Effort Peer Effects in Team Production: Evidence from Professional Football

Danny Cohen-Zada* Itay Dayag† Naomi Gershoni‡

December 3, 2020

Abstract

We exploit a unique dataset from the Israeli Professional Football Leagues which provides high-frequency, direct measures of players' effort, 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 dominant 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, complemetarity, distance.

*Ben-Gurion University of the Negev and IZA, Germany; danoran@bgu.ac.il
†Ben-Gurion University of the Negev; itayday@bgu.ac.il
‡Ben-Gurion University of the Negev; naomige@bgu.ac.il

1 Introduction

How workers determine their effort 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 their 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 may also demonstrate that efforts can be motivated not only through 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 the latter cannot be directly observed. 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.[2]

To overcome this limitation, we exploit high-frequency individual level data 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. In our setting, player and team performance depend not only on their own efforts and ability but rather also on the ability and efforts of the opponent team as well as on additional circumstances such as weather, crowd, or media attention. This further enhances the advantage of directly focusing on effort when aiming to analyse behavior.

[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] For example, 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.

Another interesting feature of our setting is that similarly to other prevalent types of work teams such as research and development groups and court litigation teams, 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. For example, in football, the number of goals is not an appropriate measure of defenders' performance. In such cases, analysing effort could be a more efficient way to measure the contribution of each team member.

Like this paper, several leading studies have chosen to investigate team production in the context of sports teams. This is because the highly detailed data that are available about such teams offer unique opportunities to explore intricate production functions.[3] 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 to create a causal relationship between their efforts. First, complementarity in production could drive a positive correlation between efforts because a worker's marginal productivity increases with the effort of his peers, thus making it optimal for him to increase his effort as well. The opposite could occur when workers are substitutes in production, in which case workers may "free ride" their teammates' efforts. Gould and Winter (2009) use measures of performance for professional baseball players to show that the type of production function (complements or substitutes) determines whether workers increase or decrease their performance when their teammates perform better. In the case of football, it is reasonable to assume that players' efforts are mostly complementary. Moreover, since players' efforts are constantly observed both by the crowd and the coach and because explicit measures of effort are readily available, free riding will be minimal even if to some extent players' efforts are substitutes. Thus, it is reasonable to hypothesize that this mechanism will create a positive relationship between teammates' efforts.

A second potential mechanism that has been considered in the literature is behavioral, namely positive effort and productivity spillovers in groups of workers can be the result of either social pressure, prosocial behavior, or shame. Contrary to complementarity in production, behavioral effects can be dominant among workers in the same location even when the task performed is an individual one. For example, cashiers' productivity was found to be positively affected by those of neighboring cashiers who can monitor their behavior (Mas and Moretti, 2009). We hypothesize that this will create a positive relationship between teammates' efforts in our football setting as

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

well.[4]

A third mechanism that has been shown to determine interdependent effort choices among coworkers is an employer's incentive scheme or more generally the institutional setting. For example, social connections between managers and their workers were found to impact productivity but only when the manager's salary is *not* determined by the team's output (Bandiera et al., 2009). In a different setting, Amodio and Martinez-Carrasco (2018) show that evaluation of workers according to average group performance leads to free riding, which might be alleviated by introducing a piece-rate pay. In football, studies have documented that players' efforts are not manifested in their market values (Wicker et al., 2013; Weimar and Wicker, 2017), and thus are not likely to be strongly related to the specific pay and reward structure.

Lastly, workers may change their effort in response to changes in peer efforts because they learn from each other and benefit from knowledge spillovers. Such positive spillovers were found specifically for teachers (Jackson and Bruegmann, 2009) and scientists (Azoulay et al., 2010) although not for professional golfers (Guryan et al., 2009). In the football setting, we expect this channel to be less influential. For one thing, similarly to Guryan et al. (2009), players' ability to learn during a game is quite limited. Furthermore, Cornelissen et al. (2017) present evidence indicating that in general the role of learning is less dominant relative to that of the other mechanisms mentioned above.[5]

The above assessment suggests that in our football setting we should expect to find positive interactions between teammates' efforts which was also the most prevalent finding both in previous studies of peer effects in the workplace (Herbst and Mas, 2015). Manski (1993) distinguishes between three effects that may generate similarity in the behavior of individuals and their peers: 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. In football, the endogenous effect would be the impact of peer contemporaneous effort on individual effort, as opposed to, for example, the impact of peer fitness or market value on individual effort, which is an exogenous effect.

Our aim is to identify the endogenous effect, which is the most interesting when considering incentivizing and managing effort in teams and the only effect that embeds a "social multiplier" in the sense that there are repeated feedbacks between teammates' efforts. Thus, even encouraging just

[4]Less directly related to our paper, in other settings when co-workers are competitors, different patterns of peer effects were found. For example, Guryan et al. (2009) use data from professional golf tournaments and find no effect of either the permanent or the instantaneous level of performance of one's peers. Yamane and Hayashi (2015) show that both positive and negative spillovers can be found among adjacent competitors in swimming competitions depending on the swimmer's relative position (whether their peers swim behind them or ahead of them).

