

# Separately measuring home-field advantage for offenses and defenses: A panel-data study of constituent channels within collegiate American football

Matthew J. McMahon<sup>1</sup> | Sarah Marx Quintanar<sup>2</sup>

<sup>1</sup>Department of Economics and Finance,  
West Chester University, West Chester,  
Pennsylvania, USA

<sup>2</sup>Department of Economics, Finance, and  
General Business, Midwestern State  
University, Wichita Falls, Texas, USA

## Correspondence

Matthew J. McMahon, Department of  
Economics and Finance, West Chester  
University, 411 Business and Public  
Management Center, 50 Sharpless St.,  
West Chester, PA 19383, USA.  
Email: [matthew.mcmahon21@gmail.com](mailto:matthew.mcmahon21@gmail.com)

## Abstract

We improve constituent-channel estimates of home-field and neutral-site advantage for collegiate American football's top division by utilizing a richer, 12-season data set and by exploiting the COVID-19 pandemic as a random shock. Novel to the literature, we separately examine points scored by each team, allowing us to identify impacts on each team's offense and defense individually. The information set provided by our model is a strict superset of that provided by the previous standard in the literature, making ours a strictly dominant modeling choice. We demonstrate this improvement theoretically and empirically. Physiologically, away-team travel distance does not impact their own score, but it increases home-team scores, consistent with the notion that defenses tire faster than offenses. There is also similar but limited evidence of this effect for neutral-site teams. Time zones may play a minor role, too. Psychologically, crowd size and density hurt away-team scores but do not impact home or neutral-site teams. The away-team effect disappears in 2020, however, indicating that the pre-2020 effect is caused by the crowd's noise, not their mere presence. We also find that increasing stadium capacity while holding crowd size constant hurts home-team scores, highlighting the importance of considering ticket demand when considering stadium expansion. Tactically, stadium

familiarity helps offenses, not defenses, while team-opponent familiarity has the opposite effect. Weather also plays a role. At median values for key variables, we find an overall home-field advantage of 4.1 points.

#### KEY WORDS

COVID, football, home-field advantage, NCAA

#### JEL CLASSIFICATION

L83, Z2, C51

## 1 | INTRODUCTION

Home-field advantage describes the well-established sports phenomenon that teams perform better when playing on their own field. The effect has been documented and studied across a wide variety of sports, from soccer (association football) to Australian-rules football (Jamieson, 2010; Reade et al., 2022; Stefani & Clarke, 1992). The overall advantage is traditionally measured as an increased point spread in favor of the home team, but more recent work has looked into identifying and quantifying the underlying channels through which that change in point spread is created (e.g., Pollard, 2008; Wang et al., 2011). For example, the home team may gain a relative advantage due to crowd size (Boudreux et al., 2017; Inan, 2020), stadium familiarity (Carmichael & Thomas, 2005; Jamieson, 2010; Pollard, 2002), or their opponent's travel fatigue (Nevill & Holder, 1999; Nichols, 2014). The COVID-19 pandemic has sparked increased interest in home-field advantage. Recent work has exploited increased attendance variation from COVID-19 mitigation policies to better identify home-field advantage magnitudes in many sports (Cross & Uhrig, 2020; Fischer & Haucap, 2021), and there is also evidence that the pandemic has changed the impact of various home-field advantage factors, shrinking both officiating bias and overall home-field advantage in professional soccer (Bryson et al., 2021; The Economist, 2020; Wunderlich et al., 2021). In collegiate American football, Dodd (2020) found the lowest raw measure of home-field advantage since 2005 through 4 weeks of play at the beginning the 2020 season, with 59.5% of home games being won by home teams, and uncertainty surrounding the impact of the pandemic caused bettors in Las Vegas to estimate a two-point decrease in home-field advantage (Bozich, 2020). Overall, suggestive evidence predicts home-field advantage in college football would diminish, but the data has not been thoroughly investigated until now (Dodd, 2020).

We use data from all 131 teams in the top level of collegiate American football for 12 seasons—including the 2020 season, which was uniquely impacted by the COVID-19 pandemic—to contribute to the home-field advantage literature in several important ways. First, we thoroughly explore constituent factors through which home-field advantage may operate, including many not previously studied within college football (Caudill & Mixon Jr., 2007; Fullagar et al., 2019; Jamieson, 2010; Wang et al., 2011). Following Stefani and Clarke (1992), we categorize these factors within three main channels: physiological (e.g., travel fatigue), psychological (e.g., crowd presence), or tactical (e.g., the team's familiarity with the stadium). Second, we use a novel empirical setup which offers greater detail than the previous literature by measuring home-field advantage in terms of the impact on the number of points scored by each team, instead of aggregating to the cross-team point spread. This allows us to identify each

factor's impact on each team's offense and each team's defense individually. Estimates of the point-spread model are recoverable from our model, while the reverse is not true, which implies that the information set provided by our model is a strict superset of that provided by the previous standard in the literature. We provide intuitive, theoretical, and empirical comparisons of the two models, using the same data, to illustrate this in Sections 4.3 and 5.2. Third, we include neutral-site games in our analysis. Thus, in addition to the advantage a team gains from playing at home rather than away, we also measure the relative (dis)advantage of playing at a neutral site. While not ignored by all, such games are omitted by most of the home-field advantage literature despite their prevalence and importance, especially in college football. Fourth, the impact of the COVID-19 pandemic and corresponding attendance policies during the 2020 season provide significantly greater exogenous variation for many key variables, such as crowd size, allowing us to more precisely estimate the effects of these variables (Moore & Brylinsky, 1993). Last, we also investigate whether the effects of these channels changed during the 2020 season.

Our results support many commonly held narratives within college football. The main physiological factor that impacts home-field advantage is travel: home teams score more points for each additional mile away teams travel to play the game, *ceteris paribus*. Away-team points are not impacted, though. This supports the popular notion that defenses tire faster than offenses. (This is because, as discussed in Section 4.3, if defenses tire first, it is defensive fatigue that is the binding constraint on how many points the offense can score). We also find some evidence suggesting travel across time zones can be disadvantageous, all else (including distance) held equal. Tactically, familiarity also plays a role. Teams tend to score more points for each additional game they have recently played in the stadium, all else equal.<sup>1</sup> This effect does not extend to the other side of the ball, however, indicating that stadium familiarity helps offenses more than defenses. On the other hand, increased team-opponent familiarity causes decreased scoring, which is in line with the intuition that defenses who better anticipate offenses' actions are more successful.

Since crowd noise disrupts on-field communication, which (in football) is more vital to the offense, the rule of thumb is that home crowds stay quiet when the home team is on offense yet get loud when the away team is on offense. Thus, increased crowd size should negatively impact the number of points away teams score but not those scored by home teams. In fact, we find increased crowd size (or density) causes away teams to score fewer points, though it does not impact home or neutral-site teams. Additionally, increased fan capacity—holding crowd size constant—negatively impacts home team scores. This serves as a caution to university administrators that expanding stadiums without considering demand could hurt team performance. To our knowledge, this is the first study to test crowd capacity in this context. Beyond the factors for which we directly control, we find no further “leftover” or “residual” home-field (or neutral-site) effects.

We also contribute to the home-field advantage literature by exploiting the random shock of the COVID-19 pandemic on college football. The COVID-19 pandemic has dramatically affected sports teams, their fans, and their revenues since March 2020. Estimates range from a 9% increase in deaths after indoor mass gatherings in the United States (Ahammar et al., 2023) to a single men's college basketball game in the first week of the pandemic in the US directly causing 55 cases and 4 deaths (Wing et al., 2021). Researchers have identified risk of spread from attendees (Dave et al., 2021) along with willingness to pay based on safety precautions (Humphreys et al., 2023). Home-field advantage in other sports has been studied in this context, most notably in European football (Cross & Uhrig, 2020; Fischer & Haucap, 2021).

<sup>1</sup>As discussed in Section 3, we define “recently played” as the number of times a team has played in the given stadium previously within the same season or at any point in the preceding two seasons.

In the context of our work, the most notable way in which the 2020 season was affected by the COVID-19 pandemic was restrictions on crowd size, which varied greatly across both location and time over the course of the season (Auerbach, 2020; Bender, 2020; SDS Staff, 2020). In professional soccer leagues, Bryson et al. (2021) find a significant decrease in the number of yellow cards issued when no crowd is present, but no other impact on home-field advantage. Both Cross and Uhrig (2020) and Fischer and Haucap (2021) find evidence of decreased home-field advantage in top leagues using exogenous variation of fan attendance driven by COVID-19 policies. This effect does not seem to be driven by referee behavior or teams' tactics.

Though these studies are relevant to the present work, it has remained unclear until now as to whether college football follows a pattern similar to professional soccer. Prior to 2020, crowd size and overall stadium noise in college football are more closely related, and we find that crowd size negatively impacts away-team scores in this period. Given the general decrease in crowds in 2020, however, each conference created its own volume limit for stadium sound systems, allowing stadium sound-system managers to pipe in “fake crowd noise” up to that limit. At least to some extent, this decouples crowd size and overall stadium noise in 2020, making crowd size in 2020 more accurately a measure of the psychological impact of fan presence rather than of the noise they create. We find no effect of crowd size on points scored by home or away teams in 2020, which suggests that this factor's effect in pre-pandemic seasons is driven by noise.

The remainder of this paper is organized as follows: Section 2 gives a background on collegiate American football, which may be particularly helpful for readers who are less familiar with the sport. Section 3 describes our data and provides summary statistics. Section 4 describes our empirical setup, including a brief intuitive discussion of our model's improvement over the standard model. We present our main empirical results in Section 5, including the impact of various home-field advantage factors via our main model specification, theory and empirics illustrating the relationship between our model and the literature's standard model, evidence of the consistency of our data and model within the literature, and the difference in various home-field advantage factors' impacts during the 2020 season due to the COVID-19 pandemic. Section 6 explores a series of robustness checks and model extensions. Last, Section 7 discusses our findings in a broader context and concludes.

## 2 | A PRIMER ON COLLEGIATE AMERICAN FOOTBALL

While college athletes in the United States are legally considered amateurs, for many sports—including football—top colleges serve a role closer to that of professional teams, with the market value of top athletes far exceeding the minimal payment in-kind (e.g., free tuition, room, and board) they receive (Garthwaite et al., 2020; Sanderson & Siegfried, 2015). The distinction between amateur and professional athletes in the context of college football is a question of legal status—not a question of skill—as has been both discussed and lampooned from federal courtrooms and academic law journals (e.g., NCAA vs. Alston, 2020; O'Bannon vs. NCAA, 2014; Rosenthal, 2017) to the popular press and mainstream social commentary (e.g., Bokat-Lindell, 2020; Connelly, 2020; Parker & Stone, 2011; Wilbon, 2011).<sup>2</sup> Indeed, the first known

<sup>2</sup>Recent evidence shows that a majority of American adults now support the payment of college student-athletes (Knoester & Ridpath, 2020), essentially recognizing them as professionals. In fact, as of 2021, college players can be paid by third parties for use of their name, image, and likeness (Hosick, 2021; Samuels, 2021), though this began after our sample period.

game of American football was played between students of two rival universities, lending a history that predates both professional football and even organized non-collegiate amateur American football by decades. Even as professional football developed and surpassed college football's nationwide popularity, few professional teams existed outside the Northeast and northern Midwest until relatively late in the 20th century, and as such college football is still significantly more popular across much of the country.

We focus our analysis on the top level of competition—the Division I Football Bowl Subdivision (or FBS, previously known as Division I-A). As it currently stands, the National Collegiate Athletic Association (NCAA) allows for FBS teams to count one win per season over a team from the Division I Football Championship Subdivision (or FCS, previously Division I-AA) toward post-season eligibility.<sup>3</sup> However, as discussed in Section 3, we omit all FBS-vs-FCS games from our analysis.

Because football is a particularly brutal sport, college players are limited to roughly one game per week. Each team has one bye week during the 13-week season, meaning they play 12 regular-season games<sup>4</sup> with the potential for post-season games.<sup>5</sup> With roughly 130 FBS teams in any given season, matchups are at a premium. Even most conferences are too large for a round-robin schedule, and thus conferences frequently subdivide into two conference divisions. In such conferences, each team plays all other teams within their division each season, with each pair of teams generally alternating which is the home team from one season to the next. Teams in such conferences also generally play one or two games against teams from the other division of their conference, often with one fixed opponent and one slot that rotates each season. In this way, most teams are not much more familiar with their non-divisional conference-mates than they are with teams in other conferences, which we take into account in our empirical specifications.

Beyond the regular season, FBS football also has a unique post-season arrangement. In addition to conference championship games and the small handful of teams for whom the post-season involves competing for a national championship, other teams who perform well during the regular season are also rewarded with a single post-season game. These games, known as “bowl games,” are generally played at a neutral site against a non-conference opponent.

While, across all levels of collegiate sports, Division I FBS football is the only one for which the NCAA does not officially crown a national champion, a formalized method for determining a widely-regarded “champion” team has been in place in some form since the 1998 season.<sup>6</sup> From 1998 through the 2013 season (which includes the first part of our sample period), the Bowl Championship Series (BCS) selection system determined which two teams played in the national championship game, as well as which other teams would play in the remaining four “BCS bowl” games.<sup>7</sup> Starting with the 2014 season, the BCS system was replaced by the College Football Playoff (CFP) system. The CFP system ranks top teams and assigns them slots

<sup>3</sup>The FCS is the second-highest level of competition. FBS teams cannot count wins over Division II or Division III teams toward bowl eligibility, however, and so we observe only one FBS-versus-DII game and no FBS-versus-DIII games in our data set. We drop the FBS-versus-DII game from all analysis.

<sup>4</sup>Given the uniquely large travel costs for Hawaii, they are permitted to play one extra game (as a home game) each season to help recoup their away-game travel costs, and thus they play a 13-game regular season.

<sup>5</sup>The possible post-season games include: conference championship games, bowl games, and (since 2014) the additional (unofficial) national championship game.

<sup>6</sup>Disputes still occur, however. Most recently, the University of Central Florida claims a national championship for their undefeated 2017 season despite not playing in the unofficial “national championship” game (Heim, 2018).

<sup>7</sup>There were only three other BCS bowl games through the 2004 season, but this predates our sample period.

in the “New Year’s Six” (NY6) bowls, two of which (rotating by year) serve as the semi-final games for the CFP National Championship Game (which is still not officially an NCAA-sanctioned championship). We empirically designate the five BCS bowls and the seven NY6 bowls (including the championship game in each) as “top” bowl games.

Within both the BCS and CFP systems, preference has always been given (albeit in various, evolving ways) to teams in “top” FBS conferences. Under the BCS system, the six “automatic qualifying” (AQ) conferences and the independent (i.e., non-conference) Notre Dame University were given preference. The CFP system gives preference to the “Power 5” (P5) conferences and Notre Dame over the remaining “Group of 5” (G5) conferences and other independent teams.<sup>8</sup> Because preferential treatment under each system is a rough reflection of past and present on-field performance, we empirically denote a team as a “top-tier” team in a given season if and only if they are in an AQ conference in that season (for BCS seasons), are in a P5 conference in that season (for CFP seasons), or are Notre Dame University (for all seasons). The remainder are all denoted as “bottom-tier” teams within the FBS.

