Economics and Human Biology 52 (2024) 101341

Contents lists available at ScienceDirect

Economics and Human Biology

journal homepage: www.elsevier.com/locate/ehb

Full length article

People inflows as a pandemic trigger: Evidence from a quasi-experimental study[☆]

Andrea Caria[a,1], Marco Delogu[b,c,*], Marta Meleddu[b,2], Giovanni Sotgiu[d,3]

[a] University of Cagliari and CRENoS, Italy
[b] DISEA and CRENoS, University of Sassari, Italy
[c] DEM, University of Luxembourg, Luxembourg
[d] University of Sassari, Department of Medicine, Surgery and Pharmacy, Italy

ARTICLE INFO

JEL classification:
I10
Z38
L83

Keywords:
COVID-19
Causality
Mobility
Difference-in-differences
Instrumental variable

ABSTRACT

Although it has been established that population density can contribute to the outbreak of the COVID-19 virus, there is no evidence to suggest that economic activities, which imply a significant change in mobility, played a causal role in the unfolding of the pandemic. In this paper, we exploit the particular situation of Sardinia (Italy) in 2020 to examine how changes in mobility due to tourism inflows (a proxy of economic activities) influenced the development of the COVID-19 pandemic. Using a difference-in-differences approach, we identify a strong causal relationship between tourism flows and the emergence of COVID-19 cases in Sardinia. We estimate the elasticity of COVID-19 cases in relation to the share of tourists to be 4.1%, which increases to 5.1% when excluding local residents. Our analysis suggests that, in the absence of tools preventing the spread of infection, changes in population density due to economic activities trigger the pandemic spreading in previously unaffected locations. This work contributes to the debate on the complex relationship between COVID-19 and the characteristics of locations by providing helpful evidence for risk-prevention policies.

1. Introduction

Recent research has established a causal relationship between population density and the outbreak of the COVID-19 virus (Carozzi et al., 2022),[4] and an extensive body of evidence strongly indicates that mobility significantly correlates with higher infection rates and increased mortality (Kraemer et al., 2020; Zhou et al., 2020; Spelta and Pagnottoni, 2021; Glaeser et al., 2022). However, it is difficult to quantify the exact impact of *changes* in mobility on the spread of COVID-19, because the pandemic has been caused by a complex interaction of various factors. Glaeser et al. (2022) highlight that while all infectious diseases can potentially transmit through human interaction, the specific relationship between mobility and contagion is shaped by the disease's inherent traits and the behaviors of travelers. Consequently, delving into the factors that motivate people's mobility becomes a critical endeavor. This paper takes advantage of tourism-related inflows to uncover the relationship between mobility and the pandemic's outbreak, while also examining the influence of tourism on the dissemination of COVID-19. Higher accessibility amplifies the likelihood of importing infections. Specifically, regions with prominent transportation hubs, where a substantial number of travelers converge, become more vulnerable to potential epidemic outbreaks (Han et al., 2021). Hence, understanding whether there is a causal relationship between tourism and COVID-19 is crucial for developing policies that

☆ We thank the participants to the DISEA seminar (Sassari, Italy), to the 14th Workshop Tourism: Economics and Management (Santiago de Chile, Chile) and to the GSSI seminar (L'Aquila, Italy). We are particularly grateful to Gianfranco Atzeni, Bianca Biagi, the editor and two anonymous referees for their insightful comments. This research was supported by Fondazione di Sardegna (Economia e Benessere tra Società e Istituzioni), Italy, grant 2022-2023.

* Correspondence to: DISEA, University of Sassari, Via Muroni 25 07100 Sassari, Italy.
E-mail addresses: andrea.caria91@unica.it (A. Caria), mdelogu@uniss.it (M. Delogu), mmeleddu@uniss.it (M. Meleddu), gsotgiu@uniss.it (G. Sotgiu).

1 Department of Economics and Business Science - University of Cagliari, Viale S. Ignazio, 84 09123 Cagliari.
2 DISEA, University of Sassari, Via Muroni 25-07100 Sassari.
3 Department of Medicine, Surgery and Pharmacy, University Of Sassari, Viale San Pietro, 43 07100 Sassari.

4 The existence of a direct relationship between density and COVID-19 cases has been questioned in recent studies. For instance, Carozzi et al. (2022) show that the unfolding of the epidemic slowed down in more dense locations because they are populated by individuals more likely to follow social norms that prevent the unfolding of the epidemic. However, there is a consensus that denser locations are more likely to experience an epidemic outbreak, see Díaz Ramírez et al. (2022).

5 Sardinia is an Italian region and the second-biggest Mediterranean island.

https://doi.org/10.1016/j.ehb.2023.101341
Received 22 March 2023; Received in revised form 17 November 2023; Accepted 9 December 2023
Available online 15 December 2023
1570-677X/

can effectively mitigate risk and strike a balance between public health and economic concerns.

The particular situation of Sardinia (Italy),[5] during the year 2020, helps to establish the existence and magnitude of a causal effect between an increase in the movement of people because of economic activity (tourism inflows) and the spread of COVID-19. Traveling to and between Italy was strictly restricted for much of 2020 to fight the spreading of the COVID-19 virus, and such policies prevented a widespread diffusion (Amuedo-Dorantes et al., 2021). The Italian government chose to lift mobility restrictions at the beginning of the summer, thereby allowing tourists to move once again. In fact, the stringent Italian policies concerning the mobility of individuals prevented the diffusion of the virus on the island, with Sardinia enjoying a virtual absence of COVID-19 infections before lifting mobility restrictions.

Sardinia is a popular tourism spot during the summer months, particularly in July and August, when tourism activity is at its highest (CRENOS, 2021).[6] Before the restrictions were lifted, Sardinia was one of the Italian regions least affected by the epidemic. However, the region experienced a sharp increase in the number of cases in August, when tourists were visiting in large numbers the island, which has also been noted by the mainstream press (see, e.g. The Guardian, Giuffrida, 2020).[7] The Italian case was at the forefront of the press debate because, from being almost virus-free at the beginning of the summer, cases soared from nearly 0 to hundreds every day in a matter of weeks.[8] According to these media reports, there appears to be a correlation between the tourist inflows and the spread of COVID-19. Nevertheless, this coincidence is not enough to claim a causal relationship between tourism inflows and the spread of the pandemic. We take advantage of information at the municipality level, similar to Armillei et al. (2021), to estimate how sudden increases in tourism-related mobility impact the pandemic's outbreak.

The lifting of mobility restrictions can be interpreted as a policy shock, allowing us to study the impact of tourist flows on COVID-19 cases in Sardinia municipalities. Noteworthy is that, in Sardinia, the movement of tourists is concentrated in seaside locations, with only a small fraction visiting the inside of the island.[9] Given these characteristics, the case of Sardinia can be seen as a natural experiment to properly analyze the effect of the re-openings and movements of people during the early diffusion of a new virus. The lack of tourists before restrictions were lifted, coupled with their preferences to visit seaside/touristic locations, allows us to use a standard difference-in-differences approach to investigate the causal relationship. As treated units, we consider seaside municipalities, which are also classified as tourist destinations by the Italian National Institute of Statistics (ISTAT, 2022). To ensure the robustness of our findings, we also broaden our criteria for treated units to include all municipalities classified as tourist destinations by ISTAT.

The difference-in-differences approach is only able to capture whether there is a difference between the treated and control units in the outbreak of the epidemic but does not provide an estimate of how many tourists are necessary to trigger the outbreak of the epidemic as it leverages on indirect measures capturing the presence of tourists.

To quantify how the infection grows depending on tourism, we rely on a continuous difference-in-differences approach.[10] For each Sardinia municipality we have information on *Arrivals* (number of tourists that visited the municipality in the specific month) and *Overnight stays* (number of arrivals multiplied by the number of days in Sardinia). We exploit this information to compute, for each Sardinian municipality, a monthly measure of tourism intensity, in terms of the local population. Overall, the analyses uncover a positive and statistically significant impact of tourism inflows on the outbreak of the pandemic, providing additional evidence that population density is a crucial ingredient in the spread of COVID-19.

We verify the validity of our main results by conducting robustness checks. First, we consider a specification of our event study which extends the treatment period. Second, we employ an Instrumental Variables (IV) approach that similarly to the continuous difference-in-differences, produces estimations using as a measure of tourist inflows either data on *Arrivals* or on *Overnight stays*.

Our paper relates to several strands of the literature. Since the pandemic's start, the complex relationship between COVID-19 and socio-demographic and economic features started to be investigated from several perspectives (Armillei et al., 2021). Interestingly, Ascani et al. (2021) found that individuals' mobility was an effective transmission channel in small local market areas. Mobility and lockdown policies limited the spread (Rader et al., 2020; Tantrakarnapa et al., 2020; Amuedo-Dorantes et al., 2021; Farzanegan et al., 2021; Moosa and Khatatbeh, 2021; Perra, 2021). Krisztin et al. (2020) and Han et al. (2021) underscore the significance of international flights, presenting evidence of the pathogen's ability to traverse considerable distances via global travel networks. Laroze et al. (2021) and Han et al. (2021) both demonstrate substantial and positive connectivity-driven spillover effects in virus transmission through commuter flow data. Notably, Glaeser et al. (2022) establish a direct correlation between COVID-19 cases per capita and mobility by analyzing zip code data from five U.S. cities. Building upon this line of thought, Gianmoena and Rios (2023), using cross-province population movements within Italian provinces, observe that a rise in infections within a specific province yields a noteworthy and statistically meaningful impact on the infections occurring in neighboring provinces.

