



Mass gathering events and the spread of infectious diseases: Evidence from the early growth phase of COVID-19

Rafael González-Val^a, Miriam Marcén^{b,*}

^a Universidad de Zaragoza & Institut d'Economia de Barcelona (IEB), Spain

^b Universidad de Zaragoza, IEDIS & GLO, Spain

ARTICLE INFO

Jel Classification:

I12

I18

J10

Keywords:

COVID-19

Mass gathering events

Cases

Deaths

Spain

ABSTRACT

This paper studies the impact on reported coronavirus 2019 (COVID-19) cases and deaths in Spain resulting from large mass gatherings that occurred from March 6 to March 8, 2020. To study these outcomes, the geographic differences in the planned pre-pandemic major events that took place on these dates were exploited, which is a quasi-random source of variation for identification purposes. We collected daily and detailed information about the number of attendees at football (soccer) and basketball matches in addition to individuals participating in the Women's Day marches across Spain, which we merged with daily data on reported COVID-19 cases and deaths at the provincial level. Our results reveal evidence of non-negligible COVID-19 cases related to the differences in the percentage of attendees at these major events from March 6 to March 8. In a typical province, approximately 31% of the average daily reported COVID-19 cases per 100,000 inhabitants between mid-March and early April 2020 can be explained by the participation rate in those major events. A back-of-the-envelope calculation suggests that this implies almost five million euros (169,000 euros/day) of additional economic cost in the health system of a typical province with one million inhabitants in the period under consideration. Several mechanisms behind the spread of COVID-19 are also examined.

1. Introduction

This work evaluates the contribution of mass gathering events to the spread of infectious diseases and their associated economic costs.¹ For small gatherings, such as those that occur in schools, evidence in the economic literature of a causal link between openings/closures of schools and several viral diseases (influenza, gastroenteritis, and chickenpox; [Adda, 2016](#)) can be found. In the epidemiological literature, [Hoang and Gautret \(2018\)](#), [Karami et al. \(2019\)](#), and [Rainey et al. \(2016\)](#) wrote systematic reviews of mass-gathering-related disease outbreaks in observational studies. The impact of mass gatherings on coronavirus 2019 (COVID-19) has been poorly understood, especially in the early stages of the pandemic ([Nunan and Brassey, 2020](#)), and more work is needed on this topic in the economic literature ([Murray, 2020](#)).²

To our knowledge, causal evidence in the early stages is limited to two studies concerning United States (US) mass sports gathering events, which show that one additional sporting event (National Basketball Association [NBA] or National Hockey League [NHL]) caused an increase in the cumulative number of COVID-19 deaths by 11% ([Ahammer et al., 2020](#)), or 7520 additional cases and 658 deaths during the first wave of the COVID-19 pandemic ([Carlin et al., 2021](#)). We examined this issue in more depth by extending the different types of mass gathering events and focusing our attention on the attendance effects that were only partially analyzed in [Carlin et al. \(2021\)](#). To examine this issue in more detail, we exploited the plausibly exogenous role of the geographic variation in pre-pandemic planned mass gathering events that took place from March 6 to March 8, 2020, on COVID-19 cases and deaths in Spain. This period occurred three days before the World Health

* Correspondence to: Facultad de Economía y Empresa, Universidad de Zaragoza, Gran Vía 2, 50005 Zaragoza, Spain.

E-mail address: mmarcen@unizar.es (M. Marcén).

¹ The World Health Organization (WHO) considers an event to be a mass gathering when “the concentration of people at a specific location for a specific purpose over a set period of time and which has the potential to strain the planning and response resources of the country or community” (WHO, 2015).

² Recent medical evidence on the transmission of the coronavirus-2019 (COVID-19), based on experimental indoor mass gathering event infection information from indoor mass gathering events, shows that infections significantly depend on the quality of the ventilation system and hygiene practices ([Moritz et al., 2021](#)).

Organization (WHO) declared a pandemic and was a week before the nationwide lockdown.³

Spain was one of several epicenters of the world COVID-19 pandemic during the spring of 2020. In April 2020, reported cases reached 7% (approximately 200,000, April 2020) and contributed to over 10% of deaths (approximately 23,000) worldwide in a country with 0.61% (47 million) of the world population (Amuedo-Dorantes et al., 2021; Redondo-Bravo et al., 2020). A year later (April 2021), those figures were much lower, with Spanish reported cases representing 2.5% (3.3 million) and 2.6% (approximately 76,000) of deaths worldwide.⁴ The early growth phase of the spread of COVID-19 in Spain was so rapid that the Spanish government had to declare a nationwide lockdown only 17 days after the first reported local case (Redondo-Bravo et al., 2020); the spread was also quite heterogeneous among regions (Kochańczyk et al., 2020). The question can be asked: “What happened during those days that could, at least in part, explain the quite rapid spread of COVID-19 in Spain?” The observational evidence points to super-spreading events as a possible source of the initial outbreaks in some particular regions (Kochańczyk et al., 2020).

Our empirical strategy addresses the challenge of studying the role of mass gathering events by exploiting the quasi-random differences in the geographical locations and in the number of attendees at pre-pandemic planned mass gathering events. The Spanish case is of interest in this context for several reasons. As mentioned previously, this country was one of the hardest hit at the beginning of the COVID-19 pandemic, with considerable differences across regions. This case guarantees a high degree of geographic variation in the mass gathering events and in the number of attendees, which we exploited for identification purposes. Not only were major sporting events considered, as in the US case (Ahammer et al., 2020; Carlin et al., 2021), but also the Women’s Day marches, when women took to the Spanish streets in early March 2020. These events are of interest since both are mass gathering events; however, the way in which people participated (sports events with assigned seating versus women’s marches with people in motion) and the kind of people participating in both of them could be quite different (with important gender differences). Most of the largest women’s marches in Europe have been located in Spain, with the feminist movement mobilizing tens of thousands of people in their street demonstrations with the approval of most of the Spanish political parties, both on the right and left.⁵

Spain provides a scenario with almost no differences in pre- and post-events. One week before the nationwide lockdown, individuals did not privately adjust their activities to avoid risk, overlooking the environmental risk (Amuedo-Dorantes et al., 2021). In the pre-lockdown week, despite the distressing news coming from China, Iran, and even Italy, the Spanish authorities believed that COVID-19 was still distant enough that it was not an immediate concern. The Spanish head of medical emergencies, in several press conferences held after January 31, 2020, claimed that “Spain will only have a handful of cases” (Tremlett, 2020) and even recommended during a press conference the participation of his own son in the massive demonstrations to mark International Women’s Day on March 8. Possible heterogeneity in the regional social

distancing measures that were applied is not an issue in the Spanish case.⁶ The strict lockdown, which is the most effective social-distancing measure to contain the spread of the virus (Amuedo-Dorantes et al., 2021; Fowler et al., 2020; Siedner et al., 2020), was unexpectedly imposed nationwide by an extraordinary meeting of the Council of Ministers on Saturday, March 14, 2020, only one week after the mass gathering events under consideration here took place. This event generated a similar disruption of mobility across all Spanish regions (Amuedo-Dorantes et al., 2021).

We compiled detailed daily information on the number of attendees at football (soccer) matches, including First and Second Division, and basketball matches, First Division, in addition to the individuals participating in the Women’s Day marches across Spain, which we merged with daily data on reported cases of COVID-19 and deaths at the province level.⁷ In our analysis, we account for time-invariant province idiosyncrasias via the use of provincial fixed effects and temporal trends by including fixed-date effects in addition to province-specific linear time trends. Using these methods, we captured pre-existing differences in testing and/or therapeutic availability differences that could also have affected differential changes in COVID-19 cases and deaths (Amuedo-Dorantes et al., 2021). Consistent with the epidemiological observational evidence mentioned above, in which mass gatherings are events that could amplify the virus transmission and potentially disrupt the host country’s response capacity, we found a relationship between attendance at mass gathering events and the COVID-19 cases by exploiting the plausibly exogenous variation in the geographical locations and in the number of attendees at pre-pandemic planned mass gathering events. Specifically, on average, approximately 31% of the average daily reported COVID-19 cases per 100,000 inhabitants from mid-March to early April 2020 could be related to the participation rate in mass gathering events during the period from March 6 to March 8, 2020, in a typical province with an average number of attendees of 2.29 per hundred inhabitants.⁸ These findings are robust across different specifications and subsamples.

Mass gathering events undoubtedly produced an effect on economic costs for the healthcare system. Assuming the tariff system established by the Government of Catalonia on time of hospitalization in hospitals with/without ICUs for patients with COVID-19 (Decree law 12/2020), a back-of-the-envelope calculation indicates that the estimated impact of the mass gathering events on COVID-19 cases generated almost five million euros (169,000 euros/daily) of additional economic costs in the healthcare system of a typical province with one million inhabitants in the period under consideration.⁹ Another way to summarize this effect is to aggregate it throughout the total healthcare system at the country level; this summary reached roughly 230 million euros, representing

⁶ In the US, huge state discrepancies in the intensity of social-distancing measures taken in March and April 2020 and in the response of the population (Amuedo-Dorantes et al., 2022; Marcén and Morales, 2021) were found, which generated additional difficulties with respect to clearly identifying the impact of mass gathering events in that country.

⁷ The Spanish territory is divided into 17 autonomous regions (NUTS 2 regions). Those regions are divided into 50 provinces (NUTS 3 regions), which is the spatial level considered in this analysis. We excluded two autonomous cities (Ceuta and Melilla) located on the African coast.

⁸ Note that comparisons with prior literature are not easy, since those studies on the US cases focus their analysis on the marginal impact of one additional sports game (Ahammer et al., 2020; Carlin et al., 2021), while we are considering the number of participants at the mass gathering events. The attendance effects presented in Carlin et al. (2021) should be taken with caution because the authors indicate they do not have information on true attendance.

⁹ Decree law 12/2020, April 10, 2020: “Decreto Ley 12/2020, de 10 de abril, por el que se adoptan medidas presupuestarias, en relación con el Sistema sanitario integral de utilización pública de Cataluña, en el ámbito tributario y en la estructura de la Administración de la Generalidad, para paliar los efectos de la pandemia generada por la COVID-19.”

³ See the WHO declaration: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020>, updated April 2021.

⁴ See the Johns Hopkins Coronavirus Resource Center data in <https://coronavirus.jhu.edu/map.html>, updated April 2021.

⁵ See EL PAÍS, March 9, 2020. <https://english.elpais.com/society/2020-03-09/womens-day-marches-in-spain-attract-mass-numbers-despite-coronavirus-fears.html>

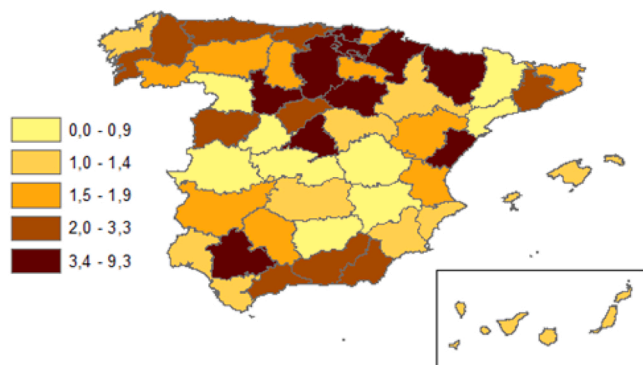


Fig. 1. % Attendees to mass gathering events (March 6–8) over total population at the provincial level. Notes: Lighter colors correspond to lower % attendees to mass gathering events (March 6–8) in each province.

approximately 0.2% of the total spending in the Spanish healthcare system (more than 100,000 million euros, including public/private healthcare systems, according to the Spanish Ministry of Health).¹⁰ While such back-of-the-envelope calculations must be cautiously interpreted, they are informative for understanding the potential magnitude of the effect of mass gathering events on the healthcare system.

