Sporting Events and Road Accidents: A Case Study from NFL Games

Tahmeed Hossain

April 2024

Abstract

This paper examines the causal impact of major sporting events, which are known to elevate alcohol consumption and traffic congestion, on the incidence of traffic accidents. Utilizing the quasi-random nature of National Football League (NFL) game scheduling and an extensive accident dataset spanning from 2017 to 2022, I exploit within-region variation over time to estimate the effect of gamedays on accident rates. My main results indicate a 15.60% increase in traffic accidents within the 0-5 mile radius of the stadiums on game days. This effect is largely driven by an increase in accidents in the hours immediately preceding and following the games. Additionally, I find no significant changes in accident rates at the county level or in wider distance ranges from the stadiums.

1 Introduction

Road safety is one of the world's greatest public health challenges, with over 3,500 daily fatalities and 50 million annual injuries on roads worldwide (WHO, 2021). Vehicular crashes rank as the eighth leading cause of death globally (Tumwesigye et al. 2016). This issue is particularly acute in the United States, which has higher road accident death rates compared to most other OECD nations[1] (Badger & Parlapiano, 2022). While many complex factors contribute to road accidents, two of the main causes are heavy drinking and traffic congestion. Approximately 37 people die every day from road accidents related to drunk driving in the US, which translates to one death every 39 minutes (Stewart, 2023). With congestion or more cars on roads, the probability of accidents occurring naturally increases due to more opportunities for collisions. Both of these factors are often intensified during major sporting events such as National Football League (NFL), Major League Baseball (MLB), National Basketball Association (NBA). For example, social drinking traditions including tailgating, binge drinking and seeking out local bars' game day beer specials are highly prevalent during professional (Toomey et al. 2008) and college (Neal and Fromme 2007) football game days. Adding to this, sporting events draw large crowds that significantly increase traffic congestion around stadiums. For example, NFL games averaged 69,389 attendees in 2022 season, heavily impacting local traffic flows (Gough, 2024). The combination of both social drinking and congested roads creates a high-risk environment for impaired driving and vehicular accidents.

In this paper, I aim to identify how major sporting events, which often lead to higher alcohol consumption and increased traffic congestion, impact the frequency of traffic accidents. The National Football League (NFL) serves as the primary case study for this study due to being the most followed sport in the US, with an average attendance of 69,389, compared to 26,808 average

[1] See Appendix Figure 1 which compares deaths per million people across the U.S., France, Japan, and other OECD nations.

in Major League Baseball (MLB) and 17,947 in National Basketball Association (NBA). Utilizing detailed accident dataset provided by Moosavi et al. (2022), I exploit the random nature of the scheduling of game days to calculate the change in accident reports attributable to football games. Intuitively, I contrast the accident reports from counties on days when NFL games are played against those on days without games, adjusting for potential day-of-the-week, monthly, and annual variations. In addition to exploiting within county variation over time, this study also takes into account variations within different distance ranges (0-5 miles, 5-10 miles, 10-20 miles, 20-50 miles, 50-100 miles) from the stadiums in multiple separate regressions. This approach is motivated by the possibility that significant population movements towards game locations, along with many individuals traveling to pubs or communal spaces to watch the games in counties not hosting the game, might lead to heightened congestion and drunk driving incidents across a wider area on game days.

My results indicate a 15.60% increase in traffic accidents within the 0-5 mile radius of the stadium, which roughly translates to 420 additional accidents during the 2021 NFL season (September to February) across the same radius around 31 stadiums. These additional accidents account for 0.74% of the total accidents during the 6 months of the 2021 NFL season. This effect, however, does not extend to the county level or to further distance ranges (5-10 miles, 10-20 miles, 20-50 miles, and 50-100 miles) from the stadiums. Analysis of hourly data supports these findings, suggesting a 60.31% increase in accident rates within the 0-5 mile radius in the three hours before the game starts and a 123.22% increase in the three hours after the game ends. On the other hand, during the 3-hour window when the game is occurring, accidents fall by 39.77%. Additionally, I observe no significant changes in accident rates during these time periods, with imprecise estimates both at the county level and in the wider distance ranges. The results indicate that the areas in close proximity to the stadium, especially during the hours immediately preceding and following NFL games, have a higher risk of traffic accidents. Additionally, I find that the game day effect on accident rates is robust and remains similar in the immediate vicinity (0-5 mile radius) of the stadium even after accounting for the size of the crowd. However, the overall level of attendance itself does not appear to have a significant impact on accident occurrences.

My estimates contribute to the literature by providing the first causal analysis of the impact of a regular sports season on traffic accidents. While Redelmeier and Stewart (2003) and Wood, McInnes and Norton (2011) have examined the association between game days and accidents, their studies are limited in scope and do not establish causality. Redelmeier and Stewart find a 41% increase in fatalities on 27 Super Bowl Sundays, while Wood, McInnes and Norton (2011) observe that closer games in football and basketball are correlated with more fatalities. In contrast, my research utilizes a more comprehensive dataset encompassing six NFL seasons and employs an identification strategy that leverages the quasi-random timing of game days to estimate the causal relationship. The paper also contributes to the growing body of literature examining the impact of large-scale sporting events on various negative externalities, especially crime. Rees and Schnepel (2009) show that violent assaults are associated with Division I-A college football game days. Lindo et al. (2018b) find a 28% increase in daily reports of rape victimization among women aged 17–24 on college football game days. Andres, Fabel and Rainer (2023) explore the incidence of violent crime during professional soccer games in Germany and find a 17 percent increase in violent crimes on gamedays. While not directly related to crime, my study contributes to this literature by demonstrating the significant impact of major sporting events on another important negative externality: traffic accidents.

2 Data

In this study, I utilize "US-Accidents", a comprehensive motor vehicle accident dataset publicly available from Moosavi et al. (2019)[2]. They constructed this dataset using data from "MapQuest Traffic" and "Microsoft Bing Map Traffic," which provide real-time updates on various traffic events, including accidents and congestion. These events were captured from sources like U.S. and state transportation departments, law enforcement, traffic cameras, and road-network sensors. Data collection was conducted at intervals of 90 seconds from 6 am to 11 pm and 150 seconds from 11 pm to 6 am (Moosavi et al. 2019). Over the period from February 2016 to March 2023, Moosavi et al. compiled a total of 7.72 million traffic accident cases. "US-Accidents" is the first dataset to cover motor vehicle accidents across the entire country, in contrast to other available datasets that typically only include accident reports from one state. This dataset also provides an extensive range of data attributes for each accident record, which includes location details, timestamps, natural language descriptions, weather data, period-of-day information, and relevant points-of-interest data such as traffic signals and stop signs. However, the attributes do not indicate whether there are any fatalities or if drugs or alcohol are involved.

To assess the dataset's reliability and determine whether it reflects real-world accident patterns, I present a series of figures analyzing accident trends over time. These visualizations highlight variations in accident occurrences across different times of the day, days of the week, and seasons, as well as accident rates adjusted per million Vehicle Miles Traveled (VMT).

