Effort Peer Effects in Team Production: Evidence from Professional Football

Danny Cohen-Zada* Itay Dayag† Naomi Gershoni‡

October 19, 2022

Abstract

We exploit a unique dataset from the Israeli Professional Football Leagues which provides high-frequency, direct measures of players' effort in order to estimate effort peer effects in a high-skill, collaborative team task. Employing two complementary identification strategies, we find robust evidence of substantial positive peer effects. Our findings highlight that effort spillovers play an important role in team production and that even a change in just one worker's effort can substantially influence team effort and thus performance. Moreover, we present suggestive evidence that behavioral considerations are a relevant mechanism for creating peer effects even in highly skilled teams of workers.

JEL Codes: J22, J44, D91, L23, L83.

Keywords: peer effects, effort, team, performance, productivity, complementarity, social pressure, contest theory, football.

*Ben-Gurion University of the Negev and IZA

†Ben-Gurion University of the Negev

‡Ben-Gurion University of the Negev and IZA.

We thank Ran Eilat, Lukas B. Freund, Tomer Ifergane, Moshe Justman, Alex Krumer, Zvika Neeman, Matthew Notowidigdo, Aner Sela, and Offer Shapir for helpful discussions and suggestions, and Hadas Stiassnie for superb research assistance. We are also grateful to the Associate Editor and the four referees for their excellent comments which helped us greatly improve the paper.

1 Introduction

How workers determine their effort level and how they can be motivated to increase effort are fundamental questions in labor economics. Both classical theory and modern efficiency wage theory posit that higher wage rates and wage structure can drive workers to exert more effort (see, e.g., Akerlof, 1982; Shapiro and Stiglitz, 1984; Akerlof and Yellen, 1990).[1] The efficient determination of wages is even further complicated as a growing share of firms report that their work is mostly performed in teams with the tendency to shift organizational structure away from hierarchies towards "networks of teams" (see, e.g., Deloitte Global Human Capital Trends, 2019). This trend has challenged organizations to change their management practices and to reward employees based on their contributions rather than their job titles. However, when production calls for collaboration and individual efforts are complementary, assessing workers' individual contributions and designing optimal incentives in teams requires exploring how teammates affect each other (Arcidiacono et al., 2017). This exploration is important for testing the different theoretical predictions and for potentially demonstrating that efforts can be motivated not only through direct monetary incentives but also by peers.

Since this is not a trivial task, despite immense interest, the empirical literature on peer effects in work teams is relatively scarce in addition to being ambivalent (Cornelissen et al., 2017). Furthermore, although economic models draw a clear distinction between agents' behavior as reflected by their effort and their outcomes as reflected by their productivity, empirical studies aiming to analyze economic behavior regularly focus on workers' productivity since in most settings effort cannot be directly observed (Malueg and Yates, 2010). While using performance (productivity) as a measure of effort might be considered a reasonable approach for low-skill jobs where variation in ability is relatively small, it is less so for high-skill jobs where it is essential to account for the fact that performance is the product of both effort and ability. Consequently, researchers commonly control for imperfect but necessary proxies of ability when they study performance or exploit panel data to account for permanent ability via individual fixed effects (see, e.g., Guryan et al., 2009; Malueg and Yates, 2010).[2]

Clearly, productivity is the most important outcome for the firm and therefore studies that address performance and ability peer effects help inform managers on optimal team or work-group composition. Directly addressing effort peer effects provides an additional dimension by which managers can potentially increase team productivity. Specifically, for a given team composition, if one worker's effort positively affects her teammate's effort, incentivizing effort on an individual level could lead to substantial productivity gains. However, if an increase in one's effort causes

[1]For empirical studies that examine the relationship between wage and effort see Fehr and Goette (2007); Raff and Summers (1987); Sandvik et al. (2020).

[2]Malueg and Yates (2010) use betting odds (BO) in tennis to identify players with similar skill levels and Guryan et al. (2009) use past scores of golfers adjusted for the difficulty of the golf course.

teammates to shirk, the policy implications could be different.

In this paper, we exploit high-frequency individual level data from the Israeli Professional Football Leagues on both effort and performance of professional football (soccer) players, in order to measure *effort* peer effects among team workers who perform a high-skill collaborative task. To the best of our knowledge, this is the first attempt to estimate peer effects using a direct measure of effort rather than proxying it by output. Using effort allows us to obtain more accurate estimates since it directly captures the player's choice following a change in peer behavior. This is especially important in a contest setting where player and team performance depend not only on their own efforts and ability but rather also on those of the opponent team as well as on additional circumstances such as weather, crowd, or media attention. Even if all of these factors were observed, winning the match (or scoring a goal) is a probabilistic outcome, namely, chance plays a role and thus output cannot be predicted with certainty for given inputs as in other production processes (see e.g. Peeters and Szymanski, 2014). Moreover, when there is complementarity in team production, an increase in peer effort may increase a worker's output even if he does not change his own effort.[3] Thus, an estimation of peer effects on performance may capture spurious correlations if effort is not controlled for. Another interesting feature of the football setting is that performance measures, such as number of goals or wins, are much less frequently observed relative to effort measures such as running speed and intensive runs. In addition, common performance measures are irrelevant for large parts of the team (e.g., goals are not an appropriate measure of defenders' performance). Other performance indicators are just difficult to measure at the individual level (e.g., winning a game or preventing an opponent goal).

These features, which are also prevalent in other types of work teams such as research and development units, court litigation teams, political lobbying groups, and marketing divisions, further enhance the advantage of directly focusing on effort when aiming to analyze individual team members' behavior or to measure their contribution. However, effort typically cannot be observed in other work environments, whereas highly detailed data are available about sports teams and offer unique opportunities to explore intricate production functions.[4] A prominent example is Arcidiacono et al. (2017) who use data on basketball team composition and performance for each possession to separately estimate the players' own productivity and their productivity spillovers in order to better understand individual contributions to the team. Sports contests also offer a substantial advantage over laboratory settings because, similarly to other workplace environments, they involve high-stake decisions of professional agents in the regular course of business.

In general, several mechanisms are expected to operate among workers in teams that would create a causal relationship between their efforts. First, complementarity or substitutability in team production could lead to either positive interactions or negative ones ("free riding") as has been

[3]The opposite would happen when there is substitutability between players.

[4]For a general discussion of this advantage see Kahn (2000).

shown by Gould and Winter (2009) for professional baseball players.[5] Second, positive spillovers among workers can be driven by behavioral mechanisms such as social pressure, prosocial behavior, or shame, even when the task performed is an individual one (Mas and Moretti, 2009; Brune et al., 2022). Third, the incentive structure set by the employer could create interdependent effort choices among coworkers, e.g., when individual rewards are determined according to group output (Bandiera et al., 2009; Amodio and Martinez-Carrasco, 2018). Lastly, in some settings, positive effort peer effects were shown to be the result of learning and knowledge spillovers (Jackson and Bruegmann, 2009; Azoulay et al., 2010). The latter two mechanisms are probably less influential in the football setting where players' efforts are not manifested in their market values (Wicker et al., 2013; Weimar and Wicker, 2017) and where, similarly to Guryan et al. (2009), players' ability to learn during a game is quite limited.[6]

The effect of the first mechanism depends on whether in football, players' efforts are complements or substitutes, which may not be obvious. While it is commonly claimed that a football team is more than the sum of its parts, implying that teammates' efforts are complementary, it is not entirely clear whether this is a general truth or an aspiration. The numerous studies and non-academic discussions aiming to identify the specific components that elevate teams to higher levels of collaboration and synergy seem to suggest that their lack, renders team production not necessarily complementary (see e.g., Brave et al., 2019; Levine, 2015; Lechner and Gudmundsson, 2012). It is also quite reasonable to expect substitutability among football players in the sense that when some players push harder and "cover more ground", they enable their teammates to rest more without substantially lowering team performance. The most pronounced example is when players exert more effort to compensate for a teammate slowing down due to injury or for a missing player following a red card (penalty).[7] Relatedly, both sports' experts and academic scholars constantly debate the contribution of superstar players to teams, asking whether it is individual performance or collaborative effort that wins games. Michael Jordan's famous saying - "There's no 'i' in team but there is in win" - expresses his view on this subject in basketball, a game which similarly to football requires coordination between teammates. More generally, substitutability between team members' efforts characterizes production processes in which only the best individual execution or the most skillful technique matters, as can frequently be the case in a football contest.[8]

[5]In the labor economics literature, the term "peer effects" is not used to describe peer effort interactions due to complementarity in production. Rather, it refers to effects created through the behavioral or learning channels which will be described below. We aim to estimate the combination of "peer effects" as defined in the labor literature and of the effects stemming from production complementarity.

[6]More generally, Cornelissen et al. (2017) present evidence indicating that the role of learning is less dominant relative to that of the other mechanisms.

[7]The fact that the popular fantasy football computer games implicitly assume perfect substitution between players' individual inputs illustrates that complementarity is not necessarily perceived as the dominant factor in team production.

[8]In more traditional, non-sports contexts, substitutability is expected in problem solving tasks which require just one worker to come up with a good idea making the efforts of his coworkers redundant (Grossman and Maggi, 2000; Garicano and Hubbard, 2012). Substitution may also be more prevalent when efforts are determined sequentially as

Still, collaboration and coordination can elevate a team's performance and even make up for the lack of individual talent, as the following commentary on Barcelona's star player's move to Paris Saint-Germain demonstrates: "Only a collective effort can replace Lionel Messi at Barcelona".[9] This quote further emphasizes that, as with other complex and high-skill team tasks, it is not straightforward to expect either complementarity or substitutability in the production of a football team.

Overall, the different mechanisms imply that we should expect to find positive interactions between teammates' efforts if football players' efforts are mostly complementary. Positive interactions between teammates have also been the most prevalent finding in previous studies of peer effects in the workplace (Herbst and Mas, 2015). According to Manski (1993), this similarity in the behavior of individuals and their peers could be attributed to three distinct effects: an endogenous effect, which is the effect of peer behavior on their own behavior; an exogenous (contextual) effect, which is the effect of peer attributes on individual behavior; and correlated effects, which refer to similarity in behavior due to exposure to common shocks or circumstances. The endogenous effect is the most interesting one when considering incentivizing and managing effort in teams as it is the only effect that embeds a "social multiplier" in the sense that there are repeated feedbacks between teammates' efforts. Thus, even encouraging just a relatively small increase in the efforts of just a few workers could generate a substantial increase in the effort of the entire team. However, the endogenous effect is also the most challenging one for empirical estimation due to the well-known "reflection problem" which is caused by the simultaneous determination of individual behavior (Manski, 1993; Bramoullé et al., 2020).[10] Consequently, most studies of peer effects focus on estimating exogenous effects or some combination of the two.

In order to isolate the endogenous effect, we employ two different estimation strategies. Although we cannot completely rule out that our estimated effects mix endogenous and exogenous (contextual) effects, we minimize the presence of the latter by controlling for any observed peer characteristics. Furthermore, an additional analysis based on the methodology proposed by Bramoullé et al. (2009) and on alternative assumptions about peer group composition supports our conclusions that teammates' *efforts* substantially and positively affect each other.

The first strategy that we implement follows Gould and Winter (2009) by relying on a combination of individual worker fixed effects and an instrumental variable (IV) approach in which peers' contemporaneous behavior is instrumented by their average behavior during the season. The fact that our data provides direct quantification of individual effort allows us to add to their findings

in swimming relay races (Neugart and Richiardi, 2013). A football match involves both simultaneous and sequential effort choices which are difficult to separate.

[9]This is a quote from Albert Ferrer, a former football player and coach, as reported by the football news website FourFourTwo on August 21st, 2021 `https://www.fourfourtwo.com/features/albert-ferrer-only-a-collective-effort-can-replace-lionel-messi-at-barcelona`.

[10]We explain this problem in more detail when we discuss our empirical strategy in section 4.2.1.

regarding *performance* and directly address peer effects in *effort*. To establish that running distance is a relevant and reliable measure of effort, we run a preliminary analysis demonstrating that this measure is positively, significantly, and robustly related to output. To estimate this relationship at the individual level, we include in the regression equation both individual, game, and section fixed effects, while also controlling for numerous observed characteristics. At the team level, we exploit the fact that each team meets each specific opponent at least twice during a season to estimate how changes in team and opponent's efforts determine the probability of a given team to win the game within each pair of teams.[11]

Beyond the fact that we can directly quantify effort, the high-frequency of our data which reports individual efforts for every five-minute segment of each game (hereinafter "section"), allows us to derive our estimates based on the events of one season, as opposed to relying on variation between seasons in a player's career as in Gould and Winter (2009). This is especially advantageous when relying on players' fixed effects in order to control for unobserved ability which might cause an overestimation of the peer effect due to correlation between teammates' abilities. Clearly, it is more plausible to assume that a player's ability is constant throughout one season rather than throughout his entire career. Another advantage is that it allows us to exploit within-game variation in team composition and effort and thus to increase the efficiency of the instrument.

In this analysis, the dependent variable is the running distance of a specific player in a given five-minute section and the main explanatory variable is the average running distance of all the other team members during the same section, which is instrumented by their average effort in the *same section* of the game throughout the season. This approach should overcome concerns that common shocks during the specific section similarly impacted all the players' efforts. In addition, our model includes individual player, game, and section number fixed effects and a rich set of controls such as fatigue (cumulative time on the field), ex-ante winning probability (based on BO), crowd size, previous section events, and the score gap. Also, to refute the concern that the correlation between players' running distance is driven by changes in the coach tactics during the game, we control for the formation of the team, namely, the number of players in each position (defenders, strikers, and midfielders) and for major events that occurred in the previous section. We believe that within-game tactical changes are usually triggered by such events or by changes in the opponent team's effort which is also controlled for in several ways. We take the stance that conditional on this rich set of control variables and fixed effects, the average seasonal effort of the specific composition of teammates impacts the running distance of a player in a specific section only through the actual instantaneous effort that these teammates exert, which implies that the exclusion restriction holds. Importantly, to ensure that the estimated effect focuses on the endogenous effect, all the observed characteristics of peers are controlled for, including age, tenure,

[11]To corroborate and validate our findings, this team level specification is also estimated using data from the Union of European Football Associations (UEFA) Champions League.

height, and aggresiveness. The main findings show that a player's effort is highly responsive to the efforts of his teammates, so that, on average, a player will increase his running distance by 78% of the increase in the average team effort. This estimated effect may still include some contextual effects stemming from unobserved peer characteristics. However, we show that these confounding factors will most likely bias the effect downwards.

To reinforce our findings, we implement a second strategy that focuses only on the sections just before and after substitutions and analyze how a gap in the running distance between the last full section of the outgoing fatigued player and the first full section of the incoming fresh player impacts the change in other players' effort between these two sections. This approach relies on a preliminary analysis that supports our premise that fatigue is an important determinant of running distance and therefore substitutions will typically generate a positive shock to effort which could plausibly be considered exogenous simply because the incoming player is less tired relative to the outgoing one. Nevertheless, correlated effects could still bias the estimated peer effects, even when we condition on observables and control for unobservables via player, game and section fixed effects. Therefore, we instrument contemporaneous effort levels of the outgoing and incoming players with their corresponding average seasonal effort when playing the same cumulative time in the game. The results indicate that even a change in the effort level of only a single player can substantially impact the efforts of his teammates. These results are robust to the exclusion of injury-induced substitutions, of substitutions in sections just before or after major events, and of substitutions which changed the formation of the team. The estimated effect is somewhat larger than the group peer effect that we estimate using the first strategy, mainly because it captures the immediate response to a change in effort rather than an average impact of changes in team composition over several sections throughout the game.