[5]They show that positive *wage* spillovers are smaller in high-skill occupations relative to low-skill ones. Since knowledge spillovers are expected to be more pronounced in high-skill occupations, this implies that learning plays a relatively small part in generating positive peer effects among coworkers.

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).[6] Consequently, most studies of peer effects focus on estimating exogenous effects or some combination of these effects.

To address this challenge, we employ two different estimation strategies. Our first strategy follows Gould and Winter (2009) by relying relies on a combination of individual worker fixed effects and an instrumental variable (IV) approach, where peers' contemporaneous behavior is instrumented by their average behavior in other games during the season. However, contrary to their study, we are able to measure peer effects in *effort* rather than in performance as our data provides direct quantification of individual effort. To establish that our measure of effort, namely running distance, is relevant and reliable, 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 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.[7]

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 to control for unobserved ability in order to avoid 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 in all of the other games during 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

[6] We explain this problem in more detail when we discuss our empirical strategy in section 3.2.1.

[7] To corroborate and validate our findings, this team level specification is also estimated using data from the European Champion's League.

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 major events that occurred in the previous section. We believe that within-game tactical changes are usually triggered by such major events. 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 results are not driven by strategic choices of the coach, we also control for the formation of the team, namely the number of players in each position (defenders, strikers, and midfielders). The main findings clearly 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 70-75% of the increase in the average team effort.

To reinforce this finding, we use a second empirical 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 that 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 obsarvables 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 (excluding the current 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.

In order to shed light on the specific mechanisms that drive these positive peer effects, we estimate heterogeneous effects by players' experience (age) and tenure. We find that senior players are less responsive to peer behavior. 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 and that peers tend to be more cooperative when their team is the underdog.

The paper proceeds as follows. In section 2, we describe the data and provide background on the setting. Section 3 first presents evidence on the relation between effort and output and then reports our main results on peer effects. Section 4 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).[8] Here, we describe only the most relevant institutional details of the league and the games, while a more detailed description is provided in Appendix section 5.1.

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 5-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.[9] Team ranking and BO are used to proxy for the relative ability of the teams in each match.

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 and then the league splits into 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).[10] The total number of players who participated in this 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.

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

[8]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).

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

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

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	0.25
	(0.431)

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.[11] Here, as well, we use game level data from the original documentation on average player-speed and on team distance rather than calculating it 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.

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.

[11] 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.

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 3 below, where we introduce the empirical methodology.

3 Empirical Estimation and Results

3.1 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 number of goals or 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. Although speed appears to be a very natural measure of effort, other aspects of effort exist that we cannot measure (e.g. level of concentration, pre-game physical and mental preparation), as well as factors other than effort which can be generally referred to as ability and may impact both speed and performance. 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.

As discussed below in more detail, 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. We add team, individual, and pair fixed effects to control for time-invariant ability and relative ability, game fixed effects to account for game specific unobservables, and section fixed effects to control for within-game trends. 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 econometricians.

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.[12] We limit our sample

[12]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 it does not make sense to use a player's distance at the game level since players do not

to include only strikers and midfielders and ignore defense players who we assume have a much lower probability of scoring a goal.[13] 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:

$$goal_{i,s,g} = \alpha_i + \zeta_s + \gamma_g + \theta avg\_speed_{i,s,g} + X'_{i,s,g}\delta + \epsilon_{i,s,g} \quad (1)$$

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 observed game,[14] 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)).[15] We also add controls 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, and the average speed of the team in section $s-1$. An additional control that was shown to be important in previous studies is the size of the crowd for each team, but this data is missing for four rounds in the season and thus we only include this control as a robustness test. In any case, since our main specification includes game fixed effects, crowd size only varies across teams in the same game.[16]

We 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 team fixed effects.[17] 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 at the player level.

necessarily play for the entire 20 sections of the game. Thus, speed is a more reliable measure at the *player* $\times$ *game* level.

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

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

[15]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.

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

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

The results of the estimation are reported in Table 2 and show that, *within* player and game, 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 *within-game* 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 2) 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 0.3 percentage points, which is an almost 30% increase in this probability (0.98%, as reported in Table 1). In column (5) 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.

Column (6) presents the same estimate at the game level and confirms that the level of effort is positively and significantly related to the goal probability.[18] Notably, the game-level 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.

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

$$win_{i,r} = \eta_p + \rho_r + \theta_1 team\_dist_{i,r} + \theta_2 opp\_dist_{i,r} + Z'_{i,r}\nu + \xi_{i,r} \tag{2}$$

where $win_{i,r}$ denotes the probability that team i wins in round r.[19] 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.

