Mechanisms of Learning in Collaborative Jobs*

Gábor Békés[1] and Bence Szabó †[2]

[1]Central European University, KRTK, CEPR
[2]Corvinus University of Budapest, Magyar Nemzeti Bank

November 18, 2025

Draft. Link to latest version

Abstract

We investigate novel mechanisms of learning in collaborative environments with granular data of European men's football, overcoming key limitations of linked employer-employee administrative data. First, we find that conditional on average quality, team composition and exposure to stars do not affect learning. Second, peer quality affects task allocation, leading standard approaches to underestimate learning effects by 30%. Third, intensive collaboration with high-quality peers drives learning. Fourth, peer effects manifest in learnable skills (passing, reactions) but not in innate abilities (speed), providing evidence against selection bias. Finally, learning persists after job changes, suggesting genuine human capital accumulation rather than temporary complementarities.

Keywords: Learning, collaboration in teams, peer effects, stars, sports data

JEL-codes: E24, J31, O33, D83, Z20

*We thank Paolo Falco, Daniel Prinz, Zsófia Bárány, Balázs Reizer and participants at workshops: Hungarian Society of Economics, Central European University, KRTK Institute of Economics, and Corvinus University of Budapest.

†Corresponding author:bence.szabo@uni-corvinus.hu

1 Introduction

Learning from colleagues is an essential source of human capital growth. Compelling evidence shows that long-run wage growth is strongly related to peer quality: workers in firms with 10% higher average coworker pay experience 0.7% higher wage growth within one year, cumulating to 2.1% over a decade (Jarosch et al., 2021). Similar patterns emerge in various contexts, from Germany (Cornelissen et al., 2017; Herkenhoff et al., 2024) to Sweden (Nix, 2020), Brazil or Italy (Arellano-Bover and Saltiel, 2023). While wage growth is driven in part by firms learning about workers' ability, human capital acquisition plays a primary role (Pastorino, 2024).

Despite this robust evidence on peer effects, the *mechanisms* through which human capital accumulation occurs remain poorly understood. Existing evidence reveals substantial heterogeneity: Mas and Moretti (2009) document peer effects among supermarket cashiers driven primarily by social pressure rather than skill transfer, while Cornelissen et al. (2017) find near-zero peer effects in knowledge-intensive occupations where learning should matter most. These disparate results likely reflect different mechanisms operating across different circumstances and unobserved learning contexts. Even when peer effects are correctly identified, we cannot discern *how* learning occurs: through observation versus active participation, through broad exposure versus intensive collaboration with specific individuals, or through acquiring general skills versus context-specific knowledge. Indeed, administrative data seldom reveal who the actual coworkers to learn from are, and which workers actually interact.

We address these challenges using granular data from professional football. This setting offers three key advantages for identifying learning mechanisms. First, we observe the production process directly and can identify actual peers and their interactions, eliminating the ambiguity about who learns from whom. Second, we can measure access to senior workers through playing time and collaboration intensity, allowing us to estimate the trade-off between learning from peers and learning by doing. Third, we observe frequently-updated market valuations and skill evolution across multiple dimensions, enabling us to distinguish human capital accumulation from selection on unobserved ability. These data allow us to document four key findings about how workers learn from peers.

The key main challenge is that we often cannot identify the actual peers from whom workers learn. In typical linked employer-employee datasets [1], peers are defined by sharing the same occupation at the same firm, without observing actual working relationships or interactions. Professional football offers critical advantages over traditional administrative data in this regard. We observe the *production process directly*, as the data cover every minute played, every pass completed, and more. Importantly, we can be certain that all players on a team interact professionally. Team size and roles are institutionally fixed, eliminating endogeneity concerns about firm size and peer exposure that complicate interpretation in administrative data. Furthermore, we track *direct peer*

[1] As in Germany (Jarosch et al., 2021) or XXX

interactions: the number and quality of passes exchanged with specific teammates, distinguishing broad exposure from intensive collaboration.

Access to senior workers matters for learning, yet this access may be constrained in equilibrium. When high-skilled workers are available, firms may assign complex tasks to them for immediate productivity, potentially reducing junior workers' access to skill-building opportunities (Gibbons and Waldman, 1999, 2006). This creates tension between 'learning from peers' through knowledge spillovers and 'learning by doing' through hands-on responsibility. The literature on mentorship and exposure to superstars shows mixed findings. Evidence from science suggests that early-career exposure to eminent mentors drives long-run success (Azoulay et al., 2019; Waldinger, 2012).[2] Yet identifying mechanisms remains challenging: does proximity to stars facilitate learning, or crowd out opportunities as stars dominate key tasks (Azoulay et al., 2019)? Without observing task allocation and collaboration patterns, researchers cannot estimate this trade-off and may systematically underestimate the gross learning effect from better peers if crowding out dampens the net effect. In our setting, playing time serves as a key metric of access to top players. Younger and less valued players spend considerably less time on the pitch and interact less frequently with top players.

The paper is also related to the measurement of human capital accumulation. If better firms systematically hire workers with higher growth potential or reveal workers' ability with more success (Pastorino, 2024), they find 'hidden gems' unobservable to the econometrician more frequently. Then, standard peer effect estimates conflate human capital accumulation with selection on unobserved ability. Here, we observe *frequently-updated market values* alongside wages. Transfer market valuations, determined by clubs, agents, and expert consensus, update at least twice per season and incorporate growth potential. Furthermore, we observe *skill evolution across multiple dimensions*, including both learnable technical abilities (passing accuracy, tactical positioning) and fixed physical attributes (speed, strength), providing a natural test for selection bias.

While football is a specialized setting, it shares key features with many high-skilled, competitive, collaborative professional environments. Players, like workers in consulting, R&D, or academia, observe peer quality with reasonable accuracy, work in teams where collaboration is essential, and face competition for scarce high-value opportunities (Fonti et al., 2023). The wage distribution of elite football players closely resembles that of top-decile earners in administrative data (Jarosch et al., 2021), and careers exhibit similar dynamics of learning and sorting.

This new dataset allows us to uncover how peer learning operates in practice. We first confirm substantial peer effects but uncover a critical omitted variable bias. Players joining teams with 10% higher average peer quality experience 3.3% higher wage growth over three years. However, this estimate substantially understates the true learning effect. Better peers reduce individual playing time by 35% and passing network centrality by nearly 50%, creating a 'learning from peers' versus

[2] Spillovers from exposure to high-quality senior figures are documented in industrial innovation (Bell et al., 2019), the fine arts (Fraiberger et al., 2018), and academic research (Li et al., 2019).

'learning by doing' trade-off. Controlling for this crowding out channel, we find that better peers increase the *gross* learning effect by approximately 30%. This mechanism helps explain why star injuries facilitate talent discovery: they create opportunities for young players to gain hands-on experience (Hoey, 2023). Our finding reconciles seemingly contradictory results in prior work: peer effects may be large in settings where complementarity dominates, and small where crowding out is severe, as in knowledge-intensive occupations (Cornelissen et al., 2017).[3]

Beyond aggregate effects, we show that learning requires active engagement, not mere proximity. Conditional on average peer quality, team composition (concentration in the wage distribution, share of superstars) does not affect learning. However, *direct collaboration intensity* matters enormously. One standard deviation more passes exchanged with top-5% colleagues yields 10% higher wage growth. This finding speaks directly to the mentorship literature: exposure to stars facilitates learning only when accompanied by genuine interaction and collaboration. Proximity alone is insufficient: apprentices must work *with* masters, not merely alongside them.

Examining skill-specific evidence, we find that better peers improve learnable technical skills such as passing accuracy or tactical awareness but have no effect on fixed physical attributes like sprinting speed or strength. This pattern provides strong evidence against pure selection bias. If better teams simply selected 'hidden gems' with higher latent ability across all dimensions, we would observe improvements in both learnable and innate traits. Instead, peer effects operate precisely where learning is possible, validating a causal interpretation. This finding connects to research on heterogeneous human capital (Gathmann and Schönberg, 2010; Yamaguchi, 2012; Lazear, 2009): not all human capital is equally malleable, and peer learning operates on acquirable capabilities rather than endowments.

Finally, we demonstrate that learning persists after job changes. For players who switch teams, the quality of peers at their *initial* team predicts skill growth even three years later at new employers. Peer effects remain significant and stable for team switchers, indicating genuine human capital accumulation rather than temporary production complementarities that would vanish upon mobility. This persistence has important policy implications: early-career investments in learning from the best yield lasting returns, even if workers eventually move to lower-ranked employers. Organizations should facilitate junior workers' exposure to high-quality mentors early, as these benefits compound over entire careers.

The paper proceeds as follows. Section 2 describes our data construction and key variables. Section 3 outlines our empirical strategy and identification approach. Section 4 presents results on team composition, task allocation and crowding out, direct collaboration, skill evolution, and persistence for switchers. In the conclusion we also discuss external validity of these results.

[3]Several sports studies document related mechanisms. Hoegele et al. (2014) examine the role of superstars in team performance, Arcidiacono et al. (2017) estimate how peer quality affects in-game performance via production complementarity, Cohen-Zada et al. (2024) examine effort spillovers in Israeli soccer, and Bilen and Matros (2023) find negative effects where superstar pressure increases mistakes.

2 Data

2.1 Overview

We start from a recently assembled dataset (Békés and Ottaviano, 2025) web scraped from different sources including transfermarkt.com, whoscored.com, that covers each event (each pass, tackle, etc.) from all matches, along with team composition and results of games in the top seven men's football leagues (Premier League in England, Ligue 1 in France, Bundesliga in Germany, Serie A in Italy, La Liga in Spain, the Eredivisie in the Netherlands, and the Portuguese Primeira Liga) over eight sporting seasons (2012-13 to 2018-19, except for Portugal that starts in 2015-2016).[4] We augment them with information on skill attributes and wages gathered from fifaindex.com, and salaries from capology.com. Linking these datasources, we can observe the evolution of players' estimated market values, wages, skill attributes, along with their in-game (at-work) interactions and other individual and team characteristics. The details of the data linking procedure can be found in the Appendix at A.6.

Using sports data in general, and player career data in particular has strong advantages over traditional administrative data, even if it comes with some shortcomings as well. A major advantage of our dataset over several usual administrative datasets with linked employer–employee information is that we can differentiate between simply having different types of coworkers in terms of productivity in close proximity or on our team vs. actively collaborating with them during work, which we can measure using the passes and the shared minutes on the pitch. Furthermore, the usual occupation categorizations are often overly general regarding the actual tasks employees have to perform and therefore require several steps to infer whether learning could actually take place meaningfully.

In our setup, work activity and during-work interactions are directly related to the human capital accumulation of the employees. This establishes a solid connection between learning and interpersonal interactions at the workplace. Our approach is supported by Werner and Dickson (2018) studying peer learning using semi-structured interviews in the Bundesliga. They find that learning from others, knowledge sharing functions via four main channels: imitating, peer communication, labor mobility, and knowledge brokers.

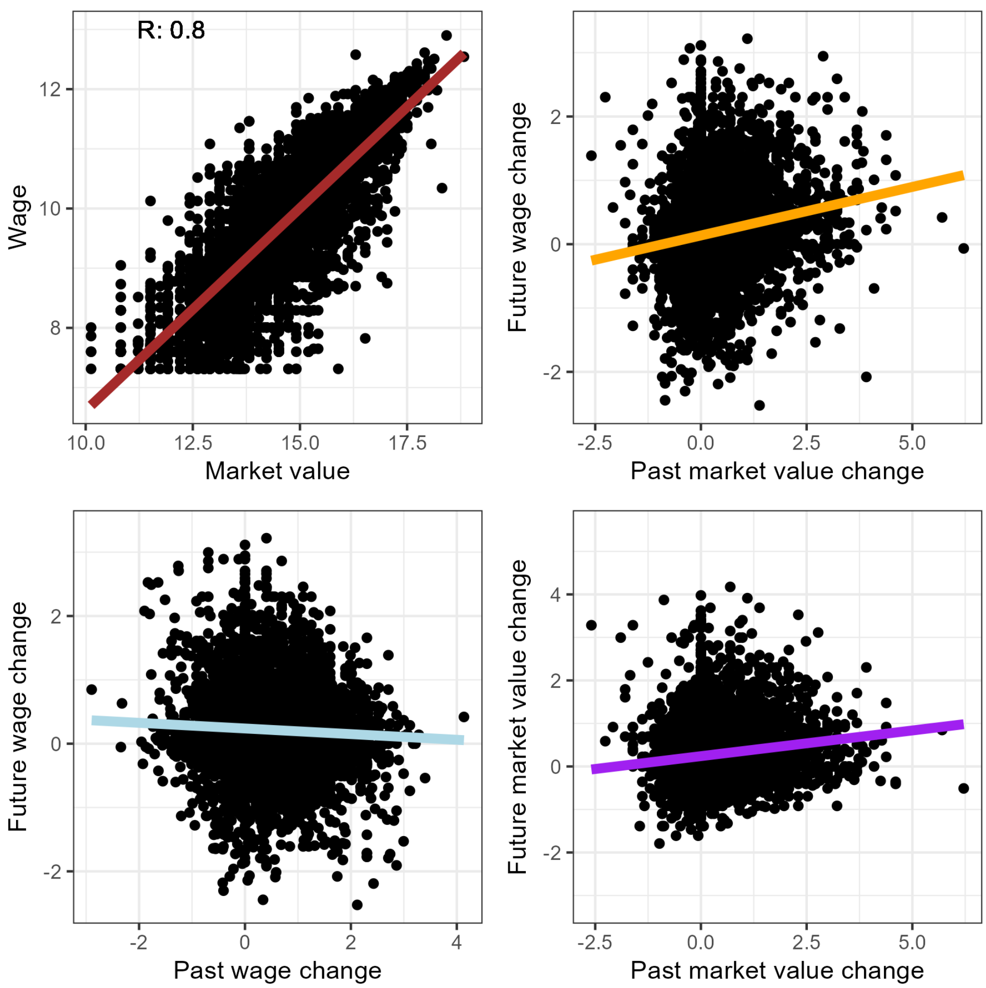
We use three metrics of human capital: wage, market value, and skill attributes highlighting different aspects of human capital. We retrieve market values from transfermarkt.com[5] Their concept of market value targets the *expected value of a player in a free market*, they augment

[4]Data quality and coverage are both very high in our datasets. Nevertheless, small data cleaning steps were needed and we discuss these issues in Appendix (A.6). By 'European football', or simply 'football' henceforth, we refer to 'association football'. The Premier League, La Liga, Serie A, and Ligue 1 are all composed of 20 teams (playing 20×19=380 games), while there are 18 teams in the Bundesliga, in the Eredivisie, and the Primeira Liga (18×17=306 games). Due to relegation and promotion, we have players from a total pool of 181 teams in the sample.

