ARTICLE IN PRESS

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

COVID-19 infections and short-run worker performance: Evidence from European football

David Butler [a], Robert Butler [a,*], Alex Farnell [b], Robert Simmons [c]

[a] Centre for Sports Economics & Law, Department of Economics, University College Cork, Ireland
[b] Department of Economics, Maynooth University, Ireland
[c] Management School, Lancaster University, United Kingdom

ARTICLE INFO

JEL codes:
J24
J44
Z20

Keywords:
OR in sport
Productivity
Performance
Football
COVID-19

ABSTRACT

COVID-19 infections represent a recurrent source of workplace absenteeism impacting labour productivity. Using a unique matched employee-employer dataset, we consider the effects of the virus on the performance of highly valuable employees when returning to work: professional footballers in the top five European leagues. This offers a window to study job scheduling and managerial decision-making. We employ a difference-in-differences (DiD) model that compares the performance of infected players to a matched control group for game tasks that require physical exertion. Results suggest that per-minute performance is unaffected upon returning to play. This is likely due to effective management of minutes on the pitch. We carry out a battery of checks on the primary results to consider causal mechanisms outside of infection that could impact the results such as lockdown breaks, clusters within squads, and scheduling effects. The findings carry an optimistic message and specifically speak to managers supervising physical labour. If appropriately managed, infected workers can return to past performance levels.

1. Introduction

As of December 2023, there were over 772 million confirmed cases of COVID and over 6.98 million deaths attributed to the disease worldwide (World Health Organisation, 2023). To date however, little is known about how organisations and managers respond to employees recovering from COVID-19. This investigation is necessary as mass infection of employees has brought about significant management challenges. In particular, the virus has brought about labour shortages (Nagurney, 2021) and presented managers with workforce planning dilemmas.

We consider how employees are selected and perform on return to the workplace following infection. Gaining empirical insights on this subject is a relevant international issue for operations research. Our analysis speaks to general topics such as job scheduling, output optimisation and performance evaluation. As we study the effect of viral infections in a labour market where high levels of physical fitness are required, our results provide a test case in managing COVID-19 for firms in related industries, where managers must solve workforce planning problems brought about by sickness.

The context of this research - professional football - is a domain where it is possible to empirically study how employees are selected, supervised, and perform on return to work following illness. This is due to the availability of rich productivity data and public disclosure of illness. While various manual occupations cannot work remotely, fit and relatively young footballers are on-site for matchday activities *and* produce an abundance of detailed performance statistics. Furthermore, it is possible to study managerial decision-making as a player's workload (minutes) must be decided by a manager from a roster of potential staff. Thus, we have a rare opportunity to directly evaluate how workers recuperate from a physically debilitating disease and how human resources are selected and managed on their return to work.

The work lost due to illness was costly even prior to the COVID-19 pandemic. The World Health Organisation (2019) reports that an average of 11.9 days per worker, per year, are lost due to illness or injury in the European Union. In the United States, the Bureau of Labor Statistics (2019) suggests 2% of working time is lost to sickness each year. For each case of COVID-19, it is estimated that 1.0 to 1.5 days of productive work are lost (Berdan et al., 2023). Yet very little is documented about the effects on productivity at the individual worker level. Most studies conducted on the topic of performance on return to work rely on survey data and self-reported productivity measures.

Our study makes several contributions. First, we evaluate managerial

* Corresponding author.
E-mail address: r.butler@ucc.ie (R. Butler).

https://doi.org/10.1016/j.ejor.2023.12.017
Received 22 May 2023; Accepted 17 December 2023
Available online 19 December 2023

Please cite this article as: David Butler et al., *European Journal of Operational Research*, https://doi.org/10.1016/j.ejor.2023.12.017

decision making within firms and employee performance, considering worker choice and workload intensity. Second, we add to nascent research considering the productivity effects of contracting COVID-19 in sport (Wagemans et al., 2021; Fischer et al., 2022; Wezenbeek et al., 2023) and expand on a sparse empirical literature concerning managerial decision-making upon recovery from the illness and return to work. Wagemans et al. (2021) show that COVID positive players' well-being, stress levels and moods diminished after contracting the virus. Wezenbeek et al. (2023) show evidence that player's aerobic performance seemed compromised after infection, but anaerobic performance seemed to be spared. Fischer et al. (2022) report a persistent deterioration in player performance for more than six months post-infection. This negatively affects productivity in the long-run and is unlike other respiratory infections, which are not reported as having the same effect. Our work builds on these papers as we estimate the effect of infection across all five top UEFA leagues (Premier League, Ligue 1, Bundesliga, Serie A and La Liga). It is also the first study to exclude goalkeepers from the estimations – a crucial differentiation on earlier empirical work given the unique aspects and requirement of this position.

Establishing relevant datasets to consider performance on return to work is challenging as illness data are protected by privacy laws. Even if data are available, an exact diagnosis is commonly not disclosed. Our setting represents a unique exception. We study the effect of a uniform infection (that has heterogeneous effects) where worker information is publicly disclosed. As such, we contribute to an emerging literature considering the effect of the virus in professional sports industries (e.g., Bryson et al., 2021a; Fischer & Haucap, 2021). We do this with a comprehensive matched employee-employer dataset that is linked to precise performance data.

Professional football is an interesting labour market. The personnel are extremely valuable, require remarkably high levels of physical and mental fitness to perform optimally and are supported by a world class medical infrastructure. Furthermore, elite football is a high-stakes environment. We are confident that workers are intensively monitored, and that strong competition results in effective incentives to encourage players to return to work as soon as possible. The competitive nature of the industry, along with the considerable financial resources at the disposal of elite clubs, means that players can be promptly replaced. The evidence shows that contracts are increasingly tied to performance and players lose out on bonuses and contract extensions by not performing well (Buraimo et al., 2015).

Given the conditions of our context, and the general medical findings so far, forming two opposing priors on the effects of infection is reasonable. The first is that COVID-19 has persistent effects on returning players. As the elite footballers we sample are young athletes that epitomise the idea of optimal physical conditioning, this would be a pessimistic result. If accurate, it would be particularly alarming given that comparable medical supports are not available for regular workers in industries requiring high levels of physical fitness. An alternative prior is that player recovery is overseen, so that players regain pre-infection levels of performance. This hypothesis represents a relatively optimistic outcome. While not discounting the potency of the virus, nor expunging the fact that COVID-19 can cause critical illness to individuals, a hypothesis that falls into this set of expectations would presuppose that institutional supports (e.g., medical, sports science/analytics) swiftly allows players to achieve continuity in their performance levels.[1]

The rest of the paper continues as follows. Section 2 discusses relevant literature emerging on COVID-19. Section 3 presents the data, beginning with our performance measures, and followed by our COVID sample and control group. Section 4 presents the empirical model. The results are presented in Section 5. Our empirical analysis is split into alternative stages of analysis: player level effects, team level effects and heterogeneous effects. Our models are all at a player performance-per minute level. Following our primary analysis, we carry out a series of robustness checks in Section 6. Here we empirically consider checks on minutes per match thresholds, checks on the treatment groups, the role of scheduling effects and the role of any temporal performance variations. Section 7 concludes the paper with a brief discussion of the findings.

2. Related literature

An important strand of COVID-19 research relevant to our priors above explores acute post-COVID-19 effects. A primary message from this literature is that while a sufferer is no longer infectious, and has largely recovered, adverse side-effects can remain. For example, Bernanke and Yellen (2020) claim that while the long-term consequences remain unknown, debilitating post infection effects can result from lingering viral infections. Lasting effects include fatigue, breathlessness, coughing, aching muscles, and pressure and heaviness in the chest (Nabavi, 2020). These early results, documenting the lingering effects of infection, could have serious consequences for employee productivity. This topic of employee illness, absenteeism and effectiveness when returning to work has been explored previously. The main findings from this literature are mostly as anticipated; illnesses/injury have a notable negative effect on employee working days and their effectiveness when returning to work. Keech et al. (1998) study the effects of influenza-like symptoms for 628 employees in a large pharmaceutical company. This was a non-epidemic context, and the analysis was based on retrospective questionnaires. They report that the flu incapacitated employees for an average of 2.5 days and reduced effectiveness for 3.5 days upon return.

Lötters et al. (2005) conduct a longitudinal study exploring the productivity of 253 workers after absences. This absenteeism was due to musculoskeletal disorders. The data was based on questionnaires measuring perceived productivity. On average, the sickness absence was 34 days and musculoskeletal disorders resulted in reduced productivity upon return for 60% of participants. 40% of those effected continued to operate at reduced levels of productivity after 12 months. Recently, Grinza and Rycx (2020) consider the causal effect of sickness absenteeism on firm productivity. They suggest that absenteeism is accentuated when workers have in-depth knowledge of tasks they must complete, operate in a highly interconnected team and are limited in their ability to replace workers that are absent. The authors find that a one percentage point increase in the absenteeism rate results in productivity losses of 0.66%. This was more pronounced for higher tenured workers.

