DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

Mass Gatherings Contributed to Early COVID-19 Mortality: Evidence from US Sports

by

Alexander AHAMMER

Martin HALLA

Mario LACKNER

Working Paper No. 2013
July 2020

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

Corresponding author: martin.halla@jku.at

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

Alexander Ahammer[a,b], Martin Halla[a,b,c,d], Mario Lackner[a,b]

[a]*Johannes Kepler University, Linz*
[b]*Christian Doppler Laboratory on Ageing, Health, and the Labor Market, Linz*
[c]*IZA, Institute for the Study of Labor, Bonn*
[d]*GÖG, Austrian Public Health Institute, Vienna*

This version July 29, 2020
(First version: June 11, 2020)

Abstract

Social distancing is important to slow the community spread of infectious disease, but it creates enormous economic and social cost. Thus, it is 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 indoor mass gatherings on COVID-19 mortality in affected US counties. One additional mass gathering increased the cumulative number of COVID-19 deaths in affected counties by 9 percent.

JEL Classification: I18, H12, I10.
Keywords: Social distancing, mass gatherings, Coronavirus Disease 2019, COVID-19.

Corresponding author: Martin Halla, Department of Economics, Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria; e-mail: martin.halla@jku.at. A previous version of this paper has been published in *Covid Economics, Vetted and Real-Time Papers*. For helpful discussions and comments we would like to thank Jochen Güntner, Brad Humphreys, Dominik Schreyer, Carl Singleton, Martin Watzinger, and participants at the Reading Online Sport Economics Seminar (ROSES). The usual disclaimer applies. Financial support from the Christian Doppler Laboratory "Aging, 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 in infections, to take pressure off healthcare systems, and ultimately to save lives. In addition to careful 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 to stop the pandemic.

One important public policy to promote physical distancing is to ban mass gatherings (Memish et al. 2019).[1] 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 *National Basketball Association* (NBA) and *National Hockey League* (NHL) games have contributed to the spread of COVID-19 in the United States.[2] Both leagues play exclusively in indoor venues, which present a high-risk setting for infectious disease transmission. Before play was suspended 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 games held between March 1 and March 11 have 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.[3] In fact, we can show that game schedules are not correlated with observable county characteristics, and that ticket sales did not systematically change until the NBA and NHL suspended play. While we do not have data on actual game attendance, we note that any no-show behavior would imply that our estimates are conservative

[1] 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.*"

[2] 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.

[3] 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.

and the true effect on COVID-19 mortality perhaps much higher.[4]

Our results suggest that one additional indoor mass gathering between March 1 and 11 in the form of an NBA or NHL game increased the cumulative number of COVID-19 cases (measured on April 30, 2020) in affected counties by at least 243 per one million population ($p < 0.01$) or 8.3 percent, and the number of COVID-19 deaths per million by 13 ($p < 0.01$) or 8.9 percent. These effects are larger in densely populated areas and in colder regions. We conclude that banning indoor mass gatherings has an enormous potential to save lives. This is especially important given that such measures are relatively easy and cheap to implement.

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, Nunan & Brassey (2020) conclude that the impact of mass gatherings on COVID-19 is still poorly understood. Until now, 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). The best available evidence suggests that multiple-day events with crowded communal accommodations are most associated with increased risk of infection (Nunan & Brassey 2020). Two papers relate sport events to local seasonal influenza mortality. Stoecker et al. (2016) show that NFL team appearances in the Super Bowl caused a 18 percent increase in flu mortality in the population over age 65, and Cardazzi et al. (2020) find that US cities which get to host a professional sports team experience an increase in local influenza mortality by an estimated 4 to 24 percent. We are not aware of any design-based estimation of the impact of any mass gatherings on the *community* spread of COVID-19.[5]

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, Friedson, Matsuzawa & Sabia (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

[4]Evidence from Belarusian soccer, which did not suspend play, indicates that sports attendance did decline somewhat in the early stages of COVID-19 (Reade et al. 2020).

[5]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.

100,000 population and deaths by 1,661. All these papers use difference-in-differences designs,[6] which Goodman-Bacon & Marcus (2020) provide a critical account of in the context of evaluating NPIs.

Extensive work has focused on previous pandemics. 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, since 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 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 percent less and made 13 percent 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 percent. 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,

[6]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.

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 presents results from in-space placebo tests. Section 7 reports different sensitivity checks. Section 8 provides concluding comments. Additional figures and tables we delegate to a Web Appendix.

2. Data

Our estimation sample consists of 38 counties which host either a NBA or a NHL venue, or both, and all their 204 neighboring counties. A venue county together with its adjacent counties we call a 'perimeter' (see Figure 1). For all affected venue and perimeter counties, we collect information on COVID-19 cases and related deaths.[7] In Appendix Figure A.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 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 our treatment variable, we use information on NBA and NHL games played between March 1 and March 11.[8] 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 canceled two games right before tip-off on March 11: Utah Jazz at Oklahoma City Thunder, where Utah player Rudy Gobert tested positive for COVID-19 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 Utah Jazz game earlier the same week.

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*.[9] To stratify our analysis, we use, among others, information on population density, climate, and the timing of SIPOs. The information

[7]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`). 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.

[8]The information on NBA games is scraped from *Basketball Reference* (see `basketball-reference.com`). Data on NHL games are collected from *Hockey Reference* (see `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.

[9]Data are available at `data.nber.org/seer-pop/desc/`.

on county land area is collected from the US Census Bureau.[10] Historical climate data on the county-level, including data on April temperatures, are collected from the *National Centers for Environmental Information*.[11] Finally, information on the introduction of SIPOs on the state level is taken from Dave, Friedson, Matsuzawa & Sabia (2020). Descriptive statistics for all variables used in our empirical analysis by county type are presented in Appendix Table A.1.

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 located in state s, affected by 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 $\text{deaths}_{c(s),v}$, is defined as the cumulative number of COVID-19 deaths in county c (measured on April 30, 2020) per one million population. The average death rate in venue counties is 211.3 with a standard deviation of 329.7 (see Appendix Table A.1).

The explanatory variable of primary interest, $\text{games}_{c(s),v}$, 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.[12] On average, there were 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:

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

where $\mathbf{X}_c$ are county-level controls, and γ_s are state fixed-effects. Our county-level controls comprise population density and the sex-race-age distribution. We refrain from estimating venue fixed effects, as we observe 75 counties which are affected by events held at more than one team or venue. We estimate model (1) with county population weights.

Our main parameter of interest is β, 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 $\text{games}_{c(s),v}$ and the error term $\varepsilon_{c(s),v}$. This identifying assumption is supported by the fact that

[10] Data are available at https://data.census.gov/cedsci.
[11] The NOAA's Climate Divisional Database is available at data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00005.
[12] Prior to that only two games on March 11 were canceled. In both cases, players were tested/suspected for COVID-19.

the number of games does not correlate with observed county characteristics (see Appendix Table A.2).

There are three potential issues in regards to the interpretation of our estimate. First, there could be anticipation effects, in the sense that people may increasingly had refrained from visiting games prior to the lockdown. This would lead to an attenuation bias, which implies that our estimate is a lower bound of the actual effect. Unfortunately we do not have data on game attendance, but we can show that ticket sales did not systematically change before suspension of play (see Appendix Figure A.2). Second, our results can only speak to US counties that either host or are adjacent to a sports venue. These are primarily urban regions, where disease may spread more rapidly than in lesser populated areas. Also, there may be spillovers to counties not in the data that we cannot account for. Although we know little about whether fans travel inter-state for sports games, secondary infections, in particular, may cause the virus to spread beyond our perimeters. This again implies that our findings are conservative estimates. Third, for now we assume that an additional game has the same effect on each county. This may be a strong assumption, because there are substantial disparities in land area and population density across counties. To account for this heterogeneity, we provide separate estimates for low and high population density areas in Section 5.

