I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17272

# High Temperatures and Workplace Injuries

Matteo Picchio
Jan C. van Ours

SEPTEMBER 2024

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17272

# High Temperatures and Workplace Injuries

**Matteo Picchio**
*Marche Polytechnic University, Ghent University, IZA and GLO*

**Jan C. van Ours**
*Erasmus School of Economics and Tinbergen Institute, IZA and CEPR*

SEPTEMBER 2024

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## High Temperatures and Workplace Injuries*

High temperatures can have a negative effect on workplace safety for a variety of reasons. Discomfort and reduced concentration caused by heat can lead to workers making mistakes and injuring themselves. Discomfort can also be an incentive for workers to report an injury that they would not have reported in the absence of heat. We investigate how temperature affects injuries of professional tennis players in outdoor singles matches. We find that for men injury rates increase with ambient temperatures. For women, there is no effect of high temperatures on injuries. Among male tennis players, there is some heterogeneity in the temperature effects, which seem to be influenced by incentives. Specifically, when a male player is losing at the beginning of a crucial (second) fourth set in (best-of-three) best-of-five matches, the temperature effect is much larger than when he is winning. In best-offive matches, which are more exhausting, this effect is age-dependent and stronger for older players.

**JEL Classification:** J24, J81, Q51, Q54

**Keywords:** climate change, temperatures, tennis, injuries, health

**Corresponding author:**
Matteo Picchio
Department of Economics and Social Sciences
Marche Polytechnic University
Piazzale Martelli 8
60121 Ancona
Italy
E-mail: m.picchio@staff.univpm.it

* We acknowledge the ERA5-Land dataset and Global Reanalysis (EAC4) dataset from the EU-funded Copernicus Climate Change Service and Atmosphere Monitoring Service. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus Information or Data it contains. Declaration of interest: none.

# 1 Introduction

Workplace injuries are not very common, but they still occur regularly. Every year, approximately 1 to 3 percent of the workforce is affected by occupational injuries.[1] The frequency of workplace injuries depends on the nature of the work, the work experience of individual workers, environmental variables including the weather, etc. With global warming, an important question is whether higher ambient temperatures lead to more workplace injuries. To the extent that the effect of temperature on work-related activities is studied, the focus is on labor productivity and working time.[2] Nevertheless, the relationship between high temperatures and injuries is simple. High temperatures may increase exhaustion and fatigue and with fatigue injuries are likely to occur more often.

We study the relationship between temperature and workplace injuries using data from professional tennis. An injury in a tennis match differs from regular workplace injuries, but there are also clear overlaps. The differences are that tennis players participate in high-intensity activities over a limited period of time. Their injuries are caused by the physical demands and movements involved in tennis, while regular workplace injuries may also be due to factors related to the equipment used in the work environment. There are also overlaps. Serious workplace injuries always lead workers to stop working, perhaps for several days or longer. Serious tennis-related injuries lead players to abandon the game, which in tennis is called "retiring". Small workplace injuries may go unnoticed as they do not cause a loss or substantial reduction in productivity. Small tennis injuries may be uncomfortable for the tennis player, but do not necessarily prevent them from playing or perhaps even winning a match. Between small and serious injuries, it is not immediately clear what will happen. Workers may decide to report the injury and stop working, or they may continue to work. The choice that workers make will depend on the consequences of that choice. Boone and van Ours (2006) argued that, whether or not a worker reports a small injury, may depend on the state of the labor market. If unemployment is high, they may fear future job loss if their employer holds them responsible for their injury. Losing a job in a slack labor market most likely means that it will take more time to find a new

[1] According to ILO-statistics, in 2020, in the US the annual non-fatal occupational injury rate was 1.2% for male workers and 1.4% for female workers. In 2022, this was 0.8% for female workers and 2.1% for male workers in Germany and 2.0% for female workers and 3.7% for male workers in France. Cross-country differences cannot be compared directly as the way the data were collected may differ. The US data are from a survey of occupational injuries, the German data are from insurance records and the French data are from the health insurance fund.

[2] See Dell et al. (2014) for a general overview of the so called "new weather economy" literature.

job. In a tight labor market, the consequences of reporting a workplace injury may be less severe. Therefore, according to Boone and van Ours (2006), in a tight labor market, the number of reported workplace accidents goes up. They found some evidence of this using cross-country time series information on workplace accidents. Boone et al. (2011) analyzed Austrian data on individual workers and firms, confirming this line of reasoning. Probst et al. (2013) found that in Italy accident under-reporting is more relevant when workers' perception of job insecurity is larger and Picchio and van Ours (2017) suggested that in Italian firms, workers on temporary jobs may be less likely than workers with permanent contracts to report a small workplace accident in fear that their contract will not be extended. Therefore, reporting an accident at work is determined not only by physical or mental damage, but also by incentives related to the consequences of reporting. Here too, there is an overlap with tennis injuries. Although tennis players may not fear the consequences of retiring from a match in terms of their position in the labor market, they may balance marginal costs and marginal benefits of continuing to play, which implies that also retiring may be influenced by incentives.

The main reason why it is interesting to study the impact of temperature on injuries using tennis data is that it is possible to investigate how incentives and timing affect injury reporting. In tennis matches, 'retiring' due to an injury may not be driven solely by physical discomfort, but also by incentives and by the moment in a match or during the tournament when the injury occurs. The incentives may indeed differ within a match between sets. If a player starts in a set with a negative score from previous sets, (s)he may be more likely to retire when temperatures are high, as the balance between marginal benefits in terms of winning and marginal costs in terms of efforts may turn negative when it is too hot. Furthermore, depending on the severity of an injury, a player may be more likely to quit playing if it is early in the match or early in the tournament, or if the stakes are not high. However, if the match is in its final games or the match is late in the tournament and the stakes are higher, the player may decide to continue playing.

A reason to study the impact of ambient temperature on tennis injuries is that tennis is a very popular game. Almost two million tennis players belong to the International Tennis Federation (ITF) and almost 90 million people play tennis in the 41 countries surveyed by the ITF (ITF, 2021). However, our findings also have external validity beyond sport itself. They may have broader public health implications for outdoor high-stake activities which combine, as in tennis, mental skills, physical strength, technical proficiency, and fine motor control to be successful. Some examples, to name a few, are army special forces, fire-fighters, emergency doctors, and professional performers like dancers. This is especially relevant when such outdoor activities are performed in situations characterized by relevant temperature variations and heat waves. Tennis data are informative about the role of incentives around heat-induced potential injuries. This insight contributes to the broader literature on physiological adaptation to extreme heat conditions. Potentially, it opens up new explorations of the trade-off between the long-term economic costs of reporting an injury and the immediate mental calculations and exertion of physical effort exacerbated by heat.

The setup of our paper is as follows. In Section 2 we present an overview of previous studies on the effects of high temperatures on regular workplace injuries and, more specifically, on tennis sports injuries. Section 3 illustrates our data sources, discusses the relationship between temperature and tennis injuries and describes the methodology of the statistical analysis. Section 4 reports and comments on the main findings. Section 5 concludes.

# 2 Previous studies on temperature and workplace injuries

## 2.1 Regular workplace injuries

Otte im Kampe et al. (2016) gave an overview of studies on how high temperatures affect injuries, not only related to the workplace, but also related to recreational, sports, and traffic. Injuries increase with high temperatures across the board. Varghese et al. (2018) focused their overview of previous studies on the relationship between high temperatures and workplace injuries. From this, they concluded that the positive effect of hot weather on injuries is due to fatigue, reduced psychomotor performance, loss of concentration and reduced alertness. According to Spector et al. (2019), workers in industries at high risk for workplace injuries are often also at high risk for heat stress (agriculture, construction). They mentioned as potential mechanisms for the positive effect of temperature on injuries: impaired balance, changes in safety behavior, muscle fatigue, dehydration, poor sleep or sleepiness. They also pointed out that complex tasks are more susceptible to high temperatures. Morrissey et al. (2021) mentioned that there are several physical (reduced physical capacity), physiological (increased core temperature, sweat rate, heart rate), and psychological (increased discomfort, reduced concentration) responses that increase the risk of workplace injuries.

Dillender (2021) and Park et al. (2021) for the US, Filomena and Picchio (2024) for Italy, and Ireland et al. (2023) for Australia found that workplace injury rates significantly increase with temperature. Dillender (2021) studied US workplace data focusing on both low and high temperatures finding that injury rates go up at both sides of the temperature distribution. For high temperatures, the main conclusion is that is easier to avoid working in high temperatures when these are rare. It is not so easy to do this when high temperatures are common. Park et al. (2021) also studied US workplace injury data. They found that high temperatures have a positive effect on workplace injuries, both for outdoor and indoor activities and for injury types that are not directly related to high temperatures – such as falling from heights. The authors attributed this to high temperatures negatively affecting cognitive performance. Filomena and Picchio (2024) studied workplace accident data for Italy finding that higher temperatures cause injury rates to increase, in particular for male workers. This effect is present for less severe injuries; fatal work accidents, for example, are not affected by temperature. Ireland et al. (2023) studied the relationship between high outside temperatures and occupational health claims in the Australian state of Victoria. Work accidents increase at high temperatures. Focusing on high-risk occupations – manual workers in outdoor industries – they found that this is mainly because heat has a negative effect on concentration levels of workers.