However, the impact of tourism movements on the spread of COVID-19 has not yet been thoroughly analyzed. Indeed, the literature concentrates on studying the impact of the pandemic on tourism in terms of its effect on arrivals in a set of countries (Karabulut et al., 2020 consider 129 countries), a single country (Della Corte et al., 2021 for Italy; (Batalha et al., 2022) for the housing market in Portugal), or top urban destinations (Anguera-Torrell et al., 2021 study 16 worldwide cities). A more limited strand of studies has tried to uncover the reverse: the effect of tourism on the spread of COVID-19. On this matter, Mallapaty (2020) considers the effect of cruise tourism on COVID-19, offering the first hint that the mobility of people for tourism reasons could be a leading driver of propagation. For instance, Farzanegan et al. (2021), studying the role of tourism in the virus outbreak, find a positive correlation between countries with higher flows and COVID-19 cases and deaths; they conclude that tourism may have facilitated the spread of the virus. Interestingly, Casini and Roccetti (2020) investigate the effect of domestic tourism flows on new infections of COVID-19 in the Italian regions by using the Italian flows in 2019 because of the unavailability of 2020 observations. Their findings suggest that tourism movements influenced the spread of COVID-19 in 2020, confirming the intuition provided by the observation of daily COVID-19 data. Armillei et al. (2021) also studied the Italian case, examining the relationship between the first wave of COVID-19 in March 2020, center-periphery dynamics, and related socio-demographic and economic factors.

The plan of the paper is as follows. The next Section describes the institutional context. Section 3 describes the data used, along with the preliminary evidence. Section 4 defines the empirical strategy used to uncover the causal relationship. Section 5 reports the empirical results. Section 6 verifies the validity of our main findings by conducting several robustness checks. Section 7 discusses the limitations of the study and identifies areas for further research. Section 8 concludes the paper.

[5] Sardinia is an Italian region and the second-biggest Mediterranean island.

[6] Also, the fact that the island has six main access points (three by sea and three by air), with limited connections in the winter season, allowed for agile access control during 2020.

[7] See Guardian, link

[8] See Pietromarchi V. at link and Matthews, at link.

[9] In 2019, the percentage of tourist attendance found in July and August was 50%; this share rose to 82% in the months between June and September, clearly signaling the presence of sun-and-sand tourism (CRENOS, 2021).

[10] Our approach mimics the one of Batalha et al. (2022). As they, we acknowledge that inference with a continuous treatment variable requires a stronger validation for the parallel trends assumption, as outlined in Callaway et al. (2021).

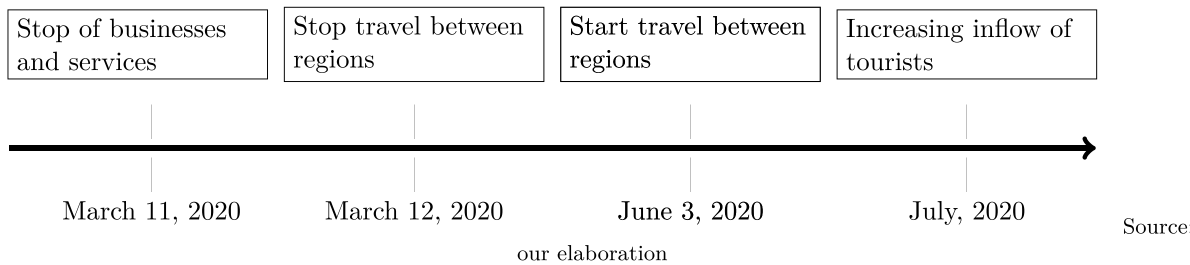


Fig. 1. Timeline of regulations in Italy during 2020.
Source: Our elaboration.

Table 1
Descriptive statistics.

	Nonseaside location			Seaside location		
	Mean	SD	Obs.	Mean	SD	Obs.
COVID-19 cases	1.08	5.72	1525	4.23	20.0	325
Arrivals	54.0	281.8	1525	3535.7	6154.4	325
Overnight stays	150.0	712.7	1525	17 118.3	31 389.9	325
Population	3183.4	8405.3	1525	10 261.7	22 024.7	325
Tourism location	64%	0.48	1525	100%	0	325
N. of locations (municipalities)		305			65	

2. Institutional context

Italy was the first European country that had to face the spread of the new COVID-19 virus in the beginning of 2020. With very little information concerning the characteristics of the disease and with a rapid spread in the number of cases, Italian policymakers implemented very rigid rules to reduce the diffusion of the virus, as noted by Vinceti et al. (2020). The northern region of Lombardy was hit first and put in lockdown on March 8, 2020. Soon, the policymakers realized how serious the situation was, and on March 11, 2020, the orders came to stop all non-necessary business and services activities. Ten days later, travel between regions was not allowed, and only in June 2020 free mobility across Italian regions was restored.[11]

Fig. 1 displays the unfolding of the regulations in Italy in 2020. These measures showed good performance in stopping the spread of the virus. On June 3, 2020, with COVID-19 cases almost confined in the northern region of Lombardy, the mobility between regions, businesses, and services started again.[12] People were free to move across the country at this stage, while international travels and arrivals were still partially regulated. We highlight that tourism flows did not restart immediately: very few people were willing to move for leisure purposes before July 2020, as discussed in Section 3.1.

3. Data

To investigate the relationship between touristic inflows and the evolution of COVID-19 cases, we retrieve data on both variables at the municipality levels.

Data on COVID-19 cases are provided by the Italian National Institute of Health (Istituto Superiore di Sanita, ISS) upon the authors' request.[13] ISS supplied information on the number of COVID-19 cases for each Sardinian municipality for each month of the year 2020. Section 3.1 describes the evolution of COVID-19 cases in Sardinia during this period.

Data on tourism flows are provided by SIRED, a data collection and processing information system provided by the Region of Sardinia (see link), from which we retrieve, for all Sardinian municipalities, information on *Overnight stays* (the number of tourists multiplied by the days of journey) and *Arrivals* (number of arrivals). In Sardinia, the bulk of tourism takes place during the summer, with the peak in the month of August (CRENOS, 2021). Importantly, the allocation of tourists among Sardinian municipalities is not random, with the vast majority deciding to visit Sardinia during the summer period in municipalities located close to the seaside (see Figs. 2b and 2c), with few individuals going to places inland. To capture the difference in tourist attractiveness, we take advantage of a dataset produced by ISTAT that classifies Sardinian municipalities according to the type of tourism. Although all seaside municipalities are classified as tourist locations, 64% of Sardinian municipalities located far from the seaside meet the requirements to be classified as touristic municipalities, as shown in Table 1. From Table 1 we can also see that seaside locations host, on average, a larger population than nonseaside locations.

Table 1 reports the descriptive statistics on the number of COVID-19 cases and tourist inflows at the municipal level, differentiating between seaside and nonseaside locations.[14] Statistics are from the municipalities in Sardinia from May 2020 to September 2020. We report the mean, standard deviation, and total non-null observations divided by nonseaside or seaside location.

Table 1 also reports that COVID-19 cases in seaside municipalities were almost four times higher than those counted in nonseaside municipalities. *Arrivals* and *Overnight stays* were also much higher in seaside municipalities than in nonseaside municipalities.

Finally, we collect information about each municipality from the 2011 Census regarding the housing situation, employment and education situation, and labor mobility, employed to use as controls in part of our empirical analyses.[15] Specifically, we retrieve a housing crowding index, the percentage ratio of unused buildings to total buildings, the share of large families, the female employment rate, the unemployment rate, an index of long-range mobility, the ratio between the area of population centers and cores to the total area, the share of illiterates, the share of adults with a high school diploma or degree, the share

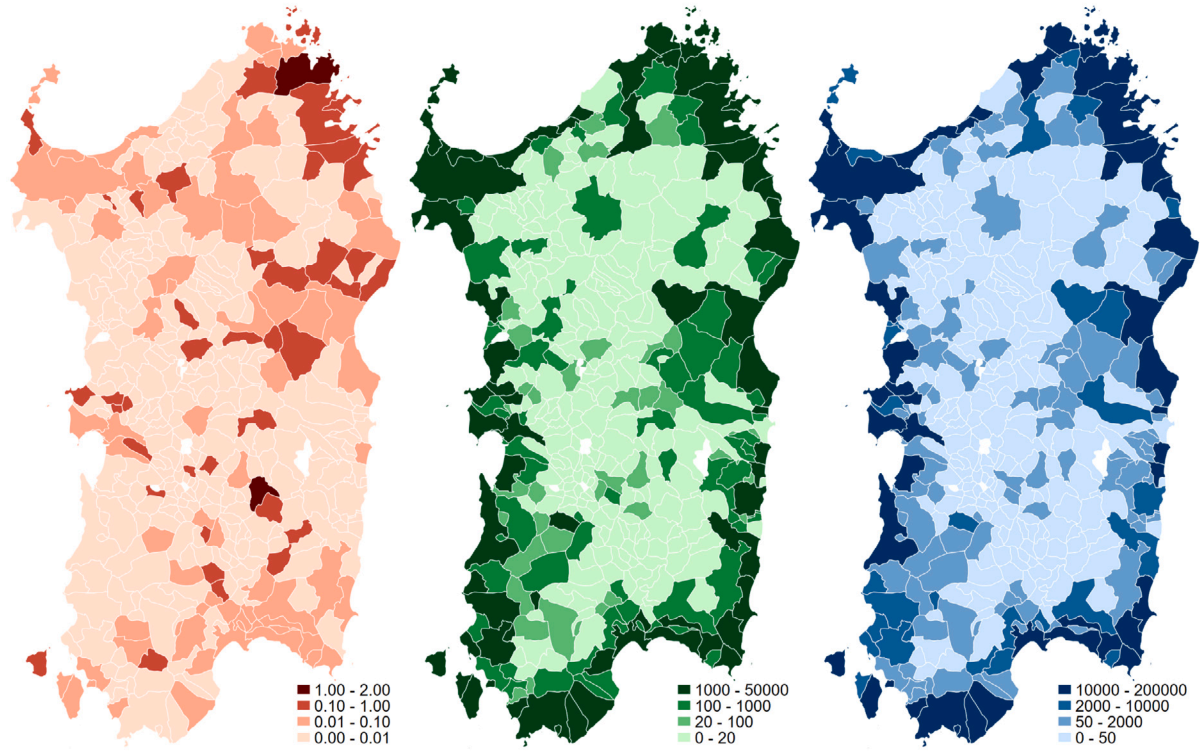
[11] See the document of the relative Italian DPCM link and of the Italian Gazzetta Ufficiale link.

