DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

When performance melts away: Heat causes mental errors in high-stakes competitions

by

Mario LACKNER
Hendrik SONNABEND

Working Paper No. 2411
October 2024

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

mario.lackner@jku.at

When performance melts away: Heat causes mental errors in high-stakes competitions *

Mario Lackner[a] and Hendrik Sonnabend[b]

[a]*Johannes Kepler University Linz (JKU)*
[b]*University of Hagen*

October 2024

Abstract

This paper examines the impairing effect of heat stress on cognitive abilities in a high-stakes setting. Building on rich play-by-play data from the National Football League (NFL) linked to variations in game-time temperature, we find that players are about 25% more likely to be sanctioned for infractions associated with mental errors in games with temperatures above 85°F (29.4°C) compared to games with lower temperatures. Furthermore, we identify situations with (i) little room to adapt to heat stress and (ii) high work intensities, as well as the players' physical constitution as channels that can explain the heat-induced decline in mental performance.

JEL Classification: Q51, J24, J81
Keywords: heat stress, mental performance, football

**Corresponding author*: Mario Lackner, Johannes Kepler University of Linz, Department of Economics, Altenbergerstr. 69, 4040 Linz, Austria. Phone: +43 70 2468 7390, email: mario.lackner@jku.at.
We are grateful to Matthias Westphal, Jan van Ours, Ulf Rinne, W. David Allen, Alex Krumer, Felix Gläser, and Alexander Gussone, participants at the EALE 2024 conference and the Reading Online Sport Economics Seminar (ROSES), and attendees at research seminars at the universities of Linz (JKU) and Hagen.

1. Introduction

Climate change, characterized by a greater frequency of weather extremes such as heatwaves and flooding, presents a massive challenge for societies (Diffenbaugh & Burke 2019, Costello et al. 2023) and the global economy (Dell et al. 2012, 2014). Besides the enormous economic costs of bearing the direct impacts on the environment, high temperatures can also negatively affect individual productivity. When the temperature exceeds the thermal comfort zone (i.e., the 'room temperature' between 64°F (18°C) and 72°F (22°C), the human body makes efforts to restore its heat balance (Heal & Park 2016). This happens by redistributing the blood flow toward the skin and secreting sweat (Ebi et al. 2021).

While reducing physical activity is a natural way to mitigate the increase in body temperature, heat stress also impairs mental performance. Among others, this is because the redistribution of the blood flow leads to a lower level of cool blood reaching the brain (Graff Zivin et al. 2018). Hence, exposure to high temperatures may lower performance not only through reduced physical activities of the workers but also by mental errors that may come along with higher accident rates (e.g., Taylor et al. 2016). These mental errors can have severe health consequences in some occupations, such as working at heights, but also threaten productivity in less risky professions when climate controls are not possible or present. This may apply to regions with little prior experience with extreme heat, like Northeastern and Central Europe, but also the Pacific Northwest, where office and government buildings are not or only partially air-conditioned (Biardeau et al. 2020), either for reasons of cost or due to legal regulations.

This study tries to improve our understanding of how high temperatures affect the cognitive abilities of workers exposed to high temperatures in an environment with high pressure to perform and significant incentives. A growing body of literature aims to estimate the causal effect of high temperatures on labor productivity and decision-making. For example, it is a consensus finding that heat does hurt productivity in general (Burke et al. 2015*b*, Zander et al. 2015, Lai et al. 2023), in manufacturing (Somanathan et al. 2021, Kabore & Rivers 2023) and among interviewers (as workers) (LoPalo 2023), politicians' speeches (Keivabu & Widmann 2024) and the performance and success of the students (Park et al. 2020, Park 2022) in particular. Moreover, heat also has a documented negative effect on workers' health (Dillender 2021, Ireland et al. 2023) and worker safety (Park & Pankratz 2021, Filomena & Picchio 2024). Finally, a branch of literature (primarily medical research) pioneered by an experimental study of telegraph operators in the British Navy (Mackworth 1946) suggests that individuals work less accurately under heat stress. However, we still know little about the channels that control this negative effect.

We use data from the North American elite football league (National Football League, NFL) to identify mental errors and link them to variations in temperature during games that we interpret as a natural experiment. In football, some violations of the playing rules, such as when players fail to find the correct position or act too early, can be unambiguously categorized as mental errors since they are penalized, e.g., by moving the ball's position before a play, and do not bring any benefit to the offending team. Hence, teams will always try to avoid these infractions. The data allow us to observe a large number of competing players with time-varying productivity, different characteristics, and well-observed performance measures at a very granular level. Moreover, they include players' height and body weight, detailed information that is rarely available in traditional firm or even register data. From this, we can calculate the body mass index (BMI), which in the NFL – unlike most other sports – is very dispersed and much more similar to the distribution of BMI in a typical corporate or public-office workforce (e.g., Harp & Hecht 2005). This is one of the reasons we are confident about the general validity of our results for less specific environments where performance pressure combined with high temperatures at the workplace meets a population of professionals with heterogeneous physical characteristics.

Furthermore, compared to settings like exams in schools and universities, where performance strongly depends on (often unobserved) preparation efforts, this dimension is less important for our measure of reduced cognitive abilities. In contrast to previous studies focusing on more conventional settings, we examine an environment with very high incentives, comparatively high wages, and a high degree of specialization. Therefore, we can also neglect heat-induced changes in labor supply and time allocation in professional football (as discussed in LoPalo 2023). We also do not have to worry about heat-induced changes in input factors and their prices (as discussed in Somanathan et al. 2021).

Our primary finding is that the propensity of professional football players to commit mental errors increases substantially when the game-time temperature surpasses 85°F. In our preferred specification, this rise amounts to 25.73% when evaluated at the sample mean. Since NFL players are highly trained and monitored agents with high incentives to perform well, this may represent a lower bound for the population-wide effect. Furthermore, we provide a whole array of robustness checks and sensitivity analyses to ensure the validity and reliability of our main result. For instance, we demonstrate that the heat effect cannot be explained by physical exhaustion as the effect is already present at the beginning of a contest. Moreover, we do not find an increase in mental errors at low temperatures, nor when the outside temperature is high, but the stadium is closed and fully air-conditioned. This second result also supports the argument for the protective effect of air conditioning systems. We also consider factors affecting heat perception, like humidity, sun exposure, and working speed.

Our heterogeneity analyses allow us to offer two valuable contributions. First, we gain novel insights into the role of physical characteristics. Our analyses reveal that extreme heat conditions negatively affect only players with an above-median body mass index (BMI). For this group, the probability of being sanctioned for infractions associated with mental errors increases by up to 50% when temperature exceeds 85°F. We provide evidence to support the assumption that this result is due to players with an unfavorable body fat percentage. Second, we identify the adaption to high temperatures, the speed of action, and breaks as important mediators of heat-induced mental errors. Specifically, it shows that the heat effect on mental performance is driven (i) by visiting teams who have to deal with a sudden increase in temperature and, on the game level, (ii) by a greater speed of the game when a trailing team must catch up, while (iii) it is not present in plays that immediately follow a break and thus the opportunity to gather oneself physically and mentally.

Due to the general advantages of sports data, such as the high incentives and standardized rules, they have previously been used to estimate the effect of temperature on performance. For example, Qiu & Zhao (2022) find that the performance of professional competitors in archery suffers as the temperature increases. This effect is more pronounced for low-performing competitors, leading to the conclusion that there is also a redistributive effect. Fesselmeyer (2021) documents that the performance of umpires in professional baseball decreases when the temperature reaches 95°F or more, meaning that they make more mistakes when calling balls and strikes as their accuracy of judgments decreases. Burke et al. (2023) and Picchio & Van Ours (2024) use data from professional tennis and find an age- and skill-specific decrease in performance as the ambient temperature increases. Compared to this branch of literature, a major advantage of our underlying setting is that we do not use performance measures to approximate productivity but focus on a group of penalties working as indicators of mental errors due to poor concentration. Therefore, we can rule out potential biases arising from the fact that heat also affects the flight characteristics of arrows or balls and playing conditions in tennis. Furthermore, team sports place greater demands on the players' coordination skills.

2. Institutional setting and background

American football is a team sport played between two teams, each aiming to score points by advancing the ball through running or passing into the opponent's end zone. The team in possession of the ball is given four 'downs' (i.e., attempts) to advance the ball 10 yards toward the end zone. In case of success, the team earns a new set of downs. A 'play' is the action taken during each attempt, from the moment the ball is 'snapped' to the end of the play.

When it comes to organized sports, the NFL is the most prolific professional football league worldwide (next to the National Basketball Association, NBA). Generating an annual revenue of $11.9 billion in 2022 (Ozanian 2023), it is considered the sports league that pays the second-highest average salaries to its athletes. In the 2022 season, each of the 32 NFL franchises was allowed to pay a salary sum of $208.2 million, regulated by the NFL salary cap.

A season is divided into two parts: A regular season of 16 games (17 games after the 2021 season) is followed by a playoff tournament culminating in one final game, crowing the season's champion ('Super Bowl').

In most cases, players are exposed to weather conditions in open-air stadiums. However, there are some stadiums where teams and spectators are protected from the elements, either because they are domed or have a retractable roof. These indoor (or retractable roof) stadiums are kept at 75°F (23.9°C). In the 2019 season, their number was 8. The remaining 23 stadiums are open-air stadiums with only shades for air conditioning. Figure A.1 in the appendix shows where the NFL teams were based in the 2022 season. The only difference compared to the 2019 situation is that two new stadiums were built in Los Angeles (hosting the LA Chargers and the LA Rams) and Las Vegas (hosting the Las Vegas Raiders).

It is important to emphasize that the NFL does not cancel or postpone games due to bad weather conditions, such as extremely high or low temperatures. The only reasons games are postponed or interrupted are hurricanes, thunderstorms, and snowstorms. Therefore, the probability of observing a game in hot-weather conditions is conditionally exogenous and only influenced by geographical localization and the time of year. The ongoing trend towards indoor arenas should not be understood as an action made by teams being aware of the adverse effects of extreme weather on players but as a business move to increase the spectators' comfort and have the opportunity to host the Super Bowl.

3. Data

We use play-by-play data for NFL regular season games for the seasons 2000 – 2019.[1] After excluding games without information on temperature, player, and other play-level details, we are left with a sample of 5,066 games (97.94% of all games). Note that we also discard international games as we cannot define the home team and rule out that other factors than heat may affect the mental performance of the players and hence bias our results. Finally, this gives us a sample of 752,917 individual plays in first, second, third, and fourth downs.

The game-time temperatures, as reported by the NFL, range between 1° (-17°) and 109° (43°) degrees Fahrenheit (Celsius). Figure A.2 in Appendix A illustrates the distribution of temperatures for all games in our sample. We combined the play-by-play data with additional weather data from weather stations near the stadium. This allows us to conduct a placebo treatment for games played in fully air-conditioned domes and closed-roof facilities.[2] We also added the audience size reported by the NFL for each game.[3]

As an indicator of mental errors, we focus on violations of the rules that are sanctioned by a penalty. However, a violation of the rules does not necessarily have to be a mistake; it could also result from deliberate and strategic decisions. For instance, a team may benefit from illegally

[1]All data, including information on games and team-season-specific rosters, is provided online at https://github.com/nflverse/nflverse-data/. We refrain from using playoff games due to the selection of better-performing teams.

[2]The weather and environmental conditions data are provided by https://github.com/ThompsonJamesBliss/WeatherData/tree/master/data.

[3]All attendance numbers are provided by Pro Football Reference at https://www.pro-football-reference.com and reflect ticket sales.

preventing the opposing team from scoring, even at the cost of a penalty. For this reason, we define a group of penalties that are clearly identifiable as a consequence of mental errors, listed in Table A.1 in Appendix A. Illegal actions such as 'illegal formation', 'illegal motion', and 'too many men' do not give the team an advantage but are sanctioned by moving the ball toward the offending team's end zone. Hence, players will always try to avoid these infractions, and violations indicate a lack of attention.[4] In Appendix B, we provide evidence that our group of penalties indeed reduces a team's performance on the play, drive, and game levels.

Note that we do not consider mental errors that could be affected by endogenous factors such as crowd noise. For instance, infractions like a 'false start' can be manipulated by the opposing quarterback, and manipulation is more complicated in the presence of a noisy home crowd (Farnell 2023). The number of spectators and their behavior is, in turn, influenced by weather conditions and, therefore, correlated with our main explanatory variable, i.e., temperature. Although we show in Appendix C that attendance does not vary significantly with high temperatures, we decided on a more conservative approach as fans' behavior may still differ.

A subgroup of penalties is rather associated with aggressive behavior than lack of attention. Examples include 'taunting' or 'unsportsmanlike conduct'. Since prior research has established a link between high temperatures and violent behavior (e.g., Anderson 2001, Jacob et al. 2007, Burke et al. 2015*a*), we pay special attention to this type of mental error in our analysis.

Table 1 presents descriptive statistics for the key variables at the play and game levels. *Heat games* are events characterized by a game-time temperature of more than 85°F (29.4°C). We will rationalize this threshold in the next section. It shows that penalties in general and our group of penalties caused by mental errors are 6 and 17% more likely in heat games than in non-heat games. We take this as the first descriptive evidence from our natural experiment in favor of an adverse effect of heat on cognitive performance. The fact that the average BMI in heat and non-heat games is virtually identical suggests that teams do not respond to high temperatures by changing the team compositions. Moreover, while the difference in 'yards to go' (i.e., the number of yards a team must advance to achieve a first down and reset the downs count) is negligible (0.058 yd / 5.31 cm), the table suggests that games in hot weather are less well attended. However, estimates from a regression model including sport-specific controls indicate that the difference can probably be explained by games in stadiums with a lower average attendance, see Appendix C.

