UNIVERSITY of York

HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

WP 18/01

Health shocks and labour market outcomes: evidence from professional football

Vincenzo Carrieri; Andrew M. Jones and Francesco Principe

January 2018

Health shocks and labour market outcomes: evidence from professional football

Vincenzo Carrieri
University of Salerno
HEDG, University of York
RWI Research Network

Andrew M. Jones
University of York
Monash University
University of Bergen

Francesco Principe*
University of Salerno
University of York

Abstract

The negative association between health shocks and labour market outcomes is a well-established finding in the health and labour economics literatures. However, due to lack of data, most of the contributions focus on elderly workers and analyse labour force participation as their main outcome. This paper uses traumatic injuries as a source of exogenous variation in professional football players' health to provide estimates of the causal impact of a health shock on two main labour market outcomes: annual net wages and the probability of renegotiating the contract between club and player. We have built a unique longitudinal dataset recording information about wages, injuries and other characteristics of the universe of professional football players in the Italian Serie A from 2009 to 2014. We employ panel fixed effects models combined with an IV strategy, which uses the average number of yellow cards received by the team as an instrument. We find that injuries reduce the net wage in the following season by around 12%. This result is mainly driven by precautionary reasons due to the club's concern about depreciation in the player's human capital rather than by a direct effect of the shock on the player's productivity.

Keywords: health shocks; top incomes; football; panel data; instrumental variables.

JEL codes: C26; D31; I1; J24; J31.

*Corresponding Author: Department of Economics and Statistics, Via Giovanni Paolo II, 84084 Fisciano (SA), Italy. E-mail: fprincipe@unisa.it

1. Introduction

Health shocks are a major concern for the labour market. In fact, health shocks might affect peoples' lives across a number of different dimensions including productivity, labour market participation, labour income and healthcare costs, resulting in reduced household disposable income. The negative association between health shocks and labour market outcomes is a well-established finding in the health and labour economics literatures, with evidence from both developed and developing countries (see Section 2 for some insights from the literature). However, due to the scarcity of relevant data, most of the contributions have focused on elderly workers and have analysed labour market participation as their main outcome.

Furthermore, to the best of our knowledge, there are no studies that analyse the effects of the health shocks on the top tail of the labour income distribution, the so-called "working super-rich". Indeed, in the last decade, this segment of the distribution has been the subject of keen economic interest for a number of reasons. Firstly, since the 1970s, top income shares started to rise in most developed countries, in particular, the English-speaking ones (Atkinson *et al.*, 2011). Secondly, top incomes are considered a reliable indicator of long run inequality across the entire distribution: analysing the income concentration at the top is useful to understand the bottom of the distribution In fact, the income concentration at the top of the distribution is highly correlated with relative poverty and other inequality measures (Leigh, 2007). Lastly, following the recent contributions of Alvaredo *et al.* (2013), a substantial change in the composition of income shares at the top of the distribution has been found. In fact, in all the countries for which composition data are available, there has been a shift from capital income to labour income as the main component of the top 1% income shares. Thus, the labour market seems to be fertile ground for the escalation of contemporary society's extreme inequalities and, understanding how the "working super-rich" respond to health shocks can be a key element of these dynamics.

This paper contributes to the literature by providing evidence of the relationship between health shocks and labour market outcomes, focusing on those in the top tail of earnings distribution. In particular, the aim of this research is to analyse traumatic injuries as exogenous variation in professional football players' health to provide estimates of the causal impact of a health shock on two main labour market outcomes: the annual net wages and the probability of renegotiating the contract between the employer (the club) and the employee (the player).

The original contributions of the paper can be summarised as follows. Firstly, it provides causal evidence of the relationship between health shocks and labour market outcomes. Secondly, it focuses on the consequences of health shocks for those on top incomes, for whom there is scant evidence so far. Lastly, it allows for a deeper analysis of the main mechanisms of the health shock, disentangling the effect mediated through the player's performance and the one generated by human capital depreciation, inducing the club, to offer a lower wage for precautionary reasons.

We have created a longitudinal dataset recording data about wages, performances, popularity and injuries of the football players of the Italian Serie A[1]. This dataset was built by merging information from several reliable sources of data. In fact, in Italy, football players represent an important share of the top income group. According to Franzini *et al.* (2016), in 2003, among the top 500 income earners, who represent the top 0.01% of the distribution, about 20% were football players or managers.

In this paper, to exploit the longitudinal nature of our dataset, we make use of a fixed effects model to sweep out individual time invariant characteristics and an instrumental variable strategy to reduce any further endogeneity concerns.

We find that a health shock has a negative impact on the net annual wage of the football players of the Italian Serie A. In particular, a 30-day injury results in a reduction of the following year's wage by about 12%. Interestingly, the results suggest that this reduction can be explained in terms of precautionary measures being taken by the club rather than through any direct effect that the injury might have had on performance. Furthermore, the injury is found to have no effect on the probability of renegotiating the contract at the end of the season for those players who are not in the final year of their contracts. However, particularly severe injuries might increase the probability of renegotiation by inducing the players to be traded to a lower level team. Finally, we find that the main factors addressed by the literature about the determinants of super-earnings, such as performance and popularity, have a significant impact on the players' wages. These results are robust to a number of model specifications and to alternative measures of health shocks.

The remainder of the paper is structured as follows. Section 2 provides some insight about the literature on health shocks' consequences on employees' outcomes and the theoretical and empirical evidence about working super-rich's wages determination. Section 3 describes the data gathering process, explaining the main variables and providing some descriptive statistics. Section 4 discusses the empirical methodology. Section 5 presents the results and the final section summarises and concludes the study.

[1] Italian Serie A is one of the five most followed football leagues in the world, alongside the Premier League (England), Bundesliga (Germany), La Liga (Spain) and Ligue 1 (France). Yearly, it has a turnover of about 1.9 billion Euros. Among the "big five" leagues, Italian Serie A is the one with the highest incidence of the wage bill on the clubs' costs, absorbing about 70% of the clubs' total earnings (Deloitte, 2017).

2. The Literature

This paper is grounded in two main strands of the literature. Firstly, the literature about the impact of health shocks on labour market outcomes. Secondly, the theoretical and empirical literature about the determinants of top incomes, in particular the working super-rich, defined as individuals who receive high earnings due to their employment.

The relationship between health shocks and labour market outcomes is well established in the economic literature, with empirical evidence provided for both developed countries (e.g., García-Gómez and López-Nicolás, 2006) and for developing countries (e.g., Wagstaff 2007, Wagstaff and Lindelow 2014, Mitra *et al.* 2016). However, due to the availability of data, most of the contributions focus on labour participation as the main outcome (e.g., Jones *et al.* 2010, Cai *et al.* 2014) and use data based on elderly workers (e.g., Bound *et al.* 1999; Disney *et al.* 2006; Lee and Kim, 2008; Trevisan and Zantomio, 2016). The main findings show that health shocks have a negative impact on several dimensions, including consumption and income. For example, García-Gómez and López-Nicolás (2006), using Spanish data from the European Community Household Panel, find that a transitory health shock has a negative impact on workers' labour income that ranges between 25% and 40%, according to the length of the incapacity period. While, García-Gómez *et al.* (2013), using Dutch hospital and tax register data, find that an acute hospital admission lowers the employment probability by 7% and results in a 5% loss of personal income two years after the shock.

However, one main challenge when analysing the effect of health shocks on labour market outcomes is dealing with the possible endogeneity of health shocks in relation to labour supply, which may result from both simultaneity and unobserved heterogeneity (Lindeboom and Kerkhofs 2009; Cai, 2010). With respect to this, in this paper, we use traumatic injuries as exogenous variation in player's health and an instrumental variables approach to reduce any further endogeneity concerns.

A second strand of literature is mainly based on the "superstars" theories offered in the seminal contributions of Rosen (1981) and Adler (1985), as the basis for disentangling the determinants of super-earnings. In fact, they both argue that superstars arise in markets characterised by imperfect substitution on the demand side and joint consumption on the supply side, which generate a demand concentration towards the better performers (i.e., sportsmen, singers, artists, etc.), who "win and take all". However, while for Rosen (1981) marginal differences in talent are magnified into huge earnings for the most talented due to the convexity of the revenue function; Adler (1985) argues that popularity is the main determinant of superstars' earnings.

From an empirical point of view, a number of researchers have made use of sports data to test these theories. For example, Lucifora and Simmons (2003) investigate “superstar” effects in the wage determination among football players in the Italian Serie A. They find earnings to be highly convex in measures of performance, after controlling for a set of personal characteristics and team fixed effects. Franck and Nuesch (2012), using data from the German Bundesliga, find that both talent and popularity significantly contribute to increasing the market value of superstars. Other studies (e.g., Mullin and Dunn 2002, Treme and Allen 2009, Treme and Allen 2011) focus on American professional sports, finding a positive effect of both measures of performance and media exposure on the entry earnings of baseball (MLB) and basketball (NBA) players.