[12] See Gazzetta Ufficiale: link.

[13] For privacy reasons, municipalities that record a positive number of cases but fewer than 5 do not report the exact number of cases. In our computations and estimations, we replace the 0-5 interval with its average, which is 2.5.

[14] Table A.1 in the Appendix provides a detailed description for each variable considered.

[15] For further information please visit the ISTAT website, see link.



(a) *Cases %* (b) *Arrivals* (c) *Overnight stays*

Fig. 2. Geographical distribution.

of employment in high-medium-skilled occupation variation, and the intercensual variation.[16]

During the summer of 2020, following the resumption of travel, Sardinia recorded a total of 1,148,720 tourists in July, August, and September. Notably, 93% of these tourists opted to lodge in seaside towns.

Panel (a) of Fig. 2 depicts the geographical distribution of COVID-19 cases (in percentage), panel (b) shows touristic arrivals, and panel (c) reports the distribution of *Overnight stays*. All figures refer to the month of August. Specifically, panel (a) shows the geographical distribution of the percentage of COVID-19 cases in August in terms of inhabitants (population of each municipality). Panel (b) depicts the tourist arrivals in each municipality in absolute value. The third one reports the *Overnight stays* in absolute value (*Arrivals* multiplied by days of stay). Fig. 2 allows us to easily spot the similarities between COVID-19 cases and touristic presence during August in Sardinia. In addition, it shows that both tourism and COVID-19 cases were concentrated within seaside municipalities.

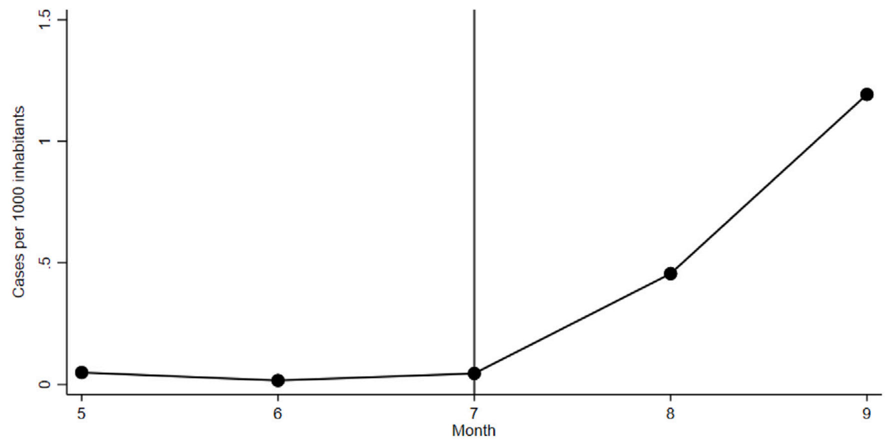
3.1. Preliminary evidence

According to the World Health Organization, COVID-19 is spread in three main ways: (i) the virus can spread from an infected person's mouth or nose through coughing, speaking, and breathing; (ii) in poorly ventilated settings where aerosols remain suspended in the air; and (iii) by touching the eyes, mouth, or noses after touching contaminated surfaces. The spread of the virus requires people contact and is facilitated by the gathering of people, especially in closed places. The strict Italian confinement policies implemented between March 2020 to June 2020 effectively stopped the spread of the epidemic. Fig. 3, shows the evolution of the number of COVID-19 cases per 1000 inhabitants in Sardinia, between May 2020 and September 2020.[17] Preventing people's movement was highly effective in reducing the diffusion of COVID-19 cases, especially in isolated regions like Sardinia, where the virus was virtually absent once mobility was restored (see Fig. 3). However, from panel (a) of Fig. 3, we see that the number of cases started increasing in July 2020, with Sardinia experiencing a substantial number of cases, in seaside locations starting from August. In panel (b) of Fig. 3, we report the evolution considering separately seaside municipalities versus nonseaside locations. The display of this graph corroborates the intuition that, after lifting the mobility restrictions, COVID-19 cases increased faster in seaside locations. Assuming that COVID-19 was absent in Sardinia, as suggested by Fig. 3, it is natural to infer that touristic inflows were a likely trigger of the spread of COVID-19 cases in Sardinia. As our primary focus is to investigate the influence of tourism on the pandemic's outbreak the benchmark econometric analysis studies the evolution of COVID-19 cases from May 2020 up to September 2020. Nevertheless, we also conduct a similar analysis that includes data for March, April, October, and November as part of our robustness checks.

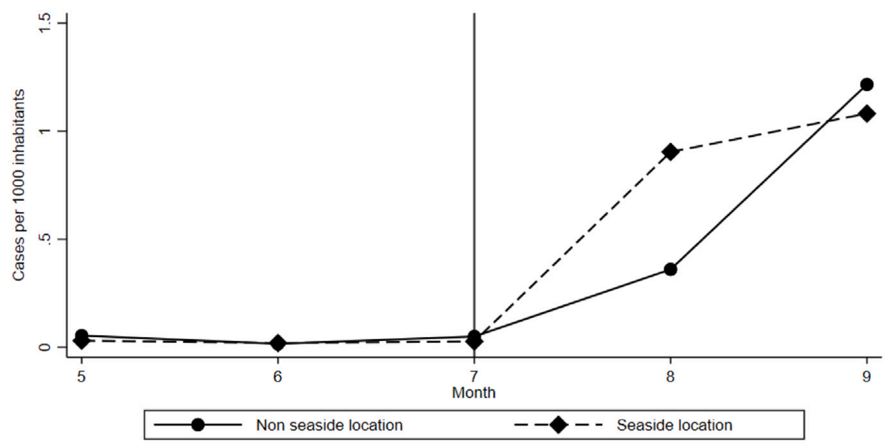
Section 5 assesses the existence of a causal link between tourist inflows and COVID-19 cases by implementing difference-in-differences estimations. The next Section defines and discusses our identification strategy.

[16] In the Appendix in Appendix A, Table A.1 provides a detailed explanation of these variables.

[17] Appendix B.1 of the appendix reports similar figures including data for the months of October and November. While witnessing a similar substantial increase in COVID-19 cases among seaside municipalities, Fig. B.1 suggests that additional spatial and dynamic effects influenced the evolution of the pandemic during the autumn of 2020.



(a) *Cases*



(b) *Cases by location*

Fig. 3. Cases evolution.
Notes: Panel (a) reports the evolution of the number of COVID-19 cases in terms of 1000 inhabitants. Panel (b) reports the number of cases (per 1000 inhabitants) differentiating between nonseaside (solid line) and seaside (dashed line) locations.

4. Empirical strategy

Uncovering whether inflows of people due to economic activities have a direct impact on the diffusion of COVID-19 cases is challenging for several reasons. First, it is not straightforward to differentiate mobility arising from other reasons (such as work) from touristic inflows. Second, locations characterized by a higher population density facilitate pandemic outbreaks, as discussed in Section 1. Nevertheless, the particular characteristics of Sardinia make these concerns less effective. In fact, as outlined in Section 2, most of the inflows to Sardinia during the summer period are because of tourism. Also, we cannot rule out the hypothesis that the virus was already present in the resident population, giving rise to a spurious correlation with tourism. However, the data reported in Section 3.1 show that COVID-19 was almost absent in Sardinia before lifting mobility restrictions.

To uncover the causal relationship between tourist inflows and COVID-19 cases, we rely on one of the most widely utilized econometric techniques for analyzing observational data: difference-in-differences. This method allows us to exploit the quasi-experimental variation in tourist inflows induced by relaxing mobility restrictions.

A crucial assumption for a difference-in-differences estimation is that the evolution of COVID-19 cases would have had parallel trends between the control and treated units. In our analysis, this means that, without lifting the mobility restrictions (i.e., absent tourism), the trend in the number of cases among *treated and non-treated* locations would have not changed. To evaluate the soundness of this identifying assumption in Section 3.1, we show that there is almost no difference in the evolution of cases before July 2020 between seaside and nonseaside locations.[18] Taking advantage of the fact that the lifting of restrictions in terms of tourist inflows affects mainly seaside municipalities, a difference-in-differences estimation should capture the average impact of tourism inflows on the number of COVID-19 cases.[19] The standard difference-in-differences approach captures the impact of tourism on the outbreak of the COVID-19 pandemic by leveraging the differential spread of tourists among seaside and nonseaside municipalities. However, it provides only an indirect measurement, which precludes us from deriving the elasticity of COVID-19 cases in response to variations in tourist inflows. To address this limitation, in line with Batalha et al. (2022) and González-Val and Marcén (2022), we employ a continuous difference-in-differences specification. Specifically, we define our continuous treatment (the *dosage*) as the share of tourists in terms of the

[18] We obtain comparable evidence when the treatment group includes all the municipalities classified as touristic by ISTAT.

