UNIVERSITY OF LIVERPOOL | Management School

Working Paper in Economics

 # 202110

Forced to Play Too Many Matches? A Deep-Learning Assessment of Crowded Schedule

Stefano Cabras
Marco Delogu
J.D. Tena

FORCED TO PLAY TOO MANY MATCHES? A DEEP-LEARNING ASSESSMENT OF CROWDED SCHEDULE.

Stefano Cabras[1], Marco Delogu[2], and J.D. Tena* [3]

[1]Universidad Carlos III, Spain

[2] Università di Sassari and CRENoS, Italy, DEM University of Luxembourg

[3] University of Liverpool, United Kingdom and Università di Sassari and CRENoS, Italy.

February 2021

Abstract

Do important upcoming or recent scheduled tasks affect the current productivity of working teams? How is the impact (if any) modified according to team size or by external conditions faced by workers? We study this issue using association football data where team performance is clearly defined and publicly-observed before and after completing different activities (football matches). UEFA Champions League (CL) games affect European domestic league matches in a quasi-random fashion. We estimate this effect using a deep learning model, a novel strategy in this context, that allows controlling for many interacting confounding factors without imposing an ad-hoc parametric specification. This approach is instrumental in estimating performance under 'what if' situations required in a causal analysis. We find that dispersion of attention and effort to different tournaments significantly worsens domestic performance before/after playing the CL match. However, the size of the impact is higher in the latter case. Our results also suggest that this distortion is higher for small teams and that, compared to home teams, away teams react more conservatively by increasing their probability of drawing. We discuss the implications of these results in the multitasking literature.

JEL CODES: TBC,

KEYWORDS: multitasking, causal analysis, deep learning, sports economics.

*We are grateful to David Forrest, Elias Carroni, Giuseppe Pulina, Giuliano Resce and Alex Krumer for helpful comments and suggestions. We also thank participants at the 11^th ESEA Conference on Sport Economics, seminars at the Universidad Carlos III de Madrid and University of Corsica-Pasquale Paoli for valuable discussions. The usual disclaimer applies.

1 Introduction

How crowded schedule affects working teams' performance is an essential question in management. Understanding this effect provides a way to increase productivity with little or no change in resources. Thanks to the increasing availability of databases, modern days have witnessed the emergence of papers estimating the productivity consequences of not focusing attention and effort on a single activity. Many authors have documented that workers' performance could be negatively affected by unplanned interruptions or contingent multitasking; the bulk of this evidence comes from randomized trial experiments (Adler and Benbunan-Fich (2012); Buser and Peter (2012)) and observational studies (Staats and Gino (2012); Cai et al. (2018); and Goes et al. (2018)). However, unlike unexpected interruptions, assessing the impact of crowded schedule is more challenging for at least two reasons. First, causal estimation using standard models is complicated as scheduled tasks can be largely anticipated and affect other management decisions. An econometric model should consider the complex relationship between treatment and other potential confounding variables to take this issue into account. A second and most important reason is the difficulty in observing workers' productivity in standard industries before and after completing scheduled tasks. Among the limited number of prior studies, Goulas and Megalokonomou (2020) and Pope and Fillmore (2015) analyse the impact of exam scheduling on students' academic grades.

Association football provides a fertile ground for this type of analysis. Two traditional arguments to support the use of football data in behavioural studies are that (unlike experiments) football involves high stakes decisions familiar to players and that performance and their determinants are clearly observed. An additional reason in our context is the fact that soccer provides crowded schedule examples. More specifically, many European football clubs are involved in more than one competition. The two most important ones are the domestic league and the UEFA Champions League (CL). They require some different (but related) abilities to survive and/or prosper. Thus, while the national leagues reward consistency in performance (league standings are determined by the total number of points in all matches through a season), the CL includes knock out stages where a small number of important matches decides the champion. Moreover, football data allow observing teams' performance before and after switching tournaments and under different external conditions defined by home and away matches.

This paper estimates the short-run effect of close CL matches on previous and subsequent team performance in the domestic league using match level information from seasons 1999/2000 to 2015/2016.[1] National league matches are affected by CL matches in a quasi-random fashion. However, in the estimation endogeneity may arise from unobservables potentially correlated with treatment variables. To

[1]Note that, in other contexts, the importance of scheduling in sports tournaments has already been explored. For example, Goller and Krumer (2020) and Krumer and Lechner (2018) analyse how playing in the middle of the week affects home advantage in football league matches, while Krumer et al. (2017) explore the positive psychological effect of playing first in short round-robin tournaments. However, this paper has a different focus as it analyses how team performance is affected by participating in distinct sports tournaments.

deal with these issues we employ a deep learning model (DL), see Schmidhuber (2015). Similar to other approaches, such as difference in difference or regression discontinuity, our causal estimation is based on projections under different hypothetical scenarios. However, a DL model does not impose the form in which predictors affect the response variable. Although DL models do not have a theoretical interpretation, variables are free to interact at different layers, and model predictions can estimate the impact of treatment performance under different hypothetical situations. Importantly, a DL approach offers a way to deal with unobservables. In our empirical estimation, it is challenging to envisage an omitted variable that is not an implicit function of any of the included model predictors. A final advantage of this approach is that it provides a whole distribution of the estimated causal impacts rather than just an average causal estimate. To preview, we find that a close CL match exerts a detrimental impact on domestic league performance. In particular, both a past and an upcoming CL match leads to performance deterioration in the domestic competition. However, the size of this effect is larger for past CL matches. We also find that external factors affect teams' strategies in this situation. In particular, away teams exhibit a higher risk averse reaction than home teams. Thus, while a CL match reduces the probability of winning in both cases, away (home) teams increase their probability of drawing (losing). Our estimation also indicates that performance deterioration is not homogeneous, with bigger clubs, defined as the ones with a better performance in the past, are less damaged by close CL matches. Overall, results suggest that dispersion of effort and attention to different tournaments can worsen performance. However, the effect of CL may depend on multiple factors such as the size of work teams and the external circumstances they face, taken into account by our estimation strategy.

The rest of the paper is structured as follows. The next section discusses the previous literature along with the hypothesis to test. Section 3 presents the data used in the analysis. Section 4 explains the use of DL models in causal analysis. Sections 5 and 6 present and discuss our results. Section 7 concludes.

2 Related literature and hypotheses

2.1 Related economic and management literature

Previous literature has traditionally explored the importance of specialization and task interruption on labour productivity. Regarding the former issue, Singh and Staats (2012) and Coviello et al. (2019) find that dealing with similar (compared to dissimilar) tasks has a significant effect on productivity in the medical and judicial environment, respectively. Staats and Gino (2012) study a Japanese bank's home loan application process line finding that productivity is worsened by switching tasks within the same day but enhanced by performing different tasks on different days.