## 2.2 Tennis sports injuries

A tennis match is recorded as retired when a match was discontinued due to illness or injury of one of the players after the match began. In search for determinants of retirement Breznik and Batagelj (2012) focused on matches of male professional tennis players studying data over the time period 1968 to 2010. Contrary to their expectations, they found that on clay courts the probability of a retired match was highest (2.8%). The second highest probability was for hard courts (2.6%). On grass, the probability of retirement was lowest (1.6%) but only about 5% of the matches was played on grass. Breznik and Batagelj (2012) also found that the proportion of retired matches was not lower at later stages of the tournament. They did find that in Grand Slams the probability of retirement was relatively high (2.4%) which was attributed to 'best-of-five' matches being played in these tournaments. Finally, they found that indoor matches have a lower retirement rate than outdoor matches (2.1 against 2.7%).

Smith et al. (2018a) analyzed 346 female tennis matches in the first four rounds of the Australian Open of 2014 to 2016. Overall, 1.4% of these matches were retired with none of them related to heat stress. Smith et al. (2018b) analyzed 340 male tennis matches in the first four rounds of the Australian Open of 2014 to 2016. Overall, 5.5% of these matches were retired (0.9% heat stress related). Of the 30 matches played with temperatures above 32.3°C four matches (13.3%) were retired of which two were due to heat stress.

As we will discuss later on, we conducted our empirical analysis separately for men and women since there may be gender-specific differences in the way tennis matches are played.[3] The differences between female and male tennis play may imply that ambient temperature may affect injuries differently for males and females and justifies to analyze them separately. First, all female tennis matches are best-of-three, while male matches can be best-of-five, with a relevant impact on its duration, its expected duration and, therefore, on playing behavior.[4] Second, rallies in women single matches are significantly longer than those in men single matches, essentially due to the fact that men tend to serve and volley more than women (Fernandez et al., 2006). This difference may be reflected in different physical exercise and effort and, in extreme ambient conditions, may generate a different likelihood of injuries. Third, previous research showed that injury occurrence may differ by sex, although previous empirical evidence is not clear-cut: Hartwell et al. (2017), using data from United States Tennis Association (USTA) Pro Circuit tournaments during 2013, and Sell et al. (2014), employing information on US Open matches, found that men were subject to a higher injury rate than women; Gescheit et al. (2019) found no difference in male and female injury incidence among Australian junior players; Gescheit et al. (2017) and McCurdie et al. (2017) detected a higher injury rate for women compared to men at 2011-2016 Australian Open and 2003-2012 Wimbledon, respectively. It is therefore interesting to dig deeper and understand if men and women cope differently with high temperatures.

Burke et al. (2023), like we, studied the relationship between temperature and retirement in tennis matches. They analyzed data from almost 180,000 professional tennis

[3]Gender differences are not always detected in tennis matches. Krumer et al. (2016) found that in the final score tennis matches between women were less tight than matches between men. However, when controlling for physiological differences in height and Body Mass Index this difference disappeared. From an analysis of stroke-by-stroke data, Paserman (2023) concluded that tennis players change their behavior in high pressure situations, i.e. they played less aggressive. This response was not gender-specific.

[4]Since 2008, only ATP Grand Slam matches are best-of-five. Up to 2007 included also the finals of some tournaments, for example Barcelona, Hamburg, Indian Wells, Miami, Monte Carlo and Rome, were best-of-five.

matches played during the period 2002-2017 with an average retirement rate of 3.1%. They found that the probability of match retirement increased by 2.1% per degree Celsius. They found no statistical different effect of temperature on the retirement rates of male and female tennis players.

# 3 Methods

## 3.1 On tennis injuries and temperature

This section presents a simple model to formalize the relationship between the probability of a player retiring, ambient temperature and incentives to keep playing. The rate by which players face a serious injury and they retire is the product of the injury arrival rate and the probability that the severity of an injury is above a certain threshold, i.e., a reservation severity below which a player continues the match and above which a player stops playing and retires. The injury arrival rate may depend on personal characteristics (age, experience), tournament characteristics (surface) and by characteristics of the opponent (strong/weak). The probability that an injury leads to retirement depends on the severity of the injury and the threshold which, in itself, depends on incentives.

The reduced form presentation of our model is as follows:

$$y_{im} = y(s(temp), e(temp), I), \tag{1}$$

where $y_{ism}$ is the probability that player $i$ retires because of an injury in match $m$, $s$ is the severity of an injury, $e$ represents effort, $I$ indicates incentives to continue playing, and $temp$ is ambient temperature. Whether a player decides to retire because of an injury depends on how serious that injury is initially and how that severity evolves during the match. The actual severity of the injury is not fully observed by the audience and the opponent player. The severity of an injury can range from a player not feeling very comfortable about some part of his or her body, for example a somewhat strained muscle in arm or leg. When an injury is serious, the player has to retire because it would not make sense to continue playing as the player would for sure lose game, set and match. When an injury is less serious, the player has to balance the expected costs and expected benefits of continuing to play.[5] The expected costs of retiring depend on the balance in

[5] Even if there is no injury at all, the player may decide to retire because the expected costs of continuing

(revenues or) utility of continuing to play or retiring on the spot. The expected utility of play continuation depends on the effort needed to keep performing and the possible benefits of winning the current game, set and match. Conditional on match win, possible wins of future matches in the tournament will be taken into account. The effect of temperature materializes in the effort needed to keep performing. Higher temperatures increase that effort. The possible benefits within a match depend on the stage of the tournament. Earlier in a tournament expected benefits will be low, later on in the tournament expected benefits will be high when there is a positive probability to win the tournament.

It may be that high temperatures affect the occurrence and severity of an injury. If high temperatures negatively affect focus and concentration of tennis players, heat may increase both incidence and severity of injuries and thus may have a positive effect on retirement. Incentives-driven retirement are not necessarily influenced by high ambient temperatures but, in combination with required effort, they may be. Thus, observed retirement is a combination of injury-driven retirement and incentive-driven retirement. As the true severity of an injury is not observed, we can only establish incentive-driven retirement by considering how high temperatures affect retirement under different incentives.

Incentive-driven retirement may materialize at the match level and within a match at the set level. At the match level, incentives to continue will be higher at a later stage of the tournament, the type of tournament (Grand Slam and others) or when playing at home. As we describe in more detail in the next section, we do not observe the elapsed time since the beginning of the match, but we do know in which set it occurred. At the set level, incentives to continue will be higher at a later stage of the match or when at the start of the set the player has an advantageous set score. In the latter case, winning the set may imply winning the match. If a player is behind in the set score, winning the set will not be enough to win the match. To do this, additional sets need to be won. The number of sets in a match may also be important for the probability of an injury occurring. If injury rates go up with temperature, an injury is more likely to materialize if a match lasts longer. If so, high temperatures have a bigger effect on injuries if a match is played over many sets rather than a few.

---

to play are higher than the expected benefits.

## 3.2 Data and sample

Our study relies on various data sources. Meteorological data come from the Copernicus Climate Change Service, which is supported by the Earth Observation Program funded by the European Union. Specifically, we utilized ERA5-Land (Muñoz-Sabater et al., 2021), a global land surface dataset spanning from 1950 to the present.[6] The dataset offers detailed granularity, providing information for each hour and day, with a horizontal resolution of $0.1° \times 0.1°$ in latitude/longitude coordinates, equivalent to approximately 9 $km^2$. As our tennis match datasets lack information regarding match times, we extracted daily temperatures recorded at 3 pm, two meters above the surface, to approximate temperatures experienced during both early and late matches on tournament days. This approach was adopted because, in ATP and WTA tournaments, only the first match of the day typically has a specified start time, while subsequent matches begin "followed by but not before" a certain time, with finals often commencing around 3 pm.

Additionally, we acquired daily data from the global greenhouse gas reanalysis (EGG4) of the Copernicus Atmosphere Monitoring Service (Inness et al., 2019). This dataset includes wind speed, particulate matter 2.5 ($PM_{2.5}$), ozone ($O_3$), and dew point temperature at the surface, recorded at 3 pm from 2003 to 2021. The horizontal resolution of this dataset is $0.75° \times 0.75°$. Given the potential influence of temperature on injury rates, which can vary depending on the combination of temperature and humidity, we conducted a sensitivity analysis using the Humidex index calculated as in Blazejczyk et al. (2012). This index, derived from ambient temperature and dew point temperature, provides insight into how hot weather is perceived by the average person.

We matched the temperature dataset with singles tennis matches sourced from two distinct repositories. Extracting match results from *Tennis-Data.co.uk*, we compiled ATP (WTA) tour season data spanning from 2003 (2007) to 2021.[7] During these years, several match details were accessible, including the match date, location, tournament series (e.g., Grand Slam, ATP/WTA 1000, etc.), surface type, indoor or outdoor setting, match round (e.g., final, semifinal, etc.), and for men, whether the match was best-of-three or best-of-five sets. Additional information encompassed the names of the winner and loser, set and game scores, players' ATP/WTA rankings and points before the tournament's

[6]For more details see https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?-tab=overview (last accessed March 14th, 2024).

[7]The dataset is available for download at http://www.tennis-data.co.uk/alldata.php (last accessed on March 15th, 2024).

commencement, and details regarding match completion or retirement by either player. After the exclusion of canceled matches,[8] our dataset comprised 84,895 matches, of which 72,385 were contested outdoors.