[19] We acknowledge that the lifting of restrictions affected the remaining municipalities to a lesser degree. It should be noted that this effect goes in the opposite direction, downgrading the estimate's magnitude.

resident population, interacting this variable with the periods where mobility was restored. Furthermore, in Section 6, we also evaluate our benchmark findings by implementing an instrumental variable estimation.

5. Results

This section presents the empirical findings. First, in Section 5.1 we evaluate whether the parallel trends assumption holds. Section 5.2 reports the results of the difference-in-differences estimations.

5.1. Panel event study

We start by considering a standard panel event study setup (Clarke and Tapia-Schythe, 2021), to investigate whether tourist locations are the ones that triggered the COVID-19 outbreak experienced in Sardinia during the summer of 2020.

First, we include in the treated group only seaside municipalities, which are a subset of the ISTAT-defined tourist destinations, see Section 3. Our analysis exploits the fact that the vast majority of tourists visit seaside municipalities, as shown in panels (b) and (c) of Fig. 2 and discussed in Section 3.1. To ensure a comprehensive analysis, we also utilize the ISTAT classification to distinguish between municipalities in Sardinia that are considered tourist destinations and those that are not. In this context, the treated group comprises municipalities classified as tourist destinations, while the remaining ones serve as the control group. To evaluate the parallel trends assumption we estimate the following specification:

$$Y_{m,t} = \alpha + \sum_{j=1}^{J} \beta_j (Lead\ j)_{m,t} + \sum_{k=1}^{K} \gamma_k (Lag\ k)_{m,k} + \Lambda X_{st} + \phi_m + \gamma_t + \epsilon_{m,t}. \quad (1)$$

On the left side of Eq. (1) $Y_{m,t}$, is the number of monthly (denoted by the subscript t) COVID-19 cases recorded for each Sardinian municipality (denoted by the subscript m). On the right side of Eq. (1) $Leads$ and $Lags$ are dummy variables capturing whether the treated municipality was a given number of periods before (j) or away (k) from the lifting of mobility restrictions. Estimating the coefficients on the Lead variables allows us to test for the existence of pre-trends. Conversely, the estimation of the coefficients of the Lag variables allows us to evaluate the existence of dynamic effects of the tourism inflows on COVID-19 cases. The first leading variable refers to May which is two months before the inflow of tourists restarted; accordingly, we have $j = 2$. In addition, we account for three post-periods, $k = 3$.[20] To estimate Eq. (1), we set the baseline to the first lag (which refers to July 2020, when tourist inflows restarted). In line with the standard approach of a panel-event study, we include both municipality and month fixed effects, ϕ_m and γ_t, respectively, in Eq. (1).

With these sets of fixed effects, the control vector, $X_{m,t}$, includes only month-changing variables. Specifically, in Eq. (1) $X_{m,t}$, represents the percentage of monthly arrivals from Sardinia, No EU Countries, and North and Centre Italy. These variables allow us to control for the difference in the spread of the virus according to tourist origin area.

In a panel event study, it is important to account for the possibility of serial correlation in the outcome variable. This issue is usually dealt with by clustering the standard errors at the unit level. Accordingly, we cluster the standard errors at the municipality level.

Fig. 4 presents the Leads and Lags estimates. The dots represent the point estimates, and the blue and black bars report 90% and 95% confidence intervals, respectively.

Fig. 4, in Panel *a*, shows the estimates obtained when we consider municipalities located along the seaside as the treated group; panel *b*, reports the estimates when the treated municipalities include all the municipalities classified as touristic by ISTAT.

We estimate a sizable impact once we consider the subset of seaside municipalities as treated units; see panel *a* of Fig. 4. In this case, we estimate a difference of + 8 cases in August for the treated group. The effect persists as we uncover a positive and weaker effect in September (+6.8 cases). The results reported in panel *b* of Fig. 4 also depict a positive association between COVID-19 cases and tourist municipalities in August and September (+3 cases for both months). The coefficients obtained for the Leads variables align with the absence of pre-trends in the outcome variable.[21]

We also perform an event study to assess the absence of pre-existing trends for the continuous treatment, specifically the monthly inflow of tourists in Sardinian municipalities (*the dosage)*. This approach closely aligns with the methodology employed by Batalha et al. (2022) and González-Val and Marcén (2022), both of whom explored related subjects using a continuous difference-in-differences specification.[22] More precisely, we replace the treatment indicator with a variable that represents the variations in monthly tourist inflows across Sardinian municipalities throughout the treatment period. For each municipality, we compute the monthly share of tourists relative to the resident population, denoted as $st_{m,t} = \left(\frac{Tourists_{m,t}}{Pop_m}\right)$. Fig. 5 illustrates the results using *Arrivals* data as a measure of tourist inflows and it aligns with the parallel trends assumption.[23] Comparable findings are also obtained when utilizing data on *Overnight stays*, see Section 3.[24]

In the following section, we delve deeper into the continuous treatment specification, where we also present the results derived from estimating both the standard difference-in-differences and the continuous specification.

5.2. Difference-in-differences

To uncover the impact of tourist inflows on the COVID-19 outbreak we estimate a standard difference-in-differences specification by employing both definitions of the treated units, specifically:

$$Y_{m,t} = \alpha + \beta Post * Treat + \Lambda X_{st} + \phi_m + \gamma_t + \epsilon_{m,t}. \quad (2)$$

On the left side of Eq. (2) $Y_{m,t}$, is the number of monthly COVID-19 cases reported in each Sardinian municipality. The subscript m denotes the municipality, while the subscript t accounts for the month. $Post$ is a binary variable that assumes a value of 1 when the current month is after June, while $Treat$ is a binary variable that indicates the treated municipalities. Thus, with our identifying assumptions, β of the interaction $Post * Treat$ captures the effect of tourism inflows on the spread of COVID-19 cases in the treated municipalities. Similar to Eq. (1) ϕ_m and γ_t are the municipalities and month-fixed effect. We also consider specification with and without the vector $X_{s,t}$ which includes the month-changing controls. Table 2 reports the estimates of our standard difference-in-differences specification.

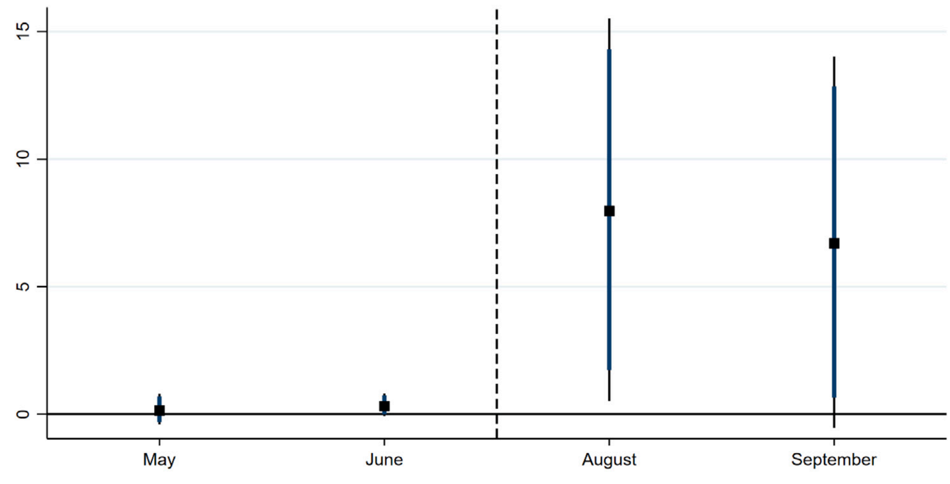
[20] The Leads and Lags variables take a value equal to 1 if the municipality is in the treated group (touristic/seaside) and the month is the k period before the event (Leads) or j periods after the event (Lags). Otherwise, these variables take a value equal to zero. For non-treated municipalities, these dummy variables always take values equal to zero.

[21] Table A.2 in the appendix reports the empirical estimates obtained considering Eq. (1).

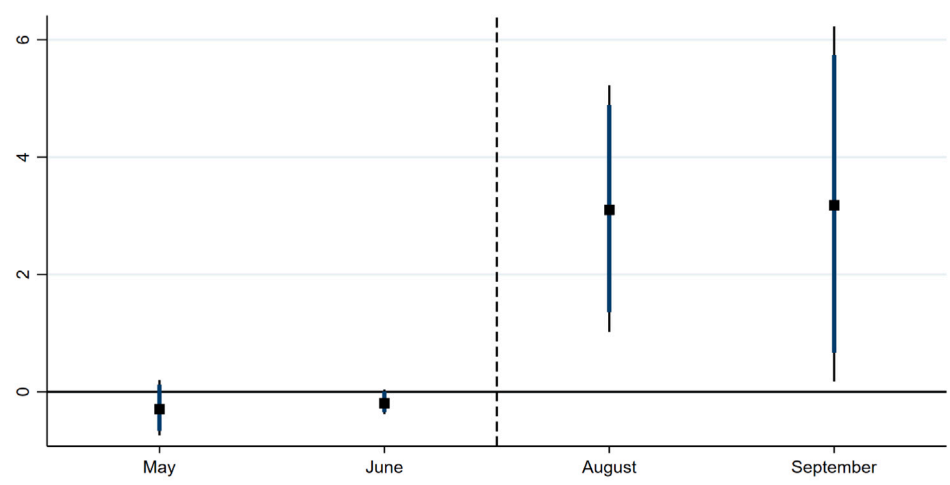
[22] Batalha et al. (2022) show that COVID-19 had a negative impact on short-term rentals. They employed a continuous difference-in-differences specification with the treatment variable being the share of short-term rental units. González-Val and Marcén (2022) explored the role of mass gatherings on the spread of COVID-19, with the treatment variable being the number of attendees per 100 inhabitants at mass gathering events.