The effect of either unplanned or scheduled interruptions was initially explored by means of experiments, Buser and Peter (2012); Abadi et al. (2015). In particular, Buser and Peter (2012) propose an experiment where participants have to perform two separate tasks: sudoku and world search puzzle. Participants were randomly allocated to two different treatments where they could either know or do not know the time allowed to perform each task in advance. The authors show that in both cases multitasking significantly lowers performance compared to sequential execution. Participants in the experiment proposed by Adler and Benbunan-Fich (2012) can discretionarily switch between *sudoku* and some tasks of shorter duration. They consider the possibility of a nonlinear effect of the number of switches on performance, finding an inverted-U relationship where medium multitaskers perform better than high- or non-multitaskers. More recent papers have estimated the impact of multitasking on individuals in high-stakes situations. Thus, Tan and Netessine (2014) estimate the impact of multitasking on performance measures in a restaurant chain. They find that workers' strategy depends on the amount of workload as they try to work more promptly and reduce their sales effort when their workload surpasses a certain threshold. Cai et al. (2018) find that unexpected interruptions due to machine breakdowns produce a decline in workers' productivity in a Chinese firm the following day. Goes et al. (2018) estimate the negative effect of multitasking on some outcome indicators in a SP 500 firm. Pope and Fillmore (2015) find that academic performance is positively associated with the time between exams. Goulas and Megalokonomou (2020) analyse the impact of STEM exam scheduling on students' scores. This is a complex relationship where fatigue and warming-up play an important role. In particular, students' performance deteriorates the lower the number of days between exams and the higher the number of days since the first exam (fatigue) but increases with the number of exams the student has taken (warm-up). They also provide evidence that the effect is not uniform across all students as higher-performing students show a higher warm-up and lower fatigue effects.

We contribute to this literature by estimating the separate impact of upcoming and previous scheduled tasks on workers' productivity. Our focus is not on individual students but on managerial decisions involving working teams in high stakes situations. However, like Goulas and Megalokonomou (2020), who identify a separate impact of exam scheduling on higher- and lower-performing students, we also study whether higher and lower performing teams are differently affected by CL matches. We also explore, for the first time, how external conditions, determined by playing at home or away, could affect estimation results.

2.2 Related sport literature

A traditional concern in the sport literature regards the importance of fatigue in influencing sport performance. In an early and highly-cited contribution, Pollard (1986) discusses the different factors that explain home advantage in English football. He finds similar levels of home advantage in matches where the away team had to travel within and beyond 200 miles. This leads him to conclude that the impact of travel fatigue is small compared with familiarity with home conditions. Subsequent papers studying this issue using different sports and methodologies have found little or no effect of fatigue on sport performance.

This literature can be generally broken down into two categories depending on whether they focus on the role of travel distance or days of rest. Regarding the first group, Oberhofer et al. (2010) and Nichols (2014) estimate the influence of travel distance on home advantage in the first tier of German football and in the National Football League (NFL), respectively. Nutting (2010) explores the same question for the National Basketball Association (NBA) whilst Carter (2017) estimates the performance advantage of the closer team to a neutral site in the National Collegiate Athletics Association (NCAA). A general result in these papers is that although travel distance contributes to a small increase in home advantage, this effect is not monotonic but decreasing. The impact of days of rest on performance has been investigated for example in Entine and Small (2008) and Scoppa (2015) for the NBA and the soccer World Cup and the European Football Championship, respectively. They find no effect of rest days on performance in either case. Moreover, Scoppa's (2015) results suggest that the effect of fatigue due to resting is now less important than it used to be before 1990, which could be explained by better conditions and preparation. Pina (2019) finds some evidence that longer rest times increase the probability that a surfer qualifies for the next round in the ASP World Tour. This effect is nonlinear and reaches the maximum at five days. Unlike these papers we consider two different competitions to explore not only the role of fatigue due to past matches but also upcoming matches. For teams and individual players, being involved in different competitions could imply not only fatigue but also the dispersion of attention and higher levels of psychological stress. In related research, Poli et al. (2015) provide comprehensive descriptive statistics on the short, medium and long-run effect of participating in the CL and the Europa League. Interestingly, the paper shows that participation in UEFA club competitions is uncorrelated with domestic league results and, in particular, similar domestic league results have been observed regardless of UEFA competition fixtures taking place within a five-day period. The present paper differs from Poli et al. (2015) in that our estimation has a causal interpretation because we control for different variables that can affect performance. Failing to control for other confounding factors does not only affect the precision of the causal estimation but, more importantly, can result in biased estimation as some of the confounding and treatment variables can be correlated. To mention two examples, this is the case if home and away treatment probabilities are correlated as most CL matches are played on similar days or if, due to strategic reasons, treated teams could perform similarly in some specific months regardless of the proximity of a CL match. Another related paper is Moffat (2020). He focuses on the impact of participating in the group stage of the UEFA Europa League on performance of teams in their national league. He proposes a regression discontinuity design, comparing the deviation of points per game with respect to the previous season of teams that participate and those that do not participate in the UEFA Europea League group stage. He finds that participation exerts a positively significant effect only for teams that do not belong to the five biggest leagues. Our research question and methodological approach are very different from Moffat (2020). While his analysis focuses on the long-term impact of knowledge and experience, this paper deals with the importance of scheduling on match performance. Our estimation approach is also different as we focus on match level data, allowing us to control for match and team-specific characteristics that cannot be observed at the season level.

2.3 Hypotheses and theoretical framework

A general result in the literature discussed in Sections 2.1 and 2.2 is that not focusing attention on a single task can cause a deterioration in performance because workers may need to acclimatise to the new task rather than to performance improvement. Krumer et al. (2017) and Krumer and Lechner (2018) propose an insightful framework to evaluate the determinants of sports contest outcomes. The basic intuition of their framework is characterised by modelling each game as an all-pay contest between two teams, i and j. We consider a simplified version of this model in which the prize valuation of winning the match is given by $v_k, k\epsilon\{i, j\}$. Both teams exert efforts x_i and x_j (with associated costs c_i and c_j respectively) to win the match and the team with the higher effort wins. If we assume $\frac{v_i}{c_i} > \frac{v_j}{c_j}$ and the mixed strategy of each team is denoted by $F_k(x), k\epsilon\{i, j\}$ a mixed strategy Nash equilibrium can be characterised by:

$$\frac{v_j}{c_j}F_i(x) - x = 0, \tag{1}$$

$$\frac{v_i}{c_i}F_j(x) = \frac{v_i}{c_i} - \frac{v_j}{c_j}(2)$$

which give team i and j's winning probabilities as:

$$p_i = 1 - \frac{c_i v_j}{c_j v_i} \text{ and } p_j = \frac{c_i v_j}{c_j v_i}. \tag{3}$$

The expected team efforts are such that each team invests a fraction $\frac{c_i v_j}{c_j v_i}$ of their valuation. According to this simple framework, the weaker team's strategy is to try to lower the effort of the stronger team by not competing with some probability. Other tournaments generate a dispersion of attention, thus reducing the team's likelihood of winning , $\frac{dp_k}{dc_k} < 0 \; k\epsilon\{i, j\}$. Notice that this framework predicts that an increase in the cost of effort by one team increases the rival team's incentives to compete for the match strongly.
A second discussion concerns the distinction of the impact of past and upcoming CL matches on domestic league performance. Playing before a CL match can distort the prize value of the domestic match, and therefore team productivity for at least two reasons. Head coaches could save some key players or even players cannot exert their maximum effort due to preparations for the next match. However, playing after a CL match could be especially disruptive as performance is also likely to be affected by injuries and physical and psychological fatigue. Therefore, based on this discussion, we offer the following two hypotheses:

- Hypothesis 1 (H1): Both a past and an approaching CL match negatively affect league performance.
- Hypothesis 2 (H2): Playing after a CL match creates a higher performance disruption than playing before it.