Our predictions and results are consistent with Gould and Winter (2009) who show that baseball players' measures of performance interact positively when they are complements in production and negatively when they are substitutes. In the football context, it is plausible to assume strong complementarity in team production and therefore to expect positive effort spillovers among players as we indeed find. In terms of magnitude, our estimated effects are larger, which could be related to the specific attributes of each type of sport but may also indicate that effort interactions are stronger than performance ones which are moderated through changes in players' effort. Arcidiacono et al. (2017) also focus on performance, providing evidence for the importance of positive spillovers among players in the production of a basketball team which is similar to football in the sense that the game requires high levels of collaborative effort. They further highlight potential heterogeneity by estimating player specific spillovers.

To shed light on the specific mechanisms that drive the positive interactions between players' efforts, we estimate their heterogeneity along several player, team, and game characteristics. While complementarity in production is probably a significant force as depicted by the finding that midfielders are the most responsive to changes in other players' effort, we also find that more senior players are less responsive to peer behavior. The latter holds both for players with more experience (proxied by age) and for players with higher tenure. Generally, complementarity in production would predict that higher tenure would either increase or not affect the tendency to cooperate. Therefore, these findings suggest that the behavioral mechanism accounts for at least part of the positive effort peer effects. This assertion is supported by the fact that team-level tenure does not significantly affect the results.

The paper proceeds as follows. In section 2, we describe the data and provide background on the setting. Section 3 presents a simplified theoretical framework describing the main forces that determine rational effort choices in a football match. Section 4 first presents evidence on the relation between effort and output and then reports our main results on how teammates' efforts affect each other. Section 5 concludes.

2 Data and Background

2.1 Data and descriptive statistics

Our data relies on publicly available records for the games that were played during the 2017/2018 season as provided by the Israeli Professional Football Leagues (IPFL).[12] Here, we describe only the most relevant institutional details of the league and the games, while a more detailed description is provided in the Appendix, sections 6.1 and 6.2.

The data were coded from documents that include a detailed description of events and actions during a game, including goals, red and yellow cards, substitutions, and injuries. A unique feature of our data is that the running distance and the number of sprints are recorded for each player *in each five-minute section* of the game. This provides an explicit and close to continuous measure of players' effort, which in turn allows us to study how players change their efforts (input), rather than performance (output), within-games. In addition, for most games, we have data on the home and away crowd, teams' rank in the league (before each round), and the BO for the specific game.[13] Team ranking and BO are used to proxy for the relative ability of the teams in each match. We have also collected data on players' characteristics including their age, height, and tenure.[14]

Table 1 describes the full data and presents summary statistics for the main samples. The league includes 14 teams that play against each other in the regular season in two rounds of 13 games each

[12]The IPFL is a private non-profit organization that is responsible for the two professional football leauges in Israel (first and second divisions), and operates in collaboration with the Israeli Football Association (IFA). It is a member of the European Professional Football Leagues (EPFL).

[13]Data on team ranking after each round were downloaded from the Channel 5 website (Israel's leading sports channel) `https://www.sport5.co.il/liga.aspx?FolderID=44`. BO for each game were obtained from the Odd-sPortal.com website `https://www.oddsportal.com/soccer/israel/ligat-ha-al-2017-2018/results/`.

[14]The source of these data is `https://www.transfermarkt.com/`. Height data was supplemented from google player cards and additional online sources for approximately 20% of the players.

and then the league splits into the top and bottom playoffs. Overall, in a season there are 240 games, where teams in the top and bottom playoffs play 36 and 33 rounds, respectively. The records that we use are available for 95% of these games (a total of 228 games or 456 team×game combinations).[15] The total number of players who participated in the specific season is 365, and they are classified into four categories by their position: strikers, midfielders, defenders, and goalkeepers. The number of players in each category is reported in the top panel of Table 1.

Table 1: Summary Statistics

Total	
Rounds	36
Games	228
Teams	14
Players	365
Strikers	88
Midfielders	137
Defenders	111
Section level	**Mean**
Number of Player×Section observations	101,295
Number of Team×Section observations	9,120
Full Sections (by Player, Share)	.97
Overtime Sections (Share)	.1
Player's Distance (only full, non-overtime sections)	.51
	(.133)
Player's Sprints (only full, non-overtime sections)	2.16
	(1.635)
Strikers and Midfielders' Goal Probability (only full, non-overtime sections)	0.0098
	(0.098)
Team Distance (only non-overtime sections)	5.60
	(.702)
Game level	**Mean**
Number of Player×Game observations	6,308
Number of Team×Game observations	456
Number of Sections Played (by Player)	16.05
	(5.84)
Player's Avg. Speed (kmph)	6.29
	(1.23)
Team Distance	106.79
	(4.165)
Home Crowd	4,747.96
	(5,355.21)
Away Crowd	1,054.54
	(1,736.62)
Red Card Probability (any team)	0.25
	(0.431)

Each game is divided into two halves. In each half, there are nine five-minute sections and

[15] We are not aware of any systematic or specific reason for missing data that could potentially impact our analysis.

an overtime section that can take either less or more than five minutes, amounting to a total of 101,295 observations on players in a specific section, of which 10% are overtime sections (section numbers 10 and 20). Approximately 3% are partial sections, meaning that the player either entered or was taken out of the game during this section due to a substitution or a red card. In most of our section-level analyses, we disregard partial and overtime sections to avoid distortions.

A key variable in our analysis is the running distance of each player in each section. This distance is measured in kilometers and rounded to units of 0.1 (100 meters or 328 feet). Table 1 shows that the average running distance for a player in a five-minute section is 510 meters (or 1673 feet). This measure is obviously impacted by the player's number of intensive runs or sprints, the average of which is 2.16 per section. Notably, there is substantial variation in the distance that players run during a section, ranging between 0 and 1 kilometers, with more than 11% of observations at 300 meters or below and more than 10% at 700 meters or above. The average team distance in a section can be calculated as the average distance per player times the number of players. However, because distances are rounded to 0.1 units, we use an independent measure for team distance as appears in the original documentation. Thus, the average team distance is approximately, but not exactly, 11 times the average player distance — 5.6 km (3.5 miles).

Another interesting statistic that we report is the probability of a personal goal in a given section. This is calculated for strikers and midfielders only, since defense players are much less likely to score goals (as this is not their main objective). For these players, there is an almost one percent chance of scoring in each specific section.

At the game level, there are 6,308 observations for individual players who play an average of 16 sections per game.[16] Here, as well, we use game level data from the original documentation on average player-speed and on team distance rather than calculations based on the section level data, which are rounded to 0.1 units and are thus distorted especially for partial and overtime sections. A player's average speed is 6.29 kmph (3.9 mph) and the total distance that the entire team runs in a game is on average 106.79 kilometers (66.36 miles).

Other important variables at the game level are the size of the home and away crowds. As expected, the home team crowd is, on average, almost five times larger than the away team crowd, and there is substantial variation in these variables. However, we note that crowd data are missing for 12% of the sample (28 games). Therefore, in our regressions we present specifications with and without this control variable to demonstrate the robustness of the results.

[16] The distribution of the total number of sections that each player played in a game is highly skewed, with almost 60% playing for the entire 20 sections. This is due to the rule that allows a maximum of three substitutions per game for each team.

2.2 Team composition

Our analysis of peer effects substantially relies on within-game changes in team composition. Therefore, we now briefly describe the institutional features that are relevant to the choice of the initial team composition and its adjustment during the game.

Each team in our data has an average pool of 26 players: 6 strikers, 10 midfielders, 8 defenders and 2 goal keepers. The 11 players that actually participate in each section of every game are chosen from this pool by the team's coach according to strategic consideration, before and during the game. In line with their tactical goals, coaches usually choose a formation (number of players in each position) that is more offensive or defensive.

However, this choice is substantially constrained by the rule which allows a maximum of three substitutions to be initiated during each specific game, and thus deters frequent adjustments to within-game occurrences and changes in conditions or expectations (as is the case, for example, in basketball). Another constraint for team composition could be a players' absence or a disability due to injury. In addition, referees may penalize players for misconduct by dismissing them from the game without replacement, so that the team continues to play with less players (such a penalty is indicated by a red card).

Although constrained, players' selection cannot be thought of as random and therefore, in the analysis, we exploit our unique high-frequency data to add multi-level fixed effects and to efficiently control for within-game events. We are also able to distinguish between compositional changes that involve a formation change and those that do not. Further specific details and identifying assumptions are discussed in Section 4 below, where we introduce the empirical methodology.

3 Theoretical Framework

To inform our empirical analysis, we present a relatively simple theoretical model that illustrates the main forces and mechanisms operating in a football match.[17] The main feature of the model is that the teams compete for a given reward which cannot be shared. Each member of each team chooses how much effort to exert in the production of the team's joint output. The outputs of the different teams determine which team wins the reward according to a known (probabilistic) contest success function (CSF). Although we focus on football, the framework is not limited to sports applications, and may fit other contexts such as an R&D patent race, where two (or more) R&D teams from different firms compete over the same innovation, and only the first to invent or to patent (depending on the legal regime) will profit from the innovative product or technology.

In our model, two teams $j \in \{A, B\}$ compete in a Tullock contest, where players simultaneously

[17]We do not model the entire league as a contest but rather focus on one specific game because our data covers one season and we are mainly interested in the interaction between individual players' efforts rather than the seasonal inputs of the teams (e.g., player acquisition, contracts, training facilities).

choose how much effort to exert. For simplicity, we assume that each team consists of two players $i \in \{1, 2\}$ whose joint production is described by a standard CES function:

$$y_j = \left(x_{1j}^{-\gamma} + x_{2j}^{-\gamma}\right)^{-\frac{1}{\gamma}} , \qquad \gamma > -1 , \tag{1}$$

where x_{1j} and x_{2j} are the efforts exerted by players 1 and 2 in team j, respectively, and y_j is team j's production. In the context of a football match, production represents outputs such as ball possession, shots, assists, and tackles, all of which contribute to the probability of scoring goals and eventually winning the game. The CES function allows us to consider different levels of complementarity among players as higher levels of γ imply a lower elasticity of substitution between players' effort, given by $\eta = \frac{1}{1+\gamma}$. In particular, $\gamma > 0$ when teammates' efforts are gross complements ($\eta < 1$), and as γ approaches infinity the production function converges to the Leontief production function $y_j = Min[x_{1j}, x_{2j}]$, where players' efforts are perfect complements ($\eta = 0$).

The game outcome depends on both teams' production, where the probability of team j winning the game, denoted π_j, is given by the following CSF:

$$\pi_j = \frac{y_j}{y_A + y_B} . \tag{2}$$

This functional form, which is widely used in the literature on contests, exhibits the feature that each team's winning probability increases in its players' efforts and decreases in the effort of the opponent players. [18]

Exerting effort is costly. Denote by c_{ij} the cost that player i of team j bears for each unit of effort he exerts. Thus, players choose their effort level to maximize the following payoff (utility) function:

$$V_{ij} = \pi_j \cdot \omega - c_{ij} \cdot x_{ij} , \tag{3}$$

where ω is the value of winning, which is assumed to be common for all players.[19]

Therefore, the first-order condition (FOC) of the maximization problem of player 1 in team A is

$$\frac{\partial V_{1A}}{\partial x_{1A}} = \frac{\omega \cdot y_B}{(y_A + y_B)^2} \left[x_{1A}{}^{-\gamma} + x_{2A}{}^{-\gamma}\right]^{-\left(\frac{1}{\gamma}+1\right)} x_{1A}^{-(\gamma+1)} - c_{1A} = 0 . \tag{4}$$

The FOCs for the other three players are symmetric. These conditions show that the optimal effort of each player responds to changes in the effort exerted by his teammate as well as by the players

[18]Furthermore, for any value of y_j, changing own effort has the largest effect on the winning probability, i.e., $\frac{\partial \pi_j}{\partial y_j}$ is maximal, when the winning probability is one-half ($y_A = y_B$). This implies that teams will gain more from an increase in their players' efforts as the opponent's production becomes more similar to their own production. We also note that we focus on winning as the desired outcome of the game and pool together the options of lose and draw. This is both because of the non-linear structure of the league score (see the Appendix for more details) and because considering three different game outcomes substantially complicates the analysis without providing any significant contribution to our understanding of the economic forces at play.

[19]We assume a linear costs structure which is prevalent in the contest theory literature.

of the opponent team.[20]

Performing some simple algebraic manipulations on the FOC (eq. 4) we obtain that

$$x_{1A} = \frac{y_B^{\frac{1}{1+\gamma}} y_A}{(y_A + y_B)^{\frac{2}{1+\gamma}}} \cdot \frac{\omega^{\frac{1}{1+\gamma}}}{c_{1A}^{(1+\gamma)}} \, . \tag{5}$$

Then, dividing Equation (5) by the symmetric expression for x_{2A} yields the following relationship between teammates' efforts:

$$\frac{x_{2j}}{x_{1j}} = \left(\frac{c_{1j}}{c_{2j}}\right)^{\frac{1}{(1+\gamma)}} \quad \forall\, j \in \{A, B\} \tag{6}$$

Thus, the ratio of efforts is inversely related to the ratio of effort costs, but less so as the elasticity of substitution decreases (γ increases). This is because when efforts are more complementary, even when one teammate's effort cost is substantially higher, his effort cannot be completely substituted by that of his peer. At the extreme case of perfect complements ($\gamma \to \infty$), $x_{1j} = x_{2j}$ regardless of players' individual costs. The intuition is straightforward — increasing just one player's effort above the effort of his peer does not change the team's production y_j and thus does not affect the probability of winning π_j.

Substituting (6) into (1) for team A we obtain:

$$y_A = x_{1A} \left[1 + \left(\frac{c_{1A}}{c_{2A}}\right)^{\frac{-\gamma}{(1+\gamma)}}\right]^{-\frac{1}{\gamma}} \, . \tag{7}$$

Then, based on (7) and (5), and since these two conditions are symmetric for team B, we can define the ratio between the teams' production as a function of the cost parameters, and denote it by α:[21]

$$\frac{y_B}{y_A} = \left[\frac{c_{1A}^{\frac{\gamma}{1+\gamma}} + c_{2A}^{\frac{\gamma}{1+\gamma}}}{c_{1B}^{\frac{\gamma}{1+\gamma}} + c_{2B}^{\frac{\gamma}{1+\gamma}}}\right]^{\frac{1+\gamma}{\gamma}} \equiv \alpha \, . \tag{8}$$

The relative production of each team decreases with the costs of its own players' effort and increases with those of the opponent. Finally, substituting (7) and (8) into (5) yields the equilibrium level of effort for player i in team j, to be

$$x_{ij} = \frac{\omega\alpha}{c_{ij}^{\frac{1}{1+\gamma}} \left(c_{1j}^{\frac{\gamma}{1+\gamma}} + c_{2j}^{\frac{\gamma}{1+\gamma}}\right) (1+\alpha)^2} \tag{9}$$

Equation (9) shows that the effort of each player is determined by the cost parameters of both

[20] It is straightforward to verify that the second-order condition for maximization holds for any parameter value.