While the data obtained from *Tennis-Data.co.uk* provides match dates, it lacks player characteristics such as age and nationality, and the set scores are not always clearly reported. These details are particularly relevant for investigating the variability of effects across player attributes and match stages. To address this shortfall, we supplemented our dataset with tennis data sourced from Jeff Sackmann's GitHub repository available at https://github.com/JeffSackmann.[9] While Sackmann's data offer more comprehensive match and player statistics compared to *Tennis-Data.co.uk*, they do not consistently include the precise match date, providing only the tournament start date. Consequently, direct integration of Sackmann's data with the daily meteorological and air pollution datasets was unfeasible. To circumvent this issue, we merged the two tennis databases using common match-level variables present in both datasets, including players' surnames, ATP rankings and points, game and set scores, surface type, and tournament round. However, not all observations perfectly aligned during the merging process, resulting in a reduced sample size of 81,204 matches, of which 69,348 were played outdoors. After the exclusion of indoor matches and the removal of observations falling within the first or last percentile of the temperature distribution (1,366 observations),[10] as well as matches with missing values for regression analysis variables, our final dataset comprised 67,728 matches, translating to 135,456 player-level observations (78,946 male and 56,510 female player-level observations).

Table 1 reports summary statistics of the variables used in multivariate statistical analysis by sex. The injury rates of men and women were about the same, approximately 1.5%. The average temperature was 20.5°C. The weather and air quality statistics are very much the same for men and women. About 40% of our observations are on female matches; of the male matches almost 25% are best-of-five. The other sample character-

[8]Some matches were canceled due to pre-match withdrawals (*walkover*), comprising approximately 0.5% of the total matches, or suspensions resulting from player disqualification (*default*), which occurred in only 2 instances.

[9]For further information about Jeff Sackmann, refer to https://www.jeffsackmann.com/ (last accessed on March 14th, 2024) for his biography or https://www.tennisabstract.com/ (last accessed on March 14th, 2024) for his website containing professional tennis results and statistics.

[10]Given the use of linear regression models, we excluded the first and last percentiles of the temperature distribution to prevent outlying observations from disproportionately affecting the estimation of the regression coefficients and reducing the robustness of our findings (Chatterjee and Hadi, 1986).

istics of men and women are similar, with two exceptions: the tournament series and the surface of the court. While 25% of the matches is from a Grand Slam tournament for both men and women, about 40% of the male matches are from ATP 250 and for the female matches only 4% are from WTA 250. As to the surface of the court, about 50% of the male matches was on hard court, while for the female matches this was almost 60%. Around 40% of the male matches was on clay courts, while about 30% of the female matches was on clay courts.

The unconditional relationship between temperatures and injuries is presented in three graphs in Figure 1. These graphs show the smoothed values of kernel-weighted local first-order polynomial regression of tennis injury rates on temperatures by gender and, for men, by best-of-five and best-of-three matches. For men playing best-of-five matches, there was an evident increase in injury rates with temperature: from about 2% with temperatures between 12°C and 26°C to 4% and more with temperatures hotter than 28°C. When looking at best-of-three matches, men displays a slight increase in injury rates, from about 1–1.5% to 2% with temperatures above 30°C. For women, the relation between temperatures and injury rates seems absent.

## 3.3 Modeling tennis injuries

Our empirical analysis is at the match level and our modeling strategy is identical to the one used in Picchio and van Ours (2024), who exploited within tournament-edition[11] variation in temperature for identification. We use a linear probability model for the injury indicator $y_{ijmte}$ of player $i$, playing against player $j$, on match date $m$ of tournament $t$ in edition $e$, which we estimate separately for males and females:

$$y_{ijmte} = \boldsymbol{\alpha} temp_{mte} + \boldsymbol{\beta} \mathbf{x}_{ijmte} + \gamma_{te} + \delta_i + \eta_j + \varepsilon_{ijmte}, \tag{2}$$

where $\gamma_{te}$ are tournament-edition fixed effects (FE), obtained by the interaction among tournament indicators and edition indicators; $\delta_i$ are player FE; $\eta_j$ are opponent FE; $\mathbf{x}_{ijmte}$ is the set of explanatory variables shown in Table 1;[12] $temp_{mte}$ is the outdoor temperature

[11] Tournament-edition almost perfectly overlaps tournament-year. The difference is due to the fact that some tournaments, for example the ATP Hamburg tournament, were played twice in the same calendar year due to scheduling changes and tournament relocation in the ATP/WTA calendar.

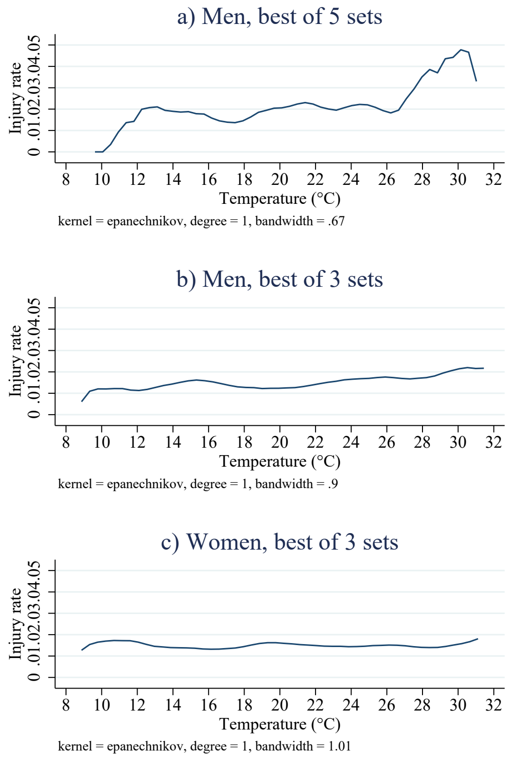
[12] When we estimate the tennis injury equation in its most general specification as described by Equation (2), we do not include the tournament series and surface, because of collinearity with the tournament-edition FE.

**Table 1:** Summary statistics

| | Men | | Women | |
|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. |
| Injury indicator | 0.016 | 0.126 | 0.015 | 0.121 |
| Temperature at 3pm (°C) | 20.698 | 4.608 | 20.327 | 4.651 |
| Wind speed at 3 pm (m/sec) | 3.182 | 1.861 | 3.165 | 1.907 |
| $PM_{2.5}$ at 3 pm (mg/kg) | 0.022 | 0.034 | 0.029 | 0.051 |
| $O_3$ at 3 pm (mg/kg) | 0.064 | 0.029 | 0.058 | 0.031 |
| Best of 5 sets | 0.240 | 0.427 | – | – |
| Age (years at tournament start) | 26.716 | 3.868 | 25.041 | 3.885 |
| \|Difference in players' ATP/WTA ranking\| | 72.016 | 111.290 | 66.166 | 95.644 |
| Sum in player's ATP/WTA ranking | 150.811 | 142.652 | 145.115 | 129.375 |
| Playing home[(a)] | 0.131 | 0.338 | 0.102 | 0.303 |
| *Round* | | | | |
| Less than quarterfinal | 0.843 | 0.364 | 0.851 | 0.357 |
| Quarterfinal | 0.090 | 0.287 | 0.086 | 0.280 |
| Semifinal | 0.044 | 0.206 | 0.043 | 0.202 |
| Final | 0.022 | 0.147 | 0.021 | 0.143 |
| *Tournament series*[(b)] | | | | |
| Grand Slam[(c)] | 0.240 | 0.427 | 0.239 | 0.426 |
| ATP/WTA 1000 | 0.226 | 0.418 | 0.390 | 0.488 |
| ATP/WTA 500 | 0.135 | 0.342 | 0.333 | 0.471 |
| ATP/WTA 250 | 0.399 | 0.490 | 0.038 | 0.191 |
| *Surface* | | | | |
| Grass | 0.138 | 0.344 | 0.121 | 0.326 |
| Clay | 0.387 | 0.487 | 0.304 | 0.460 |
| Hard | 0.475 | 0.499 | 0.575 | 0.494 |
| *Quarter* | | | | |
| January-February-March | 0.284 | 0.451 | 0.297 | 0.457 |
| April-May-June | 0.354 | 0.478 | 0.337 | 0.473 |
| July-August-September | 0.316 | 0.465 | 0.326 | 0.469 |
| October-November-December | 0.046 | 0.209 | 0.040 | 0.196 |
| *Day of the week* | | | | |
| Monday | 0.056 | 0.230 | 0.070 | 0.255 |
| Tuesday | 0.183 | 0.387 | 0.196 | 0.397 |
| Wednesday | 0.250 | 0.433 | 0.244 | 0.429 |
| Thursday | 0.186 | 0.389 | 0.182 | 0.386 |
| Friday | 0.147 | 0.354 | 0.143 | 0.350 |
| Saturday | 0.107 | 0.309 | 0.095 | 0.294 |
| Sunday | 0.071 | 0.256 | 0.070 | 0.255 |
| # of observations | 78,946 | | 56,510 | |

(a) The variable 'Playing home' is a dummy indicator equal to 1 if the nationality of the player matches the country where the match is played, 0 otherwise.
(b) These indicators are not used as covariates in models with tournament fixed effects because they are time constant within tournaments.
(c) We included in the category Grand Slam the 30 matches of the 2003 and 2004 editions of the male Masters Cup, which were exceptionally played outdoors in Houston.

**Figure 1:** Kernel-weighted local polynomial smoothing of the relation between temperature and injury rate by gender and best-of-three/best-of-five matches




*Notes*: This figure is obtained using 135,456 player-level observations: 18,914 (60,032) for men playing best-of-5(3) matches and 56,510 for women.

registered at 3 pm of the day of the match in the tournament location and, therefore, $\alpha$ is the marginal effect of temperature on the injury rate; finally, $\varepsilon_{imte}$ is an idiosyncratic error term.

