

Alexander, Ahammer; Martin, Halla; Lackner, Mario

**Working Paper**

## Mass gatherings contributed to early COVID-19 spread: Evidence from US sports

Working Paper, No. 2003

**Provided in Cooperation with:**

Christian Doppler Laboratory Aging, Health and the Labor Market, Johannes Kepler University Linz

*Suggested Citation:* Alexander, Ahammer; Martin, Halla; Lackner, Mario (2020) : Mass gatherings contributed to early COVID-19 spread: Evidence from US sports, Working Paper, No. 2003, Johannes Kepler Universität Linz, Christian Doppler Laboratory Aging, Health and the Labor Market, Linz

This Version is available at:

<https://hdl.handle.net/10419/222442>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# **Mass Gatherings Contributed to Early COVID-19 Spread: Evidence from US Sports**

by

Alexander AHAMMER

Martin HALLA

Mario LACKNER

June 2020

Corresponding author: [alexander.ahammer@jku.at](mailto:alexander.ahammer@jku.at)

---

**Christian Doppler Laboratory**  
**Aging, Health and the Labor Market**  
[cdecon.jku.at](http://cdecon.jku.at)

**Johannes Kepler University**  
**Department of Economics**  
**Altenberger Strasse 69**  
**4040 Linz, Austria**

# Mass Gatherings Contributed to Early COVID-19 Spread: Evidence from US Sports\*

Alexander Ahammer<sup>a,b</sup>, Martin Halla<sup>a,b,c,d</sup>, Mario Lackner<sup>a,b</sup>

<sup>a</sup>*Johannes Kepler University, Linz*

<sup>b</sup>*Christian Doppler Laboratory on Ageing, Health, and the Labor Market, Linz*

<sup>c</sup>*IZA, Institute for the Study of Labor, Bonn*

<sup>d</sup>*GÖG, Austrian Public Health Institute, Vienna*

June 16, 2020

## Abstract

Social distancing is important to slow the community spread of infectious disease, but it creates enormous economic and social cost. It is thus important to quantify the benefits of different measures. We study the ban of mass gatherings, an intervention with comparably low cost. We exploit exogenous spatial and temporal variation in NBA and NHL games, which arise due to the leagues' predetermined schedules, and the suspension of the 2019-20 seasons. This allows us to estimate the impact of these mass gatherings on the spread of COVID-19 in affected US counties. One additional mass gathering increased the cumulative number of COVID-19 deaths in affected counties by 11 percent.

*JEL Classification:* I18, H12, I10.

*Keywords:* Social distancing, mass gatherings, Coronavirus Disease 2019, COVID-19.

---

\**Corresponding author:* Mario Lackner, Department of Economics, Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria; e-mail: [mario.lackner@jku.at](mailto:mario.lackner@jku.at). Financial support from the Christian Doppler Laboratory "Ageing, Health and the Labor Market" is gratefully acknowledged.

# 1. Introduction

Due to the lack of vaccines and effective antiviral drugs, countries have to rely on a set of non-pharmaceutical interventions (NPI) in response to the *Coronavirus Disease 2019* (COVID-19) pandemic. The goal of these measures is to prevent a sharp peak of infections, to take pressure off healthcare systems, and ultimately to save lives. In addition to good personal hygiene and mandatory face masks, social distancing is perhaps the most important NPI (Cowling & Aiello 2020). Since maintaining physical distance inevitably creates enormous economic and social cost, it is crucial to quantify the benefits of different measures in controlling epidemics.

One important public policy to promote physical distancing is to ban mass gatherings (Memish et al. 2019).<sup>1</sup> Such events may foster the transmission of contagious disease as a result of large crowds being in close contact, often for extended periods of time. A temporary mass gathering ban is relatively cheap and easy to implement compared to, for example, school and workplace closures. In response to the spread of SARS-CoV-2, the pathogen leading to COVID-19, several prominent events have been canceled or postponed, even before widespread quarantine measures were enacted (McCloskey et al. 2020). These include religious, cultural, and sporting events.

We quantify how much *National Basketball Association* (NBA) and *National Hockey League* (NHL) games have contributed to the spread of COVID-19 in the United States.<sup>2</sup> Before the leagues suspended play on March 12, up to 12 games per league with an average audience of about 18,000 people were held per day. We analyze how much the number of games held between March 1 and March 11 has contributed to the community spread of COVID-19 in counties surrounding NBA and NHL venues. Since the game schedules were determined long before the first COVID-19 case became public, their spatial and temporal distribution should be unrelated to the initial spread of COVID-19 in the US.<sup>3</sup> In fact, we can show that game schedules are not correlated with observable county characteristics and that game attendance did not systematically change until the NBA and NHL suspended play.

Our results suggest that one additional mass gathering between March 1 and 11 in the form of a NBA or NHL game increased the cumulative number of COVID-19 cases (measured on April 30, 2020) in affected counties by at least 379 per one million population ( $p < 0.05$ ) or 13 percent, and the number of COVID-19 deaths per million by 16 ( $p < 0.05$ ) or 11 percent. These effects are larger in colder regions and in states where shelter-in-place orders (SIPOs) were implemented late. We conclude that banning mass gatherings has an enormous potential to save lives, which is especially important given that such measures are relatively easy and cheap to implement.

---

<sup>1</sup>The World Health Organization (WHO) describes a mass gathering as “a planned or spontaneous event where the number of people attending could strain the planning and response resources of the community or country hosting the event. The Olympic Games, The Hajj, and other major sporting, religious, and cultural events are all examples of a mass gathering.”

<sup>2</sup>We focus on NBA and NHL because their seasons were ongoing when COVID-19 broke out. The *National Football League* (NFL) was in offseason and *Major League Baseball* (MLB) in spring training, which involves scrimmage games in smaller ballparks held in Arizona and Florida.

<sup>3</sup>The first known case in the US was a man in Washington State who returned January 15, 2020 from Wuhan. The NBA 2019/20 schedule had been released on August 12, 2019, the NHL schedule on June 25, 2019.

Our results contribute to the literature evaluating the role of mass gatherings in the spread of infectious disease, and the benefits of social distancing more generally. In a recent survey of the literature, [Nunan & Brassey \(2020\)](#) conclude that the impact of mass gatherings on COVID-19 is still poorly understood. So far, evidence comes almost solely from case reports. For other infectious diseases there is more evidence, but mostly in the form of retrospective observational studies ([Rainey et al. 2016](#), [Hoang & Gautret 2018](#), [Karami et al. 2019](#)). We are not aware of any design-based estimation of the impact of mass gatherings on the *community* spread of infectious disease.<sup>4</sup> The best available evidence suggests multiple-day events with crowded communal accommodations are most associated with increased risk of infection ([Nunan & Brassey 2020](#)).