Given these results, we conjecture that negative productivity effects may also exist in sport. It is also reasonable to assume that illnesses, particularly those that impact upon the respiratory system, are likely to have an amplified effect on athletes. Few insights exist however on this topic. Lichter et al. (2017) is an exception. They study the impact of air pollution – a restrictor of respiratory function – on performance. Accessing match-level data on the German Bundesliga they show that increased pollution, measured by particulate matter, decreases the number of passes made at the individual and team level. An increase of one standard deviation in pollution led to a reduction in passes by 2.4%, with stronger effects associated at higher levels of pollution.

Vaudreuil et al. (2021) were one of the first to consider the impact of COVID-19 on athletes. They consider player recovery for a limited sample of 'bubble players' in the National Basketball Association, finding no long-lasting effects on player performance. Using a before-after comparison they conclude that players showed no signs of their performance deviating from a pre-COVID-19 level. In the context of professional football in Belgium, Wagemans et al. (2021) and Wezenbeek et al. (2023) evaluate the impact of COVID-19 infections between 2020 and 2021. The nearest work to our research is Fischer et al. (2022), who examine COVID infections in the *Bundesliga* and *Serie A* football

[1] At least one observation in our dataset was hospitalised for a significant period of time.

leagues. The authors consider player productivity and report that performances decline upon return from infection by an average of 6% and remain as much as 5% lower eight months after recovery. While conceptually like our research, there are several key differences; first, we consider a short time frame while focusing on a wider range of precise productivity measures that might plausibly be impacted by COVID infections. Our post-`treatment' time horizon differs from Fischer et al. (2022) who consider a period of up to a year and primarily assess passing capabilities. Second, our control group selection differs from Fischer et al. (2022). We construct a matched control group to be as comparable as possible to the treatment group to reduce confounding, while Fischer et al. (2022) consider all non-infected players as the control group. Third, we consider data from five leagues while Fischer et al. (2022) study COVID positive players from two leagues.

3. Data

To examine the effects of COVID-19 on performance, we record instances of infections among outfield players and construct a control group that is not subjected to infection. We pair these observations with match-level performance statistics.[2] This section progresses by providing a detailed description of our productivity measures and our approach to measuring productivity in this context. We then specify our COVID and control samples before detailing any additional data used for controls.

3.1. Performance data

The use of advanced performance metrics is increasingly common in sports data analysis. Examples relating to football can be found in Kharrat et al. (2020) and McHale et al. (2023). We pair a (COVID) treatment and control group with high quality performance data generated by DATA SPORTS GROUP (DSG) and STATSBOMB accessed via FBref.com. Performance data are available for domestic league matches and UEFA competitions only. There is a strong incentive for clubs and players to perform to the best of their ability in these competitions and we consider both types of competition as elite. Exhibition matches and domestic cups warrant exclusion as clubs regularly rotate squads or place less value on these competitions. We evaluate performance over a ten-match period. A less homogenous sample is present over a greater window due to the potential for suspensions, additional injuries, personnel changes, lockdowns, pre-seasons, and international breaks. For the infected players, this includes five game pre- and post-infection. It is important to note that a ten-match period is not conflated with an excessively short time period; our sample timeframe regularly involves playing matches over two months or quarter of a season.[3] In addition, the dataset was collected manually. Given that one must identify *specific performers*, at *specific time periods*, it would be computationally challenging to automate the process to extract specific advanced measures only. Consequently, collecting an entire season of performances is not feasible. We select high intensity performance traits, choosing physiological metrics that require a functioning circulatory and respiratory system. While there is now a dearth of performance statistics available, our measures are intended to capture efforts requiring stamina, endurance, sprinting and acceleration capacities.

Table 1
Measures, Definitions, and Demands.

Performance Measure	Definition*	Demands
Minutes	Number of minutes played per match	Endurance
Presses	Number of times pressure is applied to an opposing player who is receiving, carrying, or releasing the ball (within a 5-yard radius)	Sprinting & Acceleration
Progressive Distance	Total distance, in yards, a player moved the ball while controlling it with their feet towards an opponent's goal	Ball Control & Acceleration
Carries	Number of times the player controlled the ball with their feet	Ball Control, Movement Intensity (Work Rate)
Progressive Carries	Carries that move the ball towards the opponent's goal at least 5 yards, or any carry into the penalty area. Excludes carries from the defending 40% of the pitch	Ball Control, Acceleration, Agility & Balance
Dribbles	An attempt by a player to beat an opponent when they have possession of the ball.	Ball Control & Acceleration
Expected Goals (xG)	Non-Penalty Expected Goals	Positioning, Vision & Composure
Pass Completion (%)	Pass Completion percentage	Concentration & Composure

* Source: FBref.com definitions.

Past research has demonstrated the importance of physical exertion on win probabilities (Weimar & Wicker, 2017). We carefully choose performance measures as we believe the effects of COVID would be best considered in the context of activities that require high levels of physical effort exertion. While we mostly focus on in-possession statistics, we also measure presses (attempts to pressurise opponents into giving up possession) as an out-of-possession statistic which requires high intensity effort. These improve upon measures used previously. Our full set of performance measures is detailed in Table 1.

It is essential to control for position in any empirical models using these performance data as many performance measures in football will be unbalanced across roles. For example, defenders display greater progressive distances while offensive players press the opposition far more. Given the fluidity of modern football, we believe the choice of metrics is representative of a variety of higher intensity tasks and offer reasonable coverage of performance that is not biased towards any one role.

Table 2
Performance controls.

Control	Definition	Type
Age	Age of the player on the date of contracting COVID-19	Metric
Position	Defence, Midfield and Attack/Forward	Indicators
Cluster	1 if three or more players within a club contracted COVID-19 in a week where a match was played, 0 otherwise	Indicator
Injury	1 if a player suffered an extenuating or additional injury, 0 otherwise	Indicator
Europe	1 if the player's club is still involved in a European competition, 0 otherwise	Indicator
Transfer	1 if the player was transferred to another club, 0 otherwise	Indicator
Cross-Season	1 if performance was measured before/after a preseason period, 0 otherwise	Indicator
Lockdown	1 if performance was measured before/after league suspensions, 0 otherwise	Indicator
Home	1 if the match was played in a home stadium, 0 otherwise	Indicator
Opposition ELO	The ELO rating of the opposing team	Metric

[2] Twenty positive cases involved goalkeepers. This position requires a significantly different skillset, and these players are evaluated differently compared to dynamic outfield players (Berri et al., 2023). They are therefore excluded from the sample.

[3] An extended analysis acutely increases confounds that may increase noise in the data and subsequent analysis. These confounds still arise within a five-match window but are fewer, and as we show later, it is possible to generate a meaningful sample over a five-match period when we impose restrictions to account for the challenges in attaining clean data.

Table 3
Descriptive statistics.

Variable	Obs. Whole Sample	Mean	Std. dev.	Min	Max	Obs. Control	Mean	Obs. Treatment	Mean	Difference in means (p-value)
Outcomes										
Mins	3976	70.978	27.959	1	120	1738	78.694	2238	64.986	0.000
Presses	3976	11.755	9.207	0	117	1738	14.671	2238	9.491	0.000
Prog Distance	3971	86.697	74.627	0	696	1733	92.610	2238	82.118	0.000
Carries	3975	30.878	20.627	0	445	1738	32.712	2237	29.454	0.000
Prog Carry	3967	3.647	3.799	0	51	1728	3.628	2239	3.661	0.783
Dribbles	3972	1.366	1.893	0	19	1733	1.411	2239	1.331	0.182
Pass Completion%	3945	79.220	13.201	0	100	1734	77.771	2211	80.357	0.000
XG	1067	0.208	0.287	0	1.8	567	0.222	500	0.192	0.086
Covariates										
Age	4580	26.651	4.184	16	38	2290	26.952	2290	26.349	0.000
ELO	3975	1954.001	162.223	1089	2901	1737	1943.031	2238	1962.516	0.000
Europe	4580	0.404	0.491	0	1	2290	0.262	2290	0.546	0.000
Home	3941	0.499	0.500	0	1	1735	0.496	2206	0.502	0.729
Position										
Midfielder	4580	0.443	0.497	0	1	2290	0.437	2290	0.500	0.506
Defender	4580	0.384	0.486	0	1	2290	0.397	2290	0.371	0.153
Attacker	4580	0.172	0.378	0	1	2290	0.166	2290	0.180	0.286
League										
England	4580	0.205	0.404	0	1	2290	0.201	2290	0.210	0.514
France	4580	0.142	0.349	0	1	2290	0.197	2290	0.087	0.000
Germany	4580	0.166	0.372	0	1	2290	0.201	2290	0.131	0.000
Italy	4580	0.317	0.465	0	1	2290	0.201	2290	0.432	0.000
Spain	4580	0.170	0.376	0	1	2290	0.201	2290	0.140	0.000

*Note: The number of observations differ across treatment and control groups since not all players will play a complete 10 matches (consisting of 5 pre- and 5 post infection games). Within treatment and control groups, the number of observations for certain outcomes and covariates differ because certain outcomes were missing. These are few and far between however. Pass completion% is notably lower owing to players playing very few minutes in certain matches. Table 8 offers a robustness check restricting to observations where players play at least 10 min in matches. xG has fewer observations since we only observe this for forwards.

