Workforce Substitutability: Evidence from Professional Football*

Santiago Velasquez†

This version: November, 2024

Abstract

Quantitative analysis of labor production functions requires high-quality and longitudinal data on labor productivity, which is unlikely to be available. I study labor substitutability using the information on the top 5 most competitive men's football leagues in Europe. By considering individual player productivity to be the in-game performance statistics and focusing only on unpredictable injuries, I use event study approaches to causally estimate the effect of the negative exogenous shock on the injured players and their team's productivity. I further extend the analysis to understand what type of players serve as substitutes or complements. The findings suggest that injuries have long and medium-term effects on the injured player's performance. The extent to which a player's performance affects their teammates is determined by both the player's and his team's performance—specifically, high-performing players playing in an underperforming team function as complements by substantially positively influencing their teammates.

*This paper was initially written in 2020 as part of the requirements for completing my undergraduate degree in Economics at Universidad EAFIT. The thesis received an honorable mention in recognition of its effort and contribution. The original version, written in Spanish, is available here. I am deeply grateful to Arlen Guarín for his invaluable guidance and support.

†Development Innovation Lab and Universidad EAFIT. svelasquezbonilla@gmail.com

Most updated full version here.

1 Introduction

Football is the most played sport worldwide (Dvorak et al., 2004). In 2018, the five main European leagues had revenue between 1,700 and 5,400 million euros, thus representing around 75% of the total income of all the leagues of the 55 countries belonging to the Union of European Football Associations (UEFA) (Efe, 2020). As in most highly competitive sports, physical problems and injuries play a crucial role in the performance of athletes. The impact of injuries on player performance has been extensively studied in the field of sports medicine (Stubbe et al., 2015; Zuke et al., 2018) highlighting significant effects on metrics such as minutes played, various performance indicators, and return-to-competition measures, even though the latter is often broadly defined (Zuke et al., 2018). However, there is still a gap in empirical research examining how such injuries affect team productivity from an economic perspective.

This paper studies how the exogenous shocks in labor productivity affect overall firm productivity in the context of football. Football provides an ideal setting for testing hypotheses related to classical economic theories, such as the producer's theory. Football offers accurate performance metrics, unlike other areas (such as industry) where production and the contribution of factors are challenging to measure, differentiate, or quantify. These metrics can represent the firm's production, reflected in team results, and the contribution of productive factors, particularly labor, assessed through player performance in each match. Additionally, measuring team performance and player productivity over short intervals allows me to create panel-type datasets, which makes it suitable for applying econometric techniques to estimate causal effects.

I aim to quantify the contribution of labor to the production function of firms, using data from football games for the five main European Leagues: Bundesliga (Germany), La Liga (Spain), Ligue One (France), Premier League (England) and Serie A (Italy), with game-by-game information for each player, from the 2014-2015 season to 2018–2019. Taking advantage of the set of unexpected events defined by the players' injuries during competition, it is possible to identify how these injuries affect a player's contribution in the upcoming matches and how the absence of this production factor later affects the performance and production of the firm.

The players' performance decreases after an injury. Remarkably, the average game time is reduced by 42 minutes right after the injury, equivalent to a relative reduction of 85% of the minutes played. Also, the number of goals and expected goals decreases by approximately 90% in the next game. These effects are persistent for at least six more games; during this period, there is also an average decrease of about 50% in expected assists, number of shots and key passes, and new metrics such as *XGBuildUp* which measures the player's involvement in all passing sequences that lead to a shot or goal. The team's overall performance remains unaffected. My analysis finds no statistically significant impact of unexpected injuries on team performance. This may suggest that teams, on average, successfully compensate for the contributions of injured players, maintaining their overall performance level.

I exploit the heterogeneity of teams and players' differences in prior performance to unravel whether there is a subgroup of teams affected by injuries. I classify players and teams as having low and high productivity considering the distribution of the variable *XGChain*, cumulative until the date of the injury. Estimates indicate that a player's impact varies depending on the quality of the team they belong to and the skill levels of their teammates. Injuries to high-productivity players on low-productivity teams lead to a 6% to 8% decline in team performance across metrics such as key passes, *XGChain*, expected non-penalty goals (*NPXG*), shots, expected assists, expected goals, and *XGBuildUp*. Among the four groups analyzed, high-productivity players on low-productivity teams are the only group significantly affected at the 5% level.

I additionally estimated the effect of the injuries on the team-level variables, discounting the individual contribution of each player. While only the coefficients for key passes and shots are statistically significant (*XGChain* is marginally significant), showing team performance drops of 3.77% and 2.9%, respectively. The patterns across the four groups are noteworthy. For low-productivity players, the effects are minimal and close to zero. In contrast, the estimates suggest a positive effect across all variables for high-productivity players on high-performing teams. This indicates that injuries to these players are offset by contributions from highly productive substitutes or an overall increase in the efforts of remaining team members, effectively compensating for the loss. On the other hand, the impact on highly productive players in low-productivity teams is negative and significant in some cases. I finally computed cross-elasticities, suggesting that highly productive players on low-productivity teams are complementary to their teammates. In contrast, for highly productive players in highly productive teams, players are easily substituted.

This study constitutes a contribution of general interest, particularly for applied microeconomics. Particularly, my work contributes to two broader literatures: First, this study contributes to the literature on sports economics. Although this field has been studied for decades, many questions remain open (Rottenberg, 1956; Neale, 1964; Sloane, 1971; Scully, 1995; Rosen and Sanderson, 2001). Previous studies exploiting sports data have focused on labor demand, factor substitution, and wages (Caruso et al., 2017; Rosen and Sanderson, 2001). Some of the most recent works have focused on testing hypotheses related to other areas of economics, such as behavioral economics Gauriot and Page (2019);

Bryson et al. (2020). Conversely, this paper focuses on identifying the contributions of productive factors to production using exogenous events that allow the effect to be causally quantified.

The paper's second contribution lies in studying labor productivity within team settings. For instance, Bonhomme (2020) have proposed a methodology to estimate individual contributions in contexts where only aggregate outcomes are observed. Similarly, Arcidiacono et al. (2017) has previously examined the role of heterogeneous workers in influencing overall team performance, and Arcidiacono et al. (2017) and Devereux (2018) have assessed individual contributions to team outcomes. I build upon those studies and incorporate innovations such as leveraging unexpected exogenous shocks to establish causal effects and relaxing assumptions of serial independence and network homogeneity.

The rest of the paper is organized as follows: Section 2 describes the data used. Section 3 discusses the methodology, including the empirical strategy and the data-driven algorithm to identify unpredictable injuries. Section 4 discusses the results, specifically the effects of an unexpected injury on individual and team performance, as well as the heterogeneous effects and the estimation of players' substitutability or complementarity. Section 5 describes strategies to enhance the analysis in the future. Section 6 concludes.

2 Data

The data available contains match-by-match information for each player of all the teams in the five main football leagues in Europe—Premier League (England), La Liga (Spain), Serie A (Italy), Bundesliga (Germany), and Ligue One (France)—with an analysis period of five seasons, from the 2014-2015 season to the 2018-2019 season. I extracted the data automatically from the internet using web scraping methods.