### 3 | DATA

Our data set includes details regarding all games played by an FBS team from the 2009 through 2020 seasons. For each game, this includes the date, the two teams playing, whether the game was played at a neutral site, which team was the home team (if not played at a neutral site), the teams’ national rankings (if ranked), the number of points scored by each team, which team won the game, and the number of people attending the game.<sup>9,10</sup> We have identifiers for conference championship games, bowl games, “BCS” bowl games, “New Year’s Six” bowl games, BCS national championship games, college football playoff (CFP) games, and CFP national championship games. We also append specific yearly characteristics of each stadium to our main data set, including (current) stadium name, city, state, country, capacity (including whether stadiums are set up in their “expanded” formation), year built, playing surface (including type of grass/turf), latitude, longitude, and the year of the most recent expansion/renovation.<sup>11</sup> Last, we calculate and include variables such as team and opponent winning percentage from the previous season and the number of times a team has played at a stadium either earlier in the same season or at any time in the preceding two seasons.

Robustness checks and model extensions in Section 6 necessitate additional data: time-zones traveled, weather conditions, and head coach for each team. In addition to an indicator for games played in domes, our weather controls also include wind speed, temperature, and whether there was precipitation actively falling at kickoff for each game. We utilize head coach data to create a measure of familiarity between opposing head coaches.<sup>12</sup>

<sup>8</sup>AQ conferences included the ACC, Big 12, Big East, Big 10, Pac-10/Pac-12, and SEC through the 2012 season. The Big East was replaced by the AAC for the 2013 season when the Big East dropped football from being a sponsored sport. The P5 conferences include the ACC, Big 12, Big 10, Pac-12, and SEC.

<sup>9</sup>This game data is scraped from <http://www.cfbstats.com>.

<sup>10</sup>Attendance data measures actual gate attendance, not ticket sales.

<sup>11</sup>Data was compiled and cross-checked using various sources, which are listed in the source appendix.

<sup>12</sup>Weather data is exclusively from <https://collegefootballdata.com>. Head coaching data is largely from <https://collegefootballdata.com>, though it also includes data from other sources. See the source appendix for a comprehensive list.

Home-field advantage is created by many factors, such as the distance each team travels to the game, each team's familiarity with the stadium, and both overall crowd size and the relative allocation of ticket sales between the two schools' fan bases (Jamieson, 2010; Pollard, 1986; Wang et al., 2011). In distinguishing between which games should be considered a "neutral-site" game versus a "home" (or "away") game, we create two rules that consider these factors. First, a stadium is considered a home stadium for a team in a season if that team plays at least three games within that season at that stadium. For example, the University of California (Berkeley) Golden Bears' usual home stadium, known as California Memorial Stadium, underwent renovations throughout the 2011 season. The team's "home" games in that season were instead played at what is now known as Oracle Park, where the San Francisco Giants, a Major League Baseball team, play their home games.

Second, we consider stadium repetition across seasons important in designating potential "home" stadiums: if a team repeatedly plays at a specific stadium over multiple years and that stadium is not a home field for the other team, players likely experience similar benefits to playing at home. Specifically, we define a stadium to be a "home stadium" for team  $i$  when playing against opponent  $j$  if it meets the following three criteria: team  $i$  must have played at least one game at the stadium in at least three of the prior four seasons; the stadium is not a home stadium for opponent  $j$ , whether through the standard definition or through the first criterion here (for instance, rivalry games such as the "Red River Rivalry" between Oklahoma and Texas held annually at the Texas State Fair); and the game is neither a conference championship game nor a bowl game (since ticket sales are allocated equally between the two teams' fan bases for these games).

For example, the University of Arkansas Razorbacks, based in Fayetteville, played two games per season at War Memorial Stadium in Little Rock from 2009 through 2013 and one game per season there from 2014 through 2019. Thus, while the team may not have three games at War Memorial Stadium within any single season, the players are surely familiar with the stadium due to the repetition across seasons. For this reason, along with the significant alumni fan base in Little Rock, we designate these as "home" games for the University of Arkansas, though the "distance traveled" variable still accurately represents that the Razorbacks traveled nearly 200 miles to play each of these "home" games.

### 3.1 | Summary statistics

Table 1 presents the mean and standard deviation of our main specification variables. Mean and standard deviation for each variable are given separately for home, away, and neutral-site games, as well as pooled across all game locations. Team distance traveled is measured in hundreds of miles. Crowd and capacity are measured in tens of thousands. "Familiarity (2-Year)" variables measure the total number of occurrences at any previous point in the given season and any time during the two preceding seasons. FBS conference tiers and bowl tiers are defined in Section 2.

While a quick observation offers only suggestive evidence, a few trends are immediately noticeable. For example, home teams score an average of nearly 4.4 more points than away teams, indicating rudimentary evidence of a home-field advantage. These trends help motivate our further study of home-field advantage and its constituent factors.

**TABLE 1** Summary Statistics

	<b>Home</b>	<b>Away</b>	<b>Neutral Site</b>	<b>Pooled</b>
<b>Non-Indicator Variables</b>				
Team Points	30.2 (14.4)	25.8 (14.0)	28.0 (13.0)	28.0 (14.3)
Team Distance Traveled	0.00 (0.11)	5.77 (5.34)	6.28 (7.33)	3.15 (5.09)
Crowd Size	4.18 (2.75)	4.18 (2.75)	4.81 (2.29)	4.23 (2.73)
Stadium Capacity	5.36 (2.32)	5.36 (2.32)	6.38 (1.70)	5.44 (2.29)
Familiarity (2-Year)				
Team-Stadium	15.4 (2.7)	0.6 (0.5)	0.7 (1.9)	7.4 (7.6)
Opponent-Stadium	0.6 (0.5)	15.4 (2.7)	0.7 (1.9)	7.4 (7.6)
Team-Opponent	1.3 (0.9)	1.3 (0.9)	0.5 (0.9)	1.2 (0.9)
<b>Indicator Variables</b>				
FBS Conference Tiers				
Team Top × Opp. Top	0.449 (0.497)	0.449 (0.497)	0.578 (0.494)	0.459 (0.498)
Team Top × Opp. Bottom	0.112 (0.315)	0.035 (0.183)	0.092 (0.289)	0.075 (0.263)
Team Bottom × Opp. Top	0.035 (0.183)	0.112 (0.315)	0.092 (0.289)	0.075 (0.263)
Team Bottom × Opp. Bottom	0.405 (0.491)	0.405 (0.491)	0.239 (0.426)	0.392 (0.488)
Conference Championship Game	0.004 (0.061)	0.004 (0.061)	0.086 (0.280)	0.010 (0.1000)
Bowl Game	0 (0)	0 (0)	0.630 (0.483)	0.049 (0.217)
Top Bowl Game (BCS/NY6)	0 (0)	0 (0)	0.106 (0.308)	0.008 (0.091)

*Notes:* The mean and standard deviation is given for each variable at each game location, as well as pooled across all game locations. Team distance traveled is measured in hundreds of miles. Crowd and capacity are measured in tens of thousands. “Familiarity (2-Year)” variables measure the total number of occurrences at any previous point in the given season and any time during the two preceding seasons. As such, they begin with the 2011 season. “Top” bowls indicate BCS and NY6 bowls. See Section 2 for more details on FBS conference tiers and bowl games. Values denoted “0” are exactly zero, whereas values denoted “0.00” are positive but round down to zero.

## 4 | MODEL

### 4.1 | Our model

Much existing work on home-field advantage uses the cross-team point differential as the outcome variable (e.g., Cross & Uhrig, 2020; Jamieson, 2010; Wang et al., 2011). There is one main drawback to this approach, however. It is possible that home-field advantage affects offensive and defensive performances via separate channels, potentially having non-equal impacts on each side. Separating these effects is not possible when combining points scored by each team together to the point-differential level, meaning the mechanisms of these effects are not fully considered. We illustrate this issue in more detail in Sections 4.3 and 5.2. Similar to the setup of Carmichael and Thomas (2005), we separately model home-field effects for each of the two participating teams to avoid this drawback.

More formally, we can consider a participating team to be a standard firm that utilizes inputs, such as players and coaches, to produce output, which is the number of points the team scores. In that sense, the marginal product of inputs can vary based on whether a team is playing at home, away, or at a neutral site. As mentioned previously, home-field advantage can be discussed within three general categories: physiological factors (e.g., travel fatigue), psychological factors (e.g., crowd presence), and tactical factors (e.g., team's familiarity with the venue) (Pollard, 1986, 2008; Stefani & Clarke, 1992). We can thus compare the marginal impact of a given input for home teams relative to that for away teams (as well as for neutral-site teams) to better understand each component of the overall home-field advantage.

With these considerations in mind, we estimate the following model as our main specification:

$$\begin{aligned}
 \text{teampts}_{ig} = & \beta_0 + \beta_1 \cdot \text{home}_{ig} \cdot \text{teamdist}_{ig} + \beta_2 \cdot \text{home}_{ig} \cdot \text{oppdist}_{jg} + \\
 & \beta_3 \cdot \text{away}_{ig} \cdot \text{teamdist}_{ig} + \beta_4 \cdot \text{away}_{ig} \cdot \text{oppdist}_{jg} + \\
 & \beta_5 \cdot \text{ns}_{ig} \cdot \text{teamdist}_{ig} + \beta_6 \cdot \text{ns}_{ig} \cdot \text{oppdist}_{jg} + \\
 & \beta_7 \cdot \text{home}_{ig} \cdot \text{crowd}_g + \beta_8 \cdot \text{away}_{ig} \cdot \text{crowd}_g + \beta_9 \cdot \text{ns}_{ig} \cdot \text{crowd}_g + \\
 & \beta_{10} \cdot \text{teamstadfam}_{ig} + \beta_{11} \cdot \text{oppstadfam}_{jg} + \\
 & \beta_{12} \cdot \text{home}_{ig} + \beta_{13} \cdot \text{ns}_{ig} + \\
 & \beta_{14} \cdot \text{teamwinpercprevseason}_{ig} + \beta_{15} \cdot \text{oppwinpercprevseason}_{jg} + \\
 & \beta_{16} \cdot \text{teamtop}_{ig} \cdot \text{oppbottom}_{jg} + \beta_{17} \cdot \text{teambottom}_{ig} + \\
 & \beta_{18} \cdot \text{teambottom}_{ig} \cdot \text{oppbottom}_{jg} + \\
 & \beta_{19} \cdot \text{teamoppfam}_{ijg} + \beta_{20} \cdot \text{ccgame}_g + \\
 & \beta_{21} \cdot \text{bowlgame}_g + \beta_{22} \cdot \text{topbowlgame}_g + u_i + v_j + \omega_s + \varepsilon_{ijg}
 \end{aligned} \tag{1}$$

where the dependent variable  $\text{teampts}_{ig}$  measures the number of points scored by team  $i$  (against opponent  $j$ ) in game  $g$ . Indicator variables  $\text{home}_{ig}$ ,  $\text{away}_{ig}$ , and  $\text{ns}_{ig}$  respectively denote whether team  $i$  played at home, away, or at a neutral site in game  $g$ . We cross these with variables that we expect to have differential impacts depending on the team's location for the given game (i.e., whether at home, away, or at a neutral site). Our data set allows us to measure physiological factors using distance traveled by team  $i$  and opponent  $j$  ( $\text{teamdist}_{ig}$  and  $\text{oppdist}_{jg}$ , respectively; measured in hundreds of miles), psychological factors using crowd size at game  $g$

( $\text{crowd}_g$ , measured in tens of thousands of people), and tactical factors by proxying for familiarity with the stadium using the number of games each team has played in the stadium either earlier in the same season or at any point in the previous two seasons ( $\text{teamstadfam}_{ig}$  for team-stadium familiarity and  $\text{oppstadfam}_{jg}$  for opponent-stadium familiarity).<sup>13</sup> The two-season look-back window for the latter two variables precludes us from using the 2009 and 2010 seasons in our estimation. Similarly, because our data is incomplete for FCS teams, we must exclude all games involving an FCS team (i.e., unless otherwise noted, all analysis in this paper solely examines FBS-vs-FBS games) as well as the first two FBS seasons for any of the 12 teams that transitioned up to the FBS during our sample period.

Beyond what we capture via specific home-field advantage factors, there may be a further home-field (and/or neutral-site) advantage, perhaps due to other factors for which we lack data, for example. Any such “leftover”—or “residual”—home-field advantage is captured by stand-alone indicators for home and neutral-site teams ( $\text{home}_{ig}$  and  $\text{ns}_{ig}$ , respectively), relative to the away-team baseline. Additional covariates include each team’s winning percentage from the previous season ( $\text{teamwinpercprevseason}_{ig}$  and  $\text{oppwinpercprevseason}_{jg}$ ), which proxy for team  $i$ ’s and opponent  $j$ ’s respective skill levels; indicators for whether team  $i$  and opponent  $j$  are in top- or bottom-tier conferences as of game  $g$  (arranged as  $\text{teamtop}_{ig} \times \text{oppbottom}_{jg}$ ,  $\text{teambottom}_{ig}$ , and  $\text{teambottom}_{ig} \times \text{oppbottom}_{jg}$ ); and indicators for each type of postseason game (i.e., denoting conference championship games with  $\text{ccgame}_g$ , bowl games with  $\text{bowlgame}_g$ , and BCS/NY6 bowl games with  $\text{topbowlgame}_g$ ).<sup>14</sup> We also include fixed effects at the team ( $i$ ), opponent ( $j$ ), and season ( $s$ , where the set of all  $s$  is a partition of the set of all  $g$ ), leaving  $\varepsilon_{ijg}$  as the idiosyncratic error term.<sup>15</sup>

## 4.2 | How to interpret our model

Broadly speaking, the coefficient estimates in our model show the marginal impact of the given covariate on the number of points scored by team  $i$  in game  $g$ , ceteris paribus. For example, the average impact of an additional 10,000 fans attending a home-away (i.e., non-neutral-site) game is for the home team to score  $\beta_7$  more points and for the away team to score  $\beta_8$  more points, all else equal. This could equivalently be interpreted as an additional 10,000 fans giving the home team an advantage equal to the difference in these estimates, or  $\beta_7 - \beta_8$ .

The four coefficients on the terms interacting home or away dummies with distance traveled by team or opponent likely require the most thought to properly interpret. This is largely because, as discussed in Section 3, there are a small handful of games where the “home” team must travel a nonzero distance to the game. These are captured by  $\beta_1$  and  $\beta_4$ . The coefficient  $\beta_1$  captures the impact on team  $i$ ’s points scored when team  $i$  travels an additional 100 miles, given

<sup>13</sup>We look back over the current plus the previous two seasons because players must play a minimum of three seasons in college before entering the NFL, leaving the best players with experience looking back over the past 2 years. Additionally, many players—especially those who play four or the maximum of five seasons in college—do not get as much time playing during their first year.

<sup>14</sup>Conference quality and bowl games are defined in detail in Section 2.

