Labour Markets in Professional Sports

Author
Alexander Farnell

Supervisors
Professor Robert Simmons
Dr Vincent O’Sullivan

A thesis submitted toward the degree of Doctor of Philosophy

in the

Department of Economics

Lancaster University Management School

July 2021

Abstract

Labour Markets in Professional Sports

Measuring performance and quantifying outcomes can prove a difficult task in empirical economics research. Because of this, economists have often turned to the setting of professional sports to overcome these data limitations. Sports and sports data presents a unique opportunity to study the behaviour of workers, firms and supervisors, since performance can be accurately measured and compared across agents.

This thesis offers three chapters in the broad fields of labour and personnel economics, using data from professional sports to illustrate. In Chapter One, we consider the role of Head Coaches at football clubs, and whether teams can benefit from Head Coach turnover. This extends on previous work on this topic along several lines. Most notably, Head Coach turnover can either be voluntary or involuntary. In a principal-agent framework, these are theoretically two quite different events, with each producing different predictions about changes to team performance. We also use data from multiple leagues and can distinguish between a short run "bump" effect, and a longer run learning effect. Results show that teams can benefit from Head Coach turnover, particularly following a dismissal, though the result is sensitive to how we define our follow up period.

In Chapter 2, we examine the ability of baseball pitchers to switch between different tasks, by considering how their pitching performance is affected by the additional demands of having to bat and run bases. Despite the prevalence of task switching in modern day work, there is a surprising lack of empirical evidence on its effects on productivity. Baseball is an ideal setting to consider this question empirically, making use of the two-league structure of Major League Baseball. In one league, pitchers are faced with a forced task switching rule of having to both pitch and bat, while in the other, pitchers can focus on their primary job; pitching. The structure of the game of baseball, consisting of innings and a batting order, also means we can cleanly identify cases of workers switching back and forth between tasks. Our results indicate that pitchers can actually benefit from batting, but at all costs should avoid excessive fatigue after running bases.

Finally, in Chapter 3, we return to Coaches, this time in the National Football League. We examine the determinants of coaching changes at the levels of Head Coach and Coordinator.

In particular, we pay close attention to the role of the league's affirmative action policy, the Rooney Rule, on the likelihood of minority coaches being appointed to a Head Coaching role. Results suggest that the rule has been somewhat successful, since teams now appear to be hiring equally skilled black and white coaches, despite evidence that there had always been a supply of equally skilled black coaches.

Acknowledgements

I am sincerely grateful to my supervisors Rob and Vincent for their fantastic support, input and friendship along the way. I'd also like to extend a special thanks to Dakshina and Dave for their comments in the annual reviews, to Andrew and Charlotte for their thoughts and comments on the work, to all my co-authors, Rob, Vincent, Alex, Tunde, Brian and Dave, and to the whole of the Department of Economics at Lancaster, and of course, to my friends and family. This work was funded by the Economic and Social Research Council.

Declaration

1. I hereby declare that, except where explicitly stated, the contents of this thesis are original and have not been submitted for consideration for any other degree or qualification.
2. The work contained in this thesis is original, barring the following respects
   a. An earlier version of Chapter 1 is available as IZA Discussion Paper No.14104 and was submitted to a journal
   b. A version of Chapter 3 is available as a Working Paper at Lancaster University Economics Department
3. The names of institutions of co-authors are listed on the title page preceding each Chapter. Statements regarding the originality of the work and the contributions of co-authors for Chapters 1 and 3 (versions of which are available elsewhere, see note 2) are below
   a. Robert Simmons "C*ontributed to discussions over empirical strategy, model development and interpretation of results"*
   b. Vincent O'Sullivan "*Aside from the regular duties of a PhD supervisor, I contributed to the Stata programming and writing for the chapter about task switching"*
   c. Alex Bryson "*We all contributed to Special Ones in terms of data derivation, analysis and writing"*
   d. Babatunde Buraimo *"Special Ones is a joint effort by all four authors. With regard to Alex Farnell's contribution, his efforts were manifold. He worked on the organizing of the data ensuring that it was 'clean'. He also contributed to the modelling process and created a series of models including the final models that are in the manuscript. Finally, he also made substantive contributions to the final manuscript. In all, Alex's contributions were substantive and instrumental in the creation of the manuscript"*
   e. Dave Berri *"While the idea to look at NFL coaches was my own, Alex constructed and continued to update the data, ran the models and wrote the manuscript"*
4. No conflicts of interest declared.

Note: Here, 'Speical Ones' is referring to Chapter 1 (Head Coach Replacement and Football Team Performance). Special Ones is the title of a paper version of this chapter.

Contents

List of Figures

List of Tables

Head Coach Replacement and Football Team Performance

Assessing the Role of Fatigue and Task Switching on Worker Performance. Evidence from MLB Pitchers

Race and Coaching Hierarchy: An Analysis of Hiring and Firing in the National Football League

Head Coach Replacement and Football Team Performance

Alex Farnell (Lancaster University)
Alex Bryson (UCL, IZA and NIESR)
Babatunde Buraimo (University of Liverpool)
Rob Simmons (Lancaster University)

Abstract

Organisations often spend substantial sums of money and time appointing and incentivising their leaders. As such, one may expect those who lead organisations to affect their performance. Yet there is little evidence establishing a causal link between leaders and organisational performance. Using employer-employee linked data for professional football in four countries over fifteen seasons, we compare the performance of teams after they have sacked their Head Coach with spells where the Head Coach remains in post. We undertake a similar exercise for cases where the Head Coach quits. We deal with the endogeneity of Coach departures using entropy balancing to reweight teams' performance prior to the departure of a Coach so that trends in team performance prior to the departure match spells which ended with a Coach remaining in post. Consistent with theory, Head Coach quits have little or no impact on team performance whereas teams who fire their Head Coach experience small but statistically significant improvements in team performance, although this is sensitive to the way in which our follow up spell is defined. Our results lend support to the proposition that teams can benefit from Head Coach turnover, firing them when it is optimal to do so, and replacing a Head Coach during the offseason.

Keywords: managerial performance; team performance; football; entropy balancing.

JEL Classification: J63, Z21, Z22

Acknowledgements

We thank seminar participants at Lancaster University, University College Cork, the University of Sheffield, University College London, and at the Reading Online Sports Economics Seminars (ROSES) and conference participants at the European Association of Sports Economists, Groningen, the Western Economic Association International, Portland and the Colloquium on Personnel Economics in Zurich for helpful comments. We also thank Andrew McKendrick and Vincent O'Sullivan for their comments.

1. Introduction

Across a range of disciplines there is a strong prior that leaders affect performance. In military history, leaders on the battlefield are credited for victories and blamed for defeats linked to their strategies, while the tactics of leaders of political parties may be called into question following a poor election result. Economists have long maintained that the person who leads an organisation can have a substantial effect on its productivity. This is because the quality of leaders' decision-making and leaders' own productivity can have profound implications for the way the organization is run and thus the productivity of those further down the corporate hierarchy (Rosen, 1990). Lazear et al. (2015) confirm this to be true; an average boss adds roughly 1.75 times more to output than an average worker, with peer effects paling into economic insignificance relative to the effects of bosses.

It has, however, been very difficult to identify a causal impact of managers on performance outcomes because managers are not randomly assigned to organizations and changes in corporate leadership are usually endogenous. For this reason, some analysts have relied on unforeseen death or hospitalisation episodes to identify the effects of leaders on performance. Bennedsen et al. (2012) use hospitalization episodes to identify the effects of CEOs on corporate performance while Besley et al. (2011) use the sudden death of heads of state to establish the importance of leaders' education for growth in countries' gross domestic product.

In this paper we focus on the role of the Head Coach in determining sports teams' performance. Sports Head Coaches can be thought of as holding a reasonably similar role at their club as a CEO or high-ranking manager in more conventional settings. They are both hired by the board of directors and charged with overseeing the day to day running of the organisation, including managing their subordinate workers. According to Pieper et al. (2014), both tend to be on average late 40's or early 50's, work in extremely high-pressure labour markets and are accustomed to dealing with intense media scrutiny.

The role of the Head Coach can vary across sports and even within a sport across countries. But in our setting of professional football (as it is known in Europe, or soccer in North

America), they typically have the power to recruit football players to the squad, appoint their backroom and support staff, pick the team for each game, and decide on match tactics. It seems reasonable to conjecture, therefore, that Head Coaches play a crucial role in determining team performance. Yet the literature finds little evidence of a positive performance effect following a change in Head Coach. This seems somewhat surprising since hiring is costly to firms and club owners should, in principle, have the information required to ensure a good person-job match since weekly football matches provide regular updates on the quality of potential candidates.

Using a large, rich data set on Head Coaches from the top two tiers of four European countries over the seasons 2000/01 to 2014/15, we use entropy balancing to estimate the effects of a change in Head Coach on team performance measured as points achieved in league games played. In a novel step, we distinguish between how the previous coach departed the club. Namely, whether they were dismissed (a forced exit), or whether they quit (a voluntary exit). In contrast to most of the literature, we find some positive effects of a Head Coach change following a dismissal. These occur both in the short run (3-4 games after a change; a "bump" effect) and also in the longer run (10 games and onwards after a change; a learning effect), though these findings are sensitive to the way in which we define our follow up period. We find less support for a positive change in performance after a Head Coach quits, though some longer-term positive effects do emerge. We argue that this is precisely what one would anticipate from theory. Previous studies have not been able to make this distinction between dismissals and quits or, if they have, their sample sizes have been insufficient to provide the necessary statistical power to identify Coach effects.

In Section Two, we provide a theoretical grounding, review the literature on Head Coaches and football team performance, and identify the ways in which this work builds on the existing literature. In Section Three we present our data and estimation techniques. Section Four presents the results before concluding in Section Five.

2. Theory and Empirical Evidence

2.1 Theoretical Background

According to job match theory, a firm will make a hire when the match-specific surplus generated for the firm exceeds the cost of making the hire. This employment relationship will persist until the value of the job match for one or both of the parties falls below the value of an outside option (Farber, 1999). Thus, termination of the contract will occur either through dismissal by the employer (often termed "layoff"), or a quit by the worker. In football, club owners can update their information on Head Coach performance with the results from each game, which may happen once, sometimes twice a week during the football season. This provides them with an opportunity to consider Head Coach performance relative to expectations on an almost continual basis. Such performance evaluations may be harder to carry out in non-sporting settings where principals may only receive reports of executive performance in the annual financial accounts, while monitoring executive performance may prove costly. Football club owners act on this information: Bryson et al. (2021) find that dismissals accounted for over 70 per cent of all Head Coach departures and that the gap between actual team performance and expected performance (captured by betting odds) is a strong predictor of dismissals.

For the football club, the outside option comes in the form of an alternative Head Coach. If Head Coaches are heterogeneous in ability, then teams should be able to replace a departing Coach with a better one. Muehlheusser et al. (2016) confirm that there is substantial heterogeneity in Head Coach ability in the German Bundesliga, and that team performance varies according to the ability of the in-coming Coach. However, there are several reasons why owners may be unable to improve team performance through the recruitment of a new Coach. First, even if Head Coaches are heterogeneous in ability, it may prove difficult for club owners to identify these differences in ability and select the most talented. Their past performance may be attributable to factors other than ability, including luck, meaning it is not possible to infer Coach talent directly from their performances at previous clubs. Second, theory suggests inefficient hiring in talent markets whereby mediocre workers are re-hired in the face of the risk associated with appraising the talent of workers that are new to an industry

(Tervio, 2009). This market failure arises where talent is industry-specific, is only revealed on the job but once revealed, is then public information. More productive firms hire those revealed to be high-ability whereas less productive firms may be forced to experiment with untested new workers. Indeed, a number of potentially highly talented workers may never even be given an opportunity in the first place since they are seen as risky hires. Where there is insufficient discovery of new talent firms tend to re-hire workers who are known to be mediocre. Peeters et al. (2016) confirm that this market failure exists among Head Coaches in professional football in England. Third, it is uncertain just how much of the "talent" Head Coaches possess is generalisable, and how much any success was due to a team-specific effect. If there is a large job-match specific component, performing well at one club may not translate to good performance at another.

For the Head Coach, the outside option comes in the form of alternative employment. Clubs searching for a new Head Coach have three possible options: recruit from the pool of unattached Coaches, promote from within (for example, the youth team coach, or the previous assistant coach), or poach another club's Head Coach. The latter involves a Head Coach quitting their current post to take up their new job, with the recruiting club likely to have to pay a release clause to begin talks. One would assume that higher ability and/or overperforming Head Coaches would be the primary targets for recruiting clubs. However, the effect on the performance of the club losing their Head Coach is unclear. The clubs would not necessarily have planned for this event (unlike a dismissal) and seemingly had no intentions to part ways with the incumbent Coach if the job match were already optimal in the eyes of the club owner. The best the club could do is to recruit a new coach where the quality of the job match is just as good as the previous coach's was. It is therefore unclear, a priori, what impact a Head Coach quit will have on team performance.

2.2 Empirical Evidence

The question of the importance of a Head Coach to a football club is a popular source of debate in academic circles, in the media, and amongst fans. Perhaps one of the most damming pieces of evidence against the importance of the Head Coach is the well-established relationship between a team's wage bill and their finishing position in the league. Kuper and

Szymanski (2009) show that wage bills explain 89% of the variation in final standings in English football (in other words, you get what you pay for). This leaves very little in the way for other factors, including Head Coach ability, to explain team performance. Yet, we can never truly be certain of the importance of a Head Coach, partly because we have never observed football team performance in the absence of a Head Coach.

Since Head Coaches in professional football typically have the power to recruit football players to the squad and backroom support staff, pick the team for each game, and decide on match tactics, it would not be surprising to find that teams who dismiss poorly performing Coaches see performance improve with an in-coming Coach. Yet this is not what is found in most of the literature. In their review of the literature on Head Coaches and football team performance, Van Ours and Van Tuijl (2016) identify eleven studies published since 2000 analysing the period 1993-2010 spanning six countries. None of them identify a positive effect of an incoming Coach following a Coach dismissal.

However, there are some important limitations to the studies reviewed. First, with the exception of Dobson and Goddard (2011), they rely on a small number of Coach dismissal observations, and typically in a single league. Second, they tend to report changes over relatively short periods of time (usually four games) which may be insufficient to pick up performance changes if Head Coaches take some time to "make their mark". This would appear likely given the need to adjust to a new environment, alter the composition of the squad, implement tactics, and hire their own backroom staff. Third, the studies that rely on difference-in-difference estimates do not provide a convincing counterfactual to the dismissal spells.

Van Ours and Van Tuijl (2016) address some of these issues. They deploy a nearest neighbour matching strategy using the gap between team performance and expected performance (using betting odds) to match team spells ending in dismissal, to team spells (for the same team) that experienced similar runs of performance and expected performance but did not switch Head Coach. This strategy offers a much more plausible counterfactual against which to judge the performance effects of an in-coming Head Coach. They find performance improves after a dismissal, but the same improvement is observed in counterfactual cases, leading the authors to conclude that they are simply observing "a regression to the mean phenomenon" (p. 602). However, their study also suffers from small sample sizes, something that particularly affects their ability to estimate models for the subset of cases where Head Coaches quit. They also combine estimates for short and long follow-up spells without identifying the short and long-run effects of a Coach switch.

Not all the literature, however, finds no effect of Head Coach turnover. Using game-level data from 19 seasons of Danish top division football, Madum (2016) also investigates team performance after Head Coach departures using a nearest neighbour matching estimator, matching on recent performance and league ranking of the team and their opponents, but not expected performance derived from betting odds. Madum's findings contrast with most of the literature, uncovering some positive effect of an incoming coach relative to counterfactual scenarios, but the performance only improves in home games. This finding is similar to Tena and Forrest (2007) for Spain although they did not use matching methods.[1] Madum also shows that the effect is apparent only for those teams that fired Coaches (the average treatment-on-the-treated effect) but that the effect would have been absent among the non-treated, a finding that suggests team owners behave optimally when deciding whether to dismiss poorly performing Coaches. Scelles and Llorca (2020) uncover a positive effect of Head Coach change for teams in the French Ligue 1, though in some specifications they also observe an improvement in their counterfactual groups. As such, it is hard to determine which of their models to believe.

However, to say that Head Coaches do not make a great deal of difference to team performance is not the same as saying that the role of a Head Coach is unimportant. The more likely scenario, as pointed out by Goff et al. (2019) is that the person who fills the role of Head Coach is less important. Since the ability distribution of Coaches is so compressed, an incoming Coach will find it very difficult to make even a marginal improvement to performance.

[1] In contrast, Muehlheusser et al. (2016) find performance improvements among German teams are driven by away matches.

Outside of professional football, Goff et al. (2019) report estimates for the effect of both Head Coach changes and General Manager (GM) changes across the National Football League (NFL), Major League Baseball (MLB) and the National Basketball Association (NBA). They find some positive effects for changing a Head Coach, most notably in the NFL where a new Coach contributes between 0.5 and 1.2 extra wins per season (out of a 16 game season), however, their estimates fail to deal with the endogeneity of both Coach and GM departures. Effects in the other two leagues were less pronounced, while a new GM was found to have virtually zero impact on team performance. It is also worth noting that the role of the Head Coach extends beyond team performance. Player development is one such example. Yet, Bradbury (2017), for the case of MLB, and Berri et al. (2009) for the case of the NBA also find that Coaches have little impact on individual player performances, at least on average. However, Bradbury reports that hiring a new Coach is associated with gains in attendance of up to 1000 spectators per game.

2.3 Contributions

Our estimates differ somewhat from those in the literature in several respects. Most importantly, because our data are large enough, we can be confident in identifying even quite small effects, not only for dismissals, but also for quits on changes to team performance. To the best of our knowledge, we are the first paper that can make this distinction between these two theoretically different events. Quits are decisions made by agents, rather than principals, so they may be less likely to lead to improvements in team performance, at least in the longer term, since the principal was otherwise happy to keep the incumbent Coach. Second, we estimate performance outcomes over a longer period (20 games, roughly half a season) of time to establish whether any effects of a Coach change differ in the short and longer term. Below, we argue how and why these effects may differ. Third, we examine whether any performance changes differ if the Coach change occurs within season or between seasons. We also use entropy balancing to construct counterfactual spells to those ending in quits or dismissals.

We test the hypothesis that fires result in performance improvements, notwithstanding the caveats outlined above, but quits are less likely to do so. Because we track Head Coaches over long periods of time, we are able to compare and contrast short-run and longer-run performance effects, as well as effects across seasons. This distinction between shorter- and longer-run impacts is important in picking up quite separate effects of Head Coach changes on team performance. The short-run effect is the "bump" in performance that is attributable to simply making a change. There are two aspects to this. The first, often referred to by football pundits is the motivational impact of a new Coach on current players who seek to impress the new Coach to cement their place in the team. The second element that might have an immediate impact on performance is simply the fact of having made a change. Levitt (2021) finds there are happiness benefits of making life-changing decisions when determined by the toss of a coin - that is, even when the decision is made based on a random event. Analogously, it seems reasonable to assume that a simple change in Coach, regardless of the in-coming Coach's quality or the circumstances surrounding his appointment (i.e. following a quit or dismissal of the previous Coach), may result in improvements in team performance. Then again, if instead we take the view that a newly appointed Coach who follows a quit is likely to be taking charge of a better performing team (as opposed to a Coach who takes over following a dismissal), this bump may be less pronounced than after a dismissal. So often we hear media reports of players downing tools just before a Coach is dismissed, in which case a new Coach may have a greater motivational impact after a dismissal.

The longer-run impact of a change in Head Coach is likely to be a two-way process. On the one hand, Coaches will learn about the club, their new players, and the expectations and orientation of the owners. Coaches will also be able to sell unwanted players and recruit new ones via the transfer market, though this can only occur during limited windows during the year, a three-month period over summer, and during the month of January. Recent studies emphasise the importance of on-the-job learning for individual worker productivity (e.g. Gaynor et al. (2005) in the health economics literature), and especially among new hires (de Grip, 2015).[2] The other side of this process is players learning about their new Coach. Players

[2] Perhaps the most successful football club manager of all time, Sir Alex Ferguson, described the time it took to "build a club" (https://hbr.org/2013/10/fergusons-formula). Yet he was not successful in his early years as he recalled in his autobiography: "After the farewell in May 2013, the pivotal moments filled my thoughts. Winning that FA Cup third-round tie against Nottingham Forest in January 1990, in which a Mark Robins goal sent us on our way to the final when my job was supposedly on the line. Without the FA Cup [final] victory over Crystal Palace nearly four years after my arrival, grave doubts would have been raised about my

may take time to adapt to new training regimes, fitness regimes, changing tactics and systems, and in some cases the new Coach may oversee changes to lifestyles and diets. We look directly at time-variance in any performance effects after a Coaching change.

3. Data and Empirical Approach

Our data consist of all games from the top two divisions of four major European football leagues (France, Germany, Italy and Spain) over the period 2000/01 to 2014/15 for which we can precisely ascertain the start and end dates of managerial spells.[3] This period covers 273 teams, with 769 individual Coaches taking charge of games for those teams. Coaching tenures were hand-collected from *Wikipedia*, supported by online newspaper sources from each country. In line with literature such as van Ours and van Tuijl (2016), we exclude caretaker spells where an interim Coach took over management of a team prior to a permanent appointment. It could be that an interim candidate performs well enough to be given the job on a full-time basis; in this case we only consider the date from when they were permanently appointed. In aggregate, we have 1,327 fires and 533 quits, which on average lasted for 35 (std. dev. = 31) and 60 (std. dev. = 51) games respectively.

Dismissals and quits are recorded according to the Head Coach's *Wikipedia* biography and are cross checked with local media sources. If the end of a coaching spell is described as "mutual consent", we classify these cases as a dismissal. In reality, these are circumstances where the coach has been asked to leave, but the club will announce it as a joint decision, allowing the coach to "save face", such that any potential employers are not put off by their apparent dismissal. Table 1 shows the number of dismissals and quits per season, aggregated over the leagues in our data. Dismissals exceed quits and there appears to be a rising trend in

suitability for the job. We will never know how close I was to being sacked, because the decision was never forced on the United board. But without that triumph at Wembley, the crowds would have shrivelled. Disaffection might have swept the club" (Ferguson, 2013).

[3] We exclude the English leagues from our analysis since many teams in England operate with a Manager rather than a Head Coach. Typically, a Manager will be involved in the same roles as a Head Coach (coaching the team, picking the matchday squads, motivating players etc.) with the added responsibility of recruitment and overseeing the progression of youth players into the senior team. In European football, teams now typically operate with a Head Coach and a Director of Football (who plays a similar role to the General Manager is North American sports teams) who takes on the other responsibilities, typically with input from the Head Coach.

both dismissals and quits. The increased firing rate may be a consequence of growing revenue differences between league positions generally in European football.[4] This increase in reward for success was proposed by d'Addona and Kind (2014) as an explanation for increased Head Coach turnover in English football in their study covering the post-war period up to 2008.

Table 1: Frequency of Exits (by type) per season

Season	**Dismissals**	**Quits**
2000-01	70	20
2001-02	61	34
2002-03	63	26
2003-04	71	43
2004-05	63	36
2005-06	74	31
2006-07	79	39
2007-08	69	34
2008-09	95	44
2009-10	112	29
2010-11	99	38
2011-12	111	39
2012-13	99	36
2013-14	110	28
2014-15	151	56
Total	1327	533

Figure 1 shows the timing of dismissals and quits respectively as the season progresses. Time lapsed is measured monthly (as opposed to say, number of games) since the different countries and different tiers within a country have different season lengths.[5] There are large spikes in Coach departures at the end of the season (usually May, though a season occasionally extends into June). This makes sense on several counts. The off season is a period with no games other than pre-season friendlies and coincides with the summer transfer

[4] Prominent amongst the sources of revenue differences between league positions is the growth of UEFA Champions' League revenues for the top three or four teams that qualify for this competition from our four sample Leagues. These revenues have grown substantially over time prompting increased investment in playing squads by aspiring teams (Green et al., 2015).

[5] The number of teams per leagues per season varies between 18 and 24, meaning season length varies between 34 and 46 games. Due to restructuring of leagues, bankruptcy and or disqualification of clubs, season length may vary from year to year.

window. Together, these give a new appointment the best opportunity to work with their new squad and implement any changes they deem necessary. This could entail working with the current squad of players, honing their skills, developing a playing style, and making use of the transfer market to recruit new players to the team. Moreover, the off-season is when many Head Coach contracts expire or are reviewed by the board of directors, so teams wishing to dismiss their Coach may find it best to wait until contract expiry, rather than sacking mid-season which may require a substantial severance payment to the Coach.

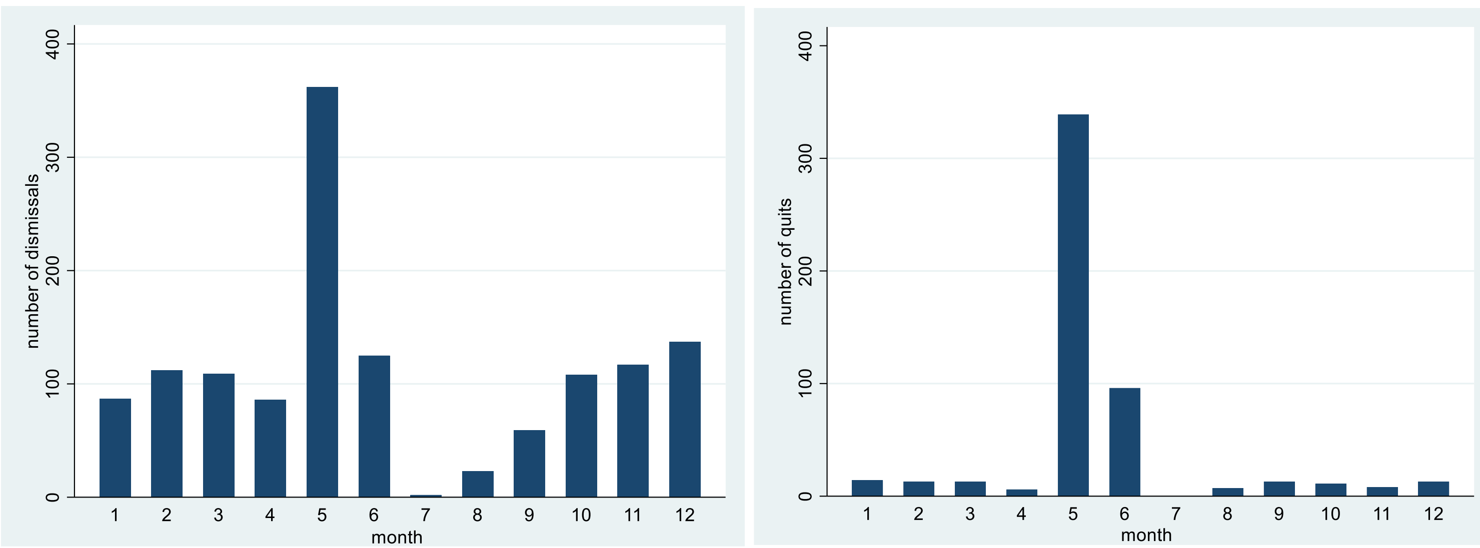


Figure 1: Frequency of Coach exits by month (1=January, 12=December)

During the season dismissals tend to peak in mid-season when some leagues have a winter break. Quits on the other hand show little pattern over time. It appears that many clubs reassess their prospects during the winter break and are more likely to fire their Head Coaches at this juncture than at other points in the season. Importantly for our analysis, the two histograms give a preliminary suggestion that the statistical processes driving Head Coach dismissals and quits could well be different.

Figure 2 shows average team performance before and after Coach changes, with dismissals and quits considered separately. We assess team performance across the whole sample, up to 20 games before a Coach change and up to 20 games after the change, with team performance being measured as Mean Points Per Game. 20 games was chosen as this is approximately half a season, though for smaller leagues with shorter seasons this will be a little longer than half a season. The blue line refers to performance during a quit spell, and the red line refers to performance during a dismissal spell.

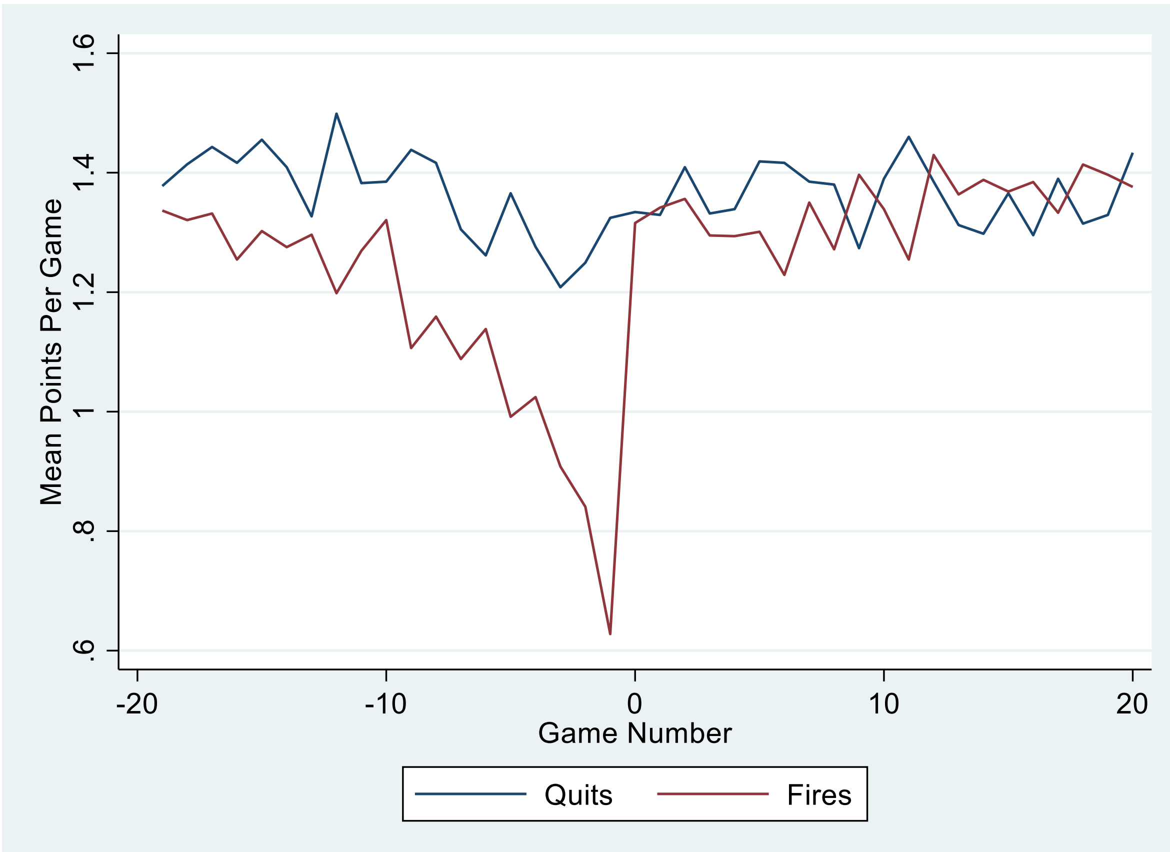


Figure 2: Points Per Game before and after a Coaching change

Prior to dismissals, team performance drops as indicated by the decline in the line representing fires as game number zero approaches. This is akin to the Ashenfelter Dip, something one needs to be mindful of when making over-time comparisons before-and-after Head Coach dismissals (Bruinshoofd and ter Weel, 2003).[6] The slight disparity between our setting of football teams and of Ashenfelter's work on participants in job training programmes, is that every team in our sample experiences a treatment at some point in time.

[6] The Ashenfelter Dip, first observed by Orley Ashenfelter (1978), describes the drop in the earnings of participants in job training programs in the year before entry. Thus, a simple before and after comparison of the effect of job training programs on earnings is likely to be overestimate the true effect.

Post-dismissal team performance recovers and stabilises at a level close to that for the pre-period. In contrast, there is less evidence of a dip in performance prior to quits, nor much of a change in performance after a quit.

As de Paulo and Scoppa (2012) and van Ours and van Tuijl (2016) argue, the recovery in team performance following a Coach firing could simply be the result of regression to the mean. The key question that we address below in more formal regression analysis is whether we can discern any causal impact of Head Coach turnover on team performance after accounting for the endogeneity of Head Coach change and other confounding factors.

Our empirical approach begins by specifying a naïve OLS regression as follows:

$$Y_{ijk} = X'_{ijk}\alpha + \beta d_{ijk} + \epsilon_{ijk} \qquad (1)$$

where the subscripts are denoted as i for team, j for game and k for season. This is our outcome model, where the dependent variable, Y_{ijk}, is points per game: teams get three points for a win, one for a draw and none for a defeat. We run models for points per game obtained for spells of the next single game through to longer outcome spells of up to 20 games. Match results and betting odds (which we make use of later) were provided by www.football-data.co.uk. d_{ijk} is our main variable of interest; a dummy variable to indicate whether there has been a coach change. Because we have two possible types of exit (quit or dismissal), we run the above specification twice to account for this, taking out coaching tenures that end in the other type of exit (i.e. we drop spells that end in a quit when analysing dismissals and vice versa). Naturally, our test that a coach change has a positive effect on performance is then a *t*-test of the null of $\beta = 0$ in equation (1). X_{ijk} is a vector of control variables which includes information on previous team performance, captured by points per game over the previous 10 fixtures, and performance relative to expected performance (called Surprise, described below). We also include opposition form, measured by the opponent's league positions, and home advantage, measured by the proportion of home games over the follow up period. To complete (1), ε_{ijk} is a random error term. Throughout our estimations, standard errors are clustered at the team level.

Following van Ours and van Tuijl (2016) we incorporate a measure of Surprise which is the difference between actual and expected performance. Performance above or below or expectations in any given match, or indeed across multiple games are likely to affect future performance.[7] Expected points in a given match is computed as:

$$E(Points) = (3 * Prob.of\ Win) + Prob.of\ Draw \qquad (2)$$

where the probabilities are derived from bookmakers' betting odds. When converting odds to probabilities, the sum of the probabilities tends to sum to greater than one. This is the overround and is how bookmakers make a profit. Therefore, the probability of each outcome is weighted by the sum of the probabilities of winning, drawing and losing the match. Surprise is then actual points minus expected points. Naturally, a Surprise value of 0 indicates that a team performed as expected in a given match, with this being reflected by the betting market. We include Surprise in the most recent game, cumulative (total) surprise over games lagged two to five and cumulative (total) surprise over games lagged six to ten to capture any longer runs of good or bad form.

The difficulty in relying on OLS estimation of Head Coach changes on team performance is that Head Coach changes are not random, posing a challenge for causal inference. Exposure to treatment (Head Coach turnover) may be related to certain covariates, which too are related to outcomes. Indeed, Head Coach changes are likely to be endogenous with respect to team performance. To put this another way, it is likely that only the poor or underperforming teams sack their coach, as is apparent in Figure 2. Consequently, we cannot infer what would have happened to a team's performance in the absence of a Head Coach change by comparing the performance of teams that did and did not make a change. De Paolo and Scoppa (2012), Van Ours and Van Tuijl (2016) and Besters et al. (2016) found positive and significant effects of Head Coach dismissals on team performance for Italian, Dutch and English football,

[7] As well as predicting future outcomes, Surprise is a determinant of Head Coach turnover. This is a point we come to during our discussions on covariate balancing.

respectively, from naïve OLS estimates only for these effects to become statistically insignificant when they compared performance with a matched comparator group.

We adopt a quite different approach to previous literature to obtain the causal impact of Head Coach changes on team performance, namely Entropy Balancing (Hainmueller, 2012), implemented by the Stata command *ebalance* (Hainmueller and Zu, 2013). This is a data pre-processing method that reweights each observation in the control group to achieve covariate balance, such that the mean, variance, and skewness (and theoretically higher moments) of the variables are equal across the treatment and control groups. The weights are chosen such that a loss function, describing the dissimilarity between the control and treatment variable distributions, is minimised. As such, we can think about Head Coach departures mimicking a random process and any selection into treatment is stripped out of the outcome equation (1). These weights are then simply used in a weighted version of the OLS regression described in (1).