4. Main Estimation Results

Our estimation results are summarized in Table 1. 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 (column 4) indicate that each additional mass gathering between March 1 and 11 increased cases by 243 per one million and deaths by approximately 13 per million population. These are substantial effects. Compared to the average case and death rates across the counties in the data, our estimates correspond to increases of 8.3 percent and 8.9 percent per game, respectively. Both are statistically significant at the 1 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 state fixed effects. In column (3), we additionally include a binary indicator capturing whether the county hosts a venue, the population density, and the shares of non-whites, people above 60 years of age, and females in the population. In column (4), we alternatively use the full interacted 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+. This is our preferred specification. When analyzing deaths in panel B, columns (3) to (4), 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.

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 black squares capture the cumulative number of games (NBA plus NHL) before the leagues suspended play on March 12, which is indicated by the red 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 3 weeks 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 sooner 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 −0.007 (−0.083) on March 13 and 13.168 (242.823) on April 30.

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 8.9 percent per game. 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. In Panel (a) we consider cases, in Panel (b) deaths. Due to the smaller sample sizes, we report 90 percent confidence intervals.

Effects are stronger in areas with high population density. This is to be expected; where people live close to one another, the risk of transmission is greatly increased. In less densely populated areas, the effect of mass gatherings is close to zero and statistically insignificant. This is true for both cases and deaths. In contrast, we do not find different effects if we split the sample by the share of Black people in the population. 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).

In terms of temperature, we find that colder areas clearly drive our effects. In counties with below-median temperatures, the effect on deaths is almost five times as high as in the baseline. This is in line with the idea that the virus replicates more easily in lower-temperature conditions, while higher temperatures, and in particular sunlight, may offer protection against infection (Slusky & Zeckhauser 2020). However, the literature has not yet reached consensus whether this is indeed the case for COVID-19. 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). Note that, because we consider indoor mass gatherings, temperature plays a lesser role for primary infections, but it may be important in determining the extent of secondary and subsequent infections. Finally, we do not find any meaningful effect heterogeneity by how early states enacted SIPO orders. We assume that this is because SIPO timing measures two countervailing effects; while SIPOs reduce disease transmission, early-adopter states may also be more strongly hit by the pandemic.

6. In-space Placebo Tests

If our estimations results indeed measure the effect of sports games on COVID-19 deaths, we should not find systematic effects on deaths in counties that are not exposed to sports games. Therefore, we suggest an in-space placebo test using continental US counties outside our sample to validate our results.[13] We think of this as randomly reassigning venues to counties not in our data (i.e., the white-colored counties in Figure 1), and repeating our estimations from above. To construct this hypothetical scenario, we identify counties that are similar in observables to our venue counties. Essentially, we approach this by estimating conditional propensities to host a venue for each county, and then look for nearest-neighbor pairs in terms of these propensity scores. We then construct perimeters around each placebo county, similar to our main sample.[14] This gives us a sample of 37 placebo venue counties, and 145 adjacent counties. The results of this exercise are in Figure 4. To show how estimates change over time, we repeat our rolling window estimates from before. We find robust zero effects on both cases and deaths.

As a second placebo test, we include the cumulative number of a team's away games in our baseline specification. This variable should have no influence on the COVID-19 transmissions in the home venue county and its perimeter. We can confirm this. If we estimate such a regression, the coefficient on the cumulative number of away games is insignificant ($p = 0.43$). We do not show this estimation output due to space constraints, but it is available upon request.

[13]Note since our treatment is defined in a cross-sectional spirit, the spatial distribution of venues is the only possible variation we can use to perform a placebo test. In particular, it is not possible to do in-time placebo tests within our estimation sample, since there is serial correlation in COVID-19 cases/deaths, and any effect we measure in these counties will always pick up the treatment effect too.

[14]Each non-venue county can appear in only one placebo perimeter.

7. Sensitivity Checks

To assess the sensitivity of our findings, we replicate our main analysis using different estimation samples, estimation weights, and inference methods. First, we modify our estimation sample by omitting each state once and estimating effects for all remaining states. This exercise is in the spirit of jacknife resampling; it should reveal whether individual states have enough leverage to drive our estimates. We obtain 31 different leave-one-out estimation samples, the estimates based on those samples are summarized in Appendix Figure A.3. Panel A shows the estimated coefficients for the cumulative number of COVID-19 cases on April 30. All estimated coefficients are statistically significant and have overlapping 95 percent confidence intervals. Panel B shows the estimated coefficients for the cumulative number of COVID-19 deaths. With one exception, all coefficients are statistically significant, and confidence intervals are again overlapping.

We want to point out two observations. First, omitting California leads to wider confidence intervals. This can be explained by the fact that a substantial number of counties in our sample are located in California (26 counties or 10.74 percent). If we reduce the sample in such a substantial manner, our estimate is still positive but not statistically significant at the 5 percent level. Second, omitting New York reduces the estimated effect size from 13 deaths (243 cases) to 7.6 deaths (109 cases). The estimated coefficients are still statistically significant at the 5.1 (deaths) and 5.4 (cases) percent levels, and the corresponding confidence intervals overlap with those resulting from the other restricted samples. Overall, we conclude that our findings are not driven by a certain state.

Second, we repeat our analysis without population-weighting the regressions. For the cumulative number of COVID-19 cases, we obtain larger coefficients with larger standard errors (see Panel A.1 in Appendix Figure A.4). For the cumulative number of COVID-19 deaths, the coefficients hardly change, but standard errors increase somewhat (see Panel A.2 in Web Appendix Figure A.4). For both outcomes, the overall pattern is comparable to our weighted baseline estimates.

Third, we present our baseline estimates with clustered standard errors at the state-level. In our baseline estimation, we refrained from clustering (Cameron & Miller 2015), since we have a rather low number of clusters and some clusters have only few observations (see Figure 1). It is reassuring that our estimates on the cumulative number of COVID-19 deaths remain highly statistically significant if we use clustered standard errors at the state-level (see Panel B.2 in Appendix Figure A.4). For the cumulative number of COVID-19 cases (see Panel B.1 in Appendix Figure A.4), clustered standard errors lead to a drop in statistical significance. The estimates are now only significant at the 14 percent level. However, deaths are clearly the more reliable and important outcome. Overall, we are confident that our estimates indeed measure a causal effect of indoor mass gatherings on COVID-19 mortality.

8. 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 8.9 percent. We conclude that banning mass gatherings is an effective NPI to slow the spread of COVID-19. We suggest that public health officials recommend canceling or postponing mass gatherings during the COVID-19 pandemic and other future pandemics.

While the benefits of banning mass gatherings are tremendous, the cost are likely low. Estimates in NBA circles suggest, for example, that each game yields an average $1.2 million in gate revenue.[15] 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. More importantly, however, the opportunity cost of banning sports games are likely much lower than those of other NPIs. While it is inevitable that some jobs may be lost in the process, we believe that the resulting human capital loss is orders of magnitude smaller than what we would expect from, for example, school or workplace closures.