Overall, there is information for 136 teams, 6,145 players, and 9,130 games, as well as multiple productivity variables at different aggregation levels, such as performance throughout the season or during a specific game, both for the team and each player. Table 1 contains descriptive statistics for the main variables of interest; columns 1 and 2 present the information about player and team performance for the entire sample, while columns 2 and 4 restrict the data to the matches to those in which the players participated. Information for all the players, regardless of whether they participated in the matches or were substitutes or not called up, is crucial to address the empirical question since substitutes constitute the team's options for replacing usual players if they are affected by injury.

I consider individual player productivity to be the in-game performance statistics. As shown in table 1, there are traditional performance metrics such as goals, assists, shots, and key passes. However, one potential limitation of these traditional metrics is that they are less variable than traditional performance metrics in other sports (Gaviria, 2000; Matano et al., 2018). For example, the number of goals scored has less variability than the number of points in a basketball game. I include modern metrics to overcome the variability issue, which allow for more precise measurement of a player's performance on the field—for example, the probability of scoring or making important plays during a game. Modern predictive models allow the computation of "Expected Goals" and "Expected Assists" that measure the probability of success of a play based on other key variables such as the distance of the shot, angle, type of play (free kick, individual play, corner kick), and part of the body used for the shot (leg, head, among others). Importantly, these measures are not conditional on the event's final result, as described in Bundesliga (2018).

Additionally, the data includes information about players' status in each match, such as their position on the field and various situational factors, including whether the player was

Table 1. Sample Characteristics

	Individual Performance		**Team Performance**	
	All Players	**With Game Time**	**All Players**	**With Game Time**
	(1)	(2)	(3)	(4)
Assists	0.040	0.083	0.923	0.948
	(0.213)	(0.299)	(0.674)	(0.687)
Expected Assists	0.036	0.074	0.059	0.115
	(0.124)	(0.170)	(1.877)	(1.890)
Shots	0.463	0.949	11.980	12.167
	(1.044)	(1.332)	(5.184)	(5.217)
Goals	0.051	0.1037	1.399	1.434
	(0.249)	(0.349)	(1.276)	(1.297)
Expected Goals	0.051	0.104	1.340	1.320
	(0.173)	(0.236)	(1.247)	(1.239)
Non-Penalty Goals	0.046	0.094	1.299	1.332
	(0.235)	(0.330)	(0.880)	(0.896)
Minutes Played	35.283	72.319	1.190	1.221
	(41.165)	(28.194)	(0.813)	(0.828)
NPGX	0.046	0.094	8.964	9.113
	(0.155)	(0.211)	(4.263)	(4.301)
Key Passes	0.350	0.718	0.364	0.353
	(0.842)	(1.091)	(0.481)	(0.478)
Yellow Cards	0.073	0.150	0.249	0.246
	(0.261)	(0.357)	(0.432)	(0.431)
Red Cards	0.002	0.004	0.388	0.401
	(0.042)	(0.060)	(0.487)	(0.490)
XGBuildUp	0.079	0.163	2.040	2.114
	(0.202)	(0.264)	(2.149)	(2.214)
XGChain	0.134	0.276	3.453	3.570
	(0.295)	(0.374)	(3.057)	(3.145)
Number of Observations	**668,412**		**335,864**	

Note: This table shows averages of key variables for players and teams. Standard deviations are in parenthesis. Columns 2 and 4 restrict the sample to players with playing time, defined as having more than zero minutes. See Appendix B for a detailed description of each variable.

injured. It also specifies the type of injury for which a player was absent. I observed various injuries in the data due to the heterogeneity among players and the long analysis period. I grouped injuries that occurred fewer than 80 times into similar categories for this analysis. Table 2 provides an overview of the types of injuries found in the data and their frequency. The main estimates include only injuries not preceded by another injury within the last eight months to avoid double-counting data due to correlated or simultaneous injuries. Column 2 in table 2 shows the frequencies for both total injuries and the number of injuries once the 8-month restriction is applied.[1]

Table 2. Description of the Injuries

Description	All the injuries (1)	Restricted injuries (2)
Hamstring injury	875	263
General injury/problems	860	262
Muscle injury	607	45
Muscle problem	544	135
Ligament injury	474	155
Knee injury	467	47
Unknown injury	448	165
Thigh problem	435	109
Muscular problem	400	124
Ankle injury	380	131
Adductor strain	362	103
Allergies	330	104
Bone injury (no fracture)	320	110
Fracture	295	103
Other	258	73
Knee problems	256	66
Muscle tear	235	59
Crash (in game)	225	79
Calf injury	214	192
Contusion	198	55
Tendon injury	170	50
Influenza	147	28
Cruciate ligament injury	141	58
Back problem	138	37
Shoulder injury	132	143
Calf problem	130	23
Abdominal/Intestinal problem	127	34
Surgery recovery	120	48
Foot injury	118	43
Achilles tendon injury	116	34
Fatigue	105	27
Wounds	87	36
Total	**9,714**	**2,941**

Note: This table shows the description and frequency of the injuries used as events in my identification strategy. To avoid double-counting events resulting from correlated or simultaneous injuries, I excluded injuries that another injury had occurred within the past eight months.

[1] The main results are robust to the inclusion of all injuries.

3 Methodology

Given the availability of data and the different performance variables mentioned above, it is possible to use data from each match for the players to correctly contrast the type of theories related to performance in team contexts and the contribution of labor to the total team output. By making use of exogenous shocks that happen to individuals in different periods, it is possible to adjust the available data to perform an event study type estimation, which allows computing the effect of said shocks on the variables of interest (Freyaldenhoven et al., 2019; Roth, 2022).

In our case, exogenous shocks are defined by the set of injuries that occur to players during the season; I aim to identify the effect of an unexpected injury on performance measures both at the individual level and the aggregated performance of the team to which the player affected by the injury belongs.

I am interested in assessing the dynamics underlying these treatment effects, which is often done by estimating an equation such as the following:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{\tau} \mu_i 1[t - E_i = \tau] + v_{i,t} \tag{1}$$

"Where $Y_{i,t}$ is a variable of interest for a unit i in period t, E_T is the period where unit i was initially treated and α_i and λ_t are the fixed effects of individual and time... The relative time $\tau = t - E_i$ included in (1) covers most of the relative periods but may exclude others" (Sun and Abraham, 2021). This methodology has the following three key identification assumptions:

- Assumption 1: Parallel trends in the variables for the periods prior to the shock.
- Assumption 2: There is no anticipatory behavior prior to treatment.

- Assumption 3: Homogeneity in the treatment effect.[2]

The first two assumptions are similar to the traditional assumption of parallel trends of the difference-in-differences model, in which it is assumed that in the absence of the event, the differences between the trends of the treated group and the control group would have remained constant. In contrast, assumption three refers to each treated cohort experiencing the effects of the treatment in the same way.

Considering the above, this methodology presents appropriate characteristics for its use in empirical studies [3]; therefore, in my study, I estimated an equation of the following form :

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{k=-6}^{6} \beta_k D_{i,t}^k + \phi_{t<-6} + \phi_{t>6} + v_{i,t} \tag{2}$$

Where α_i y λ_t are the fixed effects of individual and time, and the coefficients of ϕ are associated with the indicator variables of the periods six games before and six games after the period where the event occurs, defined as the match where the injury occurs or the team's next game if the injury occurs outside of a game. To comply with the assumptions above, injuries that were not preceded by another injury in the previous eight months were considered. Also, given the possibility that a player could experience multiple injuries that meet the above restriction, estimates were made using player-injury level data and including individual fixed effects, as mentioned in the equation, and tournament week and season fixed effects identify significant temporal variation in this context.