More formally, following the notation in Hainmueller and Zu (2013), entropy balancing can be thought of as a generalised propensity score weighting approach to form a counterfactual mean as follows:

$$\boldsymbol{E}\big(\widehat{Y(0) \mid D} = 1\big) = \frac{\sum_{\{i|D=0\}} Y_i w_i}{\sum_{\{i|D=0} w_i} \qquad (3)$$

The left hand side of equation (3) is the estimated outcome of interest in the control group, Y(0), conditional on receiving the treatment D=1. w_i is the entropy balancing weight chosen for each observation in the control sample. These weights are chosen according to the following scheme to minimise the entropy distance metric, defined by a loss function *h(.)*, describing the dissimilarity between two distributions.

$$\min_{w_i} H(w) = \sum_{i|D=0} h(w_i) \qquad (4)$$

Subject to the following constraints

$$\sum_{i|D=0} w_i c_{ri}(X_i) = m_r \qquad r \in 1, \ldots, R \tag{5}$$

$$\sum_{i|D=0} w_i = 1 \tag{6}$$

$$w_i \geq 0 \qquad for\ all\ i\ such\ that\ D = 0 \tag{7}$$

In equation (5), $c_{ri}(X_i)=m_r$ describes a set of R balance constraints imposed on the covariate moments in the reweighted control group. The second constraint is arbitrary, and the weights could sum to any constant. In the Stata procedure, R is set to 3 meaning we balance covariates on their mean, variance and skewness. Tables 2 and 3 show these moments of our treatment and control groups before and after applying our entropy balancing weights. After applying these weights, the moments of the distribution of the variables in the weighted control group are almost identical to our treatment group. Entropy Balancing has several advantages, both in a practical and an econometric sense, over more conventional weighting methods (such as Inverse Probability Weighting Regression Adjustment). From the researcher's point of view, the scheme removes the need for the continual iterative process of running a propensity score model and checking for covariate balance, not to mention the concern of mis-specifying the treatment model. Zhao and Percival (2017) also show that entropy balancing possesses the attractive property of being doubly robust, even though no treatment model is estimated, while also producing treatment effects that are within the range of observed outcomes.

The covariates we balance on are all variables that, at least in theory, should predict Head Coach departures. We follow Bryson et al. (2021) in our selection of covariates that affect departures. These covariates capture a combination of team form, coaching characteristics, and season progress. See Table 4 displaying the results of a multinomial logistic regression, analysing how these variables predict both quits and dismissals. Team form variables include mean points per game over the last 10 games, league position (where position is captured as rank across both tiers per country) and the final league position of the team in the previous season. Since owners' (and stakeholders) expectations about performance (as well as actual performance) are likely to play a role in coaching departures, we also include the lagged cumulative Surprise variables as discussed earlier. Should performance slip below some acceptable level in the eyes of the principal, which will include knowledge about opponent quality, then the team may look to replace the Head Coach (van Ours and van Tuijl, 2016). A negative Surprise value is a likely signal of a poorly performing Head Coach, and this is indeed a strong predictor of exits, with more recent runs of form appearing to matter more.

Our measures of Head Coach characteristics can be broadly split into team specific experience (affecting the quality of the job match only at the current employer), and general experience (which all teams will find of value). Our team specific measures are tenure at the current team (measured as number of games since appointment), and dummy variables indicating whether the coach was an ex-player, and whether they were an internal appointment. Increased tenure reduces the likelihood of dismissal but increases the likelihood of quitting. Ex-players and internal appointments are significantly less likely to quit (perhaps representing some emotional attachment to the club), while ex-players are also significantly less likely to be sacked, which could indicate fans tend to be more patient with ex-players, placing less pressure on the board to make a change. General measures of experience include years of experience (years since first coaching job), age and its square, the number of previous Head Coach spells, and dummy variables capturing previous successes and failures as a coach (previous promotions, previous cup winners, and a previous relegation). Age is not significantly related to dismissals, but younger coaches are much more likely to quit. Younger coaches may be more attractive to hiring clubs since it represents an opportunity for a long working relationship. Hiring an older coach may present risks such as possible retirement and concerns over possible deteriorating health, which both would detract from the job-match surplus. Previous successes as a Head Coach, along with years of experience, appear to have some protective effect against dismissals, even controlling for team performance. Finally, our measures of season progress (in line with Figure 1) include the proportion of games remaining (again, proportion is to account for differences in season length) and whether the departure occurred after the last game of the season. As the season progresses, teams are more likely to dismiss their Coach, though this occurs at a diminishing rate. At some point, a new appointment will not have enough games left in the season to make a difference, so teams just wait until the season ends. Season progress does not affect quits. Descriptive statistics of our covariates and selected outcomes are shown in Table 5.

Table 2: Entropy Balancing Moments (Dismissals)

	Mean			Variance			Skewness		
	Unweighted Treatment	**Unweighted Control**	**Weighted Control**	**Unweighted Treatment**	**Unweighted Control**	**Weighted Control**	**Unweighted Treatment**	**Unweighted Control**	**Weighted Control**
Surprise t-1	-0.4482	0.02179	-0.4481	1.076	1.438	1.075	0.9632	0.251	0.9632
Surprise t-2 to t-5	-1.097	0.08274	-1.097	4.764	5.618	4.763	0.4275	0.1144	0.4274
Surprise t-6 to t-10	-0.7296	0.09391	-0.7296	5.765	7.024	5.764	0.3676	0.09937	0.3675
Mean Points Prev 10 Games	1.069	1.401	1.069	0.1722	0.228	0.1722	0.5912	0.2315	0.5919
Position	17.1	19.99	17.09	162.7	156.6	162.7	0.2108	0.2828	0.2119
Position Squared	454.8	556	454.8	238663	315283	238643	0.954	0.8982	0.9542
Last Season Position	30.08	28.36	30.08	516.1	477.8	516.1	0.1102	0.2905	0.1105
Tenure	38.85	44.61	38.85	1126	2225	1125	1.903	2.696	1.903
Experience	11.46	11.47	11.45	63.38	59.25	63.37	0.7147	0.7602	0.7151
Age	49.22	48.42	49.21	45.65	43.21	45.65	0.3663	0.3959	0.3677
Age Squared	2468	2388	2468	468590	430630	468551	0.7298	0.754	0.731
N Prev HC Jobs	4.877	4.385	4.876	17.82	15.06	17.82	1.16	1.283	1.16
Internal Appointment	0.1367	0.1386	0.1369	0.1181	0.1194	0.1181	2.115	2.092	2.113
Previous Promotion	0.4981	0.5254	0.4981	0.2502	0.2494	0.25	0.007648	-0.1016	0.007628
Previous Cup	0.1396	0.1959	0.1398	0.1202	0.1575	0.1202	2.08	1.532	2.078
Previous Relegation	0.3184	0.2678	0.3186	0.2172	0.1961	0.2171	0.7799	1.049	0.7785
Ex Player	0.1338	0.1604	0.134	0.116	0.1347	0.1161	2.151	1.851	2.149
Proportion of Games Remaining	0.3355	0.489	0.3354	0.09008	0.07989	0.09007	0.2687	0.01582	0.2694
Proportion of Games Remaining Squared	0.2025	0.319	0.2026	0.0559	0.08237	0.05589	1.044	0.6702	1.043
Last Game of Season	0.3193	0.01488	0.3196	0.2176	0.01465	0.2175	0.7751	8.015	0.7738

Table 3: Entropy Balancing Moments (Quits)

	Mean			Variance			Skewness		
	Unweighted Treatment	**Unweighted Control**	**Weighted Control**	**Unweighted Treatment**	**Unweighted Control**	**Weighted Control**	**Unweighted Treatment**	**Unweighted Control**	**Weighted Control**
Surprise t-1	-0.08757	0.02179	-0.08753	1.379	1.438	1.378	0.415	0.251	0.4149
Surprise t-2 to t-5	-0.3931	0.08274	-0.3929	4.969	5.618	4.968	0.2405	0.1144	0.2404
Surprise t-6 to t-10	-0.08577	0.09391	-0.08579	7.511	7.024	7.508	0.2257	0.09937	0.2257
Mean Points Prev 10 Games	1.319	1.401	1.319	0.2627	0.228	0.2626	0.4913	0.2315	0.4937
Position	18.06	19.99	18.05	156.4	156.6	156.4	0.4881	0.2828	0.4909
Position Squared	482.2	556	482.1	296922	315283	296846	1.095	0.8982	1.096
Last Season Position	27.29	28.36	27.29	468.7	477.8	468.5	0.4341	0.2905	0.4353
Tenure	64.29	44.61	64.26	2925	2225	2924	2.58	2.696	2.581
Experience	12.58	11.47	12.58	71.22	59.25	71.2	0.826	0.7602	0.8274
Age	49.44	48.42	49.43	51.95	43.21	51.94	0.5006	0.3959	0.5055
Age Squared	2496	2388	2495	549883	430630	549779	0.8332	0.754	0.8376
N Prev HC Jobs	4.797	4.385	4.795	17.23	15.06	17.23	1.202	1.283	1.203
Internal Appointment	0.1038	0.1386	0.1041	0.09327	0.1194	0.09328	2.597	2.092	2.592
Previous Promotion	0.5598	0.5254	0.5595	0.247	0.2494	0.2465	-0.241	-0.1016	-0.2396
Previous Cup	0.2506	0.1959	0.2512	0.1882	0.1575	0.1881	1.151	1.532	1.148
Previous Relegation	0.2415	0.2678	0.2421	0.1836	0.1961	0.1835	1.208	1.049	1.204
Ex Player	0.1264	0.1604	0.1268	0.1107	0.1347	0.1107	2.248	1.851	2.243
Proportion of Games Remaining	0.1039	0.489	0.1037	0.05434	0.07989	0.05444	2.195	0.01582	2.2
Proportion of Games Remaining Squared	0.06501	0.319	0.06519	0.03024	0.08237	0.03035	3.003	0.6702	3.001
Last Game of Season	0.772	0.01488	0.771	0.1764	0.01465	0.1765	-1.297	8.015	-1.29

Table 4: Multinomial Logistic Regression, Determinants of Exits

VARIABLES	Dismissal	Quit
Team Performance		
Surprise t-1	-0.312***	-0.123**
	(0.035)	(0.051)
Surprise t-2 to t-5	-0.155***	-0.130***
	(0.024)	(0.036)
Surprise t-6 to t-10	-0.024	-0.035
	(0.024)	(0.035)
Mean Points Prev 10 Games	-0.957***	-0.271
	(0.189)	(0.264)
Position	-0.031***	-0.048***
	(0.011)	(0.019)
Position Squared	0.001**	0.001***
	(0.000)	(0.000)
Last Season Position	0.002	-0.000
	(0.002)	(0.003)
Coach Characteristics		
Tenure	-0.002**	0.006***
	(0.001)	(0.001)
Experience	-0.039***	-0.019
	(0.008)	(0.014)
Age	-0.007	-0.302***
	(0.057)	(0.087)
Age Squared	0.000	0.003***
	(0.001)	(0.001)
N Prev HC Jobs	0.050***	0.041*
	(0.014)	(0.022)
Internal Appointment	0.025	-0.428**
	(0.115)	(0.198)
Previous Promotion	-0.265***	0.048
	(0.076)	(0.122)
Previous Cup	-0.226**	0.275*
	(0.103)	(0.143)
Previous Relegation	0.196**	-0.266**
	(0.079)	(0.134)
Ex Player	-0.218**	-0.486***
	(0.109)	(0.181)
Season Progress		
Proportion of Games Remaining	4.983***	-0.777
	(0.652)	(1.419)
Proportion of Games Remaining Squared	-5.059***	0.373
	(0.638)	(1.425)
Last Game of Season	4.638***	5.270***
	(0.166)	(0.299)
Observations	66,157	66,157

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Outcomes *					
Mean Points per game Next 1 Game	65,998	1.391	1.293	0	3
Mean Points per game Next 5 Games	65,339	1.390	0.620	0	3
Mean Points per game Next 10 Games	64,494	1.391	0.481	0	3
Mean Points per game Next 15 Games	63,626	1.391	0.423	0.133	3
Mean Points per game Next 20 Games	62,751	1.391	0.390	0.150	2.900
Team Performance					
Surprise t-1	66,157	0.014	1.198	-2.707	2.797
Surprise t-2 to t-5	66,157	0.061	2.371	-8.269	8.277
Surprise t-6 to t-10	66,157	0.080	2.649	-9.681	9.760
Mean Points per game Prev 10 Games	66,157	1.395	0.479	0	3
Position	66,157	19.927	12.524	1	48
Last Season Position	66,157	28.377	21.872	1	66
Coach Characteristics					
Tenure (n games)	66,157	44.653	47.062	1	441
Experience (years)	66,157	11.475	7.707	0	44
Age	66,157	48.439	6.582	30.212	73.739
N Prev HC Jobs	66,157	4.395	3.888	0	23
Previous Promotion	66,157	0.525	0.499	0	1
Previous Cup	66,157	0.195	0.397	0	1
Previous Relegation	66,157	0.268	0.443	0	1
Internal	66,157	0.138	0.345	0	1
Ex Player	66,157	0.160	0.366	0	1
Season Progress					
Proportion of Games Remaining	66,157	0.484	0.285	0	0.978
Last Game of Season	66,157	0.025	0.155	0	1

* Note: The number of observations for our outcome variables decreases as we expand on the number of games for our follow up spell because our sample period ends at the 14/15 season, and so do not observe games at the start of the 15/16 season.

Our preferred variants of the entropy balanced models include team fixed effects, thus focusing on comparisons of team performance within team over time. In doing so we avoid biases in estimates of Head Coach departures arising from fixed unobservable differences across teams. We also have models which include Season Fixed Effects in models (that is, considering estimations controlling for fixed unobservable differences across different seasons). Our baseline models compare spells ending in either a Head Coach quit or dismissal (at time t=0), relative to counterfactual spells which did not end in a Head Coach departure, where we follow performance across a further 20 game period (t=1 to t=20), regardless of whether there are subsequent Head Coach changes in the period after t=0. It is arguable that football results should count when estimating the impact of a Coach dismissal or quit, even if there is subsequent Coach turnover in the outcome spell. In a later analysis, we restrict our analyses to spells of games where no subsequent Head Coach change occurs. This facilitates an assessment of the long-term performance of the initial Head Coach change, where that performance is permitted to develop. However, it is also arguable that in dropping spells with a subsequent Head Coach change, we are truncating the sample based on a potentially endogenous variable i.e. whether team owners choose to retain the Coach for another 20 games, since this will partly reflect how well the new Head coach is performing during that period. Indeed, a simple probit regression reveals that good performance (both absolute and relative to expectations), whether the team is promoted or relegated and the proportion of games remaining are all strong predictors of such spells.

Unlike most other studies which confine analysis to within-season changes in Head Coach, we allow team performance history to straddle seasons and we also include between-season Coach changes in our analysis. We test the sensitivity of this by running analyses with and without closed season Coach changes, as it is possible that those changes that occur in the closed season are qualitatively different to those that occur within season. For example, they may include a larger number of contract non-renewals. There are not enough within season quits to extract any meaningful results, hence our soul focus here will be on dismissals. Of the 1,327 total dismissals, 883 occurred within season, and the remaining 444 occurred during the offseason. Our spells of games may also include instances of teams being promoted and relegated. Of course, a win is worth three points and a draw one point in both the first and second tier, but due to the quality of opponents in the top tier, a win is likely harder to come by for promoted teams. Equally, a relegated team should find it easier to pick up points in the lower tier. This could play on owner’s expectations, and so we check the robustness of our results by taking out teams who are playing in their first season after a promotion or relegation.

4. Results

4.1 Baseline OLS Estimates

Tables 6 and 7 display our (unweighted) OLS estimates for Dismissals and Quits respectively. Table 8 shows the effect of our covariates for selected spells of 1, 5, 10 and 20 games. All the controls work in the way we would expect them too. According to the OLS estimates, there does not appear to be any benefit to teams, at least not over any sustained run of games from Head Coach turnover of either kind. It is possible that we some short term "bump" effect following a dismissal, with an additional 0.036-0.04 points per game three games into the new Coach's tenure, though this is only significant at 10%.

4.2 Entropy Balance Models

Next, we augment our baseline OLS models with the entropy balancing weights as specified by the Stata routine. Table 9 displays the results for dismissals, while Table 10 displays the results for quits. Both sets of results suggest that team performance does not significantly improve for any sustained run of games following either a dismissal or a quit. The effect of including of team fixed effects is to reduce the magnitude of the point estimates. Including team fixed effects means we are relying on spells of games within team to obtain our counterfactual spells. If these omitted differences are correlated with the tendency to change Coaches, then the estimates without team fixed effects will be biased, with the team fixed effects soaking up a great deal of the across team differences. In practical terms, any positive effects of a coaching change may be limited to a select number of teams. Results with season fixed effects are nearly identical to when they are not included in the models.

Of course, these are all average effects, but within that average will lie a range of outcomes, with some teams benefitting from changing Coach, others will indeed experience no effects, while others will likely suffer. These results lend some initial support to two lines of argument discussed in the Literature Review. One is the notion of mediocre talent, as laid out by Peeters et al. (2016). The other, as pointed out by Goff et al. (2019), is that all Head

Coaches are remarkably similar in terms of their ability. Whichever argument we take, that Head Coaches struggle to make much of a difference is consistent with both these views.

So, why change Head Coach? Even if teams know the average effect is negligible, they may be attracted by the small probability of a successful Coach change. On the other hand, this zero average effect could be consistent with the scapegoat hypothesis of fan disgruntlement and pressure (e.g. Tena and Forrest, 2007), in that a change is made simply to appease disgruntled fans, even though performance is unlikely to improve. Nevertheless, we feel that jumping to the conclusion of ‘Head Coaches make no difference’ on the basis of these results is possibly a little short sighted given our theoretical discussion outlined in Section 2. Hence, we offer alternative specifications to look at this in more depth, and outline circumstances where an incoming coach could make a difference to team performance.

Table 6: Unweighted OLS Estimates (Dismissals)

	no FE		**team FE**		**season FE**		**N**	**Adj. R2**		
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	Season FE
1	-0.007	(0.037)	0.002	(0.038)	-0.007	(0.037)	74,718	0.107	0.118	0.107
2	0.035	(0.027)	0.038	(0.027)	0.035	(0.027)	74,516	0.082	0.089	0.082
3	0.040*	(0.022)	0.036*	(0.022)	0.040*	(0.022)	74,311	0.110	0.118	0.110
4	0.030	(0.019)	0.023	(0.019)	0.031	(0.019)	74,106	0.121	0.134	0.122
5	0.012	(0.017)	0.001	(0.017)	0.012	(0.017)	73,901	0.141	0.159	0.141
6	0.005	(0.016)	-0.008	(0.015)	0.005	(0.016)	73,697	0.154	0.179	0.155
7	-0.001	(0.015)	-0.016	(0.015)	-0.001	(0.015)	73,493	0.170	0.200	0.170
8	0.002	(0.014)	-0.014	(0.014)	0.002	(0.014)	73,286	0.183	0.219	0.183
9	-0.006	(0.014)	-0.022*	(0.013)	-0.005	(0.014)	73,078	0.196	0.238	0.196
10	0.002	(0.013)	-0.015	(0.012)	0.003	(0.013)	72,869	0.208	0.256	0.208
11	0.005	(0.012)	-0.013	(0.012)	0.005	(0.012)	72,661	0.218	0.272	0.219
12	0.002	(0.012)	-0.018	(0.011)	0.002	(0.012)	72,454	0.228	0.287	0.229
13	0.007	(0.012)	-0.012	(0.011)	0.008	(0.012)	72,246	0.237	0.301	0.238
14	0.009	(0.011)	-0.011	(0.010)	0.010	(0.011)	72,039	0.245	0.314	0.245
15	0.013	(0.012)	-0.008	(0.010)	0.014	(0.012)	71,831	0.251	0.326	0.251
16	0.014	(0.011)	-0.007	(0.010)	0.015	(0.011)	71,624	0.257	0.337	0.257
17	0.019	(0.011)	-0.004	(0.010)	0.019*	(0.011)	71,418	0.261	0.348	0.262
18	0.020*	(0.011)	-0.003	(0.010)	0.020*	(0.011)	71,213	0.265	0.357	0.266
19	0.026**	(0.011)	0.003	(0.010)	0.026**	(0.011)	71,007	0.268	0.365	0.269
20	0.026**	(0.011)	0.002	(0.010)	0.026**	(0.011)	70,803	0.270	0.373	0.271

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table 7: Unweighted OLS Estimates (Quits)

	no FE		**team FE**		**season FE**		**N**	**Adj. R2**		
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	Season FE
1	-0.058	(0.047)	-0.077*	(0.046)	-0.059	(0.047)	74,038	0.107	0.118	0.107
2	-0.045	(0.039)	-0.050	(0.039)	-0.046	(0.039)	73,840	0.082	0.089	0.082
3	-0.025	(0.036)	-0.028	(0.036)	-0.026	(0.036)	73,637	0.111	0.119	0.111
4	-0.022	(0.030)	-0.024	(0.030)	-0.023	(0.030)	73,433	0.122	0.134	0.122
5	-0.020	(0.028)	-0.022	(0.027)	-0.021	(0.028)	73,229	0.142	0.159	0.142
6	-0.003	(0.026)	-0.004	(0.026)	-0.004	(0.026)	73,027	0.155	0.179	0.155
7	0.004	(0.026)	0.003	(0.025)	0.003	(0.026)	72,827	0.170	0.200	0.170
8	-0.001	(0.026)	-0.002	(0.026)	-0.001	(0.026)	72,623	0.183	0.220	0.183
9	-0.001	(0.024)	-0.003	(0.024)	-0.002	(0.024)	72,419	0.196	0.239	0.196
10	-0.010	(0.024)	-0.013	(0.023)	-0.011	(0.024)	72,217	0.208	0.256	0.208
11	-0.013	(0.022)	-0.016	(0.021)	-0.014	(0.022)	72,010	0.219	0.273	0.219
12	-0.003	(0.021)	-0.006	(0.020)	-0.004	(0.021)	71,806	0.229	0.288	0.229
13	0.002	(0.020)	-0.002	(0.019)	0.001	(0.020)	71,603	0.238	0.302	0.238
14	-0.000	(0.018)	-0.003	(0.018)	-0.001	(0.019)	71,400	0.245	0.315	0.246
15	-0.006	(0.017)	-0.009	(0.017)	-0.007	(0.018)	71,193	0.252	0.327	0.252
16	-0.005	(0.017)	-0.008	(0.016)	-0.005	(0.017)	70,990	0.258	0.338	0.258
17	-0.009	(0.017)	-0.013	(0.016)	-0.010	(0.017)	70,788	0.263	0.349	0.263
18	-0.009	(0.016)	-0.012	(0.015)	-0.010	(0.016)	70,587	0.267	0.358	0.267
19	-0.013	(0.016)	-0.017	(0.015)	-0.014	(0.016)	70,387	0.270	0.367	0.270
20	-0.013	(0.015)	-0.017	(0.014)	-0.013	(0.015)	70,182	0.272	0.375	0.273

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table 8: Effects of control variables

	Dismissals				Quits			
	Mean Points Next … Games							
VARIABLES	1	5	10	20	1	5	10	20
Mean Points Prev 10 Games	1.212***	0.598***	0.465***	0.356***	1.218***	0.600***	0.466***	0.359***
	(0.036)	(0.029)	(0.032)	(0.034)	(0.035)	(0.029)	(0.032)	(0.034)
Cumulative Surprise (t-1)	-0.115***	-0.051***	-0.039***	-0.030***	-0.117***	-0.051***	-0.039***	-0.030***
	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)
Cumulative Surprise (t-2 to t-5)	-0.112***	-0.054***	-0.040***	-0.031***	-0.113***	-0.055***	-0.041***	-0.032***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Cumulative Surprise (t-6 to t-10)	-0.126***	-0.058***	-0.045***	-0.035***	-0.127***	-0.058***	-0.045***	-0.035***
	(0.005)	(0.003)	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)
Average Opposition Position Next 1 Games	-0.024***				-0.025***			
	(0.001)				(0.001)			
Proportion of Home Games Next 1 Games	0.587***				0.589***			
	(0.010)				(0.010)			
Average Opposition Position Next 5 Games		0.003***				0.003***		
		(0.001)				(0.001)		
Proportion of Home Games Next 5 Games		0.607***				0.607***		
		(0.026)				(0.027)		
Average Opposition Position Next 10 Games			0.009***				0.009***	
			(0.001)				(0.001)	
Proportion of Home Games Next 10 Games			0.618***				0.613***	
			(0.054)				(0.054)	
Average Opposition Position Next 20 Games				0.012***				0.012***
				(0.001)				(0.001)
Proportion of Home Games Next 20 Games				0.736***				0.744***
				(0.097)				(0.097)
Observations	74,718	73,901	72,869	70,803	74,038	73,229	72,217	70,182
R-squared	0.118	0.159	0.256	0.373	0.118	0.159	0.256	0.375

Cluster robust standard errors in parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01
All models contain team Fixed Effects

Table 9: Entropy Balanced OLS (Dismissals)

	no FE		**team FE**		**season FE**		**N**	**Adj. R2**		
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	Season FE
1	-0.008	(0.043)	0.011	(0.045)	-0.016	(0.042)	65,603	0.084	0.144	0.089
2	0.032	(0.033)	0.028	(0.035)	0.030	(0.033)	65,440	0.056	0.112	0.061
3	0.048*	(0.028)	0.039	(0.030)	0.047*	(0.028)	65,275	0.074	0.137	0.079
4	0.038	(0.025)	0.028	(0.026)	0.037	(0.025)	65,110	0.073	0.143	0.078
5	0.015	(0.022)	0.005	(0.023)	0.014	(0.022)	64,944	0.086	0.160	0.090
6	0.004	(0.021)	-0.007	(0.021)	0.003	(0.020)	64,778	0.092	0.181	0.096
7	-0.007	(0.019)	-0.021	(0.019)	-0.008	(0.019)	64,612	0.100	0.198	0.106
8	-0.006	(0.019)	-0.019	(0.018)	-0.006	(0.018)	64,443	0.103	0.210	0.107
9	-0.015	(0.018)	-0.031*	(0.018)	-0.014	(0.018)	64,273	0.109	0.227	0.114
10	-0.002	(0.017)	-0.019	(0.016)	-0.001	(0.017)	64,101	0.124	0.242	0.129
11	-0.002	(0.016)	-0.016	(0.015)	-0.001	(0.016)	63,929	0.127	0.256	0.132
12	-0.006	(0.016)	-0.021	(0.014)	-0.006	(0.016)	63,755	0.134	0.269	0.139
13	0.001	(0.015)	-0.013	(0.013)	0.001	(0.015)	63,582	0.134	0.283	0.140
14	0.005	(0.015)	-0.009	(0.013)	0.005	(0.015)	63,409	0.145	0.298	0.152
15	0.005	(0.015)	-0.009	(0.013)	0.005	(0.015)	63,235	0.147	0.312	0.154
16	0.006	(0.015)	-0.009	(0.013)	0.006	(0.015)	63,059	0.147	0.320	0.155
17	0.011	(0.015)	-0.004	(0.012)	0.011	(0.014)	62,885	0.153	0.327	0.160
18	0.011	(0.014)	-0.006	(0.012)	0.011	(0.014)	62,711	0.154	0.333	0.162
19	0.017	(0.014)	0.000	(0.011)	0.017	(0.014)	62,535	0.155	0.342	0.163
20	0.016	(0.014)	-0.002	(0.011)	0.017	(0.014)	62,360	0.154	0.352	0.162

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table 10: Entropy Balanced OLS (Quits)

	no FE		**team FE**		**season FE**		**N**	**Adj. R2**		
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	Season FE
1	0.020	(0.066)	-0.031	(0.073)	0.022	(0.067)	65,048	0.152	0.281	0.157
2	-0.002	(0.050)	-0.008	(0.054)	-0.007	(0.050)	64,888	0.092	0.232	0.100
3	0.023	(0.048)	0.011	(0.051)	0.019	(0.048)	64,725	0.100	0.253	0.103
4	0.004	(0.042)	0.000	(0.043)	0.004	(0.042)	64,561	0.105	0.269	0.110
5	0.003	(0.038)	-0.010	(0.039)	0.005	(0.039)	64,396	0.113	0.279	0.117
6	0.016	(0.035)	0.005	(0.036)	0.016	(0.036)	64,232	0.116	0.301	0.119
7	0.017	(0.033)	0.003	(0.033)	0.017	(0.033)	64,069	0.126	0.326	0.130
8	0.012	(0.031)	-0.002	(0.031)	0.011	(0.032)	63,902	0.138	0.361	0.141
9	0.006	(0.030)	-0.013	(0.030)	0.006	(0.030)	63,736	0.141	0.367	0.145
10	0.000	(0.029)	-0.020	(0.029)	0.000	(0.029)	63,569	0.161	0.398	0.165
11	-0.009	(0.028)	-0.032	(0.027)	-0.008	(0.028)	63,398	0.166	0.408	0.171
12	-0.003	(0.026)	-0.024	(0.025)	-0.001	(0.026)	63,228	0.178	0.423	0.183
13	0.004	(0.026)	-0.019	(0.025)	0.006	(0.026)	63,058	0.183	0.435	0.189
14	0.005	(0.025)	-0.018	(0.024)	0.006	(0.025)	62,888	0.197	0.452	0.203
15	-0.003	(0.024)	-0.022	(0.023)	-0.002	(0.024)	62,714	0.198	0.452	0.202
16	-0.003	(0.023)	-0.022	(0.023)	-0.002	(0.023)	62,542	0.205	0.460	0.209
17	-0.008	(0.023)	-0.025	(0.022)	-0.008	(0.023)	62,371	0.210	0.470	0.214
18	-0.006	(0.022)	-0.018	(0.021)	-0.005	(0.022)	62,200	0.224	0.481	0.228
19	-0.010	(0.022)	-0.022	(0.021)	-0.008	(0.022)	62,030	0.228	0.489	0.232
20	-0.008	(0.021)	-0.022	(0.020)	-0.006	(0.021)	61,856	0.228	0.498	0.233

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

4.3 Alternative Specifications

4.3.1 No Turnover Follow Up Spells

We define a 'No Turnover' follow up spell as one where no subsequent Coaching change occurs after the initial change at t=0. In other words, we are considering teams who stick with their new Coach. Under this definition, both quits and dismissals now show evidence of positive returns after changing a Head Coach (Tables 11 and 12), a result which also holds with the inclusion of team fixed effects.[8] Notwithstanding the limitations of this approach that we outlined in Section 3, we believe there is still great value in these estimations, as we are likely capturing an upper bound of the effects of a Head Coach change. Given that these teams are likely happy with their new appointment, compared to teams who are unhappy and change Coaches again, these spells drop cases where the new Coach has been less successful.

While the results of positive effects of a Head Coach change are particularly evident for dismissals, we also observe some positive effects following a quit, though the effects occur much later in the follow up period. It could be that a new appointment following a quit takes a longer to adjust to the new club, if for example they are still appointing their backroom staff or figuring out their best team having not had time to plan unlike the situation following a dismissal. With that being said, given that in these follow up spells teams are likely happy with their new appointment, regardless of the manner of exit of the previous coach, then perhaps we should not be surprised to see longer term improvements to performance due to the learning process and adjustment period following a new appointment. This could have implications for team hiring policies and the process they go through to select a Head Coach. There is no official interview process that teams must go through, and teams often have a new appointment lined up even before they have dismissed the incumbent coach. Without taking the time to interview and carefully select candidates, it is possible that the wrong hire is made with a low job match surplus, only to be dismissed a few games later.

[8] Spells that last 20 games or fewer represents a fairly sizeable portion of our data. 34% percent of Head Coach spells are over by or on the 20^{th} game. Over 13% of coaches don't even last until the 10^{th} game. These short spells are predominantly occurring in Italy and Spain.

Table 11: Entropy Balanced OLS (Dismissals) with a No Turnover follow up spell

	no FE		**team FE**		**season FE**		**N**	**Adj. R2**		
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	Season FE
1	0.013	(0.044)	0.051	(0.045)	0.003	(0.043)	65,461	0.085	0.139	0.090
2	0.054	(0.034)	0.058	(0.035)	0.051	(0.034)	65,297	0.052	0.107	0.055
3	0.064**	(0.028)	0.056*	(0.030)	0.062**	(0.029)	65,131	0.071	0.133	0.073
4	0.071***	(0.025)	0.057**	(0.027)	0.070***	(0.025)	64,956	0.069	0.138	0.074
5	0.056**	(0.022)	0.036	(0.023)	0.055**	(0.022)	64,783	0.084	0.158	0.088
6	0.046**	(0.021)	0.024	(0.022)	0.045**	(0.021)	64,609	0.088	0.179	0.091
7	0.042**	(0.020)	0.019	(0.020)	0.042**	(0.020)	64,428	0.096	0.196	0.100
8	0.041**	(0.020)	0.016	(0.019)	0.041**	(0.020)	64,253	0.101	0.208	0.104
9	0.039**	(0.019)	0.009	(0.019)	0.039**	(0.019)	64,069	0.108	0.229	0.112
10	0.056***	(0.018)	0.030*	(0.017)	0.056***	(0.018)	63,880	0.127	0.255	0.132
11	0.061***	(0.018)	0.036**	(0.016)	0.062***	(0.018)	63,684	0.132	0.270	0.137
12	0.059***	(0.018)	0.031*	(0.016)	0.059***	(0.018)	63,494	0.142	0.285	0.147
13	0.066***	(0.017)	0.039***	(0.015)	0.066***	(0.017)	63,312	0.147	0.302	0.152
14	0.072***	(0.017)	0.045***	(0.014)	0.071***	(0.017)	63,128	0.157	0.315	0.162
15	0.081***	(0.017)	0.051***	(0.015)	0.080***	(0.017)	62,933	0.165	0.332	0.170
16	0.080***	(0.017)	0.050***	(0.014)	0.080***	(0.017)	62,743	0.165	0.339	0.172
17	0.091***	(0.017)	0.060***	(0.014)	0.090***	(0.017)	62,551	0.172	0.349	0.178
18	0.097***	(0.016)	0.063***	(0.013)	0.096***	(0.016)	62,364	0.180	0.360	0.187
19	0.110***	(0.016)	0.073***	(0.013)	0.109***	(0.016)	62,172	0.186	0.372	0.193
20	0.112***	(0.015)	0.075***	(0.012)	0.112***	(0.015)	61,980	0.189	0.383	0.195

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table 12: Entropy Balanced OLS (Quits) with a No Turnover follow up spell

	no FE		**team FE**		**season FE**		**N**	**Adj. R2**		
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	Season FE
1	0.043	(0.070)	-0.005	(0.077)	0.048	(0.071)	65,003	0.147	0.274	0.154
2	0.028	(0.053)	0.028	(0.057)	0.023	(0.052)	64,842	0.088	0.229	0.098
3	0.057	(0.049)	0.050	(0.053)	0.056	(0.049)	64,679	0.095	0.251	0.099
4	0.029	(0.043)	0.026	(0.044)	0.031	(0.043)	64,512	0.094	0.265	0.099
5	0.038	(0.040)	0.029	(0.041)	0.042	(0.040)	64,343	0.107	0.287	0.112
6	0.041	(0.038)	0.032	(0.037)	0.044	(0.038)	64,178	0.109	0.310	0.114
7	0.036	(0.035)	0.022	(0.034)	0.038	(0.035)	64,011	0.122	0.331	0.128
8	0.036	(0.033)	0.024	(0.033)	0.038	(0.034)	63,841	0.133	0.360	0.137
9	0.037	(0.032)	0.018	(0.032)	0.040	(0.032)	63,671	0.134	0.375	0.140
10	0.024	(0.031)	0.003	(0.030)	0.026	(0.031)	63,503	0.152	0.399	0.157
11	0.030	(0.029)	0.003	(0.029)	0.031	(0.029)	63,326	0.166	0.409	0.172
12	0.037	(0.027)	0.008	(0.026)	0.038	(0.027)	63,151	0.176	0.419	0.181
13	0.053*	(0.027)	0.020	(0.026)	0.055**	(0.028)	62,974	0.183	0.428	0.188
14	0.064**	(0.026)	0.027	(0.025)	0.064**	(0.027)	62,797	0.199	0.447	0.204
15	0.053**	(0.026)	0.023	(0.024)	0.054**	(0.026)	62,619	0.198	0.448	0.201
16	0.064**	(0.025)	0.031	(0.024)	0.065**	(0.025)	62,439	0.203	0.455	0.205
17	0.064**	(0.025)	0.035	(0.024)	0.064**	(0.025)	62,260	0.211	0.466	0.213
18	0.080***	(0.024)	0.050**	(0.023)	0.080***	(0.025)	62,080	0.232	0.476	0.234
19	0.085***	(0.025)	0.055**	(0.023)	0.085***	(0.025)	61,902	0.238	0.493	0.240
20	0.091***	(0.024)	0.061***	(0.022)	0.091***	(0.024)	61,722	0.238	0.501	0.240

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table 13: Within and Between Season Dismissals

	Within Season							Between Season						
	no FE		**team FE**		**N**	**Adj. R2**		**no FE**		**team FE**		**N**	**Adj. R2**	
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE
1	-0.022	(0.049)	-0.004	(0.054)	65,364	0.087	0.161	0.104	(0.075)	0.102	(0.083)	64,892	0.095	0.199
2	0.028	(0.034)	0.021	(0.037)	65,201	0.058	0.127	0.109*	(0.058)	0.091	(0.059)	64,732	0.073	0.182
3	0.024	(0.028)	0.010	(0.030)	65,036	0.070	0.147	0.137***	(0.048)	0.119**	(0.049)	64,569	0.113	0.223
4	0.016	(0.025)	0.002	(0.027)	64,871	0.081	0.164	0.133***	(0.045)	0.105**	(0.044)	64,405	0.097	0.220
5	-0.012	(0.022)	-0.025	(0.023)	64,705	0.102	0.183	0.122***	(0.039)	0.084**	(0.038)	64,240	0.092	0.213
6	-0.017	(0.020)	-0.031	(0.021)	64,539	0.117	0.210	0.108***	(0.033)	0.070**	(0.033)	64,077	0.105	0.233
7	-0.025	(0.018)	-0.040**	(0.019)	64,373	0.124	0.216	0.092***	(0.031)	0.048	(0.031)	63,914	0.110	0.246
8	-0.024	(0.017)	-0.036**	(0.018)	64,204	0.126	0.221	0.094***	(0.031)	0.049	(0.031)	63,748	0.115	0.268
9	-0.029*	(0.016)	-0.042**	(0.017)	64,034	0.137	0.234	0.083***	(0.031)	0.036	(0.030)	63,582	0.125	0.291
10	-0.022	(0.015)	-0.035**	(0.016)	63,862	0.148	0.244	0.096***	(0.030)	0.047	(0.030)	63,415	0.142	0.311
11	-0.021	(0.014)	-0.031**	(0.015)	63,690	0.154	0.256	0.098***	(0.029)	0.045	(0.028)	63,244	0.154	0.328
12	-0.024*	(0.013)	-0.034**	(0.014)	63,516	0.160	0.268	0.088***	(0.029)	0.040	(0.027)	63,074	0.151	0.333
13	-0.018	(0.014)	-0.026*	(0.014)	63,343	0.159	0.275	0.093***	(0.027)	0.043*	(0.025)	62,904	0.156	0.350
14	-0.015	(0.013)	-0.023*	(0.013)	63,170	0.171	0.291	0.089***	(0.027)	0.040	(0.025)	62,735	0.158	0.357
15	-0.014	(0.013)	-0.022*	(0.013)	62,996	0.175	0.302	0.091***	(0.028)	0.041*	(0.024)	62,562	0.165	0.378
16	-0.010	(0.013)	-0.018	(0.012)	62,820	0.177	0.307	0.082***	(0.027)	0.032	(0.024)	62,390	0.167	0.392
17	-0.004	(0.013)	-0.012	(0.011)	62,646	0.178	0.310	0.080***	(0.027)	0.033	(0.023)	62,219	0.175	0.399
18	-0.005	(0.013)	-0.015	(0.011)	62,472	0.175	0.314	0.076***	(0.026)	0.031	(0.022)	62,048	0.179	0.407
19	0.003	(0.013)	-0.006	(0.011)	62,296	0.177	0.322	0.076***	(0.026)	0.030	(0.021)	61,878	0.172	0.412
20	0.002	(0.012)	-0.008	(0.011)	62,121	0.175	0.331	0.077***	(0.026)	0.029	(0.021)	61,704	0.174	0.419

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

4.3.2 Within and Between Season changes

In Table 13, we distinguish between dismissals that took place during the season (left hand side of Table 13), and those that took place between seasons (right hand side). To obtain these results, we remove the Last Game of Season variable from the balancing scheme for within season dismissals, and Last Game of Season and Season Progress are removed from the balancing scheme for the offseason dismissal results. Focusing on mid-season dismissals, changing Head Coach appears to make little difference to team performance, while the team fixed effects model suggests that team performance actually declines somewhat over games 7-15. If we instead focus only on coaching changes during the offseason (following a dismissal), team performance improves across the whole follow up period (apart from the first game of the season). The team fixed effects variant also maintains some significance in the early part of the season. This distinction could be explained by Head Coaches being given the time to implement their new ideas and methods during pre-season training, as well as having access to the summer transfer window, where they can sell unwanted players, and bring in new players to improve their squad. Coaches who are hired mid-season are not afforded the opportunity to work with their squad without the burden of matches.

