

Lackner, Mario; Sonnabend, Hendrik

Article — Published Version

Presenteeism when employers are under pressure: evidence from a high-stakes environment

Economica

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Lackner, Mario; Sonnabend, Hendrik (2023) : Presenteeism when employers are under pressure: evidence from a high-stakes environment, *Economica*, ISSN 1468-0335, Wiley, Hoboken, NJ, Vol. 90, Iss. 358, pp. 477-507,
<https://doi.org/10.1111/ecca.12461>

This Version is available at:

<https://hdl.handle.net/10419/287876>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Presenteeism when employers are under pressure: evidence from a high-stakes environment

Mario Lackner¹ | Hendrik Sonnabend² 

¹Johannes Kepler University and Christian Doppler Laboratory, Austria

²University of Hagen, Germany

Correspondence

Hendrik Sonnabend, University of Hagen, Germany.

Email: hendrik.sonnabend@fernuni-hagen.de

Abstract

This study analyses whether the decision to work while sick can be linked to workload fluctuations. Drawing on data collected from professional football, we exploit the dynamics of a season and use additional (national and international) cup games conducted in the second half of a season as a source of exogenous variation. We find robust evidence that players are 6.3 percentage points more likely to return from injuries earlier than expected when their teams are exposed to a high workload. The effect is driven by players who are more important to their teams and those who are less vulnerable to injuries. Finally, we find that presenteeism comes at the cost of an early comeback significantly shortening the time until the next injury by approximately 16 days.

1 | INTRODUCTION

Why do employees work even when their health status gives them a reason to stay at home? This phenomenon, often referred to as ‘presenteeism’, has received growing attention in different research areas such as occupational medicine, social psychology and various fields of economics. There is evidence that, for instance, working conditions and job security, workers’ attitudes, age, health status, and companies’ absence policies are important drivers of presenteeism (Hirsch *et al.* 2017; Miraglia and Johns 2016; Arnold 2016; Lohaus and Habermann 2019).

From an employer’s perspective, apart from the context of infectious diseases—‘contagious presenteeism’ has received much attention during the COVID-19 crisis (e.g. Pichler *et al.* 2020)—presenteeism acts as a double-edged sword. Reduced productivity is better than zero productivity resulting from sickness absence; however, it bears the risk of future health consequences and longer absence times (e.g. Schultz and Edington 2007; Bergström *et al.* 2009; Hansen and Andersen 2009; Skagen and Collins 2016).

The net utility derived from presenteeism is not necessarily constant over time. This is because workers’ absence may have little consequences in ‘quiet times’, but this could change when the employer is under pressure. One may think of financial auditors in the deadline phase.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Economica* published by John Wiley & Sons Ltd on behalf of London School of Economics and Political Science.

Consequently, decision-makers might feel an incentive to overvalue the effects of individual absences relative to the potential negative consequences, and enforce presenteeism in times of high workloads.

This interplay between presenteeism and employer workload dynamics has not been examined clearly in the existing literature. In this study, we try to fill this gap using data from professional football. Specifically, we examine how quasi-random shocks on the number of matches (the workload) affect the recovery time of players (the employees). The idea is that additional games put pressure on teams to field players who were previously unavailable due to well-described medical conditions. Depending on factors such as a player's importance to his team (e.g. his relative productivity), his vulnerability to injuries, and the overall level of absenteeism, players may have an incentive to return from their absence before the scheduled time if their team needs them. In the field of sports, there have been prominent examples of a massive abuse of painkillers, and top players participating in decisive matches while being affected by the flu.¹ The 'playing hurt' culture in professional sports has received attention in sport science, sport medicine and sport sociology (e.g. Roderick *et al.* 2000; Mayer and Thiel 2018; Chen *et al.* 2019). There is reason to expect that the underlying motives for this behaviour are the same as those in less specific labour market segments. Players may want to be loyal to their club and their peers, or they want to signal their resilience and reliability to the job market. On the employer side, it may be beneficial for teams to take the risk of a secondary injury if the current incentives are sufficiently high. Conflicts may arise between players and teams when their interests do not coincide.

In general, while firm-level data on presenteeism are hardly available, and most empirical studies rely on self-reported survey data, our setting allows us to examine the nexus between workload and work-while-sick behaviour in a high-stakes field setting. Data on injuries in professional sports have been used before in studies observing issues in the fields of labour economics (Gregory-Smith 2021) and management science (Chan and Fearing 2019). Most closely related to our study, Ngo and Roberts (2021) analyse the relationship between contract status and missed games in the National Basketball Association (NBA). The authors show that for the average player, the likelihood of missing a game decreases towards the end of his contract, whereas the opposite is true for the best players (i.e. players who have been selected for the NBA All-Star Game).

Our empirical strategy is to exploit the dynamics of a standard season in European professional football. In the first step, we calculate the average recovery times from the first part of a season, where the number of games is fixed and the workload is (to a large extent) predictable. We analyse further how recovery times vary with respect to a team's number of games in the second part of the season, where teams are under pressure and the workload is far less predictable. This extra workload is virtually predetermined, as it originates from additional national and/or international cup games, which in turn result from a team's success in the first part of the season. Our estimates suggest that in the presence of a high workload, players return from their injuries significantly earlier. Specifically, players in the treatment group—depending on the definition and degree of presenteeism—have a higher probability for a reduced injury time compared to the non-treated players by 5–7 percentage points (p.p.). However, this effect is not equally strong for all types of players. In line with expectations, our analyses indicate that early returns are associated with a high (relative) productivity and a low vulnerability to injuries. Specifics of professional sports such as age and tenure do not seem to drive these results. We also find that presenteeism is associated with future costs: we estimate that the period until the next injury is shortened by 16–20 days for players exposed to a reduced recovery time. This means that the shorter the healing time, the sooner the next absence due to injury.

This study is structured as follows. In Section II, we outline a simple conceptual framework to motivate our empirical analysis. Section III presents the dataset and descriptive statistics of the absenteeism of players due to medical conditions. Sections IV and V investigate the nexus

between workload and presenteeism within our baseline models and in terms of various kinds of heterogeneity. Section VI identifies the costs of presenteeism. Robustness checks are provided in Section VII, and Section VIII concludes the study.

2 | CONCEPTUAL FRAMEWORK

To illustrate the key trade-off that we want to highlight in this study, we introduce a firm employing workers with different levels of productivity $\alpha_1 \gg \alpha_2 \gg \dots \alpha_i \dots > \alpha_{n-1} > \alpha_n$. The difference in abilities decreases in i , implying that the loss in productivity is smaller when worker $n - 1$ is absent and worker n is doing her job as compared to a situation where the most productive worker has to be replaced.²

Now consider the case where worker i is not in good health (but does not have an infectious disease) in period $t = 1$. Presenteeism means that the worker could continue working in period $t = 1$ at the expense of a lower productivity $\eta\alpha_i$, with $0 < \eta < 1$, and the risk that the disease can develop into a more serious condition that would cause absence in the next period, $t = 2$, with probability π_i . Otherwise, without presenteeism, the worker is replaced by a worker with lower productivity $i + 1$ in $t = 1$, and returns in good health in $t = 2$. Figure 1 illustrates the decision problem.

To account for the fact that the importance of a high productivity may vary between periods, payoffs in period 2 are discounted by a discount factor $0 < \delta < 1$. For instance, the company could be in troubled times in $t = 1$ due to a workload peak, or may have to deal with an already high level of overall absentees, whereas a stabilization is expected in $t = 2$. Payoffs are $Y^{NP} = \alpha_{i+1} + \delta\alpha_i$ and $Y^P = \eta\alpha_i + \delta[\pi_i\alpha_{i+1} + (1 - \pi_i)\alpha_i]$. Then for $\eta > \alpha_{i+1}/\alpha_i$, presenteeism yields higher payoffs than a sick leave if and only if

$$\delta \leq \frac{\eta\alpha_i - \alpha_{i+1}}{\pi_i(\alpha_i - \alpha_{i+1})}. \quad (1)$$

Hence presenteeism is more likely to occur in the presence of (i) high values of η , (ii) low values of δ and π_i , and (iii) the top productivity segment where workers can be less easily replaced because of the greater productivity gap. Since η and π_i depend on both individual predisposition and the disease type, these variables act as determinants of presenteeism. Finally, δ must be sufficiently small to cause presenteeism, all other things being equal. In other words, the firm places a much higher value on productivity in $t = 1$ compared to $t = 2$. Examples of periods of excessive

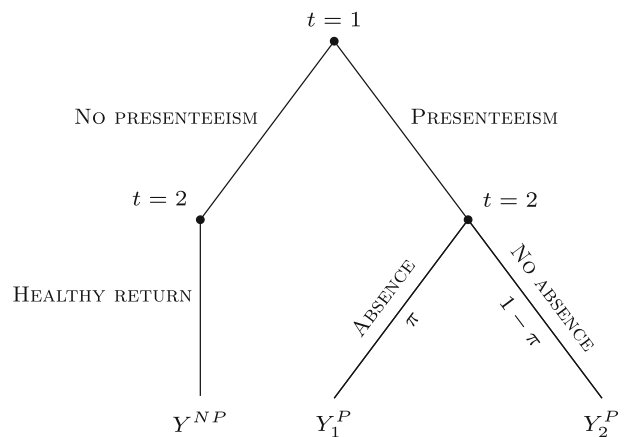


FIGURE 1 Decision on presenteeism.

workload and easing thereafter include high occupancy rates in hospitals, all kinds of business deadlines, and the decisive phase of a season in sports.

Note that the mechanisms highlighted in this simple exercise can be transferred easily to a situation where worker i has been absent due to medical conditions in period $t = 0$ and is not fully recovered in $t = 1$. A repeated absence in $t = 2$ then could mean exhaustion or that the disease recurred as the healing process was incomplete. This is exactly the scenario that we analyse in our empirical setting.

2.1 | Inefficiencies associated with presenteeism

So far, our simple model does not capture inefficiencies associated with presenteeism. Prior literature, however, has identified the risk of a worse future health status and lower mental wellbeing as negative consequences of work-while-sick behaviour, among others—not to mention the absence of co-workers in the event of a contagious disease (e.g. Miraglia and Johns 2016; Lohaus and Habermann 2019).

Two sources of inefficiencies that can be taken directly from the model are false beliefs about the risk of absence in the second period (π_i) and the present bias, meaning that firms overvalue payoffs from the first period (δ). For example, it is well known that without rest, mild infections can develop into serious coronary events. Nevertheless, there is evidence that people tend to underestimate that risk when they decide to go to work ill (e.g. Kivimäki *et al.* 2005). For companies/managers, prior research has demonstrated that present-biased preferences may lead to suboptimal investment decisions (Guenzel and Malmendier 2021; Kim and Nguyen 2022). Hence even in the case of complete information about health risks, the overvaluation of first-period payoffs may lead to an ‘underinvestment problem’ related to workers’ health.

This brings us to the third source of inefficiency: conflicts of interest between the employer and the employee. While presenteeism is preferred by the representative firm given that condition (1) holds, that does not necessarily have to be true from the worker perspective. For instance, it is reasonable to assume that income is not constant over time but depends on the worker’s health status. For the sake of simplicity, consider a situation where absence in period 2 results from a more severe and continuing sickness than in period 1, which in turn leads to a lower level of income (that may also involve transfers in case of unemployment). Then even if we assume that a healthy worker can earn more in period 1 than in period 2, due to bonuses for instance, it is reasonable to assume that there exist parameter values for which the worker prefers not to go to work ill and presenteeism would be welfare-reducing. And this situation is more likely to occur when the disutility associated with different degrees of illness is taken into account.

3 | DATA AND DESCRIPTIVE STATISTICS

We use data from the top football leagues in Germany (*Bundesliga*), Spain (*Primera Division*) and Italy (*Serie A*), covering ten seasons from 2010 to 2019. They were collected from *transfermarkt*,³ a popular German-based football website.

The focus is on player absences due to medical conditions, such as a pulled hamstring, a traumatic brain injury and gastroenteritis, among others. For the sake of simplicity, this study uses the term ‘injury’ to refer to a medical condition that causes a player to be absent from the game.

