The Willingness to Pay for a Cooler Day: Evidence from 50 Years of Major League Baseball Games

Kevin Kuruc[*] Melissa LoPalo[†] Sean O'Connor[‡]

Abstract

The climate-economy literature has documented adverse effects of extreme temperatures on well-being through mechanisms such as mortality, productivity, and conflict. Impacts due simply to discomfort are less well understood. This paper investigates individuals' valuations of weather using a revealed preference approach. Using 50 years of data, we first quantify the decline in attendance at Major League Baseball games on hot days. Leveraging this finding coupled with historically-informed assumptions on the supply curve, we infer a monetized estimate of the disutility of heat. We estimate a $1 utility loss per hour of exposure to high temperatures, implying non-trivial aggregate welfare effects.

JEL classification: Q51; Q54
Keywords: climate amenities, environmental valuation, climate change

[*]University of Texas at Austin; University of Oklahoma. Email: kevinkuruc@utexas.edu

[†]Montana State University. Email: melissa.lopalo@montana.edu

[‡]Federal Housing Finance Agency. This article was written by Sean T. O'Connor in his private capacity. No official support or endorsement by the Federal Housing Finance Agency is intended or should be inferred. Email: seantjoconnor@gmail.com

I. Introduction

A large literature in economics documents the impacts of extreme temperature on well-being: more frequent episodes of high heat are expected to harm human health (e.g. Deschenes and Moretti (2009); Barreca (2012)), make workers less productive (e.g. Adhvaryu et al. (2019); Zhang et al. (2018)), and lower the overall GDP of economies (e.g. Burke et al. (2015); Hsiang (2010)), for example. These impacts entail large losses of individual well-being. One impact that is relatively less well documented is loss of utility due to the simple fact that extreme temperatures are unpleasant. While utility losses from this discomfort may be relatively small at the individual level, on the aggregate level they may add up to large losses in welfare due to their universality.

In this paper, we identify the causal effects of weather on individuals' engagement in outdoor activities in a context where we can isolate and infer a monetized value for the disutility of heat. Major League Baseball (MLB) games provide an ideal setting: attendance has been scrupulously documented for over a century, creating a rich dataset of over 80,000 games in our sample period with substantial variation in game-day weather and local climate. We use variation in daily average game day temperature to identify the impacts of weather on attendance at games between 1950 and 2000, controlling for a variety of observable characteristics of the game, stadium-specific seasonality in attendance, and time fixed effects. The results indicate significant effects of both very hot and very cold days, with attendance declining by 14% on days over 90 degrees.

We next calculate the decline in ticket prices that would reverse the decline in attendance on hot days to estimate the change in willingness to pay for baseball. To do this, we exploit the 20th century fixed pricing behavior of MLB teams: teams have broadly only begun adjusting ticket prices for a given game according to expected or realized demand over the last couple of decades. Since fixed prices imply a horizontal supply curve, the decline in attendance on hot days represents a leftward demand shift. Given the well-documented fact that MLB teams have historically priced tickets on the inelastic region of demand (Krautmann and Berri, 2007), we take as an upper-bound that the (local) price elasticity of demand is 1.0. The unit elasticity then implies that a 14% decrease in price would offset this decrease in attendance.[1] The average of inflation-adjusted MLB ticket prices in our sample period is approximately \$18,[2] implying that the marginal consumer's willingness to pay for a ticket drops by roughly \$2.52 on a day over 90 degrees. This exercise provides information about individuals' preferences over weather, which encompass both the simple disutility from the discomfort of heat, as well as any concern the individuals may have about health consequences.

---

1 To the extent that we overestimate the price elasticity of demand, we underestimate the price decline required to compensate for the high heat.

2 Source: https://eh.net/encyclopedia/the-economic-history-of-major-league-baseball/. This is in 2020 dollars.

We complement this first approach with an alternative method for estimating how willingness to pay for baseball varies with the weather. During the 2021 season, we scraped game-level ticket price data from Seatgeek, a secondary market for tickets. We use a similar empirical strategy to estimate the impacts of game-day weather on the prices of ticket listings on Seatgeek, finding that prices fall by 9.4% on days over 80 degrees, which translates to approximately a $1.69 decline in willingness to pay for tickets. The similarity in the magnitude of these effects helps to validate our primary results.

In general, the availability of alternative outdoor activities creates a wedge between our estimate of the price decline necessary to reverse the attendance decline on hot days and the disutility of heat. We show that as long as individuals are choosing between baseball games, indoor activities, and other outdoor activities that become (absolutely) less enjoyable in the heat, the willingness to pay estimate will be weakly smaller (in absolute value) than the disutility of heat. An exception would occur if individuals have the option to engage in an outdoor activity that becomes more enjoyable in the heat relative to mild weather. In this case, the decline in willingness to pay for baseball on hot days would be partially attributable to the availability of a better outdoor option. However, we show that attendance increases at baseball games in covered stadiums on hot days, indicating that declines in baseball attendance at outdoor stadiums are unlikely to be driven by an increasing desirability of alternative outdoor activities.

In order to aggregate these effects to provide a sense of the order of magnitude of monetized welfare loss when heat waves occur, we first note that baseball games last 161 minutes on average; our attendance estimates then imply approximately $0.94 of lost value per hour of outdoor activity. Individuals in the U.S. spend approximately 30 minutes outside on hot days; we apply our disutility estimate only to these 30 minutes, conservatively not accounting for utility lost due to individuals switching to less preferred indoor activities in the heat. If each individual suffers welfare losses of $0.47 per day, we find that under a business as usual climate change scenario, annual total losses from very hot days will be on the order of $2.1 billion by 2080-2090.

This paper contributes to the prior literature in several ways. First, we contribute to the literature on valuation of non-market climate amenities by exploiting unique aspects of the market for baseball tickets. Our paper is the first to provide an estimate of the valuation of mild weather based on consumer choice in a market setting, in a sufficient statistics framework that does not require strong structural assumptions. Settings where it is possible to use well-identified estimates of behavioral change to monetize revealed preferences are rare, in part because climate is closely tied to consumption of non-market goods/leisure which are themselves difficult to value. Recent work by Chan and Wichman (2022) derives a valuation method for the effects of climate on consumer surplus for any given activity, based on changes in participation in that activity and baseline consumer surplus. Our approach estimates the direct effect of temperature on utility through well-documented market behavior rather than relying on travel cost methods to obtain consumer surplus. A second advantage of the setting here is that the market for MLB tickets is better documented and studied than nearly any other form of outdoor recreation, allowing for more credible empirical evidence of behavior change and a more robust literature from which to draw external assumptions.

More generally, previous literature has aimed to identify individuals' monetary valuation of climate through several methods. One method is to identify preferences over long-term climate by observing individuals' choices over where to live (Albouy et al. (2016); Sinha et al. (2018)). These papers estimate these preferences using cross-sectional variation in climate; however, correlation between climate and other place-specific characteristics may bias these estimates. A second method is to use survey data on self-reported happiness or life satisfaction to examine the impacts of heat on utility (Denissen et al., 2008; Feddersen et al., 2012). Finally, Baylis (2020) finds evidence from Tweeting behavior that overall moods tend to sink on extremely hot days: expressed sentiment on Twitter becomes more negative, accompanied by an increase in profanities. Baylis then compares the impacts of heat on Twitter sentiment with the impacts of changes in quarterly local wages to back out a monetary valuation of temperature, finding that individuals would be willing to pay approximately \$5 to \$12 to exchange a 30-35 degree day with a 20-25 degree day (degrees in Celsius).[3]

In addition, we add to the literature documenting changes in individuals' allocation of time according to the weather. Graff Zivin and Neidell (2014) find that individuals in the United States tend to reallocate leisure time indoors on hot days, indicating that outdoor leisure becomes relatively less valuable in the heat. Similarly, Connolly (2008) finds that rainy weather lowers the opportunity cost of labor by making leisure time less attractive, causing individuals to increase their labor supply. More recently, Chan and Wichman (2020) document the change in time spent cycling according to the weather, leveraging a relatively active outdoor leisure activity.[4] We contribute further evidence that individuals prefer to allocate time away from outdoor leisure on hot days.

The rest of the paper proceeds as follows. Section II. provides background on MLB attendance, discussing the history of determinants of what fans pay both at the box office and on secondary markets such as StubHub and Seatgeek. Section III. describes the data we use on game attendance and game day weather. Section IV. introduces our regression framework, which uses stadium-specific seasonally unusual weather variation to causally identify the nonlinear impacts of temperature on game attendance.

---

3 In addition, there is a related literature that seeks to estimate willingness to pay to avoid the health impacts of pollution, by valuing the sickness caused by air pollution in a value of statistical life framework and/or by estimating the willingness to pay for exposure-reducing technologies (see, for example, Ito and Zhang (2020); Deschenes et al. (2017)). Staying inside on a hot day may be partially thought of as a defensive instrument against health impacts of high heat, but the willingness to pay parameter we estimate encompasses the pure utility impacts of high heat as well, aside from the health impacts.

4 Chan and Wichman (2020) find more adjustments on cold days, leading them to conclude that climate change will, on net, increase outdoor recreation. Like this study, they do find a negative adjustment on very hot days as well.

Section V. presents the results of the analysis of weather and attendance. Section VI. investigates whether the effects vary significantly by time of day or by usual climate. Section VII. first translates the main results on attendance to an estimate of the disutility of heat, and then presents the results of a second exercise examining changes in Seatgeek ticket prices according to the weather. Section VIII. concludes.

II. Background: the Market for Baseball Tickets

This section describes the market for baseball tickets, focusing on the determinants of game attendance and secondary market prices and providing the foundation for the structural assumptions made in Section VII.A.

II.A Attendance and the Primary Market for Tickets

Historically, the majority of fans bought tickets in the primary market, directly from teams. Teams sold tickets at a fixed price: prices varied only according to seat location and were otherwise kept constant throughout the season.[5] Furthermore, the sports literature has consistently found that teams have historically priced their tickets in the inelastic region of demand, below price levels that would be expected to maximize revenues.[6] Under this pricing model, observed variation in attendance from game to game is thus driven by non-price-related demand-side factors.

In more recent years, many teams have moved to variable, and subsequently dynamic pricing for tickets. Teams using variable pricing still set prices before the season begins, but they vary prices by additional game-level characteristics that predictably determine demand, such as day of the week, season, and popularity of the rival team. Teams using dynamic pricing, on the other hand, vary prices over time within the season according to observed changes in demand. Courty and Davey (2020) documents year of adoption of variable and dynamic pricing for each team in MLB, showing that most teams had adopted dynamic pricing by 2016, but none had before 2009. The earliest adopter of variable pricing was the Colorado Rockies in 1997. By definition, we would expect dynamic pricing to move pricing into a more elastic region of demand, which affects the assumptions of our willingness to pay exercise. Partially due to this concern, in our main specification we restrict our sample to 1950-2000, when fixed pricing was dominant.

5 For example, season ticket holders and other fans were mailed pre-ordered tickets with fixed prices printed on them before the season began.

6 The sports literature has proposed several explanations for this: one is that cheap tickets may draw fans through the gates to buy high-margin concessions goods (Krautmann and Berri, 2007). Another is that inexpensive ticket prices may "hook" new loyal baseball fans (Ahn and Lee, 2003). Fort (2004b) shows evidence that teams may keep ticket prices low due to the incentive to attract public subsidies. Overall, the literature suggests that pricing in the inelastic region may not be inconsistent with profit maximization (Fort, 2004a).

II.B The Secondary Market for Tickets

Secondary sales have always been part of the market for MLB tickets: historically, scalpers purchased tickets in advance, especially for popular games, and then sold them on game day in front of the stadium. Scalpers attracted public distaste by creating unwanted competition for tickets in the primary market and selling them at exorbitant prices (Bhave and Budish, 2017). In the past, sports leagues aggressively moved to limit ticket scalping, even leading to anti-scalping state laws in many places, albeit with limited effectiveness (Drayer, 2011).

The market for secondary sales of baseball tickets has exploded in recent years with the creation of websites such as Ticketmaster (on the Internet starting 1996), StubHub (2000), and Seatgeek (2009).[7] These websites typically allow sellers to upload tickets in electronic format, and have proprietary algorithms that determine a recommended price for the listing. Sellers then have the ability to set their own price or take the recommended price, and can change their listing price as desired. Stubhub, for example, allows sellers to set up a notification that alerts them if market conditions have changed such that their price now falls outside of a recommended range.

Sellers on these websites are a mixture of casual fans, season pass holders, and professional ticket brokers. An estimate from one major primary ticket seller suggests that about 30 percent of major league sports tickets are sold to brokers in the primary market (GAO, 2018).[8] On the other hand, using data from e-Bay, Sweeting (2012) shows that 88% of MLB ticket sellers list only a single set of tickets to a particular game, suggesting that fans (especially season ticket holders) make up the majority of the market. Prices on secondary markets are typically significantly higher than primary market prices: Sweeting (2012) finds that tickets are listed on StubHub at about twice their face value, on average. These prices represent a much more elastic region of the demand curve for baseball tickets (Diehl et al., 2015). The same paper shows evidence that prices are similar across multiple secondary market sites, suggesting that they can be thought of as part of the same overall market. With fully dynamic prices on both the primary and second markets in recent years, and a fixed supply of seats in a stadium, these facts indicate that secondary market prices are very likely to track consumer willingness to pay for baseball.

Another major change in this market over the last 30 years has been the construction of new, well-equipped, and often smaller, stadiums. Combined with the ease of buying tickets online, this trend has meant that stadiums have steadily filled up over the past few decades. Figure A1 shows a marked

7 Teams for the most part have also embraced the secondary market: In 2007, MLB formally signed an agreement with StubHub, making the website an official outlet for ticket sales complete with the use of official team logos. MLB in return shares in the revenue from sales of tickets on StubHub.

8 With the use of software bots, brokers have a significant advantage over fans in purchasing low-price tickets on the primary market, and this rent-seeking behavior significantly eats away at the efficiency gains of secondary markets (Leslie and Sorensen, 2014).

increase in average attendance as a percent of capacity across all teams since 1990. We would expect fuller stadiums and a more robust secondary market for tickets to limit our ability to observe an impact of weather on attendance at MLB games, which we discuss further below.

III. Data

III.A MLB Attendance Data

Our data source for MLB game attendance and game characteristics is www.retrosheet.org, following Neidell (2009). Retrosheet has collected game records from 1871 through present. We compile attendance records from every game for each stadium in the U.S. from 1950-2000. Pre-1950 games are dropped because of the availability of weather data, and post-2000 games are dropped to eliminate the influence of secondary markets and more sophisticated pricing strategies by teams, which arose around the turn of the century. Before 2000, we rely on the fixed pricing behavior of teams to translate attendance declines on hot days into a monetary estimate of the disutility of heat under minimal assumptions. In all, this provides a sample of over 80,000 games.

Attendance data captures gate entry, so attending fans will be counted regardless of whether they bought their ticket on the primary or secondary market, and no-shows do not count. Table A1 provides information on each team included in the sample. The teams' home stadiums are well distributed throughout the U.S., and cool, temperate, and hot climates are well represented, as shown in Figure A2. In addition to attendance records, Retrosheet reports game-level information on day of the week, time of day, scores, and detailed play-by-play records of the progression of the game. This information allows us to control for a battery of game-level observable characteristics to improve the precision of our estimates.

Table 1 reports summary statistics on game-level characteristics. On average, around 20,170 people attend an American League game and 22,310 people attend a National League game.[9] With average stadium capacity close to 50,000, this implies that stadiums are on average less than half full. The existence of substantial leftover capacity in the average baseball game allows for variation in attendance based on game-day factors such as temperature. The average game lasts about 159-163 minutes, and nearly two out of three take place in the evening, as opposed to in the afternoon. In a small percentage of cases, more than one game takes place in the same stadium on the same day, usually due to postponement of a previous game.