4.3.3 Promotions and Relegations

As a final check, we consider the role of promotions and relegations in our estimations. Results tables can be found in the appendix, tables A1 and A2. Results are largely unchanged from our baseline specification, though the team fixed effects variant of our dismissals model when taking out promoted teams shows some evidence of a bump to performance early in the new coach's tenure. By excluding these newly promoted teams, who are likely to be lower in the table and perhaps struggling to adapt to the higher division, and thus a new coach may find it harder to have any impact on results. Hence, we see this "bump" emerge when excluding these newly promoted teams.

4.4 Robustness Checks

We test the robustness of specifying the outcome model as an OLS regression. For particularly short follow up spells, our outcome variable points per game will appear ‘lumpy’, but as we extend our follow up period, points per game will more closely resemble a continuous variable. For example, for one game, points per game can be 3,1 or 0. For two games, points per game could be 3, 2, 1.5, 1, 0.5 or 0. As we divide a larger number of games, these ‘gaps’ in the outcome variable will be filled in. To address this, we use a Poisson regression which will improve the modelling in shorter spells where the outcome is essentially a count. We also test a Generalised Linear Model, where the outcome is points share i.e. Points achieved divided by Maximum Attainable Points over said period. Results, shown in Tables A3-A6, are identical to specifying as an OLS in both cases.

5. Conclusions

Using a large, linked employer-employee data set for professional football in four countries, we are able to separate out the theoretically different effects on performance of a Coach being dismissed and a Coach quitting. Professional football setting is useful in trying to isolate the causal impact of leadership on organisational performance, partly because the industry does not usually suffer from the exogenous shocks that afflict many other industries which make it harder to attribute performance change to management. The setting also means principals who hire and fire their managers - Head Coaches in this setting - benefit from quick and frequent updating of firm performance because football teams tend to play one or two games per week during the football season.

Even though there is a strong theoretical argument to suggest that leadership changes in football could, and perhaps should make a difference, our estimates using entropy balancing fail to show any consistent gains to performance following either a dismissal or a quit, when compared to unconstrained counterfactual scenarios in which teams suffer similar runs of form but do not immediately experience Head Coach turnover. The finding is largely in keeping with other studies which suggest regression to the mean can explain the lack of sustained positive effects of Head Coach changes on football team performance.

However, we find a strong exception to this conventional result. We estimate what is likely to be an upper bound of the effect of managerial change by constraining our results to spells where performance is permitted to develop and examining the effects of a Coach change among teams who make no subsequent coaching change in the 20 games after the initial change. Using these constrained spells, we find teams can experience positive returns after a dismissal of between 0.04-0.1 points per game, and between 0.05-0.09 points per game after a quit, with the effects for quits occurring later in the follow up spell. Even though the magnitude is rather small in a sporting sense, this could well prove the difference between relegation and staying up or qualifying for a European competition or not which are undoubted signs of success. That is not to say that teams should keep hold of their new coach regardless of results. Instead, we believe this finding highlights the importance of a finding good job match in the first place, rather than continually changing coaches.

That quits and dismissals result in somewhat different performance outcomes is consistent with economic theory as laid out by Farber (1999). Dismissals are triggered by principals (team owners) rather than agents (employed coaches). The dismissal is itself triggered by poor team performance which is a signal of a bad job match. The owner uses their acquired information on the Head Coach's ability and productivity to terminate the relationship with the aim of securing a better job match with a new hire. Quits are triggered by the agent rather than the principal with the departing Coach seeking better opportunities elsewhere (which include switching to a different job as well as different employer). Given that the job match was satisfactory to the employer (team owner) without consideration of the Coach's outside options then the best the employer can do is to replace the Coach with a job match that is just as good as the previous one. Our results show that team performance is neither improved nor impaired by Head Coach succession following a quit, suggesting that job matches between teams and voluntarily departing coaches were, on average, efficient. Nevertheless, longer term performance improvements are still possible because of a learning process.

As to why we see some differing effects in the short run and in the longer run, depending on specification, we conjecture that two effects could explain this. In the short run, there is a motivational effect of a new Head Coach where players are keen to impress the new leader. Recall that most teams that fire a Head Coach do so after experiencing a decline in performance. Given that it takes time to reequip a playing squad, the existing players have a time window in which to impress the incoming Head Coach to avoid being dropped or transferred. This would explain any upturns in form we observe in some specifications. In the medium to longer term, new coaches have to learn about their new team and its infrastructure very quickly given pressures to deliver good results. Many will have studied the team's attributes from afar but will have little to no first-hand experience of working at the club. There is a quick learning process as incoming coaches discern which management practices work best for their new employers.

We note as a point for further research that our results do not entirely support the conjecture of a market for mediocre managerial talent advanced by Tervio (2009) and Peeters et al. (2016). If most coaches were mediocre then we would not observe any positive effects on team performance that we find from cases of fired coaches. It is possible that a Head Coach who appears mediocre at one club can be successful at another. Put another way, the value of a job match varies across clubs and each club has an idiosyncratic element in this value. A poorly performing club will tend to draw its hiring from the lower end of the ability distribution but such a coach can nevertheless help improve team performance.

Further work is needed to investigate heterogeneity of Head Coach effects on team performance, since coaches themselves are likely to be heterogeneous in ability (Peeters et al, 2016). Even if our estimates, and indeed estimates of past work, yield low or zero mean effects, there may well be some positive, some zero and some negative effects and it is worth probing into where and how these occur and of course whether there are systematic patterns to the positive and negative effects.

APPENDIX

Table A1: Excluding Relegations

	Dismissals							**Quits**						
	no FE		**team FE**		**N**	**Adj. R2**		**no FE**		**team FE**		**N**	**Adj. R2**	
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	*Coeff*	*s.e.*	*Coeff*			No FE	Team FE
1	-0.027	(0.048)	-0.003	(0.045)	59,933	0.089	0.163	0.042	(0.072)	-0.005	(0.068)	59,449	0.156	0.303
2	0.036	(0.035)	0.033	(0.033)	59,781	0.061	0.129	0.023	(0.054)	0.009	(0.050)	59,300	0.100	0.260
3	0.049*	(0.029)	0.039	(0.027)	59,627	0.078	0.151	0.035	(0.046)	0.016	(0.042)	59,148	0.105	0.272
4	0.045*	(0.026)	0.034	(0.024)	59,473	0.075	0.158	0.018	(0.041)	0.008	(0.036)	58,994	0.109	0.288
5	0.019	(0.023)	0.009	(0.021)	59,318	0.091	0.177	0.015	(0.037)	-0.005	(0.033)	58,840	0.117	0.298
6	0.009	(0.022)	-0.000	(0.020)	59,163	0.097	0.201	0.026	(0.034)	0.008	(0.030)	58,687	0.125	0.319
7	-0.005	(0.020)	-0.015	(0.018)	59,009	0.109	0.218	0.023	(0.032)	0.002	(0.028)	58,535	0.136	0.342
8	-0.003	(0.019)	-0.014	(0.017)	58,853	0.112	0.232	0.017	(0.030)	-0.003	(0.026)	58,380	0.149	0.379
9	-0.012	(0.019)	-0.026	(0.017)	58,696	0.119	0.249	0.010	(0.029)	-0.013	(0.025)	58,227	0.149	0.384
10	0.002	(0.018)	-0.014	(0.016)	58,537	0.133	0.263	0.005	(0.028)	-0.020	(0.024)	58,073	0.172	0.414
11	0.001	(0.017)	-0.012	(0.015)	58,378	0.136	0.275	-0.007	(0.027)	-0.032	(0.023)	57,915	0.180	0.428
12	-0.004	(0.017)	-0.016	(0.014)	58,217	0.141	0.288	-0.000	(0.026)	-0.025	(0.022)	57,757	0.192	0.443
13	0.005	(0.016)	-0.007	(0.014)	58,057	0.143	0.302	0.007	(0.025)	-0.019	(0.021)	57,600	0.196	0.455
14	0.008	(0.016)	-0.003	(0.014)	57,896	0.153	0.316	0.007	(0.025)	-0.018	(0.021)	57,443	0.214	0.473
15	0.009	(0.016)	-0.002	(0.013)	57,735	0.155	0.331	-0.001	(0.024)	-0.021	(0.020)	57,282	0.216	0.471
16	0.012	(0.015)	0.001	(0.013)	57,572	0.158	0.340	-0.001	(0.024)	-0.022	(0.020)	57,123	0.224	0.480
17	0.017	(0.015)	0.004	(0.013)	57,411	0.162	0.345	-0.006	(0.023)	-0.025	(0.019)	56,964	0.228	0.491
18	0.015	(0.015)	0.001	(0.012)	57,250	0.164	0.350	-0.005	(0.023)	-0.017	(0.019)	56,806	0.244	0.502
19	0.020	(0.014)	0.006	(0.012)	57,087	0.166	0.359	-0.008	(0.022)	-0.020	(0.019)	56,649	0.247	0.511
20	0.019	(0.014)	0.004	(0.012)	56,925	0.165	0.369	-0.007	(0.022)	-0.020	(0.018)	56,488	0.248	0.521

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table A2: Excluding Promotions

	Dismissals							Quits						
	no FE		team FE		N	Adj. R2		no FE		team FE		N	Adj. R2	
Mean Points Next ... Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*		No FE	Team FE	*Coeff*	*s.e.*	*Coeff*			No FE	Team FE
1	0.009	(0.047)	0.037	(0.045)	59,919	0.088	0.154	0.015	(0.072)	-0.051	(0.068)	59,415	0.158	0.309
2	0.049	(0.034)	0.046	(0.032)	59,766	0.065	0.123	0.014	(0.054)	-0.004	(0.051)	59,265	0.103	0.252
3	0.070**	(0.028)	0.059**	(0.027)	59,611	0.076	0.140	0.036	(0.047)	0.014	(0.042)	59,112	0.113	0.283
4	0.062**	(0.025)	0.052**	(0.024)	59,456	0.072	0.140	0.008	(0.041)	-0.002	(0.036)	58,958	0.122	0.299
5	0.039*	(0.022)	0.027	(0.021)	59,300	0.085	0.158	0.009	(0.037)	-0.016	(0.033)	58,803	0.136	0.314
6	0.025	(0.021)	0.012	(0.020)	59,144	0.089	0.178	0.025	(0.035)	0.005	(0.030)	58,649	0.136	0.336
7	0.009	(0.020)	-0.005	(0.018)	58,988	0.096	0.195	0.026	(0.033)	0.001	(0.028)	58,496	0.146	0.354
8	0.009	(0.019)	-0.005	(0.017)	58,829	0.100	0.202	0.021	(0.031)	-0.006	(0.026)	58,339	0.158	0.389
9	0.003	(0.019)	-0.015	(0.017)	58,669	0.108	0.219	0.017	(0.029)	-0.016	(0.024)	58,182	0.162	0.395
10	0.015	(0.018)	-0.003	(0.016)	58,507	0.122	0.234	0.012	(0.028)	-0.023	(0.023)	58,025	0.180	0.424
11	0.011	(0.017)	-0.003	(0.015)	58,345	0.125	0.248	0.000	(0.027)	-0.036	(0.023)	57,864	0.182	0.429
12	0.003	(0.016)	-0.012	(0.014)	58,182	0.130	0.260	0.008	(0.027)	-0.024	(0.022)	57,704	0.188	0.443
13	0.007	(0.016)	-0.007	(0.014)	58,020	0.130	0.272	0.012	(0.026)	-0.021	(0.021)	57,545	0.191	0.454
14	0.011	(0.016)	-0.004	(0.014)	57,858	0.139	0.284	0.015	(0.025)	-0.019	(0.021)	57,386	0.203	0.468
15	0.012	(0.015)	-0.002	(0.013)	57,695	0.142	0.299	0.008	(0.025)	-0.019	(0.020)	57,223	0.206	0.469
16	0.012	(0.015)	-0.003	(0.013)	57,530	0.142	0.307	0.010	(0.024)	-0.017	(0.020)	57,062	0.214	0.475
17	0.015	(0.015)	0.001	(0.013)	57,367	0.149	0.313	0.006	(0.024)	-0.020	(0.019)	56,902	0.220	0.486
18	0.015	(0.014)	-0.001	(0.012)	57,204	0.150	0.319	0.009	(0.023)	-0.013	(0.019)	56,742	0.236	0.497
19	0.021	(0.014)	0.004	(0.012)	57,039	0.152	0.328	0.005	(0.023)	-0.017	(0.019)	56,583	0.238	0.505
20	0.022	(0.014)	0.004	(0.012)	56,875	0.154	0.339	0.008	(0.023)	-0.016	(0.019)	56,420	0.239	0.514

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table A3: Poisson Regression Robustness Check

	Dismissals					**Quits**				
	no FE		**team FE**			**no FE**		**team FE**		
Mean Points Next … Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	N	*Coeff*	*s.e.*	*Coeff*	*s.e.*	N
1	-0.004	(0.033)	0.020	(0.035)	65,603	0.014	(0.049)	-0.015	(0.055)	65,048
2	0.025	(0.025)	0.024	(0.027)	65,440	-0.001	(0.037)	-0.006	(0.040)	64,888
3	0.037*	(0.021)	0.031	(0.023)	65,275	0.017	(0.035)	0.008	(0.037)	64,725
4	0.029	(0.019)	0.022	(0.020)	65,110	0.002	(0.031)	-0.000	(0.031)	64,561
5	0.012	(0.017)	0.004	(0.017)	64,944	0.002	(0.028)	-0.006	(0.029)	64,396
6	0.003	(0.016)	-0.005	(0.016)	64,778	0.011	(0.026)	0.003	(0.026)	64,232
7	-0.005	(0.015)	-0.016	(0.015)	64,612	0.012	(0.024)	0.003	(0.024)	64,069
8	-0.004	(0.014)	-0.015	(0.014)	64,443	0.008	(0.023)	-0.002	(0.023)	63,902
9	-0.011	(0.014)	-0.023*	(0.014)	64,273	0.003	(0.022)	-0.009	(0.022)	63,736
10	-0.001	(0.013)	-0.014	(0.012)	64,101	-0.001	(0.021)	-0.016	(0.021)	63,569
11	-0.001	(0.012)	-0.012	(0.011)	63,929	-0.008	(0.020)	-0.024	(0.020)	63,398
12	-0.004	(0.012)	-0.016	(0.011)	63,755	-0.004	(0.019)	-0.019	(0.019)	63,228
13	0.001	(0.012)	-0.010	(0.010)	63,582	0.002	(0.019)	-0.015	(0.018)	63,058
14	0.004	(0.011)	-0.007	(0.010)	63,409	0.002	(0.018)	-0.015	(0.017)	62,888
15	0.004	(0.011)	-0.007	(0.010)	63,235	-0.004	(0.018)	-0.017	(0.017)	62,714
16	0.005	(0.011)	-0.007	(0.010)	63,059	-0.003	(0.017)	-0.017	(0.016)	62,542
17	0.008	(0.011)	-0.003	(0.009)	62,885	-0.007	(0.017)	-0.020	(0.016)	62,371
18	0.008	(0.011)	-0.005	(0.009)	62,711	-0.006	(0.016)	-0.014	(0.015)	62,200
19	0.013	(0.011)	0.000	(0.009)	62,535	-0.008	(0.016)	-0.017	(0.015)	62,030
20	0.012	(0.010)	-0.001	(0.008)	62,360	-0.007	(0.015)	-0.017	(0.015)	61,856

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table A4: Poisson Regression Robustness Check (with a No Turnover follow up spell)

	Dismissals					**Quits**				
	no FE		**team FE**			**no FE**		**team FE**		
Mean Points Next … Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	N	*Coeff*	*s.e.*	*Coeff*	*s.e.*	N
1	0.012	(0.033)	0.049	(0.035)	65,461	0.033	(0.051)	0.005	(0.057)	65,003
2	0.041	(0.025)	0.046*	(0.027)	65,297	0.022	(0.038)	0.021	(0.041)	64,842
3	0.049**	(0.021)	0.043*	(0.023)	65,131	0.042	(0.035)	0.036	(0.037)	64,679
4	0.054***	(0.019)	0.043**	(0.020)	64,956	0.022	(0.031)	0.019	(0.031)	64,512
5	0.042**	(0.017)	0.028	(0.018)	64,783	0.028	(0.029)	0.022	(0.029)	64,343
6	0.034**	(0.016)	0.018	(0.016)	64,609	0.030	(0.027)	0.023	(0.027)	64,178
7	0.032**	(0.015)	0.015	(0.015)	64,428	0.026	(0.025)	0.016	(0.024)	64,011
8	0.031**	(0.015)	0.012	(0.015)	64,253	0.026	(0.024)	0.016	(0.023)	63,841
9	0.029**	(0.014)	0.007	(0.014)	64,069	0.027	(0.023)	0.013	(0.023)	63,671
10	0.041***	(0.013)	0.022*	(0.013)	63,880	0.017	(0.022)	0.001	(0.022)	63,503
11	0.046***	(0.013)	0.027**	(0.012)	63,684	0.021	(0.021)	0.001	(0.020)	63,326
12	0.044***	(0.013)	0.023*	(0.012)	63,494	0.027	(0.019)	0.005	(0.019)	63,151
13	0.049***	(0.012)	0.029***	(0.011)	63,312	0.038**	(0.019)	0.013	(0.019)	62,974
14	0.053***	(0.012)	0.033***	(0.011)	63,128	0.045**	(0.019)	0.018	(0.018)	62,797
15	0.059***	(0.012)	0.038***	(0.011)	62,933	0.038**	(0.018)	0.015	(0.017)	62,619
16	0.059***	(0.012)	0.037***	(0.010)	62,743	0.046***	(0.018)	0.021	(0.017)	62,439
17	0.066***	(0.012)	0.043***	(0.010)	62,551	0.045**	(0.018)	0.023	(0.017)	62,260
18	0.070***	(0.011)	0.046***	(0.010)	62,364	0.056***	(0.017)	0.033**	(0.016)	62,080
19	0.079***	(0.011)	0.052***	(0.009)	62,172	0.059***	(0.017)	0.037**	(0.016)	61,902
20	0.081***	(0.011)	0.054***	(0.009)	61,980	0.063***	(0.016)	0.041***	(0.015)	61,722

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table A5: GLM Regression of Points Share

	Dismissals					**Quits**				
	no FE		**team FE**			**no FE**		**team FE**		
Points Share Next … Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	N	*Coeff*	*s.e.*	*Coeff*	*s.e.*	N
1	-0.003	(0.014)	0.004	(0.015)	65,603	0.007	(0.022)	-0.010	(0.024)	65,048
2	0.011	(0.011)	0.009	(0.012)	65,440	-0.001	(0.017)	-0.003	(0.018)	64,888
3	0.016*	(0.009)	0.013	(0.010)	65,275	0.008	(0.016)	0.004	(0.017)	64,725
4	0.013	(0.008)	0.009	(0.009)	65,110	0.001	(0.014)	0.000	(0.014)	64,561
5	0.005	(0.007)	0.002	(0.008)	64,944	0.001	(0.013)	-0.003	(0.013)	64,396
6	0.001	(0.007)	-0.002	(0.007)	64,778	0.005	(0.012)	0.002	(0.012)	64,232
7	-0.002	(0.006)	-0.007	(0.006)	64,612	0.006	(0.011)	0.001	(0.011)	64,069
8	-0.002	(0.006)	-0.006	(0.006)	64,443	0.004	(0.010)	-0.001	(0.010)	63,902
9	-0.005	(0.006)	-0.010*	(0.006)	64,273	0.002	(0.010)	-0.004	(0.010)	63,736
10	-0.001	(0.006)	-0.006	(0.005)	64,101	0.000	(0.010)	-0.007	(0.009)	63,569
11	-0.001	(0.005)	-0.005	(0.005)	63,929	-0.003	(0.009)	-0.011	(0.009)	63,398
12	-0.002	(0.005)	-0.007	(0.005)	63,755	-0.001	(0.009)	-0.008	(0.008)	63,228
13	0.000	(0.005)	-0.004	(0.004)	63,582	0.001	(0.009)	-0.006	(0.008)	63,058
14	0.002	(0.005)	-0.003	(0.004)	63,409	0.002	(0.008)	-0.006	(0.008)	62,888
15	0.002	(0.005)	-0.003	(0.004)	63,235	-0.001	(0.008)	-0.007	(0.008)	62,714
16	0.002	(0.005)	-0.003	(0.004)	63,059	-0.001	(0.008)	-0.007	(0.008)	62,542
17	0.004	(0.005)	-0.001	(0.004)	62,885	-0.003	(0.008)	-0.008	(0.007)	62,371
18	0.004	(0.005)	-0.002	(0.004)	62,711	-0.002	(0.007)	-0.006	(0.007)	62,200
19	0.006	(0.005)	0.000	(0.004)	62,535	-0.003	(0.007)	-0.007	(0.007)	62,030
20	0.005	(0.005)	-0.001	(0.004)	62,360	-0.003	(0.007)	-0.007	(0.007)	61,856

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Table A6: GLM Regression of Points Share (with a No Turnover follow up spell)

	Dismissals					**Quits**				
	no FE		**team FE**			**no FE**		**team FE**		
Points Share Next … Games	*Coeff*	*s.e.*	*Coeff*	*s.e.*	N	*Coeff*	*s.e.*	*Coeff*	*s.e.*	N
1	0.004	(0.015)	0.017	(0.015)	65,461	0.014	(0.023)	-0.002	(0.026)	65,003
2	0.018	(0.011)	0.019*	(0.012)	65,297	0.009	(0.018)	0.009	(0.019)	64,842
3	0.021**	(0.009)	0.019*	(0.010)	65,131	0.019	(0.016)	0.017	(0.017)	64,679
4	0.024***	(0.008)	0.019**	(0.009)	64,956	0.010	(0.014)	0.009	(0.015)	64,512
5	0.019**	(0.007)	0.012	(0.008)	64,783	0.013	(0.013)	0.010	(0.014)	64,343
6	0.015**	(0.007)	0.008	(0.007)	64,609	0.014	(0.013)	0.011	(0.012)	64,178
7	0.014**	(0.007)	0.006	(0.007)	64,428	0.012	(0.012)	0.007	(0.011)	64,011
8	0.014**	(0.007)	0.005	(0.006)	64,253	0.012	(0.011)	0.008	(0.011)	63,841
9	0.013**	(0.006)	0.003	(0.006)	64,069	0.012	(0.011)	0.006	(0.011)	63,671
10	0.019***	(0.006)	0.010*	(0.006)	63,880	0.008	(0.010)	0.001	(0.010)	63,503
11	0.020***	(0.006)	0.012**	(0.005)	63,684	0.010	(0.010)	0.001	(0.010)	63,326
12	0.020***	(0.006)	0.010*	(0.005)	63,494	0.012	(0.009)	0.003	(0.009)	63,151
13	0.022***	(0.006)	0.013***	(0.005)	63,312	0.018**	(0.009)	0.007	(0.009)	62,974
14	0.024***	(0.006)	0.015***	(0.005)	63,128	0.021**	(0.009)	0.009	(0.008)	62,797
15	0.027***	(0.006)	0.017***	(0.005)	62,933	0.018**	(0.009)	0.008	(0.008)	62,619
16	0.027***	(0.006)	0.017***	(0.005)	62,743	0.021**	(0.008)	0.010	(0.008)	62,439
17	0.030***	(0.006)	0.020***	(0.005)	62,551	0.021**	(0.008)	0.012	(0.008)	62,260
18	0.032***	(0.005)	0.021***	(0.004)	62,364	0.027***	(0.008)	0.017**	(0.008)	62,080
19	0.037***	(0.005)	0.024***	(0.004)	62,172	0.028***	(0.008)	0.018**	(0.008)	61,902
20	0.037***	(0.005)	0.025***	(0.004)	61,980	0.030***	(0.008)	0.020***	(0.007)	61,722

Cluster Robust Standard Errors in Parentheses (clustered at the Team level) *p<0.1, **p<0.05, ***p<0.01

Chapter References

Ashenfelter, O. (1978), *'Estimating the Effect of Training Programs on Earnings'*, The Review of Economics and Statistics, pp 47-57

Bennedsen, M., Perez-Gonzalez, F. and Wolfenzon, D. (2012), *'Evaluating the Impact of the Boss: Evidence from CEO Hospitalization Events'*, unpublished manuscript

Berri, D.J., Leeds, M.A., Leeds, E.M. and Mondello, M. (2009), *'The Role of Managers in Team Performance'*, International Journal of Sports Finance, Vol.4, No.2, pp 75-93

Besley, T., Montalvo, J. G., and Reynal-Querol, M. (2011), '*Do Educated Leaders Matter?'*, The Economic Journal, Vol.121, No. 554, F205-227

Besters, L., Van Ours, J. and Van Tuijl, M. (2016), *'Effectiveness of In-Season Manager Changes in English Premier League Football'*, De Economist, Vol.164, No.3, pp 335-356

Bradbury, J.C. (2017), *'Hired to be Fired: The Publicity Values of Managers'*, Managerial and Decision Economics, Vol.38, No.7, pp 929-940

Bruinshoofd, A. and ter Weel, B. (2003), *'Manager to go? Performance Dips Reconsidered with Evidence from Dutch Football'*, European Journal of Operational Research, Vol.148, No.2, pp 233-246

Bryson, A., Buraimo, B., Farnell, A. and Simmons, R. (2021), *'Time to go? Head Coach Quits and Dismissals in Professional Football'*, De Economist, Vol.169, No.1, pp 81-105

D'Addona, S. and Kind, A. (2014), *'Forced Manager Turnovers in English Soccer Leagues'*, Journal of Sports Economics, Vol.15, No.2, pp 150-179

De Grip, A. (2015), *'The importance of informal learning at work*', *IZA World of Labor*

De Paola, M. and Scoppa, V. (2012), *'The Effects of Managerial Turnover: Evidence from Coach Dismissals in Italian Soccer Teams'*, Journal of Sports Economics, Vol.13, No.2, pp 152-168

Dobson, S. and Goddard, J. (2011), '*The Economics of Football'*, 2nd ed. Cambridge: Cambridge University Press

Farber, H. S. (1999), '*Mobility and Stability: The Dynamics of Job Change in Labor Markets',* Chapter 37 in O. Ashenfelter and D. Card (eds.) '*Handbook of Labor Economics',* Vol.3, pp. 2439-2483

Ferguson, A. (2013), '*Alex Ferguson: My Autobiography'*, Hodder and Stoughton

Gaynor, M., Seider, H. and Vogt, W.B. (2005), *'The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing'*, American Economic Review, Vol.95, No. 2, pp 243-247

Goff, B., Wilson, D. and Zimmer, D. (2019), *'The Effect of Management Changes on Winning in Professional Sports: Analysis Using a Dynamic Lag Adjustment Model'*, Managerial and Decision Economics, Vol.40, No.8, pp 982-992

Green, C., Lozano, F and Simmons, R. (2015), *'Rank-Order Tournaments, Probability of Winning and Investing in Talent: Evidence from Champions' League Qualifying Rules'*, National Institute Economic Review, Vol.232, No.1, R30-40

Hainmueller, J. (2012), *'Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples'*, Political Analysis, pp 25-46

Hainmueller, J. and Xu, J. (2013), *'ebalance: A Stata Package for Entropy Balancing'*, Journal of Statistical Software, Vol.54, No.7: 1-18

Kuper, S. and Szymanski, S. (2009), '*Soccernomics',* Harper Collins

Lazear, E.P., Shaw, K.L. and Stanton, C.T. (2015), *'The Value of Bosses'*, Journal of Labour Economics, Vol.33, No.4, pp 823-861

Levitt, S. (2021), *'Heads or Tails: The Impact of a Coin Toss on Major Life Decisions and Subsequent Happiness'*, The Review of Economic Studies, Vol.88, No.1, pp 378-405

Madum, A. (2016), *'Managerial Turnover and Subsequent Firm Performance: Evidence from Danish Soccer* Teams', International Journal of Sport Finance, Vol.11, No.1, pp 46-62

Muehlheusser, G., Schneemann, S. and Sliwka, D. (2016), *'The Impact of Managerial Change on Performance: The Role of Team Heterogeneity'*, Economic Inquiry, Vol.54, No.2, pp 1121-1149

Peeters, T., Szymanski, S. and Terviö, M. (2016), *'Entry, Career Dynamics and Quality in the Labour Market for Talent'*, mimeo

Pieper, J., Nüesch, S. and Franck, E. (2014), *'How Performance Expectations Affect Managerial Replacement Decisions'*, Schmalenbach Business Review, Vol.66, No.1, pp 5-23

Rosen, S. (1990), *'Contracts and the Market for Executives'*, NBER Working Paper No.3542

Scelles, N. and Llorca, L. (2020), *'Head Coach Change and Team Performance in the French Men's Football Ligue 1, 2000-2016'*, Economics Bulletin, Vol.40, No.2, pp 920-937

Tena, J. D. and Forrest, D. (2007), *'Within-Season Dismissal of Football Coaches: Statistical Analysis of Causes and Consequences'*, European Journal of Operational Research, Vol.181, No.1, pp 362-373

Terviö, M. (2009), *'Superstars and Mediocrities: Market Failure in the Discovery of Talent'*, The Review of Economic Studies, Vol.76, No.2, pp 829-850

Van Ours, J. C. and Van Tuijl, M. A. (2016), *'In-season Head Coach Dismissals and the Performance of Professional Football Teams'*, Economic Inquiry, Vol.54, No.1, pp 591-60

Zhao, Q. and Percival, D. (2017), *'Entropy Balancing is Doubly Robust'*, Journal of Causal Inference, Vol.5, No.1

Assessing the Role of Fatigue and Task Switching on Worker Performance. Evidence from MLB Pitchers.

Alex Farnell (Lancaster University)

Dave Berri (Southern Utah University)

Brian Mills (University of Texas)

Vincent O'Sullivan (Lancaster University)

Robert Simmons (Lancaster University)

Abstract

Opportunities to study how workers respond to the demands of task switching outside of a laboratory setting are rare. In this paper, we use three seasons of (pre Covid) Major League Baseball (MLB) data to see how pitchers are affected by the additional demands of having to bat and run bases. MLB is an ideal setting because of its two-league structure in which the American League has a Designated Hitter rule, allowing teams to nominate a player to bat in place of the pitcher. The National League does not (or did not, pre Covid). Under the reasonable assumption that teams in the two leagues are not selectively hiring pitchers based on their batting ability, we assess changes to pitching velocity, accuracy, and the number of walks and runs given up, in the (half) inning immediately following an at bat and/or getting on base. Results suggest that task switching in the form of batting is associated with gains across most of our performance measures, but that pitchers should avoid getting on base at all costs. The implication is that it is beneficial to stay active rather than sitting around between innings, but not to over-exert oneself. This finding is robust to within game and across league selection of pitchers, and to a placebo test when we allocate the at bat and on base to the inning before it happened. We offer explanations for these findings and suggest how these findings could inform other settings.

Keywords: Labour Productivity, Task, Baseball

JEL Classification: J24, M54, Z21, Z22

Acknowledgements

We would like to thank seminar participants at Lancaster University, Reading Online Sports Economics Seminars (ROSES), and conference participants at the Southern Economic Association, the Western Economic Association and the Scottish Economic Society for helpful advice and comments.

1. Introduction

Managers should be greatly interested in how fatigue affects the productivity of workers. Over the course of a working day, for example, workers may become mentally and / or physically fatigued, possibly leading to a loss in productivity. Hart (2004) proposes that the marginal productivity of hours worked varies over the course of the working day. In fact, at the start of the day it could be that marginal productivity actually rises as workers are "warming-up", but eventually fatigue or boredom sets in and productivity falls.

One possible source of fatigue comes from the requirement for workers to carry out multiple tasks (see for example Russ and Crews (2014)). Most, if not all jobs involve some degree of switching between different tasks. These may be job related (checking work emails, attending meetings, meeting clients etc.) or not (checking mobile phones, checking sports news etc.). But changing tasks is likely to involve switching costs, perhaps in the form of a mental adjustment to adapt to a new task, or through lost productive time when switching tasks. Indeed, a body of literature from psychology and behavioural economics (for example Buser and Peter (2012)) suggests subjects tend to struggle when faced with such demands.

