Received: 9 April 2020 | Revised: 27 October 2020 | Accepted: 30 October 2020

DOI: 10.1111/jors.12510

ORIGINAL ARTICLE

JOURNAL OF REGIONAL SCIENCE WILEY

The geography of COVID-19 and the structure of local economies: The case of Italy

Andrea Ascani | Alessandra Faggian | Sandro Montresor

Social Sciences, Gran Sasso Science Institute, L'Aquila, Italy

Correspondence
Andrea Ascani, Social Sciences, Gran Sasso Science Institute, Viale Francesco Crispi 7, 67100 L'Aquila, Italy.
Email: andrea.ascani@gssi.it

Abstract

The aim of this article is to analyze the subnational spread of COVID-19 in Italy using an economic geography perspective. The striking spatial unevenness of COVID-19 suggests that the infection has hit economic core locations harder, and this raises questions about whether, and how, the subnational geography of the disease is connected to the economic base of localities. We provide some first evidence consistent with the possibility that the local specialization in geographically concentrated economic activities acts as a vehicle of disease transmission. This could generate a core-periphery pattern in the spatiality of COVID-19, which might follow the lines of the local economic landscape and the tradability of its outputs.

KEYWORDS

COVID-19, geographical concentration, local economic structure, tradability

JEL CLASSIFICATION

R10; R11; R12

1 | INTRODUCTION

In December 2019, the Chinese city of Wuhan is hit by a rapidly growing number of pneumonia cases caused by an unknown coronavirus. The outbreak swiftly extends to other Asian countries in the following weeks, including Japan, South Korea, and Iran, raising public health concerns at the international level. By the end of February 2020, the contagion explodes in Europe, with a rapid surge of infections in Italy. On March 11th, when the global number of confirmed infections amount to 118,319 in 113 countries, the Director General of the World Health Organization (WHO) officially declares the disease the first "pandemic" caused by a coronavirus (WHO, 2020a). The fast-moving diffusion of the disease reaches a global number of 7,823,289 ascertained cases on June 15th (WHO, 2020b), confirming the very high global risk assessment given by most national and international public authorities. The new virus is initially, on January 12th 2020, called "2019 novel Coronavirus (2019-nCov)" by WHO, and successively, on February 11th, renamed "Severe Acute Respiratory Syndrome Coronavirus 2" (SARS-CoV-2) by the International Committee on Taxonomy of Viruses. On the same day, the WHO officially names the disease caused by SARS-CoV-2 as "Coronavirus Disease 2019" (COVID-19).

From an economic geography perspective, it is an intriguing fact that the subnational territorial spread of COVID-19 is very uneven. For instance, in the case of Italy, which is the first and hardest hit European country, followed by Spain at the beginning of April 2020, the great bulk of COVID-19 cases is tremendously concentrated in a specific area of the country. By March 4th 2020, the day of the first nation-wide containment measures taken by the Italian government, the top ten infected provinces (NUTS-3) account for about 80% of all confirmed cases. These locations occupy an area in Northern Italy, which spans across the regions of Lombardy, Veneto, and Emilia-Romagna, the well-known economic core of the country where most of its high value-added economic activities are concentrated. The combined GDP of these top ten provinces in 2017 was in fact about one-fifth of the national total GDP and the per capita GDP of eight of them was well above the national average. These are striking spatial features that raise questions about whether, and how, the subnational geography of COVID-19 cases is connected to the local economic structure. The issue is of paramount importance for at least two reasons. First, identifying a link between the local share of COVID-19 infections and the characteristics of the local economy may help policy makers in the definition of containment measures that will possibly have to be targeted now that (at the time of this writing) the nation-wide lockdown has been lifted. Second, in the presence of a relationship between COVID-19 cases and local economic structures, future academic investigations of the impact of COVID-19 on the economy should take into consideration that the territorial sectoral specialization may play a notable role.

The present work intends to investigate the relationship between the geography of COVID-19 and the structure of local economies, by focusing on the case of Italian provinces. Specifically, based on an economic geography-inspired conceptual argument, we hypothesize that locations endowed with economic activities characterized by high geographical concentration might be subject to relatively higher infections due to the intense physical interaction of people and commodities entailed by the agglomeration advantages of these industries. A long-standing literature, in fact, suggests that the spatial clustering of sectors is connected to the existence of increasing returns internal and external to firms, and that this gives rise to traded or untraded agglomeration benefits potentially involving frequent face-to-face interactions (e.g., Iammarino & McCann, 2006; Kaldor, 1970; Krugman, 1991a; McCann, 2013; Storper & Venables, 2004). Furthermore, spatially concentrated industries also tend to serve markets that overcome the local administrative boundaries, thus generating intense trading relationships across geographical space, with other locations within a country or even at the global level (Boschma & Iammarino, 2009; Cainelli et al., 2014; Kemeny & Storper, 2015).

From the empirical standpoint, we cannot offer an investigation of the causal effects related to these ideas, nor we intend to produce an epidemiology model. Rather, we can provide an analysis of the correlation between the local share of countrywide COVID-19 confirmed cases and a measure of provincial engagement in geographically concentrated sectors, which we call "provincial economic base." This is an indicator of Italian NUTS3 regions' (employment) involvement in industries with different geographical concentration. We work out by this geographical concentration by looking at countrywide data for all the 5-digit economic activities covered by the "Census of Industry and Services" run every ten years by the Italian National Institute of Statistics (ISTAT)—with respect to two waves: that of 2011 and, in a robustness check of our results, that of 1991. In searching for the relationship at stake, we also control for other potential factors influencing the number of local infections, such as past COVID-19 cases, population density, demographic indicators, as well as regional dummies (NUTS-2). In line with our ideas, our results suggest that the local infection of COVID-19 is positively correlated with the 2011 local economic specialization in spatially concentrated industries, ceteris paribus, and that the result holds true also when, at the cost of losing a substantial number of observations, the same specialization is worked out for 1991. Specifically, the association at stake is driven by the provincial specialization in clustered manufacturing sectors, rather than services or other activities. As we will argue, this can be due to the specific occupations, tasks, and job contexts characterizing manufacturing activities rather than services. For instance, manufacturing workers are subject to long shifts and continuative interactions at close distance with a relatively large number of coworkers, as compared to many other occupations (Caselli et al., 2020; Dingel & Neiman, 2020). Furthermore, it is also plausible that trading manufacturing output requires more human interaction than trading services, as the former consist in most cases of tangible goods that need to be physically shipped. When we examine the role of provincial trade linkages, we also find a correlation between manufacturing exports and COVID-19 infections, in line with the possibility that tradable local economic bases might be a vehicle of disease transmission.

This article is structured as follows. Section 2 provides a conceptual background to the notion that persistent spatial processes of economic clustering may be channeling COVID-19 infections toward core locations. Section 3 offers a contextual description of the emergence of COVID-19 cases in Italy and the institutional measures that have been adopted by the national government: from the lockdown decrees (issued by the President of the Council of Ministers, DPCM) of March 2020 to the decrees attenuating/relaxing the lockdown in June 2020. Section 4 presents the data used in the empirical analysis while our methodological approach is explained in Section 5. Results are presented and discussed in Section 6. Finally, we offer our concluding remarks in Section 7.

2 | CONCEPTUAL FRAMEWORK

The conceptual framework on which the present paper is based revolves around the idea that the emergence and the unfolding of the uneven geography showed by COVID-19 might be associated with the unevenness of the core-periphery pattern that results from spatial economic forces.

A plethora of scholarly contributions have documented the striking and persistent geographical concentration of economic activity within specific areas (e.g., Brakman et al., 2019; Duranton & Overman, 2005; Ellison & Glaeser, 1997; Puga, 2010). In terms of market structure, the geographical concentration of economic activity emerges in the presence of firm-level increasing returns to scale and imperfect competition (Krugman, 1991b). In fact, increasing returns encourage companies to concentrate their activities due to the cost advantages deriving from creating larger plants. Concentrating production in space, nonetheless, requires firms to face the costs of shipping goods over large geographical distance to serve other locations. Hence, to minimize transport costs, companies locate where demand is larger and/or where providers of inputs are located (Venables, 1996). This, in turn, determines a path-dependent and circular process of interconnected colocation behavior, as input producers and demand are influenced by the location of the companies operating under increasing returns.

Industry-level sources of agglomeration benefits constitute another substantial mechanism of self-reinforcing geographical clustering (Arrow, 1962; Marshall, 1890; Romer, 1987). The advantages of spatially concentrated industries, including labor market pooling, the local availability of specialized inputs and services as well as potential knowledge spillovers, tie together sector activities in space (Brakman et al., 2017; Faggio et al., 2017; Iammarino & McCann, 2006). Moreover, with the market expansion of an industry due to such agglomeration benefits, the sector division of labor increases with the emergence of new specialized companies, the development of new competences, and the additional differentiation of activities on a local basis (Young, 1928).

In this framework, learning and effectively communicating knowledge is key for firms to sustain their innovative efforts and foster their competitiveness. Along these lines a large body of work suggests that new knowledge creation and technological change have a relevant relational component, thus heavily involving the spatial clustering of specialized knowledge producers and workers (Glaeser, 1999; Hanson, 2000). One fundamental reason for this concentration stands in the essential role of face-to-face contacts to diminish coordination costs, increase trust between partners, transmit not easily codified knowledge, and reduce moral hazard and other issues related to information asymmetries (Storper & Venables, 2004). In this sense, spatially bound locations become the *loci* for a set of untraded connections facilitating interfirm interactions in terms of knowledge-related common practices and spillovers (Iammarino & McCann, 2006; Lundvall, 1992; Storper, 1995).

The geographical concentration of certain economic activities, due to the reason mentioned above, generates clear implications in terms of the spatiality of the markets that will be served by agglomerated industries (Krugman, 1991a;

Jensen & Kletzer, 2006). As a matter of fact, these markets tend to transcend the regional dimension due to the nonubiquity of producers (Gervais & Jensen, 2019; Kemeny & Storper, 2015). In this sense, clustered industries constitute the tradable sector of a specific location, intended as the aggregated portion of local economic activities that supply output to other locations, at the national or even the global level. As a result, local industries with external demand linkages play a substantial role in local economic development processes, as their output growth is driven by larger markets than the regional one (Kaldor, 1970).

Taken together, the ideas here outlined suggest that some economic behaviors and processes have an intimate spatial nature, generating path-dependent patterns of geographical concentration of economic activity at the local level. This may generate dense business and human interactions that follow patterns in line with the specific economic structure of a location. In economic geography, regions specialized in activities exhibiting a notable geographical concentration are celebrated for their agglomeration advantages and for the profitable linkages with extra-regional markets, due to the tradability of their output. In the context of the present work, however, this may also imply that locations whose internal economic base is characterized by concentrated activities might be more conducive of COVID-19 transmission. Indeed, a correlation between the former and the latter appears plausible if, as recent virology studies have proved, the transmission of COVID-19 is spurred by dense localized interactions and associated with extra local linkages linked to the tradability of local output. To the empirical testing of such a correlation is dedicated the analysis of the following sections.

2.1 | The emergence of COVID-19 cases in Italy and the institutional response

On January 31th 2020, the Italian health authorities confirmed the first two cases of COVID-19 infection in Rome, after two Chinese tourists, originally from Wuhan, tested positive for SARS-CoV-2.[1] They previously entered Italy from Malpensa Airport in Milan on January 23rd and subsequently moved to Verona and Parma, before reaching Rome on January 28.[2] The number of infected people grows to three on February 6th, when an Italian, repatriated a week earlier from Wuhan and since then quarantined in Rome, also tests positive.[3] The outbreak in the North of Italy becomes evident only some weeks later, on February 16th when a 38-year-old Italian patient, with flu-like symptoms, reports respiratory issues at the Codogno hospital, in the province of Lodi (Lombardy). It is possible that this infection is connected to an asymptomatic contact in the German city of Munich, happened around January 20th, which can also be the first European transmission of COVID-19 (Rothe et al., 2020). Therefore, the Italian patient evidently remained asymptomatic for almost a month, by potentially spreading the infection through various social interactions. For example, the first doctor who treated the 38-year-old patient tested positive on February 21st. By the end of February, the number of cases in the whole of Italy reaches 1,000 confirmed COVID-19 infections, most of them in the Northern provinces of Lodi (Lombardy), Piacenza (Emilia-Romagna), Cremona (Lombardy), Bergamo (Lombardy), and Padova (Veneto), suddenly throwing the country into a health crisis. Nonetheless, in the first weeks of diffusion of the disease, the public and policy attitude toward the declaration of emergency remained mixed, for a number of reasons: partly in the attempt to reassure the economy that the crisis can be controlled; partly because of an initial cognitive bias of decision makers in targeting the growing emergency,[4] similar to the experiences of other Western countries in the subsequent weeks; and partially because of the novelty that the infection represents in the European context in February 2020.