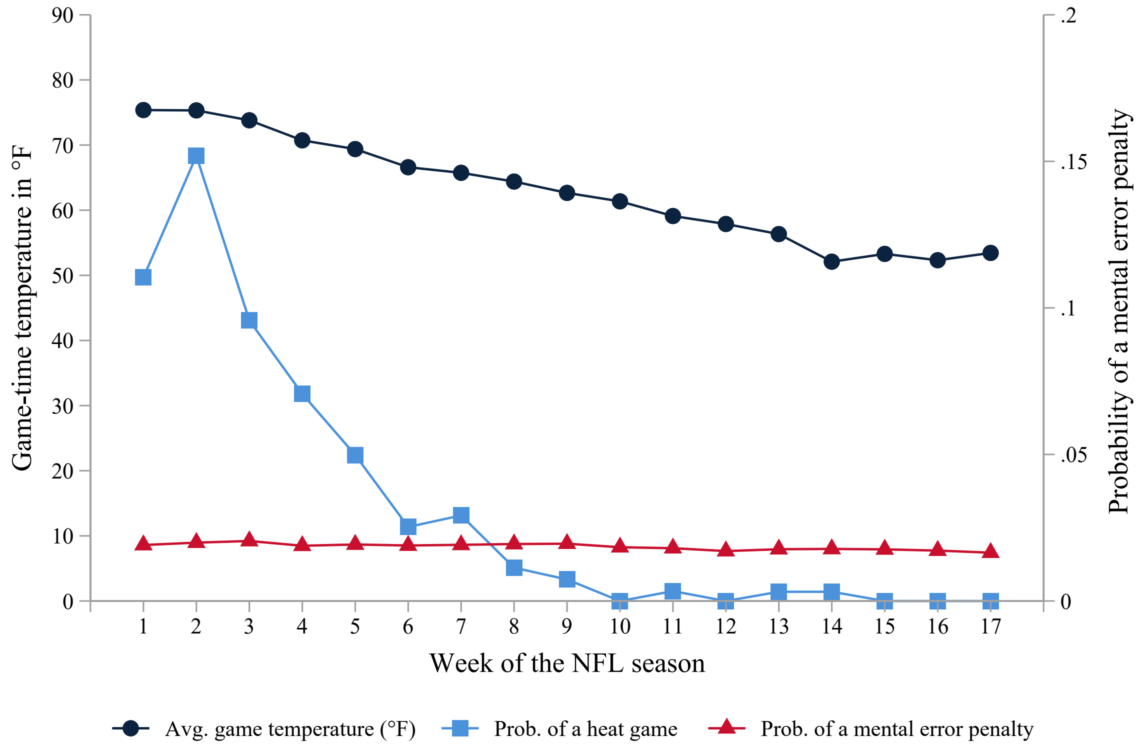
At the game level, Table 1 shows that heat games have about 2.6 fewer plays than non-heat games. While this difference is statistically significant, we do expect little practical significance. The same applies to the average time gap between plays (as a proxy for the speed of action), where the difference is below 0.7 seconds.

Figure 1 illustrates average game-time temperatures throughout a season. Additionally, it

[4]In some situations, coaches and coordinators may also have a responsibility. If this were true, however, it would merely mean an extension of the group whose cognitive performance is impaired by heat stress.

shows that heat games are rare in the second part of a season. On the contrary, the prevalence of penalties associated with infractions due to mental errors hardly changes over the course of the season.

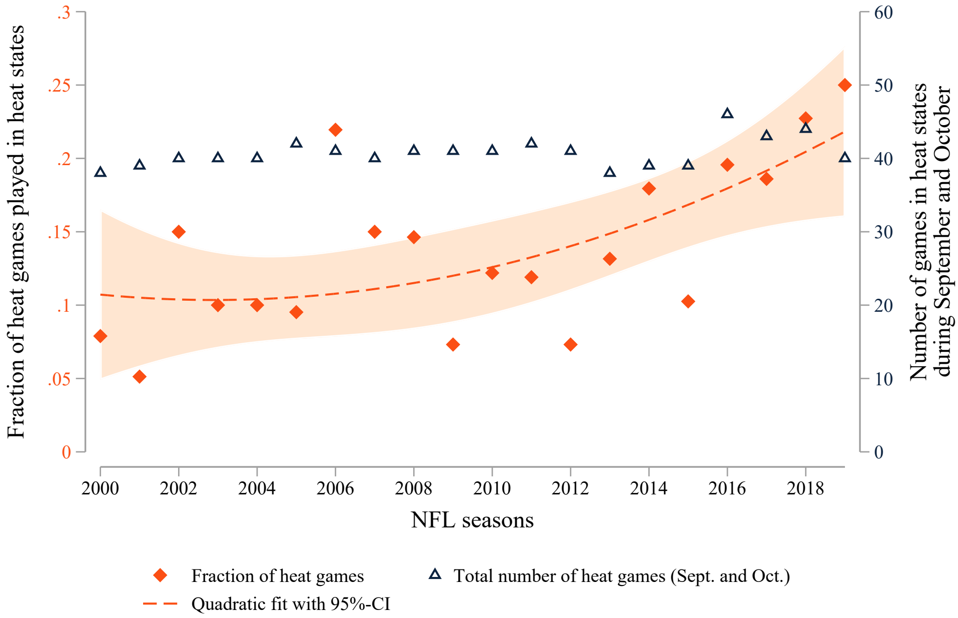
FIGURE 1 — Average temperature and the heat game and penalty probabilities throughout a regular season.



Notes: N = 5,066 NFL regular season games. Games are categorized as heat games when the reported game-time temperature in the stadium is ≥ 85°F. The NFL season features 256 regular-season games, with every team playing 16 games over 17 weeks, from mid-September to early January.

To give an idea about the prevalence of heat games over time, Figure 2 shows the share of heat games in the four states with the highest number of these games: California (CA), Florida (FL), Maryland (MD), and North Carolina (NC). The quadratic fit suggests that the share is increasing over time, and hence, it is of growing relevance to our understanding of the consequences of heat stress. This observation implies that if the NFL organizers were aware of this trend, they would not react to it, for instance, by scheduling fewer home games in the first part of a season in the states exposed most to high temperatures.

Figure 2 — Trends in the frequency of heat games in states most affected by high temperatures



Notes: N = 815 NFL games in the regular seasons 2000 – 2019 played in the states of California (CA), Florida (FL), Maryland (MD), and North Carolina (NC). Regarding the total number of games (right y-axis), we restrict to the first two months of an NFL season (September and October) when heat games are most likely.

Table 1 — Descriptive statistics

	(1) **Heat game**	(2) **Non-heat game**	*Difference*
Penalty (any)	0.085	0.080	0.005***
(1 = yes, 0 = no)	(0.279)	(0.272)	
Penalty indicating a	0.021	0.018	0.003***
mental mistake (1 = yes, 0 = no)	(0.145)	(0.135)	
Penalty associated with	0.009	0.008	0.001
aggressiveness (1 = yes, 0 = no)	(0.093)	(0.089)	
Temperature (in °F)	88.967	61.843	27.124***
	(3.472)	(15.539)	
Body Mass Index (BMI)	32.327	32.250	0.077
	(4.741)	(4.797)	
Yards to go	8.648	8.590	0.058**
	(4.141)	(4.101)	
Game attendance	64.626	67.571	−2.945***
in 1,000 spectators	(12.087)	(8.788)	
Number of plays	24,700	728,217	
Game-level			
Number of plays	146.154	148.707	−2.553***
	(11.726)	(10.890)	
Time gap	27.540	26.863	0.678***
between plays	(1.992)	(2.367)	
Number of games	169	4,897	

Notes: Basic play-level descriptive statistics (mean and standard deviation) for key variables and all weeks, NFL seasons 2000–2019. The statistical significance for the difference between both groups is based on a t-test. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level.

4. Estimation approach and main results

In the first step, we estimate a dose-response type of model of sanctions for infractions associated with mental errors, including eleven temperature bins:

$$Penalty_{i,g} = \alpha_0 + \sum_{k=1}^{6} \beta_k \cdot T_{g,k} + \sum_{k=8}^{11} \beta_k \cdot T_{g,k} \; + \; \xi' \mathbf{X}_{i,g} \; + \; \pi_h \; + \; \rho_t \; + \; \theta_a \; + \; \varepsilon_{i,g} \quad , \qquad (1)$$

where $Penalty_{i,g}$ is a binary variable equal to one if a penalty (related to mental performance) occurs in play i of game g, and zero otherwise. Each variable $T_{g,k}$ represents a binary indicator for the game-time temperature falling within interval k, which corresponds to the following temperature ranges (in °F): ≤ 45, $]45, 50]$, $]50, 55]$, $]55, 60]$, $]60, 65]$, $]65, 70]$, $]75, 80]$, $]80, 85]$, $]85, 90]$, and > 90. The omitted reference category is $]70, 75]$, which includes all games played in climate-controlled environments, such as domes or stadiums with closed roofs.

We add home-team-season fixed effects π_h to the model to control for unobserved characteristics of the home team and account for the fact that heat games are more likely in some places than others. θ_a are away-team fixed effects accounting for some time-invariant characteristics of the away team (e.g., weather conditions at training facilities, fans' travel activities, and market size). Finally, week fixed effects ρ_t control for general trends in players' propensity for mental errors over a season (see Figure 1).

The model also incorporates a large number of football-specific controls. Specifically, the vector **X** includes several game-, play- and team-specific controls, such as the number of previous wins for the home and the away team for all games before game g, the average weight and BMI of the away team, the minute of the game, the position on the field at the start of play i ('line of scrimmage'), the number of rest days for both teams before the game, the number of yards needed for a first down or a touchdown, and game attendance.[5]

Additionally, the vector includes binary variables that control for the down, the sub-quarter[6], the number of previous heat games in the season for both teams, and the score difference before the play from the perspective of the team in possession. We also include dummy variables for the weekday and (local) kick-off time. Referee-crew fixed effects control for unobserved heterogeneity in the referee crews' propensity to call certain penalties.

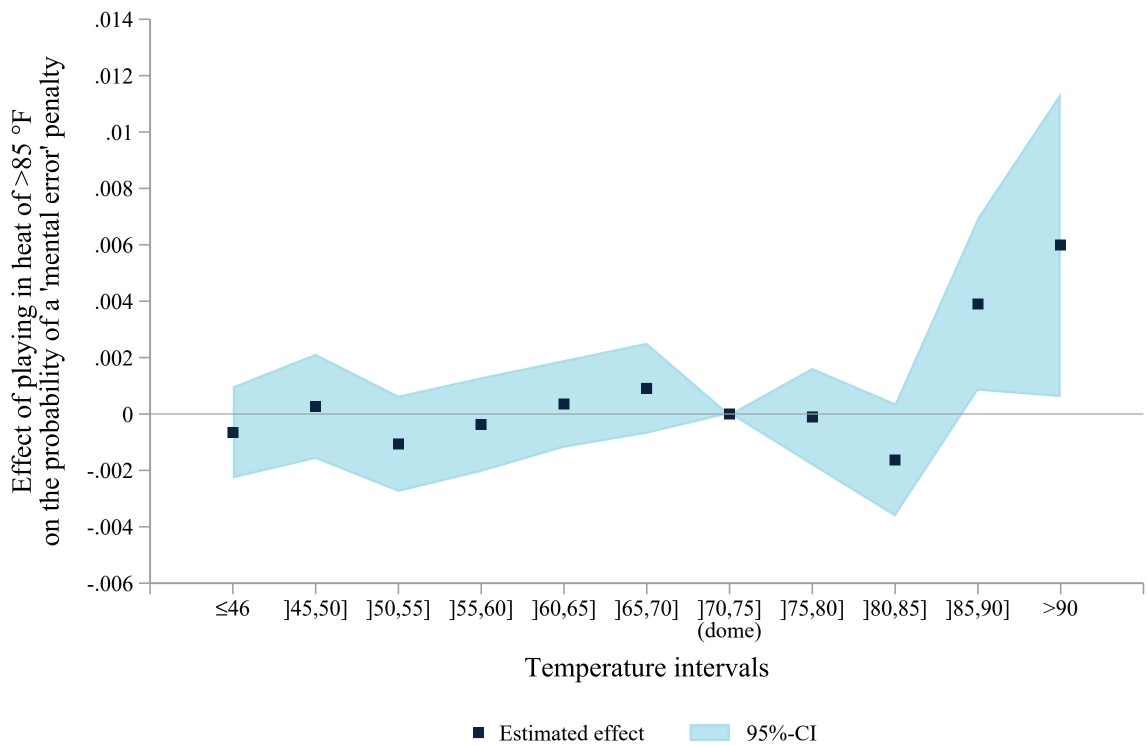
Results Figure 3 presents estimates of the β-coefficients from model (1) for the full sample. We find that the likelihood of a penalty for a mental mistake increases by approximately 0.4 percentage points when game-time temperatures are in the $]85, 90]$ range, compared to the

[5]We use the number of tickets sold as a proxy for demand and actual attendance.

[6]We split each quarter of 15 minutes into two periods with the first period covering all plays with at least 8 minutes still to play, and the second all remaining plays of the quarter. This gives us nine sub-quarters per game, including overtime as sub-quarter 9.

omitted base category of games played in the $]70, 75]$ range. When temperatures exceed 90°F, this effect rises to about 0.6 percentage points. No other significant effects are observed for temperature intervals below or above the reference category, indicating that the negative impact of heat on mental performance becomes pronounced only at temperatures above 85°F. Note that these results perfectly align with the critical values documented in Somanathan et al. (2021).

