2.8 | above | 1.5 | - | 1.53 | 3 |
| Arsenal | | 2.3 | above | 2.2 | above | 1.50 | 4 |
| Everton | | 2.1 | above | 1.7 | above | 1.43 | 5 |
| Spurs | | 1.5 | - | 2.2 | above | 1.35 | 6 |
| Man Utd | | 1.4 | - | 2 | above | 1.64 | 7 |
| Southampton | | 1.5 | - | 1.2 | - | 1.37 | 8 |
| Stoke | | 1.7 | above | 0.5 | below | 1.31 | 9 |
| Newcastle | | 1.8 | above | 1.5 | - | 1.40 | 10 |
| Crystal Palace | | 1.1 | below | 0.6 | below | 1.38 | 11 |
| Swansea | | 1 | - | 1.1 | - | 1.20 | 12 |
| Westham | | 0.9 | below | 0.6 | below | 1.32 | 13 |
| Sunderland | | 0.7 | below | 0.7 | below | 1.34 | 14 |
| Aston Villa | | 0.8 | below | 1.5 | - | 1.43 | 15 |
| Hull | | 1.8 | above | 0.5 | below | 1.12 | 16 |
| West Brom | | 1.2 | - | 0.9 | below | 1.38 | 17 |
| Norwich | | 1.2 | - | 0.8 | below | 1.38 | 18 |
| Fulham | | 1 | below | 0.9 | below | 1.43 | 19 |
| Cardiff | | 1.2 | - | 0.6 | below | 1.38 | 20 |
| Chelsea | 2014 | 3 | above | 1.9 | above | 1.62 | 1 |
| Man City | | 2.3 | above | 2.3 | above | 1.59 | 2 |
| Arsenal | | 2.1 | above | 1.5 | - | 1.38 | 3 |
| Man Utd | | 2.5 | above | 1.2 | below | 1.72 | 4 |
| Spurs | | 1.5 | - | 1.7 | above | 1.27 | 5 |
| Liverpool | | 1.6 | - | 1.5 | - | 1.58 | 6 |
| Southampton | | 2 | above | 1.6 | - | 1.29 | 7 |
| Swansea | | 2 | above | 0.9 | below | 1.47 | 8 |
| Stoke | | 1.4 | - | 1.2 | - | 1.25 | 9 |
| Crystal Palace | | 1.1 | - | 0.8 | below | 1.32 | 10 |
| Everton | | 1.3 | - | 0.9 | below | 1.36 | 11 |
| Westham | | 1.9 | above | 1.3 | - | 1.38 | 12 |