While the negative impact of multitasking on productivity is acknowledged in the literature, there is little research about how the impact depends on team size. This heterogeneity could come from two sources. Firstly, and most importantly, it is plausible to assume that cost of effort by smaller teams is relatively more affected as they have less resources to cope with many competitions. Secondly, there is a strategic effect. Thus, if a CL match increases the cost of effort of team i (the stronger team), from expression (3), the reduction in its probability of winning is lower the higher the strength of team i: v_i; while the opposite happens with an increase of effort of team j (the weaker team). However, it is more likely that in a national league match the stronger team is the one involved in the CL tournament. Accordingly, we hypothesise the following.

- Hypothesis 3 (H3): Small clubs will be more negatively affected by a recent or approaching CL match than big clubs.

We are not aware of previous papers that appraise the role of the external environment on multitasking. In very different contexts, the scientific literature generally suggests that working teams and individuals become more risk-averse under adverse situations. For example, Pauley et al. (2008) show that tolerance of risk by aviation pilots is affected by adverse weather conditions and previous exposure to these situations. Sung et al. (2019) study strategic decisions by firms listed on the New York Stock Exchange. They find that they become more conservative in downturn periods by putting more emphasis on value appropriation than value creation. In line with previous literature, our framework suggests that a more considerable increase in effort can make teams more risk-averse. Therefore, we hypothesize that teams are more risk-averse when playing under adverse environments (away matches) than when they play under favourable environments (home matches). The following hypothesis summarises this idea:

- Hypothesis 4 (H4): Dealing with two tournaments reduces home and away team performance differently. The home team will increase (decrease) the probability of losing (winning) the domestic league match. The away team will adopt a more conservative strategy by increasing the probability of drawing.

3 Data

3.1 Background

We collect data on leagues and CL matches from github at the url https://github.com/jalapic/engsoccerdata (see (Curley, 2016)). Match data are available for 11 European leagues (along with CL information).[2] The period availability differs across leagues and, to produce comparable estimation results, the dataset

[2] This classification includes the five biggest leagues: England (Premier League), France (League 1), Germany (Bundesliga), Italy (Serie A), and Spain (Primera Division). The six smaller leagues include (Belgium (First Division), Greece (Souper Ligka), Holland (Eredivisie), Portugal (Primeira Liga), Scotland (Premiership) and Turkey (Super Lig).

used here comprises league and CL matches from season 1995/21996 till season 2015-2016. The dataset contains data for 72,080 domestic league matches played by 445 different teams. National leagues follow a round-robin format in which each contestant meets all other contestants in turn. Awards depend on the final standings determined by computing the number of points; winning a match earns 3 points, drawing 1, and a loss 0. Qualification for CL depends on league standings. The number of slots given to each national league is contingent on the results obtained by national teams during the previous UEFA competitions (see, https://www.uefa.com/memberassociations/uefarankings/country//yr/2021). The maximum number of participating teams in a country is four (the English Premier League case), with the minimum equal to one (Scotland in Season 2002/2003).[3]

The format of the CL is different from national leagues. There is an initial group stage in which teams compete following a classical round-robin format. The best two teams per group advance to the knockout phase which generally starts in mid-February. Each round, up but not including the Final, is played in the form of two-legged ties. League matches are usually played at the weekend while CL matches are played midweek (usually on Tuesdays and Wednesdays).[4]

3.2 Variable definitions and descriptive statistics

Our response variable is the outcome of the league game, which can take values 1, X, 2, corresponding to home victory, draw, and away victory, respectively. To evaluate the effect of playing a CL match on the league match, we construct four treatment variables based on the number of days separating each league match and the previous and next CL match for the home and away teams. In particular, we define treatment based on a maximum separation of three-days. Although this is an ad-hoc decision, it can be justified to be a reasonable one. Thus, on the one hand, considering a time window of one day would not make any sense because no team has played a league match the day before or after playing a CL match. Likewise, a threshold equal to 2 generates only 22 matches treated out of the 72080 of our dataset. On the other hand, if we were to consider a larger threshold (think about five days), we would implicitly be assuming that most league matches are similarly affected by previous and upcoming CL matches while our intention is to identify cases in which league and CL matches are close in time and therefore clearly impacted by each other. This leaves us with only two choices, a three-day or a four-day threshold. This paper reports results with the former definition as the match affected is more clearly identified. Moreover, football head coaches are especially vocal against playing matches within three days.[5] However, results

[3]However, a country (such as England in Season 2005/2006, when Liverpool gained access thanks to its victory in the CL 2004-2005 but got only at the fifth place in the premier league) may qualify up to five clubs; this happens when the CL's former winner does not qualify among the best four clubs in his national league. When this happens, UEFA grants an additional slot to allow the title's defendant to participate in the next CL.

[4]A notable exception is the CL final, played since season 2009-2010 on Saturday.

[5]There are many examples of these complaints in the media. In November 2020, in the English Premier League, the head coach of Manchester City, Pep Guardiola, declared his team would prioritize the CL when they had to face Liverpool CP

with a four days threshold are qualitatively similar and available from the authors on request.
Based on the discussion above, the following treatment variables are defined:

1. *Homeprechampion*: this variable takes value one if the home team plays a CL match in one of the three days following a national league match. Otherwise, it takes a value equal to zero.

2. *Awayprechampion*: this variable takes value one if the away team plays a CL match in one of the three days following a national league match. Otherwise, it takes a value equal to zero.

3. *Homepostchampion*: this variable takes value one if the home team had played a CL match in one of the three days before the national league match. Otherwise, it takes a value equal to zero.

4. *Awaypostchampion*: this variable takes value one if the home team had played a CL match in one of the three days before the national league match. Otherwise, it takes a value equal to zero.

The data source contains valuable information for retrieving/constructing confounding variables to disentangle the treatments' role in national League performance. In particular, we include a categorical variable from the league match date equal to the league match month to control for the season's different phases. It is plausible to assume that team's effort correlates with the month, with the most important games played in the late spring when awards are assigned. Every season, clubs which participate in the CL are more likely to make particular investments to cope with the two competitions. We take this into account by including two confounding variables, Home CL and Away CL, that take values one when the home and away team participates in the CL that season and zero otherwise.