Given the heterogeneity of injuries that individuals may face in the sporting context, it is challenging to empirically argue that all of these comply with the assumptions above

[2]Recent studies, including Sun and Abraham (2021); Roth (2022); Freyaldenhoven et al. (2019), have introduced methodologies to address violations of these assumptions and improve identification. The validity of this assumption will be discussed later in the text.

[3]See Jacobson et al. (1993)

(mainly assumption 2). Since some of these are dynamically affected by realizations of the previous periods, i.e., an increase in the number of minutes played, or a muscular overload can affect the players' performance long before the period where treatment is evidenced in the data. A data-based algorithm was designed to deal with this problem in the cleanest way possible. The algorithm identifies, by extracting a random sample, the injuries that are most likely to be predicted in the data, thus excluding them from the analysis.

In other words, for each type of existing injury, an event study is estimated in a random sample of 20 percent of individuals with this injury (training sample). A hypothesis test is performed on the probability that the associated coefficients to the periods before treatment are jointly different from zero, and the p-value associated with the test is saved. Finally, for the remaining 80 percent, the equation is estimated again without considering the types of injuries with a p-value less than 0.2 in the exercises of the training sample.

This way, I ensure that the estimates of my model are not affected since I am left only with the types of injuries that are not very predictable according to the data. In addition, due to the use of random sampling and the design of the algorithm, which is wholly data-driven, I make sure that I deliberately select the sample to a minimum, thus avoiding bias, given the orthogonality of the randomly chosen data.

I additionally estimated heterogeneous effects by making use of the possibility of creating, through the data, categories of higher/lower productivity both at the team level and at the individual level, seeing where inside the distribution was located just before the period of injury and partitioning by the 75th percentile. With this, I can empirically evidence in which class of players, teams, or player-team combination the effects are mainly concentrated by estimating an equation like the following:

$$Y_{i,t} = \alpha_i + \lambda_t + X'_{i,t}\phi + \sum_{\pi}[\beta_{pre} D^{pre}_{i,t} I_\pi + \beta_{post} D^{post}_{i,t} I_\pi] + \phi_{t<-6} + \phi_{t>6} + v_{i,t} \tag{3}$$

Where all the variables are defined in the same way as in 2 and π represents the number of groups among which the effect is to be seen, and I_π is an indicator variable that takes the value of one of the individuals i belongs to group π or zero in another case. Unlike the event study regression, where the dynamics of the coefficient are studied, on this occasion, I group the coefficients of the periods before and after the shock to ensure that I have enough power in the data to estimate the effect of each group by not having to assess for many parameters.

One caveat regarding the event studies, when estimated using the two-way fixed-effects (TWFE) estimator, is that it can compare just-treated and already-treated units (e.g., De Chaisemartin and d'Haultfoeuille (2020); Goodman-Bacon (2021)). Under heterogeneous treatment effects, this estimator could lead to "negative weights" and severely bias the TWFE estimator. However, I have a substantial sample of never-treated units. Thus, most of the control group comprises never-treated units rather than units treated in earlier periods. This significantly reduces the influence of prohibited comparisons between already-treated units and helps mitigate potential biases in TWFE estimators. I plan to estimate Callaway and Sant'Anna (2021) in the future.

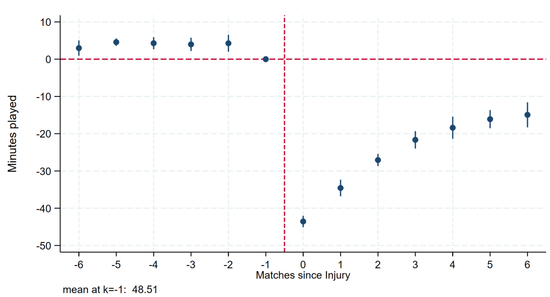
4 Results

Not all injuries are unpredictable. Figure 1a shows the coefficients associated with the exercise before applying the data-based algorithm to select the unpredictable types of injuries. Figure 1a focuses on minutes played and key passes as performance measures. When considering all types of injuries in the data, statistically significant differences exist between the parameters of the periods before the event and the excluded period. These "*pre-trends*" indicate that, compared to the period immediately before the shock, in the periods before this, the players had been having more minutes per game, which may be evidence of anticipatory behavior.

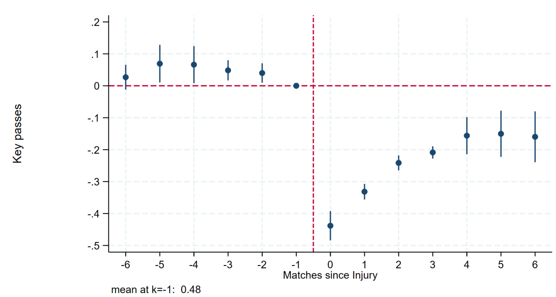
On the other hand, 1b shows additional evidence towards the "*pre-trends*". In the periods before the excluded period, the number of key passes made by a player in each game was significantly higher. However, the data reveals a trend of declining pass numbers, suggesting that some types of injuries in the database are not entirely unpredictable. Instead, some injuries affect the player partially in previous periods until he is finally absent due to the injury.

Figure 1. Estimation without applying the algorithm

(a) Minutes played



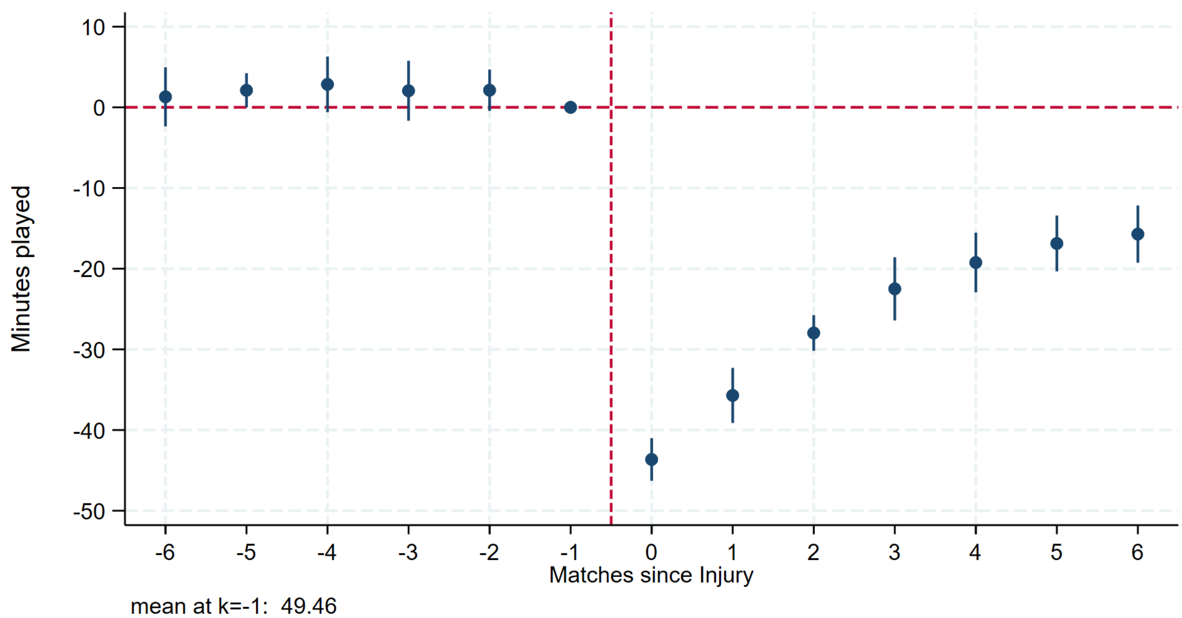
(b) Key passes



Note: This figure shows the effects of an unexpected injury on individual performance using all injuries. The results show the potential presence of anticipatory behaviors. Estimates include individual and match-week-season fixed effects, with clustered standard errors (individual, match-week-season).