[1]Cerqua and Di Stefano (2020) suggest that the virus began spreading more systematically already in the last week of January 2020.

[2]https://www.corriere.it/cronache/20_gennaio_31/virus-primi-due-casi-italia-due-cinesi-marito-moglie-italia-dieci-giorni-e365df1c-43b3-11ea-bdc8-faf1f56f19b7.shtml

[3]https://edition.cnn.com/asia/live-news/coronavirus-outbreak-02-06-20-intl-hnk/index.html

[4]https://hbr.org/2020/03/lessons-from-italys-response-to-coronavirus?fbclid=IwAR3bne1xKvxeFrk5d-34ZtbmsFq3cmzAAKJmuCYp2uCDUORinP1FFknrl4M#comment-section

Emergency measures taken by the Italian government start with the *Decreto del Presidente del Consiglio dei Ministri* (DPCM) of February 23rd. This is an executive act of the Prime Minister implementing containment measures in the most infected areas in the North of the country, namely in the regions of Lombardy and Veneto, fundamentally imposing a quarantine to the most affected municipalities. The first country-wide measures are instead included in a government decree of March 1st, which extends the length of the existing measures and provides a national framework for the containment of the disease. In consideration of the exponential growth of infections in the first days of March, a new DPCM is adopted on March 4th with more restrictive measures. This act includes, among other points, the national suspension of all public shows, social and sport events where the minimum social distancing norm of 1.5 m cannot be assured, the national closure of schools and universities until March 15th and a number of other restrictions in the whole country. Faced with the growing spread of the disease in the first days of March and a death toll above a hundred, the government announces a new DPCM during the night between March 7th and 8th, including unprecedented lockdown measures in the region of Lombardy and in 14 widely affected provinces in the regions of Piedmont (five), Emilia-Romagna (five), Veneto (three), and Marche (one), for a total of about 16 million inhabitants. These measures aim at avoiding the mobility and interaction of people inside and across these areas and include, for instance, the suspension of social gatherings, cultural, and religious activities, limitations to restaurants, just to name a few. Containment measures in the rest of the country also become more severe. The announcement of this new DPCM creates panic among the population triggering, in the night between March 7th and 8th, an attempt of people from the lockdown areas to escape to their origin locations in the Centre and South of the country before the new containment measures become effective on March 8th.[5] In front of a growing number of COVID-19 infections, exceeding tens of thousands cases on March 10th, the Prime Minister signs another DPCM on March 11th, substantially extending the lockdown to the whole country, known as *#Io resto a casa* (#I stay home). This includes strong limitations to peoples' mobility and the suspension of all retail and commercial activities with the exception of grocery stores, pharmacies, and a limited number of others for basic necessities and services, with the objectives of severely limiting social contact. On March 22nd a new DPCM blocks all nonessential and nonstrategic economic activities, and introduces new tightening measures, such as the prohibition to people to move from the municipality of residence to another.

After a month of severe lockdown, considering the slow and gradual improvement of numbers related to deaths and new infection, with the DPCM of April 10th, effective from April 14th, some retail activities are finally allowed to open. With the DPCM of April 26th, the government then launches the so-called Phase 2, starting on May 4th, which allows the movement of people within regions and the reopening of a large set of economic activities. This fundamentally represents a loosening of the lockdown measures and a gradual restart of social and economic life. Subsequently, with the DPCM of May 16th, effective on May 18th, the Italian government allows the opening of a larger set of activities such as retail and restaurants, as well as the celebration of religious events and the access to museums. In the same DPCM, the movement of workers between regions is allowed starting on June 3rd.

The previously delineated context of institutional measures has been accompanied by a number of COVID-19 infections that, not only has changed over time, but has also shown a quite irregular spread across the Italian provinces. To the number of these infections and to their geography we now turn in the next section.

3 | DATA

As of June 15th 2020, Italy ranks fifth in the world for the total number of tests performed for the detection of COVID-19, with 4,620,718 confirmed tests. Consistently, Figure 1 shows the top-10 countries by COVID-19 tests at this date, thus indicating that the Italian case is worth of investigation as the detection of the infection lies on a good number of tests.

[5]https://www.ilmessaggero.it/italia/coronavirus_milano_fuga_milano_treno_romani-5097472.html

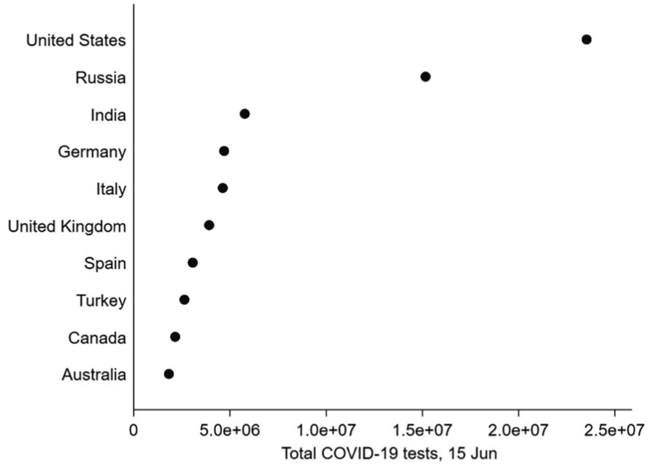

FIGURE 1 Top-10 countries by total COVID-19 tests performed as of June 15, 2020

3.1 | COVID-19 provincial data

We collected data from the Italian Ministry of Health regarding the daily number of confirmed COVID-19 infections in the 107 Italian provinces (NUTS-3) starting with February 25th 2020. This is a rather precise spatial level of data aggregation, allowing notable variation at a fine geographical scale even within regions. Every day the Ministry of Health releases these data in a daily report. A tiny minority of infections cannot be associated to provinces in these reports and thus we do not consider these cases, thus obtaining a (slight) under-representation of the outbreak.[6] Figure 2 shows the provincial share of COVID-19 confirmed cases at different points in time. It is evident from map (a) that the outbreak starts from a group of Northern provinces, including Lodi and Cremona in the Lombardy region, and rapidly spreads to other areas. Map (b) refers to the outbreak as of March 4th, which is the day of the DPCM that extends for the first time to the whole country a number of measures to contain the diffusion of the disease. Map (c) indicates the geography of COVID-19 after 12 days from the implementation of the DPCM of March 4th, suggesting that more provinces experience higher shares of infections by this date. The latter may also capture some return migration from the North to the Centre-South of Italy in the initial days of the outbreak, thus potentially spreading the infection to other areas. Map (d) reports the data for 26 April, which is 12 days after the DPCM of April 10th, effective from April 14th, allowing the reopening of some economic activities. Map (e) refers to May 16th, that is 12 days after the so-called Phase 2 is launched. Finally, map (e) illustrates the COVID-19 geography as of June 15th, that is 12 days after the movement of people between regions is allowed. In general, COVID-19 infections are quite strongly and persistently concentrated in the Northern provinces of the country during the whole lockdown regime period and also in the subsequent Phase 2 period.

3.2 | Provincial economic base

In constructing our focal explicative variable, we conceive provinces as spatial economic systems consisting of an industry mix whose performance is dependent on the local portion of geographically concentrated activities. As we said, such a portion signals, first of all, the advantages of internal and external increasing returns to scale. Furthermore, the

[6]Ongoing research also suggests that the number of unidentified infected individuals can be about four times larger than the confirmed number of cases (Pedersen & Meneghini, 2020). This would lead to an even larger underestimation of the incidence of COVID-19 cases in our data.

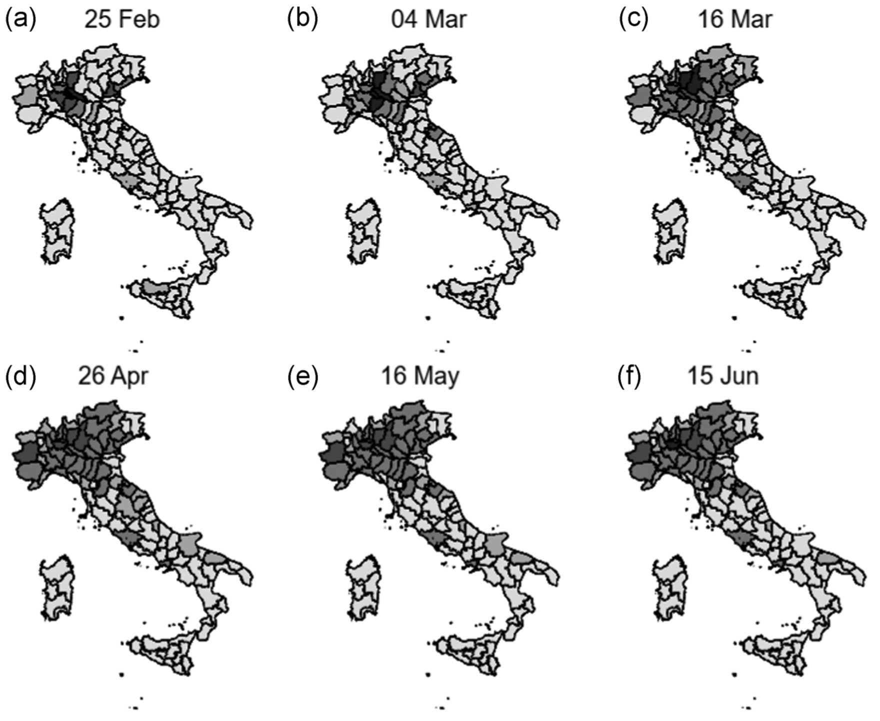

FIGURE 2 Provincial share of COVID-19 infections

spatial concentration of industries is also connected to the geography of the demand for the output produced by each provincial industry mix. In the latter sense, our idea of provincial economies echoes the export-based approach to regional economic analysis (e.g., Kaldor, 1970), whereby regional industries with external demand linkages play a substantial role in local economic development processes. Rather than classifying economic sectors in arbitrary categories to reflect their dependence on agglomeration dynamics or external demand conditions, we construct our indicator of the provincial economic base in two steps. First, we use a simple measure of industry concentration at the national level, to identify the patterns of spatial clustering revealed by each and every industry. Second, for each and every industry, we multiply this spatial concentration measure by the provincial employment level by sector and we take the provincial median of the resulting number. In so doing, we obtain an indicator of the relevance that more or less spatially concentrated activities have in each provincial economy in terms of occupation. The resulting economic base of the focal province, i, can thus be written as follows:

$$\text{Economicbase}_i = M_i(\text{HHI}_k \times E_{ik}), \tag{1}$$

where HHI is the Herfindahl–Hirschman Index of a five-digit industry k, based on employment; E is the employment of industry k in province i; M indicates the median of the resulting measure within each province i. Given the central role of this indicator in our analysis, the two steps that lead to its construction deserves a more detailed illustration.