3.2. *COVID-19 sample*

We manually construct a matched dataset for 229 players who tested positive for COVID-19 in 66 clubs across the top 5 European Leagues (Premier League, Ligue 1, Bundesliga, Serie A and La Liga). This dataset excludes false positives, goalkeepers, COVID positive players that transferred to or from leagues outside of the top 5 over our sample timeframe (as they were then not covered by our data provider), youth/reserve players and players that were relegated or promoted with their club (as precise performance stats were not available and nor would they be comparable). There is natural variation in the timing of these infections. Positive cases were recorded from the 8th of March 2020 to 26th of April 2021, and our data cover performances in the 2019/2020 and 2020/2021 seasons.

Given the lineage of COVID-19, this time spans the emergence of multiple major variants of concerns, with the exception of the Omicron variant. These cases were collated from official league and club disclosures, social media posts (where the player publicly discloses that they tested positive), and reputable media. In all cases, we checked to ensure a period of inactivity for in-season cases, explored alternative reasons for the absence (e.g., suspension), and recorded player return dates to elite competitions.[4]

Three points are worthy of addressing regarding potential sample bias. First, are the media more likely to publicise positive cases if a player is well-known? After assessing the underlying valuation estimates of the positive cases (via *transfermarkt.com*), we are confident this is not a concern. We observe considerable heterogeneity in the distribution of valuations of disclosed positive cases which are reported across the spectrum of player ability levels. Second, we can only analyse players who returned. This may raise a concern about upwardly biased results that fail to account for non-returning players. Again, this is not a concern as we record very few cases of players not returning over the sampling timeframe. Third, we understand that no players were vaccinated during this timeframe. Inoculation effects could undermine the uniformity and effect of the treatment. This is indisputable for most of our sample given the timeline of vaccine development. Additionally, we are satisfied that even when vaccines were available for a short cross-over period, the average age of this cohort meant that they were not prioritised.

3.3. *Control sample*

Even though we believe it is fair to assume that infections are a random shock to a player performance, we cannot rule out general trends in productivity. Therefore, a control group is required that is not subject to infection.

We imposed decision rules to design a comparative control group. To certify a fair representation of talent across the population of players where positive cases are observed, we selected a minimum of two players from each elite club in the top 5 leagues, and the remainder from other clubs. These players are randomly drawn but must meet a minutes threshold over the season to attain sufficient performance statistics. The control group is matched by position, so a disproportionate number of observations are not in one position. It is worth emphasising that this matching is at an exact positional level – we do not group all defenders or midfielders together for the purposes of comparison. Instead, we randomly match on exact positions given the complexity of roles within the general positional zones (e.g. right-back, defensive midfield, winger etc.).

The players selected had played at least four matches between the 1st January 2020 and 7th of March 2020, prior to the widespread outbreak of COVID-19. Several potential sampling issues arise during this timeframe. From an epidemiological standpoint, COVID-19 was circulating at low levels when we sampled. To address this concern, we assessed historical data available for the English Premier League (through fantasy football websites) and examined if a disproportionate number of players were unavailable relative to past seasons. There is no compelling evidence to suggest lower squad sizes due to illness. Second, matches in our control group took place in front of a live audience.

[4] We later consider infections contracted during lockdown breaks and off-season.

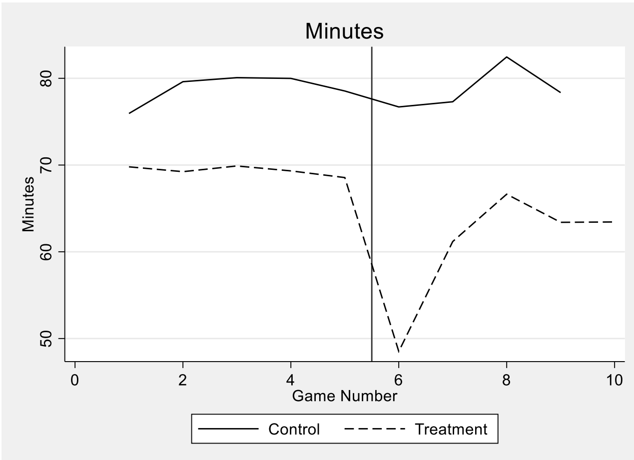

Fig. 1. Minutes.

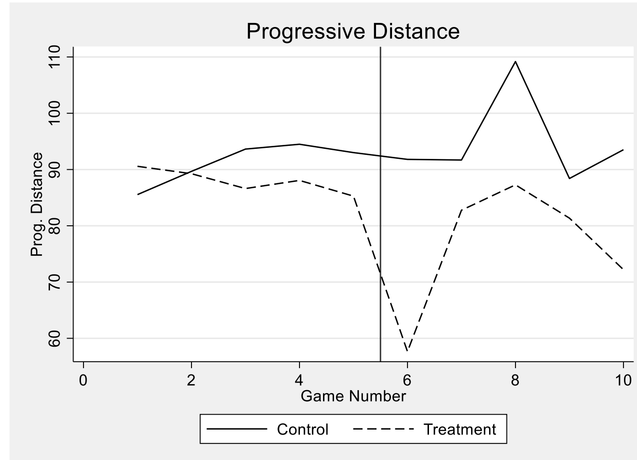

Fig. 3. Progressive distance.

Despite these drawbacks, this method represents the most cautious approach. It is critical to sample from the most recent period prior to the cessation of football to ensure as much continuity as possible in the external environment. Professional football is a fast-paced business, where expectations rapidly shift and managerial changes frequently occur (Bryson et al., 2021b). Moreover, some European leagues take a winter break. If sampling was conducted for an earlier time-period, such as prior to the new calendar year, a greater number of confounds arise to invalidate the control.

Most importantly, the control must be sampled before the virus is endemic. Sampling in a post-lockdown carries risk. To illustrate this, consider that between 31st August 2020 and 16th May 2021, 261 positive cases were recorded alone in the English Premier League (Premier League, 2021). While this included players and staff, if we assume a similar positivity rate across leagues and consider our treatment size, it is possible that positive cases could be included in our control. In summary, our control group sampling strategy represents the most prudent approach. However, we acknowledge this necessitates trade-offs.

3.4. Control variables

We control for various factors which could otherwise determine performance. These include player age, position, match location (home or away), major injuries, club outbreaks, scheduling and rotation risks through involvement in European competitions. As we later consider team level output effects, we also record the scoreline of each match. We estimate the severity of the infection by measuring the number of days of inactivity. In general, the unavailability period is relatively short; on average positive cases (excluding lockdown/pre-seasons cases) spend 23 days out (median 18). For most players, this is a relatively short time away and includes a mandatory quarantine period. Considering the personal training that players received remotely during isolation, it is unlikely that players would suffer any match fitness effects from quarantine that need to be distinguished from a causal COVID-19 effect.

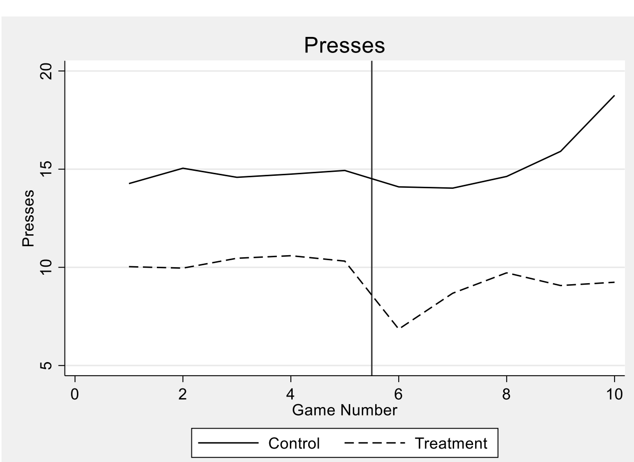

Fig. 2. Presses.

The above measures are player and club-level performance determinants. Critically, the standard of opposition is a determinant of player productivity. Since we are carrying out a cross-league analysis, we use club ELO ratings to measure opponent strength. These ratings have the advantage of giving a relative measure of quality across leagues and European competitions. The ELO ratings are based on the prior results of each club and are updated throughout the season. A list of the control variables is defined in Table 2.

Although we control for various influences on productivity, cross-country differences could persist. We include league fixed effects to account for this. One notable difference is the number of substitutes permitted across the leagues. The Premier League reverted to three substitutions per game, while the other leagues in our sample continued with five substitutions. Additionally, there are other more subtle differences between the leagues such as differing playing styles, team tactics and intensities, which could affect our observed performance statistics.

3.5. Descriptive statistics

Table 3 shows the descriptive statistics, both for the whole sample and separately by treatment and control groups. The performance statistics for the treated players are often lower than for the control group.

The extent to which we can attribute this to COVID-19 infections is the focus of our empirical strategy that follows. The covariates seem reasonably balanced across the groups, except for the Europe dummy which is likely caused by outbreaks at specific clubs still participating in European competitions (e.g., Paris Saint-Germain, Bayern Munich and Real Madrid). Italy also represents a much greater portion of the treatment, which is to be expected as this is where the first major European outbreak occurred.

Figs. 1–8 illustrate the evolution of the raw performance measures over the 10-game period, with treatment and control groups plotted separately.

Evolution of performance across infected and non-infected groups.

