Why Does COVID-19 Affect Some Cities More than Others? Evidence from Brazil*

Juan Pablo Chauvin

Research Department, Inter-American Development Bank

Abstract

This paper investigates what explains the variation in impacts of COVID-19 across Brazilian cities. I assemble data from over 2,500 cities on COVID-19 cases and deaths, population mobility, and local policy responses. I study how these outcomes correlate with pre-pandemic local characteristics, drawing comparisons with existing US estimates when possible. As in the United States, the connections between city characteristics and outcomes in Brazil can evolve over time, with some early correlations fading as the pandemic entered a second wave. Population density is associated with greater local impact of the disease in both countries. However, in contrast to the US, the pandemic in Brazil took a greater toll in cities with higher income levels – consistent with the fact that higher incomes correlate with greater mobility in Brazil. Socioeconomic vulnerabilities, such as the presence of slums and high residential crowding, correlate with higher death rates per capita. Cities with such vulnerabilities in Brazil suffered higher COVID-19 death rates despite their residents' greater propensity to stay home. Policy responses do not appear to drive these connections.

Keywords: COVID-19, coronavirus, Brazil, cities, developing countries.
JEL Codes: I18, R10, O18.

*Please address related correspondence to juancha@iadb.org. I am grateful to an anonymous reviewer for useful suggestions, and to Nicolás Herrera L., Juliana Pinillos, Julio Trecenti, and Haydée Svab for their outstanding research assistance. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

1 Introduction

The impacts of the COVID-19 pandemic can vary strikingly from one city to another within the same country (Allcott et al., 2020; Desmet and Wacziarg, 2021; Glaeser et al., 2020). I study the drivers of these differences among cities in Brazil, a country that had suffered one of the world's highest number of fatalities from the disease at the time of writing – second only to those in the United States. Combining data from multiple sources, I investigate how local variation in COVID-19 cases and deaths, residents' mobility, and local governments' policy responses correlate with pre-pandemic city characteristics at different points in time. My study focuses on the first year of the pandemic, from February 26, 2020, when the first case was reported in Brazil, until a year later, February 25, 2021.

The first part of my analysis closely replicates the specification used by Desmet and Wacziarg (2021) to study the correlates of COVID-19 cases and deaths per capita across US counties. This allows me to explore how the drivers of the local toll of the disease compare in the two countries with the highest number of deaths during the first year of the pandemic, which also have different income levels. I consider the role of population density, commuting patterns, median household income, distance to internationally connected airports, and the presence of populations that are vulnerable because of age or because they reside in nursing homes. I then introduce an additional set of covariates that may be especially relevant in low- and middle-income countries. These include informality rates, the education of the labor force, the presence of racial minorities, residential crowding, and the presence of slums (locally known as "favelas"). Last, I investigate potential mechanisms shaping these correlations by examining how local mobility and policy responses to the pandemic varied with city characteristics, using a local mobility panel based on data from over 60 million cell phones, and a containment intensity panel built from a novel data set of local COVID-19 policies.

I find that some of the correlates of the local impact of COVID-19 are very similar in Brazil and in the United States, but others are distinctly different. Among the similarities, the robust connection of the local toll of the disease with population density stands out. Following a short period at the beginning of the pandemic, in which population density had a negative correlation with deaths, the situation changed. Density became correlated with more COVID-19 deaths per capita as the pandemic continued, and this carried on throughout the rest of the study period in both countries. Other parallels include the the roles of public transportation commuting and of distance to the closest internationally connected airport. Both factors exhibit early positive correlations that later fade away.

Another similarity between the Brazilian results and those in the US is that the connections between city characteristics and COVID-19 outcomes over time look very similar regardless of whether they are measured on the same calendar dates for all cities, or whether they are measured at a fixed number of days since the epidemic's local onset, which varied from city to city. This supports the view that the differences in the local impact of COVID-19 at a given point in time are not just an artifact of heterogeneity in the time of arrival of the virus, but reflect lasting disparities in the vulnerability of cities to the pandemic (Desmet and Wacziarg, 2021).

Notwithstanding these parallels, the analysis also shows some striking differences between the United States and Brazil. These include household income and local education levels – which are associated with fewer COVID-19 deaths per capita in the United States, but more in Brazil – and the presence of racial minorities – which have a positive correlation with deaths in the United States, but no statistically significant correlation in Brazil, after controlling for the other covariates.

The most salient contrast between the two countries is the correlation between the median household income and COVID-19 deaths, which in Brazilian cities is large, positive, and statistically significant. Put differently, higher-income cities in Brazil tended to have worse COVID-19 repercussions – the opposite of the case of US counties. The mobility analysis suggests that this may be at least partially explained by the fact that richer cities maintained relatively more vibrant economic activity during this period. In Brazil, cities around the country started reopening their economies as early as in April, even as reported infections and deaths continued to escalate. Cities with higher income levels tend to feature more commercial activity, which incentivizes mobility and increases exposure of the local population to the virus. Accordingly, I find a consistent negative association between income and the proclivity to stay at home, which started to emerge in May 2020. Other results are also consistent with this interpretation. For Brazilian cities with high population density and a large share of college-educated in employment – two characteristics that tend to be associated with a vibrant local economy in both developing and high-income countries (Chauvin et al., 2017) – the correlation with the propensity to stay at home was positive at the very beginning, but shortly after became and stayed negative.

An alternative, non-exclusive explanation for the connection between income and the local impact of the disease in Brazil, is that richer cities may have been better able to identify and report COVID-19 cases and deaths (for example, due to better testing coverage, or better reporting protocols). The evidence, however, is less supportive of this interpretation. Because states are the highest subnational jurisdiction in the country, testing policies and data reporting standards are more likely to vary across than within states. Thus, if there is a positive income bias in reporting, an analysis that includes state fixed effects in the regression would likely attenuate it; instead, I find that, after incorporating state fixed effects, this connection becomes more pronounced.

I also find that characteristics associated with lower mobility are not always associated with fewer deaths per capita across Brazilian cities. In particular, variables capturing socioeconomic and demographic vulnerabilities to the pandemic – such as older populations (Levin et al., 2020), residential crowding (Ahmad et al., 2020), and the presence of favelas (Brotherhood et al., 2020) – tend to be correlated with more deaths per capita, particularly during the first wave of the pandemic, despite a higher propensity to stay at home in cities with such characteristics. This pattern is most pronounced in cities with large shares of their households located in favelas, which had a disproportionately high number of deaths per capita during the first wave, but not afterwards. Meanwhile, these cities were more likely than others to stay at home throughout the period of study. This finding resonates with the results in Sheng et al. (2021), who find that the comparatively higher infection rates found in Mumbai slums cannot be explained by differential compliance with mobility restrictions.