A challenge in estimating the effects of mass gathering events by exploiting plausibly exogenous variation in the geographical locations and in the number of attendees to pre-pandemic planned mass gathering events is that such events do not occur in a vacuum. Economic incentives for risk-avoiding behaviors by nonparticipants may play a role, as seen in the case of the 2020 Black Lives Matter protests that unexpectedly did not reignite the US community-level COVID-19 growth (Dave et al., 2020). During the Black Lives Matter protests, nonparticipants appeared to adjust their risk-taking behavior based on their perceived exposure to risk, which is in line with risk-compensating behavior (Dave et al., 2020). If this compensating risk behavior is adopted in a different way across Spanish regions, a simple correlation between the participants in the mass gathering events and reported COVID-19 cases and deaths might result in biased estimates. Although no causal link was found between mobility behavior related to daily activities in the pre-lockdown period and the spread of COVID-19 in Spain (Amuedo-Dorantes et al., 2021), the decrease in the number of participants in the Women's Day marches in comparison to prior years yielded non-negligible data, which may indicate some degree of risk-avoiding behaviors.^{11,12}

While, admittedly, some degree of compensating behavior is present, our concern is limited to whether that behavior is random across the regions or may affect our estimates. To address this concern, we analyzed the following: (1) a supplement of our primary analysis with an event study examining how COVID-19 cases and deaths responded to participation in the mass gathering events and (2) placebo regressions randomly distributing the percentage of attendees at the mass gathering events across Spanish provinces; if we are truly capturing the mass gathering event's effect, we should obtain no significant and close to

¹⁰ See https://www.mscbs.gob.es/estadEstudios/portada/docs/DATOS_SNS_A4_112020.pdf, updated April 2021.

¹¹ Some evidence on possible adjustments in daily activities to avoid COVID-19 risk when the disease became more prevalent in the US (Gupta et al., 2020) can be found, although that observed behavior could also be a response to other social-distancing measures that occurred across the US and were omitted in many early papers that limited their analysis of stay-at-home orders/business closures (Marcén and Morales, 2021).

¹² See EL PAÍS, March 9, 2020. <https://english.elpais.com/society/2020-03-09/womens-day-marches-in-spain-attract-mass-numbers-despite-coronavirus-fears.html>.

zero estimated coefficients. Additionally, since the evolution of COVID outcomes appears to be related to several geographical variables such as income, population density, health expenditure, demographic profile, and weather, among others (Armillei et al., 2021), we tested whether the participation rate was correlated with various observables, such as socio-economic and weather province characteristics. We also reinforced our analysis by including those variables as controls, which helped us mitigate omitted variable bias concerns.

Since COVID-19 is transmitted by close contact, we also explored to what extent a mass gathering event's impact could be amplified/reduced by considering three possible mechanisms: (1) participants in Women's Day marches versus those in sports events; (2) socio-economic variables; and (3) weather and location variables. First, we separate the types of mass gathering events to assess the role of the different mass gathering events on the spreading of the virus. Second, we studied the importance of socio-economic variables in the amplification of the mass gathering events' effects on the spread of COVID-19 by exploiting the differences in various socio-economic characteristics by province. Third, we examined whether or not the weather conditions prevented or amplified the COVID-19 transmission generated by the mass gathering events.

The rest of the paper is organized into several sections: Section 2 provides a description of the data sets used in the analysis; Section 3 presents the empirical strategy; our main findings are analyzed in Section 4, in which we also discuss the economic costs to the healthcare system and present the results from various identification and robustness checks; Section 5 explores some of the key mechanisms at play; and Section 6 summarizes the paper.

2. Data

2.1. Location and attendance at mass gathering events

To assess how the mass gathering events affected reported COVID-19 cases and deaths, we collected detailed data on the location, place, and number of attendees at the major planned pre-pandemic sporting events and the Women's Day marches that took place across Spain between March 6 and March 8. We considered major sports events related to football and basketball, the main form of sports entertainment in Spain. We gathered data from the First and Second Divisions of the Spanish football league (*Liga Santander* and *Liga SmartBank*) and from the major league in basketball (*Liga Endesa*). The information was double-checked against the official statistics of the organizers of those major events, *La Liga* (football) and *ACB* (basketball), and from local newspapers.¹³ All these sporting events mobilized thousands of individuals, not only inside the city in which the event took place but also in the surrounding area inside the province where the team supporters are more likely to be located, according to the Spanish Sociological Research Center.¹⁴ The sporting event with the most attendees took place in Barcelona (*Barcelona* versus *Real Sociedad*), with more than 77,000, but in the province of Madrid, six sports games took place during the period under consideration with more than 103,000 aggregate attendees (roughly 1.6% of the province inhabitants). Alongside this, we compiled publicly available data on the Women's Day marches across Spain from the Spanish Government Regional Offices (*Delegaciones de Gobierno*) at the province-level; when the Regional Offices did not disaggregate the number of

¹³ See <https://www.laliga.com/estadisticas> and <https://www.acb.com/>.

¹⁴ The Spanish Sociological Research Center (*Centro de Investigaciones Sociológicas*) included questions in its monthly values survey (*Barómetro*) about the closeness to a football team and activities that were carried out due to proximity to a team. Many of the supporters of the football teams live in the province where their team is located. Additionally, roughly 32% of respondents attended football matches of their preferred football team. See http://www.cis.es/cis/opencm/ES/1_encuestas/estudios/ver.jsp?estudio=14090.

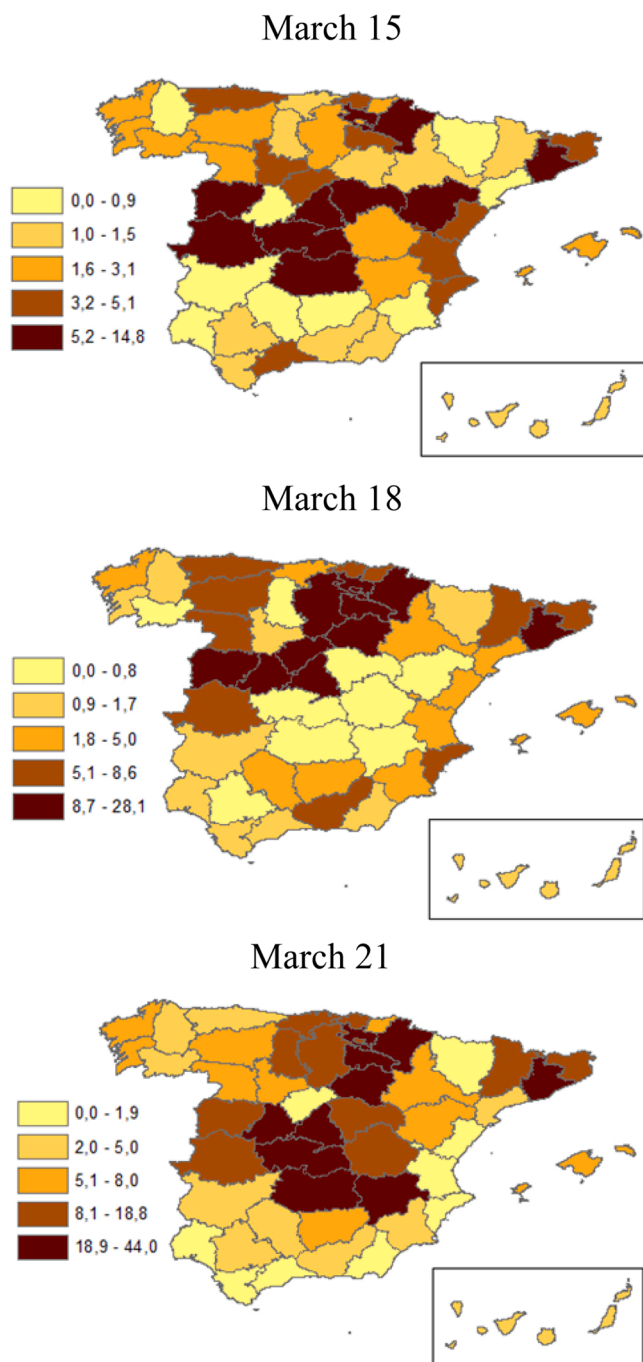


Fig. 2. Number of COVID-19 cases per 100,000 inhabitants at the provincial level. Notes: Fig. 2 shows the evolution of COVID-19 cases per 100,000 inhabitants at the province level. Lighter colors correspond to lower COVID-19 cases in each province and on each day.

participants at that administrative level, we were able to obtain that information from local newspapers. All provinces had Women’s Day marches, with the number of participants varying from 0.3% of the population in Badajoz to 6.4% of the population in Vizcaya.

Fig. 1 documents significant cross-province differences in the number of aggregate attendees at the sports events and the Women’s Day marches per hundred province inhabitants, ranging from 0.68% in

Toledo to 9.3% in Vitoria.¹⁵ The attendance rate averaged 2.29, with a standard deviation of 1.79. Spatial discrepancies are clearly observed, with lighter colors corresponding to lower levels of attendees per hundred inhabitants in the south of Spain (with the exception of the southern provinces of Sevilla, Málaga, and Almería) and higher levels in the central-north, especially in Madrid, Navarra, and in some provinces of Aragón, Castilla y León, and País Vasco. Ten provinces with more than 3.4% of attendees across the province population participating in the mass gathering events between March 6 and 8, and ten were below 0.9%.¹⁶

2.2. Reported cases and deaths

We use data on COVID-19 cases and deaths provided by the Spanish Ministry of Health and Regional Governments.¹⁷ The dataset contains information on the daily accumulated number of reported COVID-19 cases and deaths, from which we computed the daily incidence by province and date. Using province level data on the population from the Spanish Statistical Office for the year 2019, we next calculated the reported COVID-related cases and deaths per 100,000 inhabitants. Our dataset spans from February 26 (date of the first local case detected) to April 12, 2020 (five weeks after the mass gathering events). Note that the Spanish Ministry of Health changed the definition of COVID-19 confirmed cases after April 17, generating a break in the time series (Amuedo-Dorantes et al., 2021). Only cases confirmed by the Spanish authorities were included. It should be noted that, because of the shortage of diagnostic tests during the first wave, our analysis is limited to the severity of reported cases (Redondo-Bravo et al., 2020; Richterich, 2020). Roughly 45.4% of the confirmed and reported cases up to April 2020 required hospital admission (7.1% required mechanical ventilation and 4.6% ICU), and 11.9% died in Spain (Redondo-Bravo et al., 2020). This finding is not comparable to the second and subsequent waves since the testing process has been improved in Spain (Soriano et al., 2021).

Fig. 2 shows the distribution of reported COVID-19 cases across Spanish regions between March 15 and 21. This distribution corresponds to the time window in which the initial effect of the mass gathering events on the number of reported COVID-19 cases should be observed, taking into account the incubation period until symptom onset and the period between symptom onset and diagnosis (in total, a minimum of seven days).¹⁸ On average, the number of reported COVID-19 cases each day was 12.75 per 100,000 inhabitants after March 15.¹⁹ Lighter colors correspond to lower levels of case rates. This figure

¹⁵ Data on the province population corresponds to data gathered from the Spanish Statistical Office (*Instituto Nacional de Estadística*).

¹⁶ Note that individuals could attend multiple mass gathering events, in which case the ratio of attendees did not directly correspond to the percent of individuals participating in a mass gathering event. In any case, the greater the number of participants at meetings, the greater the probability of being infected or, if already infected, of spreading the disease.

¹⁷ The data at the province level were compiled by the project Escovid19data; see <https://github.com/montera34/escovid19data>.

¹⁸ From incubation to symptom onset during the first wave in Spain, an average of four to five days (maximum of 14 days; CDC, 2021) was observed, and from symptom onset to diagnosis, six days on average were noted (the minimum observed was three days; Redondo-Bravo et al., 2020). In the case of deaths, the time window expands considerably, which makes our challenge of identifying the link between the mass gathering events and the daily deaths more challenging. Since the days from symptom onset to death varied from 9 to 25 (Redondo-Bravo et al., 2020), the probable impact of mass gathering events fades over a wide period of time, a process that generates some difficulties in disentangling the impact of mass gathering events and that of other factors.

¹⁹ The daily deaths averaged 1.51 per hundred inhabitants after March 22 (a minimum of 14 days were required since COVID-19 exposure [March 6–8] to death).

Table 1
 COVID-19 cases and deaths response to the variation in attendees of mass gathering events (March 6–8).

D.V.:	(1)		(2)		(3)		(4)	
	Reported Cases				Reported Deaths			
	March 15		March 22		March 22		March 29	
Effect Since (Post Period):								
% Attendees	1.748*** (0.590)	1.740*** (0.605)	0.026 (0.059)	0.006 (0.053)				
Observations	2350	2350	2350	2350				
R-squared	0.707	0.707	0.684	0.684				
<i>In the Post Period</i>								
D.V. Mean	12.75	14.73	1.51	1.64				
D.V. Standard Deviation	14.63	15.78	1.61	1.61				
<i>For all</i>								
Province FE	Yes	Yes	Yes	Yes				
Date FE	Yes	Yes	Yes	Yes				
Province-specific linear trends	Yes	Yes	Yes	Yes				

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. The main explanatory variable is the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women’s Day marches (mean: 2.29; S.D.: 1.79). We estimate Eq. (1). Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

provides a sense of the spatial differences in the COVID-19 case rates, with higher case rates located in central-north Spain. Although this analysis is a descriptive one, the geographical pattern is quite similar to that observed in Fig. 1, in which the attendance rate is plotted. We plotted the case rates for three different days because of possible differences in testing availability for those days; the pattern does not show much variation.²⁰ Of course, more work is needed to deduce a link.