<sup>15</sup>In the interest of recent work on two-way (and higher dimensional) fixed-effects models (e.g., Imai & Kim, 2021), we also run a robustness check that eliminates season-specific fixed effects. All significant coefficient estimates from our main model yield the same sign and maintain statistical significance, and all estimates that were not previously significant still lack statistical significance. The results table is omitted for brevity.

that team  $i$  is the home team.<sup>16</sup> The coefficient  $\beta_4$  captures the impact on away team  $i$ 's points scored when their "home" opponent  $j$  travels an additional 100 miles. These covariates are essentially control variables, however, since their coefficients' interpretations are not particularly meaningful. On the other hand,  $\beta_2$  and  $\beta_3$  are coefficients of interest. The coefficient  $\beta_2$  captures the impact on the number of points scored by home team  $i$  when opponent  $j$  travels an additional 100 miles. The coefficient  $\beta_3$  captures the impact on away team  $i$ 's points scored when away team  $i$  travels an additional 100 miles.<sup>17</sup>

Importantly, this setup allows us to parse apart the effects on team  $i$ 's offense separately from those on team  $i$ 's defense, as well as separately from the effects on opponent  $j$ 's offense or defense. We achieve this by utilizing covariates specific to either team  $i$  or opponent  $j$  in conjunction with our use of two observations per game and a (solely) team- $i$ -specific dependent variable. For example, consider the various potential impacts of travel in our atypical home-away games, where both the "home" and away teams travel. From the home team's perspective—meaning when  $home_{ig}$  equals one—their own score is the product of their own home offense taking the field against their away opponent  $j$ 's defense. Of course, each side's own travel has a physiological effect only on its own side's players. Thus, the impact of the home team's travel on the home team's offense is given by  $\beta_1$ , while the impact of the away opponent's travel on the away opponent's defense is given by  $\beta_2$ , and thus the total impact on home team  $i$ 's score is  $\beta_1 + \beta_2$ . Similarly, in the other observation from the same game (i.e., when the away team is team  $i$ ), the physiological effect of the away team's travel on the away team's offense is given by  $\beta_3$ , while that of the home opponent's travel on the home opponent's defense is given by  $\beta_4$ , and thus the total impact on away team  $i$ 's score is  $\beta_3 + \beta_4$ .<sup>18</sup>

An interesting note regarding symmetry arises here, too. Readers who are accustomed to the home-field advantage literature may initially find our estimates' lack of symmetry to one another counterintuitive. However, there is no reason to expect, for example,  $\beta_2$  and  $\beta_3$  to be negatives of each other, and indeed we do not find any such relationship in practice.<sup>19</sup> More formally, we can identify the conditions under which symmetry necessarily occurs. Let a variable be "singly team-opponent symmetric" if, for any game  $g$ , the variable's value for one team equals the negative of that same variable's value for the other team.<sup>20</sup> Similarly, let two variables be "pairwise team-opponent symmetric" to each other if, for any game  $g$ , the value of one

<sup>16</sup>That is,  $\beta_1$  tells us, for example, when Arkansas plays a "home" game that is not in Fayetteville, how many additional points Arkansas scores when "home" team Arkansas travels an additional 100 miles.

<sup>17</sup>For example, suppose Cincinnati (away) plays a game at Ohio State (home). Rounding to the nearest tenth of a mile, Ohio State's stadium is exactly 100 miles from Cincinnati's stadium. The coefficient  $\beta_2$  describes how many additional points Ohio State scores from Cincinnati traveling that 100 miles, all else equal, while  $\beta_3$  represents the change in points that Cincinnati scores due to Cincinnati traveling that 100 miles, all else equal.

<sup>18</sup>For example, the 2015 game between Toledo and Arkansas was a "home" game for Arkansas, despite being held in Little Rock rather than in Arkansas's typical home stadium in Fayetteville. The impact of Arkansas's travel on Arkansas's offense is captured by  $\beta_1$ , while  $\beta_2$  captures the impact of Toledo's travel on Toledo's defense; similarly,  $\beta_3$  captures the impact of Toledo's travel on Toledo's offense, while  $\beta_4$  captures the impact of Arkansas's travel on Arkansas's defense. The total effect of both teams' travel on Arkansas's points scored is  $\beta_1 + \beta_2$ , while the total effect of both teams' travel on Toledo's points scored is  $\beta_3 + \beta_4$ .

<sup>19</sup>That is, there is no reason to expect the marginal impact of Cincinnati's travel on Ohio State's points scored to be the negative of the impact of Cincinnati's travel on Cincinnati's points scored.

<sup>20</sup>For example, if the point differential increases by some value from the home team's perspective, then naturally it must decrease by that same value from the away team's perspective. This is because, for any given game, the point differential for the away team must equal the negative of the point differential for the home team.

variable for one team equals the value of the other variable for the other team, and vice versa.<sup>21</sup> If the dependent variable is singly team-opponent symmetric and if all other components of the estimation equation (all independent variables and fixed effects, any random effects, etc.) are either team-opponent symmetric—singly and/or pairwise—or do not vary across team and opponent within a given game (such as crowd size), then for all pairwise symmetric variables, their coefficient estimates will be the negative of each other's. The existing literature largely uses point differential as the dependent variable, and hence much previous work yields team-opponent symmetric estimates.<sup>22</sup> While the right-hand side of Equation (1) meets the necessary criteria, the dependent variable—points scored by team  $i$ —is not singly team-opponent symmetric. Thus, our estimation of Equation (1) does not yield positive-negative coefficient symmetry.<sup>23</sup>

We designed our model with the intent of controlling for all confounding factors, whether directly, via inclusion of the variable itself, or indirectly, for example via proxy variables—such as lagged team win percentage (to control for reverse causality between crowd size and points scored)—or via fixed effects. If our model indeed succeeds at capturing all potential confounding effects, then the coefficient estimates for the structural components of our model, such as crowd size, can be interpreted causally. (The coefficients on proxy variables included solely to control for confounders—such as lagged team win percentage—do not necessarily have meaningful causal interpretations, though). This is in line with the literature on home-field advantage (e.g., Böheim et al., 2019; Carmichael & Thomas, 2005; Ferraresi & Gucciardi, 2023; Jamieson, 2010; Wunderlich et al., 2021) and causality in panel data models more generally (Athey et al., 2021; Baker et al., 2022; Borusyak et al., 2021; Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021; Imai & Kim, 2021). However, we acknowledge that the astute reader may be able to identify confounding factors for which we do not adequately control, which would weaken the argument for a causal interpretation of our model. We welcome such critiques whole-heartedly, and we look eagerly toward future work in the literature that more adequately alleviates such issues to capture the underlying causality more robustly.

#### 4.3 | Improvement over standard models

Our model offers significant advantages over previous models in the home-field advantage literature. First, our specification uses team  $i$ 's points scored in game  $g$  ( $\text{teampts}_{ig}$ ) as the dependent variable, while the majority of previous models use the point differential between team  $i$  and opponent  $j$  in game  $g$  ( $\text{ptdiff}_{ijg} = \text{teampts}_{ig} - \text{teampts}_{jg}$ ) as the dependent variable.<sup>24</sup> Note that the latter is an aggregated measure of the former; this mechanical difference means that we can

<sup>21</sup>For example, consider a neutral-site game between LSU and Tennessee.  $\text{TeamDist}_{ig}$  from LSU's perspective equals  $\text{OppDist}_{jg}$  from Tennessee's perspective, and  $\text{OppDist}_{jg}$  from LSU's perspective equals  $\text{TeamDist}_{ig}$  from Tennessee's perspective. Thus,  $\text{TeamDist}_{ig}$  and  $\text{OppDist}_{jg}$  are pairwise symmetric to one another.

<sup>22</sup>Not all previous work is team-opponent symmetric, however. For example, Wang et al. (2011) model is not team-opponent symmetric because it includes random effects for team  $i$  but not for opponent  $j$ .

<sup>23</sup>However, for Equation (1), replacing the dependent variable  $\text{TeamPts}_{ig}$  with points conceded by team  $i$  (i.e., points scored by opponent  $j$ )—which is pairwise-symmetric to  $\text{TeamPts}_{ig}$ —would switch all coefficient estimates with those to which they are pairwise-symmetric (e.g.,  $\beta_2$  and  $\beta_3$ ).

<sup>24</sup>Alternatively, some papers use probability models where the dependent variable is a dummy variable indicating whether team  $i$  won game  $g$ . This is essentially an indicator for if  $\text{teampts}_{ig} > \text{teampts}_{jg}$ , and is thus also an aggregate of our model, albeit a more complex one.

identify the impact of individual home-field advantage factors on the number of points scored by each individual team, while previous models can only estimate the impact on the overall difference in points. While estimation using our model could in turn be used to recover the point-differential model's estimations, the reverse is not true. Thus, the information set provided by our model is a strict superset of that provided by the traditional model.<sup>25</sup> We both show this mathematically and illustrate it using a simple empirical example with a subsample of our data in Section 5.2.

Our model also offers a significant advantage by allowing us to separately examine the effects on each team's offense and each team's defense for factors that are specific to either team  $i$  or opponent  $j$ , as discussed in Section 4.2. This increased specificity provides useful insights that cannot be seen when using an aggregated measure, such as point differential, as the dependent variable. For example, it is a common adage in football that defenses tire faster than offenses. Our results would support this adage if we were to find that away-team travel distance positively impacts home-team score. (If defenses tire first, it is defensive fatigue that is the binding constraint on how many points the offense can score.) A point-differential model would not be able to distinguish between whether away-team travel is causing home teams to score more points (supporting the adage) or causing away teams to score fewer points (neither supporting nor contradicting the adage), however.

Our main model, as well as the robustness checks and model extensions in Section 6, also improve upon previous models of collegiate football home-field advantage by exploring factors examined in the home-field advantage literature for other sports leagues.<sup>26</sup> Within college football, Wang et al. (2011) and Fullagar et al. (2019) both include distance and crowd size as covariates, though neither include any measures of familiarity. Additionally, Wang et al. (2011) include a stand-alone home dummy to capture "residual" home-field advantage, while Fullagar et al. (2019) account for time-zone travel. In the NFL, Ehrlich et al. (2021) account for distance and use dummies for full-, reduced-, and zero-crowd games, but they do not include a continuous crowd size or density measure, while Smith et al. (1997) account for distance and time zones but not crowd, familiarity, or other factors. Our main specification advances the literature within college football by introducing various measures of familiarity, including team-stadium, opponent-stadium, and team-opponent familiarity, which loosely resemble some familiarity measures used in other sports (e.g., Boudreaux et al., 2017; Fischer & Haucap, 2021; Losak & Sabel, 2021; McHill & Chinoy, 2020). In the various robustness checks and model extensions we explore in Section 6, we also introduce crowd density (e.g., Böheim et al., 2019; Inan, 2020; Schwartz & Barsky, 1977), weather (utilized, at least to some extent, by Cross & Uhrig, 2020; Fischer & Haucap, 2021; Losak & Sabel, 2021), and head-coach familiarity (perhaps novel to the entire home-field advantage literature, though Fischer & Haucap, 2021, utilize a dummy variable indicating head coaches in their first season with their team) into the college football literature.

<sup>25</sup>For example, suppose our model estimate indicates that a 10,000-person increase in crowd size leads the home team to score 2 fewer points and leads the away team to score 5 fewer points, *ceteris paribus*. If we instead aggregate the dependent variable to the point differential, the resulting estimates would only show that the gap between the home team's points and the away team's points would change by the difference of these two effects, or 3 points, but we would not know how much of that gap's change comes from a change in the home team's score and how much of it comes from a change in the away team's score.

<sup>26</sup>Caudill and Mixon Jr. (2007) only consider the Alabama-Auburn college football rivalry, so many explanatory variables discussed here are not applicable to their study.

## 5 | RESULTS

### 5.1 | Main results

Table 2 shows our main regression results. Model (1), our complete main model specification, includes fixed effects at the team, opponent, and season levels. Model (2) instead uses team-season and opponent-season fixed effects, which necessitates the omission of any covariates that are invariant at the team-season or opponent-season levels. Both models include standard errors clustered at both the team and the opponent levels.<sup>27</sup> Except the constant, row numbers for Table 2 align with the  $\beta$  coefficients' subscript numbers in Equation (1).

Rows (1)–(6) present the effects of our physiological factor: distance traveled (measured in hundreds of miles).<sup>28</sup> Row (2) indicates that, on average, the home team scores one additional point per 888 miles the away team travels, *ceteris paribus*. This supports the common football narrative that defenses tire before offenses: if defenses tire first, then it is defensive fatigue that is the binding constraint on how many points the offense can score; since we observe increased scoring from increased defensive (travel) fatigue, this implies that the defense is indeed the binding constraint, thus indicating that defenses tire first. To put this effect size in perspective, of all 6500 home-away games in our usable sample, 1104 (17.0%) of them involved the away team traveling at least 888 miles. Row (6) provides limited evidence that travel affects neutral-site teams similarly. While the point estimate in Model (1) is not statistically significant, Model (2) indicates that neutral-site teams score one additional point per 975 miles their opponent travels. This comprises 222 (19.4%) of the 1146 neutral-site observations in our usable sample. This is additional evidence that defenses indeed tire faster than offenses.

Away-team travel does not impact their own score, however, as seen in row (3). Similarly, row (5) indicates that there is no statistically significant effect of neutral-site-team travel on their own score. This is akin to a failure to reject the notion that defenses tire before offenses. That is, if defenses tire first, it is defensive fatigue that is the binding constraint on how many points the offense can score. If the offense is only slightly more fatigued than usual (say, because they traveled farther), then the defense is still more fatigued than the offense, and hence the defense's fatigue is still the constraint. In sum, through the mechanism of travel-induced fatigue, all four of our estimates of interest for distance-traveled effects are consistent with the adage that defenses tire before offenses.

Rows (7)–(9) indicate the marginal impact of crowd size (measured in tens of thousands) on points scored by home, away, and neutral-site teams. Following existing literature, we use crowd size as a proxy for loudness in the stadium, though there is potentially a psychological aspect to crowd size, too, which we discuss in more detail in Section 5.4. Because crowds know to stay quiet when the home team is on offense and to “get loud” to interrupt offensive communication when the away team is on offense, intuition suggests that crowd size should negatively impact away-team offenses and have no, or at least a less-negative, effect on home-team offenses. We indeed find that crowd size has no discernible impact on the number of points scored by the home team, while away teams tend to score fewer points when playing in front of

<sup>27</sup>Degrees of freedom are adjusted, using Stata's `reghdfe` command (Correia, 2019), according to whether each set of fixed effects is nested within the model's cluster level(s); see Gormley and Matsa (2014, 2016).

<sup>28</sup>Recall from Section 3 that a handful of games across three teams are designated as “home” games despite not being played at the “home” team’s stadium. The distance traveled by the home team or home opponent is nonzero for only these 42 observations, driving our need to include the respective covariates in rows 1 and 4, largely as controls, despite their lack of meaningful interpretation.

TABLE 2 Main results.