Building on these strands of literature, this study aims at reconciling the evidence cited above and investigating the impact of a health shock on the labour market outcomes for a specific category of working super-rich: the professional football players of the Italian Serie A. This choice is based on a number of underlying considerations. Firstly, the football market is based on short-term contracts (generally 3-4 years), generating high volumes on renegotiation and players being traded in every market window; this allows for observations based on a large variation in wages across seasons. Secondly, in Italy, the football players constitute a significant share of the top earners. In fact, in 2003, among the top 500 earners, representing the top 0.01% of the distribution, about 20% were football players or managers (Franzini *et al.* 2016). Lastly, to the best of our knowledge, there is no previous evidence investigating the consequences of health shocks on the earnings of individuals in the top tail of the distribution.

Thus, as argued by Kahn (2000): “professional sports offers a unique opportunity for labour market research. There is no research setting other than sports where we know the name, face, and life history of every production worker and supervisor in the industry". In fact, professional sports data allow a perfect match between employer (the club) and employee (the player) and to have measures to proxy individual productivity and estimate team production functions. Indeed, in recent years, the use of sports data to address health-related issues has been rising. Recent contributions include: Stoecker *et al.* (2016), who examine the impact of influenza transmission on mortality by looking at local sports team success through the participation in the playoffs and championship game (Super Bowl) of the National Football League (NFL)[2]; and Hanson *et al.* (2017), who exploit a difference-in-differences framework in order to examine the intended and unintended effects of the “Crown of the Helmet Rule”[3] on players’ injuries in the NFL.

[2] American football league.
[3] Implemented in 2013 by NFL, it aims at reducing incidence of concussions and head injuries by penalizing a player who intentionally initiates contact with another player using the top of his helmet.

3. Data and variables

We created an original dataset recording information about wages, performance, injuries and other individual characteristics of the professional players of the Italian Premier League (Serie A), using several data sources. As a starting sample, we analysed data relating to 469 players who have had at least one appearance in the 2013-2014 football season - excluding goalkeepers, following the standard approach of this kind of literature (Lucifora and Simmons, 2003), since their performance is measured differently and they have a dissimilar exposure to the risk of injury compared to the outfield players. The careers of the players were then followed over a 5-season period, from 2010-2011 to 2014-2015. This provides a longitudinal dataset of 1,585 observations. The panel is unbalanced. In fact, the relegation/promotion system[4] between Serie A and Serie B and the transfer market across national and international clubs generates a relatively large turnover of players in the league.

Data on players' yearly wage - recorded net of taxes and excluding any performance-related bonus – are taken from the annual report, published at the beginning of each football season by the most influential Italian sport newspaper, *La Gazzetta dello Sport*. Importantly, focusing only on the fixed part of the wage allows for better understanding the effect of the shock on the stock of human capital and reduces concerns about reverse causality, since the dependent variable does not include bonuses that depend on performance. Moreover, to assess the impact of the injury on the renegotiation of the contract, we built a dichotomous variable with value 1 if the contract has been renegotiated at the end of the season with the same club or a new one, in case the player has been traded during the transfer market window, or value 0 in case of no renegotiation.

Information about individual player's characteristics (i.e., birth year, position on the pitch and international appearances) and performance (such as, goals and assists) are extracted from the website *transfermarkt.com*[5]. In addition, in order not to underestimate the performance of midfielders and defenders, we collected data about each player's overall performance. These were recorded as the average of the three most read Italian sport newspapers, *La Gazzetta dello Sport, Il Corriere dello Sport* and *Tuttosport*, which rate the player's performance after every match, with a scale that ranges from 0 (poor performance) to 10 (excellent performance). In fact, most of the marks assigned by the journalists range between 4 and 8.

[4] There are 20 teams taking part in Serie A. At the end of the season, the last three are in the table are relegated to the second division (Serie B) and replaced by the first three of Serie B.
[5] A German website recording information about football statistics, results, fixtures and news. Data from transfermarkt.com have been used in previous economic studies (i.e., Bryson et al., 2013).

With respect to injuries, we collected data about any injury that occurred in the seasons analysed. In particular, all the injuries were categorised according to the terminology and classification of Mueller-Wohlfahrt *et al.* (2012). Alongside the type of injury, we collected information about the impact of the injury proxied through the days off, identified as the period of time in which the player has not been available for participating in the sports activity of the club (training and matches), and the number of official championship matches missed by the injured player. These data are taken from *footballmarkt.com*. Furthermore, since in the same season a player could incur more than one injury, to build the variable of interest, we consider the number of days off due to either muscular or traumatic injury and we include a dummy variable to control for the re-occurrence of the same injury during the season.

Concerning the clubs, we analysed the annual balance sheets as approved by the directors of the 27 clubs, matched with the players who took part in Serie A for the seasons considered. In particular, information was recorded about five balance sheet items: total wage expense, net sales, earnings before taxes (EBT) and revenues from ticketing and revenues from television rights.

Other variables include a proxy of popularity, measured through the pre-season number of Google search results obtained each year for each player[6] and several other characteristics, related to both clubs and player, which are used as controls in the estimates. These are presented, along with some summary statistics for all the variables, in the next section.

3.1 Descriptive statistics

All the variables included in our dataset, along with their mean values and standard deviations, are presented in Table 1. According to data used in this research, the average annual net wage of a player, in Italian Serie A, amounts to about 840,000 Euros, but with a very large standard deviation (891,000 Euros). Figure 1 shows the non-parametric estimate of the wage distribution for the reference football season 2013-2014. The distribution is positively skewed with a long upper tail. This supports the idea of a restricted number of players who earns huge wages compared to the rest of the distribution and it is consistent with the theories of "superstars" emerging (Rosen 1980, Adler 1985). Indeed, the Gini index within the sampled football players group is 0.47, showing a large degree of inequality even within this group of "privileged" individuals.

[6] Data refers to the same day for each player. They have been collected browsing "name-surname-team" in order to reduce any bias due to homonymy with respect to some surnames, which have higher incidence in Italy.

Furthermore, the dataset shows that more than 65% of the contracts were renegotiated every year. This reinforces the focus on a particular type of job market in which the contracts quickly reflect the changes in terms of both productivity (performance) and health shocks.

Concerning injuries, 50.5% of players included in the sample had at least one injury, of which 41.7% were muscle related, while the remaining 58.3% were caused by some kind of traumatic event, such as fractures, cruciate ligament rupture, etc. The re-occurrence rate is 10.4%. In the reference season 2013-2014, on average, a player spent 21 days off activity and missed about 4 official matches. Moreover, information about the exact date of the injury was also collected. Figure A.1 (in the Appendix) shows the percentage distribution of the injuries alongside the different months of the year. Interestingly, the highest percentage of injury was recorded in September (about 15%), which is the first month of official matches after the summer break, followed by the winter months, probably due to worse weather conditions and player fitness. This heterogeneity is controlled for through the inclusion of month fixed effects in some model specifications.

In this paper, we focus in particular on traumatic injuries since they can be interpreted as an exogenous health shock. This assumes that the probability of such an exogenous shock is not correlated with any observed or unobserved characteristics that are related to the outcomes. Figure 2 shows the non-parametric wage distribution comparing the group of those having at least one traumatic injury during the season and the control group of those who did not. The two distributions are highly similar, showing a long upper tail and an asymmetric distribution, a characteristic already observed above, when analysing the whole sample. However, as reported in Table 2 along with the relevant percentiles of wage distribution, a large dispersion emerges but the distribution is fairly similar within both groups. In fact the within-groups inequality is slightly larger among the injured players, with the Gini indices of 0.45 and 0.43, respectively.

Furthermore, to ensure that the two groups do not systematically differ due to pre-injury observable characteristics and to investigate possible concerns of selection bias, we report descriptive statistics for the relevant covariates in Table 3. Remarkably, the mean ages of the two groups do not show a large difference, being 27.9 and 27.3, respectively. The injured players on average earned 95,000 euro more than the controls and played 200 minutes less. However, even if the average wage gap between the two groups seems to be large in absolute terms, it is important to notice that since the paper analyses a sample of individuals with very high earnings, the relative size of the gap is less concerning. Indeed, a Student's t-test of the means suggests that these differences are not statistically significant at the 5% level, reinforcing the idea that such a shock may be regarded as random.

Finally, Table A.1 reports the clubs' characteristics regarding some relevant balance sheet items. The clubs have great differences in wage bills with the richest clubs paying almost 7 times more than the bottom ones. Indeed, this inequality is also reflected in the television rights distribution which is fairly skewed among the clubs. Substantial heterogeneity emerges among Serie A clubs suggesting the need for controlling for team fixed effects in the estimates.

4. Empirical strategy

We aim to obtain causal estimates related to the impact of a health shock on three dimensions: wages, probability of renegotiating the contract and productivity. Firstly, a fixed effects model is used to assess the impact of an injury on the player's wage while controlling for unobserved heterogeneity and an IV strategy is used to handle potential endogeneity concerns and random measurement errors. Secondly, a fixed effects logit model is used to assess the probability of contract renegotiation following the injury.