[5]The site's methodology is explained in detail at https://www.transfermarkt.com/market-value-definition/thread/forum/357/thread_id/3433. Transfermarkt publishes them twice in a season (for frequency see Figure A7), usually after the end of a season, and once during a season after sufficiently many games were played.

player pricing models with the community judgement, taking into account factors such as future prospects or contracting (an exhaustive list can be found in Appendix A.7). Using market value alongside wages also allows us to control for the growth potential that might not appear in wages, along with issues of bargaining, tax and accounting considerations (Poli et al., 2021). While market value reacts more sensitively and quickly to current changes in a player's performance and status by constructions, wages incorporate them with delay. As Figure 1 shows, market values and wages are highly correlated but they capture different aspects of human capital: an increase in current market value predicts higher future wages while an increase in current wages predicts the opposite.

Figure 1: Comparing market value and wage dynamics



Note: The figure shows the correlation between different aspects of market value and wage log-levels and dynamics (log-differences) for the final analysis sample.

Finally, as a third measure, we track skill attributes that provide a more direct assessment of how the individual's job-related skills evolve over the lifetime of a player. These are core competences which for other industries remain mostly unobserved, highlighting an additional advantage of using sports data. Skill ratings contain 28 elements scaled 1–100, determined for each player by experts based on advanced stats and scouting information [6]. The correlation structure of the skill ratings can be found in the Appendix at A3, we selected five main attributes for analysis: short pass, reactions, interceptions, finishing, and sprinting speed, reflecting relatively uncorrelated but important aspects of players ability.[7] We expect some skills to develop with collaboration such as short passes, but some are expected to be more unaffected by peers such as sprinting speed.

To capture peer effects, we build on the combined team-level, temporal, and playing-position level distribution of player values to create individual indicators and team-level variables. First, we use the lineup and wage information to produce the average peer quality, along with a Herfindahl—Hirschman index, and related measures. Second, we identify star players as those that have been in the top 5% and 25% of the wage distribution in the relevant top 7 leagues conditional on the season/season-half and their respective playing position, in any period of the last two years. This segmentation of the players allows us to consider different types of interactions between the elite players, very good players, and regular players (keeping in mind that players in these football leagues are incredibly talented individuals in their respective fields). Third we track the direct work interaction between players during all matches in terms of completed passes, pass combinations, and shared minutes on the pitch, enabling us to quantify and finely measure the nature of collaboration in their working environment. Collaboration is measured as the number of minutes and passes in games during regular season, while besides minutes we measure centrality in the production via the eigenvector centrality of a player in the dynamic directed network of pass interactions between teammates.

In our final dataset, we aggregate all variables to a player $\times$ half-season level to reflect the the market valuation update frequency. A half-season contains typically 16-20 games in either the Summer/Fall (August to January) or Winter/Spring (February to May).

2.2 Descriptive statistics

Let us briefly describe our dataset, please find detailed descriptive statistics, variable descriptions, and the steps of creating the final dataset are presented in Appendix Sections A.2 and A.3. Table A10 presents the descriptive statistics of our key variables.

Each player appears in the dataset once per spell at a team for spells that lasted at least two half-seasons. We required at least 6 half-seasons of future observability and valid market values, wages and pass information.

[6] A large team of data editors, reviewers and scouts at Electronic Arts used advanced stats along with extensive scouting and manual adjusments to produce a detailed rating system that represents players well in their computer game. For more details on their approach, visit https://www.espn.com/soccer/story/_/id/37456940/fifa-17-player-ratings-system-blends-advanced-stats-subjective-scouting, accessed 2025/01/23.

[7] 'Short passing' refers to passing ability in closer range, an overall important skill attribute; 'reactions' captures the mental capacity to read situations; 'interceptions' refers to the defensive ability to intercept the passes of opponents; 'finishing' refers to the offensive ability of scoring; and 'sprinting speed' captures the physical ability of running speed.

The differences between the entire panel with some baseline restrictions, and the final sample is reported in A24. The final dataset contains 5,227 observations. The distribution of wages, averaging €1.43 million, is right-skewed with a median of €0.96 million. The values range significantly from as low as €0.08 million (p1) to as high as €8.06 million (p99). There are a handful of players in the top 1% with values close or even above 20 million. The average teammates' wage also has a right-skewed distribution, with a mean of €1.55 million and a median of €1.18 million, while it ranges from €0.14 million (p1) to €7.09 million (p99). The average yearly wage in our dataset is around third of the average market value. Skill attributes are on a 1–100 scale, in our sample they are centered around 50-70 and have a quite symmetric distribution without reaching the maximum rating point.

In terms of concentration of quality, the HHI index has a smaller variation, the top 1% has 0.11 which shows the presence of stars in a squad of about 25 players. The share of top 5% players in the team (the 5% taken for the whole sample) indicates a skewed distribution where most teams rely very little on their top 5% of players, but some teams rely heavily, up to 78% (p99), with a broad dispersion (SD = 0.17). These are the star teams like Real Madrid or Manchester City. We also identify manually a select group of elite teams that are the traditional top teams in their respective leagues and regularly participate in international tournaments. [8] Share of the top 25% players in the team is more democratic, many teams have such players (on average, 27% of the sample has such a teammate). Players' total minutes on the field average 1,077 with a symmetric distribution but a very wide range.

3 Establishing peer effects

In this section we describe our empirical strategy to estimate the mechanisms of learning. We will start by documenting the relationship between wage growth and team quality.

3.1 Human capital growth and peer quality

First we show how wage growth is related to the quality (average wage) of the team. Equation (1) estimates the reationship between wage growth and team quality separately for various horizons (in our case, half-seasons):

$$w_{i,t+h} = \alpha w_{i,t} + \beta \overline{w}_{-i,t} + \gamma X_{i,t} + \omega_{\text{player chars}} + \omega_{\text{league}\times t} + e_{i,t+h} \tag{1}$$

where human capital in half-season $t+h$ is captured by individual i's log wage $w_{i,t+h}$, $\overline{w}_{-i,t}$ is the log mean wage of teammates (excluding player i) in half-season t, and wi, t is player i's own log wage in

[8]These teams are the following: Real Madrid, Barcelona, Atletico Madrid, Manchester City, Manchester United, Liverpool, Chelsea, Tottenham, Arsenal, Bayern München, Dortmund, Schalke, Juventus, AC Milan, Inter Milan, Paris Saint-Germain, Olympique Lyonnais. The selection was done on the basis of historic performance, close to Deloitte's top 20 (an annual report), but is somewhat arbitrary.

half-season t. The vector $X_{i,t}$ includes rich set of individual characteristics (age, contract maturity, injury status, spell count), $\omega_{\text{player chars}}$ captures position fixed effects (instead of occupation fixed effects) and an elite team indicator, and $\omega_{\text{league}\times t}$ denotes league $\times$ half-season fixed effects to control for country-specific and time-dependent differences as labor markets are closely but not perfectly integrated.

The sample contains 5,227 player-season observations representing players' first observation after arriving at a new team between 2012-2019 in seven top European leagues. Indeed, We considered each player-team spell only the first season-half a player joined. We also excluded players for whom we cannot observe 6 half-seasons at all (older players.)[9] We estimate this specification with cross-sectional OLS for different horizons, with standard errors clustered at the league $\times$ half-season $\times$ position level (117 clusters). As common in the literature, establishment fixed effects are not included as they capture almost all variation in peer quality, hindering identification.[10]

Table 1 presents results from Equations (1) and (2) at horizons of two, four, and six periods (one to three years).

First, in columns (1,2,3) we confirm patterns found in the german industry. Having 10% higher earning team mates is associated with a 2.7%, 3.5% and 3.8% higher wage growth over 2,4,6 half-seasons, respectively. Our estimates closely match those of Jarosch et al. (2021) They found a long run average benefit of 2.1% and for top-decile earners – closest to our group of people, a 3.5% wage growth per 10% increase in peer quality. This suggests that our setting is indeed remarkably close to high earning strata of the economy

3.2 Controlling for quality

The baseline model is reasonable for most of the economy, where there are limitations in wage growth as a function of wage. However, for high earners, a key concern is sorting *on growth expectations*: if workers with high wage growth profiles are matched with better firms, it confounds the results.

We address this concern by augmenting wages with post-arrival market values: a perceived (average) value of human capital. These valations include an option value of a worker beyond the current marginal product. In particular, the benefit of using market values beside wages is that it can capture the increase in human capital when it happens, reflecting the match quality between the team and the player, and the human capital growth of young people with higher predicted

[9] We deviated in two ways from Jarosch et al. (2021), who estimated this model as pooled cross section including all worker-year observations. We reproduced this in Table ?? the Appendix, point estimates were about 30% lower. Second, they considered fractured periods, ie workers with only a few observed years.

[10] In Appendix Table A11, we present how much of the peer quality variation is explained by different sets of fixed effects. Around 95% of the variation in peer quality is explained by team and league $\times$ half-season fixed effects. So identifying variation comes from different players across different teams, and not within teams (similarly to Jarosch et al. (2021)).

outside option value emerge sooner. We extend the basline model:

$$w_{i,t+h} = \alpha w_{i,t} + \beta \overline{w}_{-i,t} + \delta(mv_{i,t} - w_{i,t}) + \gamma X_{i,t} + \omega_{\text{player chars}} + \omega_{\text{league}\times t} + e_{i,t+h} \tag{2}$$

where $mv_{i,t}$ is player i's log market value in half-season t, measured after arrival at the new team. The term $(mv_{i,t} - w_{i,t})$ captures the log-difference between market value and wage, reflecting growth potential as perceived by expert consensus.

Market values update at least twice per season and incorporate forward-looking assessments, helping control for unobserved growth potential that wages may not capture due to rigidity or compression.

Results, presented in columns (4) to (6) of Table 1 show that previously we overestimated peer effects, as for the 4 half-season benchmark, average peer quality is associated with 3.3% higher wage growth when controlling for initial market values (Column 5), compared to 3.8% without this control (Column 2). This suggests in circumstances when expeted growth potential could vary, standard peer effect estimates may be upward-biased when growth potential is unobserved. We believe this is the case in the high-earning sectors such as consulting or banking.

Table 1: Wage evoution and peer quality over time

	h=2	h=4	h=6	h=2	h=4	h=6
		$\log(\text{wage}_{t+h})$			$\log(\text{wage}_{t+h})$	
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.538***	0.419***	0.374***	0.583***	0.471***	0.440***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)
log(teammates' mean wage)	0.277***	0.354***	0.386***	0.242***	0.313***	0.333***
	(0.032)	(0.036)	(0.035)	(0.031)	(0.034)	(0.033)
log(value / wage)				0.162***	0.190***	0.241***
				(0.015)	(0.017)	(0.019)
Observations	5,227	5,227	5,227	5,227	5,227	5,227
R^2	0.692	0.559	0.472	0.705	0.575	0.495
Within R^2	0.376	0.237	0.172	0.400	0.265	0.208

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

4 Estimating mechanisms

In this section, we explain our estimation methods and then present key results. We also present some of the results separately for a sample of junior players who were at most 23 years old in t.

4.1 Estimation methods

We investigate the mechanisms through which peer effects translate into human capital growth. Peer quality can affect growth via two main channels. The first is *learning from peers*: direct interaction with or observation of high-quality colleagues may increase productivity and labor market value. The second is *learning by doing*: better peers may reduce an individual's role in meaningful tasks, constraining hands-on experience. If these channels operate in opposite directions, standard peer effect estimates that omit task allocation will conflate the two effects and underestimate the gross learning benefit.