A major limitation of our study is that we cannot speak to outdoor mass gatherings. Recent evidence suggests that COVID-19 is primarily transmitted via aerosols (Morawska & Cao 2020, Bourouiba 2020), which implies that indoor gatherings carry a particularly high risk for infection. Outdoor events, especially when certain safety protocols (minimum physical distance, mandatory masks, etc.) are established, may be less risky. There is first evidence, for example, that the 2020 Black Lives Matter protests in the US did not lead to in increase in COVID-19 cases and deaths (Dave, Friedson, Matsuzawa, Sabia & Safford 2020). Since the two other major US sports leagues, the *National Football League* (NFL) and *Major League Baseball* (MLB), are played primarily in open arenas, an interesting avenue for future research would be to test whether disease transmission is indeed lower when our design is applied to these two leagues.

[15] See, for example, nbcsports.com/chicago/bulls/report-nba-could-lose-nearly-500-million-ticket-revenue-without-gam 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.

Bourouiba, L. (2020), 'Turbulent Gas Clouds and Respiratory Pathogen Emissions: Potential Implications for Reducing Transmission of COVID-19', *JAMA* **323**(18), 1837–1838.

Cameron, A. & Miller, D. L. (2015), 'A Practitioner's Guide to Cluster-Robust Inference', *Journal of Human Resources* **50**(2), 317–372.

Cardazzi, A., Humphreys, B. R., Ruseski, J. E., Soebbing, B. & Watanabe, N. (2020), Professional Sporting Events Increase Seasonal Influenza Mortality in US Cities, Economics Faculty Working Papers Series 49, West Virginia University.

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, Discussion Paper 13224, IZA Institute of Labor Economics.

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., Friedson, A., Matsuzawa, K., Sabia, J. & Safford, S. (2020), Black Lives Matter Protests, Social Distancing, and COVID-19, Technical Report w27408, National Bureau of Economic Research, Cambridge, MA.

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?, Working Paper 2020-26, University of Chicago, Becker Friedman Institute for Economics.

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.

Morawska, L. & Cao, J. (2020), 'Airborne transmission of SARS-CoV-2: The world should face the reality', *Environment International* **139**, 105730.

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).

Reade, J. J., Schreyer, D. & Singleton, C. (2020), Stadium Attendance Demand During the Covid-19 Crisis: Early Empirical Evidence from Belarus, SSRN Scholarly Paper ID 3651275, Social Science Research Network, Rochester, NY.

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.

Slusky, D. & Zeckhauser, R. J. (2020), Sunlight and Protection Against Influenza, Working Paper 24340, National Bureau of Economic Research.

Stoecker, C., Sanders, N. J. & Barreca, A. (2016), 'Success Is Something to Sneeze At: Influenza Mortality in Cities that Participate in the Super Bowl', *American Journal of Health Economics* **2**(1), 125–143.

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.

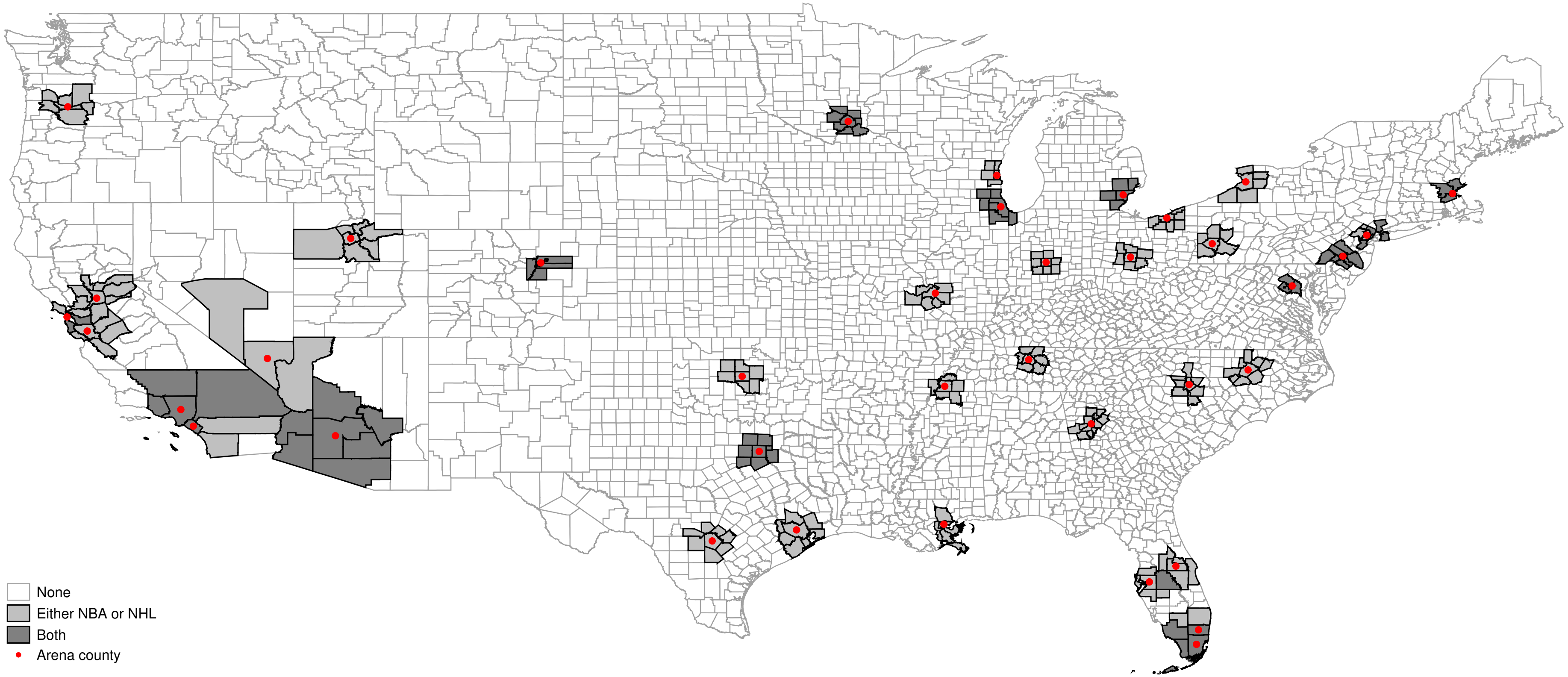
Wheeler, C. C., Erhart, L. M. & Jehn, M. L. (2010), 'Effect of School Closure on the Incidence of Influenza Among School-Age Children in Arizona', *Public Health Reports* **125**(6), 851–859.

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).