Knowing the problem of "*pre-trends*", it is vital to sort out the types of injuries that are less predictable. I performed the algorithm discussed in the methodology section to separate the injuries that influenced the players' performance in previous periods from those that did not in order to comply with the assumptions of the event studies and make valid causal estimates. Figure 2 shows the effect of unexpected injuries on the number of minutes played when applying the data-driven classification algorithm. Comparing the results from figure 2 with figure 1a, one can see that the coefficients associated with the periods before the shock are not statistically significant at 5%, nor is there an evident trend in the punctual estimators of such parameters.

Figure 2. Effect of Injuries on the number of minutes played



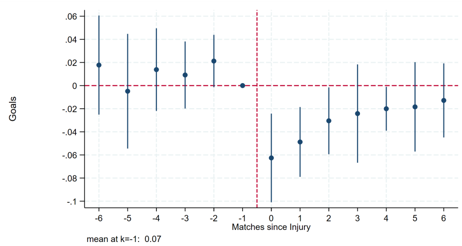
Note: This figure shows the effects of an unexpected injury on individual performance using the injuries selected with the data-driven algorithm to classify unexpected injuries. Estimates include individual and match-week-season fixed effects, with clustered standard errors (individual, match-week-season).

The periods after the event have highly significant statistical differences in the number of minutes played compared to the excluded period. For example, in period 0, the players are present on average 42 minutes less, which, compared to the average of the base period, is a relative reduction of 85% minutes played. This effect persists during the next six games, although the point estimate decreases as time passes. In other words, the players continue to play fewer minutes, although this decrease is slighter and less because as the games go by, it is likely that the players will recover from the injury.

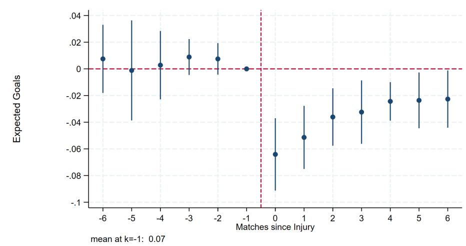
The overall performance of the injured players in the matches also decreased. Figure 3 shows the impact of injuries on different types of individual performance variables. Regardless of the metric, there is evidence that unexpected injuries negatively and significantly impact player productivity. That is, labor productivity suffers a negative shock.

Figure 3. Effect of injury on individual variables

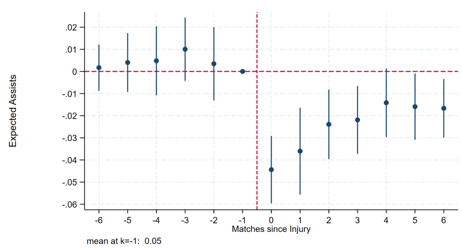
(a) Goals



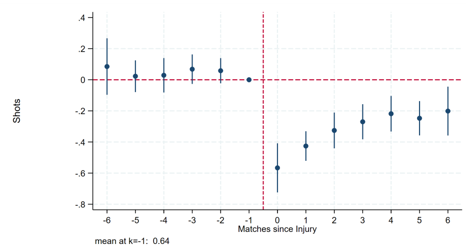
(b) Expected goals



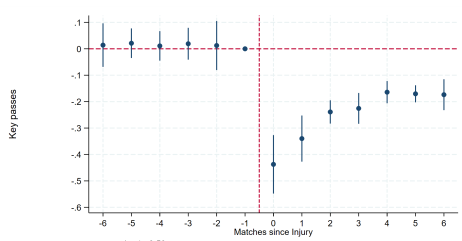
(c) Expected Assists



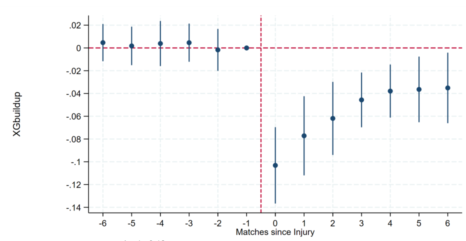
(d) Shots



(e) Key Passes



(f) XGBuildUp



Note: This figure shows the effects of an unexpected injury on individual performance using the injuries selected with the data-driven algorithm to classify unexpected injuries. There is evidence of a decrease in player's productivity measured using different performance variables. Estimates include individual and match-week-season fixed effects, with clustered standard errors (individual, match-week-season).

The number of real goals (Figure 3a) and expected goals (Figure 3b) have an immediate relative decrease of approximately 90%; however, the expected goals are more precise in terms of standard error. Differences in variability arise because, as mentioned above, variables such as goals have little variability in low-scoring sports such as football; thus, the effect of the injuries in the expected goals is persistent for more periods, although in the same way, after approximately six games, the impact stops being statistically different from zero.

The effect on other metrics of individual performance shows a similar dynamic, in which the period of the event presents the greatest impact in terms of magnitude, and this impact decreases as time passes. Note that, unlike figure 1b, in this case, none of the variables shows a trend in the periods before the shock.

Finally, Figure 3f shows the effect of unexpected injuries in the XGBuildUp. This is an innovative variable in football metrics as it quantifies the total expected goals of all the possessions in which the player was involved in the game, discounting key passes and shots during the 90 minutes, i.e., this variable gives an aggregate measure of the player's performance, with the *compliers* being the offensive players since most of these variables are measures of offensive performance.

Considering the dynamics presented by the previous results, being free of "pre-trends," I will continue analyzing results of estimates like the difference-in-differences model, where the coefficient presented is interpreted as the average of the effect in the six periods following the event[4].

Table 3 summarizes the evidence of the effect of the injuries on individual performance, similar to the graphs. All the performance measures show a negative and significant impact. For example, there is a 52.4% relative reduction in the average number of minutes played in the next six games, a 44% reduction in the probability of scoring a goal, and a

[4] The results for the rest of the variables of interest can be found in AppendixA.

52% reduction in the expected goals. Similarly, the rest of the variables show decreases of around 50%. These results show how the set of selected injuries presents behavior that allows me to identify the effect of these events on individual performance causally.