To identify these channels, we expand Equation (2) to include:

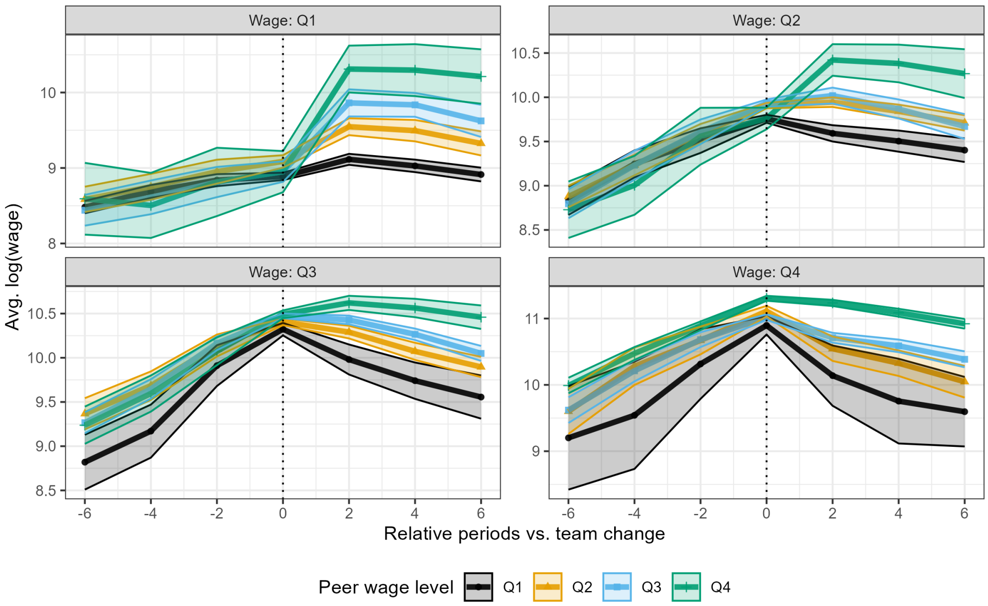
$$\begin{aligned} hc_{i,t+h} =& \alpha hc_{i,t} + \beta \overline{hc}_{-i,t} + \delta(mv_{i,t} - w_{i,t}) \\ &+ \theta_1 \text{Minutes}_{i,t} + \theta_2 \text{PassCentrality}_{i,t} \\ &+ \phi_1 \text{Composition}_{-i,t} + \gamma X_{i,t} \\ &+ \omega_{\text{player chars}} + \omega_{\text{league}\times t} + e_{i,t+h} \end{aligned} \tag{3}$$

where $hc_{i,t+h}$ is any measure of human capital incuding wage or skills, $\text{Minutes}_{i,t}$ denotes a player i's total playing time (standardized), $\text{PassCentrality}_{i,t}$ is the eigenvector centrality in the team's passing network (standardized), and $\text{Composition}_{-i,t}$ includes measures of peer distribution: Hirschman-Herfindahl Index of wages, share of top 5% players, and share of top 5-25% players. Including $\text{Minutes}_{i,t}$ and $\text{PassCentrality}_{i,t}$ allows us to estimate the direct learning effect β conditional on task allocation, separating it from the indirect effect operating through reduced opportunities.

An important threat to our identification strategy is the unobserved growth potential that could possibly relate to both better peer quality and realized human capital growth, introducing selection bias. Let us note that this bias affects not only our setup, but any peer effect estimation without a specific strategy to control for it. We address this threat in three ways. First, we use the post-arrival market value along with wage at the new team as baseline, which should reflect the match quality and the eliminiation of some of the ex ante uncertainty about growth potential. Second, we show that peer effects of their initial teams persist for switchers as well. And third, we run several robustness checks including short-term future value in the right hand side to show that even after potential 'overcontrolling', most of the peer effects remain.

To foreshadow the main results, we can check pre- and post-arrival wage trends conditional on just-after-arrival wage and peer quality quartiles, in the event-study-esque setup of Figure 2. While before arriving to the new team average wages moved reasonably parallel, afterwards they are markedly separated by the new team's peer quality. We repeat the exercise for short passing skill in Figure A5 showing somewhat similar patterns, and for sprinting speed as well in Figure A6 serving as placebo and displaying no separation in its evolution along peer quality.

Figure 2: Wage dynamics conditional on wage level and peer quality quartiles after arrival



Note: The figure shows the pre- and post-arrival wage dynamics with 95% confidence intervals, by quartiles of wage levels and peer quality, for the final analysis sample with sufficient pre-arrival periods.

4.2 Composition: exposure to stars

In this section, we examine whether the *composition* of peers, conditional on average quality, affects human capital development. We measure composition using: (i) the Hirschman-Herfindahl Index (HHI) of team wages, where higher values indicate greater concentration of talent (i.e., presence of stars); and (ii) the share of superstars (top 5% of all players by wage) and high-flyers (top 5-25%) in the lineup, where top status is determined by wage in the current or previous two half-seasons.[11]

Table 2 reports estimates from Equation (3) for $h = 6$ (three years), using the full analysis sample (5,227 observations), and for junior players (age $\leq$ 23, 2,341 observations). Column (1) shows the baseline peer effect: 10% higher average teammate wage is associated with 3.3% higher wage growth. Columns (2)-(3) add composition measures. Conditional on average peer quality, neither the HHI (Column 2), share of top 5% players nor share of top 5-25% players (Column 3) are significantly associated with wage growth. For junior players shown in Columns (4)-(6), the pattern is similar: average peer quality matters (10% better peers is associated with 4.0% wage growth), but composition does not. This result implies that learning depends on average peer quality, not the distribution of talent. Having superstars on the team, conditional on overall quality, does not provide additional learning benefits beyond what average quality already delivers.

[11] With a standard lineup of 23 players, one additional top 5% player increases the share by approximately 4.3 percentage points.

Table 2: Wage evolution and exposure to stars

	log($wage_{t+6}$)					
	Full sample			Junior		
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.440***	0.440***	0.439***	0.398***	0.397***	0.399***
	(0.021)	(0.021)	(0.021)	(0.032)	(0.032)	(0.032)
log(value / wage)	0.241***	0.241***	0.239***	0.218***	0.218***	0.218***
	(0.019)	(0.019)	(0.019)	(0.025)	(0.025)	(0.025)
log(teammates' mean wage)	0.333***	0.333***	0.290***	0.404***	0.405***	0.406***
	(0.033)	(0.033)	(0.054)	(0.047)	(0.048)	(0.083)
Team HHI of wages		0.310			-0.644	
		(0.994)			(1.86)	
Share of top 5% teammates			0.226			-0.099
			(0.180)			(0.293)
Share of top 5-25% teammates			0.099			0.028
			(0.116)			(0.177)
Observations	5,227	5,227	5,227	1,957	1,957	1,957
R^2	0.495	0.495	0.495	0.472	0.472	0.472
Within R^2	0.208	0.208	0.208	0.211	0.211	0.211

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

Using five specific skill attributes to capture human capital, Table 3 examines how composition relates to growth in short passing, reactions (tactical awareness), finishing, interceptions, and sprint speed. These regressions use the same specification as Table 2 but with skill ratings as outcomes. We find that 10% higher average peer quality is associated with 0.12 points higher short passing rating (Column 1) and 0.10 points higher reactions rating (Column 2) – both learnable technical skills. However, peer quality is not significantly associated with interceptions (Column 3), finishing (Column 4), or sprint speed (Column 5). This pattern of effects on learnable but not innate skills provides evidence against pure selection bias: if better teams simply identified 'hidden gems' with higher latent ability, we would expect peer quality to correlate with all skill dimensions, not just those acquirable through training. Importantly, the share of top players is again not significantly associated with *any* of the examined skill attributes, reinforcing that composition, conditional on average quality, does not drive learning.

Table 3: Skill evolution and exposure to stars

	Short pass$_{t+6}$ (1)	Reactions$_{t+6}$ (2)	Interceptions$_{t+6}$ (3)	Finishing$_{t+6}$ (4)	Sprint speed$_{t+6}$ (5)
Own skill	0.537***	0.374***	0.705***	0.779***	0.877***
	(0.019)	(0.017)	(0.016)	(0.015)	(0.015)
Team avg. skill	0.013	0.152***	-0.062	-0.037	-0.054
	(0.040)	(0.047)	(0.040)	(0.040)	(0.041)
log(value / wage)	0.500***	0.632***	0.173	0.434***	-0.049
	(0.123)	(0.118)	(0.167)	(0.148)	(0.151)
log(teammates' mean wage)	1.24***	1.04***	0.499	0.415	-0.056
	(0.287)	(0.304)	(0.418)	(0.381)	(0.414)
Share of top 5% teammates	-0.133	0.065	0.165	1.07	2.37
	(1.04)	(1.03)	(1.32)	(1.42)	(1.53)
Share of top 5-25% teammates	0.876	-0.025	0.816	-0.866	-0.390
	(0.666)	(0.697)	(0.850)	(0.861)	(0.862)
Observations	5,227	5,227	5,227	5,227	5,227
R^2	0.607	0.415	0.870	0.850	0.701
Within R^2	0.392	0.219	0.545	0.596	0.581

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

4.3 Learning by doing vs learning from peers

Having established that star exposure does not drive learning, we turn to the two key channels through which peer quality affects growth: direct learning from peers versus indirect effects through reduced task allocation. In many production environments such as football, the number of meaningful tasks is limited[12]. When better peers are available, individual workers may be assigned fewer instructive tasks, creating substitution between 'learning from peers' and 'learning by doing.' However, the two could be complementary if better peers create more instructive tasks. Depending on which direction dominates, standard peer effect estimates may over- or underestimate true learning effects.

Access to learning from peers is possible in training. However, intensive learning happens under pressure when actually playing. Moreover, training drills are often prepare for plays, and also favor players who spend more time on the pitch. Before we turn to estimates, let us discuss playing time. The coach decides how much players can play. This is related to players quality: players in the top quartile in their team in a given half-season will play 61%[13] more than those in the bottom quartile. Moreover, when looking at time spent together on the pitch as a proxy of collaborative experience, we find that the best people play mostly with other top players. People in the highest quartile of wage distribution will play 818 minutes per half-season with people also from the top

[12] Football's production technology—while specific in its zero-sum playing time—reflects a broader economic principle: workers compete for scarce, instructive tasks. In consulting firms, only select junior associates work closely with star partners on high-profile cases; in research labs, senior scientists control access to key projects and co-authorship. Our setting allows us to quantify this trade-off explicitly.

[13] This is the coefficient of log minutes played regressed on log wages with team x half-season, age and position dummies.

quartile, while only 472 minutes with players in the lowest quartile (see Appendix For details).

So in our case, playing time is unwanted mechanism (of peer quality on learning), we shall partial it out.

Table 4 presents estimates from Equation (3) controlling for task allocation. The baseline peer effect showed that 10% higher peer quality is associated with 3.3% wage growth. When we add controls for minutes played and passing network centrality (Columns 1 and 2), the coefficient on peer quality *increases*, implying that better peers would generate 4.4% wage growth if task allocation were held constant. The difference (4.4% vs. 3.3%) represents the *crowding out* effect: better peers reduce growth by approximately 1.1 percentage points through diminished opportunities. For junior players, this pattern is even stronger. The baseline peer effect was 0.41 log points, rising to 0.54 when controlling for minutes and centrality (Column 4 and 5).

Conditional on average peer quality, the share of top 5% or top 5-25% players on the team is not significantly related to wage growth (Columns 3 and 6). This confirms the result from Section 4.1: exposure to stars per se does not drive learning. Instead, *contribution to production* matters: one standard deviation more minutes played is associated with 22.6% higher wage growth, and one standard deviation higher passing centrality is associated 4.5% higher growth. For juniors, the association with minutes played is similar in magnitude but passing centrality is three times as important (12.6%).

Table 4: Wage evolution and contribution to production

	$\log(\text{wage}_{t+6})$					
		Full sample			Junior	
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.344***	0.343***	0.341***	0.291***	0.287***	0.286***
	(0.021)	(0.021)	(0.021)	(0.030)	(0.030)	(0.030)
log(value / wage)	0.179***	0.179***	0.176***	0.151***	0.153***	0.153***
	(0.018)	(0.018)	(0.018)	(0.023)	(0.023)	(0.023)
log(teammates' mean wage)	0.427***	0.435***	0.383***	0.519***	0.544***	0.529***
	(0.034)	(0.034)	(0.052)	(0.050)	(0.051)	(0.081)
Total minutes (Z)	0.266***	0.229***	0.226***	0.310***	0.203***	0.202***
	(0.015)	(0.024)	(0.024)	(0.023)	(0.032)	(0.032)
Pass centrality (Z)		0.043*	0.045**		0.126***	0.126***
		(0.022)	(0.022)		(0.034)	(0.034)
Minutes shared with top 5%			0.276*			0.074
			(0.145)			(0.230)
Minutes shared with top 5-25%			0.082			0.026
			(0.089)			(0.132)
Observations	5,227	5,227	5,227	1,957	1,957	1,957
R^2	0.530	0.530	0.531	0.515	0.517	0.517
Within R^2	0.263	0.264	0.265	0.275	0.278	0.278

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

Our results suggest that without conditioning on task allocation, standard estimates might underestimate learning from peers by approximately 30% (0.54 vs. 0.41 for younger players, and 0.44 vs. 0.33 overall). To quantify the decomposition of peer effects through these two channels of task allocation, we estimate auxiliary regressions where minutes played and passing centrality are the dependent variables (Appendix Table A12). These regressions reveal that 10% higher peer quality is associated with 0.35 standard deviations fewer minutes played and 0.50 standard deviations lower passing centrality. Hence, the total peer effect on wage growth can be decomposed into a direct learning effect of 4.35% and negative indirect effects through minutes (-0.80%) and centrality (-0.22%), summing to the observed 3.33% net effect. It implies that if we could increase peer quality by 10% while keeping task allocation constant, wage growth would be approximately 30% higher. This decomposition highlights that standard estimates might substantially underestimate the gross learning benefit from peers because they do not account for crowding out of learning-by-doing opportunities.

Table 5 reinforces these findings using skill attribute evolution as outcomes. The table reports estimates from Equation (3) for five skills measured at the 3-year mark. Peer quality is significantly associated with learnable skills: 10% higher peer quality correlate with 0.15 points better short passing (Column 1) and 0.11 points higher reactions (Column 2). In contrast, peer quality has no significant relationship with interceptions (Column 3), finishing (Column 4), or sprint speed (Column 5).

Passing centrality and minutes are significantly associated with short passing and reactions as well (with interceptions also improving along minutes played), but not with finishing and speed. The pattern that effects are concentrated on learnable technical abilities but absent for innate physical traits provides strong evidence that our estimates capture genuine learning rather than selection on unobserved growth potential. If better teams systematically recruited players with higher latent ability across all dimensions, we would expect correlations with all skill types, not just those acquirable through training. Finally, the share of top 5% or top 5-25% players on the team is not significantly associated with any skill attribute, consistent with earlier results: learning depends on average quality and task allocation, not exposure to superstars.