Step 1: As we said, we first construct an HHI of spatial concentration by using employment data on five-digit industries from the 2011 Census of Industry and Services undertaken by ISTAT: a census-year that covers the same number and topography of provinces affected by the COVID-19. By considering the whole economy, with the exception of primary activities, interesting differences emerge across industries. For example, Figure 3 shows the different employment distribution of sectors with relatively high and low HHI. Maps (a) and (b) refer to manufacturing sectors Ateco 20140 "Manufacture of other organic basic chemicals," which is a spatially

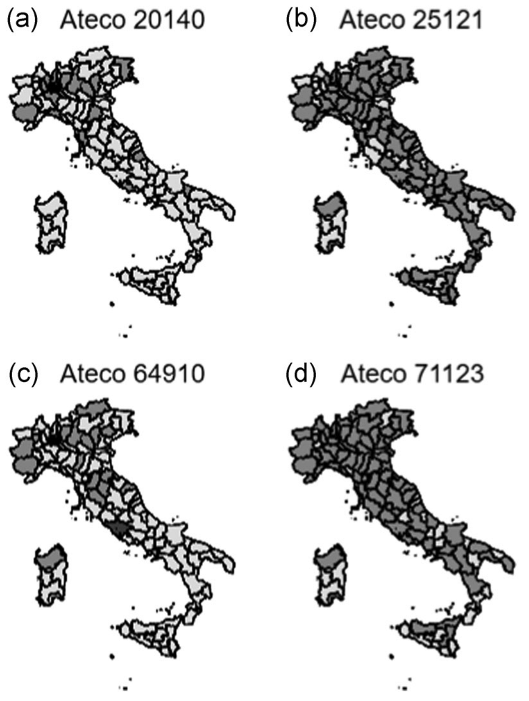

FIGURE 3 Spatial concentration of employment in different industries

concentrated activity, and Ateco 25121 "Manufacture of doors and windows of metal," which is instead evenly distributed across provinces.[7] Interestingly, the export value of Ateco 25121 in 2018 is only 8% of that of Ateco 25140 according to ISTAT data, thus suggesting that the most concentrated industry serves a spatially larger market than the least concentrated one. A similar picture can be drawn for service sectors, as exemplified by the spatial differences in the concentration of employment between Ateco 64910 "Financial leasing" and Ateco 71123 "Engineering activities and related technical consultancy."

Step 2: Second, we consider the employment level of provincial economies in each 5-digit sector to assign the national industry HHI measure to provinces. Following the idea that industry size matters to define local specialization profiles (Kemeny & Storper, 2015), we weigh industry HHIs by absolute provincial employment figures. Hence, we obtain a measure of provincial economic structure accounting for the relevance of the geographical concentration and the tradability of the local industry mix. To alleviate the influence of outlier industries within individual provinces, we consider the provincial median of our measure. In fact, there might be monopolistic industries or highly centralized state-owned service activities, such as Ateco 30400 "Manufacture of military fighting vehicles" or Ateco 53100 "Postal activities under universal service obligation": these are concentrated in one province only, which can influence the general economic base of specific provinces. Figure 4 shows our measure of provincial economic base plotted in map (b) and the geography of COVID-19 infection in map (a), both grouping provinces by eight quantiles. It is evident that the spatial distributions of the two variables have common traits based on this descriptive evidence, with the North of the country being characterized by both an economic base specialized in less spatially ubiquitous activities (i.e., more tradable) and a higher share of infections of COVID-19. While this is a mere visual association, it is suggestive of the potential existence of a link

[7]Ateco codes refer to the Italian industrial classification of 2007 and they are directly comparable to NACE codes.

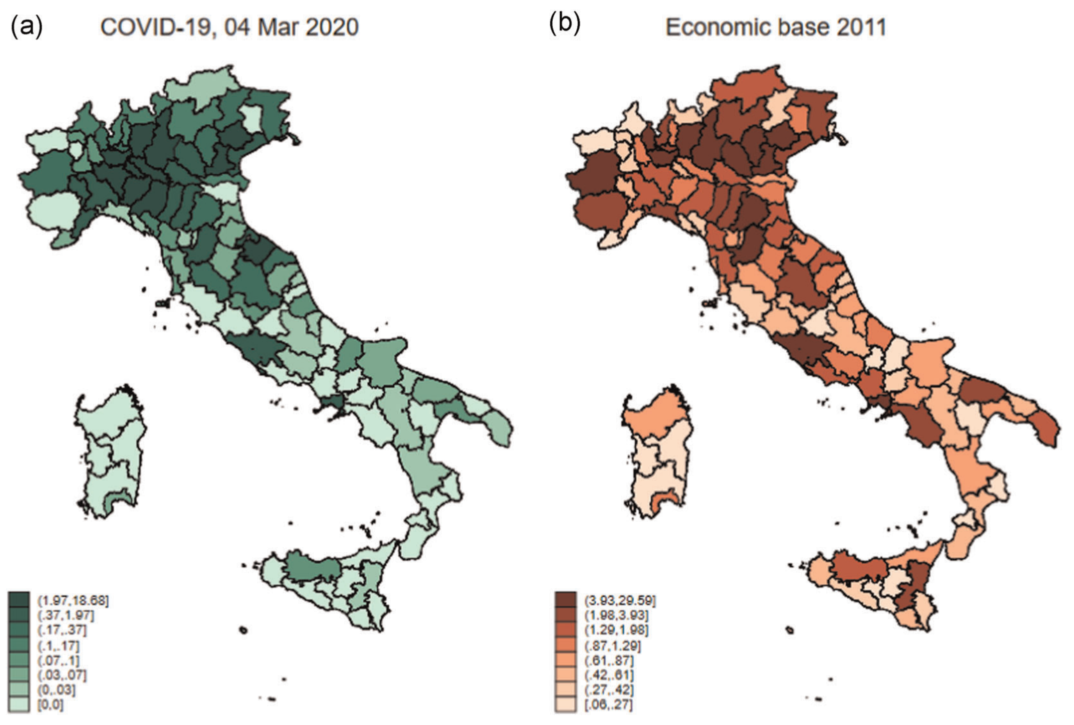

FIGURE 4 Geography of COVID-19 and provincial economic base [Color figure can be viewed at wileyonlinelibrary.com]

between the nature of local economic activities and the diffusion of the COVID-19 infections. Similarly, Table 1 shows the top ten provinces by COVID-19 cases as of March 4th 2020 (i.e., when the DPCM with the first national containment measure is adopted), together with their positioning in the distribution of our measure of provincial economic base on aggregate and also disaggregated for manufacturing and services, respectively. Again, this descriptive output suggests that a connection between the local economic base and the presence of COVID-19

TABLE 1 Top-10 provinces by COVID-19 cases as of March 4 and their economic base positioning

Province	COVID-19 cases	Economic base rank	Percentile	Manufacturing base rank	Percentile	Service base rank	Percentile
Lodi	559	80	26th	71	34th	87	19th
Bergamo	423	6	94th	5	95th	10	91th
Cremona	333	43	53th	41	62th	67	38th
Piacenza	319	51	60th	49	55th	56	48th
Padova	162	7	94th	8	93th	9	92th
Milan	145	1	99th	1	99th	1	99th
Brescia	127	5	95th	3	97th	7	94th
Pavia	126	30	72th	42	61th	36	67th
Parma	115	27	75th	22	80th	29	73th
Treviso	86	12	89th	6	94th	15	86th

cases might exist, as 6 of the 10 most infected provinces fall in the topmost distribution quartile for the aggregate and manufacturing indicators, and five for the service indicator. In an extension of the empirical analysis, we check the robustness of the economic base indicator described above by assessing the sensitivity of our estimates to alternative measures and data. First, we consider a traditional measure of specialization such as the Krugman Index. Differently from the economic base measure described above, this is a relative measure of specialization whereby the economic structure of Italian provinces is compared to the national average. Second, we reconstruct our economic base variable with data from the Census of 1991, generating an indicator with a deep time lag with respect to the outbreak of COVID-19, which however refers to a smaller number of provinces.

3.3 | Other variables

In exploring the relationship between COVID-19 cases and provincial economic base, we account for a number of other concurrent elements that can, in principle, provide alternative explanations for the geography of the outbreak across Italy. We collect these data from ISTAT for the most recent time-period possible.

First, we construct a measure of provincial population density in 2019 to capture the role of dense urbanized areas. Population density can certainly be a driver of rapid contagion, as the frequency of human social interactions is plausibly larger within more crowded spatial units (Tarwater & Martin, 2001). Social contact plays indeed a strong role in the transmission of diseases, as suggested by epidemiology studies, especially in the case of respiratory infectious agents (Wallinga et al., 2006; Meyer & Held, 2017). Second, we consider the provincial number of deaths in 2018 to account for the fact that a more rapid spread of COVID-19 may occur in provinces already afflicted by higher mortality. This could be the case, for instance, if the disease is more easily transmitted among more fragile people. Third, we take into consideration provincial health migration by collecting data for 2016 on the number of days spent in other regions' hospitals by the residents of each province. This can proxy for the capacity of local health systems to address large scale outbreaks, such as COVID-19. Next, we explicitly account for the age structure of provinces in 2018, defined as the local share of residents older than 64. While all age groups can be infected with COVID-19, some initial evidence on the diffusion of virus suggests that the aging populations can be more at risk of infections (WHO, 2020a; Zunyou and McGoogan, 2020). Furthermore, we include in our analysis the provincial proportion of male population in 2018, as males seem to be more susceptible of infections according to a number of recent scientific findings (Sun et al., 2020; Wang et al., 2020).

Besides these demographic aspects, we also account for other provincial elements that can facilitate the diffusion of the infection. These include an indicator of tourism by province, measured as the number of days in 2018 with a presence of tourists in local hotels and other touristic structures. To capture the touristic presence during the season of diffusion of COVID-19, we account for non-summer months only, as provided by ISTAT. Accounting for the touristic vocation of local economies is important in the setting of the present analysis as tourism flows can increasingly represent vectors of disease transmission (Sönmez et al., 2019). Then, we take into consideration the dynamism of the local labor market by including the provincial unemployment rate in 2018. More dynamic labor markets, in fact, can be characterized by more frequent business interactions, thus representing a potential risky environment in terms of disease contagion. Not surprisingly, one of the tightest measures taken by the Italian government with DPCM of 22 March to contain the transmission of COVID-19 is to freeze all economic activities labeled as non-essential or nonstrategic. Next, we also consider the local attraction force towards migrants, by incorporating in the analysis the provincial share of foreign residents in 2019. In fact, peoples' mobility can represent a transmission channel of infections, such that territories receiving larger portions of migrants can be more at risk of disease spread (Apostolopoulos & Sönmez, 2007; Backer et al., 2020; Herbinger et al., 2016). Finally, connected to the previous point, we also consider the presence of airports in a province, as this type of transport infrastructure can connect each territory to a national or international network of linkages through which the mobility of people is reinforced.

TABLE 2 Variables description

Variable	Measure	Year	Geography	Source
COVID-19 cases	Number of COVID-19 cases on national total	2020	Province	Ministry of Health
Economic base	Employment-weighted Herfindahl–Hirschman Index	2011	Province	ISTAT
Population density	Population divided by provincial area (sq. km)	2019	Province	ISTAT
Deaths	Log number of deaths	2018	Province	ISTAT
Health emigration	Number of days spent by residents in other regions' hospitals	2016	Province	ISTAT
Old population	Population above 64 divided by total population	2018	Province	ISTAT
Male population	Male population divided by total population	2018	Province	ISTAT
Tourism rate	Number of days with touristic presence in hotels and other touristic structures	2018	Province	ISTAT
Unemployment rate	Percentage of unemployed	2018	Province	ISTAT
Foreign residents	Number of foreign residents on total population	2019	Province	ISTAT
Airport	Dummy equal to 1 if the province has an airport	2019	Province	ISTAT

Table 2 summarizes the variables discussed in this section and Appendix A reports their summary statistics as well as their correlation matrix (Tables A1 and A2).