Initially, the raw data encompassed a total of 17,831 absences, categorized by medical diagnosis and linked with detailed information on the start and end dates as well as player and team characteristics.⁴ Since our analysis uses variations in the number of matches conducted in the cup

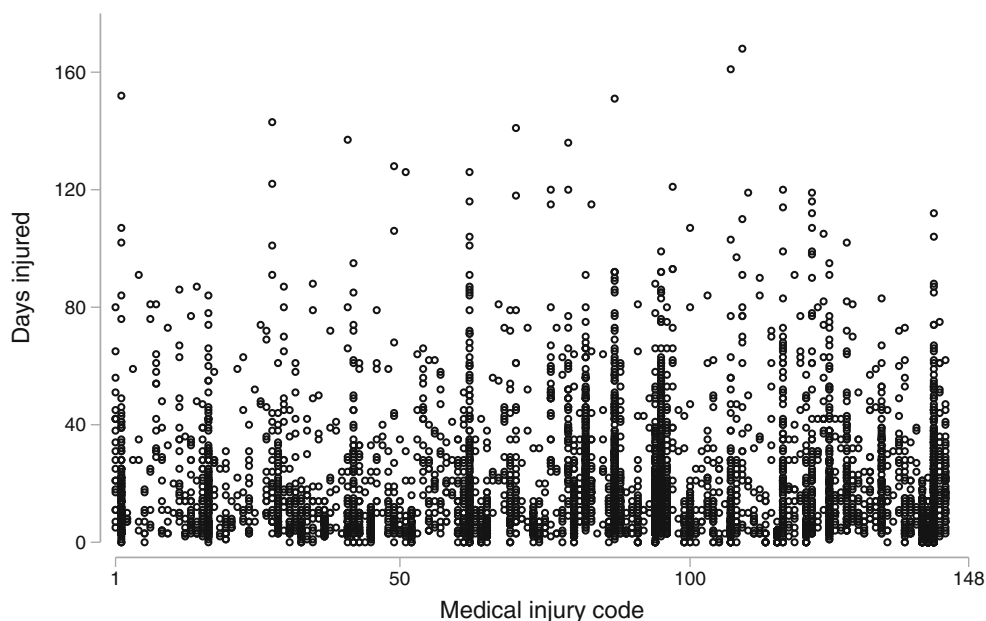


FIGURE 2 Observed recovery time in days by medical injury code. Notes : This figure illustrates the absolute injury duration (in days) for all injury categories in our final estimation sample. All injuries occurred and ended before 1 July. $N = 6429$.

games in the *second* half of a season (January to June), we calculate the average number of days that it takes for a player to recover from an injury for each category using the data on the injuries that ended within the *first* part of a season only (July to December). This is important because all absences that end within the first part of a season are not based on the decisions affected by the high workload treatment used in the empirical analysis.

As explained in the next section, we take these expected recovery times as reference points since they are exogenous in the sense that they are not affected by the workload variations under consideration. For the same reason, the data related to the players traded in the second half of a season (i.e. in January when the transfer window closes) were excluded. Finally, we are left with 6429 observations for the period January to June, assigned to 148 different types of injuries, with average duration 17.55 days of absence and standard deviation 18.51.⁵ Figure 2 plots the recovery times in days per category.⁶

We initially group the diagnoses into 21 categories according to the median recovery duration.⁷ Note that we do not use the mean, because extreme outliers may bias the results. Figure 3 illustrates the median recovery duration per injury type in ascending order, grouped into 21 categories (1 is the lowest injury severity group, and 21 is the highest injury severity group). It shows that the dispersion of recovery times is low for the majority of categories (1–20), but high for the category of the highest rank (with medians between about 60 and 180). Since this group appears to be extremely heterogeneous and therefore may bias our results, we excluded it in the analysis conducted below. Based on these 20 categories, we calculate expected recovery times for all types of injuries.

4 | EMPIRICAL STRATEGY AND MAIN RESULTS

Our empirical strategy is to exploit the dynamics of a standard season in European professional football. In football, the first half of a season (July to December, ‘the reference period’), starting

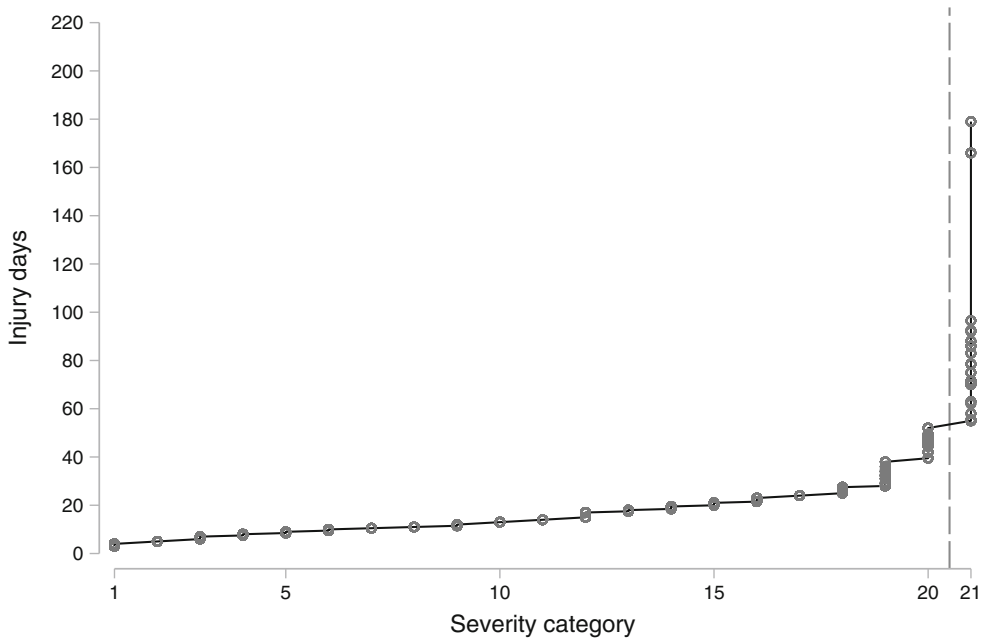


FIGURE 3 Median recovery duration in days per injury type in ascending order and grouped into 21 categories. Notes: The hollow dots indicate the median recovery duration in days per injury type in ascending order. $N = 7030$.

with a preparation period, is characterized by a fixed number of games in the national league, the first rounds of the national cup(s), and, if qualified, the group stage of European competitions (i.e. *UEFA Europa League* and *UEFA Champions League*). The workload is predictable, and teams can adjust their personnel decisions. However, in the second half of a season (January to June, ‘the relevant period’), the number of matches played varies according to a team’s success in the earlier phase of the national and international cup competitions. This is the critical phase of a season, where teams are under pressure and the workload is far less predictable.

The basic idea of the research design is to calculate the average recovery times per injury category based on the data obtained for the reference period; we then analyse how recovery times vary with respect to the number of games played by the team in the relevant period. In our setting, presenteeism means that a player returns from an injury of a certain type earlier than the predicted usual recovery duration of the specific injury. An intuitive way to model presenteeism behaviour of player j is to create a binary variable equal to 1 if the recovery time (in days) is at least one day shorter than expected for an injury of type i , and 0 otherwise. We call this variable $early_return_{i,j,k,t}$.

In order to compare teams with high and low workloads during the relevant observation period, we introduce a binary variable defined as

$$high_load_{k,t} = \begin{cases} 1, & \text{if } load_{k,t} \in Q_4, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

to indicate that the number of national and European cup games played by team k in the relevant period of season t is in the fourth quartile of the continuous variable $load_{k,t}$ (mean 2.68 and standard deviation 3.57). This variable measures the number of matches played after the initial group stages of the UEFA Champions League and the UEFA Europa League in the first half of a season. It also includes the number of matches played in the national cup competition after the first three rounds. The fourth quartile is equal to a workload of more than four games. The

maximum value of $load_{k,t}$ is 13, and the average number of games played by a team exposed to a high workload is 8.

While this definition of a high workload treatment is based on the overall distribution of games in the second half of the seasons, it also has strong theoretical foundations: since the fourth quartile of $load_{k,t}$ equals five or more games, it follows that for $high_load_{k,t} = 1$, team k still participates in one or more cup competitions. This means that our measure of workload is, to a large extent, predetermined with respect to the start and end of all injury times in our outcome sample.⁸ In essence, it indicates whether a team experiences *any* extra workload in addition to the mandatory league games during the second half of a season (which applies to 25% of all the team–season combinations in our data). Our estimates should therefore be viewed as a lower bound of the association between a high workload and presenteeism behaviour.

As we demonstrate later in this section, our main findings are robust to alternative cut-off values (3 and 5 matches). Moreover, we check the robustness of our results with a sample restricted to absences starting in the May to July period. This is because we cannot rule out the fact that our estimates may suffer from endogeneity bias in the sense that earlier returns from injuries increase a team’s workload due to an increase in performance. It shows that our main results hold and that the effects are even stronger. We explain this by the fact that injuries occurring during the ‘crunch time’ of a season are highly detrimental to team performance. Thus the incentive for presenteeism behaviour when the workload is high is greater than that observed earlier in the season.

Table 1 reports the summary statistics of injury characteristics for the high- and low-workload groups. As expected, the mean duration of injury is almost four days shorter for high-workload teams compared to low-workload teams. In other respects, we do not find relevant differences: the average severity category is basically identical, and the difference in the average starting month (0.12) is statistically significant but economically insignificant.

Table 2 presents our key variables on the team–season level. As per definition, teams with a high workload play additional national and international cup games in the relevant period from January to June. In addition, these teams perform significantly better in the autumn part of the season compared to low-workload teams. Finally, there is a difference in the overall number of injuries (2.42, significant at the 5% level), probably due to the higher number of games played. Although the difference is small, this could pose a threat to our identification strategy in case the higher number of injuries explains a potential positive association between high workload and presenteeism behaviour. We will address this issue in the heterogeneity analysis presented in Section V.

TABLE 1 Summary statistics by treatment status: injury level

	Low	High	Difference
Days	18.50 (19.24)	14.54 (15.56)	3.96***
Start month	2.83 (1.32)	2.95 (1.32)	−0.12**
End month	3.41 (1.40)	3.43 (1.39)	−0.02
Severity	9.11 (5.69)	9.31 (5.87)	−0.20
<i>N</i>	4898	1531	

Notes: All injuries in the second halves of seasons are included (January to July). $N = 6429$. Standard deviations are presented in parentheses. *t*-tests for the difference in means are presented in the third column. The variable ‘Severity’ refers to the 20 injury categories. *, **, *** indicate statistical significance at the 10%, 5%, 1% level, respectively.

TABLE 2 Summary statistics by treatment status: team–season level

	Low	High	Difference
Number of extra games	0.80 (1.21)	8.66 (2.77)	−7.86***
Rank ratio	0.58 (0.27)	0.28 (0.24)	0.30***
Injury count	12.43 (8.05)	14.84 (9.68)	−2.42**
<i>N</i>	396	103	

Notes: Descriptive statistics of the main variables at the team–season level. $N = 499$. Standard deviations are presented in parentheses. t -tests for the difference in means are presented in the third column. The variable ‘Rank ratio’ illustrates team performance by dividing the team’s current ranking by the worst possible according to the maximum number of teams in the league. *, **, *** indicate statistical significance at the 10%, 5%, 1% level, respectively.

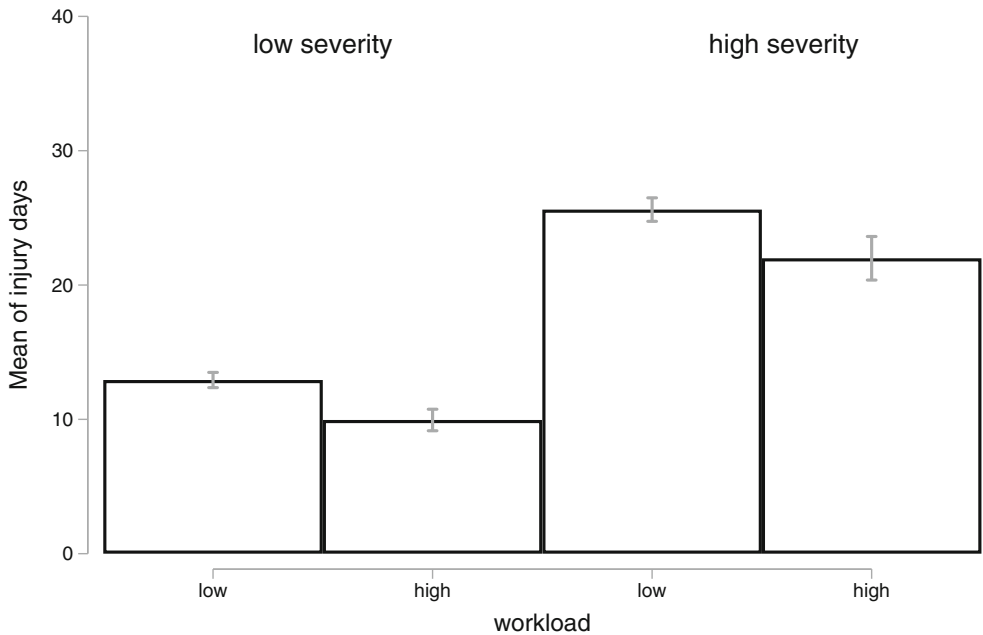


FIGURE 4 Average recovery times for high and low workloads. Notes: $N = 6429$. A high workload is defined by equation (2). Low severity means injuries of categories 1–10; high severity means categories 11–20.

Figure 4 provides an illustration of the association between workload and recovery times. We split the 20 injury severity categories into two main categories: ‘low severity’ (categories 1–10) and ‘high severity’ (categories 11–20). For both main categories, the average number of days for which an injured player is absent is lower when the club is exposed to high workload. That is, players of clubs with a high workload are more likely to return earlier than expected.