continued

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | 13 |
| West Brom | | 0.9 | below | 0.9 | below | 1.32 | |
| Leicester | | 1 | below | 0.6 | below | 1.27 | 14 |
| Newcastle | | 1.8 | above | 0.9 | - | 1.21 | 15 |
| Sunderland | | 0.9 | below | 1.1 | below | 1.34 | 16 |
| Aston Villa | | 1.1 | - | 1.1 | - | 1.29 | 17 |
| Hull | | 0.9 | below | 1 | below | 1.41 | 18 |
| Burnley | | 1 | - | 0.7 | below | 1.10 | 19 |
| QPR | | 1.8 | - | 0 | below | 1.48 | 20 |
| Leicester | | 1.9 | above | 2.1 | above | 1.32 | 1 |
| Arsenal | | 2.3 | above | 1.9 | above | 1.42 | 2 |
| Spurs | | 1.9 | above | 1.7 | above | 1.18 | 3 |
| Man City | | 2.4 | above | 1.5 | - | 1.34 | 4 |
| Man Utd | | 1.9 | above | 1.4 | - | 1.36 | 5 |
| Westham | | 1.6 | above | 1.6 | above | 1.16 | 6 |
| Liverpool | | 1.6 | - | 1.5 | - | 1.41 | 7 |
| Southampton | | 1.4 | - | 1 | below | 1.31 | 8 |
| Stoke | | 1.6 | above | 1.6 | above | 1.22 | 9 |
| Chelsea | | 1.4 | - | 0.9 | below | 1.48 | 10 |
| Everton | 2015 | 1.2 | - | 1.5 | above | 1.19 | 11 |
| Swansea | | 1.3 | - | 0.6 | below | 1.36 | 12 |
| Watford | | 1.4 | - | 1.5 | above | 1.14 | 13 |
| West Brom | | 1.1 | - | 1.3 | - | 1.30 | 14 |
| Crystal Palace | | 1.4 | - | 1.7 | above | 1.27 | 15 |
| Bournemouth | | 1 | - | 1 | - | 1.20 | 16 |
| Sunderland | | 1.1 | below | 0.4 | below | 1.45 | 17 |
| Newcastle | | 1 | - | 0.7 | below | 1.23 | 18 |
| Norwich | | 1.5 | - | 0.8 | below | 1.44 | 19 |
| Aston Villa | | 0.6 | below | 0.5 | below | 1.34 | 20 |
| Chelsea | | 2.7 | above | 2.2 | above | 1.66 | 1 |
| Spurs | | 2.6 | above | 1.6 | above | 1.33 | 2 |
| Man City | | 2.1 | above | 2.1 | above | 1.51 | 3 |
| Liverpool | | 2.3 | above | 1.9 | above | 1.55 | 4 |
| Arsenal | | 2.3 | above | 1.8 | above | 1.41 | 5 |
| Man Utd | | 1.9 | above | 2 | above | 1.39 | 6 |
| Everton | | 1.9 | above | 1.1 | below | 1.43 | 7 |
| Southampton | | 1.5 | - | 0.9 | below | 1.25 | 8 |
| Bournemouth | | 1.7 | above | 0.8 | below | 1.35 | 9 |
| West Brom | | 1.7 | - | 1.2 | - | 1.37 | 10 |
| Westham | 2016 | 1.4 | - | 0.8 | - | 1.18 | 11 |
| Leicester | | 1.8 | above | 0.3 | below | 1.35 | 12 |
| Stoke | | 1.5 | - | 0.9 | below | 1.42 | 13 |
| Crystal Palace | | 0.7 | below | 0.9 | below | 1.33 | 14 |
| Swansea | | 0.8 | below | 0.7 | below | 1.35 | 15 |
| Burnley | | 1.9 | above | 0.1 | below | 1.21 | 16 |
| Watford | | 1.4 | - | 0.8 | below | 1.30 | 17 |
| Hull | | 0.9 | below | 0.4 | below | 1.17 | 18 |
| Mid'brough | | 1.1 | - | 0.8 | below | 1.17 | 19 |
| Sunderland | | 1.1 | - | 0.4 | below | 1.22 | 20 |

*Notes*: The double columns (1) and (2) report the actual number of points that teams collected during the first 10 games in each venue (i.e. 10 home games in (1) and 10 away games in (2)). The C.I. is the confidence interval of the distribution generated by the bootstrap, which is based on the data of the first 20 games of each team during the season. Bootstrap points in (3) are calculated using the median of the bootstrap distribution. Table 4 shows the team rank in the league table by the end of the season. The hyphens in (1) and (2) are equivalent to ‘standard’, while ‘above’ is above 95% C.I. and ‘below’ is below 5% C.I.

**4.6.5 Figure 1**




*Note*: The vertical axis represents clubs' wage bills in £million. The horizontal axis represents clubs examined in the sample period. One can notice that the differences (clustering) in wages for mid table teams are thinner than those in the top 6 rankings.

**4.6.6 Table 5 Relegated and survived teams' performance in home/away scenarios (top six rankings teams are excluded)**

| | (i) | | (ii) | | (iii) | | (iv) | | (v) | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Season** | **Below/Below** | **rel** | **Standard/Below** | **rel** | **Above/Below** | **rel** | **Standard/Standard** | **rel** | **Above/Standard** | **rel** |
| 2012 | 4 | 3 | 4 | 0 | 4 | 0 | 2 | 0 | 0 | 0 |
| 2013 | 4 | 1 | 3 | 2 | 2 | 0 | 2 | 0 | 1 | 0 |
| 2014 | 4 | 1 | 4 | 2 | 1 | 0 | 3 | 0 | 3 | 0 |
| 2015 | 2 | 1 | 5 | 2 | 0 | 0 | 3 | 0 | 0 | 0 |
| 2016 | 3 | 1 | 5 | 2 | 4 | 0 | 2 | 0 | 0 | 0 |

*Note*: Under each of the 5 scenarios, there are 2 columns; the first one defines team performance at home and away fields based on our scenario criteria, under which the number of observations for teams falling within that scenario are presented. The second column is the relegation (rel), under which the number of observations for relegated teams are presented.