4 | METHODOLOGY

Our analysis of the relationship between the geography of COVID-19 cases and the local economic base of Italian provinces lies on the notion that infections can propagate within and also between spatial units (Meyer & Held, 2017; Paez et al., 2020). Furthermore, we also refer to recent medical evidence regarding the incubation period of COVID-19 and the timeline of the emergence of symptoms. By analyzing 181 confirmed cases of COVID-19, Lauer et al. (2020) find that the median incubation period is 5.1 days and that 97.5% of infected sample cases develop symptoms within 11.5 days. Other studies also provide similar estimates (e.g. Backer et al., 2020; Lai et al., 2020). Taken together, these elements provide a methodological guide in the analysis of the relationship under investigation, as explained below. Hence, we estimate the following spatial autoregressive model with generalized spatial two-stage least square methods (Kelejian & Prucha, 1998):

$$y_{i,d} = \beta y_{i,d-t} + \lambda \boldsymbol{W} y_{j,d-t} + \vartheta x_i + Z\gamma_i + \rho + \varepsilon_{i,} \quad (2)$$

where the dependent variable y is the share of country-wide COVID-19 confirmed cases in province i on day d,[8] $\boldsymbol{W}$ is a spatial weighting matrix based on inverse distance between province i and all other provinces j, x is our measure of provincial economic base, Z is a vector of control variables, and ε is the error term. Importantly, with the inclusion of regional dummies ρ we are able to control for factors operating at the level of regions (NUTS-2) that can affect the spread of COVID-19 infection. For instance, specific regional emergency measures taken by

[8]This provides an indicator of the relative relevance of the infection for each province within the national framework of the Covid-19 crisis. In a robustness check, available upon request, we use the infection rate (based on 2019 population) and the results remain stable.

local authorities can be controlled with this term. In the case of Italy, in fact, the decentralization of powers gives mandate to regional governments, rather than provinces, for many matters and for the provision of services such as health and transportation. To account for the fact that new infections can emerge where previous infections have taken place, we enter the lagged number of COVID-19 cases in the right-hand side of the equation, where t denotes a time lag in terms of days. In line with the scientific evidence reported above, we start by considering a 12-day lag to cover the largest possible span of infection. This choice is also driven by the fact that our measure of COVID-19 does not only require that a person is infected, but also that the infection is detected and reported by the health authorities. Nevertheless, we subsequently present estimates by also reducing the time lag of one-third (i.e., 4 days), to avoid that some of the measures contained in the various and frequent DPCMs described in a previous section influence the number of COVID-19 cases. The time lags will also enter the spatial lag, as the incubation period and emergence of the symptoms are plausibly the same regardless of whether the infection occurs within or across specific provincial areas.

5 | RESULTS

5.1 | Baseline regressions

We start our empirical analysis by estimating the model presented above and gradually including the various regressors. Table 3 presents this first set of results. In Column 1, we show the most parsimonious version of our model, where we include the time and spatial lags of our dependent variable along with the economic base indicator and the regional dummies. What emerges from the results is, first and foremost, that new COVID-19 infections tend to concentrate in provinces where prior cases are recorded. The coefficient on the dependent variable lagged of 12 days (i.e., February 25th) is indeed positive and statistically significant. This is not surprising and in line with both the visual evidence provided in Figure 2 and with the more general notion of infection occurring through interactions, which are clearly highly localized (Meyer & Held, 2017). Second, we cannot detect a significant relationship at the geographical level between the share of infections as of March 8th (i.e., our dependent variable) and its spatial lag fixed at February 25th. This seems to suggest that, on average, infections across provinces in Italy as a whole are not statistically significant. However, this does not necessarily mean that provinces are self-contained spatial units when it comes to the disease transmission. In fact, there might be specific vehicles and networks that may facilitate the diffusion of COVID-19 across locations, as we specify later, but our spatial lag based on inverse distance seems not to capture any of these channels. Third, and most important for the sake of our analysis, we detect a significant positive relationship between COVID-19 cases and the local economic base of provinces, ceteris paribus. As we did expect according to our conceptual background, we find evidence of the fact that provincial economies with larger employment in geographically concentrated industries in 2011 actually exhibit a larger share of COVID-19 cases as of March 8th, 2020. As we said, we refrain from attaching any causal meaning to this association, as our data and methodological setting do not permit this type of analytical inference. Nonetheless, we consider that shedding light on this potential link deserves attention in future research.

In the remaining specifications of Table 3, we enter the other covariates discussed above, which can offer other explanations of the geographical unevenness of the COVID-19 infection.[9] Surprisingly, we cannot find any stable and significant effect across columns for these other variables. One exception is represented by the presence of an airport within a province (Column 10), which seems to be facilitating the transmission of COVID-19. In a sense, this compensates for the nonsignificant role of the spatially lagged dependent variable, as airport connections clearly indicate the spatial nature of COVID-19 transmissions across provinces. Population density

[9]The number of observations changes from 107 to 106, as the data for the new Province of South Sardinia are missing for many variables.

TABLE 3 GS2SLS estimates of COVID-19 cases in Italian provinces as of March 8, 2020

Dep. Var: COVID-19 cases, Mar 8	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
COVID-19 cases, Feb 25	0.306***	0.301***	0.302***	0.304***	0.300***	0.303***	0.301***	0.301***	0.295***	0.294***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)	(0.033)
Economic base	0.134***	0.191***	0.175***	0.177***	0.175***	0.165***	0.169***	0.165***	0.152**	0.145**
	(0.039)	(0.047)	(0.062)	(0.062)	(0.062)	(0.061)	(0.062)	(0.063)	(0.064)	(0.062)
Density		−0.001**	−0.001**	−0.001**	−0.001**	−0.001*	−0.001*	−0.001	−0.001	−0.001*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Deaths			0.128	0.220	0.115	0.314	0.283	0.285	0.337	−0.066
			(0.327)	(0.417)	(0.428)	(0.435)	(0.441)	(0.441)	(0.442)	(0.461)
Health emigration				−0.000	−0.000	0.000	0.000	0.000	−0.000	0.000
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Old population					−12.827	0.775	0.296	0.433	2.979	4.447
					(10.386)	(12.579)	(12.620)	(12.627)	(12.833)	(12.504)
Male population						87.507*	82.490*	83.440*	68.464	77.896
						(47.165)	(48.624)	(48.745)	(50.856)	(49.643)
Tourism rate (no summer)							−0.029	−0.029	−0.025	−0.071
							(0.070)	(0.070)	(0.069)	(0.070)
Unemployment rate								−0.015	−0.009	−0.013
								(0.058)	(0.058)	(0.056)
Foreign residents									8.578	5.042
									(8.718)	(8.608)
Airport										0.837**
										(0.344)

(Continues)

TABLE 3 (Continued)

Dep. Var: COVID-19 cases, Mar 8	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Spat. COVID-19 cases, Feb 25	−0.474	−0.522	−0.501	−0.497	−0.284	−0.298	−0.311	−0.306	−0.365	−0.255
	(0.504)	(0.495)	(0.497)	(0.497)	(0.523)	(0.515)	(0.515)	(0.516)	(0.517)	(0.505)
Obs.	107	107	107	107	106	106	106	106	106	106
Pseudo R^2	0.68	0.69	0.69	0.69	0.70	0.71	0.71	0.71	0.71	0.73
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***Significance level: $p < .01$.

**Significance level: $p < .05$.

*Significance level: $p < .1$.

exhibits an unexpected negative sign, albeit weakly significant or irrelevant. The share of male population, also, loses its significance using these richer specifications. Importantly for the purposes of this study, the coefficient on the local economic base remains positive and significant at 1% or 5% level, thus supporting our focal argument. In terms of magnitude, the coefficient is also rather stable and appreciable. In Column 10, an increase of one unit in the provincial economic base is associated, on average and keeping all else fixed, with an increase of 0.145% points in COVID-19 infections. Put differently, a one standard deviation increase in the economic base indicator is associated with an increase of 0.53% in the provincial share of COVID-19 cases as of March 8th. An increase of 0.53%, based on the number of infections as of March 8th, amounts to about 36 new cases from the start of the COVID-19 crisis on 25th February.

The results we have obtained and illustrated above remain largely unchanged when, conscious of their possible collinearity with our focal regressor, *Economic base*, we plug in the analysis additional variables to control for the economic and demographic size of the sample provinces. Population size, in particular, could be correlated with both the local share of infections and the level of provincial employment, thus affecting the reliability of our setting. For instance, in absence of geographical concentration of industries on a national scale, the economic base indicator would only capture the employment size of a province, which is correlated with its population size and could, in turn, explain the local shares of infections. While we are aware of this potential limitation, we further discuss in Appendix B the pros and cons of the inclusion of these additional covariates in the model. Briefly, Table B1 confirms the significance and size of the effect of the provincial economic base on COVID-19 infection shares once we include the economic and demographic characteristics of provinces.

5.2 | Alternative estimations

Table 4 presents a number of alternative estimations of our econometric model. First, we propose the same model of Column 10 of Table 3, with the exception that the dependent variable now captures the share of COVID-19 cases as of March 4th, rather than March 8th. That is, we shorten the time lag from twelve days, which is the time span identified as sufficient for symptoms to appear in 97.5% of patients (e.g. Lauer et al., 2020), to 8 days. The reason for this choice is that the DPCM of March 4th introduces the first set of restrictive containment measures at the national level, as explained above, including social distancing and the suspension of schools, universities, public shows, sport events, and other serious limitations. Therefore, picking the infections as of March 4th should exclude the effect of these nation-wide policies. At the same time, the risk is that we miss some COVID-19 cases, also considering that the inclusion of infections in our data also depends on the detection and the reporting by public health authorities. The results are reported in Column 1, where we find that the correlation between the provincial economic base and the share of COVID-19 infections as of March 4th is positive and significant at the 10% level, controlling for the full set of covariates. The local male population is also positive and weakly significant, corroborating the scientific evidence that this portion of the population is more susceptible of infections (Sun et al., 2020; Wang et al., 2020). Again, the presence of airport infrastructures in a location is positively correlated with the presence of COVID-19 cases, while the spatially lagged dependent variable remains insignificant. In line with previous results, instead, the provincial infections as of February 25th are strongly and directly correlated with the local cases 8 days later. While the significance of the coefficient of interest decreases in this model, we extend our analysis by considering an 8-day lag from now on, rather than 12 days. By allowing a shorter time lag, in fact, we may be providing a safer estimate in consideration of the fact that data before February 25th are unavailable and some infections could have taken place earlier. In Column 2 of Table 4, we include an economic base indicator based on the number of firms, rather than employment, to explicitly account for the geographical concentration of businesses. The estimated coefficient is not statistically significant, thus suggesting that a firm-based measure is probably not appropriate to capture a phenomenon that spread among people (i.e., workers). In Columns 3 and 4, we replicate the employment-based and firm-based economic base indicators of Columns 1 and 2, by assigning the

TABLE 4 COVID-19 as of March 4, 2020 by different measures of economic base

Dep. Var: COVID-19 cases, March 4	(1)	(2)	(3)	(4)
COVID-19 cases, Feb 25	0.464***	0.464***	0.461***	0.462***
	(0.033)	(0.033)	(0.033)	(0.033)
Economic base (employment)	0.104*			
	(0.062)			
Economic base (firms)		1.569		
		(0.965)		
Economic base (LQ-employment)			43.894	
			(44.763)	
Economic base (LQ-firms)				28.683
				(79.915)
Density	−0.001	−0.001	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Deaths	−0.326	−0.419	−0.373	−0.163
	(0.458)	(0.488)	(0.563)	(0.609)
Health emigration	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Old population	7.535	9.425	10.089	10.120
	(12.422)	(12.380)	(12.495)	(12.675)
Male population	89.351*	91.732*	95.710*	94.198*
	(49.318)	(49.287)	(49.650)	(49.851)
Tourism rate (no summer)	−0.083	−0.083	−0.053	−0.058
	(0.070)	(0.070)	(0.071)	(0.073)
Unemployment rate	−0.018	−0.016	−0.043	−0.035
	(0.056)	(0.056)	(0.056)	(0.056)
Foreign residents	5.237	5.897	8.566	8.467
	(8.552)	(8.494)	(8.463)	(8.542)
Airport	0.742**	0.730**	0.761**	0.770**
	(0.342)	(0.343)	(0.344)	(0.346)
Spatial COVID-19 cases, Feb 25	−0.322	−0.296	−0.175	−0.243
	(0.502)	(0.501)	(0.510)	(0.507)
Obs.	106	106	106	106
Pseudo R^2	0.81	0.81	0.81	0.81
Region dummies	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***Significance level: $p < .01$.