Taking advantage of match league data from season 1992/1993, for each pair team/season, we compute the ratio between the number of points earned over the maximum number of points available. Considering an unweighted average of the number of points won by a club in each of the prior seasons, the UEFA ratings have the weakness that treats each of the preceding seasons equally, see Flores et al. (2015). More recent seasons should be more relevant to predicting the outcome of a match. Thus, our index of past-performance at time t is a weighted average of the values of these ratios. By team we give $\frac{1}{2}$ weight to the relative result during the last season, a weight equal to $\frac{1}{4}$ for the ratio from two seasons earlier, and $\frac{1}{8}$ for the corresponding ratio referring to three seasons earlier. Older seasons receive weight equal to zero. We compute those variables for both home- away teams, and we label, accordingly, *Home strength* and *Away strength.*

To control for the role of inertia (there is evidence that football team results are serially correlated; Rebeggiani and Gross (2018)), we include among the confounding variables the result obtained in the last

just 3 days ahead of a CL match. The same month, Jurgen Klopp declared to the BBC "No team playing on Wednesday night should play again on Saturday at 12:30". In the Italian Serie A, Antonio Conte, Inter Milan's head coach, shows his discontent for a calendar that obliged his team to play domestic and CL matches in July 2020.

league match for both home and away teams denoted by *Home previous match* and *Away previous match* respectively.

We also consider the average number of points earned before the league match by the home and away teams. The variables labels are *Home points* and *Away points* respectively.

Expected team's performance may be affected by other official matches played within the three days interval. We build a set of dummy variables to account for such confounding effects. The variable *Homepostleague* takes value one if the home team plays another league match on one of the three days before the subject league match. Similarly, the dummy variable *Homepreleague* takes value one if the home team plays another league match one of the coming three days. Similar dummy variables, *Homepostleague* and *Awaypostleague*, are built to account for the same confounders after a recent league match.

Apart from CL, European football teams may compete in other international contests (Europa League, UEFA Cup, Winner's cup). We retrieve match league data for those competitions from Wikipedia, taking advantage of basic web scraping techniques. Using a matching process, similar to the one described for the treatments, we include among the confoundings dummy variables taking value one if the team plays another international match within the three-day interval, always considering the four cases (*HomepreEC*, *HomepostEC*, *AwaypreEC*, *AwaypostEC*).

Table 1: Descriptive Statistics

Variable	Obs	Mean	SD	Min	Median	Max
Homeprechampion	67802	0.014	0.116	0.000	0.000	1.000
Homepostchampion	67802	0.008	0.090	0.000	0.000	1.000
Awayprechampion	67802	0.014	0.116	0.000	0.000	1.000
Awaypostchampion	67802	0.007	0.085	0.000	0.000	1.000
Home strength	67802	0.345	0.176	0.000	0.365	0.767
Away strength	67802	0.345	0.176	0.000	0.365	0.767
Home points	67802	1.357	0.534	0.000	1.300	3.000
Away points	67802	1.380	0.534	0.000	1.333	3.000
Home previous match	67802	0.851	0.848	0.000	1.000	2.000
Away previous match	67802	1.149	0.850	0.000	1.000	2.000
partchampionshome	67802	0.141	0.348	0.000	0.000	1.000
partchampionsvisitor	67802	0.141	0.348	0.000	0.000	1.000
HomepreEC	67802	0.002	0.047	0.000	0.000	1.000
HomepostEC	67802	0.016	0.125	0.000	0.000	1.000
AwaypreEC	67802	0.003	0.052	0.000	0.000	1.000
AwaypostEC	67802	0.016	0.124	0.000	0.000	1.000
Homepreleague	67802	0.096	0.294	0.000	0.000	1.000
Homepostleague	67802	0.067	0.249	0.000	0.000	1.000
Awaypreleague	67802	0.095	0.294	0.000	0.000	1.000
Awaypostleague	67802	0.067	0.250	0.000	0.000	1.000

This table reports descriptive statistics for all treated and confounding variables considered in the analysis. Notice that there are 67802 observations because the first round of every season is missing due to the inclusion of Home previous match and Away previous match.

Table 1 reports descriptive statistics for both treatments and confounding variables. The percentage of treated games is relatively small; our data contains 2625 games with at least one treatment variable equal to 1, corresponding to 3.87% of all the matches employed in the empirical analysis. Notice that the number of treated games for the "previous" case more than double that of the post case. This is due to the fact that the CL is played Tuesday and Wednesday. The opposite happens for the other EU competitions; the Europa League is usually played on Thursday; consequently we have a larger percentage of matches for which the control variable takes value 1 in the "post" case.

It is interesting to notice that the percentage of games played in the three-day window is substantial. Recall that not every CL match generates a treated match; the threshold requirement must hold. However, most CL matches take place within the three-day interval. For teams playing more than two CL matches per year, the ratio between CL matches played within three days of league matches over the total of CL matches equals 70 percent.

4 The Deep Learning Model

4.1 Model specification and estimation

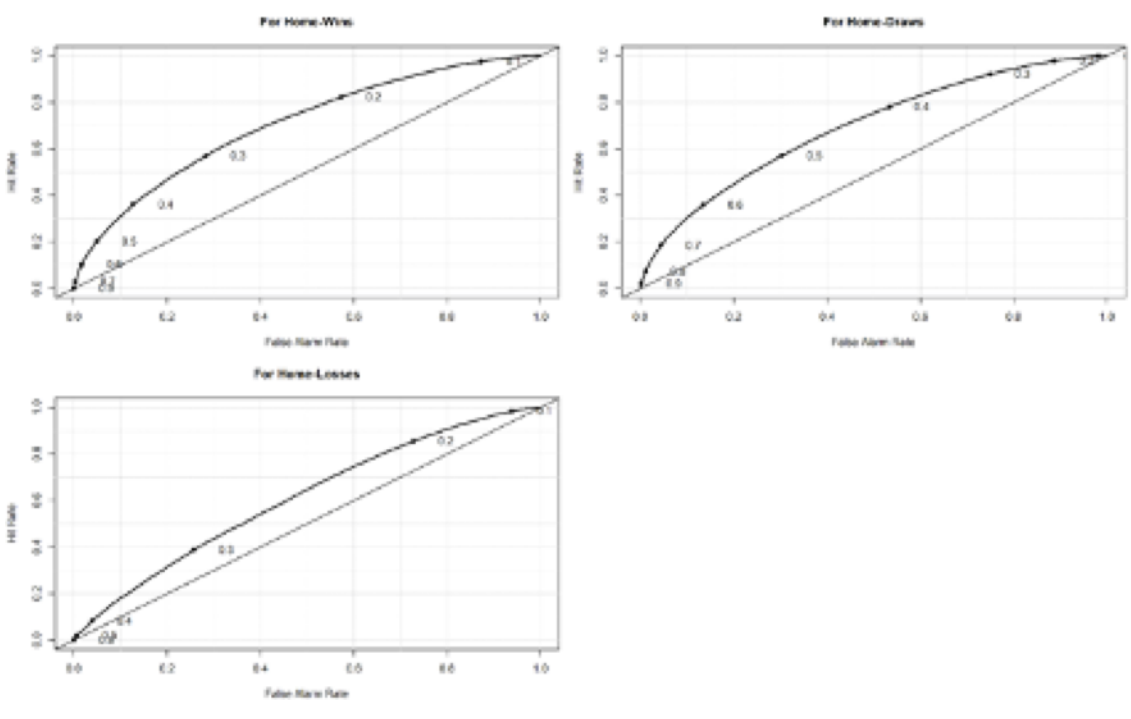
A deep learning (DL) model is a neural network with many layers of neurons, Schmidhuber (2015). Model specification follows an algorithmic rather than probabilistic approach, see Breiman et al. (2001) for the merits of both methodologies. In the model, each neuron is a deterministic function such that a neuron of a neuron is a function of a function along with an associated weight w. Essentially for a response variable Y_i, which corresponds to the outcome of the ith league match (home victory, draw and away victory) and a predictor X_i, model specification (without the bias term) becomes $Y_i = w_1 f_1 \left(w_2 f_2 \left(..w_k f_k \left(X_i\right)\right)\right)$, and a larger value of k implies a deeper network. Different stacked layers of neurons connected (i.e. dense layers) allow us to capture high non-linearities and complex interactions among variables in the model. Here, we estimate the model using a compositional function rather than an additive function (i.e. $Y_i = w_1 f_1 \left(X_i\right) + w_2 f_2 \left(X_i\right) + ... + w_k f_k \left(X_i\right)$) which is usually considered in usual regression techniques, even the most modern ones. Although less transparent, a compositional function improves the model's prediction accuracy, and thus, we eventually opted for this more complex model. A comprehensive review of DL is beyond the scope of this paper, but readers are referred to Schmidhuber (2015) for an overview.