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

[18] The regression equation includes the same basic controls as in equation (1). Obviously, when the observation is at the game level, we cannot control for the number of section and pre-section events. We should also note that adding crowd and BO controls decreases the number of observations but has no substantial impact on the results.

[19] 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).

Table 2: Player's running speed and goal probability

	Dependent Variable: Goal Probability					
	Section					Game
	(1)	(2)	(3)	(4)	(5)	(6)
Speed (kmph)	0.00282***	0.00278***	0.00288***	0.00286***		0.0403***
	(0.000513)	(0.000514)	(0.000535)	(0.000572)		(0.0103)
Number of Sprints					0.00336***	
					(0.000449)	
Observations	47,922	47,922	45,200	39,444	45,200	3,860
R^2	0.014	0.014	0.015	0.015	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—5 the unit of observation is at the *player* $\times$ *section* level, and all regressions include player, section and game fixed effects. In column 6 the unit of observation is at the *player* $\times$ *game* level, and the regression includes player and game fixed effects. They also 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. Columns 2—5 also controls for a quadratic trend in the number of sections that the player played in the game up to section s. In columns 3—5 we also add controls for the score gap, goals, red cards, and injuries in section $s-1$ as well as for the team's average speed in this section. 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 at the player level.

game fixed effects. 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 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 with a tie, and such an outcome 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 (with and without controls). 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	0.0437***	0.0426***	0.0407***	0.0707**		
	(0.0118)	(0.0118)	(0.0142)	(0.0283)		
Opponent Distance	-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						-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 home team and 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 2 using data from the UEFA Champion League. We exploit data on 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.

3.2 The interaction between teammate efforts

After establishing 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 aims to investigate, namely how the efforts of different players on the same team impact each other. 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.

As explained above, our empirical strategy targets the identification of the *endogenous* peer effects, as defined in Manski (1993), namely how a team member's effort is affected by a change in his peers' effort. By definition, this is an iterative effect that may result in a large final impact even when the initial change is relatively small, and is therefore of particular interest to employers, firms, and policymakers alike.

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). Moreover, 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). We employ several different estimation strategies to overcome these 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.

3.2.1 How group effort impacts individual effort

We begin this analysis with an estimation of 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; hence we implicitly assume linearity in the aggregation of effort. As pointed out by Gould and Winter (2009), the exclusion of player i, in addition to the use of individual fixed effects, verify that 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 due to two additional potential concerns.

First, in our setting, the identity of peers cannot be considered random because of 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 fixed effects can effectively control for the first two levels and allow us to rely only on within-game variation in team composition due to substitutions. The fact that the rules of the game allow a maximum of three substitutions limits this variation but also deters frequent within-game adjustments that increase concerns for endogeneity. To further refute these concerns 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 between compositional changes that involve a formation change and those that do not. Conditional on these controls, we argue that the composition of the team can be plausibly considered exogenous.

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 the result of the endogenous 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. In our case, this strategy would imply using the average season-level effort of player i's peers, while relying on variation in 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, this 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, excluding the current game, affects a player's current running speed only through its impact on contemporaneous peer effort.

Thus, we can use the season-level measure as an instrument for contemporaneous team effort, thereby netting out correlated effects and cleanly identifying the endogenous effect. This 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 seasons. Third, contrary to previous studies, we are able to directly measure effort and therefore estimate the interaction between players' *choices* rather than their outputs.

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

$$dist_{i,t,s,g} = \omega_i + \sigma_s + \lambda_g + \beta_1 \bar{dist}_{-i,t,s,g} + X'_{i,t,s,g}\mu + \varepsilon_{i,t,s,g} \tag{3}$$

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 1, we include individual, game, and section number fixed effects, and control for the same observable characteristics in addition to team formation. 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). As explained above, the only remaining obstacle to a causal interpretation of β_1 is that events that occur during section s in game g could cause a correlated change in the efforts of teammates.

Therefore, we will also estimate equation 3 using 2SLS and instrumenting $\bar{dist}_{-i,t,s,g}$ with the average running distance of player i's peers in section s of all games in the season *excluding* game g.[20] This approach relies *only* on the variation in peer composition in section s across games which can be plausibly considered 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.[21]