4.1 Fixed-effects model

We make use of the fixed effects model as principal specification. In fact, after performing the Hausman test, it strongly rejects the hypothesis that the random effects model provides consistent estimates (Woolridge, 2002). The main strength of the FE model is that it takes into account time-invariant observed and unobserved individual characteristics. Controlling for time-invariant unobserved heterogeneity eliminates the effect of some individual peculiarities, such as their attitude to training and genetic or physical characteristics, which might have an effect on both the health shock and the outcome.

Thus, the model to be estimated is the following Mincerian wage equation:

$$\boldsymbol{Log(W)_{it} = \beta_0 + \beta_1 X_{i,t-1} + \beta_2 P_{i,t-1} + \gamma S_{i,t-1} + \delta Season_t + u_{it}} \qquad \textbf{(1)}$$

where, the dependent variable is the logarithm of the annual player's wage in season t, net of taxes and bonuses. The coefficient of interest is $\boldsymbol{\gamma}$, where the covariate $\boldsymbol{S_{it-1}}$ represents the incidence of the health shock through the number of days off due to the reported injury. $\boldsymbol{X_{it-1}}$ represents a set of individual time-varying characteristics that are used as controls. All the variables are time lagged to reduce potential reverse causality concerns. Specifically, we include the following controls: age, the square of age, a number of experience related characteristics (e.g., the number of international appearances with either the senior national team or the under 21, a dummy for the team's captain), the total number of minutes played in the previous season. Moreover, we include dummies for the player's position on the pitch (defenders, midfielders and forwards), dummies to distinguish the players nationality (Italian, EU and extra EU) and season dummies. $\boldsymbol{P_{it-1}}$ is a set of covariates which includes proxies of talent and popularity to take into account the main explanations of superstars' earnings (Rosen 1980, Adler 1985). These are goals scored, assists served, ratings given by the newspapers and the number of Google search queries. $\boldsymbol{u_{it}}$ is the composite error term, including both the individual specific and the idiosyncratic error terms.

4.2 Instrumental variable strategy

While the fixed effects model controls for time invariant individual characteristics, unobserved idiosyncratic factors affecting both the health shock and the outcome of interest might be still an issue. Thus, to address concerns about the endogeneity of the health shock and possible measurement errors, we exploit a fixed-effects instrumental variables strategy (FE-IV).

The first stage equation is defined as follows:

$$\boldsymbol{S_{it-1} = \beta_0 + \beta_1 X_{i,t-1} + \beta_2 P_{i,t-1} + \gamma z_{i,t-1} + \delta Season_t + u_{it}} \qquad (2)$$

where the instrument **z** is assumed to satisfy ***cov(S,z)≠0*** and ***cov(u, z)=0*** so that it is correlated with the endogenous variable $\boldsymbol{S}$ and uncorrelated with the error term $\boldsymbol{u}$. In the FE-IV estimator, the fitted values of $\boldsymbol{\hat{S}}$ derived from equation (2) are plugged into the original regression equation:

$$\boldsymbol{Log(W)_{it} = \beta_0 + \beta_1 X_{i,t-1} + \beta_2 P_{i,t-1} + \gamma \hat{S}_{i,t-1} + \delta Season_t + u_{it}}$$

As an instrument we make use of average value of the yellow cards received by the team *j*, excluding those of the player *i*. This excludes the individual's own effort and creates a more pure peer effect, which is correlated with the individual "aggressiveness" put in the game but not with his own wage. Yellow cards have the advantage of measuring the severity of the fouls and should be exogenous being assigned by an external figure, the referee. Indeed, we also might have used information about the yellow cards received by the opponent teams. Yet, the fact that the model takes the football season as the unit of time, implies that in one season every team faced the others twice, thus using the opponent's yellow cards as an instrument would not show enough dispersion.

There are a number of conditions to hold for the instrument to be reliable. Firstly, the instrument has to show some variation over time. This is guaranteed by the fact that, as argued above, the trading of players generates a large turnover of players among teams. The same is also true for the managers. Thus, every season, the players play alongside different teammates and face a different team attitude, according to their team's aims and manager's strategy.

Secondly, the instrument has to be correlated with the potential endogenous variable "days off due to injury". In the framework, yellow cards can be used as a measure of the "aggressiveness" of the team. Thus, it is correlated to the probability of being injured through two different channels. On one hand, a more aggressive team might incur a higher risk due to their attitude towards the game. On the other hand, due to a mechanism of reciprocity, the opponent team might play in a tougher way in response, increasing the number of foul plays and, consequently, the probability of injuries.

Lastly, the instrument, conditional on the health shock, should not affect the individual wage. With respect to this, we make use of the average yellow cards of the player's i teammates, which may affect his game style through a "peer effect", but it is unlikely to have any correlation with his own wage.

4.3 Binary choice models

One main advantage of using sports data for labour market analysis is the availability of information about the length of the contract between the player and the team. There are two main reasons for renegotiating the player's contract. First, if the contract is close to expiry, the club might not want to risk losing a player without any monetary compensation. Second, the player may be traded to another club during one of the two transfer market periods, thus entailing the signing of a new contract. To assess the impact of the health shock on the probability of renegotiating the contract at the end of the season, we make use of two different binary choice models.

Firstly, a linear probability model (LPM) is used to exploit the longitudinal nature of the data and the IV strategy, with fixed effects to control for individual time-invariant characteristics. We estimate the following equation:

$$\boldsymbol{d_{it} = \beta_0 + \beta_1 X_{i,t-1} + \beta_2 P_{i,t-1} + \gamma S_{i,t-1} + \delta Season_t + u_{it}}$$

where $\boldsymbol{d_{it}}$ is a dichotomous variable assuming either value 1, if the contract between the player and the club has been renegotiated at the end of the season, or value 0 if it had not. The covariate $\boldsymbol{S_{it-1}}$ represents the incidence of the health shock through the number of days off due to the reported injury. The remaining set of covariates is the same as described in the equation (1). We use standard heteroscedasticity-robust standard errors and t statistics to deal with heteroscedasticity. Furthermore, to reduce any concern about the endogeneity of the health shock, we implement the

IV strategy as explained above, using the 2SLS estimator and instrumenting the injury through the average value of the team's yellow cards, excluding those of the individual *i*.

Secondly, we use a fixed-effect logit model. The FE logit allows for both the individual time invariant characteristics and the dichotomous nature of the dependent variable. Compared to the LPM, the FE logit gives a response probability that ranges from 0 to 1. However, one drawback of the FE logit models is that the interpretation of the results is somewhat cumbersome, due to the problem of computing predicted probabilities of the outcome and the marginal or discrete effects when the fixed effect is unknown. To overcome this issue, it is common practice in the empirical literature to interpret the effect in terms of odds ratio or on the conditional probability (Cameron and Trivedi, 2010). Indeed, both these methods are less intuitive when it comes to the economic interpretation. In this paper, we interpret the FE logit model coefficients in terms of odds ratios.

5. Results

5.1 Strength of the instrument

Before presenting the main results, we assess the validity of the IV approach. Table A.2 in the Appendix presents the first-stage regressions on the full set of covariates, for the different specifications. The instrument is strongly correlated with the health shock measure. Thus, results indicate that playing for a more aggressive team is associated with a higher number of days off, significant at 1% level, confirming the intuition behind the relevance of the instrument.

Furthermore, at the bottom of the Table A.2, we report the F-test results of whether the excluded instrument significantly differs from zero. The F-statistic is always above 10, which is generally adopted as "rule of the thumb" for the minimum threshold to reject the null hypothesis of a weak instrument (Staiger and Stock, 1997). In addition, Stock and Yogo (2005) provided a more formal threshold (16.38) to test the weakness of the instrument. Our instrument's F-test is also above this critical value, ranging between 18.34 and 20.45 across the different specifications.

5.2 Effects of health shock on wages

Tables 4-6 show the results of the fixed effects model described in Equation (1), with the logarithm of wage as the dependent variable. Table 4 shows the results for the full sample, Table 5 controls for muscular injuries and Table 6 shows results for a subsample in which muscular injuries are dropped. Column 1 of each table reports the results of the FE model, without team and month fixed effects. Column 2-3 include results with either team or month fixed effects, Column 4's results account for both. Moreover, to make the economic interpretation of the coefficient of interest easier and comparable, we account for 30 days variation in the number of days off due to injury. Other covariate coefficients are presented according to a change of one standard deviation, as reported in the summary statistics (Table 1). Since the dependent variable is in log format, the estimated coefficients report a percentage variation in the net annual wage associated with a one standard deviation increase of the independent variable.

Table 4 shows a negative relationship between days off due to injury and the net wage. In particular, one-month of injury is associated with a reduction of about 2.5% in the wage of the following season. The size and significance of the coefficient remains the same when controlling for team fixed effects and it slightly increases where month fixed effects are controlled for.

With respect to the other covariates, Table 4 shows that both indicators of performance and popularity are significant and have a positive effect on the annual net wage. In particular, in the baseline model (Column 1), a one-S.D. increase in goals and popularity affects the following season's wage by about 4.9% and 3.2%, respectively. Coefficients remain almost unaltered when accounting for team and month fixed effects.