Other NPIs have received more attention in the context of COVID-19. These studies differ with respect to outcomes, interventions, and geographic coverage. [Gupta et al. \(2020\)](#) demonstrate how different state- and county-level measures that aim at fostering social distancing have affected people's mobility. The authors proxy mobility with cell signal data, and find SIPOs to have the largest mobility-reducing impact. Two studies examine the impact of SIPOs on COVID-19 cases and deaths. [Dave et al. \(2020\)](#) exploit variation in SIPOs across time and all US states. Their results suggest that approximately three weeks following the adoption of a SIPO, cumulative COVID-19 cases fell by 44 percent. [Friedson et al. \(2020\)](#) focus on California, which was the first state to enact a SIPO. Using a synthetic control design, they find that California's SIPO reduced cases by 125.5 per 100,000 population and deaths by 1,661. Methodologically, all papers use a difference-in-differences approach.<sup>5</sup> [Goodman-Bacon & Marcus \(2020\)](#) provide a critical account of this estimation approach in the context of NPIs.

There is extensive evidence on previous pandemics available. However, the majority of these studies are descriptive in nature. Studying the 1918 influenza pandemic, [Markel et al. \(2007\)](#), [Bootsma & Ferguson \(2007\)](#), and [Hatchett et al. \(2007\)](#), for example, find a strong correlation between excess mortality and how early public health measures were enacted in US cities. It is difficult to infer causality from these results, however, because NPIs are not exogenous and may be enacted in response to preexisting trends in death rates. [Barro \(2020\)](#) attempts to account for this endogeneity by using the distance to army ports in Boston as instrumental variables for NPI introduction. He argues that, because the influenza spread from Boston to other US cities, the farther away cities are from Boston, the more time they had to react and implement NPIs. [Barro](#) finds no effect on overall deaths, but that the ratio of peak to average deaths decreased (i.e., a flatter curve). [Chapelle \(2020\)](#) finds a similar pattern using a difference-in-differences model exploiting differences in the timing of NPI introduction. He claims that the lack of herd immunity in subsequent years offset the initial reduction in deaths during the peak of the pandemic, which led to an overall zero effect on deaths.

For recent influenza waves, there is some suggestive evidence that school closures (e.g., [Earn 2012](#), [Wheeler et al. 2010](#)) and workplace social distancing (e.g., [Ahmed et al. 2018](#), [Miyaki et al. 2011](#)) may be associated with lower disease transmission. However, this literature consists mostly of small case

---

<sup>4</sup>A notable contribution is [Mangrum & Niekamp \(2020\)](#), who present evidence that college student travel contributed to the spread of COVID-19. Their estimates show that counties with more early spring break students had higher confirmed case growth rates than counties with fewer early spring break students.

<sup>5</sup>There are also a number of non-US studies. [Fang et al. \(2020\)](#) study the case of Wuhan (China), and [Hsiang et al. \(2020\)](#) study localities within China, France, Iran, Italy, South Korea, and the US.

studies on scheduled school closures (for example, during holidays) or single firms. [Viner et al. \(2020\)](#) conclude that school closures were largely ineffective in controlling *past* Coronavirus outbreaks (i.e., SARS and MERS).

The cost of school and workplace closures are massive. For example, [Sadique et al. \(2008\)](#) estimate that school closures in the US could cost up to £1.2 billion per week. In the early stages of COVID-19, [Alexander & Karger \(2020\)](#) find that people already traveled 9% less and made 13% fewer visits to non-essential businesses. Their preliminary evidence suggests that consumer spending for over 1 million small US business may be reduced by 40%. In a recent survey, respondents reported average wealth losses due to COVID-19 of about \$33,000 ([Coibion et al. 2020](#)). However, [Greenstone & Nigam \(2020\)](#) find that even a moderate form of social distancing (i.e., isolation of suspect cases and their family members and social distancing of the elderly) can reduce COVID-19 fatalities by almost 1.8 million over the next 6 months, amounting to economic benefits of almost \$8 trillion. Similarly, [Thunström et al. \(forthcoming\)](#) estimate the potential benefits of social distancing at around \$5.2 trillion.

The remainder of the paper is structured as follows. Section 2 describes our data sources. In Section 3, we present our estimation strategy. Sections 4 and 5 report the main results and a heterogeneity analysis. Section 6 provides concluding comments. Additional figures and tables we delegate to a web appendix.

## 2. Data

We use information on NBA and NHL games played between March 1 and March 11.<sup>6</sup> During this time span, 78 NBA games (on average about 7 per day) and 57 NHL games (on average 5 per day) were played in US venues. Both leagues suspended all remaining games for the 2019/20 season indefinitely on March 12. The NBA cancelled two games right before tip-off on March 11: Utah Jazz at Oklahoma City Thunder, where Utah player Rudy Gobert tested positive for Sars-Cov-2 prior to the game, and New Orleans Pelicans at Sacramento Kings, due to a suspected infection involving a referee who was part of the officiating crew in a game involving the Utah Jazz earlier the same week.

In our estimation sample, we focus on 38 counties which host either a NBA or a NHL venue, or both, and all their 204 neighboring counties, which we call the ‘perimeter’ (see Figure A.1 in the Web Appendix for a map). For all affected venue and perimeter counties, we collect information on COVID-19 cases and related deaths.<sup>7</sup> In Figure 1, we show the number of cases (panel a) and deaths (panel b) per million population measured on March 13 (indicated by the left scatter) and on

---

<sup>6</sup>The information on NBA games is scraped from *Basketball Reference* (see [basketball-reference.com](http://basketball-reference.com)). Data on NHL games are collected from *Hockey Reference* (see [hockey-reference.com](http://hockey-reference.com)). Since we focus on US territory, we disregard 16 NHL games played in Canada. No NBA game was played in Toronto during the relevant time span.

<sup>7</sup>The information on COVID-19 cases and deaths up to April 30, 2020 is obtained from a database maintained by *The New York Times*, which collects county-level information from reports of state and local health agencies (see [nytimes.com/interactive/2020/us/coronavirus-us-cases.html](https://nytimes.com/interactive/2020/us/coronavirus-us-cases.html)). In our main analysis, we exclude New York City. Data on COVID-19 cases and deaths are not available for city boroughs separately. Adjacent counties in New Jersey and New York are coded as affected by games in New York City.

TABLE 1 — Descriptive statistics for main variables

	By county type	
	Venue	Perimeter
Total number of games (NBA + NHL) between March 1–11	3.74 (3.06)	3.48 (2.71)
Cumulative number of COVID-19 infections <sup>†</sup>		
On March 13	9.08 (14.18)	4.92 (13.28)
On April 30	3,846.71 (4,425.77)	2,834.96 (4,792.83)
Cumulative number of COVID-19 deaths <sup>†</sup>		
On March 13	0.05 (0.20)	0.02 (0.22)
On April 30	211.30 (329.72)	139.18 (249.98)
Total county population (in mio.)	1.79 (1.79)	0.38 (0.45)
Population density	3,506.84 (3,915.39)	835.78 (1,588.35)
Population characteristics		
% female	51.20 (0.85)	50.57 (1.10)
% non-white pop.	33.50 (13.43)	17.03 (12.89)
% pop. 60+	18.46 (2.52)	21.62 (5.28)
Number of counties	38	204

Notes: Sample means with standard deviations in parentheses.