FIGURE 3 — Main results: effect of temperature on mental errors



Notes: N = 752,917 plays in NFL regular season games season 2000-2019. The base category is the temperature interval $[75, 79]$ (in °F). This category includes all games played in a dome or a stadium with a closed roof.

Since 85°F has proven to be the threshold above which mental errors increase, from now on, we will work with a simplified version of model (1):

$$Penalty_{i,g} = \beta_0 + \beta_1 \mathbb{1}[temp._g > 85] + \xi' \mathbf{X}_{i,g} + \pi_h + \rho_t + \theta_a + \varepsilon_{i,g} \quad . \qquad (2)$$

Here, our explanatory variable of interest is an indicator variable for *heat games*, meaning that the game-time temperature exceeds 85°F (29.4°C).

Table 2 presents our main results. First, it shows that the estimated β_1 of an 'empty model' with our heat game indicator being the only explanatory variable perfectly mirrors the mean differences in the raw data presented in Table 1, supporting the quasi-experimental character of the underlying natural experiment (column (1)).

On the contrary, estimates of the full model indicate that the probability of a penalty associated with a mental mistake increases by 0.44 percentage points in heat games compared to non-heat games (column (2)). Evaluated at the sample mean, this translates into a rise of 23.78%. Note that estimates based on logit regressions are virtually identical and confirm our main results (see Table A.2 in Appendix A).

A possible concern is that heat games are rare beyond week 10 of a season (Figure 1). We, therefore, repeat the exercise with a sample reduced to the first ten weeks of a season (column (3)). This is our preferred model specification. It shows that mental errors are 0.48 percentage points (or 25.73%) more likely in heat games. We take this result as our baseline estimate.

Finally, we find that game-day temperature does not explain variations in mental errors within our model when defined as a continuous variable (column (4)), indicating the absence of a linear and monotonic relationship. Alternatively, Hoffmann et al. (2002) and Ermakov & Krumer (2023) suggest using squared deviations from 57.2°F (or 14°C) – the ideal temperature for competitive sports – as a measure of heat stress. In the absence of a significant effect (column (5)), we also conclude that the relationship between heat and mental errors is not U-shaped.

4.1. Volume effects

Our main results were derived from a model where the outcome variable indicates whether a penalty associated with mental performance (that we will call 'mental error penalty') occurs in a given play. A possible concern is that the number of plays might be lower in heat games due to a slower pace and more resting time. For this reason, we provide an alternative approach where we use the total number of mental error penalties by sub-quarter (we divide each quarter of 15 minutes into two periods, with the first period covering all plays with at least 8 minutes still to play, and the second all remaining plays of the quarter) as the dependent variable. In addition, we regress the heat game indicator on the total number of plays in a sub-quarter to test whether temperature affects the game's pace.

Table 3 shows that the number of mental error penalties is 0.05 (15.98% at the sample mean) higher in games with a temperature above the 85°F threshold compared to games with a temperature below that threshold. For the number of plays, we document a decrease of 0.2 (1% at the sample mean) in heat games. Although statistically significant, the magnitude of the coefficient is small enough to be of little practical relevance. Our conclusion is that heat-related changes in the course of a game do not explain our main finding.

Table 2 — Effect of heat on mental errors: main results

	(1)	(2)	(3)	(4)	(5)
$1[Temp._g > 85]$	0.0030***	0.0044***	0.0048***		
	(0.0010)	(0.0014)	(0.0015)		
Temperature (°F)				1.087×10^{-5}	
				(0.0000)	
$(\text{Temp.} - 57.2)^2$					4.572×10^{-7}
					(0.0000)
Field location[a]		−0.0003***	−0.0003***	−0.0003***	−0.0003***
yard line 1-99		(0.0001)	(0.0001)	(0.0001)	(0.0001)
Time		−0.0014**	−0.0017*	−0.0014**	−0.0014**
(in 10 minutes)		(0.0007)	(0.0009)	(0.0007)	(0.0007)
To go (yards)		0.0002***	0.0002***	0.0002***	0.0002***
		(0.0000)	(0.0001)	(0.0000)	(0.0000)
Game attendance		0.0002	−0.0001	0.0001	0.0001
in 10,000 spectators		(0.0006)	(0.0008)	(0.0006)	(0.0006)
Away team weight[a]		0.0002	−0.0028**	0.0002	0.0001
		(0.0010)	(0.0013)	(0.0010)	(0.0010)
Away team BMI[a]		−0.0008	0.0321***	−0.0010	0.0000
		(0.0083)	(0.0114)	(0.0083)	(0.0083)
Away team rest days		0.0002*	0.0001	0.0002*	0.0001*
		(0.0001)	(0.0001)	(0.0001)	(0.0001)
Away team wins		0.0001	0.0002	0.0001	0.0001
		(0.0001)	(0.0002)	(0.0001)	(0.0001)
Home team wins		0.0002	−0.0001	0.0002	0.0002
		(0.0001)	(0.0003)	(0.0001)	(0.0001)
Home team rest days		−0.0001	−0.0001	−0.0001	−0.0001
		(0.0001)	(0.0001)	(0.0001)	(0.0001)
Home team-season fixed-effects	*no*	*yes*	*yes*	*yes*	*yes*
Week fixed-effects	*no*	*yes*	*yes*	*yes*	*yes*
Add. binary controls	*no*	*yes*	*yes*	*yes*	*yes*
R^2	0.003	0.003	0.003	0.003	0.004
N	752,917	752,917	752,917	752,917	427,445

Notes: The dependent variable is equal to 1 if play *i* results in a penalty associated with a mental mistake, and 0 otherwise; mean 0.0185, std. dev. 0.1349. The sample includes all weeks of the NFL seasons 2000-2019 except for column (3), where only weeks 1–10 are included. In column (5), we use the squared deviation from 57.2 degrees Fahrenheit, as suggested by Hoffmann et al. (2002) and Ermakov & Krumer (2023). *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered on the game level) in parentheses. [a] These variables were divided by 10 to make the estimated coefficients more readable.

TABLE 3 — Effect of heat on mental errors: volume effects

	(1) Number of penalties	(2) Number of plays
$\mathbb{1}[temp._g > 85]$	0.0546***	−0.1972**
	(0.0194)	(0.1001)
Semi-elasticity	[15.98%]	[−1.07%]
Home-team-season fixed effects	*Yes*	*Yes*
Away team fixed effects	*Yes*	*Yes*
Sub-quarter fixed effects	*Yes*	*Yes*
Add. binary controls	*Yes*	*Yes*
Mean dep. var.	0.3417	18.4371
R^2	0.0575	0.7241

Notes: N = 40,837. The dependent variables are the total number of penalties (column (1)) and plays (column (2)) in each sub-quarter. Each quarter is divided into two sub-quarters of 7 and 8 minutes. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level. Standard errors in parentheses are clustered on the game level.

4.2. Robustness

We provide a whole array of robustness checks to ensure the validity and reliability of our main results established in Table 2 (column (3)).

First, we test the sensitivity of our results to different thresholds between 80 to 95°F. Figure 4 suggests that the point estimates are constantly increasing but do not gain statistical significance until 84°F (at the 10-percent level), and they reach a plateau afterward (solid squares).

Second, humidity is an essential factor in heat stress, as it determines the perceived temperature and heat balance of the human organism. This is because when humidity is high, the air contains more moisture, making it more difficult for sweat to evaporate from the skin and cool the body. Therefore, we calculate a heat index based on Steadman (1979) and use this measure to define alternative heat-related thresholds. These new thresholds are shifted upwards on the scale since high humidity and temperatures result in a higher perceived temperature. The hollow squares in Figure 4 give a similar picture as in Figure 3: the point estimates are significantly different from zero and gradually increase above 90°F of perceived temperature.

Third, we present estimates from a model with a different set of fixed effects in Figure 5. Specifically, we use referee and stadium fixed effects instead of home-team-season fixed effects. In addition, we estimate a model similar to our baseline specification but with away-team-season fixed effects instead of the controls for game-specific away-team characteristics. Furthermore, we add more weather controls (humidity and wind speed) to the model. Note that this reduces our sample considerably because this information is only available for a limited number of games. Figure 5 shows that all point estimates are quantitatively similar to the baseline results and significant at the 5 or 1 percent level.

Fourth, when considering all penalty types in the definition of the dependent variable, we find that the point estimate is quantitatively similar but less precise (Figure 5, column (4)). This is expected because the addition of penalty types, which can also be indirectly influenced by high temperatures (e.g., by crowd size and team strategies), should result in more statistical noise.

Fifth, there might be concerns that the increase in mental errors for temperatures over 85°F is affected by games where players are exposed to extremely low temperatures. For instance, Cook & Heyes (2020) show that students at the University of Ottawa perform worse in exams when the outdoor temperature is low. Although our analysis so far does not point in this direction (see Figure 3), we re-estimate model (2) using a sample reduced to games with a temperature ≥ 50°F (≥ 10°C). Figure 5 shows that the point estimate hardly changes compared to the baseline result. We conclude that the low number of games with cold weather before week 11 per season does not bias our results.

Sixth, we account for the fact that for a subgroup of penalties we use as realizations of mental errors, a situation occurs where teams have to repeat the play. Since these plays would not exist without the penalty, they may bias our results. However, Figure 5 (sixth specification) indicates this is not true: excluding plays that directly follow a penalty does not change our results.

Seventh, we conduct a placebo test focusing on the outside temperature for indoor games. For these games, high local temperatures should not affect the propensity of mental errors as they take place in an air-conditioned environment with a temperature of 75°F (23.9°), see Section 2. Due to the low number of indoor games, we use the point estimate based on the full sample regression, including all weeks of a season (already presented in column (4) of Table 2) as a reference value. Figure 5 shows no significant effect of the placebo treatment, which contrasts with the findings of Heyes & Saberian (2019), suggesting that outdoor temperatures influence immigration judges' decisions in indoor (climate-controlled) trials.[7] As the authors point to 'mood' as a moderator of the effect, the placebo test suggests that 'bad mood' due to high outside temperatures does not play a role in indoor performances.

Finally, as another placebo test, we repeat the Figure 5 routine for 'cold weather' games, i.e., temperatures at the lower end of the distribution. Figure A.3 in Appendix A shows that 'cold weather' games do not affect a player's probability of getting sanctioned for a mental mistake, suggesting that large temperature fluctuations per se cannot explain our results.

4.3. Alternative explanations

In this part of the analysis, we explore the importance of factors other than temperature for our main results. First, exposure to sunlight could affect thermoception (i.e., how human beings

[7]Spamann (2022) demonstrates that the effect presented in Heyes & Saberian (2019) shrinks massively after correcting coding errors and extending the sample period.

perceive hot or cold stimuli in their environment) and, hence, moderate the heat effect. Moreover, sunlight may even directly impact performance in our setting if the light blinds players. We, therefore, distinguish between games with strong and weak exposure to sunlight. The latter group includes games that (i) end before 4 pm, (ii) are played indoors, and (iii) have a share of ‘clear sky’ data points of less than 60%. Figure A.4 in Appendix A shows that the heat effect is greater than the baseline result in games where players are heavily exposed to sunlight. However, even for the group of games with less sunlight exposure, the heat effect is still large and quantitatively almost identical to the baseline effect. We conclude that sunlight seems to have an additional effect on mental performance in our setting but does not cause our baseline effect.

To support this interpretation, we carry out an additional placebo test with sun exposure (instead of temperature) as our focus explanatory variable. Here, the indicator variable is equal to one if all weather data points of a game are categorized as ‘clear sky’ and zero if the share is less than 60%. Indoor games and games with a share between 60 and 100% were excluded. We find no significant effect of sun exposure, neither for the whole sample (given that weather data are available) nor for a reduced sample of games with a game temperature of more than 70°F (21°C), see Figure A.4.

Second, we turn to the types of penalties we use to identify mental errors. Specifically, we test whether one of the 22 categories of penalties (listed in Table A.1 in the Appendix) is more important for the heat effect than others. Panel *(a)* of Figure A.5 in Appendix A shows estimates of model (2) in 22 steps, omitting a different category each time (leave-one-out approach). It shows that all estimates are close to the baseline results, indicating that our selection of sanctions is not characterized by one dominant category.

Third, a growing strand of the literature investigates the effect of high temperatures on aggressive behavior and crime (Carleton & Hsiang 2016, Blakeslee & Fishman 2018, Baysan et al. 2019, Blakeslee et al. 2021, Annan-Phan & Ba 2023, Lynott et al. 2023, Colmer & Doleac 2023, Cohen & Gonzalez 2024). A consensus finding seems to be that heat stress tends to cause more crime and violent behavior. Since a subgroup of our selection of penalties is associated with aggression, evaluating their contribution to our main results is important. Therefore, Figure 6 plots estimates of the model specification presented in column (5) of Table 2: first, for a variant where the dependent variable refers to the four types of sanctions associated with aggression only, and second, for a variant where these variables are excluded. We find that the semi-elasticities for aggressive behavior are quantitatively similar to the heat effect for all penalties that are not associated with aggressiveness.