These results are robust to different model specifications, when controlling for muscular injuries (Table 5) or dropping them from the sample (Table 6). Indeed, in Table 5, the coefficient of days off due to muscular injuries is not statistically significant. This result supports the intuition that the health shock effect on wages is due to those injuries which have an underlying random component, such as the traumatic injury types.

Interestingly, it is possible to observe a larger spread between coefficients when month fixed effect are included. These results indicate that, during the wage bargaining process, variation across the season should be accounted for. In particular, some months might be associated with worse weather conditions or with a larger number of matches played due to the calendar of fixtures, which exposes players to higher stress in some periods of the season.

The results for the FE-IV estimation on the full set of controls are shown in Table 7. Making use of the teammates' yellow cards as an instrument, results in a significant increase in the coefficient of the days off variable. A one-month injury has a more negative effect on the wage, reducing it by about 12%; the coefficient is statistically significant at 5% level.

With respect to the other covariates, a one-S.D. increase in goals and popularity leads to a positive effect on the annual net wage of about 5.4% and 3.4%, respectively. The effect is slightly higher than in the FE model and the coefficients concerning goals are still significant at the 5% level while the for popularity they drop to the 10% level, showing less spread among the different model specifications, when team and/or month fixed effect are controlled for.

Moreover, results show a non-linear relationship between age and the wage. In particular, wage increases with age due to the cumulated experience and knowledge of the game until a turning point in which these time varying personal characteristics are overcome by the progressive loss of physical skills. We estimated this turning point to be at around 32 years old, higher than in the early-2000 evidence (Lucifora and Simmons, 2003) and consistent with the evidence of the elongation of players' careers observed in the last 15 years.

Table 8 reports estimates that exclude the performance-related covariates. This further specification disentangles the effect of the injury on the wage between being a direct effect and an indirect one, which is mediated by the performance. It is interesting to notice that when the model does not control for performance, the coefficient of the days off increases by around 1%, across the different specifications. Thus, it is possible to argue that the effect of the injury is mediated through performance, due to reduced physical skills, by only a small amount. Indeed, these results can be seen in a framework of human capital depreciation as showing that during the bargaining process the club is willing to offer a lower amount to the player for precautionary reasons rather than because it is concerned by the negative effect of the injury on the individual's productivity.

5.3 Effects of health shocks on renegotiation

Table 9 shows the results of the linear probability model described in Equation (2), for the dichotomous variable representing the wage renegotiations as the dependent variable. As expected, health shocks are positively associated with the probability of renegotiating the contract at the end of the season. However, the coefficient is not significant at 10% level.

Concerning the other covariates, every additional 90 minutes played increase the probability of renegotiating the contract by about 0.85 points. Essentially, the clubs are interested in extending the contracts of those players who guarantee a higher reliability in terms of both physical and tactical integrity and experience, cumulated through the larger number of matches played.

Interestingly, a 1 S.D. increase in popularity is negatively associated with the contract renegotiation. This result is indicative of a well-known dynamic in the football market. In fact, when a player's popularity increases they have an increased incentive to delay the contract's renegotiation with their own club with the aim of arriving as close as possible to the expiration date in order to threaten the club into offering them a higher wage so as not to lose the player without any monetary compensation, at the end of the contract. Furthermore, a more popular player represents for the club an important asset in terms of merchandising and image rights. Hence, the bargaining process becomes much more complex and the timing plays a central role.

These results hold for the fixed-effect logit model. In fact, both traumatic and muscular injuries are positively associated with contract renegotiation, holding the other covariates constant, but both are not significant at the 10% significance level. With respect to other covariates, an additional 90-minute match played has a positive effect in favour of contract renegotiation, which is higher by approximately 1 percentage point and a 1 S.D. increase in popularity generates an effect that is approximately 3 percentage points lower.

In a further specification, we run both models without controlling for performance and thus including in the sample also those individuals whose performance was not measurable due to the fact that they did not play any official match, because of season-long injuries.

Table 10 shows an important new finding. On one hand, the coefficients corresponding to the matches played and popularity remain significant with similar magnitudes, affecting the probability of renegotiation by approximately 1% and 3.1%, respectively. Conversely, compared to the previous specification, the injury effect is now significant at the 5% level. Specifically, a 30-day injury increases the probability of renegotiating the contract by approximately 9.8%. Results show similar patterns when considering the FE logit model, in terms of odds ratio.

Thus, from these results an interesting dynamic emerges. Injury is found to have a significant effect on the probability of renegotiation of the contract only when the injury persisted throughout the entire season, preventing the player from taking part in any official match. It follows that a player, failing to demonstrate a degree of reliability measured in terms of the number of games played, is found to be compelled to migrate to another club at the end of the season and thus to renegotiate his contract.

Overall, the results show that health shocks have a negative impact on the wages of football players. In particular, they reduce the net annual wages. However, only a residual part of this reduction is justified by the negative effect that the injury has on performance. In fact, the results suggest that the clubs tend to reduce wages for precautionary reasons, which can be understood as following a human capital depreciation model. Moreover, the injury itself does not directly increase the probability of renegotiating the contract at the end of the seasons. If a player still has some years left before the contract expires, the club does not have any option of forcing the player to renegotiate, except the option of trading him to another club. In fact, the results show that the effect of the injury on the probability of renegotiation is significant only when taking into account those players who did not play any match in a season due to long-term injuries and, thus, are more likely to be traded to a lower level club at the end of the season.

6. Robustness checks

To check the robustness of these findings, we perform a number of sensitivity analyses. Firstly, an alternative measure of health shock, namely the total number of official league matches in which the player was reported as "not available due to injury" is used to check whether the results are confirmed. In fact, it may be argued that the maximum number of days off is specific to the form of injury and not influential in given periods of the year if it does not overlap with any official match. Nevertheless, the channel through which the club decides to offer a lower wage might be based on the actual rate of participation of the player in official matches. The estimates, reported in Table 11, show that the results are robust based on this alternative health shock measure. In fact, every additional match missed negatively affects the wage for the following season by approximately 2.8%. This result is consistent for both magnitude and sign with the main results. In fact, in one month, on average, a club has scheduled about 5 Serie A matches, making this result proportional to the baseline one, in which 30-day variations in the number of days off were considered. Furthermore, this alternative model confirms the strength of the instrument, whose F-test ranges between 16.56 and 19.08, across the different specifications, as reported in the first stage F-test at the bottom of Table 11.

Secondly, we test whether the dynamics of the model are correctly specified. In fact, under multi-year contracts, the effect on the wage might have been determined by a health shock which occurred in a previous season compared to the one of the baseline model (i.e., t-2 or t-3). Thus, we run the model including the lagged values of days off, up to three seasons before. Table 12 shows that even though the number of days off in earlier seasons is negatively associated with the net wage, the coefficients are not statistically significant at conventional levels, reinforcing the idea that the negative effect is caused by the injury occurred just in the previous season with respect to the eventual contract renegotiation.

Thirdly, to check the robustness of the hypothesis about the existence of a difference between the direct effect of the health shock on the outcome and the indirect one, mediated through performance, we estimate performance measures that are not affected by injury. In particular, we proxy the non-injury affected performance by the OLS residuals of the regression of the performance's measures (i.e., grade, goal and assist) on the number of days off, while controlling for age and seasonal fixed effects. Then, we include the injury-purged variables as covariates in the

FE-IV estimation. The results, reported in Table 13, show that the magnitude and the sign of the health shock index remains unaltered compared to the baseline model. Interestingly, in this specification, the grade of the overall performance becomes statistically significant: a 1 S.D. increase in the injury-purged performance positively affects the net wage of the following year by approximately 4.7%. This strengthens the evidence in support of the existence of both a direct and indirect effect of the health shock. Furthermore, the first stage regressions reported in Table 13 show a slight improvement of the power of the instrument in this specification and the F-tests range between 19.77 and 21.98.

7. Conclusions

The relationship between health shocks and labour market outcomes is an issue of keen interest in both health and labour economics. In fact, a substantial body of literature finds a negative relationship between health shocks and several labour market dimensions. However, due to the lack of sufficient data, most of the studies focus only on elderly workers and analyse the labour market participation as the main outcome. Thus, to the best of our knowledge there is no evidence on the health-income dynamics at the top of the income distribution. Indeed, this segment has attracted the interest of scholars since in many countries data on top earners are public. Moreover, another strand of literature suggests the determinants of the earnings of the working super-rich might be different from the rest of the distribution, due to the central role played by talent differentials and mass-media exposure.

We use a unique longitudinal dataset, created by recording information about wages, performances, popularity and injuries of the football players of the Italian Serie A, spanning the football seasons 2010-2011 to 2014-2015. Traumatic injuries are accounted for as exogenous variations in the individual's health. Furthermore, the potential endogeneity of the injury is addressed through an instrumental variable strategy combined with fixed effect models to control for time invariant individual characteristics that might affect the estimates; thus our results can be interpreted as causal.