Figures and Tables (to be placed in paper)

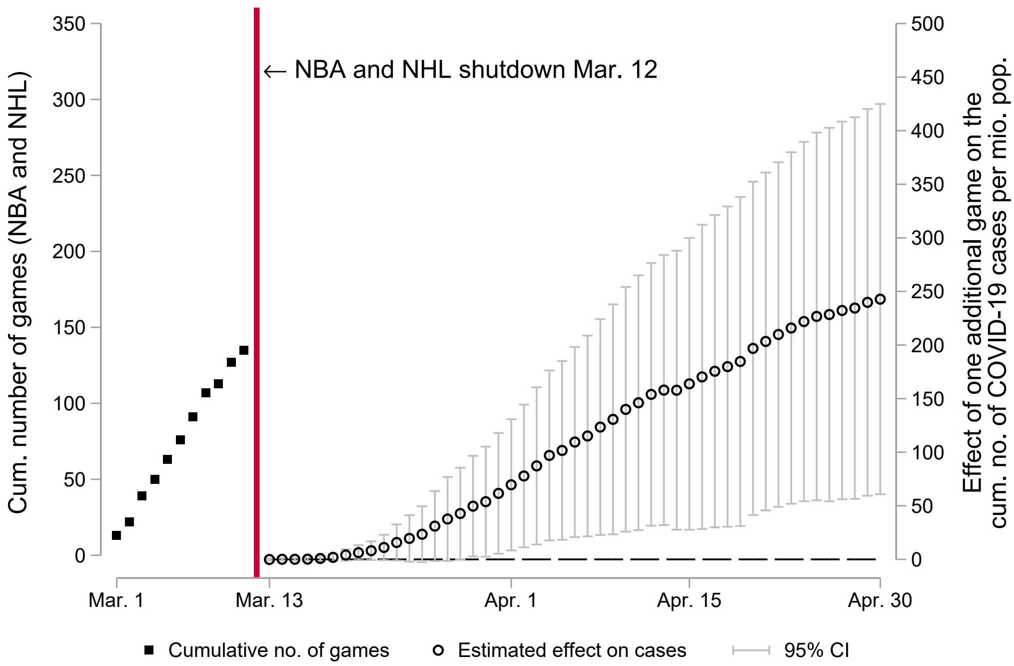
FIGURE 1 — NBA and NHL venues and adjacent countries in the United States



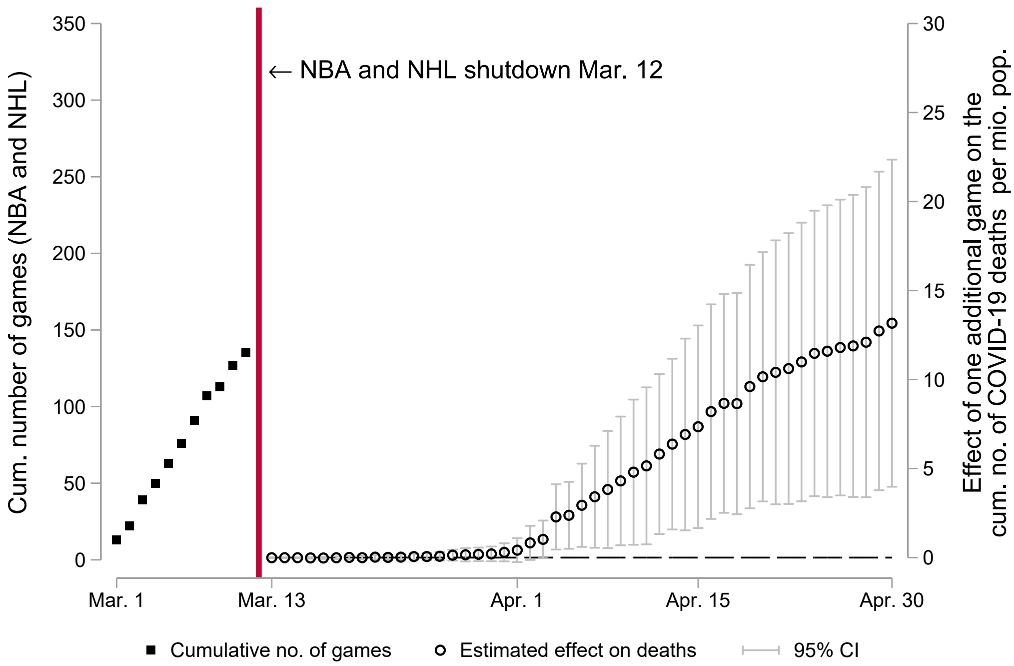
Notes: This map provides an overview on the counties we use in our analysis. Counties where venues are located are marked with red dots. The light-gray shaded counties are adjacent to either a NBA or a NHL venue, the dark-gray shaded counties are in the perimeter of both a NBA and a NHL venue.

Figure 2 — Estimated effect of sport games on confirmed deaths and cases.

(a) Reported cases per million population



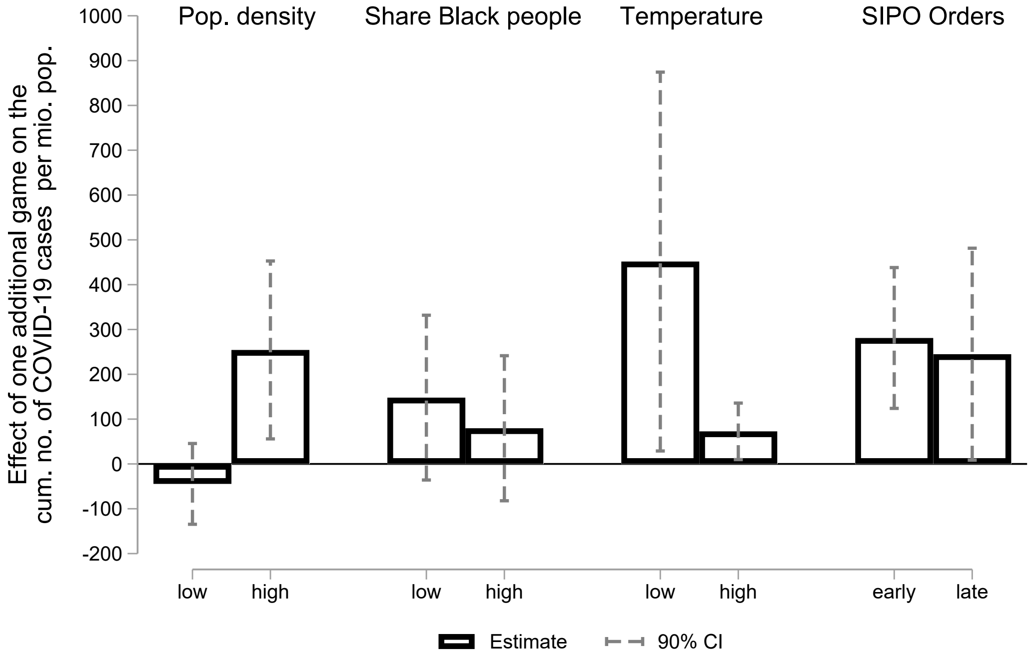
(b) Reported deaths per million population