A number of MLB teams play in covered stadiums, with some transitioning to such stadiums in recent years. These stadiums are either permanently enclosed (domed) or are temporally covered for bad

[9] These leagues have historically differed slightly in the rules they use to govern gameplay, but are both part of the umbrella organization of Major League Baseball and play against one another throughout the season.

weather (including hot days) using a retractable roof.[10] In our main specification, we drop all domed stadiums and stadiums with retractable roofs from the sample. We then later use these stadiums in a placebo test—we investigate the impacts of weather on baseball attendance in settings where baseball games are held in indoor or climate-controlled stadiums. This eliminates the influence of exposure to the weather on the decision to go to a baseball game.

III.B Weather Data

We use data on daily maximum and minimum temperature and daily precipitation derived from the PRISM weather data set from 1950-2000.[11] These data files give weather information on a 2.5 x 2.5 mile gridded basis for the contiguous United States. The data are adjusted by Wolfram Schlenker to provide a balanced panel of weather station data: missing daily station readings are filled in by the distance-weighted average of the cumulative density function of surrounding stations.

We take the average of daily maximum and minimum temperature to create our measure of daily average temperature. We then merge the weather data with baseball attendance data by taking the average of the daily average weather readings from the four surrounding weather gridpoints for each stadium, weighted by inverse distance between each gridpoint and the stadium. Figure 1 displays summary statistics of the merged MLB-weather data, showing the number of MLB games in our sample whose daily average temperature fell into each of the 11 5-degree bins we use for our regression analysis. On average, game day temperatures are warm but relatively mild (the baseball season runs from April through September), with 65-70 degrees being the most common temperature range. However, we observe thousands of games that occurred on days with average temperatures over 85 degrees or under 45 degrees.

Figure 2 displays average game attendance as a percent of the capacity of the stadium for each of five major teams as it progresses throughout the season after opening day. The figure shows a pronounced peak in attendance around the mid to late summer months, but the peak is earlier and smaller in the relatively hot home stadiums of the Texas Rangers and the Atlanta Braves. This figure provides suggestive evidence that heat has a quantitatively important impact on attendance at MLB games.

---

[10] The home stadiums of the Arizona Diamondbacks, Seattle Mariners, Texas Rangers, Houston Astros, Milwaukee Brewers, and Miami Marlins all now have retractable roofs, and were built in the past few decades.

[11] The data can be downloaded from Wolfram Schlenker's website, at http://www.columbia.edu/ ws2162/links.html

IV. Empirical Strategy

We estimate the impacts of temperature on attendance at baseball games using a semi-parametric specification. This flexibly identifies the effect of a day's average temperature falling into a certain bin on game day attendance, relative to a reference bin of 65-70 degrees Fahrenheit. We estimate the following equation:

$$y_{isdmv} = \sum_j \beta_j \cdot Exposure_{sdm}(T_j) + \theta_{sm} + \lambda_{my} + \eta precip_{sdm} + \nu X_{isdmv} + \phi_v + \epsilon_{isdmv} \tag{1}$$

where y_{isdmv} is logged total attendance at game i at stadium s on date d in month m against visiting team v. β_j is the coefficient of interest and gives the effect of daily average temperature falling in bin j on game day on attendance, relative to the reference bin of 65-70 degrees Fahrenheit. We estimate the impact of temperature falling into 10 5-degree bins: $<$45 degrees, 45-50, 50-55, 55-60, 60-65, 70-75, 75-80, 80-85, 85-90, and $>$90. The omitted bin is 65-70 degrees Fahrenheit, so the thought experiment is to compare the impact of the game day's temperature falling into the bin of interest with the impact of temperature instead falling between 65-70 degrees. θ_{sm} are stadium by month of year fixed effects (61 stadiums by 12 calendar months), which net out the average popularity of games at a team's home location at a certain time of year and therefore control for usual place-specific monthly weather conditions. Additionally, controlling for stadium rather than home team fixed effects nets out changes in attendance within teams across stadiums due to differences in capacity or stadium amenities. λ_{my} are month by year fixed effects, which net out any universal time-varying determinants of baseball attendance. $precip_{sdm}$ refers to a linear control for daily total precipitation on game day.[12] X_{isdmv} are controls for observable characteristics of a particular game that are likely important determinants of attendance. In our main specification, we control for day-of-week fixed effects, the share of the home team's last 100 games that it won, and whether the game was an afternoon or evening game.[13] Finally ϕ_v are fixed effects for the visiting team, netting out variation in game attendance due to the popularity of the rival team. Standard errors are clustered at the stadium level.

---

[12] Prior work has found that attendance at baseball games falls on rainy days, but that attendance subsequently bounces back, suggesting the presence of habit-formation in baseball attendance (Ge et al., 2020).

[13] If a team moved cities and changed names, we consider the team in the new city to be a new team. In our sample period, for example, the Kansas City Athletics moved to Oakland and became the Oakland Athletics, the Washington Senators moved to Minnesota and became the Minnesota Twins, and the Seattle Pilots moved to Milwaukee and became the Milwaukee Brewers. If a team changed names but stayed in the same city, we consider the teams to be the same. For example, the Florida Marlins became the Miami Marlins and the Anaheim Angels became the Los Angeles Angels. More information can be found in Appendix B.

IV.A Identifying Assumption

Our identifying assumption is that any unobserved determinants of attendance at a given MLB game are uncorrelated with variation in game day weather after controlling for stadium by month and month by year fixed effects. That is, instances of weather that are unusual for a certain stadium in a certain month are not correlated with attendance for any reason other than the direct effects of weather on the enjoyability of attending a baseball game.

Our identifying assumption would not be violated, for instance, by attendance rising in later months in the season as excitement for playoffs builds in a way that is correlated with temperature falling in early fall. It would, however, be violated if the expected quality of the game were to be affected by unusually hot weather. For instance, if players' performance is affected by unexpected heat, and fans change their attendance behavior in anticipation of this, this would violate our identifying assumptions. However, as we show in section V., we don't see a significant negative impact of unexpected heat on indicators of game quality such as the number of runs in a game.

V. Main Empirical Results

Figure 3 displays the results of Equation 1 for the full sample of games from 1950-2000. The results corroborate the suggestion from Figure 2 that extreme heat dissuades baseball fans from attending games. Attendance appears to be highest between 75-80 degrees, and it falls by 14 percent on days with a daily average temperature above 90 degrees Fahrenheit relative to days between 65-70 degrees. Attendance is also significantly lower on very cold days, with a 13-20 percent decline on days below 55 degrees.

Table 2 displays the results of several alternative specifications of Equation 1. The first column controls only for stadium by month and month by year fixed effects. The second column adds controls for daily precipitation, day of the week, and whether the game took place in the afternoon or evening. Finally, column 3 displays the results of the full specification, with visiting team fixed effects and controls for the home team's performance over its past 100 games. This is the same specification displayed in Figure 3. The results are stable across specifications, though the impacts of high temperature become larger and more precise as the additional controls are added in columns 2 and 3.

One mechanism for a decline in baseball attendance on hot days could be that fans expect the quality of the game itself to decline on hot days. If this were true, then part of the effect we observe on attendance may not reflect valuation of weather conditions, but rather, valuation of high quality game play. In Table A3 we investigate this possibility by taking advantage of the rich data on game play in the Retrosheet records. The table displays the results of Equation 1, but using indicators of the quality of the game as outcome variables. The results suggest that runs, and particularly home runs, become more frequent on particularly hot days. This is a well known phenomenon: the baseball flies farther on hot days due to lower air density.[14] Strike outs become relatively less common on hot days. This indicates that if number of runs is an indicator of how "exciting" a game is to watch, if anything, games improve on hot days. On the other hand, games do last slightly longer ($<$2 minutes on a mean of 161 minutes), and teams go through more pitchers on hot days. Overall, however, these results suggest fans have little reason to expect a systematically less enjoyable game on a hot day, for any reason other than the discomfort of sitting in the heat.

One limitation of the attendance data is that we are unable to observe any changes in the composition of which seats are purchased. One possibility is that some fans choose to still attend baseball games even if it's very hot, but they buy seats in more sheltered areas of the stadium. To the extent that these seats are more expensive, this behavior reveals that the fans value avoiding the disutility of heat in a way that we are unable to capture. To the extent that this occurs, we will derive an underestimate of the disutility of heat.

VI. Heterogeneity and Placebo Tests

In this section, we examine heterogeneity in impacts by game characteristics and usual climate. Figure 4 splits the sample into afternoon games and evening games. The impacts of heat are much larger for afternoon games: the effects are seen starting at an average temperature of 80-85 degrees rather than being negative and significant only for $>$90 degrees. For the same daily average temperature afternoons will be hotter, so this larger effect is to be expected. Likewise, the negative impacts of very cold temperatures appear to be larger for evening games, while attendance rises for daily average temperatures in the 70-85 degree range relative to 65-70 for evening games.

As mentioned in Section III., our dataset also spans a variety of climates, from the hot and humid home of the Florida (now Miami) Marlins to frigid game days in Minneapolis. This variety allows us to examine whether fans' adaptation to their local climate affects their valuation of mild weather. To do so, we split the sample into thirds by the stadium's annual average temperature over the whole sample period. We then run our main regression, interacting the indicators for daily average temperature falling in the bins of interest flexibly with dummies for the home stadium falling in the coolest, hottest, and middle thirds of the sample. The results are displayed in Figure A3; while the results are noisy, there is little systematic evidence of the effects differing by usual climate. Interestingly, the effects on the hottest bin are smallest for the coolest third of stadiums. Thus, there is a lack of evidence that individuals' valuation of mild weather as opposed to very warm weather attenuates as they adapt and acclimatize

[14] see Florio and Shapiro (2016), for example.

to warmer climates. One reason for the lack of differential effect here, which contrasts with evidence from the mortality literature, could be the lack of available climate control at outdoor baseball games. While the availability of air conditioning substantially mediates the temperature-mortality relationship in the U.S. (see Barreca et al. 2016), there is no such mechanism for dampening the effects of heat on the enjoyability of outdoor activities.

Finally, we leverage the subset of games played in domed or retractable stadiums for a placebo test. These games are climate controlled, eliminating our proposed mechanism for lower attendance. Any remaining effect is interpretable as the impact of heat on the enjoyability of other outdoor activities. For example, if baseball was equally enjoyable regardless of the heat, but visiting a water park or beach became much more pleasant, we would pick up a negative attendance effect in the absence of heat-related disutility at the baseball game. Instead, Figure 5 demonstrates that baseball attendance *increases* in climate controlled environments. This suggests the decline in attendance for outdoor games on hot days is in spite of other outdoor activities becoming less pleasant as well. We discuss the impacts of heat on alternative choices of outdoor activites further below.

VII. The Willingness to Pay for Baseball, the Disutility of Heat, and Welfare Loss on Hot Days

This section leverages the attendance results above and an alternative approach relying directly on pricing data to make inferences about how the willingness to pay for baseball games changes on hot days. We then make explicit under what conditions the willingness to pay for baseball games informs us about the disutility of heat. Finally, we discuss what can be learned about the aggregate welfare loss of a hot day via this channel, after optimizing adjustments take place.

VII.A Declining Attendance and the Willingness to Pay for Baseball

To generate an estimate of the change in the willingness to pay to attend a baseball game according to the weather, we must make two assumptions, consistent with previously mentioned facts about the market for baseball tickets. First, we assume that our estimates constitute a change in demand, rather than a joint quantity-price equilibrium adjustment. Equivalently, we assume a perfectly elastic supply curve at the individual game level; prices do not adjust to demand declines for a given game. As mentioned previously, baseball box office ticket prices historically have not adjusted on a game-by-game basis. Only in the mid-2000s did teams begin using dynamic pricing, wherein ticket prices would change based on realized demand for respective games. We conservatively estimate the main results on only the pre-2000 sample to avoid the possibility of dynamic pricing practices biasing our results.[15]

In order to map these quantity changes into price-equivalent changes, we additionally need an assumption on the local elasticity of demand. A natural benchmark is unit elasticity, where a revenue maximizing monopolist would price. Interestingly, as discussed in Section II., Krautmann and Berri (2007) summarizes a literature that is puzzled by consistent elasticity estimates that fall well below unity for MLB pricing. As we show below, the higher is the assumed elasticity, the *lower* our estimate. We therefore take 1.0 as a conservative upper-bound on this term for our primary estimate.

Figure 6 formalizes how these two assumptions allow us to translate our coefficient estimates into monetary values. Note that this is a log-log plot in accordance with the log-quantity estimates from Section V.. If prices are fixed with respect to weather, it must be the case that the estimated quantity change β, from Table 2, represents a demand shift. Against a horizontal pricing curve any quantity change identifies a shift in demand. To translate this leftward shift in demand into an equivalent price change, we ask by how much (log) ticket prices would need to fall to reverse the decline in attendance. Formally, the (local) log-price-change-equivalent, x_p, for any log-quantity change, β, takes the following form, where e_p^d is the price elasticity of demand.

$$x_p = \frac{1}{e_p^d}\beta \tag{2}$$

Returning to Figure 6, let the equilibrium outcome on a pleasant day be point A; the equilibrium after a demand decrease from a hot day is point B. Point C determines the price decrease that would induce an offsetting increase in (log-)quantity. The inverse of the price elasticity of demand governs the slope of this curve, and accordingly how much of a price decrease, x_p, is necessary to get to point C.

Holding quantity fixed between points A and C ensures that for the same number of consumers the value of attending the game surpasses the outside option. Assuming homogenous preferences over temperature, the price decrease generating point C compensates consumers for their lost surplus due to the high heat.[16] If point A was generated by price-temperature bundle $(p_0, 65 - 70)$ then point C would be generated by $(.86p, >90)$. Inflation-adjusted ticket prices over the sample are approximately \$18, and so this change in the willingness to pay for baseball is about \$2.52 per game.

Under a more realistic assumption of heterogeneous dislike of high heat, the analysis is less straightforward. However, as long as the variance of the surplus from baseball games is large relative to the

[15] Though we note that the bias would be downwards, as we would observe a smaller quantity effect if prices did in fact adjust where we assume they do not.

[16] To see this, consider that the marginal consumer receives 0 surplus at point A, and the price decreases enough to exactly offset the loss from temperature and return her to zero surplus. Assuming homogenous preferences for weather, this estimate of the decline in surplus due to temperature for the marginal consumer is representative.

variance of the disutility of heat, the setting is reasonably approximated by one of homogeneous disutility.[17] As the surplus associated with attending baseball games varies tremendously (think of people's taste for sporting events, their time-cost of travel, etc) and human physiology only allows for so much variance over temperature related utility, we think this is a reasonable assumption. Thus, the $2.52 estimate above approximately represents the average change in willingness to pay.

If the sports economics literature is correct that MLB teams price on the inelastic region of demand, then this estimate is a lower bound for the change in willingness to pay for baseball. Suppose instead the price elasticity of demand were $\frac{2}{3}$, which is still on the upper end of papers summarized by Krautmann and Berri (2007) that estimate the price elasticity of demand in this setting. Then $\frac{1}{e_p^d} = \frac{3}{2} > 1$ which implies prices would need to fall by $1.5 \times \beta$, or about 22%. The corresponding decline in willingness to pay would also increase by 50%, to $3.78. Figure A4 plots the range of estimates as they correspond to different elasticity assumptions.

Further, if the supply curve has a non-zero slope in practice, our empirical quantity change estimates understate the true demand shift. In the limit, a perfectly inelastic supply curve would produce no quantity change even for a large demand shift. We would erroneously conclude there is *no* preference for the cooler day. This is an additional way in which our conclusions err on the conservative side. For all of these reasons our estimate of $2-$4 dollars per attendee, per game, is likely to be a lower bound.