The results confirm a negative relationship between health shocks and labour market outcomes. The main findings can be summarised as follows. Firstly, health shocks have a casual impact on the wages of the professional football players of the Italian Serie A. In particular, having a 30-day injury reduces the wage of the following year by approximately 12%. Secondly, only a residual part of the negative effect could be explained through the reduced performance of the players after the injury. In fact, results suggest that the largest part of the coefficient is explained by a direct effect of the shock on the outcome. This result can be explained in a framework of human capital depreciation. The club has an incentive to offer a lower wage for precautionary reasons, supposing that players who experienced a severe injury might incur similar injuries in the future. Thus, the club insures itself against this risk by reducing the fixed share of the wage, which is independent by performances.

The injury itself does not affect the probability of renegotiating the contract at the end of the season. However, different model specifications show that having a season-long injury increase the probability of renegotiating the contract or being traded to another club by about 10%. This result suggests that the individuals who suffer most from a shock are those in the last year of their contract. Finally, in line with the theoretical and empirical literature about super-earnings, performance and popularity are found to have played important roles in explaining the wages of the "working super-rich".

Overall, these findings also suggest that the working super-rich are negatively affected by health shocks in terms of reduced earnings. However, the mechanisms seem to be much more complex compared to the rest of the distribution for the peculiar role played by performance and popularity in the markets in which they are involved. Furthermore, the short length of the contracts represents an additional element of uncertainty. In fact, on one hand it represents an advantage for those players who increase their performance or popularity. On the other hand, it exposes them to the risk of obtaining reduced wages in case of a health shock.

Following the focus on the top tail of the distribution, future lines of research might investigate the labour market response to health shocks of other categories included in the top tail of labour income distribution, such as the CEOs, to analyse the results in a dynamic perspective and to assess how much of the health-related income inequalities might be explained by these mechanisms.

References

Adler, M. (1985). Stardom and talent. *The American Economic Review*, *75*(1), 208-212.

Alvaredo, F., Atkinson, A. B., Piketty, T., & Saez, E. (2013). The top 1 percent in international and historical perspective. *The Journal of Economic Perspectives*, *27*(3), 3-20.

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, *49*(1), 3-71.

Bound, J., Schoenbaum, M., Stinebrickner, T. R., & Waidmann, T. (1999). The dynamic effects of health on the labor force transitions of older workers. *Labour Economics*, *6*(2), 179-202.

Bryson, A., Frick, B., & Simmons, R. (2013). The returns to scarce talent: footedness and player remuneration in European soccer. *Journal of Sports Economics*, *14*(6), 606-628.

Cai, L. (2010). The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics*, *17*(1), 77-90.

Cai, L., Mavromaras, K., & Oguzoglu, U. (2014). The effects of health status and health shocks on hours worked. *Health Economics*, *23*(5), 516-528.

Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics using stata* (Vol. 2). College Station, TX: Stata press.

Deloitte (2017). Football money league. 20th Edition. *Sports Business Group*.

Disney, R., Emmerson, C., & Wakefield, M. (2006). Ill health and retirement in Britain: A panel data-based analysis. *Journal of Health Economics*, *25*(4), 621-649.

Franck, E., & Nüesch, S. (2012). Talent and/or popularity: what does it take to be a superstar?. *Economic Inquiry*, *50*(1), 202-216.

Franzini, M., Granaglia, E., & Raitano, M. (2016). *Extreme Inequalities in Contemporary Capitalism: Should We be Concerned about the Rich?*. Springer.

García-Gómez, P., & López Nicolás, Á. (2006). Health shocks, employment and income in the Spanish labour market. *Health Economics*, *15*(9), 997-1009.

García-Gómez, P., Van Kippersluis, H., O'Donnell, O., & Van Doorslaer, E. (2013). Long-term and spillover effects of health shocks on employment and income. *Journal of Human Resources*, *48*(4), 873-909.

Hanson, A., Jolly, N. A., & Peterson, J. (2017). Safety regulation in professional football: empirical evidence of intended and unintended consequences. *Journal of Health Economics*, *53*, 87-99.

Jones, A. M., Rice, N., & Roberts, J. (2010). Sick of work or too sick to work? Evidence on self-reported health shocks and early retirement from the BHPS. *Economic Modelling*, *27*(4), 866-880.

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

Lee, J., & Kim, H. (2008). A longitudinal analysis of the impact of health shocks on the wealth of elders. *Journal of Population Economics*, *21*(1), 217-230.

Leigh, A. (2007). How closely do top income shares track other measures of inequality?. *The Economic Journal*, *117*(524).

Lindeboom, M., & Kerkhofs, M. (2009). Health and work of the elderly: subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics*, *24*(6), 1024-1046.

Lucifora, C., & Simmons, R. (2003). Superstar effects in sport: Evidence from Italian soccer. *Journal of Sports Economics*, *4*(1), 35-55.

Mitra, S., Palmer, M., Mont, D., & Groce, N. (2016). Can households cope with health shocks in Vietnam?. *Health Economics*, *25*(7), 888-907.

Mueller-Wohlfahrt, H. W., Haensel, L., Mithoefer, K., Ekstrand, J., English, B., McNally, S., ... & Blottner, D. (2012). Terminology and classification of muscle injuries in sport: a consensus statement. *British Journal of Sports Medicine*, bjsports-2012.

Mullin, C. J., & Dunn, L. F. (2002). Using baseball card prices to measure star quality and monopsony. *Economic Inquiry*, *40*(4), 620-632.

Rosen, S. (1981). The economics of superstars. *The American Economic Review*, *71*(5), 845-858.

Staiger, D., & Stock, J. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica, 65*(3), 557-586.

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In Andrews, D.W.K., & Stock, J., (Eds.), Identification and Inference for Econometric Models, Essays in Honor of Thomas Rothenberg, 80-108. *New York: Cambridge University Press.*

Stoecker, C., Sanders, N. J., & Barreca, A. (2016). Success is something to sneeze at: Influenza mortality in cities that participate in the Super Bowl. *American Journal of Health Economics*.

Trevisan, E., & Zantomio, F. (2016). The impact of acute health shocks on the labour supply of older workers: Evidence from sixteen European countries. *Labour Economics*, *43*, 171-185.

Treme, J., & Allen, S. K. (2009). Widely received: Payoffs to player attributes in the NFL. *Economics Bulletin*, *29*(3), 1631-1643.

Treme, J., & Allen, S. K. (2011). Press pass: Payoffs to media exposure among National Football League (NFL) wide receivers. *Journal of Sports Economics*, *12*(3), 370-390.

Wagstaff, A. (2007). The economic consequences of health shocks: evidence from Vietnam. *Journal of Health Economics*, *26*(1), 82-100.

Wagstaff, A., & Lindelow, M. (2014). Are health shocks different? Evidence from a multishock survey in Laos. *Health Economics*, *23*(6), 706-718.

Wooldridge, J. M. (2002) Econometric Analysis of Cross Section and Panel Data. *The MIT Press 0*, *5*(1), 5.

Tables and Figures

Table 1. Sample characteristics

Variable	*Description*	*Mean (St. dev.)*
Dependent variable		
Wage	Net earnings (pre-season values) in thousands/€	875.1 (911.7)
Log wage	Log of net earnings (pre-season values)	6.38 (0.87)
Renegotiation	Dummy for contract's renegotiation	
Individual controls		
Age	Age (years)	26.6 (4.2)
Age square	Age squared	725.8 (226.7)
Position	Dummies for defenders (40.2%), midfielder (39.9%) and forward (19.9%)	
Captain	Dummy for the team's captain	0.034 (0.181)
Minutes played	Minutes played during the season	1352.8 (1068.3)
Total international caps	Number of caps with the national team up to 2014-2015	15.75 (25.10)
Total Under-21 caps	Number of caps with the U21 national team up to 2014-2015	5.85 (8.68)
International caps	Number of caps with the national team during the season	2.01 (4.65)
Under-21 caps	Number of caps with the U21 national team during the season	0.37 (1.59)
Player's performance		
Grade	Mean grade by newspapers during the season	5.77 (0.41)
Goal	Goals scored during the season	1.93 (3.50)
Assist	Assists served during the season	1.28 (2.04)
Index of popularity		
Popularity	Google search results (million)	4.21 (9.37)
Index of health shock		
Injury	Dummies for the kind of injury	
Days-off traumatic	Number of days off due to traumatic injury	11.08 (38.77)
Days-off muscular	Number of days off due to muscular injury	4.77 (12.90)
Re-occurrence	Dummy for the re-occurrence of the same kind of injury	4.79 (7.35)
Matches off	Total number of matches missed due to any injury	4.04 (7.12)
Instrumental variable	Player/Team	
Yellow cards	Number of teammates' avg. yellow cards	1.91 (0.48)

Table 2. Wage distribution: percentiles and Gini Index

	Total	Treatment	Control
Mean	985.9	1050.9	955.8
Standard Deviation	941.3	1025.1	898.9
Minimum	30	30	30
p10	300	300	300
p25	400	400	400
p50	650	700	600
p75	1200	1200	1100
p90	2100	2400	2100
p99	4500	4900	4200
Maximum	6500	5500	6500
Gini Index	0.474	0.453	0.435