[21] Substituting 5 into 7 yields $(y_A + y_B)^2 = \frac{\omega \cdot y_B (c_{1A}^{-\frac{\gamma}{1+\gamma}} + c_{2A}^{-\frac{\gamma}{1+\gamma}})^{\frac{1+\gamma}{\gamma}}}{c_{1A} \cdot c_{2A}}$. Equation 8 is then obtained by deriving the symmetric equation for team B and dividing team A's equation by team B's.

his teammates and the opponent players, and that this relationship is influenced by the level of complementarity in production (γ). In addition, the player's effort positively depends on the value of winning (ω). This equation further implies that a player's effort could be related to his peer's effort through several different mechanisms. Next, we discuss these mechanisms and relate them to our estimation strategy.

3.1 Mechanisms and predictions

In light of the fact that our empirical strategy relies on changes in team composition and in players' fatigue within games, we first address how player i is expected to respond to such changes. In our model, these factors are captured by the cost of peer effort. For example, we expect a substitution to change the peer effort cost parameter both because the incoming player is less tired than the outgoing one, and because different players may have different effort costs (which reflect different abilities). From the FOC of team A (Equation 4) we can see that player 1 does not directly consider player 2's cost in order to determine his optimal effort level. However, player 1's optimal effort is indirectly affected through the effort chosen by player 2 (which is the effect that we aim to estimate). We therefore proceed by discussing how the efforts of each player in team A and of the opponent team (y_B) change in equilibrium when the cost of effort changes exogenously for player 2 in team A.

While it is easy to see that under perfect complementarity ($\gamma \to \infty$) teammates' efforts will exhibit a perfect positive correlation when one player's effort cost changes, in more realistic scenarios of imperfect complementarity, the analytical solution is harder to interpret. Thus, in Figures 1(a) and 1(b) we present numerical solutions for two specific values of γ demonstrating that even when complementarity levels are quite modest, team A's players' efforts mostly change in the same direction when only player 2's cost changes.[22] Furthermore, Appendix Figure A1 shows that positive peer effects may be found even when the elasticity of substitution is slightly larger than one, which implies relatively low complementarity (or high substitutability) in team production. The simulations also show that the opponent players' efforts decrease and thus y_B decreases, moving in the same direction as team A's efforts and production when c_{2A} increases above 1.[23]

Based on these simulations and under the plausible assumption that football teammates' efforts are mostly complementary, the causal effect of a change in peer effort cost and consequently in his effort on player i's effort (namely, the endogenous effect) is expected to be positive.[24] The

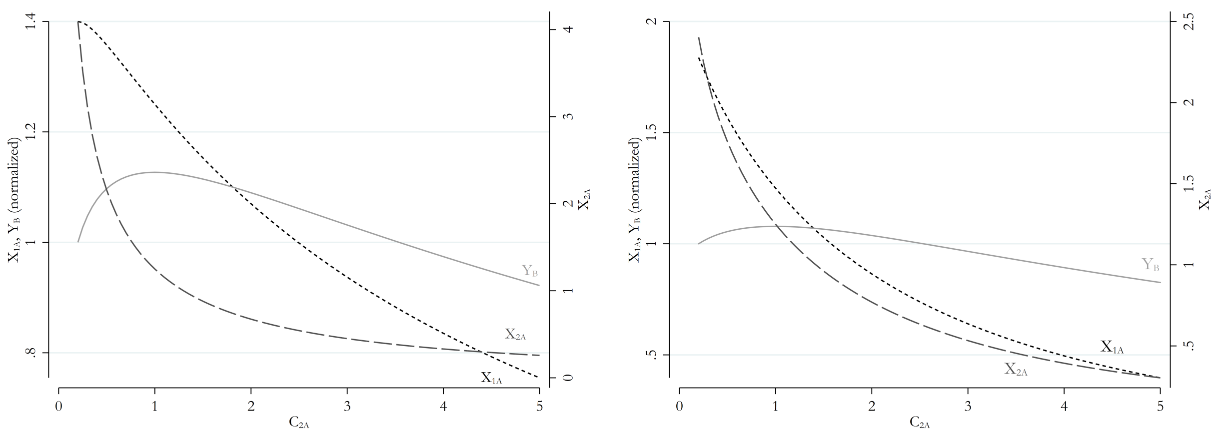
[22]In all of our simulations, we set ω to be equal to 10 and assume that the cost of effort is equal to 1 for all other players. This choice only affects the level of effort but not the relative efforts or the qualitative effect of the considered change in one player's cost.

[23]The opponent's effort is the highest at $c_{2A} = 1$, namely, when the effort costs and thus the efforts of all the players are identical. The reason for this is that, as explained above, the functional form of the CSF makes it most beneficial to increase team production when $y_A = y_B$.

[24]We also note that the specific value of γ may vary across teams and team formations, which is not depicted by our simple model. However, we will account for this option in our estimations.

Figure 1: The effect of changing one player's effort cost on other players' efforts (simulation)

(a) $\gamma = .5$ (b) $\gamma = 5$



Notes: The figures present simulation of the efforts of team A's players and the output of team B for different values of c_{2A} and for the following constant parameter values: $c_{1A} = c_{1B} = c_{2B} = 1, \omega = 10$. The value of γ is specified in the title of each figure. In order to show the changes in all the variables in the same figure we present x_{1A} and x_{2A} on a different scale and normalize the values of y_B. We note that due to this choice, the lines for x_{1A} and x_{2A} do not cross when $c_{2A} = c_{1A} = 1$ although $x_{1A} = x_{2A}$.

simulations further highlight that this effect can be decomposed into two channels of impact. The first, direct channel, stems from complimentarity in team production which makes it beneficial for player i to increase his effort when his peer's effort increases due to a reduction in cost. The second, indirect channel, is related to the CSF and operates through the effect of the peer's effort on the opponent team's effort which in turn affects player i's efforts. Our estimation should capture both channels. Controlling for the contemporaneous effort of the opponent blocks the second channel and may thus bias our estimates. Stated differently, since the opponent's effort itself is an outcome of the treatment, it is considered to be a "bad control".

On the other hand, positive interaction of peer efforts may also result from an exogenous change in the *opponent* players' effort costs which similarly affects player i and his teammate. This type of interaction is not considered to be a peer effect but rather, according to Manski's definition, a correlated effect which our estimation will aim to clean. To reduce this concern, we will control for the opponent's effort in the previous section and for the average seasonal effort of the opponent team in the same section. The first helps to account for the specific behavior of the opponent in the current game and the second addresses confounders that are related to the typical opponent's effort at different stages of the game.

Other potential sources of correlated effects which should be addressed by our estimation strategy are changes in game conditions which simultaneously affect the effort costs of all the players or the value of winning ω. For example, worsening weather conditions or fatigue in the later stages of the game may increase effort costs, and teams' rank in the league or the round number could affect the value of the game.

While our model captures mechanisms that relate players' efforts through the production function and the contest structure, it does not account for exogenous peer effects or for potential behavioral mechanisms which may underlie endogenous peer effects. One way to consider behavioral mechanisms within our simple framework is that a player's cost negatively depends on his teammate's effort due to social pressure. That such considerations operate in the context of a football team is supported by the sports economics literature which often views "team spirit" as an important element in production (see e.g., Peeters and Szymanski, 2014). Intuitively, if an increase in peer effort decreases the cost of own effort by increasing the social cost of maintaining the current effort level, then players' efforts will tend to positively affect each other regardless of the level of complementarity in production. In fact, behavioral considerations lead to the same effect as complementarity but through the negative influence of teammates' effort on the player's marginal cost rather than through a positive influence on his marginal productivity.

4 Empirical Estimation and Results

4.1 Running effort and performance

Before proceeding to our main analysis of peer effects, we first explore how changes in effort, as measured by running speed, relate to performance, as measured by either the number of goals or the game outcome. This preliminary analysis has two main objectives. First, we aim to show that running speed (or distance) is an important factor of productivity.[25] Second, the quantitative estimates of this analysis will allow us later on to approximate how a change in team effort due to interactions between teammates increases individual and team performance.

In this analysis, in order to minimize the endogeneity problem, we add multiple fixed effects and a rich set of covariates that account for observed and unobserved determinants of performance. Team, individual, and pair fixed effects control for time-invariant ability and relative ability, game fixed effects account for game specific unobservables, and section fixed effects control for within-game trends.[26] In addition, we utilize BO data to absorb any pre-game information about team ability that is observed by the booking agencies but not by the econometrician.

[25] Although speed appears to be a very natural measure of effort, players' unobserved ability and other dimensions of effort which cannot be measured (such as level of concentration or pre-game physical and mental preparation) may impact both speed and performance.

[26] While we believe that these three levels of fixed effects are highly relevant, our results are robust to the exclusion of any of them. This refutes any concerns for potential biases related to multi-level fixed effects models. In particular, the recent literature discusses potential biases in two way fixed effects models focusing on binary treatment specifications where fixed effects are estimated using both treated and un-treated units (see e.g., Imai and Kim, 2021; De Chaisemartin and d'Haultfoeuille, 2020; Gardner, 2021). These concerns only apply if treatment effects are heterogeneous which may not be the case in our setting. Moreover, the proposed adjustments for the binary treatment setting cannot be applied when treatment is continuous as in our case since none of the observations are "untreated" (i.e., treatment equals zero).

In the individual level analysis, we first use data on each section in each game in order to relate players' average speed during a section to the probability of a personal goal.[27] We limit our sample to include only strikers and midfielders and ignore defense players who we assume have a much lower probability of scoring a goal.[28] Additionally, for each player we include only full five-minute sections, namely we exclude extra-time-sections (sections 10 and 20), sections with substitutions for the incoming and outgoing players, and sections in which a player was removed from the game due to a red card. Using this sample, we estimate the following equation:

$$goal_{i,s,g} = \alpha_i + \zeta_s + \gamma_g + \theta \cdot avg\_speed_{i,s,g} + X'_{i,s,g}\delta + \epsilon_{i,s,g} \tag{10}$$

where $goal_{i,s,g}$ indicates whether player i scored a goal during section s of game g. $X_{i,s,g}$ is a vector of the observed game,[29] section, and player characteristics, including an indicator for player i's team being the home team, the rank gap between player i's team and the opponent team *before* game g, the score gap at the beginning of section s, and the number of sections that the player played in the game up to section s. The latter variable controls for fatigue and includes a quadratic term to allow for a non-linear increase in players' fatigue as they play more sections. In one specification, we also control for the ex-ante probability of a win for player i's team which is calculated based on BO (as in Buraimo and Simmons (2009)).[30] We also add controls for the average speed of the team and for the following events that may have occurred during the previous section (section $s-1$): player i scored a goal, player i's team scored a goal, the opponent team scored a goal, a player was injured, a player was given a red card. An additional control that was shown to be important in previous studies is the size of the crowd for each team. However, we only include it as a robustness test due to the fact that crowd data are missing for four rounds in the season. In any case, since our main specification includes game fixed effects, crowd size only varies across teams in the same game.[31]

We also add individual fixed effects to control for unobserved player characteristics such as ability which are held fixed over the entire season. Since only 26 players (out of 365) moved between teams during the season, the set of individual fixed effects also essentially controls for

[27] A player's speed can be easily transformed to distance by simply dividing the speed in kmph by 12 to calculate the distance per 5 minutes. This is the measure that we use in most of our analysis. The only reason that we use speed instead of distance is that players do not necessarily play for the entire 20 sections of the game. Thus, speed is a more reliable measure at the *player* $\times$ *game* level.

[28] Moreover, many of the few goals that they do score follow a free kick and thus are not impacted by their running speed.

[29] Game level variables that vary across the two teams who participate in the game will not cause multicollinearity.

[30] Although this variable carries valuable information, since the BO data are missing for one game (approximately 400 observations), in order to avoid using a selected sample, our main results are based on regressions that omit this control. However, including these controls and running the same regressions on the smaller sample yields practically identical estimates and significance levels. We also note that when game fixed effects are included, the probability of a draw (which is equal for both teams) is omitted due to perfect multicolinearity.

[31] As we include game fixed effects, we are able to control for either the home crowd or the away crowd but not for both.

team fixed effects.[32] Section fixed effects account for confounding factors that relate to the stage in the game (e.g. team fatigue), and game fixed effects account for game-specific characteristics that may impact both effort and performance. Standard errors are clustered two-way: at the player and at the game level.[33]

The results of the estimation are reported in Table 2 and show that, *within* player and conditional on section and game fixed effects, increasing running speed is substantially and robustly associated with an increase in goal probability. Although our estimation strategy accounts for most potential confounding factors, we are extra cautious about interpreting these results as causal since we cannot rule out the possibility that unobserved choices and actions taken by the player during a section are correlated with both his running speed and his goal performance. For example, it may be the case that players increase their speed when they are more concentrated in the game or more determined to succeed, and hence our estimates reflect the combined impact of speed and concentration. If so, we are indeed overestimating the impact of speed. Nonetheless, as concentration is another dimension of effort, we still have a reliable measure of how changes in effort, more broadly defined, affect performance.

In columns (1)-(4), we report the estimated coefficient of interest θ for specifications with different sets of controls, as indicated at the bottom of the table. The number of observations decreases as we add controls for previous section events (since the observations for the first section in each game are dropped) and for crowd and BO (due to missing data). Nevertheless, the estimates are remarkably stable across specifications. The quantitative interpretation of our main specification (column 3) is that a one standard deviation increase in a player's speed (1.23 kmph) increases his probability of scoring a goal during this one section by approximately 0.36 percentage points, which is an almost 37% increase in this probability (0.98%, as reported in Table 1). In columns (5) and (6), we show that a similar result is obtained when our preferred measure of effort is replaced by the number of intensive runs that a player engages in during a section or when both measures are included in the same regression. Obviously, these measures are highly correlated as the number of sprints is practically embedded within the average speed and therefore it is difficult to separate the contribution of each of these measures to the outcome.

Lastly, column (7) presents the same estimate at the game level and confirms that the level of effort is positively and significantly related to the goal probability.[34] Notably, the game-level

[32]Our results are also robust to including both team and player fixed effects.

[33]The modeling approach suggests that clustering should be applied at the level in which random unobserved shocks may lead to correlation in treatment assignment (Abadie et al., 2017). In our setting, treatment is assigned at the individual player level in the sense that each player's speed in specific games and sections is drawn from a pool of options limited by his ability and characteristics. Clustering at the player level also accounts for potential serial correlation in the data. In addition, the assignment of treatment is related to the specific game since different games have different conditions, opponents, and values of winning. As our theoretical framework clearly demonstrates, these factors may generate correlation between the potential treatment assignment of players who participate in the same game.