Finally, I find that the set of policies chosen by municipalities in response to the disease are largely uncorrelated with the city characteristics under analysis over the period of study. These results suggest that the links between city-level characteristics, the local COVID-19 impact, and mobility, are not explained by differential policy responses at the city level.

This paper contributes to the fast-growing COVID-19 literature, particularly to a branch focusing on cities. It is closest to research exploring how city characteristics relate to the impact of the pandemic (Allcott et al., 2020; Almagro and Orane-Hutchinson, 2020; Desmet and Wacziarg, 2021; Glaeser et al., 2020; Knittel and Ozaltun, 2020; McLaren, 2020). While this literature mostly focuses on the United States, my analysis provides evidence from Brazil, the second hardest-hit country during the first year of the pandemic. My work also relates to a literature that explores how COVID-19 may affect low- and middle-income countries differently from countries with higher income levels. Existing work tends to focus on country-level characteristics (Alfaro et al., 2020; Alon et al., 2020; Brown et al., 2020; Busso et al., 2020; Chauvin et al., 2020; Goldberg and Reed, 2020; Hausmann and Schetter, 2020) or local case studies (Brotherhood et al., 2020; Sheng et al., 2021). My research highlights subnational spatial variation across the territory of a large middle-income country.

2 Data

I use two types of data in this paper. The first type consists of time-invariant, pre-pandemic city characteristics obtained from multiple sources. Most of these are directly constructed from the microdata of the 2010 Brazilian demographic census, or obtained from aggregates constructed by the Brazilian Institute for Geography and Statistics (IBGE) based on data from this census year (IBGE, 2012). There are two exceptions in my main specification. One is the measure of population density – the total population living within 1 km of the average inhabitant of the city; I compute this measure using 2015 data from the Global Human Settlement Layer (Schiavina et al., 2019). The other is the share of households in the city located in slums (commonly referred to as "favelas"), which is based on 2019 estimates generated by the IBGE using multiple sources (IBGE, 2020). In addition, I include in the appendix tests of the robustness of the results to the inclusion of a number of additional regressors, in which I also use data from the Brazilian National Health System (DATASUS, 2021), and climate data from Harris et al. (2014). Appendix A provides further details on the sources and definition of each variable, and Appendix Table B1 reports summary statistics.

The second type of data consists of time-varying outcome measures. Summary statistics for these variables are reported in Appendix Table B2. The main outcomes are the number of daily COVID-19 cases and deaths, which I obtain from Brasil.io (Justen, 2021), an open data platform that collects information directly from the state-level health secretaries. In spite of their widespread use in the COVID-19 literature, cases and deaths statistics have important limitations. It is widely recognized that they likely undercount the true impact of the disease due to limited testing coverage or preferences-driven demand for testing (Baqui et al., 2020; Cintra and Fontinele, 2020). Across countries, while there is a positive correlation between income levels and COVID-19 mortality, there is no evidence that the undercounting of these deaths and cases is systematically correlated to income (Goldberg and Reed, 2020). The situation might, however, be different across cities. Testing coverage and deaths reporting protocols – which I do not observe – might be correlated with some of regressors of interest, such as income and population. I use state fixed effects in all regressions to mitigate these concerns. The COVID-19 data in this paper are produced by state health secretaries, and the testing and reporting standards are more likely to vary across than within states. In addition, I focus most of the analysis on death counts. As in prior epidemics such as Ebola and SARS, deaths are seen as relatively more reliable than cases to track the spread of the disease because it is less likely for a death than for a non-fatal infection to go unregistered (Avery et al., 2020; Maugeri et al., 2020; O'Driscoll et al., 2021). That said, it is important to keep in mind that, if measurement error were systematically associated with a given regressor in a way that were not corrected by the inclusion of other covariates and of state fixed effects in the regression, then the corresponding estimates would reflect the effects of the regressor on both COVID-19 deaths and their reporting.[1]

[1] An alternative would be to use excess deaths relative to prior years, but this is only feasible in Brazil for

The second time-varying outcome is the "Social Isolation Index" produced by the private firm InLoco (InLoco, 2021). This measure is is based on anonymized data from over 60 million cell phones, and is defined as the share of phone users in the city who stayed at home on a given date, where staying at home is defined as remaining within 450 meters of the location identified as the residence. After the pandemic began, InLoco made the index available to researchers upon request, and it has been used in multiple recent COVID-19 studies (e.g., Candido et al., 2020; Brotherhood et al., 2020).

The third outcome that I observe over time is the policy response of local governments to the crisis. I use a measure of containment intensity, built from a novel data set of local COVID-19 policies by Chauvin et al. (2021). The data set is a municipal-level panel that identifies, for a set of COVID-19 containment policies, whether each of them was in place in a given municipality on a given date. The original sources are local legislation documents obtained from public online records, from which the text of individual articles was extracted and analyzed to attribute policies to municipalities and dates. My measure assigns to each municipality-date the number of days (over the immediately prior 30-day window) in which the municipality had in place at least one of six containment policies, including workplace, commerce, travel, or public transit restrictions; lockdowns; or curfews.

Last, I choose my city definition to capture the level at which the relationships that this paper studies are likely to operate. The administrative jurisdiction typically associated with a "city" in Brazil is the municipality. This is also the level of government at which most of the data described above initially were made available. However, when considering large and small cities in the same analysis, this level is not that best suited to measure the connection between local characteristics and the impact of COVID-19, as municipalities vary greatly in those characteristics and are not necessarily comparable. There were 5,570 Brazilian municipalities at the beginning of 2020, including very rural communities, with fewer than 10,000 people, and municipalities located in large urban areas, with millions of inhabitants. Moreover, medium and large urban areas frequently contain multiple

a small set of cities. Specifically, excess death estimates are based on predictions of the expected number of deaths. These predictions are based on historical mortality data. To produce a dependable weekly panel of excess deaths, we would first need enough pre-pandemic data for each city-week cell to generate a reliable prediction of expected deaths from March 2020 onwards. This may be attainable in highly populated jurisdictions, but it is much less feasible in less-populated cities, where there are relatively few deaths in any given year, and less so in any given week. In addition, we would need comparable mortality data for the post-pandemic period. At the time of this writing, the main source of historical mortality data – the Ministry of Health's Mortality Information System (SIM) – has made available data only through 2019. As a result, existing excess death estimates in Brazil – generated at the level of the 27 states and their capital cities – use 2020 data from the Civil Registry; these data are not entirely compatible with the historical data (Marinho et al., 2020). Another alternative would be to use data on the prevalence of COVID-19 antibodies in the population, as in Buss et al. (2021) and Hallal et al. (2020). However, these data are available only for a limited, non-representative sample of cities.

municipalities, which share common labor and housing markets in spite of being ruled by separate local governments. Because these municipalities are effectively part of the same community in terms of exposure to the virus, it makes sense to group them as part of the same locality in the data. I do this by treating all municipalities that belong to the same commuting zone[2] as a single city. I also constrain my analysis to municipalities of at least 10,000 people, resulting in a total of 2,509 cities in my main sample. An exception is the analysis of the connection between city characteristics and COVID-19 containment policies, which I perform at the municipality level to match the jurisdiction of the policymakers.