Table 3. Effect of injuries on individual performance

	Minutes (1)	**Goals (2)**	**Exp. Goals (3)**	**Exp. Assists (4)**	**Shots (5)**	**Key Passes (6)**	**XGbuildup (7)**
Pre	2.087*	0.0114	0.00509	0.00480	0.0526	0.0155	0.00264
	(0.971)	(0.008)	(0.006)	(0.005)	(0.032)	(0.016)	(0.005)
Post	-25.95***	-0.0310**	-0.0364***	-0.0247***	-0.322***	-0.250***	-0.0568***
	(0.816)	(0.010)	(0.006)	(0.004)	(0.035)	(0.012)	(0.010)
Constant	43.29***	0.0548***	0.0600***	0.0421***	0.549***	0.424***	0.0959***
	(1.019)	(0.008)	(0.005)	(0.004)	(0.029)	(0.018)	(0.008)
Observations	592,617	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.339	0.124	0.225	0.142	0.297	0.235	0.159
Mean at t-1	49.46	0.070	0.070	0.050	0.640	0.500	0.120

Note: This table shows the effects of an unexpected injury on individual performance. Estimates include individual and match-week-season fixed effects, with clustered standard errors (individual, match-week-season) in parentheses. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

I additionally performed regressions where the dependent variable is team performance to answer how these exogenous shocks in the productivity of the players can affect the firm's total productivity. Table 4 shows the effect of the shock on the variables at the team level. Despite being negative, the coefficients associated with the team performance are only marginally or not significant. That is, on average, individual injuries do not seem to have a statistically significant or large effect on the aggregate playmaking variables measured in the number of key passes, *NPXG*, number of shots, goals, and expected assists, or more complex measures like *XGBuildUp* or *XGChain*.

Given the evidence of the figures and tables presented so far, I cannot reject the hypothesis that, even though an exogenous shock generated by the absence of labor affects individual productivity, the team can optimally respond to this event to compensate in some way for such a decrease and that the aggregate output is not affected. The teams manage to replace the contribution of the injured players so that the team performance, on

Table 4. Effect of injuries on team performance

	Key Passes (1)	**XGChain (2)**	**NPXG (3)**	**Shots (4)**	**Exp. Assists (5)**	**Exp. Goals (6)**	**XGBuildUp (7)**
Pre	-0.172	-0.042	-0.006	-0.243	-0.008	0.002	-0.034
	(0.117)	(0.076)	(0.007)	(0.152)	(0.009)	(0.009)	(0.067)
Post	-0.223	-0.103	-0.031*	-0.331*	-0.018	-0.021	-0.077
	(0.115)	(0.057)	(0.012)	(0.147)	(0.013)	(0.015)	(0.050)
Constant	9.052***	3.472***	1.192***	12.130***	0.920***	1.293***	2.060***
	(0.085)	(0.077)	(0.017)	(0.115)	(0.013)	(0.019)	(0.060)
Observations	592,617	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.175	0.218	0.183	0.171	0.178	0.177	0.209
Mean at t-1	9.270	3.690	1.240	12.410	0.960	1.350	2.200

Note: This table shows the effects of an unexpected injury on team performance. Estimates include individual and match-week-season fixed effects, with clustered standard errors (team, match-week-season) in parentheses. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

average, is not altered. However, my lack of power to reject this hypothesis may be due to the possibility of mixing heterogeneous effects that, when studied individually, could shed more light on the impact of these exogenous shocks on individual productivity, a heterogeneity that I intend to exploit in the next section.

4.1 Heterogeneous Effects

This section includes the results of heterogeneous effects using the double interaction between teams with high or low productivity and players with high or low productivity. To carry out such a classification, I considered the distribution of the variable *XGChain*, cumulative until the date of the injury, both individually and at the team level. Those above the 75th percentile were classified as high productivity.[5] The importance of this variable, which adds the expected goals of all the plays in which the player was involved, is that it has preferable characteristics in my context, such as the fact that it has more significant variability for players in all positions on the field, which allows better identification of the level of productivity.

Table 5 shows the heterogeneity in individual productivity. The injury's effect is mostly

[5] Given the absence of clearly defensive court indicators, most of the variables analyzed in the study mainly consider the offensive aspect of the players' performance.

Table 5. Heterogeneous effects of injuries on individual performance

	Minutes Played (1)	**Goals (2)**	**Exp. Goals (3)**	**Exp Assists (4)**	**Shots (5)**	**Key Passes (6)**	**XGbuildup (7)**
(a) Low P. - Low T.	-26.490***	-0.017*	-0.024***	-0.019**	-0.241***	-0.192***	-0.044***
	(1.391)	(0.008)	(0.004)	(0.005)	(0.031)	(0.017)	(0.007)
(b) Low P. - High T.	-19.130***	-0.020	-0.023*	-0.011	-0.180*	-0.123***	-0.044***
	(1.207)	(0.011)	(0.009)	(0.005)	(0.070)	(0.026)	(0.009)
(c) High P. - Low T.	-28.270***	-0.049**	-0.054***	-0.029***	-0.450***	-0.323***	-0.061***
	(1.135)	(0.013)	(0.007)	(0.004)	(0.039)	(0.023)	(0.011)
(d) High P. - High T.	-26.120***	-0.045**	-0.048***	-0.039***	-0.414***	-0.348***	-0.084***
	(1.037)	(0.014)	(0.009)	(0.005)	(0.038)	(0.027)	(0.013)
Constant	42.890***	0.054***	0.059***	0.041***	0.538***	0.416***	0.094***
	(1.066)	(0.009)	(0.005)	(0.004)	(0.031)	(0.018)	(0.008)
Observations	592,617	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.340	0.124	0.225	0.142	0.298	0.236	0.159
Mean at t-1	49.460	0.070	0.070	0.050	0.640	0.500	0.120
Mean (a)	44.390	0.020	0.040	0.020	0.370	0.270	0.070
Mean (b)	33.810	0.050	0.040	0.020	0.270	0.140	0.070
Mean (c)	56.140	0.070	0.090	0.070	0.840	0.730	0.130
Mean (d)	60.360	0.160	0.130	0.100	1.130	0.860	0.220

Note: This table shows the heterogeneous effects of an unexpected injury on player performance. The table focuses on the interaction between team and player productivity levels. Productivity was classified based on the variable *XGChain*, measured cumulatively up to the date of the injury, both individually and at the team level. Teams or players with productivity above the 75th percentile were categorized as high productivity. Estimates include individual and match-week-season fixed effects, with clustered standard errors (individual, match-week-season) in parentheses. Each row shows (a) low productivity player - low productivity team. (b) low productivity player - high productivity team. (c) High productivity player - low productivity team. (d) high productivity player - high productivity team. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

maintained in the four analysis groups, showing relatively similar magnitudes of the coefficients associated with the impact. However, the effect on goals and shots stops to be significant for the group comprising low-productivity players in high-productivity teams. The effect on expected assists and key passes, even though they continue to be significant, is only marginally so, probably due to a lower precision in the estimation.