**Significance level: $p < .05$.

*Significance level: $p < .01$.

industry HHI to provinces according to their (median) location quotients (LQs) as a measure of sector specialization. Similar to before, we cannot detect any statistically significant correlation between the local economic base and the presence of COVID-19 cases. One reason for this, as compared to the estimate of Column 1, can again be linked to the nature of the phenomenon under analysis. In fact, while the LQ provides a measure of the extent to which a local economy is specialized in an industry as compared to a national benchmark, the absolute employment figure used in the construction of the economic base of Column 1 offers an indication of the size of industries, in terms of workers. The latter is plausibly what matters when the object of analysis is a phenomenon affecting people directly.

5.3 | Distinguishing manufacturing from services

This section proposes the analysis of the association between the provincial economic base and the local COVID-19 infections by accounting for specific segments of economic activity. Specifically, we separately consider the different macroactivities that form our aggregate measure of economic base at the provincial level. An important motivation for this type of investigation is that local economic profiles with similar aggregate economic base indicators can hide profound differences if their industry mix is more or less dominated by one activity or another. One key distinction, in this sense, is for instance related to local economies that are dominated by services rather than manufacturing industries. Although the colocation of different types of activities is obviously an important characteristic of modern economies (e.g., Castellani et al., 2016), we are interested in disentangling potential different relationships associated with different activities and with their geographical concentration. Table 1 shows that the top provinces by COVID-19 cases are characterized by the high economic base scores on aggregate, and by slightly higher scores for the manufacturing sector as compared to services, although these differences seem marginal in the case of these top ten provinces. Nonetheless, the results in Column 1 of Table 5, suggest that the correlation between the provincial economic base and the share of COVID-19 cases is driven by the engagement of provinces in spatially clustered manufacturing activities, rather than services. This can also partially explain the reason why certain large metropolitan areas, such as Rome and Naples, have a relatively low share of infections, considering that their economy is proportionally more service-oriented. Therefore, it is possible that the agglomeration effects and the dynamics described above regarding the geographical concentration of activities are especially relevant for the case of manufacturing, where face-to-face contacts and other forms of interaction might be more frequent than in service activities. While the contact with consumers is certainly more limited in manufacturing jobs and occupations, as compared to services, in the latter category it remains a relatively low-intensity contact in terms of number of people involved and time length. While services are more prone to home-working or usually occur in offices with a lower number of workers than in a plant, most manufacturing activities require long hours with many coworkers that jointly perform complementary tasks in close vicinity. Put it differently, it is possible that manufacturing tasks and occupations may produce situations where individuals are exposed to a larger number of coworkers and for a longer and continuative time. Both these conditions seem to be rather favorable for the spread of the disease.[10,11] Indeed, initial anecdotal evidence regarding COVID-19 hotspots in the US seem also to point in this direction.[12] Therefore, it is possible that manufacturing firms occupy, on average, a larger number of employees than service firms, thus generating a relatively cramped environment where social interactions and complementarities between different tasks are also more pervasive. Finally, it is

[10]https://www.corriere.it/salute/malattie_infettive/20_maggio_16/coronavirus-dove-ci-si-ammala-piu-ecco-come-riconoscere-rischi-ed-evitarli-3762a12c-96b6-11ea-a66c-1f6181297d24.shtml?fbclid=IwAR1mSVbT_8WhEHkVLwJ76STeDj7S9xe96jc3OSB82eSlf6BUZ5d6-BwunXY

[11]https://www.repubblica.it/economia/2020/06/12/news/coronavirus_ritardi_lockdown_sindacati_bergamo_zona_rossa-259024115/

[12]https://www.theguardian.com/world/2020/may/15/us-coronavirus-meat-packing-plants-food

TABLE 5 COVID-19 cases by economic activity

Dep. Var: COVID-19 cases, March 4	(1)	(2)	(3)	(4)	(5)	(6)	(7)
COVID-19 cases, Feb 25	0.475***	0.472***	0.463***	0.472***	0.474***	0.476***	0.475***
	(0.0330)	(0.0329)	(0.0329)	(0.0329)	(0.0332)	(0.0330)	(0.0330)
Economic base							
Manufacturing	0.0006**	0.0005*		0.0005**	0.0006*	0.0007**	0.0006**
	(0.0003)	(0.0003)		(0.0002)	(0.0003)	(0.0003)	(0.0003)
Services	−0.0000						
	(0.0000)						
Energy		−0.0002					
		(0.0002)					
Water, sewage, waste		0.0014					
		(0.0059)					
Wholesale, retail, repair			0.0003				
			(0.0002)				
Transport and storage				−0.0000			
				(0.0001)			
Hotel and restaurant					−0.0004		
					(0.0008)		
Professional services						−0.0001	
						(0.0001)	
Other services							−0.0006
							(0.0008)
Spat. COVID-19 cases, Feb 25	−0.3110	−0.3165	−0.3205	−0.3277	−0.3158	−0.2943	−0.3232
	(0.4981)	(0.4964)	(0.5044)	(0.4980)	(0.4990)	(0.4974)	(0.4975)
Obs.	106	106	106	106	106	106	106
Pseudo R^2	0.81	0.81	0.81	0.81	0.81	0.81	0.81
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***Significance level: $p < .01$.

**Significance level: $p < .05$.

*Significance level: $p < .1$.

also plausible that trading manufacturing output requires more human interaction than trading services, as the former consist in most cases of tangible goods that need to be physically shipped. In the context of the present work, these remain mere hypothetical explanations of the prevalence of manufacturing in the relationship under investigation. Disentangling these specific channels is well beyond the scope of our analysis. In Column 2, for the sake of completeness, we add other activities to the regression model, namely the production and provision of energy as well as the activities related to water, sewage, and waste. Considering the high correlation between these and services (see Table A3 in Appendix A), Column 2 excludes the latter from the specification. In this case, the significance of the manufacturing economic base decreases to the 10% level. In Columns 3–7, we unpack services by different types of activities, as these can be highly heterogeneous. We consider different types of service activities separately from the other as the correlation between them is very high. In Column 3, we consider services grouped into wholesale, retail, and repair, by excluding the manufacturing indicator due to the high correlation between the two variables. In the next columns, we enter transport and storage services, hotel and restaurants, professional activities and, finally, other services. We cannot detect any significant coefficient on any of the subcategories of services, while the manufacturing economic base of provinces remains significantly correlated with local COVID-19 cases, ceteris paribus.

5.4 | Alternative measure of the provincial economic structure

As anticipated above, in this section, we consider the Krugman Index as an alternative way to detect the relevance of national industries at the province level, on which our provincial economic base is built up. As is well known, this index provides a relative measure of specialization of provincial economies with respect to the national average. Lower values of the Krugman Index indicate that a provincial economy is more similar to the national average economic structure, while higher values indicate that it is more dissimilar, thus implying specialization. Table 6 presents the results of this robustness check. In Column 1 of Table 6, we replicate the specification of Column 10 of Table 3. The Krugman Index, calculated on the basis of employment data, exhibits a positive coefficient significant at the 10% level. In line with our baseline model, this suggests that an association exists between COVID-19 cases and local specialization profiles, although its statistical significance decreases as compared to the specification with the economic base indicator. In Column 2 of Table 6, instead, we consider the Krugman Index based on the number of firms in each province, rather than employment. Similar to the previous result, we cannot detect any association with COVID-19 infections, again suggesting that the number of businesses is certainly less important that the size of industry employment for the disease to spread. In Column 3, finally, we account for the sectoral heterogeneity of provincial specialization by individually considering an employment-based Krugman Index for manufacturing and service activities. Similar to previous results, the positive association in our data is driven by the local specialization in manufacturing activities and this is strongly significant.

Overall, this sensitivity check confirms the results obtained with our focal indicator of provincial economic base and suggests that such an indicator could be a reliable measure of the structure of local economies. As an additional check, instead of using the Krugman Index as an alternative of our economic base indicator, in Appendix B, we have illustrated an estimate that keeps the same focal indicator and use the Krugman index to control for the relative specialization of provinces. Table B2 in the same Appendix B shows that results are robust to this further check.

5.5 | Early lockdown measures in March 2020

We further explore the relationship between the local economic base and the provincial share of COVID-19 infections by considering the timing of the different lockdown measure implemented with the various DPCMs discussed above at the beginning of the crisis in March 2020. The time lag adopted in this analysis is 8 days, similar to the results presented in Table 4, rather than the 12 days of Table 3. This should provide a more conservative measure of COVID-19 infections because a portion of the affected population does not show symptoms within 8 days, as previously explained. A lag of 8 days is also motivated by the frequency of policy measures in early March 2020, as a strategy to avoid potential strong overlapping in the influence of different DPCMs on the number of COVID-19 cases. However, we also run the same specifications by adopting the longest time lag possible, that is using infection data as of February 25th.

Table 7 presents the results for a number of alternative regression specifications. In Columns 1–4, we consider the DPCM of March 4th, which extends for the first time to the whole of Italy the initial containment measures to address the disease transmission. Therefore, the dependent variable in these models refers to the COVID-19 cases as of March 12th. Columns 5–8 are based on the DPCM of the night of March 7th, which initiates a severe national

TABLE 6 Alternative measure of specialization

Dep. Var: COVID-19 cases, March 8	(1)	(2)	(3)
COVID-19 cases, Feb 25	0.463*** (0.033)	0.464*** (0.033)	0.475*** (0.033)
Krugman Index (employment)	2.721* (1.483)		
Krugman Index (firms)		0.009 (0.008)	
Krugman Index (manufacturing)			0.046** (0.023)
Krugman Index (services)			−0.005 (0.007)
Density	−0.000 (0.000)	−0.000 (0.000)	−0.001 (0.000)
Deaths	−0.005 (0.421)	−0.172 (0.446)	−0.742 (0.538)
Health emigration	0.000 (0.000)	0.000 (0.000)	0.000* (0.000)
Old population	9.423 (12.533)	9.149 (12.466)	9.777 (12.281)
Male population	94.670* (49.864)	95.659* (49.595)	70.333 (50.643)
Tourism rate (no summer)	−0.066 (0.070)	−0.078 (0.070)	−0.040 (0.072)
Unemployment rate	−0.033 (0.056)	−0.018 (0.057)	−0.026 (0.057)
Foreign residents	8.108 (8.489)	6.233 (8.615)	6.011 (8.490)
Airport	0.769** (0.346)	0.751** (0.344)	0.793** (0.340)
Spat. COVID-19 cases, Feb 25	−0.252 (0.507)	−0.301 (0.506)	−0.311 (0.498)
Obs.	106	106	106
Pseudo R^2	0.806	0.808	0.814
Region dummies	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***Significance level: $p < .01$.

**Significance level: $p < .05$.

*Significance level: $p < .1$.

lockdown and generates a sudden return migration from the North to the Centre-South of the country. In this set of regressions, hence, the dependent variable refers to COVID-19 cases as of March 15th. Finally, we account for the lockdown tightening contained in the DPCM of March 11th, known as *#Io resto a casa* (#I stay home), which imposes limitations to peoples' mobility and the suspension of all retail and commercial activities with the exception of grocery stores, pharmacies, and a small number of others for basic necessities and services. For each of these three sets of regressions, we consider both the aggregate provincial economic base and the disaggregation in manufacturing and services.

The regression coefficients are in line with the previous findings and are rather stable across different specifications, regardless of the timing of the lockdown and whether the time lag is 8 days or longer (i.e., 16, 19, or 23 days in the specifications based on infections as of February 25th). The positive correlation between the provincial economic base and the share of provincial cases of COVID-19 is persistent and its magnitude slightly increases over time. In particular, the coefficients on the aggregate economic base indicator in Columns 1 and 3 imply that a one standard deviation increase in this variable is related to an increase in the provincial share of infection equal to 0.80% and 0.60%, respectively. Based on the number of cases on March 12th and on the days of the time lags in each regression (February 25th for Column 1 and March 4th for Column 3), these percentages imply an increase in COVID-19 cases of about 117 from February 25th and 70 in the 8 days from March 4th.[13] Similarly, the coefficients for March 15th in Columns 5 and 7 imply an increase of infections of 0.71% and 0.36% for a standard deviation increase in the provincial economic base. These correspond, respectively, to 172 new cases from February 25th and 67 new cases in the 8 days between March 7th and March 15th. Finally, the coefficients for March 19th (Columns 9 and 11) suggest an increase of 0.91% and 0.40%, corresponding to about 360 and 108 new cases from February 25th and March 11th, respectively. Similar to previous results, this significant association is exclusively driven by the local specialization in concentrated manufacturing sectors, ceteris paribus.