3. Empirical strategy

Our objective was to explore the extent to which mass gathering events impact the transmission of viral diseases and their health system-related economic costs. Our analysis relies on a natural experiment originating from the COVID-19 pandemic, along with the geographical discrepancies generated by the pre-pandemic planned mass gathering events in Spain. Our benchmark specification is given by the equation:

$$Y_{pt} = \beta \left(Attendees_p^{March\ 6-8} 1\{t \geq t^m\} \right) + \delta_t + \theta_p + \varepsilon_{pt} \tag{1}$$

in which Y_{pt} is the outcome (reported COVID-19 cases or deaths per 100,000 inhabitants) in province p at date t . The variation in the number of attendees is captured in $Attendees_p^{March\ 6-8}$, which gauges the number of attendees per 100 inhabitants (attendance rate) at mass gathering events in province p between March 6 and 8. Because of the lapse of time between exposure to COVID, symptom onset, diagnosis, and death, we considered several time windows. We define a dummy variable $1\{t \geq$

²⁰ Alternative indicators that describe the epidemiological situation on reported COVID-19 cases, such as a 7/14 day average number of cases, are not considered in our work. Note that the Spanish lockdown occurred only one week after the mass gathering events, so calculating the 7/14 day averages of cases could have influenced the real impact of the mass gathering effects because of the reduction in cases caused by the lockdown in a downward direction.

t^m) that takes value 1 since date t^m (the t^m considered are March 15 and March 22) and 0 otherwise.²¹ If the mass gathering events impact the pandemic’s outcomes (cases and/or deaths), we should expect β to take positive values, signaling that the higher the attendance rate, the higher the rate of reported cases/deaths associated with the pandemic.²² This specification includes a set of province and time (date) effects (δ_t , θ_p) that control for unincorporated factors that potentially affect the pandemic’s outcomes. The province fixed effects account for time-invariant dissimilarities across Spanish provinces, such as population differences (aging or density) or other traits that can be related to the COVID-19 pandemic’s outcomes (Amuedo-Dorantes et al., 2021; Aparicio Fenoll and Grossbard, 2020). The time (date) fixed effects, in addition to the set of province-specific linear trends, account for temporal evolution of pre-existing trends and varying province time factors, such as variations affecting the pandemic’s outcomes across provinces related to testing or hospital capabilities (Amuedo-Dorantes et al., 2021; Murray, 2020). ε_{pt} is the error term.²³

The validity of our estimates is based on the assumption that there is not an endogeneity problem. What matters in this situation in terms of inference purposes is the possible endogeneity with regard to the outcomes of interest. We checked this relationship in two ways: (1) conducting an event study and (2) running placebo regressions in Section 4.3. In the same section, we also tested whether the attendance rate correlated with observed province characteristics. We also reinforced our analysis by including those observed province characteristics as controls, which helped us mitigate omitted variable bias concerns.

When interpreting the estimated coefficient, some attenuation bias of the true effects of the mass gathering events could have been possible. The explanatory variable gauging the number of participants at the Women’s Day marches was a rough calculation of the true number of participants. Since this number was determined by the official Spanish government offices in each region, we expect that this would not have significantly affected our estimates because the measurement error should be in the same direction across provinces. To mitigate this concern, we repeated the analysis with different subsamples, as discussed in Sections 4.1 and 4.2. Additionally, the interpretation of the results is limited to the context of the first wave in Spain. Many of the officially reported cases of COVID-19 were severely affected (45.4% were hospitalized (Redondo-Bravo et al., 2020)). Although this discrepancy between waves is a limitation, the existing dataset allowed us to focus our attention on the impact of mass gathering events on a sample of quite severely affected officially reported cases of COVID-19. Estimates produced by Eq. (1) should be seen as a lower boundary of the true effects of mass gathering events.

4. Quantifying the impact of mass gathering events on the spread of COVID-19

4.1. Main findings

Table 1 presents the estimation for Eq. (1). This shows the impact of attendance at mass gathering events on both pandemic-related outcomes (reported cases and deaths). In the main analysis, the time span

²¹ The time span from infection to diagnosis is seven days and to death is 14 days (CDC, 2021; Lauer et al., 2020; Redondo-Bravo et al., 2020). We considered other time windows below.

²² We recognize that the use of different attendance rates (values of dosage as implicit in continuous treatment) would require an assumption of strong parallel trends (all units would experience the same evolution if given the same dose) to avoid selection bias. However, pre-trend tests commonly used to detect violations of the standard parallel trends are not useful in this setting (Callaway et al., 2021).

²³ Robust standard errors were clustered at the province level to account for any within-province correlation in the error terms.

from infection to diagnosis is fixed at seven days (effect since March 15 in column [1]) and to death at 14 days (effect since March 22 in column [3]), although we repeated the analysis considering, in addition, a time span from infection to diagnosis of 14 days (effect since March 22) in column (2) and to death at 21 days (effect since March 29) in column (4). We observed that the attendance rate significantly affected the number of reported COVID-19 cases. One standard deviation increase in the attendance rate implied a significant increase of 3.1 daily reported COVID-19 cases per 100,000 inhabitants (21% of the post-effect standard deviation).²⁴ In a typical province with an average attendance rate of 2.29%, approximately 31% (or four daily cases per 100,000 inhabitants) of the average daily reported COVID-19 per 100,000 inhabitants between mid-March and early April 2020 can be explained by the levels of participation in these major events.²⁵ The impact of the mass gathering events on the death rate is positive, albeit non-significant. As mentioned, the lapse of time from symptom onset to death varied from nine to 25 days in the first Spanish wave (Redondo-Bravo et al., 2020), preventing the possibility of finding a significant effect in this setting.²⁶ In addition, the hardest hit population in the first Spanish wave of COVID-19 were institutionalized elderly people, with 23.4% of the deaths among those older than 80 (37% for those older than 70); this age group is unlikely to participate in the mass gathering events considered here and, thus, less likely to be directly affected by participation. No significant differences after considering the alternative time lapses since exposure/infection to diagnosis/death were found in columns (2) and (4).

A comparison with prior studies on the early impact of mass gathering events on pandemic outcomes using National Basketball Association and National Hockey League (NBA and NHL, respectively) games is tricky because the studies estimate the marginal impact of an additional game (Ahammer et al., 2020; Carlin et al., 2021). It is true that Carlin et al. (2021) approximated a dose-response model in which the attendance based on ticket sales (not true attendance) is used as the main independent variable. Interestingly enough, they did not find a significant effect in all outcomes (cases and deaths) depending on the sports events. Those authors indicate that their attendance effects should be interpreted cautiously because of the imperfect proxy of attendance.²⁷ In any case, assuming the information on the average attendance (around 18,000) and considering that the US Metropolitan Statistical Areas with teams in those leagues have an average population of 4.082 million (Ahammer et al., 2020; Carlin et al., 2021) a back-of-the-envelope calculation equivalent to 0.44% of the population participating in a mass gathering event in Spain leads to an increase in the number of cases to 910 per 4.082 million inhabitants (0.77 daily cases per 100,000 inhabitants) in the period under consideration. This calculation is not close to the extra 7520 cases per MSA associated with an additional NHL/NBA suggested by Carlin et al. (2021) using game data from March 1–11.²⁸ It is worth noting that the analyses are not directly comparable since Carlin et al. (2021) use data from the database maintained by the

New York Times until May 31. By extending the impact of our mass gathering effects until the end of May, we calculated an impact of 2450 cases per 4.082 million inhabitants, a figure that is still not quite close to the findings provided by Carlin et al. (2021).²⁹ Several questions may arise at this point: (1) do differences in the spread of the COVID-19 depending on the type of mass gathering event (sports events versus demonstration) exist? (2) did the pre-March mass gathering events have any role? and (3) is the use of games rather than attendance rates inflating the effect of the mass gathering events? We discuss this in the mechanisms section.

Using the severity of the distribution of cases during the first wave in Spain, that is, 45.4% of hospital admissions and 4.6% of ICU admissions among all reported cases, we could roughly determine the health system-related economic costs of these mass gathering events. Considering the severity of the cases and our estimates, 1.8 of the four daily reported cases per 100,000 inhabitants that were related to participation in the March 6–8 mass gathering events in a typical province (with an attendance rate equal to 2.29%) were hospitalized, and 0.18 of those cases were in ICU. With the tariff system established by the Government of Catalonia (one of the main regions in Spain) addressing in-patient hospitalizations in hospitals with/without ICU (5000/43,400 euros respectively) for patients with COVID-19 (Decree law 12/2020), a back-of-the-envelope calculation suggests that the estimated impact of the mass gathering events on COVID-19 cases generated almost five million euros (roughly 169,000 euros/daily) of additional economic costs in the healthcare system of a typical province with one million inhabitants in the period under consideration (from March 15 to April 12). This calculation also incorporated the assumption that the tariff system in Catalonia coincides with that of a typical province. In any case, while these back-of-the-envelope calculations must be cautiously interpreted, the economic costs determined in this situation illustrate the importance of evaluating the risk factors associated with mass gathering events to prevent augmentation of health system problems.

4.2. Robustness checks

To assess the sensitivity of our findings, we conducted various robustness checks: (1) changing the characteristics of the specifications, (2) considering different time windows from incubation to diagnosis or death, (3) examining the impact on the excess mortality, and (4) varying the sample. Panel A in Table A1 in the Appendix shows the estimates of Eq. (1), excluding province-specific linear trends. A non-negligible impact of mass gathering events on COVID-19 cases per 100,000 inhabitants and a positive but non-significant impact on deaths can be observed. However, the magnitude of the coefficient was reduced by 18%, which may indicate that the coefficient in the specification without province trends captured some already detected differences in the province trends that caused a downward bias in our estimated impact (Amuedo-Dorantes et al., 2021).^{30,31} In Panel B, the estimates were obtained without clustering at the province level and without province trends. In this case, we were able to observe a significant impact on both pandemic outcomes. We prefer to be conservative and present all our

²⁴ This value is obtained using the calculation: Standard Deviation of the attendance rate (1.79) \times estimated coefficient on attendance rate (1.748) = 3.1 reported cases per 100,000 inhabitants, representing 21% of the S.D. in the post-effect period (since March 15 (exposure March 6–8) until diagnosis is seven days) = 3.1/14.63.

²⁵ The average attendance rate (2.29%) \times estimated coefficient on attendance rate (1.748) = 4 reported cases per 100,000, representing 31% of the average daily reported COVID-19 cases from March 15 to April 12 = 4/12.75.

²⁶ We revisited this issue as described below.

²⁷ Note that the magnitude of the impact of the attendance effects is not interpreted by those authors, surely due to these concerns (Carlin et al., 2021).

²⁸ In the prior working paper written by Carlin et al., it is observed that one extra game in the NHL/NBA leads to an increase in the number of cases by 895, which is a figure much more similar to the one estimated in this study. However, once again, those numbers are not directly comparable since in this case Carlin et al. considered all the games from January to March 2020.

²⁹ The extension of the impact has been made only for comparison with prior literature.

³⁰ An analysis of the R^2 resulting from regressions with/without province-specific time trends shows that by additionally controlling by province-specific time trends, we could account for about 14.8% points more of the variance of the COVID-19 cases per 100,000 inhabitants. The R^2 of that regression without the province-specific linear trends yielded an R^2 of 0.552, and after adding the province-specific trends, an increase in the R^2 to 0.707 was obtained.

³¹ Amuedo-Dorantes et al. (2021) showed clear differential Spanish regional trends in the pandemic's outcomes among regions when the nationwide lockdown was implemented. Province-specific linear trends accounted for these different trends in our analysis.

estimation clustering at the province level to account for any within-province correlation in the error terms.