Including player FE ($\delta_i$) and opponent FE ($\eta_j$) mitigate potential confounding factors arising from unobservable player characteristics that could correlate with both the likelihood of injury and ambient temperature. For instance, certain players may anticipate physiological discomfort in extreme temperatures and, consequently, opt out of tournaments held in regions prone to heat waves.

The inclusion of tournament-edition FE ($\gamma_{te}$) serves to account for distinctive characteristics specific to each edition of a particular tournament. Thus, these fixed effects not only capture idiosyncrasies at the tournament level, but also address their temporal evolution over successive editions. For instance, consider the Australian Open, one of the four major ATP and WTA tournaments, which is held annually in January in Melbourne. As one of the prestigious Grand Slam events, with significant prizes and public attention, the mental and physical effort exerted by players may systematically differ compared to lesser-known tournaments. Such unobservable factors may vary over time, exemplified by the evolving heat policy implemented at the Australian Open across different years. This heat policy, as a time-varying fixed effect at the tournament level, could be correlated with the prevailing temperatures experienced during the tournament, influencing players' mental and physical effort and potentially affecting injury rates. By incorporating tournament-edition fixed effects, we effectively purge our estimates from this kind of time-varying unobservables.

An alternative perspective on our identification strategy involves viewing the inclusion of tournament-edition FE as akin to a transformation of Equation (2). This transformation entails subtracting from each variable its within tournament-edition average. Consequently, our identification of the causal effect relies on the deviation of match temperature from the average temperature observed across matches within the same tournament-edition. Given player FE and opponent FE, this short-term variability is reasonably considered exogenous with respect to any potential unobservable factors influencing the likelihood of injury occurrence.

The error term $\varepsilon_{ijmte}$ may exhibit correlation across observations, particularly within tournaments and players. The former correlation can stem from the distinct characteristics inherent to each tournament, such as surface type and attendance levels. Moreover, multiple matches within the same tournament may be affected by heat waves, which can persist over several days. Concerning within-player correlation regarding the probability of injury occurrence, it may arise from individual players' unique playing styles, strengths, weaknesses, and physiological limits. Consequently, for inference purposes, we adopt the two-way cluster-robust variance estimator proposed by Cameron et al. (2011), with clustering at both the tournament and player levels. The number of clusters is sufficiently large in both dimensions for both male and female players.[13]

[13]For (wo)men, we have 91 (108) different tournaments and 971 (755) different players.

# 4 Results

## 4.1 Main results

Our main parameter estimates on the determinants of injury rates are shown in Table 2. Temperature has a significant and positive effect on the injury rate only for men. An increase of one degree Celsius in ambient temperature translates into an increase of 0.06 percentage points in the injury rate. This is quite a large effect if we relate it to the average injury rate for men, which is 1.6%. Indeed, an increase of one degree Celsius in temperature amounts to a 3.75% increase with respect to the male average injury rate. Columns (3) and (4) of Table 2 present the temperature effect for men if the sample is split into best-of-three and best-of-five matches, respectively. This kind of analysis can be conducted only for men, as women do not play best-of-five matches. Best-of-five matches typically last longer. They last at minimum three sets, whereas best-of-three matches last at maximum three sets. Hence, best-of-five matches require players the exertion of mental and physical effort for a longer period of time. With extreme heat, fatigue is exacerbated, leading more easily to physical distress and dehydration and, therefore, to higher injury rates. However, even if the effect of temperature on injury rates (per unit of time) is identical for best-of-three and best-of-five matches, the probability that a player gets injured is larger in best-of-five matches simply because these last longer and therefore injuries are more likely to materialize. In line with this, we find that the injury probability is significantly higher in best-of-five matches by 2.3 percentage points and that the temperature gradient is larger in best-of-five matches: a rise by one degree Celsius in temperature increases the injury rate by 0.08 percentage points in best-of-five matches and by 0.05 percentage points in best-of-three matches. Nevertheless, although sizeable, the difference in the temperature effect is not statistically significant ($p$-value equal to 0.576).

About other covariates, the injury rates of male tennis players are lower if the difference in the ranking of players increases, if the player plays at home or in the final match of a tournament. The injury rates of female tennis players are higher when air quality is lower (i.e. $PM_{0.5}$ is higher) and, similarly to men, the injury rates are lower in final matches of tournaments.

Hence, on average injury rates increase with temperature for men, whereas for women such an impact is absent. For both men and women, there may be heterogeneity in the

**Table 2:** Parameter estimates (multiplied by 100) of the injury equation

| | Women | Men | Men Best-of-3 | Men Best-of-5 |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Coeff. | Coeff. | Coeff. | Coeff. |
| | (Std. Err.) | (Std. Err.) | (Std. Err.) | (Std. Err.) |
| Temperature at 3 pm (°C) | -0.0278 | 0.0603*** | 0.0549* | 0.0783** |
| | (0.0222) | (0.0202) | (0.0283) | (0.0311) |
| $PM_{2.5}$ at 3 pm (mg/kg) | 3.8272* | 4.2529 | 4.0287 | -2.4081 |
| | (2.1963) | (2.7427) | (2.9849) | (6.9697) |
| $O_3$ at 3 pm (mg/kg) | 7.3361 | -1.9901 | -3.5974 | 2.4175 |
| | (4.4306) | (4.7570) | (5.8252) | (7.5680) |
| Wind speed at 3 pm (m/sec) | -0.0501 | 0.0019 | -0.0000 | 0.0164 |
| | (0.0566) | (0.0366) | (0.0364) | (0.0974) |
| Best of 5 sets | | 2.2806** | | |
| | | (1.0115) | | |
| Absolute difference players' ranking[(a)] | 0.0661 | -0.2354** | -0.2540** | -0.1512 |
| | (0.1172) | (0.0937) | (0.0998) | (0.2941) |
| Sum in player's ATP/WTA ranking | -0.1921* | 0.1478 | 0.1454 | 0.0832 |
| | (0.1133) | (0.0962) | (0.0916) | (0.3055) |
| Playing home | -0.2304 | -0.5035*** | -0.5709*** | -0.3279 |
| | (0.1775) | (0.1417) | (0.1685) | (0.3231) |
| *Round - Reference: Before quarterfinal* | | | | |
| Quarterfinal | 0.0599 | -0.1105 | -0.1917 | 0.2434 |
| | (0.2057) | (0.1876) | (0.1607) | (0.5663) |
| Semifinal | 0.2768 | -0.0377 | -0.0488 | -0.2499 |
| | (0.3097) | (0.2140) | (0.2701) | (0.5388) |
| Final | -0.7417** | -0.6150** | -0.4990 | -1.7747** |
| | (0.3345) | (0.3056) | (0.3856) | (0.6217) |
| # of observations[(b)] | 56,122 | 78,480 | 59,526 | 18,682 |
| # of players | 755 | 971 | 906 | 634 |
| # of tournaments | 108 | 91 | 88 | 15 |
| Adj. R-Square | 0.0097 | 0.0045 | 0.0023 | 0.0127 |

* $p$-value<0.10, ** $p$-value<0.05, *** $p$-value<0.01. Two-way clustered standard errors are in parenthesis; clusters are at the level of tournaments and players. All the models include player FE, opponent FE and tournament-edition FE. They also include age and dummies for the day of the week and the quarter of the year. The corresponding parameters are not reported for the sake of brevity and they are available from the authors upon request.

(a) The difference in the ranking is the absolute value of the difference in the ATP/WTA rankings divided by 100.

(b) The number of observations are smaller than those reported in Table 1 because, when estimating the model with tournament-edition, player FE and opponent FE, 466 (388) (fe)male players were not used in the estimation because singleton observations. For the same reason, the sum of observations from best-of-3 matches and best-of-5 matches used in columns (3) and (4) do not equalize the total observations for men in column (2).

temperature effects. We first studied this issue by interacting the temperature variable with several characteristics that may be relevant in determining a different physiological and mental response to high temperatures and, therefore, a different probability of retirement when temperatures rise. More in detail, we checked the temperature gradient across the following dimensions: age, surface of the tennis court, tournament round, tournament series, ATP/WTA ranking of the player and playing at home. The various parameter estimates that show the heterogeneity in the effect of temperature on injuries are presented in the appendix in Table A.1. To sum up, temperature only has a positive effect on injury rates for male players, with no heterogeneity across several tournament, match and player's characteristics. The appendix describes more extensively the heterogeneity analysis.

In addition to heterogeneity of the temperature effect by these characteristics, the temperature effect may be heterogeneous within a match, i.e. it may differ between sets and depend on whether players at the start of a set are leading, behind or whether the score in sets is balanced. We investigate this in what follows.