Table 6 shows the heterogeneity in the results at the team level. Estimates suggest that a player of a certain skill level can have a different impact depending on the type of players around him, i.e., the quality of the team to which he belongs. Injuries that affect high-productivity players in low-productivity teams decrease team productivity by around 6% to 8% when measured by key passes, *XGChain*, expected non-penalty goals (*NPXG*), shots, expected assists, expected goals, and *XGBuildUp*. The group of high-productivity players in low-productivity teams is the only of the four groups that is significantly affected at 5%. This result turns out to be highly intuitive since it is not surprising that the teams

Table 6. Heterogeneous effects of injuries on team performance

	Key Passes (1)	**XGChain** (2)	**NPXG** (3)	**Shots** (4)	**Exp. Assists** (5)	**Exp. Goals** (6)	**XGbuildup** (7)
(a) Low P. - Low T.	-0.113	-0.048	-0.016	-0.206	-0.005	-0.001	-0.042
	(0.120)	(0.061)	(0.015)	(0.127)	(0.012)	(0.019)	(0.055)
(b) Low P. - High T.	-0.150	-0.166	-0.027	-0.232	-0.017	-0.014	-0.130
	(0.211)	(0.102)	(0.040)	(0.310)	(0.035)	(0.041)	(0.063)
(c) High P. - Low T.	-0.614***	-0.270**	-0.075**	-0.754***	-0.057**	-0.076**	-0.179**
	(0.112)	(0.078)	(0.017)	(0.112)	(0.016)	(0.022)	(0.063)
(d) High P. - High T.	-0.038	0.011	-0.013	-0.152	-0.002	-0.001	-0.002
	(0.210)	(0.091)	(0.031)	(0.286)	(0.026)	(0.027)	(0.073)
Constant	9.036***	3.470***	1.191***	12.120***	0.919***	1.291***	2.060***
	(0.079)	(0.076)	(0.016)	(0.113)	(0.012)	(0.018)	(0.059)
Observations	592,617	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.175	0.218	0.183	0.172	0.178	0.177	0.209
Mean at t-1	9.270	3.690	1.240	12.410	0.960	1.350	2.200
Mean (a)	8.190	2.840	1.050	11.210	0.790	1.130	1.600
Mean (b)	10.150	4.280	1.360	13.430	1.070	1.510	2.650
Mean (c)	8.410	3.120	1.100	11.270	0.860	1.170	1.820
Mean (d)	11.790	5.600	1.690	15.390	1.330	1.850	3.520

Note: This table shows the heterogeneous effects of an unexpected injury on team performance. The table focuses on the interaction between team and player productivity levels. Productivity was classified based on the variable *XGChain*, measured cumulatively up to the date of the injury, both individually and at the team level. Teams or players with productivity above the 75th percentile were categorized as high productivity. Estimates include individual and match-week-season fixed effects, with clustered standard errors (team, match-week-season) in parentheses. Each row shows (a) low productivity player - low productivity team. (b) low productivity player - high productivity team. (c) High productivity player - low productivity team. (d) high productivity player - high productivity team. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

with the lowest productivity are the ones that depend on the most on their players, even more so if it is a player who, at the time of his injury, had a high level of productivity.[6]

The heterogeneity of the teams and players suggests that there are specific characteristics under which it is only sometimes possible to supply a shock fully and optimally in the productivity of the factors of production. Moreover, these shocks decrease the team's ability to play in a team. Finally, even though the evidence on the probability of defeat, draw, or victory is not statistically significant, Point estimator signs associated with these coefficients are intuitive for the group in question.

Table 7 the Leave-Out estimation. It estimates the effect of the event on the team-level

[6] The effects on the probability of defeat, draw, or victory were also estimated where, perhaps due to the low variability of these measures, I did not find significant effects. (See Appendix A).

variables, discounting the individual contribution of each player. Although only the coefficients associated with key passes and shots (*XGChain* marginally significant) turn out to be statistically significant, with a drop in the total performance of the team of around 3.77% and 2.9%, respectively, it is worth mentioning the signs and magnitudes of the four groups in question. First, for groups of players with low productivity, the relative effects are much smaller and close to zero compared to those with high productivity. Then, for the most productive players in productive teams, estimates suggest a positive sign in all the variables, suggesting that the injuries of productive players are replaced mainly by players who are also highly productive, which means that, by discounting his contribution, the rest of the players have a positive effect due to the entry of a substitute player and/or the increase in the contribution of the rest of the players.

Table 7. *Leave-Out* estimation of the effects of injuries on team performance

	Key Passes	**XGChain**	**NPXG**	**Shots**	**Exp. Assists**	**Exp. Goals**	**XGbuildup**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
(a) Low P. - Low T.	0.079	0.022	0.007	0.035	0.014	0.023	0.002
	(0.116)	(0.053)	(0.013)	(0.117)	(0.011)	(0.016)	(0.049)
(b) Low P. - High T.	-0.027	-0.099	-0.006	-0.052	-0.006	0.009	-0.087
	(0.198)	(0.102)	(0.040)	(0.295)	(0.032)	(0.042)	(0.062)
(c) High P. - Low T.	-0.290**	-0.153*	-0.030	-0.305**	-0.028	-0.022	-0.118
	(0.096)	(0.070)	(0.017)	(0.073)	(0.015)	(0.020)	(0.056)
(d) High P. - High T.	0.310	0.153	0.035	0.262	0.038	0.047	0.082
	(0.231)	(0.087)	(0.032)	(0.329)	(0.030)	(0.027)	(0.068)
Constant	8.620***	3.310***	1.137***	11.580***	0.878***	1.232***	1.965***
	(0.074)	(0.066)	(0.014)	(0.087)	(0.008)	(0.016)	(0.054)
Observations	592,617	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.171	0.214	0.179	0.170	0.171	0.174	0.205
Mean at t-1	8.780	3.490	1.180	11.770	0.910	1.270	2.080
Mean (a)	7.920	2.740	1.020	10.840	0.760	1.100	1.530
Mean (b)	10.010	4.170	1.320	13.170	1.050	1.470	2.580
Mean (c)	7.680	2.890	1.020	10.430	0.800	1.080	1.690
Mean (d)	10.930	5.220	1.570	14.250	1.230	1.710	3.300

Note: This table shows the *Leave-Out* estimation. This table shows the effects of an unexpected injury on the rest of the team's performance, excluding the player's individual contribution. Estimates include individual and match-week-season fixed effects, with clustered standard errors (team, match-week-season) in parentheses. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

On the other hand, the impact on productive players in low-productivity teams is negative and significant in some cases. In this case, it gives an intuition of the complementarity of this productive player over his teammates, i.e., discounting the decrease in individual productivity. The rest of the players have worse performance since the affected player enhanced the performance of his teammates on the field.

Table 8. Cross-Elasticities

	Key Passes (1)	XGChain (2)	NPXG (3)	Shots (4)	Exp. Assists (5)	Exp. Goals (6)	XGbuildup (7)
			Panel A - Individual Elasticities				
(a) Low P. - Low T.	-0.014	-0.011	-0.009	-0.005	-0.020	-0.035	-0.002
(b) Low P. - High T.	0.003	0.043	0.009	0.006	0.011	-0.011	0.054
(c) High P. - Low T.	**0.085**	**0.105**	**0.052**	**0.055**	**0.086**	0.035	**0.149**
(d) High P. - High T.	**-0.070**	**-0.079**	**-0.056**	**-0.050**	**-0.078**	**-0.074**	**-0.065**
			Panel B - Aggregate Elasticities				
(a) Low P. - Low T.	-0.413	-0.308	-0.298	-0.144	**-0.759**	**-0.959**	-0.037
(b) Low P. - High T.	0.220	**1.478**	0.289	0.286	0.556	-0.398	1.973
(c) High P. - Low T.	**0.898**	**1.319**	**0.656**	**0.678**	**0.983**	0.415	**1.931**
(d) High P. - High T.	**-0.891**	**-1.085**	**-0.738**	**-0.633**	**-0.959**	**-0.973**	**-0.979**

Notes: In Panel A, the individual cross-elasticities are calculated as the division between the relative effect of the team-level *Leave-Out Means* estimate on the relative effect of the individual estimate. Elasticities greater than 0.05 in absolute value are in bold. In Panel B, the aggregate cross-elasticities are calculated as the coefficient associated with the effect of the injury on the *Leave-Out Means* estimate on the coefficient of the individual estimate. Elasticities greater than 0.6 in absolute value are in bold.