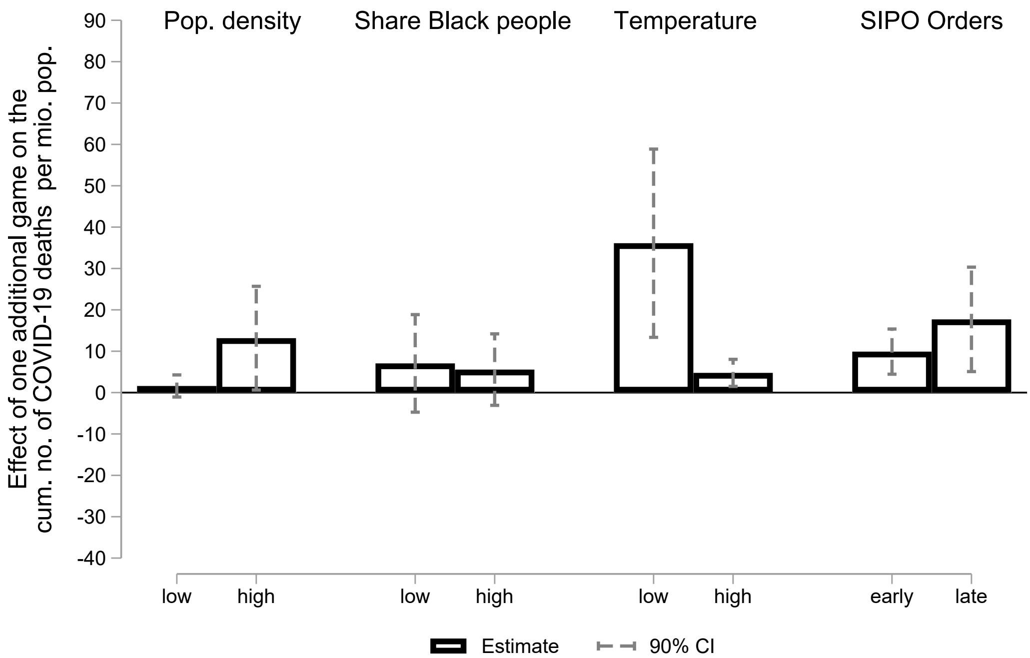
Notes: The squares indicate the cumulative number of games (NBA plus NHL) before suspension of the leagues (red vertical line). The hollow circles give the estimated effect of one additional game on the cumulative number of COVID-19 cases (Panel A) and deaths (Panel B) on each day between March 13 and April 30. Each estimate comes from a separate regression, with the dependent variables being measured on different days, and the control variables being the same as in column (4) of Table 1.

Figure 3 — Treatment effect heterogeneity

(a) Reported cases per million population



(b) Reported deaths per million population



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, Friedson, Matsuzawa & Sabia (2020). States without statewide SIPO regulations (MA, MN, OK, TN, and UT) are coded 0.

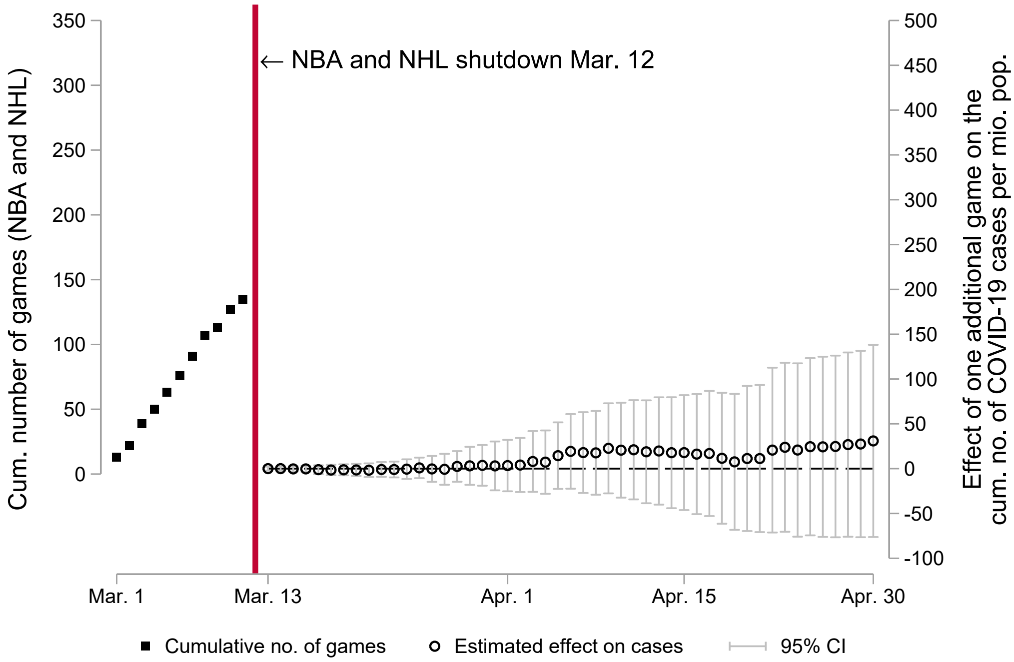
Table 1 — Impact of pre-scheduled mass gatherings on COVID-19 infections and death rate

	(1)	(2)	(3)	(4)
Panel A. Cumulative number of confirmed COVID-19 infections per million population[a]				
Cum. number of games	243.745	305.815**	283.996**	242.823***
	(208.635)	(123.462)	(111.576)	(92.313)
Population density			0.380**	0.493***
			(0.150)	(0.157)
Perc. non-white population			60.827**	
			(23.731)	
Perc. population aged 60+			−102.025*	
			(52.980)	
Perc. population female			67.810	
			(363.291)	
Sex-race-age distribution[d]	No	No	No	Yes
State fixed effects	No	Yes	Yes	Yes
Venue county (1 = yes, 0 = no)	No	No	Yes	Yes
Panel B. Cumulative number of COVID-19 deaths per million population[b]				
Cum. number of games	13.349	15.227**	13.817**	13.168***
	(11.871)	(6.389)	(5.380)	(4.659)
Population density			0.013	0.024**
			(0.008)	(0.010)
Confirmed COVID-19 cases[c]			2.940*	2.765*
			(1.496)	(1.655)
Perc. non-white population			2.317*	
			(1.365)	
Perc. population aged 60+			−1.140	
			(2.601)	
Perc. population female			22.435	
			(15.474)	
Sex-race-age distribution[d]	No	No	No	Yes
State fixed effects	No	Yes	Yes	Yes
Venue county (1 = yes, 0 = no)	No	No	Yes	Yes

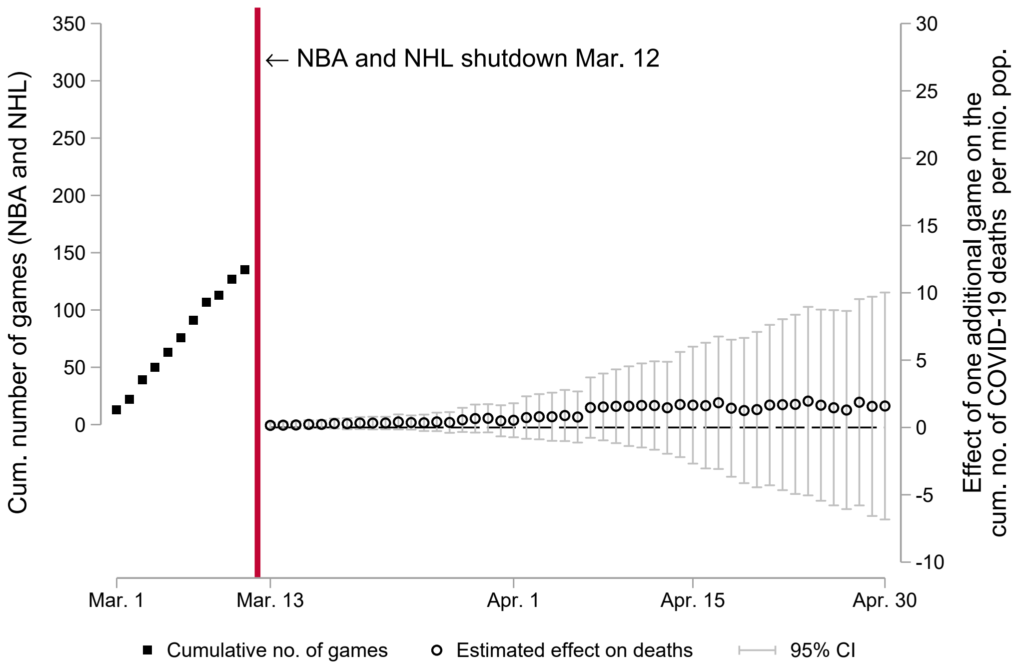
Notes: The number of observations is 242. Robust standard errors are presented in parentheses, stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications are weighted for the total population in a particular county. [a] The dependent variable in Panel A is the number of confirmed COVID-19 infections per million inhabitants on April 30, 2020 with a mean of 2,940.56 (std. dev. 4,690.66). [b] 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 147.91 (std. dev. 262.45). The number of games measures all NBA and NHL games which took place between March 1 and March 12. [c] The number of confirmed COVID-19 infections per county and 1,000,000 county residents on March 13. [d] 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+.