	(1)	(2)
	TeamPts	TeamPts
Home <sub>ig</sub> × TeamDist <sub>ig</sub>	0.980 (0.878)	1.073 (0.658)
Home <sub>ig</sub> × OppDist <sub>jg</sub>	0.113*** (0.037)	0.098*** (0.037)
Away <sub>ig</sub> × TeamDist <sub>ig</sub>	0.005 (0.042)	0.031 (0.044)
Away <sub>ig</sub> × OppDist <sub>jg</sub>	-1.485*** (0.538)	-1.859** (0.847)
NS <sub>ig</sub> × TeamDist <sub>ig</sub>	-0.030 (0.066)	-0.019 (0.060)
NS <sub>ig</sub> × OppDist <sub>jg</sub>	0.089 (0.063)	0.103* (0.056)
Home <sub>ig</sub> × Crowd <sub>g</sub>	-0.127 (0.121)	-0.078 (0.103)
Away <sub>ig</sub> × Crowd <sub>g</sub>	-0.257** (0.109)	-0.133 (0.093)
NS <sub>ig</sub> × Crowd <sub>g</sub>	0.060 (0.236)	0.013 (0.206)
TeamStadiumFamiliarity <sub>ig</sub>	0.179** (0.075)	0.114** (0.056)
OppStadiumFamiliarity <sub>jg</sub>	0.068 (0.065)	0.019 (0.065)
Home <sub>ig</sub>	-0.091 (1.352)	0.535 (1.196)
NS <sub>ig</sub>	-0.330 (1.456)	0.024 (1.327)
TeamWinPerc_PrevSeason <sub>ig</sub>	5.997*** (1.012)	
OppWinPerc_PrevSeason <sub>jg</sub>	-4.780*** (0.967)	
TeamTopConf <sub>ig</sub> × OppBottomConf <sub>jg</sub>	3.793** (1.577)	
TeamBottomConf <sub>ig</sub>	1.080 (1.350)	
TeamBottomConf <sub>ig</sub> × OppBottomConf <sub>jg</sub>	3.463** (1.499)	

TABLE 2 (Continued)

	(1) TeamPts	(2) TeamPts
TeamOppFamiliarity <sub>ijg</sub>	−0.352** (0.149)	−0.261* (0.139)
ConfChampGame <sub>g</sub>	1.675 (1.299)	0.114 (1.242)
BowlGame <sub>g</sub>	0.622 (0.844)	0.405 (0.719)
TopBowlGame <sub>g</sub>	1.216 (1.222)	−0.146 (1.263)
Constant	23.576*** (1.846)	27.371*** (1.061)
Fixed effects	Team, Opp, and Season	Team-Season and Opp-Season
N	14,146	14,146

Note: Standard errors in parentheses. Distances are given in hundreds of miles (100 mi = 160.9 km). Crowd size is given in tens of thousands. All standard errors are clustered at the team and opponent levels.

\* $p < .10$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

larger crowds, at least according to Model (1). The point estimate suggests that an additional 38,875 fans equates to the visiting team scoring an average of 1 fewer point. There are 3063 home-away games (47.1%) with at least 38,875 fans in our usable sample. Interestingly, we find no statistically significant effect for neutral-site teams.

Rows (10) and (11) show the tactical advantage of familiarity with the stadium. Teams tend to score an extra 0.179 points for each additional game they have played in the stadium either earlier in the same season or within the previous two seasons, all else equal. This translates into one additional point per 5.6 games played there. Given that the average number of home games per team per season in our sample is slightly greater than 6, this equates to more than 2 additional points for the home team in each game. However, a team's score is not impacted by their opponent's familiarity with the stadium. This result suggests that the stadium-familiarity effect helps offenses more than it helps defenses. Rows (12) and (13) show that there is no evidence of any "residual" (i.e., leftover) unexplained home-field or neutral-site advantage beyond what we detect through measuring explicit factors.

While the remaining covariates are largely intended to control for other factors, they also offer some interesting interpretations. Each team's winning percentage from the previous season appears to be a good proxy for team-specific skill.<sup>29</sup> Additionally, our controls for whether each team is in a "top" FBS conference are a significant predictor of score.<sup>30</sup> Jointly, the results in rows (16)–(18) indicate that offenses tend to score points similarly regardless of their own

<sup>29</sup>Due to endogeneity concerns regarding the dependent variable (points scored) in season  $s - 1$  impacting that season's winning percentage, which is then a (lagged) independent variable in season  $s$ , we also modify our main specification to instead proxy for team-specific skill using player recruiting ratings. Our results do not change, regardless of whether we use total or mean (current) team-level recruit ratings. This exercise is presented in full in Appendix B.

<sup>30</sup>"Top" conferences include automatically qualifying conferences in the BCS era, Power 5 conferences in the playoff era, and Notre Dame University in both eras. See Section 2 for more information.

conference status, while bottom-conference defenses tend to allow roughly 3.5 to 4 more points per game than top-conference defenses do, regardless of the offense's conference level. Put another way, this suggests that the performance gap between conference levels is on the defensive side of the ball.

Familiarity between the two teams seems to play a role. For each additional game that two teams play against one another either earlier in the same season or at any point in the two previous seasons, points scored per team tends to drop by an average of 0.352 points. This result implies that familiarity helps defenses more than offenses, though further analysis would be required to determine the channels through which this works (such as cross-year returns to scale for game-planning, familiarity/experience with predicting the opponent's in-game coaching decisions, or defensive-player familiarity with opposing offensive schemes). We also explore other measures of team-opponent familiarity as robustness checks in Sections 6.3.2 and 6.3.3.

The importance of the game itself rarely impacts the number of points scored. There is no evidence that teams score a different number of points in conference championships or "lower-tier" bowl games than they do in regular-season games. This suggests that if game importance does elevate play on the field for such games, it does so in a way that is relatively balanced between offense and defense. While we find no evidence that points scored in "top-tier" (BCS/NY6) bowl games are different from those scored in lower-tier ones, as shown in row (22), the sum of the coefficients for rows (21) and (22) in Model (1) indicates that teams tend to score roughly 1.84 more points in top-tier bowl games than in regular-season games, *ceteris paribus*, though the statistical significance of this disappears in Model (2).

Recall that the main distinction for Model (2) is the use of team-season and opponent-season fixed effects rather than separate team, opponent, and season fixed effects. This allows for more precise control in yearly shocks to a given team, though it does not allow for team-specific cross-season effects (such as branding and recruitment trends) or season-specific cross-team effects (such as one might expect from the COVID-19 pandemic in 2020) to be measured explicitly. The results in Model (2) largely corroborate those in our main specification, aside from the three previously noted exceptions. Broadly speaking, our results are extremely robust to various model extensions and tweaks. We discuss many such robustness checks in Section 6.

In sum, we find that many factors have differing impacts on home, away, and neutral-site teams' scores. We construct a back-of-the-envelope calculation, taking each statistically significant coefficient estimate at face value and using medians for covariate values when needed (e.g., for the distance traveled by away teams), using only the factors in Table 2 that relate to home-field advantage (i.e., rows 1–13). This simplistic calculation indicates that home-field advantage increases home teams' scores by 3.2 points—0.5 via opponent distance traveled and 2.7 via team-stadium familiarity—and decreases away teams' scores by 0.9 points—solely via crowd size. Together, this increases the relative home-away score differential by 4.1 total points. This estimate aligns with the literature (Jamieson, 2010).

## 5.2 | Improvement over standard models: Theory and empirics

The following brief theoretical exercise both shows our model advances explicitly and provides a common language that allows for the translation of empirical results from our main model specification (Model (1) in Table 2) into the point-differential realm used in the literature. That is, by demonstrating that our model provides a strict superset of information (relative to

that provided by the standard model), we will, in the process, derive the relationship between any given covariate's marginal impact on points scored and its corresponding marginal impact on the point differential. This in turn allows for an apples-to-apples comparison of our main results from Section 5.1 to the literature. (We then make this comparison in Section 5.3 to demonstrate that our results are consistent with the broader literature).

Consider for illustrative purposes the following pair of simplified models. Without loss of generality or expandability, they include solely the impact of crowd size and a “leftover” home dummy, ignore neutral-site games, and exclude any fixed effects, random effects, or clustering of error terms. Equation (2) represents a simplified version of our main specification, while Equation (3) represents a simplified version of the standard model from the literature (e.g., from Wang et al., 2011):

$$\text{teampts}_{ig} = \alpha_0 + \alpha_1 \cdot \text{home}_{ig} \cdot \text{crowd}_g + \alpha_2 \cdot \text{away}_{ig} \cdot \text{crowd}_g + \alpha_3 \cdot \text{home}_{ig} + \mu_{ig} \quad (2)$$

$$\text{ptdiff}_{ijg} = \gamma_0 + \gamma_1 \cdot \text{home}_{ig} \cdot \text{crowd}_g + \gamma_2 \cdot \text{crowd}_g + \gamma_3 \cdot \text{home}_{ig} + \nu_{ijg} \quad (3)$$

where  $\text{ptdiff}_{ijg}$  is the point differential between team  $i$  and opponent  $j$  in game  $g$ . Formally, Equation (4) represents this relationship between point differential (ptdiff) and points scored (teampts) by team  $i$  and by opponent  $j$  in game  $g$ :

$$\text{ptdiff}_{ijg} = \text{teampts}_{ig} - \text{teampts}_{jg}. \quad (4)$$

In addition to the difference in dependent variables, also notice that the covariates accompanying  $\alpha_2$  in Equation (2) include both a dummy variable for if team  $i$  is the away team in game  $g$  and the crowd size in game  $g$ , analogous to the setup in our main specification, while  $\gamma_2$  in Equation (3) is accompanied solely by the covariate for the crowd size in game  $g$ , as in Wang et al. (2011) specification. As a result, even if the dependent variables were identical across Equations (2) and (3),  $\alpha_2$  and  $\gamma_2$  would still have different interpretations.

Given this setup, our goal of relating the marginal impact in our model to that in the literature standard model is now achieved by relating the corresponding coefficients in Equations (2) and (3). More specifically, the home-field advantage literature refers to  $\gamma_1$  as the home-field advantage of increased crowd size, and thus our goal is to show  $\gamma_1$  as a function of the  $\alpha$  coefficients from Equation (2). In doing so, the resulting mathematical relationship is given by

$$\gamma_1 = 2 \cdot (\alpha_1 - \alpha_2). \quad (5)$$

The full mathematical proof used to arrive at Equation (5) is detailed in Appendix C.

Critically, Equation (5) contains two coefficient estimates that result from estimating Equation (2), yet only one that results from estimating Equation (3). Thus,  $\gamma_1$  is recoverable from our novel model's estimates, yet dimensionality leaves our model's estimates unrecoverable when estimating the standard model from the literature.<sup>31</sup> Thus, our model provides strictly more information.

We can demonstrate the relationship in Equation (5) numerically via a brief empirical example. Doing so also showcases our model's strict increase in information provided.

<sup>31</sup>While other relational equations similar to Equation (5) can also be found, the unidirectional recoverability still holds for the full system of equations.

We estimate Equations (2) and (3), adding only team and opponent fixed effects. For simplicity, we use only data for all regular-season in-conference games from the Pac-12 Conference in 2010.<sup>32</sup> The estimation results can be seen in Models (1) and (2), respectively, in Table 3. (Note that since this example is meant only to demonstrate the coefficient relationships across models, we will ignore statistical significance in our discussions here).

When the crowd increases by 10,000 people, Model (1) indicates that the home team scores 0.224 more points and the away team scores 1.650 more points, all else equal. This means that the home-minus-away point differential has increased by the difference in these estimates, which is –1.426 points. Twice that difference is –2.852, which is the coefficient in the top row for Model (2). This demonstrates how we can use the estimates from Model (1) to recover those from Model (2) using Equation (5). The reverse is not possible, however: looking exclusively at Model (2), we cannot tell by how much the home and away teams' points each individually change; we can only tell by how much the difference in their points changes. This illustrates how Model (1) necessarily provides more information than Model (2) solely by using points scored, rather than the point differential, as the dependent variable, making the former a strictly dominant modeling choice (assuming that the additional information is not undesirable).

Generalizing, Equation (5) represents the relationship between home-field advantage effects in our main specification and those in the existing literature. In Section 5.3, we test the analogous version of the righthand side of Equation (5) for each of our home-field advantage factor results we found in our main results in Table 2, and we compare these test results to the effects found in the existing literature.

TABLE 3 Comparison of models – Different dependent variables.

	(1)	(2)
	TeamPts	PtDiff
Home <sub>ig</sub> × Crowd <sub>g</sub>	0.224 (1.206)	–2.852 (2.119)
Away <sub>ig</sub> × Crowd <sub>g</sub>	1.650 (1.206)	
Crowd <sub>g</sub>		1.426 (1.640)
Home <sub>ig</sub>	11.137 (8.695)	22.274* (11.823)
Constant	20.346** (8.286)	–11.137 (11.267)
N	90	90

Note: Standard errors in parentheses. Crowd size is given in tens of thousands. Model (1) uses points scored by team *i* as the dependent variable, while Model (2) uses the differential between points scored by team *i* and points scored by opponent *j*. Both models include both team and opponent fixed effects. Limited subsample used for simplicity/illustrative purposes.

\**p* < .10; \*\**p* < .05; \*\*\**p* < .01.

<sup>32</sup>This is an arbitrary subsample; any subset of games in the data could be used to illustrate this point.

### 5.3 | Consistency: Comparing our data, model, and results to the literature

In order to establish consistency of our novel model design, we make two sets of comparisons. First, we estimate a commonly utilized model from the literature using our data to see if our data yields similar estimates as in the original literature. This establishes data consistency: our data is not anomalous within the literature. Second, we also compare the results of our main specification to those found in the existing literature. This demonstrates consistency of our model specification within the broader literature.

Since the true marginal impact of each home-field advantage factor varies tremendously across sports, we replicate previous work on home-field advantage within the FBS level of collegiate American football. Wang et al. (2011) examine FBS football, and while their model omits some factors from our main specification, their model is the most similar to ours. We mirror their empirical specification as closely as possible to provide an apples-to-apples comparison of our data to theirs, holding all else—including model specification—constant. If our replication yields similar results, this would suggest that our data is valid.

Table 4 is designed to allow for a quick comparison between Wang et al. (2011) original results and our replication. Models (1)–(3) of Table 4 are taken directly from the basic, limited, and full models from Wang et al. (2011) Table 3, respectively. Models (4)–(6) of Table 4 show the results of our estimation of the same three specifications, respectively, using our data. The dependent variable is the point spread, which equals team  $i$ 's points minus opponent  $j$ 's points within game  $g$ . While ideally we would exactly replicate their empirical setup, we must make two slight modifications: we do not have penalty data and thus must omit that covariate, and we also introduce season-level fixed effects, which Wang et al. (2011) do not require since their data cover solely the 2008 season. Note that we follow their error term structure by also including team-level (but not opponent-level) random effects and by omitting any clustering of standard errors, both of which differ from the standard approach we use throughout the remainder of our analysis (unless otherwise noted).

Our replication results largely mirror those found by Wang et al. (2011), with a few exceptions. For those exceptions, however, our results still align with the broader home-field advantage literature. On the margin, our results for the home indicator coefficients mirror theirs for the basic and limited models, and while our full model's point estimate is closer to zero, it is still positive and statistically significant. Our estimated increase (decrease) in the point differential when the team (opponent) is a member of a top-tier conference is in the ballpark of the increase (decrease) they find in the 2008 season. Their full model estimation does not find that attendance is a significant predictor of point differential from the away team's perspective ( $p = .763$ ), and they do not find a statistically significant difference in this effect for home teams ( $p \geq .9995$ ). However, our full replication model estimates indicate that larger crowds negatively and statistically significantly impact score differential for away teams and that this effect is (economically) significantly larger—to the point that the sign flips—from the home team perspective.<sup>33</sup>

There are two reasons that these few discrepancies are not a concern for data consistency. First, our results still align with much of the broader college-football home-field advantage literature (e.g., Fullagar et al., 2019). Second, these discrepancies are all such that we find statistical

<sup>33</sup>Our inclusion of season-level fixed effects means that the constant in our models is the baseline solely for the 2009 season, and so it is not necessarily comparable to their constant.