Next, we document the average treatment effect of having a high workload on the probability of an early return from an injury i of category c by estimating the equation

$$early_return_{i,j,k,t,c} = \beta_0 + \beta_1 high_workload_{k,t} + \phi' X_{i,j,k,t,c} + \xi_k + \pi_t + \rho_c + \varepsilon_{i,j,k,t,c}, \quad (3)$$

where $X_{ij,k,t,c}$ is a vector of injury-, team- and player-specific characteristics, including the win ratio and the relative league position of team k after the reference period, the team's total market value, and a dummy indicating the market value quartile of player j .⁹ Furthermore, we control for the team–season specific number of injuries in the reference period and the average injury length (in days). To account for the fact that an early return of player j might also be affected by his ‘substitutability’ within the team, we control for overall squad size of team k as well as the number of injuries occurred (*NoIP*) and the available players (*APP*) assigned to player j 's position during the time he is recovering from injury i .¹⁰ We also add a binary variable that indicates whether or not a player holds the citizenship of the country where the league is based. The tenure of player j with team k is used as a proxy for loyalty.¹¹ In addition, we include the number of matches played by team k in the relevant period below the cut-off value of the *workload* dummy variable to account for the workload in the reference period. Finally, a binary variable indicates whether or not team k participates in a European competition in season t . Table 3 presents the summary statistics of the main variables of interest.

In addition, we include a set of dummies for the starting month and the category c of the injury, ρ_c . Team fixed effects (ξ_k) account for unobserved heterogeneity across teams, such as differences in financial resources. This is important as these differences may affect the quality of medical treatments and hence recovery times. In other words, team fixed effects ensure that our estimation strategy relies on variations in our workload measure within teams over time. Season fixed effects (π_t) control for general developments in football, such as injury trends and improvements in medical care.

Table 4 presents the results for different specifications of the linear probability model defined in equation (3). Control variables for the player and the team are added in columns (4)–(6). We estimate that the probability of an early return increases by 11.5 p.p. in the simplest model and 6.3 p.p. in the full model specification if a team's workload in the relevant period is high. Since there exists a considerable number of players with only one observation in our sample, we will refrain from estimating the player fixed effects model and refer to model (6) as our preferred specification.¹² In addition, in column (7) we report coefficient estimates from an alternative specification where the dependent variable is the ratio between actual and expected recovery times in per cent. We find that on balance, injured players from teams exposed to a high workload return approximately 19.6 p.p. earlier than expected.

TABLE 3 Summary statistics

Variable	Mean	S.D.	Min	Max
Injury length (days)	17.55	18.51	0	168
Tenure (years)	2.93	2.68	0	24
NoIP (injured players for j 's position)	0.64	0.84	0	6
APP (available players for j 's position)	8.98	3.18	0	31
European competition (1 = yes, 0 = no)	0.42	—	0	1
Average injury length (reference period)	24.70	13.11	3.00	165.50
Number of injuries (team–season level, reference period)	16.98	11.03	0	61
Team market value (EUR mill.)	196.59	198.90	24.02	1135.90
Player market value (EUR mill.)	8.35	13.61	0.03	180
Win percentage (reference period)	0.40	0.19	0	0.95
Squad size	35.99	6.71	25	82
Nationality of league (1 = yes, 0 = no)	0.55	—	0	1
Workload in reference period	4.98	3.24	1	11

Notes: Descriptive statistics of main variables. $N = 6429$.

TABLE 4 Effect of treatment on probability of returning from injury earlier than the reference mean

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High workload ^a (1 = yes, 0 = no)	0.115*** (0.017)	0.112*** (0.018)	0.057*** (0.021)	0.049** (0.020)	0.056** (0.023)	0.063*** (0.024)	19.591*** (7.115)
Nationality ^b (1 = yes, 0 = no)				−0.002 (0.014)	−0.002 (0.014)	−0.005 (0.014)	−2.069 (4.084)
Tenure				0.001 (0.003)	0.001 (0.003)	0.000 (0.003)	−1.258 (0.791)
Player market value				0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.952*** (0.146)
Team value					0.000 (0.000)	0.000 (0.000)	0.027 (0.032)
Win percentage in reference period					−0.008 (0.112)	0.037 (0.118)	−4.515 (42.241)
Rank ratio					−0.005 (0.066)	0.011 (0.070)	−5.421 (22.004)
Roster size					−0.001 (0.001)	0.001 (0.002)	0.740 (0.573)
Competing in UEFA (1 = yes, 0 = no)					0.028 (0.041)	0.022 (0.041)	10.376 (13.056)
Number of workload games in reference					−0.007 (0.007)	−0.007 (0.007)	−2.571 (2.110)
Number of injuries at player's position						0.037*** (0.008)	12.941*** (2.080)
Available players in relevant position						−0.010*** (0.004)	−3.185*** (1.020)
Average injury duration in reference period						−0.001 (0.001)	−0.252 (0.163)
Number of injuries in reference period						0.002 (0.001)	0.695** (0.328)
Additional controls ^c	No	Yes	Yes	Yes	Yes	Yes	Yes
Start month fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Team fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
R ²	0.010	0.036	0.085	0.090	0.091	0.097	0.097

Notes

^a Binary variable indicating a high workload for a team–season as defined in equation (2).

^b Binary variable equal to 1 if the injured player holds citizenship of the country where the league is based, 0 otherwise.

^c Additional controls include age, position and injury category, as well as season dummies. In addition, we estimate team and injury start-month fixed effects.

N = 6429. Robust standard errors, clustered at the team–season level, are presented in parentheses. The dependent variable in columns (1)–(6) is equal to 1 if the injury is reported to be over before the mean duration in the reference sample, and 0 otherwise. Mean 0.496, standard deviation 0.500. The dependent variable in column (7) is equal to the ratio between actual and expected recovery times in per cent, defined by $-(actual_duration/expected_duration) - 1 \times 100$, with mean −33.223 and standard deviation 141.006.

*, **, *** indicate statistical significance at the 10%, 5%, 1% level, respectively.

Table 4 also shows that an early return is more likely to occur when there is a shortage of players to fill in player j 's position, meaning that it is less easy to replace the absent player. One additional player decreases the probability by 1 p.p., whereas another injured player substituting that position is associated with an increase in the probability by 3.7 p.p.

Considering our simple model presented in Section II, the empirical results confirm the theoretical predictions: presenteeism is more likely to occur when employers place more weight on the present than future payoffs, and when the employees can be less easily replaced. We will elaborate on this implication and further predictions derived from the theoretical model presented in Sections V and VI.

Although our *early_return* variable presents an intuitive way to measure presenteeism, a concern might be that it is too imprecise. Specifically, it does not account for the widely dispersed distribution of deviations from the median injury duration. However, it makes a difference if a player recovers one day earlier than expected from a first-degree strain or an Achilles tendon rupture. We therefore propose a measure of *relative injury duration* for each injury category c , defined by

$$\Delta_{i,c} = \frac{\text{duration}_{i,c}}{\text{median_duration}_c} \times 100,$$

where $\text{duration}_{i,c}$ is the duration (in days) of an injury i of category c suffered in the relevant period (January to June) of season t , and median_duration_c is the expected length of injuries for that category calculated from the reference period. We prefer using the median over the mean of injury duration, as outliers will certainly affect the precision of our presenteeism measure.

To account for potential non-linear effects, we use interval dummies for the relative injury duration:

$$\text{diff}_s = \begin{cases} 1, & \text{if } \Delta_{i,c} \leq s/100, \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

for each $s \in \{5, 10, 15, \dots, 90, 95, 100\}$. Thus, for instance, $\text{diff}_5 = 1$ implies that the recovery time for injury i was at least 5% shorter than the mean in the reference category c .

Next, using this vector of dependent variables, we re-estimate model (3) for each interval, which yields 20 estimates for β_1 . The results are illustrated in Figure 5. We find that a high workload in the relevant period significantly increases (at the 5% level) the probability of an early return for all thresholds between 10% and 70%. The point estimate is approximately 6–7 p.p. for intervals below the 65% threshold, which decreases gradually thereafter. For very high values of Δ , which indicates a reduction in the recovery time by 75% or more, $\hat{\beta}_1$ is close to 0 and not statistically significant.

We conclude that the players belonging to teams ‘under pressure’, who were injured in the second half of a season, have a higher tendency to return earlier than expected. As this relation also holds true for very early returns for which a full recovery seems unlikely, we present our results as evidence for presenteeism.

In the preceding analysis, we used a cut-off value of > 4 games (which is the fourth quartile of the overall distribution of extra games) to define a high workload. To test the robustness of this definition, we use alternative cut-offs of > 3 and > 5 . The results are presented in Figure A3 of the Appendix. Overall, the main effects (as presented in Figure 5) are confirmed qualitatively and quantitatively. We conclude that the exact cut-off at > 4 games is not driving our results.

Another issue, mentioned earlier in this section, is that the cut-off criterion for defining a high workload is not fully predetermined, because it includes the cup matches held between February and early April. Consequently, some matches result from prior success in the first two rounds of

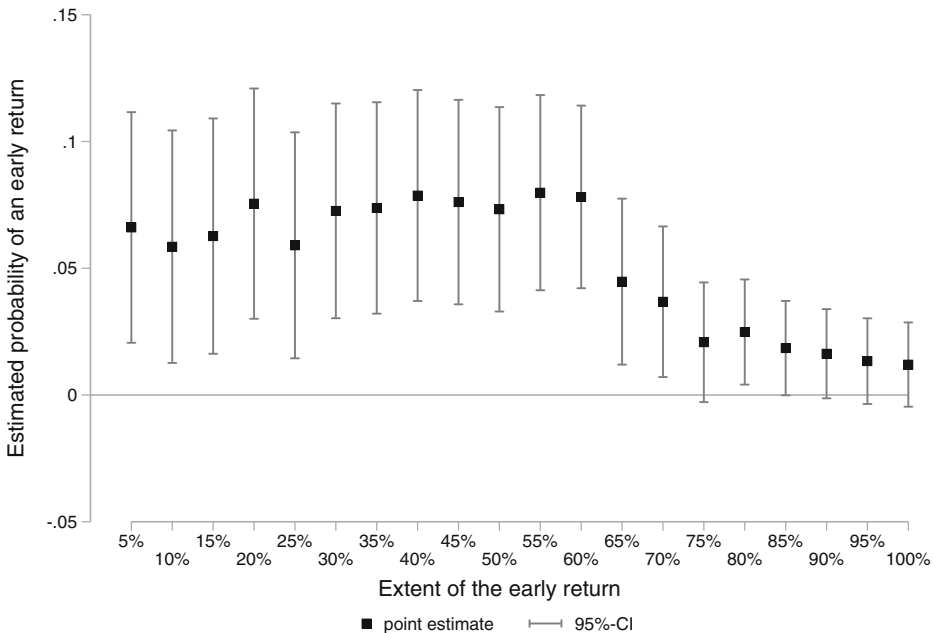


FIGURE 5 High workload affects the probability of an earlier return from a sick absence due to injury. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 6429$.

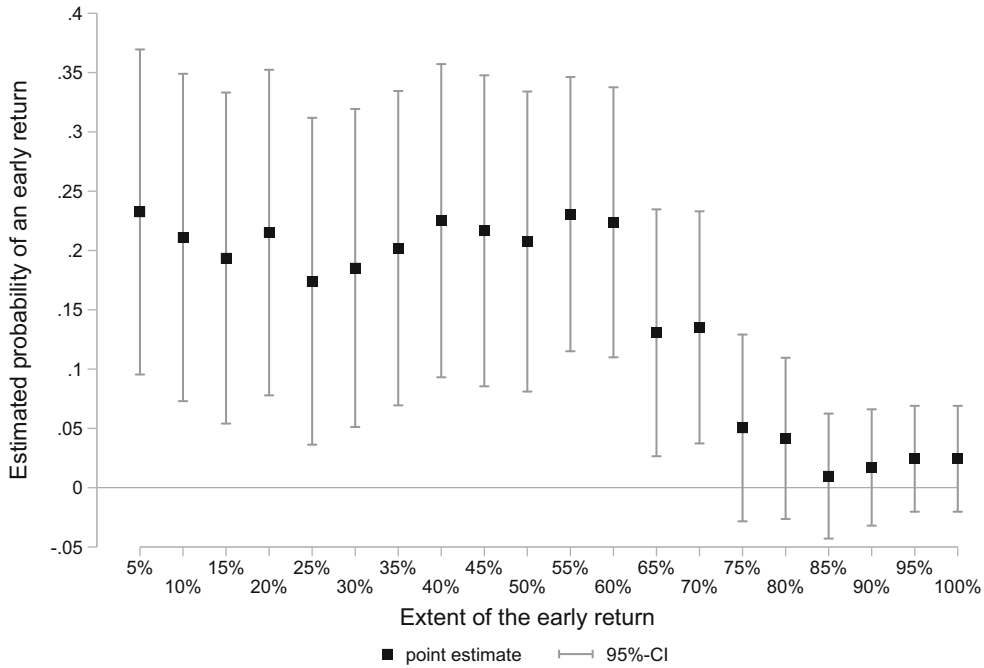


FIGURE 6 High workload affects the probability of an earlier return from a sick absence due to injury: reduced sample. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. Only the injuries that occurred in May or later are included. $N = 761$.

the play-off stage. To ensure that our main findings are not affected by an endogeneity bias, we check the robustness of our findings with a sample restricted to the injuries that occurred in May or later. Figure 6 suggests that for this sample, the presenteeism effects are even stronger. This is not surprising, because the incentives for an early return are highest in the ‘crunch time’ of a season when it is of foremost importance to succeed in national and international competition. Referring to the conceptual framework, we expect the discount factor δ to be the highest for these periods. Moreover, the sample includes (by definition) shorter and, on average, less severe injuries from which the players recovered before July.