<sup>†</sup> Per one million population.

April 30 (the right scatter) for each venue county, grouped by state, in our data. Additionally, we compute the average number of cases and deaths across each set of neighboring counties. The highest increases are in Essex County, NJ; Orleans Parish, LA; and Suffolk County, MA.

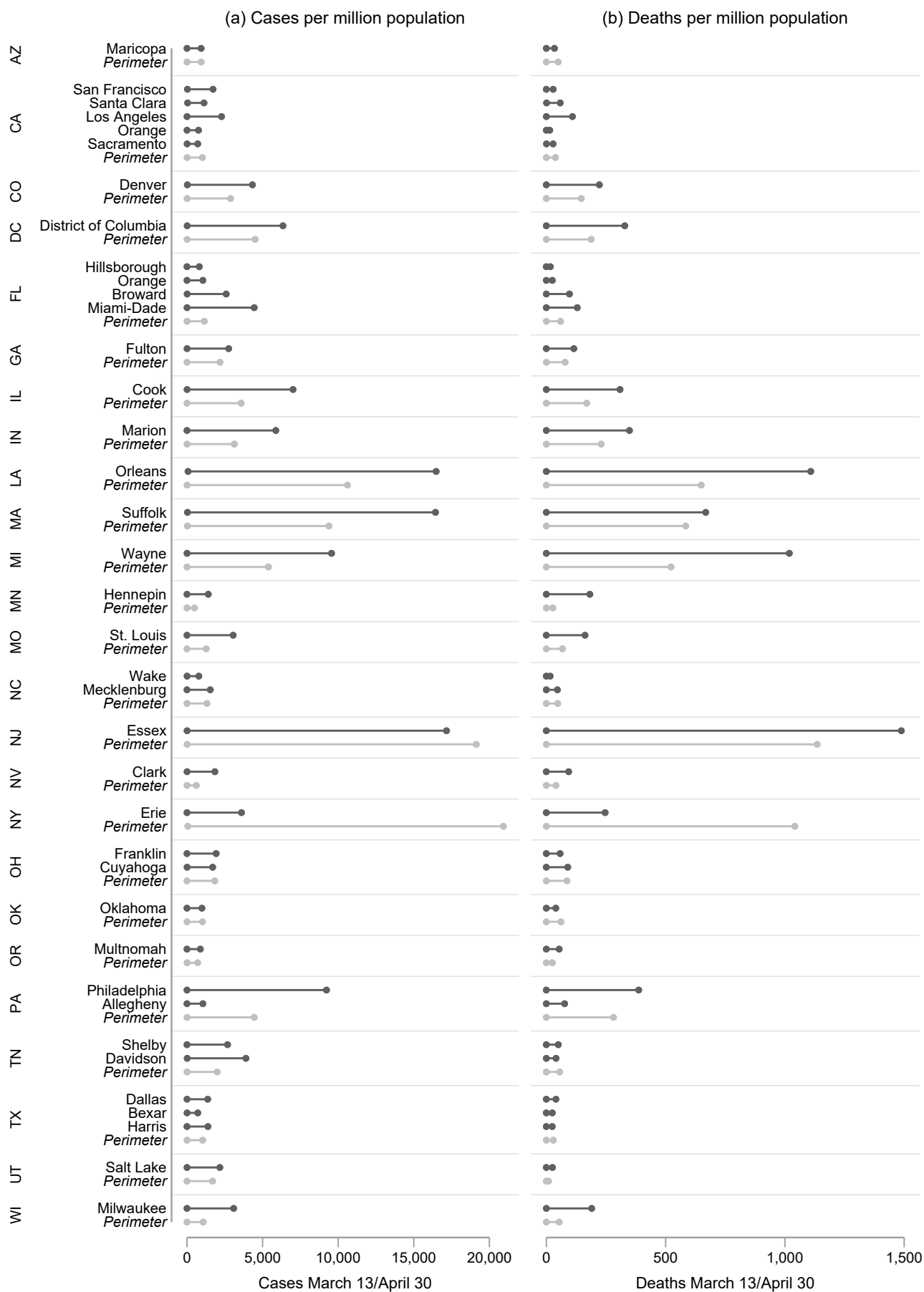
To generate covariates, we collect county-level data on population by age, sex, and ethnicity from the 2016 US census provided by the *National Bureau of Economic Research*.<sup>8</sup> To stratify our analysis, we use, among others, information on population density, climate, and the timing of SIPOs. The information on county land area is collected from the US Census Bureau.<sup>9</sup> Data on historical county climate, including data on April temperature, are collected from the *National Centers for Environmental Information*.<sup>10</sup> Finally, information on the introduction of SIPOs on the state level is taken from Dave et al. (2020). Descriptive statistics for all variables used in our empirical analysis by county type are presented in Table 1.

<sup>8</sup>Data are available at [data.nber.org/seer-pop/desc/](https://data.nber.org/seer-pop/desc/).

<sup>9</sup>Data are available at <https://data.census.gov/cedsci>.

<sup>10</sup>The NOAA's Climate Divisional Database is available at [data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C000005](https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C000005).

FIGURE 1 — Change in reported COVID-19 cases and deaths per venue and perimeter county over time



Notes: This figures displays the evolution of the number of reported COVID-19 cases (a) and deaths (b) for each venue county in our data. The left dot is always the number of cases per million population on March 13, while the right dot is the number of cases per million population on April 30. Additionally, we calculate averages of cases and deaths from March 13 and April 30 over the counties adjacent to NBA or NHL venues in each state, which we call the ‘perimeter.’



### 3. Estimation Strategy

In our estimation analysis, we aim to explain the cumulative number of COVID-19 infections and deaths in a given county  $c$  in state  $s$  adjacent to, or hosting, venue  $v$ . Our sample comprises two types of counties, those which host an NBA or NHL venue (hereafter venue county) and those adjacent to a venue county. This sample definition provides us with a clear match between each county  $c$  and venue  $v$ . The dependent variable, COVID-19 deaths $_{c,v(s)}$ , is defined as the cumulative number of COVID-19 deaths in county  $c$  (measured on April 30, 2020) per one million population. The mean of the death rate in venue counties is 211.3 with a standard deviation of 329.7 (see Table 1).

The explanatory variable of primary interest, games $_{c,v(s)}$ , varies across venues and measures the cumulative number of games (NBA and NHL) at venue  $v$  between March 1 and 11. Starting from March 12, both leagues suspended their seasons and all games were canceled.<sup>11</sup> There were on average 12.3 games with considerable variation to exploit. The number of games varies between 0 and 16, with a standard deviation of 3.52. This set-up translates to the following estimation model:

$$\text{COVID-19 deaths}_{c,v(s)} = \beta \cdot \text{games}_{c,v(s)} + \mathbf{X}_c \delta + \sum \gamma_s + \varepsilon_{c,v(s)}, \quad (1)$$

where  $\mathbf{X}_c$  are county-level controls, and  $\gamma_s$  are venue-state fixed-effects. Our county-level controls comprise population density and the sex-race-age distribution.