### 4.6.7 Table 6 Top six teams' performance in home/away scenarios

| | (i) | | (ii) | | (iii) | | (iv) | |
|---|---|---|---|---|---|---|---|---|
| **Season** | **Above/Above** | **Top** | **Above/Below** | **Top** | **Above/Standard** | **Top** | **Standard/Above** | **Top** |
| 2012 | 4 | 3 | 0 | 0 | 2 | 1 | 0 | 0 |
| 2013 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 0 |
| 2014 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 |
| 2015 | 5 | 3 | 0 | 0 | 2 | 1 | 0 | 0 |
| 2016 | 6 | 4 | 0 | 0 | 0 | 0 | 0 | 0 |

*Note*: Under each of the 4 scenarios, there are 2 columns; the first one defines team performance at home and away fields based on our scenario criteria, under which the number of observations for teams falling within that scenario are presented. The second column is the promotion (Top), under which the number of observations for teams promoted to the 4 four rankings are presented.

## 5 Concluding Remarks

The notion of rational economic decisions has been challenged by behavioural economics. Much of that literature is based on artificial experiments, but sports fields introduce a real-life setup for scientific research. One of the earliest contributions was by Staw and Hoang (1995), who tested their earlier behavioural hypotheses with regard to decision-making processes (particularly agents' tendencies to escalate on their commitments) through the NBA. There are several ways in which decision-making can drift from rationality. We investigated sunk costs and discriminatory behaviour within a football context, particularly the English Premier League. In addition, we explored how athletes could be privileged when playing at their home venues.

This thesis contributes to the literature from three different perspectives. First, it contributes to the sunk cost (or escalation of commitment) literature (e.g. Arkes & Blumer, 1985; Staw & Hoang, 1995). These studies show behavioural biases, namely the need to avoid feeling like resources have been wasted and escalating commitments, particularly in social and sporting setups. We utilised the unique features of European football, especially from a financial perspective, to test for those biases. The rationale was that managers should play their most effective players regardless of their cost. However, when a player's transfer fee escalates, it may induce the manager's need to not feel resources have been wasted on him, in addition to the pressure from fans to play the team stars (i.e. high cost players). Hence, those players may get more playing time even after controlling for their performance. Moreover, we explored the frequent manager turnover within clubs to test whether the sunk cost effect could be mitigated. The rationale of this conjecture was that a self-justification explanation of the escalation of commitment situation or sunk cost effect was removed when a new manager replaced the one who was involved in signing the players (i.e. who bore the sunk cost account).

Our empirical findings supported the sunk cost hypothesis. The transfer fee was shown to be positively related to playing time.

Second, it contributes to the literature on labour discrimination. It is complicated to measure employees' productivity with precision. However, assessing investment managers' performance is one of the few fields for such measurement, and another is the sporting field. In any case, discrimination is usually empirically measured through wage equations. We introduced a new discriminatory variable associated with player playing time in the English Premier League, which we call the nationality effect. The basic idea was that managers favoured players who shared a common cultural background with them, be it nationality, language or even a particular playing style. Additionally, clubs' sources (e.g. managers, agents, scouts) who are responsible for searching and recruiting players may be embedded with social networks, and hence seek players who resemble the local standards even if equal or even better foreign talent is available in the market. Therefore, we conjectured that players who shared a common cultural background with their manager would be allocated extra playing time per season relative to those who did not, beyond what could be attributed to mere performance. The nationality effect was supported by our empirical findings. For instance, British players playing under a British manager got an extra 100 minutes of playing time per season.