5 | FACTORS THAT AFFECT THE RELATIONSHIP BETWEEN WORKLOAD AND PRESENTEEISM

Our main results presented in the previous section indicate that a high workload causes presenteeism behaviour in professional football. However, it is reasonable to assume that the nexus between workload and presenteeism is affected by employee (player) and employer (club) characteristics. For instance, our theoretical model predicts that it is more likely that presenteeism occurs among the top players and players less vulnerable to injuries. In the following analysis, we focus on these characteristics.

5.1 | Individual and team (absolute) ability

We use information on a player’s market value in season t as a proxy variable for individual ability.¹³ These market values represent the expert estimates of possible transfer fees based on past performances and the market situation. As a general rule of thumb, better players have higher market values.¹⁴ For all the observations in our sample, the average market value of the observed player at the time of injury is 8.34 million euros, with standard deviation 13.58.

There are two countervailing forces that could shape how the importance or value of a player for a team affects presenteeism behaviour. From a team’s perspective, the most able players are also the most important and productive ones, who will be missed the most in case of their absence. This speaks in favour of an early return (see Section II). On the contrary, players with high market values are assets to their teams and require protection. That is, in a worst-case scenario, an inadequate recovery time may result in medium- or long-term negative effects due to chronic ailments. These adverse health effects are a threat not only to a player’s availability and performance, but also to possible future transfer revenues.

In the empirical analysis, we use a median split to categorize (absolute) player values as either high or low. Furthermore, by analogy with equation (4), 10 interval dummies are used as dependent variables in our regression model ($s \in \{5, \dots, 50\}$), which correspond to a 5–50% shorter recovery time compared to the reference group. The results are illustrated in Figure 7. Although we can confirm the significant positive impact of workload on the probability of returning earlier from injury, the point estimates for the high and low market value groups show a negligible difference. We therefore conclude that absolute market value is not a driving factor of presenteeism in our setting.

In addition to the individual productivity of a player, the overall ability of a team may also have an impact on presenteeism. For instance, it might be the case that a high-budget team with a large number of top players has a better ability to replace an injured player. Therefore a median split is used to categorize the teams into low and high categories of (absolute) market values.¹⁵ Our findings are illustrated in Figure A4 of the Appendix. It shows that the estimated coefficients are very similar across groups, suggesting that presenteeism does not vary by team market values.

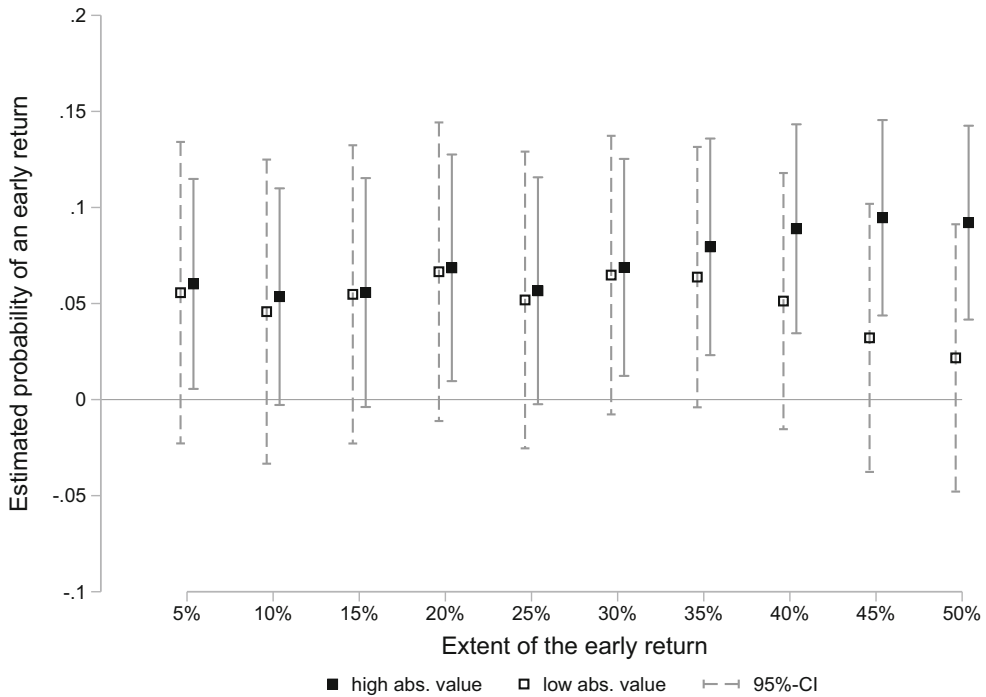


FIGURE 7 Effect heterogeneity: absolute market value of a player. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 6429$. Solid (hollow) squares represent estimates for the group of players with a high (low) absolute market value. Players were categorized using a median split.

5.2 | Relative market values

The foregoing analyses suggest that the absolute market values—for the individual players or for their teams—do not significantly impact the workload effect on presenteeism behaviour. Absolute market values, however, might not tell the whole story. The theoretical considerations in Section II suggest that it is rather the relative importance of an absent employee relative to her colleagues that causes the employer to demand a return to work while the employee is still recovering.

In our setting, we define the relative ability of player j belonging to team k in season t as

$$relative_ability_{j,k,t} = \frac{market_value_{j,k,t}}{team_market_value_{k,t}}. \quad (5)$$

Further, the players are divided at the median of $relative_ability_{j,k,t}$. We then run the same regressions as in the previous section for the players who are above or below the median of relative ability. The results are presented in Figure 8. We find a significant and sizeable workload effect on the probability of an early return for players in the high-ability group, but not for those in the low-ability group. For instance, players with a high relative ability are, on average, about 10–14 p.p. more likely to reduce the expected recovery time by half when their team is under pressure. On the contrary, there is no early return for players from the low-ability group, even when their teams are exposed to a high workload, suggesting that these players can be replaced more easily.

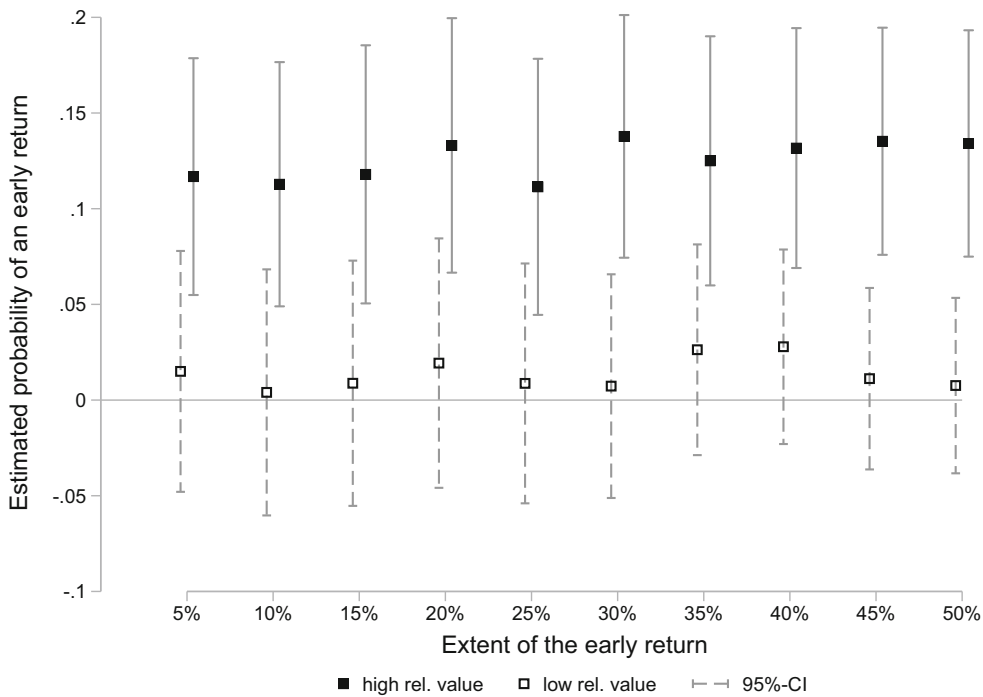


FIGURE 8 Effect heterogeneity: relative market value. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 6429$. Solid (hollow) squares represent estimates for the group of players with a high (low) relative market value. Players are divided at the median of *relative_ability_{j,k,t}* defined by equation (5), which is 0.031.

As a further refinement, we follow the idea that professional football teams, similar to most conventional enterprises, have employees in key positions. In the case of football, these key players are starters and represent the first line of employees. To investigate whether the players with or without such a prominent role exhibit a similar relationship between presenteeism and workload, we define an indicator for key players. In particular, we sort players according to their market values within their teams. Then we define ‘important players’ as those players whose ranking position is equal to or below the median ranking position. This median ranking position is 8 for the overall sample. Figure A5 in the Appendix confirms the prior results: we estimate a significant positive effect of a high workload on the probability of returning early from an injury for important players. In contrast, for the group of less important players, the estimated β_1 is not different from zero. The point estimates range from 0.04 to 0.

Similar to other labour markets, the division of labour and specialization is also present in football. Although the degree of specialization can vary across positions and players (for a recent study on this subject, see Kempa 2022), we account for this issue by introducing team-specific rankings based on market values for the four main player positions. A player is then defined as a *key player* if he is the top goalkeeper or forward player, or if he is among the top four defenders or top five midfielders. Apart from the great importance for their teams, the internal competition and the need to protect their status as regular players may work as additional incentives for these players. Figure 9 indicates that even when we use *key player* as a moderator in our model, the estimates are quite similar to the foregoing specifications.

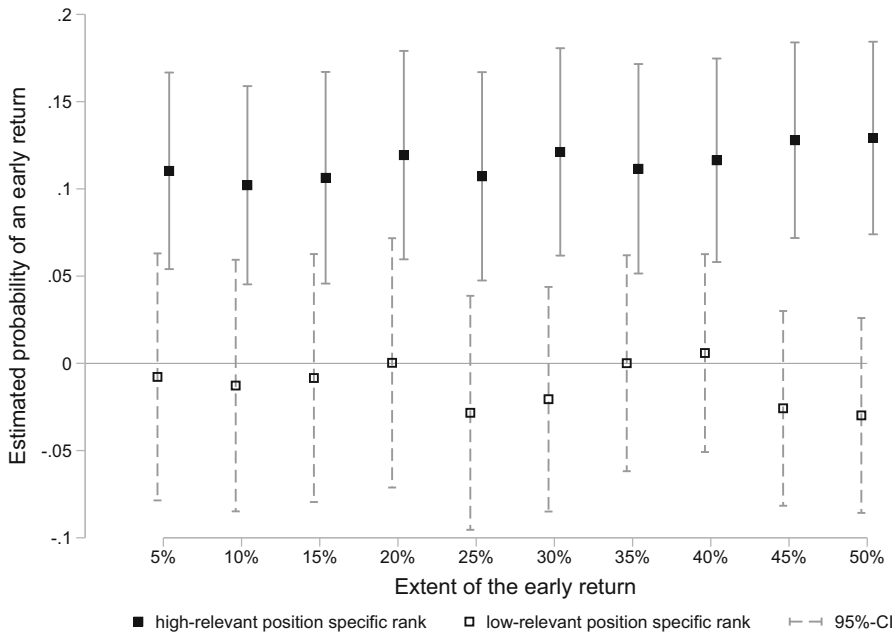


FIGURE 9 Effect heterogeneity: key players. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 6429$. Solid squares represent estimates for the group of *key players*, that is, the top players per position identified by their market values. Hollow squares represent estimates for the group of non-key players.

5.3 | Injuries: personal history and the overall injury level

Following the conceptual framework, in this subsection we test the hypothesis that presenteeism is more likely to occur when the risk of absence in the next period is sufficiently low. We operationalize this idea by using a player's injury history to proxy his 'vulnerability'. Therefore the sample is split along the median of the player-specific 'number of injuries per number of seasons' ratio. For example, this ratio would be 2/5 for an individual who had two injuries in five seasons. We then re-estimate model (3) with interval dummies defined by equation (4).