The vertical line separates the pre- and post-infection periods, with 1–5 representing pre-infection and 6–10 representing post-infection. Noticeable in most of the figures is a performance decline observed at

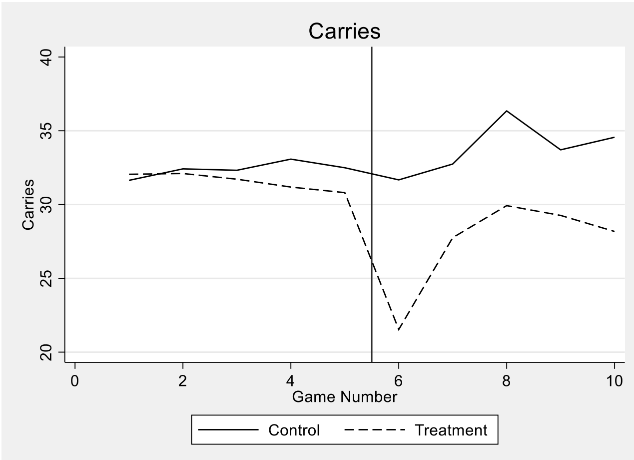

Fig. 4. Carries.

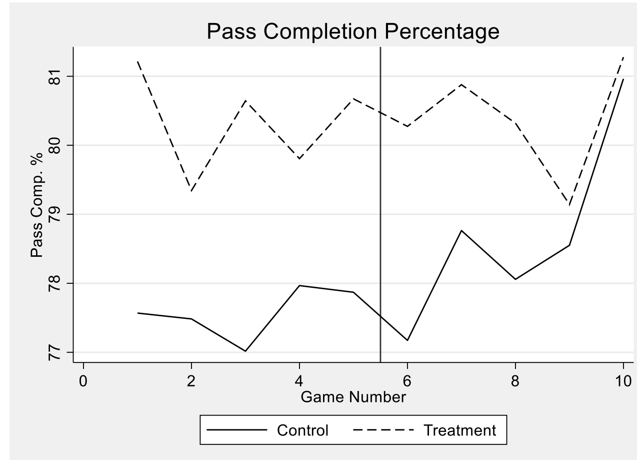

Fig. 7. Pass completion%.

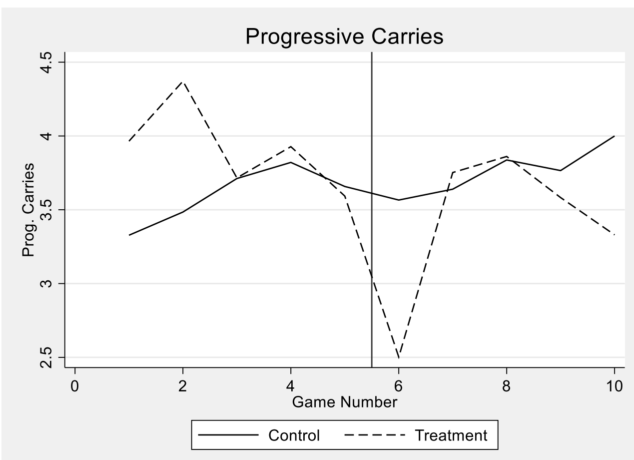

Fig. 5. Progressive carries.

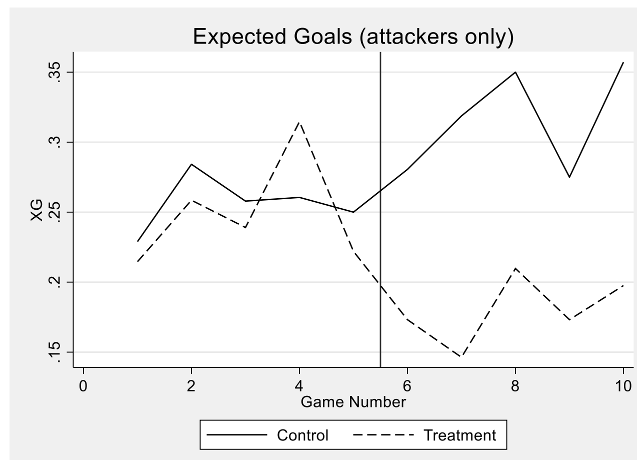

Fig. 8. Expected goals.

game 6 (first game back post-infection), followed by a recuperation period of two to three games. After this period, performance appears to revert to its pre-infection levels. Pass Completion Percentage and Expected Goals, and, to an extent, Dribbles display less of a discernible pattern - both pre- and post-infection - suggesting much of the variation in these performance measures is driven by noise rather than any

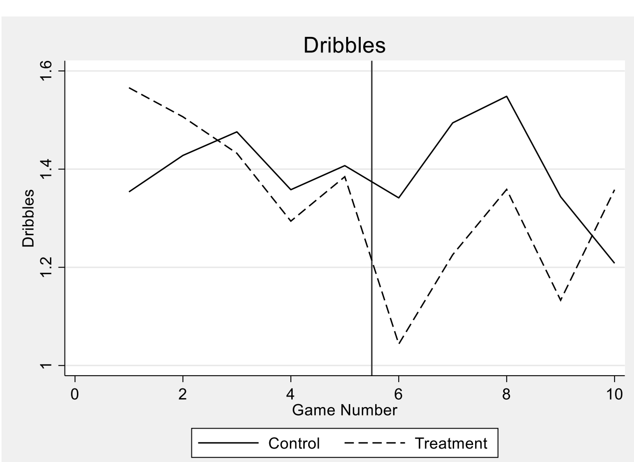

Fig. 6. Dribbles.

variables that we can explicitly control for. Note also the sharp decline in minutes played by the treated group in games immediately after return. Since the raw figures are likely to be heavily affected by the number of minutes played, our modelling approach will continue by using per-minute outcomes. On average, infected players play almost 9 minutes fewer per game, averaged over the 5 games post infection. However, this masks a great deal of variation. In the first game after infection, infected players play 22 fewer minutes than their non-infected counterparts, and in the subsequent matches, continue to play around 8–9 minutes fewer.

4. Empirical model

To model the effects of COVID-19 infections on footballers' performances, we employ a differences-in-differences approach as follows:

$$Perf_{it} = \alpha + \beta_T Treat_i + \beta_P Post_t + \beta_{TP}(Treat * Post)_{it} + \boldsymbol{\beta X} + u_{it} \quad (1)$$

Perf is the per-minute performance for player i in game t.[5] We observe performance over a 10-game period for each player. For the infected group of players, this captures five games before and after infection, while for the control group, this is a 10-game period before any recorded infections. *Treat* is a dummy variable equal to one if player i has been infected, zero otherwise, and *Post* is a dummy variable equal to one

[5] An exception is the pass completion percentage since this is already a relative statistic.

Table 4
Baseline specification.

VARIABLES	(1) Presses per min	(2) Prog. Dist. per min	(3) Carries per min	(4) Prog. Carry per min	(5) Dribbles per min	(6) XG per min	(7) Pass Completion%
Panel A: No Controls							
Treat	−0.031***	0.187***	0.069***	0.017***	0.006***	0.002*	2.754***
	(0.011)	(0.073)	(0.019)	(0.004)	(0.002)	(0.001)	(0.820)
Post	0.004	0.063	0.020**	0.002	0.005	0.001	0.554
	(0.008)	(0.046)	(0.010)	(0.002)	(0.004)	(0.000)	(0.645)
Treat * Post	−0.007	−0.167**	−0.026	−0.006	−0.007*	−0.002*	−0.510
	(0.011)	(0.078)	(0.021)	(0.004)	(0.004)	(0.001)	(0.847)
N	3976	3966	3975	3965	3966	1057	3945
R-Squared	0.009	0.003	0.007	0.011	0.002	0.009	0.010
Panel B: Including Controls							
Treat	−0.032***	0.082	0.021	0.010**	0.006***	0.002	1.250
	(0.010	(0.074)	(0.019)	(0.004)	(0.002)	(0.001)	(0.759)
Post	0.000	0.068	0.015	0.000	0.004	0.001	0.438
	(0.008)	(0.044)	(0.009)	(0.002)	(0.004)	(0.000)	(0.620)
Treat * Post	−0.004	−0.160**	−0.019	−0.004	−0.006	−0.002*	−0.289
	(0.011)	(0.076)	(0.021)	(0.004)	(0.004)	(0.001)	(0.836)
N	3940	3932	3939	3931	3932	1047	3909
R-Squared	0.085	0.053	0.074	0.063	0.039	0.033	0.109

Cluster robust standard errors in parentheses (clustered at the player level).
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$
Panel B Controls: Age, Age Squared, Opposition ELO, Europe, Home, Position Indicators, League Indicators.

indicating the second half of the 10-game period. Our coefficient of interest is then β_{PT}, which is the effect of the interaction between *Treat* and *Post.* The value of the interaction between *Treat* and *Post* equals one for the treated group in the post-infection period, zero otherwise.

Within the vector $\boldsymbol{X}$ we include covariates that explain performance. These include player characteristics (age, age squared and position dummies), game characteristics (opposition ELO rating, a European involvement dummy, and a home game dummy) and league fixed effects. To complete (1), u_{it} is a random error term. Standard errors are clustered at the player level.

We also extend Eq. (1) to examine the recovery from COVID-19. Rather than including a pre-post indicator, we model games using dummies and interact these with the treatment indicator as follows:

$$Perf_{it} = \alpha + \beta_T Treat_i + \beta_G Game(2-10) + \beta_{TG}(Treat * Game(2-10)) + \boldsymbol{\beta X} + u_{it} \quad (2)$$

where games 1–5 are the pre-infection period, and games 6–10 are the post-infection period.