However, there has been little in the way of empirical research from natural settings to understand how task switching effects productivity and performance. This, in part, is due to the lack of detailed worker level productivity data, since it can be difficult to define productivity in many occupations. Even if accurate productivity measures are available, it is rare to observe them on a frequent enough basis to track changes over short spaces of time.

To address these shortcomings in measurement, we use a particularly rich micro-level dataset containing accurate and comparable measures of worker performance and indicators of task switching. The industry is professional baseball, Major League Baseball (MLB), and the workers under consideration are starting pitchers. Pitching involves a great level of physical exertion, and so the cumulative effect of pitching over the course of a game will likely impede a pitcher's ability to perform. Indeed, we show that there is a general decline in performance over the course of a game as measured by velocity, command, and walks and runs given up.

Of note for this study on task switching is the two-league structure of MLB. The American League has a so called "Designated Hitter Rule", meaning that one player, usually the pitcher, is exempt from batting and a "Designated Hitter" takes their spot in the batting order. Whereas in the National League, pitchers must bat, and if successful at bat, run bases too. Our identification strategy relies on the fact that pitchers are observed playing in both leagues because of interleague play, and thus the same pitchers are exposed to treatment and control games throughout a season. Pitchers are specialists who have built a career based on their pitching. Batting and running bases are outside of their main skill set, so it is conceivable that these activities could impede their ability to carry out their primary role of pitching.

Contrary to expectations, however, we find a largely positive effect of previously batting, with gains to velocity, and crucially, giving up fewer runs to the opposition. However, for those pitchers who have a successful at bat and get on base, their subsequent pitching performance declines. The implication is that pitchers should stay active between innings (batting), but to no get too fatigued or distracted (running bases). Results hold when testing the robustness of our results against pitcher selection and ability, and when subjected to a placebo check.

The remainder of the paper is structured as follows. Section 2 offers a review of the literature on the effects of fatigue and task switching on productivity. Section 3 offers an overview of baseball and MLB. Section 4 describes the data and measures of task switching in our setting, followed by an overview of the model to be estimated in Section 5. Section 6 presents the results with and Section 7 concludes our work by discussing implications of these findings and where results may lie with respect to other labour markets.

2. Theory & Literature Review

We contribute to a number of strands of literature with a particular focus on the effects of fatigue and task switching on performance. Even though our focus is on baseball, we believe our findings are generalisable not only to other sports, but also to more general labour market settings, particularly jobs that involve carrying out physical tasks.

2.1 Fatigue

Work examining the effects of fatigue on performance tends to focus on the association between hours worked and output. For example, Pencavel (2015) considers the case of munition factory workers during the First World War in Britain. In this setting, exogenous variation in hours worked was largely driven by the demand for shells on the front line. He finds that up to about 48 hours of work per week, the rise in output was proportional to hours worked, but working beyond this this point causes the marginal product of hours to diminish, with maximum output being achieved at about 63 hours. Collewet and Sauermann (2017) also uncover diminishing marginal productivity of hours for workers in a Dutch call centre. For a 1% increase in hours worked, output only rose by 0.9%. The effect remained significant even when controlling for employee and shift characteristics (e.g. shorter night or weekend shifts, and hence more productive shifts), though the magnitude of the effects fell somewhat.

Not all research, however, finds evidence of this negative association. In fact, Lu and Lu (2017) find the opposite to be true. Their Difference in Differences strategy uses state variations in the abolishment of mandatory overtime for nurses, focusing on one particular sector: nursing homes. They find that the number of deficiency citations (a measure of poor service quality) increased by almost 22% in treated states. This change was not related to fatigue however, and was instead related to changes in staffing composition, with nursing homes decreasing the hours of permanent staff, and increasing the hours of contract nurses.

Crocker and Horst (1981) found no evidence of a decline in marginal value product (measured by daily earnings) associated with daily hours of work for citrus fruit pickers in California, though poor environmental conditions (ozone air pollution) did lead to a drop in earnings.[9] This does raise a potentially important distinction between mental and physical fatigue. Fruit picking is unlikely to be mentally challenging but is likely to be physically demanding, while other occupations may involve the opposite or indeed an interaction of the two. This interaction is important, as Marcora et al. (2009) show that mental fatigue can impair physical performance and limits short term endurance through perception of higher effort.

Turning to the sports economics literature, research on fatigue and performance is confined mainly to looking at the role of rest days between fixtures, rather than within game fatigue which would be more akin to the effect of extended hours in a more general labour market setting. Scoppa (2013) exploits variations in a team's rest days due to TV scheduling in the FIFA World Cup and UEFA European Football Championships. In more recent tournaments (1990's and onwards), rest days were found to have no impact on team performance even when controlling for team quality factors.[10] Entine and Small (2008) consider the role of rest days in explaining home court advantage in the National Basketball Association (NBA). Scheduling may be such that away teams are required play on successive days, a possible contributing factor to the observed 61% home win rate in the NBA. Their results suggest that the home team scored on average 3.24 points more than the away team, of which a small (0.31), though significant, portion can be attributed to the limited number of rest days. Moreover, visiting teams with back-to-back games were an estimated 1.77 points worse off than a fully rested

[9] It is unclear however, whether this was due to fatigue or simply a reduction in worker effort.
[10] Scoppa proposes that the reason tournaments before 1990 were affected by rest days was because the athletic preparation by teams and players was significantly worse than it is in modern day football.

visiting team. Other notable work examines the effect of travel distance on performance, namely Oberhofer et al. (2010) for the German Bundesliga and Nichols (2014) for the National Football League (NFL). In both cases, more travel is associated with declines in performance, while the latter also finds that direction of travel is important.

Work examining within game fatigue is mainly confined to the sports medicine literature. Rampinini et al. (2009), studying Italy's Serie A football league is a good example. They find players who covered more distance in the first half not only ran less (at various intensities) in the second half, but also saw a decline in the number of successful short passes. They concluded that match related fatigue influences both physical and technical output.

There is a well-established literature examining muscular fatigue of baseball pitchers. However, many of these studies suffer from small sample sizes and are limited to laboratory setups rather than observing data from the real world. Escamilla et al. (2007) observed that both pitch velocity and pitching mechanics (the position of the pitcher's torso) were significantly different between the first and last two innings pitched before a pitcher said they were unable to continue. In this setup, pitchers were throwing between 105 and 135 pitches and so results may only be applicable to starting pitchers. In a video analysis of MLB pitchers at Spring Training, Murray et al. (2001) found that pitch velocity decreased by 5mph, while leg rotation, knee angle and forces exerted on the shoulder were all significantly different between the first and last innings. Finally, Lyman et al. (2002) report that high pitch counts amongst youth baseball pitchers are associated with a higher self-reported incidence of elbow and shoulder pain. This was particularly evident for curveballs and sliders, types of pitches that place high loads on these joints.

2.2 Task Switching

In addition to fatigue, other studies have investigated the role of task switching and multitasking, each distinct behaviour, on productivity. Multitasking involves doing different tasks at the same time, while task switching involves doing different tasks sequentially, and evidence from Buser and Peter (2012) shows that this distinction is important. In their experiment, they randomly allocate participants into three groups; one group multitasking, one task switching at a time determined by the experiment, and a final group task switching at their own convenience.[11] Results suggest that subjects who multitasked perform worse than those

[11] In their experiment, the tasks included a Sudoku puzzle and a word search game

who task switched, while surprisingly, being able to pick when to switch tasks was associated with worse performance.

It is unclear however, how such experimental evidence translates into the real world because of the different nature of the tasks involved. Jobs involving multitasking or task switching are now synonymous with modern day work, and thus it should be of great interest to managers to understand how (or if) it affects productivity. From relatively low skilled occupations such as supermarket assistants to higher skilled jobs such as teachers and physicians, all roles will require workers to carry out different tasks. Sports too offers several examples of players having to do different tasks. In football (soccer) and rugby for example, players are constantly switching between attacking and defending whenever ball possession changes, while in cricket and baseball, players are required to both field and bat.

Theoretically, Aral et al. (2012) suggest that task switching has ambiguous effects on productivity. On the one hand, an effective ability to task switch could allow workers to smooth their output during lulls in workload, while skill complementarities across tasks should benefit productivity. On the other hand, carrying out multiple tasks could cause delays and force the prioritisation of more important tasks. Switching between tasks is also associated with mental congestions and increased errors (see for example Rubinstein et al. (2001) or Kiesel et al. (2010)).

Turning to the industry specific evidence, Coviello et al. (2015) use a sample of Italian judges specialising in labour disputes who receive randomly assigned cases. Naturally, some of these cases are more complex and so take longer to complete. Their results suggest that judges respond to an increase in future workloads by juggling more cases in the present. In particular, a 1% exogenous increase in workload increases the duration of trials by between 3 and 6 days, while judges would need to increase their effort by between 1.1% and 1.4% to maintain the same length of trials. A similar result is reported by Aral et al. (2012) using data on project outputs at an IT firm. They find that task switching increased total output, but this came at a cost of each project taking longer to complete. Singh (2014) studied physicians processing time, throughput rate and output quality from a hospital emergency department, and presents mixed evidence on the benefits of task switching (which in this setting refers to treating and attending to patients with different ailments). He finds that up to a value of about four patients per hour, task switching helps to reduce the time taken to process patients and reduces idle time. However, beyond this point, task switching eventually leads to fewer detected diagnoses and increases the likelihood of patients re-visiting the hospital within 24 hours.

Why then is there a need to re-visit this topic, and what are the benefits of using sports, specifically baseball data to address it? First, a common issue in assessment of performance in non-sports settings is that it can prove difficult to compare performance across different workers and across different firms. Moreover, performance on any one task may encapsulate several dimensions e.g. quantity of output, quality of output, or some combination of the two. In baseball, however, performance metrics are easily comparable across workers (in our case, pitchers) and firms (in our case, teams). Even though a pitch has several dimensions of quality, each provides a very clean assessment of performance, meaning pitches can be objectively assessed. Furthermore, the inherent structure of a game of baseball consisting of innings and a batting order makes it easy to identify a player's different roles. As such, this clear structure makes it easier to identify changes to performance in response to task switching. Perhaps most importantly, is that we are considering a high stakes setting where decisions have real and sizeable effects on outcomes of matches.

3. Industry Context: Baseball & Major League Baseball (MLB)

Baseball is a team sport played between two opposing teams, with each team sequentially batting and fielding. The game proceeds when a pitcher (one of nine positions on the defensive, or fielding team), standing on the pitcher's mound, throws to the batter, standing on the home plate. The batter continues to be pitched at until one of three possible outcomes: following three strikes[12], getting on base (either via hitting the ball into play, a walk, hit by pitch, or catcher's interference) or hitting a home run. The aim of the batter is to score runs by advancing around three bases and back to home plate, while the pitcher should aim to prevent the batter from reaching base or advancing.

An entire game consists of 9 innings, during which each team plays both offense and defence, and the team with the most runs at this point wins the game.[13] Each inning itself consists of two half innings; a top (first) and bottom (second) half. In the top half, the home team pitches and the away team bats, and vice versa for the bottom half. A half inning consists of three outs (three players from the batting team getting out).

[12] See section 4.1 for a full definition of a strike

[13] If the game is tied at the end of 9 innings, additional innings are played until one team is ahead at the end of a given inning.

Major League Baseball consists of 30 teams (29 from the United States and one Canadian team) who play 162 games over the course of the regular season, spanning from early April until late September. This represents an intense schedule for the teams and the players, with games taking place on a far more frequent basis than other major global sports leagues.[14] The thirty teams are split into the American League (AL), founded in 1901 and the National League (NL), founded in 1876. Since 1903, these leagues have cooperated to run a single season ending championship (the World Series), but only in 2000 did the leagues merge into a single organisation. Each league is further split into 3 divisions (East, Central and West). The winners of each division along with two wildcards from each league (teams with the best remaining Win-Loss records) go on to play in a 10-team postseason knock out tournament, culminating in the World Series, pitting the winner of the AL against the winner of the NL.

Of the 162 regular season games, the current scheduling rules are that teams play 142 games against teams from the same league. These intra-league games consist of 76 games against teams within the same division and 66 games against teams from other divisions but in the same league. The remaining 20 games are inter-league games.

The rules and regulations across the two leagues are virtually identical. There is one exception, however, crucial to our analysis in identifying performance changes due to task switching. The AL operates under the Designated Hitter (DH) rule, allowing teams in the AL to nominate a player, the DH, to fill out the batting order in place of one player in the batting order. This is the DH's only role, and they do not fill any position on defence. Pitchers are customarily poor hitters, and so it is usually them who are replaced by the DH in the batting order. The NL on the other hand, does not use this rule.[15] As such, in the NL we observe pitchers having to both pitch (their primary role) and bat to attempt to advance round bases. Whereas in the AL, pitchers only pitch; they do not bat. MLB is rare in this regard of having a such a major rule difference being applied to its teams. Minor rule differences do exist in professional football (soccer), for example, with different competitions allowing different numbers of substitutes to be used since the advent of Covid-19. But to have a different rule applied to teams in the same

[14] Of the other major global sporting leagues, teams in the National Football League play 16 games over a 4 month period between early September and late December, teams in the National Basketball League play 82 games over the 7 months from October to April, while European football (soccer) leagues run from August to May with teams playing in the region of 34-38 games.
[15] During interleague play (i.e. an AL vs NL team), the rule is operational if the game is played at an AL ballpark.

competition is a rare, albeit novel opportunity to study the demands of task switching.[16] Other baseball leagues, such as high school leagues and collegiate level baseball usually adopt some variation of the rule, so it is rare that pitchers are required to bat. The Central League, one of two leagues in Japan's Nippon Professional Baseball league is the other notable exception where pitchers are required to bat.

In MLB, the rule was originally adopted by the AL in 1973 as an experiment in the face of low offensive output. The rationale was that if pitchers were poor hitters and fans value offensive output, then this was bad news for team owners who may suffer from declining attendances. Thus, the removal of a poor hitter (the pitcher) from the batting line up would help boost attendances (Domazlicky and Kerr (1990)). The DH rule has often been a source of debate between baseball traditionalists and those who want the game to be modernised, providing a fruitful source of discussion in the media, especially when high-profile pitchers get injured batting or running bases (see for example Cassavell, 2016). Though, until recently, the NL has rarely considered this a realistic option since their last vote in 1980. The rule was not adopted on this occasion because the owner of the Philadelphia Phillies, away on a fishing trip at the time of the vote, had instructed his vice-president, Bill Giles, to vote on his behalf. However, due to a slight amendment to the introduction date of the rule, Giles was unsure how his owner would have wanted him to vote, and so, being unable to contact his owner, the Phillies abstained, and the AL and NL continued to have different rules.

In order for the DH rule to create a valid counterfactual for pitchers task switching or not, our approach requires that pitchers are (as good as) randomly affected by this rule i.e. randomly allocated to the two leagues. To put this another way, we require that teams are not selecting pitchers based on their batting ability, and only hiring based on pitching ability. We believe this to be a valid assumption, since it is unlikely that teams would hire a pitcher based on their ability to bat, a skill which pitchers rarely practice right throughout their high school and college career. Instead, teams select pitchers on their primary skillset, pitching.[17] Incidentally, average batting statistics show pitchers are somewhat worse hitters compared to other positions, as demonstrated by Table 1 below, though perhaps not as different as one may

[16] During the 2020 season, the NL approved use of the DH for first time as the MLB season was affected by the Covid-19 pandemic. The season was restricted to 60 games between July and October, and in an effort to prevent excessive fatigue during this period, the NL allowed a DH to replace the pitcher in the batting order. Our study period ends at the 2019 season however, and our results are not affected by this change. It also appears likely that the NL will adopt the DH rule as part of the new Collective Bargaining Agreement which will come into force ahead of the 2022 season, and thus the rules across the two leagues will be harmonised.
[17] Our results are robust to dropping the best pitchers in terms of their batting statistics from the sample.

anticipate. We don't see this as much an issue however, because of the argument outlined above; namely that pitchers are specialists and are hired to pitch. There are, of course, some rare exceptions to this assumption. Pitchers tend to move across to the NL later in their careers (when their ability is declining), while those pitchers who are good hitters are more valuable to NL teams. This could play a role in AL to NL trade negotiations.

Table 1: Batting Statistics by Position (2017-19)

Statistic	Non-Pitcher	Pitcher
Batting Average	0.256 (0.032)	0.247 (0.036)
Expected Batting Average	0.253 (0.027)	0.246 (0.031)
Slugging	0.433 (0.073)	0.411 (0.072)
Expected Slugging	0.428 (0.071)	0.410 (0.064)
Weighted On-Base Average	0.324 (0.039)	0.312 (0.039)
Expected Weighted On-Base Average	0.328 (0.037)	0.317 (0.035)

Standard Deviations in parentheses
Batting Average is determined by dividing a players hits by their total at-bats
Slugging (percentage) is calculated as the number of total bases divided by the number of at-bats
Weighted On-Base Average is a version of On-Base percentage accounting for how a player reached base, weighted by the relative values of each event
Expected Outcomes attempt to remove defence quality and ballpark effects
Statistics are for players with at least 200 plate appearances per season
Individual player statistics were sourced from Baseball Savant (www.baseballsavant.mlb.com)

4. Data

We examine pitch-by-pitch data for regular season MLB games for the seasons 2017, 2018 and 2019, sourced from Baseball Savant (www.baseballsavant.mlb.com). We begin our analysis in 2017 to avoid conflating changes in pitcher performance with changes in pitch measurement. Before 2017, different technology was used to record the pitch characteristics. Our analysis period ends at the 2019 season, because of the Covid-19 affected 2020 season where the season length was shortened to 60 games and teams were subjected to many temporary rule changes, including the temporary adoption of a universal DH rule. The data are nevertheless very large, with 7290 games and approximately 2.1 million individual pitches. The data include various characteristics of each pitch, most importantly to our work, velocity and location, as well as information about the outcome of each play (e.g. score, players on base). This information is captured by Trackman, a high accuracy tracking system introduced to all ballparks in 2015, replacing the camera based PITCHf/x system. Using these data, we are able to construct various outcomes of pitcher performance and define measures of both in game fatigue and task switching.

We limit our analysis to starting pitchers, a limitation that reduces our sample to about 1.3 million individual pitches. Primarily, we limit our analysis to starting pitchers because relief and closing pitchers rarely get a chance to bat or get on base, so there are few observed counterfactual opportunities. Moreover, only starting pitchers are likely to reach high enough pitch counts to be affected by severe fatigue.

4.1 Pitcher Performance

Baseball is well known for producing a multitude of statistics for evaluating player performance. Key to this study, however, is choosing outcomes that are independent (as much as possible) of the batter or luck in batting outcomes, but reflective of underlying pitching performance. One obvious choice is pitch velocity, because fatigued pitchers will not be able to throw as hard as a fully fit pitcher (Suchomel et al. (2014)). Velocity is also the outcome of choice in many sports science studies on pitcher fatigue (particularly those studying injury risk amongst pitchers e.g. Bushnell et al. (2010) and Keller et al. (2016)).

Our preferred specifications rely on samples restricted to fastballs to limit the effect to which strategy affects the results. Pitchers may purposely throw a slower pitch, such as a changeup or a curveball, after a sequence of fastballs with the aim of deceiving the batter, provoking them to swing too early and induce a bad contact. This drop in velocity is not necessarily indicative of a drop in performance. Over half of the 1.3 million pitches are categorised as a fastball, leaving us with just under 760,000 observations in the fastball sample. Figure 1 charts how likely pitchers are to throw a fastball as the game progresses. While the first pitch is almost certain to be a fastball, very quickly the probability drops to around 55-60%. Given this relative stability, our results should not be driven by pitch selection. The type of pitch is classified with the algorithm used by Statcast.[18]

We also use the location of the pitch as a measurable outcome of pitching performance, as there is a requirement to throw to certain locations in order to be successful: the strike zone. The strike zone, as defined by the Major League Baseball Rulebook is *"that area over home plate the upper limit of which is a horizontal line at the midpoint between the top of the shoulders and the top of the uniform pants, and the lower level is a line at the hollow beneath the kneecap"*. Figure 2 is the accompanying diagram (Official Baseball Rules, 2018)

[18] Specifically, Four-Seam Fastballs (code FF), Two-Seam Fastballs (FT), Sinker (SI) and Cutters (FC) are classed as fastballs.

Figure 1: Probability of throwing a fastball

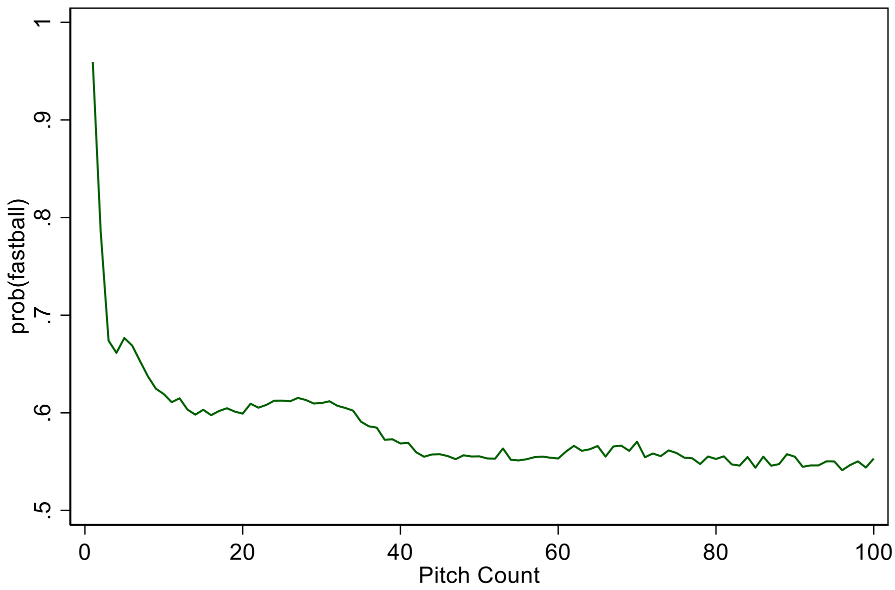
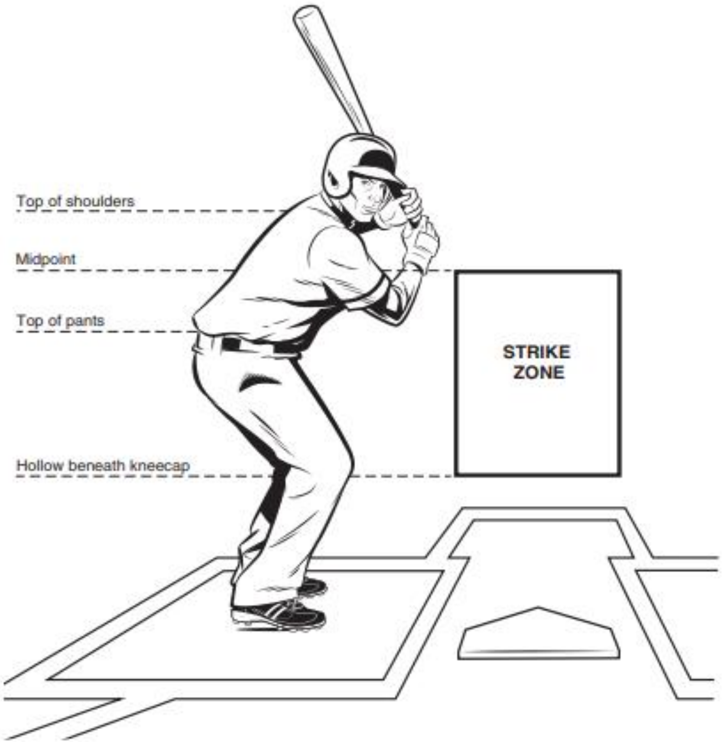


Figure 2: Definition of the Strike Zone



We define Pitch Location as the straight-line distance from the centre strike zone, calculated using the horizontal and vertical coordinates of the ball as it crosses home plate. A good pitch is considered to go through the edges of the strike zone, though not all pitches are intended to be thrown fully within this strike zone. Since our location definition could include both strikes and balls (a pitch thrown outside of the strike zone) depending on exactly how far from the centre the pitch is, we also use two binary variables to accompany this definition. The first of which, Strike, is equal to one if the pitch thrown is a strike. The second, Edge, is equal to one if the location of the pitch is within 1.5 inches either side of the edges of the strike zone. It may be advantageous for pitchers to throw pitches outside the strike zone with the intention of inducing weak contact by the batter, as pitches near the centre of the strike zone are more easily put into play by the batter. With the diameter of a baseball being 3 inches (so a radius of 1.5 inches) and Trackman measuring the location of the ball from its centre, any point of the baseball that just touches the edge of the strike zone will still be a strike. Whether this is called a strike by the umpire is a different story (see for example Mills (2014)) but having a pitcher who can throw that accurately is a sign of good performance.

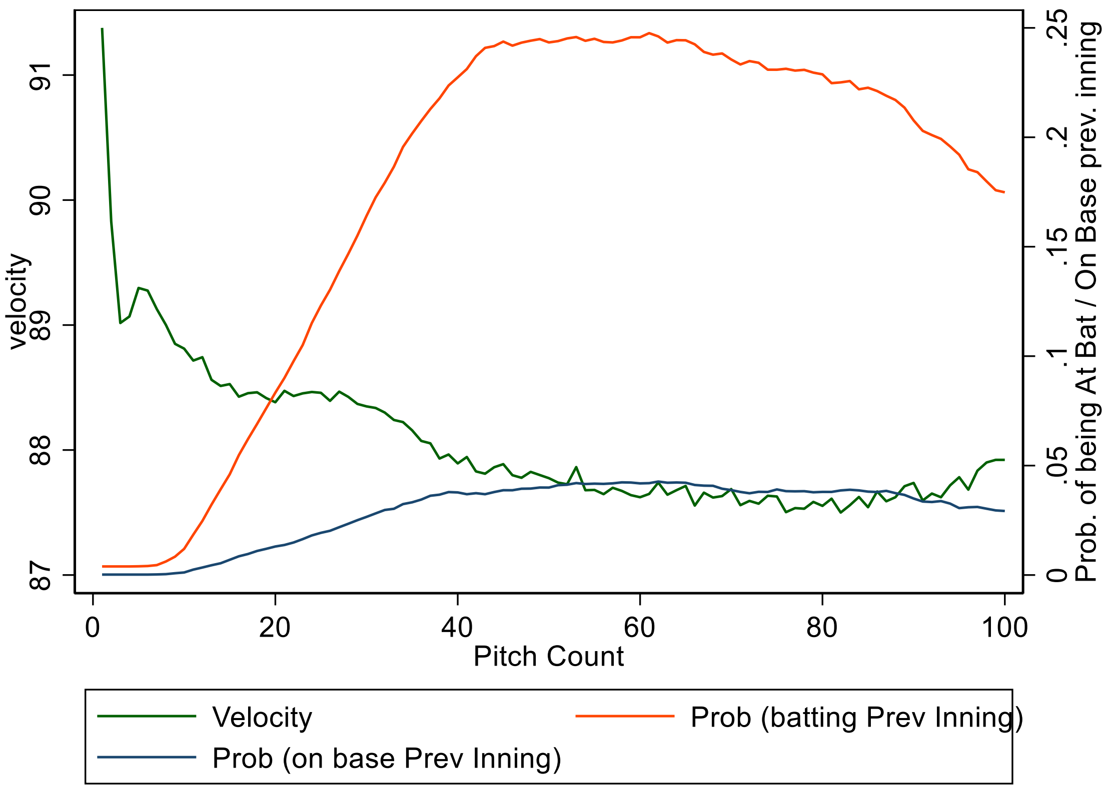
We also analyse several more objective measures of performance. Namely, whether a pitcher gives up a Walk (four pitches outside the strike zone not swung at by the batter, and the batter is awarded a first base), considered a very bad pitching outcome, whether the pitcher strikes out the batter (throws three strikes), considered a good pitching outcome, and the number of runs given up (opposition score).

4.2 Fatigue & Task Switching

To model the work done by a pitcher, we use a simple cumulative pitch count and its squared value. Our definition of task switching comes from pitchers having to bat and/or get on base during a game. In the most basic form, we define task switching using two binary variables (Pitcher Prev On Base and Pitcher Prev At Bat) identifying pitchers who have previously been at bat or on base at any point in the game up to their current pitch.

However, a drawback of this definition is that we may confuse the effects of task switching with a more general end of game fatigue effect. As an example, consider a pitcher pitching in the bottom of the 6th inning may. They might not have batted since the 3rd inning, but this would be treated as equivalent as a pitcher who batted more recently in the top of the 6th inning. It is unlikely that pitching in the 6th inning would be affected by batting in the 3rd, but it is conceivable that batting in the immediate past could have a more serious effect. As such, our preferred definition of task switching considers only task switching (previously at bat or previously on base) that occurred in the previous (half) inning.[19] This narrower definition should identify the immediate effects of task switching, if they exist, rather than potentially picking up a more general fatiguing effect due to extended play. Figure 3 graphs the how probability of batting in the previous inning (orange line, RHS scale) and the probability of getting on base (blue line, RHS scale) varies as a game progresses, along with the average velocity (green line, LHS scale). Whether we can discern any causal association between these variables is the question of the analysis that follows.

Figure 3: Average Velocity and Probability of Batting and Getting on Base



Of course, an at bat can result in several different outcomes, and what happens whilst at bat is a likely determinant of the subsequent pitching performance, rather than just batting per se. Certain outcomes are likely to involve a great deal more physical effort, such as sprinting to first base, while other outcomes may be less strenuous. As such, in Section 6.3 we offer an

[19] Defining task switching with **half** innings is key here. A pitcher pitching in the bottom of the (e.g.) 6th inning may have task switched in the top of the 6th, but a pitcher pitching in top of the 6th would have task switched in the bottom of the 5th inning.

analysis breaking down the result of the at bat into more granular events, focusing on singles, strikeouts, walks and field outs, to examine differential effects by batting outcome.

4.3 Descriptive Statistics

Table 2 below shows the descriptive statistics. Panel A is for all pitches thrown by starting pitchers, while Panel B is restricted to fastballs. The average point at which the starting pitcher is replaced is around pitch 89, with a maximum value of 134. Please see Appendix Table A1 for a breakdown of these statistics by league.

Table 2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Panel A: All Pitches (N=1,291,074)				
Pitch Count	46.72	27.93	1	134
Velocity (mph)*	88.11	5.88	40.90	101.90
Location - distance from centre of SZ**	1.14	0.63	0.00	11.32
Strike	0.46	0.50	0	1
Edge	0.18	0.38	0	1
Walk	0.02	0.14	0	1
Strikeout	0.06	0.23	0	1
Opposition Score	1.12	1.44	0	11
Prev At Bat	0.32	0.47	0	1
Prev At Bat (prev inning)	0.17	0.38	0	1
Prev On Base	0.07	0.26	0	1
Prev On Base (prev inning)	0.03	0.17	0	1
Balls	0.88	0.97	0	4
Strikes	0.89	0.82	0	2
Panel B: Fastballs (N=757,605)				
Pitch Count	44.94	28.25	1	134
Velocity (mph)$^{+}$	91.99	2.92	57.30	101.90
Location - distance from centre of SZ^{++}	1.06	0.55	0.00	9.69
Strike	0.47	0.50	0	1
Edge	0.19	0.39	0	1
Walk	0.02	0.15	0	1
Strikeout	0.04	0.20	0	1
Opposition Score	1.06	1.42	0	11
Prev At Bat	0.31	0.46	0	1
Prev At Bat (prev inning)	0.17	0.37	0	1
Prev On Base	0.07	0.26	0	1
Prev On Base (prev inning)	0.03	0.16	0	1
Balls	0.92	1.01	0	4
Strikes	0.82	0.82	0	2

Note: number of observations for velocity and location differ
* 1,285,793 ** 1,285,620 $^{+}$ 757,433 $^{++}$ 757,390

5. Estimation

Our model of pitch quality is as follows:

$$\begin{aligned} PitchQuality_{igt} &= \beta_0 + \beta_1 Pitch\ Count_{igt} + \beta_2 Pitch\ Count_{igt}^2 + \beta_3 PrevAtBat_{igt} \\ &+ \beta_4 PrevAtBat * Pitch\ Count_{igt} + \beta_5 PrevOnBase_{igt} + \beta_6 PrevOnBase \\ &* Pitch\ Count_{igt} + \beta \boldsymbol{X} + PictherFE + BatterFE + MonthFE \\ &+ BallparkFE + YearFE + \epsilon_{igt} \end{aligned}$$

such that we compare performance pre- and post-switching tasks, with pitchers in the AL, or strictly speaking, pitchers playing at AL ballparks, acting as the control group. The subscripts refer to pitcher *i*, in game g, playing for team *t*. The outcome Pitch Quality is one of several measures discussed previously, namely, Pitch Velocity, Pitch Location, Strike (0,1), Edge Pitch (0,1), Walk (0,1), Strikeout (0,1) and Opposition Runs given up. Velocity is measured at the point of release. It is possible to measure velocity at various points along the trajectory of a pitch, but these could be affected by other variables such as wind conditions, air pressure, spin etc. and as such velocity at the point of release would be the most comparable across pitches.

Prev At Bat and Prev On Base are the task switching variables and can be defined either for any point over the game up the current pitch or, our preferred definition, restricted to just the previous (half) inning. Month Fixed Effects are potentially important in explaining temperature variations across the season, where in hotter months pitchers may fatigue quicker, and could also explain a general decline in performance over the course of a season. We also control for the possible differing effects by ballparks, with different altitudes, air pressures, wind conditions etc. all possibly playing a role in the observed pitching outcomes.

Within the vector **X**, we include the number of balls and strikes that the pitcher has thrown during the current plate appearance (known as the count). These are important factors to consider since different counts are associated with favourable outcomes for either the batter or the pitcher, and thus may be associated with different levels of mental pressure. When a pitcher is faced with allowing a walk, pitchers are more likely to throw strikes down the centre, particularly fastballs. Though, when pitchers are in charge of the at bat (e.g. 0-2 count), they can be slightly riskier and aim for the extremities of the strike zone, attempting to get the batter to swing and strike out. For opposition runs, we also include indicators of whether a runner is currently standing on 1st, 2nd or 3rd base.

There are two possible issues that threaten our estimation. One is that our assumption of NL teams hiring pitchers based only on their pitching ability does not hold. The second is of within game selection of pitchers i.e. when the starting pitcher is replaced by a reliever, a point in the game when the starter is too fatigued to be effective. Teams usually have a collection of several starters that they rotate through each because of the limited number of rest days between fixtures.

It is likely that pitches we observe later in games, or in later innings, belong to pitchers who are better at dealing with the effects of fatigue, and/or simply having a good game. We address both these possibilities in our Robustness Checks in Section 6.5. To deal with the former, we exclude pitchers with the best batting statistics, which acts as a proxy for their batting ability. For the latter, we offer regressions including the lagged average inning velocity as a predictor. The starting pitcher will be pulled at some point in the game, usually when fatigue sets in and prevents them from pitching as well. By including lagged inning velocity, we can control for pitchers finishing an inning strongly and being more likely to allowed to carry on into the next inning. We also restrict the timeframe of innings over which we consider our estimations. This allows us to consider ranges of the game both where starting pitchers should not yet have been pulled, and also have a reasonable high probability of having task switched (in line with Figure 3).

6. Results

6.1 Velocity

We first present results from the velocity regressions in Table 3. It is clear that higher pitch counts are associated with declining velocity, albeit at a declining rate. Each pitch loses around 0.05-0.06 mph in velocity. To put this another way, after about 16-20 pitches, velocity has dropped by 1 mph. The squared term indicates a turning point of around 76 pitches, varying slightly by specification. This is slightly lower value than the pitch count at which starting pitchers tend to be pulled on average, which is approximately 89. For fastballs however, each additional pitch does not see the same decline in velocity (between 0.019 and 0.01 drop in velocity per fastball pitched). Though again, this occurs at a declining rate, given the positive squared term. The decline in velocity is likely capturing the gradual decline due to fatigue as the game progresses.