[34]The regression equation includes the same basic controls as in equation (1). Obviously, when the observation

estimate approximately equals the section-level estimate times the average number of sections that a player plays in a game, which is 16, as reported in Table 1.

Table 2: Player's running speed and goal probability

	Dependent Variable: Goal Probability						
	Section						Game
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Speed (kmph)	0.0028***	0.0028***	0.0029***	0.0029***		0.0012**	0.040***
	(0.00051)	(0.00051)	(0.00053)	(0.00058)		(0.00050)	(0.0103)
Number of Sprints					0.0034***	0.0030***	
					(0.00045)	(0.00044)	
Observations	47,922	47,922	45,200	39,444	45,200	45,200	3,860
R^2	0.014	0.014	0.015	0.015	0.016	0.016	0.169
Controls		√	√	√	√	√	√
Pre-section Controls			√	√	√	√	
Crowd & BO Controls				√			
Player FE	√	√	√	√	√	√	√
Section FE	√	√	√	√	√	√	
Game FE	√	√	√	√	√	√	√

Notes: The table presents estimates at the player level. In columns 1—6 the unit of observation is at the *player* × *section* level, and all regressions include player, section and game fixed effects. In column 7 the unit of observation is at the *player* × *game* level, and the regression includes player and game fixed effects. Columns 2—7 include controls for player i's team being the home team and for the rank gap between player i's team and the opponent team before game g. In columns 2—6 we also control for a quadratic trend in the number of sections that the player played in the game up to section s, and in columns 3—6 we add controls for the score gap, goals, red cards, team's average speed, and injuries in section $s-1$. Column 4 includes controls for crowd size (which are missing for 12% of the games) and BO (which are missing for one game). Goalkeepers and defense players are excluded from the sample. Standard errors are clustered by player and game.

Next, we test how team effort changes the outcome of a game by estimating a game level specification of the following form:

$$win_{i,r} = \eta_p + \rho_r + \theta_1 team\_dist_{i,r} + \theta_2 opp\_dist_{i,r} + Z'_{i,r}\nu + \xi_{i,r} \tag{11}$$

where $win_{i,r}$ denotes the probability that team i wins in round r.[35] In order not to consider each game twice, within each pair of teams that compete against each other, p, we randomly define the team whose name is first according to the Hebrew alphabetical order as team i.[36]

In light of the fact that the outcome of the game does not vary at the game level, we do not

is at the game level, we cannot control for the number of section and pre-section events. We also note that adding crowd and BO controls decreases the number of observations but has no substantial impact on the results.

[35] As in our theoretical framework, we do not consider a draw as a favorable outcome due to the non-linear structure of the league score (see Appendix for more details).

[36] We obtain very similar results when we run the same specification on a sample with two observations per game (one for each team) including *team* × *pair* fixed effects and pair level clustered standard errors. However, since the standard errors are slightly smaller in this estimation, we present the more conservative approach in the paper.

use game fixed effects in this specification. Instead, we include two sets of fixed effects that should account for any relevant unobserved game-level factors that may remain in the error term even after controlling for our set of observed variables (including BO and crowd size which are included in some specifications). Round fixed effects, denoted by ρ_r, eliminate any bias stemming from characteristics that relate to the date of the game, such as weather, the competition (league) stage and status, and even the political or social climate. In addition, we use pair fixed effects, η_p, to control for unobserved qualities of the teams and of the specific match between teams. Thus, our estimation relies on variation in the team's effort when it faces the same opponent in different rounds during the season. As mentioned above, teams are matched with the same opponent between two to four times during a season. However, because 5% of the games are missing from our data, when we employ this strategy four games are omitted from the sample (singeltons).

The main explanatory variable of interest is $team\_dist_{i,r}$, which is the aggregate running distance of team i's players in round r. However, because the outcome of the game obviously depends on both teams' efforts, we include the aggregate distance of the opponent team in the equation as well. We expect the probability of team i to win the game to positively depend on its own effort and negatively on its opponent's. However, although this is the most flexible and easily interpretable way to explore the role of effort in a game, there is some concern that the high level of positive correlation between the teams' efforts will result in opposite signed coefficients that do not reveal the true effects. Therefore, we also use alternative specifications where the explanatory variable is either the ratio or the difference between these distances. In addition, we show that the coefficient of team i's distance is stable when omitting the running distance of the opponent team from the equation. In most of the estimations, we also control for the pre-game ranking difference between the two teams, the ex-ante probabilities of a win and a draw (calculated as previously from BO data), and for the game being a home game for team i.

The results in columns (1)-(4) of Table 3 indicate that the probability of a win increases with own effort and decreases with the opponent's effort. Similarly, the results in column (5)-(6) suggest that the probability of a win increases with a *relative* increase in effort. In column (6), in addition to using the distance ratio, we control for the sum of both teams' running distance, which is a measure for the intensity of the game (or its pace). It should be noted that although the probability of a win is expected to be symmetric due to the way we choose team i, more than 20% of the games end in a tie which could be correlated with the pace of the game. Nevertheless, column (6) indicates that the point estimate of the total distance variable is very close to zero and statistically insignificant.

Table 3 further shows that the estimates are remarkably stable when different sets of controls are added to the equation in columns (2) and (3), and even increase in column (4) when we exclude games in which red cards were issued and consequently distances were aggregated over a different number of players for each team. According to the estimate in column (2), increasing a team's aggregate running distance by one kilometer (0.62 miles), for a given opponent's effort, is associated with a 4.26 percentage point increase in the probability of a win. This amounts to a more than 10 percent increase relative to the approximately 40% average winning rate. In Appendix Table A1, we present additional robustness tests. First, in columns (1)-(2), we replace the pair fixed effects with two separate sets of fixed effects for the team and the opponent for the specifications that are presented in columns (2) and (6) of Table 3. Then, in column (3), instead of using the ratio between the distances of the two groups, we use the difference between them and show that the estimates remain stable. Lastly, in columns (4)-(5) we omit the opponent's distance and show that a team's own distance is positively and significantly associated both with the probability of winning and with the number of goals scored, regardless of the opponent's effort.

Table 3: Teams' running distance and game outcomes

	Dependent Variable: Team Winning Probability					
	(1)	(2)	(3)	(4)	(5)	(6)
Team Distance (km)	0.0437***	0.0426***	0.0407***	0.0707**		
	(0.0118)	(0.0118)	(0.0141)	(0.0283)		
Opponent Distance (km)	-0.0490***	-0.0484***	-0.0432**	-0.0853***		
	(0.0157)	(0.0157)	(0.0187)	(0.0308)		
Distance Ratio					4.719***	4.874***
					(1.268)	(1.361)
Total Distance (km)						-0.00301
						(0.00605)
Observations	224	224	180	152	224	224
R^2	0.693	0.703	0.731	0.759	0.703	0.703
Controls		√	√	√	√	√
Crowd & BO Controls			√			
Pair FE	√	√	√	√	√	√
Round FE	√	√	√	√	√	√

Notes: The table presents estimates at the *team* $\times$ *game* level, where team i is selected according to the Hebrew alphabetical order of the teams' names. All regressions include pair and round fixed effects. In columns 1—4, the main explanatory variable is the aggregate distance that team i's players run during the game. We control for the same measure for the opponent team. In columns 5—6, the main explanatory variable is the ratio between the aggregate distance of team i and its opponent, where in column 6 we also control for the sum of these distances. Columns 2—6 also include controls for the home team and for the rank gap between the competing teams before game g. Column 3 includes controls for crowd size which are missing for 12% of the games and for BO which are missing for one game. In column 4, the sample is restricted to games without red cards. Standard errors are clustered at the pair level.

To conclude the discussion on how effort, as measured by running distance, relates to performance, we repeat the estimation of equation 11 using data from the UEFA Champions League. We exploit data for two consecutive seasons (between 2017-2019), which increases the within pair variation in effort and thus also allows us to estimate specifications with both separate pair and year fixed effects and alternatively with *pair* $\times$ *year* fixed effects. The results in Appendix Table

A2 confirm that our findings are not unique to the Israeli league in that the estimates are very similar in magnitude and robust to the various alternative specifications.

4.2 The interaction between teammates' efforts

After demonstrating that there is a strong association between players' running effort and performance, both at the individual and team levels, we turn to the main question that this paper addresses, namely how the efforts of different players on the same team impact each other. These impacts are defined in Manski (1993) as *endogenous* peer effects, and they are of particular interest to employers, firms, and policymakers alike in light of the fact that they are iterative effects that may result in a large final impact even when the initial change is relatively small.[37] Therefore, contrary to most peer effects studies, we do not focus only on how individuals are impacted by their group but rather also analyze how they may be affected by a change in just one of its members' behavior.

While caveats can be pointed out in almost every empirical study that aims to identify causal effects based on observational data, it has been argued that causality is especially problematic in peer effects studies without random assignment (see, e.g., Angrist, 2014). Yet, even if peers are randomly assigned, peer efforts could be correlated due to exposure to common shocks that cause all workers to either increase or decrease effort (see, e.g., Guryan et al., 2009). Furthermore, addressing such correlated effects is still insufficient if the aim is to disentangle the effect of peer effort (endogenous peer effects) from the effect of peer characteristics (exogenous or contextual peer effects).

We employ several different estimation strategies to overcome the inherent obstacles and to increase the credibility of our conclusions. We discuss the weaknesses of each of the different applied methodologies and how we cope with them, and show that the findings remain stable across methods.

4.2.1 How group effort impacts individual effort

We begin by estimating how group effort impacts individual effort. We rely on within-game changes in team composition and relate player i's running distance to the running distance of his teammates (excluding goalkeepers), conditional on i's ability as captured by individual player fixed effects. As in most studies, we use the average of the team excluding player i as the measure for peer behavior, thus implicitly assuming linearity in the aggregation of effort. As pointed out by Gould and Winter (2009), the exclusion of player i, in addition to using individual fixed effects, verify that

[37] As noted above, in the labor literature, peer effects usually refer to interactions among workers' behavior that do not result from complementarity or substitutability in production. However, when we discuss identification, we use the terminology in Manski (1993) to describe the total effect of peer effort on a worker's effort which includes all potential channels of impact.

our estimation does not suffer from the reflection problem (Manski, 1993), in the sense that the outcome variable — individual players' efforts — cannot be aggregated to generate peers' average efforts (the explanatory variable). However, this strategy may still yield biased estimates of the interaction between players' efforts due to three additional factors.

First, in our setting, as in many studies of intra-group interactions, the identity of peers cannot be considered random due to three levels of endogenous selection: the pool of players that was chosen for the season, the composition of players who take the field in the beginning of each game, and substitutions which are initiated during the game. Game and player fixed effects can effectively control for the first two levels. Game fixed effects control for any factor that may affect the line up that coaches choose before the game begins, and thus shift the focus to within-game variation in team composition due to substitutions. The pool of potential peers is also fixed per player and is thus accounted for by our player fixed effects, which also eliminate concerns for an "exclusion bias" stemming from the fact that each player's pool of potential peers does not include himself (Guryan et al., 2009; Fafchamps and Caeyers, 2020). Notably, the individual fixed effects prevent upwards bias even if coaches typically match players with similar speed levels since identification relies on within-player variation.[38] Therefore, our main obstacle is the potential endogeneity of the within-game variation in peer composition. However, as pointed out by Blume et al. (2015), endogenous formation of social structures will only bias the estimated peer effects if information that may impact the player's behavior is also available at the stage when the composition of his peers is determined. Hence, this potential bias should be relatively small in the football setting where the rules allow each team a maximum of three substitutions during the game, and deter frequent adjustments to changing conditions or to any new information that is revealed during the game. Nonetheless, we control for all observable events that occurred before the relevant section and for intermediate game results (score gap), in addition to including section fixed effects. We also control for the formation of the team, which reflects the tactical choices of the coach, and distinguish compositional changes that involve a formation change from those that do not. Although we cannot entirely refute concerns for additional unobserved sources of systematic peer selection which also affect players' choices of effort, we can assess that they are minimal based on the stability of our estimates, as discussed in more detail below.

Second, even if team members were to be chosen randomly, common shocks that simultaneously affect all players may lead to correlation in effort which is not part of the peer effects that we aim to estimate. One strategy that has often been used to avoid this problem was to employ measures of permanent or lifetime behavior of peers instead of contemporaneous ones. We exploit this strategy and use the average season-level effort of player i's peers, while relying on variation in

[38] See Gould and Winter (2009) for a similar discussion of endogenous team composition in baseball where teams tend to acquire "strong batters at the expense of good pitchers" which leads to a negative correlation in performance between the two groups of players within teams.

team composition across games and sections within each game. Then, the coefficient on peer effort would measure how teammates' *typical* effort impacts player i's effort in a specific section, and hence the effect on player i's effort would not stem from the events of the section. A general problem with this solution is that in some settings the interpretation of the coefficient of interest is changed such that the estimates reflect a combination of both the exogenous and the endogenous peer effects. This is the case when the permanent behavior of peers directly affects worker i's current behavior, in addition to affecting it indirectly via the current behavior of his peers. However, in our setting, we do not expect the permanent behavior of peers to directly impact the current behavior of player i. A direct effect would mean that player i chooses to run faster because his peers *usually* tend to run fast during the season, regardless of their *current* running speed. We believe that this scenario is highly unlikely because even if a player would increase his speed in expectation of a faster game, such behavior would not be maintained if his teammates do not meet this expectation. It is therefore plausible to assume that teammates' season-level effort at the specific section affects a player's current running speed only through its impact on contemporaneous peer effort. Thus, the season-level measure can be used as an instrument for contemporaneous team effort. This instrument satisfies the exclusion restriction and allows us to net out any correlated effects.

However, as discussed by Blume et al. (2015); Bramoullé et al. (2020) among others, addressing correlated effects is insufficient for identifying the endogenous peer effects. Thus, our third challenge is to disentangle endogenous effects from contextual effects stemming from peer attributes which are both correlated with the instrument and affect player i's effort. Methods for separating the two types of effects rely on either variation in groups size (Boucher et al., 2014; Lee, 2007) or on an overlapping network structure (Bramoullé et al., 2009). Failing to do so would mean that we estimate a combination of the two effects, and since the direction of the contextual effects is not obvious a priori, it is unclear whether this is an overestimation or an underestimation of the endogenous peer effects.[39] Given our rich set of controls and fixed effects, and the fact that we are able to directly control for several observed peer characteristics, our estimates are expected to be biased only to the extent that unobserved peer characteristics that change within a given game and within a given team are correlated with peers' seasonal effort and with the instantaneous effort of player i. Nevertheless, we will infer the direction and size of the contextual effects that these characteristics generate and derive bounds for the endogenous effect (based on Altonji et al. (2005a)).

To further investigate this point, we will refine our definition of peers, utilizing knowledge of the specific role of each player in the team and assuming that each player's peers do not include all of his teammates but rather only teammates in specific positions. Under this alternative structure, the groups of peers only partially overlap each other, which allows us to instrument each player's

[39] For example, if a player's peers are more experienced, he may trust them more and decrease his effort level, or conversely, he may try harder to impress them and actually put in more effort.

contemporaneous peer effort by the permanent effort of *the peers of his peers* as suggested by Bramoullé et al. (2009). Assuming intransitivity, i.e., that the peers of player i's peers are not his peers, we can effectively avoid the estimation of contextual effects .