TABLE 4 Replicating Wang et al. (2011).

	Wang et al. results			Our results		
	(1) PtDiff	(2) PtDiff	(3) PtDiff	(4) PtDiff	(5) PtDiff	(6) PtDiff
Home <sub>ig</sub>	9.645*** (1.038)	7.735*** (0.822)	5.944*** (1.879)	8.156*** (0.317)	5.475*** (0.240)	1.715*** (0.518)
WinPercDiff <sub>ijg</sub>		49.465*** (1.402)	49.372*** (1.409)		41.786*** (0.367)	41.425*** (0.368)
TeamTopConf <sub>ig</sub>		7.374*** (1.160)	7.594*** (1.243)		12.315*** (0.380)	11.780*** (0.399)
OppTopConf <sub>fg</sub>		-7.392*** (1.082)	-7.320*** (1.170)		-12.338*** (0.341)	-11.450*** (0.368)
PenaltiesDiff <sub>ijg</sub>			0.142 (0.117)			
Crowd <sub>g</sub>			-0.077 (0.256)			-0.470*** (0.076)
Home <sub>ig</sub> × Crowd <sub>g</sub>			0.006 (0.319)			0.756*** (0.091)
AwayDist <sub>g</sub>			-0.159 (0.105)			-0.055* (0.032)
Home <sub>i</sub> × AwayDist <sub>g</sub>			0.338** (0.145)			0.126*** (0.044)
Constant	-4.745*** (1.240)	-3.849*** (0.812)	-2.823** (1.358)	-4.454*** (0.827)	-2.753*** (0.466)	-0.682 (0.560)
Seasons	2008	2008	2008	2009–20	2009–20	2009–20
Season fixed effects	No	No	No	Yes	Yes	Yes
N	1332	1332	1332	16,140	16,140	16,140

Note: Standard errors in parentheses. Distances are given in hundreds of miles. Crowd size is given in tens of thousands. All models include team-level random effects.

\* $p < .10$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

significance for effects where Wang et al. (2011) did not. Given that we have more than 10 times as many observations, and thus have much greater statistical power, it is quite plausible that our estimation can detect effects that Wang et al. (2011) could not. Overall, the similarity between Wang et al. (2011) results and our results when using their models indicates broad similarities between their data and our data. Thus, there is evidence that our data is not anomalous.

Having validated our data, we now compare the results of our main specification—Model (1) in Table 2—to Wang et al. (2011) original results, to the results of our replication of Wang et al. (2011) model, and to the rest of the home-field advantage literature more broadly. However, given the differences in our new empirical setup, we must mathematically translate from

our model's coefficient estimates into the same measures used in the literature. We do so using Equation (5), as discussed in Section 5.2.

First, we consider the home-field advantage gained by the distance traveled by the away team. Given the notation in our main specification—Equation (1)—this effect is given by  $2 \cdot (\beta_2 - \beta_3)$ . The resulting point estimate is 0.216, which is statistically significant at the 5% level ( $p = .034$ ). Thus, by the literature's interpretation, for each hundred miles the away team travels, the home team gains an overall 0.216-point advantage. This translates into a 1-point advantage per 463 miles traveled. To put this in perspective, the away team travels at least 463 miles in 3272 (47.7%) of the 6500 home-away games in our usable sample. Using our replication results in Model (6) of Table 4, row (10) illustrates the analogous effect: the home team gains an overall 0.126-point advantage for each 100 miles that the away team travels ( $p = .004$ ). This translates into a 1-point advantage per 796 miles traveled, which occurs in 1430 (22.0%) relevant games in our usable sample. Though the point estimates are different, these estimates are qualitatively similar.

In the literature, Wang et al. (2011) find a statistically significant point estimate of 0.338 for the same marginal impact of away-team travel on home-field advantage. Distance traveled by the away team has been found to play a similarly pivotal role in home-field advantage in other studies of college football (e.g., Fullagar et al., 2019, though they use indicators for distance traveled by quartile rather than a continuous measure), professional football (Nichols, 2014), and European soccer (Cross & Uhrig, 2020). Thus, our results for distance traveled are in line with the literature.

We can similarly consider the home-field advantage gained by the crowd size, given by  $2 \cdot (\beta_7 - \beta_8)$  from Equation (1). The resulting point estimate is 0.260, though it is not statistically significant ( $p = .169$ ). Thus, while Table 2 indicates that crowd does not impact home-team score and negatively impacts away-team score, the overall result of the crowd's impact is that the home-field advantage is unaffected.<sup>34</sup> Interestingly, in our replication results in Model (6) of Table 4, row (8) indicates that the impact of crowd size on home-field advantage is 0.756 ( $p < .0005$ ). This translates into a 1-point increase in home-field advantage per 13,220 fans in attendance, which occurs in 5714 (87.9%) games within our usable home-away sample. Regardless, given the contradiction between separate and joint effects, we consider this “mild” evidence of a home-field advantage based on crowd size in our main specification.

In the college football literature, Wang et al. (2011) find a point estimate of 0.006 for the same marginal impact of crowd size on home-field advantage, though this point estimate is not statistically significant, while Fullagar et al. (2019) find that an additional 10,000 fans “resulted in a ~6-point [total] advantage” in score differential for the home team. Fischer and Haucap (2021) use COVID-19 mitigation techniques to identify the channel by which crowds in European football impact home-field advantage. They determine that the effect is primarily psychological and that smaller crowds result in lower home-field advantage. Other researchers have investigated the impact of the crowd on other variables; for example, Reade et al. (2022) link smaller crowds to less referee bias in European professional soccer.

<sup>34</sup>We posit that the large standard error for the impact on home-team score is the driving factor behind the lack of significance for the joint test, which is potentially due to underlying heterogeneity in the effect for home teams due to some variable that we are not accounting for—though what this variable might be is an open question for future research.

Using median covariate values for a back-of-the-envelope calculation, we estimate an overall home-field advantage of 4.1 points. Translating this into Wang et al. (2011) model's terms via Equation (5) results in an estimate of 8.2 points, which closely aligns with the 9-point estimate that Wang et al. (2011) find.

## 5.4 | The impact of the COVID-19 pandemic on home-field advantage factors

An additional advantage of our data is the ability to exploit differences in the 2020 season to detect the impact of the COVID-19 pandemic.<sup>35</sup> For each constituent factor of home-field advantage, we separately examine its pre-2020 and 2020 effects to test for whether that factor's marginal effect changed during to the pandemic.<sup>36</sup> We ensure our specification's validity by considering the assumptions it implicitly makes<sup>37</sup> and by running a series of placebo tests,<sup>38</sup> which are both detailed in Appendix B. Overall, both approaches provide confidence in our model specification. Hence, we present our pre-2020 and 2020 effects in Table 5.<sup>39</sup>

The most striking result in Table 5 is the change in the physiological effects of travel for the 2020 season.<sup>40</sup> Rows (2) and (3) indicate that there is no longer a “travel benefit” for the home team (where increased travel distance by the away team leads to more points for the home team) in 2020, but instead a “travel cost” for the away team, where the away team scores fewer points the farther they travel. This seems plausible if all teams experience stress with close-contact game-play during the COVID-19 pandemic.<sup>41</sup> This “cost” is ostensibly higher for

<sup>35</sup>Importantly, we are not testing for whether COVID-19 itself directly caused these differences; rather, we are testing whether the 2020 season, inclusive of all its numerous changes—whether due to COVID-19 itself, to the many varying COVID-19 mitigation policies, to the stress induced by the pandemic more broadly, or to other related changes—saw differences in how the various constituent factors impacted home-field advantage in collegiate American football. We consider these to all be “impacts of the COVID-19 pandemic” in a general sense. This is in line with the home-field advantage literature for other sports (e.g., Cross & Uhrig, 2020; Destefanis et al., 2022; Ehrlich et al., 2021; Fischer & Haucap, 2021).

<sup>36</sup>We generalize by referring to only home and away teams throughout this section because so few neutral-site games were played in 2020. This small sample size means that we cannot draw many meaningful conclusions from the pandemic's effects on neutral-site games. However, we empirically distinguish neutral-site games, mainly for control purposes. For those 36 games—Points scored: mean = 29.2, med = 28, stdev = 12.2, min = 3, max = 56. Crowd size: mean = 7087.9, med = 4744, stdev = 6991.74, min = 0, max = 24,000; exactly 25% of these games had 0 fans in attendance.

<sup>37</sup>Cross and Uhrig (2020) use a similar, though more limited, “decoupling” method to tease out crowd effects in the European soccer leagues that resumed play during 2020. They discuss how the validity of their specification rests on a handful of assumptions, many of which are also applicable here.

<sup>38</sup>Each placebo test treats a different season (referred to as that model's “placebo season”) as if it were the one impacted by the pandemic. That is, each placebo test mirrors the model shown in Table 5 except that it separates a different season—instead of the 2020 season—for the differentiated effects shown in the righthand column. We can then compare the placebo regressions to the results in Table 4 to test the validity of the model (Destefanis et al., 2022; Ferraresi & Gucciardi, 2023).

<sup>39</sup>Estimates for covariates that remain unchanged from our main specification are omitted from the table for brevity; they all matched our main results in sign and statistical significance.

<sup>40</sup>“Home” teams travel in only a handful of observations, and thus the covariates in rows 1 and 4 act more as controls rather than providing estimates with any meaningful interpretation. (See Section 3 for more information).

<sup>41</sup>In fact, NIH researchers publicly support the idea that athletes experienced negative performance consequences related to the pandemic: <https://www.ncbi.nih.gov/pmc/articles/PMC7201218/>.

TABLE 5 Pre-2020 versus 2020 home-field advantage factors.

	(1)	
	TeamPts	
	Pre-2020	2020
Home <sub>ig</sub> × TeamDist <sub>ig</sub>	-0.416 (1.446)	1.730* (0.985)
Home <sub>ig</sub> × OppDist <sub>jg</sub>	0.116*** (0.039)	-0.112 (0.139)
Away <sub>ig</sub> × TeamDist <sub>ig</sub>	0.013 (0.040)	-0.236* (0.132)
Away <sub>ig</sub> × OppDist <sub>jg</sub>	-1.682 (1.606)	-1.880** (0.722)
NS <sub>ig</sub> × TeamDist <sub>ig</sub>	-0.056 (0.066)	0.110 (0.244)
NS <sub>ig</sub> × OppDist <sub>jg</sub>	0.126** (0.063)	-0.468** (0.234)
Home <sub>ig</sub> × Crowd <sub>g</sub>	-0.130 (0.117)	-0.108 (0.976)
Away <sub>ig</sub> × Crowd <sub>g</sub>	-0.226** (0.109)	-0.865 (1.254)
NS <sub>ig</sub> × Crowd <sub>g</sub>	0.128 (0.237)	3.741* (1.909)
TeamStadiumFamiliarity <sub>ig</sub>	0.161** (0.080)	0.317* (0.163)
OppStadiumFamiliarity <sub>jg</sub>	0.082 (0.069)	-0.035 (0.202)
Home <sub>ig</sub>	0.631 (1.471)	-4.577 (3.200)
NS <sub>ig</sub>	-0.408 (1.522)	-3.398 (4.870)
N	14,146	14,146

Note: Standard errors in parentheses. Distances are given in hundreds of miles. Crowd size is given in tens of thousands. This specification modifies our main specification only by separately measuring the pre-2020 and 2020 effects for the covariates shown here. The corresponding coefficient estimates are in the left and right columns, respectively.

\* $p < .10$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

traveling teams since they have additional exposure to the public due to their travels. Future research into these potential source mechanisms is certainly warranted. Regardless of the mechanism, however, these results clearly indicate that the pandemic impacted the effect of distance traveled (in American collegiate football, at least), opposing the literature's standard of not

interacting distance traveled with the pandemic indicator (Cross & Uhrig, 2020; Ehrlich et al., 2021; Fischer & Haucap, 2021).<sup>42</sup>

The crowd potentially creates both noise and a psychological impact that benefit the home team. Prior to 2020, both of these effects were married to crowd size. Researchers often considered these two affects jointly, even when using creative methods to identify home-field advantage (e.g., Caudill & Mixon Jr., 2007, exploit differences in each team's fanbase's attendance to measure the crowd's impact). In 2020, however, crowd size was decoupled from noise due to the NCAA allowing stadium sound systems to pipe in artificial crowd noise (Auerbach, 2020).<sup>43</sup> As a result, crowd size more directly proxies for the psychological impact of fan presence in 2020. Thus, assuming that fan noise does not play a discernible role beyond the sound system's artificial crowd noise, estimates for the impact of crowd size in 2020 should reflect solely the psychological home-field advantage caused by fans, whereas the pre-2020 estimate should capture both the noise disruption and psychological effects.

Indeed, Table 5 indicates that, prior to 2020, larger crowds cause away teams to score fewer points. However, this effect is no longer statistically significant in the 2020 season. This provides evidence that crowds create a home-field advantage through noise rather than through a psychological channel. While interesting, we caution against interpreting this evidence of causality too strongly, given that the pandemic likely impacted the psychological atmosphere of the game in multiple ways. That said, this result aligns with similar work examining other sports. For example, Reade et al. (2022) meta-analysis of European soccer leagues corroborates our results here, and Ehrlich et al. (2021) find no difference between partial-attendance and full-attendance games in the NFL.

There is no evidence of any “residual” (i.e., leftover) home-field or neutral-site effects either prior to or during the 2020 season. Interestingly, the fixed effect for the 2020 season is not statistically significantly different from zero (i.e., from the baseline 2011 season;  $p = .139$ ) or from the fixed effect for any other season in our sample ( $.175 \leq p \leq .340$  for each comparison). This provides some evidence that any difference in home-field advantage in the 2020 season is explained by the individual factors that we include as pre-2020 and 2020 covariates in our model.

A back-of-the-envelope calculation of the total home-field advantage using pre-2020 and 2020 medians for each variable indicates that home-field advantage prior to 2020 increases home teams' scores by 3.1 points—0.5 via opponent distance traveled and 2.6 via team-stadium familiarity—and decreases away-team scores by 0.9 points—all via crowd size. In 2020, home-field advantage increases home teams' scores by 4.8 points—all via team-stadium familiarity—and decreases away-team scores by 1 point—all via away-team distance traveled. Thus, overall home-field advantage jumps 1.8 points, from 4 points in previous seasons to 5.8 points in 2020. Of the many moving parts here, the largest is the 2.2-point increase in the impact of team-stadium familiarity on the home team's score. We interpret this as familiarity with the stadium and setting being particularly beneficial during the pandemic. This is intuitive given how the pandemic amplified the abnormality of what were previously normal situations.

<sup>42</sup>We are not aware of any existing literature that examines how distance traveled by neutral-site teams impacts points scored differently during the 2020 season. For example, during their respective COVID-restricted seasons, the NBA and NHL played all games in neutral-site bubbles (i.e., no team traveled), many studies remove neutral-site games (e.g., Boyle et al., 2021) or under certain circumstances treat them as home-away games (e.g., Brandon & Mehmet, 2021), and the MLB and European soccer leagues did not have neutral-site games.