VII.B The Willingness to Pay for Baseball and the Disutility of Heat

The exercise above produces an estimate for the decline in willingness to pay for baseball on hot days. How does this translate to an estimate of the disutility of heat? This conceptual leap is straightforward in the simple case where individuals have two possibilities for leisure time: indoor leisure and attending an outdoor baseball game. In this case, on a hot day, demand for indoor leisure shifts out by the exact same amount that demand for baseball games shifts inwards, and the magnitude of that shift, as measured by our calculation, clearly represents the disutility of heat. In the real world, individuals have a number of outdoor activities to choose from, and the "disutility of heat" depends on that choice. For instance, running a marathon or going for a bike ride would plausibly involve a higher disutility of heat, while going to the beach may involve a lower disutility of heat. Among truly heat-exposed outdoor activities, baseball may have a relatively low disutility of heat, due to its sedentary nature, the fact that many seats are shaded, and due to the availability of cold refreshments.

Another, related consideration is the relationship between our estimates, the disutility of heat, and

[17] This is because the random variable determining attendance—baseball surplus plus heat disutility—will have a variance approaching the variance of baseball surplus in the case where this term is much larger. Summing two normal random variables, for example, results in a summed variance that will be very near the larger term if it is indeed much larger.

the availability of alternative outdoor activities on hot days. We calculate the magnitude of the price drop that would reverse the decline in attendance on hot days, which will in general be mediated by potential attendees' best available alternative activities. On a hot day, many individuals are likely to substitute from outdoor to indoor activities, but there may also be substitution towards other outdoor activities that are less demanding in the heat. This substitution across outdoor activities may lessen or exaggerate the decline in attendance at baseball games, affecting the size of our disutility estimate.

Here we formalize the impact of available alternatives on our disutility of heat estimate. Let $u_{bb,70}$, $u_{bb,90}$ be the monetized utility that a fan gets from baseball (*bb*) on days that are 65-70 and over 90 degrees, respectively; $u_{other,70}$, $u_{other,90}$ is the utility they get from whatever activity is their best available alternative in the heat.

The following two equations describe the relationship between the monetized utility of baseball and the alternative activity in mild and hot weather. We focus on the marginal fan on the hot day, such that they would be indifferent between activities if the cost of a baseball game were decreased by \$2.52. Therefore, this fan can be thought of as the "last" attendee brought back into the stadium with the price drop and we can infer that they prefer their outside option by exactly \$2.52.

$$u_{bb,70} - u_{other,70} = X \tag{3}$$

$$u_{bb,90} - u_{other,90} = -2.52 \tag{4}$$

The first equation is agnostic regarding the preference ordering between the alternative activity and baseball on a mild day: it could be that the fan would have gone to the baseball game on a mild day, so that X is a positive number, or that the fan would have preferred the alternative activity on a mild day, such that X is a negative number. Note that we do not assume here that the best available alternative activity in mild temperatures is the same as the best available alternative activity in hot temperatures. Equation (4) holds following our empirical inference that \$2.52 is the price drop necessary to return this fan to indifference (so they must prefer their other activity by exactly \$2.52 on hot days).

Subtracting (4) from (3) yields:

$$(u_{bb,70} - u_{bb,90}) + (u_{other,90} - u_{other,70}) = X + 2.52 \tag{5}$$

and subsequently:

$$(u_{bb,70} - u_{bb,90}) = X + 2.52 - (u_{other,90} - u_{other,70}) \tag{6}$$

The lefthand side of this equation then describes the decline in the utility of attending the baseball game in the heat, which is the disutility of heat parameter that we seek to estimate empirically. The 2.52 on the righthand side is our empirical estimate. Therefore, our empirical estimate differs from the disutility of heat, in general, by $X - (u_{other,90} - u_{other,70})$. Does this imply an over- or underestimate the disutility of heat? Consider four cases. First, suppose the other activity is an indoor activity whose enjoyability is unaffected by the weather. Then $u_{other,70} - u_{other,90} = 0$, and our estimates differ from the disutility of heat only by X. For a fan that was just on the margin of going to the baseball game or going indoors in mild temperatures, so that X is 0, our estimates pick up exactly the monetary disamenity of the hotter day. To the extent that X is positive, so that the individual strictly prefers baseball to going indoors on a mild day, we underestimate the disutility of heat.

A second potential case is a fan that would have gone to the baseball game on a mild day (so that X is weakly positive), but due to the heat switched to an activity that got absolutely more pleasant due to hot temperatures (so that $u_{other,90} - u_{other,70}$ is positive, and we overestimate the disutility of heat). An example could be a water-based outdoor activity such as a water park. However, it seems to us that there are very few outdoor activities that are absolutely more pleasant on days with an average temperature over 90 than they are on a mild day. In suggestive support of this, the evidence shown in Figure 5 suggests that an indoor baseball game (in a domed stadium or one with a retractable roof) becomes more preferable on average compared to outside options on hot days. This is evidence that attendance declines on hot days in open stadiums are not driven by an increasing appeal of other outdoor activities - if that were the case, we would expect attendance declines for indoor stadiums as well. Furthermore, most summer baseball games take place in the evening hours, and the effects of heat on attendance persist even for these later games, as shown in Figure 4. While individuals may plausibly substitute towards more preferable outdoor activities such as swimming in the afternoon hours, this is unlikely to be the individual's outside option between 7:00-10:00 pm on a weekday. Therefore, we do not expect this second case to have an important influence on our estimates.

A third possibility is that a $2.52 decline in ticket prices would bring in a fan who would have attended a baseball game on a mild day, but absent any price decline would have chosen an alternative activity on a hot day that got less pleasant due to the heat, but to a lesser extent than baseball (in other words, $u_{other,70} - u_{other,90} < u_{bb,70} - u_{bb,90}$). In this case, $u_{other,70} - u_{other,90}$ is a positive number, and we underestimate the disutility of heat by $X + u_{other,70} - u_{other,90}$. Finally, a fourth possibility is a fan that would have strictly preferred a different activity (perhaps a more active one such as hiking or jogging) on a mild day to baseball, such that X is a negative number, but the disutility of heat for the alternative activity is so large that the $2.52 decline in ticket prices induces them to switch to baseball on a hot day. In this case, it's theoretically ambiguous whether $X - (u_{other,90} - u_{other,70})$ is a positive or negative number and therefore whether we underestimate or overestimate the disutility of heat for baseball. However, if the strength of the individual's preference for the alternative activity were larger than the disutility of heat of that activity, it's unlikely the individual would be close enough to the margin of going to the baseball game on a hot day to be induced to switch by a $2.52 price drop. To see this, note that the decline in the level of attendance on a hot day means that there are enough people in the first three cases to induce to switch back to baseball with a price drop less than or equal to the disutility of heat without needing to compensate someone in this fourth category more than the disutility of heat to attend the baseball game.

The relationship between the empirical estimate we derive from our exercise and the true "disutility of heat" in this context will depend on what combination of these four cases is closest to the margin of attending or not attending a baseball game on a hot day. However, for the range of most common and likely best available alternative activities, our empirical estimate is likely weakly lower than the disutility of heat.

As a final note, in this section we focused on the change in willingness to pay for baseball on a hot day, despite the fact that in Section V. we estimated even larger impacts of cold days on attendance. Following the same procedure, we would estimate a willingness to pay for warmer temperatures in the range of $2.29-$3.67. However, we place less emphasis on this result due to the fact that we think the setting here is likely to produce a relatively large estimate of the disutility of cold. Given the fact that baseball is a sedentary activity, attendees are likely to feel cold in a temperature range where they may feel particularly comfortable participating in a more active outdoor activity. For example Greenstone et al. (2019) finds that the optimal temperature to run a marathon is in the 40's, which is a temperature range where we estimate very large negative effects on baseball attendance.

VII.C Seatgeek Ticket Prices and the Willingness to Pay for Baseball

In this section we exploit secondary market ticket price data directly for a much smaller sample of games to estimate the change in the willingness to pay for baseball under fewer structural assumptions. While box office ticket prices have historically been fixed with respect to variation in game-level characteristics, for the last 15+ years it has been possible to purchase tickets on online secondary markets, such as StubHub, Ticketmaster, and Seatgeek. On these secondary markets individuals update prices frequently and appear to optimize at the game level, rather than the season level (Sweeting, 2012). The existence of secondary markets, and generally more ways to buy and sell tickets with flexible prices, means tickets are allocated more efficiently. This amounts to making the market look more like one with perfectly inelastic supply—demand changes will then be reflected in price changes. Therefore, this setting is one where it makes more sense to translate a change in ticket prices directly into a measure of the change in willingness to pay for baseball.[18]

[18] To our knowledge, only Paul and Weinbach (2013) has previously examined the impacts of weather on baseball ticket prices, and we are the first to control for place-specific seasonality to separate out the causal effect.

During the 2021 MLB season, we scraped daily price data of listings on Seatgeek.com for each MLB game.[19] The data contain information on average, median, minimum, and maximum prices for ticket offerings on Seatgeek for each game, as well as the overall count of listings. We pulled price information for listings left available on game day for each game, as well as for each day in the prior week. According to the Sports Business Journal, Seatgeek held 19.5% of the secondary market for Major League Baseball tickets in 2021 (McCormick, 2022). As mentioned in Section II., previous evidence finds that tickets on different websites tend to be comparable in price and move together (Sweeting, 2012).

The Seatgeek price data can be similarly merged with weather information using the PRISM daily climate data. This is the same underlying weather data source as we used for the attendance data, but without the corrections to create a balanced panel undertaken by Schlenker (the corrected dataset as of the time of this writing runs only through 2019). To conduct the merge, we interpolated the weather readings from the surrounding gridpoints in the weather data. Table A2 displays summary statistics of the ticket prices. The ticket prices have a strong rightward skew: the average average price is around \$445, while the average median is only \$80. Accordingly, we drop the top 1% of prices in our main analysis to reduce the influence of outliers. Ticket prices decrease steadily in the days leading up to game day, while the overall number of listings declines. This matches with evidence from Sweeting (2012) that prices fall over time, as the opportunity cost of not holding the ticket in the next period declines for the seller.[20] Figure A5 shows how median prices and counts of listings evolved throughout the 2021 season. Prices tend to be higher at the beginning and end of the season, and on weekends. The count of listings increased throughout the season.

For our primary analysis, we restrict the sample to only game day listing prices so that each game contributes only one observation to the dataset and so that prices contain as much information about responses to game day weather as possible. Given that tickets are "perishable," most of the activity on the secondary market will have completed by this time, and the price will be a function of remaining inventory given the past realizations of demand. Weather forecasts will have been available for several days, presumably growing more precise as game day approaches.

To investigate the impacts of game day weather on prices, we run a slightly altered version of Equation 1:

$$y_{isdmv} = \sum_j \beta_j \cdot Exposure_{sdm}(T_j) + \theta_{sm} + \eta precip_{sdm} + \nu X_{isdmv} + \phi_v + \epsilon_{isdmv} \tag{7}$$

[19] These sites do not make the necessary historical data available to perform this analysis for the entire period the secondary market has been active.

[20] As the paper mentions, theoretically consumers should rationally wait to buy tickets later, when ticket prices are lower, given this phenomenon. But customers might have costs of waiting, partially due to the investment required to attend a game, that could prevent this behavior from smoothing out the decline in prices as a game approaches.

The fixed effects specification differs for this exercise due to the one-year sample period. y_{isdmv} is logged average listing price for game i at stadium s on date d in month m against visiting team v. β_j is the coefficient of interest and gives the effect of daily average temperature falling in bin j on game day on prices, relative to the reference bin of 65-70 degrees Fahrenheit. For this data sample, we continue to use 5-degree bins, but the highest bin is >80 instead of >95. Due to the proliferation of domed stadiums and stadiums with retractable roofs in cities with warm climates in recent years, along with the limited sample period, there are too few observations above 85 degrees to use to estimate the effect separately. θ_{sm} are stadium by month of year fixed effects. X_{isdmv} are controls for observable characteristics of a particular game: day of week fixed effects, and whether the game was an afternoon or evening game. Finally ϕ_v are fixed effects for the visiting team. Standard errors are clustered at the stadium level. In this one year period, there were only 23 stadiums in the sample, raising concerns that there are too few clusters to accurately estimate clustered standard errors. To address this concern, we calculate wild cluster bootstrapped standard errors for the estimates (Cameron et al., 2008). We implement the calculations with 999 replications using Stata's boottest module (Roodman, 2021).

Figure 7 displays the results of Equation 7 on the full 2021 season sample. The results show clear declines in ticket prices at warmer temperatures, with average ticket prices falling by about 9.4% on days over 80 degrees.

Given our assumption in the previous section of an elasticity of demand of 1, our results using ticket prices are directly comparable to the results in the previous section using shifts in attendance. The results for ticket prices are notably larger, given that the highest bin in the ticket price analysis is only >80: in that bin, we see little effect on attendance, but substantial effects on ticket prices. Using the same prevailing average ticket prices as with the attendance exercise, a 9.4% decline in ticket prices amounts to a $1.69 estimate of willingness to pay for mild weather rather than weather > 80 degrees. In the next section, we discuss potential mechanisms for the differences in these effect sizes.

VII.D Robustness

Given that we are using game-level summary statistics of listing prices as our outcome variable, rather than individual seat-level data, one concern is that changing average ticket prices may partially reflect changing composition of seat offerings on Seatgeek. One example of this concern: if many fans that are just on the margin of purchasing or not purchasing a ticket to attend a game are likely to buy cheap, bleacher-type seats on Seatgeek, and these marginal fans are especially likely to be dissuaded by high game day temperatures, then a large negative impact on average prices of listings on Seatgeek could reflect these "nosebleed" seats being left unsold, as opposed to especially large changes in demand for baseball tickets. In this example, the underlying mechanism for the effect on prices is still reduced demand, but the magnitude of the effect does not necessarily correspond to the quantity of potential attendees dissuaded by heat. Sweeting (2012) shows evidence that the composition of available seats on StubHub does not change as game day approaches, reducing this concern.[21]

To investigate the potential contribution of these composition effects directly in our data, in Table 3, we examine the impacts of the weather on alternative summary statistics of listing prices. Column 1 displays the same specification as Figure 7, with log average price as the outcome variable. Column 2 looks at the impact of weather on the minimum price among the Seatgeek offerings on game day. If cheap tickets are more likely to go un-purchased on hot days relative to average or expensive tickets, this could drive down minimum prices to a greater degree than average prices. The effects in this specification are larger, but the impacts are not statistically distinguishable and are qualitatively similar. Model 3 uses the maximum ticket price as the outcome variable: the price of the most expensive ticket offering seems little affected by temperature, but the coefficients are noisy. Column 4 uses the median price as the outcome variable, with nearly identical results to column 1. Finally, in column 5 we examine the impacts of temperature on the number of listings remaining on Seatgeek on game day. Unsold listings clearly increase substantially on very hot days, with nearly 23 percent more listings available on days greater than 80 degrees than on game days between 65 and 70. This increase in listings may reflect a combination of more fans looking to sell their tickets and fewer fans willing to buy when the game day is forecasted to be hot. Listings similarly increase on cold days, although the effect is not statistically significant.

The interpretation that both decreases in demand and increases in supply can be thought of as reductions in desire to attend baseball games may be complicated by the presence of professional ticket brokers on these secondary markets. Professional brokers may have an incentive to lower their price in anticipation of lower demand (willingness to pay) once very high temperatures make their way into the forecast, but the mechanism is no longer as direct since the brokers are not deciding whether to attend the game themselves. It is presumably the case that brokers play a bigger role in the market for tickets in high-demand games, so one way to gain suggestive evidence on the role of brokers is remove especially high-demand games from the sample. In column 6, we display the same specification from column 1, except we exclude games that had attendance over 95% of the stadium's capacity according to the Retrosheet data. Over 11% of games in our 2021 sample had missing attendance data, so the results become significantly noisier in this specification, but the coefficient is qualitatively similar.