Table 3. Within group characteristics

	Treatments				**Controls**				**Mean difference**	**p-value**[1]
	Mean	Standard Dev.	Min	Max	Mean	Standard Dev.	Min	Max		
Wage	1050.9	1025.1	30	5500	955.7	898.9	30	6500	95.2	0.11
Caps	22.49	8.95	1	38	24.5	9.51	0	38	2.01	0.005
Minutes	1651.9	814.0	8	3230	1811.6	904.9	0	3643	159.6	0.004
Age	27.9	4.1	19	40	27.3	3.8	18	39	0.6	0.01
Goal	2.56	4.14	0	28	2.52	3.68	0	29	0.04	0.84
Assist	1.72	2.29	0	12	1.67	2.15	0	14	0.05	0.72
Grade	5.76	0.43	3.33	6.75	5.78	0.39	3.67	6.7	0.02	0.33
Popularity	0.45	0.89	0.003	9.20	0.39	0.87	0.002	9.66	0.06	0.28

[1] p-value for the Student t-test of means comparison. H_0 diff=0; H_a diff≠0

Table 4. Effects of health shock and other covariates on annual net (log)wage. Fixed effects estimates.

	(1)	(2)	(3)	(4)
Age	0.8971***	0.9031***	0.8796***	0.8856***
	0.0639	*0.0640*	*0.0640*	*0.0641*
Age sq.	-0.0150***	-0.0151***	-0.0148***	-0.0149***
	0.0011	*0.0011*	*0.0011*	*0.0011*
Goal	0.0491***	0.0486**	0.0507***	0.0502***
	0.0190	*0.0190*	*0.0189*	*0.0189*
Assist	-0.0024	-0.0031	-0.0045	-0.0053
	0.0144	*0.0144*	*0.0144*	*0.0144*
Grade	0.0181	0.0188	0.0176	0.0184
	0.0149	*0.0149*	*0.0148*	*0.0148*
Popularity	0.0321**	0.0330**	0.0315**	0.0325**
	0.0160	*0.0160*	*0.0159*	*0.0159*
Days-Off traum.	-0.0250**	-0.0250**	-0.0283**	-0.0284**
	0.0118	*0.0118*	*0.0118*	*0.0118*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	1197	1197	1197	1197

***, **, * indicate significance at 1%, 5% and 10%, respectively

[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Robust Standard Errors in *italics*.

Table 5. Effects of health shock and other covariates on annual net (log)wage. Fixed effects estimates. Days off due to muscular injury as control.

	(1)	(2)	(3)	(4)
Age	0.8924***	0.8984***	0.8786***	0.8847***
	0.0642	*0.0643*	*0.0642*	*0.0643*
Age sq.	-0.0149***	-0.0150***	-0.0148***	-0.0149***
	0.0011	*0.0011*	*0.0011*	*0.0011*
Goal	0.0487**	0.0482**	0.0505***	0.0501***
	0.0190	*0.0190*	*0.0189*	*0.0189*
Assist	-0.0029	-0.0036	-0.0046	-0.0054
	0.0145	*0.0145*	*0.0144*	*0.0144*
Grade	0.0175	0.0183	0.0175	0.0182
	0.0149	*0.0149*	*0.0149*	*0.0149*
Pop	0.0321**	0.0331**	0.0315**	0.0325**
	0.0160	*0.0160*	*0.0159*	*0.0159*
Days-Off traum.	-0.0230*	-0.0230*	-0.0276**	-0.0277**
	0.0120	*0.0120*	*0.0121*	*0.0121*
Days-Off musc.	0.0007	0.0007	0.0002	0.0002
	0.0008	*0.0008*	*0.0008*	*0.0008*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	1197	1197	1197	1197

***, **, * indicate significance at 1%, 5% and 10%, respectively

[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Robust Standard Errors in *italics*.

Table 6. Effects of health shock and other covariates on annual net (log)wage. Fixed effects estimates. Days off due to muscular injury dropped.

	(1)	(2)	(3)	(4)
Age	0.8051***	0.8059***	0.8000***	0.8015***
	0.0761	*0.0765*	*0.0760*	*0.0764*
Age sq.	-0.0134***	-0.0134***	-0.0134***	-0.0134***
	0.0013	*0.0013*	*0.0013*	*0.0013*
Goal	0.0586**	0.0586**	0.0572**	0.0571**
	0.0248	*0.0249*	*0.0248*	*0.0248*
Assist	0.0097	0.0096	0.0070	0.0068
	0.0189	*0.0190*	*0.0189*	*0.0190*
Grade	0.0397**	0.0398**	0.0401**	0.0403**
	0.0191	*0.0192*	*0.0191*	*0.0191*
Popularity	0.0329	0.0331	0.0321	0.0324
	0.0204	*0.0205*	*0.0204*	*0.0205*
Days-off traum.	-0.0217*	-0.0218*	-0.0281**	-0.0283**
	0.0132	*0.0132*	*0.0136*	*0.0137*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	848	848	848	848

***, **, * indicate significance at 1%, 5% and 10%, respectively

[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Robust Standard Errors in *italics*.

Table 7. Effects of health shock and other covariates on annual net (log)wage.
Fixed effects-IV estimates.

	(1)	(2)	(3)	(4)
Days-off	-0.1211**	-0.1204**	-0.1258**	-0.1250**
	0.0563	*0.0558*	*0.0598*	*0.0592*
Age	0.9795***	0.9861***	0.9776***	0.9839***
	0.0856	*0.0857*	*0.0859*	*0.0859*
Age sq.	-0.0161***	-0.0161***	-0.0160***	-0.0161***
	0.0014	*0.0014*	*0.0014*	*0.0014*
Goal	0.0538**	0.0540**	0.0533**	0.0535**
	0.0214	*0.0214*	*0.0216*	*0.0216*
Assist	0.0043	0.0038	0.0033	0.0027
	0.0167	*0.0167*	*0.0169*	*0.0168*
Grade	0.0276	0.0286*	0.0282	0.0293*
	0.0169	*0.0169*	*0.0172*	*0.0172*
Popularity	0.0345*	0.0353*	0.0351*	0.0359*
	0.0183	*0.0183*	*0.0185*	*0.0185*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	1099	1099	1099	1099
First Stage F-test	*20.23*	*20.45*	*18.34*	*18.62*

***, **, * indicate significance at 1%, 5% and 10%, respectively
[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Standard Errors in *italics*.

Table 8. Effects of health shock and other covariates on annual net (log)wage.
Fixed effects-IV results. Performance not controlled for.

	(1)	(2)	(3)	(4)
Days-off	-0.1303**	-0.1309**	-0.1362**	-0.1367**
	0.0581	*0.0578*	*0.0620*	*0.0615*
Age	1.0168***	1.0246***	1.0141***	1.0217***
	0.0882	*0.0884*	*0.0885*	*0.0887*
Age sq.	-0.0167***	-0.0168***	-0.0167***	-0.0168***
	0.0014	*0.0014*	*0.0015*	*0.0015*
Popularity	0.0334*	0.0343*	0.0341*	0.0350*
	0.0186	*0.0186*	*0.0188*	*0.0188*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	1099	1099	1099	1099
First Stage F-test	*19.67*	*19.80*	*17.74*	*17.97*

***, **, * indicate significance at 1%, 5% and 10%, respectively
[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Standard Errors in *italics*.

Table 9. Effects of health shock and other covariates on renegotiation.
Binary choice models estimates.

	Linear probability model				Fixed effects logit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Days-off	0.079	0.079	0.054	0.057	1.082	1.083	1.056	1.059
	0.073	*0.073*	*0.075*	*0.075*	*1.08*	*1.09*	*0.72*	*0.76*
Age	0.052	0.046	0.042	0.036	1.053	1.047	1.043	1.036
	0.111	*0.103*	*0.099*	*0.100*	*0.47*	*0.45*	*0.42*	*0.36*
Age sq.	0.001	0.002	0.001	0.002	1.002	1.002	1.002	1.002
	0.001	*0.002*	*0.002*	*0.002*	*0.86)*	*0.97*	*1.01*	*1.05*
Caps	0.008*	0.009*	0.008*	0.008*	1.008*	1.009*	1.008*	1.008*
	0.005	*0.005*	*0.005*	*0.005*	*1.80*	*1.76*	*1.71*	*1.76*
Goal	-0.037	-0.037	-0.039	-0.039	0.963	0.963	0.961	0.961
	0.028	*0.027*	*0.026*	*0.026*	*-1.34*	*-1.41*	*-1.53*	*-1.52*
Assist	0.002	0.002	-0.003	-0.003	1.002	1.002	0.997	0.997
	0.022	*0.022*	*0.022*	*0.022*	*0.09*	*0.10*	*-0.15*	*-0.13*
Grade	0.020	0.020	0.024	0.023	1.021	1.020	1.024	1.023
	0.022	*0.021*	*0.021*	*0.021*	*0.93*	*0.93*	*1.13*	*1.09*
Pop.	-0.040*	-0.041*	-0.037*	-0.038*	0.960*	0.960*	0.963*	0.963*
	0.024	*0.021*	*0.020*	*0.020*	*-1.70*	*-1.94*	*-1.85*	*-1.87*
Team FE	No	Yes	No	Yes	No	Yes	No	Yes
Month FE	No	No	Yes	Yes	No	No	Yes	Yes
N	1099	1099	1099	1099	1099	1099	1099	1099

***, **, * indicate significance at 1%, 5% and 10%, respectively

[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Robust Standard Errors in *italics*.