Moving from left to right in Table 3 we move towards our preferred specifications; in columns 4-6 using previously at bat / on base in the previous inning, and then in the final three columns restricting the sample to fastballs, where we can rule out any strategic effects. Even with the inclusion of pitcher and batter fixed effects in column 8 and then month, ballpark and year fixed effects in column 9, we observe that batting in the previous inning contributes positively to velocity, adding roughly 0.1 mph to the release speed of fastballs. Each additional pitch thrown after this gradually reduces in velocity. However, the magnitude of the interaction with pitch count is, in many specifications, extremely small compared to the uninteracted Prev At Bat, and thus the effect would seem to be long lived. In our preferred specifications 8 and 9 for example, the models suggest that between 80 and 90 pitches are required for the initial positive effect to be wiped out. Given that in these models we measure task switching across innings (which on average last around 16 pitches, std.dev=6), and that pitchers on average last around 89 pitches (std.dev=18), this interaction effect pales into sporting insignificance. There is no significant additional effect from being on base in all but one of the specifications.

Table 3: Velocity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES				All Pitches				Fastballs	
Pitch Count	-0.053***	-0.048***	-0.048***	-0.050***	-0.049***	-0.048***	-0.016***	-0.009***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Pitch Count Squared	0.470***	0.362***	0.358***	0.446***	0.358***	0.354***	0.147***	0.036***	0.032***
(coeffs x1000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Prev On Base	-0.022	0.044	0.043						
	(0.069)	(0.063)	(0.063)						
Pitch Count * Prev On Base	-0.000	-0.001	-0.001						
	(0.001)	(0.001)	(0.001)						
Prev At Bat	0.443***	0.049	0.068**						
	(0.036)	(0.033)	(0.034)						
Pitch Count * Prev At Bat	-0.003***	-0.002***	-0.002***						
	(0.001)	(0.001)	(0.001)						
Prev On Base (prev inning)				-0.234**	-0.061	-0.073	0.008	0.023	0.013
				(0.096)	(0.088)	(0.088)	(0.061)	(0.037)	(0.037)
Pitch Count * Prev On Base (prev inning)				0.003**	0.001	0.001	-0.000	-0.000	-0.000
				(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prev At Bat (prev inning)				0.306***	0.053	0.069*	0.119***	0.091***	0.105***
				(0.041)	(0.038)	(0.038)	(0.026)	(0.016)	(0.016)
Pitch Count * Prev At Bat (prev inning)				-0.001**	-0.000	-0.000	-0.001*	-0.001***	-0.001***
				(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Balls	0.653***	0.673***	0.673***	0.652***	0.673***	0.673***	0.021***	0.011***	0.011***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.004)	(0.002)	(0.002)
Strikes	-0.449***	-0.475***	-0.474***	-0.450***	-0.474***	-0.474***	0.404***	0.404***	0.404***
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.004)	(0.003)	(0.003)
Times through order	-0.450***	-0.232***	-0.240***	-0.438***	-0.230***	-0.237***	-0.158***	-0.040***	-0.045***
	(0.014)	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.009)	(0.007)	(0.007)
Constant	89.758***	89.551***	89.555***	89.732***	89.556***	89.558***	92.191***	92.002***	92.003***
	(0.019)	(0.018)	(0.018)	(0.019)	(0.018)	(0.018)	(0.012)	(0.007)	(0.007)
Observations	1,285,793	1,285,788	1,285,788	1,285,793	1,285,788	1,285,788	757,433	757,414	757,414
Pitcher & Batter FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
Month, Ballpark & Year FE	NO	NO	YES	NO	NO	YES	NO	NO	YES
R-squared	0.021	0.195	0.196	0.020	0.195	0.196	0.017	0.647	0.651

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

6.2 Other Outcomes

Next, we focus on our other indicators of pitching performance. Table 4 below displays the regression results for our three indicators of pitch location, namely Distance from the Centre of the Strike Zone (labelled Location), and our two binary variables Strikes and Edge Pitches. We analyse the latter two outcomes using a Linear Probability Model because of the desire the include the various Fixed Effects in the models, which surpasses the need for a Probit / Logit regression. Pitch location is possibly a very noisy indicator of pitcher performance, since such small differences in location can determine success or failure, but still has value in that pitches around the edge or corners of the strike zone are considered harder to hit.[20]

The effect of a higher pitch count is that pitch location gets further away from the centre of the strike zone. This result has two possible interpretations. Either that these pitches further away from the centre of the strike zone are better pitches, in that they are still within the confines of the strike zone but getting closer to the edges, or that they are now worse pitches, since they now lie outside the strike zone. This is where the analysis of Strikes and Edge Pitches is useful, and columns 3-6 in Table 4 show the latter case to be true. Higher pitch counts reduce the probability of throwing both strikes and edge pitches, indicating a lack of command or control as the game progresses.

On the effects of task switching, there appears to be very little in the way of any effect, positive or negative, from batting and running bases. The strongest predictors of locational outcomes are the number of balls and strikes (the count). A higher ball count is associated with pitches getting closer to the centre. These would be regarded as safer pitches since the pitcher does not want to give up a walk. While a higher strike count means the pitcher can afford to throw riskier pitches, with the aim of hitting the extremities of the strike zone making it harder for batters to know whether to swing and risk bad contact, or not swing and risk an out.

We next turn our attention to Table 5, where we consider Walks, Strikeouts and Opposition Score.[21] Walks occur when a pitcher throws four pitches called as balls by the umpire (i.e. outside the strike zone and not swung at by the batter), and in turn the batter is awarded a first base. These are considered a bad outcome for pitchers and is something they should look to

[20] The fastball specifications are potentially important for explaining locational outcomes too, since the types of pitches thrown over the course of a game could change resulting in pitches getting closer to the centre.
[21] In Table 5, we drop the number of pre-play balls in the count from the walk model, since a walk is awarded after 4 balls, so the pitcher must be on 3 balls for a walk to occur. Equally, we drop the number of strikes from the strikeout model since there must already be 2 strikes in the count for a strikeout to be possible.

avoid. Pitchers appear marginally less likely to give up walks after batting, which could offset the increased likelihood of giving up walks as the game progresses. This result, however, is only significant at the 10% level, and drops out of significance when using the fastball sample. Interestingly, we observe that pitchers are less likely to strikeout after batting in the previous inning. This would be considered a bad outcome for pitchers. While it is possible that this somewhat contradictory finding is simply a product of noise, we believe there is a valid explanation behind it in the context of task switching. If the act of batting keeps the pitcher active between innings (for example, preventing them from stiffening up between innings), then physical output / performance may improve (i.e. higher velocity). If the task switching is a mental task however, then this may have implications for the pitcher's decision making. The net result could be less successful outcomes, despite improved physical performance. Moreover, effort does not necessarily have to translate into improved performance. Velocity (and location to a lesser extent) is (are) directly controllable by the pitcher, but whether the pitcher strikes out the batter is also dependent on the effort and performance of the batter.

Finally, we focus on runs given up in columns 5 and 6 of Table 5. In the mind of the pitcher, their most likely objective function is to try and minimise opposition runs. The outcome variable here is the score of the opposition (batting) team measured after each pitch (rather than before a pitch). Countering the increased runs given up as the game progresses is the negative effect of batting in the previous inning. Results suggest that pitchers give up between 0.23 and 0.25 fewer runs when pitching in the inning immediately after their at bat. However, working in the opposite direction is the positive effect of previously getting on base i.e. giving up more runs after getting on base. These two effects cancel each other out to some extent, thus we test the relative size of these two coefficients in each regression. We test the null hypothesis that the sum of these coefficients is equal to zero, with results being displayed in the row labelled Test (p val). We can reject this null hypothesis at conventional levels of significance. So, pitchers who only bat and fail to get on base give up fewer runs, while the overall effect for those pitchers who do get on base, is on average, still that they give up fewer runs, just not to the same extent had they have not got on base. The implication here is that pitchers should keep active between the innings pitched i.e. bat, but to not get too fatigued by running the bases. Getting on base would appear to fatigue a pitcher and their pitching performance suffers as a result. But simply being active and not sat on the side-lines in between innings is beneficial to their subsequent pitching performance.

Table 4: Locational Outcomes	(1)	(2)	(3)	(4)	(5)	(6)
	Location		Strike (0,1)		Edge (0,1)	
VARIABLES	All Pitches	Fastballs	All Pitches	Fastballs	All Pitches	Fastballs
Pitch Count	0.170*	0.211**	-0.369***	-0.526***	-0.148***	-0.147**
(coeffs x1000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pitch Count Squared	-0.004***	-0.004***	0.005***	0.007***	0.001***	0.001**
(coeffs x1000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Prev On Base (prev inning)	-0.001	0.019	-0.003	-0.012	-0.006	-0.013
	(0.010)	(0.012)	(0.008)	(0.011)	(0.006)	(0.008)
Pitch Count * Prev On Base (prev inning)	-0.000	-0.000**	0.000	0.000**	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Prev At Bat (prev inning)	-0.006	-0.003	-0.001	0.001	0.003	0.005
	(0.004)	(0.005)	(0.004)	(0.005)	(0.003)	(0.004)
Pitch Count * Prev At Bat (prev inning)	0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Balls	-0.107***	-0.085***	0.034***	0.029***	0.010***	0.007***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
Strikes	0.158***	0.109***	-0.063***	-0.060***	-0.017***	-0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Times through order	0.021***	0.017***	-0.014***	-0.015***	-0.001	-0.001
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Constant	1.062***	1.019***	0.513***	0.522***	0.189***	0.197***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Observations	1,285,615	757,371	1,291,069	757,586	1,291,069	757,586
Pitcher & Batter FE	YES	YES	YES	YES	YES	YES
Month, Ballpark & Year FE	YES	YES	YES	YES	YES	YES
R-squared	0.055	0.045	0.019	0.021	0.003	0.004

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Walks, Strikeouts and Opposition Score	(1)	(2)	(3)	(4)	(5)	(6)
	Walk (0,1)		Strikeout (0,1)		Opposition Score	
VARIABLES	All Pitches	Fastballs	All Pitches	Fastballs	All Pitches	Fastballs
Pitch Count	0.000***	0.001***	0.001***	0.001***	0.025***	0.025***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pitch Count Squared	-0.000***	-0.000***	0.000***	0.000	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Prev On Base (prev inning)	0.000	0.002	0.000	0.001	0.185***	0.157***
	(0.002)	(0.003)	(0.004)	(0.004)	(0.020)	(0.026)
Pitch Count * Prev On Base (prev inning)	-0.000	-0.000	-0.000	0.000	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Prev At Bat (prev inning)	-0.002*	-0.002	-0.006***	-0.006***	-0.276***	-0.282***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.009)	(0.011)
Pitch Count * Prev At Bat (prev inning)	0.000	0.000	0.000***	0.000***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
On 1st					-0.077***	-0.075***
					(0.002)	(0.003)
On 2nd					0.048***	0.051***
					(0.003)	(0.004)
On 3rd					0.131***	0.137***
					(0.004)	(0.005)
Times through order	-0.010***	-0.012***	-0.023***	-0.020***	0.517***	0.510***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)	(0.005)
Balls			0.031***	0.024***	-0.009***	-0.009***
			(0.000)	(0.000)	(0.001)	(0.001)
Strikes	0.014***	0.015***			-0.022***	-0.028***
	(0.000)	(0.000)			(0.001)	(0.002)
Constant	0.009***	0.013***	0.042***	0.032***	-0.506***	-0.496***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)
Test (p val)					0.000	0.000
Observations	1,291,069	757,586	1,291,069	757,586	1,291,069	757,586
Pitcher & Batter FE	YES	YES	YES	YES	YES	YES
Month, Ballpark & Year FE	YES	YES	YES	YES	YES	YES
R-squared	0.012	0.014	0.030	0.027	0.273	0.283

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.3 Singles, Walks, Strikeouts and Field Outs

From the previous section, we have uncovered a generally, though not uniform, positive effect on pitching performance for pitchers who have previously been at bat. We have also observed an additional effect of getting on base working in the opposite direction, albeit not always significant. To this point, the definition of an At Bat and On Base has considered them to be binary events. However, batting and getting on base are more than just binary events. For example, whilst batting a batter may swing and miss at three strikes and get an out, they could be awarded a walk to first base without swinging at all, they could hit a pitch into play and sprint to first base and so on. All these events are likely to induce different physical and mental responses. As such, we continue by exploring the importance of what happens at bat, and if the pitcher does make it to base, whether it matters the way that happens (walk, hit etc.). Specifically, we focus on four batting outcomes: Singles and Walks (resulting in the batter getting to first base, but a Single likely involving more effort), and Strikeouts and Field Outs (resulting in the batter getting an out).[22]

From Table 6, it appears to matter, first if, and second how pitchers got to base. Getting a single, a fairly strenuous activity involving sprinting 90 yards from home plate to first base, is associated with a drop in velocity, though only significant at the 10% level. Perhaps more notable, getting a single is associated with giving up more runs in the following half inning when pitching, which follows from the positive Prev On Base coefficients in columns 5 and 6 of Table 5. If the pitcher gets on base via a walk, then there is less of an impact on their subsequent pitching performance. This is perhaps not surprising given that a walk is less strenuous than getting a single. However, in cases where the pitcher does not get to base (strikeouts and field outs), velocity improves, certainly for the latter, which follows from results in Table 3, while on average 0.2 and 0.25 fewer runs are given up in the following half inning after these events, which follows from the negative Prev At Bat coefficients from columns 5 and 6 of Table 5.

[22] Statcast lists a total of 32 different outcomes following a plate appearance, however, some are so rare that we would gain very little by examining them. These four outcomes (singles, strikeouts, walks and field outs) are four outcomes that are a combination of the most common and interesting events to examine.

Table 6: Splitting the result of an At Bat PREVIOUS BATTING EVENT		(1) Velocity	(2) Location	(3) Strike (0,1)	(4) Edge (0,1)	(5) Walk (0,1)	(6) Strikeout (0,1)	(7) Opposition Score
Single	All Pitches	-0.196*	0.006	0.003	-0.009	-0.002	-0.007	0.090***
		(0.113)	(0.013)	(0.010)	(0.008)	(0.003)	(0.005)	(0.026)
	Fastballs	-0.024	0.020	-0.010	-0.011	-0.000	-0.007	0.089***
		(0.048)	(0.015)	(0.014)	(0.011)	(0.004)	(0.006)	(0.033)
Walk	All Pitches	0.145	-0.019	-0.013	0.001	-0.006	-0.013	0.077*
		(0.190)	(0.022)	(0.018)	(0.014)	(0.005)	(0.008)	(0.044)
	Fastballs	0.189**	0.021	-0.025	-0.003	-0.003	-0.010	-0.014
		(0.080)	(0.025)	(0.023)	(0.018)	(0.007)	(0.009)	(0.055)
Strikeout	All Pitches	0.015	-0.013**	0.003	0.004	-0.005***	-0.009***	-0.207***
		(0.052)	(0.006)	(0.005)	(0.004)	(0.001)	(0.002)	(0.012)
	Fastballs	0.057***	-0.013*	0.011*	0.005	-0.005**	-0.005**	-0.219***
		(0.022)	(0.007)	(0.006)	(0.005)	(0.002)	(0.003)	(0.015)
Field Out	All Pitches	0.124**	-0.000	-0.001	0.002	-0.002	-0.009***	-0.253***
		(0.060)	(0.007)	(0.006)	(0.004)	(0.002)	(0.003)	(0.014)
	Fastballs	0.083***	0.010	-0.004	0.006	-0.002	-0.010***	-0.253***
		(0.025)	(0.008)	(0.007)	(0.006)	(0.002)	(0.003)	(0.017)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Each coefficient is a result of a separate regression. All regressions include Pitcher, Batter, Month, Ballpark and Year FE

6.4 Interleague Play

As a final demonstration of the benefits of task switching, we offer an analysis of the performance of pitchers during interleague (IL) games. To do so, we first restrict analysis to pitchers pitching away from home to remove any familiarity effects associated with playing at home i.e., home advantage. By doing so we compare performance of away pitchers in intra-league games (AL@AL or NL@NL) to their performance in inter-league games (AL@NL or NL@AL).[23] The analysis has two aspects to it since the use of the DH rule is determined by the identity of the home team. Thus, IL games played in AL (NL@AL) ballparks will use the DH rule, and NL pitchers who are used to having to task switch will now only be required to pitch since their position in the batting order can now be filled by the DH. In IL games played at NL ballparks (AL@NL) however, the DH rule will not be active, so AL pitchers who are used to only pitching are now required to bat as well. We can now analyse the effects of dropping a familiar task and also of adopting an unfamiliar task. Results are shown in Table 7 below.

[23] The use of the @ symbol is how matchups are denoted in MLB. It is quite literally saying the away team playing ‘at’ the venue of the home team.

Table 7: Interleague Play	FE	Fastballs	Outcome	Prev At Bat (prev inning)	Prev On Base (prev inning)	Observations
Panel A	Pitcher, Batter	No	Velocity	-0.201	0.560*	314,973
AL@NL	+Month, Ballpark, Year	No		-0.127	0.442	314,973
Pitcher is *adopting batting*	Pitcher, Batter	Yes		0.007	0.017	182,232
in the IL games	+Month, Ballpark, Year	Yes		0.039	-0.094	182,232
	Pitcher, Batter	No	Opposition	-0.034	0.075	315,959
	+Month, Ballpark, Year	No	Score	-0.023	0.030	315,959
	Pitcher, Batter	Yes		-0.071*	0.005	182,254
	+Month, Ballpark, Year	Yes		-0.064	-0.037	182,254
Panel B	Pitcher, Batter	No	Velocity	0.104*	-0.004	322,115
NL@AL	+Month, Ballpark, Year	No		0.108**	-0.014	322,115
Pitcher is giving *up batting in the*	Pitcher, Batter	Yes		0.045**	0.072	192,492
IL games	+Month, Ballpark, Year	Yes		0.046**	0.066	192,492
	Pitcher, Batter	No	Opposition	-0.200***	0.175***	323,749
	+Month, Ballpark, Year	No	Score	-0.197***	0.172***	323,749
	Pitcher, Batter	Yes		-0.205***	0.140***	192,602
	+Month, Ballpark, Year	Yes		-0.204***	0.138***	192,602

The results raise an interesting asymmetry, in that it appears to matter if pitchers are accustomed to switching between pitching and batting or not. Generally, AL pitchers playing in IL games at NL ballparks (Panel A) do not suffer any significant adverse effects, either in terms of reduced velocity or runs given up following batting compared to playing away from home but in other AL ballparks where they are not required to bat. However, NL pitchers throw faster pitches and give up fewer runs after batting in away games within their own league, compared to the case where they are not required to bat in away games played in AL ballparks. In other words, batting appears to be beneficial, but only to those pitchers who are accustomed to taking on the additional demands associated with task switching.

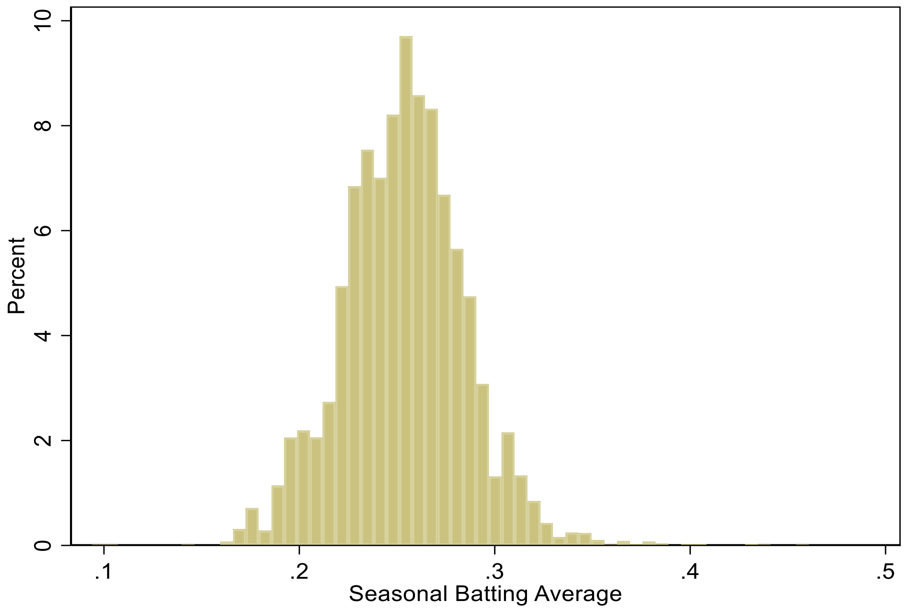
6.5 Robustness Checks

6.5.1 Selection of Pitchers to Leagues

To this point, our analysis has rested on the assumption that pitchers are not hired by teams based on their batting ability, and instead are hired only on their pitching ability. By making this assumption, we can say that pitchers are as good as randomly allocated to the two leagues, and in turn, randomly affected by the Designated Hitter forcing some to task switch. In the main, we believe this to be a perfectly reasonable assumption. Pitchers tend to be poor hitters and it is a skill that they rarely (if at all) train from high school all the way up to and including their professional careers. There are of course exceptions to this rule, albeit rare exceptions; some pitchers may be good hitters, and exceptional batting ability could play a role in AL to NL trade negotiations, where these pitchers will be of greater value to teams in the NL.

To test this assumption, we check the robustness of our results to excluding the best pitchers in terms of their batting statistics, with results shown in the Appendix Table A2. Specifically, we exclude any pitcher whose seasonal batting average was above 0.300.[24] This value was chosen as it is widely considered to a benchmark for very good batting. It has also been shown to be an important reference point that baseball players try to reach (Pope and Simonsohn (2011) and Tanji (2021)).[25] Figure 4 shows the histogram for pitcher's batting averages in each of the three seasons under consideration. Dropping players whose seasonal batting average is above 0.300 reduces the number of individual pitches under consideration by approximately 83,000, and cuts 117 pitchers from the sample, or 133 pitcher-season combinations. Regression results using this reduced sample are extremely similar to the results as shown in Table 3 (for Velocity) and Table 5 (for Opposition Score), meaning we can be confident that our results are unlikely to be driven by teams in the NL selecting pitchers based on their batting ability.

Figure 4: Histogram of Pitchers' Seasonal Batting Average



[24] Batting Average is calculated by dividing a player's total hits by his total at-bats, producing a statistic between 0.000 and 1.000 (reported to 3dp).

[25] Notice in Figure 4 the 'dip' at 0.300 and the bunching just above 0.300. This is precisely the reference point. Batters who are close to this point near the end of the season will try to finish with a BA of just greater than 0.300. The work by Tanji shows this is a reference point not motivated by monetary incentives.

6.5.2 Within Game Selection

The second threat to our identification comes from a selection bias arising from some pitchers being able to last longer in games. These pitchers may be of better ability, or just simply better able to deal with the effects of fatigue. The effect could be pitcher specific (across games) or pitcher-game specific (within game i.e. a pitcher is just having a good game). Either way, it could be that the positive effect we observe from batting in the previous inning is just a by-product of us observing these more robust pitchers lasting longer in games before replacement. This point is highlighted by the apparent upturn in velocity in Figure 3, where average pitch velocity increases slightly after around pitch 80. What we are likely observing here is observations coming from pitchers who are lasting longer in games before replacement because they are less fatigued.

We address this concern with two separate approaches. First, in Appendix Table A3 we present regressions similar to those in Table 3, but with the addition of the average velocity from the previous inning. The logic of including this variable is that poor performance in the previous inning should be associated with being pulled. Alternatively, pitchers who finish the last inning strongly will be more likely to be allowed to continue. The effect of batting in the previous inning remains positive and highly significant even with the inclusion of the lagged average inning velocity.

Our second, preferred, check involves restricting the inning numbers over which we consider our estimations. By removing later innings, we (roughly speaking) remove any pitchers who are having a very good game and lasting longer than usual, while by removing early innings, we exclude the early part of the game where there is a very low probability that pitchers have been given an opportunity to bat (see Figure 3). For reference, in the unrestricted sample, the last inning a starting pitcher appeared in was the 9^{th} inning, though usually, starting pitchers last until around the 6^{th} inning before being pulled. By running our models with the various inning restrictions in place (see Appendix Table A4), our results stay largely intact. This is especially true when we remove the first inning from consideration, a period of the game when pitchers are unlikely to appear at bat since they tend to be placed at the bottom of the batting order. Of note for our concerns over in game selection of pitchers, however, is that the removal of later innings does not dampen the effect of previously batting.

6.5.3 Placebo Check

As a final robustness check, we carry out a placebo style test, with results shown in Appendix Table A5. This test has 2 aims; first, and most importantly, to check that we do not find an effect of batting and / or getting on base when there should not be one. Second, we can rule out any anticipatory effects of pitchers who are up to bat in the forthcoming inning. To do this, we assign the at bat / on base to the inning before it happened. For example, a pitcher pitching in the bottom of the third inning having batted in the top of the third would ordinarily be assigned a value of 1 for the task switching variable Prev At Bat (prev. inning) when pitching in the bottom of the third. In the placebo check, we also (falsely) assign an at bat to the pitches thrown in the bottom of the second inning i.e. before they actually batted, and these are indicated with the addition of FALSE after the variable name. We repeat the procedure for getting on base as well.

Given these variables will be occurring earlier in games than the true at bat / on base, one may expect these coefficients to be positive, simply picking up the effects of a period of the game where pitchers are as fatigued. Thus, our placebo test involves comparing the coefficients of the FALSE and the true task switching variables. The row labelled "Test of equality (p value)" is testing the null hypothesis that Prev At Bat (prev inning) is equal to Prev At Bat FALSE. In our preferred fastball specifications, the null of equality of these coefficients is rejected. This gives us additional confidence in our results, in that we are finding a significant effect from task switching when we would expect to find one.

7. Discussion & Conclusion

Attempting to quantify the effects of task switching on short term (in our setting, that translates to within game) fatigue and productivity is not a straightforward task, not least due to difficulties in defining and comparing performance. Using play by play data from three seasons of MLB, we can overcome this difficulty and have shown task switching in the form of batting in the previous (half) inning results in largely beneficial effects on pitching performance.

In our preferred specifications, relying on fastballs and the inclusion of pitcher, batter, month, ballpark and year fixed effects, the average fastball velocity increased by up to 0.1 mph, and pitchers gave up 0.25 fewer runs in the half inning immediately following the at bat. At first this result may seem counterintuitive under the prior assumption that switching between batting and pitching may incur a switching cost and additional physical exertion whilst batting. However, we would not be the first empirical paper to find evidence that *some* task switching can be beneficial to performance. Namely, Singh (2014) found that up to about four patients per hour, physicians performance improved for each additional patient. Only after this point did the extra demands from task switching hinder performance. Moreover, if we are to assume that having to switch tasks within games creates a more challenging working environment, then according to Hommel et al. (2012), there is both behavioural and neuroscientific evidence that when faced with increasing difficulty of tasks, subjects increase their effort to compensate for and overcome that challenge. This phenomenon has also been shown experimentally by Srna et al. (2018). Our results are robust to accounting for pitcher ability, and to the consideration of within game selection of pitchers.

However, this improvement in performance after task switching was not uniform across all our outcomes. Pitchers were less likely to strike out the batter after batting, which we believe could highlight the importance of distinguishing between mental and physical effects due to task switching. Some outcomes also show a decline in performance after the pitcher gets on base, though the effect is not large enough to outweigh the initial positive effect of batting. So, on average, the overall effect from task switching remains positive. The practical implication for baseball teams and baseball managers is that they should be keen for their pitchers to go and bat, just to keep their minds active and not sit around on the side-lines, but at all costs should tell them to avoid them getting on base.

As for how we can explain these results in a baseball setting, it is possible that the switch between pitching and batting offers pitchers an opportunity to recuperate both mentally and physically. It could be for example, that batting acts as a distraction from the core task. A pitcher between innings but not batting would have more time to dwell on any previous mistakes which in their mind would eventually lead to replacement. Batting could simply reduce any mental stress associated with pitching. There could also be a physical reason; if pitchers begin to stiffen up whilst between innings, then batting may help loosen their joints and muscles in preparation for pitching. Running bases, however, may result in too much physical exertion overall, particularly if pitchers are left on base at the end of the half inning just before they are required to switch immediately back to pitching. Further work would be required to identify the channel of causality.

This does raise an interesting dilemma for MLB teams in the future. It appears increasingly likely that the NL will permanently adopt the DH rule following its temporary use for the Covid-19 affected 2020 season, especially with the current collective bargaining agreement (the agreement between players and the league) coming to an end in December 2021. On the one hand, pitcher performance may be harmed if the task switching element of a pitcher's game is taken away by the NL's adoption of the DH rule, since we have shown that task switching is beneficial to their performance. On the other hand, the removal of the requirement for pitchers to bat would remove a level of complication and strategy for coaches in the NL to consider. The 'Double Switch' is one such example of this.

As for how generalisable these results are to other sports and indeed other industries, there is certainly scope to abstract away from baseball. Cricket would provide an interesting sporting parallel. While other sports such as football (soccer) and rugby do involve players carrying out different roles (i.e. attacking and defending), the sequential nature of these tasks is not as well defined as in baseball. More generally, a scenario where temporarily moving away from one's main task would fit the same story. For example, an academic researching their work may attend seminars in a different area. Other occupations, particularly those involving manual or physical labour, may more be more applicable to our results from the baseball setting.

APPENDIX

Table A1: Descriptive Statistics by League

Variable	National League Mean	American League Mean	Difference in means (p val)
	N=650,229	N=640,845	
Pitch Count	46.587	46.847	0.000
Velocity (mph)*	88.229	87.987	0.000
Location - distance from centre of SZ**	1.137	1.147	0.000
Strike	0.462	0.456	0.000
Edge	0.179	0.179	0.430
Walk	0.020	0.020	0.294
Strikeout	0.056	0.054	0.000
Opposition Score	1.100	1.138	0.000
Prev At Bat	0.630	0.000	0.000
Prev At Bat (prev inning)	0.338	0.000	0.000
Prev On Base	0.145	0.000	0.000
Prev On Base (prev inning)	0.056	0.000	0.000
Balls	0.877	0.887	0.000
Strikes	0.891	0.890	0.366

Note: number of observations for Velocity and Location differ
* 647,102 (NL), 638,691 (AL)
** 647,011 (NL), 638,609 (AL)

Table A2: Removing Pitchers with BA>0.300	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Velocity				Opposition Score			
VARIABLES	All Pitches		Fastballs		All Pitches		Fastballs	
Pitch Count	-0.048***	-0.048***	-0.009***	-0.009***	0.024***	0.024***	0.023***	0.023***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pitch Count Squared	0.000***	0.000***	0.000***	0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Prev On Base (prev inning)	-0.075	-0.076	0.010	0.015	0.199***	0.199***	0.168***	0.170***
	(0.092)	(0.092)	(0.039)	(0.039)	(0.021)	(0.021)	(0.027)	(0.027)
Pitch Count * Prev On Base (prev inning)	0.001	0.001	-0.000	-0.000	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Prev At Bat (prev inning)	0.069*	0.085**	0.084***	0.099***	-0.279***	-0.279***	-0.293***	-0.292***
	(0.039)	(0.040)	(0.017)	(0.017)	(0.009)	(0.009)	(0.011)	(0.011)
Pitch Count * Prev At Bat (prev inning)	-0.000	-0.000	-0.001***	-0.001***	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Balls	0.684***	0.684***	0.014***	0.014***	-0.006***	-0.006***	-0.006***	-0.006***
	(0.006)	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Strikes	-0.477***	-0.477***	0.412***	0.411***	-0.023***	-0.022***	-0.027***	-0.027***
	(0.007)	(0.007)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)
Times through order	-0.229***	-0.230***	-0.036***	-0.035***	0.503***	0.504***	0.495***	0.496***
	(0.016)	(0.016)	(0.007)	(0.007)	(0.004)	(0.004)	(0.005)	(0.005)
Constant	89.622***	89.622***	92.052***	92.051***	-0.492***	-0.492***	-0.483***	-0.482***
	(0.019)	(0.019)	(0.008)	(0.008)	(0.004)	(0.004)	(0.005)	(0.005)
Observations	1,150,287	1,150,287	678,316	678,316	1,155,136	1,155,136	678,475	678,475
Pitcher & Batter FE	YES	YES	YES	YES	YES	YES	YES	YES
Month, Ballpark & Year FE	NO	YES	NO	YES	NO	YES	NO	YES
R-squared	0.193	0.193	0.647	0.649	0.260	0.264	0.270	0.275

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3: Including Lagged Inning Velocity	(1)	(2)	(3)	(4)
	Velocity			
VARIABLES	All Pitches		Fastballs	
Pitch Count	-0.016***	-0.016***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
Pitch Count Squared	0.000***	0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Prev On Base (prev inning)	-0.013	-0.022	0.064*	0.055
	(0.088)	(0.088)	(0.037)	(0.037)
Pitch Count * Prev On Base (prev inning)	0.000	0.000	-0.001*	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Prev At Bat (prev inning)	0.096**	0.123***	0.154***	0.175***
	(0.039)	(0.039)	(0.016)	(0.016)
Pitch Count * Prev At Bat (prev inning)	-0.000	-0.000	-0.001***	-0.001***
	(0.001)	(0.001)	(0.000)	(0.000)
Balls	0.721***	0.721***	0.016***	0.017***
	(0.006)	(0.006)	(0.002)	(0.002)
Strikes	-0.394***	-0.394***	0.409***	0.409***
	(0.007)	(0.007)	(0.003)	(0.003)
Lagged Inning Velocity	0.361***	0.354***	0.233***	0.224***
	(0.003)	(0.003)	(0.001)	(0.001)
Times through order	-0.223***	-0.228***	0.036***	0.031***
	(0.016)	(0.016)	(0.007)	(0.007)
Constant	56.521***	57.192***	70.806***	71.642***
	(0.274)	(0.276)	(0.117)	(0.117)
Observations	1,035,694	1,035,694	594,485	594,485
Pitcher & Batter FE	YES	YES	YES	YES
Month, Ballpark & Year FE	NO	YES	NO	YES
R-squared	0.198	0.199	0.661	0.664

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A4: Limiting the timeframe over which we consider the effect of Batting / Getting on base

Inning restriction	Prev At Bat	Prev On Base	Prev At Bat (fastballs)	Prev On Base (fastballs)
1-7	0.057	-0.047	0.090***	0.026
1-6	0.084**	-0.057	0.080***	0.011
1-5	0.090**	-0.053	0.025	0.034
1-4	0.091*	0.119	-0.026	0.140***
1-3	0.187***	0.196	-0.032	0.218***
1-2	0.251***	0.119	-0.075**	-0.140
2-7	0.073*	-0.034	0.148***	0.045
2-6	0.091**	-0.038	0.146***	0.034
2-5	0.093**	-0.030	0.108***	0.060
2-4	0.082	0.159	0.090***	0.182***
3-6	0.117**	-0.158	0.176***	-0.054
4-6	0.204**	0.131	0.216***	-0.090

*** p<0.01, ** p<0.05, * p<0.1

Table A5: Placebo Test	(1)	(2)	(3)	(4)
	Velocity			
VARIABLES	All Pitches		Fastballs	
Pitch Count	-0.049***	-0.048***	-0.008***	-0.009***
	(0.001)	(0.001)	(0.000)	(0.000)
Pitch Count Squared	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Prev On Base (prev inning)	-0.062	-0.077	0.026	0.025
	(0.088)	(0.088)	(0.037)	(0.037)
Pitch Count * Prev On Base (prev inning)	0.001	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Prev At Bat (prev inning)	0.056	0.082**	0.082***	0.105***
	(0.038)	(0.038)	(0.016)	(0.016)
Pitch Count * Prev At Bat (prev inning)	-0.000	-0.000	-0.001***	-0.001***
	(0.001)	(0.001)	(0.000)	(0.000)
Prev At Bat (prev inning) FALSE	0.003	0.025	-0.024***	-0.001
	(0.016)	(0.016)	(0.007)	(0.007)
Prev On Base (prev inning) FALSE	0.013	0.008	0.022*	0.022*
	(0.031)	(0.031)	(0.013)	(0.013)
Balls	0.673***	0.673***	0.011***	0.012***
	(0.005)	(0.005)	(0.002)	(0.002)
Strikes	-0.474***	-0.474***	0.404***	0.403***
	(0.006)	(0.006)	(0.003)	(0.003)
Times through order	-0.230***	-0.237***	-0.040***	-0.039***
	(0.015)	(0.015)	(0.007)	(0.007)
Constant	89.556***	89.556***	92.004***	92.000***
	(0.018)	(0.018)	(0.007)	(0.007)
Test (p val)	0.177	0.142	0.000	0.000
Observations	1,285,788	1,285,788	757,414	757,414
Pitcher & Batter FE	YES	YES	YES	YES
Month, Ballpark & Year FE	NO	YES	NO	YES
R-squared	0.195	0.196	0.647	0.649

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Chapter References

Aral, S., Brynjolfsson, E. and Van Alstyne, M. (2012), *'Information, Technology and Information Worker Productivity'*, Information Systems Research, Vol.23, No.3, Part 2, pp 849-867