A final concern with using the 2021 Seatgeek data is that the Covid-19 pandemic had a major impact

[21] Sweeting (2012) also tests the hypothesis that late buyers are more price sensitive by reasoning that sellers of more expensive tickets would have to slash their prices more dramatically than holders of cheap tickets to attract buyers in the days before a game. However, he finds no evidence of this, suggesting that the composition of buyers does not change leading up to a game.

on operations for the season. This could impact prices on Seatgeek: many stadiums capped attendance at some percent of capacity, potentially restricting supply, and fans may have still been reluctant to attend crowded events, limiting demand. To the extent that Covid-related trends in attendance and ticket purchasing behavior interact with game day weather, this could create an external validity concern. For example, if games are regularly sold out to the limited capacity, where they would not have been sold out with full capacity, this could limit our ability to observe the impacts of heat on ticket prices, biasing our effects downward. On the other hand, if fans are on more "on the fence" about attending baseball games due to Covid, they may be more sensitive to weather impacts, causing us to estimate larger effects than in a typical season. Ultimately, this is an external validity issue, although the fact that we estimate similar effects using 2021 ticket prices as we get for 1950-2000 baseball attendance ameliorates this concern.

VII.E Comparing the Two Methods for the Willingness to Pay for Baseball

As discussed above, the two methods for estimating how willingness to pay for baseball varies with the weather give quantitatively different results, albeit within the same order of magnitude. In this section we discuss potential reasons for this divergence, presenting empirical evidence where possible.

One reason that estimates derived from ticket prices could be larger than estimates derived from changes in attendance has to do with our assumption of unit elasticity of demand for our sample period. As mentioned previously, to the extent that demand is more inelastic than we assume, our estimate from that exercise is a lower bound.

A second reason that the ticket price estimate may be larger could have to do with the existence of well-established, robust secondary markets for tickets. If fans that do not want to attend a game on a hot day, but have a ticket for the game, are able to resell their ticket to a customer that is less dissuaded by heat, that would reduce the impacts of heat on overall attendance, pushing these estimates towards 0. At the same time, that phenomenon would lead to an increase in tickets supplied in secondary markets, pushing prices downwards and estimates of effects on prices further away from 0. Therefore, the existence of secondary markets may be pushing the two sets of estimates in opposite directions. It is partially due to this that we limit our attendance sample to the years before StubHub was founded in 2000, when the secondary market would have been much smaller than it is today. However, in order to get a sense of the significance of this effect, in column 4 of Table 2, we estimate the effect on attendance for the sample of games 1950 through 2019, including the years of growth for the secondary market. The effects of heat and cold are both significantly smaller in this specification, lending credence to the idea that the existence of a secondary market limits the decline in attendance during extreme temperature events.

VII.F Aggregating the Welfare Losses of Hot Days

The central results of this paper can be used to generate a back of the envelope estimate for the aggregate costs of a hotter day with some additional assumptions. In Section VII.A, we discussed how our estimate translates into an estimate of the disutility of heat that a fan experiences at a baseball game. How well this approximates total welfare loss on hot days depends on how people change their behavior according to the heat: many individuals will adapt their behavior to avoid the disutility of heat (for example, by going inside), so that their welfare loss is smaller, for example. In light of this, we conservatively estimate aggregate welfare losses only using time that individuals spend outside on hot days, not including welfare loss due to individuals spending less time outdoors on hot days.

We first convert the game-level estimates to per hour estimates recognizing that the typical baseball game is about 161 minutes. The \$2.52 per game monetized cost then corresponds to \$0.94 per hour. Then, assuming the monetized damages to baseball spectators is similar to other outdoor activities, we can extend these estimates using time use surveys. The average time spent outdoors for US residents on over-90-degree days is just under 30 minutes, which implies that the monetary value of welfare loss per person-day of exposure to over 90 degrees is \$0.47.[22] It is then conceptually straightforward to aggregate the annual welfare loss from days over 90 degrees in the U.S.: we can simply multiply localized estimates of the count of days over 90 degrees, by the local population, by the per-person-day welfare loss of \$0.47 cents. To implement this, we use the 1980-2000 average annual count of days over 90 degrees for each grid cell in the United States, and overlay it with gridded population estimates from the Socioeconomic Data and Applications Center (SEDAC) for 2020. We then sum our estimate of annual welfare loss from days over 90 degrees for each grid cell in the contiguous United States, arriving at a total of \$175 million per year.[23] If we use a price elasticity of demand of $\frac{2}{3}$ for the disutility of heat estimate, this scales up by 50% to \$263 million.

As noted earlier, about 2/3 of the games in our main sample take place in the evening, making daily average temperature an overestimate of experienced temperature. It may also be of interest to draw conclusions about welfare losses based only on the sample of games that start in the afternoon, to capture welfare loss in the hottest portion of the day. As shown in Figure 4, we see negative effects on attendance at temperatures warmer than 80 degrees for afternoon games. A calculation of annual

[22] Taken from the American Time Use Survey.

[23] Note again that this does not consider utility lost due to the fact that many individuals substitute from outdoor activities to less-preferred indoor activities on hot days. The average time spent outdoors on 65-70 degree days is 40 minutes, implying that the lost outdoor time is 10 minutes per person. If we suppose the value of these minutes was something between \$0.00 (or they would not have been outside when it was pleasant) and \$0.94 per hour (or they would have remained outside with the hot weather), say \$0.47 under the assumption that the value is uniformly distributed, then this utility loss would scale up our estimates by a third. On the other hand, if there is substitution towards outdoor activites with a disutility of heat lower than that of baseball on hot days, we may overestimate welfare loss by applying our estimate of the disutility of heat to all outdoor time.

welfare loss from days above 80 degrees under historical average weather distributions suggests that welfare would be $1.52 billion higher each year if all days above 80 were replaced with days between 65 and 70 degrees.[24] This estimate is so much larger due to the fact that exposure to days with an average temperature between 80-90 degrees is much more common than exposure to days above 90 degrees, and also because people spend more time outside on these days.

A natural extension of this exercise is to use projected weather data under different warming scenarios to estimate how welfare loss from heat waves will intensify as climate change worsens. We undertake this exercise with caution: as discussed above, we de-emphasize our estimates of the utility loss of cold given the nature of the activity we study, and a full accounting of the costs and benefits of climate change would incorporate the fact that fewer very cold days will occur, which will have some benefits for outdoor recreation (see, for example, Chan and Wichman (2020)). However, focusing solely on the impacts of the increased count of hot days on welfare, we calculate that under an SSP5-RCP 8.5 warming scenario (SSP5-8.5 implies a world of rapid economic growth with high greenhouse gas emissions), annual welfare loss from 90 degree days will be $2.13 billion by 2080-2090, compared to a counterfactual scenario in which those days are replaced with 65-70 degree days.[25] If estimates from afternoon games are used, the welfare loss is $3.09 billion per year. On the other hand, under the SSP2-4.5 scenario (an intermediate scenario for carbon emissions), annual welfare loss will be $782 million ($2.07 billion using afternoon game estimates).

VIII. Conclusion

In this paper, we establish a new setting for examining the impacts of heat on individual utility and leisure time. We first estimate the causal impacts of weather on MLB attendance, showing that preferences for engaging in this passive, outdoor leisure activity significantly decline on very hot and very cold days. Next, we leverage these well-identified estimates to directly estimate fans' valuation of mild game-day weather using two methods. First, given fixed pricing of baseball tickets and findings from the literature that pricing is in the inelastic region of demand, we use the attendance estimates to back out an equivalent price change that would induce the same impact on quantity of tickets demanded, finding that fans would pay $2.52 to replace a game day temperature over 90 degrees with one between

[24] To create these alternative estimates, we multiply the count of days in each bin under historical weather distributions by our estimate of per person-per day welfare loss (calculated by multiplying our coefficients on each bin as shown in Figure 4 by the average ticket price, dividing by the average length of a game to get a per minute damage, and finally multiplying by the average time in minutes spent outside on days in each bin).

[25] We use the average of the 35 global climate models (GCM's) from the Coupled Model Intercomparison Model 6 (CMIP6) to calculate the projected count of 90 degree days. The data are downloaded from the NASA Earth Exchange (NEX) Global Daily Downscaled Projections (GDDP).

65 and 70 degrees. Second, we use secondary ticket pricing data from Seatgeek to directly estimate the impacts of heat on ticket prices, finding effects of a larger but remarkably similar magnitude.

The results speak to an understudied implication of climate change. Heat waves are known from an established literature to pose significant risks to human health, productivity, and economic output, but they also cause a widespread increase in disutility due to the discomfort of hot weather. The estimates derived from this paper suggest that these impacts may add up to significant welfare losses.

References

Adhvaryu, Achyuta, Namrata Kala, and Nyshadham Nyshadham (2019) "The light and the heat: Productivity co-benefits of energy-saving technology," *Review of Economics and Statistics*, 1–36.

Ahn, Seung C and Young H Lee (2003) "Life-Cycle Demand for Major League Baseball," in *Western Economics Association International Conference, Denver, CO*.

Albouy, David, Walter Graf, Ryan Kellogg, and Hendrik Wolff (2016) "Climate amenities, climate change, and American quality of life," *Journal of the Association of Environmental and Resource Economists*, 3 (1), 205–246.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro (2016) "Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century," *Journal of Political Economy*, 124 (1), 105–159.

Barreca, Alan I (2012) "Climate change, humidity, and mortality in the United States," *Journal of Environmental Economics and Management*, 63 (1), 19–34.

Baylis, Patrick (2020) "Temperature and temperament: Evidence from Twitter," *Journal of Public Economics*, 184, 104161.

Bhave, Aditya and Eric Budish (2017) "Primary-Market Auctions for Event Tickets: Eliminating the Rents of 'Bob the Broker'?"Technical report, National Bureau of Economic Research.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel (2015) "Global non-linear effect of temperature on economic production," *Nature*, 527, 235–239.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller (2008) "Bootstrap-based improvements for inference with clustered errors," *The review of economics and statistics*, 90 (3), 414–427.

Chan, Nathan W and Casey J Wichman (2020) "Climate Change and Recreation: Evidence from North American Cycling," *Environmental and Resource Economics*, 76 (1), 119–151.

——— (2022) "Valuing Nonmarket Impacts of Climate Change: From Reduced Form to Welfare," *Environmental and Resource Economics*.

Connolly, Marie (2008) "Here comes the rain again: Weather and the intertemporal substitution of leisure," *Journal of Labor Economics*, 26 (1), 73–100.

Courty, Pascal and Luke Davey (2020) "The impact of variable pricing, dynamic pricing, and sponsored secondary markets in major league baseball," *Journal of Sports Economics*, 21 (2), 115–138.

Denissen, Jaap JA, Ligaya Butalid, Lars Penke, and Marcel AG Van Aken (2008) "The effects of weather on daily mood: a multilevel approach.," *Emotion*, 8 (5), 662.

Deschenes, Olivier, Michael Greenstone, and Joseph S Shapiro (2017) "Defensive investments and the demand for air quality: Evidence from the NOx budget program," *American Economic Review*, 107 (10), 2958–89.

Deschenes, Olivier and Enrico Moretti (2009) "Extreme weather events, mortality, and migration," *The Review of Economics and Statistics*, 91 (4), 659–681.

Diehl, Mark A, Joel G Maxcy, and Joris Drayer (2015) "Price elasticity of demand in the secondary market: Evidence from the National Football League," *Journal of Sports Economics*, 16 (6), 557–575.

Drayer, Joris (2011) "Examining the effectiveness of anti-scalping laws in a United States market," *Sport Management Review*, 14 (3), 226–236.

Feddersen, John, Robert Metcalfe, and Mark Wooden (2012) "Subjective well-being: Weather matters; climate doesn't."

Florio, John and Ouisie Shapiro (2016) "How Will Rising Temperatures Change Baseball?," https://www.theatlantic.com/science/archive/2016/10/how-will-rising-temperatures-change-to-baseball/502778/.

Fort, Rodney (2004a) "Inelastic sports pricing," *Managerial and decision economics*, 25 (2), 87–94.

——— (2004b) "Subsidies as incentive mechanisms in sports," *Managerial and Decision Economics*, 25 (2), 95–102.

GAO (2018) "Event Ticket Sales: Market Characteristics and Consumer Protection Issues,"Technical report, Government Accountability Office.

Ge, Qi, Brad R Humphreys, and Kun Zhou (2020) "Are fair weather fans affected by weather? Rainfall, habit formation, and live game attendance," *Journal of Sports Economics*, 21 (3), 304–322.

Graff Zivin, Joshua and Matthew Neidell (2014) "Temperature and the Allocation of Time: Implications for Climate Change," *Journal of Labor Economics*, 32 (1), 1–26.

Greenstone, Michael, Faraz Hayat, and Michael Galperin (2019) "Technical Support Document: The Marathon, The Climate and Your Race Against Time."

Hsiang, Solomon M. (2010) "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America," *Proceedings of the National Academy of Sciences USA*, 107 (5), 15367–72.

Ito, Koichiro and Shuang Zhang (2020) "Willingness to pay for clean air: Evidence from air purifier markets in China," *Journal of Political Economy*, 128 (5), 1627–1672.

Krautmann, Anthony C and David J Berri (2007) "Can we find it at the concessions? Understanding price elasticity in professional sports," *Journal of Sports Economics*, 8 (2), 183–191.

Leslie, Phillip and Alan Sorensen (2014) "Resale and rent-seeking: An application to ticket markets," *Review of Economic Studies*, 81 (1), 266–300.

McCormick, Bret (2022) "Dealmaking, the full return of live events, and league and team business up for grabs are positioning sports ticketing for years of growth," *Sports Business Journal*, https://www.sportsbusinessjournal.com/Journal/Issues/2022/02/28/In-Depth/Ticketing.aspx.

Neidell, Matthew (2009) "Information, avoidance behavior, and health: The effect of ozone on asthma hospitalizations," *Journal of Human resources*, 44 (2), 450–478.

Paul, Rodney J and Andrew P Weinbach (2013) "Determinants of dynamic pricing premiums in Major League Baseball," *Sport Marketing Quarterly*, 22 (3), 152.

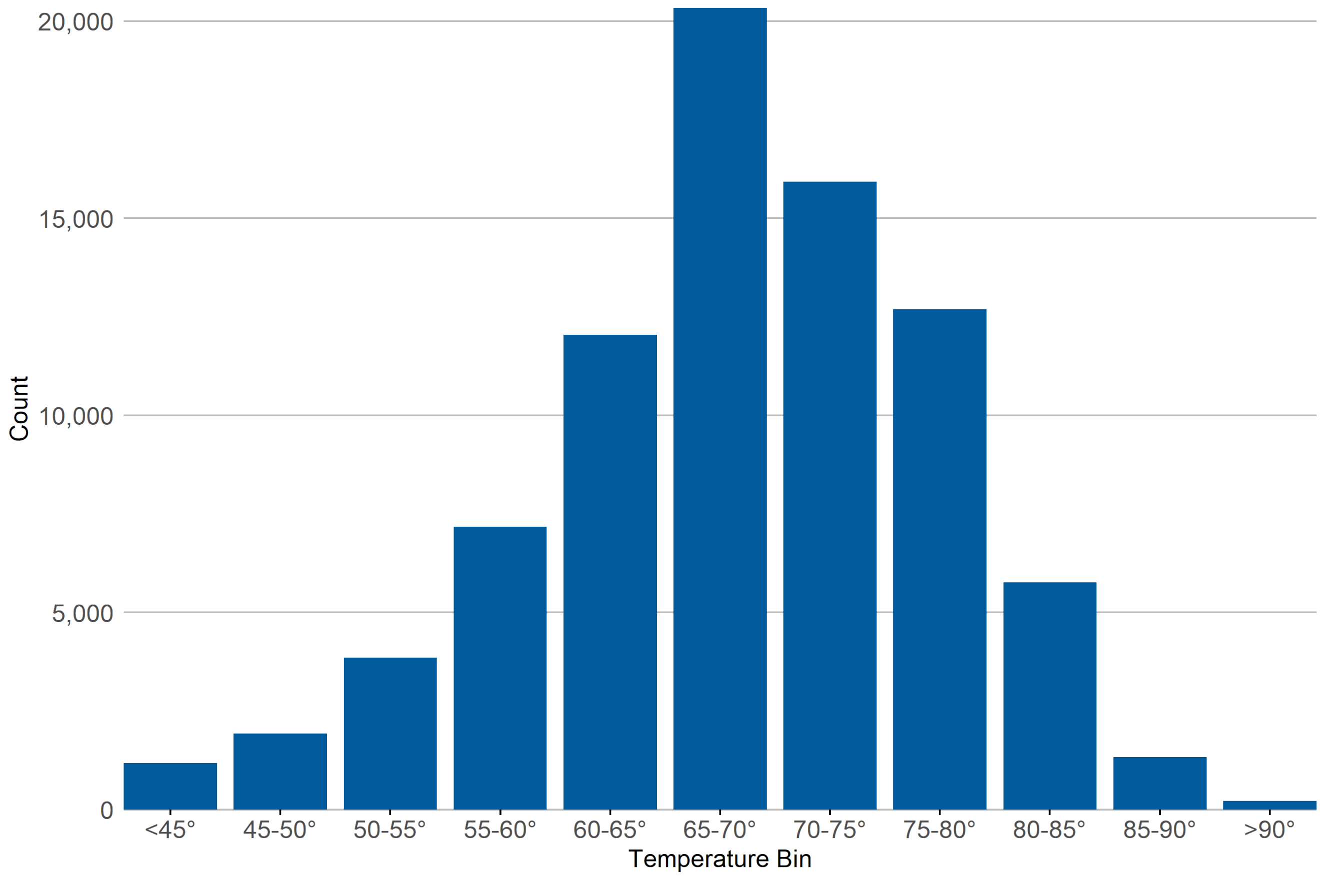
Roodman, David (2021) "BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed."