Our basic IV approach is very similar to the one used in Gould and Winter (2009) but with several important advantages. First, our high-frequency data allows using within season and game variation, which can increase the first-stage correlation, especially as our instrumental variable is calculated for a specific section in the game throughout the season. Second, controlling for unobserved characteristics that are fixed over time via player fixed effects is more compelling when the data comes from one season rather than several ones. Third, contrary to previous studies, we are able to directly measure effort and therefore estimate the interaction between players' *choices* rather than their outputs which may depend on additional factors.

Formally, we use the following equation to estimate the interaction between the efforts of player i and his teammates:

$$dist_{i,t,s,g} = \omega_i + \sigma_s + \lambda_g + \beta_1 d\bar{i}st_{-i,t,s,g} + X'_{i,t,s,g}\mu + \varepsilon_{i,t,s,g} \tag{12}$$

where the outcome variable is the running distance of player i of team t in section s of game g, and the main explanatory variable is the average running distance of his teammates in the same section. As in Equation 10, we include individual, game and section number fixed effects. X is a vector of covariates which includes the same observable characteristics as in equation 10 in addition to team formation, and the average of the following peer characteristics: tenure, age, height, and aggressiveness (measured by the seasonal probability that a player is issued a red or a yellow card during a game section).[40] Thus, β_1 measures how deviations from player i's seasonal average effort relate to changes in peer effort, within a given game, game stage (serial number of the section) and formation (more or less defensive or offensive), and conditional on all observed game and peer characteristics. The latter ensures that exogenous peer effects which are related to these specific characteristics are cleaned from β_1. In some of the specifications we also control for the opponent team's effort.

Still, events that occur during section s in game g could cause a correlated change in the efforts of teammates, which prevents a causal interpretation of β_1. Therefore, we will also estimate equation 12 using 2SLS and instrumenting $d\bar{i}st_{-i,t,s,g}$ with the average running distance of player i's peers in section s of all games in the season, namely, the peer's permanent effort.[41] This approach relies *only* on the variation in peer composition in section s across games which can be plausibly considered

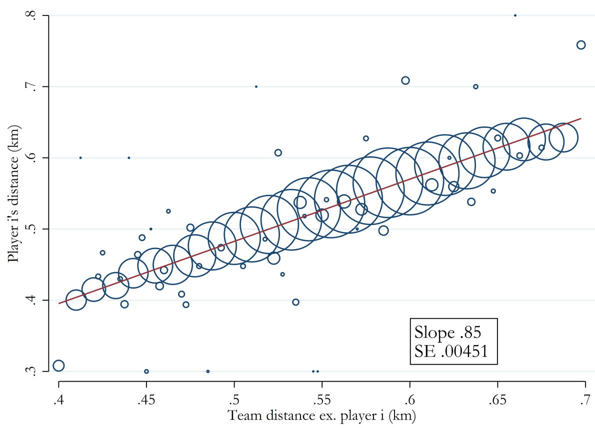
[40] We note that the same characteristics of player i are already controlled for through the individual fixed effects.

[41] If there was a substitution during section s of game g we include both the incoming and outgoing players' permanent effort in the calculation of the IV and use weights to account for the fact that each of them played a partial section. The weights that we use are set to 0.5 because we do not observe the specific number of minutes that each of these players played during the section.

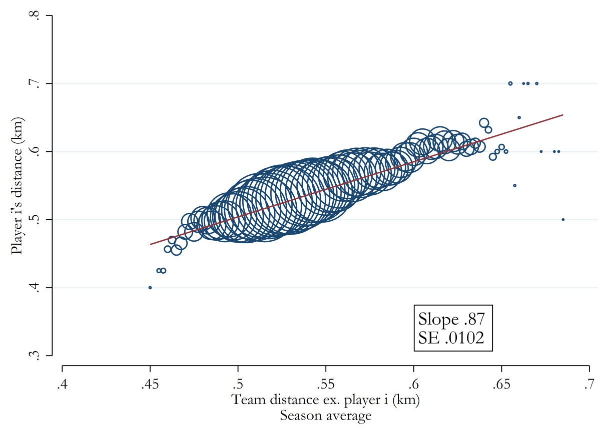
arbitrary conditional on the rich set of controls and multi-level fixed effects. As explained above, we expect the exclusion restriction to hold because teammates' effort levels in other games are unlikely to impact a player's current effort, unless through their positive correlation with teammates' current effort.[42] We further evaluate these assumptions and the validity of the instrument below.

Figure 2: Teammates' and Player's Efforts

(a) OLS



(b) Reduced form



Notes: Panel (a) plots the raw data with a linear fitted line for the relationship between player i's distance and the average distance of his teammates in a given section (excluding the goalkeeper). In panel (b), the x-axis variable is our suggested instrument — the seasonal average running distance of player i's teammates in the same section.

Panel (a) of Figure 2 presents the raw data correlation between the outcome variable and the endogenous explanatory variable $\bar{dist}_{-i,t,s,g}$ with a linear fitted line. In panel (b), we plot the reduced form relationship between the outcome variable and the instrument. Both plots show a substantial positive association which may be indicative of positive peer effects, but cannot be interpreted as causal because the confounding factors that we listed above are not accounted for. Thus, we next turn to regression results.

We first report the results of the OLS estimation in panel (a) of Table 4. As it is likely overestimating the effort interactions among peers, in panel (b) we present the results of the 2SLS estimation of the model, where peers' instantaneous effort is instrumented for by their permanent effort. The F-statistic on the excluded instrument in the first stage is presented at the bottom of the table and, as expected, indicates a very strong correlation between the instrument and the endogenous variable, which easily satisfies even the strict criteria that was recently suggested by Lee et al. (2020). In both panels, standard errors are clustered as in section 4 above at the player and game level.[43]

[42]This correlation is obviously expected to be monotonic.

[43]The same explanation for choice of clustering level applies to this estimation except that the treatment here is peer effort rather than own effort. Since each individual player has a unique set of peers, it is reasonable to assume that treatment assignment is correlated across observations for the same player.

As expected, the IV estimates are lower than the OLS ones. The estimated effect is positive, highly significant, and remarkably stable across specifications. In each of the panels, the point estimates remain practically the same as more controls are added or when games with red cards are excluded, implying that none of these characteristics and events systematically affect the interaction between teammates' efforts.[44] The only exception is in column (4), where we control for the opponent team's distance in the same section. However, as explained in section 3.1, this specification yields an underestimation since some of the impact of teammates' efforts on each other is channeled through the response of the opponent team. This makes the opponent's simultaneous effort a "bad control". At the same time, because teams respond to changes in each other's effort, not controlling for exogenous changes in opponent effort may bias the estimated effects. Therefore, in column (5) we control for the opponent's effort in section $s-1$ and for the opponent's seasonal average effort in the same section, both of which cannot be affected by our main explanatory variable. Controlling for the opponent's distance in the preceding section eliminates confounding factors that relate to within game changes in the behavior or tactics of the other team although with a lag of one section to avoid bias. At the same time, the opponent's average distance in the same section throughout the season provides a measure for the average fitness of the players and for their typical game style. For example, this measure will account for the fact that certain teams tend to "open fast" but later slow down due to fatigue. It should also be noted that the game fixed effects provide an additional control for elements of the opponent's strategy which remain constant throughout the game sections.

Quantitatively, the IV estimates indicate that an increase of 100 meters (328 feet) in the average distance that a player's teammates run during a given section will cause an increase of approximately 78 meters (255 feet) in the player's running distance during the same section. This effect is about 92% of the one estimated by OLS in panel (a), implying that a relatively small fraction of the correlation between teammates' running speed is due to within-section occurrences (namely, due to correlated effects). To further refine this conclusion, we should keep in mind that this model estimates local average treatment effects, which in our context means that it measures the interaction between own and teammate efforts when their contemporaneous effort is positively correlated with their permanent effort.

As explained above, although we control for observed peer characteristics, our estimated effect may still capture a combination of the endogenous influence of peer effort and the exogenous effects of unobserved peer characteristics that are correlated with the effort of both player i and his peers. However, the results in Table 4 indicate that the potential presence of such characteristics should not substantially affect the estimates and their interpretation. These results clearly show that the estimated peer effect is very stable across specifications with different controls. As we add

[44] In column (7) the sample is restricted to games in which no red cards were issued to player i's team in order to avoid distortions due to changes in the number of players on the field.

Table 4: Teammates' and Player's Efforts

	Dependent Variable: Player i's running distance						
							No Red cards
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A - OLS							
Peer distance	0.853***	0.855***	0.856***	0.501***	0.852***	0.851***	0.850***
	(0.00820)	(0.00829)	(0.00849)	(0.0144)	(0.00848)	(0.00880)	(0.00860)
Panel B - IV approach							
Peer distance	0.735***	0.780***	0.784***	0.428***	0.787***	0.781***	0.795***
	(0.0173)	(0.0152)	(0.0157)	(0.0398)	(0.0149)	(0.0164)	(0.0160)
F-statistic on excluded instrument	817.88	832.73	783.39	383.12	820.65	673.03	652.35
Observations	80,444	80,444	75,886	75,886	75,886	66,251	57,503
Controls		√	√	√	√	√	√
Pre-section Controls			√	√	√	√	√
Opponent Controls				√	√	√	√
Crowd & BO Controls						√	

Notes: The table presents OLS estimates in panel A and 2SLS estimates in panel B. The unit of observation is player × section. All regressions include player, section and game fixed effects. The main explanatory variable is the aggregate distance that all players except player i run during the section (excluding goalkeepers). In panel B it is instrumented by these players' average running distance in section s throughout the season. Columns 2—7 include controls for the peers' average characteristics: age (quadratic), height, tenure, and their aggressiveness (proxied by their average number of red and yellow cards in the season), and also for player i's team being the home team, the rank gap between the team and the opponent before game g, and a quadratic trend in the number of sections that player i played up to section s. In columns 3—7, we also add controls for the score gap, goals, red cards, and injuries in section $s-1$, and for the team's average speed in this section. The opponent's running distance is controlled for in columns 4—7 where in column 4 we control for the opponent players' average distance in section s of game g, and in columns 5—7, we control for this distance in section $s-1$ of the current game in addition to the opponent's average distance in section s throughout the season. Column 6 includes controls for crowd size (which are missing for 12% of the games) and BO (which are missing for one game). In column 7, the sample is restricted to games without red cards for player i's team. Goalkeepers are excluded from the sample. Standard errors are clustered by player and game.

controls, it only slightly *increases* from 0.735 in column (1) to 0.787 in column (5) where all the controls are included (except for crowd and betting odds which are missing for a substantial share of the sample). In particular, adding controls for observed peer characteristics hardly increases the estimated effect. For example, an identical specification as in column (5) but without these peer characteristics yields an estimated effect of 0.775, namely, the point estimate increases only by 0.012. As argued by Altonji et al. (2005a,b) and by Oster (2019), such stability indicates that if there is any bias due to omitted variables it is quite small. Moreover, they claim that the direction of the change in the point estimates as more controls are included can be informative about the direction in which unobservables are likely to shift the estimated effect. Thus, the results imply that the estimated effect presents a lower bound of the true peer effect. Furthermore, we implemented the method proposed by Altonji et al. (2005a) to assess the maximal potential omitted variable bias under the assumption that the amount of selection on unobservables is not larger than the amount of selection on observables. We found that the upper bound for the endogenous peer effect is about twice our estimated effect.[45] Therefore, even if some of the selection on unobservables is not in the magnitude or direction that is assumed, the estimated endogenous effects should still remain positive and meaningful.[46]

To support this conclusion, we implement an alternative approach following Bramoullé et al. (2009). They show that endogenous and contextual effects can be separated if some of the team players could be thought of as peers of player i while others are not in the sense that they cannot directly affect his behavior. While it is hard to claim that some players in the team are not related to player i at all, it is quite plausible to assume that defense players have little direct interaction with strikers (offense players). On the other hand, midfielders interact with both defense and offense players, and practically serve as mediators.[47] Thus, in the following estimation we assume that strikers' and defenders' peers are only midfielders, and that defenders and strikers are not each other's peers. Limiting our sample to strikers and defenders, we are then able to estimate a model where peer effort is instrumented by the seasonal effort of the *peers of player i's peers*. Namely, if player i is a striker, the endogenous explanatory variable is the average distance that the midfielders run in the same section, and it is instrumented by the season average distance that the defenders who are currently on the field run in the same section. In this case, the only way that the average seasonal distance of defenders may affect the contemporaneous running distance of the strikers is through its affect on midfielders' contemporaneous distance. Under the assumption that defenders are not direct peers of the strikers, even if they possess unobserved traits that affect their peer's distance, they will only impact the midfielders and not the strikers. Therefore, the effect that we estimate is clean of any potential contextual effects.

Table 5 presents the results of this alternative estimation. A comparison of the point estimates in Table 5 to the main estimates in Table 4, indicates that they are generally larger (except for the last column) which suggests that the latter might underestimate the pure endogenous effects. This could be due to the fact that defenders and strikers do not directly affect each other's effort (as we

[45] Oster (2019) argues that this procedure is likely to overstate the size of the bias because it relies on a relatively extreme assertion that including unobservables would have produced an R-squared of one. However, the alternative approach that she suggests was not formally developed for a 2SLS estimation.

[46] In addition to the estimation of potential bias, we also evaluate the validity of our IV by splitting it into several instruments, each calculating the average peer effort on a different subset of games in the season, and running an over-identification test. We split games according to the stage in the season (playoffs or not) interacted with whether the game is a home game or not, and thus create 4 sub-groups of games for each team. We then run the same specifications but with these four instruments instead of just one. P-values for the over-identification test range between 0.21 and 0.44 for our different specifications, implying that we cannot reject the joint null hypothesis that the instruments are valid.

[47] For a similar approach using proximity to define peers see Brune et al. (2022). The heterogeneity analysis by position in table 7 somewhat supports our structural assumptions by showing that midfielders react to average teammates' effort more than strikers and defenders. The explanation for this could be that they respond to changes in *any* of their teammates' efforts while strikers and defenders only respond to a subset of their teammates.