The DL model can also be interpreted as a maximum a posteriori estimation of $Pr\left(Y|X, Data\right)$ for Gaussian process priors, Polson et al. (2017). However, because of its complexity, the whole distribution $Pr\left(Y|X, Data\right)$ cannot be evaluated but only its mode. This implies that the causal effects estimated here are those which have the maximum probability density given the observed data.

In this setting Y_i and X_i can be scalars or vectors. In our case, Y is the scalar random categorical variable of results, with three categories corresponding to home victory, draw and away victory respectively while, as we will explain, X is a vector of dimension 2369 which includes the predictors defined in the data section with categorical variables encoded into columns of 0s and 1s. We randomly split 20% of the data for validation. Taking advantage of the data we estimate vectors $w_1, \ldots, w_k$ for our DL specification. The estimation requires evaluation of a multidimensional gradient, which cannot be jointly evaluated for all observations, because of its dimensionality and complexity. Recall that the derivative of a composite function is the product of the derivatives of the inner functions (i.e. the chain rule $\left(f \circ g\right)' = \left(f' \circ g\right)' * g'$), evaluated through a tensor product for computational feasibility. Such a tensor product is evaluated for a batch of observations. We take advantage of the open source software known as Google Tensor Flow (Abadi et al. (2015) running on a NVidia GPU. There are different optimization algorithms to estimate the ws and we use the Adaptive Subgradient Methods (ADAGRAD) (Duchi et al. (2011)) in order to minimise the categorical cross entropy (remind that the response is categorical with three levels) loss function, i.e.

ws are estimated to minimize $L(y, \hat{y}) = -\sum_{i=0}^{n=67802} (y_i log(\hat{y}_i))$ which represents the differences between observed y_i and the prediction from the net $y_i = \hat{w}_1 f_1 (\hat{w}_2 f_2 (...\hat{w}_k f_k (X_i)))$. We have two dense layers (each with 30 nodes), separated by a normalization batch layer. The last layer is the output with the usual softmax activation function (that is the usual multinomial logistic function). We have around 2369 (i.e. weights) to be updated. Of course, most weights will be zero as they do not contribute to the gradient of the loss function and this avoids overfitting. Furthermore, in order to achieve stability in the estimation, we introduced a normalization batch between the two sets of hidden layers (Ioffe and Szegedy (2015)). Normalization batch is the usual operation of variable standardization (i.e. mean zero and variance one) applied to weights communicating two sets (layers) of all connected neurons. It is proved that this operation allows for better stability in the gradient of the whole function $Y|X$ estimated with the DL model. Figure 1 shows the ROC curve for the three possible match outcomes. In the three cases, the areas under the ROC are significantly larger than a random allocation according to Mann-Whitney U statistics and the usual Wilcox test.

Figure 1: ROC CURVES



ROC curves are based on prediction thresholds (interior numbers) for evaluation model fit with AUC.

4.2 The Deep Learning Model in causal analysis

There is an underlying prediction model for the potential outcome in every casual analysis based on the prediction output approach.This is the case, for instance, in standard approaches such as Difference in Difference and Regression Discontinuity. In both cases, homogeneity among groups being compared is achieved by a linear regression model. However, a linear regression fixes the relation between the response and the predictors beforehand (either intervention variables and/or confounding variables). Using the approach suggested here, we avoid fixing such relations beforehand.

Therefore, once the DL model is specified for the variables defined in Section 3, a causal effect estimation of a recent (previous or future) CL match on the league game outcome is not different from other more standard approaches. More specifically, if we denote the outcome of the ith match and the treatment whose effect we are interested to estimate by Y_i and D_i, the Average Treatment Effect for that treatment is defined as $ATE_i = E_{\pi(Y_i|X_i=x_i,Data)}\left(Y_i\left(D_i = d_0\right) - Y_i\left(D_i = d_c\right)|X_i = x_i\right)$, where $D_i = d_o$ is the observed treatment value and $D_i = d_c$ represent the counterfactual that it could have taken. For instance, $D_I = d_o$ represent the what if situation, e. g. what if on match i, the home team would not have to play a CL match within 3 days. Similarly, another ATE_i can be defined where the observed event is that in match i the home team will not play a CL match within 3 days while it would play it under the estimated counterfactual. This implies that two causal effects can be estimated for each treatment variable. One would correspond to an observed treated compared to a hypothetical untreated team and another to an observed untreated team compared to a hypothetical treated one.

Note that ATE is defined upon the expectation of the random variable Y_i conditional on all information on confounding and treatment variables defined in Section 3. Thus, the predictive model $Y|X, D$ (that for sake of simplicity, we refer to it as $Y|X$, where X includes D) is an approximation of the Bayesian predictive distribution (the one that appears on the sub index of the expectation).

This approach is an alternative to one that matches teams that, say, will play an CL match within 3 days with identical teams who are not subject to that treatment. The application of the propensity score methodology is only possible if there is a region of common support between the two groups of match-teams. Because of the large set of involved confounders, such a common region does not exist. Basically, the team description used to make prediction with DL is so refined that it is unfeasible to match sets of teams with exactly the same confounders. Moreover, in order to satisfy the strong ignorability assumption, required for reliable inference, we have to account for all possible collected confounding variables (i. e. elements of vector X_i) along with their interactions in the predictive model for Y_i.