3 The Spread of COVID-19 in Brazil

The first confirmed COVID-19 case in Brazil was reported in São Paulo on February 26, 2020. Over the next few days new cases emerged, primarily in the most populated and internationally connected cities (Chauvin, 2021). Figure 1 reports cross-city averages that illustrate the evolution of cases and deaths in the weeks and months that followed. Almost three weeks after the first case, on March 17, the first death was confirmed, and the states of São Paulo and Rio de Janeiro declared a state of emergency. The next day, and for the next two weeks, mobility in Brazilian cities dropped dramatically, with almost 50% of phone users choosing to stay at home on March 18 (third column in Figure 1). Nevertheless, the virus continued to spread toward smaller cities and the interior of the country.

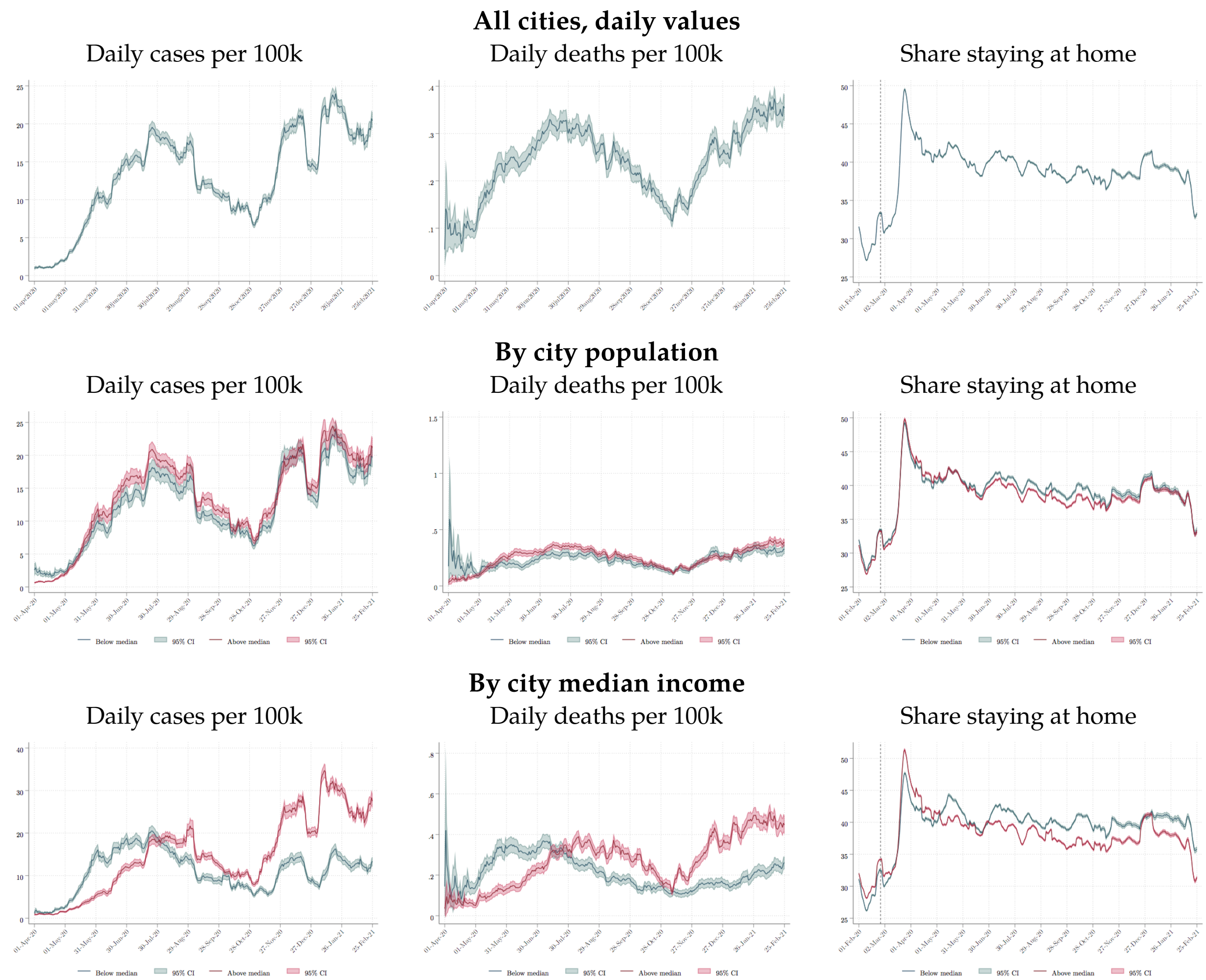
The first wave only peaked toward the end of July 2020, when the average city reported close to 20 cases and 0.32 deaths per 100,000 people. However, some states and cities had already started reopening as early as April, and establishments such as restaurants, bars and gyms were allowed to operate, albeit with restrictions, in some of the largest cities by mid-July. Over the following months, the share of people staying at home dropped, and remained relatively stable at around 40% – higher than in the pre-pandemic period but lower than in March. By the beginning of November, a second wave had started to take hold in many cities around the country and reached its peak by the end of the period of analysis, one year after the first case announcement. By then, Brazil had suffered 303,462 confirmed COVID-19 deaths, making it at that time the country with the second-highest number of deaths worldwide, only behind the United States (Roser et al., 2021). There was only a small and short-lived increase in the propensity of the population to stay at home in late November and early December of 2020. After that, mobility continued to increase in

[2] I use commuting zones ("arranjos populaicionais") as defined by the Brazilian Institute for Geography and Statistics (IBGE, 2016). These zones are groupings of geographically contiguous municipalities associated with the same urban area through work- and study-related commuting.

the average Brazilian city, even as COVID-19 cases and deaths continued to mount.