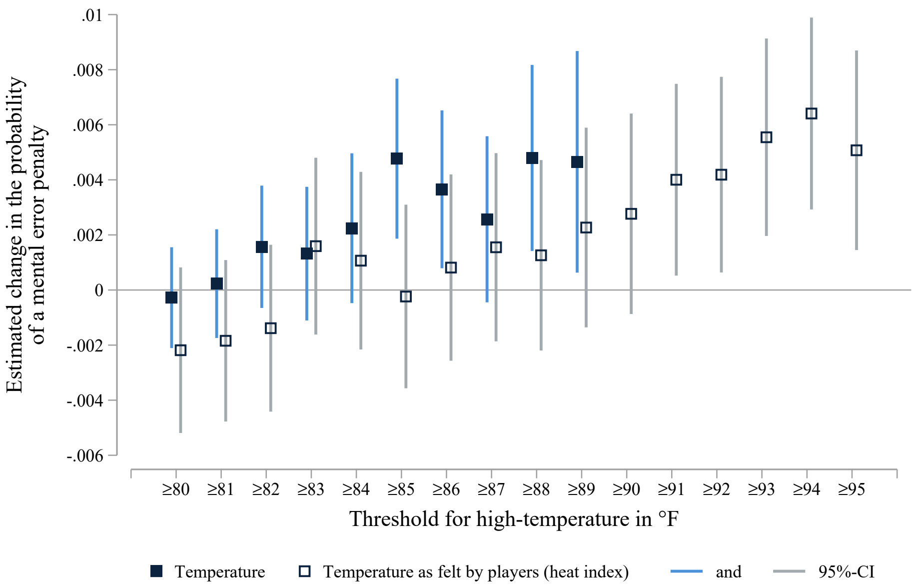
Fourth, there might be a connection between a player’s position and his risk of getting sanctioned. Therefore, we follow a leave-one-out jackknife approach to examine the sensitivity of our main findings regarding each position in American football. Table A.5 in Appendix A suggests that the estimated effect of heat games on the probability of a mental error is not driven by a certain position group.

Fifth, there could be a concern that heat games in our sample are located in four states (California, Florida, Maryland, and North Carolina) and that other state characteristics may explain our results. Therefore, we repeat the leave-one-out jackknife routine, focusing on the state in which the game is played. Panel *(c)* of Figure A.5 in Appendix A illustrates the results. We conclude that the heat effect does not depend on an individual state.

Finally, another possible explanation for our main result could be the adverse effect of heat on referees' performance, which might affect the quality of their decisions. This could lead to an increase in penalties if referees are making more and poorer calls. Unfortunately, we do not have direct measures of referee decision quality in our data. However, in the NFL, teams can challenge referee decisions, with each team allowed up to three challenges per game. A higher rate of successful challenges and overturned decisions would indicate poorer decision quality and weaker referee performance.

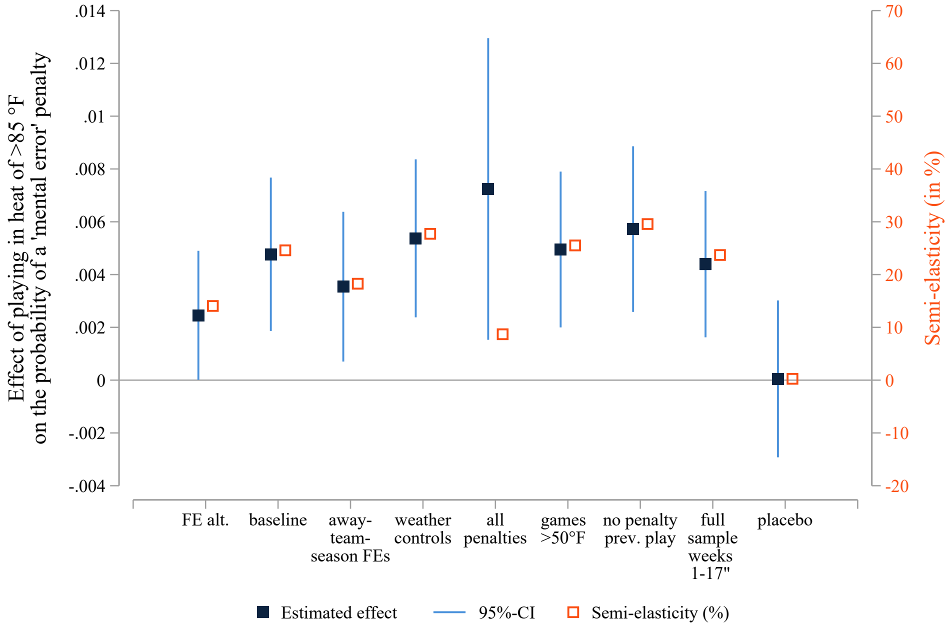
To investigate this, we estimate model 2 using a new dependent variable, which is 1 if a contested decision is upheld and 0 if it is overturned. The sample includes 4,308 contested penalties. The results, presented in Table A.3 in Appendix A, reveal no significant correlation between high temperatures and the quality of referee decisions, either with or without fixed effects. This suggests that referees are unlikely to be a significant factor in the main result.

FIGURE 4 — Effect of heat games on the propensity to make mental errors - alternative thresholds and heat-index



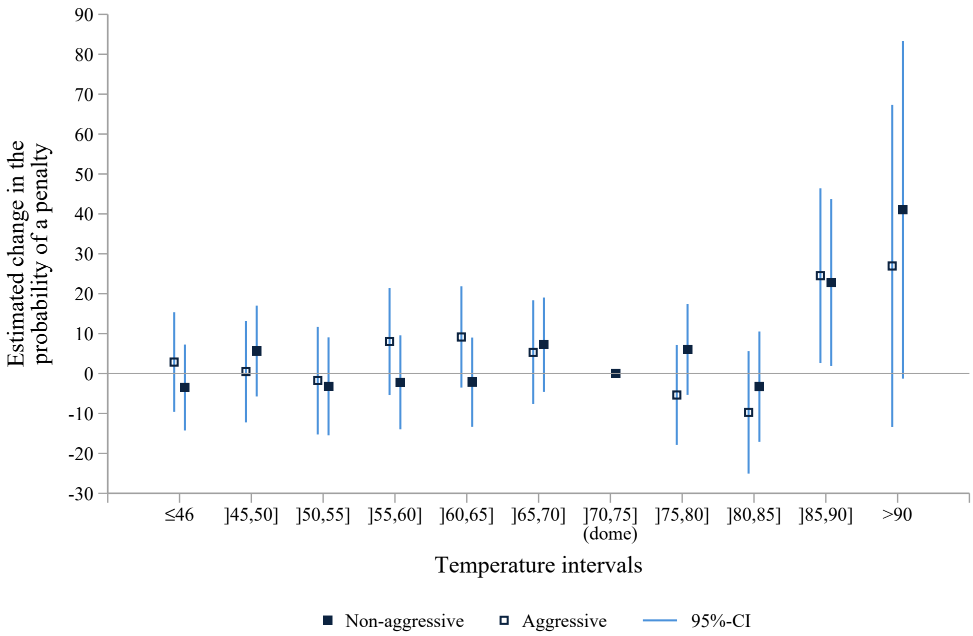
Notes: Estimated effect of heat on the probability that a penalty associated with a mental mistake occurs in play *i* with changing heat thresholds. Each estimate was derived by estimating model 2 with a varying heat indicator (dark blue) and an indicator for temperature as experienced by players as a combination of heat and humidity (heat index). Only weeks 1–11 of season 2000–2019 are considered.

Figure 5 — Sensitivity analysis I



Notes: Estimated effect of heat (game-time temperature >85°F) on the propensity of mental errors. Each square represents estimates of model (2) for a different sub-sample or alternative model specification. For estimate 1, we include the following binary control variable: Stadium fixed effects, week fixed effects, sub-quarter, referee, down, weekday, and game-time fixed effects. For estimates 2 and 4-9, we include the following binary control variables: home-team-season, week, away-team, sub-quarter, referee, down, weekday, number of heat-games, and game-time fixed effects. For estimate 3 we include all fixed effects as for estimate 2, but with away-team-season fixed effects instead of away-team fixed effects.

Figure 6 — Sensitivity analysis IV: aggressive behavior



Notes: Estimated effect of heat (game-time temperature on the propensity of mental errors. The regression model is equivalent to model 1. The penalties associated with aggressiveness are *Roughing the Passer*, *Taunting*, *Unnecessary Roughness*, and *Unsportsmanlike Conduct*. Penalties not associated with aggressiveness are listed in Table A.1.

5. Effect heterogeneity

This section aims to uncover channels that further explain the effect of heat on mental errors. We start with the type of task, the work intensity, and the adaption to high temperatures.

5.1. Tasks, coordination, and adaption

First, regarding the type of task, Figure 7 suggests that there is virtually no difference between sanctions awarded to teams with possession (offense) and defending teams (defense). This means that heat stress equally impairs both types of tasks.

Second, we examine differences in the work intensity. The length of work stretch might affect a team's vulnerability to heat-caused mental error. For this reason, we identify plays that directly follow an intermission. Intermissions occur due to timeouts, injuries, a change of possession, and the start of a new period (quarter). These shorter or longer breaks allow players to coordinate and regain focus in the face of high temperatures. In line with expectations, we find that the effect of heat on mental performance is only present in plays with a high work intensity and absent in plays directly following a break (Figure 7).

Finally, an issue raised in previous studies is the ability to adapt to weather conditions. Extreme heat can be shocking for individuals without prior experience, whereas those used to high temperatures might be less affected. We follow the approach proposed by Fesselmeyer (2021) and identify games where (i) the visiting team, (ii) the home team, or (iii) both teams experienced a large (> 10°F or > 12°C) or moderate (≤ 10°F) change in game-time temperature since the previous game $g-1$. Figure 8 presents the results from that sample split. We find that the probability of getting sanctioned for infractions associated with mental errors increases significantly by approximately 8 percentage points (a semi-elasticity of about 40%) when both teams are exposed to 'heat shocks'. The heat effect is absent for games between teams experiencing a change in temperature of less than 10°F. When we investigate home and away teams separately, it shows that this finding is completely driven by visiting teams who have to deal with a sudden increase in temperature.

5.2. The role of body mass index (BMI)

Another channel through which heat impairs productivity and mental errors is physical conditioning. The body mass index (BMI) is defined as the body weight divided by the square of the body height and is commonly used to define overweight. Moreover, it can be used to assess a person's vulnerability to adverse heat effects (Gildner & Levy 2021). As noted above, within the world of sports, professional football is an ideal testbed to study heat effects for different segments of the BMI distribution because the BMI of NFL players (see Figure A.6 in Appendix A) — compared to the vast majority of other sports — is closer to the BMI distribution of the general population in the United States (Sun et al. 2022).

Estimates of model (2) for four different dependent variables defined by quartiles of the BMI distribution reveal that only 'mental error penalties' by players with a BMI of 31.83 or higher (defined as obese according to the WHO) are affected by heat, see Figure 9.[8] This group has a 40% to 50% higher probability of being sanctioned for infractions associated with mental errors when the game-time temperature exceeds 85°F compared to games with lower temperatures.

A caveat is that NFL players can have the same weight but differ significantly in their body fat percentage on their position. For instance, while 'linemen' tend to be heavier with higher body fat percentages, 'linebackers', and 'tight ends' need a different balance of strength, size, and agility, so that they might have a moderate body fat percentage. Although we already provided evidence in Section 4.3 that our basic result is not driven by a single position group, here we aim to examine the difference between weight and body fat more deeply.

First, we repeat the previous routine with an alternative formula for BMI proposed by Trefethen (2013). Here, the exponent in the denominator is changed from 2 to 2.5 to better reflect how weight relates to height (combined with a scaling factor). The results, illustrated in Figure A.7 in Appendix A, suggest similar effect sizes.

Second, we sort all penalized players into 20 proportional weight categories (ranging from 153 to 390 pounds). After that, we group players according to the median height within their weight category. We expect that relatively taller players have less body fat than equally weighted but shorter players and, hence, should be less vulnerable to the adverse effects of high temperatures (as suggested by Figure 9). The results are presented in Table 4. We document a positive and significant effect of heat on penalties associated with mental errors for below-median-height players, i.e., short players in their weight category. Specifically, short players are 0.26 percentage points (approximately 32% at the sample mean) more likely to receive that kind of sanction in heat games compared to non-heat games. On the contrary, there is no significant association between high temperatures and the likelihood of a penalty for above-median-height players.

5.3. *Timing and the importance of performance*

As a last step, we will focus on two game-specific factors that may moderate the effect of heat on mental performance. First, we take into account that football games are divided into four quarters of fifteen minutes, with short two-minute breaks for changing sides after the first and third quarters and a longer 12-minute halftime break. Figure 10 shows that the heat effect is positive and statistically significant only in the first and third quarters. We explain this finding with a 'heat shock' that sets in when players are exposed to high outdoor temperatures after leaving the fully air-conditioned parts of the stadium (e.g., the locker room). After that, they

[8] Note that we rely on sanctions that can be attributed to a single player.

adapt to the heat. Hence, we interpret this as further evidence for the adverse effect of 'heat shocks' as presented in Section 5.1.