In Figure 1, I present a temporal analysis of accident occurrences from the "US-Accidents" dataset. In Panel A, we can see the hourly fluctuation in accident numbers, highlighting peaks during morning and afternoon hours, coinciding with typical office commute times.

In panel B, the hourly trends between weekdays and weekends are compared. During weekdays, peaks are evident at 7-8 AM and 4-5 PM, again aligning with office commute hours. On weekends, early morning hours see fewer accidents, possibly due to people sleeping in or staying at home, while the frequency gradually increases, perhaps because of activities like shopping and social events. Interestingly, during the late-night hours of 12-4 AM and 7-11 PM, weekends report higher accident rates compared to weekdays, which may reflect increased nighttime socializing and leisure activities that extend into the later hours.

In Panel C, I show the distribution of accidents by day of the week. This figure reveals the highest frequency of accidents on Fridays, followed by other weekdays, in contrast to significantly lower rates on weekends. Panel D depicts the seasonal patterns in accident occurrences, with a significant increase observed from November to December, suggesting a rise in accidents during the winter months and holiday season.

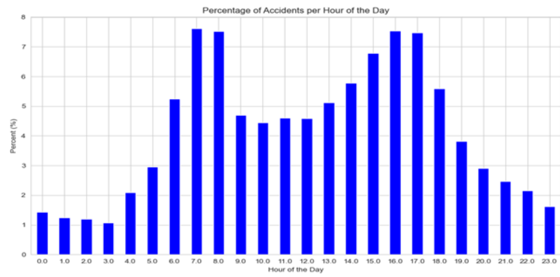
In Figure 2a, I present a similar analysis for accident rates adjusted per million Vehicle Miles Traveled (VMT) using NHTSA 2022 survey. Panel A shows that the highest rates of accidents per mile travelled occur in the very early hours, around midnight to 3 AM, perhaps due to factors like reduced visibility, driver fatigue, or a higher rate of impaired driving. This finding contrasts with the observations from Panel A of Figure 1, which shows the highest raw accident counts during typical rush hours.

Panel B of Figure 2 highlights a pronounced increase in the accident rate per million VMT in December, followed by April, suggesting specific seasonal risks associated with these months for the year 2022. The month-to-month variation in this graph is more pronounced compared to the raw seasonal trend of 2016-2022 observed in Panel B of Figure 1, where the spread of accidents across months is more uniform, likely due to the averaging effect of pooling1 multiple years' data.

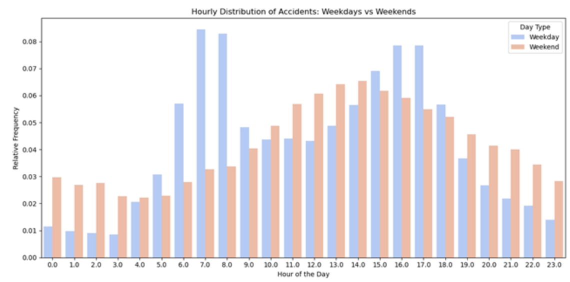
Panel C of Figure 1 suggests that Fridays are the most accident-prone days, possibly due to higher traffic volumes as people may travel for weekend plans. However, the normalized graph in panel C of Figure 2 shows that Thursdays have a higher accident rate per million VMT, indicating

[2]https://arxiv.org/abs/1906.05409

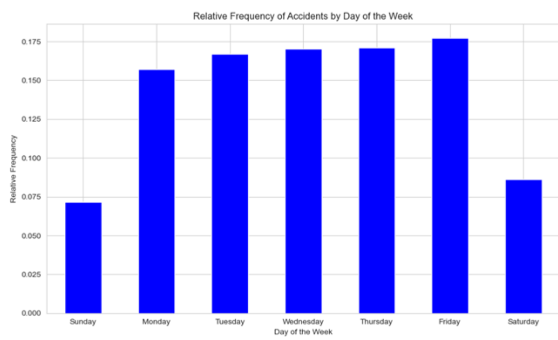
Figure 1: Distribution of Accidents across Time



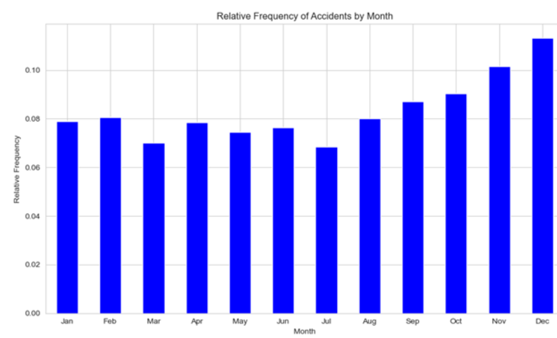
(a) Hourly Fluctuations of Accidents



(b) Hourly Fluctuations of Accidents: Weekdays vs Weekends



(c) Accidents across Days of the Week



(d) Monthly distribution of accidents

that per mile traveled, Thursdays are the riskiest, followed by Wednesday and Tuesday.

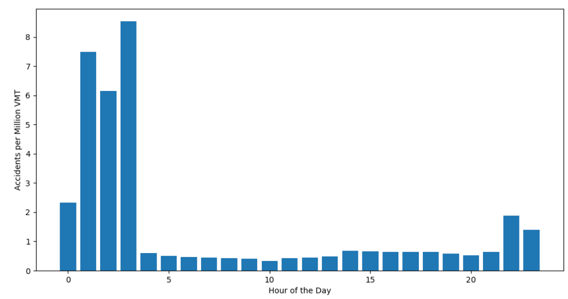
The above figures show that the observed accident patterns closely match real-world expectations. Accidents peak during commuting hours, with the highest frequencies occurring on Fridays. Seasonal increases in November-December and April likely reflect winter driving hazards, holiday travel, and increased road activity in spring.

I combine the accident data with the data on football games from the sports-reference website[3]. These data include the game dates for each team, the location of the game (home or away), and the game's result. This combined dataset allows for the analysis of accidents occurring on or around NFL football game days, with a specific focus on counties that host NFL team stadiums. I achieve this by producing a dataset on a county-by-day basis, which includes accident reports and indicators signifying whether the day is a game day for the county's affiliated team. There are 33 counties in total, distributed across 23 different states, as indicated in appendix table 1. The table also provides information about the teams considered in the analysis, their home stadiums and geographic information.