Table 5: Skill evolution and role on the team

	Short pass$_{t+6}$ (1)	Reactions$_{t+6}$ (2)	Interceptions$_{t+6}$ (3)	Finishing$_{t+6}$ (4)	Sprint speed$_{t+6}$ (5)
Own skill	0.504***	0.328***	0.698***	0.775***	0.876***
	(0.020)	(0.016)	(0.017)	(0.015)	(0.015)
Team avg. skill	0.051	0.190***	-0.054	-0.032	-0.054
	(0.039)	(0.047)	(0.040)	(0.038)	(0.042)
log(value / wage)	0.335***	0.389***	0.040	0.336**	-0.112
	(0.120)	(0.114)	(0.169)	(0.146)	(0.150)
log(teammates' mean wage)	1.52***	1.08***	0.755*	0.637	0.039
	(0.283)	(0.302)	(0.413)	(0.385)	(0.394)
Total minutes (Z)	0.318**	1.51***	0.632***	0.449*	0.148
	(0.160)	(0.163)	(0.223)	(0.263)	(0.227)
Minutes shared with top 5%	0.275	0.764	0.221	0.392	1.53
	(0.840)	(0.817)	(1.05)	(1.17)	(1.24)
Minutes shared with top 5-25%	0.516	0.159	0.209	-1.06*	-0.198
	(0.526)	(0.476)	(0.671)	(0.624)	(0.685)
Pass centrality (Z)	1.10***	0.257*	0.339	0.254	0.316
	(0.174)	(0.154)	(0.207)	(0.246)	(0.240)
Observations	5,227	5,227	5,227	5,227	5,227
R^2	0.630	0.461	0.871	0.851	0.702
Within R^2	0.426	0.280	0.550	0.599	0.582

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

4.4 Direct collaboration with top peers

Our granular passing network data allow us to distinguish between broad exposure to talented peers and intensive direct collaboration. Passing represents the most direct form of teamwork in football: players who exchange more passes should practice together, communicate frequently, and have greater opportunity for knowledge transfer. We differentiate collaboration intensity with three peer types: top 5% players (superstars), top 5-25% players (high-flyers), and the remaining players.

For direct collaboration, we further expand the specification to:

$$\begin{aligned} hc_{i,t+h} =& \alpha hc_{i,t} + \beta \overline{hc}_{-i,t} + \delta(mv_{i,t} - w_{i,t}) \\ & + \theta_1 \text{Minutes}_{i,t} + \theta_2 \text{PassCentrality}_{i,t} \\ & + \lambda_1 \text{PassesTop5}_{i,t} + \lambda_2 \text{PassesTop5-25}_{i,t} + \lambda_3 \text{PassesRest}_{i,t} \\ & + \gamma X_{i,t} + \omega_{\text{player chars}} + \omega_{\text{league}\times t} + e_{i,t+h} \end{aligned} \tag{4}$$

where $\text{PassesTop5}_{i,t}$, $\text{PassesTop5-25}_{i,t}$, and $\text{PassesRest}_{i,t}$ measure the standardized number of passes exchanged with top 5%, top 5-25%, and remaining teammates, respectively, in terms of wage distribution conditional on position and time. This specification identifies whether learning occurs through intensive collaboration with specific high-quality peers rather than mere exposure. Note that Equation (4) is estimated on a restricted sample containing only teams with all three player types in their lineup, reducing the sample to approximately 1,700 observations (one-third of the main sample). We estimate Equations (4) with the way as the baseline specification (3).

Equation (4) includes standardized measures of pass counts with each peer type. This analysis requires teams to have all three player types in their lineup, which reduces the sample to approximately one-third of the main sample (1,644 observations). This restriction ensures we can separately identify collaboration effects with each peer type without confounding by team composition.

Table 6 presents results for wage growth in three years. Column (1) replicates the baseline specification on this restricted sample. Notably, for this subsample of players on elite teams with superstars, the peer quality effect is similar: 10% higher average peer quality is still associated with 4.4% wage growth, conditional on task allocation. Interestingly, passing centrality is not significantly associated with wage growth in this subsample, suggesting that on star-studded teams, overall centrality matters less than specific collaboration patterns.

Column (2) adds direct collaboration to peer quality. Conditional on peer quality, one standard deviation more passes with top 5% players is associated with around 10% higher wage growth, and we estimate almost the same for top 5-25% players. The peer quality coefficient remains stable ($\beta = 0.44$), indicating that direct collaboration effects are distinct from average quality effects.

In Column (3), we include minutes played and passing centrality as well. Strikingly, once we control for total playing time, the direct collaboration coefficients become statistically insignificant, and the minutes coefficient is large and significant. This suggests that on elite teams, minutes played captures most variation in learning opportunities: players who get more playing time naturally interact more with top teammates, and this total exposure drives learning. The *distribution* of passes across peer types, conditional on total minutes, provides limited additional information.

Columns (4)-(6) repeat the analysis for junior players (653 observations in this restricted sample). The pattern is similar, with one notable difference: for juniors, collaboration with top 5-25% players (high-flyers) is more important than collaboration with top 5% superstars. In Column (5), passing with top 5-25% players has a coefficient of 0.14, larger than the insignificant coefficient on passing with top 5% players. This suggests that for younger players, learning may be more effective from slightly less elite but still very high-quality peers.

Table 6: Wage evolution and direct collaboration with top peers

	log($wage_{t+6}$)					
	Full sample			Junior		
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.285***	0.313***	0.283***	0.225***	0.255***	0.225***
	(0.035)	(0.035)	(0.035)	(0.044)	(0.045)	(0.044)
log(value / wage)	0.202***	0.239***	0.201***	0.176***	0.216***	0.177***
	(0.031)	(0.032)	(0.031)	(0.040)	(0.041)	(0.041)
log(teammates' mean wage)	0.444***	0.437***	0.430***	0.495***	0.465***	0.506***
	(0.065)	(0.071)	(0.067)	(0.118)	(0.116)	(0.121)
Pass centrality (Z)	-0.041		-0.044	0.057		0.043
	(0.033)		(0.039)	(0.062)		(0.064)
Total minutes (Z)	0.280***		0.267***	0.245***		0.239***
	(0.034)		(0.035)	(0.062)		(0.064)
Pass count with top 5% (Z)		0.098***	0.039		0.053	-0.019
		(0.033)	(0.030)		(0.067)	(0.057)
Pass count with top 5-25% (Z)		0.113***	0.012		0.144**	0.011
		(0.034)	(0.040)		(0.068)	(0.056)
Pass count with top 25-100% (Z)		0.010	-0.025		0.103*	0.047
		(0.033)	(0.032)		(0.054)	(0.050)
Observations	1,644	1,644	1,644	653	653	653
R^2	0.471	0.449	0.472	0.427	0.406	0.428
Within R^2	0.239	0.208	0.241	0.233	0.204	0.234

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

4.5 Skill development

To corroborate the results on wage evolution, Table 7 examines skill attribute development using the same sample and specification. Columns (1)-(5) report results for the five skill attributes at $h = 6$. Direct collaboration with top 5% players is significantly associated with short passing improvement (Column 1) and reactions improvement (Column 2), but not with finishing, interceptions, or sprint speed (Columns 3-5).

This pattern reinforces two key conclusions. First, learning occurs through skills directly related to passing, the mode of collaboration we measure. Second, the absence of effects on innate physical traits (sprint speed) or less directly related skills (finishing) provides further evidence against selection bias. If elite teams simply recruited players with better latent growth potential, we would expect correlations across all skill dimensions, not just those directly tied to measured collaboration.

Table 7: Skill evolution and direct collaboration with top peers

	Short pass$_{t+6}$ (1)	Reactions$_{t+6}$ (2)	Interceptions$_{t+6}$ (3)	Finishing$_{t+6}$ (4)	Sprint speed$_{t+6}$ (5)
Own skill	0.536***	0.351***	0.737***	0.804***	0.810***
	(0.040)	(0.023)	(0.032)	(0.024)	(0.025)
Team avg. skill	0.031	0.241**	-0.003	0.070	-0.031
	(0.067)	(0.093)	(0.061)	(0.080)	(0.094)
log(value / wage)	0.314	0.659***	-0.278	0.134	-0.018
	(0.228)	(0.193)	(0.268)	(0.272)	(0.275)
log(teammates' mean wage)	1.01**	0.226	0.981*	0.435	0.633
	(0.472)	(0.767)	(0.589)	(0.556)	(0.661)
Pass count with top 5% (Z)	0.598***	0.772***	0.395*	0.194	0.013
	(0.210)	(0.217)	(0.235)	(0.343)	(0.290)
Pass count with top 5-25% (Z)	0.631***	0.728***	0.255	0.627*	0.608*
	(0.205)	(0.223)	(0.266)	(0.355)	(0.313)
Pass count with top 25-100% (Z)	0.164	0.257	0.185	0.207	-0.020
	(0.157)	(0.181)	(0.254)	(0.298)	(0.307)
Observations	1,644	1,644	1,644	1,644	1,644
R^2	0.628	0.422	0.893	0.864	0.698
Within R^2	0.414	0.247	0.608	0.628	0.515

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

4.6 Persistence of peer effects: human capital growth of switchers

A critical question is whether the peer effects we document reflect genuine human capital accumulation or temporary production complementarities specific to particular teams. If learning were merely complementarity—e.g., playing style matches or chemistry that enhances joint production, we would expect peer effects to vanish when players switch teams. In contrast, if workers accumulate transferable skills, initial peer quality should predict growth even after job changes.

We test this by comparing 'stayers' (players who remain with their initial team through $t + 6$) to 'switchers' (players who leave before $t + 6$). Switchers provide an additional test: they did not match well enough to remain long-term, yet if peer learning is genuine, they should still benefit from early exposure to high-quality teammates at their subsequent employers.

Table 8 reports estimates from Equation (3) separately for stayers and switchers. Column (1) reports the estimates for all players at horizons $h = 6$ (three years), the previous coefficient of $\beta = 0.44$ on peer quality. Column (2) shows the estimates only for those who remained for three years, and Columns (3-4) for those who left earlier. The peer effect is strongest for those who stayed all three years ($\beta = 0.50$), but even for switchers the relationship is strong and statistically significant ($\beta = 0.41$ and $\beta = 0.44$ for those leaving after one and two years, respectively). This is expected: players who can remain with high-quality teams continue benefiting from those peers over multiple periods. Still, learning from initial peers transfers to new teams: players carry skills acquired early in their tenure to subsequent employers.

Table 8: Wage evolution of stayers vs. switchers after one and two years

	$\log(\text{wage}_{t+6})$			
	Full sample	Stayed for 3 years	Stayed for 2 years	Stayed for 1 year
	(1)	(2)	(3)	(4)
log(wage)	0.343***	0.309***	0.307***	0.378***
	(0.021)	(0.028)	(0.047)	(0.032)
log(value / wage)	0.179***	0.149***	0.177***	0.207***
	(0.018)	(0.022)	(0.039)	(0.032)
log(teammates' mean wage)	0.435***	0.499***	0.438***	0.414***
	(0.034)	(0.047)	(0.066)	(0.048)
Total minutes (Z)	0.229***	0.146***	0.196***	0.313***
	(0.024)	(0.032)	(0.050)	(0.048)
Pass centrality (Z)	0.043*	-0.005	0.057	0.065
	(0.022)	(0.033)	(0.047)	(0.051)
Observations	5,227	2,081	1,404	1,742
R^2	0.530	0.629	0.521	0.487
Within R^2	0.264	0.274	0.224	0.293

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

To investigate further, Table 9 presents analogous results for short passing skill evolution (full results for all five skills in Appendix Table A14). The pattern mirrors wage results: peer quality at the initial team significantly predicts short passing skill growth for both stayers (Column 2: $\beta = 1.58$) and switchers (Columns 3-4: $\beta = 1.77$ and $\beta = 1.43$ respectively for those who stayed one and two years), with coefficients not statistically different from each other. This skill-level evidence further confirms that learning is genuine and persistent: players who worked with better peers in their first period show lasting improvement in specific learnable skills, even after changing teams.

Table 9: Short-pass evolution of stayers vs. switchers after one and two years

	Short pass_{t+6}			
	Full sample	Stayed for 3 years	Stayed for 2 years	Stayed for 1 year
	(1)	(2)	(3)	(4)
Short pass	0.504***	0.527***	0.491***	0.479***
	(0.020)	(0.035)	(0.026)	(0.027)
Teammates' avg. short pass	0.053	0.089	-0.019	0.082
	(0.039)	(0.061)	(0.077)	(0.060)
log(value / wage)	0.340***	0.157	0.348*	0.531***
	(0.119)	(0.202)	(0.178)	(0.147)
log(teammates' mean wage)	1.67***	1.58***	1.43***	1.77***
	(0.211)	(0.408)	(0.417)	(0.338)
Total minutes (Z)	0.324**	0.112	0.119	0.537**
	(0.160)	(0.240)	(0.316)	(0.262)
Pass centrality (Z)	1.09***	0.878***	1.34***	1.15***
	(0.173)	(0.255)	(0.297)	(0.247)
Observations	5,227	2,081	1,404	1,742
R^2	0.630	0.651	0.638	0.632
Within R^2	0.426	0.427	0.422	0.425

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

The persistence of peer effects for switchers provides three additional insights. First, it corroborates genuine skill accumulation rather than temporary complementarity. Second, it mitigates selection concerns: if better teams merely identified 'hidden gems' whose latent ability was revealed ex-post, we would not expect *initial* team quality to predict growth for players whose matches failed (i.e., switchers). Third, it has policy implications: organizations should invest in junior workers' exposure to high-quality mentors early, even if those workers eventually leave, because accumulated skills persist across employers.

5 Robustness

The main identification threat is that better teams might be more successful at uncovering 'hidden gems': players with great unobserved growth potential. They could be detected only by teams with more advanced scouting resources – which might be the teams with better peer quality. While this is a common issue in peer effect studies, our setup is less vulnerable as we observe the perceived market value incorporating growth potential. As we use market value measured just after the player arrives to a new team, the growth potential might be appropriately captured already. Furthermore, we have demonstrated the persistence of peer effects for switchers, and how they relate only to learnable skills.