In Table A2, we consider alternative time windows from incubation to diagnosis by considering the effect of the mass gathering events since March 13 and 18 (five and 10 days after the March 8 gatherings) in columns (1) and (2). The time windows from incubation to deaths were also extended considering the impact since March 18 and 25 in columns (3) and (4). Our findings did not change substantially. A positive and significant impact on reported cases but not on reported deaths can be seen. The non-significant estimates on deaths could be due to a failure in the measurement of deaths derived from COVID-19. This finding would be problematic in our analysis if related deaths were higher in areas in which authorities know there have been larger gatherings. Since some authors point to excess mortality as the most reliable measure of the harshness of the pandemic, we repeated the analysis considering excess mortality as a dependent variable (Armillei et al., 2021; Michelozzi et al., 2020). In Spain, excess mortality is estimated by the National Epidemiology Centre using a moving average model. Daily registered deaths are compared with the number of deaths in the same day (± 3 days) in the last 10 years. Unfortunately, these estimates are available only at the NUTS 2 regional level. Results presented in Table A3 provide additional evidence on the non-significant impact of the mass gathering events on deaths using the excess mortality. This finding is in line with Carlin et al. (2021), who also found a non-significant impact of attendance rates on deaths for certain sporting events. However, those estimates should be viewed with caution since information on excess mortality is not available at the province level.

Tables A4 and A5 report the point estimates after changing the main sample by extending/reducing a week in the sample, by reducing the sample until March 31, and by excluding Madrid (the hardest-hit Spanish province in the first wave in absolute numbers of cases and deaths). The results after extending the sample should be viewed with caution because of the change by Spanish authorities in the COVID-19 data collection method that disrupts the pandemic's outcome series by April 17 (Amuedo-Dorantes et al., 2021). In any case, it is reassuring that both extending and reducing the sample did not alter our conclusions concerning a positive and significant effect of the attendance rate on reported COVID-19 cases per 100,000 inhabitants. The exclusion of Madrid (Table A5) yielded similar conclusions. One standard deviation increase in the attendance rate implies a significant increase of 2.8 (3.1 with Madrid) in daily reported COVID-19 cases per 100,000 inhabitants (20% and 21%, with Madrid, of the post-effect standard deviation).

4.3. Identification

4.3.1. Event-study

Economic incentives for risk-avoiding behaviors by nonparticipants in mass gathering events may play a compensating role (Dave et al., 2020).³² In the Spanish situation, while admittedly some degree of risk avoiding-behavior occurred because of the decrease in the number of participants in the Women's Day marches in comparison to the previous year, our concern in this case was limited to whether that behavior was random across the regions. What matters in terms of inference purposes is the possible endogeneity with regard to the pandemic's outcomes of interest. To address this issue, we conducted an event study for each of the pandemic's outcomes (reported cases and deaths) that could reveal biases caused by reverse causality and/or voluntary precautions (Goodman-Bacon and Marcus, 2020). This process also allowed us to gauge whether or not any impact of the mass gathering events on the pandemic's outcomes predated the expected diagnosis (death) dates after exposure to COVID-19 in the mass gathering events. Using a methodology similar to the work that considered continuous treatment

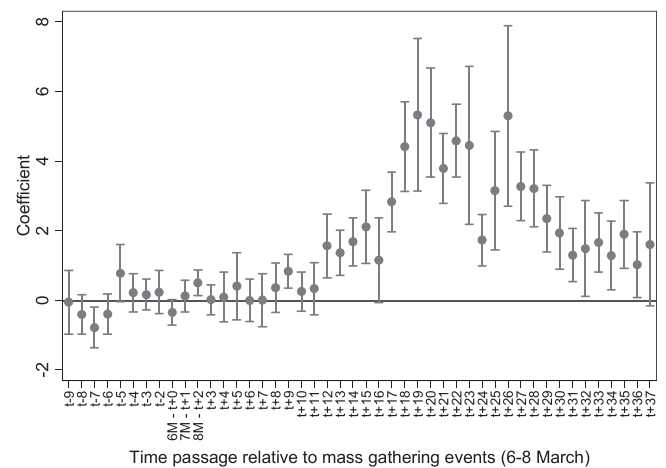


Fig. 3. Identification, event study of COVID-19 cases. Notes: Robust standard errors are clustered at the province level. Significant at the 5% level; see Table B1.

variables (Clemens et al., 2018; Goodman-Bacon, 2018), the event-study model is defined as shown below:

$$Y_{pt} = Attendees_p^{March\ 6-8} \left[\sum_{r=-9}^{-2} \tau_r 1\{t - t^{March-6} = r\} + \sum_{r=0}^{37} \rho_r 1\{t - t^{March-6} = r\} \right] + \mu_t + \theta_p + \varepsilon_{pt} \quad (2)$$

in which Y_{pt} represents the outcome (COVID-19 cases or deaths per 100,000 inhabitants) in province p at date t . The variation in the number of attendees is captured in $Attendees_p^{March\ 6-8}$, which gauges the attendance rate. Pre/post treatments are defined by dummy variables $1\{t - t^{March-6} = r\}$ that measure the time relative to the first day that the mass gathering event took place (March 6). The reference period in all event studies was taken as the period before the event occurred, when $r = -1$, which normalizes to zero the estimates of τ_r and ρ_r on that specific event day. The coefficients of interest, τ_r and ρ_r , gauge the relationship (covariate-adjusted) between the pandemic's outcomes and the attendance rate in the nine days leading up to the mass gathering events under consideration here and the days that followed. The τ_r are falsification tests that detect the relationship between the attendance rate and the pandemic's outcomes before the mass gathering events existed. The coefficients ρ_r capture the effects of the attendance rate on the pandemic's outcomes. The estimates will equal zero if attendance rates have affected the pandemic's outcomes equally across provinces. If we are really gauging the effect of the mass gathering events, non-immediate significant daily effects of the mass gathering events on the pandemic's outcomes should have been observed, since a time gap between exposure or infection and diagnosis/death was noted. This specification included a set of province fixed effects, province-specific linear time trends, month fixed effects, and day of the week fixed effects (μ_t, θ_p) that control for unincorporated factors that could potentially affect the pandemic's outcomes. Since the incorporation of date-fixed effects is not possible in this specification, the post-event coefficients may be biased because they incorporated differences in the temporal evolution of the pandemic. The longer the post-event time period was, the more possible composition changes in the groups may have affected our estimates.

Event-study estimates are presented in Figs. 3–4. No evidence of the existence of significant differential pre-trends was found, strongly supporting the assumption of no pre-trends in both reported cases and deaths. This finding is in line with the lack of regional differences observed in the mobility patterns before the nationwide lockdown (Amuedo-Dorantes et al., 2021). Post-event coefficients cannot be clearly different from zero until March 18 (12 days after March 6) for reported cases (10 days after the mass gathering events, which coincided

³² Unexpectedly, the 2020 Black Lives Matter did not reignite the US community-level COVID-19 growth rate (Dave et al., 2020).

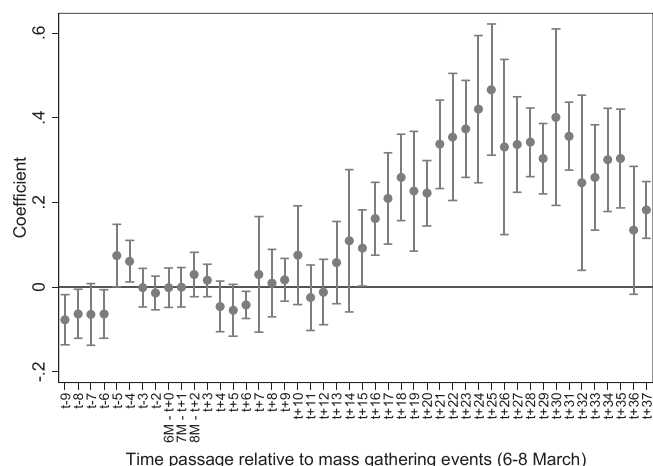


Fig. 4. Identification, event study on COVID-19 deaths. Notes: Robust standard errors are clustered at the province level. Significant at the 5% level; see Table B1.

Table 2
Identification: Placebo regressions.

D.V.:	(1)		(2)		(3)		(4)	
	Reported Cases		Reported Deaths		Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 15	March 22	March 22	March 29	March 22	March 29
% Attendees	-0.017 (0.468)	-0.007 (0.497)	-0.001 (0.067)	0.001 (0.051)				
Observations	2350	2350	2350	2350				
<i>For all</i>								
Province FE	Yes	Yes	Yes	Yes				
Date FE	Yes	Yes	Yes	Yes				
Province-specific linear trends	Yes	Yes	Yes	Yes				

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12 2020 (mean: 0.76; S.D.: 1.33). Bootstrap coefficients and standard errors are derived from 2000 placebo regressions of Eq. (1). In each replication, robust standard errors are clustered by province. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

with the average four to five days of incubation to symptom onset [CDC, 2021] plus six days to diagnosis (Redondo-Bravo et al., 2020) and until March 21 for reported deaths. The unexpected lockdown implemented nationwide in Spain also provided a quasi-random experiment in this setting to gauge the effectiveness of this kind of stringent social distance measure in mitigating the effect of mass gathering events on the spread of the virus. As can be seen, by March 24–25, the estimated coefficients began to decrease in magnitude (10 days after the nationwide lockdown), signaling a decrease in the spread of COVID-19 generated by the mass gathering events.³³ These results reinforce our findings that participation in the mass gathering events can be strongly associated with increases in the pandemic’s outcomes, especially in terms of reported cases.³⁴

³³ Our findings were similar after running the event study with the excess mortality in Fig. B2.

³⁴ We re-ran the event study following the methodology proposed by de Chaisemartin and D’Haultfoeuille (2020), although this method did not allow us to multiply the pre-event variables by the % attendees (dosage). Results were qualitatively similar as pre-trends were not detected. We observed a statistically significant impact of the attendance rates on reported cases from March 18–26 and evidence of some statistically significant impact on reported deaths, but for a short period of time (March 21 and 22); see Fig. B3.

Table 3

Mechanisms #1: Sports events vs. Women’s Day marches, COVID-19 cases and deaths response to the variation in the attendees to mass gathering events (March 6–8).

	(1)	(2)	(3)	(4)
PANEL A: % Attendees by Mass Gathering Event type (Sports vs. Women’s Day Marches)				
D.V.:	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 22	March 29
% Attendees to sport events	1.736* (0.927)	0.934 (1.033)	-0.033 (0.096)	-0.002 (0.068)
% Attendees to Women’s Day Marches	1.767** (0.797)	3.024*** (0.698)	0.119 (0.118)	0.019 (0.047)
Observations	2350	2350	2350	2350
R-squared	0.707	0.708	0.685	0.684
PANEL B: Including other sport events from February 20				
% Attendees to sport events (March 6–8)	1.401 (1.190)	1.147 (1.454)	0.078 (0.096)	0.054 (0.056)
% Attendees to Women’s Day Marches	1.579* (0.789)	3.144*** (0.791)	0.181 (0.128)	0.050 (0.039)
% Attendees to other sport events	0.345 (0.486)	-0.220 (0.634)	-0.114** (0.048)	-0.058* (0.033)
Observations	2350	2350	2350	2350
R-squared	0.708	0.708	0.687	0.685
PANEL C: Dummy for having had a match in the province (March 6–8)				
Having a match	2.882 (1.830)	-1.080 (1.746)	-0.394 (0.246)	-0.165 (0.188)
% Attendees to Women’s Day Marches	2.084* (1.140)	3.250*** (0.844)	0.120 (0.100)	0.022 (0.052)
Observations	2350	2350	2350	2350
R-squared	0.706	0.708	0.686	0.685
<i>In the Post Period</i>				
D.V. Mean	12.75	14.73	1.51	1.64
D.V. Standard Deviation	14.63	15.78	1.61	1.61
<i>For all</i>				
Province FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Province-specific linear trends	Yes	Yes	Yes	Yes

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. Panel A incorporates as the main explanatory variables the % attendees to mass gathering events over total population at province level by type of event (major sports events including First and Second Division football (soccer) matches and ACB basketball matches (mean: 0.84; S.D.: 1.31) and Women’s Day marches (mean: 1.44; S.D.: 1.00)). We estimate Eq. (1) separating the main explanatory variable by mass gathering event type. Panel B adds as the main explanatory variables the % attendees to mass gathering sports events over total population at province level from February 20 (excluding those incorporated into the events that occurred during March 6–8). Panel C introduces a dummy case for having had a match in the province (March 6–8). We estimate Eq. (1) separating the main explanatory variable by mass gathering event type. Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

4.3.2. Placebo regressions

Since our analysis exploited geographical differences in attendance rates at the mass gathering events, we can test whether the estimated coefficients presented in Table 1 are significantly different from what may be generated with a random variation of the dosage. Thus, from

Table 4

Mechanisms #2: Impact of % attendees depending on socioeconomic variables, COVID-19 cases and deaths response to the variation in the attendees to mass gathering events (March 6–8).