[23] We acknowledge that Callaway et al. (2021) proved that pre-trends tests used for the binary case are not necessarily useful for the continuous case. A stronger version of this assumption is required. Units treated with the same dose should follow the same evolution.

[24] Results using the latter measure are available upon request.



(a) *Seaside locations*



(b) *Touristic locations*

Fig. 4. Panel event study estimates.

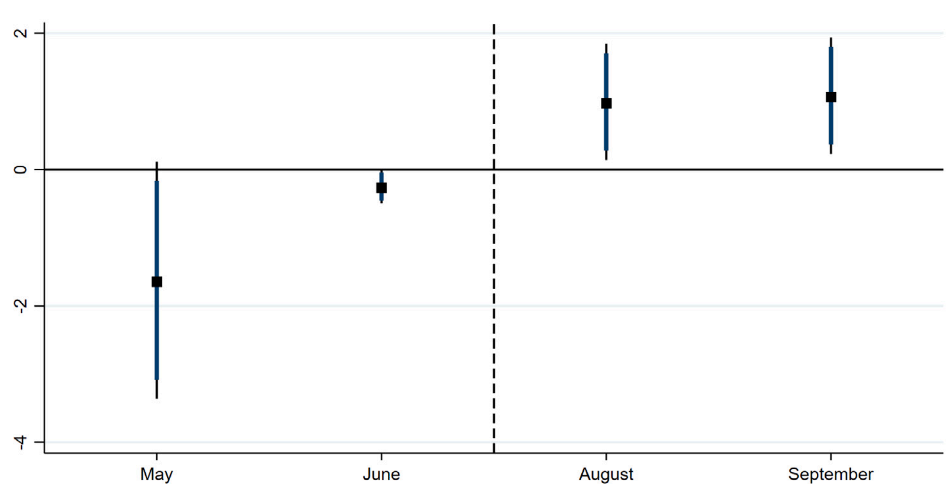


Fig. 5. Continuous treatment.

Table 2
Standard difference-in-differences: linear specification.

	(1) Seaside	(2) Touristic	(3) Seaside	(4) Touristic
Treatment	5.016**	2.210***	4.552**	2.321***
	(2.440)	(0.768)	(2.311)	(0.834)
Observations	1850	1850	1850	1850
Controls	no	no	yes	yes
Municipality fixed effects	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes

Notes: We report the results using seaside locations as the treatment group (Columns (1) and (3)) and tourist locations (Columns (2) and (4)). The time span for the study ranges from May 2020 to September 2020. Standard errors in parentheses are clustered at the municipality level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3
Standard difference-in differences: log-linear specification.

	(1) Seaside	(2) Touristic	(3) Seaside	(4) Touristic
Treatment	0.332***	0.321***	0.280***	0.316***
	(0.0982)	(0.0500)	(0.0958)	(0.0516)
Observations	1850	1850	1850	1850
Controls	no	no	yes	yes
Municipality fixed effects	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes

Notes: We report the results using seaside locations as the treatment group (Columns (1) and (3)) and tourist locations (Columns (2) and (4). The time span for the study ranges from May 2020 to September 2020. Standard errors in parentheses are clustered at the municipality level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results presented in Table 2 offer compelling evidence of a positive causal link between tourism inflows and the outbreak of COVID-19. When considering solely the seaside municipalities in the treated group, we observe a substantial effect, resulting in an increase of approximately +4.5 to +5 cases (as shown in columns (1) and (3) of Table 2). When we expand the treated group to encompass all municipalities classified as tourist destinations, the estimated effects, while somewhat reduced in magnitude, remain highly statistically significant, at around +2.2 to +2.3 cases (as shown in columns (2) and (4) of Table 2).

The specifications presented in Table 2 indeed indicate a positive causal relationship between tourist inflows and the COVID-19 outbreak in Sardinia. However, interpreting the exact magnitude of this impact can be challenging. To simplify the interpretation, we adopt a slightly modified version of Eq. (2). In this new specification, similar to Glaeser et al. (2022), we do not use the absolute number of cases as the left-hand side (LHS) variable. Instead, we employ the logarithm of the monthly number of cases at the municipal level. To ensure continuity in our analysis and preserve observations with zero values, we add 1 to the number of cases before taking the logarithm.[25] Notably, this adjustment allows us to interpret the estimate of the interaction term as an elasticity, effectively measuring the percentage increase or decrease in COVID-19 cases in tourist destinations during the treatment period. The results of this supplementary analysis are presented in Table 3. The findings strengthen the evidence of a positive causal impact of tourist inflows on the COVID-19 outbreak.

These additional estimates also indicate that there was a statistically significant higher number of COVID-19 cases in tourist municipalities during the treatment period. Specifically, results suggest that tourism was associated with a larger number of cases, with an impact ranging from $e^{0.28} = 32\%$ to $e^{0.32} = 39\%$, depending on the definition of *treated* units, when including month-changing controls.[26]

[25] Zeros are high prevalent in our dataset representing 82.14% of the total observations.

[26] We remain cautious in interpreting these estimates as average treatment effects. In fact, our sample includes a large percentage of zeros (82%), and the 11% of observations record a positive number of cases stored as 2.5, see Section 3. It is likely, that the magnitude of the estimated effect is sensitive to this rounding.

This first set of estimations provides significant evidence of the direct influence of tourist inflows on the COVID-19 outbreak experienced in Sardinia during the summer of 2020. Nevertheless, the binary indicators used heretofore to examine the impact of tourism on the occurrence of COVID-19 cases are indirect measures that provide only an average effect for treated units. These indicators do not account for the variations in tourist inflows among different destinations in Sardinia. This can be important as there is significant variation in the number of tourists received by different municipalities, even among seaside municipalities. To address this heterogeneity, we employ a continuous difference-in-difference estimation.

Similar to related studies (Batalha et al., 2022; González-Val and Marcén, 2022) we replaced our binary treatment indicator with a continuous treatment variable. This approach enables us to consider the variation in tourist inflows among the treated municipalities and estimate the direct impact of tourist inflows on COVID-19 cases. The treatment variable, denoted as st_{tm} and introduced in the preceding section, represents monthly tourists in Sardinian municipalities relative to the local population.[27] The dependent variable, the number of monthly COVID-19 cases reported in each Sardinian municipality is taken in logs, as in the estimations reported in Table 3.

For completeness, we present the new specification below, which involves a slight modification of Eq. (2), as follows:

$$ln\left(Y_{m,t}+1\right) = \alpha + \beta Post \cdot ln\left(st_{m,t}\right) + \phi_m + \gamma_t + \epsilon_{m,t}. \tag{3}$$

We estimate two different specifications of Eq. (3) for each of our two measures of tourist inflows (i.e., *Arrivals* and *Overnight Stays*). The initial specification (Columns 1 and 3 in Table 4) which accounts for Sardinian residents as part of tourist inflows, thus considering those who opted to spend their holidays within the region. However, it can be argued that internal tourists should not be factored in when assessing the impact of tourist inflows on COVID-19 cases. Hence, we present the results of an alternative specification (columns 2 and 4 in Table 4) where the calculation of $st_{m,t}$ excludes local residents.

[27] To ensure the inclusion of all observations, we added ones to st_{tm} before taking logarithms.

Table 4
Continuous difference-in-differences.

	Arrivals		Overnight stays	
	(1) Overall inflows	(2) External inflows	(3) Overall inflows	(4) External inflows
Treatment	0.411***	0.518***	0.202***	0.236***
	(0.120)	(0.154)	(0.0545)	(0.0788)
Observations	1850	1850	1850	1850
Municipality fixed effects	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes

Notes: In Columns (1) and (2), we calculate the share of tourists using monthly *Arrival* data. In Columns (3) and (4), we calculate the share of tourists using *Overnight stays* data. Columns (1) and (3) include the entire number of tourists, whereas Columns (2) and (4) exclude locals from the calculation of these shares. The time span ranges between May 2020 and September 2020. Standard errors in parentheses are clustered at the municipality level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results presented in Table 4 provide additional evidence supporting a causal link between tourist inflows and the COVID-19 outbreak. Notably this new specification provides an approximate estimate of the elasticity of COVID-19 cases to the share of tourists, resulting in a figure of 4.1% when internal tourists are included (see Column 1). Interestingly, this estimate rises to 5.1% when we calculate $st_{m,t}$ by excluding local residents, considering our tourism measure only for individuals arriving from outside the island.

6. Robustness checks

In this section, we enhance the robustness of our findings by first, extending the time span under analysis and considering an instrumental variables approach.

In the first set of robustness checks (see Appendix B.1), we extend the time span to encompass additional months: specifically, March, April, October, and November. The primary estimates are based on a shorter time frame (i.e., from May to September). This selection is consistent with the core focus of our study, which is to investigate the impact of tourist inflows on the outbreak of the pandemic. Indeed, on one hand, including March and April in the analysis may inadvertently capture some spurious pre-trends that are mitigated by the implementation of lockdown policies. On the other hand, the trajectory of COVID-19 cases during the months of October and November is likely influenced by spatial and dynamic factors that could potentially confound the influence of tourism (Glaeser et al., 2022). Accounting for the role of these effects is beyond the scope of this study, which primarily aims to emphasize the impact of tourism inflows on the pandemic outbreak. However, in the robustness check, by employing this wider time span, we repeat the analyses presented in the previous sections for the benchmark period.