Sinha, Paramita, Martha L Caulkins, and Maureen L Cropper (2018) "Household location decisions and the value of climate amenities," *Journal of Environmental Economics and Management*, 92, 608–637.

Sweeting, Andrew (2012) "Dynamic pricing behavior in perishable goods markets: Evidence from secondary markets for major league baseball tickets," *Journal of Political Economy*, 120 (6), 1133–1172.

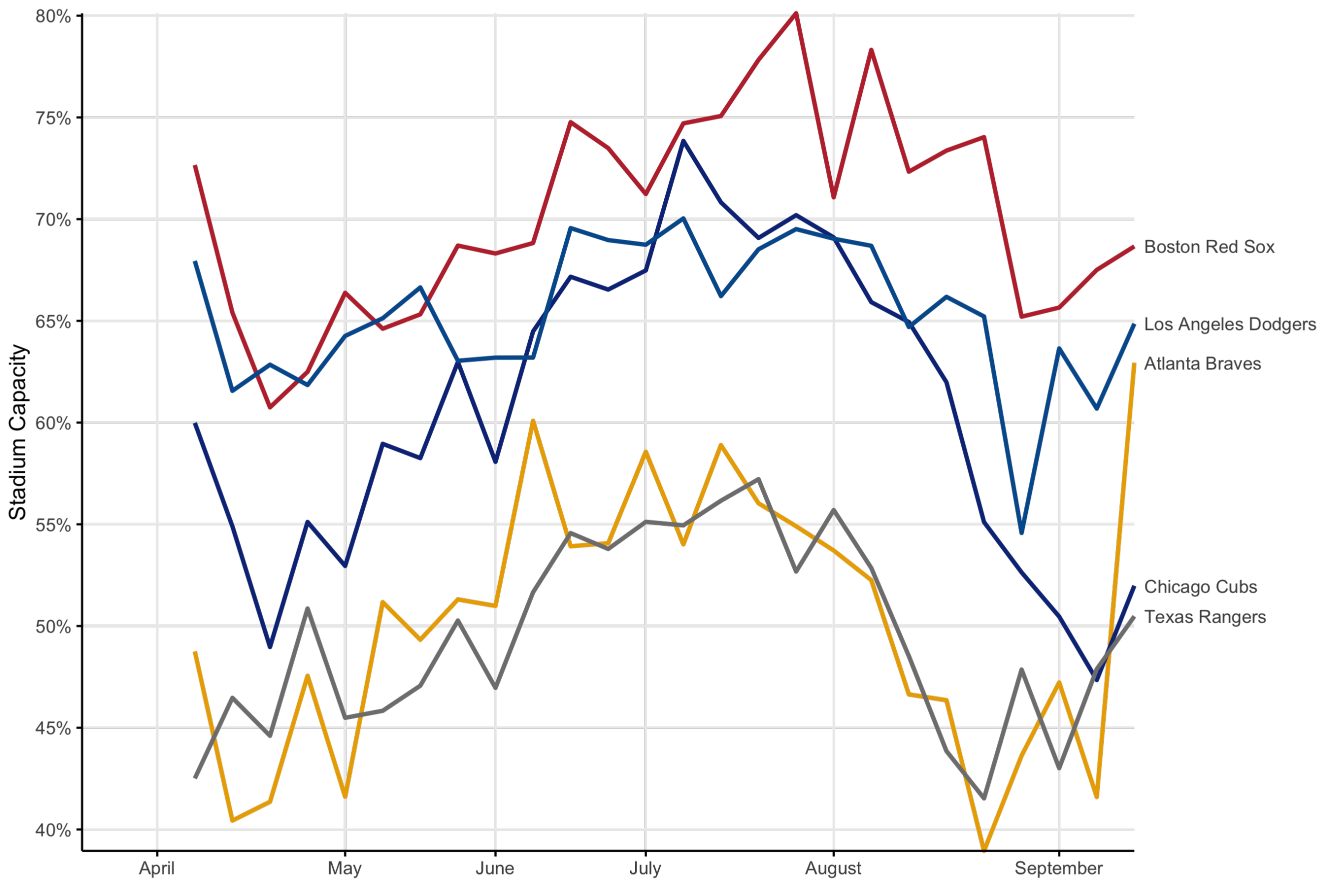
Zhang, Peng, Olivier Deschênes, Kyle Meng, and Junjie Zhang (2018) "Temperature Effects on Productivity and Factor Reallocation: Evidence from a Half Million Chinese Manufacturing Plants," *Journal of Environmental Economics and Management*, 88, 1–17.

Figure 1: NUMBER OF MLB GAMES IN EACH 5-DEGREE TEMPERATURE BIN, 1950-2000



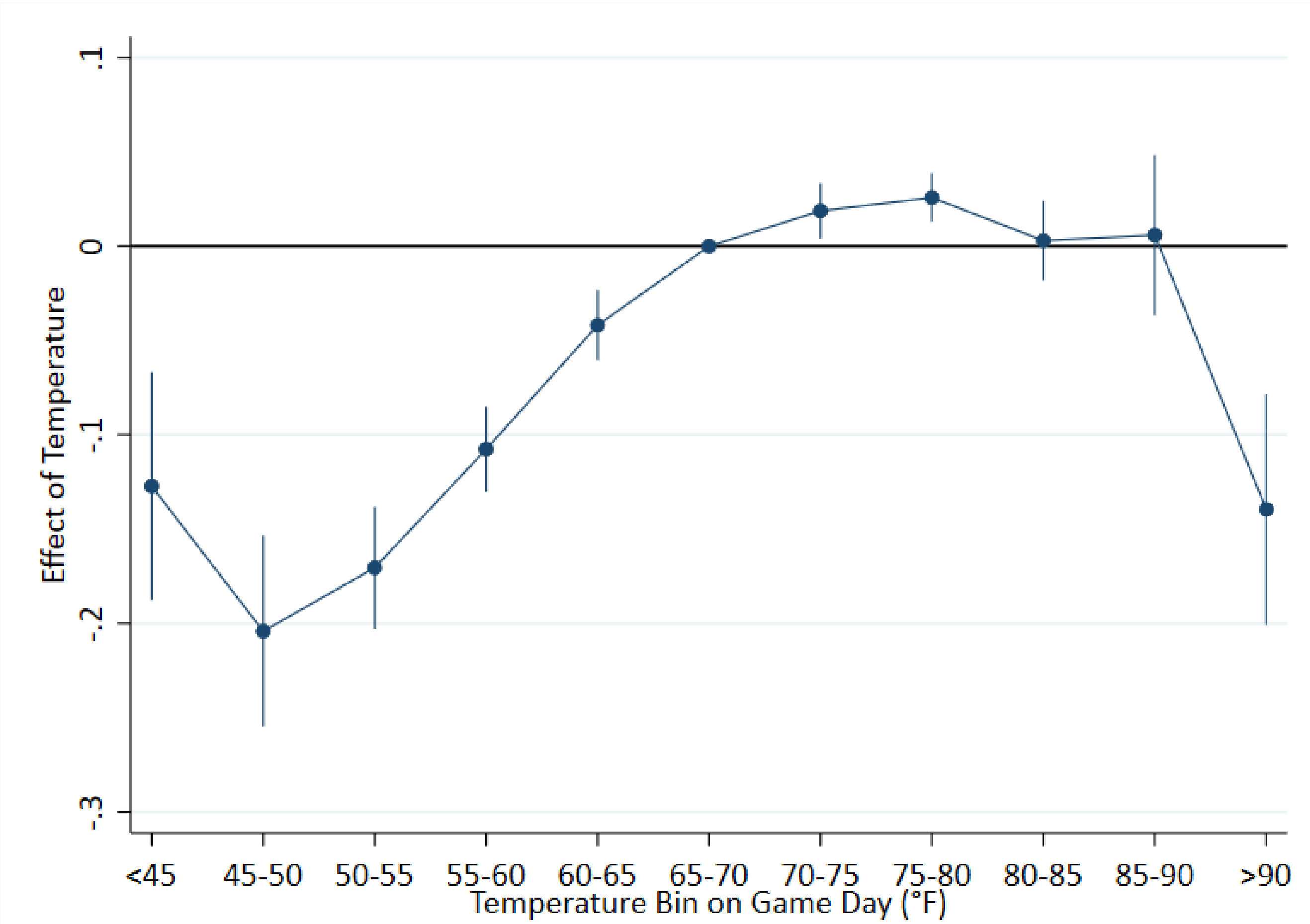
Note: This figure displays the count of MLB games that had average game-day temperatures in each of the 11 bins we use in our regression analysis.

Figure 2: ATTENDANCE SPIKES IN SUMMER MONTHS, BUT LESS SO IN HOT PLACES



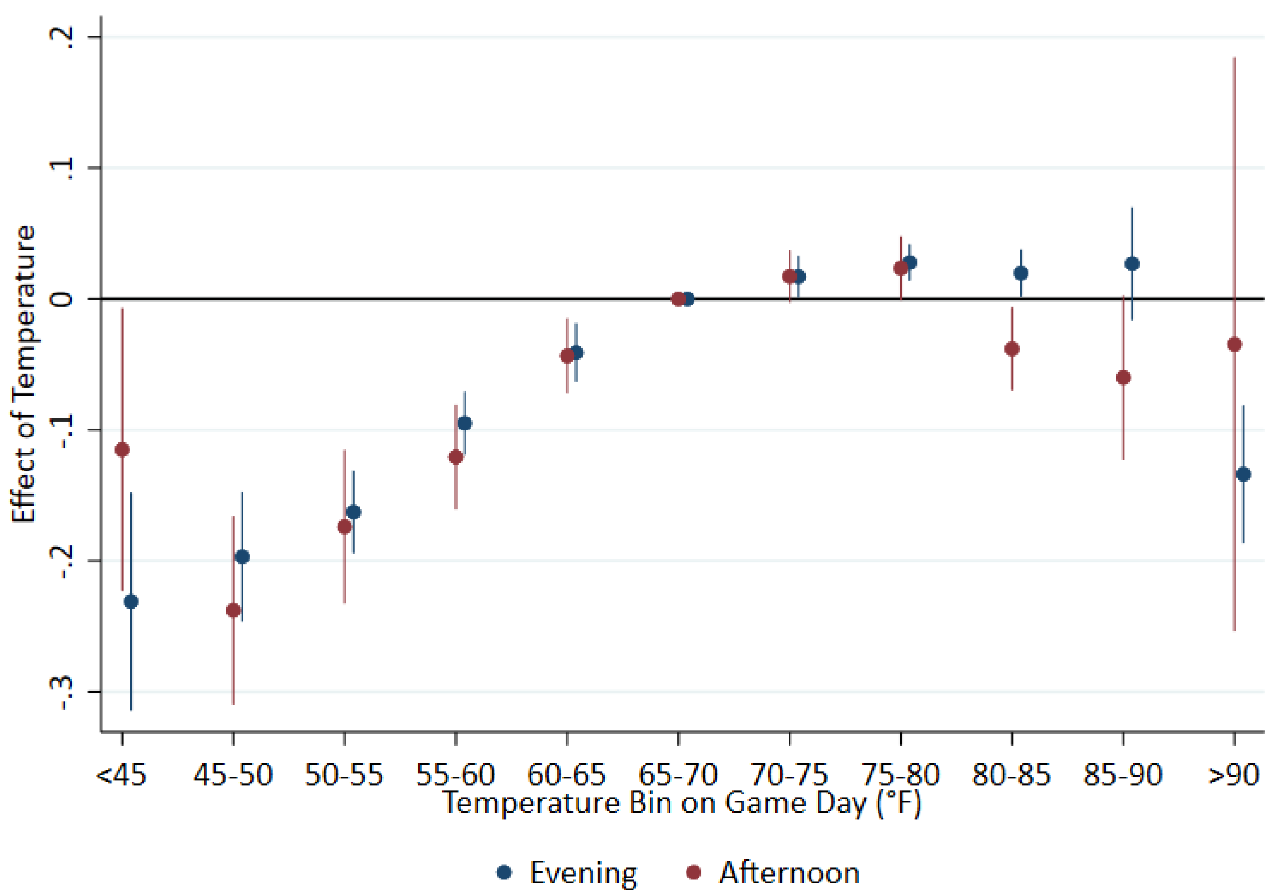
Note: This figure displays average attendance as a percent of ballpark capacity for home games for five major teams over the course of the average season.

Figure 3: BASEBALL ATTENDANCE FALLS SHARPLY AT EXTREMELY HOT AND COLD TEMPERATURES



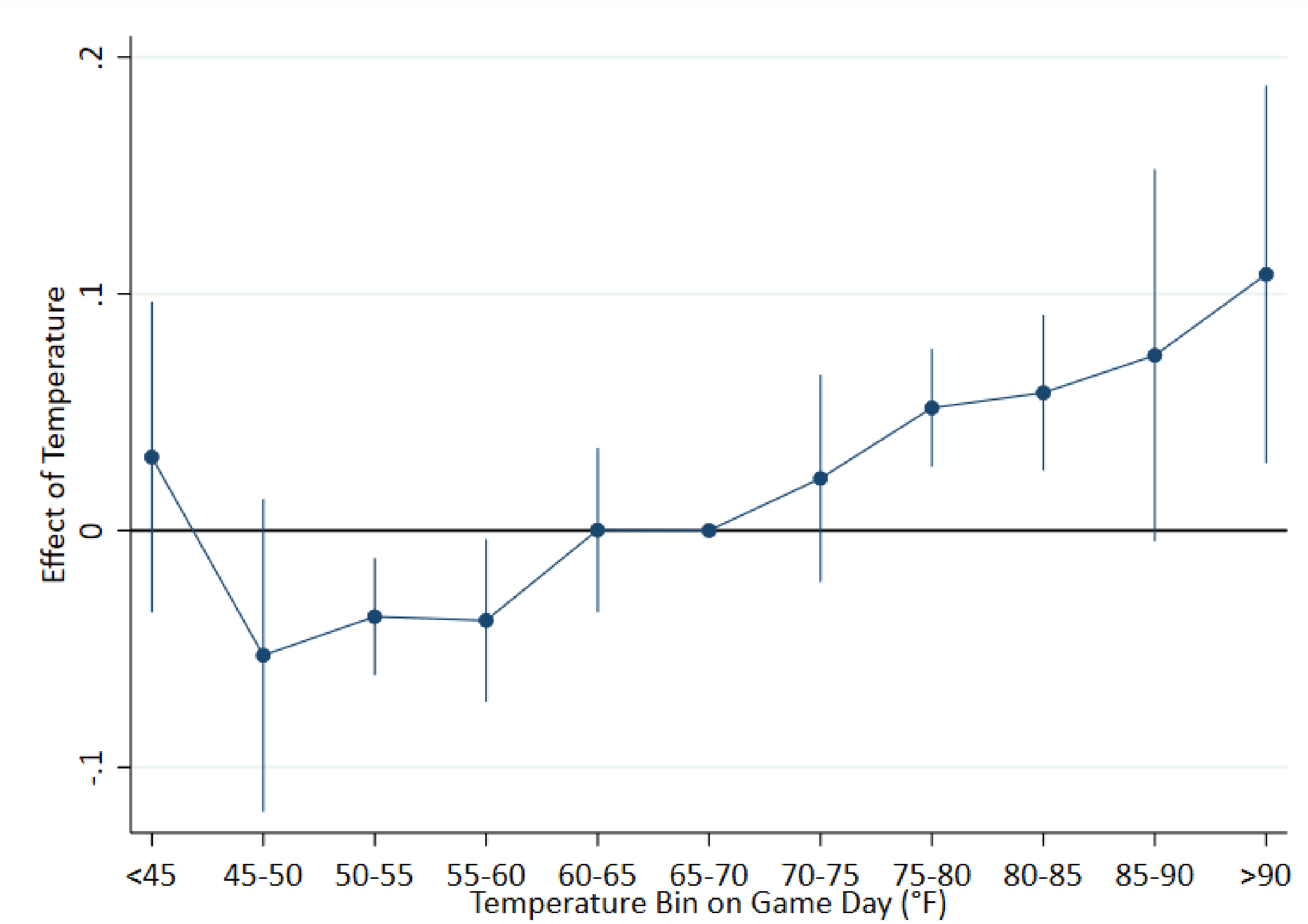
Note: This figure displays the results of Equation 1, run on the sample of MLB games from 1950 to 2000. The outcome variable is game-level log attendance, and the independent variables of interest are indicators for daily average temperature falling into the temperature bin of interest. The regressions also include stadium by month fixed effects, visiting team fixed effects, month by year fixed effects, and controls for daily precipitation, the share of the last 100 games the home team won, indicators for day of the week, and an indicator for whether the game took place in the day or evening. Standard errors are clustered by stadium. Point estimates and 95% confidence intervals are shown.