Figure 10 presents the estimates for β_1 . Consistent with our expectations from theory, it shows that the treatment effect of a high workload is restricted to the players who are less vulnerable to injuries. Therefore players with several prior injury issues do not return to their teams before the median time of healing, even when the team is under pressure. We conclude that for vulnerable players, the benefits of an early return do not outweigh the risks involved.

Figure A6 in the Appendix complements our prior findings regarding players injured at the same time. While our main results in Table 4 provide evidence that the number of additional injuries at the focused player's position increases the likelihood of an early return, the figure demonstrates that the high workload effect is equivalent for teams with a low and high overall number of injuries.

Moreover, studies such as Godøy and Dale-Olsen (2018) demonstrate positive peer effects in absenteeism. In our setting, this would suggest that there might be spillover effects from the teammates who returned early from their absence due to injuries, and this can exert peer pressure on the focused player to do likewise. However, we do not observe such behaviour in our data. That is, (unreported) results indicate that the number of previous early returns in a team per season does not affect subsequent presenteeism behaviour. This not only casts doubt on the role

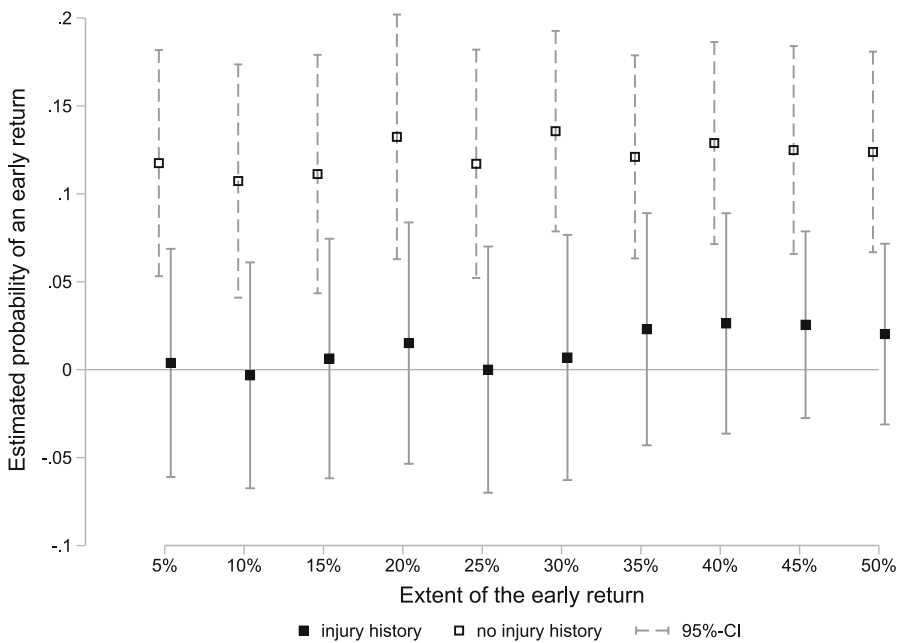


FIGURE 10 Effect heterogeneity: vulnerability. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 6429$. Solid (hollow) squares represent estimates for the group of players with a high (low) number of past injuries. Players are divided at the median of the injuries per season ratio, which is 1.

of peer effects in our setting, but also suggests that differences among coaches in handling injured players are not driving our results.

5.4 | Age, tenure and contract length

Finally, we present complementary estimates of the model in equation (3) with age, tenure and contract length as moderators.

First, splitting the sample into young and old players (median age 27) suggests that age alone is not the predictor of an early return when the workload is high (see Figure 11). As we have seen in the previous subsection, this might be because it does not capture vulnerability.

Second, Figure 12 indicates that the treatment effect is virtually the same for players with long and short tenures. This suggests that we should not overemphasize the importance of loyalty and ‘contract compliance’ in the present context.

Third, prior research suggests that the actual contract length can affect presenteeism and absenteeism behaviour, especially in professional sports. For instance, Ngo and Roberts (2021) find that contract status has an impact on absenteeism in the NBA (excluding top players). Despite the fact that the ‘free agent’ status is less common in European football than in the professional North American leagues, and players with existing contracts are typically transferred for a fee, there might be incentives for the clubs to ‘protect their assets’ near the end of the contract when the probability of a transfer increases. Similarly, it could be the case that players do not want to risk upcoming transfers. Unfortunately, data on the characteristics of player contracts are not publicly available in European football. However, we do have information on contract lengths for a subsample of players.¹⁶ For these individuals, we find that workload-induced presenteeism

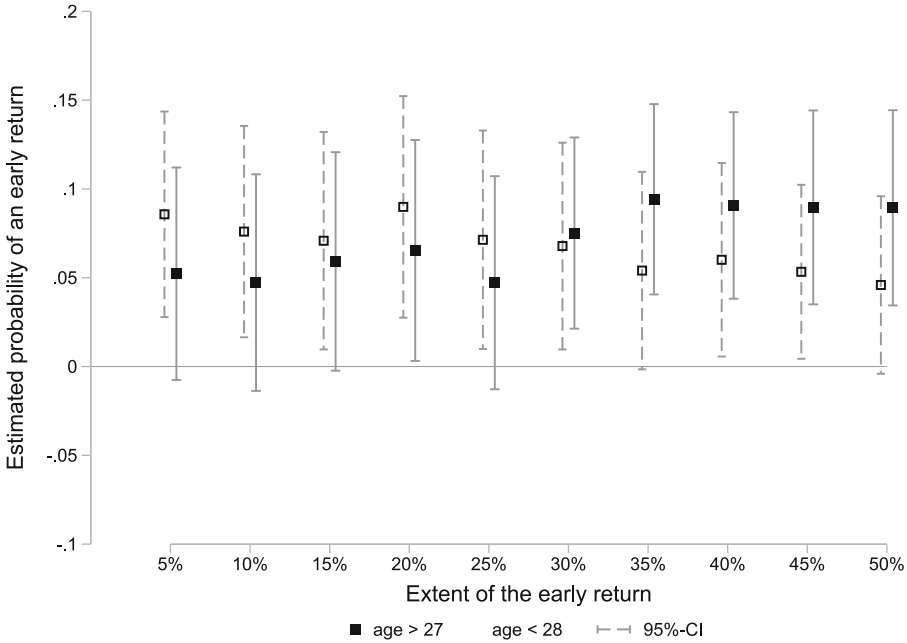


FIGURE 11 Effect heterogeneity: age. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 6429$. Solid (hollow) squares represent estimates for the group of old (young) players. Players were categorized using a median split (median age 27).

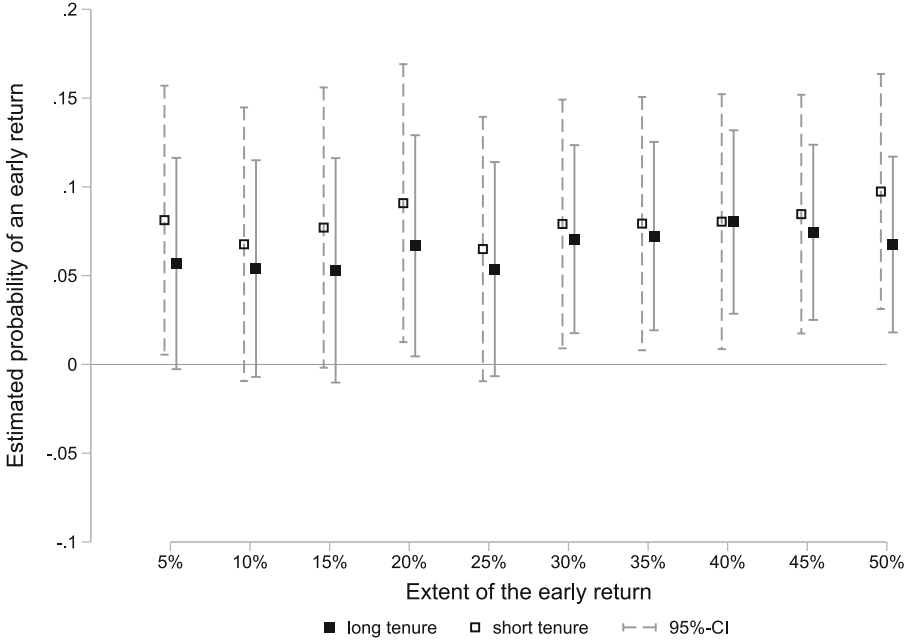


FIGURE 12 Effect heterogeneity: tenure. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 6429$. Solid (hollow) squares represent estimates for the group of long-tenured (short-tenured) players. Players were categorized as *short* with 0 or 1 seasons with the observed team, and categorized as *long* with more than 1 year with the team.

behaviour is present in the group of players with long-term contracts, but not in the group of players with short-term contracts (see Figure A1 in the Appendix). This may point to a conflict between team interests and players' career concerns. However, due to the limited and possibly selective sample, this result should be treated with caution.

6 | COST AND CONSEQUENCES OF AN EARLIER RETURN

Our main results suggest that there is presenteeism behaviour in professional football caused by additional workload. Consistent with theoretical considerations, we find that the effect is driven by those players who are the most important for their teams. In addition to a high workload, we find the number of injuries among co-players at the same position to be another driver of an early return.

While we document strong incentives for presenteeism when the teams are under pressure, we already discussed its negative side in Section II. An early return can be inefficient when teams are present-biased or underestimate its expected health consequences. We also expect presenteeism to be welfare-reducing when the health costs borne by the player were not, or only inadequately, taken into consideration.

Obviously, it is beyond our data to measure the pain associated with an incomplete healing and a possible long-term chronic disability. However, we can estimate the risk of subsequent injuries, including re-injuries and new injuries. If presenteeism significantly shortens the time until the next absence, then these costs are borne not only by the players themselves (e.g. in the form of reduced earnings and career opportunities), but also by the team that needs to compensate for the absent players.

Our empirical approach is to use the full sample of Section IV and link injuries in the relevant period to subsequent injuries. Since our data do not allow us to identify correctly the censoring for players with only one observation, we exclude these players. Hence the results will indicate the consequences of an early return conditional on re-injuring. The average player in the final sample ($N = 4996$) suffers 5.05 subsequent injuries. The mean time between events is 189.16 days (standard deviation 260.591, and maximum 3135), and the mean duration of these injuries is 27.08 days (standard deviation 43.59).

We estimate a variant of model (3) where the dependent variable is the time spent in good health after a prior injury, that is, the time between the actual and the subsequent injury. Figure 13 illustrates the results. We find that the time spent in good health for players who returned at least 5% earlier than expected is reduced by approximately 16 days. The effect is stable across intervals, suggesting that the negative health consequences are relevant for all grades of presenteeism.¹⁷

Table 5 presents estimates from a model including three binary variables equal to 1 if the recovery time was 5–25%, 30–50% or more than 50% shorter than expected, and 0 otherwise (reference category: no early return). Results from linear regressions show that an early return reduces the time until the next absence: by approximately 24 days for a low degree, 17 days for a medium degree, and 18 days for a high degree of presenteeism (column (3)).

As we use a duration measure as the dependent variable, we also specify a Cox proportional hazard model to check the robustness of our estimates derived from OLS. The results are presented in column (4) of Table 5. We find that an early return increases the risk of subsequent injury. Coding long subsequent injury times (i.e. longer than 365 days) as censored does not change our result qualitatively.

Taken together, the analysis suggests that the negative effect of workload-induced presenteeism on a worker's health is remarkably strong in our sample. An approximately one-month reduction in the six-month period in good health raises the question of whether actors in this labour market deal adequately with the issue.

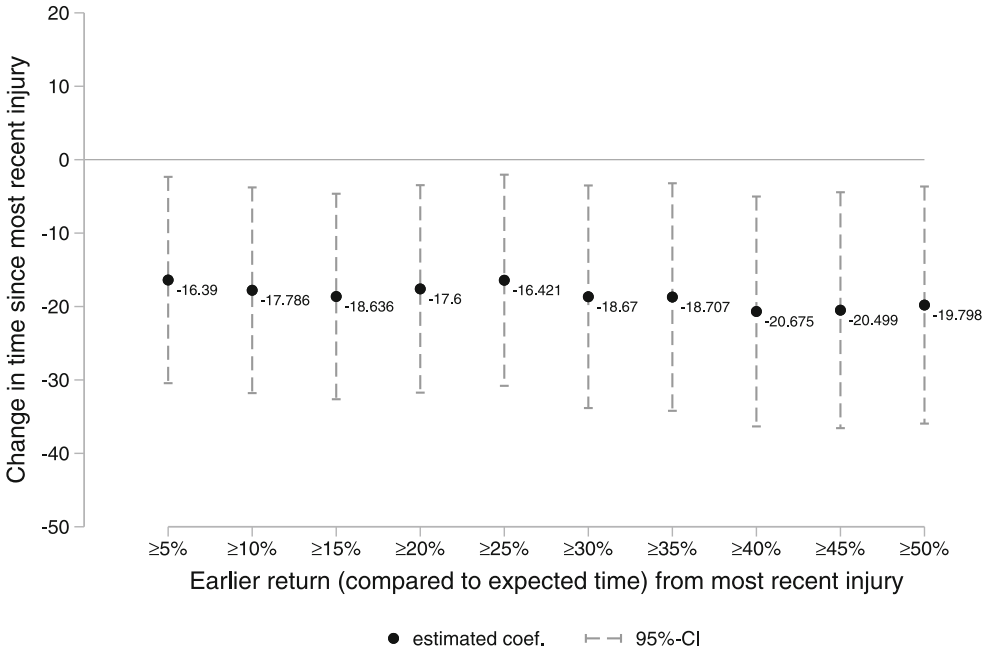


FIGURE 13 Earlier returns and time before next injury. *Notes:* Each point represents an estimate from a separate regression, using the number of days between two injuries as the dependent variable. The variable of interest is equal to 1 if the preceding injury time ended at least 5%, 10%, 15%, ..., 50% earlier than expected, 0 if no early return was observed. The number of observations decreases along the x -axis for each estimate. The full sample of 4996 injuries is used for the first estimate of a 5% early return. Standard errors are clustered at the player level.