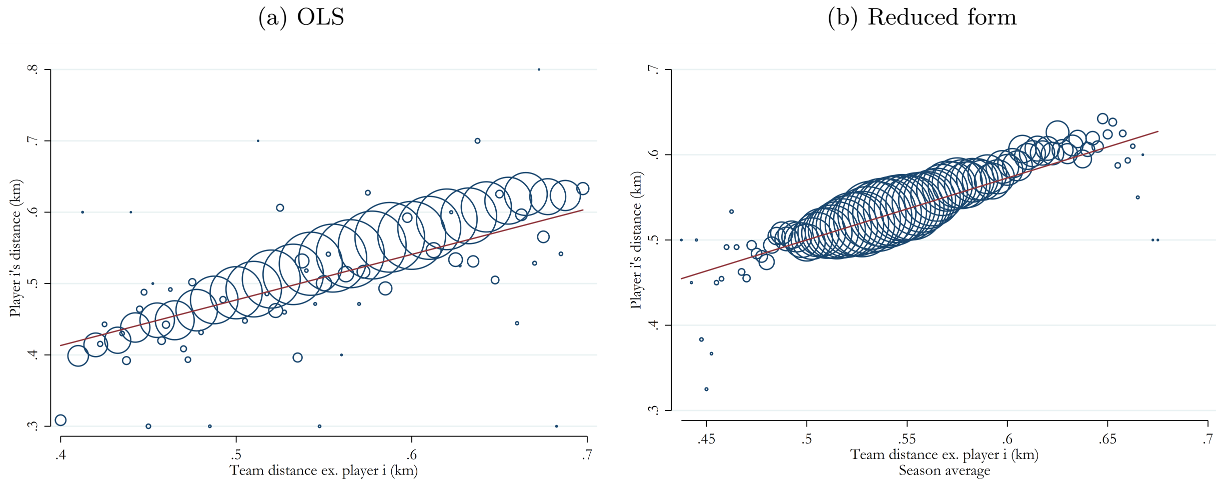
Panel (a) of Figure 1 presents the raw data correlation between the outcome variable and the endogenous explanatory variable $\bar{dist}_{-i,t,s,g}$ with a linear fitted line. In panel (b), we plot the reduced form relationship between the outcome variable and the instrument. Both plots show a substantial 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 and thus next, we turn to regression results.

We first report the results of the OLS estimation in Table 4. The estimated effect is remarkably stable across the different specifications, and indicates that a player will increase his running speed by approximately 82% of the increase in the average peer effort. The point estimates slightly

[20] 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.

[21] This correlation is obviously expected to be monotonous.

Figure 1: Teammates' and Player's Efforts



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 of all the other games in the season.

decrease when we add player fixed effects in column (2), but then remain practically the same as more controls are added to the model, implying that none of these characteristics and events systematically affect the interaction between teammates' efforts. In column (6), we also show that a similar estimate is obtained when 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.

However, as explained above, the OLS estimation is likely to overestimate the effort peer effects and therefore in Table 5 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). The point estimates decrease but remain positive and highly significant across specifications. Quantitatively, the IV estimates indicate that an increase of 100 meters (324 feet) in the average distance that a player's teammates run during a given section will cause an increase of 76 meters (250 feet) in the player's running distance during the same section. These results also imply that a relatively small fraction of the correlation between teammates' running speed is due to within-section occurrences. That is, about 93% of the combined effect can be attributed to the endogenous effect (calculation is based on the specification in column 3 in Tables 4 and 5). To further refine this conclusion, we should keep in mind that this model estimates local average treatment effects, which in our context means

Table 4: Teammates' and Player's Efforts - OLS with Player Fixed Effects

	Dependent Variable: Player's running distance					
						No Red cards
	(1)	(2)	(3)	(4)	(5)	(6)
Team distance (ex. player i)	0.836***	0.813***	0.816***	0.819***	0.824***	0.813***
	(0.00850)	(0.00847)	(0.00851)	(0.00870)	(0.00915)	(0.00917)
Observations	80,444	80,444	80,444	75,886	66,251	60,941
R^2	0.568	0.580	0.589	0.581	0.584	0.587
Controls			√	√	√	√
Pre-section Controls				√	√	
Crowd & BO Controls					√	
Player FE	√	√	√	√	√	√
Section FE		√	√	√	√	√
Game FE		√	√	√	√	√

Notes: The table presents OLS estimates. The unit of observation is player × section. All regressions include player fixed effects and columns 2—6 also add 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). Columns 3—6 include controls for player i's team being the home team, the rank gap between player i's team and the opponent team before game g, and a quadratic trend in the number of sections that the player played up to section s. In columns 4—5, 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. 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 are excluded from the sample. Standard errors are clustered at the player level.

that it measures the interaction between own and teammate efforts when their contemporaneous effort is positively correlated with their permanent effort. Overall, the findings imply that when teams are engaged in a collaborative task, one of the most important determinants of individual effort is that which is exerted by his peers.