Lastly, it makes a contribution to the literature that is associated with home field advantage in sports. It examines the persistence of that advantage in the Premier League and how it can affect team performance. Additionally, we explored how the home advantage could predict team success or failure, from the perspective of promotion to elite European football tournaments or relegation to the lower tier in the English league system. We conjectured that playing at home could positively affect team performance even after controlling for other relevant team characteristics. Our results established a highly significant HF effect of 0.46 points per game. We employed a bootstrap approach to test whether home advantage could predict team success and failure. The proposed predictive power of home over away games was supported by our empirical findings. Teams who outperformed in their early home games stayed in the Premier League even if they performed poorly in away games. Moreover, a standard home performance combined with above average away performance did not suffice to secure a top four ranking, and a standard team performance at home that was associated with a below average performance away from home increased the probability of relegation to the Championship by about 40%.

## 5.1 Future Research

An interesting avenue for future research emerges from our results in Chapters 2 and 3. Based on our findings, researchers could investigate similar effects in other football leagues such as the La Liga in Spain or the Bundsliga in Germany, especially with more detailed player wages data availability with respect to the Bundsliga in particular. It would be interesting to test for the same effects but with a recent dataset that encompasses the surge in TV rights deal values for Premier League games, and accordingly higher player wages.

Further insights can also be gained from Chapter 4 results. This could be through investigating the home field effect from the perspective of team goal scoring, as well as the football betting markets.

## References


Agnew, G. A., & Carron, A. V. (1994). Crowd effects and the home advantage. *International Journal of Sport Psychology, 25*(1), 53–62.

Allen, M. S., & Jones, M. V. (2014). The home advantage over the first 20 seasons of the English Premier League: Effects of shirt colour, team ability and time trends. *International Journal of Sport and Exercise Psychology, 12*(1), 10–18.

Arabzad, S. M., Tayebi Araghi, M. E., Sadi-Nezhad, S., & Ghofrani, N. (2014). Football match results prediction using artificial neural networks: The case of Iran Pro League. *Journal of Applied Research on Industrial Engineering*, *1*(3), 159–179.

Arkes, H. R. (1996). The psychology of waste. *Journal of Behavioral Decision Making, 9*(3), 213–224.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organisational Behaviour and Human Decision Processes, 35*(1), 124–140.

Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals?. *Psychological Bulletin*, *125*(5), 591.

Arrow, K. (1973). The theory of discrimination. *Discrimination in Labor Markets*, *3*(10), 333.

Ayton, P., & Arkes, H. (1998). Call it quits. *New Scientist*, (2135), 40–43.

Barnett, V., & Hilditch, S. (1993). The effect of an artificial pitch surface on home team performance in football (soccer). *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *156*(1), 39–50.

Bazerman, M. H., Beekun, R. I., & Schoorman, F. D. (1982). Performance evaluation in a dynamic context: A laboratory study of the impact of a prior commitment to the ratee. *Journal of Applied Psychology*, *67*(6), 873.

Becker, G. The Economics of Discrimination, 2nd edition. Chicago: University of Chicago Press, 1971.

Bell, A., Brooks, C., & Markham, T. (2013). The performance of football club managers: Skill or luck? *Economics & Finance Research*, *1*(1), 19–30.

Bodvarsson, Ö. B., & Brastow, R. T. (1999). A test of employer discrimination in the NBA. *Contemporary Economic Policy*, *17*(2), 243–255.

Bowen, M. G. (1987). The escalation phenomenon reconsidered: Decision dilemmas or decision errors? *Academy of Management Review*, *12*(1), 52–66.

Bray, S. R. (1999). The home advantage from an individual team perspective. *Journal of Applied Sport Psychology*, *11*(1), 116–125.

Bray, S. R., Law, J., & Foyle, J. (2003). Team quality and game location effects in English professional soccer. *Journal of Sport Behavior*, *26*(4).

Bray, S. R., Martin, K. A., & Widmeyer, W. N. (2000). The relationship between evaluative concerns and sport competition state anxiety among youth skiers. *Journal of Sports Sciences*, *18*(5), 353–361.

Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review*, *17*(1), 39–61.