	(1)	(2)	(3)	(4)
PANEL A: Above/Below the mean GDP per capita				
D.V.:	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 22	March 29
% Attendees (1)	0.397 (0.783)	0.680 (0.898)	-0.122 (0.139)	-0.039 (0.067)
% Attendees x Province above the mean (2)	1.261 (0.769)	0.989 (0.788)	0.138 (0.119)	0.042 (0.052)
R-squared	0.708	0.708	0.685	0.684
p-value (1)+(2)= 0	0.0048	0.0059	0.7906	0.9570
PANEL B: Above/Below the mean population density				
% Attendees (1)	1.655 ** (0.674)	1.924 *** (0.663)	0.048 (0.064)	0.031 (0.059)
% Attendees x Province above the mean (2)	0.398 (0.948)	-0.792 (0.655)	-0.095 (0.092)	-0.106 (0.079)
R-squared	0.708	0.708	0.685	0.685
p-value (1)+(2)= 0	0.0355	0.0874	0.6249	0.4322
PANEL C: Above/Below the mean share of employment in the service sector				
% Attendees (1)	1.811 *** (0.576)	2.126 *** (0.604)	0.028 (0.070)	0.025 (0.061)
% Attendees x Province above the mean (2)	-0.188 (0.824)	-1.147 * (0.668)	-0.009 (0.079)	-0.056 (0.066)
R-squared	0.708	0.708	0.684	0.685
p-value (1)+(2)= 0	0.0623	0.1458	0.7755	0.6840
PANEL D: Above/Below the mean share of population with college over working age population (aged 16–64)				
% Attendees (1)	2.453 ** (1.017)	2.835 * (1.431)	0.016 (0.181)	0.053 (0.058)
% Attendees x Province above the mean (2)	-0.715 (0.895)	-1.110 (1.152)	0.010 (0.149)	-0.048 (0.051)
R-squared	0.708	0.708	0.684	0.684
p-value (1)+(2)= 0	0.0031	0.0017	0.6660	0.9193
PANEL E: Above/Below the mean of physicians per 1000 individuals				
% Attendees (1)	0.411 (0.620)	0.246 (0.595)	-0.327 *** (0.101)	0.013 (0.078)
% Attendees x Province above the mean (2)	1.254 ** (0.610)	1.400 ** (0.555)	0.331 *** (0.088)	-0.007 (0.065)
R-squared	0.708	0.708	0.687	0.684
p-value (1)+(2)= 0	0.0046	0.0061	0.9518	0.9040
PANEL F: Above/Below the mean share of population + 65				
% Attendees (1)	2.299 *** (0.488)	1.393 ** (0.660)	-0.008 (0.060)	-0.077 (0.047)
% Attendees x Province above the mean (2)	-0.965 (0.691)	0.607 (0.763)	0.059 (0.077)	0.145 *** (0.043)
R-squared	0.708	0.707	0.685	0.686
p-value (1)+(2)= 0	0.0563	0.0188	0.4904	0.1458
PANEL G: Above/Below the mean share of foreign population				
% Attendees (1)	1.442 * (0.722)	1.857 ** (0.861)	0.008 (0.081)	0.068 (0.045)
% Attendees x Province above the mean (2)	0.631 (0.739)	-0.242 (0.737)	0.037 (0.082)	-0.128 *** (0.048)
R-squared	0.708	0.707	0.685	0.686
p-value (1)+(2)= 0	0.0004	0.0036	0.5167	0.2392
<i>In the Post Period</i>				
D.V. Mean	12.75	14.73	1.51	1.64
D.V. Standard Deviation	14.63	15.78	1.61	1.61
<i>For all</i>				
Observations	2350	2350	2350	2350
Province FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Province-specific linear trends	Yes	Yes	Yes	Yes

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. The main explanatory variable is the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women's Day marches (mean: 2.29; S.D.: 1.79). We estimate Eq. (1). Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table 5

Mechanisms #3: Impact of % attendees depending on weather and location variables, COVID-19 cases and deaths response to the variation in the attendees to mass gathering events (March 6–8).

	(1)	(2)	(3)	(4)
PANEL A: Above/Below the mean annual temperature				
D.V.:				
	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 22	March 29
% Attendees (1)	1.792 *** (0.603)	1.865 *** (0.573)	0.040 (0.063)	0.006 (0.053)
% Attendees x Province above the mean (2)	-0.528 (0.695)	-1.499 ** (0.583)	-0.176 * (0.098)	-0.005 (0.057)
R-squared	0.708	0.708	0.686	0.684
p-value (1)+(2)= 0	0.0582	0.5476	0.2358	0.9867
PANEL B: Above/Below the mean annual precipitation				
% Attendees (1)	1.399 (1.154)	2.014 (1.317)	0.057 (0.121)	0.069 (0.079)
% Attendees x Province above the mean (2)	0.436 (1.001)	-0.343 (1.033)	-0.039 (0.101)	-0.079 (0.069)
R-squared	0.708	0.707	0.685	0.685
p-value (1)+ (2)= 0	0.0013	0.0015	0.7123	0.8266
PANEL C: Above/Below the mean annual relative air humidity				
% Attendees (1)	-0.407 (0.547)	-0.045 (0.645)	-0.159 (0.114)	0.039 (0.094)
% Attendees x Province above the mean (2)	2.307 *** (0.608)	1.911 *** (0.625)	0.197 * (0.100)	-0.036 (0.074)
R-squared	0.711	0.709	0.687	0.684
p-value (1)+ (2)= 0	0.0002	0.0004	0.5337	0.9454
PANEL D: Above/Below the mean annual sunny hours				
% Attendees (1)	1.906 *** (0.472)	1.910 *** (0.483)	0.047 (0.066)	0.001 (0.049)
% Attendees x Province above the mean (2)	-1.797 ** (0.672)	-1.935 *** (0.555)	-0.239 ** (0.097)	0.059 (0.070)
R-squared	0.710	0.709	0.688	0.685
p-value (1)+ (2)= 0	0.8627	0.9615	0.0865	0.5054
PANEL G: Dummy for provinces located in the Spanish coast				
% Attendees (1)	1.844 *** (0.638)	1.962 *** (0.650)	0.075 (0.075)	0.013 (0.056)
% Attendees x Province above the mean (2)	-0.485 (0.612)	-1.113 (0.725)	-0.246 ** (0.097)	-0.036 (0.055)
R-squared	0.708	0.708	0.688	0.684
p-value (1)+ (2)= 0	0.0025	0.2739	0.1379	0.7400
<i>In the Post Period</i>				
D.V. Mean	12.75	14.73	1.51	1.64
D.V. Standard Deviation	14.63	15.78	1.61	1.61
<i>For all</i>				
Observations	2350	2350	2350	2350
Province FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Province-specific linear trends	Yes	Yes	Yes	Yes

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. The main explanatory variable is the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women’s Day marches (mean: 2.29; S.D.: 1.79). We estimate Eq. (1). Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

actual data, we used randomly distributed attendance rates across provinces (a random draw without replacement) and then re-estimated Eq. (1) for both pandemic-related outcomes with these random attendance rates. We repeated this exercise 2000 times, which yielded a large number of different estimates of the coefficients.³⁵ From these 2000 replications, we were able to obtain the bootstrap average and standard deviation of all coefficients. The results are reported in Table 2 and should be compared with those obtained in Table 1 using the actual attendance rates by province. Random attendance rates led to no significant effect (and quite close to zero) on pandemic outcomes and, more importantly, standard errors are far from those obtained in Table 1, which indicates that the point estimates in Table 1 are significantly far from the distribution of the placebo estimates. Thus, the estimates shown in Table 1 appear to be gauging the effect of the mass gathering events on pandemic outcomes.

4.3.3. Adding province variables

The existence of unmeasured factors that could be related to both COVID-19 outcomes and attendance rates could threaten the validity of our identification strategy. The evolution of COVID-19 outcomes has been suggested to be associated with several variables varying across regions/countries (Armillei et al., 2021). To mitigate the possible concerns about confounding from omitted variables, we demonstrated that the attendance rate was not related to various observables characteristics at the provincial level that have been suggested to have an impact on COVID-19 outcomes (Armillei et al., 2021; Basellini and Camarda, 2021); see Table D1. We included the GDP per capita, population density, employment in the service sector, population with college degree, physicians' rate over population, rate of the elderly population, foreign population rate, four weather variables (average annual temperature, average annual precipitation, average annual relative air humidity, and average annual sunny hours), and a location variable capturing whether the province is located on the coast as a measure of such confounders.³⁶ Only a marginally significant association (at the 10% level) was found with the GDP per capita, the variable measuring the employment in the service sector, and the share of elder population. This finding suggests that omitted variables issues were not an important concern in our analysis. In any case, we incorporated all of them as controls in Table D2. Controlling for all those province-level characteristics did not alter our main conclusions. With respect to the additional controls, our findings are in line with those of Basellini and Camarda (2021), whose findings underscore the importance of socio-economic and demographic factors across regional areas of the study of a pandemic. This comparison provided additional evidence on the validity of our analysis.

5. Mechanisms at play: Mass gathering event type, pre-March 6–8 events, and other observable characteristics at the province level

Thus far, our findings support a robust and positive impact of attendance rates at mass gathering events on COVID-19 cases. To assess whether the observed impacts were amplified/reduced depending on the mass gathering event type and by other factors, we differentiated three possible mechanisms at work: (1) Women's Day marches versus sporting events, (2) socio-economic variables, and (3) weather and geographical locations.

Being aware of the finding that some of the major events considered in these cases occurred in the same cities with only one day of difference, the analysis of their possible separate impact on the pandemic's outcomes should be viewed with caution. Panel A in Table 3 extends the

model presented in Eq. (1), separating the main explanatory variable by mass gathering type (Women's Day marches versus sporting events). It can clearly be observed that Women's Day marches acted as super-spreader events. The evidence on the impact of sports events is not so clear. Focusing on column (1), we observe a marginally significant positive impact of the sports events but not statistically different from that obtained in the Women's Day marches (p-value 0.9812). However, the significant impact of the sports events on reported cases is not maintained when the time that elapses from incubation to diagnosis ranges to March 22. With respect to deaths, we observe only positive coefficients (although non-significant) in the case of the attendance rates at marches; sports events do not appear to have significantly increased deaths (and even coefficients were close to zero) as indicated by the reported deaths in columns (3) and (4). Thus, it can be argued that differences exist in the spread of COVID-19, depending on the mass gathering event type.

We also explored whether pre-March 6–8 events had any role, following Carlin et al. (2021). To assess this issue, we have incorporated sports events from February 20 in Panel B of Table 3. Women's Day marches still matter, and the pre-events do not significantly increase the reported cases. We then studied whether attendance rates rather than the number of games or having a match were the routes through which the transmission occurred. Distinguishing among the number of games/having a match and the attendance rate may not matter in a setting with, for example, little variation in the number of available seats (or attendance) across sports events, but in our case, this difference was an important issue. For example, in Vitoria, around 5% of the population participated in two sporting events, whereas in Madrid, 1.6% of the province's population participated in six events. Because of the small variation in the number of games per province, with the exception of Madrid (but our results were robust with respect to its exclusion), we opted for inclusion of a dummy variable that captured whether or not a match was held in a province. The estimated coefficient on that dummy was negative even in some of the specifications of Panel C. Thus, what appears to matter is the attendance rate rather than just having a game.

From that point onward, we expanded the possible mechanisms at play, including data on several socio-economic, weather, and geographical location variables at the province level, which are included as controls in Table D2. Our estimations are reported in Tables 4 and 5. To lessen the possibility of those possible channels being interpreted as amplifying/reducing the effect of the attendance rate, we concentrated on provinces above/below the mean value of each of these extra channels. Instead of describing events one by one, we focused our attention on the two mechanisms at play that appeared to have mattered. Those estimates should satisfy two conditions: (1) significant differences of the attendance rate impact among provinces above/below the mean of each extra covariate should exist, and (2) the estimates should be statistically significant. Taking these conditions into consideration, we found that the rate of physicians per 1000 inhabitants significantly amplified the impact of the attendance rate in those provinces above the mean rate of physicians per 1000 inhabitants, but this was not detected in the case of those provinces below the mean. This finding may simply have reflected a better possibility of a person being tested because of residing in a province with high expenditures in the healthcare system.³⁷ However, it could also indicate that the higher the number of physicians, the higher the likelihood of spreading the virus in those provinces with high attendance rates, given that 20% of the reported cases during the first wave were physicians (Redondo-Bravo et al., 2020). The other channel amplifying the spread of the virus associated with attendance at mass gathering events was the relative air humidity. The impact of the

³⁵ In Fig. C1, we report the empirical distribution of the 2000 estimated coefficients for each of the pandemic's outcomes in the placebo regressions.