Still, we perform five additional robustness checks to bound the problem and corroborate our results. First, we compare peer effects for players from countries with a top league vs. the rest of the world. If selection bias is related strongly to differential scouting success, we expect heterogeneity in this regard. Second, we study the differences by playing positions: defenders, midfielders, and forwards require different sets of abilities to succeed, of which some can be easier to scout than others. Third, we look at somewhat 'overcontrolled' regressions with one-halfseason forward-looking future value growth (capturing any sudden previously unrealized gains in market value) and team success as controls, to show that even with those variables in the regression potentially inducing downward bias in the estimates, the core results are qualitatively similar. Fourth, we reestimate the main regressions for elite and non-elite teams. Fifth, we split the sample by player wage tertiles to see whether the relationship crucially depends on some parts of the distribution.

The estimates can be found in the Appendix, corroborating robustness along these dimensions. We find no large differences between the top 7 countries and the rest of the world (Table A15). The regressions by positions (Tables A16, A17) show that the point estimates for the three non-goalkeeper positions are close to the pooled sample, however, for goalkeepers we do not find any significant partial correlation between wage growth and peer quality. If there was a strong selection along unobserved growth potential, we expect that it would affect goalkeepers as well, strengthening the validity of our results. The regressions with team success and one-period forward-looking future market value change (Table A18) suggest that peer effects are bounded by 0.3 from below, the remaining relationship is still large and statistically significant. Peer effects are similar and strong both within elite and non-elite teams (Table A19), and along the player's wage distribution as well (Table A20).

One additional concern is the role of the manager quality. It is a time-varying team-specific factor that could be correlated with peer quality as well as player development. For instance, better managers are more effective at attracting and retaining higher-quality players, directly increasing the average peer quality. They may also be better at improving player skills. To mitigate this confounding effect, we created a manager quality metric, the point average achived over two half-seasons: 3 and 4 half-seasons before the current one. We see no change in results as seen in Table A21. Please note that the variable of elite teams also captures management talent.

6 Conclusions

This paper investigates the mechanisms through which workers learn from peers in collaborative environments, using elite male European football as a setting with exceptional measurement advantages. We observe actual collaboration, frequently-updated market values that bypass wage rigidity, and detailed skill attributes —- features rarely available in traditional administrative data.

Our analysis reveals five key findings about workplace learning. We document substantial peer quality effects on wage growth, but standard estimates underestimate the true learning effect because better peers reduce individual playing time, creating a trade-off between learning from peers and learning by doing that dampens estimated peer effects. While wages may suffer from measurement limitations due to rigidity and infrequent updates, market values that reflect expert consensus and update frequently help address these concerns. Learning requires more than proximity to talent: conditional on average quality, learning is driven by active engagement through direct collaboration, with interaction intensity rather than mere exposure determining learning gains. Peer effects manifest primarily in learnable skills like passing accuracy and tactical positioning rather than innate physical attributes like speed and strength, providing evidence against pure selection stories. Finally, learning persists after team switches, confirming genuine human capital accumulation rather than temporary complementarities specific to particular teams.

The paper used very granular data from a specific industry and setting. How does it generalize to other settings? While our setting is special, it plausibly represents high-skilled, competitive, and at the same time, collaborative professional job environments, evidenced by the similarities to the top decile earning group of Jarosch et al. (2021). Similarly to other settings, workers observe each others value quite well and able to find colleagues to learn from. Critically, we can observe a great deal in terms of values and the mechanism of learning. One critical limitation could be the zero sum nature of football – in terms of playing time, one player's gain is, by definition, another's loss. The exact nature is indeed specific to football, but the presence of scarce resources is not. The economic channel of workers competing for scarce, high-value, instructive tasks is a common feature in many high-skilled, collaborative jobs. For example, in a law firm or consultancy, only few junior associates are given primary access to the star partner or able to work on a high-profile case.

Which aspects of our findings should generalize more confidently? The importance of direct collaboration over mere proximity appears robust—learning happens through working together, not just being near talented colleagues. The persistence of learning after job changes suggests genuine human capital accumulation operates broadly. The distinction between learnable and fixed skills may matter less in knowledge work, but the principle that peer effects require realistic learning opportunities should hold.

For organizations designing training programs or team composition, our results suggest three practical insights. First, simply hiring star performers may not maximize learning if junior workers lack opportunities to engage actively. Second, policies that send workers to learn from the best early in careers can have lasting benefits, even if those workers eventually move to different roles or organizations. Third, the intensity of collaboration matters more than the breadth of exposure—organizations should structure work to facilitate direct interaction between senior and junior staff, not merely co-location.

Via documenting these mechanisms with granular production data, we contribute to understanding how workplace learning actually occurs in collaborative settings, providing evidence that complements and extends findings from administrative data studies where such detailed observation remains infeasible.

References

Arcidiacono, P., Kinsler, J., and Price, J. (2017). Productivity Spillovers in Team Production: Evidence from Professional Basketball. *Journal of Labor Economics*, 35(1):191–225.

Arellano-Bover, J. and Saltiel, F. (2023). Differences in on-the-job learning across firms. Technical report, CAGE working paper no. 670.

Azoulay, P., Fons-Rosen, C., and Graff Zivin, J. S. (2019). Does science advance one funeral at a time? *American Economic Review*, 109(8):2889–2920.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J. (2019). Who Becomes an Inventor in America? the Importance of Exposure to Innovation. *The Quarterly Journal of Economics*, 134(2):647–713.

Bilen, E. and Matros, A. (2023). The queen's gambit: Explaining the superstar effect using evidence from chess. *Journal of Economic Behavior & Organization*, 215:307–324.

Békés, G. and Ottaviano, G. I. (2025). Cultural homophily and collaboration in superstar teams. *Management Science*, 71(10):8149–8168.

Cohen-Zada, D., Dayag, I., and Gershoni, N. (2024). Effort peer effects in team production: Evidence from professional football. *Management Science*, 70(4):2355–2381.

Cornelissen, T., Dustmann, C., and Schönberg, U. (2017). Peer Effects in the Workplace. *American Economic Review*, 107(2):425–456.

Fonti, F., Ross, J.-M., and Aversa, P. (2023). Using sports data to advance management research: A review and a guide for future studies. *Journal of Management*, 49(1):325–362.

FourFourTwo (2023). fifa 22 player ratings explained. Technical report.

Fraiberger, S. P., Sinatra, R., Resch, M., Riedl, C., and Barabási, A.-L. (2018). Quantifying reputation and success in art. *Science*, 362(6416):825–829.

Gathmann, C. and Schönberg, U. (2010). How general is human capital? a task-based approach. *Journal of Labor Economics*, 28(1):1–49.

Gibbons, R. and Waldman, M. (1999). A theory of wage and promotion dynamics inside firms*. *The Quarterly Journal of Economics*, 114(4):1321–1358.

Gibbons, R. and Waldman, M. (2006). Enriching a theory of wage and promotion dynamics inside firms. *Journal of Labor Economics*, 24(1):59–107.

Herkenhoff, K., Lise, J., Menzio, G., and Phillips, G. M. (2024). Production and learning in teams. *Econometrica*, 92(2):467–504.

Hoegele, D., Schmidt, S. L., and Torgler, B. (2014). Superstars as Drivers of Organizational Identification: Empirical Findings from Professional Soccer. *Psychology & Marketing*, 31(9):736–757.

Hoey, S. (2023). One's pain is another's gain - early career exposure and later labor market outcomes. Technical report, University of Liverpool.

Jarosch, G., Oberfield, E., and Rossi-Hansberg, E. (2021). Learning From Coworkers. *Econometrica*, 89(2):647–676.

Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. *Journal of Political Economy*, 117(5):914–940.

Li, W., Aste, T., Caccioli, F., and Livan, G. (2019). Early coauthorship with top scientists predicts success in academic careers. *Nature Communications*, 10(1):5170.

Mas, A. and Moretti, E. (2009). Peers at Work. *American Economic Review*, 99(1):112–145.

Nix, E. (2020). Learning spillovers in the firm. Technical report, USC.

Pastorino, E. (2024). Careers in firms: The role of learning about ability and human capital acquisition. *Journal of Political Economy*, 132(6):1994–2073.

Poli, R., Besson, R., and Ravenel, L. (2021). Econometric Approach to Assessing the Transfer Fees and Values of Professional Football Players. *Economies*, 10(1):4.

Prockl, F. and Frick, B. (2018). Information precision in online communities: Player valuations on www.transfermarkt.de. *International Journal of Sport Finance*, 13(4):319–335.

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in nazi germany. *The Review of Economic Studies*, 79(2):838–861.

Werner, K. and Dickson, G. (2018). Coworker knowledge sharing and peer learning among elite footballers: Insights from german bundesliga players. *Sport Management Review*, 21(5):596–611.

Winkler, W. E. (1990). String comparator metrics and enhanced decision rules in the Fellegi-Sunter model of record linkage. In *Proceedings of the Section on Survey Research Methods*, pages 354–359, Alexandria, VA. American Statistical Association.

Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics*, 30(1):1–53.

A Appendix

A.1 Validity of key metrics

Before we go into the details on variables, let us discuss the validity of our key football specific metrics: the transfer value by Transfermarkt and the Fifa skills measures.

Transfer value

There is no clear measure of human capital, we do not know what a person is capable of at a given point in time. The best way is to collect as much observation as possible and make an informed prediction. The TM transfer value is not an objective measure built on test score, but a socially constructed perception of human capital. In order words, it is an average of many people's mental model of linking observed charactersitics to a value of human capital.

Despite the wisdom of crowd argument, there may be evaulation biases but the sports economics literature considers it a very good measure (Prockl and Frick, 2018).

We use the post-arrival TM value to control for a player's growth potential, thereby partialling out the causal effect of peers. The assumption is that this control variable cleanly absorbs any selection bias related to 'hidden gems'. One potential critique is some anticipation effect of the impact of joining the club. If a manager (coach) is well known to improve young players, their value would rise already 'pricing in' value increase. It is actually hard to predict this, many young talented players end up on the fringes or sold. It is hard to sign the bias (if any), but it may be present in many working places, and in our setting we get closer to causality by including this variable.

The EA Sports FIFA skills

For the EA Sports Fifa games data, the values are carefully measured by the game's makers based on very detailed statistics as explained by FourFourTwo (2023). As noted in the paper, a key benefit is that it measures a great deal of skills including more innate skills that only need practice as well as more learnable ones, where peers matter.

A.2 Variable descriptions

Our key datasets are the following:

- Player careers: players and team in every half-season with relevant demographics
- Player valuations: TM market valuation for a player in every half-season. (Average when multiple values are available.)
- Player wages and skill ratings: EA Sports' FIFA games include wage information and in addition the scores of rating points of players describing offensive and defensive skill levels such as short passes, tackles, sprinting speed etc.

- Players and squads: information of squads for every match for each team in every half-season (list of players who were ever listed in the lineup as starter or substitute.), providing time spent together, injuries etc.
- Events: number of passes between any two players

Let us define the key variables of the analysis.

- **Market value** (TransferMarkt): mean market value of the player in the half-season, without evaluation the last valid evaluation is assigned to the player
- **Teammates' mean wage**: the mean of the wages of the teammates' of the player based on the data registered in EA Sport's FIFA in the half-season, measured in log-units
- **Teammates' mean market value**: the mean of the market value of the teammates' of the player based on their mean market value in the half-season, measured in log-units
- **Total minutes**: total minutes spent on the pitch in the half-season by the player, measured in Z-scores (standardized with mean and standard deviation)
- **Pass eigenvector centrality**: based on the total pass network of the team in the half-season, calculates the eigenvector centrality of the player (proxies the importance of the player in the team's overall play), measured in Z-scores (standardized with mean and standard deviation)
- **Top 5% of players**: within the position, the player is or in the last 2 years was in the top 5% of the wage distribution in the top 7 leagues
- **Top 25% of players**: within the position, the player is or in the last 2 years was in the top 25% of the wage distribution in the top 7 leagues
- **Seniors vs. juniors**: players are split by age at 23 years of age
- **Pass count with top 5% or 25% of players** (numeric): number of passes with top 5% or 25% players in the half-season, measured in Z-scores (standardized with mean and standard deviation)
- **Minutes shared with top 5% or 25% of players out of total** (numeric): fraction of minutes shared with top 5% or 25% of players out of total shared minutes with all players
- **Hirschman–Herfindahl Index**: HHI index of the wages of the squad members in the half-season
- **Share of top 5% or 25% players on team**: share in terms of number of individuals based on the wage distribution

Fixed-effects in the regressions:

- **League $\times$ half-seasons**: league indicators interacted with the half-seasons
- **Position**: broad position of the player such as goalkeeper, defender, midfielder, forward
- **Player age**: in years
- **Elite team**: manually identified a select group of elite teams that are the traditional top teams in their respective leagues and regularly participate in international tournaments. These teams are the following: Real Madrid, Barcelona, Atletico Madrid, Manchester City, Manchester United, Liverpool, Chelsea, Tottenham, Arsenal, Bayern München, Dortmund, Schalke, Juventus, AC Milan, Inter Milan, Paris Saint-Germain, Olympique Lyonnais. The selection was done on the basis of historic performance, close to Deloitte's top 20 (an annual report).
- **Injury**: the player suffering an injury, identified by not listed as available on the squad
- **Spell count**: the count of teams the player has played on
- **Contract maturity**: years left on the contract

Skill attribute rating variables: we select from a set of highly correlated variables one that is representative of the overall skill group.