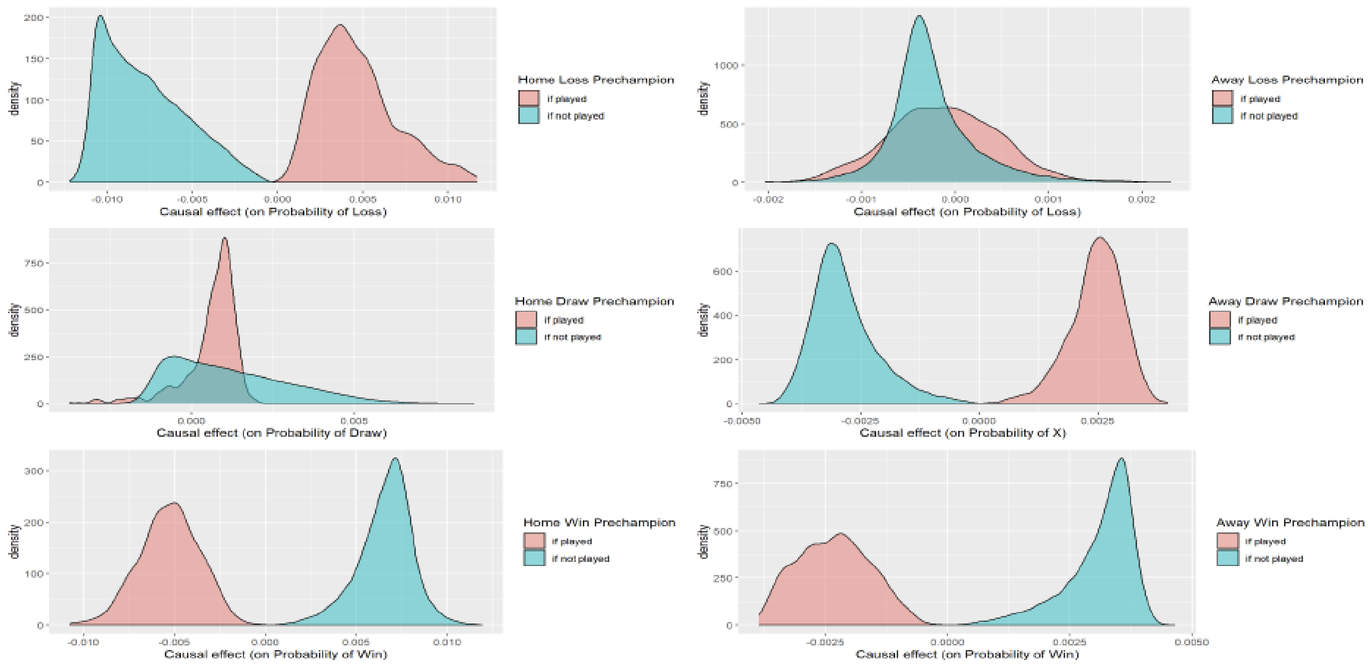
Finally, to draw sensible conclusions, it is necessary to have a reliable prediction model to capture all relations between the response and the predicting variables (Shmueli et al. (2010)). The deep-learning predictive model we are using belongs to the state of the art in model prediction capable of dealing with high dimensional data and large datasets. This model is subsequently used to construct counterfactual outcomes and, for this purpose, we need an adequate prediction model rather than an interpretable one, which would restrict the analyst to specify beforehand the relation between the response variable Y_i and the possible causes X_i (see Shmueli et al. (2010) for a discussion on this point). Taking advantage of such a large set of confounders prevents us from comparing our results with any propensity score model.

5 Empirical results

As explained in the previous section, we estimate the separate impact of four treatment variables on match performance distinguishing between two cases based on what would have happened (1) if a treated unit were untreated or (2) if an untreated unit were treated. For each match and each of these cases, the hypothetical counterfactual was estimated using the DL model described in Section 4. Therefore, the approach employed allows us to estimate the causal impacts' whole distribution rather than just reporting mean estimated values.

We start our analysis with the estimation of the causal impact of a forthcoming CL match on a league game. Figure 2 reports these estimation results for the three match outcomes (win, draw and loss) for the home and away teams. Results indicate that a forthcoming CL match significantly distorts both home and away performance but the effect size is relatively small. In particular, on average, the probability of a home win (loss) decreases (increases) by about 0.5pp-1pp and the probability of away draw (loss) increases (decreases) by around 0.3pp. These results are consistent with hypothesis H1. However, there are important differences in the estimated impact of an upcoming CL match for home and away teams that suggest a different strategic behaviour in these situations. Thus, while the probability of winning decreases in both cases, home and away teams offset this effect by increasing their losing and drawing probabilities. Our findings are consistent with hypothesis H4, and it could be explained by a more conservative approach by the away team probably motivated by performing under a more adverse external environment.

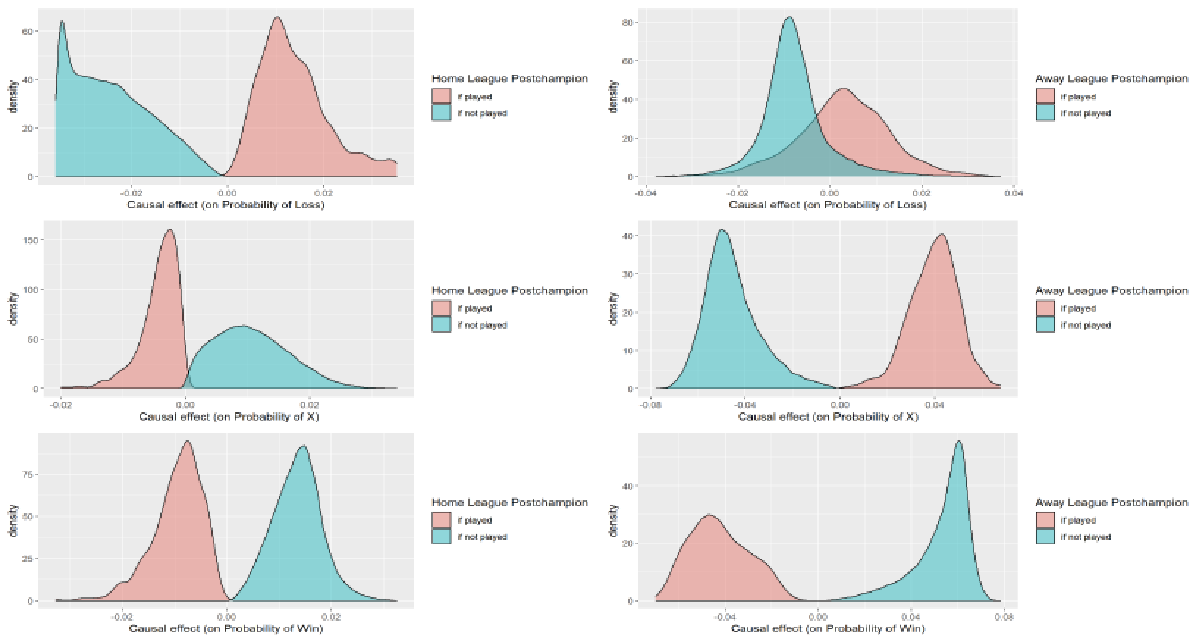
Figure 2: Causal impact of a forthcoming (within three days) CL match on a league match for the home and away teams



Each figure contains two estimates that compare observed responses (either 'played' or 'not played') with the one predicted with a DL model under a hypothetical counterfactual scenario.

Now we turn our attention to the impact of a recent CL match on team performance in the domestic league. Figure 3 reports estimation results. Interestingly, we observe a similar qualitative impact of a CL match on the home and away team's different match outcome probabilities. In particular, for the home team, not playing a previous CL match can decrease the probability of losing and increase the probability of winning by an average of 2pp and 1pp, respectively. For the away team, non-playing a CL match decreases the probability of drawing and increases the probability of winning by an average of 5pp in both cases. Therefore, a remarkable difference with the previous estimation is that a past CL match exerts a higher negative impact on league performance than a future one. This agrees with hypothesis H2 as it suggests that after a CL game, not only diversification of attention and human resources but also fatigue play a role in explaining performance deterioration.