### 4.2 Heterogeneity by set

The performance of tennis players may differ between sets. A general phenomenon within tennis matches is the response to losing. Malueg and Yates (2010) studied best-of-three tennis matches between equally skilled players and found that in the first set there was an equal win probability. The behavior in the second set was influenced by the outcome of the first set. First round winners exerted greater effort than first round losers. Banko et al. (2016) found that female tennis players who lost the first set were more likely than male players to perform worse in the second set. This was attributed to female tennis players being more likely to feel the pressure to succeed in the second set. A similar effect may materialize in the third set after the second one was lost. However, this effect was smaller since the effect of being behind disappeared. After all, a third set is played only when the score is balanced after two sets. De Paola and Scoppa (2017) also concluded that women were more likely than men to play poorly in the second set after losing the first set, whereby the gender differences in response to losing were stronger in high-stake matches.[14]

[14]The discouragement effect is not limited to within matches. Iqbal and Krumer (2019), analyzing Davis Cup tennis matches, concluded that the outcome of a match is influenced by the outcome of previous matches. There was an asymmetry between matches that started when the Davis Cup team was ahead or

In tennis the duration of a match in terms of the number of sets can vary. A best-of-three match can be decided after two sets or after three sets. Women only play best-of-three matches. For women, 68% of the matches is decided after two sets (2-0) and 32% after three sets (2-1). For men, in the best-of-three matches 66% is decided after two sets and 34% after three sets. For men, in the best-of-five matches, 51% is decided after three sets (3-0), 31% after four sets (3-1) and 18% after five sets (3-2). In best-of-three matches, only about 1 out of 3 matches lasts more than two sets both for men and women; in best-of-five matches, about half lasts more than three sets.

In our analysis of heterogeneity by set, for men we distinguish between best-of-three matches and best-of-five matches. This allows us to interpret parameter estimates of the temperature effect taking differences in incentives into account. The third set in a best-of-three match is always the last set in the match; the third set in a best-of-five match could be the last set or could be the entrance to the last two sets. Table 3 presents the parameter estimates of best-of-three matches; Table 4 shows the parameter estimates of best-of-five matches.

As shown in Table 3, for both men and women, the injury rate in the first set is low, about 0.3%. The temperature has no significant effect on the injury rate. In the second set, the injury rate, conditional on playing the 2nd set, is about three times as high, on average about 0.9%. We distinguish between second sets where the player has an advantage of 1-0 and second sets where the player has a disadvantage of 1-0. For the temperature effects for women, this distinction is irrelevant. Irrespective of being 1-0 up or 1-0 down, temperature has no effect on injury rates of women. For men, this is different. There is no temperature effect in the second set if the player is 1-0 up, but if he is 1-0 down, temperature has a significant positive effect on retirement. We attribute this to an disincentive effect. As indicated before, previous studies found that losing the first set causes a disincentive to put in effort in the second set. Apparently, high temperatures exacerbate the disincentive effect. The third set in a best-of-three match is played only if the score is 1-1 after two sets. In the third set, the injury rate, conditional on playing the third set, is somewhat lower than in the second set. Temperatures do not affect injury rates, neither for women nor for men. In the third set, there is no disincentive effect from the scores in previous sets probably because, by definition, the third set will only be played if there is a balanced score after two sets.

behind. When behind, players were discouraged having fewer incentives to exert effort and were therefore more likely to lose.

The main conclusion from Table 3 is that in best-of-three matches there is no effect of temperatures on injury rates for women. For men there is a temperature effect, but only in the second set and only for the player that starts the second set having lost the first set.

Table 4 presents additional parameter estimates for best-of-five matches. Best-of-five matches on average last significantly longer than best-of-three matches and the physical exhaustion from a much longer duration may be age dependent. For this reason, we report the analysis for the whole ATP sample, but also by age after interacting temperature variables by age dummies for being above or below the median age.[15,16]

Panel a shows that in the first set, temperature does not have an effect on injury rates. Panel b shows that this is also the case in the second set. Differently from the estimates in Table 3, the score after one set does not matter for the temperature effects. Perhaps, it is too early in the match to experience a temperature related disincentive effect. After all, it could be that after the second set the match continues for another three sets. Different from the best-of-three matches losing the second set does not imply losing the match. In panel c a distinction is made between the temperature effect according to the score balance at the start of the 3rd set: 2-0, 1-1 or 0-2. If a player has an advantage, temperature does not influence his injury rate in the 3rd set. If the score is balanced, temperature has a positive effect on the injury rate, but it is significant only at 10%; if the player is 0-2 behind at the start of the third set, temperature has a positive but insignificant effect on the injury rate. Hence, the temperature effect in the third set of best-of-five matches is also weak.

Panel d of Table 4 shows that temperature has no effect on the injury rate in the fourth set if the score is to the advantage of the players (2-1). However, if the player is 2-1 behind after three sets, there is a significant positive temperature effect on the injury rates. With 1-2 down, losing the fourth set implies losing the match. This significant effect is even more intense for older players. If the fourth set is won, a fifth set must be played. If the score is 2-1 up, winning the fourth set implies winning the match. If the fourth set is lost, the match is not yet decided until also the fifth set is played. Clearly, the incentive of stopping the game is higher when the player is 1-2 down and this influences the relationship between temperature and retirement. Older players seem to react more importantly to the incentive of stopping the game when the player is 1-2 down.[17] This

[15] In our dataset, the median age in best-of-five matches is 26.6 years.

[16] We also run the same heterogeneity analysis by age for men and women in best-of-three matches. We found that the heterogeneity by set does not depend on age. Table A.2 in the appendix reports these results.

[17] The difference in the temperature gradient between older and younger players, with a $p$-value equal to

may be due to their experience or to their worse physical conditions.[18]

Finally, panel e of Table 4 shows the effect of temperature on the injury rate in the 5th set. Here, no distinction can be made between scores at the start of the fifth set as this is by default 2-2. The effect is positive, though not significantly different from zero. Nevertheless, the point estimate of the temperature effect is quite high. The lack of significance may be explained by a too small statistical power of this heterogeneity exercise.

**Table 3:** Best-of-three sets: Heterogeneity of the temperature effect by sets (estimates multiplied by 100)

| | Men | Women |
|---|---|---|
| *a) Impact on the probability of injury in the 1st set* | | |
| Temperature at 3 pm (°C) | 0.0120 | -0.0059 |
| | (0.0145) | (0.0110) |
| # of observations | 59,526 | 56,122 |
| # of players | 906 | 755 |
| # of tournaments | 88 | 108 |
| Fraction of injuries in the 1st set (%) | 0.348 | 0.253 |
| *b) Impact on the probability of injury in the 2nd set, conditional on playing the 2nd set and depending on the result at the start of the 2nd set (reference: winning the 1st set)* | | |
| Temperature at 3 pm (°C) | 0.0082 | -0.0148 |
| | (0.0186) | (0.0192) |
| Temperature at 3 pm (°C) $\times$ losing the 1st set | 0.0414** | -0.0249 |
| | (0.0189) | (0.0236) |
| # of observations | 59,110 | 55,834 |
| # of players | 905 | 755 |
| # of tournaments | 88 | 108 |
| Fraction of injuries in the 2nd set, conditional on playing the 2nd set (%) | 0.937 | 0.981 |
| *c) Impact on the probability of injury in the 3rd set, conditional on playing the 3rd set* | | |
| Temperature at 3 pm (°C) | 0.0447 | 0.0218 |
| | (0.0321) | (0.0287) |
| # of observations | 20,352 | 18,070 |
| # of players | 647 | 547 |
| # of tournaments | 88 | 108 |
| Fraction of injuries in the 3rd set, conditional on playing the 3rd set (%) | 0.634 | 0.753 |

* $p$-value<0.10, ** $p$-value<0.05, *** $p$-value<0.01. Two-way clustered standard errors are in parenthesis; clusters are at the level of tournaments and players. All the models include player FE, opponent FE and tournament-edition FE. They also include all the covariates used in the baseline specification. The corresponding parameters are not reported for the sake of brevity and they are available from the authors upon request.

0.0578, is barely significant.

[18]An alternative explanation of why the temperature effect depends on being 1-2 down at the start of the fourth set may be that the player is already a little bit injured in the third set, and that is why the third set was lost. Then, the negative score at the start of the fourth set is the result of a forthcoming revealed injury and retirement, and not the other way around.