- **Short pass**: short pass ability, proxies passing ability
- **Finishing**: scoring ability, proxies shooting
- **Interceptions**: defensive attribute, ability to intercept a ball
- **Reactions**: general attribute, how fast a player responds to a situation around him
- **Sprinting speed**: how fast a player can sprint

A.3 Descriptive statistics

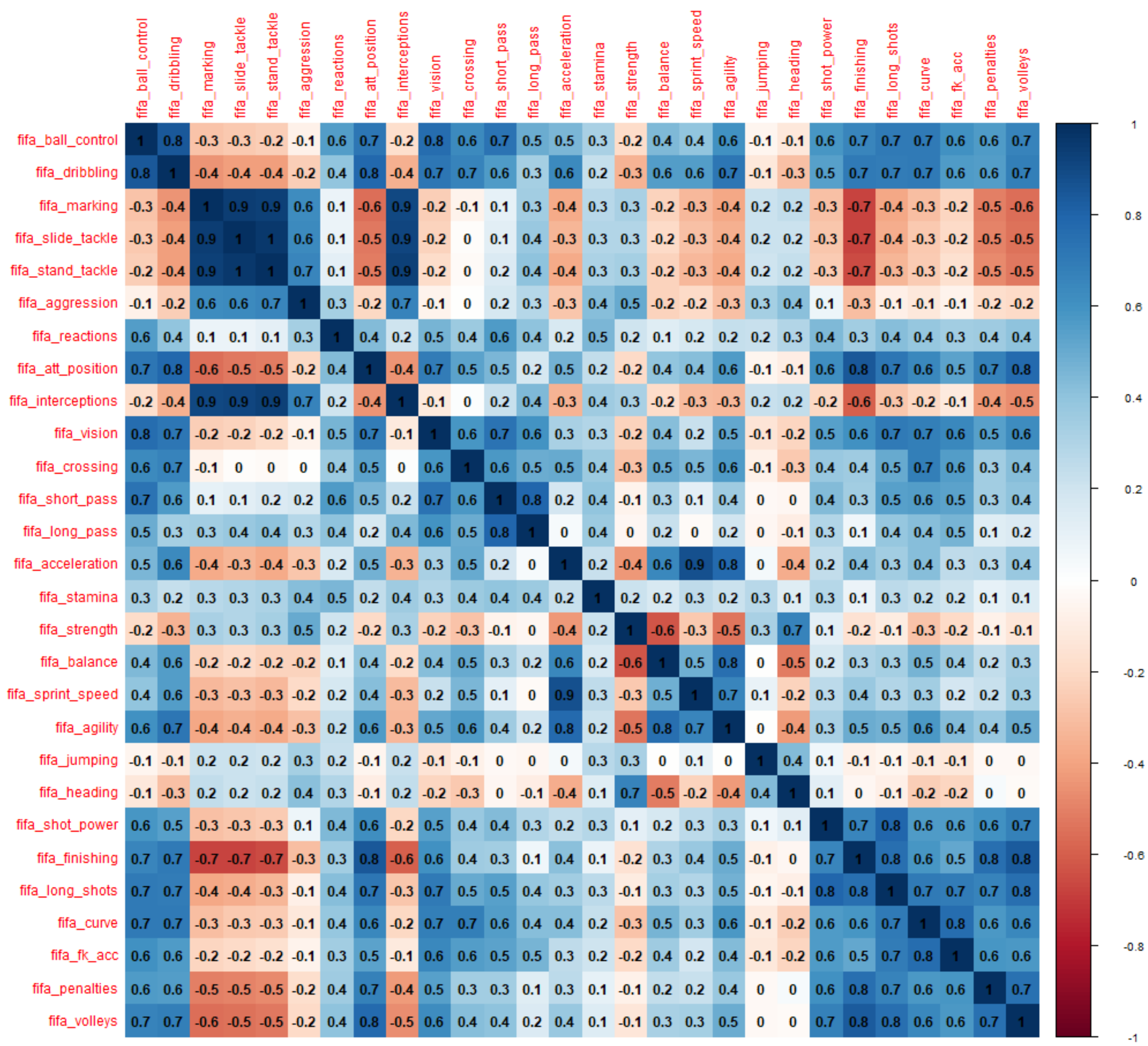
Table A10: Desriptive statistics of the key numeric variables

	N	Mean	SD	Min	P1	P50	P99	Max
Yearly wage (in mn EUR, EA-FIFA)	5227	1.43	1.63	0.08	0.08	0.96	8.06	20.80
log(yearly wage)	5227	13.61	1.12	11.26	11.26	13.78	15.90	16.85
TM market value (in mn EUR)	5227	4.46	7.48	0.03	0.10	2.00	35.00	150.00
log(TM market value)	5227	14.49	1.32	10.13	11.51	14.51	17.37	18.83
log(market value/wage)	5227	4.83	0.80	1.39	2.88	4.83	6.91	8.70
Short pass	5227	67.92	8.62	23.00	44.00	68.00	85.00	91.00
Reactions	5227	67.82	7.98	35.00	48.00	68.00	84.00	96.00
Interceptions	5227	53.16	20.11	10.00	14.00	60.00	84.00	91.00
Finishing	5227	54.27	17.39	11.00	19.00	58.00	84.00	94.00
Speed	5227	71.75	10.25	31.00	42.00	73.00	92.00	96.00
Teammates' mean wage (in mn EUR)	5227	1.55	1.43	0.10	0.14	1.18	7.09	10.98
log(teammates' mean wage)	5227	9.98	0.82	7.60	7.92	10.03	11.82	12.26
Teammates' mean value (in mn EUR)	5227	4.48	5.72	0.24	0.33	2.42	31.49	52.55
log(teammates' mean value)	5227	14.80	0.99	12.38	12.71	14.70	17.27	17.78
HHI of team (wage)	5227	0.06	0.01	0.04	0.04	0.05	0.11	0.27
Share of top 5% players in team (wage)	5227	0.07	0.17	0.00	0.00	0.00	0.78	0.96
Share of top 5-25% players in team (wage)	5227	0.27	0.22	0.00	0.00	0.23	0.79	0.95
Total minutes	5227	1076.55	649.03	1.00	13.00	1054.00	2465.96	3520.00
Pass eigenvector centrality	5227	0.44	0.29	0.00	0.00	0.43	1.00	1.00
Pass count top 5% (wage)	1718	43.59	48.00	0.00	0.00	29.37	225.04	425.00
Pass count top 5-25% (wage)	4516	38.16	45.00	0.00	0.00	26.20	199.73	1298.00
Pass count top 25-100% (wage)	5162	25.60	24.10	0.00	0.00	19.57	109.79	232.00
Minutes shared with top 5% (wage)	5227	0.09	0.21	0.00	0.00	0.00	0.91	1.00
Minutes shared with top 5-25% (wage)	5227	0.31	0.26	0.00	0.00	0.27	0.88	1.00
Player age	5227	23.97	3.52	18.00	18.00	24.00	32.00	37.00
Elite team	5227	0.10	0.31	0.00	0.00	0.00	1.00	1.00
Injury	5227	0.06	0.24	0.00	0.00	0.00	1.00	1.00
Spell rank	5227	1.78	1.00	1.00	1.00	1.00	5.00	8.00
Contract maturity	5227	3.06	1.32	0.00	1.00	3.00	6.00	7.00

Note: All players, all teams. Selected variables.

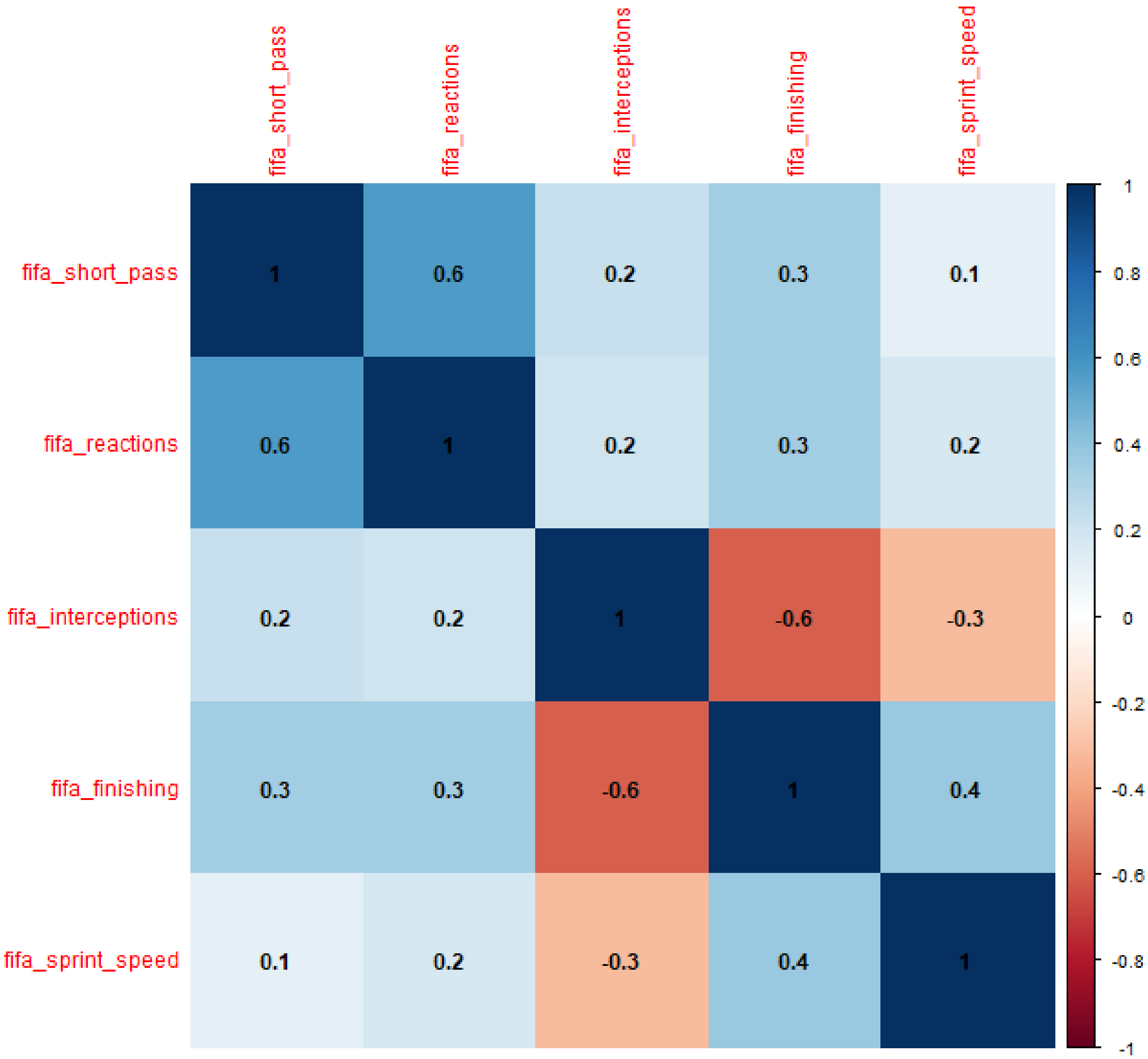
The correlation matrix of the FIFA skill ratings in the sample is the following:

Figure A3: Correlation between EA FIFA skill ratings



Note: The figure shows the correlation structure of FIFA attributes for the final sample.

Figure A4: Correlation between the selected EA Sport's FIFA skill ratings



Note: The figure shows the Pearson correlations of the selected FIFA attributes for the final sample.

A.4 Additional results

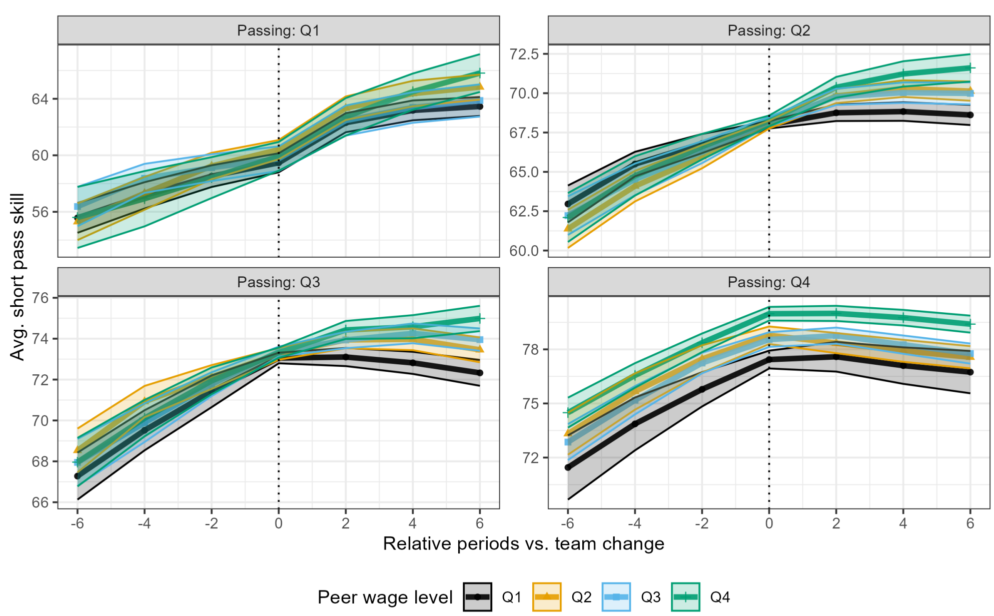
A.4.1 Fixed effects

Table A11: Fixed effects and explained variation in peer quality

	log(teammates' mean wage)			
	(1)	(2)	(3)	(4)
Observations	5,227	5,227	5,227	5,227
R^2	0.500	0.719	0.909	0.946
league X half-season fixed effects	✓	✓	✓	✓
position fixed effects		✓		
player age fixed effects		✓		
elite team fixed effects		✓		
spell count fixed effects		✓		
own injury fixed effects		✓		
contract maturity fixed effects		✓		
managername fixed effects			✓	
teamid_tm fixed effects				✓