Figure 3: Causal impact of a previous (within three days) CL match on a league match for the home and away teams



Each figure contains two estimates that compare observed responses (either 'played' or 'not played') with the one predicted with a DL model under a hypothetical counterfactual scenario.

An important advantage of our analysis is that it is possible to estimate causal effects conditional on any particular group of matches. Therefore, we explore the role of different confounding variables to explain differences in CL games' effect. In particular, estimated impacts were not significantly different across different months. Furthermore, there are no significant differences between the five big leagues (English Premier League, Spanish Primera Division, Italian Serie A, German Bundesliga, and French League 1) and the rest. A plausible explanation for this lack of significance is that, although the level of effort exerted in playing a football match could differ depending on the month of the year and the type of league, this affects the two competing teams similarly. The only variables that seem to play a role in explaining the heterogeneity of the estimated causal effects are *strength* and *points*. Tables 2 and 3 show the estimated average effect of not playing a CL match for teams with low, medium and high strength and points values. We employ the 0.25, 0.5 and 0.75 quantiles of the distribution of these two variables in the sample to do this. Results indicate that the positive effect of not playing CL matches are higher and more significant for teams with poorer performance in previous seasons and the current one. Such finding corroborates hypothesis H3 and it is in line with the idea that most powerful clubs typically have more quantity and quality of players in their squad to deal with different competitions.

Table 2: Causal impact of not playing the European Champion League conditional on strength. Mean values

	Home team					
	Before a CL match			After a CL match		
	Low strength	Medium strength	High strength	Low strength	Medium strength	High strength
Home loss Prob	-0.009 (0.002) **	-0.008 (0.002) **	-0.005 (0.003) *	-0.03 (0.006) **	-0.03 (0.007) **	-0.02 (0.01) *
Draw Prob	0.002 (0.002)	0.001 (0.002)	-0.0001 (0.001)	0.01 (0.005) **	0.01 (0.005) **	0.006 (0.005)
Home win prob	0.007 (0.001) **	0.007 (0.001) **	0.005 (0.003) *	0.01 (0.004) **	0.01 (0.004) **	0.01 (0.006) *
	Away team					
	Before a CL match			After a CL match		
	Low strength	Medium strength	High strength	Low strength	Medium strength	High strength
Away win Prob	-3.5e-04 (0.0003)	-3.2e-04 (0.0004)	-6.5e-05 (0.0006)	-0.008 (0.005) *	-0.009 (0.006)	-0.005 (0.01)
Draw Prob	-0.003 (0.0007) **	-0.003 (0.0007) **	-0.002 (0.001) *	-0.05 (0.01) **	-0.05 (0.01) **	-0.04 (0.02) **
Away loss prob	0.003 (0.0006) **	0.003 (0.0006) **	0.002 (0.001) *	0.06 (0.01) **	0.06 (0.01) **	0.04 (0.02) **

Low, medium and high are defined at the 0.25, 0.5 and 0.75 quantiles of the strength variable. ** (*) Significant at 5% (10%)

Table 3: Causal impact of not playing the European Champion League conditional on points. Mean values

	Home team					
	Before a CL match			After a CL match		
	Low strength	Medium strength	High strength	Low strength	Medium strength	High strength
Home loss Prob	-0.009 (0.002) **	-0.008 (0.002) **	-0.005 (0.003) *	-0.03 (0.006) **	-0.03 (0.007) **	-0.02 (0.01) *
Draw Prob	0.002 (0.002)	0.001 (0.002)	-0.0001 (0.001)	0.01 (0.005) **	0.01 (0.005) **	0.006 (0.005)
Home win prob	0.007 (0.001) **	0.007 (0.001) **	0.005 (0.003) *	0.01 (0.004) **	0.01 (0.004) **	0.01 (0.006) *
	Away team					
	Before a CL match			After a CL match		
	Low strength	Medium strength	High strength	Low strength	Medium strength	High strength
Away win Prob	-3.5e-04 (0.0003)	-3.2e-04 (0.0004)	-6.5e-05 (0.0006)	-0.008 (0.005) *	-0.009 (0.006)	-0.005 (0.01)
Draw Prob	-0.003 (0.0007) **	-0.003 (0.0007) **	-0.002 (0.001) *	-0.05 (0.01) **	-0.05 (0.01) **	-0.04 (0.02) **
Away loss prob	0.003 (0.0006) **	0.003 (0.0006) **	0.002 (0.001) *	0.06 (0.01) **	0.06 (0.01) **	0.04 (0.02) **

Low, medium and high are defined at the 0.25, 0.5 and 0.75 quantiles of the strength variable. ** (*) Significant at 5% (10%)

6 Discussion

We estimate CL matches' impact on league performance finding a negative and significant effect, albeit of small size (Figures 2 and 3). This result is consistent with most papers discussed in Section 2. However, the present paper focuses on situations where the different tasks schedule was previously set. In this sense, our results are consistent with Goulas and Megalokonomou (2020), as they show that students who have more time between exams perform better. Our data allows us to disentangle the productivity of teams before and after they deal with another task. According to H1, dispersion of attention due to an upcoming task (match) negatively affects current performance. However, we also find that, in line with H2, a recent match affects teams' productivity more negatively than an upcoming match. The reason is that a recent past match could be additionally affected by fatigue or injures.

We also explore the importance of playing at home or away and the team size on explaining the effect of participating in two competitions. Regarding the former, according to H3, we find that away teams behave in a more conservative way when they face a CL match by increasing the probability of drawing. This is a novel result that we cannot compare to other papers in the multitasking literature to the best of our knowledge. However, papers in other fields already find that firms become more conservative in downturn periods ((Sung et al., 2019)). When we turn our attention to the heterogeneous impact of playing close matches across teams finding that consistently with H4 bigger teams, those that perform better in the past, are relatively less affected. This result is consistent with Goulas and Megalokonomou (2020) who, in a very different context, found that higher-performing students can deal better with the fatigue generated by multiple exams.

The reported estimation can be of interest to different scientific disciplines. More specifically, the dissimilar reaction of home and away teams could help behavioural economists to understand how favourable and adverse environments can influence working teams' performance. Furthermore, the higher impact of a past CL match on performance compared to a future one may attract the attention of management researchers looking to understand the effect of dealing with different activities at different completion stages. However, the interest of these results for decision-makers in the football industry may still remain unclear, at least at first glance. Thus, an expected reduction of 0.1 points for an away match (5pp increase in the probability of drawing together with a similar reduction in the probability of winning) can be considered a tiny amount to capture managers' attention. However, our findings become relevant once one realizes that a tiny margin of points determines domestic league awards quite often (such as winning a title or qualifying for the next edition of the CL). Therefore, even a small distortion on domestic performance could have dramatic consequences for the season's outcome. Besides, according to Figures 2 and 3, the overall impact of an CL match could diverge across a wide range of values. In particular, Tables 2 and 3 show that small teams could have more problems dealing with more than one competition than big teams. Results in the present paper would also be particularly relevant to understand the potential effect of the future new format of the CL with more matches to play.