3.2.2 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 how a change in individual effort impacts his peers' effort, 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. We proceed by presenting descriptive evidence to establish that fatigue is a substantial determinant of players' efforts and thus substitutions provide

Table 5: Teammates' and Player's Efforts - IV Approach

	Dependent Variable: Player's Running Distance					
						No Red cards
	(1)	(2)	(3)	(4)	(5)	(6)
Team distance (ex. player i)	0.917***	0.714***	0.757***	0.762***	0.765***	0.773***
	(0.0128)	(0.0187)	(0.0188)	(0.0198)	(0.0209)	(0.0203)
F-statistic on Excluded Instrument	23449.13	5885.87	6186.28	5328.56	4686.63	4913.89
Observations	80,444	80,444	80,444	75,886	66,251	60,941
Controls			√	√	√	√
Pre-section Controls				√	√	
Crowd & BO Controls					√	
Player FE	√	√	√	√	√	√
Section FE		√	√	√	√	√
Game FE		√	√	√	√	√

Notes: The table presents 2SLS estimates. The unit of observation is player × section. All regressions include player fixed effects and columns 2—6 add 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) instrumented by these players average running distance in section s in all games throughout the season excluding game g. Columns 3—6 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, and a quadratic trend in the number of sections that the player played up to section s. In columns 4—5, we 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. 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. Standard errors are clustered at the player level.

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.

First, in Figure 2, we show that the average effort of individual players generally decreases as the game progresses.[22] 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 that 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 clearly establishes that fatigue is a central determinant of effort.

We find additional support to this claim in Figure 3 that presents the seasonal number of substitutions by the number of sections in the game, and clearly demonstrates that very few substitutions are initiated during the first half of the game when players are relatively fresh. During the second

[22]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.

half, the spread of substitutions over the sections is quite uniform, especially between sections 14 and 18, where most substitutions take place.[23] 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.[24]

Therefore, it is expected that the incoming player will run faster than the outgoing player. Figure 4 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.[25] 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 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.[26]

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).[27] This clear pattern indicates that, as expected, fatigue is a substantial determinant of players' effort, and hence implies that substitutions create a quasi-exogenous shift in individual effort.

We build on this shift and test whether the running distance of the team members (excluding goalkeepers) changes in accordance with the distance gap between the incoming and outgoing players. Still, other factors could be involved such as within-game events, the strategic plan of the coach, 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 factors:

$$\Delta dist_{i,k,s,g} = \tau_s + \phi_g + \beta_2 \Delta dist\_gap_{k,s,g} + W'_{i,s,g}\kappa + \psi_{i,k,s,g} \tag{4}$$

where the main explanatory variable is the distance gap between the incoming and outgoing players in substitution k. The dependent variable is the gap between the pre- and post-substitution k

[23]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.

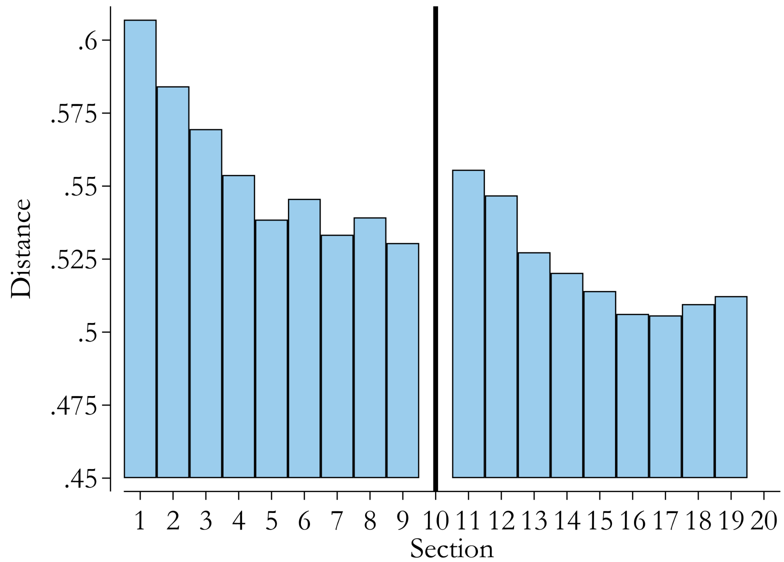
[24]The dashed line in Figure 3 shows that the number of substitutions following an injury is only 85 out of a total of 1335 and that they are spread quite evenly across all game sections, except for an arbitrary spike in section eight.

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

[26]Moreover, these two limitations are exacerbated by the fact that our data on distance is rounded in each section to 0.1 units.

[27]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 2: 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 thus are omitted.

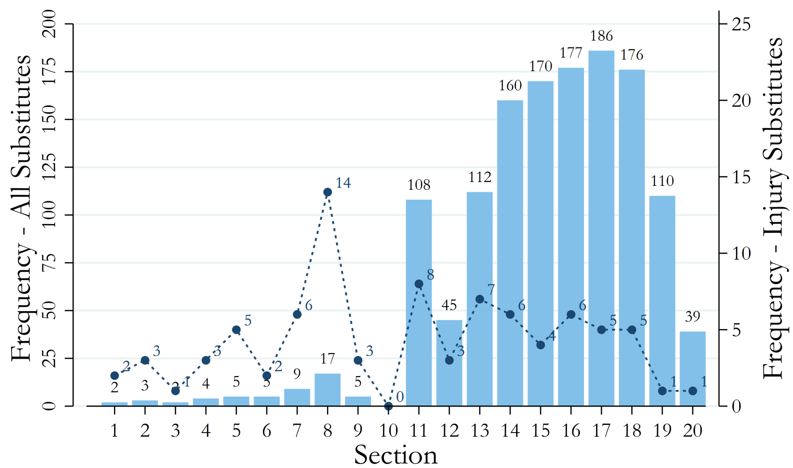
distances for player i who plays for the team that initiated substitution k but did not participate in it. Thus, β_2 captures the relationship between a change of one team member's effort due to the substitution and the consequent, simultaneous change in player i's effort. Because our outcome measure is the *difference* in effort, we are not concerned that it reflects the ability or the average effort level of the player. As in all our previous regressions, we also include section number and game fixed effects and a rich set of controls for game, player, and previous section characteristics.[28]