Our difference-in-differences analysis requires important assumptions to be satisfied. Most notably, performance in the treated and control groups should follow parallel trends in the pre-treatment period (Angrist & Pischke, 2008). Trends in performance are expected be the same across the treated and control groups in the absence of COVID-19 infection. This assumption is likely satisfied, as there is no credible argument to suggest that performance outcomes would systematically diverge in the absence of infection. We must also assume that individuals do not self-select into treatment. Again, this is non-contentious; individuals do not seek out COVID-19. We also carry out several robustness and auxiliary checks, including estimating the models on observations where players play more than 10 minutes per game, and we also check for differences by the intensity of treatment, measured by number of days absent. We detail these checks in Section 6.

5. Results

The results are presented in three main areas: player level effects, team level effects and heterogeneous effects.

5.1. Player level effects

We begin by presenting our baseline estimates in Table 4, where the outcomes are the per minute statistics.[6] Panel A presents the models without controlling for any additional factors. Panel B includes controls for player and game characteristics as discussed. For brevity, the effects of these controls are shown in Appendix Table A2. They perform largely as expected.[7] Our difference-in-difference coefficient (*Treat*Post*) is negative across all measures, although it is only significant at the 5% level in the Progressive Distance model. For Dribble and xG models, the *Treat*Post* coefficient is significant at the 10% level, with small magnitudes of effects. The rest of our performance measures do not show significant effects. An infected player performs no differently to a non-infected player, on average, in the 5 games post recovery. Including the control variables produces similar conclusions. The point estimates reduce somewhat in magnitude, and the effect on Dribbles is not significant. The significance of the *Treat* variable shows that per minute performance of infected players is higher (with the exception of presses) than the control group. This is not an issue for a difference-in-differences model however, which only requires that the pre-treatment trends in performance are the same.

Of course, these 5 game averages might mask a great deal of game-by-game variation. From Figs. 1–8, we observe a sharp drop in many performance measures in raw terms. Hence, in Table 5 we present the models estimating Eq. (2) with per minute outcomes on a game-by-game level. The results show little variation in per minute outcomes across the 5-game follow up period. There is limited evidence that progressive distance and progressive carries per minute are lower in the first game after recovery, though these results are only significant at 10%. There is little evidence to suggest that recovered players are performing differently, nor is there much evidence of a recovery period.

[6] We also include models for the raw performance outcomes and minutes as an outcome. See Appendix A1.

[7] Age has no effect on performance. Player performance is lower when playing against stronger opponents, but higher when playing at home and when a team is involved in a European competition. Our position controls perform as expected, and we do not observe any major differences across the leagues.

Table 5
Game by game analysis.

VARIABLES	(1) Presses per min	(2) Prog. Dist. per min	(3) Carries per min	(4) Prog Carry per min	(5) Dribbles per min	(6) XG per min	(7) Pass Completion%
Treat	−0.056***	0.202	0.022	0.017**	0.014**	0.005	2.122*
	(0.012)	(0.207)	(0.028)	(0.008)	(0.007)	(0.005)	(1.236)
Game 2 * Treat	0.011	−0.095	0.011	−0.002	−0.009	−0.005	−1.844
	(0.012)	(0.216)	(0.033)	(0.009)	(0.007)	(0.005)	(1.414)
Game 3 * Treat	0.020	−0.207	−0.005	−0.015*	−0.011	−0.003	0.073
	(0.016)	(0.227)	(0.038)	(0.009)	(0.007)	(0.004)	(1.488)
Game 4 * Treat	0.043***	−0.106	0.000	−0.010	−0.012	−0.003	−1.796
	(0.015)	(0.227)	(0.043)	(0.010)	(0.007)	(0.005)	(1.423)
Game 5 * Treat	0.046	−0.207	−0.013	−0.009	−0.010	−0.005	−0.735
	(0.031)	(0.218)	(0.029)	(0.009)	(0.007)	(0.005)	(1.376)
Game 6 * Treat	0.018	−0.386*	−0.030	−0.018*	−0.010	−0.003	−0.739
	(0.016)	(0.222)	(0.033)	(0.010)	(0.007)	(0.005)	(1.539)
Game 7 * Treat	0.026**	−0.098	−0.021	−0.004	−0.023*	−0.006	−1.351
	(0.013)	(0.228)	(0.033)	(0.009)	(0.013)	(0.004)	(1.566)
Game 8 * Treat	0.015	−0.269	−0.035	−0.005	−0.012*	−0.004	−1.008
	(0.027)	(0.222)	(0.033)	(0.010)	(0.007)	(0.005)	(1.665)
Game 9 * Treat	0.038*	−0.171	−0.013	−0.011	−0.009	−0.006	−2.753
	(0.023)	(0.233)	(0.044)	(0.010)	(0.009)	(0.005)	(2.027)
Game 10 * Treat	−0.015	−0.283	0.068	−0.010	−0.009	−0.006	−0.697
	(0.032)	(0.241)	(0.073)	(0.011)	(0.009)	(0.005)	(3.387)
Constant	0.297*	3.885***	0.992***	0.135**	0.059*	0.005	99.637***
	(0.151)	(1.300)	(0.312)	(0.063)	(0.035)	(0.011)	(11.947)
Observations	3940	3932	3939	3931	3932	1047	3909
R-squared	0.088	0.057	0.075	0.067	0.044	0.052	0.111

Cluster robust standard errors in parentheses (clustered at the player level).
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$
Models include controls for Age, Age Squared, Opposition ELO, Europe, Home, Position Indicators, League Indicators.

These results support the more optimistic of the two priors we propose. Even though the effects of COVID-19 may indeed be debilitating, it appears that clubs are able to effectively manage the workload of infected players upon their return. Playing fewer minutes post-infection is most likely a decision made by management, along with input from the medical staff. Game time can be actively managed, so players do not suffer any adverse productivity effects following the illness. These results suggest that such a strategy works well in mitigating any enduring effects from COVID-19 infections. It is worth noting that clubs are experienced at managing their players back from illness and injury – from a managerial perspective, clubs will likely have fine-tuned recovery procedures in place.

5.2. Team level effects

As many of our results on difference-in-difference coefficients at the player level demonstrate lack of statistical significance, and those that are statistically significant are practically inconsequential, we do not expect to find any output effects at a team level from players returning to action after infections. That said, this remains a pertinent line of inquiry following Lichter et al. (2017). We address this by estimating a model similar to Eq. (1) but instead of individual performance statistics, we use team outcomes (win, draw or loss) and performances (number of goals scored, number of goals conceded, and goal difference).[8] As controls, we include opposition strength and a home team dummy. The results are shown in Table 6. As expected, the difference-in-differences coefficient is not significant, indicating that although teams may field a recovered player, there is no significant change in team outcomes.[9]

There are various explanations why we find no team level effects. First, football is increasingly characterised by systematised routines to safeguard against the problem of player absenteeism. Second, elite clubs now have deep squads and replacements are readily available. Third, in the event of significant club-level outbreaks it was common that organisers would suspend matches. Taken together, these likely explain why we fail to see a performance decline at team level .

5.3. Heterogeneous effects

We would like to observe the severity of a COVID-19 infection, or the exact variant type, since this could quite plausibly affect the performance of a player upon their return. Unfortunately, the severity of cases is not known, at least not in a form to allow systematic analysis. However, we can proxy severity from the length of time that a player is absent for. This measure is imperfect, since various events can interfere with the recovery period. For example, whether the infection was contracted during the lockdown break, or whether it occurred cross-season. In these cases, a player would be absent for longer for reasons other than their infection. We turn to these issues in the Robustness checks section. We split a player's treatment status according to the number of days absent. Specifically, we indicate if they are absent for 15 days or less (reference category), 15–30 days, 31–45 days, and longer than 45 days.[10] We then interact each treatment indicator with the *Post* dummy. The results in Table 7 show that performance per minute is no different, regardless of the length of the absence. This confirms that regardless of the severity of the infection, effective management of minutes on return to play can alleviate negative consequences of an infection.

6. Robustness checks

To determine the validity of our results, we conduct a battery of

[8] Note, we cannot adopt the approach of Lichter et al. (2017), as we do not observe performance measures for teammates of our treated and control. The best approximation to team level effects is to consider final outcomes.

[9] Modelling Goals Scored and Conceded as a Negative Binomial regression does not change results.

[10] Fifteen days is chosen as the minimum since a positive COVID-19 test typically mandated a circa 14-day quarantine period.

Table 6
Team level effects.