³⁶ See a detailed description of the definition and sources of these variables in Table D1.

³⁷ The number of physicians included all the physicians working in both the public and private sectors.

attendance rate was amplified in those provinces having high relative air humidity.³⁸ The evidence on the importance of humidity for the spread of COVID-19 was mixed, but some works pointing to a possible positive correlation have been published (Ganegoda et al., 2021).³⁹

6. Conclusions

In this paper, we estimated the impact of mass gathering events on the COVID-19 pandemic's outcomes (cases and deaths). Specifically, by exploiting the geographical variation in the pre-pandemic planned mass gathering events and the variation in their attendance, we demonstrated that 31% of the average daily reported COVID-19 cases per 100,000 inhabitants between mid-March and early April 2020 could be explained by participation in the major events that took place from March 6 to March 8, 2020. A back-of-the-envelope calculation suggested that this finding implies almost five million euros (169,000 euros/daily) of additional economic costs in the healthcare system of a typical province with one million inhabitants in the period under consideration or 230 million euros for the entire country. Identification checks support a plausibly exogenous variation of our explanatory variables, and robustness checks confirm the reliability of our estimates.

Although many of the social distancing measures, such as lockdowns, business closures, and school closures, varied over time depending on the intensity of COVID-19 in each country/region, the mass gathering event bans were maintained over longer periods of time.⁴⁰ A key concern of policymakers regarding mass gatherings during the COVID-19 pandemic was the increased risk of transmission of contagious infections as a result of large numbers of people in close contact for extended periods of time. Consistent with the observational epidemiological evidence, we presented evidence to support this concern as well as the economic costs in an early stage of the pandemic. Our findings evaluating the risks associated with a mass gathering event, especially in the early stages of virus transmission, highlight that public health authorities must cancel or postpone mass gathering events during future pandemics.

CRedit authorship contribution statement

Conceptualization: 50% Rafael González-Val and 50% Miriam Marcén. Data curation: 50% Rafael González-Val and 50% Miriam Marcén. Formal analysis: 50% Rafael González-Val and 50% Miriam Marcén. Funding acquisition: 50% Rafael González-Val and 50% Miriam Marcén. Investigation: 50% Rafael González-Val and 50% Miriam Marcén. Methodology: 50% Rafael González-Val and 50% Miriam Marcén. Project administration: 50% Rafael González-Val and 50%

³⁸ This finding should be viewed with caution since we used the average annual data at the province level for the weather conditions.

³⁹ Some economic evidence exists that the expansion of transportation networks caused an increase in the spread of viral diseases (Adda, 2016). We examined whether or not this process was really a route for the spread of the virus associated with the mass gathering events. We assessed the role of public transportation in the amplification of the effect of the mass gathering events on the spread of COVID-19 by exploiting the differences in the number of underground/metro travelers by province (Source: INE (Spanish Statistical Office)). Because no available data could be found on other public transportation at the province level (bus passengers are only available at the NUTS 2 region level), we preferred to include this analysis in the Appendix. As can be seen in Table A6, this kind of public transportation does not appear to play a role in amplifying the impact of the mass gathering events. These estimates should be interpreted with caution because we are using average annual data on travelers from 2019, which may not coincide with the specific dates of March 6–8, 2020. Moreover, note that only seven cities in Spain have a metro/underground network (Barcelona, Bilbao, Madrid, Málaga, Palma, Sevilla, and Valencia).

⁴⁰ See a review of social distancing measures in COVID-19 Health Systems Response Monitor in <https://www.covid19.healthsystem.org/mainpage.aspx>.

Miriam Marcén, Resources: 50% Rafael González-Val and 50% Miriam Marcén, Software: 50% Rafael González-Val and 50% Miriam Marcén, Supervision: 50% Rafael González-Val and 50% Miriam Marcén, Validation: 50% Rafael González-Val and 50% Miriam Marcén, Visualization: 50% Rafael González-Val and 50% Miriam Marcén, Roles/Writing - original draft: 50% Rafael González-Val and 50% Miriam Marcén, Writing - review & editing: 50% Rafael González-Val and 50% Miriam Marcén.

Funding

The authors acknowledge support from the Spanish Ministerio de Ciencia e Innovación and Agencia Estatal de Investigación, MCIN/AEI/10.13039/501100011033 (grant number: PID2020-114354RA-I00), Gobierno de Aragón (AETRE research group) and European Regional Development Fund (ERDF).

Acknowledgements

Comments received from two anonymous referees have improved the version originally submitted. All remaining errors are ours.

Appendix

See Tables A1-A6.

Table A1

Robustness checks #1: Replication of Table 1 excluding province-specific linear trends and without clustering.

	(1)	(2)	(3)	(4)
PANEL A: Excluding Province-specific linear trends				
D.V.:	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 22	March 29
% Attendees	1.477 ** (0.650)	1.409 ** (0.682)	0.098 (0.069)	0.094 (0.063)
Observations	2350	2350	2350	2350
R-squared	0.552	0.551	0.536	0.535
PANEL B: Without clustering				
% Attendees	1.477 *** (0.229)	1.409 *** (0.272)	0.098 *** (0.025)	0.094 *** (0.031)
Observations	2350	2350	2350	2350
R-squared	0.552	0.551	0.536	0.535
<i>In the Post Period</i>				
D.V. Mean	12.75	14.73	1.51	1.64
D.V. Standard Deviation	14.63	15.78	1.61	1.61
<i>For all</i>				
Province FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Province-specific linear trends	No	No	No	No

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. Both panels incorporate as the main explanatory variable the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women's Day marches (mean: 2.29; S.D.: 1.79). We estimate Eq. (1). Robust standard errors are clustered at the province level and reported in parentheses in Panel A. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A2

Robustness checks #2: Replication of Table 1 changing the post-period effect.

D.V.:	(1)	(2)	(3)	(4)
	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 13	March 18	March 18	March 25
% Attendees	1.353 *** (0.500)	2.328 *** (0.772)	0.073 (0.056)	-0.020 (0.058)
Observations	2350	2350	2350	2350
R-squared	0.706	0.710	0.685	0.684
<i>In the Post Period</i>				
D.V. Mean	12.14	13.78	1.34	1.60
D.V. Standard Deviation	14.37	15.03	1.55	1.61
<i>For all</i>				
Region FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Region-specific linear trends	Yes	Yes	No	No

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). The post-period summary statistics for the D.V. are measured since March 13, 18, and 25 in each corresponding column. The main explanatory variable is the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women’s Day marches (mean: 2.29; S.D.: 1.79). We estimate Eq. (1). Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A3

Robustness checks #3: Replication of Panel A in Table 1 (columns (3) and (4)) changing the D.V. with/without region-specific linear trends and clustering.

D.V.:	(1)	(2)	(3)	(4)
	Excess Mortality (Region level)		Excess Mortality (Region level)	
Effect Since (Post Period):	March 22	March 29	March 22	March 29
% Attendees	0.155 (0.426)	0.069 (0.170)	0.155 (0.116)	0.069 (0.115)
Observations	799	799	799	799
R-squared	0.821	0.821	0.821	0.821
<i>In the Post Period</i>				
D.V. Mean	2.56	2.59	2.56	2.59
D.V. Standard Deviation	2.58	2.48	2.58	2.48
<i>For all</i>				
Region FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Region-specific linear trends	Yes	Yes	No	No

Notes: The dependent variable (D.V.) in columns (1) to (4) is defined as excess mortality per 100,000 inhabitants at the NUTS 2 region level (not available at the province level) from February 26 to April 12, 2020 (mean: 1.30; S.D.: 2.21; Source MoMo). The post-period summary statistics for the D.V. are measured since March 22, and 29 in each corresponding column. The main explanatory variable is the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women’s Day marches (mean: 2.37; S.D.: 1.28). We estimate Eq. (1). Robust standard errors are clustered at the region level and reported in parentheses in columns (1) and (2). *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A4

Robustness checks #4: Replication of Panel A in Table 1 extending/reducing a week in the sample and cutting the sample at March 31.

D.V.:	(1)	(2)	(3)	(4)
	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 22	March 29
PANEL A: Sample from February 26–19 April 2020				
% Attendees	2.001 *** (0.695)	2.045 *** (0.719)	0.059 (0.077)	0.011 (0.052)
Observations	2700	2700	2700	2700
R-squared	0.674	0.673	0.630	0.630
<i>In the Post Period</i>				
D.V. Mean	12.34	13.74	1.44	1.50
D.V. Standard Deviation	14.31	15.22	1.58	1.57
PANEL B: Sample from February 26–5 April 2020				
% Attendees	1.181 *** (0.271)	1.436 *** (0.456)	0.000 (0.055)	0.020 (0.056)
Observations	2000	2000	2000	2000
R-squared	0.737	0.738	0.721	0.721
<i>In the Post Period</i>				
D.V. Mean	12.94	15.93	1.50	1.73
D.V. Standard Deviation	14.44	15.91	1.63	1.64
PANEL C: Sample from February 26–31 March 2020				
% Attendees	0.482 * (0.273)	1.348 *** (0.427)	0.005 (0.061)	0.079 (0.094)
Observations	1750	1750	1750	1750
R-squared	0.733	0.736	0.703	0.704
<i>In the Post Period</i>				
D.V. Mean	11.59	16.01	1.27	1.77
D.V. Standard Deviation	13.50	15.94	1.57	1.81
<i>For all</i>				
Province FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Province-specific linear trends	Yes	Yes	Yes	Yes

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level. The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level. The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. All panels incorporate as the main explanatory variable the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women’s Day marches (mean: 2.29; S.D.: 1.79). We estimate Eq. (1). Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A5
Robustness checks #5: Replication of Panel A in Table 1 excluding Madrid.

D.V.:	(1)	(2)	(3)	(4)
	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 22	March 29
% Attendees	1.582 ** (0.608)	1.763 *** (0.609)	0.011 (0.057)	0.022 (0.053)
Observations	2303	2303	2303	2303
R-squared	0.710	0.710	0.678	0.678
<i>In the Post Period</i>				
D.V. Mean	12.51	14.53	1.46	1.60
D.V. Standard Deviation	14.58	15.78	1.59	1.61
<i>For all</i>				
Province FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Province-specific linear trends	Yes	Yes	Yes	Yes

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 7.97; S.D.: 12.90). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.73; S.D.: 1.31). The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. The main explanatory variable is the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women's Day marches (mean: 2.26; S.D.: 1.80). We estimate Eq. (1). Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A6
Additional Mechanisms #1: Transmission inside public transportation (underground/metro) considering metro passengers in 2019 in each capital city.

D.V.:	(1)	(2)	(3)	(4)
	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 22	March 29
% Attendees without metro	1.409 * (0.776)	1.768 ** (0.717)	-0.007 (0.068)	0.040 (0.065)
% Attendees *metro passengers	0.184 (0.219)	-0.016 (0.078)	0.017 (0.017)	-0.018 (0.017)
Observations	2350	2350	2350	2350
R-squared	0.708	0.707	0.685	0.685
<i>In the Post Period</i>				
D.V. Mean	12.75	14.73	1.51	1.64
D.V. Standard Deviation	14.63	15.78	1.61	1.61
<i>For all</i>				
Province FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Province-specific linear trends	Yes	Yes	Yes	Yes

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. The main explanatory variable is the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women's Day marches (mean: 2.29; S.D.: 1.79). We estimate Eq. (1) but separate the main explanatory variable among provinces with/without metro/underground. Only seven cities in Spain have a metro/underground network. Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

See Table B1 Figs. B2, B3 and C1.

Table B1

Identification #1: Event study on COVID-19 cases and deaths response to the variation in the attendees to mass gathering events (March 6–8).