To further refute 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 address 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.

To address this shortcoming, we use an IV approach which is similar to the one that was presented in section 3.2.1, 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

[28]In light of the fact that the main outcome variable is the difference in player i's distances, we do not include player level fixed effects in the estimation.

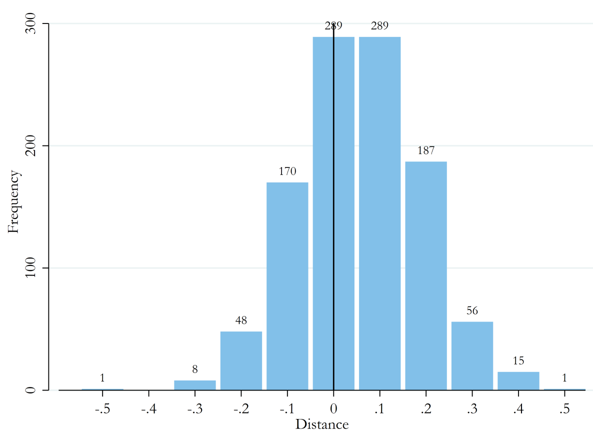
Figure 3: Number of substitutions by section



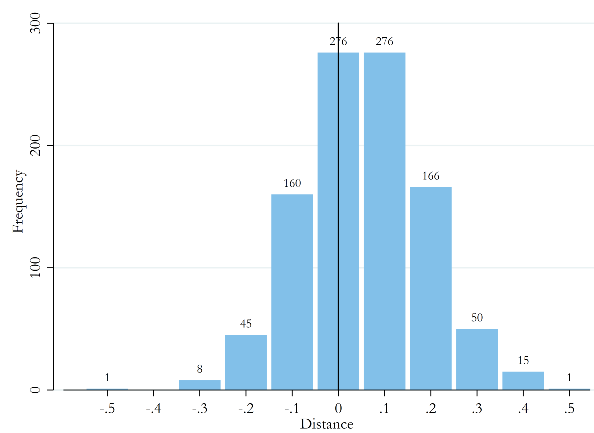
Notes: The figure presents the number of substitutions that were initiated in each section throughout the entire season. Goalkeeper substitutions are excluded from the sample. The bars correspond to the left hand side Y-axis and present counts of all types of substitutions. The dashed line shows the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

Figure 4: The distribution of the distance gap between players in substitutions

(a) All $+/-1$



(b) Non-injury, $+/-1$



Notes: The figure presents the distribution of the gap in running distance between the incoming and outgoing players of each substitution, excluding goalkeeper substitions. 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) only reports the gaps for substitutions that were not induced by injury.

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 in other games in the season.[29] 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 using all the sections which were his first in all the other games in the season. 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 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, clearly, 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.

Before proceeding to the results of our estimation, in Figure 5 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.[30]

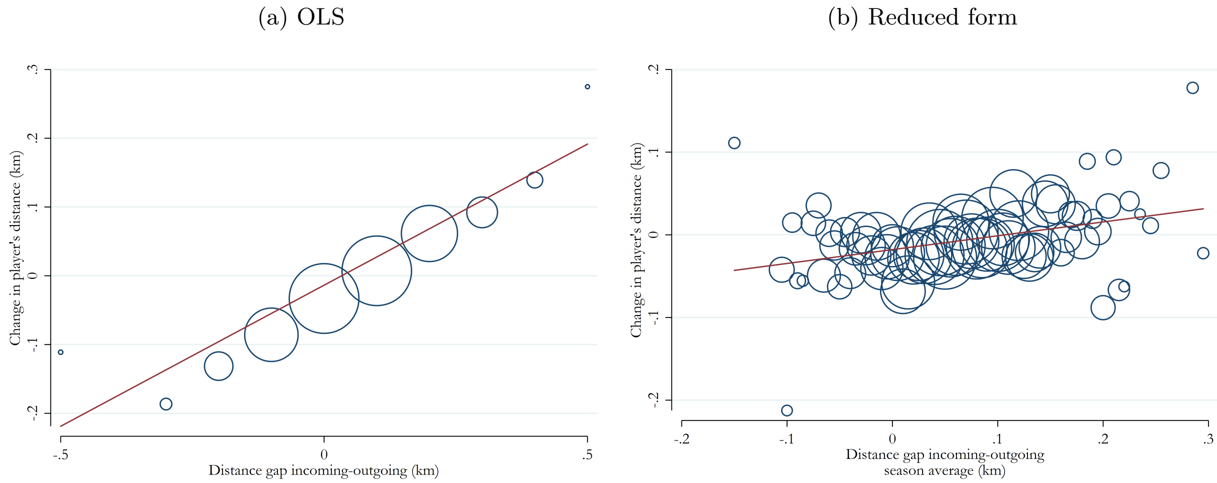
Table 6 presents the estimated impact of the distance gap in the substitution on the change in the other players' effort using both the OLS and IV strategies and for four different samples. We use different sample restrictions that refer to various within-game events and intermediate outcomes because previous studies in sports economics document that such events could impact players' behavior (Deutscher and Schneemann, 2017; Schneemann and Deutscher, 2017). All of the sixteen resulting point estimates are positive and very precisely estimated. The estimated impacts are also of a similar magnitude across samples.