Figure 4 — Placebo test: Reassigning teams to non-venue counties and applying game schedule

(a) Reported cases per million population



(b) Reported deaths per million population



Notes: These figures presents estimates from an in-space placebo test, where we reassign venues to counties not in our data that are closest in terms of observables to the venue counties, and apply the leagues' game schedules to the matched placebo counties. We use the following observable characteristics: total population, population density, share of females, share of nonwhites, and share of elderly. See Section 6 for more information. The squares indicate the cumulative number of games (NBA plus NHL) before suspension of the leagues (red vertical line). The hollow circles give 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 the dependent variables being measured on different days, and the control variables being the same as in column (4) of Table 1.

Web Appendix

This Web appendix (not for publication) provides additional material discussed in the unpublished manuscript "Mass Gatherings Contributed to Early COVID-19 Mortality: Evidence from US Sports" by Alexander Ahammer, Martin Halla, and Mario Lackner.

TABLE A.1 — Descriptive statistics for main variables

	By county type	
	Venue	Perimeter
Total number of games (NBA + NHL) between Mar. 1 - 11	3.74	3.42
	(3.06)	(2.62)
Cumulative number of COVID-19 infections†		
On March 13	9.08	4.90
	(14.18)	(13.28)
On April 30	3,846.71	2,771.76
	(4,425.77)	(4,729.53)
Cumulative number of COVID-19 deaths†		
On March 13	0.05	0.02
	(0.20)	(0.22)
On April 30	211.30	136.10
	(329.72)	(247.11)
Total county population (in mio.)	1.79	0.38
	(1.79)	(0.45)
Population density	3,506.84	822.51
	(3,915.39)	(1,583.92)
Population characteristics		
% female	51.20	50.57
	(0.85)	(1.10)
% non-white pop.	33.50	16.93
	(13.43)	(12.80)
% pop. 60+	18.46	21.68
	(2.52)	(5.33)
Number of counties	38	204

Notes: Sample means with standard deviations in parentheses.

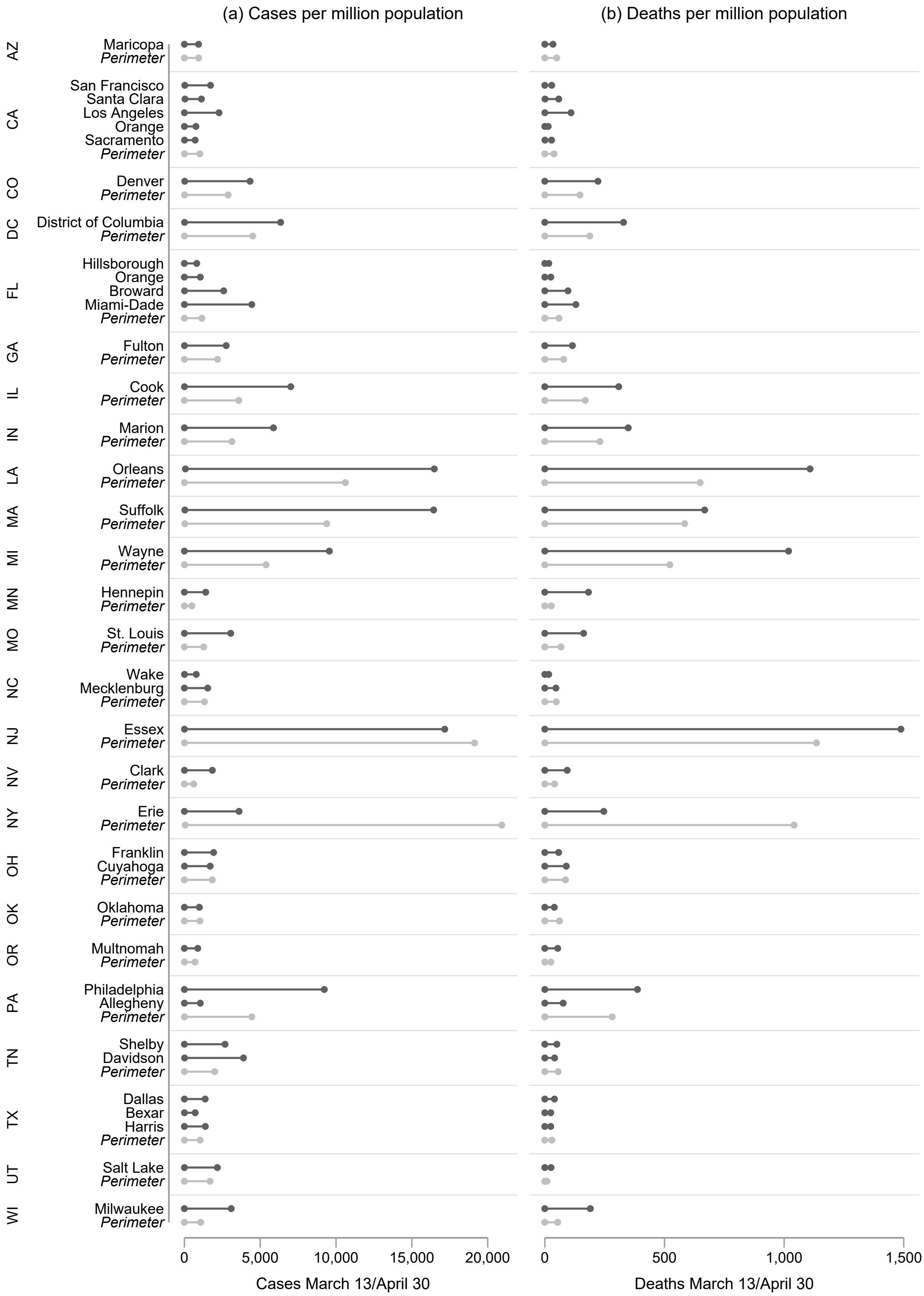
† Per one million population.