Figure 4: Effects of Heat on Attendance Larger for Afternoon Games



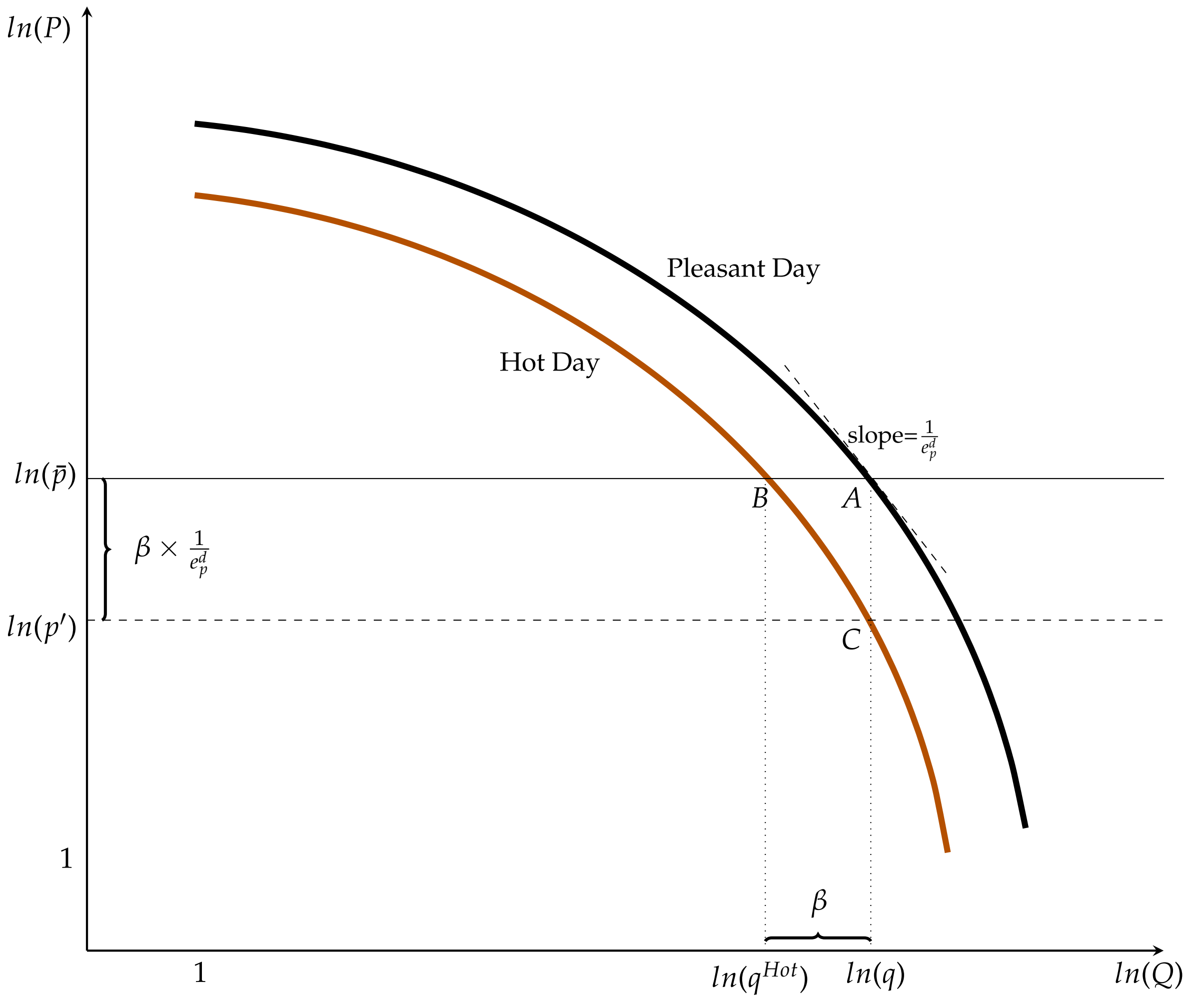
Note: This figure displays the results of Equation 1, run separately for the sample of afternoon games (red) and evening games (blue). The outcome variable is game-level log attendance, and the independent variables of interest are indicators for daily average temperature falling into the temperature bin of interest. The regressions also include stadium by month fixed effects, visiting team fixed effects, month by year fixed effects, and controls for daily precipitation, the share of the last 100 games the home team won, and indicators for day of the week. Standard errors are clustered by stadium. Point estimates and 95% confidence intervals are shown.

Figure 5: PLACEBO EXERCISE: ATTENDANCE RISES ON HOT DAYS AT COVERED STADIUMS



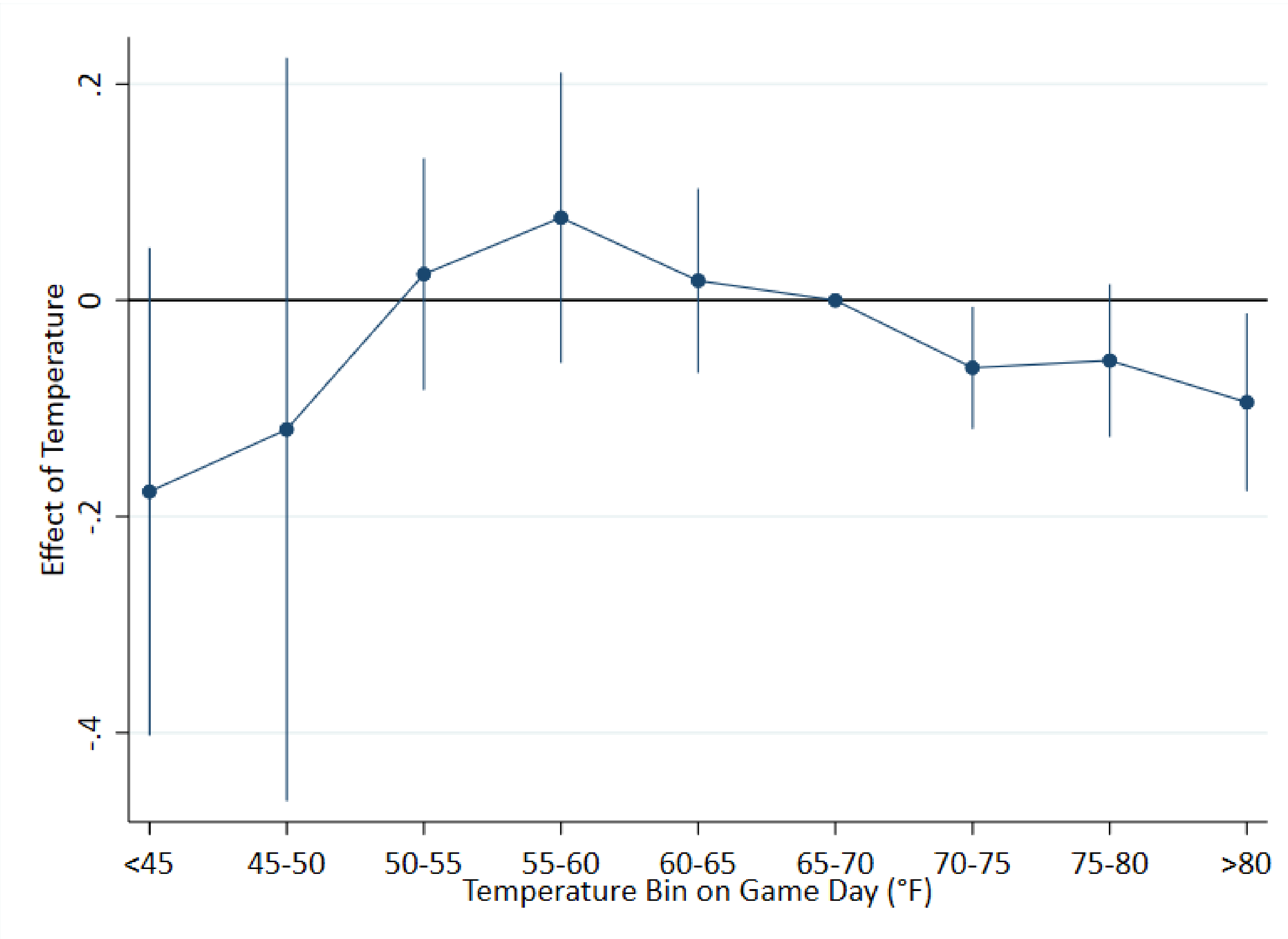
Note: This figure displays the results of Equation 1, run on the sample of MLB games from 1950 to 2000 that took place in stadiums that are either domed or have retractable roofs. The outcome variable is game-level log attendance, and the independent variables of interest are indicators for daily average temperature falling into the temperature bin of interest. The regression also includes stadium by month fixed effects, visiting team fixed effects, month by year fixed effects, and controls for daily precipitation, the share of the last 100 games the home team won, indicators for day of the week, and an indicator for whether the game took place in the day or evening. Standard errors are clustered by stadium. Point estimates and 95% confidence intervals are shown.

Figure 6: ELASTICITY DIAGRAM



Note: Assume that the supply curve for tickets at the game level is perfectly elastic, i.e., that ticket prices are fixed at the game level (but may vary at the team, season, or even team-month-year level). Then the β we estimate as attendance falls from $ln(q) \rightarrow ln(q^{Hot})$ is demand-driven, as pictured. To compute the price-increase-equivalent, we need to multiply the demand change by the slope of the demand curve in the log-log space, i.e., $\frac{\partial ln(P)}{\partial ln(Q)}$. This is equal to the inverse of the price elasticity of demand, $\frac{1}{e_p^d}$. Most of the estimates from the sports literature are well below 1. This implies we should multiply β by *at least* 1 to obtain the price-increase-equivalent, hence the lower-bound framing of using β directly as a WtP estimate.

Figure 7: Seatgeek average ticket prices fall at hot temperatures



Note: This figure displays the results of Equation 7, run on the sample of game-day Seatgeek listing prices from May-October 2021. The outcome variable is game-level log average listing price, and the independent variables of interest are indicators for daily average temperature falling into the temperature bin of interest. The regressions also include stadium by month fixed effects, away team fixed effects, controls for daily precipitation, indicators for day of the week, and indicators for whether the game took place during the day or evening. Standard errors are clustered by stadium and calculated using the wild cluster bootstrap procedure from Cameron et al. (2008). Point estimates and 95% confidence intervals are shown.

Table 1: SUMMARY STATISTICS: GAME AND STADIUM METRICS

	American League			National League		
	Mean	SD	N	Mean	SD	N
Attendance (Thousands)	20.17	12.43	42,403	22.31	13.09	40,063
Capacity (Thousands)	49.84	11.84	42,403	49.15	10.45	40,048
Percent Full	0.42	0.27	42,403	0.46	0.26	40,048
Stadium Age (Years)[1]	31.51	24.77	42,403	25.25	19.93	40,063
Game Duration (Minutes)	163.11	28.66	42,403	159.20	27.91	40,063
Night	0.64	0.48	42,403	0.61	0.49	40,063
Double Header[2]	0.08	0.27	42,403	0.07	0.25	40,063

[1]Calculated as the difference between the year a game takes place and the first year of stadium operation.

[2]Share of games that were played on days when more than one game occurred, usually due to prior inclement weather causing a game postponement.

Note: This table presents summary statistics for stadium and game-related control variables by league.

Table 2: Main Results: Baseball Attendance Falls Sharply at Extremely Hot and Cold Temperatures

	(1)	(2)	(3)	(4)	(5)
<45	-0.163	-0.118	-0.127	-0.097	-0.184
	(0.030)	(0.029)	(0.031)	(0.020)	(0.033)
45-50	-0.218	-0.201	-0.204	-0.159	-0.218
	(0.025)	(0.024)	(0.026)	(0.021)	(0.027)
50-55	-0.174	-0.163	-0.171	-0.129	-0.180
	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)
55-60	-0.106	-0.102	-0.108	-0.083	-0.115
	(0.012)	(0.013)	(0.012)	(0.010)	(0.011)
60-65	-0.038	-0.034	-0.042	-0.033	-0.044
	(0.012)	(0.012)	(0.009)	(0.008)	(0.010)
70-75	0.020	0.020	0.019	0.012	0.020
	(0.009)	(0.008)	(0.007)	(0.006)	(0.008)
75-80	0.021	0.028	0.026	0.016	0.029
	(0.008)	(0.008)	(0.007)	(0.006)	(0.007)
80-85	0.013	0.011	0.003	0.002	0.005
	(0.013)	(0.012)	(0.011)	(0.008)	(0.011)
85-90	0.017	-0.004	0.006	-0.003	0.009
	(0.018)	(0.020)	(0.022)	(0.015)	(0.021)
>90	-0.102	-0.134	-0.140	-0.102	-0.129
	(0.047)	(0.042)	(0.031)	(0.031)	(0.029)
Fixed Effects	Yes	Yes	Yes	Yes	Yes
Game Day Controls	No	Yes	Yes	Yes	Yes
Team Controls	No	No	Yes	Yes	Yes
Years	Pre-2000	Pre-2000	Pre-2000	1950-2019	Pre-2000
Exclude Sellouts	No	No	No	No	Yes
Obs.	75,787	75,787	74,034	108,977	70,129

Standard errors in parentheses

Note: This table displays the results of Equation 1, run on the sample of MLB games from 1950 to 2000. The outcome variable is game-level log attendance, and the independent variables of interest are indicators for daily average temperature falling into the temperature bin of interest. The regression in the first column includes stadium by month and month by year fixed effects. Column 2 adds controls for daily precipitation, indicators for day of the week, and an indicator for whether the game took place in the afternoon or evening. Column 3 adds visiting team fixed effects and controls for the share of the last 100 games the home team won. Columns 4 and 5 display the same specification as column 3, except column 4 contains the sample from 1950-2019, and column 5 excludes games where attendance was greater than 95 percent of stadium capacity. All standard errors are clustered by stadium.