The first row of the table presents the results of the estimation when we use the largest possible sample. This sample includes 1,064 substitutions in column (1) and 1,030 substitutions in column (3). The number of substitutions and thus observations in our IV analysis is slightly smaller because for some of the players that participate in the substitutions we cannot use the suggested IV since these players never played the same number of sections in other games in the season. The effects that we estimate in columns (3)—(4) are approximately 40% of the magnitude of the effect that is estimated in columns (1)—(2). 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 that participate in the substitution

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

[30] 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)).

Figure 5: Changes in Teammates' and Player's Efforts - Substitutions



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.

comply, namely that the actual distance gap is positively correlated with the effort patterns that these players displayed throughout the season. In the second row, we show that the results hold for a sample that excludes substitutions that were preceded or followed by a score change or a red card for either team.[31] This helps refute 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 may substitute a striker with a defense player and also instruct the entire team to slow down and play more defensively. This will cause a negative distance gap in the substitution and also a negative gap for all of the players in the team, which will 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, in our case it helps ensure that the effect that we find 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. Last, 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 and find that the point estimates are similar and even slightly larger.

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

Taken together, these results demonstrate that even when just one player increases his effort, the efforts of his peers and hence of the entire team may increase substantially. To illustrate how a change in an individual player's effort is expected to indirectly affect game outcomes through the effect on his peers, we conduct a back-of-the-envelope calculation in which we apply our estimates of how team effort is affected and how it is related to the team's winning probability. This calculation is merely indicative of the potential magnitude of this effect. We assume an increase of 10% in one player's speed, which is on average 6.29 kmph, namely the player runs additional 0.629 kilometers per hour or 0.052 kilometers per section. As, on average, a player plays 16.05 sections per game, this implies an increase of 0.84 kilometers per game. Table 6 reports that an increase of one kilometer in an individual player's running distance causes an average increase of between 0.173 and 0.251 kilometers in the running distance of each of his nine teammates. Thus, an increase of 0.84 kilometers will lead each of them to increase their running distance by between 0.15 and 0.21 kilometers and therefore the overall running distance of his teammates will increase by between 1.35 and 1.89 kilometers per game. According to column 2 of Table 3, this change will increase the probability of winning the game by 5.75 to 8.05 percentage points, which amounts to a substantial increase of between 14.4% and 20.1% in the average probability to win (approximately 40%).

3.3 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 peer effects in our context: complementarity in production 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 with different levels of experience (age), and tenure. 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.[32] The findings indicate that players who are more experienced (above the 75th percentile of age) or more senior in the team respond less to changes in their peer effort, although

[32] Nonetheless, we note that heterogeneity patterns are consistent across the two estimation strategies except for different significance levels.

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

	Dependent Variable: Difference in Player's Running Distance			
	OLS		2SLS	
	(1)	(2)	(3)	(4)
All				
Distance gap (incoming-outgoing)	0.471***	0.468***	0.178***	0.179***
	(0.00949)	(0.00952)	(0.0249)	(0.0242)
F-statistic on Excluded Instrument			2081.3	2108.7
Observations	*9,055*	*9,010*	*8,764*	*8,719*
No red cards or goals				
Distance gap (incoming-outgoing)	0.454***	0.457***	0.208***	0.227***
	(0.0159)	(0.0160)	(0.0422)	(0.0395)
F-statistic on Excluded Instrument			966.3	1078.2
Observations	*3,428*	*3,401*	*3,350*	*3,323*
Non-injury substitutions				
Distance gap (incoming-outgoing)	0.472***	0.468***	0.173***	0.175***
	(0.0103)	(0.0103)	(0.0255)	(0.0250)
F-statistic on Excluded Instrument			1984.5	1966.5
Observations	*8,473*	*8,464*	*8,182*	*8,173*
No formation-change				
Distance gap (incoming-outgoing)	0.484***	0.484***	0.251***	0.243***
	(0.0152)	(0.0155)	(0.0433)	(0.0446)
F-statistic on Excluded Instrument			687.7	612.4
Observations	*4,703*	*4,676*	*4,558*	*4,531*
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 in the substitution. The outcome variable is this difference for player i who was on the field both before and after the substitution. Columns (1) and (2) present the results of an OLS estimation; columns (3) and (4) present the results of a 2SLS estimation where the endogenous explanatory variable is instrumented by the difference in the season average running distance (excluding game g) 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 and columns (2) and (4) 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 section that precedes the substitution, the team's average speed in this section, and a quadratic trend in the number of sections that the player played up to section s. 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 at the player level.