Second, plays vary in importance to teams and, hence, in the incentives to perform. We proxy these incentives using NFL data on expected in-game win probabilities for each play and the team in possession.[9] More precisely, expected win probabilities were divided into three categories: high (win probabilities above 0.6 and below 0.95), medium (win probabilities between 0.3 and 0.6), and low (win probabilities below 0.3 and above 0.05).[10]

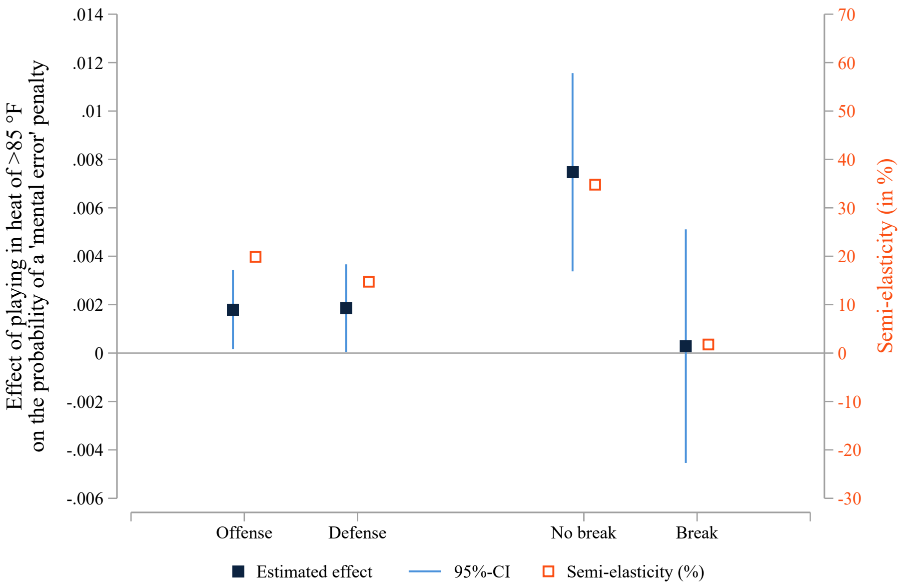
We find that heat-induced mental errors are primarily present in situations of low winning probabilities for the team in possession, see Figure 11. A candidate explanation for this pattern is the stress resulting from being behind and the need to find a way to catch up. That is, trailing teams need to speed up the pace of their offensive actions at the expense of being more prone to heat-induced errors. On the other side, the defense must also adapt to the greater speed of the game. In fact, the average time between the start of two plays (in seconds) for the three groups of win probabilities is 24.5 (low), 26.7 (medium), and 27.8 (high). In other words, the game's speed is more than two seconds, on average, faster when the win probability for the team in possession is the lowest. Referring again to Section 5.1, this finding is consistent with the observation that heat-induced mental errors are more prevalent when the work intensity is high.

---

[9] Information on the underlying statistical model can be found at https://www.pro-football-reference.com/about/win_prob.htm.

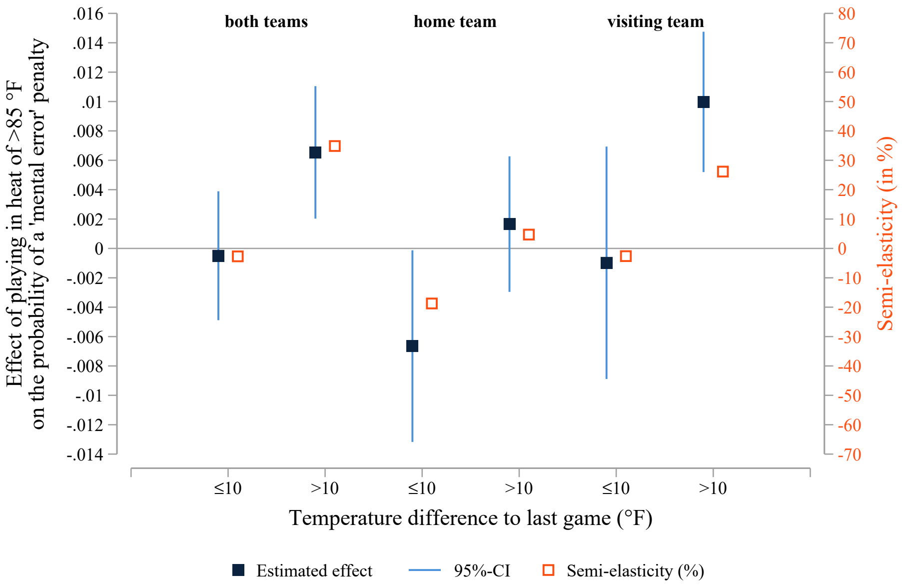
[10] We do not consider plays when the game is very close to being decided at win probabilities ≤ 0.5 and ≥ 0.95. In these situations, players' incentives to focus and avoid mental errors are potentially very low.

FIGURE 7 — Effect of heat games on the propensity to make mental errors - tasks and work intensity



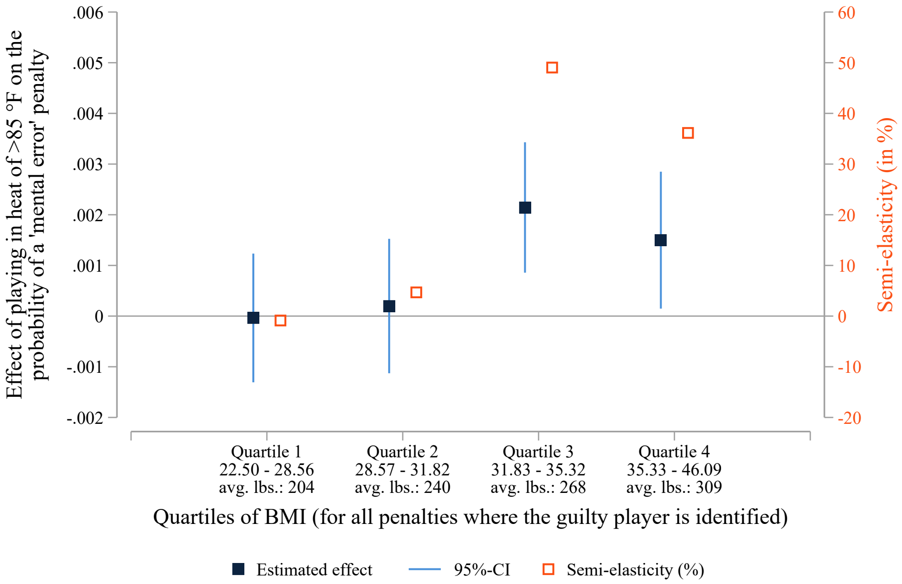
Notes: Estimated effect of heat (game-time temperature >85 °F) on the propensity of mental errors. Estimates are based on model (2) using different dependent variables or sub-samples. For the first estimate, the dependent variable is 1 if the play results in a mental error penalty on offense, 0 otherwise. For the second estimate, the dependent variable is 1 if the play results in a mental error penalty on defense, 0 otherwise. All standard errors are clustered on the game-level.

FIGURE 8 — Effect of heat games on the propensity to make mental errors - temperature differences between previous and actual game



Notes: Estimated effect of heat (game-time temperature >85 °F) on the propensity of mental errors. Estimates are based on model (2) using different sub-samples. All standard errors are clustered on the game level.

FIGURE 9 — Effect of heat games on the propensity to make mental errors - the role of the body mass index (BMI)



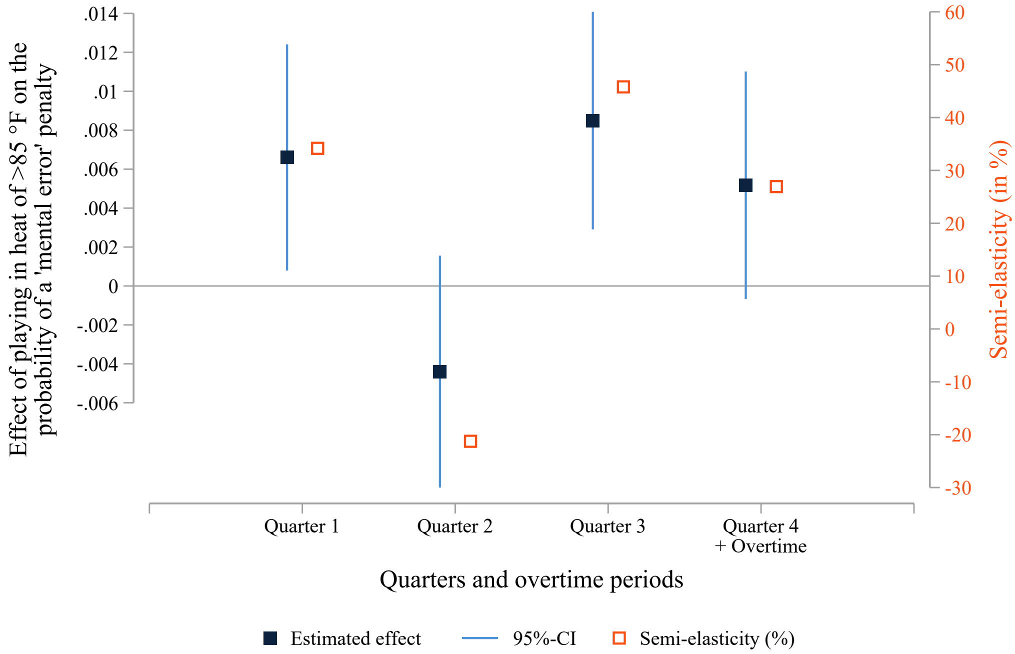
Notes: Estimates of model (2) for four different samples defined by quartiles of the BMI distribution. Each estimate is derived by a separate regression of model 2. The dependent variable for each estimation is equal to 1 if the observed play results in a mental error penalty by a player within the indicated BMI interval, 0 otherwise. The sample is reduced to sanctions that can be attributed to a player.

TABLE 4 — Effect of heat on mental errors: height-weight relationship

	(1) Rel. short players	(2) Rel. tall players
$\mathbb{1}[temp._g > 85]$	0.0040***	0.0007
	(0.0012)	(0.0008)
Semi-elasticity	[30.11%]	[12.42%]
Home-team-season fixed-effects	*Yes*	*Yes*
Visiting team fixed-effects	*Yes*	*Yes*
Sub-quarter fixed-effects	*Yes*	*Yes*
Add. binary controls^a	*Yes*	*Yes*
Mean dep. var.	0.0133	0.0060
R^2	0.7241	0.0575

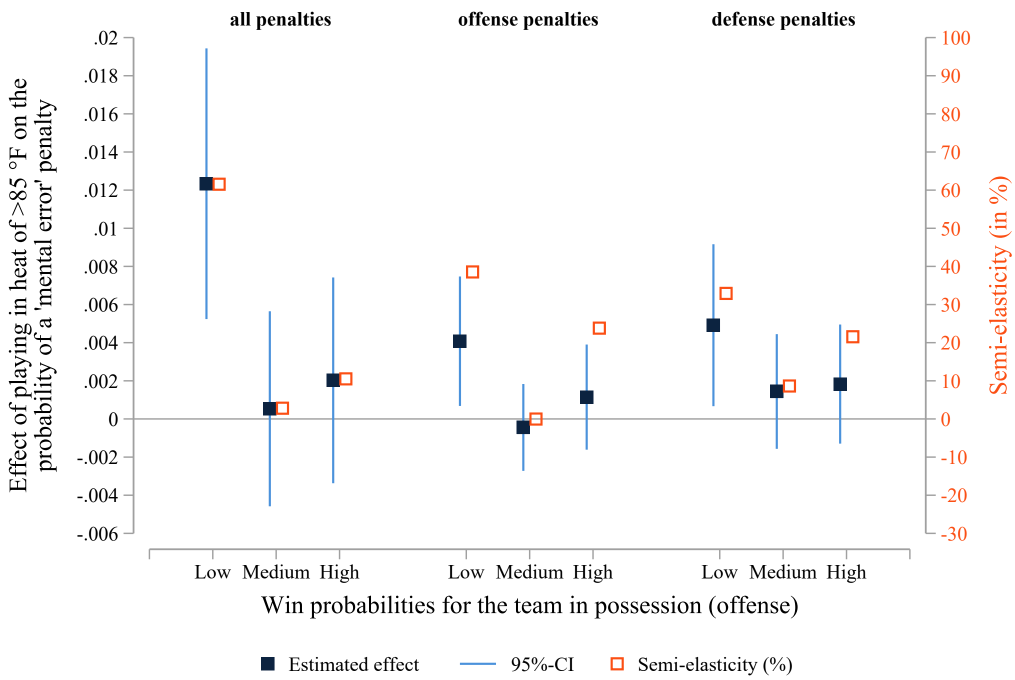
Notes: N = 427,445. The dependent variable is equal to one if a penalty (associated with mental performance) occurs in play i of game g, and zero otherwise. The sample is split into players below (column (1)) and above (column (2)) the median height in their weight category. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level. Standard errors in parentheses are clustered on the game level. ^{a} Additional controls as presented in Section 4.

FIGURE 10 — Effect of heat games on the propensity to make mental errors - timing



Notes: Estimated effect of heat (game-time temperature >85 °F) on the propensity of mental errors. Estimates are based on model (2) using different sub-samples. All standard errors are clustered on the game-level.

FIGURE 11 — Effect of heat games on the propensity to make mental errors - importance to perform



Notes: Estimated effect of heat (game-time temperature >85 °F) on the propensity of mental errors. Estimates are based on model (2) using different sub-samples. All standard errors are clustered on the game-level. Expected win probabilities for the team in possession are provided by the NFL and calculated based on a model presented at https://www.pro-football-reference.com/about/win_prob.htm.

6. Conclusion

This article uses elite American football as a testbed for studying the causal effect of heat stress on agents' cognitive abilities. Specifically, we investigate whether players are more prone to mental errors when exposed to high temperatures. We find a 25.73% increase in the probability of being sanctioned for infractions associated with mental errors when the game-time temperature exceeds 85°F (29.4°C) compared to games with lower temperatures. This also applies to a group of offenses associated with aggressive behavior.

Further analyses identify moderators of the heat effect. We find heat-induced mental errors to be mainly present in situations with (i) little room to adapt to heat stress and (ii) high work intensity. Moreover, we take advantage of the fact that the BMI distribution of NFL players – unlike most other sports – is more similar to the BMI distribution of the general population in the United States. Our results indicate that the group of players in the top segment of the distribution, i.e., individuals with a BMI of 30 or more, are vulnerable to more mental errors in heat games.