In a first inspection, we visualize the evolution of the number of COVID-19 cases per 1000 inhabitants in Sardinia, between March 2020 and November 2020. Basically, we produce a figure equivalent to Fig. 3. Interestingly, the results (Fig. B.2) somewhat align with our benchmark findings, but the number of monthly COVID-19 cases per 1000 inhabitants in nonseaside locations is higher for October and November 2020. Afterward, we replicate the panel event study presented in Eq. (1). During October and November, we observe higher estimates. When restricting the treatment group to encompass only seaside locations, these estimates no longer exhibit statistically significant differences from zero at the significance level. In the final examination of the extended time span, we perform the standard difference-in-differences estimations outlined in Section 5.2 with the wider time span. Notably, the estimates of the model of Eq. (2) are bigger in magnitude and still highly statistically significant (Tables B.1 and B.2 in Appendix). These results suggest that considering a wider time span may amplify the role of tourism in the outbreak of the pandemic by absorbing dynamic and spatial effects.

In the second robustness check, we consider an instrumental variable (IV) specification to evaluate the robustness of our findings to possible endogeneity concerns (see Appendix B.2 of the Appendix).[28] Indeed, data on tourist inflows may also include people traveling for work or other purposes. Plus, the choice of a tourist destination could be influenced by the presence and spread of COVID-19 in those municipalities. However, considering that being a seaside municipality does not have a direct effect on the spread of COVID-19 cases, we can address endogeneity concerns by using a dummy variable as an instrument for tourist inflows. This dummy variable takes the value of one for seaside municipalities and zero for others. The IV estimates reported in Tables B.3 and B.4 confirm the role of tourist inflows in the outbreak of the COVID-19 pandemic. In line with the results of the continuous difference-in-difference specification, we obtain lower estimates once we consider *Overnight stays* to capture the number of tourists. The lower magnitude obtained when using this variable is likely due to the fact that the same individual is counted a number of times equal to the days he/she spent in Sardinia.

7. Discussion

This study highlights a statistically significant relationship between tourism and increased COVID-19 cases in tourist municipalities. Specifically, We first perform a standard difference-in-differences approach which shows that there is substantial evidence of a significant impact especially when focusing on seaside municipalities in the treated group. When the treated group includes all municipalities classified as tourist destinations, the estimated effects remain statistically significant. Subsequently, employing a continuous difference-in-differences approach yields additional evidence that strengthens the causal link between tourist inflows and the COVID-19 outbreak. Importantly, this approach provides an approximate estimate of the elasticity of COVID-19 cases with respect to the share of tourists. Specifically, it yields a 4.1% increase in cases when including internal tourists. Interestingly, this estimate increases to 5.1% when we exclude local residents, focusing solely on individuals arriving from outside the island.

Nonetheless, the estimates might be sensitive to data characteristics. In the standard difference-in-differences analysis, it is possible that the binary treatment indicators fall short of capturing variations in tourist inflows across diverse destinations in Sardinia, including seaside municipalities. To address this variability, we adopted a continuous difference-in-difference approach, replacing the binary treatment indicator with a continuous treatment variable. The results from the robustness check (see Section 6) suggest that extending the time frame under examination could amplify the significance of tourism's influence on the pandemic outbreak by incorporating dynamic and spatial effects. Unfortunately, we cannot assess the significance of these effects in this study due to the absence of data on internal mobility between seaside and non-seaside areas. We emphasize that this represents a critical area

[28] Batalha et al. (2022) conducts a similar robustness check when studying the impact of COVID-19 pandemic on short-time rentals.

Table A.1
Descriptive statistics controls.

	Mean	sd	Min	Max	Count	Definition
Sired data						
Tourists from Sardinia	0.12	0.22	0	1	1850	Percentage of monthly arrivals from Sardinia
Tourists not from EU	0.0036	0.032	0	1	1850	Percentage of monthly arrivals from EU countries
Tourists from the north of Italy	0.088	0.16	0	1	1850	Percentage of monthly arrivals from Northern Italy
Tourists from the center of Italy	0.034	0.070	0	0.79	1850	Percentage of monthly arrivals from Center Italy
Istat data						
Population	4426.9	12 268	99	154 267	1850	Population in 2019
Census data						
Housing crowding index	0.28	0.30	0	2	1850	Percentage ratio of occupied dwellings with less than 40 sq m and more than 4 components or with 40–59 sq m and more than 5 components or with 60–79 sq m and more than 6 components to total occupied dwellings
Unused vs. total buildings ratio	7.54	5.33	0	28.8	1850	Percentage ratio of unused buildings to total buildings
Large families	1.26	0.74	0	4.20	1850	Percentage ratio of the number of households with 6 and more members to total households
Female employment	27.6	4.67	15.4	42.6	1850	Percentage ratio of employed females to total residents females aged 15 years and older
Unemployment	18.9	4.90	6.10	40	1850	Percentage ratio of resident population 15 years and older seeking employment to resident population 15 years and older in employment
Long-range mobility	4.24	2.28	0.60	15.6	1850	Percentage ratio of resident population commuting daily for work or study and taking more than 60 minutes to resident population commuting daily for work or study
Area incidence centers and cores	2.64	4.40	0.20	57.1	1850	Percentage ratio of the area of towns and settlements to the total area (sq. km.)
Illiterates	1.64	0.91	0	6.20	1850	Percentage ratio of illiterate to the resident population aged 6 and older
Adults with a High School or degree	37.0	7.65	19.1	65.9	1850	Percentage ratio of resident population 25–64 years old with high school diploma or college degree to resident population 25–64 years old
High-medium skilled employment	20.7	5.94	7.60	50.2	1850	Percentage ratio of those employed in high-medium skilled occupations (legislators, entrepreneurs, high executives, scientific and highly skilled intellectual occupations; technical Occupations) to the total employed
Annual intercensual variation	−0.35	0.97	−2.80	3.90	1850	Geometric mean of annual intercensual variation
Municipality area in sq.km.	64.9	61.9	2.47	547.0	1850	Municipality area in sq.km.

for future research. It is important to note that seaside municipalities accommodate a substantial number of temporary Sardinian workers during the summer months, potentially contributing to the spread of the pandemic when they return to their permanent residences. This aspect warrants further exploration and calls for additional research to comprehensively investigate its implications.

8. Conclusions

We investigated the relationship between mobility and the COVID-19 pandemic by examining the changes in tourist inflows in Sardinia (Italy) during the summer of 2020. When the Italian government lifted restrictions on mobility and allowed tourists to visit the island, the number of COVID-19 cases in Sardinia sharply increased, suggesting a causal relationship between tourist inflows and the spread of the virus. Our analyses support the existence of this relationship while also providing a quantitative assessment.

This paper makes several contributions to the literature on the relationship between mobility and the spread of COVID-19. First, we provide evidence that changes in mobility facilitated the virus's spread during the pandemic's initial stages. Our analysis reveals a positive and statistically significant influence of tourism inflows on the pandemic's outbreak, further underscoring the pivotal role of mobility for tourism reasons in the spread of COVID-19. Our study corroborates prior findings, such as those presented by Mallapaty (2020), Farzanegan et al. (2021), and Casini and Roccetti (2020), while augmenting the existing body of knowledge demonstrating that tourist inflows can trigger the spread of COVID-19.

Understanding the trade-off between COVID-19 diffusion and tourism is crucial because it has significant implications for public health, the economy, and society as a whole. Governments and policymakers face a difficult balancing act in managing the trade-off. On the one hand, they must protect public health and prevent the spread of COVID-19 by implementing measures such as border closures, quarantine requirements, and social distancing measures. On the other hand, they must support their tourism industries and minimize the economic damage caused by the pandemic.

Table A.2
Panel event study.

	(1) Seaside	(2) Touristic	(3) Continuous
May	0.194	−0.273	−1.624*
	(0.304)	(0.239)	(0.884)
June	0.369	−0.172	−0.249**
	(0.224)	(0.106)	(0.125)
August	8.020**	3.122***	0.992**
	(3.813)	(1.070)	(0.434)
September	6.749*	3.201**	1.083**
	(3.702)	(1.538)	(0.435)
Observations	1850	1850	1850
Municipality fixed effects	yes	yes	yes
Month fixed effects	yes	yes	yes
Controls	yes	yes	yes

Notes: We report the results of the Leads and Lags estimates from the panel event study specification using as the treatment group seaside locations (Column 1), tourist locations (Column 2) and employing the continuous specification with the share of tourists (Column 3). The time span for the study ranges from May 2020 to September 2020. Standard errors in parentheses are clustered at the municipality level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Our study has important implications for policymakers, especially in terms of the likely appearance of a new virus able to trigger a pandemic. Given the lack of vaccines and reliable methods for identifying infected individuals in the period considered by our study, our findings suggest that caution is necessary for allowing tourism flows, which may undo previous containment efforts and lead to the continued spread of the pandemic. Moreover, our estimates of the causal effect of tourism on the spread of COVID-19 allow us to approximately quantify the risk of balancing public health and tourism-related economic activities. The relevance of this implication can be weighted in terms of the finding of Díaz Ramírez et al. (2022). They show that a lower health system capacity together with higher population density led to a higher level of mortality. Given this relationship, the decision to allow or restrict people's movements during the pandemic cannot be disentangled from critical factors such as the severity of the outbreak, the capacity of local healthcare systems, and the availability of effective containment measures. Ultimately, the trade-off between COVID-19 and tourism requires a delicate balancing of public health and economic considerations. Our findings emphasize the importance of carefully weighing the risks and benefits of changes in population density due to people's movement, including modifications such as increased travel or reduced social distancing. This is particularly important in the context of a pandemic, where even small changes can have significant consequences.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A

A.1. Data

Table A.1 provides a detailed explanation for each variable considered in our empirical analysis.[29]

A.2. Panel event study: Estimates

See Table A.2.