Our main parameter of interest is  $\beta$ , which captures the impact of an additional mass gathering due to a NBA or NHL game on the cumulative number of COVID-19 deaths. Given that the game schedules were determined long before the first COVID-19 case became public, there should be no correlation between games $_{c,v(s)}$  and the error term  $\varepsilon_{c,v(s)}$ . This identifying assumption is supported by the fact that the number of games does not correlate with observed county characteristics (see Appendix Table A.1). A potential issue for the interpretation of our estimate would be anticipation effects, in the sense that people may have increasingly refrained from visiting games prior to the lockdown. This would lead to an attenuation bias and our results being a lower bound of the actual effect. However, we can show that game attendance did not systematically change before suspension of play (see Appendix Figure A.2).

### 4. Main Estimation Results

Our estimation results are summarized in Table 2. We estimate the model in equation (1) both on the cumulative number of COVID-19 cases (panel A) and deaths (panel B) per one million population. We find a significant positive effect of the number of mass gatherings on both of these outcomes. Our most conservative estimates indicate that each additional mass gathering between March 1 and 11 increased cases by 379 per one million (column 4) and deaths by approximately 16 per million population (column 3). These are substantial effects. Compared to the average case and death rates

---

<sup>11</sup>Prior to that only two games on March 11 were cancelled. In both cases, players were tested/suspected for COVID-19.

TABLE 2 — Impact of pre-scheduled mass gatherings on COVID-19 infections and death rate

	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Cumulative number of confirmed COVID-19 infections per million population<sup>a</sup></i>					
Cum. number of games	485.679** (208.756)	452.777** (202.890)	383.526** (183.365)	379.147** (174.000)	380.524** (174.385)
Population density			0.277* (0.141)	0.314* (0.178)	0.320* (0.178)
Perc. non-white population			69.612*** (17.147)		
Perc. population aged 60+			-0.356 (30.150)		
Perc. population female			-288.475 (536.027)		
Days since SIPO in place <sup>c</sup>					47.235 (35.065)
Sex-race-age distribution <sup>e</sup>	No	No	No	Yes	Yes
Venue-state fixed effects	No	Yes	Yes	Yes	Yes
Venue county (1 = yes, 0 = no)	No	No	Yes	Yes	Yes
<i>Panel B. Cumulative number of COVID-19 deaths per million population<sup>b</sup></i>					
Cum. number of games	29.470*** (11.067)	23.597** (9.767)	15.825** (7.064)	16.350** (6.598)	16.410** (6.610)
Population density			0.004 (0.007)	0.012 (0.009)	0.012 (0.009)
Confirmed COVID-19 cases <sup>d</sup>			4.635*** (0.517)	4.633*** (0.589)	4.656*** (0.599)
Perc. non-white population			2.997*** (0.717)		
Perc. population aged 60+			2.419* (1.316)		
Perc. population female			1.541 (9.605)		
Days since SIPO in place <sup>c</sup>					2.858 (2.126)
Sex-race-age distribution <sup>e</sup>	No	No	No	Yes	Yes
Venue-state fixed effects	No	Yes	Yes	Yes	Yes
Venue county (1 = yes, 0 = no)	No	No	Yes	Yes	Yes

*Notes:* The number of observations is 242. Robust standard errors are presented in parentheses, stars indicate significance: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . <sup>a</sup> The dependent variable in Panel A is the number of confirmed COVID-19 infections per million inhabitants on April 30, 2020 with a mean of 2,993.83 (std. dev. 4,742.62). <sup>b</sup> The dependent variable in Panel B is the death rate, defined as the number of COVID-19 deaths per million population on April 30, 2020 with a mean of 150.50 (std. dev. 264.61). The number of games measures all NBA and NHL games which took place between March 1 and March 12. <sup>c</sup> The number of days Shelter-in-Place Orders were active on April 30 2020 in the observed county; source: [Dave et al. \(2020\)](#). <sup>d</sup> The number of confirmed COVID-19 infections per county and 1,000,000 county residents. <sup>e</sup> The sex-race-age distribution is defined as a set of 16 variables capturing the share of the total population of sex  $g$ , of race  $h$ , and in age-group  $i$ , where  $h$  is white and non-white, and  $i$  is 0 – 19, 20 – 39, 40 – 59, 60+.