Since our setting involves highly trained and monitored agents with high incentives to perform well, we are aware that our results may represent a lower bound for the population-wide effect. For this reason, and because mental errors may have more severe consequences in other labor market segments such as construction and emergency medical care, our study demonstrates how urgent it is for societies to invest in measures to protect workers against heat, both on a technical and legal level. This is particularly true given global warming, with rising average temperatures and more extreme heat events across the globe. For instance, our results stress the importance of rest breaks at work for mitigating heat-induced mental errors and raising awareness towards vulnerable groups.

Additionally, if climate control measures are considered environmentally unsustainable due to their high energy consumption, it may be beneficial to explore programs aimed at reducing obesity. By improving physical fitness, such programs have the potential to enhance individuals' resilience to heat stress, thereby contributing to a safer and more productive work environment.

References

Anderson, C. A. (2001), 'Heat and violence', *Current Directions in Psychological Science* **10**(1), 33–38.

Annan-Phan, S. & Ba, B. A. (2023), 'Hot Temperatures, Aggression, and Death at the Hands of the Police: Evidence from the US', *Journal of Urban Economics* **142**, 103592.

Baysan, C., Burke, M., González, F., Hsiang, S. & Miguel, E. (2019), 'Non-economic factors in violence: Evidence from organized crime, suicides and climate in Mexico', *Journal of Economic Behavior & Organization* **168**, 434–452.

Biardeau, L. T., Davis, L. W., Gertler, P. & Wolfram, C. (2020), 'Heat exposure and global air conditioning', *Nature Sustainability* **3**(1), 25–28.

Blakeslee, D., Chaurey, R., Fishman, R., Malghan, D. & Malik, S. (2021), 'In the heat of the moment: Economic and non-economic drivers of the weather-crime relationship', *Journal of Economic Behavior & Organization* **192**, 832–856.

Blakeslee, D. S. & Fishman, R. (2018), 'Weather shocks, agriculture, and crime: Evidence from India', *Journal of Human Resources* **53**(3), 750–782.

Burke, M., Hsiang, S. M. & Miguel, E. (2015*a*), 'Climate and conflict', *Annual Review of Economics* **7**(1), 577–617.

Burke, M., Hsiang, S. M. & Miguel, E. (2015*b*), 'Global non-linear effect of temperature on economic production', *Nature* **527**(7577), 235–239.

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

Carleton, T. A. & Hsiang, S. M. (2016), 'Social and economic impacts of climate', *Science* **353**(6304), aad9837.

Cohen, F. & Gonzalez, F. (2024), 'Understanding the link between temperature and crime', *American Economic Journal: Economic Policy* **16**(2), 480–514.

Colmer, J. & Doleac, J. L. (2023), 'Access to guns in the heat of the moment: more restrictive gun laws mitigate the effect of temperature on violence', *Review of Economics and Statistics (forthcoming)* .

Cook, N. & Heyes, A. (2020), 'Brain freeze: outdoor cold and indoor cognitive performance', *Journal of Environmental Economics and Management* **101**, 102318.

Costello, A., Romanello, M., Hartinger, S., Gordon-Strachan, G., Huq, S., Gong, P., Kjellstrom, T., Ekins, P. & Montgomery, H. (2023), 'Climate change threatens our health and survival within decades', *The Lancet* **401**(10371), 85–87.

Dell, M., Jones, B. F. & Olken, B. A. (2012), 'Temperature shocks and economic growth: Evidence from the last half century', *American Economic Journal: Macroeconomics* **4**(3), 66–95.

Dell, M., Jones, B. F. & Olken, B. A. (2014), 'What do we learn from the weather? The new climate-economy literature', *Journal of Economic Literature* **52**(3), 740–798.

Diffenbaugh, N. S. & Burke, M. (2019), 'Global warming has increased global economic inequality', *Proceedings of the National Academy of Sciences* **116**(20), 9808–9813.

Dillender, M. (2021), 'Climate change and occupational health: Are there limits to our ability to adapt?', *Journal of Human Resources* **56**(1), 184–224.

Ebi, K. L., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., Honda, Y., Kovats, R. S., Ma, W., Malik, A. et al. (2021), 'Hot weather and heat extremes: health risks', *The Lancet* **398**(10301), 698–708.

Ermakov, S. & Krumer, A. (2023), 'Saturday in the stadium: on higher attendance on Saturdays in Norwegian Eliteserien soccer league', *European Sport Management Quarterly* **23**(6), 1800–1818.

Farnell, A. (2023), 'False start? An analysis of NFL penalties with and without crowds', *Journal of Sports Economics* **24**(6), 695–716.

Fesselmeyer, E. (2021), 'The impact of temperature on labor quality: Umpire accuracy in major league baseball', *Southern Economic Journal* **88**(2), 545–567.

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

Gildner, T. E. & Levy, S. B. (2021), 'Intersecting vulnerabilities in human biology: Synergistic interactions between climate change and increasing obesity rates', *American Journal of Human Biology* **33**(2), e23460.

Graff Zivin, J., Hsiang, S. M. & Neidell, M. (2018), 'Temperature and human capital in the short and long run', *Journal of the Association of Environmental and Resource Economists* **5**(1), 77–105.

Harp, J. B. & Hecht, L. (2005), 'Obesity in the National Football League', *Jama* **293**(9), 1058–1062.

Heal, G. & Park, J. (2016), 'Reflections—temperature stress and the direct impact of climate change: a review of an emerging literature', *Review of Environmental Economics and Policy* **10**(2), 347–362.

Heyes, A. & Saberian, S. (2019), 'Temperature and decisions: evidence from 207,000 court cases', *American Economic Journal: Applied Economics* **11**(2), 238–265.

Hoffmann, R., Ging, L. C. & Ramasamy, B. (2002), 'The socio-economic determinants of international soccer performance', *Journal of Applied Economics* **5**(2), 253–272.

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

Jacob, B., Lefgren, L. & Moretti, E. (2007), 'The dynamics of criminal behavior: Evidence from weather shocks', *Journal of Human Resources* **42**(3), 489–527.

Kabore, P. & Rivers, N. (2023), 'Manufacturing output and extreme temperature: Evidence from Canada', *Canadian Journal of Economics/Revue Canadienne d'Économique* **56**(1), 191–224.

Keivabu, R. & Widmann, T. (2024), 'The effect of temperature on language complexity: Evidence from seven million parliamentary speeches', *iScience* **27**(6), 110106.

Lai, W., Qiu, Y., Tang, Q., Xi, C. & Zhang, P. (2023), 'The effects of temperature on labor productivity', *Annual Review of Resource Economics* **15**(1), 213–232.

LoPalo, M. (2023), 'Temperature, worker productivity, and adaptation: evidence from survey data production', *American Economic Journal: Applied Economics* **15**(1), 192–229.

Lynott, D., Corker, K., Connell, L. & O'Brien, K. (2023), 'The effects of temperature on prosocial and antisocial behaviour: A review and meta-analysis', *British Journal of Social Psychology* **62**(3), 1177–1214.

Mackworth, N. H. (1946), 'Effects of heat on wireless operators', *British Journal of Industrial Medicine* **3**(3), 143.

Ozanian, M. (2023), 'NFL National Revenue Was Almost $12 Billion In 2022', *Forbes* . 11 July 2023.
URL: *https://www.forbes.com/sites/mikeozanian/2023/07/11/nfl-national-revenue-was-almost-12-billion-in-2022/*

Park, J. & Pankratz, N. (2021), 'Temperature, workplace safety, and labor market inequality', *IZA Discussion Paper No. 14560* .

Park, R. J. (2022), 'Hot temperature and high-stakes performance', *Journal of Human Resources* **57**(2), 400–434.

Park, R. J., Goodman, J., Hurwitz, M. & Smith, J. (2020), 'Heat and learning', *American Economic Journal: Economic Policy* **12**(2), 306–339.

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

Qiu, Y. & Zhao, J. (2022), 'Adaptation and the distributional effects of heat: Evidence from professional archery competitions', *Southern Economic Journal* **88**(3), 1149–1177.

Somanathan, E., Somanathan, R., Sudarshan, A. & Tewari, M. (2021), 'The impact of temperature on productivity and labor supply: Evidence from Indian manufacturing', *Journal of Political Economy* **129**(6), 1797–1827.

Spamann, H. (2022), 'Comment on "Temperature and Decisions: Evidence from 207,000 Court Cases"', *American Economic Journal: Applied Economics* **14**(4), 519–528.

Steadman, R. G. (1979), 'The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science', *Journal of Applied Meteorology and Climatology* **18**(7), 861–873.

Sun, J.-Y., Huang, W.-J., Hua, Y., Qu, Q., Cheng, C., Liu, H.-L., Kong, X.-Q., Ma, Y.-X. & Sun, W. (2022), 'Trends in general and abdominal obesity in us adults: Evidence from the national health and nutrition examination survey (2001–2018)', *Frontiers in Public Health* **10**, 925293.

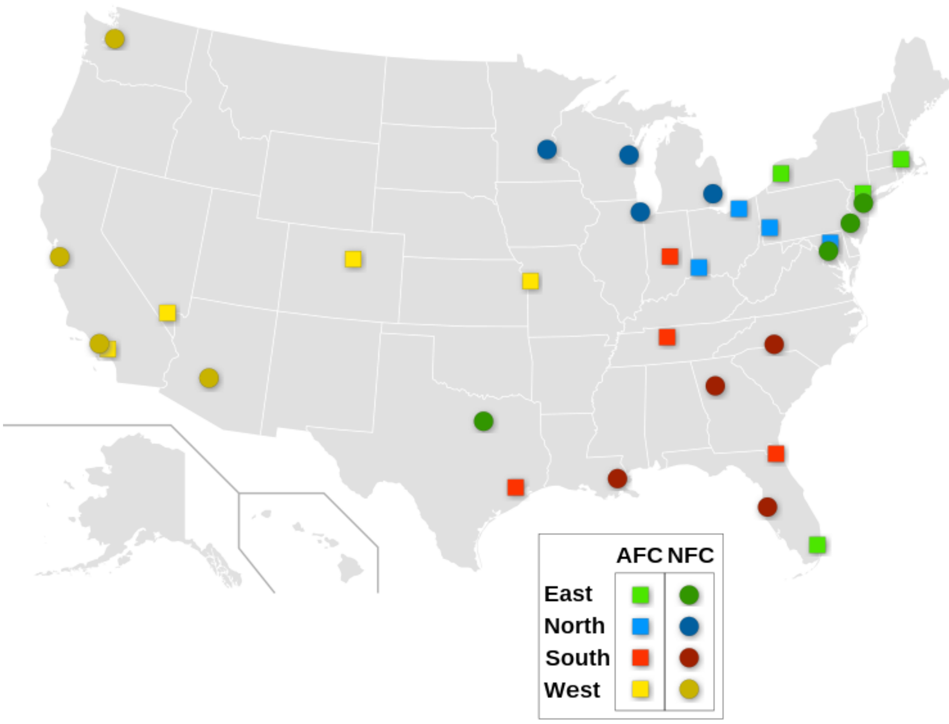
Taylor, L., Watkins, S. L., Marshall, H., Dascombe, B. J. & Foster, J. (2016), 'The impact of different environmental conditions on cognitive function: a focused review', *Frontiers in Physiology* **6**, 372.

Trefethen, N. (2013), 'Letter to the Editor'. The Economist. Accessed: 2024-06-07.
URL: *https://www.economist.com/letters/2013/01/05/on-obesity-gun-control-syria-bankers-marriage*

Zander, K. K., Botzen, W. J., Oppermann, E., Kjellstrom, T. & Garnett, S. T. (2015), 'Heat stress causes substantial labour productivity loss in Australia', *Nature Climate Change* **5**(7), 647–651.

A. Additional Figures and Tables

FIGURE A.1 — NFL team locations in 2022



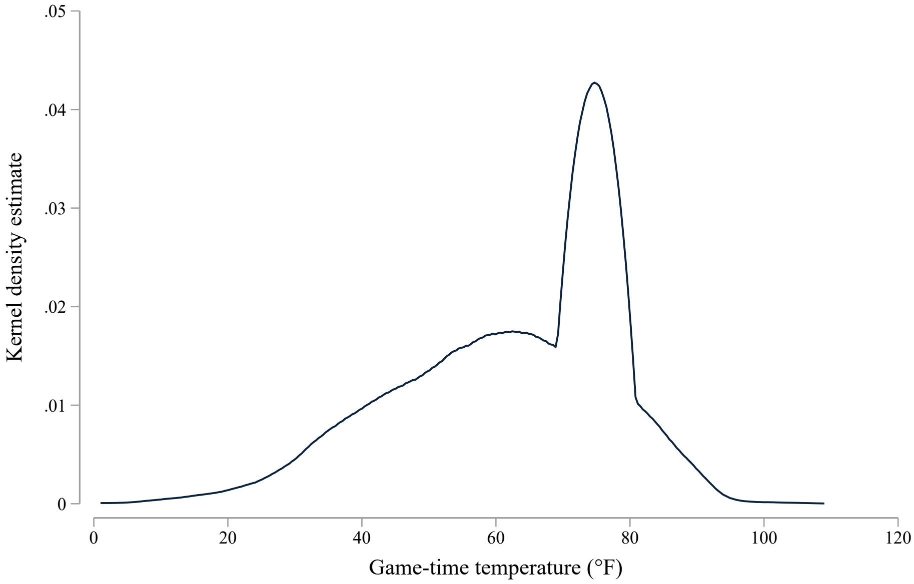
Notes: Taken from Wikipedia.org (07.11.2023).