**Table 4:** Best-of-five sets: Heterogeneity of the temperature effect by sets and age (estimates multiplied by 100)

| | All<br>Coeff.<br>(Std. Err.)<br>(1) | Young[(a)]<br>Coeff.<br>(Std. Err.)<br>(2) | Old[(a)]<br>Coeff.<br>(Std. Err.)<br>(3) | Test age<br>difference<br>$p$-value<br>(4) |
|---|---|---|---|---|
| *a) Temperature effect in the 1st set* | | | | |
| Temperature at 3 pm (°C) | 0.0093 | 0.0152 | 0.0032 | 0.3713 |
| | (0.0140) | (0.0161) | (0.0149) | |
| # of observations | | | | 18,682 |
| # of players | | | | 634 |
| Injury rate in the 1st set | | | | 0.177 |
| *b) Temperature effect in the 2nd set conditional on playing the 2nd set and depending on the result and the start of the 2nd set (reference: winning the 1st set)* | | | | |
| Temperature at 3 pm (°C) | -0.0276 | -0.0180 | -0.0385 | 0.4950 |
| | (0.0326) | (0.0346) | (0.0378) | |
| Temperature at 3 pm (°C) × losing the 1st set | 0.0012 | 0.0198 | 0.0067 | 0.2652 |
| | (0.0026) | (0.0277) | (0.0264) | |
| # of observations | | | | 18,612 |
| # of players | | | | 633 |
| Injury rate in the 2nd set, conditional on playing the 2nd set (%) | | | | 0.633 |
| *c) Temperature effect in the 3rd set conditional on playing the 3rd set and depending on the result at the start of the 3rd set (reference: winning 2-0 at the start of the 3rd set)* | | | | |
| Temperature at 3 pm (°C) | 0.0026 | 0.0201 | 0.0328 | 0.6665 |
| | (0.0310) | (0.0366) | (0.0317) | |
| Temperature at 3 pm (°C) × drawing 1-1 at the start of the 3rd set | 0.0399* | 0.0398 | 0.0405* | 0.9514 |
| | (0.0024) | (0.0249) | (0.0238) | |
| Temperature at 3 pm (°C) × losing 2-0 at the start of the 3rd set | 0.0229 | 0.0194 | 0.0251 | 0.7523 |
| | (0.0430) | (0.0432) | (0.0442) | |
| # of observations | | | | 18,366 |
| # of players | | | | 631 |
| Injury rate in the 3rd set, conditional on playing the 3rd set (%) | | | | 0.718 |
| *d) Temperature effect in the 4th set conditional on playing the 4th set and depending on the result at the start of the 4th set (reference: winning 2-1 at the start of the 4th set)* | | | | |
| Temperature at 3 pm (°C) | 0.0221 | 0.0343 | 0.0130 | 0.7024 |
| | (0.0487) | (0.0509) | (0.0620) | |
| Temperature at 3 pm (°C) × losing 2-1 at the start of the 4th set | 0.1272** | 0.1054* | 0.1434*** | 0.0578 |
| | (0.0536) | (0.0550) | (0.0541) | |
| # of observations | | | | 8,960 |
| # of players | | | | 510 |
| Injury rate in the 4th set, conditional on playing the 4th set (%) | | | | 0.867 |
| *e) Impact on the probability of injury in the 5th set, conditional on playing the 5th set* | | | | |
| Temperature at 3 pm (°C) | 0.0831 | 0.0849 | 0.0833 | 0.986 |
| | (0.0742) | (0.0901) | (0.0819) | |
| # of observations | | | | 3,230 |
| # of players | | | | 382 |
| Injury rate in the 5th set, conditional on playing the 5th set (%) | | | | 0.774 |

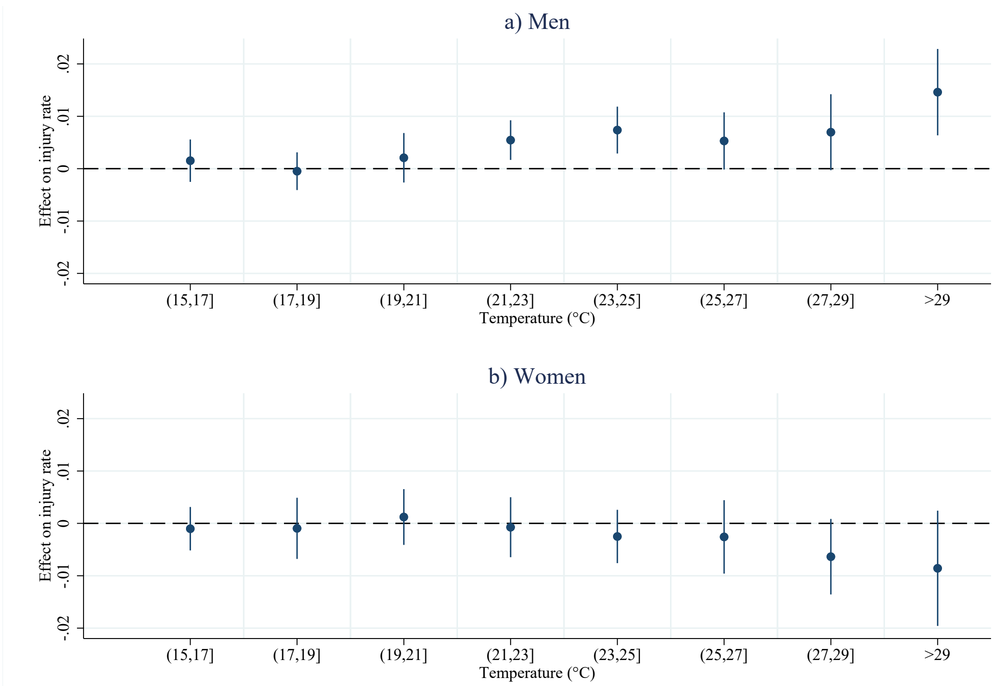
* $p$-value<0.10, ** $p$-value<0.05, *** $p$-value<0.01. Standard errors clustered at player level are in parenthesis. All the models include player FE, opponent FE and tournament-edition FE. They also include all covariates used in the baseline specification. The corresponding parameters are not reported for the sake of brevity and they are available from the authors upon request.

[(a)] 'Young' ('Old') means that the player is below (above) the median age, which is 26.6 years for players who played best-of-five matches.

## 4.3 Sensitivity analysis

As a first robustness check, we assessed if the main findings are affected by too strict parametric assumptions about the functional relationship between injury rate and temperature. Figure 2 shows the relation between temperature and the injury rate if we allow it to be nonlinear by using a piecewise constant specification. It shows that for men the temperature effect becomes much stronger for temperatures above 29°C. The injury rate with temperatures above 29°C is more than 1 percentage point higher than when temperatures are below 21°C. This is quite a large effect if we relate it to the average injury rate for men, which is 1.6%. For women, the lack of a temperature effect is also confirmed when departing from the linearity of the functional relationship between temperature and injury rate.

**Figure 2:** Non linear effect of ambient temperatures on injury rate




*Notes*: The reference category is temperature below 15°C. The number of observations is the same as those reported in Table 2. Segments are 95% confidence intervals computed with two-way clustered standard errors; clusters are at the level of tournaments and players.

In a second sensitivity analysis, we replaced ambient temperature with an alternative measure that is related to how temperature affects the human body. This measure accounts for air humidity and therefore the capability of human body to dissipate excess heat: the Humidex index calculated as in Blazejczyk et al. (2012). This index combines ambient temperature and dew point temperature coming from the Copernicus Atmosphere Monitoring Service. Panel a) of Table 5 displays the results, confirming the baseline findings: the warmer, the higher the probability of retirement for men, while there is no effect for women.

Third, we checked whether COVID-19, with its lockdown and suspension of professional tennis tournaments, may have played a role in affecting the relationship between ambient temperature and injury rate. After the outbreak of the COVID-19 pandemic, all professional tennis tournaments were suspended (the last match in our dataset before the outbreak of the COVID-19 pandemic was played on 8 March 2020) and restarted in August 2020. In addition to this long suspension, training was inhibited by COVID-19 lockdown measures and other restrictions. These may have affected the physical and mental conditions of players which, if coupled with particular temperature shocks after the COVID-19 break, may have generated spurious correlation between temperature and injury rates. Therefore, we checked the robustness of out main results by dropping from our sample the matches played in 2020 after the outbreak of the COVID-19 pandemic: we removed all the matches played from the moment in which the ATP/WTA tournaments restarted in August 2020 until the end of 2020. Panel b) of Table 5 shows that the results are unchanged after dropping these matches.

Finally, we replicated the main analysis using indoor matches. Since in indoor matches air conditioning and climate control ensure optimal physiological conditions by regulating temperature, humidity and other environmental factors, we do not expect to detect a significant impact of ambient (external) temperature on injury rate. Therefore, this kind of exercise may be seen as a placebo test. Panel c) of Table 5 shows that indeed ambient temperature does not affect the retirement probability in indoor matches.

# 5 Conclusions

We studied the effect of ambient temperatures on injury rates in professional tennis matches and derived findings at two levels: the overall match level and within individual matches. About the former, we found that temperature affects injury rates for men but not for women. Injury rates are higher in best-of-five matches, which are only played by men, and seem to be affected not only by physical circumstances, such as temperature, but also

**Table 5:** Sensitivity analyses (parameter estimates $\times$ 100)

| | (1)<br>Men<br>Coeff.<br>(Std. Err.) | (2)<br>Women<br>Coeff.<br>(Std. Err.) |
|---|---|---|
| *a) Using Humidex index instead of ambient temperatures* | | |
| Humidex index at 3 pm (°C) | 0.0472*** | -0.0127 |
| | (0.0146) | (0.0153) |
| # of observations | 78,480 | 56,122 |
| *b) Removing post COVID-19 lockdown matches (August-December 2020)* | | |
| Temperature at 3 pm (°C) | 0.0619*** | -0.0222 |
| | (0.0206) | (0.0216) |
| # of observations | 77,586 | 55,126 |
| *c) Using indoor matches* | | |
| Temperature at 3 pm (°C) | 0.0282 | 0.0028 |
| | (0.0502) | (0.1017) |
| # of observations | 17,096 | 5,430 |

* $p$-value<0.10, ** $p$-value<0.05, *** $p$-value<0.01. Two-way clustered standard errors are in parenthesis; clusters are at the level of tournaments and players. All the models include player FE, opponent FE and tournament-edition FE. They also include all covariates used in the baseline specification. The corresponding parameters are not reported for the sake of brevity and they are available from the authors upon request.

by incentives. When playing at home men are less likely to retire. Both men and women are less likely to suffer from an injury in finals. When analysing if the temperature gradient depends on further player's characteristics, like age, ranking, playing at home, and match characteristics, as surface, round and tournament series, we find that the effect of temperature on injuries is not heterogeneous.

Our findings are of direct interest for health policy making since about 90 million people play tennis worldwide. They highlight the importance of enlarging the use of extreme heat policies, which some professional tennis tournaments already have, of avoiding scheduling matches during the warmest hours of the day, and of having guidelines and education on safe playing conditions for players and their staff, especially for amateur players who potentially have lower fitness levels and may be more affected by extreme conditions. Moreover, insights gained from our study clearly have external validity beyond the game of tennis itself, especially in regions characterized by relevant temperature variations and heat waves. They have broader public health implications for outdoor high-stake activities which combine, as in tennis, mental skills, physical strength, technical proficiency, and fine motor control to be successful.