Buser, T. and Peter, N. (2012), *'Multitasking'*, Experimental Economics, Vol.15, No.4, pp 641-655

Bushnell, B.D., Anz, A.W., Noonan, T.J., Torry, M.R. and Hawkins, R.J. (2010), *'Association of Maximum Pitch Velocity and Elbow Injury in Professional Baseball'*, The American Journal of Sports Medicine, Vol.38, No.4, pp 728-732

Cassavell, A.J. (2016), *'Universal Truth? Execs talk NL DH possibility'*, from MLB News, https://www.mlb.com/news/might-dh-be-in-national-league-s-future-c162124094 (Accessed 03/06/2020)

Collewet, M. and Sauermann, J. (2017), *'Working Hours and Productivity'*, Labour Economics, Vol.47, pp 96-106

Coviello, D., Ichino, A. and Perisco, N. (2015), *'The Inneficiency of Worker Time Use'*, Journal of the European Economic Association, Vol.13, No.5, pp 906-947

Crocker, T.D. and Horst, R.L. (1981), *'Hours of Work, Labour Productivity, and Environmental Conditions: A Case Study'*, The Review of Economics and Statistics, Vol.63, No.3, pp 361-368

Domazlicky, B.R. and Kerr, P.M. (1990), *'Baseball Attendance and the Designated Hitter'*, The American Economist, Vol.34, No.1, pp 62-68

Entine, O.A. and Small, D.S. (2008), *'The role of Rest in the NBA Home-Court Advantage'*, Journal of Quantitative Analysis in Sports, Vol.4, No.2

Escamilla, R.F., Barrentine S.W. and Fleisig G.S. (2007), *'Pitching Biomechanics as a Pitcher Approaches Muscular Fatigue during a Simulated Baseball Game'*, The American Journal of Sports Medicine, Vol.35, No.1, pp 23-33

Hart, R.A. (2004), *'The Economics of Overtime Working'*, Cambridge University Press

Hommel, B., Fischer, R., Colzato, L.S., van den Wildenberg, W.P. and Cellini, C. (2012), *'The Effect of fMRI (noise) on Cognitive Control'*, Journal of Experimental Psychology: Human Perception and Performance, Vol.32, No.2, pp 290-301

Keller, R.A., Marshall, N.E., Guest, J.M., Okoroha, K.R., Jung, E.K. and Moutzouros, V. (2016), *'Major League Baseball Pitch Velocity and Pitch Type Associated With the Risk of Ulnar Collateral Ligament Injury'*, Journal of Shoulder and Elbow Surgery, Vol.25, No.4, pp 671-675

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A.M. and Koch, I. (2010), *'Control and Interference in Task Switching – A Review'*, Psychological Bulletin, Vol.136, No.5, pp 849-874

Lyman S., Fleisig, G.S., Andrews, J.R. and Osinski, E.D. (2002), *'Effect of Pitch Type, Pitch Count, and Pitching Mechanics on Risk of Elbow and Shoulder Pain in Youth Baseball Pitchers'*, The American Journal of Sports Medicine, Vol.30, No.4, pp 463-468

Lu, S.F. and Lu, L.X. (2017), *'Do Mandatory Overtime Laws Improve Quality? Staffing Decisions and Operational Flexibility of Nursing Homes'*, Management Science, Vol.63, No.11, pp 3566-3585

Marcora, S.M., Staiano, W. and Manning, V. (2009), *'Mental Fatigue Impairs Physical Performance in Humans'*, Journal of Applied Psychology, Vol.106, No.3, pp 857-864

Mills, B. (2014), *'Expert Workers, Performance Standards, and On-the-Job Training: Evaluating Major League Baseball Umpires'*, Available at SSRN

Murray, T.A., Cook, T.D., Werner, S.L., Schlegel, T.F. and Hawkins, R.J. (2001), *'The Effects of Extended Play on Professional Baseball Pitchers'*, The American Journal of Sports Medicine, Vol.29, No.2, pp 137-142

Nichols, M.W. (2014), *'The Impact of Visiting Team Travel on Game Outcome and Biases in NFL Betting Markets'*, Journal of Sports Economics, Vol.15, No.1, pp 78-96

Oberhofer, H., Philippovich, T. and Winner, H. (2010), *'Distance Matters in Away Games: Evidence from the German Football League'*, Journal of Economic Psychology, Vol.31, No.2, pp 200-211

Official Baseball Rules (2018). Available Online http://mlb.mlb.com/documents/0/8/0/268272080/2018_Official_Baseball_Rules.pdf (Accessed 15/09/2020)

Pencavel, J. (2015), *'The Productivity of Working Hours'*, The Economic Journal, Vol.125, No. 589, pp 2052-2076

Pope, D. and Simonsohn, U. (2011), *'Round Numbers as Goals Evidence From Baseball, SAT Takers, and the Lab'*, Psychological Science, Vol.22, No.1, pp 71-79

Rampinini, E., Impellizzeri, F.M., Castagna, C., Coutts, A.J., and Wisløff, U. (2009), *'Technical Performance During Soccer Matches of the Italian Serie A League: Effect of Fatigue and Competitive Level'*, Journal of Science and Medicine in Sport, Vol.12, No.1, pp 227-233

Rubinstein, J.S., Meyer, D.E. and Evans, J.E. (2001), *'Executive Control of Cognitive Processes in Task Switching'*, Journal of Experimental Psychology: Human Perception and Performance, Vol.27, No.4, pp 769-797

Russ, M. and Crews, D.E. (2014), *'A Survey of Multitasking Behaviors in Organizations'*, International Journal of Human Resource Studies, Vol.4, No.1, pp 137-153

Scoppa, V. (2013), *'Fatigue and Team Performance in Soccer: Evidence From the FIFA World Cup and the UEFA European Championship'*, Journal of Sports Economics, Vol.16, No.5, pp 482-507

Singh, D. (2014), *'Does Multitasking Improve Performance? Evidence from the Emergency Department'*, Manufacturing & Service Operations Management, Vol.16, No.2, pp 168-183

Srna, S., Schrift, R.Y. and Zauberman, G. (2018), *'The Illusion of Multitasking and its Positive Effect on Performance'*, Psychological Science, Vol.29, No.12, pp 1942-1955

Suchomel, T.J. and Bailey, C.A. (2014), *'Monitoring and Managing Fatigue in Baseball Players'*, Strength & Conditioning Journal, Vol.36, No.6, pp 39-45

Tanji, R. (2021), *'Reference Dependence and Monetary Incentives: Evidence from Major League Baseball'*, Discussion Papers in In Economics and Business, No. 20-23

Race and Coaching Hierarchy: An Analysis of Hiring and Firing in the National Football League.

Alex Farnell (Lancaster University)

Dave Berri (Southern Utah University)

Vincent O'Sullivan (Lancaster University)

Robert Simmons (Lancaster University)

Abstract

Despite its best efforts, the National Football League (NFL) has long been criticised for the lack of minority leadership amongst its teams. Recent hires and non-hires have only served to heighten this criticism. To assess this, we use a new, rich and unique dataset to examine the relationship between race and coaching hierarchy in the NFL. Our results indicate that young, experienced and well performing coordinators are likely to be promoted to Head Coach while older and poorly performing coaches are more likely to be fired. Conditional on reaching the position of Head Coach, race does not seem to play a role in either promotions or firings. In the post Rooney Rule era (post 2003) however, black coordinators are marginally more likely to be promoted than previously. Black Head Coaches on the other hand, are neither more nor less likely to find a job at the same level. The Rooney Rule has been successful to the extent that teams now consider (and ultimately appoint) equally skilled black coordinators to Head Coaching jobs, despite our evidence suggesting that equally skilled black coordinators had always been available.

Keywords: racial discrimination, NFL, coaches

JEL Classification: J71, Z21, Z22

Acknowledgements

We thank seminar participants at Lancaster University and at the European Sports Economics Association (ESEA) in Gijon for helpful advice and comments.

1. Introduction & Background

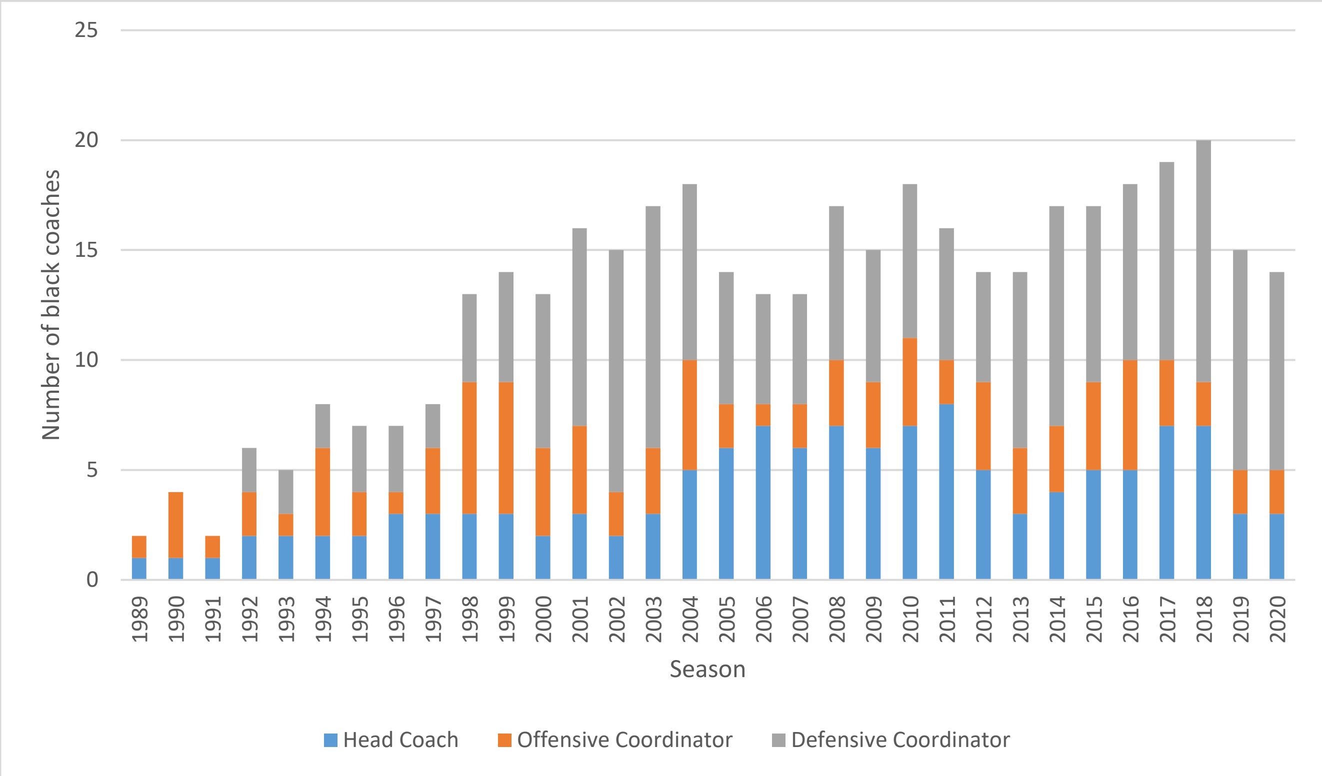
The 2021 hiring cycle in the National Football League (NFL) once again drew considerable attention after several minority candidates were seemingly overlooked for Head Coaching positions. Perhaps most notably, the minority candidate Eric Bienemy, Offensive Coordinator of the Kansas City Chiefs was overlooked for a second successive year, despite heading up one of the league's most productive offenses, and coaching arguably the league's best Quarterback, Patrick Mahomes. This was despite the league's continued attempts to promote diversity on coaching staffs through its affirmative action policy, the Rooney Rule (more on this in Section 3) and even incentivising teams with draft picks for making minority hires.

Despite marked progress over the last 30 years in this regard, the last few seasons have highlighted that there is still a long way to go. Since 2017, teams have gone from employing a joint high number of black Head Coaches (seven) to a joint low since the inception of the Rooney Rule in 2003. Ahead of the 2021 season, this left just Brian Flores (of the Miami Dolphins) and Mike Tomlin (Pittsburgh Steelers) as the league's only black Head Coaches.[26] Figure 1 charts how the composition of minority candidates in the top three coaching positions (Head Coach, and Offensive and Defensive Coordinators) has changed since 1989, when Art Shell became the NFL's first black Head Coach.

The prospects for minority coaches do seem to be improving over said period, and a couple of notable years can be picked from this timeframe. The 2003 season saw the introduction of Rooney Rule, aimed at increasing minority representation amongst Head Coaches. The rule, named after former Pittsburgh Steelers owner Dan Rooney, requires teams interview at least one minority candidate for the role of Head Coach. While its success is still widely debated by academics, analysts, journalists and even coaches themselves, one thing that is certain from Figure 1; consistently fewer black Head Coaches were in jobs in the years before the rule was introduced. This perhaps hints at some degree of success. Another notable event is the 2007 Super Bowl (the end of the 2006 season), which was the first to be contested between two minority coaches, where Tony Dungy's Indianapolis Colts beat Lovie Smith's Chicago Bears.

[26] This was as of the end of the 2020 season, following the customary rounds of firing. During the off season, the Houston Texans hired David Culley, leaving the number of black Head Coaches for the 2021 season at 3, plus Ron Rivera at Washington who is of Hispanic origin, and Robert Saleh, hired by the New York Jets, who is of Lebanese descent.

Figure 1: Number of Black Coaches per season in NFL, 1989- 2020



While relatively few coaching positions are held by black (and other minority) candidates, the composition of the playing staff who are black is an entirely different story. Around 70% of players in the NFL are black, a proportion which is higher when considering defensive positions (this is one possible explanation as to the greater numbers of Black DCs shown in Figure 1).[27] Herein lies the root of the widespread attention and criticism – in a league of predominantly black players, and players presumably making ideal coaching candidates, how come so few coaches are black?[28]

Of course, the argument is not quite as clear cut as that. The lack of black coaches in the NFL is not itself a sign of discrimination. Only if minority coaches face different barriers to entering the coaching profession, or face differential treatment by employers, can it be claimed that discrimination is present. These are arguments that we explore in this paper.

[27] The Institute for Diversity and Ethics in Sports (TIDES) 2018 article, reports that around 70% of NFL players are black, and 27% are white.

[28] Interestingly however, very few players go on to become top level coaches (only 27% in our sample). Being a good player is not a guarantee of being a good coach. Many coaches start their career after a failed college career, perhaps due to lesser ability or injury.

Detecting discrimination is notoriously difficult. It is rare to find such accurate and objective measures of worker and firm performance that are required for an assessment of discrimination. This offers studies using sports data a major advantage. As noted by Kahn (2000), there is no other industry where we know the name, face and performance of every worker (players), firm (teams) and supervisor (coaches) in the industry. Moreover, NFL teams have an easily identifiable coaching hierarchy, allowing for clear assessments of promotions and demotions.

While the matter in the NFL is interesting in its own right, the issues we discuss are certainly not limited to just the NFL, or sports more generally. Literature on sports Head Coaches likens their role to that of a Chief Executive Officer (CEO) in a regular firm (e.g. Pieper et al. (2014)). Both Head Coaches and CEOs tend to be of a similar age, and can cope with intense pressure and scrutiny, particularly from the media. Moreover, both are appointed by owners and/or directors who will ultimately decide when their employment should be terminated. A similar concern also exists about the lack of black and minority representation amongst CEOs and top executives: only five of the Fortune 500 firms are headed up by a black CEO (Fortune (2020)).

We contribute to the literature in a number of ways. Our newly assembled dataset covers NFL coaches over the last 30 seasons, allowing us to analyse more years of data than many previous studies. The larger sample size also allows us to control for team specific trends. Moreover, we are not only interested in the reasons behind coaching promotions, but also the causes of a variety of other types of exits, including firings. We also include alternative measures of coaching performance at the coordinator level.

2. Theory & Previous Literature

2.1 Theoretical Background

A number of theories lie at the heart of the issues that we will test empirically, including theories of promotion, job separation, discrimination, leadership changes and the role of affirmative action policies.

Lazear and Rosen (1981) consider (internal) promotions as an incentive device where rewards depend on rank among a group of workers. Yet, several authors (including for example Baker et al. (1988)) are sceptical about promotions acting as a pure incentive device, because promotions often involve a change of job responsibilities and require new skills, possibly leading to sub optimal job assignments (sometimes referred to as The Peter Principle (Peter and Hull (1969)). This may well be true in a sporting context, given the extra and more varied responsibilities taken on by higher-level coaches (although many assistant coaches go on to make superb coaches). Possibly more closely related to the sporting context is to consider the role of signalling in promotions, whereby outside firms use promotions as an imperfect signal of ability (see for example Waldman (1984)).

After hiring a worker, the quality of the firm-employee match will alter over time, with factors such as age and performance (probably relative to some expected performance) determining the quality of that match. When the quality of the match falls below the value of an outside option, either party may look to terminate employment (Gielen and van Ours (2006)). The employers outside option is in the form of another worker. Given that in a sporting setting, performance is relatively easy to observe, a principal should be able to improve their team's performance by changing Head Coaches (Bryson et al. (2021, a)). Ilmakunnas et al. (2005) also argue that job separation is likely to improve firm productivity by means of bringing in new ideas and knowledge. Bosses also have an important role in determining the performance of their subordinate workers (Rosen (1982), Lazear et al. (2015)). In a sporting setting, Muehlheusser et al. (2018) show that team performance does vary according to the quality of the incoming Coach, though several papers show that the average effect of appointing a new Coach is negligible (van Ours and van Tuijl (2016), Goff et al. (2019) to name just a few).

Of interest in the market for hiring and firing of NFL coaches is the large body of literature on discrimination. If teams have a desire to maximise profits or wins, then the presence of some non-discriminatory owners along with the presence of equally skilled minority workers, should mean discriminatory practises will fall over time (Groothuis and Hill (2004)). A team could simply not afford to not employ a good minority coach over an 'average' non-minority coach for example, as the potential risk of failing in competitions increases.

It is possible that the lack of minority representation we observe amongst NFL coaches today is due to discrimination much further back in a coach's career, perhaps even during their playing careers, rather than due to discriminatory hiring by NFL teams at the top levels of coaching. Pitts and Yost (2013) for example, find evidence that players are stacked into different positions according to their race as they transition from High School to College, leading to an abundance of white players at Quarterback (a central playing position with a high level of influence on the game).

Racial stacking has its roots in the theory of Centrality, whereby non-minorities tend to gravitate to roles with a higher level of influence, which in sports tends to mean a central playing position. In applying the theory of Centrality to Baseball teams, Grusky (1963) finds that players who played in central positions (thus having higher levels of interaction with other team members) were far more likely to become field managers than players who played in non-central positions. A similar situation was identified by Latimer and Mathes (1985) in their survey of college football Head Coaches. Black coaches largely played in more peripheral positions (particularly Running Back and Defensive Back), while central positions such as Center, Quarterback and Guard were the least occupied by Black players. This seems to have fed through to coaching too, with Black coaches tending to coach peripheral positions.

As such, the lack of Black coaches we see in NFL today is not necessarily due to hiring discrimination by NFL teams, but perhaps due to barriers (which may be discriminatory) preventing Black coaches coming through the ranks in the first place. Anti-discrimination laws exist to prevent discriminatory hiring and firing, and although it is difficult to prove, critics may argue that despite the existence of such laws, discrimination is still present. This is often cited as a motivation to implement affirmative action policies, which go a step beyond anti-discrimination laws by actively supporting members of minority groups that have been or still are discriminated against. The Rooney Rule is one such example in the case of employment opportunities, but affirmative action policies also exist for example, in many college application processes. Previous literature has examined the outcomes and economic features of such policies in a variety of settings. An excellent review of both theoretical treatments and empirical studies can be found in Holzer (2007).

Many studies of discrimination focus on pay as an outcome. Unfortunately, the salaries of NFL coaches are not publicised, at least not with any degree of accuracy. Nevertheless, work on gender pay gaps would seem to suggest that once observable characteristics (notably job role) are controlled for, this pay gap tends to disappear. As such, the observed pay gap is usually interpreted as unequal hiring and/or promotions to top positions, making our focus on discrimination in promotions and other types of movement a more sensible choice in this setting.

2.2 Previous Sports Literature

As demonstrated by Madden (2004), African American coaches tended to outperform white coaches between the years 1990 and 2002 (pre Rooney Rule). Her work shows that even when controlling for differences in team quality, African American coaches had better regular season records, and consequently were more likely to make the post-season playoffs, at all stages in their career.[29] She argues that this is consistent with the view that African American coaches were held to higher standards by teams and so had to be better, more able coaches in the first place before being hired as Head Coach, thus contributing to their better average records.

Madden and Ruther (2011) take this analysis one stage further, to analyse whether the implementation of the Rooney Rule saw the performance advantage of African American head coaches disappear. The rule should force NFL teams to consider equally coaches, regardless of their race, ultimately leading to more comparable performance records. Comparing the 13 seasons prior to, and the 7 seasons post Rooney Rule, the authors find that both the difference in number of wins and the probability of reaching the postseason is no longer significant after 2003. They also find that black or African American Defensive Coordinators are insignificantly less likely to be promoted to a Head Coach role before the Rooney Rule, and faced the same treatment after the rule.[30]

Work by Solow et al. (2011) explores the transitions from coordinator role to Head Coach in more depth. Using a logit regression where the outcome variable equals 1 if a coordinator is promoted, they find that strongly performing, more experienced and younger coordinators are more likely to be promoted to a Head Coach role. Teams also appear to favour hiring Offensive Coordinators, although this result is only marginally significant. There are no significant differences in the likelihood of being promoted from a coordinator position to a Head Coach depending on race. Their results and interpretations remain unchanged when using a Cox proportional hazards model instead. Solow et al. also split their sample up into a pre- and post-Rooney Rule period in order to analyse its impacts on promotions. Their results suggest that no significant change was observed of the likelihood of a minority coach being promoted after the implementation of the rule.

Fearful that increases in the number of minority Head Coaches working in NFL was just due to changing unobservable social factors (e.g. changes in racial sentiment) that coincided with the introduction of the Rooney Rule, DuBois (2015) uses a Difference in Differences specification to compare hiring trends in the treated group of NFL Head Coaches to the control group of NFL coordinators and College Head Coaches (where the rule did not apply). DuBois

[29] African American coaches on average won 1.9 more regular season games than white coaches when controlling for team quality.
[30] In the period under analysis in Madden & Ruther (2011), there was never a black Offensive Coordinator promoted to a Head Coach role, hence the analysis could only be carried out on Defensive Coordinators.

finds that a minority candidate is between 19-21% (depending on the control group) more likely to fill a Head Coaching vacancy in the post rule period.

A major drawback of all the work mentioned above, is the relatively few numbers of years of data post-Rooney Rule. It is possible that the reason behind the insignificance of race on coaching moves is simply a lack of time since implementation for any statistical result to show, even if teams' behaviour is changing as a result of the rule. With several additional seasons of data, it is possible the significance of the result may change. We also include more coach specific variables and are able to control for team specific effects in some of our specifications.

3. The NFL, and the labour market for coaches

The NFL currently consists of 32 teams, split into two conferences of 16 teams. [31] Within each conference, teams are split into four divisions, where the winner of each, along with wildcard entrants from each conference (teams with the best remaining records) qualify for the post season, a knock-out style tournament culminating in the Super Bowl.[32] Qualifying for the post season is quite often seen as a minimum requirement for most teams, though if teams are going through a rebuild period, then expectations may be more lenient.

The coaching structure of NFL teams makes it an ideal setting to study promotions and firings. Between teams, while the exact responsibilities of the staff may vary slightly, they more or less fulfil the same duties. The Head Coach oversees day-to-day coaching activities, sets the overall playing philosophy, is responsible for in game personnel changes, and is very much the public face of the team, with a far greater media presence than other coaches. They tend to work very closely with the General Manager (GM) on decisions such as draft picks and roster decisions, while keeping the wage spending within the annual salary cap. The GM is also responsible for hiring and firing the coaching staff. Below the Head Coach are the coordinator roles. An Offensive Coordinator will typically manage all offensive plays, devise offensive game plans and strategies, and head up the team of offensive positional coaches. Defensive Coordinators will fill similar roles but on the Defensive side of the ball. Exactly who calls the plays during matches may vary across teams and will likely depend on the background and specialities of

[31] There have been 32 teams since the most recent expansion in 2002, which saw the Houston Texans added. Other expansions during our sample period of 1989-2020 occurred in 1995, when the Carolina Panthers and Jacksonville Jaguars were added, and in 1996 (Baltimore Ravens). Other teams over the period have changed name and or location.

[32] The 2020 season saw the playoffs expand to 14 teams, rather than 12 as previous. This meant an extra wildcard slot for each division, but fewer teams receiving a first-round bye.

the Head Coach. In almost all years, all teams employ this trio of coaches, although occasionally the Head Coach fills one of the coordinator roles.

It is not uncommon to see a well performing coordinator promoted to a Head Coach, either internally or externally. During our sample, 135 out of 693 coordinator coaching spells (or 135 out of 1892 coordinator-seasons) have resulted in promotion. Many of them are very successful and enjoy a prolonged spell(s) as a Head Coach, whereas others drop back down into a lower coaching rank or leave coaching completely (180 out of 212 Head Coaching spells end in such a manner). Coaching moves can also occur between the NFL and the hugely lucrative college sector. Coaches who work for the top college teams can very earn large salaries, potentially at least as big as Head Coaches working in the NFL. For example, the highest paid college coach is Nick Saban at Alabama on a reported $8m per year, similar to the estimated salaries for the top earning coaches in the NFL. Moreover, attendances at some college games are regularly upwards of 80-90 thousand. Because of the lucrative nature of this sector, we model these moves as equivalent to moves to the NFL.[33]

A major feature of the labour market for NFL coaches is the Rooney Rule, named after Dan Rooney, former Pittsburgh Steelers owner and chair of the NFL's diversity committee. Its implementation followed the sacking of two high profile Black coaches; Tony Dungy from the Tampa Bay Buccaneers despite his overall winning record, and Dennis Green from the Minnesota Vikings despite his first losing season in 10 years. The rule, introduced for the 2003 season, requires that teams hiring a new Head Coach must interview at least one minority candidate. [34] DuBois (2015) describes the rule as a "soft" affirmative action policy, designed to change the composition of the candidate pool, rather than who is ultimately employed. There are rare circumstances where the rule will not apply, for example if an assistant coach's contract guarantees them the Head Coach's job should it become available. Other industries and sports have also implemented a similar type of rule. Most notably in English Football where both the England national team and all clubs in divisions 2, 3 and 4 of the footballing pyramid must now adopt a similar approach when appointing a new Head Coach (BBC (2018)).

In this paper, we use the Rooney Rule to compare outcomes of minority Coaches before and after its implementation. In particular, we focus on whether minority Coaches are more likely

[33] Robustness checks (see section 5.5) reveal this definition does not impact on results
[34] The Rooney Rule was altered in 2009. The rule now covers all senior positions including GMs, but there is no rule that covers the coordinator roles at this stage. Further, the rule was extended to cover all minorities, not just African American Coaches.

to be employed as a Head Coach, including both transitions from coordinator to Head Coach, and Head Coaches who stay on the same level, given their performance and human capital. We also extend the analysis of Madden and Ruther (2011) to compare the performance of coaches pre- and post- Rooney Rule.

4. Data & Methodology

4.1 Data

Our new dataset has been collected and assembled entirely by hand, and consists of all individuals who held a top coaching position (i.e. Head Coach, Offensive Coordinator or Defensive Coordinator) at an NFL team between 1989 to the end of the 2020 season. The uniqueness of this study lies in the data; our sample period gives us a generous number of observations before and after the implementation of the Rooney Rule in 2003 and includes a richer set of variables than previous work. Previous work has only included 5-6 years of data post implementation and has lacked variables capturing team characteristics. We have excluded any individual who held their position on an interim or temporary basis, because by definition, their exit is already determined. In the case that an interim coach performs well enough in their role that they are given the job permanently, we only consider the period after they were given the role full time.

Our main source for the data is the website Pro Football Reference (https://www.pro-football-reference.com/). From here, we obtain the entire coaching history for all teams in the NFL i.e. who filled the positions of Head Coach, Offensive Coordinator and Defensive Coordinator. We collect various details on each coach, including past coaching spells, used to construct our experience variable, and their age. This website also contains the end of season records for each club which we use to construct the performance measures for the coaches. For Head Coaches, performance is measured using the win-loss percentage during the regular season (with draws counting as half a win). For the coordinator roles, we use the seasonal percentile rank of total points scored and total yards (which we can split into passing and rushing yards) scored and conceded, for offensive and defensive coordinators, respectively.[35] A benefit of considering these alternative measures alongside points scored is that we can rule out any contamination effects of defensive contributions to attacking outputs. A percentile rank assigns the highest scoring and highest conceding team a value of 1, the lowest scoring and conceding

[35] By points, we mean in game points i.e. 6 for a touchdown, 1 for a successful point after attempt, 2 for a successful two point attempt, 3 for a field goal, and 2 for a safety.

team 0, and the mean scoring and conceding a score of 0.5. Following previous studies (including for example Fee et al. (2006)), using the percentile rank of performance rather than the raw number is preferred, since their distributions will be stable over time, allowing for more meaningful comparisons across seasons and reducing the influence of outliers.

While the main contribution of a coach to an NFL team lies in their on-field success, their complete contribution probably goes beyond this. Player development and overseeing rebuilds for example also form a large part of their job description, though these may be difficult to quantify. Coaches may be afforded a season or two grace period in which they are given the opportunity to build a squad, develop and implement new play calls etc. and as such we include an interaction of performance and tenure.

We include a dummy variable for race, with the variable taking the value 1 if the coach is a minority ethnicity, which we restrict to black coaches in the sample, and 0 otherwise. We exclude from the sample coaches who are of mixed race for two reasons. [36] First, there are very few mixed race coaches in the NFL, while the second is due to the expansion of the Rooney Rule to cover all minority candidates in 2009. The race of the coach was coded using publicly available information, following Fort et al. (2008) who *"suspect there is no bias in a dichotomous, researcher assessed measure of race"*.

A coach's experience is measured using the number of years they have held one of the three top coaching positions under consideration, up to and including their most recent year, but excluding any career gaps. This can be thought of as a measure of general human capital and is entered separately for NFL and college roles. Tenure measures the number of seasons that a coach spends at their current team, in their current role. We also include a dummy variable that is equal to 1 if the coach previously played football, 0 otherwise. We include a dummy variable identifying if the coaching change coincided with the team changing their GM, given the GM is ultimately responsible for hiring and firing the coaching staff. Finally, postRR is a variable identifying seasons after the implementation of the Rooney Rule.

In total, our data cover 32 seasons, in which 431 individual coaches held roles at NFL teams. Many of these coaches held more than one role across the 32 seasons and held positions at

[36] A total of four coaches who were of mixed race and are excluded from the sample. Namely, former Titans OC Norm Chow, of Asian-American descent, Tom Flores (former Raiders and Seahawks HC), Juan Castillo (former Eagles DC) and Ron Rivera (former Bears and Chargers DC, Panthers HC, and current Washington Football team HC), all of whom are of Hispanic descent.

multiple teams. We view each coaching spell as a series of individual seasons, resulting in 2878 coach-season observations, and in doing so modelling time as discrete.

4.2 Methodology

Our methodology is straightforward. We start by identifying coaching exits at the end of the season, which we will later break down into different types of exits. These are Promotions (from Coordinator to Head Coach, either internally or externally), a Sideways move (moving to the same role on a different team), or a Downwards move (dropping down the coaching hierarchy and or dropping out of the sample altogether). Kopkin (2014) suggests that modelling time as discrete is appropriate in this context since the end of a season is the most common and sensible time to make coaching changes. General Managers will have the greatest time to search for, interview, and hire the new Head Coach, while the new coach(es) will have the longest time to implement new training, play books etc. In the simplest form, we model exits (of any type) for coach *i* at the end of a season *t* using Logit regressions as follows:

$$Prob(Exit_{it} = 1|\boldsymbol{X_{it}}) = (1 + \exp(-\alpha - \boldsymbol{\beta X_{it}}))^{-1}$$

This is estimated separately on Head Coaches, Offensive Coordinators and Defensive Coordinators. Having controlled for performance, human capital etc., any differences in exit probabilities by race may hint at some discriminatory exits. Although such a claim can of course never be concrete.

However, it is unlikely that promotions, sideways and downwards coaching moves are a result of similar circumstances. As such, we extend the analysis to a Multinomial Logit regression to model the different types of coaching moves. In this setting, each season can end in one of four outcomes; specifically No Exit or Exit, where Exit is further split into Upwards (i.e. a Promotion), Sideways and Downward coaching moves, defined by where (or if) the coach next appears in the data. Let *k=1…4* denote the 4 possible end of season outcomes, then the probability that the type of exit of coach *i* in season *t* is

$$Prob(Exit_{it} = k) = \frac{\exp{(b_0^k + \boldsymbol{\beta^k X_{it}})}}{\sum_{l=1}^{4} \exp(b_o^l + \boldsymbol{\beta^l X_{it}})}$$

This methodology assumes that each observation is independent of one another. However, as explained by Holmes (2011) and Kopkin (2014) (in the context of college football), this assumption is questionable as certain team and/or coach characteristics may make coaching departures more or less likely. Due to the large number of coaches, it is infeasible to include coach fixed effects, although our larger sample size means including team fixed effects for the 32 teams is possible, somewhat dampening this concern. The inclusion of Team Fixed Effects controls for unobservable team characteristics that do not change over time but may differ across teams. Another key assumption when using a multinomial logit regression is the Independence of Irrelevant Alternatives (IIA) assumption. IIA postulates that the probability of one outcome should be independent of the probability of another. In our case for example, the probability of a promotion should be independent of the probability of a sideways move. This seems a reasonable assumption to make, and it seems implausible that this would be violated.

5. Results

5.1 Descriptive Statistics

Table 1 shows the frequency of each of our types of failure, while Table 2 shows the descriptive statistics for the uninteracted variables, split by role in panels B, C and D.

Table 1: Frequency of type of movement

Type of Move	Frequency	Percent
No Exit	1973	68.55
Promotion	135	4.69
Sideways	204	7.09
Downwards	566	19.67
Total	2878	100.00

Table 2: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A: All Coaches					
Tenure	2878	3.10	2.83	1	26
Age	2878	50.38	8.31	28	80
Black	2878	0.14	0.35	0	1
NFL experience	2878	7.48	6.04	0	36
College experience	2878	2.87	4.47	0	25
Played	2878	0.28	0.45	0	1
postRR	2878	0.59	0.49	0	1
GMchange	2878	0.12	0.32	0	1
Panel B: Head Coaches					
Tenure	986	4.22	3.79	1	26
Age	986	51.20	7.38	31	72
Black	986	0.13	0.34	0	1
NFL experience	986	10.14	6.45	0	36
College experience	986	2.89	4.59	0	25
Win Loss percentage	986	0.50	0.19	0	1

Panel C: Offensive Coordinators					
Tenure	934	2.43	1.90	1	12
Age	934	48.40	8.51	28	71
Black	934	0.10	0.30	0	1
NFL experience	934	5.70	4.85	0	28
College experience	934	2.82	3.99	0	21
Points for percentile	934	0.49	0.31	0	1
Yards for percentile	934	0.49	0.31	0	1
Pass yards for percentile	934	0.50	0.31	0	1
Rush yards for percentile	934	0.49	0.31	0	1
Panel D: Defensive Coordinators					
Tenure	958	2.61	1.94	1	13
Age	958	51.46	8.68	31	80
Black	958	0.20	0.40	0	1
NFL experience	958	6.49	5.72	0	34
College experience	958	2.90	4.79	0	24
Pts agn percentile	958	0.50	0.31	0	1
Yards agn percentile	958	0.50	0.31	0	1
Pass yards against percentile	958	0.50	0.31	0	1
Rush yards against percentile	958	0.50	0.31	0	1

By considering the roles separately, we can tell that, on average, Offensive Coordinators tend to be slightly younger than other coaches, while Head Coaches tend to be more experienced (at least in terms of coaching years in the NFL), and spend longer at one team. More Defensive Coordinator positions are filled by minority coaches than Offensive Coordinator positions, which is in line with descriptions in Section 1.

Before progressing with the more detailed regression models, we first present a simple before and after comparison of the probability of observing a black Head Coach in the NFL, pre and post Rooney Rule. As shown by table 3, the probability of observing a black Head Coach in any season after and including 2003 increases by 10%. While this could hint towards some success of the rule, it tells us nothing about any coaching characteristics that make this more likely. This is what we go on to address in the regression analyses that follow.