TABLE A.1 — List of all penalty types used as indicators for mental errors

Defensive 12 On-field	3.04	Ineligible Downfield Pass	2.68
Defensive Delay of Game	0.67	Invalid Fair Catch Signal	0.18
Defensive Offside	25.10	Offensive 12 On-field	0.63
Def. Too Many Men on Field	0.99	Off. Too Many Men on Field	0.29
Delay of Game	6.63	Offside on Free Kick	2.53
Delay of Kickoff	0.03	Roughing the Passer	12.66
Illegal Formation	6.32	Taunting	2.31
Illegal Forward Pass	1.04	Unnecessary Roughness	21.18
Illegal Motion	1.82	Unsportsmanlike Conduct	5.61
Illegal Procedure	0.12		
Illegal Shift	3.06		
Illegal Substitution	1.55		
Ineligible Downfield Kick	1.58		

Notes: Penalty types and relative frequency used as indicators for mental errors. All categories as defined in our data and regulated by NFL rules during the 2000–2019 season. If two or more penalties were called during a play, only the penalty that was enforced was considered.

Figure A.2 — Distribution of game-time temperatures as reported by the NFL



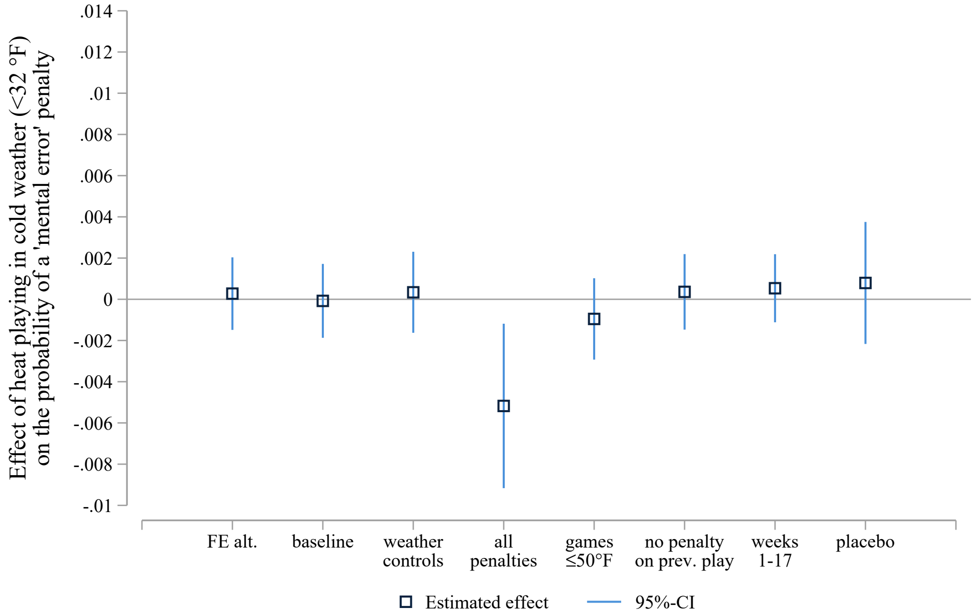
Notes: Kernel density estimate of game-time temperature in °F.

Table A.2 — Effect of heat on mental errors: LOGIT estimates

	(1)	(2)
$1[\text{Temp.}_g > 85]$	0.0043***	0.0048***
	(0.0013)	(0.0014)
Full set of controls	*yes*	*yes*
N	427,445	752,917

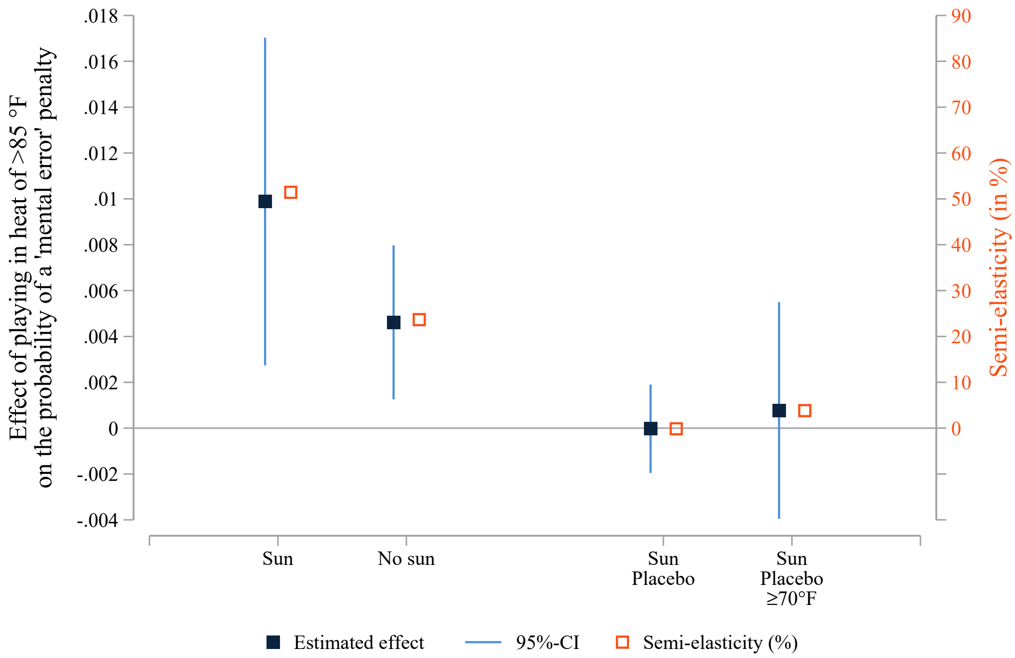
Notes: Estimates derived from LOGIT models. The dependent variable is equal to 1 if play i results in a penalty associated with a mental mistake, and 0 otherwise. The sample for columns (2)) includes all weeks of the NFL seasons 2000-2019. For column (1), only weeks 1–10 are included. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level and 1-percent level, respectively. Robust standard errors (clustered on the game level) in parentheses. All controls as in Table2 are included.

Figure A.3 — Effect of cold-weather games on the propensity to make mental errors



Notes: Estimated effect of cold temperatures on the probability that a penalty associated with a mental mistake occurs in play i. Only weeks 11–17 of NFL season 2000 – 2019 are considered. Cold weather games are defined as games with an official game-time temperature of < 32°F.

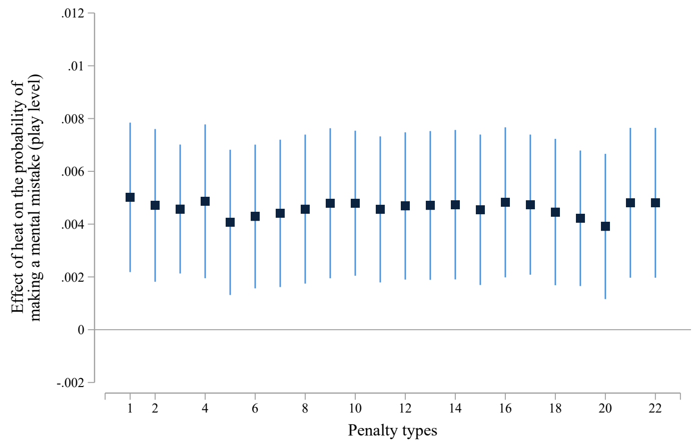
Figure A.4 — Sensitivity analysis II - other weather components



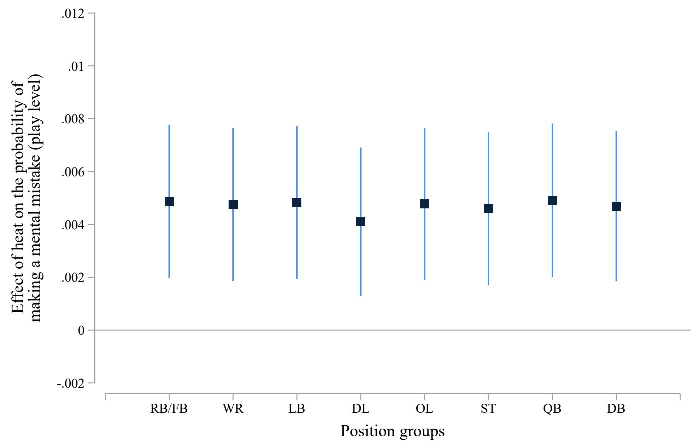
Notes: Estimated effect of heat (game-time temperature >85 °F) on the propensity of mental errors. Each square represents estimates of model (2) for a different sub-sample.

Figure A.5 — Sensitivity checks: The role of penalty categories, individual states, and positions.

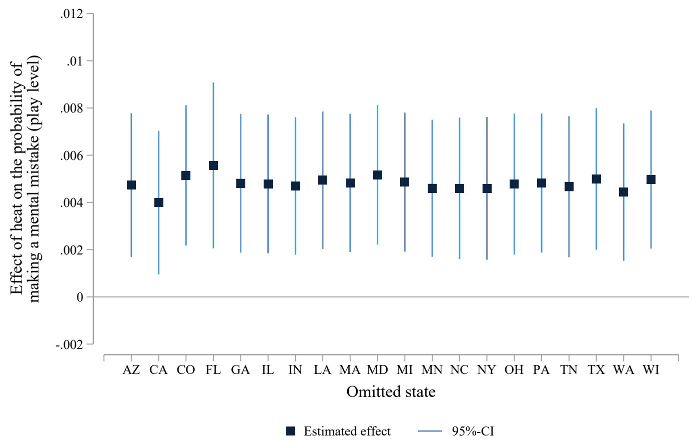
(a) Penalty category



(b) States



(c) States



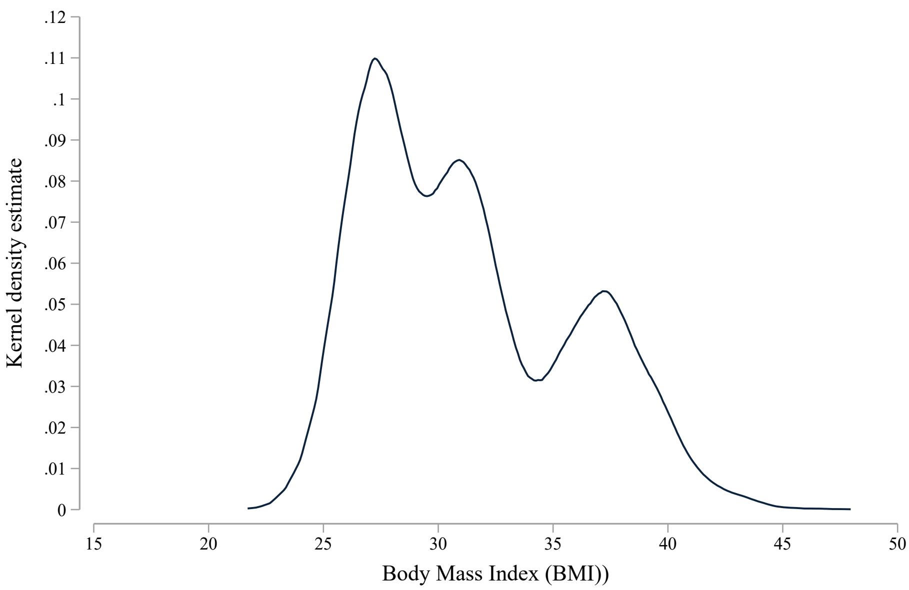
Notes: Estimated effect of heat on the probability of a mental mistake in play i. Each square in panel *(a)* represents estimates of model (2) for a different sub-sample. One of the 22 penalty types was omitted from the sample for each estimate. For panel *(b)*, we define our dependent variable as 1 if a 'mental error penalty' occurred in play i, and 0 if no penalty occurs or the penalty refers to a player of the excluded position group. "RB-FB" denotes Running Backs and Fullbacks, "WR" denotes Wide Receivers, "LB" denotes Linebackers (including Outside Linebackers, Insider Linebackers, and EDGE players), "DL" denotes Defensive Line players (including Nose Tackles, Defensive tackles, and Defensive Ends), "OL" indicates Offensive Line players (including Offensive tackles, Guards, and Centers), "ST" denotes Special Team players (including Kickers and Long-snappers), "QB" denotes Quarterbacks, and "DB" denotes Defensive Backs (including Cornerbacks and Safeties). For panel *(c)*, we step-wise omit the corresponding state from the overall sample of NFL games during seasons 2000-2019.

Table A.3 — Alternative explanation: incorrect referee decisions

	(1)	(2)	(3)	(4)
Temperature (°F)	0.0001		−0.0002	
	(0.0005)		(0.0006)	
$1[Temp._g > 85]$		0.0224		0.0518
		(0.0383)		(0.0403)
Home team-season fixed-effects	*no*	*no*	*yes*	*yes*
Week fixed-effects	*no*	*no*	*yes*	*yes*
Add. binary controls	*no*	*no*	*yes*	*yes*
R^2	0.0822	0.0823	0.1467	0.1470
N	4,308	4,308	4,308	4,308

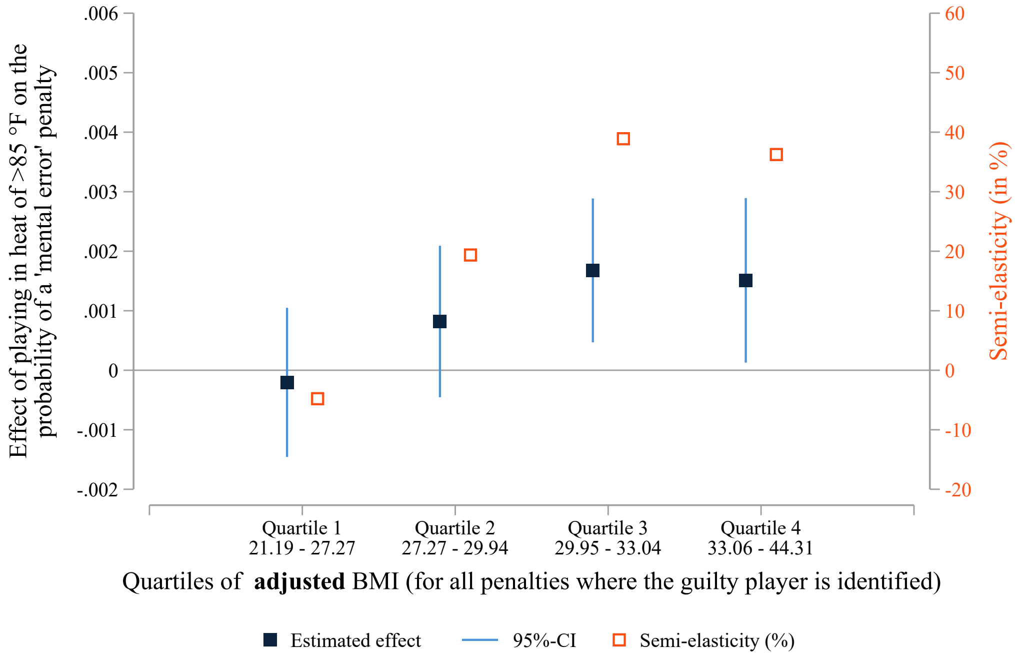
Notes: The dependent variable is equal to 1 if the challenged referee decision was upheld, and 0 if it was overturned (mean 0.396). *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered on the game level) in parentheses.