Table 5: Teammates' and Player's Efforts - "Peers-of-Peers" IV approach

	Dependent Variable: Player i's running distance					
						No Red cards
	(1)	(2)	(3)	(4)	(5)	(6)
Peer distance	0.738***	0.786***	0.807***	0.809***	0.828***	0.790***
	(0.0502)	(0.0454)	(0.0448)	(0.0446)	(0.0502)	(0.0446)
F-statistic on excluded instrument	123.51	129	126.93	124.13	119.02	113.49
Observations	50,034	50,034	47,206	47,206	41,273	35,869
Controls		√	√	√	√	√
Pre-section Controls			√	√	√	√
Opponent Controls				√	√	√
Crowd & BO Controls					√	

Notes: The table presents 2SLS estimates. The unit of observation is player × section. All regressions include player, section and game fixed effects. The main explanatory variable is the aggregate distance that all midfielders and players in the same position as player i (excluding himself) run during the section. Columns 2—6 include controls for the peers' average characteristics: age (quadratic), height, tenure, and their aggressiveness (proxied by their average number of red and yellow cards in the season), and also for player i's team being the home team, the rank gap between the team and the opponent before game g, and a quadratic trend in the number of sections that player i played up to section s. In columns 3—6, we also add controls for the score gap, goals, red cards and injuries in section $s-1$, and for the team's average speed in this section. In columns 4—6, we control for the opponent team's distance in section $s-1$ of the current game in addition to the opponent's average distance in section s throughout the season. Column 5 includes controls for crowd size (which are missing for 12% of the games) and BO (which are missing for one game). In column 6, the sample is restricted to games without red cards for player i's team. Goalkeepers and midfielders are excluded from the sample. Standard errors are clustered by player and game.

assume here), and hence including them as each others' peers (as in Table 4) biases the estimates downwards. A second reason for this underestimation may be that the exogenous peer effects that we were not able to clean from our main estimates in Table 4 are negative. This second explanation perfectly aligns with our previous analysis of bounds for the endogenous effects.

Overall, our findings imply that when teams are engaged in a collaborative task, an important determinant of individual effort is that which is exerted by his peers.

4.2.2 Can just one individual player impact his teammates' effort?

As endogenous peer effects embed a social multiplier, if they exist in our setting, we should expect to see a substantial effect on team effort even when just one player changes his effort. To estimate such a change, we focus on within-game substitutions, which usually replace a fatigued player with a fresh one, thus creating a positive shock to the effort level of this specific individual unit of the team. Because each team is allowed to initiate a maximum of three substitutions in each game, the coach will usually replace only a single player in a given section, thus leaving the composition of peers constant. We exploit this structure to test whether the pre- and post-substitution effort of peers positively relates to the difference in effort between the incoming and outgoing players. A substantial advantage of this approach is that the "treatment" in each substitution is provided by the players who participate in the substitution while the effect is measured on a separate set of players, the ones that remain on the field. This structure avoids the usual symmetry in peer effect models where each individual simultaneously receives and provides treatment.

We proceed by presenting descriptive evidence to support the notion that fatigue is a substantial determinant of players' efforts and thus substitutions provide an opportunity to rely on a quasi-exogenous individual increase in effort to study its impact on team effort. After this preliminary analysis, we describe the identification strategy that additionally builds on an instrumental variable approach. As with our first identification strategy, we will not be able to completely rule out that our estimates capture some contextual effects in addition to the endogenous effect since the incoming and outgoing players may differ on other dimensions except fatigue. We control for such confounding factors to the extent that they are observed or captured by our fixed effects.

Initial evidence for the importance of fatigue are presented in Figure 3 which shows that the average effort of individual players generally decreases as the game progresses.[48] One exception is in the sections immediately after the half-time break in which players rest. In addition, there is a very subtle increase in the last two sections of the game which can be attributed to the fact that by then at least some of the tired players have been replaced. Indeed, this tiny increase disappears when we restrict the sample to players who remain on the field for the entire game. Since there is no reason to expect this negative trend to result from some strategic choice or any other consideration, it indicates that fatigue is a central determinant of effort.

Figure 4 provides additional support to this claim by clearly demonstrating that very few substitutions are initiated during the first half of the game when players are relatively fresh. During the second half, the spread of substitutions over the sections is quite uniform, especially between sections 14 and 18, where most substitutions take place.[49] These are also the sections in which players' average effort levels are the lowest throughout the game, which again supports fatigue as a reason for substitutions.[50]

Therefore, it is expected that the incoming player will run faster than the tired outgoing player. Figure 5 confirms this by presenting the distribution of the running distance gap between these two players (excluding goalkeepers) both for all the substitutions and for the substitutions that were not injury driven.[51] A positive gap means that the incoming player runs more, or exerts more effort, than the outgoing player. To calculate this gap, we subtract the outgoing player's distance in the section that preceded the substitution from the distance of the incoming player during his first full section immediately after the substitution. We do not use the actual substitution section both

[48]We exclude goalkeepers from the sample because their running distance is substantially lower than all other players. This restriction does not change the portrayed trend over sections and only affects the average effort level.

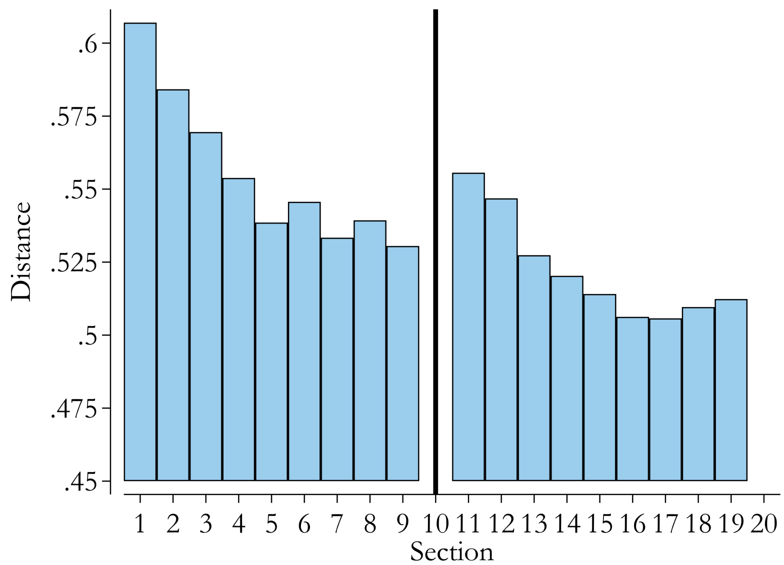
[49]Section 11 shows a spike in substitutions because coaches use the halftime break to change team composition; consequently there are relatively few substitutions in the 12th section.

[50]The dashed line in Figure 4 shows that the number of substitutions following an injury is only 85 out of a total of 1335 and that they are spread quite evenly across all game sections, except for an arbitrary spike in section eight.

[51]Goalkeepers are rarely substituted and thus only three substitutions are omitted.

because the incoming and outgoing players are only playing a partial section and because in some cases the substitution by itself could slightly change the measurements as it causes the game to pause.[52] The figure shows that the distribution has a positive mean and that it is skewed towards positive gaps. Specifically, while more than 50% of these gaps are positive (between 51 and 57 percent), only around 20% are negative (between 21 and 24 percent).[53] This clear pattern supports our premise that fatigue substantially modifies players' effort, and hence implies that substitutions create a quasi-exogenous shift in individual effort.

Figure 3: Players' average running distance by section



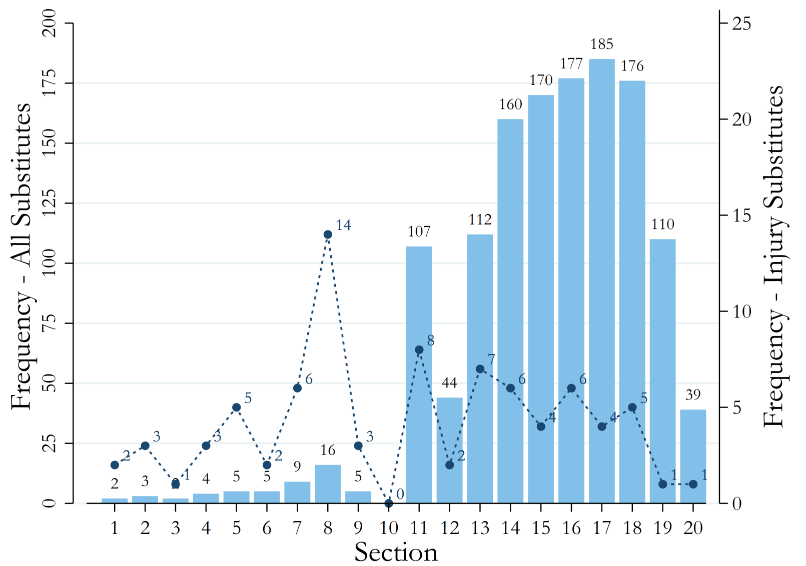
Notes: The figure presents the average running distance in each section for players that played full sections, excluding goal keepers. Sections 10 and 20 are the overtime sections that vary in duration and are thus omitted.

We build on this shift and test whether the running distance of the team members (excluding the goalkeepers) changes in accordance with the distance gap between the incoming and outgoing players. But other factors could also be involved such as within-game events, the strategic plan of the coach, the difference in characteristics between the incoming and outgoing players, game-level characteristics such as crowd size or the opponent's competitiveness, and the stakes of the game. Therefore, in our baseline specification below, we control for all observed characteristics and pre-substitution events, and also use game and section fixed effects to account for unobserved

[52]These two limitations are exacerbated by the fact that our data on distance is rounded in each section to 0.1 units.

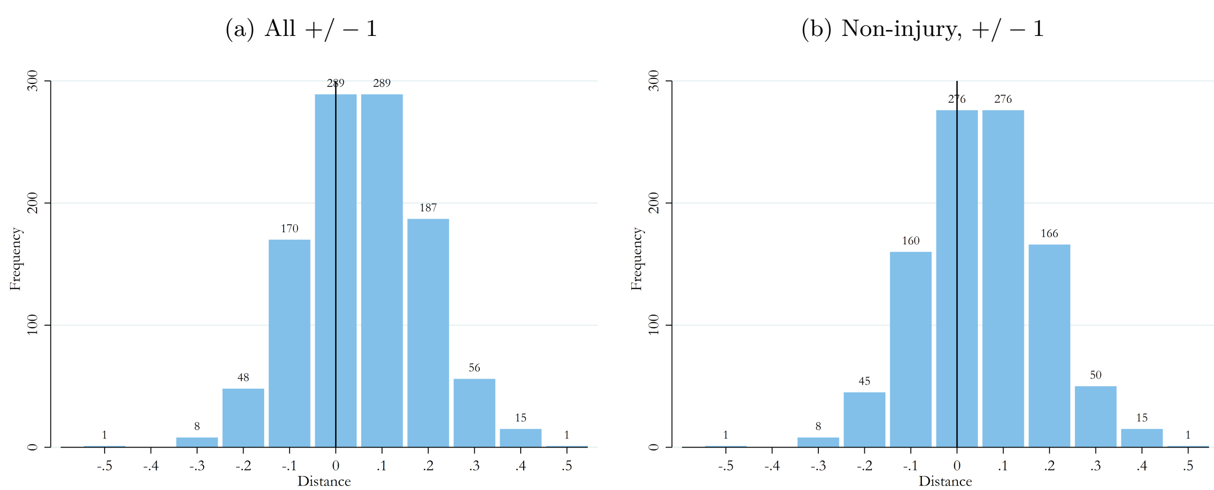
[53]Because the distances are rounded to 0.1 units, a calculated gap of zero indicates that the difference between the players was either positive or negative but closer to zero relative to values of 0.1 or -0.1.

Figure 4: Number of substitutions by section



Notes: The figure presents the number of substitutions that were initiated in each section throughout the entire season. Goalkeeper substitutions are excluded from the sample. The bars correspond to the left-hand side Y-axis and present counts of all types of substitutions. The dashed line shows the number of injury-induced substitutions and corresponds to the left-hand side axis.

Figure 5: The distribution of the distance gap between players in substitutions



Notes: The figure presents the distribution of the gap in running distance between the incoming and outgoing players of each substitution, excluding goalkeeper substitutions. In panels (a) and (b), the gap is calculated using the distances in one section pre- and post the section of the substitution. Panel (b) reports the gaps only for substitutions that were not induced by injury.

factors:

$$\Delta dist_{i,k,s,g} = \tau_s + \phi_g + \beta_2 \cdot \Delta dist\_gap_{k,s,g} + W'_{i,s,g}\kappa + \psi_{i,k,s,g} \tag{13}$$

where the main explanatory variable is the distance gap between the incoming and outgoing players in substitution k which occurred in section s of game g. The dependent variable is the gap between the pre- and post-substitution k distances for player i who plays for the team that initiated substitution k but did not participate in it. Thus, β_2 captures the relationship between a change of one team member's effort due to the substitution and the consequent simultaneous change in player i's effort. Because our outcome measure is the *difference* in effort, we are not concerned that it reflects the ability or the average effort level of the player.[54] As in all our previous regressions, we also include section number and game fixed effects and a rich set of controls for game, player, opponent, and previous section characteristics. We also control for the difference in observed characteristics (age, tenure, height, and aggressiveness) between the players that participate in the substitution in order to clean the exogenous peer effects that these changes may generate.

To further address concerns that the main explanatory variable is correlated with the type of substitution, its timing, or its aim, we estimate the same specification on several different subsamples that focus on specific substitution scenarios such as injury-induced substitutions. However, this analysis cannot completely rule out that unobserved pre- or post-substitution events similarly affect the running distance of both the incoming player and the other players, namely, that correlated effects drive our results.

To cope with this shortcoming, we use an IV approach which is similar to the one that was presented in section 4.2.1, namely instrumenting a player's contemporaneous effort with his seasonal effort. This time, however, we introduce a novel adjustment to this IV that directly captures the key element in our substitution based analysis, namely, players' level of fatigue. To accomplish this, the average effort in the season is calculated based only on sections in which the player's cumulative number of sections on the field is as in the specific substitution. In particular, for the outgoing player, we look at the cumulative number of sections played before the substitution section and calculate his season average based on sections with the same cumulative number throughout the season.[55] For the incoming player, the first section post-substitution is always his first full section in the game and thus we calculate his average seasonal running distance in all instances in which he played a first full section in the game. We note that even for the outgoing player the cumulative number of sections played for a specific player is not necessarily equal to the serial number of the section in the game because it could be the case that this player was the incoming player in a previous substitution. The IV is the difference between these two averages and we interpret

[54] For the same reason, we do not include player level fixed effects in this estimation.

[55] For example, if the substitution took place during the twelfth section and the outgoing player played from the beginning of the game, we calculate his average running distance in all instances in which he played a full section that was also his eleventh section in the game.