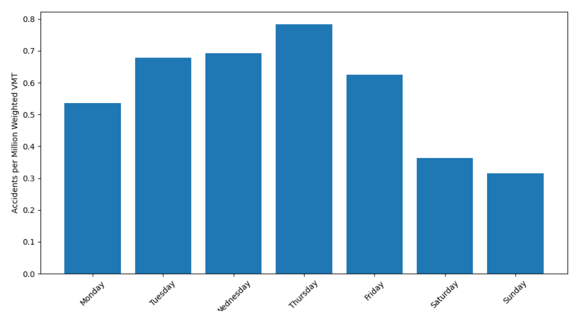
The primary variable of interest in this analysis is the daily incidence of accidents within each county. Additionally, in an alternative specification, the impact of NFL game days on neighboring areas at varying distances from the stadium is examined. These distances are categorized into bins of 0-5 miles, 5-10 miles, 10-50 miles, and 50-100 miles from each stadium. Consequently, in this specification, the dataset is structured at the distance bin-by-day-by-stadium level, meaning

[3]https://www.sports-reference.com/

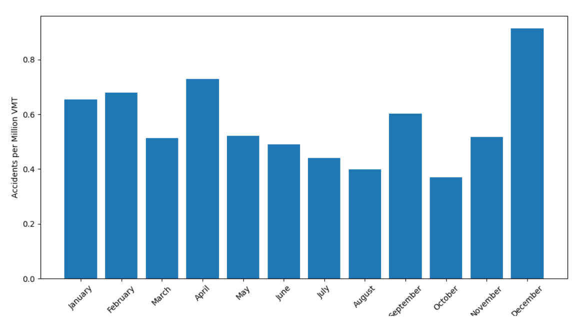
Figure 2: Distribution of Accidents across Time per Million Vehicle Miles Travelled (VMT)



(a) Hourly Fluctuations of Accidents Per Million VMT



(b) Accidents across Days of the Week per Million VMT



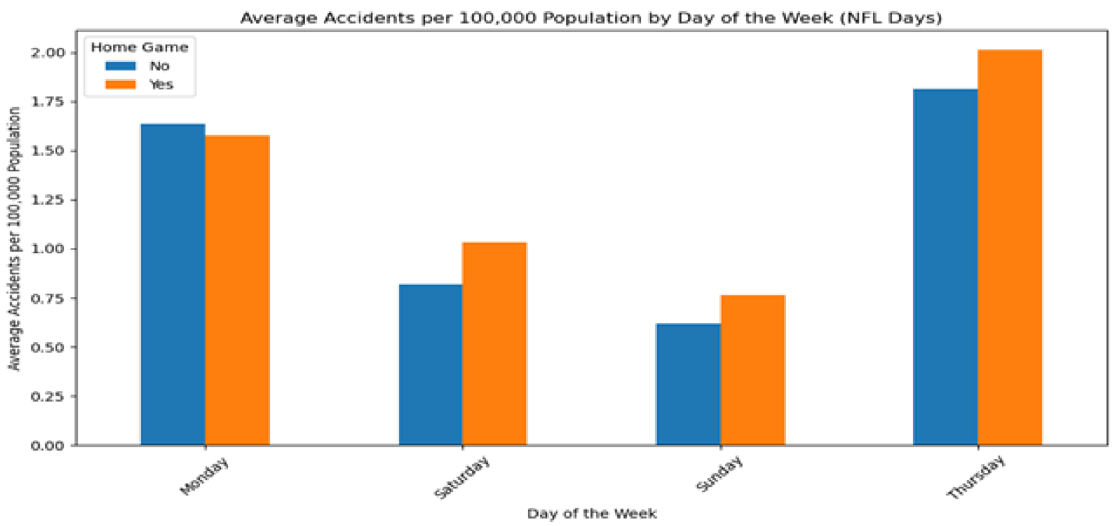
(c) Monthly Distribution of Accidents per Million VMT

that for each stadium and each day, accident data is separately recorded for different geographic distance bins. For instance, in the case of the 0-5 mile distance bin, the dataset includes accident counts within a 5-mile radius around each of the 31 stadiums, with a separate dataset created for each distance bin.

My primary dataset for the study encompasses 37,479 daily observations across 31 panels (e.g., counties, 0-5 mile distance bins). Within this dataset, there are 5,363 panel-day observations corresponding to Sundays without NFL games and 1,423 panel-day observations for Sundays with NFL games. Additionally, the dataset includes 303 panel-day observations for NFL games held on weekdays, which include 115 Monday games, 115 Thursday games, and 69 Saturday games.

Using the NFL gameday data and the US-Accident data, I provide a preliminary overview of accidents across different days of the week in figure 3. The figure shows the daily accident rates during weeks with a local home game versus weeks without one in counties that host NFL games. From the figure it is apparent that apart from Monday game days, accident rates are marginally higher on other game days of the week per 100,000 population.

Figure 3: Average Accident per 100,000 Population by Day of the Week (NFL Days)



In the main regression analysis, I also control for weather conditions. I incorporate average daily temperature and daily precipitation data obtained from the National Centers for Environmental Information (NCEI) as the weather controls. To construct the weather control dataset, I identify the weather stations that measure the relevant weather variables on a daily basis throughout the sample period. From this subset of stations, I select the weather station that is geographically closest to each stadium in the study.

Finally, I incorporate a set of binary indicators into my regression models to account for the potential impact of public holidays and other notable days on accident rates. These indicators include Labor Day, Thanksgiving, Christmas, New Year's Eve, 4th of July, Christmas Eve, Columbus Day, Easter (both Eastern and Western), Juneteenth, Labor Day Weekend, Martin Luther King, Jr. Day, Memorial Day, New Year's Day, Thanksgiving Eve, Valentine's Day, Veterans Day, and Washington's Birthday. I construct the data for these holiday indicators using Google.

3 Methodology

I estimate the effects of football games played by NFL teams using two different approaches. In the first approach, I exploit the within county variation over time. My model relies on the assumption that the fluctuations in accident reports across different days of the week in weeks without football games provide a reasonable counterfactual for estimating the changes expected on game days, after accounting for inherent differences due to years, months, days of the week, and similar factors. To explain more precisely, consider a hypothetical situation: a football game is scheduled in Tarrant County on a Sunday in January 2024. The counterfactual scenario — the predicted number of accidents had the game not taken place — is estimated by examining the accident rates in the same county on Sundays in January 2024 when there are no games or when the team plays away.

Since the accidents data is in the form of counts and a large portion of these counts are zero, I use a Poisson model for estimating the effect of NFL game days on the number of daily accident reports, conditional on whether or not the day is a game day, county fixed effects, day-of-week, month, and year fixed effects, and also conditional on public holidays and weather conditions. My baseline specification is based on the following equation:

$$E[\text{Accident}_{ct}|\text{Gameday}_{ct}, V_c, \text{time}_{dmy}, X_{ct}] = \exp(\beta\text{Gameday}_{ct} + V_c + \text{time}_{dmy} + \gamma X_{ct}) \quad (1)$$

Here, Accident_{ct} is the number of accidents reported, at county c, occurring on day t; Gameday_{ct} is a binary variable that is equal to one if county c has a game on day t (i.e., when there is a home game); County fixed effects, denoted as V_c, are used to account for time-invariant differences among different counties. Incorporating these fixed effects ensures that the derived effects are attributed to changes occurring over time within a specific county, rather than to differences that exist across multiple counties (Dr. Pattison suggested addressing incidental parameter problem when including fixed effects in non-linear model. I don't think its an issue since many time periods per county and did not see similar papers addressing it). The time variable time_{dmy} includes day-of-the-week, month, and year fixed effects. Day-of-week fixed effects account for the fact that a large majority of the games we analyze are held on Sundays, a day which significantly differs from weekdays and Saturday in terms of traffic patterns and congestion. The year fixed effects account for long run time trends in the number of accidents reported, while the month fixed effects account for any seasonal effects. The vector X_{ct} includes indicators for different public holidays such as Labor Day, Thanksgiving, Christmas, New Year's Eve. Controlling for holidays is important since traffic patterns go through systematic changes during these holidays. X_{ct} also includes weather controls for average daily temperature and daily precipitation, calculated as 24-hour averages for each location on both game days and non-game days. Finally, I calculate sandwiched standard error estimates clustered at the county level.