Table 10. Effects of health shock and other covariates on renegotiation.
Binary choice models estimates. Performance not controlled for.

	Linear probability model				Fixed effects logit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Days-off	0.098**	0.104**	0.069	0.075	1.103*	1.109*	1.072	1.079
	0.046	*0.047*	*0.049*	*0.050*	*2.13*	*2.25*	*1.40*	*1.50*
Age	0.200**	0.189**	0.189**	0.179**	1.221*	1.208*	1.209*	1.196*
	0.083	*0.079*	*0.074*	*0.076*	*2.42*	*2.38*	*2.55*	*2.36*
Age sq.	-0.001	-0.001	-0.001	-0.001	0.999	0.999	0.999	0.999
	0.001	*0.001*	*0.001*	*0.001*	*-0.91*	*-0.84*	*-0.93*	*-0.80*
Caps	0.010***	0.010***	0.009***	0.010***	1.010***	1.010***	1.010***	1.010***
	0.002	*0.001*	*0.001*	*0.002*	*5.97*	*6.15*	*5.91*	*5.95*
Pop.	-0.031*	-0.032**	-0.028**	-0.029**	0.969*	0.969*	0.972*	0.972*
	0.017	*0.014*	*0.014*	*0.014*	*-1.86*	*-2.18*	*-2.01*	*-2.05*
TeamFE	No	Yes	No	Yes	No	Yes	No	No
Month Fe	No	No	Yes	Yes	No	No	Yes	No
N	1527	1527	1527	1527	1527	1527	1527	1527

***, **, * indicate significance at 1%, 5% and 10%, respectively

[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Robust Standard Errors in *italics*.

Table 11. Robustness check: effects of health shock and other covariates on annual net (log)wage. Alternative health shock's measure. Fixed effects-IV estimates.

	(1)	(2)	(3)	(4)
Matches-off	-0.0282**	-0.0294**	-0.0289**	-0.0301**
	0.0134	*0.0141*	*0.0140*	*0.0147*
Age	1.0003***	1.0117***	0.9996***	1.0109***
	0.0930	*0.0956*	*0.0933*	*0.0960*
Age sq.	-0.0162***	-0.0164***	-0.0162***	-0.0163***
	0.0015	*0.0015*	*0.0015*	*0.0015*
Goal	0.0472**	0.0467**	0.0467**	0.0462**
	0.0220	*0.0223*	*0.0222*	*0.0225*
Assist	0.0098	0.0096	0.0093	0.0090
	0.0177	*0.0179*	*0.0177*	*0.0179*
Grade	0.0355*	0.0369**	0.0361*	0.0376**
	0.0184	*0.0187*	*0.0187*	*0.0190*
Popularity	0.0313*	0.0322*	0.0316*	0.0326*
	0.0187	*0.0189*	*0.0188*	*0.0190*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	1099	1099	1099	1099
First Stage F-test	*19.08*	*17.59*	*17.72*	*16.56*

***, **, * indicate significance at 1%, 5% and 10%, respectively

[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Standard Errors in *italics*.

Table 12. Robustness check: effects of health shock and other covariates on annual net (log)wage. Lagged health shock's values included. Fixed effects estimates.

	(1)	(2)	(3)	(4)
Age	0.9356***	0.9371***	0.9253***	0.9280***
	0.1319	*0.1313*	*0.1335*	*0.1325*
Age sq.	-0.0155***	-0.0155***	-0.0153***	-0.0154***
	0.0022	*0.0022*	*0.0022*	*0.0022*
Goal	0.0451*	0.0453*	0.0449*	0.0454*
	0.0256	*0.0256*	*0.0251*	*0.0251*
Assist	0.0266	0.0265	0.0252	0.0250
	0.0172	*0.0170*	*0.0171*	*0.0170*
Grade	0.0446**	0.0446**	0.0449**	0.0449**
	0.0206	*0.0206*	*0.0204*	*0.0204*
Popularity	0.0583*	0.0584*	0.0565*	0.0567*
	0.0336	*0.0336*	*0.0330*	*0.0331*
Days-off	-0.0333*	-0.0331*	-0.0348*	-0.0346*
	0.0186	*0.0185*	*0.0187*	*0.0186*
Days-off$_{t-2}$	-0.0220	-0.0220	-0.0201	-0.0201
	0.0215	*0.0215*	*0.0215*	*0.0213*
Days-off$_{t-3}$	-0.0365	-0.0364	-0.0351	-0.0349
	0.0242	*0.0242*	*0.0241*	*0.0241*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	664	664	664	664

***, **, * indicate significance at 1%, 5% and 10%, respectively

[a] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [b] Standard Errors in *italics*.

Table 13. Robustness check: Effects of health shock and other covariates on annual net (log)wage. Injury-purged variables included. Fixed effects estimates.

	(1)	(2)	(3)	(4)
Days-off	-0.1218**	-0.1207**	-0.1258**	-0.1246**
	0.0544	*0.0536*	*0.0574*	*0.0564*
Age	0.9777***	0.9841***	0.9755***	0.9817***
	0.0818	*0.0817*	*0.0819*	*0.0817*
Age sq.	-0.0159***	-0.0160***	-0.0159***	-0.0160***
	0.0013	*0.0013*	*0.0014*	*0.0014*
Goal_purged	0.1150***	0.1150***	0.1149***	0.1148***
	0.0324	*0.0322*	*0.0326*	*0.0323*
Assist_purged	0.0826***	0.0820***	0.0821***	0.0815***
	0.0297	*0.0295*	*0.0297*	*0.0295*
Grade_purged	0.0467**	0.0476**	0.0478**	0.0486**
	0.0202	*0.0201*	*0.0208*	*0.0206*
Popularity	0.0337*	0.0346*	0.0342*	0.0351*
	0.0180	*0.0179*	*0.0181*	*0.0181*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	1099	1099	1099	1099
First Stage F-test	*21.46*	*21.98*	*19.77*	*20.34*

***, **, * indicate significance at 1%, 5% and 10%, respectively

[a] Covariates concerning purged Goal, Assist and Grade refers to OLS residuals of the covariate regression on days off, age, age sq. and season fixed effects. [b] The model includes the full set of controls: dummy for the team captain; number of caps in the league during the season; number of senior national team caps during the season; number of national under-21 team caps during the season; dummy for the reoccurrence of the same injury; season fixed effects. [c] Standard Errors in *italics.*