Brown, E., Spiro, R., & Keenan, D. (1991). Wage and nonwage discrimination in professional basketball: do fans affect it?. *American Journal of Economics and Sociology*, *50*(3), 333–345.

Burdekin, R. C., Hossfeld, R. T., & Smith, J. K. (2005). Are NBA fans becoming indifferent to race? Evidence from the 1990s. *Journal of Sports Economics*, *6*(2), 144–159.

Camerer, C. F., & Weber, R. A. (1999). The econometrics and behavioural economics of escalation of commitment: A re-examination of Staw and Hoang’s NBA data. *Journal of Economic Behavior & Organisation, 39*(1), 59–82.

Carmichael, F., & Thomas, D. (2005). Home-field effect and team performance: Evidence from English premiership football. *Journal of Sports Economics*, *6*(3), 264–281.

Carmichael, F., Thomas, D., & Ward, R. (2000). Team performance: The case of English premiership football. *Managerial and Decision Economics*, *21*(1), 31–45.

Carnibella, G., A. Fox, K. Fox, J. McCann, J. Marsh, and P. Marsh (1996). Football violence in Europe: A report to the Amsterdam group. *The Social Issues Research Centre*.

Carron, A. V., Loughhead, T. M., & Bray, S. R. (2005). The home advantage in sport competitions: Courneya and Carron's (1992) conceptual framework a decade later. *Journal of Sports Sciences*, *23*(4), 395–407.

Chiappori, P. A., & Salanié, B. (2002). Testing contract theory: A survey of some recent work.

Clarke, S. R., & Norman, J. M. (1995). Home ground advantage of individual clubs in English soccer. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *44*(4), 509–521.

Conlon, E. J., & Parks, J. M. (1987). Information requests in the context of escalation. *Journal of Applied Psychology*, *72*(3), 344.

Courneya, K. S., & Carron, A. V. (1990). Batting first versus last: Implications for the home advantage. *Journal of Sport and Exercise Psychology*, *12*(3), 312–316.

Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport & Exercise Psychology*, *14*(1).

Crowder, M., Dixon, M., Ledford, A., & Robinson, M. (2002). Dynamic modelling and prediction of English Football League matches for betting. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *51*(2), 157–168.

Davis, M. A., & Bobko, P. (1986). Contextual effects on escalation processes in public sector decision making. *Organisational Behavior and Human Decision Processes*, *37*(1), 121–138.

Dicken, P. (1998). Global Shift, 3rd edition. London: Paul Chapman.

Dixon, M., & Robinson, M. (1998). A birth process model for association football matches. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *47*(3), 523–538.

Dobson, S., & Goddard, J. (2003). Persistence in sequences of football match results: A Monte Carlo analysis. *European Journal of Operational Research*, *148*(2), 247–256.

Dobson, S., & Goddard, J. A. (2009). *The economics of football*. Cambridge: Cambridge University Press.

Downward, P., & Jones, M. (2007). Effects of crowd size on referee decisions: Analysis of the FA Cup. *Journal of Sports Sciences*, *25*(14), 1541–1545.

Edwards, J., & Archambault, D. (1989). The home field advantage. In J. H. Goldstein (Ed.), Sports, games, and play: Social and psychological viewpoints (2nd ed., pp. 333–370). Hillsdale, NJ: Lawrence Erlbaum.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*(4), 380.

Eugster, M. J., Gertheiss, J., & Kaiser, S. (2011). Having the second leg at home-advantage in the UEFA Champions League knockout phase? *Journal of Quantitative Analysis in Sports*, *7*(1).

Fox, F. V., & Staw, B. M. (1979). The trapped administrator: Effects of job insecurity and policy resistance upon commitment to a course of action. *Administrative Science Quarterly*, 449–471.

Frick, B., & Deutscher, C. (2009). Salary determination in the German “Bundesliga”: A panel study. IASE Conference Paper 0811. *International Association of Sports Economists*.

Goddard, J., & Asimakopoulos, I. (2003). *Modelling football match results and the efficiency of fixed-odds betting*. Working Paper, Department of Economics, Swansea University.