Table 7: Heterogeneous Effects

	Dependent Variable: Player's Running Distance			
	(1)	(2)	(3)	(4)
Characteristic:	Age$\geq$75p	Tenure$\geq$2	Team tenure$\geq$2	Odds
Team distance (ex. player i)	0.767***	0.772***	0.754***	0.754***
	(0.0220)	(0.0188)	(0.0191)	(0.0186)
Team dist. $(-i)\times$ Char.	-0.0311	-0.0451**	0.00814	
	(0.0302)	(0.0193)	(0.00539)	
Team dist. $(-i)\times$ Win				-0.00138
				(0.00524)
Team dist. $(-i)\times$ Loss				0.00904*
				(0.00483)
Player FE	$\surd$	$\surd$	$\surd$	$\surd$
Section FE	$\surd$	$\surd$	$\surd$	$\surd$
Game FE	$\surd$	$\surd$	$\surd$	$\surd$
Observations	79,380	78,756	78,756	80,090

Notes: The table presents 2SLS estimates. The unit of observation is player $\times$ 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 excluding game g. All regressions include player, game and section fixed effects in addition to 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, a quadratic trend in the number of sections that the player played up to section s. In each column, the specification includes an interaction term of the main explanatory variable and a dummy variable that indicates the characteristic specified in the column title: age above 75th percentile, tenure of two seasons or more, median tenure in the team of 2 seasons or more, and the odds of winning and losing according to BO data. Goalkeepers are excluded from the sample. Standard errors are clustered at the player level.

this differentce is only significant for tenure. Therefore, it is plausible to infer that the insignificant difference between younger and older players is driven by differences in tenure rather than general experience or direct age effects. The point estimate for senior players is approximately 6 percent smaller. 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, except that they produced a higher point estimate for players in the "underdog" team. Since these factors are not expected to affect the degree of complementarity in team production, this finding again indicates that a behavioral mechanism creates at least some of the peer effects.

4 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 which also capture ability. Using a unique high-frequency dataset from the Israeli Professional Football Leagues 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 identify 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 implies that behavioral considerations play an important role in the generation of peer effects in teams.

Although caution must be exercised in generalizing the findings of this study to other labor markets, it is plausible to expect similar peer effects to operate when collaborative team tasks are involved. 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 be the key to exploiting effort peer effects to increase the productivity of teams. This demonstrates why it can be beneficial for organizations to encourage social interaction among coworkers and to invest in activities that help forge team spirit.

References

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.

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.

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.

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.

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.

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.

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.

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.

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

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.

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.

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.

Yamane, S. and R. Hayashi (2015). Peer effects among swimmers. *The Scandinavian Journal of Economics 117*(4), 1230–1255.

5 Appendix

5.1 Basic Rules of Football and the Israeli League

A football match is played between two teams, each with 11 players on the field. It lasts 90 minutes, which are 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, and to prevent 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.

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.

5.2 Appendix Tables

Table A1: Teams' running distance and game outcomes

	Dependent Variable:					
	Team winning probability					Goals
	(1)	(2)	(3)	(4)		(5)
Team Distance	0.0391***			0.0190**		0.0704***
	(0.0106)			(0.00933)		(0.0237)
Opponent Distance	-0.0476***					
	(0.0126)					
Distance Ratio		4.653***				
		(1.143)				
Total Distance		-0.00388				
		(0.00470)				
Distance Diff.			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 the game. In column 1, the main explanatory variable is the aggregate distance that team i's players run during the game and 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 – Champion league

	Dependent Variable: Team i winning probability					
	(1)	(2)	(3)	(4)	(5)	(6)
Team Distance	0.0619***	0.0579***	0.0689***	0.0580***		
	(0.0147)	(0.0152)	(0.0182)	(0.0152)		
Opponent Distance	-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						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 2017-2018. Estimates are at the *team* × *game* level, where team i is selected according to the 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 home team and 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 BO (which are missing for 1 game). In column 4, the sample is restricted to games without red cards. Standard errors are clustered at the pair level.