Figure 1. Kernel density estimate of annual net wages

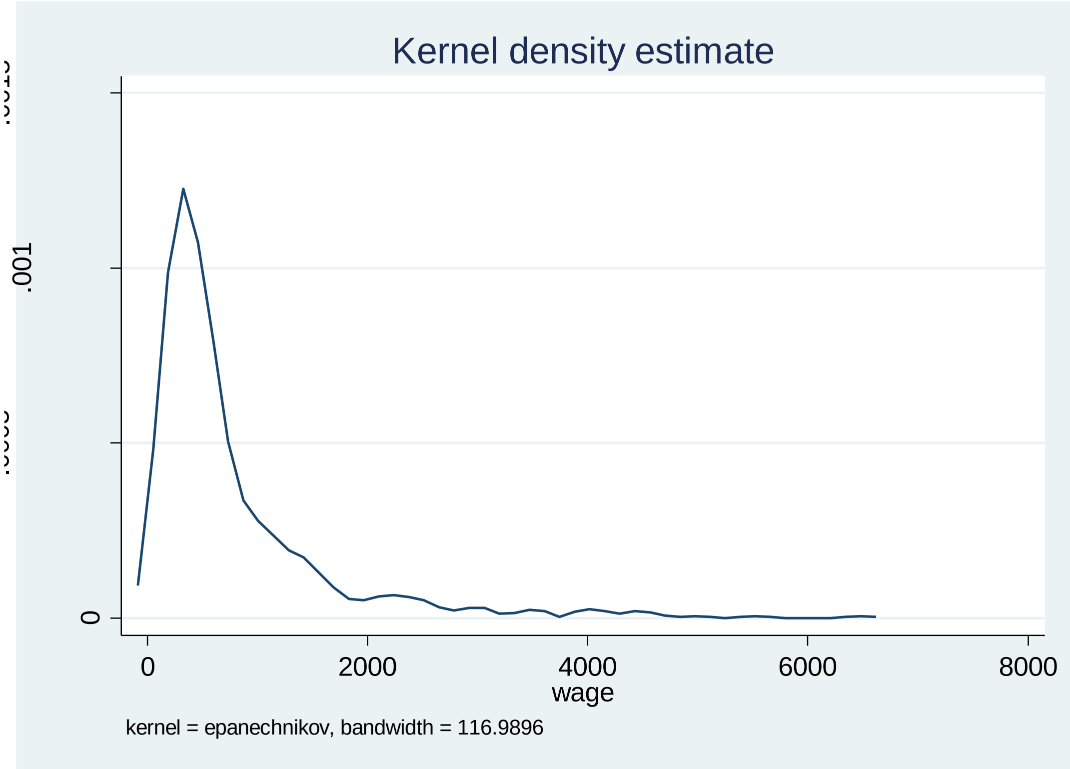
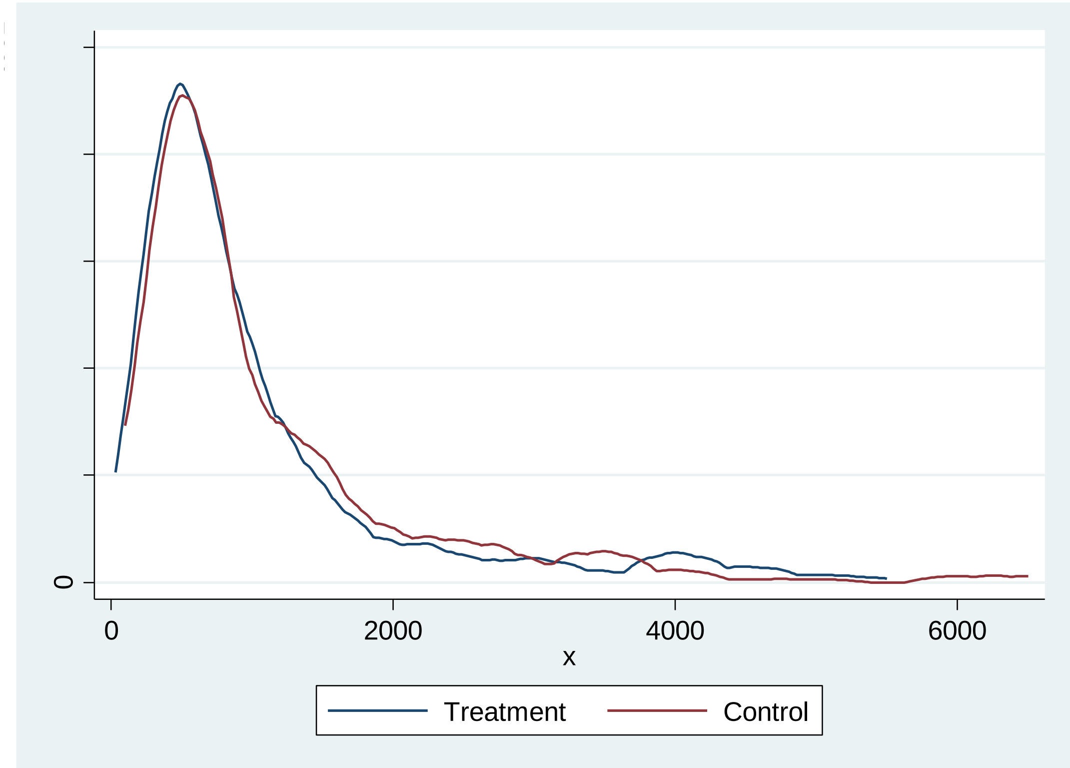


Figure 2. Kernel density estimate of annual net wages: groups comparison



Appendix

Figure A.1 Injuries percentage by month

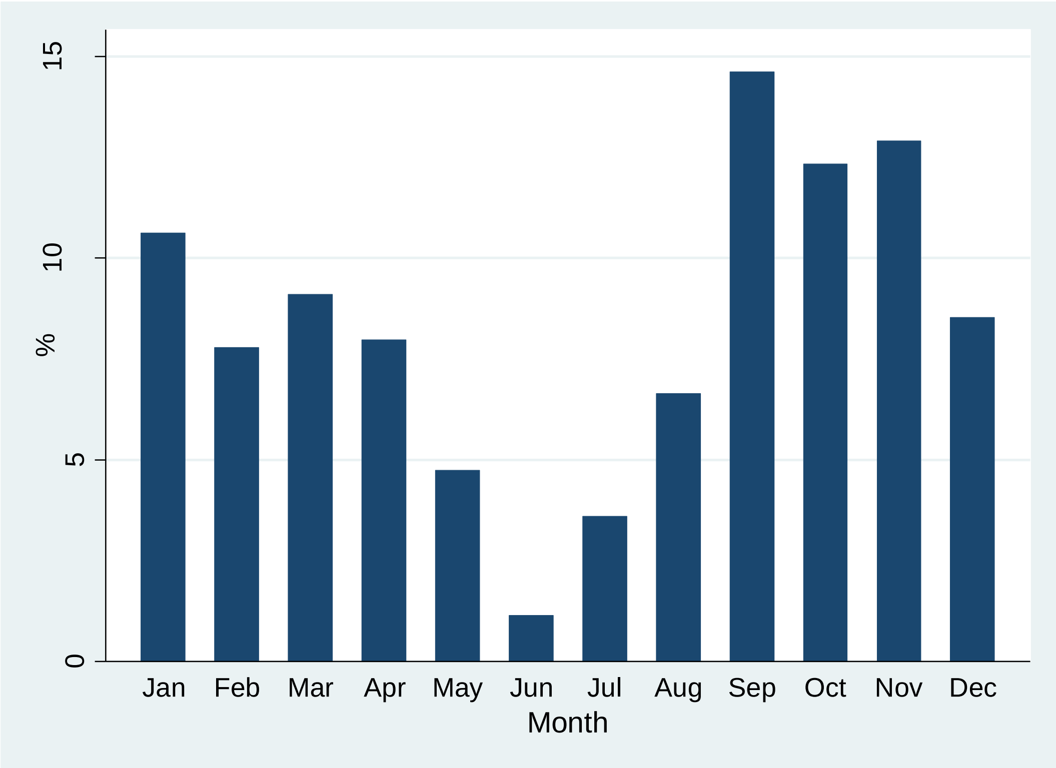


Table A.1 Teams' characteristics. Summary statistics and percentile distribution.

	Wage expense	Net Sales	EBT	Ticketing	Television
Mean	19,656	90,720	-8,630	10,838	53,341
Standard Deviation	13,519	68,682	29,247	10,511	35,240
min	6,9	22,450	-93,767	1,516	7,610
p10	7,85	36,880	-45,919	2,213	25,164
p25	9,1	42,834	-14,040	3880	29,870
p50	13	56,312	-1,745	5,014	34,499
p75	28	116,446	3,636	15,134	70,744
p90	42,95	212,419	12,438	31,017	115,010
p99	47,6	272,404	84,582	38,051	163,478
max	47,6	272,404	84,582	38,051	163,478

[a] Values expressed in thousands of Euros

Table A.2 First stage regressions.

	(1) Days-off	(2) Days-off	(3) Days-off	(4) Days-off
Age	0.7763***	0.7770***	0.7214**	0.7220**
	0.2963	*0.2997*	*0.2947*	*0.2981*
Age sq.	-0.0100**	-0.0100*	-0.0093*	-0.0093*
	0.0051	*0.0051*	*0.0050*	*0.0051*
Under 21	-0.0252	-0.0272	-0.0256	-0.0278
	0.0472	*0.0477*	*0.0467*	*0.0472*
International	0.0096	0.0098	0.0100	0.0101
	0.0157	*0.0159*	*0.0154*	*0.0156*
Captain	-0.1979	-0.1936	-0.2266	-0.2223
	0.3099	*0.3096*	*0.3149*	*0.3145*
Caps	-0.0610***	-0.0613***	-0.0587***	-0.0590***
	0.0083	*0.0084*	*0.0087*	*0.0087*
Goal	0.0113	0.0129	0.0063	0.0079
	0.0851	*0.0851*	*0.0848*	*0.0849*
Assist	0.0385	0.0389	0.0269	0.0271
	0.0629	*0.0632*	*0.0637*	*0.0640*
Grade	0.0749	0.0760	0.0774	0.0787
	0.0793	*0.0795*	*0.0786*	*0.0788*
Popularity	0.0356	0.0355	0.0400	0.0398
	0.0690	*0.0693*	*0.0690*	*0.0694*
Reoccurrence	0.2779	0.2741	0.2859	0.2819
	0.2080	*0.2088*	*0.2111*	*0.2119*
Yellow cards	0.5388***	0.5510***	0.5140***	0.5266***
	0.1595	*0.1585*	*0.1572*	*0.1563*
Team FE	No	Yes	No	Yes
Month FE	No	No	Yes	Yes
N	1099	1099	1099	1099
F-statistic	*20.23*	*20.45*	*18.34*	*18.62*

***, **, * indicate significance at 1%, 5% and 10%, respectively