Goddard, J., & Asimakopoulos, I. (2004). Forecasting football results and the efficiency of fixed-odds betting. *Journal of Forecasting*, *23*(1), 51–66.

Goddard, J., & Wilson, J. O. (2009). Racial discrimination in English professional football: Evidence from an empirical analysis of players' career progression. *Cambridge Journal of Economics*, *33*(2), 295–316.

Gwartney, J., & Haworth, C. (1974). Employer costs and discrimination: The case of baseball. *Journal of Political Economy*, *82*(4), 873–881.

Hamilton, B. H. (1997). Racial discrimination and professional basketball salaries in the 1990s. *Applied Economics*, *29*(3), 287–296.

Heckman, J. J. (1998). Detecting discrimination. *Journal of Economic Perspectives*, *12*(2), 101–116.

Hirst, P., & Thompson, G. (1996). Globalisation: Ten frequently asked questions and some surprising answers. *Soundings-London-Lawrence and Wishart*, 47–66.

Holmes, P. (2011). New evidence of salary discrimination in major league baseball. *Labour Economics*, *18*(3), 320–331.

Jamieson, J. P. (2010). The home field advantage in athletics: A meta-analysis. *Journal of Applied Social Psychology*, 40, 1819–1848.

Kahn, L. M. and P. D. Sherer (1988). Racial differences in professional basketball players' compensation. *Journal of Labor Economics*, *6*(1), 40–61.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society,* 263–291.

Keefer, Q. A. (2017). The sunk-cost fallacy in the National Football League: Salary cap value and playing time. *Journal of Sports Economics*, *18*(3), 282–297.

Koch, J. V., & Vander Hill, C. W. (1988). Is there discrimination in the "Black Man's Game"? *Social Science Quarterly*, 83–94.

Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy, 89*(5), 841–864.

Lebovic, J. H., & Sigelman, L. (2001). The forecasting accuracy and determinants of football rankings. *International Journal of Forecasting*, *17*(1), 105–120.

Leeds, D. M., Leeds, M. A., & Motomura, A. (2015). Are sunk costs irrelevant? Evidence from playing time in the National Basketball Association. *Economic Inquiry*, *53*(2), 1305–1316.

Lehmann, E. E., & Schulze, G. G. (2008). What does it take to be a star?–The role of performance and the media for German soccer players. *Applied Economics Quarterly, 54*(1), 59–70.

Lucifora, C., & Simmons, R. (2003). Superstar effects in sport: Evidence from Italian soccer. *Journal of Sports Economics*, *4*(1), 35–55.

McGovern, P. (2002). Globalisation or internationalisation? Foreign footballers in the English league, 1946-95. *Sociology*, *36*(1), 23–42.

Moroney MJ. (1956). Facts from Figures, 3rd edn. Penguin: London.

Neale, W. C. (1964). The peculiar economics of professional sports. *The Quarterly Journal of Economics*, *78*(1), 1–14.

Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise*, *3*(4), 261–272.

Nevill, A. M., & Holder, R. L. (1999). Home advantage in sport. *Sports Medicine, 28*(4), 221–236.

Nevill, A. M., Newell, S. M., & Gale, S. (1996). Factors associated with home advantage in English and Scottish soccer matches. *Journal of Sports Sciences*, *14*(2), 181–186.

Northcraft, G. B., & Neale, M. A. (1986). Opportunity costs and the framing of resource allocation decisions. *Organisational Behavior and Human Decision Processes*, *37*(3), 348–356.

Pedace, R. (2008). Earnings, performance, and nationality discrimination in a highly competitive labor market as an analysis of the English professional soccer league. *Journal of Sports Economics*, *9*(2), 115–140.

Pollard, R. (1986). Home advantage in soccer: A retrospective analysis. *Journal of Sports Sciences*, *4*(3), 237–248.

Pollard, R. (2002). Evidence of a reduced home advantage when a team moves to a new stadium. *Journal of Sports Sciences*, *20*(12), 969–973.

Pollard, R. (2006). Worldwide regional variations in home advantage in association football. *Journal of sports sciences*, *24*(3), 231–240.