The spread of the disease evolved quite similarly in cities with populations above and below the median, with only small differences (Figure 1, second panel). Smaller cities reported, on average, fewer cases and deaths per capita than in larger cities during the heights of the first and second wave. They also had a larger share of people staying at home, particularly after the first wave subsided, and before the second wave started. However, the progression of cases, deaths, and mobility over time in smaller cities closely tracked the progression in the larger ones.

Figure 1: Local COVID-19 Cases and Deaths and Mobility Over Time (Cross-City Averages)



Notes: This figure plots cross-city averages and 95% confidence intervals of key outcomes of interest calculated directly from the raw data. All outcomes are expressed as seven-day moving averages of the original daily values. The first panel considers the full sample of cities (N=2,509). The second panel plots these figures separately for the subsamples of cities above and below the median city population. The third panel plots these figures for the subsamples above and below the cross-city median household income.

Wider differences emerge when comparing cities above and below the cross-city median household income (Figure 1, third panel). By May of 2020, richer cities had, on average, significantly fewer cases and deaths per capita than cities below the median. This situation changed over the following weeks and months. The incidence of cases and deaths became the same in the two different income tiers around the peak of the first wave. After that, richer cities reported, on average, increasingly higher numbers of cases and deaths per capita than cities with lower incomes. These changes clearly correspond with changes in the mobility of the populations. During the period from the beginning of the pandemic (February 2020) until May 2020, people in higher-income cities had either a similar or a higher average propensity to stay at home than those in lower-income cities. But after that point, the trends change. From then onward, people in cities with above-median-level incomes were consistently less likely to stay home than those in cities below the median. This shows not only that some city characteristics relate more than others to differences in the local impact of COVID-19, but also that the connection between local characteristics and pandemic outcomes can vary over time. The regression results reported in the following section provide a formal exploration of these differences.

4 Covariates of the Local Impact of COVID-19 across Cities

In this section I investigate the connection between pre-pandemic city characteristics and COVID-19 cases and deaths. To facilitate comparisons with equivalent US estimates, I replicate the specifications in Desmet and Wacziarg (2021). Specifically, I use OLS to estimate the equation:

$$Y_{it}^{l} = \alpha_{0t}^{l} + \sum_{j=1}^{k} \beta_{jt}^{l} X_{ij} + \varepsilon_{it}^{l} \tag{1}$$

where Y_i^l is the outcome of interest l measured for city i on date t, X_{ij} are city-level regressors measured before the pandemic started, and ε_{it}^l is a city-level error term corresponding to outcome l on date t. The coefficient of interest β_{jt}^l estimates the effect of regressor j on outcome of interest l on date t, keeping all other regressors constant. All regressors are standardized.

4.1 Results at the End of the First Year of the Pandemic

I start by looking at the one-year mark since the pandemic began. The first two outcomes I consider are the cumulative count of cases and the cumulative count of deaths attributed to COVID-19. In the main specification I use the inverse hyperbolic sine (IHS) transformation of these variables to include in the estimation cities with zero cases and/or zero deaths

(Bellemare and Wichman, 2020). Table 1 reports these results for the outcomes measured on February 25, 2021, one year after the first reported case. The Appendix (Table B3) reports these results using the logarithmic instead of the IHS transformation, with virtually identical results.[3]

Table 1: City Characteristics and COVID-19 Toll as of February 25, 2021

	(1) IHS cases	(2) IHS cases	(3) IHS deaths	(4) IHS deaths
Panel A: Baseline regressors				
Population (Ln)	0.90***	0.88***	0.91***	0.89***
	(0.02)	(0.02)	(0.02)	(0.03)
Avg. population within 10km (Ln)	0.07**	0.08***	0.15***	0.16***
	(0.03)	(0.03)	(0.03)	(0.03)
Commuting time (Ln)	-0.05**	-0.04**	0.05*	0.03
	(0.02)	(0.02)	(0.03)	(0.02)
Share of people aged 60+	-0.03	-0.02	0.08***	0.15***
	(0.03)	(0.03)	(0.02)	(0.03)
Nursing home residents per 10k pop (IHS)	-0.01	-0.02	-0.01	-0.01
	(0.02)	(0.02)	(0.01)	(0.01)
Km to closest airport connecting to hot spots (Ln)	0.03	0.02	-0.05	-0.03
	(0.04)	(0.04)	(0.05)	(0.05)
Median household income p/c (Ln)	0.42***	0.46***	0.28***	0.45***
	(0.04)	(0.06)	(0.06)	(0.08)
Panel B: Full specification				
Informality rate		0.04**		0.02
		(0.02)		(0.02)
College graduates employment share		0.07**		-0.00
		(0.03)		(0.03)
Black and mixed population share		-0.01		0.10
		(0.05)		(0.07)
Average persons per room (Ln)		0.12**		0.23***
		(0.05)		(0.06)
Share of households located in favelas		-0.01		0.02
		(0.02)		(0.02)
Observations	2,509	2,509	2,509	2,509
R^2	0.80	0.81	0.78	0.78
R^2 of the per-capita specification	0.38	0.39	0.30	0.32

Notes: OLS regressions at the city level. All regressors are standardized. The R^2 of the per capita specification comes from regressions in which the outcomes are directly expressed in per capita terms, and population is excluded from the regressor set. All regression include a constant and state fixed effects. Robust standard errors clustered at the state level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

[3]One year after the first case was detected in Brazil, COVID-19 had reached almost all corners of the country, which explains the virtually identical results obtained with the IHS and logarithm transformations on February 25, 2021. However, this was not the case in the early months of the pandemic. I use IHS in the main specification to be able to compare the evolution of the estimated effects over time keeping the sample of cities constant.

Columns 1 and 3 of Table 1 report, respectively, the estimated effects of the baseline set of regressors on total cases and deaths. I chose these variables to match as closely as possible the main specification in Desmet and Wacziarg (2021), and, thus, to facilitate comparisons (all subsequent references to comparable US results refer to this paper, unless otherwise specified).[4] This involves including in all specifications the logarithm of population, which acts as a scaling factor and allows for the other coefficients to be interpreted as effects on cases and deaths *per capita*.[5] In the full specification (columns 2 and 4) I add as regressors a set of potential determinants of COVID-19 cases and deaths that, on the basis of the existing literature, are likely to be more relevant in low- and middle-income than in high-income countries.

I first look at the role of population density. I follow De la Roca and Puga (2017) in using as a measure of effective density the number of people living within 1 km of the average city dweller. Urban economists have traditionally pointed to the faster spread of contagious disease as a possible source of agglomeration diseconomies, particularly in the developing world (Henderson, 2002; Glaeser, 2009; Bryan et al., 2020). Early COVID-19 studies did not find a connection between density and the local incidence of the disease in cities in the United States, Spain or the United Kingdom (Carozzi et al., 2020). This, however, changed as the pandemic progressed. Desmet and Wacziarg (2021) show that, after the first few weeks, the correlation between the local COVID-19 toll and effective density across US counties became increasingly positive and significant.