Outcome	(1) result *(Loss=1, Draw=2, Win=3)*	(2) result *(Loss=1, Draw=2, Win=3)*	(3) result *(Loss=1, Draw=2, Win=3)*	(4) result *(Loss=1, Draw=2, Win=3)*	(5) Own Score	(6) Own Score	(7) Opp Score	(8) Opp Score	(9) Goal Diff	(10) Goal Diff
Estimator VARIABLES	Ordered Logit	Ordered Logit	Ordered Probit	Ordered Probit	Poisson	Poisson	Poisson	Poisson	OLS	OLS
Treat	0.356***	0.386***	0.207***	0.226***	0.240***	0.256***	−0.013	−0.040	0.399***	0.463***
	(0.079)	(0.080)	(0.000)	0.048	(0.033)	(0.034)	(0.036)	(0.037)	(0.081)	(0.077)
Post	0.015	0.038	0.008	0.016	−0.023	−0.021	−0.027	−0.039	0.003	0.024
	(0.093)	(0.094)	(0.057)	0.057	(0.043)	(0.043)	(0.044)	(0.044)	(0.098)	(0.092)
Treat * Post	−0.111	−0.099	−0.061	−0.05	−0.032	−0.029	−0.015	−0.032	−0.045	−0.033
	(0.123)	(0.124)	(0.075)	(0.075)	(0.054)	(0.054)	(0.058)	(0.058)	(0.128)	(0.120)
ELO		−0.002***		−0.001***		−0.001***		0.001***		−0.004***
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Home		0.398***		0.231***		0.193***		−0.198***		0.563***
		(0.060)		0.036		(0.025)		(0.028)		(0.058)
Constant					0.342***	2.429***	0.293***	−2.314***	0.067	7.101***
					(0.025)	(0.164)	(0.026)	(0.147)	(0.057)	(0.352)
Observations	3971	3939	3971	3939	3972	3940	3973	3941	3972	3940
R-squared									0.009	0.129

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

robustness checks. These include a restriction on the number of minutes played, and a host of checks on the treatment group.

6.1. Minutes threshold & league checks

We estimate the difference-in-differences model in Eq. (1) with a restriction that players must play at least 10 minutes in each match, otherwise they are dropped from the sample. Such checks are necessary as (i) it can take players time to integrate with the rhythm of a match and (ii) fatigue does not affect performance from the first minute played. The results are shown in Panels A and B of Table 8. These results support and, potentially even strengthen, our previous claims that recovered players perform no different, statistically, to non-infected players. Moreover, compared to our baseline results in Table 5, Progressive Carries per minute now no longer demonstrates any significant difference after infection between the two groups. This hints that previous significant results might be driven by players playing few minutes. Our conclusion remains that players do not suffer any long-term effects of COVID-19, when effectively managed on return. Also, to crosscheck our results with those of Fischer et al. (2022), we restrict our sample to only the Bundesliga and Serie A. Results in Panel C of Table 8 show that all of our measures are now insignificant, again adding confidence to our results.

6.2. Checks on the treatment group

There are numerous confounds affecting a player's recovery period from COVID-19 infection. These include whether a player contracted COVID-19 during a lockdown break, or during the off-season (between the end of the 2019–20 season and start of the 2020–21 season). These confounds directly relate to our results in Section 5.3 exploring how the length of recovery period affected post-infection performances. On the one hand, players would have a longer recovery time before the next

Table 7
Treatment effect heterogeneity.

VARIABLES	(1) Presses per min	(2) Prog. Dist. per min	(3) Carries per min	(4) Prog. Carry per min	(5) Dribbles per min	(6) XG per min	(7) Pass Completion%
Post	0.004	0.076	−0.011	−0.001	0.002	0.000	0.183
	(0.008)	(0.077)	(0.029)	(0.004)	(0.002)	(0.001)	(0.952)
Treat (15–30)	0.013	0.184*	0.038	0.001	0.000	0.002	1.778
	(0.013)	(0.108)	(0.040)	(0.010)	(0.004)	(0.002)	(1.188)
Treat (31–45)	0.019	0.046	−0.029	−0.001	0.001	0.001	−0.058
	(0.015)	(0.149)	(0.036)	(0.011)	(0.005)	(0.002)	(1.722)
Treat (>45)	0.056***	0.126	−0.003	−0.004	−0.002	0.001	−0.248
	(0.013)	(0.094)	(0.032)	(0.007)	(0.003)	(0.001)	(1.018)
Treat (15–30) * Post	−0.002	−0.179	−0.007	0.002	−0.007*	−0.003	−0.331
	(0.011)	(0.123)	(0.049)	(0.007)	(0.004)	(0.002)	(1.239)
Treat (31–45) * Post	−0.006	−0.182	0.053	−0.010	−0.004	−0.002	2.704
	(0.016)	(0.131)	(0.040)	(0.008)	(0.004)	(0.002)	(2.219)
Treat (>45) * Post	−0.011	−0.071	0.026	0.000	0.002	0.000	−0.039
	(0.012)	(0.095)	(0.031)	(0.005)	(0.004)	(0.001)	(1.102)
Constant	0.217	3.729***	0.974***	0.142**	0.061*	0.002	99.290***
	(0.157)	(1.300)	(0.309)	(0.062)	(0.034)	(0.011)	(11.938)
Observations	3940	3932	3939	3931	3932	1047	3909
R-squared	0.090	0.054	0.076	0.062	0.039	0.035	0.112

Cluster robust standard errors in parentheses (clustered at the player level).
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$
Models include controls for Age, Age Squared, Opposition ELO, Europe, Home, Position Indicators, League Indicators.

Table 8
Minutes threshold & league checks.

VARIABLES	(1) Presses per min	(2) Prog. Dist. per min	(3) Carries per min	(4) Prog Carry per min	(5) Dribbles per min	(6) XG per min	(7) Pass Completion%
Panel A: No Controls							
Treat * Post	−0.001	−0.089	−0.026*	−0.001	−0.002	−0.001	−0.316
	(0.007)	(0.061)	(0.016)	(0.003)	(0.002)	(0.001)	(0.813)
Observations	3794	3784	3793	3783	3784	997	3787
R-squared	0.028	0.002	0.008	0.012	0.004	0.008	0.009
Panel B: Including Controls							
Treat * Post	0.002	−0.078	−0.019	0.001	−0.001	−0.001	−0.082
	(0.007)	(0.059)	(0.015)	(0.003)	(0.002)	(0.001)	(0.793)
Observations	3762	3754	3761	3753	3754	990	3755
R-squared	0.168	0.105	0.132	0.084	0.118	0.058	0.132
Panel C: Including Controls (Bundesliga and Serie A only)							
Treat * Post	0.015	−0.146	−0.028	0.000	−0.003	−0.000	−1.412
	(0.010)	(0.093)	(0.022)	(0.005)	(0.002)	(0.001)	(1.054)
Observations	1806	1801	1806	1802	1803	449	1801
R-squared	0.148	0.082	0.070	0.058	0.113	0.039	0.108

Cluster robust standard errors in parentheses (clustered at the player level).
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$
The models only include observations in which players played at least 10 min in a given match. Panel B&C Controls: Age, Age Squared, Opposition ELO, Europe, Home, Position Indicators, League Indicators.

game, and as such their performance may appear to be better compared to an infection that occurs out of this window. Alternatively, the longer a player is out for, the more likely it is they lose match sharpness and performance might suffer as a result. Other confounds might include transferring to another team or picking up another injury whilst recovering from COVID-19, though these are far less common.

To assess these possibilities, we split the *Treat* variable into four further categories, namely whether a player's COVID-19 infection occurred during the lockdown break, during the summer break, whether it coincided with transferring to a new team, and whether they picked up another injury during their recovery. In each case, we interact the treatment status with the *Post* indicator. The results are shown in Appendix Table A3 and point to very little difference compared to the baseline case of infection occurring in the absence of these confounds. The exception to this is that presses and progressive distance per minute are significantly lower if infection occurred during the lockdown break.

Performance might also suffer to a greater extent if players are infected as part of an outbreak within squads. We check whether 'clusters' of infections had more severe effects on player performances. Our observations, and their performance, may not be wholly independent of one another as several cases could occur within a team unit. A cluster is defined as three or more players on the same team testing positive in a match week. We carry out a check only on the treated group of players, by interacting a *cluster* dummy variable with the *post* indicator. Results should thus be interpreted as a comparison to infections that did not occur as part of a cluster. The results are shown in Appendix Table A4 and demonstrate there is no significant difference to performance whether infections occurred as part of a cluster or otherwise.

6.3. Matches threshold

We consider the implications of changing our pre- and post- infection periods to demonstrate that our results are not dependent on the length of these periods. In particular, results in Table A5 demonstrate the effects of extending the post-treatment period to seven games (Panel A), restricting the pre-infection period to 3 games (Panel B), restricting the post-infection period to 3 games (Panel C) and restricting both the pre- and post- infection periods to 3 games (Panel D). In each case, we run the models with and without the 10-minute thresholds in place. Regardless of the number of games pre- and post- infection, results and interpretations are nearly identical to our baseline models.

6.4. Scheduling

One may be concerned that the match as a unit of analysis fails to account for COVID effects decaying as time passes. There will be variation in the number of matches players can possibly play over a given period. For example, some players may take two weeks to play 5 matches, while others will take 2 months. Furthermore, on the return of football post-lockdown, matches were scheduled closer together to expedite the completion of seasons (with the exception of Ligue 1). These schedule changes could regulate performance and match-time.

To consider this, we accessed data on the date of each match, for all players. This has allowed us to ascertain 'rest days' between matches for each player in the treatment and control. As expected, these naturally vary across players based on scheduling demands. *Rest days* is a highly skewed measure (owing to the confounds discussed in Section 6.2), but the median number of rest days is 7, both in our pre- and post-infection periods. Since the calculation of rest days involves comparing game t with game t-1, in these models we will lose game 1 from our analysis since we have no game data prior to this game. We run the models with and without a 10-minute threshold in place, and results are shown in Appendix Table A6.[11] The results are virtually identical to those in Table 5 (Panel B) and Table 9 (Panel B), adding confidence to our initial results.