<sup>43</sup>Stadium sound-system managers were given decibel limits for sound-system noise, independent of crowd size. These decibel limits varied by conference.

## 6 | MODEL EXTENSIONS AND ROBUSTNESS CHECKS

In this section, we explore several model extensions and robustness checks. In addition to those shown here, Appendix B also contains additional robustness checks. The tables in this section show only the relevant new or changed covariates; all other covariates from our main specification are included in the regressions but are omitted from the tables for brevity. All our main results from the previous section remain qualitatively the same throughout this section unless otherwise noted.

Results in this section indicate that our main results are broadly robust to many forms of model misspecification. We also identify a handful of interesting secondary findings regarding variables that impact team scoring.

### 6.1 | Physiological effects

#### 6.1.1 | Time zones

Given evidence from our main results that there is a physiological cost to travel, we additionally account for the direction and number of time zones traveled by each team in Model (1) presented in Table 6 (Charest et al., 2022; Fullagar et al., 2019; McHill & Chinoy, 2020; Smith et al., 1997). In Model (2), we additionally separate these measures by home, away, and neutral-site team indicators.

Changing time zones seems to affect points scored, all else (including distance traveled) held equal, only under certain circumstances. When home, away, and neutral-site games are pooled together, as seen in Model (1), teams traveling west score on average 1.1 fewer points per time zone traveled. By disaggregating this effect, Model (2) indicates that the sole statistically-significant driving factor is away teams, who score 1.2 fewer points per time zone traveled west; home and neutral-site teams do not see a statistically significant difference in points scored when traveling west across time zones.<sup>44</sup> These results are consistent with the literature studying home-field advantage in college football; for example, Fullagar et al. (2019) find “no clear disadvantage” for teams traveling east across at least one time zone, while those traveling at least one time zone westward saw “a ~7.5-point disadvantage,” *ceteris paribus*. We interpret our results as evidence that westward time-zone travel fatigues offenses, but not defenses, for away teams. Given that the literature has shown that mental fatigue, such as jet lag from crossing time zones (Spitzer et al., 1997, 1999; Waterhouse et al., 2003), hinders memory recall (Weingarten & Collop, 2013; Zee et al., 2010), we hypothesize that our results could be explained by the notion that offensive playcalling requires extremely quick memory recall, whereas defensive players largely react, which does not require the same level of recall ability.

Interestingly, Model (2) also indicates that neutral-site teams tend to score an additional 3.3 points for each time zone their opponent travels eastward, *ceteris paribus*. On one level, this stands at odds with the previous result; that is, this is evidence that travel harms defenses, allowing offenses to score more points. However, while the previous result indicated that only offenses were hurt when traveling westward, this result indicates that only defenses are hurt

<sup>44</sup>Note that the only three observations where a “home team” travels across time zones were the University of New Mexico’s only three home games in 2020, which were played in Las Vegas, Nevada, due to New Mexico’s state COVID-19 restrictions banning football games.

**TABLE 6** Robustness Checks: Effects of Travel Across Time Zones (“TZs”)

	(1)		(2)	
	TeamPts		TeamPts	
	West	East	West	East
TeamTZsTraveled <sub>ig</sub>	−1.129*	0.100		
	(0.618)	(0.622)		
OppTZsTraveled <sub>jg</sub>	0.864	−0.219		
	(0.689)	(0.524)		
Home <sub>ig</sub> × TeamTZsTraveled <sub>ig</sub>		6.589		
		(6.255)		
Home <sub>ig</sub> × OppTZsTraveled <sub>jg</sub>		1.011	−0.887	
		(0.719)	(0.542)	
Away <sub>ig</sub> × TeamTZsTraveled <sub>ig</sub>		−1.236*	0.069	
		(0.628)	(0.648)	
Away <sub>ig</sub> × OppTZsTraveled <sub>jg</sub>		2.431		
		(6.497)		
NS <sub>ig</sub> × TeamTZsTraveled <sub>ig</sub>		−0.619	0.244	
		(1.180)	(1.203)	
NS <sub>ig</sub> × OppTZsTraveled <sub>jg</sub>		0.310	3.305**	
		(1.088)	(1.319)	
N	14,146		14,146	

Standard errors in parentheses *Notes:* Both models modify our main specification only by adding the terms shown here; this table omits all other covariates for brevity. For each of the two models presented here, the left column denotes interaction between each row's listed covariate and an indicator denoting westward time-zone travel while the right are those covariates interacted with an eastward time-zone travel indicator. “TZsTraveled” represents the number of time zones traveled (by the team or opponent, as specified). Note that there are only three instances where a home team (away opponent) travels westward across time zones and no instances where a home team (away opponent) travels eastward across time zones.

\* $p < 0.10$ ,

\*\* $p < 0.05$ ,

\*\*\* $p < 0.01$

when traveling eastward. The reason behind this disparity is an open question and warrants future investigation.

## 6.2 | Psychological effects

### 6.2.1 | Crowd size

Crowd size is measured in absolute terms (tens of thousands of people) in our main specification, but crowd size as a percent of stadium capacity—a measure known as crowd density—has also been shown to impact team performance (Ehrlich et al., 2021; Inan, 2020; Schwartz & Barsky, 1977). Models (1) and (2) of Table 7 include both the absolute and density measures, while Models (3) and (4) include solely crowd density. Given that stadium sound-system

TABLE 7 Robustness Checks: Crowd Effects

	(1) TeamPts	(2) TeamPts	(3) TeamPts	(4) TeamPts
Home <sub>ig</sub> × Crowd <sub>g</sub>	-0.367** (0.147)	-0.309** (0.150)		
Away <sub>ig</sub> × Crowd <sub>g</sub>	-0.102 (0.132)	0.010 (0.120)		
NS <sub>ig</sub> × Crowd <sub>g</sub>	0.161 (0.308)	0.266 (0.299)		
Home <sub>ig</sub> × CrowdDensity <sub>g</sub>	0.029** (0.012)	0.036** (0.014)	0.008 (0.010)	0.015 (0.010)
Away <sub>ig</sub> × CrowdDensity <sub>g</sub>	-0.015 (0.010)	-0.023* (0.012)	-0.018** (0.008)	-0.017* (0.009)
NS <sub>ig</sub> × CrowdDensity <sub>g</sub>	-0.007 (0.020)	-0.006 (0.022)	0.008 (0.016)	0.015 (0.018)
Sample Used	Full	Pre2020	Full	Pre2020
N	14,146	13,078	14,146	13,078

Standard errors in parentheses. Notes: Crowd size is given in tens of thousands. CrowdDensity measures crowd size as a percent of crowd capacity. All models modify our main specification only as shown here; this table omits all other covariates for brevity.

\* $p < 0.10$ ,

\*\* $p < 0.05$ ,

\*\*\* $p < 0.01$

managers were allowed to pipe in artificial crowd noise in the 2020 season, Models (2) and (4) omit the 2020 season from the analysis.

Note that for Models (1) and (2), the marginal effect of increasing crowd size, *ceteris paribus*, is not given directly by the coefficient estimates. Since both crowd size (in tens of thousands) and crowd density are included, the marginal impact of increasing the crowd by 10,000 fans while holding the capacity (and all else) constant is given by the sum of the coefficient on absolute crowd size and the product of the coefficient on crowd density times the fraction 100 divided by the capacity. For Models (3) and (4), the coefficient estimate for crowd density indicates the marginal effect of increasing the crowd density by one percentage point, *ceteris paribus*.

All specifications indicate that crowd size negatively impacts away teams while having no consistent discernible effect on home or neutral-site teams.<sup>45</sup> For Models (1) and (2) respectively, at our median capacity of 50,183 seats, increasing crowd size by 10,000 fans leads away teams to score 0.402 ( $p = .019$ ) and 0.446 ( $p = .028$ ) fewer points, all else equal. This corresponds to a 1-point decrease in away-team score for every additional 24,878 fans given the Model (1) estimates. Models (3) and (4) corroborate this result, indicating that a 1-percentage-point increase in crowd size leads to the away team scoring 0.018 and 0.017 fewer points, respectively.

<sup>45</sup>The only exception is that we find that increasing crowd size by 10,000 leads home teams to score 0.399 more points ( $p = .065$ ), *ceteris paribus*, in Model (2).

We can also utilize Models (1) and (2) to estimate the effect of increasing stadium capacity on the margin while holding crowd size constant.<sup>46</sup> To our knowledge, this is completely novel to the literature. These estimates respectively indicate that, for our median crowd density of 81.091% and median capacity of 50,183, increasing capacity by 1 (empty) seat leads to home teams scoring 0.0000471 ( $p = .015$ ) and 0.0000574 ( $p = .011$ ) fewer points, *ceteris paribus*. Interestingly, this indicates that higher-capacity stadiums—when the extra seats are not filled—negatively impact the home team's score.<sup>47</sup> A back-of-the-envelope calculation using Model (1) estimates and median values shows that an increase of 21,211 empty seats leads to the home team scoring 1 fewer point.

Interestingly, there is also mild evidence that adding empty seats also impacts points scored by the visiting team. Model (2) indicates that away teams score an additional 0.0000370 points for each empty seat added to the stadium's capacity ( $p = .056$ ) at median levels, which is 1 additional point per 27,003 empty seats, though this estimate is not statistically significant in Model (1). Neither model yields a statistically significant effect for neutral-site teams.

## 6.3 | Tactical effects

### 6.3.1 | Weather

Table 8 adds weather-related variables to our main specification.<sup>48</sup> All weather covariates other than “Indoors” are also multiplied by an “Outdoors” dummy (which is omitted from the table for brevity). To our knowledge, we are the first to examine the effects of weather on team performance in American football.<sup>49</sup>

Model (1) indicates that teams score fewer points in the rain than when there is no precipitation, all else equal. This effect disappears in Model (2), however. Both models indicate that increased wind speed (measured in MPH) causes teams to score fewer points given either no precipitation or rain. Additionally, in both models, the magnitude of this wind speed effect is

---

<sup>46</sup>This effect is given by the multiplying the estimate of the coefficient on crowd density by the negative of the crowd density value, and then dividing by the capacity.

<sup>47</sup>Of course, this is a partial equilibrium result. More generally, the question of how stadium expansion impacts points scored depends on how it affects the many different intermediary factors that in turn impact points scored. Among myriad other factors, many facets of stadium renovation/expansion, such as total capacity (Coates & Humphreys, 2007; Depken, 2001; El Hodiri & Quirk, 1975; Price & Sen, 2003; Simmons, 2006), number/share of each type of seats (e.g., luxury suites/box seats, bleachers, student sections, season tickets) (Allan & Roy, 2008; Barajas et al., 2019; Chmait et al., 2020; Dobson & Goddard, 1992; Forrest et al., 2005; Kringsstad et al., 2018), quality of any given type of seats (Diehl et al., 2015; Drayer et al., 2012; Fort, 2004), and the “excitement” of a newly renovated (or newly constructed) stadium itself (Coates & Humphreys, 2005; Rishe & Mondello, 2003), have been shown to impact ticket demand (see Schreyer & Ansari, 2022, for a thorough review of the ticket demand literature), and thus by extension would also impact points scored by the home team. Other, yet-unstudied intermediary channels may also exist. For example, stadium expansion may drive increased alumni contributions, which could in turn be used to improve the home team's points scored (e.g., by hiring a more-skilled coaching staff). Thus, our evidence of the negative impact of adding empty seats is just a single piece of the much larger puzzle that athletic administrators must fully examine when considering stadium expansion.

<sup>48</sup>Recall from Section 3 that our weather data represents conditions at kickoff.

<sup>49</sup>There are literatures that examine the impact of weather on performance relative to the point-spread and total-points betting markets (Borghesi, 2007; Coleman, 2017; Kuester & Sanders, 2011; Salaga & Howley III, 2022), on attendance (Falls & Natke, 2016; Popp et al., 2019), and on player aggression (Craig et al., 2016).

TABLE 8 Robustness checks: Tactical effects – Weather.

	(1)	(2)				
	TeamPts		TeamPts			
	Pooled	Pooled	TC <sub>ig</sub>	TH <sub>jg</sub>	TC <sub>jg</sub>	TH <sub>jg</sub>
Indoors <sub>g</sub>	−0.393 (1.070)	1.305 (1.286)				
Rain <sub>g</sub>	−4.732* (2.445)	−2.866 (2.924)				
Snow <sub>g</sub>	−15.697 (13.904)	−14.401 (13.766)				
NoPrecip <sub>g</sub> × Windspeed <sub>g</sub>	−0.064** (0.026)	−0.063** (0.026)				
Rain <sub>g</sub> × Windspeed <sub>g</sub>	−0.247*** (0.083)	−0.247*** (0.084)				
Snow <sub>g</sub> × Windspeed <sub>g</sub>	0.103 (1.171)	0.099 (1.178)				
NoPrecip <sub>g</sub> × Temp <sub>g</sub>	0.017* (0.009)		0.051*** (0.018)	0.058*** (0.017)	0.029* (0.016)	0.036*** (0.012)
Rain <sub>g</sub> × Temp <sub>g</sub>	0.104*** (0.036)		0.098* (0.053)	0.105** (0.046)	0.086* (0.050)	0.108*** (0.041)
Snow <sub>g</sub> × Temp <sub>g</sub>	0.627 (0.611)		0.636 (0.611)			
N	10,178	10,174				

Note: Standard errors in parentheses. All covariates other than “Indoors” are also multiplied by an “Outdoors” dummy (which is omitted for brevity). The baseline constant for each model is outdoors with no precipitation. Windspeed is measured in MPH; temperature is measured in degrees Fahrenheit. In Model (2), temperature by precipitation type is additionally split along two other dimensions, indicated by the column headers. The temperature in game g is “too cold” for team i (labeled as TC<sub>ig</sub>) if it is below team i’s mean home-game temperature; TH<sub>ig</sub> similarly labels when it is “too hot” for team i. Likewise, TC<sub>jg</sub> and TH<sub>jg</sub> respectively indicate when game g’s temperature is below or above opponent j’s mean home-game temperature. All games played in the snow were colder than both teams’ average home-game temperatures. Both models modify our main specification only by adding the terms shown here; this table omits all other covariates for brevity. Weather data includes only 10,178 of the 14,146 games in our usable sample. We see no reason to believe that weather would correlate with whether this data is missing, so we do not suspect that this biases our results. The 4 observations where game temperature equals team or opponent mean temperature are excluded from Model (2). These are the 2 home-away games played in Hawaii for which we have weather data.  
\* $p < .10$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

roughly four times larger when it is raining ( $p = .023$  and  $p = .024$ ). In general, there are no effects of snow, likely due to the extremely small sample size of snow games.

The effect of temperature (measured in degrees Fahrenheit) is separated only by precipitation category in Model (1). Ignoring snow, the results here indicate that teams score more when it is hotter, ceteris paribus, and that this temperature effect is six times larger in the rain than when there is no precipitation ( $p = .013$ ). This result makes intuitive sense: heat causes fatigue, and if (as the common narrative asserts) defenses generally tire faster than offenses, in large

part because they run more, then this heat-induced fatigue would cause defenses to tire even more quickly, allowing offenses to score more points.