I also find density to be positively associated with cases and deaths in Brazil one year after the first case was reported. While the coefficient on cases becomes statistically non-significant after controlling for labor-market and living-condition characteristics, the effect on deaths remains positive and significant. Including additional controls – such as the

[4]Due to data availability, the Brazilian specification does not include a measure of social capital, and it uses the average commuting time instead of the share of people who commute using public transportation. In addition, it uses the share of the population aged 60 and older (instead of 70 and older) because in Brazil – as in other developing countries – younger cohorts have a higher likelihood of dying from COVID-19 than in rich countries (Chauvin et al., 2020). Furthermore, the main specification in Desmet and Wacziarg (2021) does not include state fixed effects. In the regression without fixed effects (reported in Appendix Table B4), the results remain largely the same, even though the statistical significance of the estimates varies in a few cases. I prefer to use the within-state variation in the main specification to address potential measurement differences related to the geographically disparate rollout of testing, as well as state-level differences in data collection and reporting protocols.

[5]The R^2 of these regressions is very high: around 0.8 in the main specification, and as high as 0.97 for cities with populations of 100,000 or higher. This is partly driven by the strong correlation between city population and the local number of COVID-19 cases and deaths. As in Desmet and Wacziarg (2021), I also report the R^2 of the same regressions using the IHS of cases and deaths *per capita* as dependent variables, and excluding the log population from the regressor set. This is informative of the contribution of the other explanatory variables to the overall variation in the local impact of COVID-19.

pre-pandemic stock of health equipment, pre-pandemic health workers per capita, and local weather measures (Appendix Table B5) – does not change this result. The effect does shrink and become non-significant when I restrict the sample to cities of at least 100,000 inhabitants (Appendix Table B6), suggesting that the difference is not coming from comparing medium and large cities, but from comparing small cities with the rest.

Second, I look at the effects of the local average commuting time. While this variable is negatively associated with cases, the association with deaths is positive though not statistically significant after including labor-market and living-condition controls. These results are very similar to those in Desmet and Wacziarg (2021) for the share of commuters using public transportation in the US, which may indicate that the Brazilian and the US measures broadly capture the same phenomenon.[6] In the 2017 Origin-Destination survey of São Paulo, for example, 93% of individuals with commutes longer than 30 minutes used motorized means of transportation, and 60% of this same group used public transportation (Metro de Sao Paulo, 2017). Across the metropolitan São Paulo area, the share of people using public transportation grows sharply with the length of the commute (Appendix Figure B1).[7]

Next, I consider variables related to the presence of populations that are relatively more vulnerable to the virus because of their age. This includes the share of the population aged 60 or older, and the number of nursing home residents per 10,000 inhabitants. I find that having an older population is associated with having more COVID-19 deaths, but not with the total number of cases, in line with the well-known positive connection between age and the risk of dying from the disease (Levin et al., 2020). Meanwhile, I find a *negative* albeit smaller association between COVID-19 deaths and the quantity of nursing home residents – which is not statistically significant in the baseline specification, but becomes significant after including labor-market and living-condition controls. This contrasts with the results for US counties, which in the case of deaths are negative for the share of the population aged 75 or older, and positive for the percentage of nursing home residents in the population.[8] One interpretation is that these two variables are capturing similar variation (i.e., the size of the age-vulnerable population), but in Brazil the nursing homes variable contains less information due to the markedly lower prevalence of these

[6] Harris (2020) also finds a strong, positive association between subway use and the spread of COVID-19 in New York City, but Almagro and Orane-Hutchinson (2020) show, in the same context, that these effects are not statistically significant after controlling for workers' occupations.

[7] Only a few cities have publicly available origin-destination transportation surveys, which prevents me from using this variable directly in the regressions.

[8] Desmet and Wacziarg (2021) also find a statistically significant connection between these variables and the number of COVID-19 cases in the United States; this differs from the finding that emerges from the regression analyzing the situation in Brazil.

institutions in the South American country relative to the United States.

The connection between city-level income and the local COVID-19 toll differs sharply between Brazil and the United States. While richer cities in the United States tend to experience fewer cases and deaths per capita, in Brazil it is the opposite; cities with higher median household income levels experienced more COVID-19 cases and deaths per capita. This result is apparent in the raw data (Figure 1) and in most of the specifications considered (Table 1, Appendix Tables B3, B4, B5, and B6). One potential explanation is that richer cities test more intensely, and thus report more infections than relatively poorer cities even if the are no real underlying differences. Another possible (non-exclusive) explanation is that, in a context in which circulation restrictions are hard to enforce, and a large share of the population lives hand-to-mouth, higher-income cities that support relatively more more economic activity also generate more human interactions, leading, in turn, to more infections, and ultimately more deaths.

While both accounts may partially reflect reality, the evidence appears more consistent with the economic activity story. First, COVID-19 monitoring and policy are largely carried out at the state level. Thus, testing protocols and intensity are likely to differ more across than within states. The introduction of state fixed effects, however, does not reduce but strengthens the household income effect. In the specification without state fixed effects (Appendix Table B4) the point estimates are smaller, and only statistically significant for cases and not for deaths, though still positive and of the same order of magnitude. Second, both the raw data (Figure 1) and the multivariate regression estimates discussed below (Section 5) show that in richer cities the share of the population staying at home was significantly smaller than the share in poorer cities, at least after the first few weeks of the pandemic. This can be viewed as reflecting higher intensity of economic activity in higher-income cities.

Moving beyond the baseline specification, I introduce a set of additional variables that have been highlighted in the literature as potential drivers of the local COVID-19 toll (Panel B, Table 1). US-based research has found a connection between worse local COVID-19 outcomes and lower schooling levels, or the presence of racial minorities (Brown and Ravallion, 2020; Benitez et al., 2020; Wiemers et al., 2020). These factors could play an even more important role in Brazil, given the country's long history of racial disparities and inequalities in access to higher education. In addition, multiple studies have suggested that the higher prevalence of labor informality (Alfaro et al., 2020; Busso et al., 2020; Hausmann and Schetter, 2020) and of urban slums (Brotherhood et al., 2020) may be linked to the faster spread of the disease in developing countries. The analysis shows that one year after the outbreak of COVID-19 in Brazil the number of local reported cases – but not local deaths – is greater among cities that have higher informality rates or a greater share of college graduates in employment. I do not find a statistically-significant connection between COVID-19 outcomes and the share of black or mixed race people in the population, or with the share of households living in favelas.