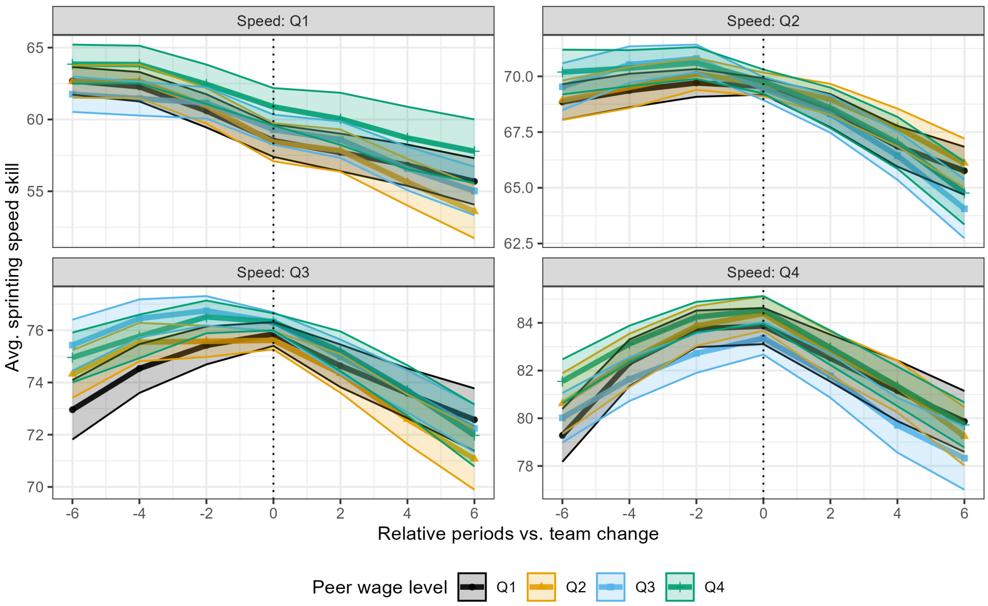
A.4.2 Event study designs

Figure A5: Short pass skill dynamics conditional on skill level and peer quality quartiles after arrival



Note: The figure shows the pre- and post-arrival passing dynamics with 95% confidence intervals, by quartiles of wage levels and peer quality, for the final analysis sample with sufficient pre-arrival periods.

Figure A6: Sprinting speed skill dynamics conditional on skill level and peer quality quartiles after arrival



Note: The figure shows the pre- and post-arrival sprinting speed dynamics with 95% confidence intervals, by quartiles of wage levels and peer quality, for the final analysis sample with sufficient pre-arrival periods.

A.4.3 Contribution to production

Table A12: Minutes played and centrality in the passing network vs. peer effects in wages

	$\log(\text{wage}_{t+6})$ (1)	(2)	Total minutes (Z) (3)	Pass centrality (Z) (4)
log(wage)	0.440***	0.343***	0.361***	0.333***
	(0.021)	(0.021)	(0.021)	(0.023)
log(value / wage)	0.241***	0.179***	0.235***	0.188***
	(0.019)	(0.018)	(0.017)	(0.016)
log(teammates' mean wage)	0.333***	0.435***	-0.353***	-0.495***
	(0.033)	(0.034)	(0.031)	(0.032)
Total minutes (Z)		0.229***		
		(0.024)		
Pass centrality (Z)		0.043*		
		(0.022)		
Observations	5,227	5,227	5,227	5,227
R^2	0.495	0.530	0.315	0.296
Within R^2	0.208	0.264	0.090	0.083

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

Decomposition of the total difference into the positive effect of learning and the negative of less opportunities:

$$\frac{\Delta\mathbb{E}[W_{t+6}|PQ,X]}{\Delta PQ} = 0.333 = \underbrace{\beta}_{0.435} + \underbrace{\gamma}_{0.229} \underbrace{\frac{\Delta\mathbb{E}[M|PQ,X]}{\Delta PQ}}_{-0.353} + \underbrace{\delta}_{0.043} \underbrace{\frac{\Delta\mathbb{E}[C|PQ,X]}{\Delta PQ}}_{-0.495}, \tag{5}$$

where W denotes wage, PQ denotes peer quality, X denotes control variables, M denotes minutes, and C denotes pass centrality.

A.4.4 Direct collaboration with top peers

Table A13: Short passing skill evolution, and direct collaboration with top peers by age

	Short pass$_{t+6}$					
	Full sample			Junior		
	(1)	(2)	(3)	(4)	(5)	(6)
Short pass	0.530***	0.536***	0.528***	0.320***	0.334***	0.320***
	(0.042)	(0.040)	(0.041)	(0.030)	(0.030)	(0.030)
log(value / wage)	0.178	0.314	0.181	0.669***	0.800***	0.675***
	(0.224)	(0.228)	(0.224)	(0.208)	(0.214)	(0.211)
log(teammates' mean wage)	0.937**	1.01**	0.896*	2.32***	2.14***	2.32***
	(0.426)	(0.472)	(0.454)	(0.669)	(0.699)	(0.683)
Teammates' avg. short pass	0.072	0.031	0.065	0.119	0.083	0.117
	(0.067)	(0.067)	(0.068)	(0.096)	(0.097)	(0.096)
Pass centrality (Z)	0.653**		0.568**	1.19**		1.09**
	(0.273)		(0.269)	(0.466)		(0.475)
Total minutes (Z)	0.742***		0.597**	0.504		0.396
	(0.243)		(0.260)	(0.444)		(0.475)
Pass count with top 5% (Z)		0.598***	0.322*		0.451	0.093
		(0.210)	(0.183)		(0.392)	(0.333)
Pass count with top 5-25% (Z)		0.631***	0.140		0.822**	0.195
		(0.205)	(0.229)		(0.394)	(0.391)
Pass count with top 25-100% (Z)		0.164	-0.074		0.400	0.052
		(0.157)	(0.154)		(0.318)	(0.324)
Observations	1,644	1,644	1,644	653	653	653
R^2	0.635	0.628	0.636	0.601	0.590	0.601
Within R^2	0.425	0.414	0.427	0.329	0.310	0.329

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

A.4.5 Persistence

Table A14: Skill evolution of stayers vs. switchers after one and two years

	Short pass$_{t+6}$ (1)	Reactions$_{t+6}$ (2)	Interceptions$_{t+6}$ (3)	Finishing$_{t+6}$ (4)	Sprint speed$_{t+6}$ (5)
Own skill	0.503***	0.330***	0.697***	0.777***	0.876***
	(0.020)	(0.016)	(0.017)	(0.015)	(0.015)
Team avg. skill	0.050	0.190***	-0.054	-0.029	-0.056
	(0.039)	(0.043)	(0.040)	(0.038)	(0.041)
log(value / wage)	0.331***	0.376***	0.025	0.322**	-0.111
	(0.118)	(0.112)	(0.168)	(0.145)	(0.151)
log(teammates' mean wage)	1.72***	1.53***	0.968***	0.344	0.362
	(0.254)	(0.313)	(0.326)	(0.296)	(0.314)
Switch in year 1	-1.02	2.78	-0.117	-3.64	2.55
	(1.86)	(2.02)	(3.51)	(2.67)	(3.02)
Switch in year 2	2.30	7.69***	5.27*	3.67	3.23
	(2.00)	(2.04)	(2.68)	(3.07)	(2.90)
Total minutes (Z)	0.216	1.38***	0.463**	0.384	0.112
	(0.156)	(0.159)	(0.219)	(0.259)	(0.225)
Pass centrality (Z)	1.11***	0.274*	0.377*	0.283	0.314
	(0.171)	(0.152)	(0.203)	(0.250)	(0.239)
log(teammates' mean wage) × Switch in year 1	-0.022	-0.420**	-0.187	0.283	-0.300
	(0.186)	(0.201)	(0.347)	(0.271)	(0.308)
log(teammates' mean wage) × Switch in year 2	-0.372*	-0.921***	-0.704***	-0.413	-0.367
	(0.198)	(0.201)	(0.264)	(0.302)	(0.281)
Observations	5,227	5,227	5,227	5,227	5,227
R^2	0.637	0.474	0.873	0.852	0.702
Within R^2	0.438	0.298	0.557	0.600	0.583

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

A.5 Robustness checks

A.5.1 Results by player nationality

Table A15: Wage evolution of players from the Top-7 countries vs. the others

	log(wage$_{t+6}$)					
	top 7 countries			Other countries		
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.447***	0.342***	0.321***	0.459***	0.371***	0.267***
	(0.026)	(0.027)	(0.049)	(0.034)	(0.033)	(0.058)
log(value / wage)	0.249***	0.186***	0.199***	0.253***	0.187***	0.231***
	(0.023)	(0.022)	(0.043)	(0.032)	(0.030)	(0.052)
log(teammates' mean wage)	0.353***	0.455***	0.343***	0.302***	0.400***	0.522***
	(0.039)	(0.042)	(0.091)	(0.055)	(0.053)	(0.112)
Total minutes (Z)		0.248***	0.286***		0.197***	0.258***
		(0.029)	(0.047)		(0.036)	(0.060)
Pass centrality (Z)		0.013	-0.036		0.088**	-0.056
		(0.024)	(0.046)		(0.035)	(0.062)
Pass count with top 5% (Z)			0.008			0.072*
			(0.039)			(0.043)
Pass count with top 5-25% (Z)			-0.021			-0.0004
			(0.041)			(0.056)
Pass count with top 25-100% (Z)			-0.051			0.042
			(0.049)			(0.038)
Observations	3,161	3,161	939	2,066	2,066	705
R^2	0.503	0.535	0.491	0.506	0.544	0.505
Within R^2	0.208	0.260	0.231	0.215	0.276	0.260

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

A.5.2 Results by position

Table A16: Wage evolution of players by positions

	log(wage$_{t+6}$)				
	Non-GK	GK	D	M	F
	(1)	(2)	(3)	(4)	(5)
log(wage)	0.440***	0.582***	0.417***	0.434***	0.465***
	(0.021)	(0.068)	(0.040)	(0.036)	(0.040)
log(value / wage)	0.241***	0.149**	0.214***	0.276***	0.237***
	(0.019)	(0.063)	(0.027)	(0.035)	(0.036)
log(teammates' mean wage)	0.333***	0.132	0.395***	0.267***	0.328***
	(0.033)	(0.112)	(0.064)	(0.048)	(0.056)
Observations	5,227	397	1,861	1,650	1,716
R^2	0.495	0.621	0.513	0.532	0.478
Within R^2	0.208	0.267	0.213	0.205	0.208

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

Table A17: Wage evolution by position

	log(wage$_{t+6}$)					
	Defender		Midfielder		Forward	
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.417***	0.329***	0.434***	0.351***	0.465***	0.340***
	(0.040)	(0.042)	(0.036)	(0.037)	(0.040)	(0.035)
log(value / wage)	0.214***	0.161***	0.276***	0.214***	0.237***	0.157***
	(0.027)	(0.027)	(0.035)	(0.034)	(0.036)	(0.036)
log(teammates' mean wage)	0.395***	0.475***	0.267***	0.251***	0.328***	0.434***
	(0.064)	(0.098)	(0.048)	(0.074)	(0.056)	(0.084)
Minutes shared with top 5%		0.149		0.346		0.300
		(0.264)		(0.215)		(0.254)
Minutes shared with top 5-25%		-0.022		0.277*		-0.028
		(0.151)		(0.147)		(0.146)
Total minutes (Z)		0.172***		0.252***		0.228***
		(0.036)		(0.052)		(0.046)
Pass centrality (Z)		0.075*		-0.005		0.111**
		(0.040)		(0.046)		(0.043)
Observations	1,861	1,861	1,650	1,650	1,716	1,716
R^2	0.513	0.544	0.532	0.562	0.478	0.526
Within R^2	0.213	0.263	0.205	0.257	0.208	0.280

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

A.5.3 Overcontrolling

Table A18: Wage evolution with overcontrolled regressions

	log(wage$_{t+6}$)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(wage)	0.440***	0.422***	0.478***	0.343***	0.325***	0.374***	0.356***
	(0.021)	(0.021)	(0.020)	(0.021)	(0.021)	(0.020)	(0.020)
log(value / wage)	0.241***	0.227***	0.283***	0.179***	0.166***	0.210***	0.198***
	(0.019)	(0.019)	(0.020)	(0.018)	(0.018)	(0.020)	(0.020)
log(teammates' mean wage)	0.333***	0.225***	0.299***	0.435***	0.332***	0.407***	0.299***
	(0.033)	(0.033)	(0.032)	(0.034)	(0.034)	(0.033)	(0.033)
log(team points)		0.366***			0.368***		0.376***
		(0.036)			(0.037)		(0.036)
Dlog(value)$_{t+1,t}$			0.299***			0.188***	0.199***
			(0.035)			(0.035)	(0.034)
Total minutes (Z)				0.229***	0.202***	0.212***	0.184***
				(0.024)	(0.023)	(0.024)	(0.023)
Pass centrality (Z)				0.043*	0.072***	0.041*	0.071***
				(0.022)	(0.021)	(0.022)	(0.021)
Observations	5,227	5,227	5,217	5,227	5,227	5,217	5,217
R^2	0.495	0.505	0.504	0.530	0.541	0.534	0.545
Within R^2	0.208	0.225	0.222	0.264	0.281	0.269	0.286

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

A.5.4 Results by teams' elite status

Table A19: Estimates by teams' elite status

	log($wage_{t+6}$)					
		elite teams			non-elite teams	
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.353***	0.287***	0.308***	0.442***	0.343***	0.279***
	(0.054)	(0.051)	(0.054)	(0.023)	(0.022)	(0.044)
log(value / wage)	0.197***	0.115**	0.125*	0.240***	0.179***	0.209***
	(0.054)	(0.054)	(0.065)	(0.020)	(0.019)	(0.035)
log(teammates' mean wage)	0.410***	0.525***	0.493***	0.325***	0.425***	0.393***
	(0.114)	(0.112)	(0.142)	(0.037)	(0.038)	(0.094)
Total minutes (Z)		0.192***	0.216***		0.235***	0.300***
		(0.064)	(0.075)		(0.027)	(0.045)
Pass centrality (Z)		0.061	-0.022		0.042*	-0.061
		(0.072)	(0.077)		(0.025)	(0.046)
Pass count with top 5% (Z)			0.191***			-0.006
			(0.059)			(0.032)
Pass count with top 5-25% (Z)			-0.057			0.024
			(0.045)			(0.042)
Pass count with top 25-100% (Z)			-0.015			-0.024
			(0.047)			(0.042)
Observations	546	546	481	4,681	4,681	1,163
R^2	0.491	0.535	0.544	0.427	0.468	0.411
Within R^2	0.184	0.254	0.279	0.202	0.259	0.227

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

A.5.5 Results by players' wage tertiles

Table A20: Estimates by players' wage tertiles

	log($wage_{t+6}$)					
	1st tertile		2nd tertile		3rd tertile	
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.385***	0.278***	0.406***	0.299***	0.566***	0.443***
	(0.048)	(0.049)	(0.067)	(0.066)	(0.047)	(0.046)
log(value / wage)	0.193***	0.136***	0.248***	0.174***	0.341***	0.276***
	(0.028)	(0.027)	(0.033)	(0.032)	(0.041)	(0.041)
log(teammates' mean wage)	0.328***	0.437***	0.317***	0.462***	0.306***	0.391***
	(0.060)	(0.059)	(0.060)	(0.063)	(0.047)	(0.046)
Total minutes (Z)		0.177***		0.235***		0.237***
		(0.049)		(0.039)		(0.034)
Pass centrality (Z)		0.131**		0.063*		-0.019
		(0.052)		(0.035)		(0.030)
Observations	1,735	1,735	1,744	1,744	1,748	1,748
R^2	0.357	0.402	0.316	0.375	0.461	0.494
Within R^2	0.103	0.165	0.092	0.170	0.195	0.245

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity.