Appendix B. Robustness checks

In this Section, we evaluate the robustness of our baseline estimations. The next Section considers a wider time span, while Appendix B.2 exploits an instrumental variable strategy to tackle possible endogeneity concerns.

B.1. Wider time range

Fig. B.1 reproduces the graphs of Fig. 3, extending the analysis to include the months of March, April, October, and November. While a significant increase in COVID-19 cases among seaside locations remains evident in the aftermath of the end of mobility restrictions, it is likely that the number of cases relative to the population became higher among nonseaside locations after September 2020. This graph suggests that the subsequent evolution of COVID-19 cases was likely influenced by both dynamic and spatial effects, as discussed in Section 6.

The following Figure reproduces the graphs of Fig. 4 encompassing a wider time range. Notably, the magnitude of the effect is higher for the two additional months included in the treatment period. However, these estimates are slightly below statistical significance at the canonical level when the treated group includes only seaside municipalities.

Finally, Tables B.1 and B.2 present estimations of Eqs. (1) and (2) with the extended time range. The estimates are both statistically significant and higher in magnitude. This increase in magnitude can be attributed to the likely presence of dynamic and spatial effects, as discussed in Section 6.

B.2. Instrumental variable approach

In this robustness check, we consider instrumental variable specifications to evaluate the robustness of our benchmark findings to possible endogeneity concerns, which we already discussed in Section 6. As long as being a seaside municipality does not have a direct effect on the spread of COVID-19 cases, we can use a dummy variable that takes the value of one for seaside municipalities and zero otherwise as an instrumental variable for tourism inflows. The IV analysis consists of two equations. Eq. (4) represents the first stage, which relates the instrument to the endogenous variable. Eq. (5) is the second stage, utilizing the exogenous part of the endogenous variable to estimate its effect on the variable of interest. For completeness, we also provide Eq. (6) which presents the reduced form, which relates the instrument and the variable of interest. Below we report all the equations:

$$Inflows_{m,t} = \alpha + \beta Seaside + \Lambda X_{m,s,t} + \theta_s + \gamma_t + \epsilon_{m,t}, \tag{4}$$

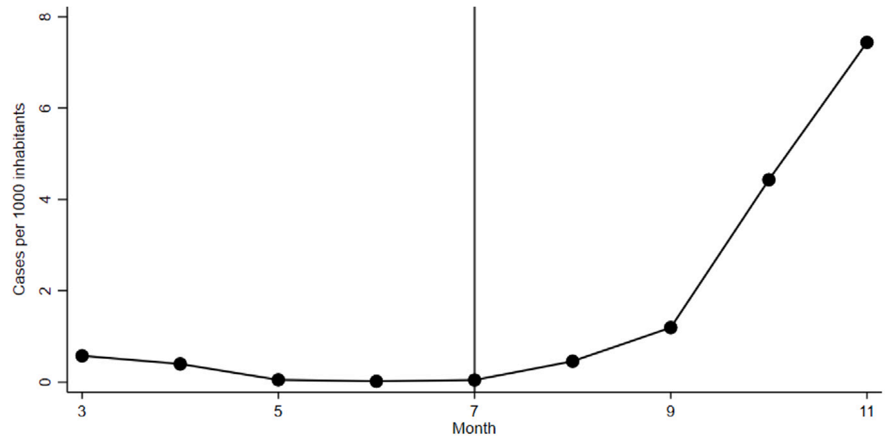
$$Y_{m,t} = \alpha + \beta \widehat{Inflows}_{m,t} + \Lambda X_{m,s,t} + \theta_s + \gamma_t + \epsilon_{m,t}, \tag{5}$$

$$Y_{m,t} = \alpha + \beta Seaside + \Lambda X_{m,s,t} + \theta_s + \gamma_t + \epsilon_{m,t}. \tag{6}$$

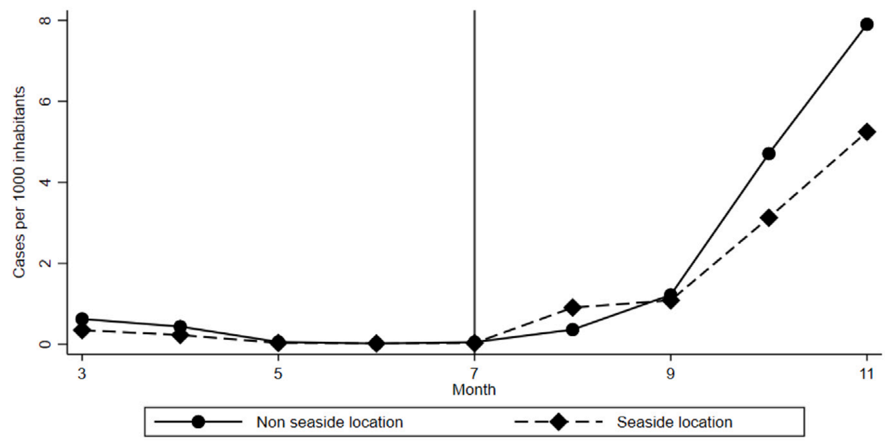
We employ two specifications of the IV model described earlier. In the first specification, we include tourist inflows relative to the month of June 2020. As detailed in Section 2, travel resumed at the beginning of June. However, during this period, the number of tourists was significantly lower compared to the same month in previous years and subsequent summer months.[30] In the second specification, aligning

[29] We include here explanations of variables employed in Appendix B.2 of the appendix.

[30] SIRED recorded only 76,340 arrivals for the month of June, while the number of tourists increased to 331,471 in July and exceeded 500,000 in August. Notably, the majority of June's inflow likely occurred towards the end of the month.



(a) *Cases*



(b) *Cases by location*

Fig. B.1. Cases evolution.
Notes: In panel (a) of the figure we report the evolution of the number of cases per 1000 inhabitants. In panel (b) we report the number of cases per 1000 inhabitants differentiating between nonseaside and seaside locations.

Table B.1
Difference-in-differences — wide time.

	(1) Seaside	(2) Touristic	(3) Seaside	(4) Touristic
Treatment	11.55**	10.55***	12.42**	10.66***
	(5.477)	(2.210)	(5.782)	(2.239)
Observations	3330	3330	3330	3330
Controls	no	no	yes	yes
Municipality fixed effects	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes

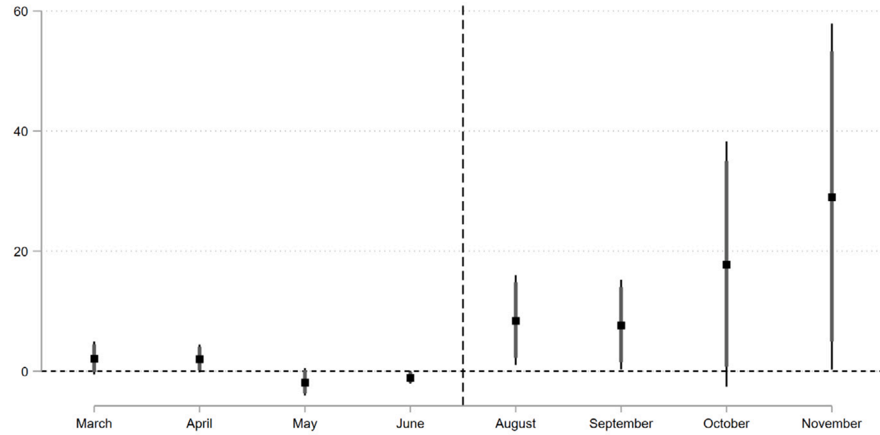
Notes: We report the results using as treatment group the (1 and 3) seaside locations (2 and 4) tourist locations. The time span ranges between March 2020 and November 2020. Standard errors in parentheses are clustered at the municipality level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

with our continuous difference-in-differences approach, we consider only tourist inflows after June 2020. Both specifications include month (γ) and district-fixed effects (θ).[31]

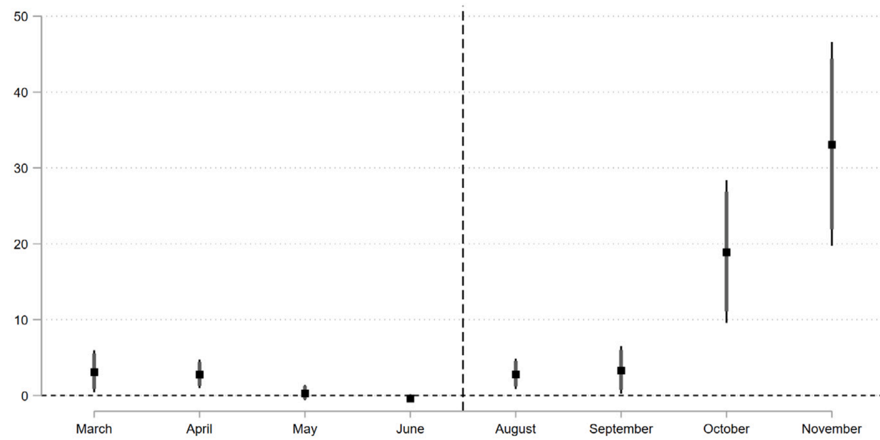
In our IV specification, similar to the continuous difference-in-differences estimations, we produce estimations using as a measure of tourist inflows either data on *Arrivals* or on *Overnight stays*. We also include variables capturing the monthly variation of tourists accounting for their origin, the municipality population size, and other variables capturing municipality heterogeneity regarding housing, demographic, employment and education situation, and mobility for each municipality (X).[32] In Table B.3, we report the estimations when using an

[31] It is worth noticing that the reduced form of the second specification almost coincides with our difference-in-differences specification.