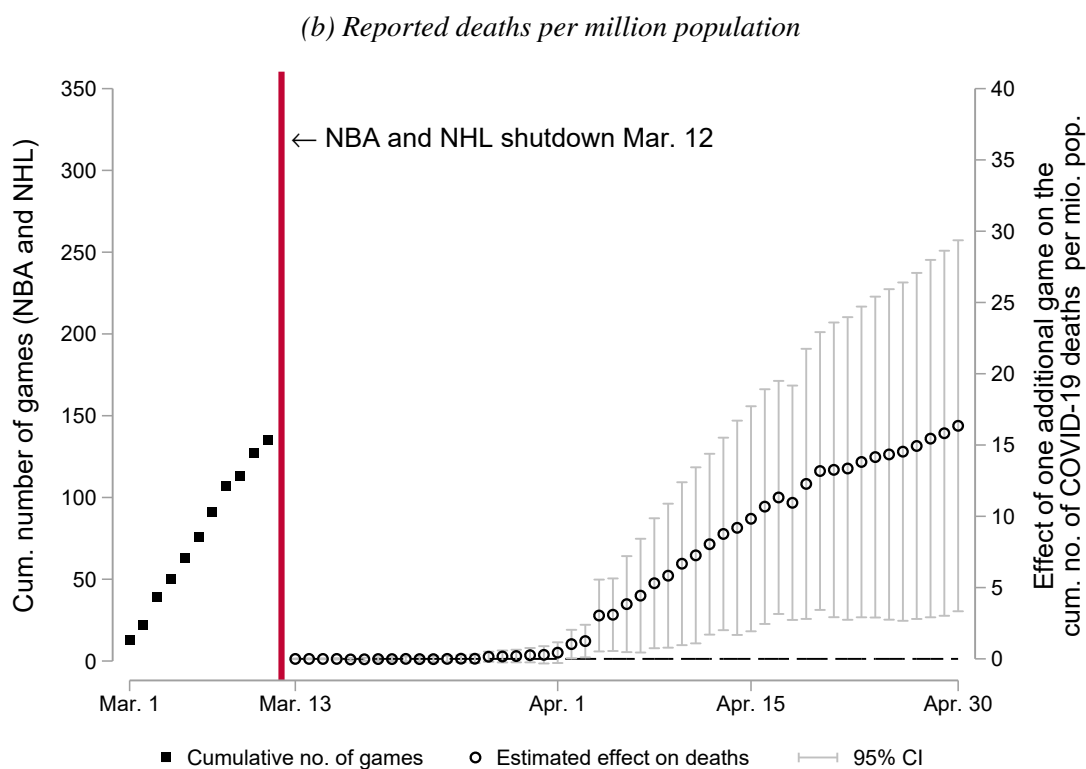
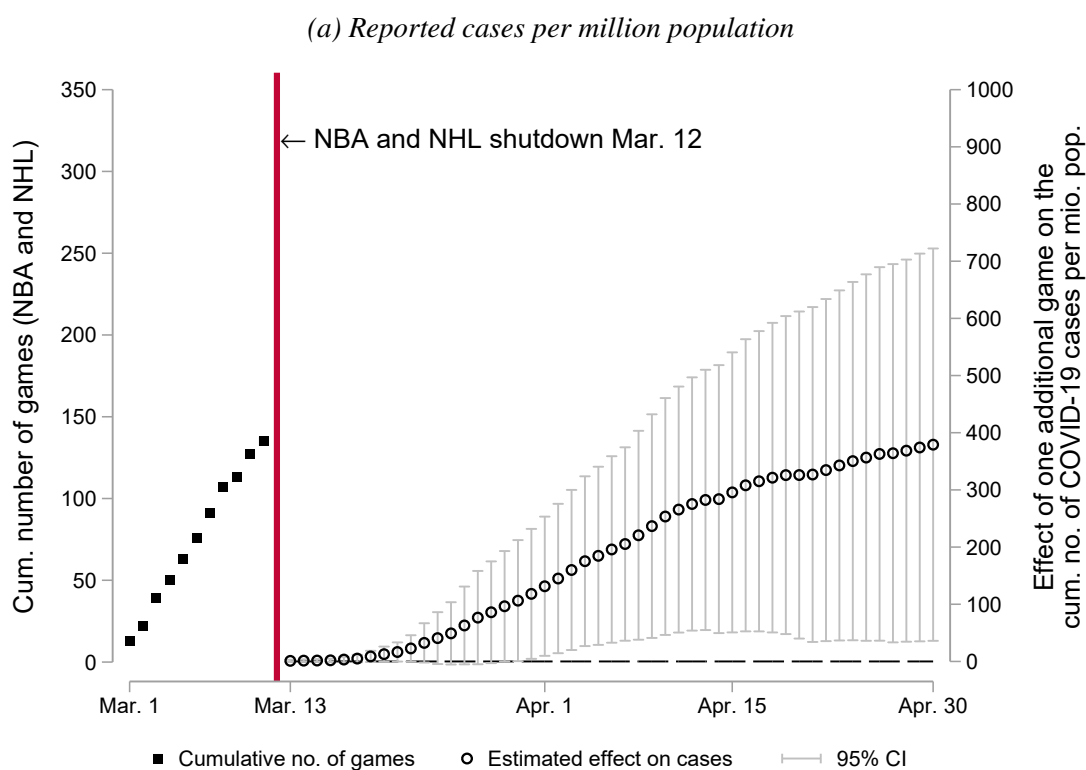
across the counties in the data, our estimates correspond to increases of 13 percent and 11 percent per game, respectively. Both are statistically significant at the 5 percent level.

These findings are robust across different specifications. In column (1), we show the unconditional relationship between cases/deaths and games. In column (2), we introduce venue-state fixed effects. In column (3), we additionally include a binary indicator capturing whether the county hosts one or multiple venues, the population density, and the shares of females, people above 60 years of age, and non-whites in the population. In column (4), we alternatively use the full sex-race-age distribution, defined as a set of 16 variables capturing the share of sex  $g$ , of race  $h$ , and in age-group  $i$  in the population; where  $h$  is white or non-white, and  $i$  is 0–19, 20–39, 40–59, or 60+. In column (5), we

control also for the number of days a statewide SIPO had been in place in state the observed county  $c$  is located in. When analyzing deaths in panel B, columns (3) to (5), we additionally control for the number of confirmed cases by March 13. Our covariates do not have causal interpretations, hence we refrain from interpreting them. However, we note that the negative sign on “Days since SIPO in place” may simply point towards the fact that states with high early case and death counts had to introduce SIPOs sooner.

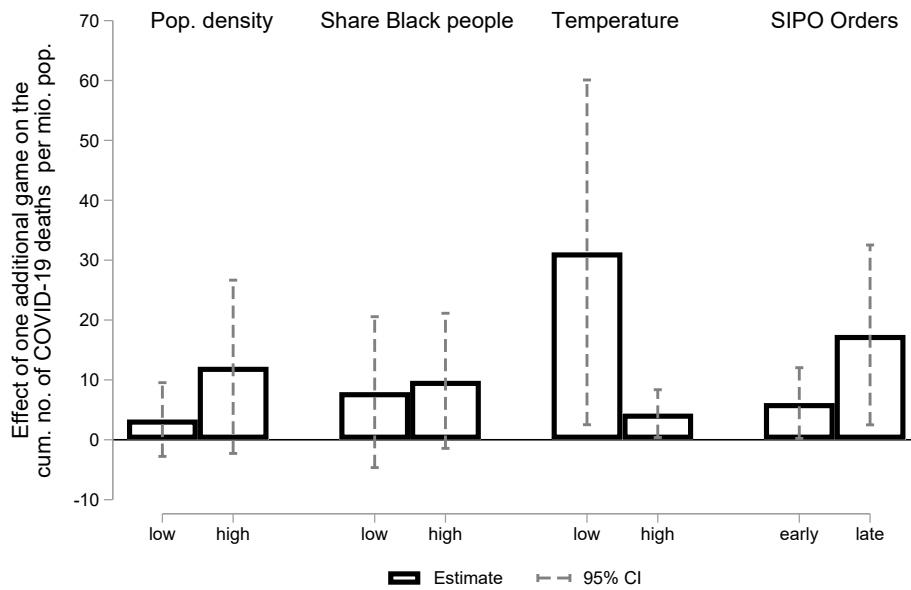
In Figure 2, we provide an overview on the dynamics underlying these effects. The horizontal axis measures time from March 1 to April 30. The squares capture the cumulative number of games (NBA plus NHL) before the leagues suspended play, indicated by the red vertical line. The hollow circles measure the estimated effect of an additional game on the cumulative number of COVID-19 cases (Panel A) and deaths (Panel B) on each day between March 13 and April 30. Each estimate comes from a separate regression, with the dependent variable being measured on different days. The right-most estimate is our baseline. A priori, we expect effects to be strongest around day 14 after the shutdown. This is precisely what we find. The effect of games starts to pick up around March 28 and increases at a decreasing rate since then. This is true for both cases and deaths. Furthermore, we see that cases respond earlier than deaths, which makes sense given the natural lag between diagnosis and death. In terms of magnitudes, estimates for COVID-19 deaths (cases) range between  $-0.003$  ( $0.996$ ) on March 13 and  $16.350$  ( $379.147$ ) on April 30.

FIGURE 2 — Estimated effect of non-canceled games on confirmed deaths and cases.