Finally, Table 8 presents an estimate of both individuals' crossed elasticities, that is, how each player is affected by the absence due to injury of another, and aggregate elasticities, whose interpretation is the aggregate effect on the team. Overall, coefficients are smaller when looking at the groups of less productive players and can change significantly depending on the variable. However, when it comes to highly productive players on unproductive teams, a robust positive sign reinforces the idea of complementarity between this class of players and their teammates. In contrast, for good players in productive teams, the negative sign reflects the substitution that may exist where most of the workforce is highly productive.

5 Future work

This article lays the groundwork for a research agenda on the individual contribution of factors of production to aggregate output in team contexts. Additional steps for the future are the download and use of defensive data, which allows us to have a comprehensive measure of player performance and to calculate cross-role interactions within the field of play, analyzing, for example, the effect on offensive productivity of the loss of defensive players.

Likewise, future work should focus on calculating confidence intervals for the elasticities presented in this report, either through a *Bootstrap-type* estimation or a two-stage estimation that calculates the denominator and numerator of such elasticities together. Finally, it is in the plans to carry out robustness exercises on the inherent assumptions of the *event study*, particularly the assumption of homogeneity of the effect.

6 Conclusion

This document provides new evidence on the effect of a decrease in individual labor productivity on total productivity in team contexts. Using exogenous events determined by unforeseen injuries to football players, I observed a significant decrease in individual performance in the periods after the shock, measured through different productivity metrics such as goals, expected assists, or modern measures of game creation. At the same time, no effect is found on the aggregate performance of the team. However, by dividing the sample into teams and players according to their level of productivity, it is found that productive players on less productive teams have a negative relative effect of approximately 6%-8% on total shots, expected assists, anticipated goals, and generation of team play.

When calculating the cross-elasticities for the four types of groups identified, I see how it is not entirely conclusive for the groups of low-productive players due to values close to zero and changes in signs depending on the analyzed variable. However, for players highly productive in low-productivity teams, a positive sign indicates the complementarity of this type of player. In contrast, the most productive players in highly effective teams show a negative sign that indicates possible substitutability between this player and his teammates, which by definition belongs to a team with a higher proportion of productive players.

References

Arcidiacono, Peter, Josh Kinsler, and Joseph Price, "Productivity spillovers in team production: Evidence from professional basketball," *Journal of Labor Economics*, 2017, *35* (1), 191–225.

Bonhomme, Stéphane, "Teams: Heterogeneity, Sorting and Complementarity," *A preliminary draft of the work presented at the 2020 World Congress of the Econometric Society*, 2020.

Bryson, Alex, Peter Dolton, J James Reade, Dominik Schreyer, and Carl Singleton, "Experimental effects of an absent crowd on performances and refereeing decisions during Covid-19," *Available at SSRN 3668183*, 2020.

Bundesliga, "xG Stats Explained: The Science Behind Sportec Solutions' Expected goals model," https://www.bundesliga.com/en/bundesliga/news/expected-goals-xg-model-what-is-it-and-why-is-it-useful-sportec-solutions-3177 2018.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of econometrics*, 2021, *225* (2), 200–230.

Caruso, Raul, Marco Di Domizio, and Domenico Rossignoli, "Aggregate wages of players and performance in Italian Serie A," *Economia Politica*, 2017, *34* (3), 515–531.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American economic review*, 2020, *110* (9), 2964–2996.

Devereux, Kevin, "Identifying the value of teamwork: Application to professional tennis," Technical Report, Working Paper Series 2018.

Dvorak, Jiri, Astrid Junge, Toni Graf-Baumann, and Lars Peterson, "Football is the most popular sport worldwide," *American journal of sports medicine*, 2004, *32* (Suppl 1), 3S–4S.

Efe, "Las cinco grandes ligas europeas, cada vez más poderosas económicamente," https://www.mundodeportivo.com/futbol/internacional/20200117/472930369346/las-cinco-grandes-ligas-europeas-cada-vez-mas-poderosas-economicamente.html Jan 2020.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M Shapiro, "Pre-event trends in the panel event-study design," *American Economic Review*, 2019, *109* (9), 3307–38.

Gauriot, Romain and Lionel Page, "Fooled by Performance Randomness: Overrewarding Luck," *Review of Economics and Statistics*, 2019, *101* (4), 658–666.

Gaviria, Alejandro, "Is soccer dying? A time series approach," *Applied Economics Letters*, 2000, *7* (4), 275–278.

Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," *Journal of econometrics*, 2021, *225* (2), 254–277.

Jacobson, Louis S, Robert J LaLonde, and Daniel G Sullivan, "Earnings losses of displaced workers," *The American economic review*, 1993, pp. 685–709.

Matano, Francesca, Lee F Richardson, Taylor Pospisil, Collin Eubanks, and Jining Qin, "Augmenting adjusted plus-minus in soccer with FIFA ratings," *arXiv preprint arXiv:1810.08032*, 2018.

Neale, Walter C, "The peculiar economics of professional sports," *The quarterly journal of economics*, 1964, *78* (1), 1–14.

Rosen, Sherwin and Allen Sanderson, "Labour markets in professional sports," *The economic journal*, 2001, *111* (469), F47–F68.

Roth, Jonathan, "Pretest with caution: Event-study estimates after testing for parallel trends," *American Economic Review: Insights*, 2022, *4* (3), 305–322.

Rottenberg, Simon, "The baseball players' labor market," *Journal of political economy*, 1956, *64* (3), 242–258.

Scully, Gerald W, *The market structure of sports*, University of Chicago Press, 1995.

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

Stubbe, Janine H, Anne-Marie MC van Beijsterveldt, Sissi van der Knaap, Jasper Stege, Evert A Verhagen, Willem Van Mechelen, and Frank JG Backx, "Injuries in professional male soccer players in the Netherlands: a prospective cohort study," *Journal of athletic training*, 2015, *50* (2), 211–216.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of econometrics*, 2021, *225* (2), 175–199.

Zuke, William A, Avinesh Agarwalla, Beatrice Go, Justin W Griffin, Brian J Cole, Nikhil N Verma, Bernard R Bach, and Brian Forsythe, "The lack of standardized outcome measures following lower extremity injury in elite soccer: a systematic review," *Knee surgery, sports traumatology, arthroscopy*, 2018, *26* (10), 3109–3117.