7. Discussion & conclusion

Our results offer a positive outlook regarding the performance of footballers following a COVID-19 infection. Contracting COVID-19 prevented footballers from carrying out their task for a short period but any decline in performance is typically short-lived. After controlling for the number of minutes played during matches, any observed performance declines are mostly not statistically significant. The practical implication of any significant effects is inconsequential. This result is consistent with the optimistic prior - footballers post-infection

[11] We also experimented using Rest days squared as an additional control variable, but results were unchanged.

performances are not adversely affected by COVID-19.

Players complete high intensity tasks shortly after returning. We conjecture that the best explanation for this result is that players minutes are managed upon their return - just as minutes would be overseen for any other injury.[12] This implies successful managerial decision-making strategies insofar as it concerns worker selection and management of working time and workload. In appealing to this explanation, it is important to acknowledge that the productivity of the individual is not fully autonomous; a strategic recovery plan is likely co-developed between medical staff, managers, and players.

Consistent with the literature, we find that infections are understandably not beneficial to workplace performance as players are incapacitated for a period of time. However, players do not suffer any per minute productivity losses in their recovery phase. This finding is consistent with Vandreuil (2021) but differs from Fischer et al. (2022) who cover fewer leagues and assess alternative measures of performance.

An obvious question is how comparable are COVID-19 infections to other injuries? We cannot precisely answer this due to the enforced quarantine period on positive COVID cases. One cannot distinguish how long it takes a player to recover as they are prevented from performing for a set period regardless of their medical status. Any attempt to compare infections to other injuries (e.g. groin or hamstring injuries) would prove futile due to this constraint.

Are there general practical implications of these results for operations research? The findings speak to managers supervising physically demanding jobs and are especially pertinent for industries facing work scheduling challenges. These types of occupations, while not requiring the same level of fitness as elite professional sport, are ubiquitous. Obvious examples include roles in construction, fitness industries and public services. The resultant policy implication is the need to train managers on strategies to reintegrate tired or recovering workers effectively. Our analysis would suggest that with appropriate schedule management an employee can optimise their output.

Several limitations of our study are noteworthy. The question of external validity remains. While there are many industries outside of professional sport that require physical fitness, we are keen to note the exceptionalism of our setting. Also, we cannot measure the potency or variant of infection and can only offer proxy measures. Consequently, we cannot fully distinguish the toll of the virus. While we introduce controls to explain performance there is almost an infinite potential set of in-match contexts that can regulate minutes. It is not feasible to control for each of these.

The findings do raise numerous future questions. With more data on infections and vaccination, one can assess the impact of inoculation on recovered performances. One notable medical finding is that COVID-19 infections can cause scar tissue, which can lead to an insufficient ability to build up mucus coating in the lungs post-infection. This could render players susceptible to future lung infections. We only assess elite professional leagues. Evaluating lower leagues would be advantageous to check if the findings generalise to players contracted to clubs with limited medical resources. We do not consider the impact of COVID-19 by sex or race. Both topics could be explored. Considering our question in other sports, particularly those that rely solely on individual performance, is a potential avenue for future research.

Acknowledgements

The authors wish to express their thanks to seminar participants at University College Dublin Geary Institute, NUI Maynooth, University College Cork, Lancaster University, University of Sheffield and four anonymous reviewers.

Appendix

Appendix Table A1
Raw performance measures.

VARIABLES	(1) Minutes	(2) Presses	(3) Progressive Distance	(4) Carries	(5) Progressive Carry	(6) Dribbles	(7) Pass Completion%	(8) xG
Treat * Post	−9.360***	−1.366**	−16.347***	−5.216***	−0.519**	−0.174	−0.289	−0.065*
	(1.749)	(0.558)	(4.475)	(1.192)	(0.217)	(0.107)	(0.836)	(0.034)
Observations	3940	3940	3932	3939	3931	3932	3909	1047
R-squared	0.139	0.151	0.132	0.151	0.062	0.095	0.109	0.090

Cluster robust standard errors in parentheses (clustered at the player level).
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$
All models include controls for Age, Indicators, League Indicators Age Squared, Opposition ELO, Europe, Home, Position.

Appendix Table A2
Control Variables.

VARIABLES	(1) Presses per minute	(2) Prog. Distance per minute	(3) Carries per minute	(4) Prog. Carry per minute	(5) Dribbles per minute	(6) xG per minute	(7) Pass Completion %
Age	−0.007	−0.088	−0.014	−0.003	−0.001	0.000	−1.300
	(0.010)	(0.084)	(0.022)	(0.004)	(0.002)	(0.001)	(0.854)
Age Squared	0.000	0.001	0.000	0.000	0.000	−0.000	0.026*
	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.016)
Opposition ELO	0.000**	−0.001***	−0.000***	−0.000***	−0.000	−0.000**	−0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Europe	−0.013	0.266***	0.141***	0.019***	−0.003	0.000	5.257***
	(0.010)	(0.058)	(0.018)	(0.004)	(0.002)	(0.001)	(0.700)
Home	0.005	0.078**	0.025**	0.005**	0.003	0.002***	1.049***

(continued on next page)

[12] This explanation has been specifically stated in the post-match press conferences by managers.

Appendix Table A2 (*continued*)

VARIABLES	(1) Presses per minute	(2) Prog. Distance per minute	(3) Carries per minute	(4) Prog. Carry per minute	(5) Dribbles per minute	(6) xG per minute	(7) Pass Completion %
	(0.005)	(0.040)	(0.010)	(0.002)	(0.002)	(0.001)	(0.358)
DEF	−0.099***	0.092	−0.009	−0.019***	−0.017***		2.945***
	(0.009)	(0.060)	(0.018)	(0.004)	(0.002)		(0.730)
FWD	−0.026**	−0.509***	−0.159***	−0.025***	0.003	0.002***	−6.432***
	(0.013)	(0.082)	(0.022)	(0.004)	(0.003)	(0.000)	(0.955)
France	0.019	−0.016	0.017	−0.002	0.003	−0.001	0.885
	(0.020)	(0.094)	(0.028)	(0.005)	(0.003)	(0.001)	(1.157)
Germany	0.004	0.041	0.003	0.009	−0.001	0.001	−1.489
	(0.010)	(0.091)	(0.020)	(0.007)	(0.003)	(0.001)	(1.096)
Italy	0.013	0.109	0.041*	0.006	0.001	0.000	0.219
	(0.009)	(0.082)	(0.023)	(0.005)	(0.002)	(0.001)	(0.896)
Spain	−0.001	−0.096	−0.041**	−0.005	0.003	−0.001	−2.096*
	(0.012)	(0.100)	(0.020)	(0.004)	(0.004)	(0.001)	(1.145)
Constant	0.284*	3.879***	0.982***	0.136**	0.058*	0.004	99.502***
	(0.150)	(1.293)	(0.308)	(0.062)	(0.035)	(0.011)	(11.899)

Cluster robust standard errors in parentheses (clustered at the player level).
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$.

Appendix Table A3
Covid recovery confounds.

VARIABLES	(1) Presses per minute	(2) Prog. Distance per minute	(3) Carries per minute	(4) Prog. Carry per minute	(5) Dribbles per minute	(6) xG per minute	(7) Pass Completion %
Treat (lockdown) * Post	−0.059**	−0.353**	−0.067	−0.012	−0.005	0.002	−0.066
	(0.023)	(0.173)	(0.049)	(0.009)	(0.007)	(0.003)	(1.726)
Treat (summer break) * Post	0.002	0.004	−0.029	0.004	−0.006	−0.003*	−1.640
	(0.014)	(0.104)	(0.029)	(0.005)	(0.004)	(0.002)	(1.031)
Treat (other injury) * Post	−0.013	−0.061	0.064	−0.015	0.002		2.468
	(0.077)	(0.570)	(0.162)	(0.030)	(0.023)		(5.644)
Treat (transfer) * Post	0.019	−0.084	0.091	0.003	−0.009	0.000	1.878
	(0.029)	(0.216)	(0.061)	(0.011)	(0.009)	(0.003)	(2.157)
Observations	3940	3932	3939	3931	3932	1047	3909
R-squared	0.081	0.058	0.077	0.063	0.046	0.049	0.116

Cluster robust standard errors in parentheses (clustered at the player level).
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$
All models include controls for Age, Indicators, League Indicators Age Squared, Opposition ELO, Europe, Home, Position.

Appendix Table A4
Cluster of infections.

VARIABLES	(1) Presses per minute	(2) Prog. Distance per minute	(3) Carries per minute	(4) Prog. Carry per minute	(5) Dribbles per minute	(6) xG per minute	(7) Pass Completion%
Cluster * Post	−0.022	0.164	−0.052	0.003	−0.004	−0.002	−0.102
	(0.021)	(0.144)	(0.046)	(0.009)	(0.005)	(0.003)	(1.332)
Observations	2205	2205	2204	2205	2205	481	2178
R-squared	0.053	0.033	0.050	0.048	0.057	0.035	0.086

Robust standard errors in parentheses.
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$
All models include controls for Age, Indicators, League Indicators Age Squared, Opposition ELO, Europe, Home, Position.