Table 3: Probability of observing a black Head Coach before and after the Rooney Rule

VARIABLES	Black HC
postRR	0.104***
	(0.023)
Observations	986

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5.2 All Exits

Table 4 presents the results of the logit estimations, which have been estimated separately for the three different coaching roles. Each column for the coordinator roles represents a different performance measure, as indicated by the column headers. Results are displayed as marginal effects, calculated at the variable means and standard errors are clustered at the coach level.

Common across all roles is that the longer spent at a team, the more likely an exit is to occur, although this occurs at a diminishing rate given the negative coefficient on the squared term, with a turning point estimated to be between about 19 seasons for Head Coaches, 7-8 seasons for Offensive Coordinators and between 6-7 seasons for Defensive Coordinators, depending on specification. Given these turning points lie well beyond the average tenure for all roles, for the most part this effect is probably linear in nature.

The performance measures enter with the expected sign, indicating that worse performance is associated with higher likelihood of exit (for HC's and OC's, low values of performance are bad outcomes, whereas for DC's, low values are good outcomes, hence the opposing signs). Win-loss percentages, points scored and points conceded are amongst the strongest predictors of exits.[37] Interestingly however, variables capturing yardage suggest that Offensive and Defensive Coordinators may be valued differently on their outputs, with total yardage mattering on offense, but only rushing yards mattering on defence.

Particularly evident for Head Coaches is the effect of consistent poor performance over a number of seasons, with each additional season further contributing to the likelihood of dismissal, as shown by the negative and significant interaction of win loss percentage and tenure. This could indicate some leniency at the start of their tenure, perhaps if the coach has

[37] We also checked to see if Points for and Points against could explain Head Coach departures, but neither could significantly explain Head Coach exits; only win loss percentage was important.

been tasked with a re-build of the roster. The effect is less pronounced for the coordinator roles. Coaches of teams who reach the postseason (as either a division winner or a wildcard) face lower exit probabilities. However, this is only significant for the coordinator roles, and not for Head Coaches which is likely due to win loss percentages being able to explain a great deal of the variation in post season qualification (the correlation between win loss percentage and post season qualification is about 0.78).

There are no differences in the probability of exiting pre and post Rooney Rule, other than in the OC(1) specification, though this is only significant at 10%, and neither does it differ depending on the coach's race. By itself, race is largely insignificant in explaining exits, other than in specification OC(1), where black Offensive Coordinators are more likely to exit. In both OC(1) and OC(2), black Offensive Coordinators appear to be treated more harshly with regards to poor performance in points scored and yards gained, respectively, since their interactions with race enter significantly.

As for the variables capturing a coach's Human Capital, there are very few significant results. Previous experience at either NFL or college teams are not significant factors in explaining exits, neither is having played professionally before entering coaching. This is perhaps somewhat surprising as one may expect that experience could protect a coach from dismissal, particularly during periods of poor performance, though would be consistent with findings in Bryson et al. (2021, b), whose work on football (soccer) coaches finds that years of experience matter very little in protecting against dismissals. Older age increases the probability of exiting, but only significantly for Head Coaches.

Table 4: All Exits

Role	HC (1)	OC (1)	OC (2)	OC (3)	OC (4)	DC (1)	DC (2)	DC (3)	DC (4)
Performance Measure	Win Loss Percentage	Points for	Yards for	Pass yards for	Rush yards for	Points against	Yards against	Pass yards against	Rush yards against
Tenure	0.090***	0.139***	0.126***	0.126***	0.103***	0.108***	0.107***	0.102***	0.118***
	(0.014)	(0.028)	(0.028)	(0.026)	(0.030)	(0.032)	(0.032)	(0.030)	(0.028)
Tenure Squared	-0.002***	-0.009***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Black	-0.162	0.224**	0.161	0.166	0.081	-0.031	0.001	0.020	-0.019
	(0.138)	(0.101)	(0.098)	(0.117)	(0.094)	(0.142)	(0.111)	(0.095)	(0.111)
Age	0.005**	0.003	0.003	0.004*	0.003	0.003	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Played	0.029	0.017	0.024	0.016	0.028	0.013	0.019	0.021	0.029
	(0.023)	(0.033)	(0.033)	(0.034)	(0.033)	(0.036)	(0.036)	(0.036)	(0.037)
NFL Experience	-0.001	0.001	0.001	0.000	0.001	0.003	0.003	0.003	0.002
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
College Experience	0.002	0.003	0.003	0.004	0.003	-0.003	-0.003	-0.004	-0.004
	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Post Season	-0.027	-0.136***	-0.180***	-0.214***	-0.216***	-0.089**	-0.160***	-0.193***	-0.159***
	(0.039)	(0.034)	(0.030)	(0.028)	(0.029)	(0.040)	(0.035)	(0.033)	(0.034)
Performance	-0.558***	-0.152*	-0.162**	-0.062	-0.110	0.295***	0.144	0.048	0.167**
	(0.120)	(0.091)	(0.079)	(0.080)	(0.091)	(0.097)	(0.101)	(0.099)	(0.082)
Black * Performance	0.255	-0.311**	-0.222*	-0.189	-0.038	0.062	0.053	-0.006	0.068
	(0.236)	(0.140)	(0.126)	(0.143)	(0.137)	(0.146)	(0.134)	(0.133)	(0.110)
Tenure * Performance	-0.065***	-0.054**	-0.031	-0.044*	-0.009	0.008	0.018	0.022	-0.006
	(0.021)	(0.027)	(0.024)	(0.022)	(0.025)	(0.029)	(0.032)	(0.028)	(0.025)
postRR	-0.029	0.055*	0.048	0.046	0.045	-0.010	-0.008	-0.007	-0.008
	(0.024)	(0.032)	(0.032)	(0.032)	(0.032)	(0.033)	(0.032)	(0.032)	(0.032)
Black * postRR	0.078	-0.062	-0.045	-0.056	-0.058	0.077	0.054	0.062	0.057
	(0.080)	(0.109)	(0.111)	(0.106)	(0.106)	(0.089)	(0.088)	(0.083)	(0.084)
GMchange	0.115***	0.202***	0.204***	0.206***	0.221***	0.119***	0.122***	0.135***	0.129***
	(0.029)	(0.040)	(0.040)	(0.041)	(0.040)	(0.042)	(0.043)	(0.043)	(0.043)
Team FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	986	934	934	934	934	958	958	958	958

Cluster robust standard errors in parentheses (clustered at the Coach level) *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is Exit (0,1)

5.3 Upwards, Sideways and Downwards Moves

Table 6 displays the results from multinomial logit regressions, where the outcome variable from table 3 (Exit) is now split into No Exit (which was chosen as the base outcome), Upward Moves (i.e. a promotion from coordinator to Head Coach), Sideways Moves (coordinator to coordinator or Head Coach to Head Coach) and Downward Moves. We believe these results are more informative than in Section 5.2, given the likely different circumstances leading to a promotion, firing etc. The results shown are marginal effects. For reference, table 5 shows the frequency of our different outcomes. An upward or sideways move must occur in the following season, such that the coach is not unemployed for a period during the season, otherwise this is classed as a downward move.

Table 5: Frequency of outcomes by role

Outcome / Role	Head Coach	Offensive Coordinator	Defensive Coordinator	Total
No Exit	774	569	630	1973
Upwards	0	72	63	135
Sideways	32	81	91	204
Downwards	180	212	174	566
Total	986	934	958	2878

The results in table 6 have combined Offensive and Defensive Coordinators into one group to overcome the relative rarity of coordinator promotions. The variable in panel B named Points captures the percentile rank of points scored for Offensive Coordinators and percentile rank of points against for Defensive Coordinators, but now the latter has been rescaled such that higher values imply better performance. We also include a dummy variable equal to one if the observation relates to an Offensive Coordinator, in line with Solow et al. (2011) to check for possible preferences for hiring offensive coaches. There is no upward move equation estimated for Head Coaches, since they are already at the top of the coaching ladder so can only move Sideways or Downwards. In order to accurately estimate standard errors, we were unable to include team fixed effects in the Head Coach specification, because for several teams we never observe a Sideways coaching movement. Specifications for our other coordinator performance variables are available in the appendix (Table A1)

As previously, more seasons at one team tend to increase the likelihood of exiting but at a diminishing rate, though tenure plays no role at all in explaining coordinator promotions. Younger, more experienced (in the NFL), and better performing coordinators are more likely to be promoted to a Head Coaching role. Teams do not appear to be showing any preference towards hiring offensive coordinators for Head Coaching roles, though they are significantly less likely to be retained by their current team even when conditioning on performance. Older Head Coaches and coordinators are likely to drop down coaching levels and/or leave coaching altogether. Good performance unsurprisingly protects all coaches from losing their job, shown by the positive and highly significant probabilities on No Exit. Not making the post-season playoffs, however, is likely to result in a downward coaching move.

The effects of age are confirmed by other studies on sports coaches, for example Solow et al. (2011) for the case of NFL coordinators, Wangrow et al. (2018) for the case of NBA (basketball) Head Coaches, and Bryson et al. (2021, b) for football (soccer) Head Coaches. Younger coaches are likely to have a higher job match surplus, while older coaches may suffer from deteriorations in job match surplus. Older coaches may be able to demand a higher salary given their experience and previous success, while with age comes increased risks of health-related issues and concerns about retirement. Indeed, a few coaches in the sample have stepped down due to poor health and or retirement. A further explanation as to why teams may wish to employ younger coaches can be found by drawing on findings from human capital theory (Malone et al. (2012)). Young coaches have an incentive to invest heavily in acquiring new coaching skills, which are rewarded when appointed as Head Coach. This incentive to invest in new skills falls over time, while skills may also depreciate over time.

This trend of young, reasonably experienced and well performing coaches being promoted is evident, particularly recently in the NFL. Of the 8 Head Coaches appointed ahead of the 2019 season, five were aged 40 or younger when appointed, four of whom would be NFL Head Coaches for the first time, yet all could boast several years of previous experience in other roles. This has been labelled in some media circles as the 'Sean McVay effect'. At his time of appointment at the L.A. Rams, McVay was the youngest ever NFL Head Coach, and his apparent success as a young offensive mind at the Rams subsequently led to several teams copying the Rams' strategy. Some even resorted to hiring McVay's assistants (Zac Taylor hired by the Bengals, Matt LaFleur hired by the Packers, and Brandon Staley hired by the Chargers).

We can also see the effect of a team changing its GM. Coaches are more likely to drop down coaching levels when the GM changes, perhaps highlighting the new GM's desire to bring in their own coaching staff. More likely however, is that the owners (who themselves hire and fire the GM) decide the whole coaching and scouting teams have been performing below standard and decide on a complete overhaul of these positions. This does of course raise potential endogeneity concerns, but is not something we address in this paper.

Now that coaching moves are considered separately, we are able to comment on the success of the Rooney Rule by considering the Black * postRR coefficient. Both promotions from coordinator positions and sideways moves for Head Coaches would potentially be affected by the interview requirement, so it is important to consider both of these effects. Results suggest that the probability of a black Coordinator being promoted after the rule increases by around 0.074, significant at 10%, pointing toward some degree of success of getting more minority coaches into top coaching roles. This result also holds up in our alternative specifications in the appendix with alternative Coordinator performance measures. We can also see however, that in the post rule period, Head Coaches are significantly less likely to find a job on the same level, but there is no difference by race. It appears therefore, that teams are now considering the current pool of black Coordinators rather than just ex-Head Coaches available on the market to fill current Head coaching vacancies.

That black Coordinators are more likely to be promoted to a Head Coach role post Rooney Rule is a finding that contrasts the results of and Solow et al. (2011), who demonstrate no significant differences in probability of promotion by race. It could be that we are capturing some longer term effect of the rule, which is being picked up in the additional seasons in our data, but with that said, the effect shown here is still only marginally significant.

Table 6: Multinomial Logistic Regression for different types of Exits

Panel A: Head Coaches

OUTCOME	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Win Loss Percentage	Black* WLP	Tenure* WLP	postRR	Black* postRR	GM change
No Exit	-0.074***	0.002***	0.029	-0.006**	-0.001	0.002	-0.002	0.071	0.530***	-0.086	0.056***	0.016	0.009	-0.119***
	(0.014)	(0.001)	(0.146)	(0.002)	(0.022)	(0.003)	(0.002)	(0.047)	(0.117)	(0.254)	(0.021)	(0.023)	(0.097)	(0.028)
Sideways	0.028***	-0.001*	0.142***	-0.001	-0.006	0.001	0.001	0.023	0.033	-0.241**	-0.023*	-0.026**	-0.052	0.010
	(0.010)	(0.001)	(0.041)	(0.001)	(0.012)	(0.001)	(0.002)	(0.019)	(0.080)	(0.095)	(0.013)	(0.013)	(0.034)	(0.014)
Downwards	0.045***	-0.001*	-0.171	0.006***	0.006	-0.003	0.001	-0.094**	-0.563***	0.328	-0.033	0.010	0.043	0.109***
	(0.012)	(0.001)	(0.130)	(0.002)	(0.021)	(0.002)	(0.002)	(0.048)	(0.114)	(0.209)	(0.021)	(0.022)	(0.091)	(0.025)
Team FE	NO													
Observations	986													

Panel B: Coordinators

OUTCOME	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Points	Black* Points	Tenure* Points	postRR	Black* postRR	GM change	OC=1
No Exit	-0.120***	0.008***	-0.064	-0.002	-0.006	-0.003	-0.000	0.118***	0.246***	0.106	0.028	-0.024	-0.061	-0.155***	-0.071***
	(0.021)	(0.002)	(0.059)	(0.002)	(0.024)	(0.002)	(0.003)	(0.026)	(0.069)	(0.099)	(0.023)	(0.022)	(0.063)	(0.030)	(0.021)
Promotion	0.014	-0.001	-0.042	-0.005***	0.012	0.003*	-0.000	0.007	0.073*	-0.032	0.006	-0.005	0.074*	0.036*	-0.000
	(0.013)	(0.001)	(0.050)	(0.001)	(0.015)	(0.002)	(0.002)	(0.012)	(0.040)	(0.060)	(0.012)	(0.013)	(0.042)	(0.020)	(0.013)
Sideways	0.066***	-0.004***	0.059	-0.003***	0.024	0.004**	0.003	-0.068***	0.049	-0.013	-0.026**	-0.001	-0.029	0.037**	-0.003
	(0.016)	(0.002)	(0.036)	(0.001)	(0.016)	(0.002)	(0.002)	(0.018)	(0.048)	(0.065)	(0.012)	(0.014)	(0.034)	(0.019)	(0.013)
Downwards	0.040***	-0.003	0.047	0.010***	-0.030	-0.003*	-0.002	-0.057**	-0.368***	-0.060	-0.008	0.031*	0.015	0.082***	0.074***
	(0.015)	(0.002)	(0.048)	(0.001)	(0.020)	(0.002)	(0.002)	(0.024)	(0.062)	(0.074)	(0.018)	(0.019)	(0.051)	(0.022)	(0.018)
Team FE	YES														
Observations	1,892														

Cluster robust standard errors in parentheses (clustered at the Coach level)

*** p<0.01, ** p<0.05, * p<0.1

5.4 Evaluating NFL Coaching Performance

In light of these findings, we now turn to examining what effect, if any, the Rooney Rule has had on coaching performance. As explained by Holzer (2007), a common critique of affirmative action policies is that they could lead to firms employing lower qualified (perhaps lower quality) workers, creating a sort of reverse discrimination. In this sense, rather than creating equal opportunities, affirmative action may actually lead to equal outcomes. To test this critique, in an extension of work by Madden and Ruther (2011), using our additional years of data post Rooney Rule, we analyse how the relative performance of black Coaches has changed post Rooney Rule. In table 7, we show the results from several Generalized Linear Models (GLM) to examine changes to Win Loss percentages for Head Coaches, and for our selection of performance metrics for coordinators. A GLM is appropriate here since the dependent variable lies between 0 and 1. The equation in the second column for Head Coaches estimates a Logit model for the likelihood of making the post season playoffs, with results displayed as marginal effects. All standard errors are clustered at the coach level.

Table 7: Coaching performance pre and post Rooney Rule

Outcome VARIABLES	HC Win Loss Percentage GLM	HC Post Season (0,1) Logit	Coordinators Points GLM	Coordinators Yards GLM	Coordinators Passing Yards GLM	Coordinators Rushing Yards GLM
Black	0.090***	0.322***	0.020	0.055	0.064	-0.003
	(0.027)	(0.073)	(0.047)	(0.044)	(0.050)	(0.041)
postRR	-0.007	-0.025	-0.017	-0.008	-0.009	0.001
	(0.019)	(0.041)	(0.021)	(0.021)	(0.020)	(0.020)
Black * postRR	-0.089*	-0.340***	-0.023	-0.053	-0.045	-0.041
	(0.050)	(0.110)	(0.054)	(0.055)	(0.060)	(0.050)
Age	-0.003	-0.003	-0.001	-0.002	-0.001	-0.002
	(0.002)	(0.004)	(0.002)	(0.002)	(0.001)	(0.002)
Played	-0.026	-0.095**	-0.009	0.012	-0.017	0.041
	(0.020)	(0.047)	(0.024)	(0.026)	(0.025)	(0.025)
NFL experience	0.007***	0.010**	0.004*	0.006**	0.003	0.005*
	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)
College experience	0.000	-0.003	0.003	0.002	0.002	-0.000
	(0.002)	(0.006)	(0.003)	(0.002)	(0.002)	(0.002)
Constant	0.578***		0.513***	0.561***	0.505***	0.587***
	(0.076)		(0.076)	(0.075)	(0.066)	(0.075)
Observations	986	986	1,892	1,892	1,892	1,892

Cluster robust standard errors in parentheses (clustered at the Coach level)
*** p<0.01, ** p<0.05, * p<0.1

The results confirm the findings by Madden and Ruther (2011), in that any performance advantage of black Head Coaches disappeared post Rooney Rule. The magnitudes and opposite signs of the coefficients on Black and Black*postRR in both the win percentage and post season qualification equations shows this disappears almost completely. The decline in win-loss percentage of about 9% equates to just under 1.5 wins per season. As a result of these fewer regular season wins, the resulting probability of making the playoffs also declines. This performance decline is consistent with the idea that the Rooney Rule is encouraging / forcing teams to consider equally skilled Black and White candidates for Head Coaching positions, though how much of this result is driven by unobserved heterogeneity remains unclear, but this could be an avenue for future work. Yet, prior to the implementation of the rule, it would appear that only the best black coaches were hired, and consequently, the average performance of black Head Coaches pre Rooney Rule was higher. A key point to note here though is that these results are conditional on the candidates making it to a high enough coaching level to be considered for these Head Coaching vacancies. Even if we are detecting some improvements in the opportunities for black coaches post Rooney Rule, we cannot say anything about the opportunities at the grass roots or entry levels of coaching.

We observe no performance decline post Rooney Rule for the coordinator roles. Performance of black and white coordinators was statistically the same both before and after the Rooney Rule implementation in 2003. Of course, the rule did not cover coordinators, so we would not expect it to differ post its implementation. However, given that we also have no statistical significance on the uninteracted Black variable, it is possible that teams were considering equally skilled candidates, regardless of race, unlike the case we demonstrate for Head Coaches. This could shed some light on a comment from Solow et al. (2011), but never fully explored. Quite rightly, they claim that it is not clear if the Rooney Rule was the sole reason behind the increase in the number of minority Head Coaches (i.e. guaranteeing black coordinators an interview to showcase their ability where they were previously being overlooked), or whether minority coordinators were developing into better coaches in the same period. The results here appear to lend support to the former hypothesis, in that minority coordinators have always been of equal ability, and the rule has simply allowed them to showcase their ability to potential employers during interviews, which perhaps they were not able to do pre–Rooney Rule.

The age of a coach is not a determinant of performance, however more experience in the NFL (though, not experience in College football) does show some association to improved performance, particularly for Head Coaches. The link between performance and experience could be due to experienced coaches accumulating more skills and coaching ability over their career; an extension of an on the job learning type argument (see for example Gaynor et al. (2005) in the health economics literature). However, this finding could also just be reflecting a selection effect where only the best coaches stay in a job and or find new jobs.

Being a former player is largely insignificant, though Head Coaches who are former players are less likely to make the playoffs. There is certainly no evidence of former players making better coaches, a result that contrasts with the findings by Goodall et al. (2011), (albeit using different methodology in a different sport, and this is by no means the focus of our paper) who compare the outcomes of NBA coaches who were former players against those who never played. They find that teams who hire a former player see an improvement in win percentage and perform better in the playoffs than teams who hire a coach who never played, and this finding is exaggerated when the coach was a star player. Of course, the skills required to be a good coach and a good player are not necessarily the same, so not all good players will make good coaches, although we are not wishing to make any claims about causality here.

5.5 Robustness Checks

We discuss the results from two robustness checks in the following section. The full results tables for these can be found in the appendix.

The first check involves dropping coach-year observations where the coach retires (Table A2). In the previous analyses, such exits were classed as a downward coaching move since they do not re-appear in the sample. However, a retirement is likely a different outcome to a downward move. The cases were easy to identify, usually via media publications or press statements where the coach announced their intention to retire at some point. On some occasions following their decision to retire, the coach came out of retirement (e.g. Bruce Arians has twice announced his retirement only to return). These cases are still classified as a retirement, on the grounds that their initial decision to retire was, at the time, deemed a permanent decision. This eliminates 53 coach year observations from the downward moves equations. The results and main findings remain unchanged, though age drops out of significance for the No Exit outcome for Head Coaches.

Finally, we check the robustness of our definition of exits to college teams (Table A3). So far, we have classed an equivalent title at college on equal footing to an NFL team. In reality, some colleges would likely be seen as a downward step, particularly colleges with lesser reputations.

We adjust our definition of sideways and downwards moves to college teams according to the prestige of some colleges. In particular, we use the top 25 ranked college teams, based on CollegeChoice rankings, as colleges that are maintained on an equal footing to NFL teams, whereas moves to other college teams will be considered as a downward move.[38] Using this changed definition, in the main, results are unchanged, however, we do observe some slight differences in the significance of the race variable and its interactions. Namely, sideways moves for black coordinators now enter significantly, while black Head Coaches are now significantly less likely to find a job on the same level post Rooney Rule, though results for black coordinator promotions remain unchanged.

6. Conclusions

We have investigated the relationship between race and organizational (specifically coaching) hierarchy in the NFL, a setting where this issue is never far from the headlines. Using a new, unique dataset considering the top two levels of coaching staffs dating back to 1989, we examine movements between Coordinator and Head Coach positions, and out of these positions altogether.

Results suggest that teams favour employing younger, more experienced and better performing coordinators to fill a Head Coaching vacancy, while teams are more likely to fire older and poorly performing coaches. A coach's race has little impact on exits, although black coordinators are marginally more likely to be promoted to a Head Coaching role after the implementation of the Rooney Rule. It would appear this somewhat hindered moves for current NFL Head Coaches moving to another team, but there were no significant differences identified between Black and White HCs.

An analysis of Head Coach and coordinator performance pre- and post-Rooney Rule reveals two interesting findings. First, teams do now appear to be considering equally skilled black coaches to fill the Head Coach role, as shown by the win percentages of black and white Head Coaches equalising after the implementation of the rule. Second, when we consider coordinator performance, the performance across our four measures is statistically the same for both white and black coordinators, pre- and post-Rooney Rule. Taken together, this would imply that while a skilled supply of coordinators, regardless of race, has always been available to teams, the Rooney Rule seems to have forced to teams to consider and then hire equally skilled

[38] Rankings available at https://www.collegechoice.net/rankings/best-football-schools/, with colleges ranked on 4 categories; on field success, alumni success, game day experience, and culture and influence.

candidates at the Head Coach position. In spite of this, the success of the Rooney Rule is still an open source of debate, particularly when a strong black candidate is seemingly overlooked.

The adoption of very similar policies in a variety of other settings, both sporting (e.g. the English Football League (EFL), the Football Association, the English Cricket Board) and non-sporting (e.g. Facebook) suggest that these other industries look favourably on the outcomes of the rule, or at the very least, believe it has a positive PR value. A survey carried out by Kilvington (2018) on British Asian coaches working in English football (soccer), shows a slight favouritism towards the policy being introduced in the EFL. Support was not universal though, as some coaches referred to the policy as *'tokenistic'*, with coaches pointing out that hiring tends to be as a result of networks, connections and recommendations in the industry, which essentially renders the policy redundant. This could lead to so called 'sham interviews' being conducted just to tick a box and avoid punishment by the league. To date though, only one NFL team has ever been found to be in violation of the Rooney Rule; the Detroit Lions in 2003.

So, the Rooney Rule is probably best described as a small step in the right direction. One potential avenue could be to improve opportunities for minority coaches lower down the coaching ladder, particularly for coaches with offensive backgrounds. A Denver Post (2017) analysis highlights this issue, that between 2007 and 2017, of the 147 Offensive Coordinator job openings, 110 were filled by former quarterback coaches (in line with Foreman et al. (2018) on teams' preference for coaches with central position experience). Of these 110, only five were filled by a black coach, with Hue Jackson alone filling three of them. Several authors (Foreman et al. (2018), Solow et al. (2011)) suggest that the reason we observe so few minority coaches in top coaching positions is due to barriers early on in careers, maybe even during the coach's high school and college playing days (Pitts and Yost (2013)) and so prospective coaches never gain the experiences required for career development. So, while the Rooney Rule has been marginally successful in ensuring opportunities for minority coaches at the very top of the coaching hierarchy, the next step to ensure equal access and opportunity for minority coaches is to target grassroots levels.

As a final point, the NFL can only do so much to help improve prospects and opportunities for minority coaches; they have even begun to incentivise teams to make minority hires by awarding them with additional draft picks. However, if there is no willingness by team owners and executives to make minority hires, then the NFL policies will be redundant. This is a much harder problem to solve, but interview and recruitment training, is one such possibility.

APPENDIX

Table A1: Alternative Performance Measures for Coordinators

Panel A: Total Yards

	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Yards	Black*Yards	Tenure* Yards	postRR	Black* postRR	GM change	OC=1
No Exit	-0.121***	0.008***	-0.071	-0.002	-0.013	-0.004*	-0.000	0.179***	0.183***	0.086	0.020	-0.022	-0.053	-0.157***	-0.069***
	(0.021)	(0.002)	(0.068)	(0.002)	(0.024)	(0.002)	(0.003)	(0.023)	(0.065)	(0.090)	(0.021)	(0.023)	(0.066)	(0.030)	(0.021)
Promotion	0.012	-0.001	-0.010	-0.004***	0.011	0.003*	-0.000	0.016	0.080**	-0.089*	0.008	-0.005	0.072*	0.038*	-0.001
	(0.014)	(0.001)	(0.043)	(0.001)	(0.015)	(0.002)	(0.002)	(0.012)	(0.038)	(0.050)	(0.011)	(0.013)	(0.041)	(0.020)	(0.013)
Sideways	0.067***	-0.004**	0.040	-0.003***	0.023	0.004**	0.003	-0.076***	0.060	0.021	-0.027**	-0.001	-0.026	0.037*	-0.004
	(0.016)	(0.002)	(0.040)	(0.001)	(0.016)	(0.002)	(0.002)	(0.016)	(0.043)	(0.062)	(0.011)	(0.014)	(0.033)	(0.019)	(0.013)
Downwards	0.042***	-0.003*	0.041	0.009***	-0.022	-0.003	-0.002	-0.119***	-0.323***	-0.018	-0.001	0.027	0.007	0.083***	0.073***
	(0.015)	(0.002)	(0.056)	(0.001)	(0.020)	(0.002)	(0.002)	(0.021)	(0.054)	(0.072)	(0.017)	(0.018)	(0.053)	(0.022)	(0.018)
Team FE	YES														
Observations	1,892														

Table continued on next page

Panel B: Passing Yards

VARIABLES	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Passing Yards	Black* Passing Yards	Tenure* Passing Yards	postRR	Black* postRR	GM change	OC=1
No Exit	-0.120***	0.008***	-0.047	-0.002	-0.008	-0.003	0.000	0.218***	0.085	0.048	0.031	-0.021	-0.053	-0.164***	-0.072***
	(0.021)	(0.002)	(0.074)	(0.002)	(0.024)	(0.002)	(0.003)	(0.022)	(0.065)	(0.092)	(0.019)	(0.022)	(0.063)	(0.030)	(0.021)
Promotion	0.015	-0.002	0.014	-0.004***	0.012	0.003*	-0.000	0.030***	0.045	-0.154***	0.010	-0.004	0.074*	0.030	-0.000
	(0.013)	(0.001)	(0.039)	(0.001)	(0.015)	(0.002)	(0.002)	(0.011)	(0.039)	(0.054)	(0.012)	(0.013)	(0.042)	(0.019)	(0.012)
Sideways	0.064***	-0.004**	0.003	-0.003***	0.023	0.004**	0.003	-0.079***	0.037	0.088	-0.020	-0.003	-0.025	0.040**	-0.003
	(0.017)	(0.002)	(0.045)	(0.001)	(0.017)	(0.002)	(0.002)	(0.015)	(0.045)	(0.066)	(0.013)	(0.014)	(0.033)	(0.019)	(0.013)
Downwards	0.042***	-0.002	0.030	0.010***	-0.027	-0.003*	-0.003	-0.168***	-0.167***	0.018	-0.021	0.027	0.004	0.094***	0.076***
	(0.016)	(0.002)	(0.055)	(0.001)	(0.020)	(0.002)	(0.002)	(0.021)	(0.052)	(0.072)	(0.016)	(0.018)	(0.050)	(0.022)	(0.018)
Team FE	YES														
Observations	1,892														

Panel C: Rushing Yards

VARIABLES	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Rushing Yards	Black* Rushing Yards	Tenure* Rushing Yards	postRR	Black* postRR	GM change	OC=1
No Exit	-0.102***	0.008***	0.013	-0.002	-0.016	-0.003	0.000	0.202***	0.138**	0.002	-0.003	-0.016	-0.069	-0.166***	-0.069***
	(0.019)	(0.002)	(0.066)	(0.002)	(0.024)	(0.002)	(0.003)	(0.023)	(0.060)	(0.091)	(0.017)	(0.023)	(0.065)	(0.030)	(0.022)
Promotion	0.021	-0.002	-0.122**	-0.005***	0.010	0.003*	0.000	0.023**	0.050	0.082	0.000	-0.007	0.094**	0.031	0.002
	(0.014)	(0.002)	(0.057)	(0.001)	(0.015)	(0.002)	(0.002)	(0.012)	(0.037)	(0.071)	(0.010)	(0.013)	(0.040)	(0.019)	(0.013)
Sideways	0.058***	-0.004***	0.068**	-0.003***	0.026	0.004**	0.003	-0.069***	-0.023	-0.034	-0.009	-0.002	-0.032	0.036*	-0.007
	(0.014)	(0.001)	(0.032)	(0.001)	(0.017)	(0.002)	(0.002)	(0.017)	(0.041)	(0.051)	(0.010)	(0.015)	(0.034)	(0.019)	(0.013)
Downwards	0.024	-0.002	0.040	0.010***	-0.020	-0.004*	-0.003	-0.156***	-0.164***	-0.050	0.011	0.026	0.007	0.099***	0.073***
	(0.017)	(0.002)	(0.058)	(0.001)	(0.020)	(0.002)	(0.002)	(0.022)	(0.054)	(0.075)	(0.017)	(0.019)	(0.055)	(0.023)	(0.018)
Team FE	YES														
Observations	1,892														

Cluster robust standard errors in parentheses (clustered at the Coach level)

*** p<0.01, ** p<0.05, * p<0.1

Table A2: Dropping Retired Coaches

Panel A: Head Coaches

VARIABLES	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Win Loss Percentage	Black* WLP	Tenure* WLP	postRR	Black* postRR	GM change
No Exit	-0.083***	0.002**	-0.012	-0.003	0.007	0.003	-0.002	0.080*	0.444***	-0.010	0.066***	0.013	0.025	-0.122***
	(0.015)	(0.001)	(0.133)	(0.002)	(0.021)	(0.003)	(0.002)	(0.046)	(0.114)	(0.251)	(0.023)	(0.023)	(0.089)	(0.027)
Sideways	0.029***	-0.001*	0.144***	-0.001	-0.005	0.001	0.001	0.024	0.029	-0.244**	-0.024*	-0.026**	-0.054	0.009
	(0.010)	(0.001)	(0.043)	(0.001)	(0.012)	(0.001)	(0.002)	(0.020)	(0.083)	(0.098)	(0.014)	(0.013)	(0.035)	(0.015)
Downwards	0.054***	-0.001	-0.132	0.004**	-0.001	-0.004*	0.001	-0.105**	-0.473***	0.255	-0.042*	0.014	0.029	0.112***
	(0.013)	(0.001)	(0.113)	(0.002)	(0.021)	(0.003)	(0.002)	(0.048)	(0.116)	(0.211)	(0.024)	(0.022)	(0.080)	(0.023)
Team FE	NO													
Observations	966													

Panel B: Coordinators

VARIABLES	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Points	Black* Points	Tenure* Points	postRR	Black* postRR	GM change	OC=1
No Exit	-0.124***	0.008***	-0.068	-0.000	-0.009	-0.003	-0.000	0.123***	0.218***	0.107	0.034	-0.029	-0.055	-0.156***	-0.073***
	(0.022)	(0.002)	(0.058)	(0.002)	(0.025)	(0.002)	(0.003)	(0.026)	(0.069)	(0.099)	(0.023)	(0.023)	(0.063)	(0.030)	(0.021)
Promotion	0.014	-0.001	-0.042	-0.005***	0.012	0.003**	-0.000	0.007	0.074*	-0.034	0.007	-0.005	0.076*	0.037*	-0.000
	(0.014)	(0.002)	(0.050)	(0.001)	(0.015)	(0.002)	(0.002)	(0.013)	(0.041)	(0.061)	(0.012)	(0.013)	(0.043)	(0.020)	(0.013)
Sideways	0.067***	-0.004***	0.059	-0.003**	0.024	0.004**	0.002	-0.069***	0.047	-0.012	-0.026**	-0.003	-0.028	0.039**	-0.002
	(0.016)	(0.002)	(0.037)	(0.001)	(0.017)	(0.002)	(0.002)	(0.018)	(0.049)	(0.066)	(0.013)	(0.014)	(0.034)	(0.019)	(0.013)
Downwards	0.042***	-0.003	0.051	0.008***	-0.028	-0.004*	-0.002	-0.061**	-0.339***	-0.061	-0.014	0.037**	0.007	0.080***	0.075***
	(0.016)	(0.002)	(0.047)	(0.001)	(0.020)	(0.002)	(0.003)	(0.024)	(0.061)	(0.075)	(0.017)	(0.019)	(0.049)	(0.022)	(0.018)
Team FE	YES														
Observations	1,859														

Cluster robust standard errors in parentheses (clustered at the Coach level)

*** p<0.01, ** p<0.05, * p<0.1

Table A3: Redefining Exits to College Teams by prestige

Panel A: Head Coaches

VARIABLES	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Win Loss Percentage	Black* WLP	Tenure* WLP	postRR	Black* postRR	GM change
No Exit	-0.073***	0.002***	0.012	-0.006**	0.000	0.002	-0.002	0.072	0.535***	-0.054	0.057***	0.016	0.031	-0.118***
	(0.014)	(0.001)	(0.148)	(0.002)	(0.022)	(0.003)	(0.002)	(0.047)	(0.117)	(0.257)	(0.021)	(0.023)	(0.100)	(0.028)
Sideways	0.027***	-0.001*	0.165***	-0.001	-0.005	0.001	0.000	0.029	0.012	-0.292***	-0.021*	-0.023*	-0.088**	0.005
	(0.010)	(0.001)	(0.041)	(0.001)	(0.012)	(0.001)	(0.002)	(0.019)	(0.078)	(0.098)	(0.012)	(0.012)	(0.042)	(0.014)
Downwards	0.046***	-0.001*	-0.177	0.006***	0.005	-0.003	0.002	-0.101**	-0.547***	0.347*	-0.035*	0.007	0.058	0.113***
	(0.012)	(0.001)	(0.133)	(0.002)	(0.021)	(0.002)	(0.002)	(0.048)	(0.112)	(0.208)	(0.021)	(0.022)	(0.093)	(0.026)
Team FE	NO													
Observations	986													