One possible issue with my empirical approach is that it only considers home games as treatment. This focus on home games might bias the counterfactuals, as days when away games occur are treated as part of the control group. Intuitively, there is a possibility that fans might congregate in pubs or other places, causing more accidents in their home area. To address this concern, I extend the analysis to include away games as part of treatment and show the effects of both home games and away games in the same regression. Hence, in the extended analysis, the control group is now only the non-gamedays.

In assessing the causal effect of game days on the incidence of traffic accidents, one challenge is the potential for spillover effects into adjacent areas. On game days, many individuals travel to various venues such as pubs or communal spaces to watch the games, which could lead to congestion and increase in drunk driving. Consequently, an increase in traffic accidents can be seen not only in the county where the game is held but also in neighboring counties. If the observed rise in accidents in the vicinity of the stadium is primarily due to population inflow, and if this inflow results in increased traffic in the areas from which these crowds originate, our estimates may understate the overall impact. Such a scenario is plausible, where sharp population movements towards the game location might increase the incidence of accidents in the originating counties.

To address these possibilities, my second approach to estimating the effects of accidents is running multiple separate regressions for different distance ranges from the stadium following Marcus (2019) and Nguyen (2019).

$$E[\text{Accident}_{rt}|\text{Gameday}_{rt}, V_r, \text{time}_{dmy}, X_{rt}] = \exp(\beta\text{Gameday}_{rt} + V_r + \text{time}_{dmy} + \gamma X_{rt}) \quad (2)$$

In this equation, Accident_{rt} denotes the count of accidents reported within distinct radius ranges (such as 0-5 miles, 5-10 miles, 10-20 miles, 20-50 miles, and 50-100 miles) from the stadium on day t. Gameday_{rt} is a binary variable that equals 1 for all locations within a specified radius range r from the NFL stadium where a game is being played on day t, and 0 otherwise.; V_r denotes fixed effects specific to the radius range r, capturing characteristics unique to that radius range from the stadium. Each regression is conducted separately for the specified radius ranges (e.g., 0-5 miles, 5-10 miles, etc.), thus analyzing the impact of game days on accident rates within those distinct distances from the stadium.

4 Results

4.1 Main Results

Despite initial indications of a positive impact of game days on accidents from Figure 1 in the data section, the baseline results of my county level analysis suggest no such effects. Table 1 presents the results corresponding to Equation (1). Column (1), representing the baseline model, incorporates fixed effects for county, day of the week, month, and year. In column 2, I further extend the model to controls for weather conditions and holidays. The estimates from both the specifications are insignificant, suggesting that a home game does not have an impact on accident rates at the county level.

Table 1: Estimated Effects of Home Game Days on Accidents (County Level)

	(1)	(2)
Home gameday	-0.008	0.005
	(0.031)	(0.033)
Observations	37,479	37,479
Number of counties	31	31
County FE	YES	YES
Date FE	YES	YES
Holiday controls	NO	YES
Weather controls	NO	YES

*Statistically significant at 10% level; ** at 5% level; *** at 1% level.
Notes: The outcome variable is daily accident counts per county. Date FE includes year, month and day of the week fixed effects. Population is used as an exposure variable. Standard errors in the parentheses are corrected for clustering at the county level.

While results from Table 1 indicate that game days seemingly have no effect on accident rates at the county level, it is possible that accidents may increase within closer proximity to the stadium, or that an uptick in one area could be neutralized by a downturn elsewhere. Moreover, the specific placement of a stadium—especially if it's situated near the border of another county—could lead to increased accidents in adjacent counties as well. Therefore, I further explore the effects of football games within various distance ranges from the stadium (such as 0-5 miles, 5-10 miles, 10-20 miles, 20-50 miles, 50-100 miles, and 100-200 miles) through separate regressions. This approach allows for comparison of how accident rates change in close proximity to the stadium versus farther away which is not possible in county-level analysis.

Table 2 presents the results of this distance bins approach, corresponding to Equation (2). Columns 1 through 5 show estimates for the effects at different distances from the stadium. All models incorporate fixed effects for county, day of the week, month, and year, in addition to controls for weather conditions and holidays. The estimate from column 1 suggests that a home game increases the number of accidents by 15.60% within the 0-5 miles radius of the stadium. However, results from other columns suggest that these effects do not persist farther away from the stadium, as I find insignificant results for other distance bins from the stadiums.

Table 2: Estimated Effects of Home Game Days on Accidents (Different Distance Ranges from Stadiums)

	0-5 Miles	**5-10 Miles**	**10-20 Miles**	**20-50 Miles**	**50-100 Miles**
Home gameday	0.145***	0.024	0.006	0.006	-0.026
	(0.039)	(0.064)	(0.063)	(0.031)	(0.023)
Observations	37,479	37,479	37,479	37,479	37,479
Number of geographic units	31	31	31	31	31
County FE	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES
Holiday control	YES	YES	YES	YES	YES
Weather controls	YES	YES	YES	YES	YES

*Statistically significant at 10% level; ** at 5% level; *** at 1% level.

Notes: The outcome variable is daily accident counts per geographic unit (county or 0-5, 5-10, 10-20, 20-50 miles distance ranges from stadiums). Date FE includes year, month and day of the week fixed effects. Population is used as an exposure variable. Standard errors in the parentheses are corrected for clustering at the geographic unit level.