	(1) Reported Cases	(2) Reported Deaths
9 days before the event x % Attendees	-0.057 (0.456)	-0.077 ** (0.029)
8 days before the event x % Attendees	-0.405 (0.281)	-0.063 ** (0.029)
7 days before the event x % Attendees	-0.784 *** (0.291)	-0.064 * (0.036)
6 days before the event x % Attendees	-0.392 (0.289)	-0.063 ** (0.029)
5 days before the event x % Attendees	0.781 * (0.410)	0.074 ** (0.037)
4 days before the event x % Attendees	0.215 (0.276)	0.061 ** (0.024)
3 days before the event x % Attendees	0.165 (0.222)	-0.001 (0.023)
2 days before the event x % Attendees	0.236 (0.309)	-0.014 (0.020)
The day of the event x % Attendees	-0.347 * (0.182)	-0.001 (0.023)
1 day after the event x % Attendees	0.121 (0.225)	-0.000 (0.023)
2 days after the event x % Attendees	0.500 *** (0.183)	0.030 (0.026)
3 days after the event x % Attendees	0.018 (0.215)	0.016 (0.019)
4 days after the event x % Attendees	0.094 (0.356)	-0.045 (0.030)
5 days after the event x % Attendees	0.409 (0.480)	-0.055 * (0.030)
6 days after the event x % Attendees	-0.001 (0.302)	-0.042 ** (0.016)
7 days after the event x % Attendees	0.005 (0.379)	0.030 (0.068)
8 days after the event x % Attendees	0.360 (0.354)	0.009 (0.040)
9 days after the event x % Attendees	0.836 *** (0.239)	0.017 (0.025)
10 days after the event x % Attendees	0.251 (0.279)	0.076 (0.058)
11 days after the event x % Attendees	0.334 (0.374)	-0.025 (0.038)
12 days after the event x % Attendees	1.564 *** (0.454)	-0.012 (0.038)
13 days after the event x % Attendees	1.368 *** (0.325)	0.058 (0.048)
14 days after the event x % Attendees	1.682 *** (0.343)	0.109 (0.084)
15 days after the event x % Attendees	2.110 *** (0.523)	0.092 ** (0.045)
16 days after the event x % Attendees	1.157 * (0.608)	0.162 *** (0.043)
17 days after the event x % Attendees	2.832 *** (0.427)	0.209 *** (0.054)
18 days after the event x % Attendees	4.419 *** (0.644)	0.259 *** (0.051)
19 days after the event x % Attendees	5.337 *** (1.095)	0.227 *** (0.070)
20 days after the event x % Attendees	5.111 *** (0.782)	0.222 *** (0.038)
21 days after the event x % Attendees	3.793 *** (0.498)	0.337 *** (0.052)
22 days after the event x % Attendees	4.590 *** (0.521)	0.355 *** (0.075)
23 days after the event x % Attendees	4.453 *** (1.131)	0.374 *** (0.057)
24 days after the event x % Attendees	1.730 *** (0.367)	0.420 *** (0.087)
25 days after the event x % Attendees	3.155 *** (0.846)	0.466 *** (0.077)
26 days after the event x % Attendees	5.304 *** (1.292)	0.331 *** (0.103)

(continued on next page)

Table B1 (continued)

	(1) Reported Cases	(2) Reported Deaths
27 days after the event x % Attendees	3.276 *** (0.491)	0.337 *** (0.056)
28 days after the event x % Attendees	3.218 *** (0.551)	0.343 *** (0.040)
29 days after the event x % Attendees	2.351 *** (0.479)	0.304 *** (0.041)
30 days after the event x % Attendees	1.939 *** (0.519)	0.401 *** (0.104)
31 days after the event x % Attendees	1.302 *** (0.377)	0.357 *** (0.040)
32 days after the event x % Attendees	1.491 ** (0.684)	0.247 ** (0.103)
33 days after the event x % Attendees	1.667 *** (0.424)	0.259 *** (0.062)
34 days after the event x % Attendees	1.291 ** (0.491)	0.301 *** (0.061)
35 days after the event x % Attendees	1.898 *** (0.486)	0.304 *** (0.058)
36 days after the event x % Attendees	1.029 ** (0.471)	0.135 * (0.075)
37 days after the event x % Attendees	1.607 * (0.879)	0.182 *** (0.034)
Observations	2350	2350
R-squared	0.713	0.670
Province FE	YES	YES
Day of the week FE	YES	YES
Month FE	YES	YES
Province-specific linear trends	YES	YES

Notes: The dependent variable (D.V.) in column (1) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in column (2) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). We estimate Eq. (2) with the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women’s Day marches (mean: 2.29; S.D.: 1.79). Robust standard errors are clustered at the province level and reported in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

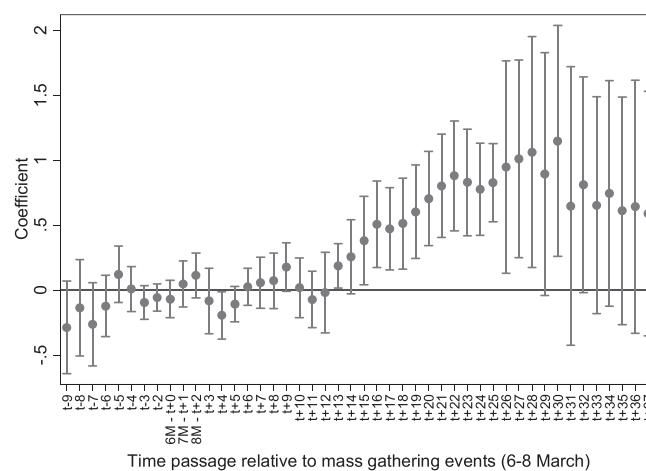


Fig. B2. Identification #2: Event study on excess mortality, Notes: Robust standard errors are clustered at the NUTS 2 region level. Significant at the 5% level.

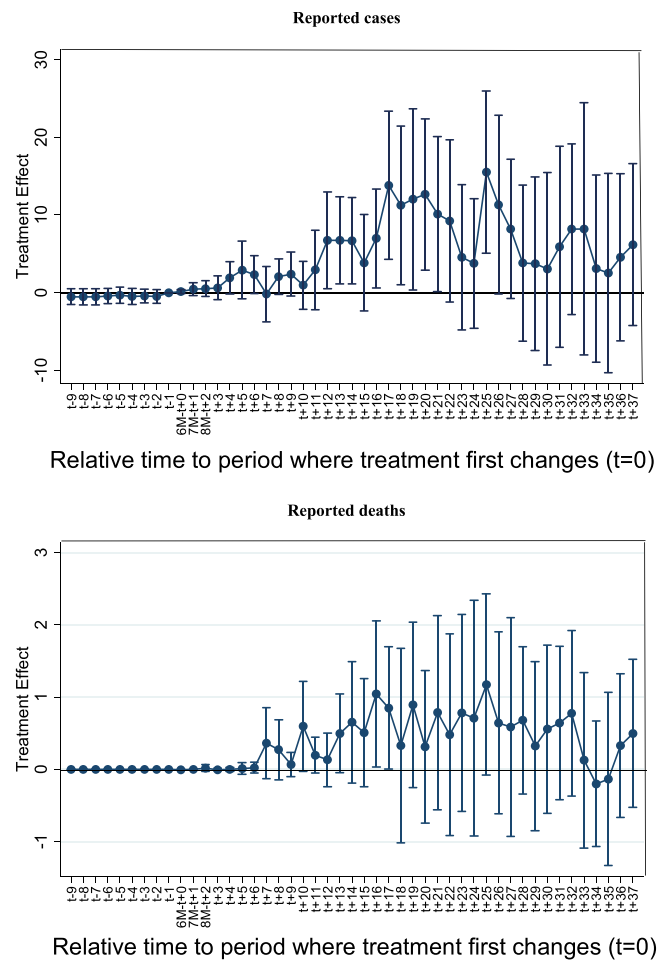


Fig. B3. Identification #3: DID, from last period before treatment changes ($t = -1$) to t .

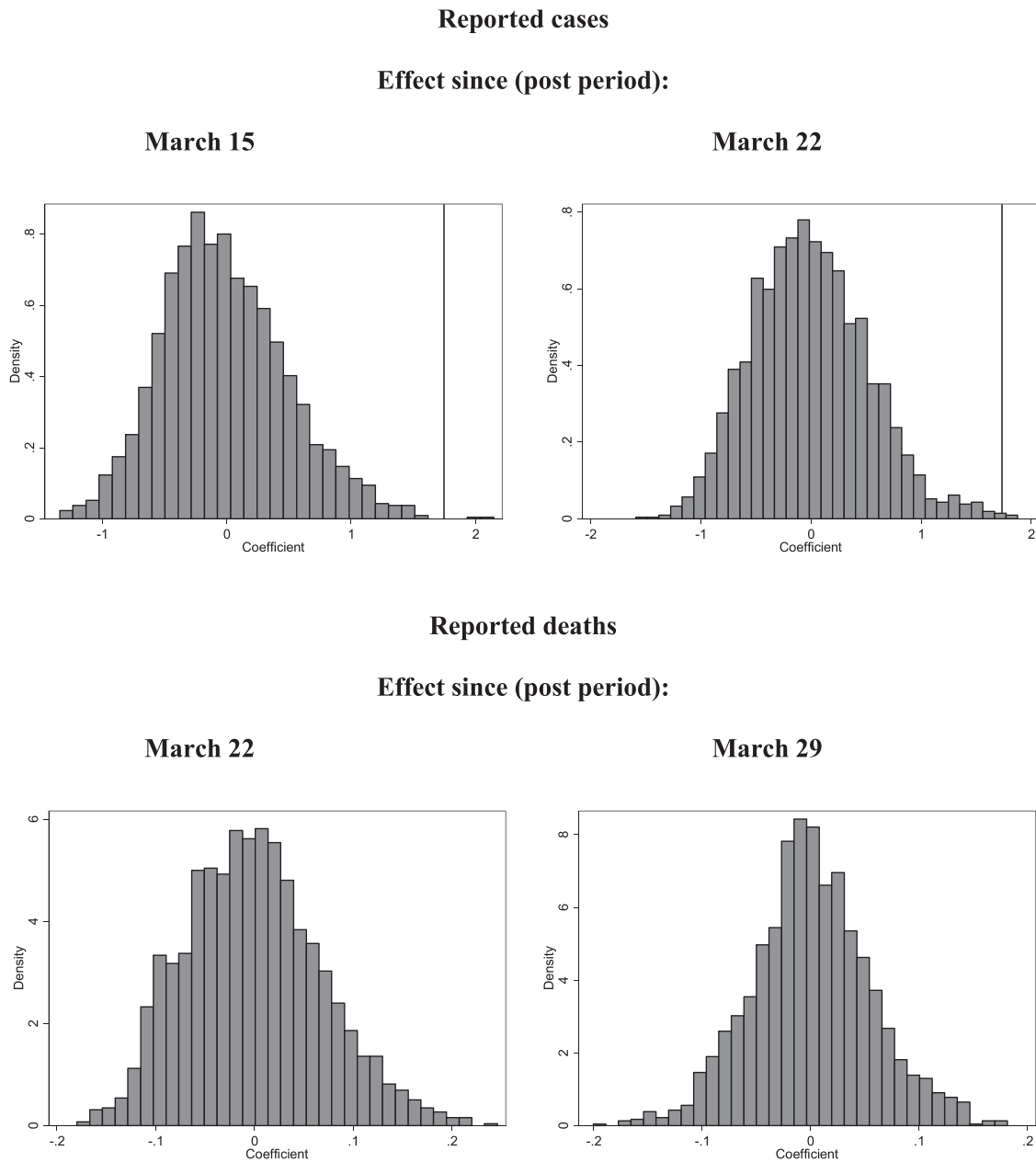


Fig. C1. Distribution of placebo estimates. This appendix shows the distribution of the estimated coefficients from 2000 placebo regressions using random attendance rates. Vertical black lines represent estimated significant coefficients using the actual attendance rates, obtained from Table 1, columns (1)–(4). The mass of density of the placebo estimated coefficients is far to the left of the estimated values using the actual attendance rates.

See [Table D1](#) and [D2](#).

Table D1

Balancing regression: Correlation of the % attendees to mass gathering events (March 6–8) with observed province characteristics.

	(1)
GDP per capita	0.156 * (0.092)
Population density	0.000 (0.002)
Share of employment in services	-8.269 * (4.599)
Pop with college over working age population (aged 16–64)	-2.271 (11.703)
Physicians per 1000 people	0.748 (0.480)
Share of population aged 65 and above	-11.684 * (6.885)
Share of foreign population	-6.637 (6.852)
Annual average temperature	0.037 (0.076)
Annual average precipitation	0.001 (0.001)
Average relative humidity	-0.040 (0.070)
Annual average sunny hours	-0.001 (0.002)
Coast dummy	0.359 (0.400)
Observations	50
R-squared	0.530

Notes: The dependent variable (D.V.) in column (1) is defined as % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women's Day marches (mean: 2.29; S.D.: 1.79). Data on the observed province characteristics are obtained from INE and IVIE for the last period available before March 6–8 at the province level (2018 GDP per capita, 2019 population density, 2019 share of employment in services, 2013 population with college over working age population, 2019 physicians per 1000 people, 2019 share of population aged 65 and above, 2019 share of foreign population, 2013 annual average temperature, 2013 annual average precipitation, 2013 average relative humidity, 2013 annual average sunny hours). Robust standard errors are reported in parentheses. * ** Significant at the 1% level, * * Significant at the 5% level, * Significant at the 10% level.