Panel B: Coordinators

VARIABLES	Tenure	Tenure Squared	Black	Age	Played	NFL Exp.	College Exp.	Post Season	Points	Black* Points	Tenure* Points	postRR	Black* postRR	GM change	OC=1
No Exit	-0.123***	0.008***	-0.078	-0.002	-0.006	-0.003	0.000	0.118***	0.229***	0.129	0.032	-0.024	-0.061	-0.154***	-0.072***
	(0.021)	(0.002)	(0.059)	(0.002)	(0.024)	(0.002)	(0.003)	(0.026)	(0.069)	(0.098)	(0.022)	(0.022)	(0.063)	(0.030)	(0.021)
Promotion	0.017	-0.001	-0.018	-0.005***	0.011	0.004**	-0.001	0.007	0.095**	-0.051	0.001	-0.006	0.069*	0.027	-0.004
	(0.013)	(0.001)	(0.046)	(0.001)	(0.014)	(0.001)	(0.002)	(0.012)	(0.038)	(0.057)	(0.011)	(0.013)	(0.039)	(0.020)	(0.012)
Sideways	0.067***	-0.004***	0.070**	-0.003**	0.021	0.003**	0.002	-0.063***	0.065	-0.020	-0.029**	0.006	-0.038	0.034*	-0.004
	(0.016)	(0.002)	(0.036)	(0.001)	(0.016)	(0.002)	(0.002)	(0.017)	(0.047)	(0.063)	(0.012)	(0.014)	(0.033)	(0.019)	(0.013)
Downwards	0.039**	-0.003*	0.026	0.009***	-0.026	-0.004**	-0.001	-0.063***	-0.390***	-0.058	-0.004	0.024	0.030	0.093***	0.080***
	(0.015)	(0.002)	(0.049)	(0.001)	(0.020)	(0.002)	(0.002)	(0.024)	(0.063)	(0.076)	(0.018)	(0.018)	(0.052)	(0.022)	(0.018)
Team FE	YES														
Observations	1,892														

Cluster robust standard errors in parentheses (clustered at the Coach level)

*** p<0.01, ** p<0.05, * p<0.1

Chapter References

Baker, G.P., Jensen, M.C. and Murphy, K.J. (1988), *'Compensation and Incentives: Practice vs. Theory'*, The Journal of Finance, Vol. 43, No.3, pp 593-616

BBC (2018), *'Football Association adopts 'Rooney Rule' as Martin Glenn outlines 2018 plans'*, https://www.bbc.co.uk/sport/football/42623021 (Accessed 22/05/2019)

Bryson, A., Buraimo, B., Farnell, A. and Simmons, R. (2021, a), *'Special Ones? The Effect of Head Coaches on Football Team Performance'*, IZA Discussion Paper No. 14104

Bryson, A., Buraimo, B., Farnell, A. and Simmons, R. (2021, b), *'Time to go? Head Coach Quits and Dismissals in Professional Football'*, De Economist, Vol.169, No.1, pp 81-105

Denver Post (2017), '*Climbing a slippery mountain: Why NFL's black offensive coaches struggle to become a head coach, coordinator'*, https://www.denverpost.com/2017/09/11/nfl-black-coaches/amp/ (Accessed 19/08/19)

DuBois, C. (2015), *'The Impact of "Soft" Affirmative Action Policies on Minority Hiring in Executive Leadership: The Case of the NFL's Rooney Rule'*, American Law and Economics Review, Vol.18, No.1, pp 208-233

Fee, C.E., Hadlock, C.J. and Pierce, J.R. (2006), *'Promotions in the Internal and External Labor Market: Evidence from Professional Football Coaching Careers'*, The Journal of Business, Vol.79, No.2, pp 821-850

Foreman, J., Soebbing, B., Seifried, C. and Agyemang, K (2018), *'Examining Relationships between Managerial Advancement and Centrality, Race, and the Rooney Rule'*, International Journal of Sport Management, Vol.19, pp 315-338

Fort, R., Lee, Y.H. and Berri, D. (2008), *'Race, Technical Efficiency and Retention'*, International Journal of Sports Finance, Vol.3, No.2, pp 84-97

Fortune (2020), *'The Number of Black CEOs in the Fortune 500 Remains Very Low'*, https://fortune.com/2020/06/01/black-ceos-fortune-500-2020-african-american-business-leaders/ (Accessed 24/06/21)

Gaynor, M., Seider, H. and Vogt, W.B. (2005), *'The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing'*, American Economic Review, Vol.95, No.2, pp 243-247

Gielen, A.C. and van Ours, J.C. (2006), *'Why do Worker-Firm matches dissolve'*, IZA Discussion Paper No. 2165

Goff, B., Wilson, D. and Zimmer, D. (2019), *'The Effect of Management Changes on Winning in Professional Sports: Analysis Using a Dynamic Lag Adjustment Model'*, Managerial and Decision Economics, Vol.40, No.8, pp 982-992

Goodall, A.H., Kahn, L.M. and Oswald, A.J. (2011), *'Why do leaders matter? A study of expert knowledge in a superstar setting'*, Journal of Economic Behaviour & Organisation, Vol.77, No.3, pp 265-284

Groothuis, P.A. and Hill, J.R. (2004), *'Exit Discrimination in the NBA: A Duration Analysis of Career Length'*, Economic Inquiry, Vol.42, No.2, pp 341-349

Grusky, O. (1963), *'The Effects of Formal Structure on Managerial Recruitment: A Study of Baseball Organization'*, Sociometry, Vol.26, No.3, pp 345-353

Holmes, P. (2011), *'Win or Go Home: Why College Football Coaches Get Fired'*, Journal of Sports Economics, Vol.12, No.2, pp 157-178

Holzer, H.J. (2007), *'The economic impact of affirmative action in the US'*, Swedish Economic Policy Review, Vol.14, No.1, pp 41-71

Ilmakunnas, P., Maliranta, M. and Vainiomäki, J. (2005), *'Worker turnover and productivity growth'*, Applied Economics Letters, Vol.12, No.7, pp 395-398

Kahn, L.M. (2000), *'The Sports Business as a Labor Market Laboratory'*, Journal of Economic Perspectives, Vol.14, No.3, pp 75-94

Kilvington, D. (2018), *'Does English football warrant the Rooney Rule? Assessing the thoughts of British Asian coaches'*, Sport in Society, Vol.22, No.3, pp 432-448

Kopkin, N.A. (2014), *'You're Fired!: The Impact of Race on the Firing of Black Coaches in Major College Football'*, The Review of Black Political Economy, Vol.41, No.4, pp 373-392

Latimer, S.R. and Mathes, S.A. (1985), *'Black College Football Coaches' Social, Educational, Athletic and Career Pattern Characteristics'*, Journal of Sport Behaviour, Vol.8, No.3, pp 149-162

Lazear, E.P. and Rosen, S. (1981), *'Rank-Order Tournaments as Optimum Labor Market Contracts'*, Journal of Political Economy, Vol.89, No.5, pp 841-864

Lazear, E.P., Shaw, K.L. and Stanton, C.T. (2015), *'The Value of Bosses'*, Journal of Labour Economics, Vol.33, No.4, pp 823-861

Madden, J.F. and Ruther, M. (2011), *'Has the NFL's Rooney Rule Efforts "Leveled the Field" for African American Head Coach Candidates?'*, Journal of Sports Economics, Vol.12, No.2, pp 127-142

Madden, J.F. (2004), *'Differences in the success of NFL coaches by race, 1990-2002'*, Journal of Sports Economics, Vol.5, No.1, pp 6-19

Malone, K.D., Couch, J.F. and Barrett, J.D. (2012), *'The Market for NFL Coaches and Managers'*, Chapter 11 in Quinn K.G. *'The Economics of the National Football League'*, Springer Publishing, p191

Muehlheusser, G., Schneeman, S. Sliwka, D. and Wallmeier, N. (2018), *'The Contribution of Managers to Organizational Success: Evidence from German Soccer'*, Journal of Sports Economics, Vol.19, No.6, pp 786-819

NFL (2019), *'Aspiring Coaches, GMs frustrated with NFL's diversity issue'*, http://www.nfl.com/news/story/0ap3000001038346/article/aspiring-coaches-gms-frustrated-with-nfls-diversity-issue (Accessed 27/09/19)

Peter, L.J. & Hull, R. (1969), *'The Peter Principle'*, Souvenir Press, London

Pieper, J., Nüesch, S. and Franck, E. (2014), *'How Performance Expectations Affect Managerial Replacement Decisions'*, Schmalenbach Business Review, Vol.66, No.1, pp 5-23

Pitts, J.D. and Yost, D.M. (2013), *'Racial Position Segregation in Intercollegiate Football: Do Players become more Racially Segregated as they Transition from High School to College?'*, The Review of Black Political Economy, Vol.40, No.2, pp 207-230

Rosen, S. (1982), *'Authority, Control, and the Distribution of Earnings'*, The Bell Journal of Economics, Vol.13, No.2, pp 311-323

Solow, B.L., Solow, J.L. and Walker, T.B. (2011), *'Moving on up: The Rooney Rule and minority hiring in the NFL'*, Labour Economics, Vol.18, No.3, pp 332-337

Van Ours, J. C. and Van Tuijl, M. A. (2016), *'In-season Head Coach Dismissals and the Performance of Professional Football Teams'*, Economic Inquiry, Vol.54, No.1, pp 591-60

Waldman, M. (1984), *'Job Assignments, Signalling and Efficiency'*, The RAND Journal of Economics, Vol.15, No.2, pp 255-267

Wangrow, D.B., Schepker, D.J. and Barker III, V.L. (2018), *'Power, performance and expectations in the dismissal of NBA coaches: A survival analysis study'*, Sports Management Review, Vol.21, No.4, pp 333-346

Holistic Conclusions and Implications

This thesis has offered three contributions at the intersection of Sports, Labour and Personnel Economics. Throughout, the thesis highlights the unique opportunity afforded to empirical economics work by using sports data, whereby the contributions and outputs of workers, firms and supervisors can be accurately assessed. Perhaps more importantly, it is also a setting where the aims of participants and the incentives that drive them are clear (players want to win, usually), they operate in the real world and stakes are typically high (Palacios-Huerta, 2014).

Of course, the natural critique of using sports data (as with any other data) to inform economics topics more broadly is one of external validity; a professional athlete is not representative of a worker in a more conventional labour market. For example, sports tasks involve utilising fine motor skills, frequent travel, unusual types of pressure and media exposure, all of which may be uncommon features of work for most people (Bar-Eli et al. (2020)). Yet, as Papps (2020) points out, as well-trained and highly skilled workers, athletes are possibly more representative of a modern labour market than ever before, more so than fruit pickers and supermarket cashiers. Even the wage inequalities that exist in sport are increasingly becoming a feature of other labour markets. In this sense, it is perhaps less of a stretch than at first glance to see the similarities between sports and other industries.

Sports is an industry that deserves special attention from economists, not just because of the opportunities from the superb data that it gives us. It is a globally important industry that supports numerous jobs, attracts vast sums of public and private investment and captures the interest and contributes to the wellbeing of numerous sports fans (Bryson et al. (2015), Kavetsos and Szymanski (2010)).

In Chapter 1, we discussed how football Head Coaches contribute to the success of football teams. There are theoretical justifications to believe that coaches are in a position to make a difference to the on-field success of their team, though this is caveated with arguments surrounding the mediocrity of or similarly talented coaches. In fact, much of the literature would tend to support the view that coaches fail to make any difference to team performance (van Ours and van Tuijl (2016)). However, much of the work has failed to capture a number of potentially important distinctions. Primarily, this surrounds the nature of the exit of the previous coach. In a principal-agent framework, either the coach (agent) or the owner (principal) can terminate employment, and these two separate events result in quite different theoretical predictions. We also distinguish between short-run (motivational bump) and long-run (learning process) effects of coach turnover, while our work makes use of football matches from four separate countries.

The results of this work show that the distinction between quits and dismissals may be important, though findings are sensitive to the way in which the follow up period is defined. In cases where a team experiences no subsequent coaching change, team performance does indeed improve after a dismissal, however, team performance also improves after a quit, albeit to a lesser extent and occurring only in the longer term. These results are interpreted as an upper bound of the effect of a coaching change, and we take this as loose evidence to suggest that teams can benefit from Head Coach turnover, as long as their new appointment is the 'right' person for the job. Since there is no official interview process that must take place prior to the appointment of a new Head Coach, it is possible that teams make mistakes when selecting their new coach (often the replacement is in place even before the incumbent is dismissed). This work would point to the importance of finding a good job-match, rather than continually changing coaches, though the effect is still rather modest.

In Chapter 2, we considered the effects of task switching on worker performance. Experimental evidence would point towards subjects struggling when faced with the demands of task switching (for example, Buser and Peter (2012)), a feature synonymous with modern day work. Major League Baseball offers a unique opportunity to study how subjects respond in a more natural setting (outside of a laboratory) and makes for a convincing empirical design by virtue of its two-league structure, whereby players in the two leagues face different rules. In particular, pitchers in the American League can specialise on their primary task; pitching (this is what their career has been built upon). Whereas pitchers in the National League are required to both pitch and bat. Given that pitchers are usually pretty poor batters, and that batting is a skill which they rarely practice, we can make the assumption that players are randomly assigned to the two leagues, and thus, randomly affected by the task switching requirement.

Our results suggest that, somewhat surprisingly given the unfamiliarity of the task, a pitcher's subsequent pitching performance actually improves after batting in the previous inning. We observe gains in velocity, reduced likelihood of allowing walks, and giving up fewer runs to the opposition in the subsequent inning. However, for those pitchers who had a successful at bat and got on base, their subsequent pitching performance declined. We believe this highlights an important distinction between the mental and physical implications of task switching. Moreover, these findings could have implications for the upcoming seasons of MLB, where it appears likely that the Designated Hitter rule will be made uniform across the two leagues. Aside from offering a potentially interesting follow up study in a few years' time, the removal of the requirement for pitchers to bat could mean that pitching performance in the National League declines. However, the adoption of the rule will also mean a removal of a layer of strategy for coaches in the National League, and if fans value offensive output, then an addition of a better hitter could lead to increased attendances (Domazlicky and Kerr (1990)).

Finally, in Chapter 3, we examined transitions between and out of the top two levels of coaching staffs for teams in the National Football League. Our analysis finds that teams are more likely to promote coordinators who are younger, have more years of coaching experience, and are in charge of more successful offensive or defensive units, while older and poorly performing coaches are more likely to be dismissed. Of particular importance in this chapter was the focus on the coach's race, and whether the league's affirmative action policy, the Rooney Rule, had fulfilled its aim of getting more minority candidates into top coaching positions. The findings hint at some moderate success of the rule. Minority coordinators are more likely to be promoted after the implementation of the rule in 2003, though the result was only significant at the 10% level. Teams are less likely to (re)hire a Head Coach who had previously been working at the same level, though this did not differ by the race of the coach. However, we also present strong evidence suggesting that teams are now appointing equally skilled white and minority coaches. Before 2003, minority Head Coaches had a higher win-loss percentage, consistent with the view that teams held minority coaches to higher standards, and thus, they had to be better coaches (Madden (2004), Madden and Ruther (2011)). This performance difference no longer exists after the introduction of the rule, suggesting that teams are now considering, and ultimately appointing, equally skilled white and minority coaches. This was despite there always appearing to have been a supply of skilled coordinators, regardless of race, indicated by the equivalent performance records of coordinators pre- and post-2003. It would appear that the Rooney Rule has allowed coaches a chance to showcase their skills at interview, where previously they may have been overlooked.

However, these results are conditional on coaches getting to coordinator positions to begin with. At present, we know very little about the transitions of coaches at the lower end of the coaching hierarchy and any possible barriers facing minority coaches from getting into these positions at the start of their career. This is a topic that merits further analysis.

Chapter References

Bar-Eli, M., Krumer, A. and Morgulev, E. (2020), *'Ask Not What Economics can do for Sports – Ask What Sports can do for Economics'*, Journal of Behavioural and Experimental Economics, Vol.89, 101597

Bryson, A., Frick, B. and Simmons, R. (2015), *'Sports Economics: It May Be Fun But What's The Point?'*, National Institute Economic Review, Vol.232, No.1, pp R1-R3

Buser, T. and Peter, N. (2012), *'Multitasking'*, Experimental Economics, Vol.15, No.4, pp 641-655

Domazlicky, B.R. and Kerr, P.M. (1990), *'Baseball Attendance and the Designated Hitter'*, The American Economist, Vol.34, No.1, pp 62-68

Kavetsos, G. and Szymanski, S., (2010), '*National well-being and international sports events'*, Journal of Economic Psychology. Vol.31, No.2, pp 158-171

Madden, J.F. (2004), *'Differences in the success of NFL coaches by race, 1990-2002'*, Journal of Sports Economics, Vol.5, No.1, pp 6-19

Madden, J.F. and Ruther, M. (2011), *'Has the NFL's Rooney Rule Efforts "Leveled the Field" for African American Head Coach Candidates?'*, Journal of Sports Economics, Vol.12, No.2, pp 127-142

Palacios-Huerta, I. (2014), *'Beautiful Game Theory'*, Princeton: Princeton University Press

Papps, K.L. (2020), *'Sports at the Vanguard of Labor Market Policy'*, IZA World of Labor

Van Ours, J. C. and Van Tuijl, M. A. (2016), *'In-season Head Coach Dismissals and the Performance of Professional Football Teams'*, Economic Inquiry, Vol.54, No.1, pp 591-60

Consolidated References

Aral, S., Brynjolfsson, E. and Van Alstyne, M. (2012), *'Information, Technology and Information Worker Productivity'*, Information Systems Research, Vol.23, No.3, Part 2, pp 849-867

Ashenfelter, O. (1978), *'Estimating the Effect of Training Programs on Earnings'*, The Review of Economics and Statistics, pp 47-57

Baker, G.P., Jensen, M.C. and Murphy, K.J. (1988), *'Compensation and Incentives: Practice vs. Theory'*, The Journal of Finance, Vol. 43, No.3, pp 593-616

Bar-Eli, M., Krumer, A. and Morgulev, E. (2020), *'Ask Not What Economics can do for Sports – Ask What Sports can do for Economics'*, Journal of Behavioural and Experimental Economics, Vol.89, 101597

BBC (2018), *'Football Association adopts 'Rooney Rule' as Martin Glenn outlines 2018 plans'*, https://www.bbc.co.uk/sport/football/42623021 (Accessed 22/05/2019)

Bennedsen, M., Perez-Gonzalez, F. and Wolfenzon, D. (2012), *'Evaluating the Impact of the Boss: Evidence from CEO Hospitalization Events'*, unpublished manuscript

Berri, D.J., Leeds, M.A., Leeds, E.M. and Mondello, M. (2009), *'The Role of Managers in Team Performance'*, International Journal of Sports Finance, Vol.4, No.2, pp 75-93

Besley, T., Montalvo, J. G., and Reynal-Querol, M. (2011), '*Do Educated Leaders Matter?'*, The Economic Journal, Vol.121, No. 554, F205-227

Besters, L., Van Ours, J. and Van Tuijl, M. (2016), *'Effectiveness of In-Season Manager Changes in English Premier League Football'*, De Economist, Vol.164, No.3, pp 335-356

Bradbury, J.C. (2017), *'Hired to be Fired: The Publicity Values of Managers'*, Managerial and Decision Economics, Vol.38, No.7, pp 929-940

Bruinshoofd, A. and ter Weel, B. (2003), *'Manager to go? Performance Dips Reconsidered with Evidence from Dutch Football'*, European Journal of Operational Research, Vol.148, No.2, pp 233-246

Bryson, A., Buraimo, B., Farnell, A. and Simmons, R. (2021), *'Time to go? Head Coach Quits and Dismissals in Professional Football'*, De Economist, Vol.169, No.1, pp 81-105

Bryson, A., Buraimo, B., Farnell, A. and Simmons, R. (2021), *'Special Ones? The Effect of Head Coaches on Football Team Performance'*, IZA Discussion Paper No. 14104

Bryson, A., Frick, B. and Simmons, R. (2015), *'Sports Economics: It May Be Fun But What's The Point?'*, National Institute Economic Review, Vol.232, No.1, pp R1-R3

Buser, T. and Peter, N. (2012), *'Multitasking'*, Experimental Economics, Vol.15, No.4, pp 641-655

Bushnell, B.D., Anz, A.W., Noonan, T.J., Torry, M.R. and Hawkins, R.J. (2010), *'Association of Maximum Pitch Velocity and Elbow Injury in Professional Baseball'*, The American Journal of Sports Medicine, Vol.38, No.4, pp 728-732

Cassavell, A.J. (2016), *'Universal Truth? Execs talk NL DH possibility'*, from MLB News, https://www.mlb.com/news/might-dh-be-in-national-league-s-future-c162124094 (Accessed 03/06/2020)

Collewet, M. and Sauermann, J. (2017), *'Working Hours and Productivity'*, Labour Economics, Vol.47, pp 96-106

Coviello, D., Ichino, A. and Perisco, N. (2015), *'The Inneficiency of Worker Time Use'*, Journal of the European Economic Association, Vol.13, No.5, pp 906-947

Crocker, T.D. and Horst, R.L. (1981), *'Hours of Work, Labour Productivity, and Environmental Conditions: A Case Study'*, The Review of Economics and Statistics, Vol.63, No.3, pp 361-368

D'Addona, S. and Kind, A. (2014), *'Forced Manager Turnovers in English Soccer Leagues'*, Journal of Sports Economics, Vol.15, No.2, pp 150-179

De Grip, A. (2015), *'The importance of informal learning at work*', *IZA World of Labor*

De Paola, M. and Scoppa, V. (2012), *'The Effects of Managerial Turnover: Evidence from Coach Dismissals in Italian Soccer Teams'*, Journal of Sports Economics, Vol.13, No.2, pp 152-168

Denver Post (2017), '*Climbing a slippery mountain: Why NFL's black offensive coaches struggle to become a head coach, coordinator'*, https://www.denverpost.com/2017/09/11/nfl-black-coaches/amp/ (Accessed 19/08/19)

Dobson, S. and Goddard, J. (2011), '*The Economics of Football'*, 2nd ed. Cambridge: Cambridge University Press

Domazlicky, B.R. and Kerr, P.M. (1990), *'Baseball Attendance and the Designated Hitter'*, The American Economist, Vol.34, No.1, pp 62-68

DuBois, C. (2015), *'The Impact of "Soft" Affirmative Action Policies on Minority Hiring in Executive Leadership: The Case of the NFL's Rooney Rule'*, American Law and Economics Review, Vol.18, No.1, pp 208-233

Entine, O.A. and Small, D.S. (2008), *'The role of Rest in the NBA Home-Court Advantage'*, Journal of Quantitative Analysis in Sports, Vol.4, No.2

Escamilla, R.F., Barrentine S.W. and Fleisig G.S. (2007), *'Pitching Biomechanics as a Pitcher Approaches Muscular Fatigue during a Simulated Baseball Game'*, The American Journal of Sports Medicine, Vol.35, No.1, pp 23-33

Farber, H. S. (1999), '*Mobility and Stability: The Dynamics of Job Change in Labor Markets',* Chapter 37 in O. Ashenfelter and D. Card (eds.) '*Handbook of Labor Economics',* Vol.3, pp. 2439-2483

Fee, C.E., Hadlock, C.J. and Pierce, J.R. (2006), *'Promotions in the Internal and External Labor Market: Evidence from Professional Football Coaching Careers'*, The Journal of Business, Vol.79, No.2, pp 821-850

Ferguson, A. (2013), '*Alex Ferguson: My Autobiography'*, Hodder and Stoughton

Foreman, J., Soebbing, B., Seifried, C. and Agyemang, K (2018), *'Examining Relationships between Managerial Advancement and Centrality, Race, and the Rooney Rule'*, International Journal of Sport Management, Vol.19, pp 315-338

Fort, R., Lee, Y.H. and Berri, D. (2008), *'Race, Technical Efficiency and Retention'*, International Journal of Sports Finance, Vol.3, No.2, pp 84-97

Fortune (2020), *'The Number of Black CEOs in the Fortune 500 Remains Very Low'*, https://fortune.com/2020/06/01/black-ceos-fortune-500-2020-african-american-business-leaders/ (Accessed 24/06/21)

Gaynor, M., Seider, H. and Vogt, W.B. (2005), *'The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing'*, American Economic Review, Vol.95, No.2, pp 243-247

Gielen, A.C. and van Ours, J.C. (2006), *'Why do Worker-Firm matches dissolve'*, IZA Discussion Paper No. 2165

Goff, B., Wilson, D. and Zimmer, D. (2019), *'The Effect of Management Changes on Winning in Professional Sports: Analysis Using a Dynamic Lag Adjustment Model'*, Managerial and Decision Economics, Vol.40, No.8, pp 982-992

Goodall, A.H., Kahn, L.M. and Oswald, A.J. (2011), *'Why do leaders matter? A study of expert knowledge in a superstar setting'*, Journal of Economic Behaviour & Organisation, Vol.77, No.3, pp 265-284

Green, C., Lozano, F and Simmons, R. (2015), *'Rank-Order Tournaments, Probability of Winning and Investing in Talent: Evidence from Champions' League Qualifying Rules'*, National Institute Economic Review, Vol.232, No.1, R30-40

Groothuis, P.A. and Hill, J.R. (2004), *'Exit Discrimination in the NBA: A Duration Analysis of Career Length'*, Economic Inquiry, Vol.42, No.2, pp 341-349

Grusky, O. (1963), *'The Effects of Formal Structure on Managerial Recruitment: A Study of Baseball Organization'*, Sociometry, Vol.26, No.3, pp 345-353

Hainmueller, J. (2012), *'Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples'*, Political Analysis, pp 25-46

Hainmueller, J. and Xu, J. (2013), *'ebalance: A Stata Package for Entropy Balancing'*, Journal of Statistical Software, Vol.54, No.7: 1-18

Hart, R.A. (2004), *'The Economics of Overtime Working'*, Cambridge University Press

Holmes, P. (2011), *'Win or Go Home: Why College Football Coaches Get Fired'*, Journal of Sports Economics, Vol.12, No.2, pp 157-178

Holzer, H.J. (2007), *'The economic impact of affirmative action in the US'*, Swedish Economic Policy Review, Vol.14, No.1, pp 41-71

Hommel, B., Fischer, R., Colzato, L.S., van den Wildenberg, W.P. and Cellini, C. (2012), *'The Effect of fMRI (noise) on Cognitive Control'*, Journal of Experimental Psychology: Human Perception and Performance, Vol.32, No.2, pp 290-301

Ilmakunnas, P., Maliranta, M. and Vainiomäki, J. (2005), *'Worker turnover and productivity growth'*, Applied Economics Letters, Vol.12, No.7, pp 395-398

Kahn, L.M. (2000), *'The Sports Business as a Labor Market Laboratory'*, Journal of Economic Perspectives, Vol.14, No.3, pp 75-94

Kavetsos, G. and Szymanski, S., (2010), '*National well-being and international sports events'*, Journal of Economic Psychology. Vol.31, No.2, pp 158-171

Keller, R.A., Marshall, N.E., Guest, J.M., Okoroha, K.R., Jung, E.K. and Moutzouros, V. (2016), *'Major League Baseball Pitch Velocity and Pitch Type Associated With the Risk of Ulnar Collateral Ligament Injury'*, Journal of Shoulder and Elbow Surgery, Vol.25, No.4, pp 671-675

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A.M. and Koch, I. (2010), *'Control and Interference in Task Switching – A Review'*, Psychological Bulletin, Vol.136, No.5, pp 849-874

Kilvington, D. (2018), *'Does English football warrant the Rooney Rule? Assessing the thoughts of British Asian coaches'*, Sport in Society, Vol.22, No.3, pp 432-448

Kopkin, N.A. (2014), *'You're Fired!: The Impact of Race on the Firing of Black Coaches in Major College Football'*, The Review of Black Political Economy, Vol.41, No.4, pp 373-392

Kuper, S. and Szymanski, S. (2009), '*Soccernomics',* Harper Collins

Latimer, S.R. and Mathes, S.A. (1985), *'Black College Football Coaches' Social, Educational, Athletic and Career Pattern Characteristics'*, Journal of Sport Behaviour, Vol.8, No.3, pp 149-162

Lazear, E.P. and Rosen, S. (1981), *'Rank-Order Tournaments as Optimum Labor Market Contracts'*, Journal of Political Economy, Vol.89, No.5, pp 841-864

Lazear, E.P., Shaw, K.L. and Stanton, C.T. (2015), *'The Value of Bosses'*, Journal of Labour Economics, Vol.33, No.4, pp 823-861

Levitt, S. (2021), *'Heads or Tails: The Impact of a Coin Toss on Major Life Decisions and Subsequent Happiness'*, The Review of Economic Studies, Vol.88, No.1, pp 378-405

Lu, S.F. and Lu, L.X. (2017), *'Do Mandatory Overtime Laws Improve Quality? Staffing Decisions and Operational Flexibility of Nursing Homes'*, Management Science, Vol.63, No.11, pp 3566-3585

Lyman S., Fleisig, G.S., Andrews, J.R. and Osinski, E.D. (2002), *'Effect of Pitch Type, Pitch Count, and Pitching Mechanics on Risk of Elbow and Shoulder Pain in Youth Baseball Pitchers'*, The American Journal of Sports Medicine, Vol.30, No.4, pp 463-468

Madden, J.F. (2004), *'Differences in the success of NFL coaches by race, 1990-2002'*, Journal of Sports Economics, Vol.5, No.1, pp 6-19

Madden, J.F. and Ruther, M. (2011), *'Has the NFL's Rooney Rule Efforts "Leveled the Field" for African American Head Coach Candidates?'*, Journal of Sports Economics, Vol.12, No.2, pp 127-142

Madum, A. (2016), *'Managerial Turnover and Subsequent Firm Performance: Evidence from Danish Soccer* Teams', International Journal of Sport Finance, Vol.11, No.1, pp 46-62

Malone, K.D., Couch, J.F. and Barrett, J.D. (2012), *'The Market for NFL Coaches and Managers'*, Chapter 11 in Quinn K.G. *'The Economics of the National Football League'*, Springer Publishing, p191

Marcora, S.M., Staiano, W. and Manning, V. (2009), *'Mental Fatigue Impairs Physical Performance in Humans'*, Journal of Applied Psychology, Vol.106, No.3, pp 857-864

Mills, B. (2014), *'Expert Workers, Performance Standards, and On-the-Job Training: Evaluating Major League Baseball Umpires'*, Available at SSRN

Muehlheusser, G., Schneeman, S. Sliwka, D. and Wallmeier, N. (2018), *'The Contribution of Managers to Organizational Success: Evidence from German Soccer'*, Journal of Sports Economics, Vol.19, No.6, pp 786-819

Muehlheusser, G., Schneemann, S. and Sliwka, D. (2016), *'The Impact of Managerial Change on Performance: The Role of Team Heterogeneity'*, Economic Inquiry, Vol.54, No.2, pp 1121-1149

Murray, T.A., Cook, T.D., Werner, S.L., Schlegel, T.F. and Hawkins, R.J. (2001), *'The Effects of Extended Play on Professional Baseball Pitchers'*, The American Journal of Sports Medicine, Vol.29, No.2, pp 137-142

NFL (2019), *'Aspiring Coaches, GMs frustrated with NFL's diversity issue'*, http://www.nfl.com/news/story/0ap3000001038346/article/aspiring-coaches-gms-frustrated-with-nfls-diversity-issue (Accessed 27/09/19)

Nichols, M.W. (2014), *'The Impact of Visiting Team Travel on Game Outcome and Biases in NFL Betting Markets'*, Journal of Sports Economics, Vol.15, No.1, pp 78-96

Oberhofer, H., Philippovich, T. and Winner, H. (2010), *'Distance Matters in Away Games: Evidence from the German Football League'*, Journal of Economic Psychology, Vol.31, No.2, pp 200-211

Official Baseball Rules (2018). Available Online http://mlb.mlb.com/documents/0/8/0/268272080/2018_Official_Baseball_Rules.pdf (Accessed 15/09/2020)

Palacios-Huerta, I. (2014), *'Beautiful Game Theory'*, Princeton: Princeton University Press

Papps, K.L. (2020), *'Sports at the Vanguard of Labor Market Policy'*, IZA World of Labor

Peeters, T., Szymanski, S. and Terviö, M. (2016), *'Entry, Career Dynamics and Quality in the Labour Market for Talent'*, mimeo

Pencavel, J. (2015), *'The Productivity of Working Hours'*, The Economic Journal, Vol.125, No. 589, pp 2052-2076

Peter, L.J. & Hull, R. (1969), *'The Peter Principle'*, Souvenir Press, London

Pieper, J., Nüesch, S. and Franck, E. (2014), *'How Performance Expectations Affect Managerial Replacement Decisions'*, Schmalenbach Business Review, Vol.66, No.1, pp 5-23

Pitts, J.D. and Yost, D.M. (2013), *'Racial Position Segregation in Intercollegiate Football: Do Players become more Racially Segregated as they Transition from High School to College?'*, The Review of Black Political Economy, Vol.40, No.2, pp 207-230

Pope, D. and Simonsohn, U. (2011), *'Round Numbers as Goals Evidence From Baseball, SAT Takers, and the Lab'*, Psychological Science, Vol.22, No.1, pp 71-79

Rampinini, E., Impellizzeri, F.M., Castagna, C., Coutts, A.J., and Wisløff, U. (2009), *'Technical Performance During Soccer Matches of the Italian Serie A League: Effect of Fatigue and Competitive Level'*, Journal of Science and Medicine in Sport, Vol.12, No.1, pp 227-233

Rosen, S. (1982), *'Authority, Control, and the Distribution of Earnings'*, The Bell Journal of Economics, Vol.13, No.2, pp 311-323

Rosen, S. (1990), *'Contracts and the Market for Executives'*, NBER Working Paper No.3542

Rubinstein, J.S., Meyer, D.E. and Evans, J.E. (2001), *'Executive Control of Cognitive Processes in Task Switching'*, Journal of Experimental Psychology: Human Perception and Performance, Vol.27, No.4, pp 769-797

Russ, M. and Crews, D.E. (2014), *'A Survey of Multitasking Behaviors in Organizations'*, International Journal of Human Resource Studies, Vol.4, No.1, pp 137-153

Scelles, N. and Llorca, L. (2020), *'Head Coach Change and Team Performance in the French Men's Football Ligue 1, 2000-2016'*, Economics Bulletin, Vol.40, No.2, pp 920-937

Scoppa, V. (2013), *'Fatigue and Team Performance in Soccer: Evidence From the FIFA World Cup and the UEFA European Championship'*, Journal of Sports Economics, Vol.16, No.5, pp 482-507

Singh, D. (2014), *'Does Multitasking Improve Performance? Evidence from the Emergency Department'*, Manufacturing & Service Operations Management, Vol.16, No.2, pp 168-183

Solow, B.L., Solow, J.L. and Walker, T.B. (2011), *'Moving on up: The Rooney Rule and minority hiring in the NFL'*, Labour Economics, Vol.18, No.3, pp 332-337

Srna, S., Schrift, R.Y. and Zauberman, G. (2018), *'The Illusion of Multitasking and its Positive Effect on Performance'*, Psychological Science, Vol.29, No.12, pp 1942-1955

Suchomel, T.J. and Bailey, C.A. (2014), *'Monitoring and Managing Fatigue in Baseball Players'*, Strength & Conditioning Journal, Vol.36, No.6, pp 39-45

Tanji, R. (2021), *'Reference Dependence and Monetary Incentives: Evidence from Major League Baseball'*, Discussion Papers in In Economics and Business, No. 20-23

Tena, J. D. and Forrest, D. (2007), *'Within-Season Dismissal of Football Coaches: Statistical Analysis of Causes and Consequences'*, European Journal of Operational Research, Vol.181, No.1, pp 362-373

Terviö, M. (2009), *'Superstars and Mediocrities: Market Failure in the Discovery of Talent'*, The Review of Economic Studies, Vol.76, No.2, pp 829-850

Van Ours, J. C. and Van Tuijl, M. A. (2016), *'In-season Head Coach Dismissals and the Performance of Professional Football Teams'*, Economic Inquiry, Vol.54, No.1, pp 591-60

Waldman, M. (1984), *'Job Assignments, Signalling and Efficiency'*, The RAND Journal of Economics, Vol.15, No.2, pp 255-267

Wangrow, D.B., Schepker, D.J. and Barker III, V.L. (2018), *'Power, performance and expectations in the dismissal of NBA coaches: A survival analysis study'*, Sports Management Review, Vol.21, No.4, pp 333-346

Zhao, Q. and Percival, D. (2017), *'Entropy Balancing is Doubly Robust'*, Journal of Causal Inference, Vol.5, No.1