A limitation of this paper is that results are specific to the football industry. However, football is particularly useful to analyse, for example, the importance of past and upcoming tasks, and the external environment in high-stakes multitasking situations. A second limitation concerns our causal estimation which, as is common in this type of analysis, could be affected by omitted variable bias. However, using DL approach lessens this concern; it is difficult to think of an omitted variable that is not a function of the 2369 interactions of variables that our estimation includes. To mention two examples, although there is no variable capturing match importance, the interaction of month with the number of home- and away-team points and month is a good proxy to control this unobservable. We do not have a variable to deal with traditional rivalries in some matches. However, we are implicitly controlling for this fact given that the DL model considers each interaction of curent and historical performance for each pair of teams.

7 Concluding remarks

While there is enough consensus that workers' productivity is negatively affected by unexpected task interruptions or by the arrival of contingent tasks, there is less research about how different scheduled tasks influence working teams at different stages of the production process and the way they could be affected by external factors. This paper studies this issue in the context of soccer by estimating the impact of close CL matches on team performance using match-level data from 11 European leagues. This paper relies on a deep learning (DL) model to produce casual estimates. In this context, DL is a novel approach that reveals itself particularly useful for predicting team performance in hypothetical situations necessary to produce casual estimates.

Our results confirm that both previous and future CL matches negatively affect team performance in domestic leagues. However, probably due to fatigue, this deterioration is higher after a CL match. We also find that this estimation depends on past team performance and whether teams play at home or away. More specifically, clubs with poorer recent performance are more likely to be adversely influenced by CL matches. Also, away (home) teams react to a decrease in the probability of winning induced by a CL match by increasing their probability of drawing (losing).

This paper provides first evidence that: (1) crowded schedule could affect current performance (2) productivity in this situation could be affected by the external envirnoment. Our findings may guide managers' strategy when teams have to cope with different activities.

Future research should study these issues in other settings. There is a lack of research on the potential impacts of shorter scheduled tasks as well as their long-run effect at different stages of the production process. Moreover, studying the influence of covid-19 on multitasking productivity looks like an exciting research avenue. Again, sports data seem an optimal candidate to conduct the proposed research.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., et al. (2015). Tensorflow: Large-scale machine learning on heterogeneous systems.

Adler, R. F. and Benbunan-Fich, R. (2012). Juggling on a high wire: Multitasking effects on performance. *International Journal of Human-Computer Studies*, 70(2):156–168.

Breiman, L. et al. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, 16(3):199–231.

Buser, T. and Peter, N. (2012). Multitasking. *Experimental Economics*, 15(4):641–655.

Cai, X., Gong, J., Lu, Y., and Zhong, S. (2018). Recover overnight? work interruption and worker productivity. *Management Science*, 64(8):3489–3500.

Carter, K. E. (2017). Relative home-court advantage: The impact of travel on team production when one team is closer than its opponent to a neutral game site. *Managerial and Decision Economics*, 38(1):76–91.

Coviello, D., Ichino, A., and Persico, N. (2019). Measuring the gains from labor specialization. *The Journal of Law and Economics*, 62(3):403–426.

Curley, J. (2016). engsoccerdata: English soccer data 1871-2016. r package version 0.1.5.

Duchi, J., Hazan, E., and Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of machine learning research*, 12(7).

Entine, O. and Small, D. S. (2008). The role of rest in the nba home-court advantage. *Journal of Quantitative Analysis in Sports*, 4(2).

Flores, R., Forrest, D., De Pablo, C., and Tena, J. (2015). What is a good result in the first leg of a two-legged football match? *European Journal of Operational Research*, 247(2):641–647.

Goes, P. B., Ilk, N., Lin, M., and Zhao, J. L. (2018). When more is less: Field evidence on unintended consequences of multitasking. *Management Science*, 64(7):3033–3054.

Goller, D. and Krumer, A. (2020). Let's meet as usual: Do games played on non-frequent days differ? evidence from top european soccer leagues. *European journal of operational research*, 286(2):740–754.

Goulas, S. and Megalokonomou, R. (2020). Marathon, hurdling, or sprint? the effects of exam scheduling on academic performance. *The BE Journal of Economic Analysis & Policy*, 20(2).

Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR.

Krumer, A. and Lechner, M. (2018). Midweek Effect On Soccer Performance: Evidence From The German Bundesliga. *Economic Inquiry*, 56(1):193–207.

Krumer, A., Megidish, R., and Sela, A. (2017). First-mover advantage in round-robin tournaments. *Social Choice and Welfare*, 48(3):633–658.

Moffat, J. (2020). The impact of participation in pan-european competition on domestic performance in association football. *European Sport Management Quarterly*, 20(4):440–457.

Nichols, M. W. (2014). The impact of visiting team travel on game outcome and biases in nfl betting markets. *Journal of Sports Economics*, 15(1):78–96.

Nutting, A. W. (2010). Travel costs in the nba production function. *Journal of Sports Economics*, 11(5):533–548.

Oberhofer, H., Philippovich, T., and Winner, H. (2010). Distance matters in away games: Evidence from the german football league. *Journal of Economic Psychology*, 31(2):200–211.

Pauley, K., O'Hare, D., and Wiggins, M. (2008). Risk tolerance and pilot involvement in hazardous events and flight into adverse weather. *Journal of safety research*, 39(4):403–411.

Pina, G. (2019). Task scheduling and performance: Evidence from professional surf tournaments. *Journal of Economic Psychology*, 75(PB).

Poli, R., Benson, R., and Ravenel, L. (2015). Impact of uefa club competitions on domestic league results. *Football Observatory, International Centre for Sports Studies, University of Neuchatel.*

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

Polson, N. G., Sokolov, V., et al. (2017). Deep learning: A bayesian perspective. *Bayesian Analysis*, 12(4):1275–1304.

Pope, D. and Fillmore, I. (2015). The impact of time between cognitive tasks on performance: Evidence from advanced placement exams. *Economics of Education Review*, 48(C):30–40.

Rebeggiani, L. and Gross, J. (2018). Chance or ability? the efficiency of the football betting market revisited.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61:85–117.

Scoppa, V. (2015). Fatigue and team performance in soccer: Evidence from the fifa world cup and the uefa european championship. *Journal of Sports Economics*, 16(5):482–507.

Shmueli, G. et al. (2010). To explain or to predict? *Statistical science*, 25(3):289–310.

Singh, K. C. D. and Staats, B. R. (2012). Accumulating a portfolio of experience: The effect of focal and related experience on surgeon performance. *Manufacturing and Service Operations Management*, 14.

Staats, B. R. and Gino, F. (2012). Specialization and variety in repetitive tasks: Evidence from a japanese bank. *Management science*, 58(6):1141–1159.

Sung, J. K., Park, J., and Yoo, S. (2019). Exploring the impact of strategic emphasis on advertising versus r&d during stock market downturns and upturns. *Journal of Business Research*, 94:56–64.

Tan, T. F. and Netessine, S. (2014). When does the devil make work? an empirical study of the impact of workload on worker productivity. *Management Science*, 60(6):1574–1593.