5.6 | Later lockdown regime and the launch of Phase 2

While we detect a stable and positive association between the local economic base and the provincial cases of COVID-19 at the beginning of the crisis, this evidence remains so far limited to the very early days of the infection outbreak. If our hypothesis that local economies characterized by geographically concentrated industries can be more conducive of infection is plausible, we should then observe the same positive relationship persisting during and after the lockdown regime of the Italian society and economy. Therefore, we run a new set of regressions where we consider the provincial infections at three additional key dates, namely: April 3rd, April 26th, May 16th, and June 15th. These dates are justified by the DPCM measures taken 12 days before.[14] As summarized in Section 3, the DPCM of March 22nd introduces new tightening measures, including the closure of all economic activities that are not classified as strategic or essential, and the prohibition to people to move from the municipality of residence to another. With DPCM of April 10th, effective from April 14th, some retail activities are allowed to open, while with the DPCM of April 26th the government launches the so-called Phase 2, starting on May 4th, which fundamentally represents a loosening of the lockdown measures and a gradual restart of social and economic life. Finally, on June 3rd, the movement of people across NUTS-2 regions is again allowed.

Table 8 presents the results for a set of regressions using the economic base measure defined above (Columns 1–4) and set of regressions with the Krugman Index of specialization (Columns 5–8). Both indicators exhibit positive and statistically significant coefficients for all the key dates considered. This implies that the association between the local economic structure and the provincial share of COVID-19 does not only characterize the early

[13]These numbers are calculated as differences in cumulative cases between the contemporaneous day and the day of the lag.

[14]We consider a 12-day lag in these regressions as there is no overlapping between the time span of these DPCMs.

TABLE 7 COVID-19 cases by timing of lockdowns

Dep. Var:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	COVID-19 Cases, Mar 12 (8 days after DPCM of Mar 4; first national measures)				COVID-19 Cases, Mar 15 (8 days after DPCM of Mar 7; national lockdown and escape from the North)				COVID-19 Cases, 19 Mar (8 days after DPCM of 11 Mar; lockdown tightening)			
Economic base	0.2211***		0.1678***		0.1973***		0.1010***		0.2521***		0.1095***	
	(0.0668)		(0.0445)		(0.0632)		(0.0360)		(0.0544)		(0.0174)	
Econ. base, manufacturing		0.0010***		0.0008***		0.0009***		0.0006***		0.0011***		0.0005***
		(0.0003)		(0.0002)		(0.0003)		(0.0002)		(0.0003)		(0.0001)
Econ. base, services		−0.0000		−0.0000		−0.0000		−0.0000		−0.0000		0.0000
		(0.0000)		(0.0000)		(0.0000)		(0.0000)		(0.0000)		(0.0000)
COVID-19 cases, Feb 25	0.1630***	0.1816***			0.1073***	0.1245***			0.0583**	0.0785***		
	(0.0354)	(0.0356)			(0.0335)	(0.0336)			(0.0288)	(0.0287)		
Sp. COVID-19 cases, Feb 25	−0.2704	−0.2591			−0.2193	−0.2116			−0.1326	−0.1270		
	(0.5424)	(0.5367)			(0.5128)	(0.5065)			(0.4413)	(0.4325)		
COVID-19 cases, Mar 4			0.5516***	0.5654***								
			(0.0408)	(0.0395)								
Sp. COVID-19 cases, Mar 4			0.2162	0.1275								
			(0.5432)	(0.5286)								

TABLE 7 (Continued)

Dep. Var:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	COVID-19 Cases, Mar 12 (8 days after DPCM of Mar 4; first national measures)				COVID-19 Cases, Mar 15 (8 days after DPCM of Mar 7; national lockdown and escape from the North)				COVID-19 Cases, 19 Mar (8 days after DPCM of 11 Mar; lockdown tightening)			
COVID-19 cases, Mar 7							0.4266***	0.4299***				
							(0.0263)	(0.0252)				
Sp. COVID-19 cases, Mar 7							0.1905	0.1251				
							(0.2615)	(0.2536)				
COVID-19 cases, Mar 11											0.6834***	0.6809***
											(0.0212)	(0.0194)
Sp. COVID-19 cases, Mar 11											0.1359	0.0628
											(0.1853)	(0.1710)
Obs.	106	106	106	106	106	106	106	106	106	106	106	106
Pseudo R^2	0.66	0.66	0.85	0.86	0.63	0.64	0.88	0.89	0.67	0.68	0.97	0.97
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***Significance level: $p < .01$.

**Significance level: $p < .05$.

TABLE 8 COVID-19 cases during lockdown and Phase 2

Dep. Var: COVID-19 Cases	(1) Apr 3	(2) Apr 26	(3) May 16	(4) Jun 15	(5) Apr 3	(6) Apr 26	(7) May 16	(8) Jun 15
COVID-19 cases, Feb 25	0.017	0.012	0.010	0.008	0.014	0.009	0.007	0.005
	(0.018)	(0.013)	(0.013)	(0.014)	(0.023)	(0.020)	(0.021)	(0.021)
Economic base	0.275***	0.295***	0.315***	0.323***				
	(0.034)	(0.025)	(0.025)	(0.026)				
Krugman Index					2.020**	1.883**	1.909**	1.918**
					(1.023)	(0.907)	(0.938)	(0.959)
Spat. COVID-19 cases, Feb 25	0.019	0.072	0.071	0.070	0.203	0.270	0.283	0.286
	(0.276)	(0.206)	(0.204)	(0.208)	(0.349)	(0.310)	(0.320)	0.328)
Obs.	106	106	106	106	106	106	106	106
Pseudo R^2	0.80	0.87	0.88	0.88	0.67	0.71	0.71	0.71
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***Significance level: $p < .01$.

**Significance level: $p < .05$.

days of the outbreak, but it shows persistency throughout the whole lockdown regime and also during the period of lightening of the containment measures (i.e., Phase 2).

5.7 | Provincial economic base with deep lags

To reinforce our empirical setting, we reconstruct our economic base variable with data from the Census of 1991, thus generating an indicator with a deep time lag (i.e., 29 years) with respect to the outbreak of COVID-19. The structure of local economies in 1991 can potentially be less endogenous than the 2011 and, hence, this check can provide additional validity to our previous results. One caveat is that the number of provinces in 1991 is lower than in 2011 and this implies the loss of some observations.

We consider five specific key dates, ranging from the early days of the outbreak through the lockdown regime until Phase 2, thus covering the whole available time dimension to the current date and offering a wide robustness check in terms of the timing of the infection. We also consider the aggregate economic base and its manufacturing and service components, similar to Table 5. Results are presented in Table 9. Overall, the main outcomes emerging from the previous sections hold also in this exercise. The variable for the aggregate economic base of provinces is characterized by positive and significant coefficients (5%–10% level) across the retained specifications. Furthermore, in line with our prior results, the positive relationship is substantially driven by the local specialization in concentrated manufacturing activities, rather than in services. We interpret this additional piece of evidence as supportive of the idea that the geography of COVID-19 is connected to the economic geography of the country in terms of the structure of the local economic systems.

TABLE 9 Deep lags, Census 1991

Dep. Var: COVID-19 Cases	(1) Mar 4	(2) Mar 4	(3) Apr 3	(4) Apr 3	(5) Apr 26	(6) Apr 26	(7) May 16	(8) May 16	(9) Jun 15	(10) Jun 15
COVID-19 cases, Feb 25	0.693***	0.684***	0.109**	0.101**	0.075*	0.068*	0.067*	0.059	0.063	0.056
	(0.063)	(0.063)	(0.045)	(0.044)	(0.039)	(0.038)	(0.040)	(0.039)	(0.041)	(0.040)
Economic base 1991	0.838**		0.243*		0.222*		0.227*		0.230*	
	(0.336)		(0.140)		(0.121)		(0.124)		(0.127)	
Economic Base 1991, manufacturing		0.852***		0.411*		0.410**		0.429**		0.437**
		(0.298)		(0.230)		(0.198)		(0.203)		(0.208)
Economic Base 1991, services		−0.000		0.019		0.014		0.014		0.015
		(0.054)		(0.024)		(0.020)		(0.021)		(0.021)
Density	−0.000	−0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Deaths	0.958*	0.791*	0.862***	0.911***	1.037***	1.082***	1.105***	1.149***	1.126***	1.169***
	(0.491)	(0.452)	(0.321)	(0.309)	(0.277)	(0.265)	(0.285)	(0.273)	(0.291)	(0.279)
Health emigration	0.000**	0.000**	−0.000	−0.000	−0.000	−0.000	−0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Old population	6.678	5.284	1.620	1.148	6.898	6.226	8.815	8.175	9.155	8.535
	(13.302)	(13.285)	(8.855)	(8.909)	(7.632)	(7.656)	(7.862)	(7.877)	(8.024)	(8.038)
Male population	133.7***	142.6***	69.57*	71.49**	54.56*	56.02*	53.68*	55.27*	54.583*	56.256*
	(51.604)	(51.776)	(35.990)	(35.902)	(31.019)	(30.852)	(31.951)	(31.742)	(32.612)	(32.392)
Tourism rate (no summer)	−0.072	−0.077	−0.070	−0.082	−0.062	−0.076	−0.065	−0.079	−0.069	−0.083*
	(0.074)	(0.073)	(0.052)	(0.055)	(0.044)	(0.047)	(0.046)	(0.048)	(0.047)	(0.049)
Unemployment rate	−0.000	−0.022	0.005	−0.002	0.008	−0.000	0.006	−0.002	0.005	−0.003
	(0.061)	(0.062)	(0.044)	(0.043)	(0.038)	(0.037)	(0.039)	(0.038)	(0.039)	(0.039)

(Continues)

TABLE 9 (Continued)

Dep. Var: COVID-19 Cases	(1) Mar 4	(2) Mar 4	(3) Apr 3	(4) Apr 3	(5) Apr 26	(6) Apr 26	(7) May 16	(8) May 16	(9) Jun 15	(10) Jun 15
Foreign residents	−2.361	0.075	10.937	11.343	11.314*	11.946*	11.722*	12.332**	11.763*	12.354*
	(9.847)	(9.900)	(6.967)	(7.101)	(6.005)	(6.102)	(6.185)	(6.278)	(6.313)	(6.407)
Airport	1.298***	1.067***	0.482*	0.419*	0.331	0.273	0.310	0.253	0.329	0.272
	(0.387)	(0.381)	(0.261)	(0.250)	(0.225)	(0.215)	(0.232)	(0.221)	(0.236)	(0.226)
Spatial COVID-19 cases, Feb 25	2.169	1.553	3.259***	2.920***	3.518***	3.170***	3.809***	3.446***	3.917***	3.547***
	(1.462)	(1.473)	(1.019)	(1.036)	(0.878)	(0.890)	(0.904)	(0.916)	(0.923)	(0.935)
Obs.	93	93	93	93	93	93	93	93	93	93
Pseudo R^2	0.76	0.77	0.74	0.74	0.79	0.79	0.79	0.79	0.79	0.79
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***Significance level: $p < .01$.

**Significance level: $p < .05$.

*Significance level: $p < .1$.