TABLE A.2 — Correlation of the number of NBA and NHL games between March 1 and March 11 with observed county characteristics

	Venue counties (1)	All counties (2)
Population density	−0.0005	0.0003
	(0.0003)	(0.0002)
Share male, white, 20–39	20.727	−1.849
	(13.567)	(2.621)
Share male, white, 40–59	11.795	−1.915
	(12.780)	(3.889)
Share male, white, 60+	16.726	−2.498
	(13.428)	(2.055)
Share male, non-white, <20	56.351**	0.620
	(23.373)	(3.615)
Share male, non-white, 20–39	7.285	−1.188
	(12.108)	(3.317)
Share male, non-white, 40–59	15.308	−7.565
	(17.065)	(5.209)
Share male, non-white, 60+	17.383	1.712
	(13.664)	(5.211)
Share female, white, <20	34.337	−2.874
	(26.370)	(4.476)
Share female, white, 20-39	14.738	−1.212
	(12.368)	(2.145)
Share female, white, 40–59	23.455	−1.070
	(14.448)	(2.063)
Share female, white, 60+	17.258	−0.919
	(12.644)	(2.826)
Share female, non-white, <20	−28.196	−6.000
	(17.411)	(4.258)
Share female, non-white, 20–39	25.471	−2.203
	(16.127)	(2.887)
Share female, non-white, 40–59	26.287**	6.390*
	(10.810)	(3.402)
Share female, non-white, 60+	15.691	−4.494
	(14.903)	(4.478)
Max. Temperature April	−0.107	0.146
	(0.071)	(0.200)
State fixed effects	No	Yes
Venue county	No	Yes
Number of Observations	38	242

Notes: In this table we test whether the cumulative number of NBA and NHL games between March 1 and March 12 is correlated with observable county characteristics. In column (1) we only consider the 38 venue counties, in column (2) the sample consists of all 242 venue and perimeter counties. For the latter, we cluster standard errors on the state level. Stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

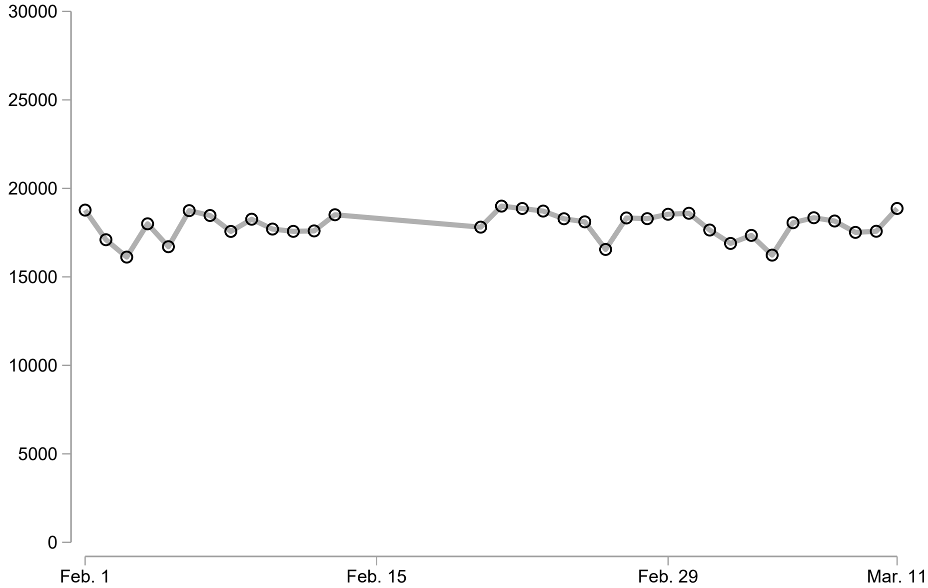
Figure A.1 — Change in reported COVID-19 cases and deaths per venue and perimeter county over time



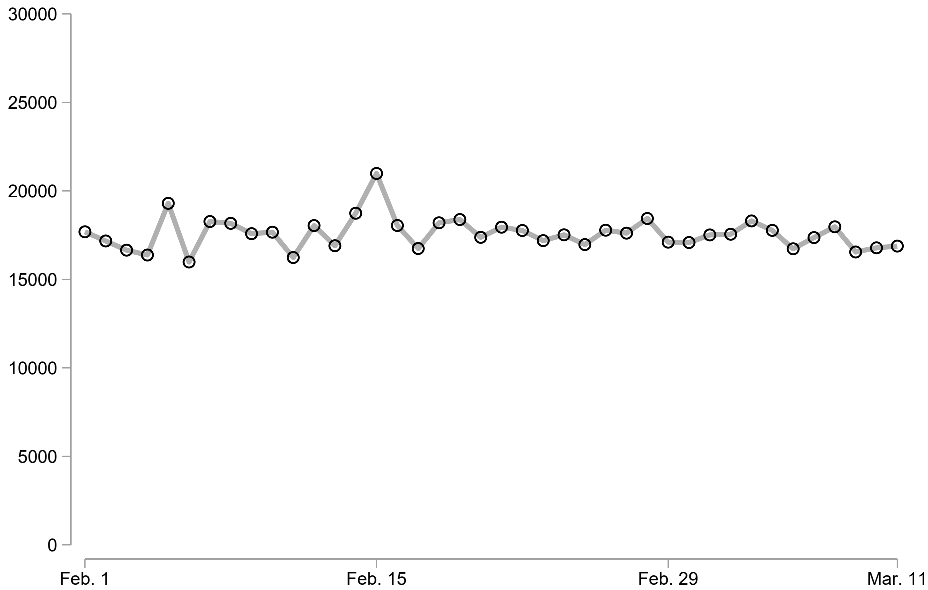
Notes: This figures displays the evolution of the number of reported COVID-19 cases (a) and deaths (b) for each venue county in our data. The left dot is always the number of cases per million population on March 13, while the right dot is the number of cases per million population on April 30. Additionally, 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.'

Figure A.2 — Average ticket sales by NBA and NHL game between February 1 and March 11