Notes: The squares capture the cumulative number of games (NBA plus NHL) before suspension of the leagues (red vertical line). The hollow circles measure the estimated effect of one additional game on the cumulative number of COVID-19 cases (Panel A) and deaths (Panel B) on each day between March 13 and April 30. Each estimate comes from a separate regression, with dependent variables measured on different days, and the control variables as in column (4) of Table 2 are used.

FIGURE 3 — Treatment effect heterogeneity



*Notes:* We replicate our results in sub-samples defined by the median of the respective stratification variable. Estimates base on county level population density are based on the split along the median of the population density distribution of all 242 counties in our data. The county level share of African American population was calculated using 2016 US census data provided by the National Bureau of Economic Research. The third sample split is based on the maximum temperature in April for the 20 most recent years, 1998 – 2019. *low* indicates counties below the median of this long-term temperature median, *high* indicates above the median. Finally, we split along the median of days statewide SIPO regulations were in place by April 30, source: [Dave et al. \(2020\)](#). States without statewide SIPO regulations (MA, MN, OK, TN and UT) are coded 0.

## 5. Heterogeneity

So far, we have established that mass gatherings in early March increased COVID-19 deaths in counties surrounding NBA and NHL venues by 11 percent. Additionally, we are interested whether there is heterogeneity in these effects by county characteristics. In Figure 3, we therefore stratify our sample by population density, ethnic composition (measured by the share of Black people in the population), average temperature, and policy responsiveness (i.e., when SIPOs were first introduced). We split each variable by its sample median and repeat our regressions from above.

Effects tend to be stronger in counties with a denser population, but the difference to less densely-populated counties is not significant. Splitting the sample by the share of Black people in the population gives a similar picture. This is surprising, given that early reports in the medical literature suggest that Black people tend to be affected more strongly by COVID-19 than other ethnic groups (e.g., [Yancy 2020](#)). Splitting by temperature, we find that colder areas clearly drive our effects. In counties with below-median temperatures, the effect on deaths is almost twice as high than in the baseline. This is in line with the idea that the virus replicates more easily in lower-temperature conditions. However, the literature has not yet reached consensus whether this is indeed the case. While some early reports from China document a negative correlation between temperature and COVID-19 spread (e.g., [Wang et al. 2020](#)), others find no such (or even a positive) connection (e.g., [Yao et al. 2020](#),

Ma et al. 2020). Finally, we use the time since statewide SIPO orders were enacted as a measure of policy responsiveness. Here we find that mass gatherings have the largest effects in counties situated in late-adopter states. This is perhaps because the virus can spread more easily without SIPOs in place.

## 6. Policy Conclusions

In this paper, we present estimates for the impact of mass gatherings in the form of NBA or NHL games on the community spread of COVID-19. We find that one additional game increased the cumulative number of COVID-19 deaths in affected US counties by 11 percent. We conclude that banning mass gatherings is an effective NPI to slow the spread of COVID-19.

Common estimates in NBA circles, for example, suggest that each game yields an average \$1.2 million in gate revenue.<sup>12</sup> This figure comprises all game-day revenue, including tickets and concessions, but excludes revenues from TV and sponsoring deals, and the resulting consumer surplus. The latter two components might not be lost if games are played without audience. Thus, even if the full 82 game season in 2020 would have been canceled, this would not have exceeded \$100 million in losses, which is considerably less than what estimates suggest school closures in the wake of COVID-19 cost. Since we now know that most of the season is merely postponed instead of canceled completely, we expect the league to recover at least part of these losses.

We suggest that public health officials recommend canceling or postponing mass gatherings during COVID-19 and future pandemics.

---

<sup>12</sup>See, for example, [nbcsports.com/chicago/bulls/report-nba-could-lose-nearly-500-million-ticket-revenue-wi](https://www.nbcsports.com/chicago/bulls/report-nba-could-lose-nearly-500-million-ticket-revenue-wi) accessed June 9, 2020.

## References

- Ahmed, F., Zviedrite, N. & Uzicanin, A. (2018), 'Effectiveness of workplace social distancing measures in reducing influenza transmission: A systematic review', *BMC Public Health* **18**(1), 518.
- Alexander, D. & Karger, E. (2020), Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior, Report, Federal Reserve Bank of Chicago.
- Barro, R. J. (2020), Non-Pharmaceutical Interventions and Mortality in U.S. Cities during the Great Influenza Pandemic, 1918-1919, Working Paper 27049, National Bureau of Economic Research.
- Bootsma, M. C. J. & Ferguson, N. M. (2007), 'The effect of public health measures on the 1918 influenza pandemic in U.S. cities', *Proceedings of the National Academy of Sciences* **104**(18), 7588–7593.
- Chapelle, G. (2020), The medium-run impact of non-pharmaceutical Interventions: Evidence from the 1918 influenza in US cities, in B. W. di Mauro & C. Wyplosz, eds, 'Covid Economics, Vetted and Real-Time Papers', Vol. 18, CEPR Press.
- Coibion, O., Gorodnichenko, Y. & Weber, M. (2020), 'The Cost of the COVID-19 Crisis: Lockdowns, Macroeconomic Expectations, and Consumer Spending', *Covid Economics: Vetted and Real-Time Papers* (29), 1–49.
- Cowling, B. J. & Aiello, A. E. (2020), 'Public Health Measures to Slow Community Spread of Coronavirus Disease 2019', *Journal of Infectious Diseases* **221**(11), 1749–1751.
- Dave, D. M., Friedson, A. I., Matsuzawa, K. & Sabia, J. J. (2020), When Do Shelter-in-Place Orders Fight COVID-19 Best? Policy Heterogeneity Across States and Adoption Time, Working Paper 27091, National Bureau of Economic Research.
- Earn, D. J. (2012), 'Effects of School Closure on Incidence of Pandemic Influenza in Alberta, Canada', *Annals of Internal Medicine* **156**(3), 173.
- Fang, H., Wang, L. & Yang, Y. (2020), Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-NCOV) in China, NBER Working Paper 26906, National Bureau of Economic Research, Cambridge, MA.
- Friedson, A. I., McNichols, D., Sabia, J. J. & Dave, D. (2020), Did California's Shelter-in-Place Order Work? Early Coronavirus-Related Public Health Effects, Working Paper 26992, National Bureau of Economic Research.
- Goodman-Bacon, A. & Marcus, J. (2020), Using Difference-in-Differences to Identify Causal Effects of COVID-19 Policies, Working paper, Vanderbilt University.
- Greenstone, M. & Nigam, V. (2020), 'Does Social Distancing Matter?', *Covid Economics: Vetted and Real-Time Papers* (7), 1–22.
- Gupta, S., Nguyen, T. D., Rojas, F. L., Raman, S., Lee, B., Bento, A., Simon, K. I. & Wing, C. (2020), Tracking Public and Private Responses to the COVID-19 Epidemic: Evidence from State and Local Government Actions, NBER Working Paper 27027, National Bureau of Economic Research, Cambridge, MA.
- Hatchett, R. J., Mecher, C. E. & Lipsitch, M. (2007), 'Public health interventions and epidemic intensity during the 1918 influenza pandemic', *Proceedings of the National Academy of Sciences* **104**(18), 7582–7587.