TABLE 5 Effect of an earlier return from previous injury on time to follow-up injury

	OLS			Cox model
	(1)	(2)	(3)	
Return 1 (1 if return \in [5%–25%], 0 otherwise)	–12.111 (11.298)	–27.103** (11.105)	–24.808** (11.182)	0.124*** (0.046)
Return 2 (1 if return \in [30%–50%], 0 otherwise)	–16.489* (9.664)	–20.286* (10.417)	–17.347* (10.460)	0.085** (0.042)
Return 3 (1 if return > 50%, 0 otherwise)	–22.191*** (8.045)	–23.185*** (8.578)	–17.995** (8.625)	0.126*** (0.040)
Additional binary controls	Yes	Yes	Yes	Yes
Team last injury dummies	Yes	Yes	Yes	Yes
Player dummies	No	Yes	Yes	No
End-month previous injury	No	No	Yes	No
R^2	0.207	0.487	0.493	—
N		4996		4996

Notes: Robust standard errors clustered at the player level are presented in parentheses. The dependent variable in columns (1)–(3) observes to the number of days between the observed injury and the end of the preceding injury. Additional binary controls include season and age dummies as well as dummies indicating the severity category of the preceding injury. Column (4) reports the coefficients derived from a Cox proportional hazard survival model, stratified by the team to which the player was signed when the preceding injury ended. *, **, *** indicate statistical significance at the 10%, 5%, 1% level, respectively.

7 | ROBUSTNESS

7.1 | Injury categories and infectious diseases

In our analysis, injuries were divided into 20 categories to ensure a balanced number of observations across injury types. To test the robustness of our results, we repeat the analysis for 9, 15, 24 and 30 injury categories; see Figure A2 in the Appendix. We find that our main results from Section IV hold.¹⁸

Another issue is that there are medical diagnoses in our sample that refer to infectious diseases that can spread in teams. The number of observations, however, is small (713 in total), and excluding these observations from our sample does not affect our results.

7.2 | Competitions for national teams

The two most prestigious competitions for national teams, the FIFA World Cup and the UEFA European Championship, are important for players and teams. These tournaments not only shorten the recovery time for players but also are also regarded as career milestones. Miklós-Thal and Ullrich (2016), for instance, show that players adapt their league performances in the run-up to such an event. Hence it might be the case that the decision for injured players to return to play could be affected. Although we use season fixed effects to account for heterogeneity across years, we want to test additionally whether our main results presented in Figure 5 are driven by international tournaments. For this reason, Figure A7 in the Appendix illustrates the estimates based on a sample restricted to season without international tournaments. It shows that our main findings can be confirmed, indicating that international tournaments are not the driving force behind our estimates.

7.3 | Player fixed effects

We also run an alternative specification to our empirical model defined by equation (3), including player fixed effects. The results are presented in Figure A8 of the Appendix. Due to the high number of players with only one injury in our data, the effects are estimated with less precision. Our main results, however, are confirmed.

7.4 | Excluding national cup games

Finally, our definition of *high_workload* pertains to national and European cup matches. Since national cups differ in terms of competition formats across leagues, Figure A9 in the Appendix presents estimates from an alternative definition using European competition games only. The effects are estimated less precisely and are smaller in size. However, as our main results are confirmed, we conclude that the specifics of the national cup competitions are not driving our findings.

8 | CONCLUSION AND DISCUSSION

This study contributes to the small but growing literature of the analysis of work-while-sick behaviour. Using data collected from professional football, we investigate the effect of an increased team workload on the propensity for players to return from absence due to injury or sickness earlier than expected. We propose a novel measure of presenteeism that allows us to predict the expected time of absence and to measure precisely the extent of presenteeism. Our

estimates suggest that the probability of returning earlier than expected from an injury increases by 6.3 p.p. when the team is exposed to a high workload.

Furthermore, the setting allows us to study the potential moderators of the nexus between workload and presenteeism behaviour. As predicted by our simple model, we find that the effect is driven by players with a higher relative productivity. These players are more important to their teams and therefore harder to replace. Also in line with our expectations, we find that the players who are more vulnerable to injuries seem to be protected against an early comeback to competition. The external validity of our findings is supported by the fact that sport-specific factors are not the drivers of our results: neither the age nor the tenure of players—two characteristics in professional sports that typically differ from standard work environments—appears to affect presenteeism behaviour.

These observations may suggest that teams balance the pros and cons of an early return in a way such that presenteeism occurs when the ratio of benefits to costs is the greatest. However, we document that this behaviour has serious consequences for players' health. Specifically, it shows that the time in good health between two injury events is reduced by about one month for players who had an early comeback due to a high team workload. We find this effect to be large given that the mean time between two absences, on average, is about six months. This may indicate inefficiencies related to presenteeism. First, our findings suggest that teams have a general idea of how to deal with the return of injured players in troubled times, but may underestimate the actual health risks that exist even for the less vulnerable. Second, if teams are aware of the risks, then the massive reduction in the amount of time for which the player is available to the team caused by early comebacks in high-workload periods raises concerns about a structural underinvestment in players' health.

Finally, the question arises why workers in general, and football players in particular, agree to presenteeism when there are substantial health consequences accompanied by income losses. Possible explanations are related to career concerns and job market signalling (Crichton *et al.* 2011; Markussen 2012; Ngo and Roberts 2021), reciprocal behaviour (e.g. Charness and Kuhn 2011), direct pressure by the employer, and the fear of losing income due to the low levels of sick payments or even unemployment (e.g. Olsson 2009; Hirsch *et al.* 2017; Callison and Pesko 2022). However, there might be other and more indirect reasons for a (tacit) agreement on presenteeism. These are work attitudes and social norms (e.g. Godøy and Dale-Olsen 2018), a (perceived) lack of replacement (e.g. Lohaus and Habermann 2019), and peer pressure in the sense that workers want to prevent a workload increase for their colleagues (Skåtun and Skåtun 2004; Barmby *et al.* 2016).

ACKNOWLEDGMENTS

We thank Christian Deutscher and numerous seminar and conference participants (SOLE 2022, EALE 2022, ESPE 2022, ROSES 2021, ESEA 2021) for helpful discussions and comments. In addition, we thank Erik Scheil, Houda Ben Said and Julian Herwig for their help with the data collection. Open Access funding enabled and organized by Projekt DEAL.

NOTES

¹ In professional basketball, Michael Jordan's 'flu game' as well as Dirk Nowitzki's 'fever game' in the 1997 and 2011 NBA Finals still receive attention from the fans and the media (see, for example, Sherrington 2021).

² We do not model explicitly the way replacements work. For instance, when worker $i + 1$ is doing the job of the absent worker i , this could mean that the workload for $i + 1$ increases or that the worker is missing elsewhere. However, the crucial element for our conceptual framework is that there is heterogeneity in 'substitutability'. See Pauly *et al.* (2002) for a more comprehensive model of absent team members, and Skåtun and Skåtun (2004) for an analysis of workload shifting among employees.

³ See <https://www.transfermarkt.de>, accessed 2 January 2023.

⁴ The recovery time ends when the player is available again and part of the team. Note that this does not necessarily imply that players are fielded in the next match.

- ⁵ Note that we restrict to data from the seasons 2010/11 to 2018/19 for our analysis. Additionally, observations from the first part of the 2019/20 season were used to calculate expected recovery times only. Observations from the second part of that season were excluded from the sample due to the COVID-19 pandemic. Consult the Appendix for a detailed description of the data preparation.
- ⁶ Table A1 in the Appendix specifies the medical diagnoses.
- ⁷ The number of categories may appear arbitrary, yet it ensures a balanced number of observations across injury types. Nevertheless, we conduct robustness checks with different numbers of categories in Section VII.
- ⁸ Qualification for the knock-out stage of a European cup competition results from performances in the reference period and is therefore independent of the injuries that occurred during the relevant period.
- ⁹ The relative league position is defined as the ranking position divided by the number of teams. The first quartile of the market value distribution serves as the reference category.
- ¹⁰ The four major positions used in the analysis are goalkeeper, defender, midfielder and forward.
- ¹¹ Galizzi and Boden (2003), for instance, show that the absence of workers with job-related injuries is affected by tenure. Arnold (2016) reports that the number of sickness presenteeism days increases with tenure, whereas individuals show less presenteeism behaviour after a job change within the first year.
- ¹² In Section VII, we present the results of a model with player fixed effects. Our main results still hold.
- ¹³ Market value data were collected from <https://www.transfermarkt.de>.
- ¹⁴ Note that the *transfermarkt* market values have been used in prior studies such as Krumer and Lechner (2018). Since market values are also affected by the age of a player, which in turn may not play a role in his actual ability, we conduct the analysis again with age-adjusted values. That is, we use the residuals $\mu_{j,k,t}$ derived from estimating the model $\text{market_value}_{j,k,t} = \alpha_0 + \alpha'_1 \text{AGE} + \alpha'_2 \mathbf{X}_{j,k,t} + \mu_{j,k,t}$, where *AGE* is a vector of age interval dummies, and $\mathbf{X}_{j,k,t}$ is a vector of player–team characteristics, including position, league and season. The results remain unchanged and are available on request.
- ¹⁵ We prefer the median over the mean, because only a few outliers at the top end of the distribution can affect mean market values disproportionately.
- ¹⁶ Contract status data are available on the *transfermarkt* website for a selection of players only. The subsample includes 1038 observations, which corresponds to about 16% of the total sample. The median of the remaining term is 3 years.
- ¹⁷ Note that repeating this exercise separately for the groups of high- and low-ability players as introduced in Section 5.2 shows that the former does not have more time in good health after high-workload-induced early comebacks. From this, we conclude that the top players do not benefit from better medical treatments.
- ¹⁸ In addition, results also hold for a sample restricted to severity categories with at least 250 observed injuries per category. This exercise ensures that injury types with few observations do not bias our results.

ORCID

Hendrik Sonnabend  <https://orcid.org/0000-0003-0667-8595>

REFERENCES

- Arnold, D. (2016). Determinants of the annual duration of sickness presenteeism: empirical evidence from European data. *Labour*, **30**(2), 198–212.
- Barmby, T., Sessions, J. G. and Zangelidis, A. (2016). Looking after number two? Competition, cooperation and workplace interaction. *Journal of Economic Behavior & Organization*, **131**, 166–82.
- Bergström, G., Bodin, L., Hagberg, J., Aronsson, G. and Josephson, M. (2009). Sickness presenteeism today, sickness absenteeism tomorrow? A prospective study on sickness presenteeism and future sickness absenteeism. *Journal of Occupational and Environmental Medicine*, **51**(6), 629–38.
- Callison, K. and Pesko, M. F. (2022). The effect of paid sick leave mandates on coverage, work absences, and presenteeism. *Journal of Human Resources*, **57**(4), 1178–208.
- Chan, T. C. and Fearing, D. (2019). Process flexibility in baseball: the value of positional flexibility. *Management Science*, **65**(4), 1642–66.
- Charness, G. and Kuhn, P. (2011). Lab labor: what can labor economists learn from the lab? In O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 4. Amsterdam: Elsevier, pp. 229–330.
- Chen, Y., Buggy, C. and Kelly, S. (2019). Winning at all costs: a review of risk-taking behaviour and sporting injury from an occupational safety and health perspective. *Sports Medicine—Open*, **5**(1), 1–21.
- Crichton, S., Stillman, S. and Hyslop, D. (2011). Returning to work from injury: longitudinal evidence on employment and earnings. *Industrial and Labor Relations Review*, **64**(4), 765–85.
- Galizzi, M. and Boden, L. I. (2003). The return to work of injured workers: evidence from matched unemployment insurance and workers' compensation data. *Labour Economics*, **10**(3), 311–37.
- Godøy, A. and Dale-Olsen, H. (2018). Spillovers from gatekeeping—peer effects in absenteeism. *Journal of Public Economics*, **167**, 190–204.