Table D2
Additional provincial variables COVID-19 cases and deaths response to the variation in the attendees to mass gathering events (March 6–8).

D.V.:	(1) (2)		(3) (4)	
	Reported Cases		Reported Deaths	
Effect Since (Post Period):	March 15	March 22	March 22	March 29
% Attendees without metro	1.610 * ** (0.551)	2.084 * ** (0.700)	0.023 (0.060)	-0.001 (0.050)
Post x GDP per capita	0.297 (0.254)	0.457 (0.329)	0.001 (0.019)	-0.000 (0.017)
Post x Population density	0.005 (0.007)	0.005 (0.006)	-0.000 (0.001)	-0.000 (0.001)
Post x Share of employment in services	24.199 (16.525)	15.734 (17.272)	2.205 * * (1.006)	2.260 * * (1.031)
Post x Pop with college over working age population	33.079 (38.316)	30.755 (40.411)	6.599 * * (3.194)	6.614 * * (3.189)
Post x Physicians per 1000 people	-1.483 (0.890)	-0.717 (1.102)	-0.162 * * (0.078)	-0.167 * * (0.080)
Post x Share of population aged 65 and above	-5.892 (20.578)	-17.853 (26.150)	-0.232 (1.988)	-0.154 (2.013)
Post x Share of foreign population	-29.073 (20.619)	-35.867 (23.664)	-1.103 (1.361)	-1.058 (1.341)
Post x Annual average temperature	-0.119 (0.247)	-0.080 (0.252)	-0.011 (0.021)	-0.012 (0.021)
Post x Annual average precipitation	-0.004 (0.003)	-0.003 (0.003)	-0.000 (0.000)	-0.000 (0.000)
Post x Average relative humidity	0.597 * * (0.244)	0.556 * * (0.255)	0.040 * * (0.016)	0.041 * * (0.016)
Post x Annual average sunny hours	0.003 (0.005)	0.002 (0.005)	0.000 (0.000)	0.000 (0.000)
Post x Coast dummy	0.437 (1.341)	0.805 (1.319)	0.145 (0.098)	0.143 (0.096)
Observations	2350	2350	2350	2350
R-squared	0.713	0.715	0.687	0.687
<i>In the Post Period</i>				
D.V. Mean	12.75	14.73	1.51	1.64
D.V. Standard Deviation	14.63	15.78	1.61	1.61
<i>For all</i>				
Province FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Province-specific linear trends	Yes	Yes	Yes	Yes

Notes: The dependent variable (D.V.) in columns (1) and (2) is defined as COVID-19 cases per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 8.15; S.D.: 12.97). The D.V. in columns (3) and (4) is defined as COVID-19 deaths per 100,000 inhabitants at the province level from February 26 to April 12, 2020 (mean: 0.76; S.D.: 1.33). The post-period summary statistics for the D.V. are measured since March 15, 22, and 29 in each corresponding column. The main explanatory variable is the % attendees to mass gathering events over total population at province level, including major sports events (First and Second Division football (soccer) matches and ACB basketball matches) and Women’s Day marches (mean: 2.29; S.D.: 1.79). The post dummy variables take value 1 after March 9 and 0 otherwise. Robust standard errors are clustered at the province level and reported in parentheses. * * * Significant at the 1% level, * * Significant at the 5% level, * Significant at the 10% level.

References

Adda, J., 2016. Economic activity and the spread of viral diseases: evidence from high frequency data. *Q. J. Econ.* 131 (2), 891–941. <https://doi.org/10.1093/qje/qjw005>.
 Ahammer, A., Martin, H., Lackner, M., 2020. Mass gatherings contributed to early COVID-19 spread: Evidence from US sports (No. 2003).
 Amuedo-Dorantes, C., Marcén, M., Morales, M., Sevilla, A., 2022. COVID-19 School Closures and Parental Labor Supply in the United States (No. 13827; IZA WP).

Amuedo-Dorantes, C., Borra, C., Rivera-Garrido, N., Sevilla, A., 2021. Early adoption of non-pharmaceutical interventions and COVID-19 mortality. *Econ. Hum. Biol.* 42, 101003 <https://doi.org/10.1016/j.ehb.2021.101003>.
 Aparicio Fenoll, A., Grossbard, S., 2020. Intergenerational residence patterns and Covid-19 fatalities in the EU and the US*. *Econ. Hum. Biol.* 39, 100934 <https://doi.org/10.1016/j.ehb.2020.100934>.
 Armillei, F., Filippucci, F., Fletcher, T., 2021. Did Covid-19 hit harder in peripheral areas? The case of Italian municipalities. *Econ. Hum. Biol.* 42, 101018 <https://doi.org/10.1016/J.EHB.2021.101018>.
 Basellini, U., Camarda, C.G., 2021. Explaining regional differences in mortality during the first wave of Covid-19 in Italy. *Popul. Stud.* <https://doi.org/10.1080/00324728.2021.1984551>.
 Callaway, B., Goodman-Bacon, A., Sant’Anna, P.H.C., 2021. Difference-in-Differences with a Continuous Treatment.
 Carlin, P.R., Minard, P., Simon, D.H., Wing, C., 2021. Effects of large gatherings on the COVID-19 epidemic: Evidence from professional and college sports. *Econ. Hum. Biol.* 43, 101033 <https://doi.org/10.1016/j.ehb.2021.101033>.
 CDC, 2021, February 16. Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19). Centers for Disease Control and Prevention. (<https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html>).
 de Chaisemartin, C., D’Haultfoeuille, X., 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *Am. Econ. Rev.* 110 (9), 2964–2996. <https://doi.org/10.1257/AER.20181169>.
 Clemens, M.A., Lewis, E.G., Postel, H.M., 2018. Immigration restrictions as active labor market policy: evidence from the Mexican Bracero exclusion. *Am. Econ. Rev.* 108 (6), 1468–1487. <https://doi.org/10.1257/aer.20170765>.
 Dave, D., Friedson, A., Matsuzawa, K., Mcnichols, D., Redpath, C., Sabia, J.J., 2020. Risk Aversion, Offsetting Community Effects, and COVID-19: Evidence from an Indoor Political Rally * Risk Aversion, Offsetting Community Effects, and COVID-19: Evidence from an Indoor Political Rally.
 Fowler, J.H., Hill, S.J., Remy, R., Obradovich, N., 2020. The Effect of Stay-at-Home Orders on COVID-19 Cases and Fatalities in the United States. In *Arxiv.Org*.
 Ganegoda, N.C., Wijaya, K.P., Amadi, M., Erandi, K.K.W.H., Aldila, D., 2021. Interrelationship between daily COVID-19 cases and average temperature as well as relative humidity in Germany. *Sci. Rep.* 11 (1), 1–16. <https://doi.org/10.1038/s41598-021-90873-5>.
 Goodman-Bacon, A., 2018. Public insurance and mortality: evidence from medicaid implementation. *J. Political Econ.* 126 (1), 216–262. <https://doi.org/10.1086/695528>.
 Goodman-Bacon, A., Marcus, J., 2020. Using Difference-in-Differences to Identify Causal Effects of COVID-19 Policies (No. 1870; DIW Berlin Discussion Paper). Elsevier BV. <https://doi.org/10.2139/ssrn.3603970>.
 Gupta, S., Nguyen, T., Rojas, F.L., Raman, S., Lee, B., Bento, A., Simon, K., Wing, C., 2020. Tracking Public and Private Responses to the COVID-19 Epidemic: Evidence from State and Local Government Actions (No. 27027; NBER WORKING PAPER). <https://doi.org/10.3386/w27027>.
 Hoang, V.T., Gautret, P., 2018. Infectious diseases and mass gatherings. *Current Infectious Disease Reports*. Current Medicine Group LLC, p. 1. <https://doi.org/10.1007/s11908-018-0650-9>.
 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*. Cambridge University Press, pp. 1035–1046. <https://doi.org/10.1017/dmp.2018.161>.
 Kočańczyk, M., Grabowski, F., Lipniacki, T., 2020. Super-spreading events initiated the exponential growth phase of COVID-19 with R_0 higher than initially estimated. *R. Soc. Open Sci.* 7 (9), 200786 <https://doi.org/10.1098/rsos.200786>.
 Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., Zheng, Q., Meredith, H.R., Azman, A.S., Reich, N.G., Lessler, J., 2020. The incubation period of coronavirus disease 2019 (CoVID-19) from publicly reported confirmed cases: Estimation and application. *Ann. Intern. Med.* 172 (9), 577–582. <https://doi.org/10.7326/M20-0504>.
 Marcén, M., Morales, M., 2021. The intensity of COVID-19 nonpharmaceutical interventions and labor market outcomes in the public sector. *J. Reg. Sci.* 61 (4), 775–798. <https://doi.org/10.1111/jors.12535>.
 Michelozzi, P., De’Donato, F., Scortichini, M., Pezzotti, P., Stafoggia, M., De Sario, M., Costa, G., Noccioli, F., Riccardo, F., Bella, A., Demaria, M., Rossi, P., Brusaferrero, S., Rezza, G., Davoli, M., 2020. Temporal dynamics in total excess mortality and COVID-19 deaths in Italian cities. *BMC Public Health* 20 (1), 1–8. <https://doi.org/10.1186/S12889-020-09335-8/TABLES/1>.
 Moritz, S., Gottschick, C., Horn, J., Popp, M., Langer, S., Klee, B., Purschke, O., Gekle, M., Ihling, A., Zimmermann, F.D.L., Mikolajczyk, R., 2021. The risk of indoor sports and culture events for the transmission of COVID-19. *Nat. Commun.* 12 (1), 1–9. <https://doi.org/10.1038/s41467-021-25317-9>.
 Murray, E.J., 2020. Epidemiology’s time of need: COVID-19 calls for epidemic-related economics. *J. Econ. Perspect.* 34, 105–120. <https://doi.org/10.1257/JEP.34.4.105>.
 Nunan, D., Brassey, J., 2020. What is the evidence for social distancing during global pandemics? A rapid summary of current knowledge. In *Oxford COVID-19 Evidence Service*.
 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), e0160378. <https://doi.org/10.1371/journal.pone.0160378>.
 Redondo-Bravo, L., Moros, M.J.S., Sanchez, E.V.M., Lorusso, N., Ubago, A.C., Garcia, V. G., Villanueva, P.S., Azon, A.P., Bescos, J.G., Boone, A.L.D., Ibanez, A.F., Fernandez, B.A., Riutort, A.N., Castell, M.S., Duran, J.G., Gallo, D.N., Moreno, M.L.

- R., de la Fuente, A.B., Raymundo, L.J.V., Larrauri, A., 2020. The first wave of the COVID-19 pandemic in Spain: characterisation of cases and risk factors for severe outcomes, as at 27 April 2020. *Eurosurveillance* 25 (50), 1–13. <https://doi.org/10.2807/1560-7917.ES.2020.25.50.2001431>.
- Richterich, P., 2020. Severe underestimation of COVID-19 case numbers: effect of epidemic growth rate and test restrictions. *MedRxiv*. <https://doi.org/10.1101/2020.04.13.20064220>.
- Siedner, M.J., Harling, G., Reynolds, Z., Gilbert, R.F., Venkataramani, A., Tsai, A.C., 2020. Social distancing to slow the U.S. COVID-19 epidemic: an interrupted time-series analysis. *MedRxiv*. <https://doi.org/10.1101/2020.04.03.20052373>.
- Soriano, V., Ganado-Pinilla, P., Sanchez-Santos, M., Gómez-Gallego, F., Barreiro, P., de Mendoza, C., Corral, O., 2021. Main differences between the first and second waves of COVID-19 in Madrid, Spain. *Int. J. Infect. Dis.* 105, 374–376. <https://doi.org/10.1016/j.ijid.2021.02.115>.
- Tremlett, G., 2020. How did Spain get its coronavirus response so wrong? *The Guardian*. (<https://www.theguardian.com/world/2020/mar/26/spain-coronavirus-response-analysis>).
- WHO, 2015. In: Endericks, T., McCloskey, B., Vincent, E., Llamas, A., Berns, S. (Eds.), *Public Health for Mass Gatherings: Key Considerations*. World Health Organization.