A.5.6 Results with manager quality

Table A21: Wage evolution and manager's past performance

	h=2		h=4		h=6	
			$\log(\text{wage}_{t+h})$			
	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	0.530***	0.531***	0.421***	0.420***	0.402***	0.401***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
log(value / wage)	0.160***	0.161***	0.212***	0.210***	0.272***	0.270***
	(0.021)	(0.021)	(0.023)	(0.024)	(0.027)	(0.027)
log(teammates' mean wage)	0.260***	0.279***	0.349***	0.325***	0.329***	0.302***
	(0.046)	(0.047)	(0.044)	(0.045)	(0.054)	(0.056)
Manager avg. score (lag 3-4)		-0.058		0.074		0.082
		(0.043)		(0.052)		(0.055)
Observations	2,415	2,415	2,415	2,415	2,415	2,415
R^2	0.734	0.735	0.613	0.614	0.491	0.491
Within R^2	0.367	0.368	0.253	0.254	0.179	0.180

Note: Standard errors clustered at position x league x half-season level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fixed effects include: league x halfseason, age, position, injury, elite team, spell count, contract maturity. This table has fewer observations as we require two prior observable years for manager performance.

A.6 Data cleaning steps: sample selection

Sample selection steps were the following:

- We start from the histories of player values (N = 1,007,727). We discard all observations without a valid market value: meaning before the first valuation (mostly U17 or U19 player histories) as we complete the market value histories with past values if it is not available. (638,776 rows remain)
- Keep only those players who have at least 11 players with a valid valuation on their team. (418,462 rows remain)
- Keep those that are:
  - in the relevant seasons of 2012/2013-2019/2020
  - Top 7 European football leagues
  - at least 18 years of age
  - data cleaning: has valid team id, has less than 50 players to interact with, and non-negative minutes
- 42,868 observations remain, out of which:
  - for each player keep the first valid half-season observation for each spell at their team where they spent at least two half-seasons (10,256 obs remain)

  - keep those with valid minutes or passing centrality information (10,081 obs. remain)
  - keep those with valid team information: points and manager known (9,886 obs. remain)
  - keep those in the 2013–2019 seasons that have 6 half-seasons of lead market value, skill, and wage observations (5,610 obs remain)
  - Exclude goalkeepers (5,227 obs. remain)

- Final sample contains 5,227 observations.

For matching the TransferMarket–WhoScored, the EA Sports' FIFA observations, and the Capology observations, we relied on player name, team name (also league), and season information. First we linked on the exact matches, and then as player and team names might have differed in the different sources (special characters, third names, etc.), we used the 'jarowinkler' function of the RecordLinkage package to clean the names and produce string similarity indices, based on Jaro and Winkler's algorithm Winkler (1990). From the rest, we matched those where player name similarities were higher than 0.95, in case of ties the one with higher similarity in the team name was selected. For the rest, we only matched when both player name's and team name's maximal similarity referred to the same observation in each of the datasources. Others were discarded. Table A22 reports the success rates by league and seniority of the player. In the EA Sport's FIFA dataset almost all players could be found, while in the Capology data only around 80%.

Table A22: Data matching success rate by league and seniority

			in_fifa_id	in_cap_id
league	player_type	N	mean	mean
england_premier-league	junior	1051	1.00	0.79
	senior	5766	1.00	0.86
france_li	junior	1736	1.00	0.75
	senior	4867	1.00	0.81
germany_bundesliga	junior	1600	1.00	0.84
	senior	4735	1.00	0.88
italy_serie-a	junior	1546	1.00	0.80
	senior	5942	1.00	0.78
netherlands_eredivisie	junior	1986	1.00	0.77
	senior	3121	1.00	0.76
portugal_liga-nos	junior	587	1.00	0.78
	senior	2428	1.00	0.79
spain_laliga	junior	1172	1.00	0.85
	senior	5716	1.00	0.84

Note: Data matching success rates based on leagues and player types.

Capology salaries and EA-FIFA wages are very similar, with a regression slope of 0.98 and R-squared of 0.653. Column (2) reports on the EA-FIFA market values and the TransferMarkt market value's relationship, which are a bit shifted (constant being quite large), but the slope still being around 0.86.

Table A23: Capology salaries vs. FIFA wages, TM market values vs. FIFA market values

	log(Capology salary) (1)	log(TM market value) (2)
Constant	-0.780***	2.03***
	(0.083)	(0.165)
log(wage in EA FIFA)	0.976***	
	(0.008)	
log(value in EA FIFA)		0.855***
		(0.011)
Observations	29,305	42,214
R^2	0.653	0.671

Note: Standard errors clustered at the player level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Finally, Table A24 compares the overall sample with relevant seasons, leagues, and variables available to the final analysis sample. The final sample, due to focusing on younger, less established players, is slightly lower in market value and salary, and play with teammates on average lower value.

Two-sided t-tests between the two samples are as follows:

Table A24: Observable characteristics and selection into final sample

	base (N=37026)		final sample (N=5227)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
TM market value (in mn EUR)	6.8	12.1	4.5	7.5	−2.4***	0.1
log(TM market value)	14.8	1.4	14.5	1.3	−0.3***	0.0
Yearly wage (in mn EUR, EA-FIFA)	1.8	2.2	1.4	1.6	−0.4***	0.0
log(yearly wage)	13.8	1.1	13.6	1.1	−0.2***	0.0
Short pass	66.8	14.2	67.9	8.6	1.1***	0.1
Reactions	70.7	7.9	67.8	8.0	−2.9***	0.1
Interceptions	54.6	21.9	53.2	20.1	−1.4***	0.3
Finishing	51.7	20.0	54.3	17.4	2.6***	0.3
Speed	68.1	13.5	71.8	10.2	3.6***	0.2
Teammates' mean value (in mn EUR)	6.3	8.0	4.5	5.7	−1.8***	0.1
log(teammates' mean value)	15.1	1.1	14.8	1.0	−0.3***	0.0
Teammates' mean wage (in mn EUR)	1.7	1.6	1.6	1.4	−0.1***	0.0
log(teammates' mean wage)	10.0	0.9	10.0	0.8	0.0**	0.0
Team HHI of market values	0.1	0.0	0.1	0.0	0.0***	0.0
Team HHI of wages	0.1	0.0	0.1	0.0	0.0***	0.0
Share of top 5% players in team	0.1	0.2	0.1	0.2	0.0***	0.0
Share of top 5-25% players in team	0.3	0.2	0.3	0.2	0.0***	0.0
Total minutes	879.6	629.4	1076.6	649.0	197.0***	9.6
Pass eigenvector centrality	0.4	0.3	0.4	0.3	0.0***	0.0
Pass count top 5%	44.6	50.8	48.6	51.9	4.0**	1.4
Pass count top 5-25%	31.0	33.0	37.7	38.7	6.7***	0.6
Pass count top 25-100%	17.1	19.1	24.0	23.0	6.9***	0.3
Minutes shared with top 5%	0.1	0.2	0.1	0.2	0.0***	0.0
Minutes shared with top 5-25%	0.4	0.3	0.3	0.3	0.0***	0.0
Player age	26.3	4.2	24.0	3.5	−2.3***	0.1
Elite team	0.1	0.4	0.1	0.3	0.0***	0.0
Injury	0.1	0.3	0.1	0.2	0.0***	0.0
Spell rank	1.6	1.0	1.8	1.0	0.2***	0.0
Contract maturity	2.6	1.3	3.1	1.3	0.4***	0.0

Note: The table displays the differences between the initial relevant panel dataset and the final analysis sample in terms of observables.

A.7 Most important factors for TM market values

Via the website https://www.transfermarkt.com/market-value-definition/thread/forum/357/thread_id/3433 the following factors are considered when producing the market value of a player.

1. Most important factors:

    - Future prospects

- Age
- Performance at the club and national team
- Level and status of the league, both in sporting and financial terms
- Reputation/prestige
- Development potential
- League-specific features
- Marketing value
- Number & reputation of interested clubs
- Performance potential
- Experience level
- Injury susceptibility
- Different financial conditions of clubs and leagues
- General demand and 'trends' on the market
- General development of transfer fees
- External factors such as the coronavirus pandemic and its consequences

2. Individual modalities:

    - Transfers via an option to buy/obligation to buy
    - Loan fee
    - Only part of transfer rights acquired
    - Exit clause
    - Buyback option
    - Player swap deal
    - Contract length
    - Resale participation
    - Bonus payments
    - Improvement of financial balance

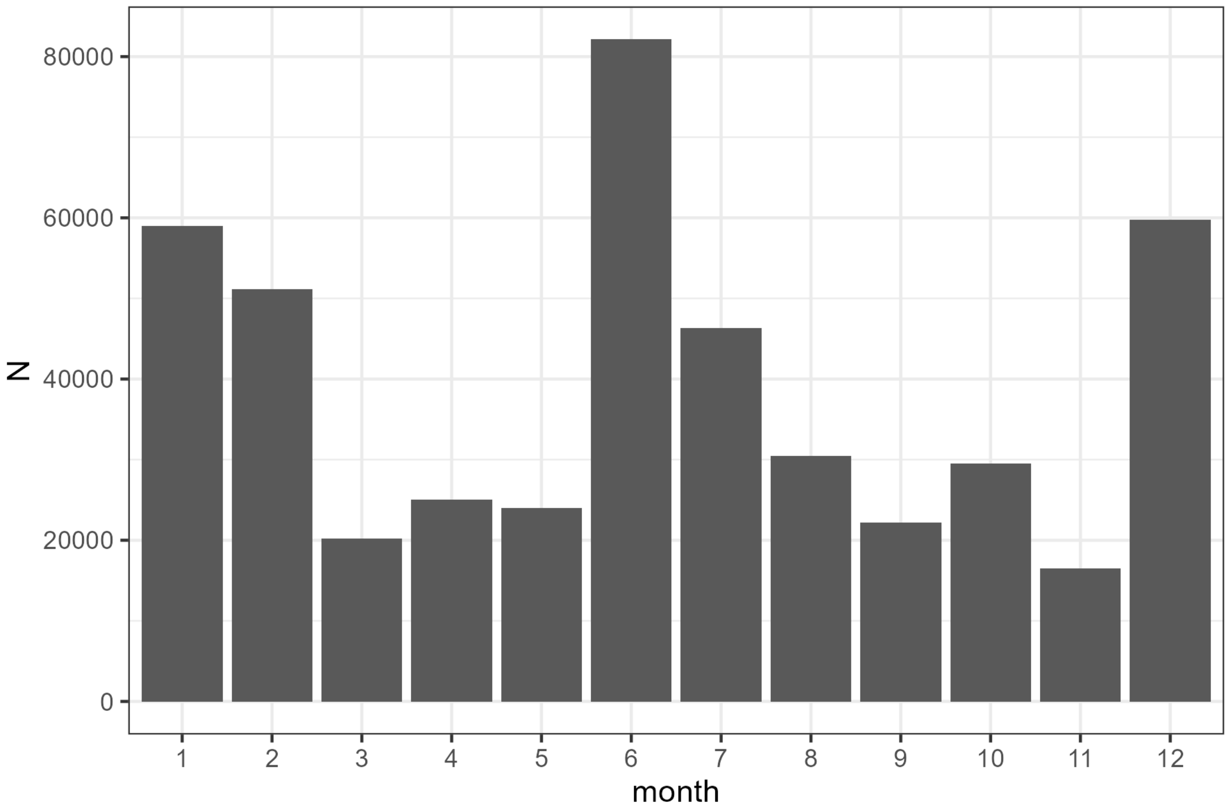
3. Situational conditions:

    - Pressure situations such as competitive, success or financial pressure, etc.
    - Will/desire/interests of the player
    - Club does not sell to highest bidder

- Player goes on strike or similar
- High salary
- Club wants to sell player

The frequency of updates in the market value is displayed below, showing that it is most prevalent in the Winter and in the Summer:

Figure A7: Frequency of market value updates



Note: The figure shows the number of market value evaluations per month in the TransferMarkt original data.