- Gregory-Smith, I. (2021). Wages and labor productivity: evidence from injuries in the national football league. *Economic Inquiry*, **59**(2), 829–47.
- Guenzel, M. and Malmendier, U. (2021). Behavioral corporate finance: the life cycle of a CEO career. *Oxford Research Encyclopedia of Economics and Finance*; available online at <https://doi.org/10.1093/acrefore/9780190625979.013.457> (accessed 2 January 2023).
- Hansen, C. D. and Andersen, J. H. (2009). Sick at work—a risk factor for long-term sickness absence at a later date? *Journal of Epidemiology & Community Health*, **63**(5), 397–402.
- Hirsch, B., Lechmann, D. S. and Schnabel, C. (2017). Coming to work while sick: an economic theory of presenteeism with an application to German data. *Oxford Economic Papers*, **69**(4), 1010–31.
- Kempa, K. (2022). Task-specific human capital and returns to specialization: evidence from association football. *Oxford Economic Papers*, **74**(1), 136–54.
- Kim, H. T. and Nguyen, Q. (2022). An exploration on the nexus between managers' present bias and corporate investment. *Applied Economics Letters*, **29**(2), 150–4.
- Kivimäki, M., Head, J., Ferrie, J. E., Hemingway, H., Shipley, M. J., Vahtera, J. and Marmot, M. G. (2005). Working while ill as a risk factor for serious coronary events: the Whitehall II study. *American Journal of Public Health*, **95**(1), 98–102.
- Krumer, A. and Lechner, M. (2018). Midweek effect on soccer performance: evidence from the German Bundesliga. *Economic Inquiry*, **56**(1), 193–207.
- Lohaus, D. and Habermann, W. (2019). Presenteeism: a review and research directions. *Human Resource Management Review*, **29**(1), 43–58.
- Markussen, S. (2012). The individual cost of sick leave. *Journal of Population Economics*, **25**(4), 1287–306.
- Mayer, J. and Thiel, A. (2018). Presenteeism in the elite sports workplace: the willingness to compete hurt among German elite handball and track and field athletes. *International Review for the Sociology of Sport*, **53**(1), 49–68.
- Miklós-Thal, J. and Ullrich, H. (2016). Career prospects and effort incentives: evidence from professional soccer. *Management Science*, **62**(6), 1645–67.
- Miraglia, M. and Johns, G. (2016). Going to work ill: a meta-analysis of the correlates of presenteeism and a dual-path model. *Journal of Occupational Health Psychology*, **21**(3), 261–83.
- Ngo, P. T. H. and Roberts, S. (2021). There is no 'I' in team: career concerns, risk-taking incentives, and team outcomes. *Journal of Economics and Management Strategy*, **30**(1), 122–38.
- Olsson, M. (2009). Employment protection and sickness absence. *Labour Economics*, **16**(2), 208–14.
- Pauly, M. V., Nicholson, S., Xu, J., Polsky, D., Danzon, P. M., Murray, J. F. and Berger, M. L. (2002). A general model of the impact of absenteeism on employers and employees. *Health Economics*, **11**(3), 221–31.
- Pichler, S., Wen, K. and Ziebarth, N. R. (2020). COVID-19 emergency sick leave has helped flatten the curve in the United States: study examines the impact of emergency sick leave on the spread of COVID-19. *Health Affairs*, **39**(12), 2197–204.
- Roderick, M., Waddington, I. and Parker, G. (2000). Playing hurt: managing injuries in English professional football. *International Review for the Sociology of Sport*, **35**(2), 165–80.
- Schultz, A. B. and Edington, D. W. (2007). Employee health and presenteeism: a systematic review. *Journal of Occupational Rehabilitation*, **17**(3), 547–79.
- Sherrington, K. (2021). Dirk Nowitzki's 'Fever Game': Mavs win Game 4 to even series vs. Heat at 2–2. *Dallas Morning News*, 10 June.
- Skagen, K. and Collins, A. M. (2016). The consequences of sickness presenteeism on health and wellbeing over time: a systematic review. *Social Science & Medicine*, **161**, 169–77.
- Skåtun, D. R. and Skåtun, J. D. (2004). The impact of turnout on turning up: the complementarity of attendance among co-workers. *Journal of Economics*, **83**(3), 225–42.

How to cite this article: Lackner, M. and Sonnabend, H. (2023). Presenteeism when employers are under pressure: evidence from a high-stakes environment. *Economica*, **90**(358), 477–507. <https://doi.org/10.1111/ecca.12461>

APPENDIX

Data description

We use all injuries entries available at <https://www.transfermarkt.de> for the seasons 2010/11 to 2019/20. Due to the fact that the second half of the 2019/20 season was influenced strongly by

TABLE A1 Medical injury codes

Code	Diagnosis	Code	Diagnosis
1	Achilles tendon irritation	75	Knee contusion
2	Abdominal discomfort	76	Knee problems
3	Abdominal muscle strain	77	Laceration wound
4	Adductor problems	78	Lateral ligament injury
5	Angina	79	Leg injury
6	Ankle injury	80	Ligament inflammation in the knee
7	Ankle problems	81	Ligament injury
8	Ankle sprain	82	Ligament stretch
9	Appendix surgery	83	Lumbago
10	Arm injury	84	Lumbar vertebrae problems
11	Back contusion	85	Meniscus injury
12	Back injury	86	Meniscus tear grade 2
13	Back problems	87	Metacarpal fracture
14	'Locked' vertebrae	88	Metatarsal contusion
15	Blow	89	Minor blemish
16	Bronchitis	90	Muscle bundle tear
17	Bruise	91	Muscle contusion
18	Bruise on the ankle	92	Muscle fatigue
19	Calf hardening	93	Muscle hardening
20	Calf injury	94	Muscle injury
21	Calf muscle tear	95	Muscle strain
22	Calf strain	96	Muscle tear
23	Capsule injury	97	Muscular problems
24	Capsule tear	98	Nasal fracture
25	Collarbone fracture	99	Nasal injury
26	Common cold	100	Neck injury
27	Concussion	101	Nose surgery
28	Conservation	102	Outer ligament tear
29	Contracture	103	Lateral collateral ligament rupture
30	Contusion	104	Lateral collateral ligament injury
31	Contusion on the knee	105	Overstretching
32	Cruciate ligament strain (grade 1)	106	Patellar tendon irritation
33	Cruciate ligament strain (grades 2 and 3)	107	Pelvic injury
34	Dental surgery	108	Pubic inflammation
35	Distortion on the ankle	109	Pubic irritation
36	Elbow injury	110	Rib contusion
37	Eye injury	111	Rib fracture
38	Facial injury	112	Shin contusion
39	Fatigue fracture	113	Shin injury
40	Fever	114	Shoulder injury
41	Finger fracture	115	Sick
42	Finger injury	116	Sports hernia
43	Flesh wound	117	Sprain

(Continues)

TABLE A1 (Continued)

Code	Diagnosis	Code	Diagnosis
44	Flu	118	Sprained ankle
45	Flu-like infection	119	Stomach problems
46	Foot contusion	120	Strain
47	Foot injury	121	Strain in the thigh and buttock muscles
48	Fracture	122	Stress reaction of the bone
49	Fracture of the arm	123	Surgery
50	Gastroenteritis/stomach flu	124	Syndesmosis ligament tear
51	Groin injury	125	Tendonitis
52	Groin problems	126	Tendon irritation
53	Groin strain	127	Thigh injury
54	Hand fracture	128	Thigh muscle tear
55	Hand injury	129	Thigh problems
56	Head injury	130	Thigh strain
57	Heel injury	131	Toe fracture
58	Heel problems	132	Toe injury
59	Hip contusion	133	Tonsillitis
60	Hip problems	134	Torn ankle ligament (grade 1)
61	'Dead leg' (quadriceps contusion)	135	Torn ankle ligament (grade 2)
62	Infection	136	Torn ankle ligament (grade 3)
63	Inflammation	137	Torn knee ligament
64	Inflammation in the knee	138	Torn ligament
65	Inguinal hernia	139	Torn ligament ankle joint
66	Injury to the abdominal muscles	140	Torn muscle fibre (grade 1)
67	Injury to the ankle	141	Torn muscle fibre (grades 2 and 3)
68	Injury to the leg flexor muscle	142	Torn muscle fibre in the adductor area
69	Inner ligament injury	143	Torn tendon
70	Inner ligament strain	144	Training deficit
71	Inner ligament stretching knee	145	Viral disease
72	Inner ligament tear (grade 1)	146	Wound
73	Inner ligament tear (grades 2 and 3)	147	Wrist fracture
74	Inner ligament tear knee	148	Zygomatic fracture

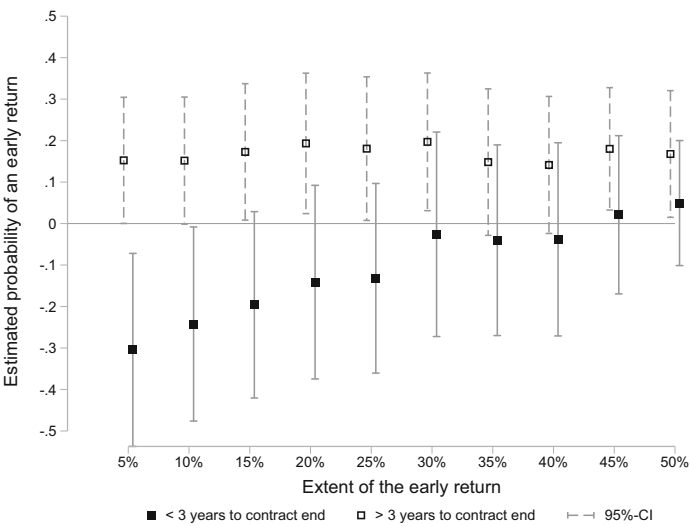


FIGURE A1 Effect heterogeneity: contract length. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of an early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 1038$.

the COVID-19 pandemic, we discarded observations from the year 2020. This also applies to injuries that start before the pandemic in January 2020, because we expect the players' comeback to be affected by the worldwide shutdown of sports events.

Our data cover the major football leagues in Germany (*Bundesliga*), Spain (*Primera Division*) and Italy (*Serie A*). We do not use data from the UK and France as the website lists implausibly low injury numbers per year for these leagues.

Furthermore, we also excluded injury codes that are associated with a maximum duration of more than 365 days. These very severe injuries are unlikely to involve a potential decision for an early return within the relevant time span of 6 months from January to June. We also dropped injury codes where we observe fewer than 10 cases overall, to avoid large prediction errors for the expected injury duration. All injuries categorized as 'Unknown injury' as well as 133 injury spells coded as 'Virus infection' were discarded.

After cleaning the data and checking for implausible intervals (such as when the end date was before the start date), we end up a sample of 7108 observations in the reference periods (autumn season) and 6429 observations in the relevant period (spring season). All injury types observed in our final sample are tabulated in Table A1.

Additional figures

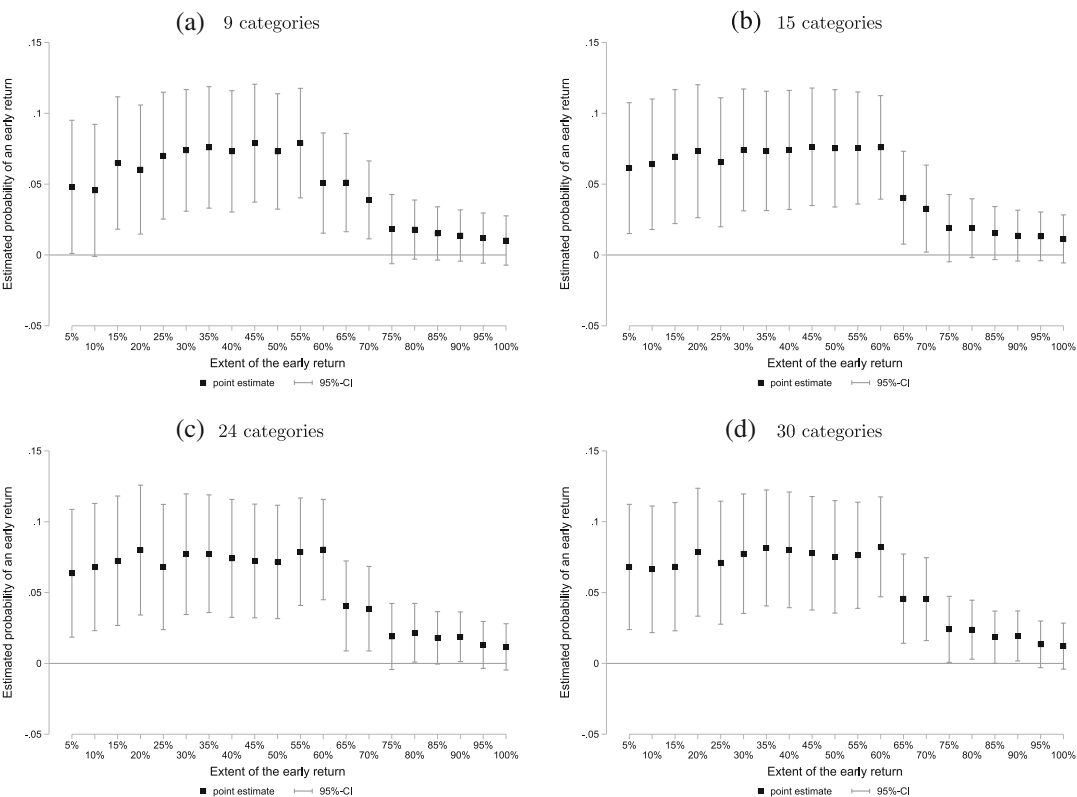


FIGURE A2 Robustness check: different numbers of injury categories.