5.8 | Exporting and importing activities

Finally, we investigate whether provinces specialized in tradable activities still experience higher shares of COVID-19 infections, by using provincial international trade data for 2018 taken from ISTAT. As discussed above, economic sectors that are spatially concentrated tend to serve markets that transcend the local dimension. This also suggests that increasing returns to scale, internal, and external to firms, and tradability are interconnected elements (e.g., Kemeny & Storper, 2015; Krugman, 1991a). Table 10 presents a set of regressions where we examine the role of the provincial shares of exports and imports of manufacturing goods. We exclude the economic base indicator from this regression, due to the high correlation coefficient with the trade measures (see Table A4 in Appendix A), which further corroborates the notion that geographical concentration and tradability are intimately associated. We also enter exports and imports separately because of their high correlation (see Appendix A). In Columns 1 and 2, we present the estimated coefficients for exports and imports by considering the world as a trade partner. Columns 3 and 4 consider the EU-28 as a trade partner, which is the main exporting and importing area for Italian firms. Finally, Columns 5 and 6 analyze trade linkages with China, which supposedly is the origin country of COVID-19. The setting of these regressions is similar to that of Table 4, with the dependent variable capturing the provincial share of COVID-19 cases as of March 4th and with a time lag of 8 days only.

The results suggest that the relationship between trade and COVID-19 emerges in connection with total provincial manufacturing exports, in line with our conceptual framework, and not with imports. In terms of

TABLE 10 COVID-19 cases and trade linkages in manufacturing

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var: COVID-19 cases, March 4						
Exports to the World	44.223**					
	(17.186)					
Imports from the World		12.010				
		(9.131)				
Exports to EU-28			69.974***			
			(21.449)			
Imports from EU-28				11.419		
				(8.210)		
Exports to China					14.482	
					(8.874)	
Imports from China						8.729
						(8.978)
COVID-19 cases, Feb 25	0.468***	0.460***	0.467***	0.461***	0.466***	0.449***
	(0.032)	(0.033)	(0.032)	(0.033)	(0.033)	(0.036)
Spatial COVID-19 cases, Feb 25	−0.383	−0.315	−0.362	−0.315	−0.333	−0.336
	(0.494)	(0.505)	(0.484)	(0.504)	(0.503)	(0.512)
Obs.	106	106	106	106	106	106
Pseudo R^2	0.82	0.81	0.82	0.81	0.81	0.81
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***Significance level: $p < .01$.

**Significance level: $p < .05$.

geography, export linkages to China are not correlated with COVID-19, while the correlation is very strong in the case of provinces with tighter links with EU-28 markets. This may be due to the fact that trading within the EU involves transport modes requiring more frequent human interactions.

6 | CONCLUDING REMARKS

Since December 2019, the outbreak of COVID-19 has rapidly thrown the world into an unexpected high-risk global health crisis, with a total number of 7,823,289 confirmed cases on June 15th (WHO, 2020b). Most countries have adopted containment measures, also including the complete lockdown of the population and most economic activities. As a result, the expectations about a new global economic crisis are rather solid. The present article offered an economic geography perspective of analysis on the subnational spread of the COVID-19 infection in the case of Italy. This is based on the striking and intriguing unevenness of COVID-19 at the local level, where the most infected provinces represent a disproportionate share of national GDP, suggesting that the infection hits economic core locations harder. These clear spatial features raise questions about whether, and how, the subnational geography of COVID-19 cases is connected to the local economic structure.

In this sense, we developed an analysis of the relationship between the geography of COVID-19 and the structure of local economies, by hypothesizing that locations specialized in economic activities that are characterized by high geographical concentration might be subject to relatively higher infections due to the physical contacts entailed by the agglomeration advantages characterizing these industries. The spatial agglomeration of economic activity, in fact, implies the existence of localized traded or untraded advantages, which potentially involve frequent and dense face-to-face interactions (Krugman, 1991a; Storper & Venables, 2004). Moreover, the high geographical concentration of industry also relates to the spatiality of its demand (i.e., tradability). Hence, local economies hosting clustered sectors are prone to generating trading relationships across geographical space, at the national and international level.

Although they cannot be interpreted in causal terms, our results suggest that there is a positive association between the geography of COVID-19 and the economic base of Italian provinces, measured using sector data at the five-digit level taken from the 2011 (and 1991) Census of Industry and Services. This relationship is robust to the inclusion of a large number of covariates, such as previous infections, population density, demographic factors, and others, as well as regional dummies. Moreover, we find that the relationship under investigation is mostly driven by the local specialization in geographically concentrated manufacturing activities, rather than services. This could be due to the long-time shifts, close vicinity, and continuative interactions between a relatively large number of employees performing manufacturing tasks, as compared to services. In line with this evidence, the positive correlation persists also in the case of provincial manufacturing exports and COVID-19 infections. This descriptive evidence is consistent with the possibility that the geographical concentration of economic activity in specific areas of the country acts as a vehicle of disease transmission, thus generating a core-periphery pattern in the geography of COVID-19, which might follow the lines of the local economic landscape and the tradability of its output.

Some limitations of this study should also be acknowledged. As mentioned above, we cannot produce causal evidence on the relationship under analysis, partly due to the early stage of the COVID-19 outbreak and the ongoing containment measures at the time of this writing. Although we consider a large number of covariates and regional dummies in our empirical setting, we cannot rule out the potential bias connected to provincial unobserved heterogeneity. Moreover, the data on COVID-19 confirmed cases might also suffer from weaknesses related to how infections are detected and reported within each regional health system (although regional dummies might capture this). As we discussed above, the available information probably underestimates the real incidence of COVID-19. This might generate a lower bound picture of the situation. In general, however, national data should be more harmonized than cross-country data, given the existence of national protocols and guidelines in testing and detecting symptoms.

Although the evidence produced in this article is descriptive in nature and relatively preliminary, it can offer important implications for policy and research. First, clarifying the link between the incidence of COVID-19 infections and the characteristics of the local economy may help policymakers in the definition of containment measures that will possibly have to be targeted now that (at the time of this writing) the nation-wide lockdown has been lifted. In this sense, social distancing measures and reinforced containment checks could be reintroduced in areas susceptible of more frequent transmission along the lines described in this study. Considering that these areas may also represent the economic core of a national economy, a strong public financial support in favor of these locations should definitely accompany eventual containment measures. Second, the recognition that the subnational geography of COVID-19 does not follow a random pattern, but may instead be associated with specific economic profiles, can be a point for the involvement of regional policy, at the national and EU level, in the design of targeted support tools for local recovery and entrepreneurship (Pini & Rinaldi, 2020). In this sense, considering that the nation-wide lockdown hits all regions irrespective of the local severity and spread of COVID-19 infections, peripheral and/or stagnating regions that have more limited recovery capacity should also be supported through tailor-made development interventions (Barca et al., 2012). Finally, in the presence of a relationship between COVID-19 cases and local economic structures, future academic investigations of the impact of COVID-19 on the economy should take into consideration that the territorial profile of sector specialization may play a notable role.

ACKNOWLEDGMENTS

The authors would like to thank the participants of the AISRe Conference 2020.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ORCID

Andrea Ascani http://orcid.org/0000-0003-4738-3144
Alessandra Faggian https://orcid.org/0000-0002-2479-2933
Sandro Montresor https://orcid.org/0000-0002-1195-901X

REFERENCES

Apostolopoulos, Y., & Sönmez, S. (2007). *Population mobility and infectious disease*. Springer.
Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies*, *29*, 155–173.
Backer, J. A., Klinkenberg, D., & Wallinga, J. (2020). Incubation period of 2019 novel coronavirus (2019-nCov) infections among travelers from Wuhan, China, 20-28 January 2020. *Euro Surveillance*, *25*(6):2000062. https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062
Barca, F., McCann, P., & Rodriguez-Pose, A. (2012). The case for regional development intervention: Place-based versus place-neutral approaches. *Journal of Regional Science*, *52*(1), 134–152.
Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography*, *85*(3), 289–311.
Brakman, S., Garretsen, H., & Marrewijk, C. (2019). *An introduction to geographical and urban economics. A spiky world*. (3rd ed.). Cambridge University Press.
Brakman, S., Garretsen, H., & Zhao, Z. (2017). Spatial concentration of manufacturing firms in China. *Papers in Regional Science*, *96*, 179–214.
Cainelli, G., Di Maria, E., & Ganau, R. (2014). An explanation of firms' internationalization modes, blending firm heterogeneity and spatial agglomeration: Microevidence from Italy. *Environment and Planning A: Economy and Space*, *46*(4), 943–962.
Caselli, M., Fracasso, A., & Traverso, S. (2020). *Mitigation of risks of Covid-19 contagion and robotisation: Evidence from Italy*. Covid Economics Vetted and Real Time Papers, Issue 17, CEPR Press.
Castellani, D., Meliciani, V., & Mirra, L. (2016). The determinants of inward foreign direct investment in business services across European regions. *Regional Studies*, *50*(4), 671–691.
Cerqua, A., & Di Stefano, R. (2020). *When did coronavirus arrived in Europe?* Mimeo, Sapienza University of Rome.
Dingel, J., & Neiman, B. (2020). *How many jobs can be done at home*? NBER Working Paper No. 26948.

Duranton, G., & Overman, H. (2005). Testing for localization using micro-geographic data. *Review of Economic Studies*, *72*(4), 1077–1106.

Ellison, G., & Glaeser, E. (1997). Geographic concentration in U.S. manufacturing industries: A dartboard approach. *Journal of Political Economy*, *105*(5), 889–927.

Faggio, G., Silva, O., & Strange, W. (2017). Heterogeneous agglomeration. *Review of Economics and Statistics*, *99*(1), 80–94.

Gervais, A., & Jensen, B. (2019). The tradability of services: Geographic concentration and trade costs. *Journal of International Economics*, *118*, 331–350.

Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics*, *46*, 254–277.

Hanson, G. (2000). Firms, workers and the geographic concentration of economic activity. In G. Clark, M. Gertler, & M. Feldman (Eds.), *The Oxford handbook of economic geography* (pp. 477–494). OUP.

Herbinger, K. H., Beissner, M., Berens-Riha, N., Bretzel, G., Alberer, M., Nothdurft, H. D., Schunk, M., von Sonnenburg, F., & Löscher, T. (2016). Spectrum of imported infectious diseases: a comparative prevalence study of 16,817 German travelers and 977 immigrants from the tropics and subtropics. *American Journal Tropical Medicine and Hygiene*, *94*(4), 757–766.

Iammarino, S., & McCann, P. (2006). The structure and evolution of industrial clusters: Transactions, technology and knowledge spillovers. *Research Policy*, *35*(7), 1018–1036.

Jensen, J. B., & Kletzer, L. (2006). Tradable services: understanding the scope and impact of services offshoring. In L. Brainard, & S. M. Collins (Eds.), *Offshoring White-Collar Work – Issues and Implications*. Brookings Trade Forum.

Kaldor, N. (1970). The case for regional policy. *Scottish Journal of Political Economy*, *17*(3), 337–348.

Kelejian, H. H., & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, *17*, 99–121.

Kemeny, T., & Storper, M. (2015). Is specialization good for regional economic development? *Regional Studies*, *49*(6), 1003–1018.

Krugman, P. (1991a). *Geography and Trade*. Published jointly by Leuven University press and the MIT Press.

Krugman, P. (1991b). Increasing returns and economic geography. *Journal of Political Economy*, *99*, 483–499.

Lai, C. C., Shih, T. P., Ko, W. C., Tang, H. J., & Hsueh, P. R. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges. *International Journal of Antimicrobial Agents*, *55*(3):105924. https://doi.org/10.1016/j.ijantimicag.2020.105924

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. *Annals of Internal Medicine*, *172*, 577–582. https://doi.org/10.7326/M20-0504

Lundvall, B. A. (1992). *National systems of innovation: Toward a theory of innovation and interactive learning*. Frances Pinter.

Marshall, A. (1890). *The principles of economics*. History of economic thought books from McMaster University archive for the history of economic thought.

McCann, P. (2013). *Modern urban and regional economics* (2nd ed.). Oxford University Press.

Meyer, S., & Held, L. (2017). Incorporating social contact data in spatio-temporal models for infectious disease spread. *Biostatistics*, *18*(2), 338–351.