(a) NBA



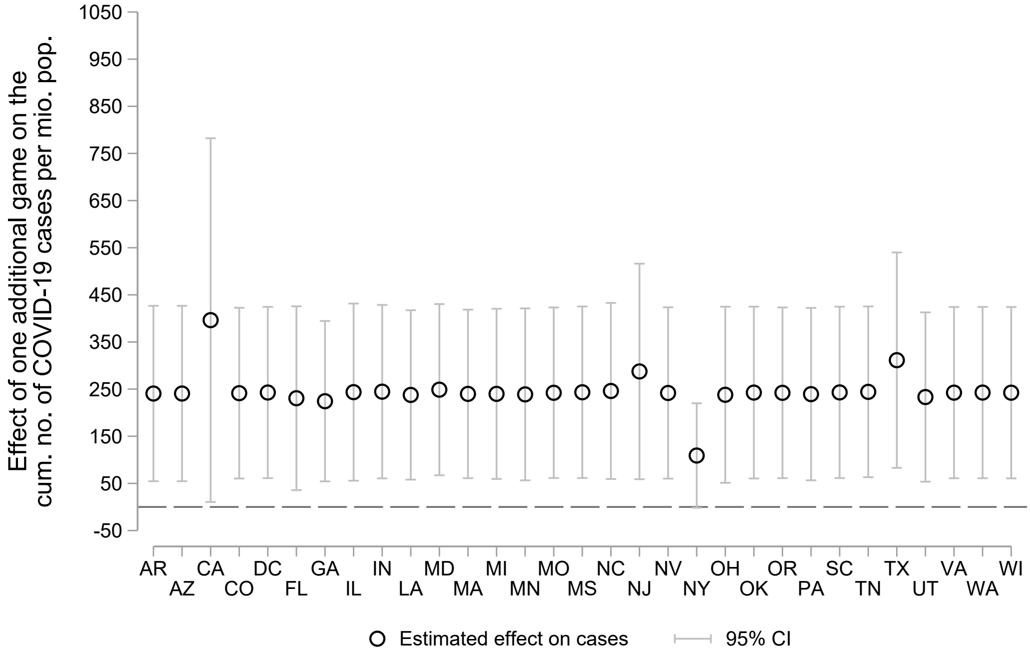
(b) NHL



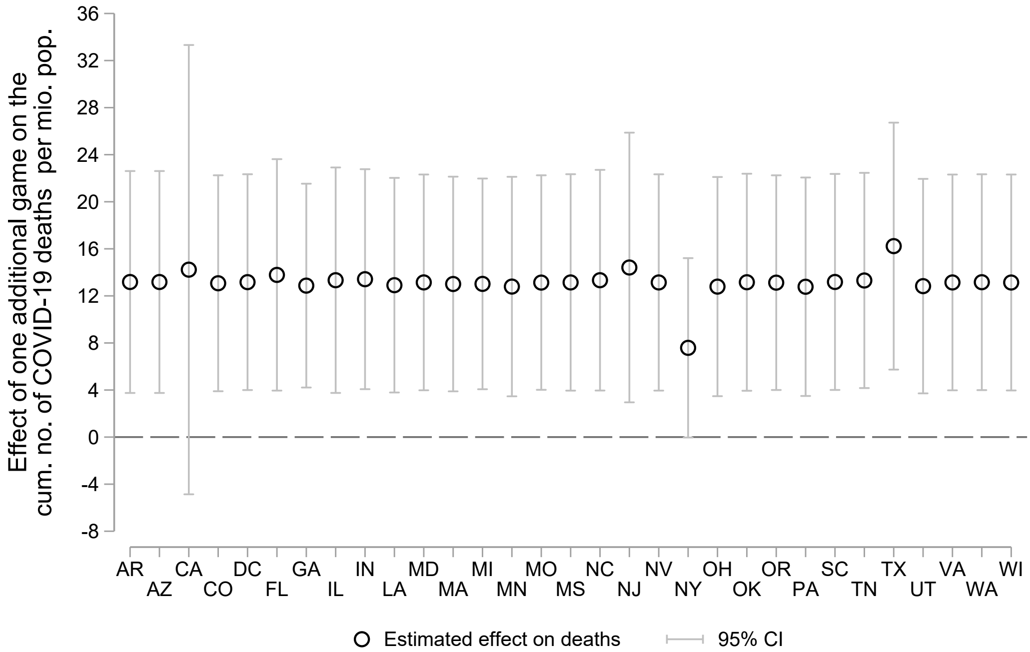
Notes: This figure displays the average number of tickets sold for all NBA and NHL games held between February 1 and March 11 in venues located in the US. On March 12, the NBA canceled two games before tip-off. After March 11, both leagues suspended their seasons indefinitely.

FIGURE A.3 — Robustness check: Leave-one-out sampling

(a) Reported cases per million population

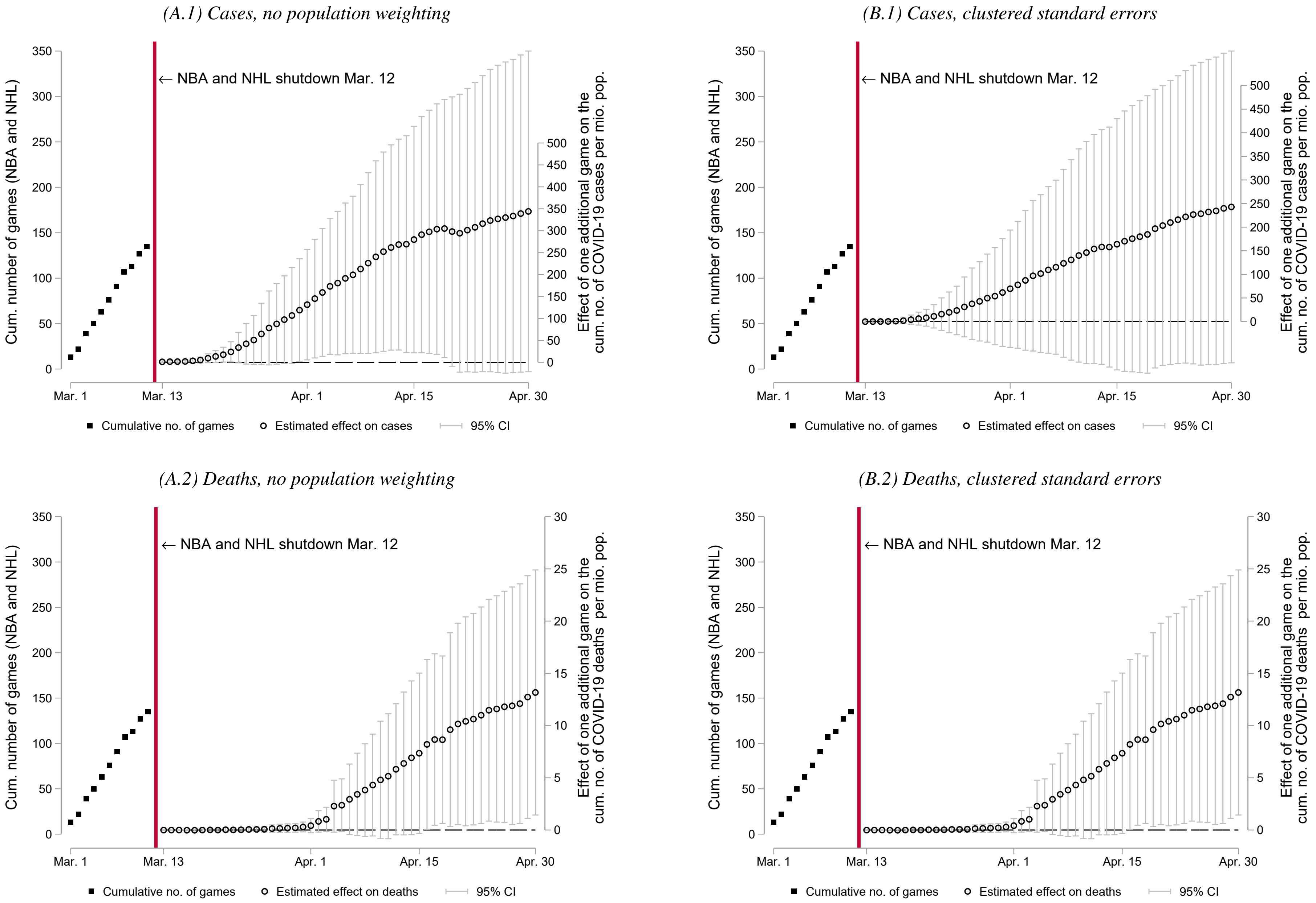


(b) Reported deaths per million population



Notes: For this robustness check we estimate our model 31 times, each time leaving out one of the US states. For each left-out state on the horizontal axis, we then provide the estimated effect on COVID-19 cases (panel A) and deaths (panel B). In particular, the hollow circles represent the estimated effect of one additional game on the cumulative number of cases in Panel A and deaths in Panel B by April 30. The control variables are the same as in column (4) of Table 1.

Figure A.4 — Robustness check: No population weights (left) and clustered standard errors (right)



Notes: These figures show how our regression results change when we use no population weights (left-hand side) and clustered standard errors (right-hand side), for both cases (Panels A.1 and A.2) and deaths (Panels B1. and B.2). The squares indicate the cumulative number of games (NBA plus NHL) before suspension of the leagues (red vertical line). The hollow circles give the estimated effect of one additional game on the cumulative number of COVID-19 cases and deaths 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, and the control variables being the same as in column (4) of Table 1.