- Hoang, V. & Gautret, P. (2018), 'Infectious Diseases and Mass Gatherings', *Current Infectious Disease Reports* **20**(11), 44.
- Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., Druckenmiller, H., Huang, L. Y., Hultgren, A., Krasovich, E., Lau, P., Lee, J., Rolf, E., Tseng, J. & Wu, T. (2020), The Effect of Large-Scale Anti-Contagion Policies on the Coronavirus (COVID-19) Pandemic, Working paper, UC Berkeley, Berkeley, CA.
- Karami, M., Doosti-Irani, A., Ardalan, A., Gohari-Ensaf, F., Berangi, Z., Massad, E., Yeganeh, M. R., Asadi-Lari, M. & Gouya, M. M. (2019), 'Public Health Threats in Mass Gatherings: A Systematic Review', *Disaster Medicine and Public Health Preparedness* **13**(5–6), 1035–1046.
- Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., Yan, J., Niu, J., Zhou, J. & Luo, B. (2020), 'Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China', *Science of The Total Environment* **724**, 138226.
- Mangrum, D. & Niekamp, P. (2020), College Student Contribution to Local COVID-19 Spread: Evidence from University Spring Break Timing, Unpublished manuscript, Vanderbilt University and Ball State University.
- Markel, H., Lipman, H. B., Navarro, J. A., Sloan, A., Michalsen, J. R., Stern, A. M. & Cetron, M. S. (2007), 'Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic', *Journal of the American Medical Association* **298**(6), 644–654.
- McCloskey, B., Zumla, A., Ippolito, G., Blumberg, L., Arbon, P., Cicero, A., Endericks, T., Lim, P. L. & Borodina, M. (2020), 'Mass gathering events and reducing further global spread of COVID-19: A political and public health dilemma', *Lancet* **395**(10230), 1096–1099.
- Memish, Z. A., Steffen, R., White, P., Dar, O., Azhar, E. I., Sharma, A. & Zumla, A. (2019), 'Mass gatherings medicine: Public health issues arising from mass gathering religious and sporting events', *Lancet* **393**(10185), 2073–2084.
- Miyaki, K., Sakurazawa, H., Mikurube, H., Nishizaka, M., Ando, H., Song, Y. & Shimbo, T. (2011), 'An effective quarantine measure reduced the total incidence of influenza A H1N1 in the workplace: Another way to control the H1N1 flu pandemic', *Journal of Occupational Health* **53**(4), 287–292.
- Nunan, D. & Brassey, J. (2020), What is the evidence for mass gatherings during global pandemics? A rapid summary of best-available evidence, Evidence Service, Centre for Evidence-Based Medicine, Nuffield Department of Primary Care Health Sciences, University of Oxford.
- Rainey, J. J., Phelps, T. & Shi, J. (2016), 'Mass Gatherings and Respiratory Disease Outbreaks in the United States – Should We Be Worried? Results from a Systematic Literature Review and Analysis of the National Outbreak Reporting System', *PLoS ONE* **11**(8).
- Sadique, M. Z., Adams, E. J. & Edmunds, W. J. (2008), 'Estimating the costs of school closure for mitigating an influenza pandemic', *BMC Public Health* **8**(1), 135.
- Thunström, L., Newbold, S. C., Finnoff, D., Ashworth, M. & Shogren, J. F. (forthcoming), 'The benefits and costs of using social distancing to flatten the curve for COVID-19', *Journal of Benefit-Cost Analysis* pp. 1–27.
- Viner, R. M., Russell, S. J., Croker, H., Packer, J., Ward, J., Stansfield, C., Mytton, O., Bonell, C. & Booy, R. (2020), 'School closure and management practices during coronavirus outbreaks including COVID-19: A rapid systematic review', *Lancet Child & Adolescent Health* **4**(5), 397–404.