Paez, A., Lopez, F., Menezes, T., Cavalcanti, R., & da Rocha Pitta, M. G. (2020). A spatio-temporal analysis of the environmental correlates of COVID-19 incidence in Spain. *Geographical Analysis*. Advance online publication, https://doi.org/10.1111/gean.12241

Pedersen, M. G., & Meneghini, M. (2020). *Quantifying undetected COVID-19 cases and the effects of containment measures in Italy*. Mimeo, University of Padova.

Pini, M., & Rinaldi, A. (2020). Nuova imprenditorialità mancata e perdita di occupazione: Prime valutazioni sugli effetti della pandemia sul sistema produttivo italiano. *EyesReg*. Advance online publication, *10*(3).

Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science*, *50*(1), 203–219.

Ribeiro, H. V., Sunahara, A. S., Sutton, J., Perc, M., & Hanley, Q. S. (2020). City size and the spreading of COVID-19 in Brazil. *PLoS One*, *15*(9), e0239699. https://doi.org/10.1371/journal.pone.0239699

Romer, P. M. (1987). Growth based on increasing returns due to specialization. *The American Economic Review*, *77*(2), 56–62.

Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Froeschl, G., Wallrauch, C., Zimmer, T., Thiel, V., Janke, C., Guggemos, W., Seilmaier, M., Drosten, C., Vollmar, P., Zwirglmaier, K., Zange, S., Wölfel, R., & Hoelscher, M. (2020). Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine*, *382*, 970–971. https://doi.org/10.1056/NEJMc2001468

Sönmez, S., Wiitala, J., & Apostolopoulos, Y. (2019). How complex travel, tourism, and transportation networks influence infectious disease movement in a borderless world. In D. Timothy (Ed.), *Handbook of globalization and tourism*. Edward Elgar.

Sun, P., Lu, X., Xu, X., Sun, W., & Pan, B. (2020). Understanding of COVID-19 based on current evidence. *Journal of Medical Virology*, *92*, 548–551. https://doi.org/10.1002/jmv.25722

Stier, A. J., Berman, M. G., & Bettencourt, L. M. A. (2020). *COVID-19 attack rate increases with city size, Covid-19-e-pre-print*, Ithaca, NY: Cornell University.
Storper, M. (1995). The resurgence of regional economies, ten years later: The region as a nexus of untraded interdependencies. *European Urban and Regional Studies*, *2*, 191–221.
Storper, M., & Venables, A. (2004). Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography*, *4*(4), 351–370.
Tarwater, P., & Martin, C. (2001). Effects of population density on the spread of disease. *Complexity*, *6*(6), 29–36.
Venables, A. (1996). Equilibrium locations of vertically linked industries. *International Economic Review*, *37*, 341–359.
Wallinga, J., Teunis, P., & Kretzschmar, M. (2006). Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American Journal of Epidemiology*, *164*(10), 936–944.
Wang, W., Tang, J., & Wei, F. (2020). Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCov) in Wuhan, China. *Journal of Medical Virology*, *92*(4), 441–447.
WHO (2020a). *Coronavirus disease 2019 (COVID-2019)*. Situation Report 51. Geneva, Switzerland: World Health Organization.
WHO (2020b). *Coronavirus disease 2019 (COVID-2019)*. Situation Report 147. Geneva, Switzerland: World Health Organization.
Young, A. (1928). Increasing returns and economic progress. *Economic Journal*, *38*(152), 527–542.
Zunyou, W., & McGoogan, J. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China. *Journal of the American Medical Association*, *323*(13), 1239–1242.

How to cite this article: Ascani A, Faggian A, Montresor S. The geography of COVID-19 and the structure of local economies: The case of Italy. *J Regional Sci*. 2021;61:407–441. https://doi.org/10.1111/jors.12510

APPENDIX A

TABLE A1 Descriptive statistics

Variable	Obs	Mean	*SD*	Min	Max
Share of COVID-19, Mar 4	107	0.9346	2.789	0	18.683
Share of COVID-19, Mar 8	107	0.9346	2.329	0	14.370
Economic base	107	2.01	3.612	0.06	29.595
Population density	107	269.77	382.251	36.99	2,616.675
Deaths	107	8.43	.656	6.973	10.625
Health emigration	107	24,758.38	20,548.9	1,819	126,000
Old population	106	0.236	.024	0.174	0.291
Male population	106	0.488	.005	0.476	0.504
Tourism rate	107	2.811	4.092	0.33	31.793
Unemployment rate	107	10.977	5.906	2.893	27.625
Foreign residents	106	0.082	.034	0.022	0.185
Airport	107	0.327	.471	0	1

TABLE A2 Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Economic base	1.00									
(2) Population density	0.61	1.00								
(3) Deaths	0.73	0.51	1.00							
(4) Health emigration	0.46	0.34	0.64	1.00						
(5) Old population	−0.18	−0.22	−0.31	−0.34	1.00					
(6) Male population	−0.13	−0.14	−0.25	−0.13	−0.54	1.00				
(7) Tourism rate	0.05	−0.06	−0.02	−0.08	0.03	−0.08	1.00			
(8) Unemployment rate	−0.18	−0.01	−0.01	0.35	−0.44	0.15	−0.33	1.00		
(9) Foreign residents	0.33	0.15	0.18	−0.11	0.25	−0.10	0.17	−0.66	1.00	
(10) Airport	0.37	0.23	0.45	0.22	−0.22	−0.14	0.17	0.11	0.08	1.00

TABLE A3 Matrix of correlations between economic-based indicators by economic activity

`	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Manufacturing	1.00								
(2) Services	0.71	1.00							
(3) Energy	0.35	0.82	1.00						
(4) Water, sewage, waste	0.61	0.90	0.74	1.00					
(5) Wholesale, retail, repair	0.81	0.95	0.60	0.86	1.00				
(6) Transport and storage	0.39	0.87	0.98	0.82	0.68	1.00			
(7) Hotel and restaurants	0.75	0.97	0.71	0.85	0.96	0.77	1.00		
(8) Professional services	0.72	0.99	0.80	0.87	0.94	0.84	0.95	1.00	
(9) Other services	0.71	0.98	0.81	0.94	0.93	0.87	0.94	0.95	1.00

TABLE A4 Matrix of correlations between economic base and trade variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Economic base	1.00							
(2) Economic base (manufacturing)	0.89	1.00						
(3) Exports to the World	0.86	0.95	1.00					
(4) Imports from the World	0.94	0.82	0.87	1.00				
(5) Exports to the EU-28	0.76	0.93	0.96	0.75	1.00			
(6) Imports from the EU-28	0.91	0.79	0.86	0.99	0.74	1.00		
(7) Exports to China	0.89	0.90	0.92	0.92	0.82	0.91	1.00	
(8) Imports from China	0.85	0.77	0.82	0.91	0.71	0.90	0.87	1.00

APPENDIX B

Controlling for the economic and demographic size of provinces

An important aspect in the empirical setting of our analysis is, as usual, the potential omitted variable bias. In principle, it could be argued that two good candidates to attenuate this bias could be the economic and demographic sizes of provinces, both likely to be related to the local economic structure and, although with not yet decisive evidence, to the share of national infections (Ribeiro et al., 2020; Stier et al., 2020).

To control for these factors, we include in our model the provincial GDP in 2017 expressed in constant PPP (base year 2015), taken from the OECD, and the provincial population in 2019, taken from ISTAT. Both variables are taken at the most recent available year. We also consider the dependent variable at different relevant dates to check the stability of our estimates. The results of this exercise are presented in Table B1. The coefficient of the economic base indicator remains positive and significant across the presented specifications, thus reassuring us of

TABLE B1 Checking for economic and demographic size

Dep. Var: COVID-19 cases	(1) Mar 8	(2) Apr 3	(3) May 16	(4) Jun 15
COVID-19 cases, Feb 25	0.289***	0.014	0.009	0.006
	(0.032)	(0.017)	(0.012)	(0.013)
Economic base	0.487***	0.571***	0.590***	0.597***
	(0.172)	(0.091)	(0.065)	(0.066)
GDP	−0.000**	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Population	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Density	−0.001**	−0.001***	−0.001***	−0.001***
	(0.001)	(0.000)	(0.000)	(0.000)
Deaths	−0.523	−0.097	0.028	0.023
	(0.533)	(0.282)	(0.200)	(0.206)
Health emigration	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Old Population	3.111	−2.320	1.489	1.589
	(12.541)	(6.624)	(4.714)	(4.834)
Male population	48.858	15.763	6.454	5.993
	(50.439)	(26.640)	(18.960)	(19.444)
Tourism rate (no summer)	−0.072	−0.045	−0.038	−0.040
	(0.070)	(0.037)	(0.026)	(0.027)
Unemployment rate	−0.047	−0.022	−0.009	−0.009
	(0.057)	(0.030)	(0.022)	(0.022)
Foreign residents	8.129	4.873	4.188	4.198
	(8.729)	(4.610)	(3.281)	(3.365)
Airport	0.840**	0.417**	0.298**	0.318**
	(0.339)	(0.179)	(0.127)	(0.131)

(Continues)

TABLE B1 (Continued)

Dep. Var: COVID-19 cases	(1) Mar 8	(2) Apr 3	(3) May 16	(4) Jun 15
Spat.COVID-19 cases, Feb 25	−0.147	0.102	0.143	0.141
	(0.497)	(0.263)	(0.187)	(0.192)
Obs.	106	106	106	106
Pseudo R^2	0.736	0.819	0.904	0.903
Region dummies	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.

***$p < .01$.

**$p < .05$.

TABLE B2 Controlling for provincial specialization

Dep. Var: COVID-19 cases	(1) Mar 8	(2) Apr 3	(3) May 16	(4) Jun 15
COVID-19 cases, Feb 25	0.294***	0.017	0.010	0.008
	(0.033)	(0.018)	(0.013)	(0.014)
Economic base	0.145**	0.275***	0.315***	0.323***
	(0.062)	(0.034)	(0.025)	(0.026)
Krugman Index	2.157	1.169	0.933	0.919
	(1.486)	(0.811)	(0.600)	(0.613)
Density	−0.001*	−0.001**	−0.001***	−0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Deaths	−0.066	0.209	0.254	0.248
	(0.461)	(0.252)	(0.186)	(0.190)
Health emigration	0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Old population	4.447	−0.130	4.187	4.298
	(12.504)	(6.826)	(5.047)	(5.155)
Male population	77.896	39.804	28.008	27.526
	(49.643)	(27.102)	(20.039)	(20.466)
Tourism rate (no summer)	−0.071	−0.049	−0.044	−0.046
	(0.070)	(0.038)	(0.028)	(0.029)
Unemployment rate	−0.013	0.008	0.020	0.019
	(0.056)	(0.031)	(0.023)	(0.023)
Foreign residents	5.042	3.002	2.974	2.994
	(8.608)	(4.699)	(3.475)	(3.549)

TABLE B2 (Continued)

Dep. Var: COVID-19 cases	(1) Mar 8	(2) Apr 3	(3) May 16	(4) Jun 15
Airport	0.837**	0.400**	0.274**	0.293**
	(0.344)	(0.188)	(0.139)	(0.142)
Spat.COVID-19 cases, Feb 25	−0.255	0.019	0.071	0.070
	(0.505)	(0.276)	(0.204)	(0.208)
Obs.	106	106	106	106
Pseudo R^2	0.725	0.798	0.884	0.884
Region dummies	Yes	Yes	Yes	Yes

Note: Standard errors are in parenthesis.
***Significance level: $p < .01$.
**Significance level: $p < .05$.
*Significance level: $p < .1$.

the fact that the local economic and demographic size of provinces do not affect the relationship under analysis. Nonetheless, one concern with these augmented models is that the local economic size can plausibly be an outcome of our measure of the local economic base. In fact, regional economies with a stronger presence of industries that are subject to agglomeration benefits, can reasonably be the largest regional economies, as their production (i.e., economic size) is literally the output of their local activities. Not surprisingly, we detect a multicollinearity issue, given the high correlation coefficient between the economic base and the GDP (>0.9). Regarding population as a control variable, the concurrent presence of provincial density in the model (with population as the numerator) might also pose some concerns regarding its inclusion in the model. These are both issues to consider in interpreting the results of Table B1.