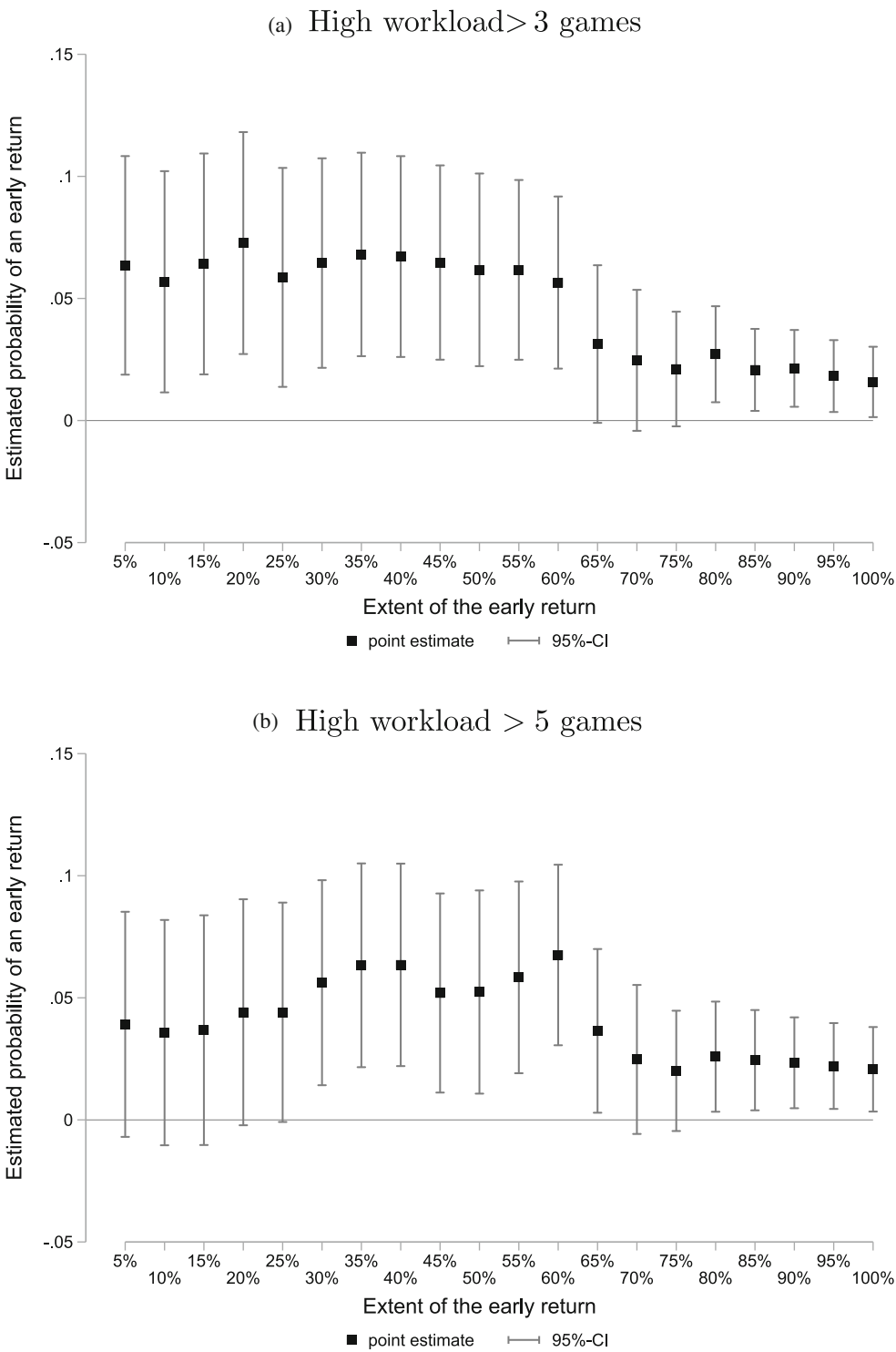


FIGURE A3 Robustness check: alternative cut-off values for *high_workload*. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. *N* = 6429.

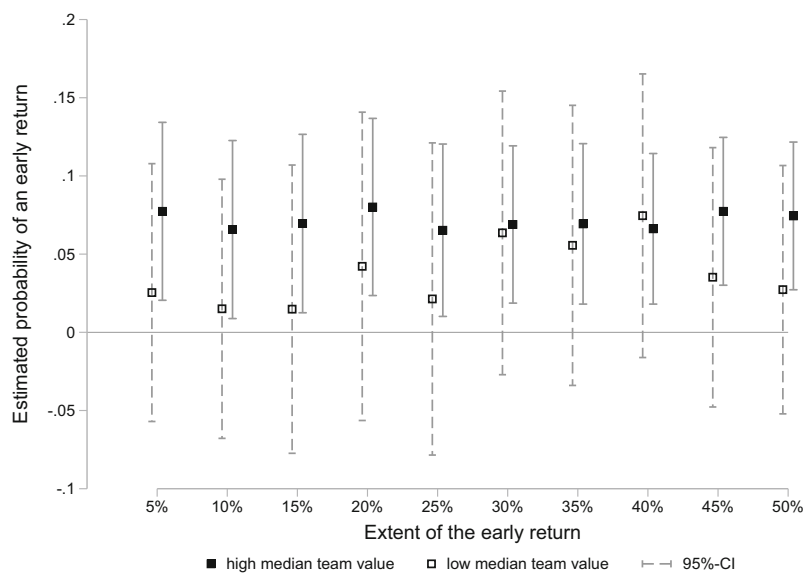


FIGURE A4 Effect heterogeneity: high-budget versus low-budget teams. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 6429$. Solid (hollow) squares represent estimates for the group of high-budget (low-budget) teams. Teams were categorized using a median split.

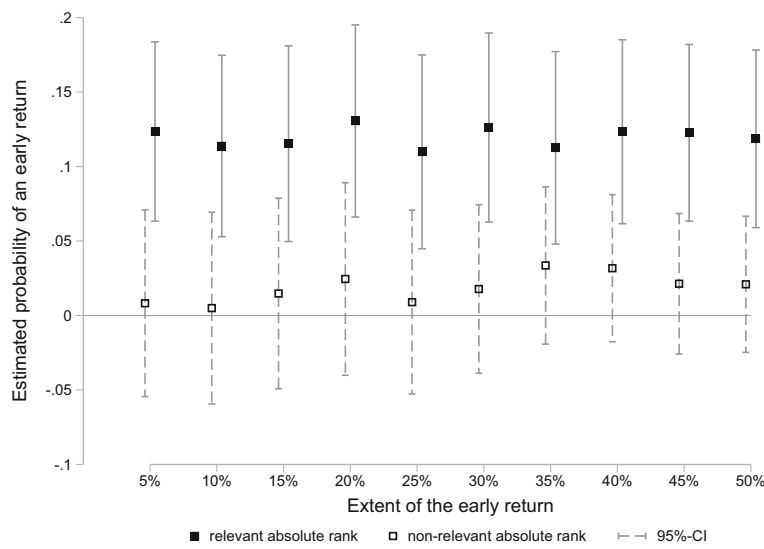


FIGURE A5 Effect heterogeneity: high-skilled versus low-skilled players. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 4429$. Solid (hollow) squares represent estimates for the group of high-skilled (low-skilled) players. Players were categorized according to their market value ranking position within their teams (threshold 8).

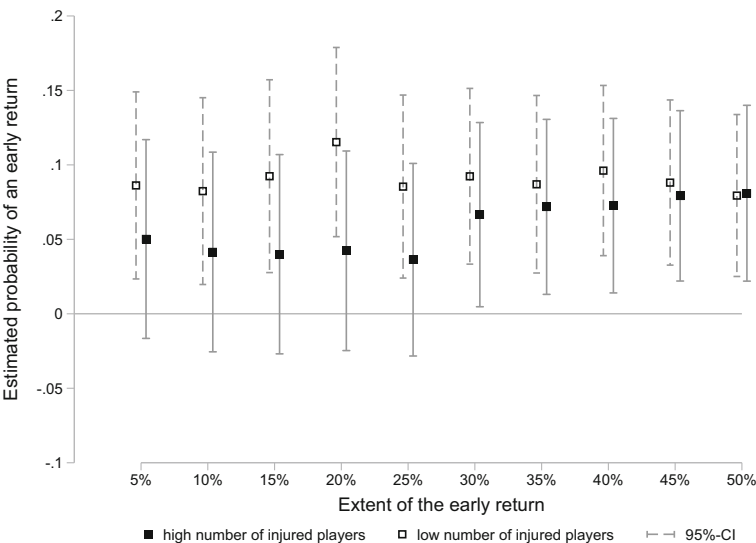


FIGURE A6 Effect heterogeneity: team-specific number of injuries. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 4429$. Solid (hollow) squares represent estimates for the group of teams with a high (low) absolute number of injured players at the time when the focused player is absent. Teams were categorized using a median split.

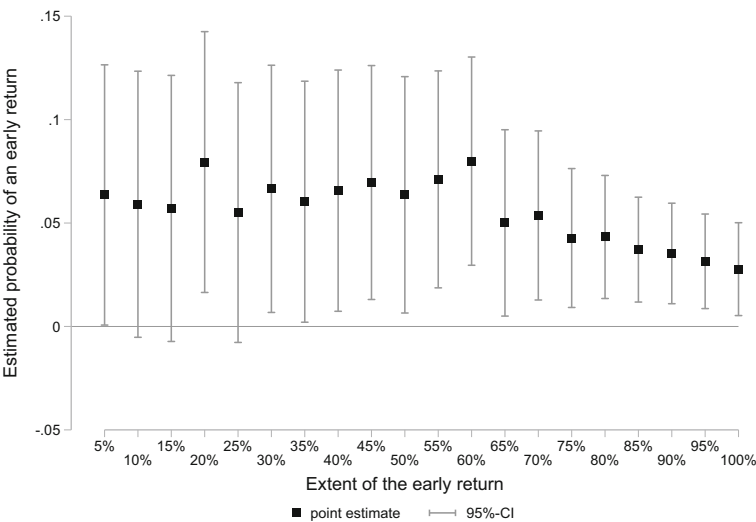


FIGURE A7 High workload affects the probability of returning from injury earlier—no seasons with international tournaments. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 3606$. Only injuries that started in May or later are included. The sample does not include seasons followed by the FIFA World Championship or the UEFA European Championship.

FIGURE A8 High workload affects the probability of returning from injury earlier—player fixed effects. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 5330$. Sample is reduced due to dropping singleton observations. Only injuries that started in May or later are included.

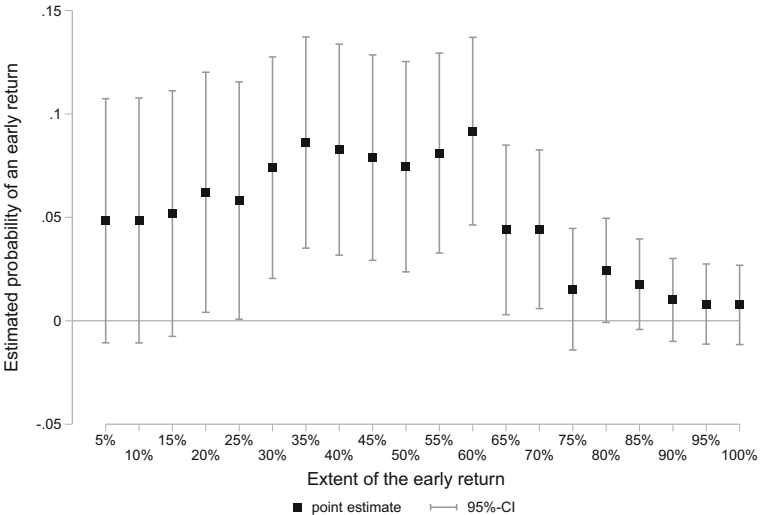


FIGURE A9 High workload affects the probability of returning from injury earlier—no national cup games included. *Notes:* Each square represents the point estimate for the binary variable *high_workload* and different dependent variables related to the extent of the early return. Full model specification as defined by equation (3) and similar to column (6) in Table 4. $N = 4429$. The variable *high_workload* was calculated without national cup competitions.