Appendix Table A5
Changing pre- and post- infection periods.

	(1) Presses per minute	(2) Prog. Distance per minute	(3) Carries per minute	(4) Prog. Carry per minute	(5) Dribbles per minute	(6) xG per minute	(7) Pass Completion%
Panel A1 (7 games post)							
Treat * Post	−0.007	−0.168**	−0.029	−0.005	−0.006	−0.002*	−0.862

(*continued on next page*)

Appendix Table A5 (*continued*)

	(1) Presses per minute	(2) Prog. Distance per minute	(3) Carries per minute	(4) Prog. Carry per minute	(5) Dribbles per minute	(6) xG per minute	(7) Pass Completion%
	(0.011)	(0.074)	(0.019)	(0.004)	(0.004)	(0.001)	(0.809)
Panel A2 (7 games post with 10 mins restriction)							
Treat * Post	0.001	−0.089	−0.025*	−0.000	−0.001	−0.001	−0.617
	(0.007)	(0.058)	(0.015)	(0.003)	(0.002)	(0.001)	(0.770)
Panel B1 (3 games pre)							
Treat * Post	−0.017	−0.112	−0.015	−0.000	−0.004	−0.002*	−0.341
	(0.014)	(0.077)	(0.023)	(0.004)	(0.004)	(0.001)	(0.921)
Panel B2 (3 games pre with 10 mins restriction)							
Treat * Post	−0.003	−0.081	−0.019	0.003	−0.000	−0.001*	−0.248
	(0.007)	(0.066)	(0.017)	(0.004)	(0.002)	(0.001)	(0.880)
Panel C1 (3 games post)							
Treat * Post	−0.004	−0.132	−0.030	−0.002	−0.007	−0.001	−0.133
	(0.011)	(0.084)	(0.019)	(0.004)	(0.005)	(0.001)	(0.842)
Panel C2 (3 games post with 10 mins restriction)							
Treat * Post	0.004	−0.069	−0.019	0.001	−0.001	−0.000	−0.077
	(0.007)	(0.064)	(0.015)	(0.003)	(0.002)	(0.001)	(0.811)
Panel D1 (3 games pre and 3 post)							
Treat * Post	−0.017	−0.084	−0.025	0.002	−0.004	−0.001	−0.175
	(0.014)	(0.085)	(0.021)	(0.005)	(0.005)	(0.001)	(0.925)
Panel D2 (3 games pre and 3 post with 10 mins restriction)							
Treat * Post	−0.002	−0.072	−0.019	0.003	0.000	−0.001	−0.227
	(0.007)	(0.070)	(0.016)	(0.004)	(0.002)	(0.001)	(0.891)

Appendix Table A6
Inclusion of rest days as a control.

VARIABLES	(1) Presses per minute	(2) Prog. Distance per minute	(3) Carries per minute	(4) Prog. Carry per minute	(5) Dribbles per minute	(6) xG per minute	(7) Pass Completion%
Panel A: No minute's threshold restriction							
Treat * Post	−0.011	−0.129*	−0.020	−0.003	−0.005	−0.002**	−0.068
	(0.012)	(0.071)	(0.023)	(0.004)	(0.004)	(0.001)	(0.860)
Rest Days	0.000	−0.000	0.000	−0.000	0.000	0.000	−0.002
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.012)
Observations	3463	3457	3462	3454	3455	916	3434
R-squared	0.080	0.079	0.074	0.067	0.039	0.077	0.109
Panel B: 10-minute inclusion restriction							
Treat * Post	−0.002	−0.093	−0.024	0.000	−0.001	−0.001	0.051
	(0.007)	(0.063)	(0.017)	(0.003)	(0.002)	(0.001)	(0.822)
Rest Days	0.000	−0.000	0.000	−0.000	0.000	0.000	0.002
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.012)
Observations	3305	3299	3304	3296	3297	867	3298
R-squared	0.167	0.108	0.130	0.085	0.117	0.067	0.132

Cluster robust standard errors in parentheses (clustered at the player level).
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$
All models include controls for Age, Age Squared, Opposition ELO, Europe, Home, Position Indicators, League Indicators.

References

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics*. Princeton University Press.

Bernanke, B. S., & Yellen, J. L. (2020). *Former Fed Chairs Bernanke and Yellen testified on COVID-19 and response to economic crisis*.

Berri, D., Butler, D., Rossi, G., Simmons, R., & Tordoff, C. (2023). Salary determination in professional football: Empirical evidence from goalkeepers. *European Sport Management Quarterly*, 1–17.

Bryson, A., Buraimo, B., Farnell, A., & Simmons, R. (2021b). Time to go? Head coach quits and dismissals in professional football. *De Economist, 169*, 81–105.

Bryson, A., Dolton, P., Reade, J. J., Schreyer, D., & Singleton, C. (2021a). Causal effects of an absent crowd on performances and refereeing decisions during COVID-19. *Economics Letters, 198*, Article 109664.

Buraimo, B., Frick, B., Hickfang, M., & Simmons, R. (2015). The economics of long-term contracts in the footballers' labour market. *Scottish Journal of Political Economy, 62* (1), 8–24.

Berdan, C., Charumilind, S., Craven, M., Lamb, J., & Singhal, S.. One billion days lost: How COVID-19 is hurting the US workforce. McKinsey & Company. https://www.mckinsey.com/industries/healthcare/our-insights/one-billion-days-lost-how-covid-19-is-hurting-the-us-workforce.

Bureau of Labor Statistics. (2019). 4.2 million workers have illness-related work absences in January 2018. https://www.bls.gov/opub/ted/2018/4-point-2-million-workers-have-illness-related-work-absences-in-january-2018.htm.

Fischer, K., Reade, J., & Schmal, W. (2022). What cannot be cured must be endured: The long lasting effect of a COVID-19 infection on workplace productivity. *Labour Economics, 79*, Article 102281.

Fischer, K., & Haucap, J. (2021). Does crowd support drive the home advantage in professional football? Evidence from German ghost games during the COVID-19 pandemic. *Journal of Sports Economics, 22*(8), 982–1008.

Grinza, E., & Rycx, F. (2020). The impact of sickness absenteeism on firm productivity: New evidence from Belgian matched employer–employee panel data. *Industrial Relations: A Journal of Economy and Society, 59*(1), 150–194.

Kharrat, T., McHale, I. G., & Peña, J. L. (2020). Plus–minus player ratings for soccer. *European Journal of Operational Research, 283*(2), 726–736.

Keech, M., Scott, A. J., & Ryan, P. J. J. (1998). The impact of influenza and influenza like illness on productivity and healthcare resource utilization in a working population. *Occupational Medicine, 48*(2), 85–90.

Lichter, A., Pestel, N., & Sommer, E. (2017). Productivity effects of air pollution: Evidence from professional soccer. *Labour Economics, 48*, 54–66.

Lötters, F., Meerding, W. J., & Burdorf, A. (2005). Reduced productivity after sickness absence due to musculoskeletal disorders and its relation to health outcomes. *Scandinavian Journal of Work, Environment & Health, 31*(5), 367–374.

McHale, I. G., & Holmes, B. (2023). Estimating transfer fees of professional footballers using advanced performance metrics and machine learning. *European Journal of Operational Research, 306*(1), 389–399.

Nabavi, N. (2020). Long covid: How to define it and how to manage it. *British Medical Journal, 370*. Available at https://www.bmj.com/content/370/bmj.m3489.

Nagurney, A. (2021). Supply chain game theory network modeling under labor constraints: Applications to the COVID-19 pandemic. *European Journal of Operational Research, 293*(3), 880–891.

Premier League (2021) Latest statement on results of COVID-19 tests. Available online: https://www.premierleague.com/news/1814863.

Vaudreuil, N. J., Kennedy, A. J., Lombardo, S. J., & Kharrazi, F. D. (2021). Impact of COVID-19 on recovered athletes returning to competitive play in the NBA "bubble. *Orthopaedic Journal of Sports Medicine, 9*(3), Article 23259671211004531. PMID: 33855099.

Wagemans, J., Catteeuw, P., Vandenhouten, J., Jansen, J., De Corte, X., Ceusters, C., et al. (2021). The impact of Covid-19 on physical performance and mental health—A retrospective case series of Belgian male professional football players. *Frontiers in Sports and Active Living, 3*, Article 803130.

Wezenbeek, E., Denolf, S., Bourgois, J. G., Philippaerts, R. M., De Winne, B., Willems, T. M., et al. (2023). Impact of (long) COVID on athletes' performance: A prospective study in elite football players. *Annals of Medicine, 55*(1), Article 2198776.

Weimar, D., & Wicker, P. (2017). Moneyball revisited: Effort and team performance in professional soccer. *Journal of Sports Economics, 18*(2), 140–161.

World Health Organisation. (2019). Absenteeism from work due to illness, days per employee per year. Available at https://gateway.euro.who.int/en/indicators/hfa_411-2700-absenteeism-from-work-due-to-illness-days-per-employee-per-year/visualizations/#id=19398.

World Health Organisation (2023) WHO Coronavirus (COVID-19) Dashboard. Available at https://covid19.who.int/.