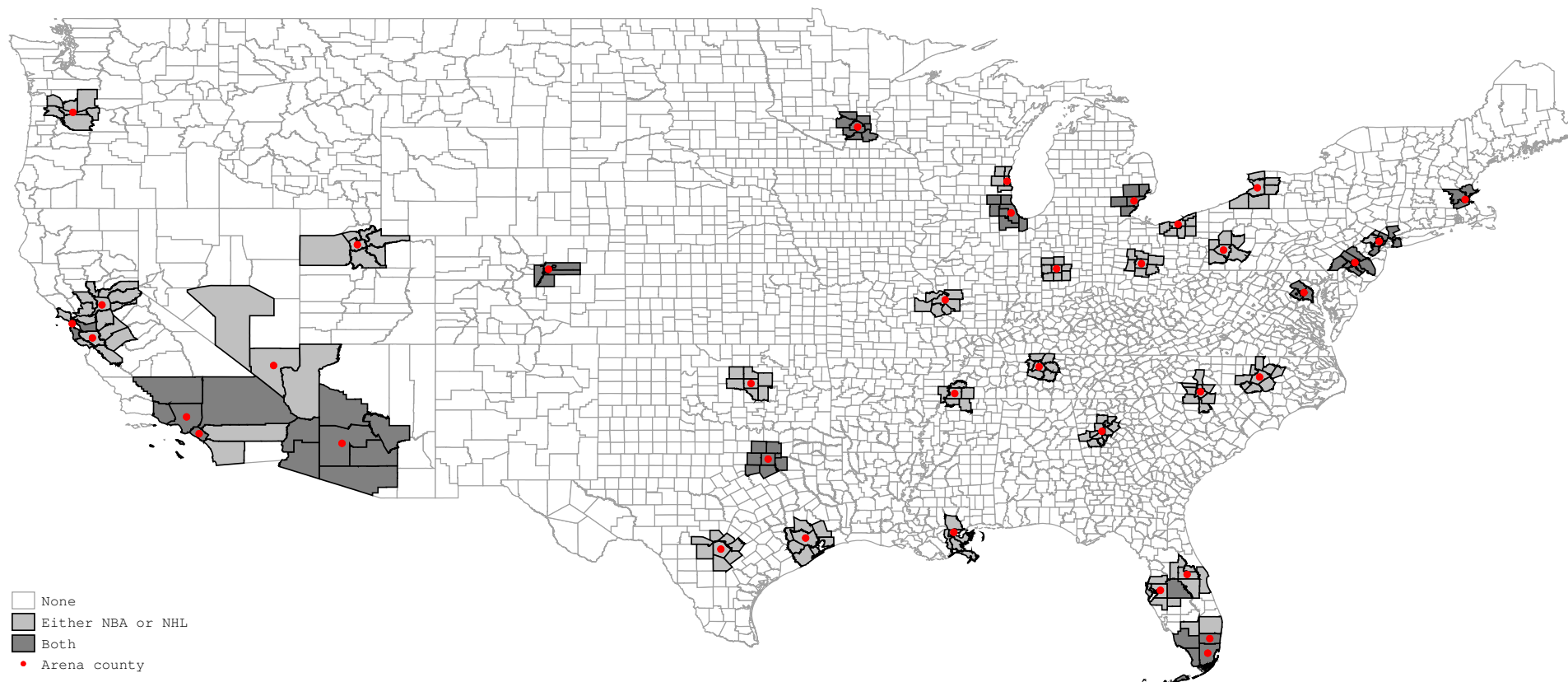


- Wang, J., Tang, K., Feng, K., Lin, X., Lv, W., Chen, K. & Wang, F. (2020), High Temperature and High Humidity Reduce the Transmission of COVID-19, SSRN Scholarly Paper ID 3551767, Social Science Research Network, Rochester, NY.
- Wheeler, C. C., Erhart, L. M. & Jehn, M. L. (2010), 'Effect of School Closure on the Incidence of Influenza Among School-Age Children in Arizona', *Public Health Reports* **125**(6), 851–859.
- Yancy, C. W. (2020), 'COVID-19 and African Americans', *Journal of the American Medical Association* **323**(19), 1891–1892.
- Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H. & Wang, W. (2020), 'No association of COVID-19 transmission with temperature or UV radiation in Chinese cities', *European Respiratory Journal* **55**(5).

# Appendix

This appendix provides additional tables and figures.

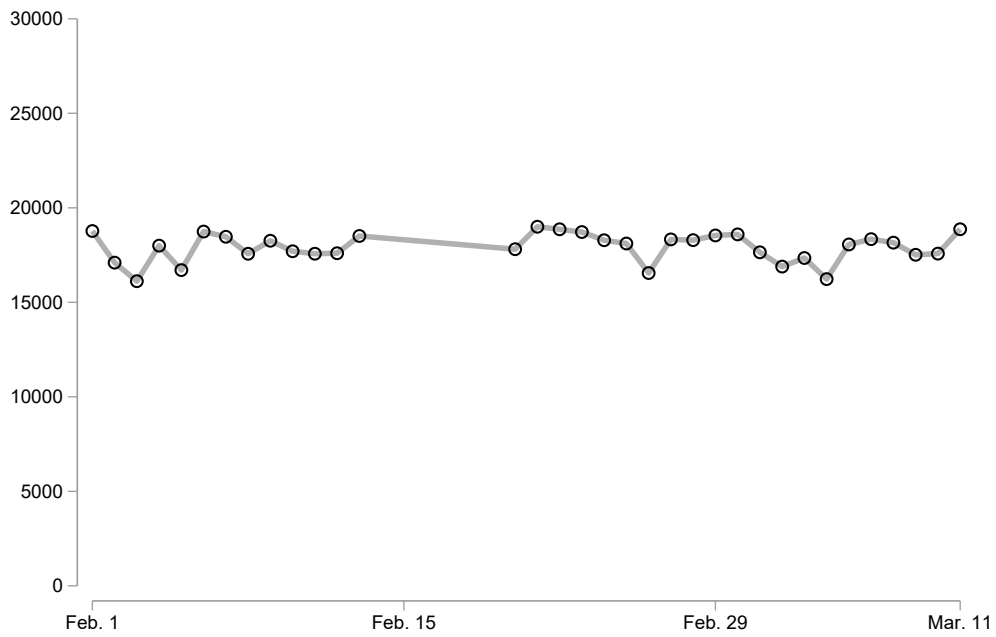
FIGURE A.1 — NBA and NHL venues and adjacent counties in the United States



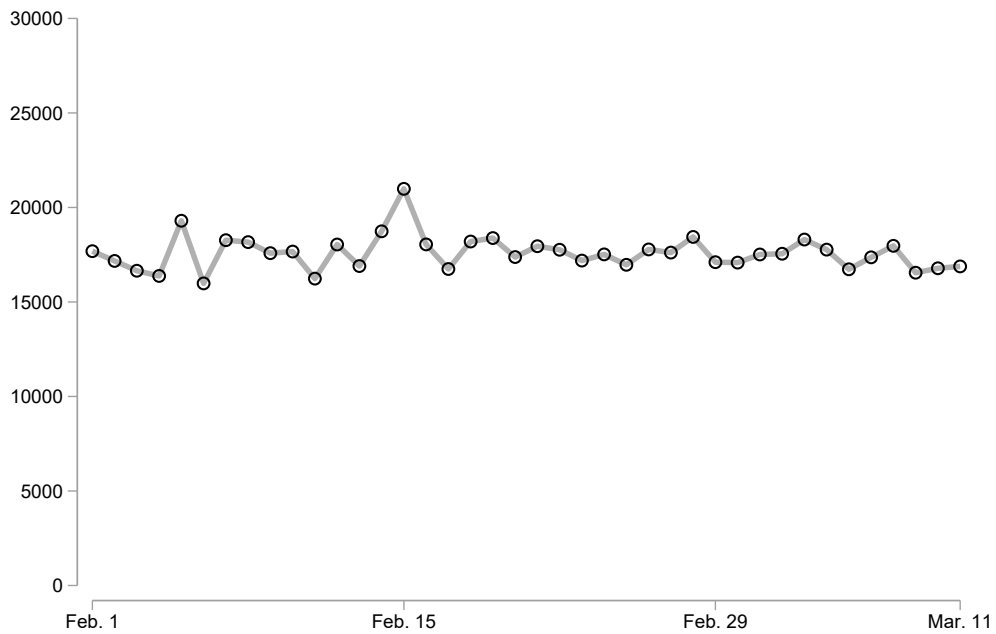
*Notes:* This map provides an overview on the counties we use in our analysis. The light-gray shaded areas are counties that are adjacent to either a NBA or a NHL venue, the dark-gray shaded counties are in the perimeter of both a NBA and a NHL venue. Counties where venues are located are marked with red dots.

FIGURE A.2 — Average absolute attendance by NBA and NHL game between February 1 and March 11

(a) NBA



(b) NHL



Notes: This figure displays the average attendance of all NBA and NHL games held between February 1 and March 11 in venues located in the US. On March 12, the NBA cancelled two games before tip-off. After March 11 both leagues suspended the seasons indefinitely.

TABLE A.1 — Correlation of the number of NBA and NHL games between March 1 and March 11 with observed county characteristics

	(1)	(2)
Population density	0.000** (0.000)	0.000 (0.000)
Share female	-0.146 (0.160)	0.249 (0.160)
Share non-white	0.008 (0.013)	-0.003 (0.013)
Share elderly	-0.029 (0.030)	-0.050 (0.035)
Max. Temp. April <sup>a</sup>	-0.021 (0.018)	0.092 (0.081)
Venue-state fixed effects	No	Yes
Venue county (1 = yes, 0 = no)	No	Yes
Number of Observations	242	242

*Notes:* The number of observations is 242. Robust standard errors are presented in parentheses, stars indicate significance: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The dependent variable is the number of all NBA and NHL games which took place between March 1 and March 12 in the observed county. Independent variables correspond to those use in specification (4) of Table 2. <sup>a</sup> measures the monthly average maximum temperature in April, based on the 20 most recent years, 1998 – 2019.