In addition to precipitation category, Model (2) also splits the temperature effect along two additional dimensions, indicated by the column headers. The temperature at game  $g$  is “too cold” for team  $i$  (labeled as  $TC_{ig}$ ) if it is below team  $i$ ’s mean home-game temperature;  $TH_{ig}$  similarly labels when it is “too hot” for team  $i$ . Likewise,  $TC_{jg}$  and  $TH_{jg}$  respectively indicate when game  $g$ ’s temperature is below or above opponent  $j$ ’s mean home-game temperature.<sup>50</sup>

Ignoring snow games, temperature positively impacts points scored in Model (2), as it did in Model (1). In the rain, this effect size is indistinguishable across any pairwise comparison. Absent any precipitation, however, temperature has heterogeneous effects. Comparing columns (3) to (4) and (5) to (6) yields no statistical significance for either, indicating that the marginal effect of temperature on score is independent of whether it is too cold or too hot for the defense. However, comparing columns (3) to (5) and (4) to (6), the former is statistically significantly larger within each pair. Thus, the marginal impact of temperature on points scored is greater when it is too cold for the offense (relative to when it is too hot for the offense).

Intuitively, this aligns with our previous results. When it is too cold for the offense, the lower temperature likely interferes with game/player mechanics, such as throwing the ball, resulting in a lower score. As it heats up toward a more typical range, the offense loosens up, and so we see a large marginal impact here—there is no associated increase in fatigue when temperature increases within this range. On the other hand, when it is too hot for the offense, they get increasingly fatigued from the heat, which is not a mechanics-based issue. The offense still continues to score more points as temperatures rise here, though, since their increased fatigue is outpaced by the defense’s increase in fatigue. That is, because defenses tire before offenses, the increased temperatures hurt the defense more, even though the offense is also tiring from the heat. Thus, while the marginal impact of temperature is positive when it is too hot for the offense, it is still smaller in magnitude than the marginal effect when it is too cold for the offense.<sup>51</sup>

### 6.3.2 | Team-opponent familiarity – Conferences and divisions

Beyond our main specification—which accounts for team-opponent familiarity using the total number of times that the given team and opponent have played one another at any point so far in the current season or in the preceding two seasons—the models in Table 9 account for team-opponent familiarity in a variety of other ways, including changes in covariates and in the level at which errors are clustered (as indicated in the table). Regardless of these changes, however, all six models show a statistically significant, negative relationship between points scored and team-opponent familiarity.<sup>52</sup>

Models (1)–(2) use the same measure as—and have statistically-significant point estimates quite similar to—our main specification. They only vary from our main specification in terms of the levels of error clustering. Models (3) and (5) replace the original measure with the same-

<sup>50</sup>All games played in the snow were colder than both teams’ mean home-game temperatures.

<sup>51</sup>Consider the following intuitive example. Kentucky and Virginia’s mean temperatures are closest to the data set’s mean game-time temperature of 65°. Temperature will have a greater impact on each team’s points scored when they play in Chicago, where it is too cold for both teams, than when they play in Miami, where it is too hot for both teams.

<sup>52</sup>All estimates omitted from Table 8 corroborate our main results except that the negative coefficient indicating that away teams score fewer points when their home opponents travel farther is no longer statistically significant in Models (2) and (6). However, since “home” opponents travel in only a handful of observations, as previously discussed, this is not a concern.

TABLE 9 Robustness checks: Tactical effects – Familiarity, by conference/division.

	(1) TeamPts	(2) TeamPts	(3) TeamPts	(4) TeamPts	(5) TeamPts	(6) TeamPts
TeamOppFamiliarity <sub>ijg</sub>	−0.352** (0.141)	−0.352* (0.179)				
SameConf <sub>ijg</sub>			−0.824** (0.376)	−0.514 (0.460)	−0.824* (0.473)	−0.514 (0.544)
SameDiv <sub>ijg</sub>				−0.457 (0.300)		−0.457 (0.370)
Errors clustered by						
Team <sub>i</sub> and Opp <sub>j</sub>	Yes	Yes	Yes	Yes	Yes	Yes
TeamConf <sub>ig</sub> and OppConf <sub>jg</sub>	Yes	Yes	No	No	Yes	Yes
TeamDiv <sub>ig</sub> and OppDiv <sub>jg</sub>	No	Yes	No	No	No	Yes
N	14,146	14,146	14,146	14,146	14,146	14,146

Note: Standard errors in parentheses. All models modify our main specification only as listed here. This table omits some covariates for brevity.

\* $p < .10$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

conference dummy, and each indicates, with statistical significance, that teams score fewer points when they and their opponents are more familiar with each other. In Models (4) and (6), the marginal effect of being in the same conference is not statistically significant. Likewise, the marginal effect of being in the same division—relative to being in the same conference, but not the same division—is not statistically significant (where divisions are a subset of conferences, and each conference has at most two divisions in a given year). However, the sum of both point estimates—which measures the marginal effect of being in the same division relative to not being in the same conference—is again negative and is indeed statistically significant for both models ( $p = .0097$  and  $p = .0638$ ).

### 6.3.3 | Team-opponent familiarity – Head coaching familiarity

Table 10 accounts for team-opponent familiarity along another dimension: head coach familiarity. Model (1) defines familiarity as the total number of times the given team's head coach and opponent's head coach have played against one another (as head coaches) at any point in the preceding two seasons or earlier within the current season. Model (2) includes both our original familiarity measure and the new head coach measure. One limitation of the coaching data we use is that it only accounts for coaching familiarity as head coaches; it does not include familiarity either of current assistant coaches or of current head coaches when they were serving as assistant coaches. Given that essentially every head coach served as an assistant coach beforehand, we can consider our measure's estimate as an underestimate of the true impact of head-coaching familiarity.

Results presented in Table 10 omit covariates not directly relevant to the current context, since results within Models (1) and (2) are consistent with our main results. The newly introduced measure for head-coaching familiarity is not statistically significant in either model,

TABLE 10 Robustness checks: Tactical effects – Familiarity, via coaching.

	(1)	(2)
	TeamPts	TeamPts
TeamOppFamiliarity <sub>ijg</sub>		-0.376** (0.174)
TeamOppHeadCoachFamiliarity <sub>ijg</sub>	-0.129 (0.181)	0.047 (0.209)
N	14,146	14,146

Note: Standard errors in parentheses. Both models modify our main specification only as shown here; this table omits all other covariates for brevity.

\* $p < .10$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

though our original familiarity measure does remain negative and statistically significant even when both measures are included. If there exists a true effect of coaching familiarity, the limitations of our coaching data are causing us to underestimate that true effect by enough that we do not detect it.

To our knowledge, we are the first to examine the effects of coach familiarity on team performance or home-field advantage for any sport. Future work should look to gather more-detailed data on assistant coaches—or at least assistant coaching data for those who later become head coaches—to better tease out any potential effects of coaching familiarity.

## 7 | DISCUSSION AND CONCLUSION

Using data for all 131 Division I FBS collegiate American football teams from the 2009 to 2020 seasons, we estimate the impact of specific physiological, psychological, and tactical factors to better explain the mechanisms of home-field advantage. In doing so, we add to the literature in a number of ways. First, relative to the NFL and to other sports, such as soccer, college football is under-studied in the literature. In correcting this, we use a significantly richer data set—in terms of both variables of interest and number of seasons—than existing college football studies. Second, we construct a novel empirical model that offers several important advantages over the conventional model(s) in the literature. Third, we include neutral-site games. While much (though not all) of the literature omits such games, their particularly large prevalence within college football lets our data set offer a unique look at the “neutral-site (dis)advantage” of each constituent factor. Fourth, the COVID-19 pandemic and its myriad associated policies during the 2020 season create significantly greater exogenous variation for many key variables, such as crowd size, which allows us to more precisely estimate these variables’ effects. Last, in considering the validity of the previous point, we also test whether the home-field advantage provided by each individual factor changes during the 2020 season. This analysis also leads to several interesting corollary findings.

Our empirical specification is novel in that, by using team points scored as the dependent variable, it allows us to identify each factor’s effect on each team’s offense and each team’s defense individually. Furthermore, the information set provided by our model is a strict super-set of that provided by the traditional model: estimates of the conventional point-differential

model are recoverable from our model, while the reverse is not true. We provide theoretical and empirical comparisons to illustrate this difference. Given these advantages, we propose that our points-scored model be adopted as the new standard across the home-field advantage literature.

Broadly speaking, our results are intuitive, and they provide empirical evidence for many home-field-advantage narratives commonly espoused by coaches, players, and sportscasters. We find robust evidence to support the idea of travel fatigue: when away teams travel a greater distance, home teams score more points, and we also find limited similar evidence regarding time-zone travel. These results lend support to the notion that defenses tire faster than offenses: the physiological impact of travel causes traveling-team defenses to tire more quickly, thus allowing opposing offenses to score more points. We find additional evidence for this notion when we extend our model to account for weather: increased temperatures exhaust defenses even more quickly, allowing offenses to score more points, and this effect is larger when the temperature is greater than the defense's mean home-game temperature.

Another common narrative is that crowd noise adversely affects offenses by hampering communication (whereas defenders are more “reactive” and thus require less on-field communication, especially at the collegiate level). For typical home-away games, home crowds and stadium sound-system managers know to keep the volume down when their team is on offense and to “get loud” when their opponent is on offense. Thus, we would expect to see crowd size—which intuitively proxies for loudness—negatively impact away-team scores, while having no (or at least less of an) effect on home and neutral-site teams. As expected, we find that crowd size does not impact home or neutral-team scores, while away teams tend to score 1 fewer point per additional 38,875 fans in attendance, *ceteris paribus*.

More closely considering the impact of the crowd, we are able to estimate the impact of increasing stadium capacity while holding attendance constant. Using median levels, we find that home teams score 1 fewer point per extra 21,211 additional empty seats in the stadium, holding all else (including crowd size) constant. We find mixed evidence that increased capacity leads to away teams scoring more, while we find no evidence that it impacts neutral-site teams' scores. This serves as a caution to universities that ticket demand is a crucial factor when considering stadium expansion.

From a tactical perspective, teams score more points when they are more familiar (within recent seasons) with the stadium, *ceteris paribus*. Team score is not impacted by their opponent's familiarity with the stadium, however. This suggests that stadium familiarity helps offenses more than defenses. Conversely, team-opponent familiarity over the same timeframe leads to lower scores, which suggests that repeated exposure to the same opponent helps defenses more than offenses. This result is robust to a variety of specifications regarding familiarity. Further investigation is needed to clarify whether this is due to cross-year returns to scale for game-planning, experience with predicting the opponent's in-game coaching decisions, defensive player familiarity with opposing offensive schemes, or other channels.

Prior to the 2020 season, intuition suggests that crowd noise should correlate closely with crowd size. During the pandemic, however, many locations imposed crowd size restrictions, which varied both geographically (by city, county, and state) and temporally, fluctuating throughout the season for many locations. Stadium sound systems were permitted to fill the void of crowd noise to some extent, though conferences set their own sound-system decibel limits. Thus, a decoupling of crowd size and noise in the stadium could plausibly lead to crowds having a different impact—or even no impact at all—in 2020. Our results show that while away teams score fewer points when playing in front of larger crowds prior to 2020, this effect disappears during the 2020 season. While this indeed supports the notion that crowds impact games

via noise rather than by creating a psychological advantage, we caution that this causal interpretation should be taken with a grain of salt, given the many ways in which the pandemic changed the psychological atmosphere of games. Future work should look to include data on actual noise levels during games to gain a more robust vantage of the channels through which crowds impact home-field advantage.

Our investigation into the various factors influencing the overall home-field advantage in college football offers many lessons. Where our data overlaps with conventional narratives discussed by commentators and other sports personnel, we find support for these widely-held beliefs. Front offices and university officials can also learn from our results: while crowd size certainly breeds an advantage, be wary of expanding too much and adding seats that sit unused. League administrators—who tend to value competitive balance—of all sports may also be interested to learn that crowd and stadium size can impact home-field advantage. Perhaps they should consider such factors as a team's stadium capacity and average attendance, rather than solely balancing home and away games, when creating league schedules. Future work in the academic literature on competitive balance could benefit from considering these points, too.

For home-field advantage researchers, however, our examination offers important blueprints for future work. Namely, utilizing the number of points scored by each team creates more valuable insights than aggregating to point differentials, separately investigating home-field advantage by factor is critical, and the factors' impacts can differ from season to season, as we see for some of the constituent factors during the COVID-19 pandemic. This blueprint is largely transferable to other sports, even though the various effect sizes themselves likely differ.

## ACKNOWLEDGMENTS

The authors would like to thank Moiz Bhai, Justin Roush, Daniel Vielhaber, Larry Vielhaber, and anonymous reviewers for their helpful comments and suggestions.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## REFERENCES

- Ahammer, A., Halla, M. & Lackner, M. (2023) Mass gatherings contributed to early COVID-19 mortality: Evidence from US sports. Discussion paper, Department of Economics, Johannes Kepler University Linz, Austria.
- Allan, G. & Roy, G. (2008) Does television crowd out spectators? New evidence from the Scottish Premier League. *Journal of Sports Economics*, 9(6), 592–605.
- Athey, S., Bayati, M., Doudchenko, N., Imbens, G. & Khosravi, K. (2021) Matrix completion methods for causal panel data models. *Journal of the American Statistical Association*, 116(536), 1716–1730.
- Auerbach, N. (2020) How college football's artificial crowd noise DJs are juicing up empty stadiums. <https://theathletic.com/2151075/2020/10/21/artificial-crowd-noise-college-football-stadiums/> [Accessed 28th May 2021].
- Baker, A.C., Larcker, D.F. & Wang, C.C. (2022) How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2), 370–395.
- Barajas, A., Shakina, E. & Gasparetto, T. (2019) At the stadium or at home: the effect of broadcasting matches. *Sport, Business and Management: An International Journal*, 9(5), 495–505.
- Bender, B. (2020) Big Ten football is back: everything to know on and off field in 2020. <https://www.sportingnews.com/us/ncaa-football/news/big-ten-football-returns-ohio-statebuckeyes-covid-19-impact-college-football-playoff/1bd90ine6ex4p1hsj1707ztiet> [Accessed 7th June 2021].
- Böheim, R., Grübl, D. & Lackner, M. (2019) Choking under pressure—evidence of the causal effect of audience size on performance. *Journal of Economic Behavior & Organization*, 168, 76–93.