Within individual matches, we detected heterogeneity of the temperature effect across sets and depending on whether players are in front or behind at the beginning of a set.

Our results strongly suggest that when a male player is losing at the beginning of a crucial (second) fourth set in (best-of-three) best-of-five matches, the temperature effect is much larger than when he is winning. In best-of-five matches, which are more exhausting, this effect is age-dependent and stronger for older players. There are at least two explanations for this. When losing, a player may give up more easily as soon as he feels some heat-related physical discomfort because the chances of winning are low. If a player is behind in set-score, he realizes that he has more sets to win to win the match. Therefore, the match is expected to last longer. So, when a player is behind in score, expected marginal benefits go down and expected marginal costs go up, which makes it rational to reduce effort in some case to zero. An alternative explanation is that heat-related physical discomfort may have started exerting its effects before the start of the crucial second or fourth set and may jointly explain why the player is losing and the decision to retire before the end of the match. However, a rational player who realizes that losing the previous set or sets is related to an injury will not wait for the next set to stop playing. The temperature effect of injuries when behind in score materializes in the penultimate set, i.e., the second set in best-of-three and the fourth set in best-of-five matches. Because the set-score at the start of a set mainly matters for the crucial penultimate set, we attach more weight to the first explanation. Some of the temperature-induced injuries are caused by disincentives to play continuation. Another way of looking at this result is that when there is more at stake, i.e. a player is winning at the beginning of a crucial set, injury reporting with higher temperatures is less likely than when there is less at stake, i.e. a player is losing at the beginning of a crucial set. This is in line with previous research which investigated the relation between injury reporting behavior and perceived job insecurity in standard workplaces. For example, Probst et al. (2013) found that in Italy and the US when job insecurity is high, i.e. there is more at stake, under-reporting is also high and registered injury rates tend to be lower.

Our main contribution to the broader literature on physiological adaptation to extreme heat conditions is the finding that economic agents respond to incentives around heat-induced potential injuries. Accounting for incentives is challenging. Our solution is to study within-match specifics. We detect a trade-off between the long-term economic costs of reporting an injury and the immediate exertion of physical effort exacerbated by heat. Our current research is not able to dig deeper on the relation between injury reporting and incentives beyond the relevance of winning at the beginning of a crucial set in tennis. We leave this for future research. What we can conclude from our research is that workers are more likely to ignore discomfort from warm temperatures and not report workplace injuries when there is more at stake. Not timely reporting an accident or continuing to be exposed to a hazardous condition, like extreme heat, may generate greater problems for the worker later on. By better understanding the relation between incentives and injury reporting behavior, public health institutions and organizations may design more effective workplace safety policies based on incentives.

# References


Alexander, S., N. Naaz, and S. Fernandes (2022). The incidence of injuries across various tennis surfaces: A systematic review. *ITF Coaching & Sport Science Review 30*(88), 39–44.

Banko, L., E. M. Leeds, and M. A. Leeds (2016). Gender differences in response to setbacks: Evidence from professional tennis. *Social Science Quarterly 97*(2), 161–176.

Blazejczyk, K., Y. Epstein, G. Jendritzky, H. Staiger, and B. Tinz (2012). Comparison of UTCI to selected thermal indices. *International Journal of Biometeorology 56*(3), 515–535.

Boone, J. and J. C. van Ours (2006). Are recessions good for workplace safety? *Journal of Health Economics 25*(6), 1069–1093.

Boone, J., J. C. van Ours, J.-P. Wuellrich, and J. Zweimüller (2011). Recessions are bad for workplace safety. *Journal of Health Economics 30*(4), 764–773.

Breznik, K. and V. Batagelj (2012). Retired matches among male professional tennis players. *Journal of Sports Science and Medicine 11*, 270–278.

Burke, M., V. Tanutama, S. Heft-Neal, M. Hino, and D. Lobell (2023). Game, sweat, match: Temperature and elite worker productivity. NBER Working Paper 31650.

Cameron, A., J. Gelbach, and D. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Chatterjee, S. and A. S. Hadi (1986). Influential observations, high leverage points, and outliers in linear regression. *Statistical Science 1*(3), 379–393.

De Paola, M. and V. Scoppa (2017). Gender differences in reaction to psychological pressure: Evidence from tennis players. *European Journal of Work and Organizational Psychology 26*(3), 444–456.

Dell, M., B. F. Jones, and B. A. Olken (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature 52*(3), 740–98.

Dillender, M. (2021). Climate change and occupational health. *Journal of Human Resources 56*(1), 184–224.

Fernandez, J., A. Mendez-Villanueva, and B. M. Pluim (2006). Intensity of tennis match play. *British Journal of Sports Medicine 40*, 387–391.

Filomena, M. and M. Picchio (2024). Unsafe temperatures, unsafe jobs: The impact of weather conditions on work-related injuries. *Journal of Economic Behavior & Organization 224*, 851–875.

Gescheit, D. T., S. J. Cormack, R. Duffield, S. Kovalchik, T. O. Wood, M. Omizzolo, and M. Reid (2017). Injury epidemiology of tennis players at the 2011-2016 Australian Open Grand Slam. *British Journal of Sports Medicine 51*(17), 1289–1294.

Gescheit, D. T., S. J. Cormack, R. Duffield, S. Kovalchik, T. O. Wood, M. Omizzolo, and M. Reid (2019). A multi-year injury epidemiology analysis of an elite national junior tennis program. *Journal of Science and Medicine in Sport 22*(1), 11–15.

Hartwell, M. J., S. M. Fong, and A. C. Colvin (2017). Withdrawals and retirements in professional tennis players: An analysis of 2013 United States Tennis Association Pro Circuit tournaments. *Sports Health 9*(2), 154–161.

Inness, A., M. Ades, A. Agustí-Panareda, J. Barré, A. Benedictow, A.-M. Blechschmidt, J. J. Dominguez, R. Engelen, H. Eskes, J. Flemming, V. Huijnen, L. Jones, Z. Kipling, S. Massart, M. Parrington, V.-H. Peuch, M. Razinger, S. Remy, M. Schulz, and M. Suttie (2019). The CAMS reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics 19*(6), 3515–3556.

Iqbal, H. and A. Krumer (2019). Discouragement effect and intermediate prizes in multi-stage contests: Evidence from Davis Cup. *European Economic Review 118*, 364–381.

Ireland, A., D. Johnston, and R. Knott (2023). Heat and worker health. *Journal of Health Economics 91*, 102800.

ITF (2021). *ITF global tennis report*. London: International Tennis Federation.

Krumer, A., M. Rosenboim, and O. M. Shapir (2016). Gender, competitiveness, and physical characteristics: Evidence from professional tennis. *Journal of Sports Economics 17*(3), 234–259.

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

Martin, C. and J. Prioux (2016). Tennis playing surfaces: The effects on performance and injuries. *Journal of Medicine and Science in Tennis 21*(1), 11–18.

McCurdie, I., S. Smith, P. H. Bell, and M. E. Batt (2017). Tennis injury data from The Championships, Wimbledon, from 2003 to 2012. *British Journal of Sports Medicine 51*(7), 607–611.

Morrissey, M. C., G. J. Brewer, W. J. Williams, T. Quinn, and D. J. Casa (2021). Impact of occupational heat stress on worker productivity and economic cost. *American Journal of Industrial Medicine 64*(12), 981–988.

Muñoz-Sabater, J., E. Dutra, A. Agustí-Panareda, C. Albergel, G. Arduini, G. Balsamo, S. Boussetta, M. Choulga, S. Harrigan, H. Hersbach, B. Martens, D. G. Miralles, M. Piles, N. J. Rodríguez-Fernández, E. Zsoter, C. Buontempo, and J.-N. Thépaut (2021). ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data 13*(9), 4349–4383.

Otte im Kampe, E., S. Kovats, and S. Hajat (2016). Impact of high ambient temperature on unintentional injuries in high-income countries: A narrative systematic literature review. *BMJ open 6*(2), e010399.

Park, R., N. Pankratz, and A. Beher (2021). Temperature, workplace safety, and labor market inequality. IZA Discussion Paper No. 14560, IZA Institute of Labor Economics, Bonn, Germany.

Paserman, M. D. (2023). Gender differences in performance in competitive environments? Evidence from professional tennis players. *Journal of Economic Behavior and Organization 212*, 590–609.

Picchio, M. and J. C. van Ours (2017). Temporary jobs and the severity of workplace accidents. *Journal of Safety Research 61*, 41–51.

Picchio, M. and J. C. van Ours (2024). The impact of high temperatures on performance in work-related activities. *Labour Economics 87*, 102509.

Probst, T. M., C. Barbaranelli, and L. Petitta (2013). The relationship between job insecurity and accident under-reporting: A test in two countries. *Work & Stress 27*(4), 383–402.

Romano, J. P. and M. Wolf (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association 100*(469), 94–108.

Romano, J. P. and M. Wolf (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica 73*(4), 1237–1282.

Sell, K., B. Hainline, M. Yorio, and M. Kovacs (2014). Injury trend analysis from the US Open Tennis Championships between 1994 and 2009. *British Journal of Sports Medicine 48*(7), 546–551.