Pollard, R., & Pollard, G. (2005). Long-term trends in home advantage in professional team sports in North America and England (1876–2003). *Journal of sports sciences*, *23*(4), 337–350.

Preston, I., & Szymanski, S. (2000). Racial discrimination in English football. *Scottish Journal of Political Economy*, *47*(4), 342–363.

Ramchandani, G. M., & Wilson, D. J. (2010). Home advantage in the Commonwealth Games. *South African Journal of Sports Medicine*, *22*(1), 8–11.

Reep C, Pollard R, Benjamin B. (1971). Skill and chance in ball games. *Journal of the Royal Statistical Society, Ser. A*, 131, 581–585.

Rosen, S. (1981). The economics of superstars. *The American Economic Review*, *71*(5), 845–858.

Rottenberg, S. (1956). On choice in labor markets. *ILR Review*, *9*(2), 183–199.

Schaubroeck, J., & Williams, S. (1993). Type A behavior pattern and escalating commitment. *Journal of Applied Psychology*, *78*(5), 862.

Schindler, M. (1998). Consequences of perceiving oneself as responsible for obtaining a discount: Evidence for smart-shopper feelings. *Journal of Consumer Psychology*, 7, 371–392.

Schwartz, B., & Barsky, S. F. (1977). The home advantage. *Social Forces*, *55*(3), 641–661.

Scully, G. W. (1974). Pay and performance in major league baseball. *The American Economic Review*, *64*(6), 915–930.

Sloane, P. J. (1971). Scottish journal of political economy: the economics of professional football: The football club as a utility maximiser. *Scottish Journal of Political Economy*, *18*(2), 121–146.

Sloane, P. J. (2006). Rottenberg and the economics of sport after 50 years: An evaluation. IZA Discussion Papers, no. 2,175

Sloane, P. J. (2015). The economics of professional football revisited. *Scottish Journal of Political Economy*, *62*(1), 1–7.

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*, *17*(3), 426–436.

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organisational Behaviour and Human Performance*, *16*(1), 27–44.

Staw, B. M., & Fox, F. V. (1977). Escalation: The determinants of commitment to a chosen course of action. *Human Relations*, *30*(5), 431–450.

Staw, B. M., & Hoang, H. (1995). Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball. *Administrative Science Quarterly*, 474–494.

Staw, B. M., & Ross, J. (1978). Commitment to a policy decision: A multi-theoretical perspective. *Administrative Science Quarterly*, 40–64.

Staw, B. M., & Ross, J. (1980). Commitment in an experimenting society: A study of the attribution of leadership from administrative scenarios. *Journal of Applied Psychology*, *65*(3), 249.

Staw, B. M., & Ross, J. (1987). Knowing when to pull the plug. *Harvard Business Review*, *65*(2), 68–74.

Szymanski, S. (2000). A market test for discrimination in the English Professional Soccer Leagues. *Journal of Political Economy, 108*(3), 590–603.

Szymanski, S., & Zimbalist, A. S. (2006). *National pastime: How Americans play baseball and the rest of the world plays soccer*. Brookings Institution Press.

Terry, P. C., Walrond, N., & Carron, A. V. (1998). The influence of game location on athletes' psychological states. *Journal of Science and Medicine in Sport*, *1*(1), 29–37.

Thaler, R. (2015). *Misbehaving*. Great Britain. Penguin Books.

Volkamer, M. (1971). Zurur aggressivitat in Konkumenz—orientierten sozialen. *Sport Wissenschaft*, *1*, 68–76.

Vroom, V. R. Work and motivation. New York: Wiley, 1964.

Wallace, M. (1988). Labor market structure and salary determination among professional basketball players. *Work and Occupations*, *15*(3), 294–312.

Whyte, G. (1986). Escalating commitment to a course of action: A reinterpretation. *Academy of Management Review*, *11*(2), 311–321.

Wicklund, R., & Brehm, J. (1976). Perspective on cognitive dissonance. Hillsdale. NJ: Erlbaum.

Wilson, D. P., & Ying, Y. H. (2003). Nationality preferences for labour in the international football industry. *Applied Economics*, *35*(14), 1551–1559.