Next, based on hourly rates of accidents (instead of daily), I examine if there's an increase in accidents in the hours leading up to a game, during the game itself, and in the hours following the game's conclusion. To explore this, I construct two 3-hour windows before the game starts (lead 2 and lead 1), the 3-hour window covering the game's duration (lag 0), and two 3-hour windows after the game ends (lag 1 and lag 2). In this specification, the control group consists of hours outside these designated 15 hours. I also extend this model by incorporating fixed effects that interact day of the week with specific hours of the day instead of just accounting for day of the week fixed effects. Hence, I essentially compare the specified time windows on game days against equivalent time windows on non-game days. The results are presented in figure 4 and table 3. We can see from figure 2b and column 2 of table 3 that within the 5-10 mile radius from the stadium, coefficients for all the 3-hour windows are insignificant with large standard errors. I find similar insignificant coefficients with large standard errors for the distance ranges of 5-10 miles, 10-20 miles, 20-50 miles, and 50-100 miles (Table 2). However, as we can see in figure 2a and column 1 of Table 4, for the 0-5 mile radius, home games have large, positive, and significant effects (at the 5% level) on accidents in the immediate 3-hour intervals before and after home games commence and conclude. Specifically, accidents rise by approximately 60.31% and 123.22% during the 3-hour window before the game starts and the 3-hour window after the game ends, respectively. On the other hand, during the 3-hour game window, accidents fall by 39.77%, compared to non-game days. Finally, in the wider time windows encompassing 3-6 hours before the start of the game and 3-6 hours after the game ends, the figures show no significant effect on accident rates. A joint test of significance for the coefficients lead 1, lag 0 and lag 1 yield a p-value of 0.0028, suggesting that there are significant changes in accident rates related to game days during at least one of these 3 hour periods. Finally, aggregate impact during the 9-hour window of the game days is 0.768, which translates to a 115% increase in accidents.

These findings are consistent with the daily analysis, where positive effects on accidents were observed only within the 0-5 mile radius from the stadium. The results highlight that the areas closest to the stadium, especially during the 3-hour periods preceding and following NFL games, are most prone to accidents. This increase in risk is likely driven by increased congestion and alcohol consumption by fans.

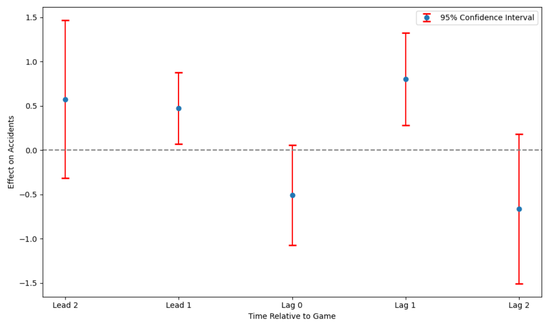
Table 3: Estimated Impact of Home Games on Hourly Accident Rates (Different Distance Ranges from Stadiums)

	0-5 miles	**5-10 miles**	**10-20 miles**	**20-50 miles**	**50-100 miles**
Lead 2	0.577	0.008	-0.163	-0.141	0.178
	(0.455)	(0.327)	(0.207)	(0.201)	(0.130)
Lead 1	0.472**	-0.159	0.281	0.096	-0.040
	(0.206)	(0.228)	(0.240)	(0.210)	(0.129)
Lag 0	-0.507*	-0.210	0.171	-0.341	-0.152
	(0.288)	(0.269)	(0.194)	(0.260)	(0.157)
Lag 1	0.803***	0.174	0.201	-0.035	-0.106
	(0.266)	(0.255)	(0.267)	(0.211)	(0.124)
Lag 2	-0.663	-0.0884	-0.120	-0.308	-0.285*
	(0.430)	(0.300)	(0.213)	(0.268)	(0.151)
Aggregate Impact (Lead 1 + Lag 0 + Lag 1)	0.768**	-0.195	0.653	-0.281	-0.298
	(0.313)	(0.480)	(0.576)	(0.472)	(0.284)
Observations	899,496	899,496	899,496	899,496	899,496
Number of geographic units	31	31	31	31	31
County FE	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES
Weather Control	YES	YES	YES	YES	YES
Hour-by-Dow FE	YES	YES	YES	YES	YES

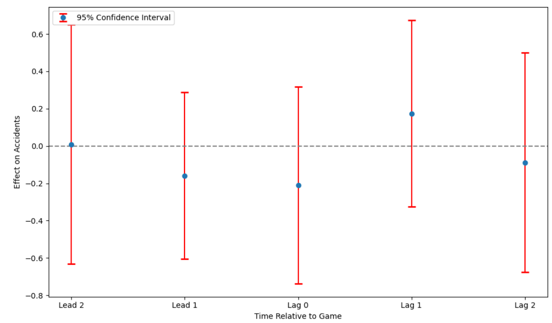
*Statistically significant at 10% level; ** at 5% level; *** at 1% level.

Notes: The outcome variable is accident counts per geographic unit (county or 0-5, 5-10, 10-20, 20-50, 50-100 miles distance ranges from stadiums) per 3 hours. Date FE includes year, month and day of the week fixed effects. Population is used as an exposure variable. Standard errors in the parentheses are corrected for clustering at the geographic unit level.

Figure 4: Impact of NFL Games on Traffic Accidents



(a) Impact of NFL Games on Traffic Accidents (0-5 Mile Radius)



(b) Impact of NFL Games on Traffic Accidents (5-10 Mile Radius)

4.2 Heterogeneous Effects: Home vs. Away

In section 4.1, all the models have been used to show the effect of home games, where away game days were part of the control group, in addition to non-game days. This might cause a downward bias in results if fans congregate in pubs or other places during away games too, causing more accidents in their home area. To address this concern, I show the effect of both home game days and away game days in Table 4. Here, the control group now includes only the non-game days. In column 1, I present the estimates of the impact of both home games and away games on accidents for the county-level analysis, while subsequent columns (2 to 6) examine these effects within varying distances from the stadium (0-5 miles, 5-10 miles, 10-20 miles, 20-50 miles, 50-100 miles).

Table 4: Estimated Effects of Home and Away Game Days on Accidents

	0-5 miles	**5-10 miles**	**10-20 miles**	**20-50 miles**	**50-100 miles**
Home gameday	0.138***	0.008	-0.010	0.001	-0.029
	(0.045)	(0.060)	(0.058)	(0.029)	(0.027)
Away gameday	-0.032	-0.078	-0.077*	-0.021	-0.020
	(0.049)	(0.047)	(0.039)	(0.042)	(0.052)
Observations	37,479	37,479	37,479	37,479	37,479
Number of geographic units	31	31	31	31	31
County FE	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES
Holiday controls	YES	YES	YES	YES	YES
Weather controls	YES	YES	YES	YES	YES

*Statistically significant at 10% level; ** at 5% level; *** at 1% level.

Notes: The outcome variable is daily accident counts per geographic unit (0-5, 5-10, 10-20, 20-50, 50-100 miles distance ranges from stadiums). Date FE includes year, month and day of the week fixed effects. Population is used as an exposure variable. Standard errors in the parentheses are corrected for clustering at the geographic unit level.

The results indicate that for all specifications, the effect of away game days on accidents is small and insignificant. Hence, away game days do not bias the results seen in Table 2. On the other hand, the effect of home games remains insignificant for all distance ranges apart from the

0-5 mile radius around the stadium. For the 0-5 mile distance range, home games cause a 14.79% increase in accidents, which is slightly lower than the effect seen in Table 2 (15.60%).