Smith, M. T., M. Reid, S. Kovalchik, T. Wood, and R. Duffield (2018a). Heat stress incidence and matchplay characteristics in Women's Grand Slam Tennis. *Journal of Science and Medicine in Sport 21*, 666–670.

Smith, M. T., M. Reid, S. Kovalchik, T. Wood, and R. Duffield (2018b). Heat stress incident prevalence and tennis matchplay performance at the Australian Open. *Journal of Science and Medicine in Sport 21*, 467–472.

Spector, J. T., Y. J. Masuda, N. H. Wolff, M. Calkins, and N. Seixas (2019). Heat exposure and occupational injuries: Review of the literature and implications. *Current Environmental Health Reports 6*, 286–296.

Varghese, B. M., A. Hansen, P. Bi, and D. Pisaniello (2018). Are workers at risk of occupational injuries due to heat exposure? A comprehensive literature review. *Safety Science 110*, 380–392.

# Appendix A: Effect heterogeneity

We studied the heterogeneity of the temperature effect by interacting the temperature variable with characteristics that may be relevant in determining a different physiological and mental response to high temperatures and therefore a different probability of retirement when temperatures rise. The various parameter estimates that show the heterogeneity in the effect of temperature on injuries are presented in Table A.1.

Panel a shows the heterogeneity of the temperature effects by age. Older players may react differently to high temperatures than younger players, determining a different probability of injury when it is warm. On the one hand, older players may physiologically suffer more from high temperatures and are therefore more subject to a heat-related injury. On the other hand, they are more experienced and they may be more able to change their game and adapt their strategies to cope with heat, for instance by paying more attention to conserving energies, strategic rest, hydration and nutrition. We find that there is no significant heterogeneity by age from the statistical point of view, although the point estimates suggest that for both men and women the temperature effect is smaller for older players.

In panel b we report the effect heterogeneity in the temperature effect by surface of the tennis court. Clay courts induce slower, higher ball bounce and higher loss of horizontal speed, increasing the duration of a rally compared to a faster surface such as grass (Fernandez et al., 2006; Martin and Prioux, 2016). Alexander et al. (2022) provided an overview of the incidence of injuries on various tennis surfaces, i.e., hard, clay and grass courts, finding no differences in injury rates across court surfaces. Nonetheless, since different surfaces are related to different energy consumption and, in hot weather, this difference may be magnified, the temperature effect on injuries may be milder on surfaces requiring less physical effort, like on grass. We find that there is no heterogeneity by surfaces, although for men it seems that heat-related injuries are less likely on grass, as expected.

Panels c and d show that there is no heterogeneity in the temperature effects by tournament round and by tournament series. The importance of a match, as measured either by the round during the tournament or by the relevance of the tournament, does therefore not affect heat-related injuries.

In panel e, we check if the ranking of the player matters. Higher ranking players typically are in a better shape and are more skilled. If they are in a better shape, they may cope better with heat-related physical stress. If they are more skilled, they have more degrees of freedom to adapt their game and strategies in order to cope with heat. The point estimates in panel e are in line with these hypotheses. However, the differences in the temperature gradient are not statistically significant.

Finally, in panel f we seek to understand if playing a home match matters for the temperature gradient. Playing home may affect the behavior of tennis players with reflexes on the heat-impact on the injury rate. On the one hand, the support of the home crowd may boost the performances above the limit and, when the weather is too warm, this may result in overheating and a higher probability of retirement. On the other hand, home players may be more familiar with the court, be subject with less fatigue of travel and have access to training facilities, resulting in a higher capabilities to cope with extreme ambient conditions. Panel f shows that, while for men the temperature effect is more important at home although not significantly, when women play a home match temperature has a negative and significant effect on injury rates.

**Table A.1:** Heterogeneity of the temperature effect on injury rate (estimates multiplied by 100)

| | Men | Women |
|---|---|---|
| *a) Effect by age (age > 27.6 years is the reference)[(a)]* | | |
| Temperature at 3 pm (°C) | 0.0294 | -0.0491 |
| | (0.5365) | (0.4246) |
| Temperature at 3 pm (°C) × age ≤ 23.9 years | 0.0536 | 0.0210 |
| | (0.1199) | (0.9231) |
| Temperature at 3 pm (°C) × 23.9 years ≤ age < 27.6 years | 0.0497 | 0.0368 |
| | (0.1199) | (0.7672) |
| *b) Effect by surface (grass is the reference)* | | |
| Temperature at 3 pm (°C) | 0.0238 | 0.0228 |
| | (0.9191) | (0.9780) |
| Temperature at 3 pm (°C) × clay | 0.0284 | -0.0383 |
| | (0.9191) | (0.9580) |
| Temperature at 3 pm (°C) × hard | 0.0534 | -0.0664 |
| | (0.7143) | (0.8681) |
| *c) Effect by tournament round (before quarterfinal is the reference)* | | |
| Temperature at 3 pm (°C) | 0.0634** | -0.0193 |
| | (0.0130) | (0.8681) |
| Temperature at 3 pm (°C) × quarterfinals or later | -0.0162 | -0.0512 |
| | (0.9191) | (0.4985) |
| *d) Effect by tournament series (non Grand Slam tournament is the reference)* | | |
| Temperature at 3 pm (°C) | 0.0541 | -0.0241 |
| | (0.1109) | (0.8681) |
| Temperature at 3 pm (°C) × Grand Slam tournament | 0.0183 | -0.0105 |
| | (0.9191) | (0.9780) |
| *e) Effect by player's ATP/WTA ranking (in the top 20 is the reference)* | | |
| Temperature at 3 pm (°C) | 0.0425 | -0.0401 |
| | (0.5135) | (0.5914) |
| Temperature at 3 pm (°C) × between 20 and 50 | 0.0139 | 0.0020 |
| | (0.9191) | (0.9780) |
| Temperature at 3 pm (°C) × above 50 | 0.0276 | 0.0236 |
| | (0.8561) | (0.9181) |
| *f) Effect on home advantage (not playing home is the reference)* | | |
| Temperature at 3 pm (°C) | 0.0560** | -0.0186 |
| | (0.0190) | (0.8681) |
| Temperature at 3 pm (°C) × playing home | 0.0331 | -0.0933** |
| | (0.3477) | (0.0180) |
| # of observations | 78,480 | 56,122 |
| # of players | 971 | 755 |
| # of tournaments | 91 | 108 |

* $p$-value<0.10, ** $p$-value<0.05, *** $p$-value<0.01. In parenthesis, we report Romano and Wolf's (2005a; 2005b) step-down adjusted $p$-values robust to multiple hypothesis testing, with two-way clusters at the level of tournaments and players (1,000 bootstrap replications). All the models include player FE, opponent FE and tournament-edition FE. They also include all the covariates used in the baseline specification. The corresponding parameters are not reported for the sake of brevity and they are available from the authors upon request.

(a) 23.9 and 27.6 years are 33rd and 66th percentiles of the age distribution.

**Table A.2:** Best-of-three sets: Heterogeneity of the temperature effect by sets and age (estimates multiplied by 100)

| | Men | | Women | |
|---|---|---|---|---|
| | Young[(a)] | Old[(a)] | Young[(a)] | Old[(a)] |
| *a) Impact on the probability of injury in the 1st set* | | | | |
| Temperature at 3 pm (°C) | 0.0175 | 0.0067 | -0.0109 | -0.0009 |
| | (0.0160) | (0.0144) | (0.0118) | (0.0125) |
| # of observations | 59,526 | | 56,122 | |
| # of players | 906 | | 755 | |
| # of tournaments | 88 | | 108 | |
| Fraction of injuries in the 1st set (%) | 0.348 | | 0.253 | |
| *b) Impact on the probability of injury in the 2nd set, conditional on playing the 2nd set and depending on the result at the start of the 2nd set (reference: winning the 1st set)* | | | | |
| Temperature at 3 pm (°C) | 0.0191 | -0.0028 | -0.0133 | -0.0162 |
| | (0.0198) | (0.0204) | (0.0200) | (0.0227) |
| Temperature at 3 pm (°C) × losing the 1st set | 0.0414** | 0.0427** | -0.0236 | -0.0261 |
| | (0.0199) | (0.0188) | (0.0243) | (0.0240) |
| # of observations | 59,110 | | 55,834 | |
| # of players | 905 | | 755 | |
| # of tournaments | 88 | | 108 | |
| Fraction of injuries in the 2nd set, conditional on playing the 2nd set (%) | 0.937 | | 0.981 | |
| *c) Impact on the probability of injury in the 3rd set, conditional on playing the 3rd set* | | | | |
| Temperature at 3 pm (°C) | 0.0553 | 0.0345 | 0.0056 | 0.0385 |
| | (0.0347) | (0.0366) | (0.0313) | (0.0328) |
| # of observations | 20,352 | | 18,070 | |
| # of players | 647 | | 547 | |
| # of tournaments | 88 | | 108 | |
| Fraction of injuries in the 3rd set, conditional on playing the 3rd set (%) | 0.634 | | 0.753 | |

* $p$-value$<$0.10, ** $p$-value$<$0.05, *** $p$-value$<$0.01. Two-way clustered standard errors are in parenthesis; clusters are at the level of tournaments and players. All the models include player FE, opponent FE and tournament-edition FE. They also include all the covariates used in the baseline specification. The corresponding parameters are not reported for the sake of brevity and they are available from the authors upon request.

[(a)] 'Young' ('Old') means that the player is below (above) the median age, which is 26.5 (24.7) years for (wo)men who played best-of-three matches.