- Bokat-Lindell, S. (2020) Should College Athletes Be Allowed to Get Paid? With a new law, California says yes. The N.C.A.A. says it's unconstitutional. <https://www.nytimes.com/2019/10/01/opinion/california-student-athletes-paid.html> [Accessed 9th April 2021].
- Borghesi, R. (2007) The home team weather advantage and biases in the NFL betting market. *Journal of Economics and Business*, 59(4), 340–354.
- Borusyak, K., Jaravel, X. & Spiess, J. (2021) Revisiting event study designs: robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.
- Boudreaux, C.J., Sanders, S.D. & Walia, B. (2017) A natural experiment to determine the crowd effect upon home court advantage. *Journal of Sports Economics*, 18(7), 737–749.
- Boyle, E., Mangan, S. & McDermott, L. (2021) The impact of match location and crowd attendance during COVID-19 on technical performance in elite Gaelic football. *Science Performance and Science Reports*, 1.
- Bozich, R. (2020) Vegas Speaks: Will home-field advantage be home-field advantage in COVID-19 college football? <https://www.wdrb.com/sports/bozich-will-home-fieldadvantage-be-home-field-advantage-in-covid-19-college-football/articleb67c57ce-f2b8-11eaacd0-1f4835e410ca.html> [Accessed 17th Jan 2021].
- Brandon, J. & Mehmet, D. (2021) Cold weather teams in the National Football League and home-field advantage. *Proceedings of International Mathematical Sciences*, 3(1), 10–24.
- Bryson, A., Dolton, P., Reade, J.J., Schreyer, D. & Singleton, C. (2021) Causal effects of an absent crowd on performances and refereeing decisions during Covid-19. *Economics Letters*, 198, 109664.
- Callaway, B. & Sant'Anna, P.H. (2021) Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- Carmichael, F. & Thomas, D. (2005) Home-field effect and team performance: evidence from English premier-ship football. *Journal of Sports Economics*, 6(3), 264–281.
- Caudill, S.B. & Mixon, F.G., Jr. (2007) Stadium size, ticket allotments and home field advantage in college foot- ball. *The Social Science Journal*, 44(4), 751–759.
- Charest, J., Cook, J.D., Bender, A.M., Walch, O., Grandner, M.A. & Samuels, C.H. (2022) Associations between time zone changes, travel distance and performance: a retrospective analysis of 2013–2020 National Hockey League Data. *Journal of Science and Medicine in Sport*, 25(12), 1008–1016.
- Chmait, N., Robertson, S., Westerbeek, H., Eime, R., Sellitto, C. & Reid, M. (2020) Tennis superstars: The relationship between star status and demand for tickets. *Sport Management Review*, 23(2), 330–347.
- Coates, D. & Humphreys, B.R. (2005) Novelty effects of new facilities on attendance at professional sporting events. *Contemporary Economic Policy*, 23(3), 436–455.
- Coates, D. & Humphreys, B.R. (2007) Ticket prices, concessions and attendance at professional sporting events. *International Journal of Sport Finance*, 2(3), 161.
- Coleman, B.J. (2017) Team travel effects and the college football betting market. *Journal of Sports Economics*, 18(4), 388–425.
- Connelly, B. (2020) A better future for college football players? Here's how to make it happen. <https://www.espn.com/college-football/story//id/29578928/a-better-future-collegefootball-players-here-how-make-happen> [Accessed 9th April 2021].
- Correia, S. (2019) REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. *Statistical Software Components*, S457874.
- Craig, C., Overbeek, R.W., Condon, M.V. & Rinaldo, S.B. (2016) A relationship between temperature and aggres-sion in NFL football penalties. *Journal of Sport and Health Science*, 5(2), 205–210.
- Cross, J. & Uhrig, R. (2020) Do fans impact sports outcomes? A COVID-19 natural experiment. *Journal of Sports Economics*, 24, 15270025221100204.
- Dave, D., McNichols, D. & Sabia, J.J. (2021) The contagion externality of a superspreading event: The Sturgis Motorcycle Rally and COVID-19. *Southern Economic Journal*, 87(3), 769–807.
- Depken, C.A. (2001) Fan loyalty in professional sports: an extension to the National Football League. *Journal of Sports Economics*, 2(3), 275–284.
- Destefanis, S., Addesa, F. & Rossi, G. (2022) The impact of COVID-19 on home advantage: a conditional order-m analysis of football clubs' efficiency in the top-5 European leagues. *Applied Economics*, 54(58), 6639–6655.
- Diehl, M.A., Maxcy, J.G. & Drayer, J. (2015) Price elasticity of demand in the secondary market: evidence from the National Football League. *Journal of Sports Economics*, 16(6), 557–575.

- Dobson, S.M. & Goddard, J.A. (1992) The demand for standing and seated viewing accommodation in the English Football League. *Applied Economics*, 24(10), 1155–1163.
- Dodd, D. (2020) With near-empty stands, college football's home-field advantage on pace to be worst in 15 years. <https://www.cbssports.com/college-football/news/with-neareempty-stands-college-footballs-home-field-advantage-on-pace-to-be-worst-in-15-years/> [Accessed 15th Dec 2020].
- Drayer, J., Rascher, D.A. & McEvoy, C.D. (2012) An examination of underlying consumer demand and sport pricing using secondary market data. *Sport Management Review*, 15(4), 448–460.
- Ehrlich, J., Potter, J. & Sanders, S. (2021) The effect of attendance on home field advantage in the National Football League: a natural experiment. *Sport Management*, 58.
- El Hodiri, M. & Quirk, J. (1975) Stadium capacities and attendance in professional sports. In: *Management science applications to leisure-time operations*. Amsterdam: North-Holland Publishing Company, pp. 246–262.
- Falls, G.A. & Natke, P.A. (2016) College football attendance: a panel study of the Football Championship Subdivision. *Managerial and Decision Economics*, 37(8), 530–540.
- Ferraresi, M. & Gucciardi, G. (2023) Team performance and the perception of being observed: experimental evidence from top-level professional football. *German Economic Review*, 24(1), 1–31.
- Fischer, K. & Haucap, J. (2021) Does crowd support drive the home advantage in professional football? Evidence from German ghost games during the COVID-19 pandemic. *Journal of Sports Economics*, 22(8), 982–1008.
- Forrest, D., Simmons, R. & Buraimo, B. (2005) Outcome uncertainty and the couch potato audience. *Scottish Journal of Political Economy*, 52(4), 641–661.
- Fort, R. (2004) Inelastic sports pricing. *Managerial and Decision Economics*, 25(2), 87–94.
- Fullagar, H.H., Delaney, J., Duffield, R. & Murray, A. (2019) Factors influencing home advantage in American collegiate football. *Science and Medicine in Football*, 3(2), 163–168.
- Garthwaite, C., Keener, J., Notowidigdo, M.J. & Ozminkowski, N.F. (2020) Who profits from amateurism? Rent-sharing in modern college sports. Discussion paper, National Bureau of Economic Research.
- Goodman-Bacon, A. (2021) Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- Gormley, T.A. & Matsa, D.A. (2014) Common errors: how to (and not to) control for unobserved heterogeneity. *The Review of Financial Studies*, 27(2), 617–661.
- Gormley, T.A. & Matsa, D.A. (2016) Playing it safe? Managerial preferences, risk, and agency conflicts. *Journal of Financial Economics*, 122(3), 431–455.
- Heim, M. (2018) NCAA recognizes UCF's national championship in addition to Alabama's. <https://www.al.com/sports/2018/08/ncaarecognizesucfsnational.html> [Accessed 18th Jan 2023].
- Hosick, M.B. (2021) NCAA adopts interim name, image and likeness policy. <https://www.ncaa.org/news/2021/6/30/ncaa-adopts-interim-name-image-and-likenesspolicy.aspx> [Accessed 19th Jan 2023].
- Humphreys, B.R., Wagner, G.A., Whitehead, J.C. & Wicker, P. (2023) Willingness to pay for policies to reduce health risks from COVID-19: evidence from U.S. professional sports. *Health Economics*, 32, 218–231.
- Imai, K. & Kim, I.S. (2021) On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis*, 29(3), 405–415.
- Inan, T. (2020) The effect of crowd support on home-field advantage: evidence from European football. *Annals of Applied Sport Science*.
- Jamieson, J.P. (2010) The home field advantage in athletics: a meta-analysis. *Journal of Applied Social Psychology*, 40(7), 1819–1848.
- Knoester, C. & Ridpath, B.D. (2020) Should college athletes be allowed to be paid? A public opinion analysis. *Sociology of Sport Journal*, 38(4), 399–411.
- Kringstad, M., Solberg, H.A. & Jakobsen, T.G. (2018) Does live broadcasting reduce stadium attendance? The case of Norwegian football. *Sport, Business and Management: An International Journal*, 8(1), 67–81.
- Kuester, D.D. & Sanders, S. (2011) Regional information and market efficiency: the case of spread betting in United States college football. *Journal of Economics and Finance*, 35, 116–122.
- Losak, J.M. & Sabel, J. (2021) Baseball home field advantage without fans in the stands. *International Journal of Sport Finance*, 16(3), 148–162.
- McHill, A.W. & Chinoy, E.D. (2020) Utilizing the National Basketball Association's COVID-19 restart "bubble" to uncover the impact of travel and circadian disruption on athletic performance. *Scientific Reports*, 10(1), 1–7.

- Moore, J.C. & Brylinsky, J.A. (1993) Spectator effect on team performance in college basketball. *Journal of Sport Behavior*, 16(2), 77.
- NCAA vs. Alston, S. (2020) National Collegiate Athletic Association v. Shawne Alston, et al. United States Court of Appeals for the Ninth Circuit.
- Nevill, A.M. & Holder, R.L. (1999) Home advantage in sport. *Sports Medicine*, 28(4), 221–236.
- Nichols, M.W. (2014) The impact of visiting team travel on game outcome and biases in NFL betting markets. *Journal of Sports Economics*, 15(1), 78–96.
- O'Bannon vs. NCAA, e. (2014) Edward O'Bannon, et al. v. National Collegiate Athletic Association; Electronic Arts Inc.; and Collegiate Licensing Company. United States District Court for the Northern District of California.
- Parker, T. & Stone, M. (2011) Crack baby athletic association. *South Park*, 15(5).
- Pollard, R. (1986) Home advantage in soccer: a retrospective analysis. *Journal of Sports Sciences*, 4(3), 237–248.
- Pollard, R. (2002) Evidence of a reduced home advantage when a team moves to a new stadium. *Journal of Sports Sciences*, 20(12), 969–973.
- Pollard, R. (2008) Home advantage in football: a current review of an unsolved puzzle. *The Open Sports Sciences Journal*, 1(1), 12–14.
- Popp, N., Simmons, J., Watanabe, N. & Shapiro, S. (2019) Why didn't they show up? An examination of factors influencing no show rates at college football games. In Sport Marketing Association Conference, Illinois River, Chicago, IL. vol. 5, p. 1571334127042. Available at: <https://static1.squarespace.com/static/54358491e4b0e09faedc4a1b>
- Price, D.I. & Sen, K.C. (2003) The demand for game day attendance in college football: an analysis of the 1997 Division 1-A season. *Managerial and Decision Economics*, 24(1), 35–46.
- Reade, J.J., Schreyer, D. & Singleton, C. (2022) Eliminating supportive crowds reduces referee bias. *Economic Inquiry*, 60(3), 1416–1436.
- Rishe, P.J. & Mondello, M.J. (2003) Ticket price determination in the National Football League: a quantitative approach. *Sport Marketing Quarterly*, 12(2), 72–79.
- Rosenthal, G.J. (2017) College play and the FLSA: why student-athletes should be classified as employees under the fair labor standards act. *The Hofstra Labor and Employment Law Journal*, 35, 133.
- Salaga, S. & Howley, W., III. (2022) The impact of weather on betting outcomes and market behaviour in the NCAA football totals market. *Applied Economics Letters*, 1–4.
- Samuels, D. (2021) Nick Saban came with an incredibly strong NIL recruiting pitch to Texas HS coaches: An Alabama player is on the verge of becoming one of the first NILera players to make \$1 million, and Saban dropped that nugget in the perfect recruiting venue. <https://footballscoop.com/news/nick-saban-came-with-an-incredible-nil-recruitingpitch-to-texas-hs-coaches> [Accessed 19th Jan 2023].
- Sanderson, A.R. & Siegfried, J.J. (2015) The case for paying college athletes. *Journal of Economic Perspectives*, 29(1), 115–138.
- Schreyer, D. & Ansari, P. (2022) Stadium attendance demand research: a scoping review. *Journal of Sports Economics*, 23(6), 749–788.
- Schwartz, B. & Barsky, S.F. (1977) The home advantage. *Social Forces*, 55(3), 641–661.
- SDS Staff, T. (2020) Florida AD shares SEC's crowd noise policy for 2020. <https://www.saturdaydownsouth.com/florida-football/florida-ad-shares-secs-crowdnoise-policy-for-2020/> [Accessed 7th June 2020].
- Simmons, R. (2006) Chapter 8: The demand for spectator sports. In: Andreff, W. & Szymanski, S. (Eds.) *Handbook on the economics of sport*. Cheltenham, UK: Edward Elgar Publishing.
- Smith, R.S., Guilleminault, C. & Efron, B. (1997) Orcadian rhythms and enhanced athletic performance in the National Football League. *Sleep*, 20(5), 362–365.
- Spitzer, R., Terman, M., Terman, J. & Williams, J. (1997) Columbia jet lag scale. In: *Biometrics research*. New York: NY State Psychiatric Institute.
- Spitzer, R.L., Terman, M., Williams, J.B., Terman, J.S., Malt, U.F., Singer, F. et al. (1999) Jet lag: clinical features, validation of a new syndrome-specific scale, and lack of response to melatonin in a randomized, double-blind trial. *American Journal of Psychiatry*, 156(9), 1392–1396.
- Stefani, R. & Clarke, S. (1992) Predictions and home advantage for Australian rules football. *Journal of Applied Statistics*, 19(2), 251–261.

- The Economist, M. (2020) Graphic detail: empty stadiums have shrunk football teams' home advantage. <https://www.economist.com/graphic-detail/2020/07/25/emptystadiums-have-shrunk-football-teams-home-advantage> [Accessed 6th Oct 2021].
- Wang, W., Johnston, R. & Jones, K. (2011) Home advantage in american college football games: a multilevel modelling approach. *Journal of Quantitative Analysis in Sports*, 7(3), 1-23.
- Waterhouse, J., Nevill, A., Edwards, B., Godfrey, R. & Reilly, T. (2003) The relationship between assessments of jet lag and some of its symptoms. *Chronobiology International*, 20(6), 1061-1073.
- Weingarten, J.A. & Collop, N.A. (2013) Air travel: effects of sleep deprivation and jet lag. *Chest*, 144(4), 1394-1401.
- Wilbon, M. (2011) College athletes deserve to be paid. <https://www.espn.com/collegesports/story//id/6778847/college-athletes-deserve-paid> [Accessed 9th April 2021].
- Wing, C., Simon, D.H. & Carlin, P. (2021) Effects of large gatherings on the COVID-19 epidemic: evidence from professional and college sports. *Economics & Human Biology*, 43, 101033.
- Wunderlich, F., Weigelt, M., Rein, R. & Memmert, D. (2021) How does spectator presence affect football? Home advantage remains in European top-class football matches played without spectators during the COVID-19 pandemic. *PLoS One*, 16(3), e0248590.
- Zee, P.C., Wang-Weigand, S., Wright, K.P., Jr., Peng, X. & Roth, T. (2010) Effects of ramelteon on insomnia symptoms induced by rapid, eastward travel. *Sleep Medicine*, 11(6), 525-533.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** McMahon, M. J., & Quintanar, S. M. (2024). Separately measuring home-field advantage for offenses and defenses: A panel-data study of constituent channels within collegiate American football. *Southern Economic Journal*, 90(4), 1060-1098. <https://doi.org/10.1002/soej.12682>

Copyright of Southern Economic Journal is the property of Wiley-Blackwell and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.