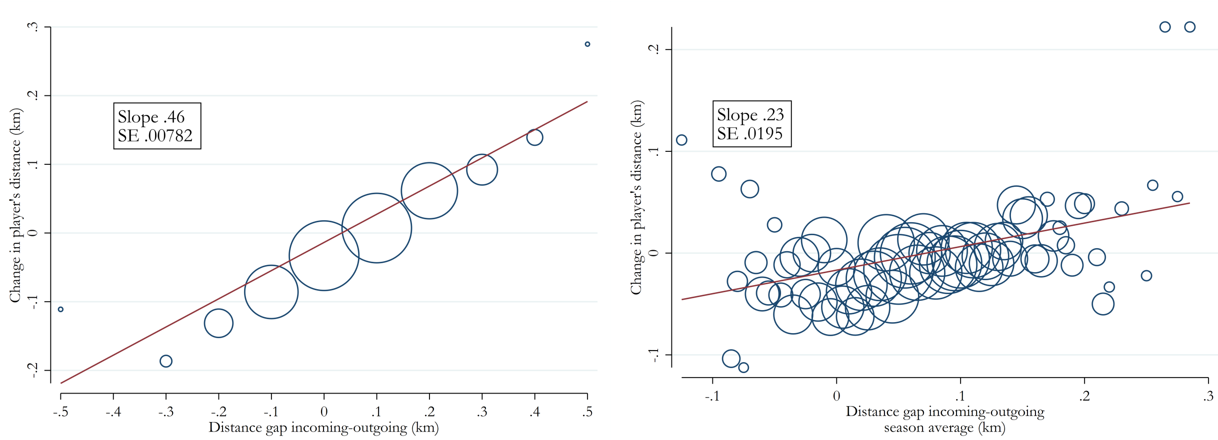
it as the gap that would be expected between the incoming and outgoing players if their effort was determined only based on their natural fatigue, regardless of the specific events that surround the specific substitution. Therefore, this difference is not expected to directly affect changes in the contemporaneous effort of their teammates pre- or post-substitution and hence the exclusion restriction holds.

It is possible, however, that differences in some unobserved characteristics of the incoming and outgoing players, which are not controlled for in our regressions, are correlated with the difference in their running distance and also affect the change in the running distance of player i. Namely, while our proposed strategy clears any correlated effects, it cannot address all types of exogenous peer effects, and our estimated peer effects will thus present a combination of the endogenous and part of the exogenous effects. As with our first IV strategy, this problem is mitigated by the fact that our estimates remain stable across specifications with and without controls.

Before proceeding to the results of our estimation, in Figure 6 we plot the raw data for the OLS specification and for the reduced form relationship, namely the correlation between the outcome variable and the instrument. Both plots show positive slopes.[56]

Figure 6: Changes in Teammates' and Player's Efforts - Substitutions

(a) OLS   (b) Reduced form



Notes: Panel (a) plots the raw data with a linear fitted line for the relationship between the distance gap between the incoming and outgoing players in a substitution and the corresponding change in other players' running distance. In panel (b), the x-axis variable is our suggested instrument — the seasonal average running distance between the incoming and the outgoing players, which is calculated based only on sections in which the player's cumulative number of sections on the field is as in the specific substitution. The sample includes 1,064 substitutions.

Table 6 presents the estimated impact of the distance gap in the substitution on the change in the other players' effort using both the OLS and IV strategies and for four different samples. We use

[56]The OLS figure (panel (a)) has relatively few dots because the values of each player's running distance are rounded to 0.1 units (the variable on the X-axis is a difference between two players), as opposed to when we look at averages over games or over players (as in panel (b)).

different sample restrictions that refer to various within-game events and intermediate outcomes because it has been documented that such events could impact players' behavior (Deutscher and Schneemann, 2017; Schneemann and Deutscher, 2017). We show specifications with and without controls for both strategies, and for the 2SLS estimation we also show the results with the two versions of opponent effort controls that were used in columns (5) and (6) of Table 4. All of the twenty four resulting point estimates are positive and very precisely estimated. The estimated impacts are also of a similar magnitude across samples. Generally, the IV estimates are approximately 40-50% smaller than the OLS estimates. These effects should be interpreted as the effect of a player's effort on each of his peer's effort for cases in which the players who participate in the substitution comply, namely that the actual distance gap is positively correlated with the effort patterns that these players displayed throughout the season.

The first row of the table presents the results of the estimation when we use the largest possible sample, which includes 1,064 substitutions. The magnitude of the estimated effects is not very sensitive to the inclusion of the various controls, except for a substantial decrease when we control for the contemporaneous change in the opponent's effort in column (5). As explained in section 4.2.1, since changes in players' effort are expected to affect the opponent players' efforts, controlling for their efforts in this way may bias the estimates downwards. Thus, in column (6), we control for the opponent's effort in the section before the substitution and for the difference in the seasonal average effort of the opponent between the same sections in the game. In this specification the estimated effects remain stable.

In the second row, we show that the results hold for a sample that excludes substitutions preceded or followed by a score change or a red card for either team.[57] This helps alleviate concerns that substitutions are related to major events in the game that could change the strategic choices of effort for the entire team. For example, one might think that after the team scores a goal, the coach might substitute a striker with a defense player and also instruct the entire team to slow down and play more defensively. This would cause a negative distance gap in the substitution and also a negative gap for all of the players in the team, which would result in a spurious positive coefficient that cannot be causally interpreted. Alternatively, substitutions which were followed by dramatic events may increase correlated peer effects and bias the estimates of the endogenous peer effects. Therefore, although conditioning the sample on post-substitution events is generally not advisable, it can somewhat assure that the effect is not driven by these events. In addition, to ascertain that our results are not specifically related to injury-induced substitutions, in the third row we limit the sample to non-injury substitutions, which yields identical findings. Lastly, to rule out the possibility that general changes in the game strategy lead both to the substitution and to the change in teammates' speed, we limit the sample to substitutions that do not affect the team's formation. The results indicate that the point estimates are similar and even slightly larger.

[57] We look at three sections before and after each substitution.

Table 6: Changes in Teammates' and Player's Efforts - Substitutions

	Dependent Variable: Difference in Player's Running Distance					
	OLS		2SLS			
	(1)	(2)	(3)	(4)	(5)	(6)
All						
Distance gap (incoming-outgoing)	0.471***	0.481***	0.235***	0.250***	0.110***	0.249***
	(0.0190)	(0.0221)	(0.0461)	(0.0486)	(0.0304)	(0.0487)
F-statistic on Excluded Instrument			217.3	197.8	217.9	201.7
Observations	*9,055*	*7,951*	*9,055*	*7,951*	*7,951*	*7,951*
No red cards or goals						
Distance gap (incoming-outgoing)	0.449***	0.463***	0.248***	0.254***	0.113**	0.257***
	(0.0305)	(0.0329)	(0.0778)	(0.0799)	(0.0523)	(0.0770)
F-statistic on Excluded Instrument			87.1	61.9	95.0	65.7
Observations	*3,015*	*2,688*	*3,015*	*2,688*	*2,688*	*2,688*
Non-injury substitutions						
Distance gap (incoming-outgoing)	0.472***	0.476***	0.234***	0.249***	0.110***	0.246***
	(0.0205)	(0.0235)	(0.0472)	(0.0497)	(0.0307)	(0.0496)
F-statistic on Excluded Instrument			213.9	188.3	207.8	191.7
Observations	*8,473*	*7,449*	*8,473*	*7,449*	*7,449*	*7,449*
No formation-change						
Distance gap (incoming-outgoing)	0.484***	0.490***	0.336***	0.320***	0.210***	0.314***
	(0.0323)	(0.0369)	(0.0755)	(0.0851)	(0.0456)	(0.0854)
F-statistic on Excluded Instrument			54.2	44.6	58.7	43.9
Observations	*4,703*	*4,150*	*4,703*	*4,150*	*4,150*	*4,150*
Controls		√		√	√	√
Opponent Controls					√	√
Section FE	√	√	√	√	√	√
Game FE	√	√	√	√	√	√

Notes: The table presents estimates at the player level for the sections just before and after a substitution. The main explanatory variable is the difference in the running distance between the incoming and the outgoing player. The outcome variable is this difference for player i who was on the field both before and after the substitution. Columns 1—2 present the results for an OLS estimation; columns 3—6 present the results for 2SLS estimations where the endogenous explanatory variable is instrumented by the difference in the season average running distance between the incoming and outgoing players when they play the same number of sections in a specific game. All regressions include section and game fixed effects while columns 2 and 4—6 also include controls for player i's team being the home team, the rank gap between player i's team and the opponent team prior to game g, the score gap, goals, red cards and injuries in the section that precedes the substitution, the team's average speed in this section, a quadratic trend in the number of sections that the player played up to section s, and the differences in characteristics (age, tenure, height, and aggressiveness) between the incoming and outgoing players. In column 5, we control for the difference in the opponent players' average distance in the sections before and after the substitution. In column 6, we control for the opponent's distance in the section that preceded the substitution in addition to the opponent's average difference in distance between sections $s + 1$ and $s - 1$ throughout the season. The sample in each panel is different and the number of observations in each of these samples is shown at the bottom of each panel. Goalkeeper substitutions and goalkeepers are excluded from the sample. Standard errors are clustered by player and game.

Taken together, these results demonstrate that even when just one player increases his effort, the efforts of his peers and hence the effort of the entire team may substantially increase. Moreover, in Table 6, the estimated effect of a change in *one* peer player's running distance is larger than would be expected based on the estimated effect in Table 4, of a change in the average of *nine* peers since it is approximately one third of the latter (rather than one ninth). This seeming discrepancy can be reconciled by two main explanations.[58] First, when we focus on just the specific sections immediately before and after the substitution, the substitution strategy captures the immediate impact of a change in team composition, whereas in the first IV strategy we exploit all sections in each game and measure average differences in effort across sections with different peer compositions. Since there is a maximum of only three substitutions per game (for each team), the same composition of players could persist for many sections before or after a substitution. This implies that the estimates in Table 4 capture the average effect of a change in team composition over all of these sections. This average effect could be much smaller than the immediate effect of a substitution if the response to a change in peer effort diminishes over time. Second, the substitution strategy exploits a subset of changes in peer effort which are used in our first IV strategy, and it is possible that these substitution-induced changes have particularly large impacts. One prominent reason could be that substitutions carry stronger behavioral effects because the changes that they create in peer effort are more salient.

Lastly, we discuss how a change in an individual player's effort is expected to affect *game outcomes* through its effect on his peers. For this purpose, we conduct a back-of-the-envelope calculation in which we apply our estimates of how team effort is affected by an individual player's effort and how, consequently, the overall change in team effort is related to the team's winning probability. This calculation is merely indicative of the potential magnitude of this effect. First, to evaluate the increase in team effort due to an increase in the effort of an individual player, we rely on the estimates reported in Column (6) of Table 6. These estimates suggest that an increase of one kilometer in an individual player's running distance causes an average increase of between 0.246 and 0.314 kilometers in the running distance of each of his nine teammates and thus an increase of between 3.2 and 3.8 kilometers in total team distance (the estimated change in each peer's distance times nine peers plus the initial one kilometer increase in the 10th player's distance). This is also the range for the multiplier effect namely, the overall increase in team effort divided by the initial change in one player's effort.[59] Then, we assume an increase of 10% in one player's speed, which is on average 6.29 kmph, namely the player runs an additional 0.629 kilometers per hour or 0.052

[58]A further explanation could be that assuming linearity in the aggregation of peer effort may be less plausible for large changes.

[59]Since players choose their level of effort simultaneously, our estimated effect captures the total equilibrium change in each player's effort except for the incoming player whose change in running distance is instrumented by his typical running distance. This calculation slightly underestimates the multiplier since it does not take into account the return effect on the incoming player who initially changed the running distance by one kilometer, but can also respond to the consequent changes in his teammates' running distance.

kilometers per section. As, on average, a player plays 16.05 sections per game, this implies an increase of 0.84 kilometers per game. Based on our estimated multiplier, this initial increase will lead to a total increase of between 2.7 and 3.2 kilometers in the aggregate running distance of the team during the game. According to column 2 of Table 3, this change will increase the probability of winning the game by 11.5 to 13.6 percentage points, which amounts to a substantial increase of around 30% in the average probability to win (approximately 40%).

4.3 Heterogeneous effects

After establishing that there is a positive interaction between teammates' efforts, it is interesting to check whether their intensity varies by player, team, or game characteristics. Such an analysis will also facilitate a more informed discussion of the two main alternative mechanisms that potentially underlie the endogenous effects in our context: complementarity in production (supplemented by the contest structure) and behavioral considerations such as social pressure, prosocial behavior or shame.

At the player level, we compare the main effects that we find across players in different positions, and with different levels of experience (age), tenure or both. We divide players into two groups of above and below the 75th percentile of experience. As tenure in the current team is counted by season, we consider players with a tenure of at least two seasons (including the current one) as senior. At the team level, we separate teams by the median tenure of their players being above or below this two-season threshold. In addition, we use BO data to define for each team in each game which outcome is most probable ex-ante — win, lose, or tie.

Table 7 presents the results of these estimations based on the IV specification which measures the impact of the team on the individual. We do not present estimates for the substitution based approach because the much smaller sample size limits our ability to efficiently estimate heterogeneous effects.[60] Column (1) shows that defenders and strikers respond less than midfielders to changes in peer effort which is in line with the latter's role as mediators[61] and therefore suggests that production complementarities play an important role. The difference in the estimated effect is substantial and significant for defenders while the difference for strikers is smaller and not statistically significant at conventional levels. It is likely that the estimates are noisier for strikers since there are fewer players in this position than in other positions. However, the p-value is still quite low at 0.14.

In columns (2)—(4), we test whether players who are more experienced (above the 75th percentile of age) or more senior (play in the team for at least 2 seasons) respond differently to peer effort. The findings indicate that these two factors reduce the estimated effect of peer distance,

[60]Nonetheless, heterogeneity patterns are consistent across the two estimation strategies except for different significance levels.

[61]This idea is also discussed in section 4.2.1

Table 7: Heterogeneous Effects

	Dependent Variable: Player's Running Distance					
	(1)	(2)	(3)	(4)	(5)	(6)
Peer distance	0.850***	0.814***	0.804***	0.831***	0.783***	0.794***
	(0.0238)	(0.0178)	(0.0157)	(0.0182)	(0.0158)	(0.0158)
Peer dist. × Defender	-0.125***					
	(0.0289)					
Peer dist. × Striker	-0.0533					
	(0.0362)					
Peer dist. × (Age≥ p75)		-0.0875***		-0.0891***		
		(0.0293)		(0.0301)		
Peer dist. × (Tenure≥ 2)			-0.0454**	-0.0460**		
			(0.0183)	(0.0184)		
Peer dist. × (Team tenure≥ 2)					0.0159	
					(0.0315)	
Peer dist. × (Prob. Win≥50%)						-0.0191
						(0.0163)
Peer dist. × (Prob. Loss≥50%)						-0.0187
						(0.0164)
Observations	75,886	74,882	74,291	74,291	75,886	75,552

Notes: The table presents 2SLS estimates. The unit of observation is player × section. The main explanatory variable is the aggregate distance that all players except player i run during the section (excluding goalkeepers) instrumented by these players' average running distance in section s in all games throughout the season. In each column, the specification includes interaction terms of the main explanatory variable and indicators for: age above 75th percentile, tenure of two seasons or more, median tenure in the team of 2 seasons or more, probability of winning or losing above 50% according to BO data, or players' positions. All regressions include player, game and section fixed effects. In addition, we control for peers' average characteristics: age (quadratic), height, tenure, and their aggressiveness (proxied by their average number of red and yellow cards in the season), for player i's team being the home team, for the rank gap between the team and the opponent before game g, for a quadratic trend in the number of sections that player i played up to section s, for the score gap, goals, red cards, injuries and both teams' average distance in section $s-1$, and for the opponent's average distance in section s throughout the season. Goalkeepers are excluded from the sample. Standard errors are clustered by player and game.

even when both are included in the regression in column (4). These findings point towards social pressure as a probable mechanism for creating peer effects since players who are more confident in their professional status or place in the team, probably respond less to social considerations. Nonetheless, even the most senior players have a substantial response to peer effort, which suggests that complementarity is still an important force.