Figure A.6 — Distribution of BMI for all players



Notes: Kernel density estimate of body mass index (BMI) for all players in a team during seasons 2000–2019 and 2021–2022.

FIGURE A.7 — Effect of heat games on the propensity to make mental errors - the role of the BMI (alternative measure)



Notes: Estimates of model (2) for four different samples defined by quartiles of the BMI distribution. We use the alternative formula for the BMI proposed by Trefethen (2013). The sample is reduced to sanctions that can be attributed to a player.

Web Appendix

B. The relationship between penalties and performance

In American football (like in most other sports), inputs are transformed into intermediate and end products. Players and coaches invest efforts and skills to produce yards, points, and, ultimately, wins. To investigate how a restriction in cognitive abilities translates into performance, we examine the correlation between our group of penalties associated with mental errors (see Section 4) and success at the play, drive, and game levels.

Play level In the first step, we estimate a variant of model (2) in which success in a play is explained by a penalty associated with mental error and awarded either to the team with possession (offense) or the defending team (defense) in the previous play.

The results are presented in Table B.1. We find a significant negative correlation between 'mental error penalties' in play $t-1$ and all three indicators of success in play t for the team with possession: A penalty of this kind is associated with approximately half a yard less gained (column 1), a decrease in the *expected point added* (EPA) by 0.09 (column 3) and a decrease in the win probability by 1.7 percentage point (column 5). In situations when the defending team is sanctioned with a 'mental error penalty' in play $t-1$, the team with possession benefits from an increase of 0.7 (column 2) in the net yards gained and an increase of 0.5 in the EPA (column 4). Winning probabilities are not affected (column 6).

TABLE B.1 — The association of 'mental error penalties' and success: play-level

	Net yards gained play i		*Δ EPA from $(i-1)$ to i*		*Δ win probability $(i-1)$ to i*	
	(1)	(2)	(3)	(4)	(5)	(6)
Own penalty $(i-1)$	−0.5495*		−0.0854***		−0.0156***	
(offense)	(0.3221)		(0.0185)		(0.0025)	
Opponent penalty $(i-1)$		0.6660**		0.4537***		0.0000
(defense)		(0.2623)		(0.0121)		(0.0017)
Previous down indicator	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Full set of controls[a]	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Mean dep. var.	39.0906		-0.0993		0.0115	
R^2	0.2509	0.2509	0.3008	0.3026	0.1333	0.1332

Notes: N = 645,415; *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Standard errors–clustered on the game level–in round parentheses. All plays during the NFL regular seasons 2000 − 2019 are included. The dependent variable for columns (1) and (2) is the yards gained in the actual play for the team in possession. Since negative consequences are also taken into account, we use the term "net" yards. The dependent variable in columns (3) and (4) is the *expected point added* (EPA) by play i compared to the EPA after play $(i-1)$. For columns (5) and (6), the dependent variable is the increase in win probabilities between play i and $(t-1)$. The NFL calculates win probabilities according to the method described at https://www.pro-football-reference.com/about/win_prob.htm.
[a] The full set of control variables is equivalent to our main estimation model (2).

Drive level In American football, a 'drive' is a series of plays by the offensive team, starting when they gain possession of the ball and continuing until they score, turn the ball over to the other team, or the period ends. Since the offensive team aims at advancing the ball down the field to score points, the success of a drive can be measured by the number of yards gained and the points scored.[11]

A 'mental error penalty' may not only impair a team's performance at the play level but also adversely affect the overall drive. For our analysis, we aggregate the data at the drive level (110,000 observations) and explain the offense's performance with two binary indicators, which are equal to 1 if the offensive team (defending team) is sanctioned by at least one 'heat stress penalty' during the drive, and zero otherwise, in a regression framework. Table B.2 indicates that drives with one or more 'heat stress penalties' for the team with possession are characterized by one play less (column 1), 11 fewer yards gained – also in terms of field position – (columns 2 and 3), a 16 percentage points higher probability to result in a punt, i.e., deliberately giving up possession (column 4), and a 12 percentage points lower scoring probability (column 5). Estimates for 'heat stress penalties' for the defending team are similar but, as expected, have a different sign.

Table B.2 — The association of penalties and success: drive-level

	(1) **Plays**	(2) **Yards gained**	(3) **Field position**	(4) **Punt**	(5) **Offensive score**
Any penalties offense	−0.987***	−10.611***	10.723***	0.168***	−0.117***
(1 = yes, 0 = no)	(0.054)	(0.467)	(0.469)	(0.009)	(0.008)
Any penalties defense	0.797***	11.154***	−11.122***	−0.146***	0.140***
(1 = yes, 0 = no)	(0.049)	(0.386)	(0.386)	(0.008)	(0.008)
Drive-level controls[a]	*yes*	*yes*	*yes*	*yes*	*yes*
Possession team-season FEs	*yes*	*yes*	*yes*	*yes*	*yes*
Defensive team-season FEs	*yes*	*yes*	*yes*	*yes*	*yes*
N	111,669	111,629	111,629	111,629	111,629
Mean dep. var.	5.857	26.843	43.716	0.468	0.338
R^2	0.041	0.094	0.210	0.084	0.098

Notes: *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Standard errors–clustered on the game level–in round parentheses.
The dependent variables are the total number of plays per drive (column (1)), the yards gained per drive (column (2)), the yard line at the end of the drive (column (3)), a binary variable indicating that a drive ends with a 'punt' (column (4)), and a binary variable equal to 1 if the team in possession was able to score and 0 if it was not or if the defense could score (column (4)). [a] Control variables include the expected win probability for the team in possession and the score of both teams at the beginning of the drive. Further controls are the field position at the beginning of the drive and drive number, game-time, sub-quarter, and referee fixed effects.

Game level Finally, Panel *A* of Table B.3 presents the results for regressing the absolute and relative number of 'mental error penalties' on game outcomes. It shows that a team's winning probability correlates positively with the opponent's total number of penalties (significant at the 10-percent level) and negatively (without statistical significance) with its own total number of

[11] Teams score through touchdowns (six points), i.e., when the ball is carried into the opposing team's end zone or caught in the end zone by an offensive player. After a touchdown, the scoring team can score additional points (one or two). Alternatively, teams can score through a field goal, which means kicking the ball through the goalposts during a play from anywhere on the field.

penalties (column 1). When we use the difference in penalties (own - opponent) as an explanatory variable, the estimated coefficient implies a reduction by 0.75 percentage points with each 'mental error penalty' the focused team commits more than its opponent (column 3). Results are similar when we use a continuous variable as the dependent variable (final score difference) and when we control for ex-ante winning probabilities based on betting odds (Panel *B*). Since betting odds are unavailable for the whole sample, the estimates in Panel *B* are less precise.

TABLE B.3 — The association of penalties and success: game-level (all regular season games)

	(1) **Game win**	(2) **Win margin**	(3) **Game win**	(4) **Win margin**
A. Full sample of games				
Own penalties	−0.0041	−0.1461		
	(0.0060)	(0.1672)		
Opponent penalties	0.0113*	0.4258**		
	(0.0064)	(0.1897)		
Penalty difference			−0.0075*	−0.2791**
(own - opponent)			(0.0044)	(0.1282)
Team-season FEs	*yes*	*yes*	*yes*	*yes*
Opponent-season FEs	*yes*	*yes*	*yes*	*yes*
Own and Opponent's wins FEs	*yes*	*yes*	*yes*	*yes*
R^2	0.44	0.47	0.44	0.47
Mean dep- var.	2.39	0.57	−2.2207	2.39
B. Games with betting odds available				
Own penalties	−0.0073	−0.4298**		
	(0.0073)	(0.1976)		
Opponent penalties	0.0115	0.5059**		
	(0.0081)	(0.2350)		
Penalty difference			−0.0093*	−0.4652***
(own - opponent)			(0.0053)	(0.1529)
Betting odds	0.0037***	−0.0022	0.0038***	−0.0021
(implied win probability)	(0.0010)	(0.0294)	(0.0010)	(0.0293)
Team-season FEs	*yes*	*yes*	*yes*	*yes*
Opponent-season FEs	*yes*	*yes*	*yes*	*yes*
Own and Opponent's wins FEs	*yes*	*yes*	*yes*	*yes*
R^2	0.46	0.49	0.46	0.49
Mean dep. var.	2.22	0.57	2.22	0.57

Notes: For panel *A*, N = 5,066. For a reduced number of games used for panel *B*, betting odds information is available, N = 3,408. The dependent variable for columns (1) and (3) is equal to 1 if the observed team wins the game, and 0 otherwise. Eight games that ended in a tie were excluded. The dependent variable for columns (2) and (4) is the final score difference. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors in round parentheses.

C. Weather and game attendance

In this section, we examine whether attendance varies with temperature. Fans can influence the game in general and on penalties associated with mental errors in particular; hence, our motivation is that a smaller number of spectators in 'heat games' (i.e., games with a temperature greater than 85°F / 29.4°C) could present a threat to identification. Our empirical approach is to estimate a game-level model where we regress various temperature measures on absolute and relative attendance (attendance rate).

The results are presented in Table C.1 (absolute attendance) and Table C.1 (relative attendance). We find a significant effect of 'heat games' on absolute and relative attendance. For temperature as a continuous measure, it shows that attendance even increases by 22 individuals if temperature increases by one degree Fahrenheit (column (2) in Table C.1). However, estimates for five-degree interval dummies suggest that this result can be partly explained with a lower attendance in cold weather games: Our estimates suggest that absolute attendance is reduced by approximately 1,000 spectators (or 1.72 percentage points of capacity) if the temperature is below 47°F (8.3°C) (columns (3)). In conclusion, we find no indication that high temperatures affect attendance.

TABLE C.1 — The association of absolute game attendance and temperature (absolute attendance)

	(1)	(2)	(3)
$1[\text{Temp.}_g > 85]$	−0.1120		
	(0.3429)		
Temperature (°F)		0.0224***	
		(0.0057)	
Temp. ≤ 46			−1.0644***
			(0.2598)
Temp.]46, 50]			−0.1936
			(0.2771)
Temp.]50, 55]			−0.2175
			(0.2586)
Temp.]55, 60]			−0.3588
			(0.2578)
Temp.]60, 65]			−0.0557
			(0.2382)
Temp.]65, 70]			−0.1166
			(0.2323)
Temp.]70, 75]			*omitted base category*
Temp.]75, 80]			−0.2867
			(0.2502)
Temp.]80, 85]			−0.0896
			(0.2863)
Temp.]85, 90]			−0.0597
			(0.3642)
Temp. > 90			−0.3213
			(0.9644)
Additional controls	*yes*	*yes*	*yes*
Team-season FEs	*yes*	*yes*	*yes*
Opponent-season FEs	*yes*	*yes*	*yes*
N	5,056	5,056	5,056
Mean of dep. var.	67.48	67.48	67.48
R^2	0.93	0.93	0.93

Notes: The dependent variable is the absolute attendance in 1,000 spectators. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Standard errors clustered on the team-season level in round parentheses.

Table C.2 — The association of absolute game attendance and temperature (relative attendance)

	(1)	(2)	(3)
$1[\text{Temp.}_g > 85]$	−0.0011		
	(0.0047)		
Temperature (°F)		0.0004***	
		(0.0001)	
Temp. ≤ 46			−0.0172***
			(0.0033)
Temp.]46, 50]			−0.0050
			(0.0036)
Temp.]50, 55]			−0.0051
			(0.0035)
Temp.]55, 60]			−0.0067*
			(0.0036)
Temp.]60, 65]			−0.0019
			(0.0032)
Temp.]65, 70]			−0.0020
			(0.0033)
Temp.]70, 75]			*omitted base category*
Temp.]75, 80]			0.0006
			(0.0032)
Temp.]80, 85]			−0.0009
			(0.0039)
Temp.]85, 90]			0.0007
			(0.0050)
Temp. > 90			−0.0006
			(0.0134)
Additional controls	*yes*	*yes*	*yes*
Team-season FEs	*yes*	*yes*	*yes*
Opponent-season FEs	*yes*	*yes*	*yes*
N	5,056	5,056	5,056
Mean of dep. var.	0.94	0.94	0.94
R^2	0.85	0.85	0.85

Notes: The dependent variable is the relative attendance (observed attendance over maximum capacity). *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Standard errors clustered on the team-season level in round parentheses.