4.3 Heterogeneous Effects: Wins vs. Losses

Next, I investigate whether game outcomes (win versus loss) affect accident rate differentially. To do this, I introduce an interaction term between the home game day variable and a loss indicator in my models represented by equations 1 and 2. The coefficient of the home game day variable captures the effect of wins, while the interaction term represents the differential effect of losses relative to wins. The results in Table 5 show that, at the county level and at different distance levels (except for 0-5 miles range), wins and losses have no significant effect on accidents. However, within the 0-5 mile radius of the stadium, home game wins result in a substantial 21.89% increase in accidents, significant at the 1% level. On the other hand, effect of a home loss relative to non-gamedays is (0.198 - 0.149=0.049), which translates to 5.02% increase in accidents. However, this is not significant even at the 10% level.

Table 5: Estimated Heterogeneous Effects of Wins and Losses on Accidents

	County	**0-5 Miles**	**5-10 Miles**	**10-20 Miles**	**20-50 Miles**	**50-100 Miles**
Home gameday	0.037	0.198***	0.035	-0.018	-0.041	-0.036
	(0.031)	(0.049)	(0.062)	(0.068)	(0.046)	(0.046)
Home gameday X loss	-0.097*	-0.149**	-0.042	0.026	0.082*	0.015
	(0.051)	(0.058)	(0.057)	(0.038)	(0.048)	(0.068)
	(0.047)	(0.046)	(0.082)	(0.070)	(0.031)	(0.038)
Observations	37,479	37,479	37,479	37,479	37,479	37,479
Number of geographic units	31	31	31	31	31	31
County FE	YES	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES	YES
Weather Control	YES	YES	YES	YES	YES	YES

*Statistically significant at 10% level; ** at 5% level; *** at 1% level.

Notes: The outcome variable is daily accident counts per geographic unit (0-5, 5-10, 10-20, 20-50, 50-100 miles distance ranges from stadiums). Date FE includes year, month and day of the week fixed effects. Population is used as an exposure variable. Standard errors in the parentheses are corrected for clustering at the geographic unit level. Coefficients for the variable Loss is calculated as: Home gameday + home gameday X loss = Loss.

Building on both the temporal patterns of accidents around game times (table 3) and the heterogeneous effects of game outcomes observed in the daily analysis (table 5), I now examine how hourly accident patterns differ based on game outcomes. I extend the hourly specification by adding interactions between the time window indicators and a loss indicator. The results are presented in table 6. Within the 0-5 mile radius of stadiums, I find strong heterogeneous effects in the pre-game period but not after games. Specifically, in the 3 hours immediately preceding kickoff (Lead 1), wins are associated with a 146.7% increase in accidents, significant at the 1% level. However, the effect differs between wins and losses, with losses leading to 73.76% (exp(-1.338)-1) fewer accidents compared to wins in the pre-game period, though still resulting in a considerable 73% increase in accidents compared to non-game days. This suggests that pre-game activities like tailgating and social gatherings are more intense when fans are optimistic about their team's prospects. The post-game period shows an increase of 97.98% in accidents during the first three hours after wins (Lag 1), but with no differential effect following losses. During the game itself (Lag 0) and in the second three-hour window after games (Lag 2), there are no significant differences in accident patterns between wins and losses. Consistent with the main analysis, these effects are largely confined to areas within 5 miles of stadiums, as coefficients for wider distance ranges show no significant patterns.

Table 6: Estimated eterogeneous Effects of Wins and Losses on Hourly Accident Rates (Different Distance Ranges from Stadiums)

	0-5 miles	**5-10 miles**	**10-20 miles**	**20-50 miles**	**50-100 miles**
Lead 2	0.909	0.419	-0.037	-0.352	0.051
	(0.587)	(0.434)	(0.248)	(0.334)	(0.222)
Lead 2 X Loss	-0.897	-1.065	-0.351	0.452	0.250
	(0.814)	(0.645)	(0.444)	(0.405)	(0.380)
Lead 1	0.903***	0.110	0.196	-0.103	-0.124
	(0.220)	(0.310)	(0.195)	(0.141)	(0.192)
Lead 1 X Loss	-1.338***	-0.505	0.192	0.356	0.172
	(0.420)	(0.394)	(0.462)	(0.328)	(0.363)
Lag 0	-0.465	0.120	0.153	-0.147	-0.308
	(0.409)	(0.268)	(0.277)	(0.317)	(0.235)
Lag 0 X Loss	-0.060	-0.940	0.073	-0.451	0.309
	(0.767)	(0.497)	(0.317)	(0.403)	(0.308)
Lag 1	0.683**	0.512	0.322	-0.418	-0.095
	(0.301)	(0.414)	(0.316)	(0.257)	(0.158)
Lag 1 X Loss	0.325	-0.828	-0.242	0.688	0.015
	(0.343)	(0.630)	(0.409)	(0.324)	(0.294)
Lag 2	-0.427	0.036	-0.411	-0.203	-0.232
	(0.501)	(0.393)	(0.273)	(0.398)	(0.249)
Lag 2 X Loss	-1.275	-0.255	0.554	-0.219	-0.192
	(0.883)	(0.378)	(0.414)	(0.449)	(0.341)
Observations	899,496	899,496	899,496	899,496	899,496
Number of geographic units	31	31	31	31	31
County FE	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES
Weather Control	YES	YES	YES	YES	YES
Hour-by-Dow FE	YES	YES	YES	YES	YES

*Statistically significant at 10% level; ** at 5% level; *** at 1% level.

Notes: The outcome variable is accident counts per geographic unit (county or 0-5, 5-10, 10-20, 20-50, 50-100 miles distance ranges from stadiums) per 3 hours. Date FE includes year, month and day of the week fixed effects. Population is used as an exposure variable. Standard errors in the parentheses are corrected for clustering at the geographic unit level.

4.4 Heterogeneous Effects: Game Importance & Attendance

Similar to the heterogeneous effects of game outcomes, one might expect that the relationship between game days and accidents could differ based on the prominence of the team or the significance of the game. In this vein, we would expect particularly important games to have more pronounced effects. I explore these ideas using proxies for game prominence, including playoff games and key rivalries such as Packers vs. Bears, Cowboys vs. Commanders, and Steelers vs. Ravens, defined as prominent games[4]. However, I don't find any differential effect of prominent games from regular games, as the interaction between home gameday and prominent gamedays remains small and insignificant in all the specifications in table 7. Surprisingly, the effect of prominent games compared to non-gamedays is also insignificant, even though the magnitude is large. This lack of significance may be due to insufficient statistical power, since the relatively small sample size of prominent games (173) in the dataset could limit the ability to detect meaningful differences. For comparison, the number of non-prominent games is 1554, while the number of non-game days is 35,752.