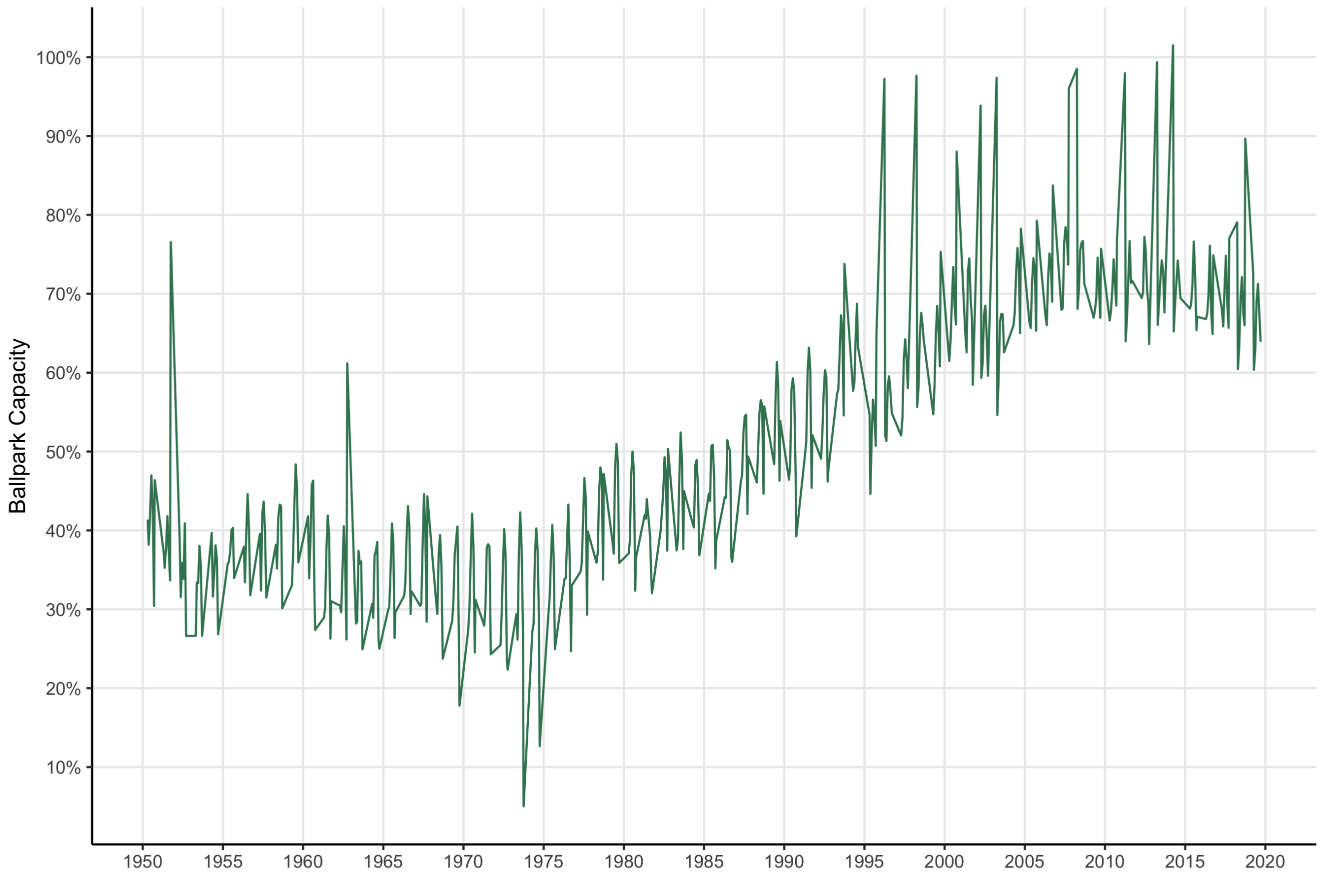
Table 3: MAIN RESULTS: TICKET PRICES

	(1) Avg. Price	(2) Min Price	(3) Max Price	(4) Median Price	(5) Listings	(6) No Sellouts
<45	-0.177	-0.295	0.545	-0.189	0.191	-0.280
	(0.07)	(0.17)	(0.44)	(0.10)	(0.17)	(0.15)
	[-0.611, 0.030]	[-0.648, 0.183]	[-0.667, 3.448]	[-0.493, 0.052]	[-0.122, 0.481]	[-0.671, 0.087]
45-50	-0.119	-0.156	-0.009	-0.197	0.112	-0.111
	(0.57)	(0.35)	(0.98)	(0.16)	(0.55)	(0.57)
	[-0.558, 0.254]	[-0.487, 0.132]	[-0.836, 0.988]	[-0.513, 0.056]	[-0.382, 0.573]	[-0.565, 0.301]
50-55	0.024	0.076	0.035	0.037	-0.286	0.094
	(0.62)	(0.26)	(0.84)	(0.52)	(0.06)	(0.13)
	[-0.076, 0.137]	[-0.065, 0.212]	[-0.293, 0.359]	[-0.080, 0.181]	[-0.587, 0.006]	[-0.038, 0.218]
55-60	0.076	0.107	0.131	0.053	-0.120	0.084
	(0.35)	(0.19)	(0.30)	(0.67)	(0.45)	(0.26)
	[-0.060, 0.232]	[-0.046, 0.283]	[-0.129, 0.396]	[-0.120, 0.272]	[-0.460, 0.169]	[-0.051, 0.236]
60-65	0.018	0.076	0.035	0.023	-0.048	0.027
	(0.72)	(0.06)	(0.80)	(0.59)	(0.76)	(0.59)
	[-0.067, 0.114]	[-0.002, 0.158]	[-0.229, 0.308]	[-0.066, 0.117]	[-0.267, 0.204]	[-0.072, 0.129]
70-75	-0.062	-0.089	0.056	-0.067	0.118	-0.050
	(0.03)	(0.01)	(0.53)	(0.03)	(0.09)	(0.08)
	[-0.120, -0.008]	[-0.156, -0.023]	[-0.134, 0.247]	[-0.123, -0.011]	[-0.016, 0.265]	[-0.114, 0.008]
75-80	-0.056	-0.119	0.133	-0.076	0.252	-0.044
	(0.13)	(0.01)	(0.29)	(0.10)	(0.01)	(0.30)
	[-0.131, 0.019]	[-0.199, -0.038]	[-0.110, 0.389]	[-0.167, 0.016]	[0.088, 0.420]	[-0.123, 0.038]
80-85	-0.094	-0.139	0.087	-0.114	0.225	-0.065
	(0.03)	(0.01)	(0.52)	(0.03)	(0.05)	(0.14)
	[-0.184, -0.008]	[-0.243, -0.032]	[-0.202, 0.383]	[-0.223, -0.011]	[0.004, 0.460]	[-0.162, 0.020]
N	1575	1575	1575	1575	1575	1342

Note: This table displays the results of Equation 7, run on the sample of game-day Seatgeek listing prices from May-October 2021. The outcome variables are game-day log average listing price (columns 1 and 6), log minimum price (column 2), log maximum price (column 3), log median price (column 4), and logged listings count (column 5). The independent variables of interest are indicators for daily average temperature falling into the temperature bin of interest. The regressions also include stadium by month fixed effects, away team fixed effects, and controls for daily precipitation, indicators for day of the week, and indicators for whether the game took place during the day or evening. Standard errors are clustered by stadium and calculated using the wild cluster bootstrap procedure from Cameron et al. (2008). The table shows the coefficient, p value, and 95% confidence interval for each estimate.

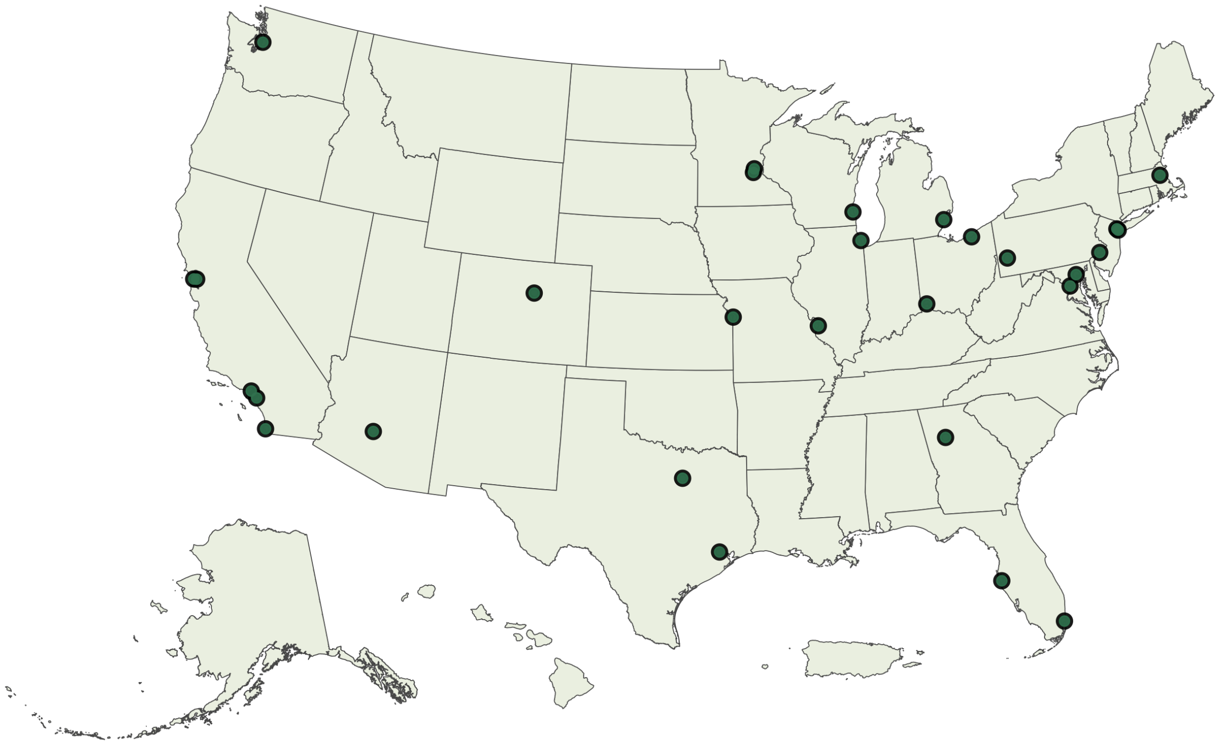
A Appendix A

Figure A1: SUMMARY STATISTICS: STADIUMS ARE MORE FULL THAN IN PAST



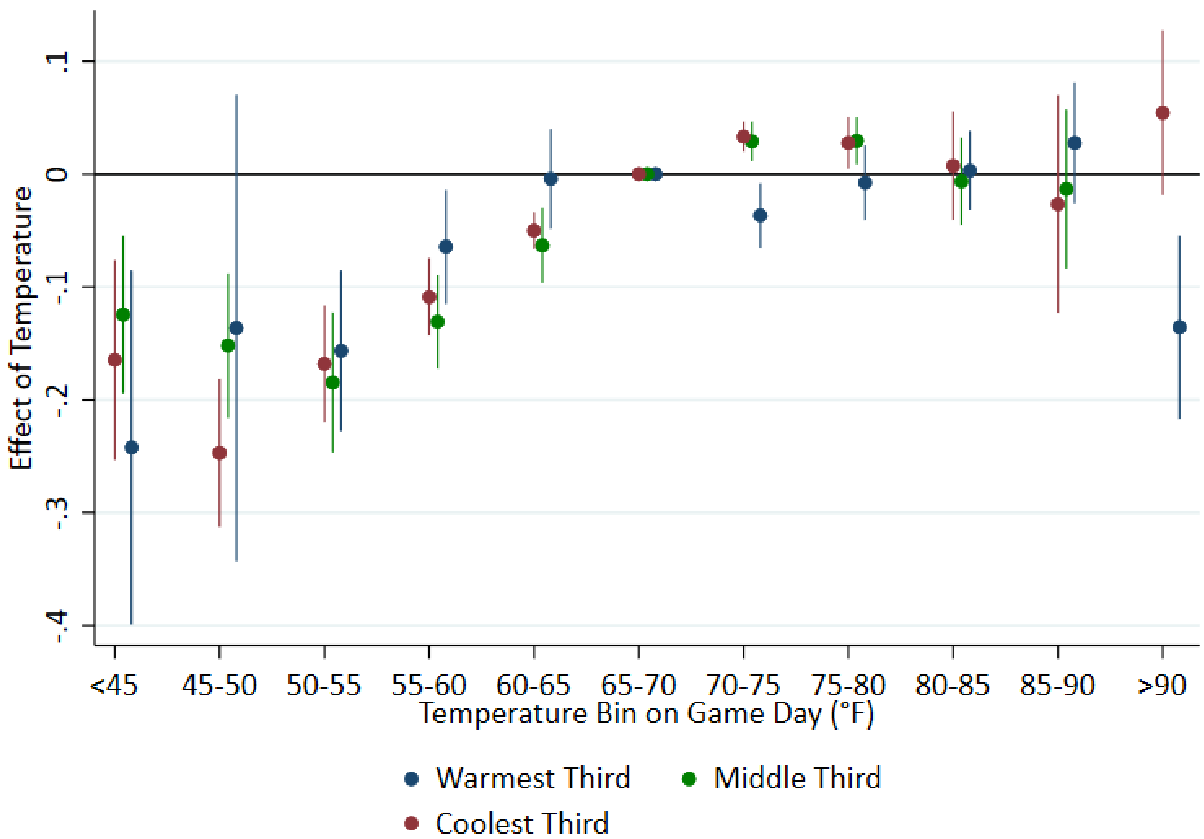
Note: This figure displays attendance as a percent of baseball stadium capacity from 1950 through 2019 using data from `Retrosheet.org`. The large spikes are generally associated with the first game of each season, which is the most well-attended game of the year for most teams.

Figure A2: CITIES WITH MLB STADIUMS



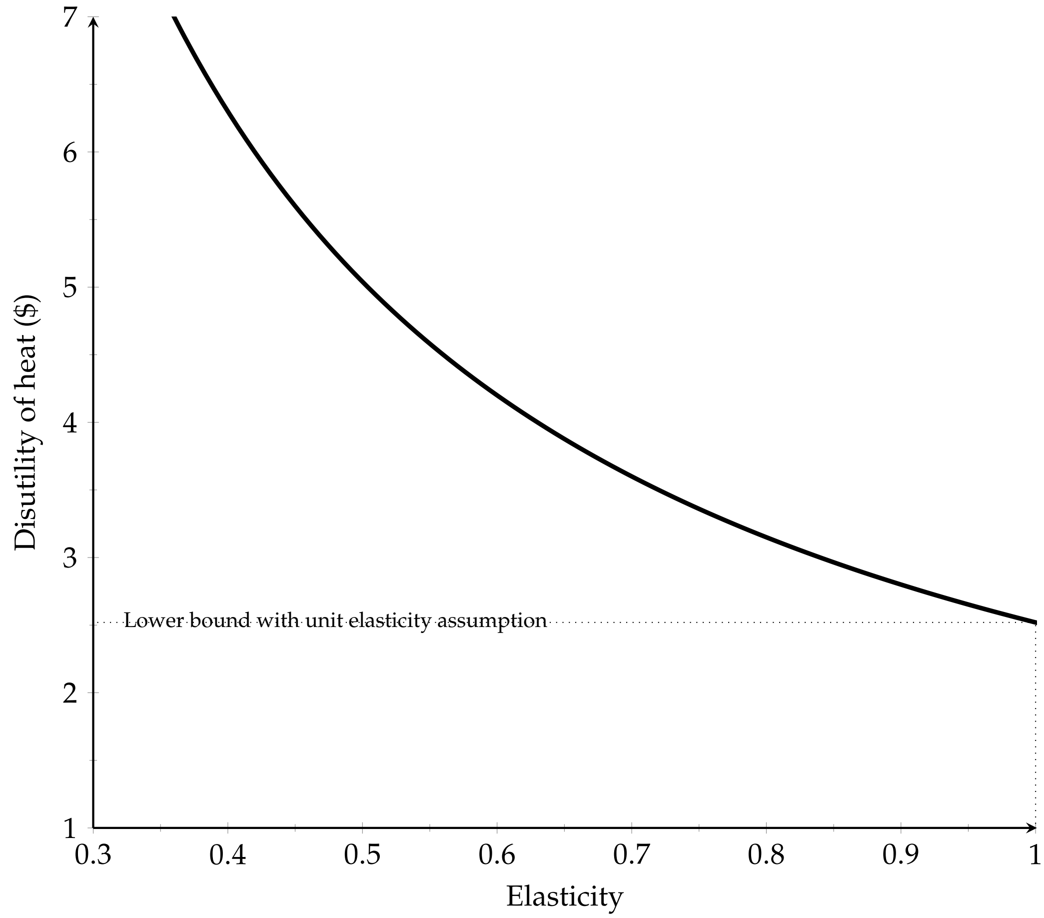
Note: This figure displays the location of the MLB stadiums in our sample. Some large cities, such as New York City and Chicago, have multiple stadiums operating at the same time. For a complete list of teams in the sample and their years of operation, see Table A1.