Throughout all specifications used in the analysis I find a strong and statistically significant connection between greater residential crowding – as captured by the average persons per room in local households – and higher rates of COVID-19 cases and deaths. This is in line with prior evidence showing that the number of people per residence is linked to worse COVID-19 outcomes across US counties (Ahmad et al., 2020; Desmet and Wacziarg, 2021).

4.2 Changes in the Effects of Local COVID-19 Covariates over Time

I turn now to investigating how the connection between local characteristics and the toll of COVID-19 varies over time. I focus from this point onward exclusively on the results for COVID-19 deaths because, as discussed earlier, death measurements are broadly seen as less susceptible to measurement error than case measurements (Avery et al., 2020; Maugeri et al., 2020; O'Driscoll et al., 2021). Results for cases are reported in the Appendix for completeness (Figure B2).

Following the approach taken by Desmet and Wacziarg (2021), I estimate equation 1 using two alternative sample definitions. The first consists in cross-sections for each date between April 1, 2020, and February 25, 2021, using the cumulative COVID-19 deaths at each date as the dependent variable. The results for the regressions with the full set of covariates – corresponding to column 4 in Table 1 – are reported in Figure 2.[9] The second sample definition identifies the onset of the epidemic in each city,[10] and estimates equation 1 at each day elapsed since the onset. The results for COVID-19 deaths with this specification are reported in Appendix Figure B3.

The results show that in Brazil, as in the United States, the associations between city characteristics and the local COVID-19 death toll vary over time. The covariates can be classified into two broad categories based on the time progression of their effects on deaths per capita.

[9] I report estimates starting on April 1, 2020, because deaths attributed to COVID-19 in March were still few and concentrated in a handful of cities. The corresponding results for COVID-19 cases are reported in Appendix Figure B2.

[10] The onset is defined as the day the city crosses a minimum threshold, defined as 0.5 per 100,000 deaths. Because in this specification municipalities with zero cases/deaths are not considered, the dependent variables are expressed in logarithms (as opposed to IHS).

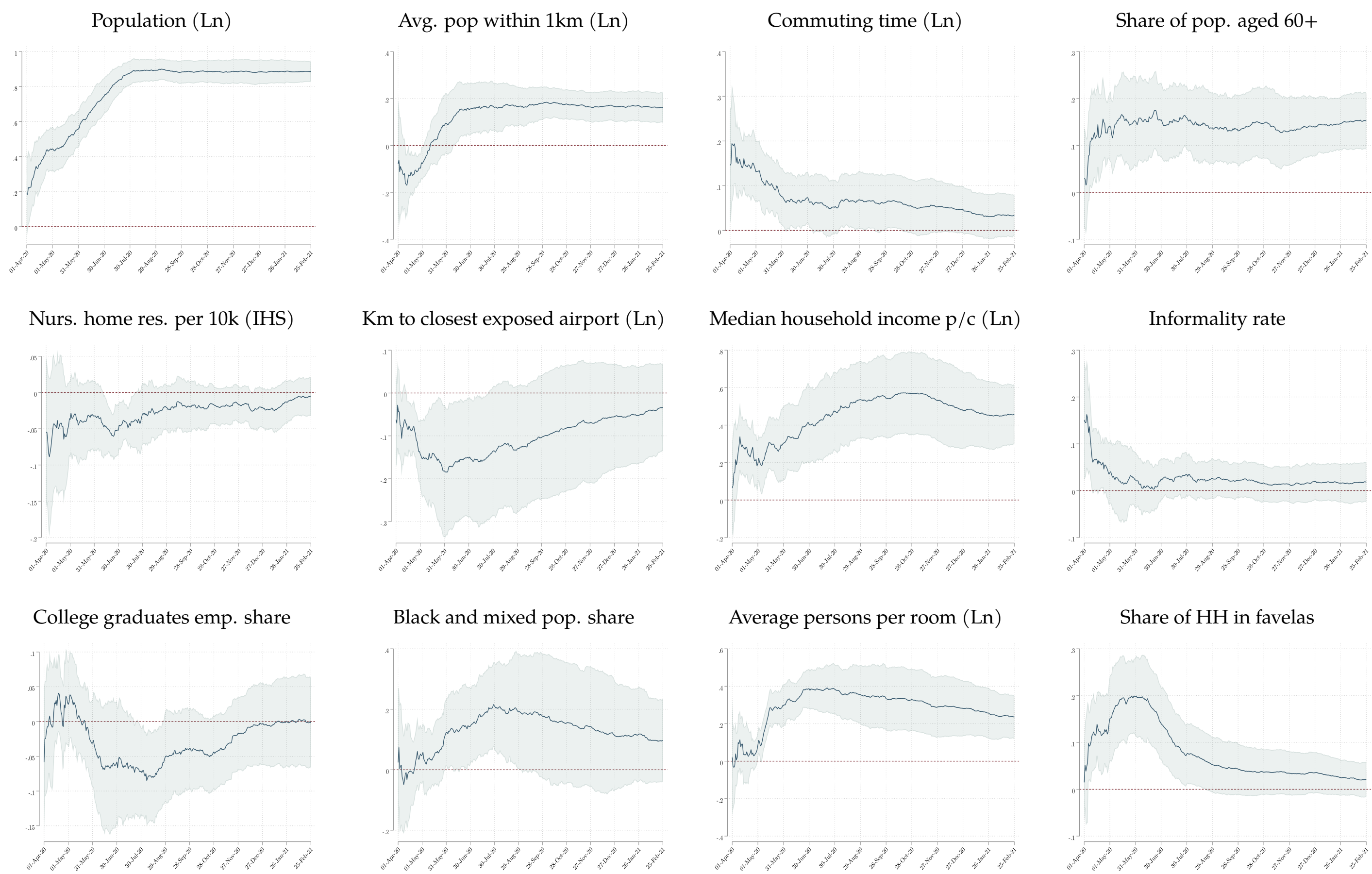
The first category consists of a set of city characteristics for which, holding all other covariates constant, the association with deaths started to raise in the early weeks of the first wave – around April and May of 2020 – and either continued to grow or stabilized after that. This subset consists of city characteristics that are typically associated with higher local productivity and a more vibrant local economy, including household income per capita, the population within 1 km of the average city dweller (my measure of population density), and the share of college graduates in the economy (Chauvin et al., 2017). While higher household income was associated with higher rates of COVID-19 deaths throughout the period of study, both population density and college share effects were initially associated with fewer deaths. Early in the pandemic, population density coefficients turned, and they remained associated with more COVID-19 deaths throughout. By contrast, the negative college share coefficients (showing that the greater the college-educated share of the population, the lower the rate of COVID-19 deaths) gradually dwindled, falling to zero during the second wave.