The team and game level factors, on the other hand, do not seem to influence the estimated effects. Although it is plausible to assume that the experience that team members gain playing together for the same team will influence the degree of their complementary, teams with a median player tenure of at least two seasons do not show stronger effort interactions among players. This may be explained by an offsetting effect of decreased social pressure among teammates in wellestablished teams compared to newly-formed ones.[62] Lastly, the expected game outcomes also do not seem to affect the extent of peer influence.

5 Conclusion

This paper presents a first attempt to explore peer effects in the workplace using a direct measure of effort rather than proxying it by performance measures. Using a unique high-frequency dataset on the effort and performance of each player, we estimate how a player's choice of effort is dependent on the efforts of his coworkers who collaborate with him on the same task. We employ two different identification strategies to detect the effect of changes in peer effort on a player's effort. The first strategy exploits multi-level fixed effects in addition to instrumenting the effort of one's coworkers with their average effort in the same stage of the game during the entire season. Thus, for a given stage of the task, variation stems entirely from the composition of the coworkers. The second strategy focuses on the sections just before or after substitutions and analyzes how a gap in the running distance between the incoming and outgoing players impacts the change in the other players' effort. This gap is instrumented by the players' corresponding average seasonal effort when playing the same cumulative time in the game.

The results indicate that not only do group efforts substantially impact individual efforts but also that a change in only one player's effort can substantially impact the effort of the entire team. These results are robust to many different specifications and samples. In addition, we showed that players who are more experienced and especially more senior in the team tend to respond significantly weaker when their teammates are working more intensively. This suggests that behavioral considerations play a role in the generation of peer effects in teams.

Although caution must be exercised in generalizing the findings of this study to other labor market settings, it is plausible to expect similar peer effects to operate when teams perform a collaborative task and its success is determined through competition with opponent teams. A potential implication of the strong peer effects that we found is that managers should determine workers' compensation not only by their direct contribution to output but also according to their effort. This may be even more efficient in environments where effort is strongly related to group performance and when individual performance is rarely observed, difficult to quantify, or when common individual performance measures are irrelevant for large parts of the team. Finally, our results suggest that teammate social connections and obligations to each other may contribute to positive peer effects in effort, and indicates why it can be beneficial for organizations to encourage social interaction among coworkers and to invest in activities that help forge team spirit.

[62] Relatedly, Brune et al. (2022) find that positive motivational effects are only present when peers working nearby are not the worker's friends.

References

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge (2017). When should you adjust standard errors for clustering? *Working Paper*.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics 97*(4), 543–569.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics 105*(2), 255–283.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005a). An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. *Journal of Human Resources 40*(4), 791–821.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005b). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy 113*(1), 151–184.

Amodio, F. and M. A. Martinez-Carrasco (2018). Input allocation, workforce management and productivity spillovers: Evidence from personnel data. *The Review of Economic Studies 85*(4), 1937–1970.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics 30*, 98–108.

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

Azoulay, P., J. S. Graff Zivin, and J. Wang (2010). Superstar extinction. *The Quarterly Journal of Economics 125*(2), 549–589.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica 77*(4), 1047–1094.

Blume, L. E., W. A. Brock, S. N. Durlauf, and R. Jayaraman (2015). Linear social interactions models. *Journal of Political Economy 123*(2), 444–496.

Boucher, V., Y. Bramoullé, H. Djebbari, and B. Fortin (2014). Do peers affect student achievement? evidence from canada using group size variation. *Journal of Applied Econometrics 29*(1), 91–109.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*(1), 41–55.

Bramoullé, Y., H. Djebbari, and B. Fortin (2020). Peer effects in networks: A survey. *Annual Review of Economics 12*, 603–629.

Brave, S. A., R. A. Butters, and K. A. Roberts (2019). Uncovering the sources of team synergy: Player complementarities in the production of wins. *Journal of Sports Analytics 5*(4), 247–279.

Brune, L., E. Chyn, and J. Kerwin (2022). Peers and motivation at work evidence from a firm experiment in Malawi. *Journal of Human Resources 57*(4), 1147–1177.

Buraimo, B. and R. Simmons (2009). A tale of two audiences: Spectators, television viewers and outcome uncertainty in Spanish football. *Journal of Economics and Business 61*(4), 326–338.

Cornelissen, T., C. Dustmann, and U. Schönberg (2017). Peer effects in the workplace. *American Economic Review 107*(2), 425–56.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–96.

Deutscher, C. and S. Schneemann (2017). The impact of intermediate information on sabotage in tournaments with heterogeneous contestants. *Managerial and Decision Economics 38*(2), 222–237.

Fafchamps, M. and B. Caeyers (2020). Exclusion bias and the estimation of peer effects. *Working Paper*.

Fehr, E. and L. Goette (2007). Do workers work more if wages are high? evidence from a randomized field experiment. *American Economic Review 97*(1), 298–317.

Gardner, J. (2021). Two-stage differences in differences. *Working paper*.

Garicano, L. and T. N. Hubbard (2012). Learning about the nature of production from equilibrium assignment patterns. *Journal of Economic Behavior & Organization 84*(1), 136–153.

Gould, E. D. and E. Winter (2009). Interactions between workers and the technology of production: Evidence from professional baseball. *The Review of Economics and Statistics 91*(1), 188–200.

Grossman, G. M. and G. Maggi (2000). Diversity and trade. *American Economic Review 90*(5), 1255–1275.

Guryan, J., K. Kroft, and M. J. Notowidigdo (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics 1*(4), 34–68.

Herbst, D. and A. Mas (2015). Peer effects on worker output in the laboratory generalize to the field. *Science 350*(6260), 545–549.

Imai, K. and I. S. Kim (2021). On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis 29*(3), 405–415.

Jackson, C. K. and E. Bruegmann (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics 1*(4), 85–108.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of Economic Perspectives 14*(3), 75–94.

Lechner, C. and S. V. Gudmundsson (2012). Superior value creation in sports teams: Resources and managerial experience. *M@n@gement 15*(3), 284–312.

Lee, D. L., J. McCrary, M. J. Moreira, and J. Porter (2020). Valid t-ratio inference for IV. *Working Paper*.

Lee, L.-F. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics 140*(2), 333–374.

Levine, B. (2015). Measuring team chemistry with social science theory. `https://community.fangraphs.com/measuring-team-chemistry-with-social-science-theory/`.

Malueg, D. A. and A. J. Yates (2010). Testing contest theory: Evidence from best-of-three tennis matches. *The Review of Economics and Statistics 92*(3), 689–692.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542.

Mas, A. and E. Moretti (2009). Peers at work. *American Economic Review 99*(1), 112–45.

Neugart, M. and M. G. Richiardi (2013). Sequential teamwork in competitive environments: Theory and evidence from swimming data. *European Economic Review 63*, 186–205.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Peeters, T. and S. Szymanski (2014). Financial fair play in european football. *Economic Policy 29*(78), 343–390.

Raff, D. M. and L. H. Summers (1987). Did Henry Ford pay efficiency wages? *Journal of Labor Economics 5*(4, Part 2), S57–S86.

Sandvik, J., R. Saouma, N. Seegert, and C. Stanton (2020). Employee responses to compensation changes: Evidence from a sales firm. *Working Paper*.

Schneemann, S. and C. Deutscher (2017). Intermediate information, loss aversion, and effort: Empirical evidence. *Economic Inquiry 55*(4), 1759–1770.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review 74*(3), 433–444.

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

Wicker, P., J. Prinz, D. Weimar, C. Deutscher, and T. Upmann (2013). No pain, no gain? Effort and productivity in professional soccer. *International Journal of Sport Finance 8*(2), 124–139.

6 Appendix

6.1 Basic Rules of Football

A football match is played between two teams, each with 11 players on the field. It lasts 90 minutes, divided into into two halves of 45 minutes each. At the end of each half, teams continue to play for a few more minutes, known as extra or stoppage time, as determined by the referee to compensate for halts in the game due to goals, substitutions, fouls, or injuries. The aim of each team is to score goals, namely to kick (or gore) the ball into the net between the goalposts, while preventing the opponent team from scoring goals. There are two goals, one on each side of the field, and each team is assigned to a specific goal. One player in each team is positioned as the team's goalkeeper. The team that scores more goals wins the game. If both score the same number of goals the game outcome is a draw.

The team's coach determines the identity of the 11 players that take the field at the beginning of the game, and can then initiate up to three substitutions during the game. Players are not allowed to intentionally or recklessly use physical contact to hinder the opponent players' efforts, or to touch the ball with their hands. If they do so, the referees call a foul and penalize players according to the severity of the foul. The lowest degree of penalty is a free kick for the opponent, the medium degree is a yellow card, and the highest degree is a red card. Two yellow cards for the same player within the same game automatically lead to a red card, resulting in the immediate removal of this player from the game, without the ability of his team to replace him. Thus, his team must continue the rest of the game with one less player, which is a significant disadvantage. In addition, a card would grant the opponent with a free kick or a penalty kick.

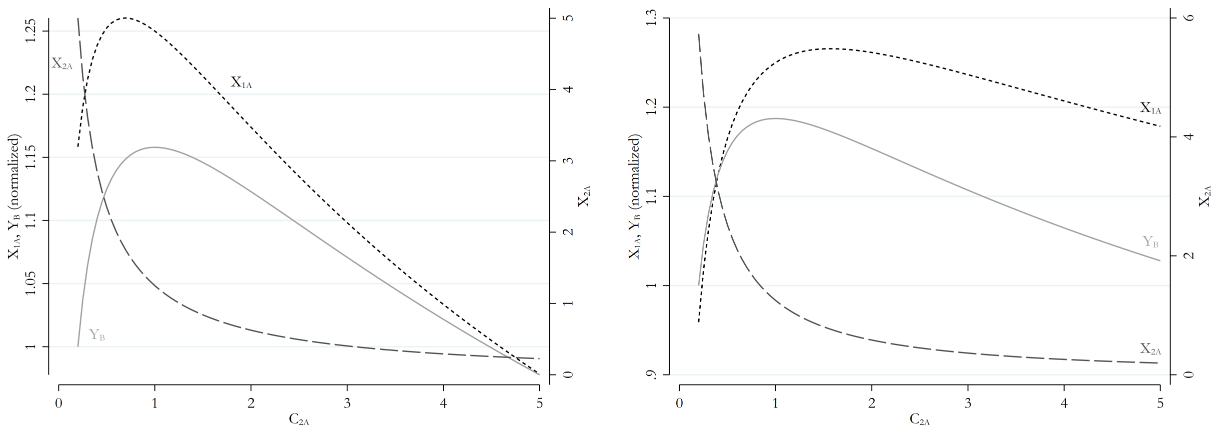
6.2 The Israeli League

The Israeli Professional League is composed of 14 teams, each pair of teams playing against each other between two to four times during the season. In the first 26 rounds, all 14 teams play each other twice, once in each team's home stadium. After that, the league is divided into two houses. In the upper house, the six leading teams proceed to play against each other twice, once in each team's stadium, while the eight remaining teams play each other only once to determine which teams will be demoted to the lower level league in the next season. Therefore, there are six teams that end up playing 36 rounds of games and eight teams that play only 33 rounds. For each game, the team receives points according to the outcome: 3 points for a win, 1 for a draw, and 0 for a loss. The teams' rank in the league is determined by their aggregate score in each round.

6.3 Appendix Tables and Figures

Figure A1: The effect of changing one player's effort cost on players' efforts (simulation)

(a) $\gamma = .1$ (b) $\gamma = -.1$



Notes: The figures present simulation of the efforts of team A players and the output of team B for different values of c_2A and for the following constant parameter values: $c_{1A} = c_{1B} = c_{2B} = 1, \omega = 10$. The value of γ is specified in the title of each figure.

Table A1: Teams' running distance and game outcomes

	Dependent Variable:				
	Team winning probability				Goals
	(1)	(2)	(3)	(4)	(5)
Team Distance (km)	0.0399***			0.0190**	0.0704***
	(0.0106)			(0.00933)	(0.0237)
Opponent Distance (km)	-0.0476***				
	(0.0126)				
Distance Ratio		4.653***			
		(1.143)			
Total Distance (km)		-0.00388			
		(0.00470)			
Distance Diff. (km)			0.0441***		
			(0.0118)		
Observations	228	228	224	224	224
R^2	0.501	0.500	0.703	0.679	0.652
Controls	√	√	√	√	√
Team FE	√	√			
Opponent FE	√	√			
Pair FE			√	√	√
Round FE	√	√	√	√	√

Notes: The table presents estimates at the *team* × *game* level, where team i is selected according to the Hebrew alphabetical order of the teams' names. All regressions include round fixed effects. Columns 1 and 2 in addition include team and opponent fixed effects, and columns 3—5 include pair fixed effects (as in Table 3). In columns 1—4, the outcome variable is an indicator for team i winning the game and in column 5 the outcome variable is the number of goals that team i scored. In column 1, the main explanatory variable is the aggregate distance that team i's players run during the game. We control for the same measure for the opponent team. In column 2, the main explanatory variable is the ratio between the aggregate distance of team i and its opponent and we also control for the sum of these distances. In column 3, we replace the main explanatory variables with the difference between the aggregate distances of the team and the opponent, and in columns 4—5 we only include the aggregate distance of team i without controlling for the distance of the opponent. Standard errors are clustered at the pair level.

Table A2: Teams' running distance and game outcomes – UEFA Champions league

	Dependent Variable: Team i winning probability					
	(1)	(2)	(3)	(4)	(5)	(6)
Team Distance (km)	0.0619***	0.0579***	0.0689***	0.0580***		
	(0.0147)	(0.0152)	(0.0182)	(0.0152)		
Opponent Distance (km)	-0.0640***	-0.0567***	-0.0698***	-0.0567***		
	(0.0147)	(0.0149)	(0.0186)	(0.0150)		
Distance Ratio					6.319***	6.319***
					(1.494)	(1.495)
Total Distance (km)						0.000649
						(0.00609)
Observations	248	248	179	248	248	248
R^2	0.614	0.620	0.590	0.623	0.621	0.621
Controls		√	√	√	√	√
Pair FE	√	√	√		√	√
Year FE	√	√	√		√	√
Pair X Year FE				√		

Notes: Data from UEFA Champions League seasons 2017-2019. Estimates are at the *team* $\times$ *game* level, where team i is selected according to the alphabetical order of the teams' names. All columns include pair and year fixed effects except for column 4 which includes *pair* $\times$ *year* fixed effects instead. In columns 1—4, the main explanatory variable is the aggregate distance that team i's players run during the game. We control for the same measure for the opponent team. In columns 5—6, the main explanatory variable is the ratio between the aggregate distance of team i and its opponent, where in column 6 we also control for the sum of these distances. Columns 2—6 also control for being the home team in game g. In column 3, the sample is restricted to games without red cards. Standard errors are clustered at the pair level.