Figure A3: No Evidence of Systematic Differences by Usual Climate



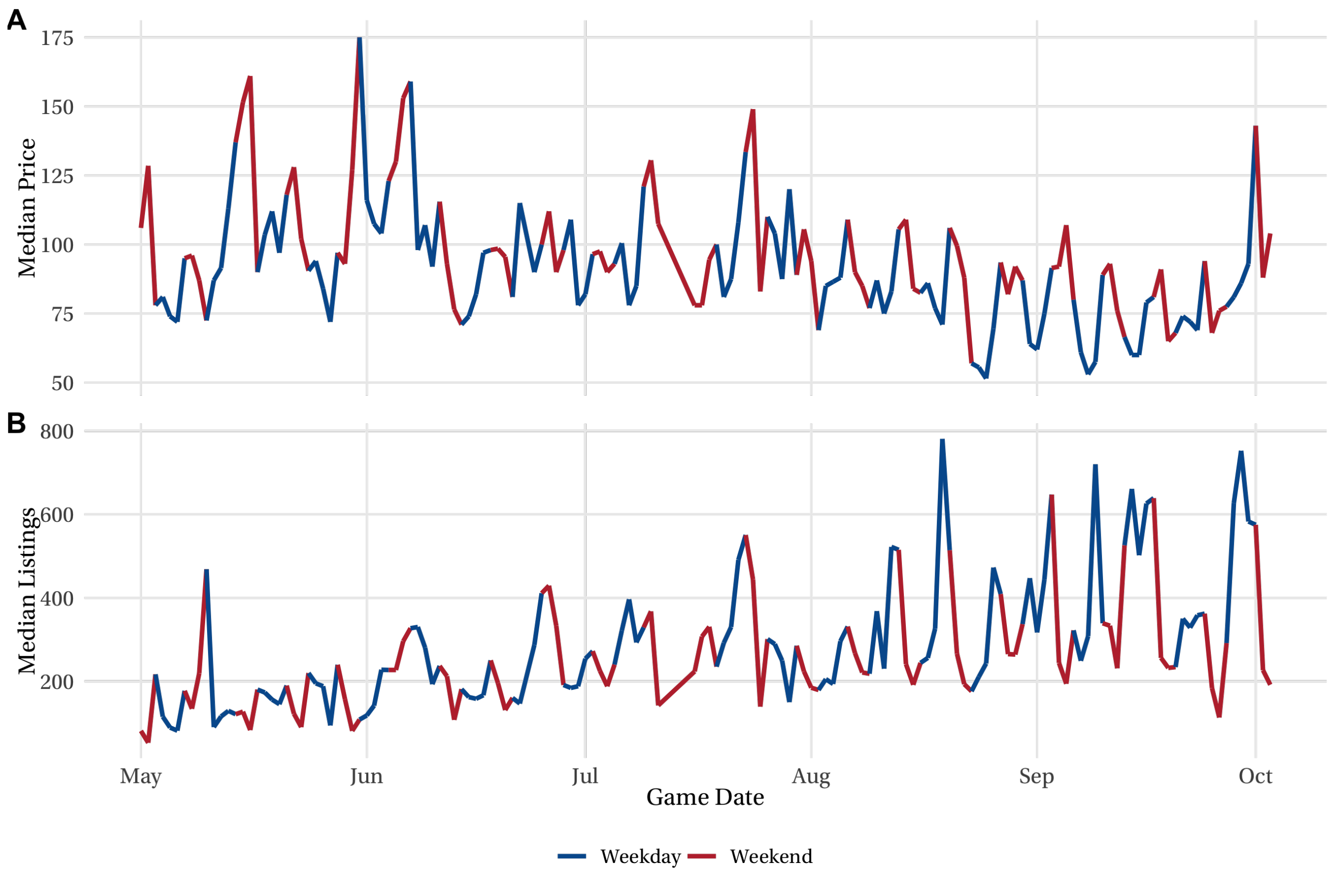
Note: This figure displays the results of Equation 1, run on the sample of MLB games from 1950 to 2000. The outcome variable is game-level log attendance, and the independent variables of interest are indicators for daily average temperature falling into the temperature bin of interest. These indicators are interacted flexibly with indicators for the stadium being in the warmest, coolest, or medium third of stadiums in the sample based on annual average temperature throughout the sample period. The regressions also include stadium by month fixed effects, visiting team fixed effects, month by year fixed effects, and controls for daily precipitation, the share of the last 100 games the home team won, indicators for day of the week, and an indicator for whether the game took place in the day or evening. Standard errors are clustered by stadium. Point estimates and 95% confidence intervals are shown.

Figure A4: DISUTILITY OF HEAT INCREASES WITH (IN)ELASTICITY



Note: Estimate of the disutility of heat at a baseball game as it changes with assumed price elasticity of demand. Baseline results use a unit elasticity. Krautmann and Berri (2007) surveys a literature documenting that MLB teams price on the inelastic range of demand. If so, the implied monetary willingness to pay is larger than the $2.52 in the baseline estimate.

Figure A5: SEATGEEK SUMMARY STATISTICS: MEDIAN PRICE AND LISTINGS COUNT



Note: This figure displays median prices (Panel A) and median listing counts (Panel B) on Seatgeek from May through October of 2021, averaged by day of the season.

Table A1: Teams Included in Sample

	N. Home Games	First Year	Last Year
West			
Arizona Diamondbacks	1,780	1998	-
Colorado Rockies	2,146	1993	-
Los Angeles Angels	4,636	1961	-
Los Angeles Dodgers	4,872	1958	-
Oakland Athletics	4,036	1968	-
San Diego Padres	3,976	1969	-
San Francisco Giants	4,786	1958	-
Seattle Mariners	3,403	1977	-
Seattle Pilots	74	1969	1969
South			
Atlanta Braves	4,173	1966	-
Baltimore Orioles	4,925	1954	-
Houston Astros	4,585	1962	-
Miami Marlins	2,131	1993	-
Tampa Bay Devil Rays	1,775	1998	-
Texas Rangers	3,754	1972	-
Washington Nationals	1,207	2005	-
Washington Senators	734	1950	1960
Washington Senators	778	1961	1971
Northeast			
Boston Braves	193	1950	1952
Boston Red Sox	5,320	1950	-
Brooklyn Dodgers	547	1950	1957
New York Giants	520	1950	1957
New York Mets	4,341	1962	-
New York Yankees	5,172	1950	-
Philadelphia Athletics	304	1950	1954
Philadelphia Phillies	5,197	1950	-
Pittsburgh Pirates	5,200	1950	-
Midwest			
Chicago Cubs	5,238	1950	-
Chicago White Sox	5,137	1950	-
Cincinnati Reds	5,255	1950	-
Cleveland Guardians	5,122	1950	-
Detroit Tigers	5,283	1950	-
Kansas City Athletics	942	1955	1967
Kansas City Royals	3,948	1969	-
Milwaukee Braves	926	1953	1965
Milwaukee Brewers	3,864	1970	-
Minnesota Twins	4,609	1961	-
St. Louis Browns	251	1950	-
St. Louis Cardinals	5,383	1950	-

Table A2: SEATGEEK PRICE AND LISTING SUMMARY STATISTICS

	Game Day			One Day Before			Two Days Before		
	Mean	Median	N	Mean	Median	N	Mean	Median	N
Average Price	445.70	90.50	1,992	278.23	94.00	1,991	264.25	98.00	1,980
Median Price	79.20	69.00	1,992	79.91	71.00	1,991	81.13	73.00	1,980
Lowest Price	32.52	25.00	1,992	30.18	24.00	1,991	30.36	24.50	1,980
Highest Price	2,626.66	466.50	1,992	2,755.85	590.00	1,991	3,232.35	683.50	1,980
N. Listings	349.99	235.00	1,992	488.42	360.00	1,991	537.72	402.50	1,980

Note: This table presents summary statistics for SeatGeek ticket listing data from the 2021 MLB season. Data was accessed using SeatGeek's API beginning on May 1, 2021 and ending on October 3.

Table A3: BASEBALL GAME QUALITY AND TEMPERATURE

	(1) Duration	(2) Runs	(3) Home Runs	(4) Strike Outs	(5) Pitchers
<45	0.336	-1.129	-0.653	0.354	-0.223
	(1.192)	(0.281)	(0.115)	(0.176)	(0.083)
45-50	-0.499	-0.724	-0.501	0.195	-0.151
	(0.888)	(0.199)	(0.081)	(0.135)	(0.067)
50-55	-0.351	-0.534	-0.342	0.196	-0.086
	(0.483)	(0.147)	(0.050)	(0.092)	(0.035)
55-60	-0.793	-0.508	-0.249	0.159	-0.139
	(0.447)	(0.110)	(0.043)	(0.073)	(0.039)
60-65	-0.656	-0.206	-0.098	0.155	-0.061
	(0.404)	(0.079)	(0.024)	(0.056)	(0.027)
70-75	0.844	0.255	0.105	-0.202	0.069
	(0.373)	(0.081)	(0.029)	(0.041)	(0.032)
75-80	1.401	0.437	0.194	-0.366	0.145
	(0.470)	(0.110)	(0.036)	(0.067)	(0.037)
80-85	1.787	0.595	0.282	-0.505	0.197
	(0.518)	(0.142)	(0.046)	(0.079)	(0.044)
85-90	1.898	0.746	0.416	-0.633	0.247
	(0.919)	(0.216)	(0.078)	(0.123)	(0.087)
>90	1.639	1.208	0.228	-1.056	0.546
	(1.733)	(0.311)	(0.081)	(0.287)	(0.133)
Obs.	74,034	74,034	74,034	74,034	74,034
Mean of Dep.	161	9	2	11	5

Standard errors in parentheses

Note: This figure displays the results of Equation 1, run on the sample of MLB games from 1950 to 2000. The dependent variables are indicators of the quality of game play, and the independent variables of interest are indicators for daily average temperature falling into the temperature bin of interest. The regressions also include stadium by month fixed effects, visiting team fixed effects, month by year fixed effects, and controls for daily precipitation, the share of the last 100 games the home team won, indicators for day of the week, and an indicator for whether the game took place in the day or evening. Standard errors are clustered by stadium.

B Data Appendix

In this Appendix, we provide more information on the construction and contents of the weather and Seatgeek ticket price data used in the main analysis.

PRISM Data For the analysis of the effects of weather on Seatgeek ticket prices, we use the PRISM weather dataset, developed by Oregon State University's PRISM Climate Group. These data are compiled from monitoring stations across the U.S. and are spatially interpolated to a 4km resolution using the PRISM model, which accounts especially for the effects of elevation on the spatial distribution of temperature and precipitation. The dataset provides information on daily maximum and minimum temperatures for each gridpoint, which we average to approximate the daily average temperature. It also contains daily total precipitation, which is the sum of rain and melted snow for the day. We download PRISM data using the prism package in R.

Schlenker's Fine-Scaled Weather Data Set For the analysis of the effects of weather on attendance, we source our weather data from Wolfram Schlenker's Daily Weather for Contiguous United States dataset, which is derived from the PRISM dataset. This dataset provides daily minimum and maximum temperature temperature and total precipitation at the 2.5 mile resolution. This dataset creates a balanced panel of monitoring stations by interpolating missing station observations with the distance-weighted average of the cumulative density function of surrounding stations. This dataset is available from 1900-2019.

Weather Projections from CMIP6 We use the average of the 35 global climate models (GCM's) from the Coupled Model Intercomparison Model 6 (CMIP6). The data are downloaded from the NASA Earth Exchange (NEX) Global Daily Downscaled Projections (GDDP). The NEX-GDDP is downscaled to improve the resolution of output from the GCM's, allowing for analysis at the 0.25 degree latitude/longitude resolution. Daily maximum and minimum near surface air temperatures are provided at the daily level, which we average to form a measure of daily average temperature. We then calculate a count of days above 90 degrees for each grid cell for each year from 2080-2090, and then create a decadal average of the yearly count. We calculate these averages for two climate change scenarios from the new scenario framework introduced in the IPCC Sixth Assessment Report. SSP5-8.5 is a high-emissions scenario with no mitigation policy, in which economic growth is fueled by fossil fuels. This scenario produces warming of 4.4 degrees Celsius by end of century. SSP2-4.5 is a middle-of-the-road scenario in which emissions begin to fall by the mid 21st century, and socioeconomic factors are stable. This scenario produces warming of 2.7 degrees Celsius by 2100.

Retrosheet Data We download data on game-level attendance and game play from retrosheet.org using the retrosheet package in R. We make a few adjustments to the retrosheet data to derive a consistent set of team-stadium observations. First, we code any team that changed its name but remained in the same city to be a single team.

- The Florida Marlins became the Miami Marlins; we code them as the same team.
- The California Angels became the Anaheim Angels in 1996 and subsequently the Los Angeles Angels of Anaheim in 2005; we code them as the same team.

If a team moved cities and changed names, we consider the team in the new city to be a new team.

- The Kansas City Athletics moved to Oakland and became the Oakland Athletics in 1968; we code these as different teams.

- The Washington Senators moved to Minnesota and became the Minnesota Twins in 1961. A new Washington Senators team sprung up in its place, from 1960-1971, before moving to Dallas-Fort Worth and becoming the Texas Rangers in 1972. The original Washington Senators, replacement Washington Senators, Minnesota Twins, and Texas Rangers are each coded as different teams in our sample.
- The Seattle Pilots moved to Milwaukee and became the Milwaukee Brewers in 1970; they are coded as different teams.

We combine a list of stadiums from Retrosheet with park configuration data from Seamheads. Seamheads.com provides data on capacity and type of cover for each stadium by year. We use this information in combination with Retrosheet to form our measure of attendance as a share of capacity as well as for our sample restriction to only open-air stadiums. We drop some one-off venues from our dataset: the Oakland Athletics played 6 games in Las Vegas in 1996, the Atlanta Braves played 1 game at Fort Bragg in 2016, the Tampa Bay Rays played 3 games at the Ballpark at Disney's Wide World in 2007, and the Pittsburgh Pirates played 1 game at BB&T Ballpark at Bowman Field in 2017. Our results are robust to excluding or including these venues from the sample.

Seatgeek Data We accessed the SeatGeek.com API on a daily basis at 8:00A.M. Central Time starting on May 1, 2021 and ending on October 3, 2021, the final day of the 2021 regular season. Each access of the API collected data on that day and the following two days' MLB games. The variables collected include:

- Home and away teams
- Date of the API request
- Game date
- Game start time
- Geographic information, including the city, state, latitude, and longitude of where the game was to take place
- The number of ticket listings
- Mean, median, minimum, and maximum ticket price for seats available during the game

The data are at the game-date level, not the individual ticket level. It should also be noted that the number of ticket listings does not equal the number of tickets available; for example, a listing for four seats next to each other would be counted as one listing.