Another subset of variables in this category is related to the presence of populations that are particularly vulnerable to the disease because of their age (the share of the population aged 60 or older, and the number of nursing home residents per 10,000 people). The connection between having an older population and having more COVID-19 deaths per capita became apparent in mid-April, and the effects remained large and statistically significant thereafter. Cities with a larger presence of nursing-home residents had fewer deaths per capita in the early months of the pandemic, but this difference disappeared in the second wave.

The second category includes socioeconomic characteristics of cities – other than the age of the population – that the literature has linked to increased vulnerability to the pandemic. For these variables I find, keeping all other regressors constant, a strong association with per capita COVID-19 deaths in the early weeks of the pandemic. However, this association then wanes and, in most cases, disappears after the peak of the first wave. This is the case for average commuting time (which, as previously discussed, is linked to greater use of public transportation), proximity to internationally connected airports,[11] informality rates, the share of black and mixed race people in the population, the average persons per room, and the share of households located in favelas. Among these socioeconomic variables, only the one measuring residential crowding remains statistically significant a year after the first case had been detected in the country. This is because the connection with per capita deaths does not drop as rapidly as the connections with other variables.

[11] Figure 2 reports the coefficients for *distance* to the closest internationally connected airport. Here, a negative coefficient implies that cities that were closer to airports saw more deaths per capita during the first wave, holding other covariates constant.

Figure 2: OLS Estimates of Correlations between Cumulative COVID-19 Deaths and City Characteristics



Notes: This figure plots daily OLS estimates from multivariate regressions, for the period between April 1, 2020, and February 25, 2021, with data from 2,509 Brazilian cities. The dependent variable is the inverse hyperbolic sine (IHS) transformation of the total COVID-19 deaths reported in that municipality from the beginning of the pandemic until the correspondent date. All regressors are standardized. All estimations include a constant and state fixed effects. The shaded area shows 95% confidence intervals constructed from robust standard errors clustered at the state level.

These results suggest that in more vulnerable places the pandemic hit particularly early and hard. The rapid decrease in deaths after the first wave could well reflect some form of localized herd immunity. For instance, a study examining the presence of antibodies among the population of Manaus – the largest city in the Amazon, and a city in which more than 50% of households are located in favelas – estimates that around 66% of the population had already been infected by June 2020 (Buss et al., 2021).[12] However, the results could also reflect other factors, such as more aggressive preventive behavior (e.g., increased physical isolation) in response to more salient levels of infections and deaths.

The specification that accounts for timing differences in the local onset of the pandemic, reported in Appendix Figure B3, yields virtually the same results – albeit less precisely measured in a few cases. This suggests that in Brazil, as in the United States, differences in the local impact of the pandemic are not just a reflection of timing, but also of more structural differences in vulnerability. Put differently, it does not seem to be the case that all cities will be equally affected by the pandemic sooner or later, but rather that some cities will in the end suffer a significantly higher toll.

5 City Characteristics, Mobility and Policy

The results discussed in Section 4 reflect both the direct effect of city characteristics on COVID-19 outcomes, and any response by local authorities and residents to the perceived risk, or the actual impact of the virus. These responses may themselves vary with local pre-pandemic characteristics. An effective response in a vulnerable community could, in principle, mitigate or even reverse city characteristics' effects. To investigate the role of the endogenous responses of the local population to the threat of COVID-19, I look at the connection between pre-pandemic characteristics and both mobility and policy responses across cities.

5.1 Propensity to Stay at Home

The literature has documented a connection between mobility and the increase of cases and deaths in the United States (Glaeser et al., 2020), Brazil (Chauvin et al., 2021), and other countries (Cho, 2020; Fang et al., 2020), but the connection between pre-pandemic city characteristics and mobility behavior has received less attention. Figure 3 reports estimates of equation 1 obtained from daily cross-section, multivariate regressions using the same

[12] If lower infection rates after June reflected a sizable immune population, this protection appears to have been transitory; Manaus went on to experience an even more dramatic second wave during the first few weeks of 2021.

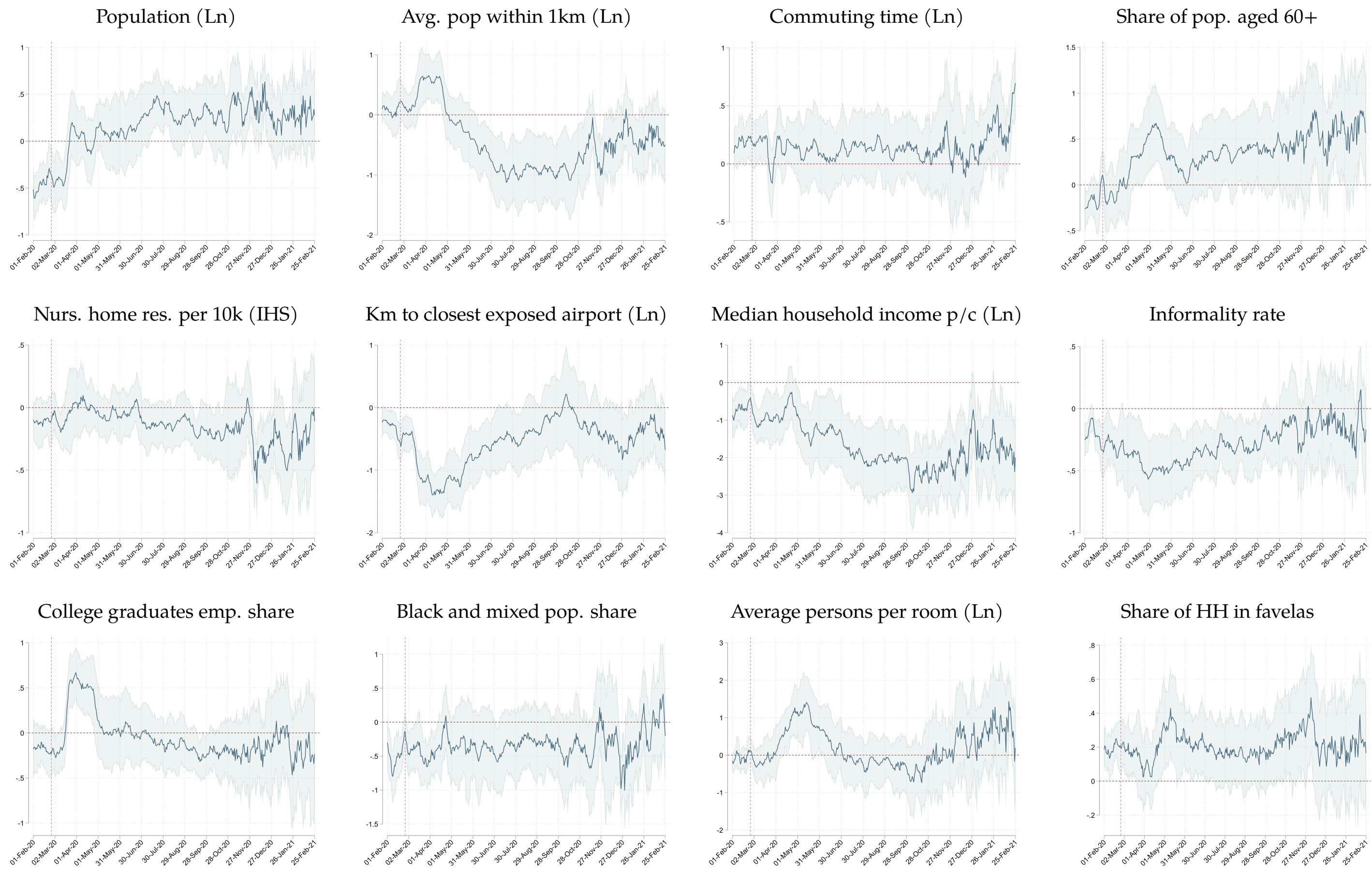
set of city-level, pre-pandemic regressors as before, and where the dependent variable is the seven-day moving average of InLoco's Social Isolation Index.