A Appendix Tables and Figures

Appendix Table A1. Effect of injuries on individual performance, other performance measures

	Assists (1)	**Yellow Card (2)**	**Red Card (3)**	**NPG (4)**	**NPGX (5)**	**XGChain (6)**
Pre	-0.002	0.022	0.002	0.014	0.007	0.010
	(0.005)	(0.012)	(0.001)	(0.008)	(0.006)	(0.008)
Post	-0.030***	-0.047**	0.001*	-0.028**	-0.033***	-0.097***
	(0.005)	(0.013)	(0.0004)	(0.010)	(0.007)	(0.012)
Constant	0.052***	0.082***	0.001	0.050***	0.055***	0.163***
	(0.005)	(0.011)	(0.001)	(0.008)	(0.005)	(0.009)
Observations	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.065	0.062	0.010	0.108	0.215	0.222
Mean at t-1	0.060	0.100	0.000	0.060	0.060	0.190

Note: This table shows the effects of an unexpected injury on individual performance. Estimates include individual and match-week-season fixed effects, with clustered standard errors (individual, match-week-season) in parentheses. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

Appendix Table A2. Effect of injuries on team performance, other performance measures

	Tie (1)	**Victory (2)**	**Defeat (3)**	**Goal Difference (4)**	**Scored Goals (5)**	**Conceded Goals (6)**
Pre	-0.007	0.025	-0.019	0.047	0.003	-0.044
	(0.020)	(0.013)	(0.017)	(0.052)	(0.034)	(0.039)
Post	-0.008	0.014	-0.006	-0.018	-0.050	-0.032
	(0.022)	(0.019)	(0.027)	(0.059)	(0.033)	(0.049)
Constant	0.253***	0.369***	0.378***	0.028	1.403***	1.375***
	(0.022)	(0.020)	(0.023)	(0.062)	(0.042)	(0.042)
Observations	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.022	0.086	0.071	0.118	0.110	0.060
Mean at t-1	0.250	0.390	0.370	0.100	1.470	1.370

Note: This table shows the effects of an unexpected injury on team performance. Estimates include individual and match-week-season fixed effects, with clustered standard errors (team, match-week-season) in parentheses. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

Appendix Table A3. Heterogeneous effects of injuries on individual performance, other performance measures

	Assists	**Yellow Card**	**Red Card**	**NPG**	**NPGX**	**XGChain**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
(a) Low P. - Low T.	-0.025***	-0.051**	0.001	-0.015	-0.023**	-0.070***
	(0.005)	(0.016)	(0.001)	(0.008)	(0.005)	(0.009)
(b) Low P. - High T.	-0.017*	-0.021	0.001	-0.017	-0.021*	-0.067***
	(0.007)	(0.019)	(0.001)	(0.010)	(0.009)	(0.014)
(c) High P. - Low T.	-0.038***	-0.053***	0.001	-0.039**	-0.045***	-0.116***
	(0.007)	(0.008)	(0.001)	(0.012)	(0.007)	(0.012)
(d) High P. - High T.	-0.039***	-0.048**	0.001	-0.046**	-0.048***	-0.141***
	(0.007)	(0.014)	(0.001)	(0.015)	(0.010)	(0.018)
Constant	0.051***	0.081***	0.001	0.048***	0.054***	0.160***
	(0.004)	(0.011)	(0.001)	(0.008)	(0.005)	(0.010)
Observations	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.065	0.062	0.010	0.108	0.216	0.223
Mean at t-1	0.060	0.100	0.000	0.060	0.060	0.190
Mean (a)	0.020	0.100	0.000	0.020	0.030	0.100
Mean (b)	0.020	0.070	0.010	0.050	0.040	0.120
Mean (c)	0.080	0.100	0.000	0.060	0.080	0.230
Mean (d)	0.150	0.100	0.000	0.140	0.120	0.380

Note: This table shows the heterogeneous effects of an unexpected injury on player performance. The table focuses on the interaction between team and player productivity levels. Productivity was classified based on the variable *XGChain*, measured cumulatively up to the date of the injury, both individually and at the team level. Teams or players with productivity above the 75th percentile were categorized as high productivity. Estimates include individual and match-week-season fixed effects, with clustered standard errors (individual, match-week-season) in parentheses. Each row shows (a) low productivity player - low productivity team. (b) low productivity player - high productivity team. (c) High productivity player - low productivity team. (d) high productivity player - high productivity team. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

Appendix Table A4. Heterogeneous effects of injuries on team performance, other performance measures

	Tie (1)	**Victory** (2)	**Defeat** (3)	**Goal Difference** (4)	**Scored Goals** (5)	**Conceded Goals** (6)
(a) Low P. - Low T.	-0.006	0.017	-0.012	0.025	-0.027	-0.052
	(0.026)	(0.024)	(0.025)	(0.060)	(0.044)	(0.038)
(b) Low P. - High T.	-0.031	0.028	0.003	-0.014	-0.031	-0.017
	(0.022)	(0.034)	(0.039)	(0.117)	(0.065)	(0.081)
(c) High P. - Low T.	-0.007	-0.001	0.008	-0.087	-0.082	0.005
	(0.018)	(0.024)	(0.029)	(0.081)	(0.041)	(0.056)
(d) High P. - High T.	0.000	0.018	-0.018	-0.022	-0.067	-0.045
	(0.034)	(0.008)	(0.031)	(0.049)	(0.038)	(0.065)
Constant	0.253***	0.369***	0.378***	0.026	1.401***	1.375***
	(0.022)	(0.020)	(0.024)	(0.062)	(0.040)	(0.042)
Observations	592,617	592,617	592,617	592,617	592,617	592,617
R-Squared	0.022	0.086	0.071	0.119	0.110	0.060
Mean at t-1	0.250	0.390	0.370	0.100	1.470	1.370
Mean (a)	0.250	0.320	0.430	-0.290	1.230	1.520
Mean (b)	0.310	0.430	0.270	0.490	1.670	1.180
Mean (c)	0.280	0.270	0.450	-0.250	1.270	1.510
Mean (d)	0.180	0.610	0.210	1.000	2.030	1.030

Note: This table shows the heterogeneous effects of an unexpected injury on team performance. The table focuses on the interaction between team and player productivity levels. Productivity was classified based on the variable *XGChain*, measured cumulatively up to the date of the injury, both individually and at the team level. Teams or players with productivity above the 75th percentile were categorized as high productivity. Estimates include individual and match-week-season fixed effects, with clustered standard errors (team, match-week-season) in parentheses. Each row shows (a) low productivity player - low productivity team. (b) low productivity player - high productivity team. (c) High productivity player - low productivity team. (d) high productivity player - high productivity team. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

B Appendix One

This appendix displays is a list of the variables used in the analysis and a brief description:

- **Assists**: The number of assists made by a player or team, which are passes leading directly to a goal.
- **Expected Assists**: The expected number of assists based on the quality of chances created.
- **Shots**: The total number of shots attempted by a player or team.
- **Goals**: The total number of goals scored by a player or team.
- **Expected Goals (xG)**: A metric that estimates the likelihood of a shot resulting in a goal based on various factors such as shot angle and distance.
- **Non-Penalty Goals**: Goals scored excluding penalty kicks.
- **Minutes Played**: The total number of minutes played by a player during matches.
- **NPGX**: A combination metric representing Non-Penalty Goals and Expected Goals.
- **Key Passes**: Passes that lead to a shot, showing a player's ability to create scoring opportunities.
- **Yellow Cards**: The number of yellow cards received by a player.
- **Red Cards**: The number of red cards received by a player, leading to ejection from the match.
- **XGBuildUp**: A measure of involvement in build-up play leading to scoring opportunities, excluding direct assists and goals.
- **XGChain**: A measure of a player's involvement in all passing sequences that lead to a shot or goal.