Table 7: Estimated Heterogeneous Effects of Prominent Games and Non-prominent Games on Accidents

	County	**0-5 Miles**	**5-10 Miles**	**10-20 Miles**	**20-50 Miles**	**50-100 Miles**
Home gameday	0.004	0.144***	0.029	0.005	0.021	-0.033
	(0.033)	(0.039)	(0.068)	(0.068)	(0.033)	(0.025)
Home gameday X prominent	0.017	0.007	-0.035	0.007	-0.097*	0.056
	(0.065)	(0.106)	(0.056)	(0.067)	(0.041)	(0.071)
Observations	37,479	37,479	37,479	37,479	37,479	37,479
Number of geographic units	31	31	31	31	31	31
County FE	YES	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES	YES
Weather Control	YES	YES	YES	YES	YES	YES

*Statistically significant at 10% level; ** at 5% level; *** at 1% level.
Notes: The outcome variable is daily accident counts per geographic unit (0-5, 5-10, 10-20,20-50, 50-100 miles distance ranges from stadiums). Date FE includes year, month and day of the week fixed effects. Population is used as an exposure variable. Standard errors in the parentheses are corrected for clustering at the geographic unit level. Coefficients for the variable Prominent gameday are calculated as: Home gameday + home gameday X prominent = Prominent gameday.

To see how important a factor congestion is in explaining accident rates, I now expand the analysis by including in my regression models the overall level of stadium attendance for each game. To do this, I first calculate the demeaned attendance at games for each stadium by subtracting the average attendance from the actual attendance for each season. To facilitate interpretation, I scale attendance by dividing it by 10,000. I then interact the home gameday dummy variable with the demeaned attendance variable. Thus, the coefficient of home gameday represents the effect of a home gameday on accidents when the attendance is at its average level for that stadium. The coefficient of the interaction term indicates how the effect of a home gameday on accidents changes with a 10,000 increase in attendance above the stadium's average[5]. The underlying hypothesis is that a higher number of attendees at a sporting event may lead to an increased likelihood of accidents. Additionally, professional football games often attract a significant number of away-team supporters who travel to the host city, potentially contributing to accidents due to their

[4]Other games defined as prominent games: Packers vs. Vikings, Cowboys vs. Giants, Patriots vs. Colts, 49ers vs. Seahawks, Giants vs. Eagles, Broncos vs. Chiefs, Patriots vs. Jets, 49ers vs. Rams, Raiders vs. Steelers, Browns vs. Ravens, Colts vs. Texans, and Seahawks vs. Rams.

[5]Attendance at stadiums on non-gamedays, which is zero, is not factored into the average calculation. Since the home gameday is a dummy variable that takes the value 0 on non-gamedays, the interaction between demeaned attendance and gameday is zero anyway during non-gamedays.

Table 8: Estimated Effects of Stadium Attendance on Home Gamedays on Accidents

	0-5 Miles	**5-10 Miles**	**10-20 Miles**	**20-50 Miles**	**50-100 Miles**
Home gameday	0.145***	0.025	0.006	0.006	-0.025
	(0.039)	(0.064)	(0.063)	(0.032)	(0.023)
Home gameday X demeaned attendance	0.0112	-0.00701	-0.0293	-0.0781	-0.0237
	(0.035)	(0.031)	(0.039)	(0.093)	(0.043)
Observations	37,479	37,479	37,479	37,479	37,479
Number of geographi	31	31	31	31	31
County FE	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES
Holiday FE	YES	YES	YES	YES	YES
Weather Control	YES	YES	YES	YES	YES

*Statistically significant at 10% level; ** at 5% level; *** at 1% level.

Notes: The outcome variable is daily accident counts per geographic unit (county or 0-5, 5-10, 10-20, 20-50 miles distance ranges from stadiums). Date FE includes year, month and day of the week fixed effects. Population is used as an exposure variable. Standard errors in the parentheses are corrected for clustering at the geographic unit level.

unfamiliarity with the local roads and driving conditions.

Table 8 presents the results of this analysis. Across all specifications apart from the 0-5 mile radius, the findings indicate that the game day effect remains largely unchanged when accounting for the level of attendance at games. The coefficients for the game day variable are significant only within the 0-5 mile radius around the stadium, with a magnitude similar to that seen in Table 1. Surprisingly, the results reveal that the level of attendance itself does not appear to have a substantial impact on accident rates. The coefficients for the interaction terms are consistently small in magnitude and statistically insignificant across all specifications.

Discussion and Conclusion

NFL games, the most widely followed sports event in the US, are associated with large alcohol consumption among the spectators and significant traffic congestion due to large crowds, both of which create a high-risk environment for impaired driving and vehicular accidents. In this study, I examine the causal impact of NFL game days on traffic accidents, using a comprehensive accident dataset that spans six NFL seasons. Employing an identification strategy that leverages the quasi-random timing of game days, the analysis not only considers variations within counties over time but also examines impacts across different proximities to stadiums—from 0-5 miles up to 50-100 miles—using multiple regressions.

Findings from this paper suggest that NFL home games have significant impact on accident rates within the 0-5 mile radius of the stadium, with a 15.60% increase on game days. An hourly analysis of the data shows that accidents rise by 60.31% and 123.22% in the 3-hour windows before and after the game, respectively, within the same 0-5 mile distance range from the stadium. However, both the daily and 3-hour interval effects dissipate farther away from the stadiums and are not seen at the county level. Additionally, in the wider time windows of 3-6 hours before the start of the game and 3-6 hours after the game ends, I find no significant effects on accident rates, even within the 0-5 mile distance range. These findings suggest that the risk of accidents is highest during the immediate periods before and after the game, possibly due to peak traffic congestion and alcohol consumption by fans during these time frames near the stadium. The absence of significant effects during the broader time windows may also indicate return to normal traffic patterns and a decrease in stadium-related activity.

Unsurprisingly, away games do not appear to influence accident rates, as areas near away teams' stadiums likely don't see any significant increase in traffic and alcohol sales, compared to areas surrounding the home teams' stadium on game days. Regarding heterogeneous effects, game outcomes also play an important role, with victories leading to a 21.89% increase in accidents within the 0-5 mile radius, while losses result in a 13.84% smaller effect compared to wins. This is likely due to the heightened intensity of pre-game activities such as tailgating and social gatherings when fans are optimistic about their team's chances, as evidenced by the significantly larger increase in accidents before wins compared to losses in the pre-game period.

Interestingly, accounting for game attendance does not change the overall effect of game days on accidents, and attendance itself does not seem to impact accident rates. In addition, I also do not find the prominence of the games themselves having a notable impact on accidents. This suggests that factors such as traffic patterns and fan behavior might be more important than the actual number of spectators in explaining higher accident rates near stadiums on gameday.

My findings have important policy implications for traffic management and public safety. Given the significant increase in accident rates near the stadium during the hours immediately before and after the game, local authorities and stadium management should take mitigating actions to decrease the likelihood of accidents in those areas and hours. These can include increasing public transportation options to and from stadiums on game days, temporarily closing roads or diverting traffic near the stadiums, and reducing alcohol sales or implementing strict regulations.