[32] Specifically, we include as controls the percentage of monthly arrivals from Sardinia, from No EU Countries, from North and Centre Italy, and the population and the municipality area in sq. kilometers in 2019; we also include information from the 2011 census data: a housing crowding index, the percentage ratio of unused buildings to total buildings, the incidence of



(a) *Seaside locations*



(b) *Touristic locations*

Fig. B.2. Panel event study estimates.

Table B.2
Difference-in-differences log-lin — wide time.

	(1) Seaside	(2) Touristic	(3) Seaside	(4) Touristic
Treatment	0.338***	0.433***	0.344***	0.442***
	(0.0903)	(0.0580)	(0.0907)	(0.0597)
Observations	3330	3330	3330	3330
Controls	no	no	yes	yes
Municipality fixed effects	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes

Notes: We report the results using as treatment group the (1 and 3) seaside locations (2 and 4) tourist locations. The time span ranges between May 2020 and September 2020. Standard errors in parentheses are clustered at the municipality level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

OLS specification (Columns (1) and (3)) and a Poisson model (Columns (2) and (4)). An instrumental variable estimation must satisfy two requirements: (1) relevance and (2) the exclusion restriction. The first requirement, relevance, means that there should be a strong correlation between the instrument and the variable of interest. This condition is usually checked by reporting the first stage F-statistic, in terms of the instrument, see the third row of Tables B.3 and B.4. We obtain for both cases a value higher than 35 which is well above the rule of thumb of 10 indicating the relevance of the instrument.[33] The second condition requires that the instrument *per se* should not influence the number of COVID-19 cases apart from the mediated effect through touristic inflows. It is difficult to think of another mechanism in place

large families, the female employment rate, the unemployment rate, an index of long-range mobility, the ratio between the area of population centers and cores to the total area, the incidence of illiterates, the incidence of adults with a high school diploma or degree, and he incidence of employment in high-medium-skilled occupations' intercensual variation. See Table A.1 for detailed information.

[33] Following Andrews et al. (2019) we computed the first-stage effective F-statistic, see Olea and Pflueger (2013), and compared them with the critical values. Those additional tests, available upon request, are in support of the relevance condition.

Table B.3
Impact of touristic inflows on COVID-19 cases: IV estimation.

	(1) OLS	(2) POISSON	(3) OLS	(4) POISSON
Arrivals	0.000509* (0.000305)	0.000113*** (0.0000287)		
Overnight stays			0.000105* (0.0000633)	0.0000266*** (0.00000584)
Observations	1850	1850	1850	1850
Adjusted R^2	0.250		0.241	
KP F First stage	34.38		34.25	
District fixed effects	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes
Controls	yes	yes	yes	yes

Notes: Columns 2 and 4 report average marginal effect. The time span ranges between May 2020 and September 2020. Standard errors in parentheses are clustered at the municipality level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4
Impact of touristic inflows on COVID-19 cases: IV estimation — instrument seaside after June.

	(1) OLS	(2) POISSON	(3) OLS	(4) POISSON
Arrivals	0.000741** (0.000341)	0.000110*** (0.0000275)		
Overnight stays			0.000148** (0.0000696)	0.0000260*** (0.00000556)
Observations	1850	1850	1850	1850
Adjusted R^2	0.270		0.253	
KP F First stage	43.26		42.25	
Month fixed effects	yes	yes	yes	yes
Controls	yes	yes	yes	yes

Notes: Columns 2 and 4 report average marginal effect. The time span ranges between May 2020 and September 2020. Standard errors in parentheses are clustered at the municipality level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

apart from tourist inflows leading seaside municipalities to experience a larger growth in COVID-19 cases. Similar results are obtained when we exclude June inflows from the estimations, see Table B.4.

References

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.
Andrews, I., Stock, J.H., Sun, L., 2019. Weak instruments in instrumental variables regression: Theory and practice. Annu. Rev. Econ. 11, 727–753.
Anguera-Torrell, O., Vives-Perez, J., Aznar-Alarcón, J.P., 2021. Urban tourism performance index over the COVID-19 pandemic. Int. J. Tour. Cities.
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.
Ascani, A., Faggian, A., Montresor, S., Palma, A., 2021. Mobility in times of pandemics: Evidence on the spread of COVID-19 in Italy's labour market areas. Struct. Change Econ. Dyn. 58, 444–454.
Batalha, M., Gonçalves, D., Peralta, S., Dos Santos, J.P., 2022. The virus that devastated tourism: The impact of COVID-19 on the housing market. Reg. Sci. Urban Econ. 95, 103774.
Callaway, B., Goodman-Bacon, A., Sant'Anna, P.H.C., 2021. Difference-in-Differences with a Continuous Treatment. Papers 2107.02637, arXiv.org.
Carozzi, F., Provenzano, S., Roth, S., 2022. Urban density and COVID-19: Understanding the US experience. Ann. Reg. Sci. 1–32.
Casini, L., Roccetti, M., 2020. A cross-regional analysis of the COVID-19 spread during the 2020 Italian vacation period: Results from three computational models are compared. Sensors 20 (24), 7319.
Clarke, D., Tapia-Schythe, K., 2021. Implementing the panel event study. Stata J. 21 (4), 853–884.
CRENOS, 2021. Economia della Sardegna. 28° Rapporto 2021. Arkadia Editore srl.
Della Corte, V., Doria, C., Oddo, G., 2021. The Impact of COVID-19 on International Tourism Flows to Italy: Evidence from Mobile Phone Data. Bank of Italy Occasional Paper 647.
Díaz Ramírez, M., Veneri, P., Lembcke, A.C., 2022. Where did it hit harder? Understanding the geography of excess mortality during the COVID-19 pandemic. J. Reg. Sci. 62 (3), 889–908.
Farzanegan, M.R., Gholipour, H.F., Feizi, M., Nunkoo, R., Andargoli, A.E., 2021. International tourism and outbreak of coronavirus (COVID-19): A cross-country analysis. J. Travel Res. 60 (3), 687–692.
Gianmoena, L., Rios, V., 2023. The diffusion of COVID-19 across Italian provinces: A spatial dynamic panel data approach with common factors. Reg. Stud. 1–21.
Glaeser, E.L., Gorback, C., Redding, S.J., 2022. JUE insight: How much does COVID-19 increase with mobility? Evidence from New York and four other US cities. J. Urban Econ. 127, 103292.
González-Val, R., Marcén, M., 2022. Mass gathering events and the spread of infectious diseases: Evidence from the early growth phase of COVID-19. Econ. Hum. Biol. 46 (C), http://dx.doi.org/10.1016/j.ehb.2022.101140.
Han, X., Xu, Y., Fan, L., Huang, Y., Xu, M., Gao, S., 2021. Quantifying COVID-19 importation risk in a dynamic network of domestic cities and international countries. Proc. Natl. Acad. Sci. 118 (31), e2100201118.
Karabulut, G., Bilgin, M.H., Demir, E., Doker, A.C., 2020. How pandemics affect tourism: International evidence. Ann. Tour. Res. 84, 102991.
Kraemer, M.U., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D.M., Group, O.C.D.W., Du Plessis, L., Faria, N.R., Li, R., et al., 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 368 (6490), 493–497.
Krisztin, T., Piribauer, P., Wögerer, M., 2020. The spatial econometrics of the coronavirus pandemic. Lett. Spatial Resour. Sci. 13, 209–218.
Laroze, D., Neumayer, E., Plümper, T., 2021. COVID-19 does not stop at open borders: Spatial contagion among local authority districts during England's first wave. Soc. Sci. Med. 270, 113655.
Mallapaty, S., 2020. What the cruise-ship outbreaks reveal about COVID-19. Nature 580 (7801), 18–19.
Moosa, I.A., Khatatbeh, I.N., 2021. The density paradox: Are densely-populated regions more vulnerable to COVID-19? Int. J. Health Plan. Manag. 36 (5), 1575–1588.
Olea, J.L.M., Pflueger, C., 2013. A robust test for weak instruments. J. Bus. Econom. Statist. 31 (3), 358–369.
Perra, N., 2021. Non-pharmaceutical interventions during the COVID-19 pandemic: A review. Phys. Rep..
Rader, B., Scarpino, S.V., Nande, A., Hill, A.L., Adlam, B., Reiner, R.C., Pigott, D.M., Gutierrez, B., Zarebski, A.E., Shrestha, M., et al., 2020. Crowding and the shape of COVID-19 epidemics. Nat. Med. 26 (12), 1829–1834.
Spelta, A., Pagnottoni, P., 2021. Mobility-based real-time economic monitoring amid the COVID-19 pandemic. Sci. Rep. 11 (1), 13069.
Tantrakarnapa, K., Bhopdhornangkul, B., Nakhaapakorn, K., 2020. Influencing factors of COVID-19 spreading: A case study of Thailand. J. Public Health 1–7.
Vinceti, M., Filippini, T., Rothman, K.J., Ferrari, F., Goffi, A., Maffeis, G., Orsini, N., 2020. Lockdown timing and efficacy in controlling COVID-19 using mobile phone tracking. EClinicalMedicine 25, 100457.
Zhou, Y., Xu, R., Hu, D., Yue, Y., Li, Q., Xia, J., 2020. Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: A modelling study using mobile phone data. Lancet Digit. Health 2 (8), e417–e424.