Future research can look into differentiating the effects of alcohol consumption from traffic congestion by incorporating data on alcohol sales. This would give us a better understanding of the relative contributions of drunk driving and traffic congestion to the observed increase in accidents and help policymakers tailor their policies to be more effective, giving more importance to mitigating the factor that is more influential in explaining the rise in accidents. Additionally, using average traffic data in each area, future research could compare the relative increase in drivers on the road to the relative increase in accidents to examine whether the accident rate is elastic or inelastic with respect to traffic volume. The findings from this analysis could explain whether the increase in accidents is just normal traffic effect or if football fans exhibit different driving behaviors compared to the general population.

References

Andres, L., Fabel, M., & Rainer, H. (2023). How much violence does football hooliganism cause? *Journal of Public Economics*, 225, 104970. https://doi.org/10.1016/j.jpubeco.2023.104970

Badger, E., & Parlapiano, A. (2022, November 27). The exceptionally American problem of rising roadway deaths. *The New York Times*. https://www.nytimes.com/2022/11/27/upshot/road-deaths-pedestrians-cyclists.html

Global status report on road safety 2018. Geneva: World Health Organization; 2018. Licence: CC BY-NC-SA 3.0 IGO.

Gough, C. (2024, February 6). Average regular season attendance in the NFL. *Statista*. https://www.statista.com/statistics/249372/average-regular-season-attendance-in-the-nfl/

Lindo, J. M., Siminski, P., & Swensen, I. D. (2018). College party culture and sexual assault. *American Economic Journal: Applied Economics*, 10(1), 236–265. https://doi.org/10.1257/app.20160031

Marcus, M. (2021). Going beneath the surface: Petroleum pollution, regulation, and health. *American Economic Journal: Applied Economics*, 13(1), 1–37. https://doi.org/10.1257/app.20190130

Moosavi, S., Samavatian, M. H., Parthasarathy, S., & Ramnath, R. (2019). A countrywide traffic accident dataset (arXiv:1906.05409). arXiv. https://doi.org/10.48550/arXiv.1906.05409

National Center for Statistics and Analysis. (2023, June). Alcohol-impaired driving: 2021 (Traffic Safety Facts. Report No. DOT HS 813 450). National Highway Traffic Safety Administration.

National Center for Statistics and Analysis. (2024, April). Overview of motor vehicle traffic crashes in 2022 (Traffic Safety Facts Research Note. Report No. DOT HS 813 560). National Highway Traffic Safety Administration.

Neal, D. J., & Fromme, K. (2007). Event-level covariation of alcohol intoxication and behavioral risks during the first year of college. *Journal of Consulting and Clinical Psychology*, 75(2), 294–306. https://doi.org/10.1037/0022-006X.75.2.294

Nguyen, H.-L. Q. (2019). Are credit markets still local? Evidence from bank branch closings. *American Economic Journal: Applied Economics*, 11(1), 1–32. https://doi.org/10.1257/app.20170543

Redelmeier, D. A., & Stewart, C. L. (2003). Driving fatalities on super bowl sunday. *The New England Journal of Medicine*, 348(4), 368–369. https://doi.org/10.1056/NEJM200301233480423

Rees, D. I., & Schnepel, K. T. (2009). College football games and crime. *Journal of Sports Economics*, 10(1), 68–87. https://doi.org/10.1177/1527002508327389

Stewart, T. (2023, April). Overview of motor vehicle traffic crashes in 2021 (Report No. DOT HS 813 435). National Highway Traffic Safety Administration.

Toomey, T. L., Erickson, D. J., Lenk, K. M., & Kilian, G. R. (2008). Likelihood of illegal alcohol sales at professional sport stadiums. *Alcoholism: Clinical and Experimental Research*, 32(11), 1859–1864. https://doi.org/10.1111/j.1530-0277.2008.00770.x

Wood, S., McInnes, M. M., & Norton, D. A. (2011). The bad thing about good games: The relationship between close sporting events and game-day traffic fatalities. *Journal of Consumer Research*, 38(4), 611–621. https://doi.org/10.1086/660164

World Health Organization. (2018). Global status report on road safety 2018. https://www.who.int/publications/i/item/9789241565684

Appendix

Appendix Table 1: List of NFL Teams, Their Home Stadiums, and Geographic Information

NFL Team	Stadium	City	County	State
Cincinnati Bengals	Paul Brown Stadium	Cincinnati	Hamilton	OH
Arizona Cardinals	University of Phoenix Stadium	Glendale	Maricopa	AZ
Atlanta Falcons	Georgia Dome	Atlanta	Fulton	GA
Indianapolis Colts	Lucas Oil Stadium	Indianapolis	Marion	IN
New York Giants	MetLife Stadium	East Rutherford	Bergen	NJ
Chicago Bears	Soldier Field	Chicago	Cook	IL
Dallas Cowboys	Cowboys Stadium	Arlington	Tarrant	TX
San Francisco 49ers	Levi's Stadium	Santa Clara	Santa Clara	CA
Green Bay Packers	Lambeau Field	Green Bay	Brown	WI
Miami Dolphins	Sun Life Stadium	Miami Gardens	Miami-Dade	FL
Denver Broncos	Sports Authority Field at Mile High	Denver	Denver	CO
Carolina Panthers	Bank of America Stadium	Charlotte	Mecklenburg	NC
Kansas City Chiefs	Arrowhead Stadium	Kansas City	Jackson	MO
Buffalo Bills	Ralph Wilson Stadium	Orchard Park	Erie	NY
Cleveland Browns	FirstEnergy Stadium	Cleveland	Cuyahoga	OH
Houston Texans	Reliant Stadium	Houston	Harris	TX
Minnesota Vikings	TCF Bank Stadium	Minneapolis	Hennepin	MN
Washington Redskins	FedEx Field	Landover	Prince George's	MD
New England Patriots	Gillette Stadium	Foxborough	Norfolk	MA
New Orleans Saints	Louisiana Superdome	New Orleans	Orleans	LA
Jacksonville Jaguars	EverBank Field	Jacksonville	Duval	FL
Baltimore Ravens	M and T Bank Stadium	Baltimore	Baltimore	MD
Philadelphia Eagles	Lincoln Financial Field	Philadelphia	Philadelphia	PA
Tennessee Titans	Nissan Stadium	Nashville	Davidson	TN
Oakland Raiders	Oakland Coliseum	Oakland	Alameda	CA
Detroit Lions	Ford Field	Detroit	Wayne	MI
San Diego Chargers	Qualcomm Stadium	San Diego	San Diego	CA
Los Angeles Rams	Los Angeles Memorial Coliseum	Los Angeles	Los Angeles	CA
Pittsburgh Steelers	Heinz Field	Pittsburgh	Allegheny	PA
Seattle Seahawks	CenturyLink Field	Seattle	King	WA
Las Vegas Raiders	Allegiant Stadium	Paradise	Clark	NV
Tampa Bay Buccaneers	Raymond James Stadium	Tampa	Hillsborough	FL
New York Jets	MetLife Stadium	East Rutherford	Bergen	NJ

Appendix Figure 1: **Deaths from Road Accidents, per Million People (Collected from New York Times)**