The mobility results are consistent with those for COVID-19 deaths in Section 4.2. In particular, I find that for city characteristics associated with larger local economic activity, the mobility effects track the deaths effects. Controlling for all other covariates, the propensity to stay at home started to drop before the peak of the first wave in high-density and high-income cities – as the deaths effects also escalated – and remained lower than the average throughout the period of study. Meanwhile, cities with a higher share of college graduates – workers who are more likely to be able to work from home, or to afford temporary job separations – had a higher propensity to stay at home for a short period during the first wave. In the weeks that followed, these cities experienced fewer deaths per capita than the average. However, the trend was short-lived, and, as the propensity to stay at home dropped again, highly educated cities became statistically indistinguishable from others in terms of both mobility and COVID-19 death rates. Overall, this evidence supports the interpretation that the economic dynamism of cities – which is in turn associated with higher mobility and human interactions – was a key driver of the local impact of COVID-19 in Brazil during the first year of the pandemic.

Cities whose socioeconomic and demographic characteristics made them more vulnerable to COVID-19 experienced higher death tolls – particularly during the first wave – *in spite of* their residents' greater propensity to stay at home. In this period, people were more likely to stay put in cities with a higher share of the population aged 60 or older, those with more persons per room on average, and those located closer to international airports. I also find this to be the case for cities with a high share of households living in favelas, and in this case the effect is sustained through most of the first year of the pandemic. In other words, holding other regressors constant, cities with large presence of favelas suffered disproportionately more COVID-19 deaths per capita during the first wave, even though their populations were less mobile than the populations in other cities. This is in line with Sheng et al. (2021), who find that the sharply higher infection rates in slums relative to non-slum areas in Mumbai, India, cannot be attributed to differences in compliance with government-imposed mobility restrictions.[13]

[13]Other variables, which had short, statistically significant associations with per capita deaths – including the average commuting time, the share of black and people of mixed races in the population, and the number of nursing home residents – had no significant association with mobility for most of the period of study. An exception to the pattern is the informality rate, which is associated with a lower propensity to stay at home for most of the period of study.

Figure 3: OLS Estimates of Correlations Between the Share of the Population Staying at Home and City Characteristics



Notes: This figure plots daily OLS estimates from multivariate regressions, for the period between April 1, 2020, and February 25, 2021, with data from 2,509 Brazilian cities. The dependent variable is the seven-day moving average of InLoco's Social Isolation Index. All regressors are standardized. All estimations include a constant and state fixed effects. The shaded area shows 95% confidence intervals constructed from robust standard errors clustered at the state level.

5.2 Policy Responses

People's reaction to the observed or anticipated threats posed by the pandemic could have been shaped by local containment policies. To explore this mechanism, I use the "containment intensity" measure derived from the Chauvin et al. (2021) data.

Relative to the preceding sections of the paper, my analysis of policies has important differences. First, the sample size is significantly smaller than in the analysis above. The policy data were collected for a sample of 501 municipalities that are representative of all municipalities in Brazil. Furthermore, the data cover a shorter span of time (March 2020 to October 2020). Second, the geographic units of observations are not directly comparable to those used in the previous sections. While I previously grouped municipalities that are part of the same commuting zone, and I treated them as part of a single entity, here I perform the policies analysis at the municipality level, which is the level at which policy decisions are made. To gauge how these differences may affect the results, I replicate the COVID-19 deaths regressions in Figure 2, and the regressions for mobility in Figure 3, using the 501 municipalities for which I have policy data (Appendix Figures B4 and B5, respectively). One might expect the results of municipality-level regressions and commuting-zone-level regressions to differ, given that COVID-19 outcomes in a given municipality can be affected by the policies and private responses of other municipalities in the same commuting zone. However, I find that, in the municipal sample, the connections between local characteristics and the evolution of both COVID-19 and mobility outcomes over time are very similar to those the main specifications – if less precisely measured.[14]

Figure 4 shows daily OLS estimates of equation 1 (with the same set of regressors as before, but computed at the municipality level), using as the dependent variable the number of days that the municipality had at least one of six possible containment policies in place over the prior 30 days. For a couple of regressors, I find an effect on the proclivity of municipal governments to implement containment policies at the very beginning of the pandemic. Holding all other regressors constant, municipalities whose populations include a larger share of those aged 60 or older were less likely to implement such policies, and municipalities with a larger shares of racial minorities were more likely to implement them. However, these correlations between pre-pandemic characteristics and local policy choices are short-lived, and only those of the population older than 60 are statistically significant. For most of the regressors the estimates are close to zero and / or not statistically significant for most of the observed period. Overall, these results suggest that the connections between

[14]Two exceptions are the estimates for the informality rate and the distance to the closest internationally connected airport. I do find, in the municipal sample, an effect on mobility similar to that in the main sample for these variables. However, I do not find a statistically significant effect on deaths over the period of analysis.

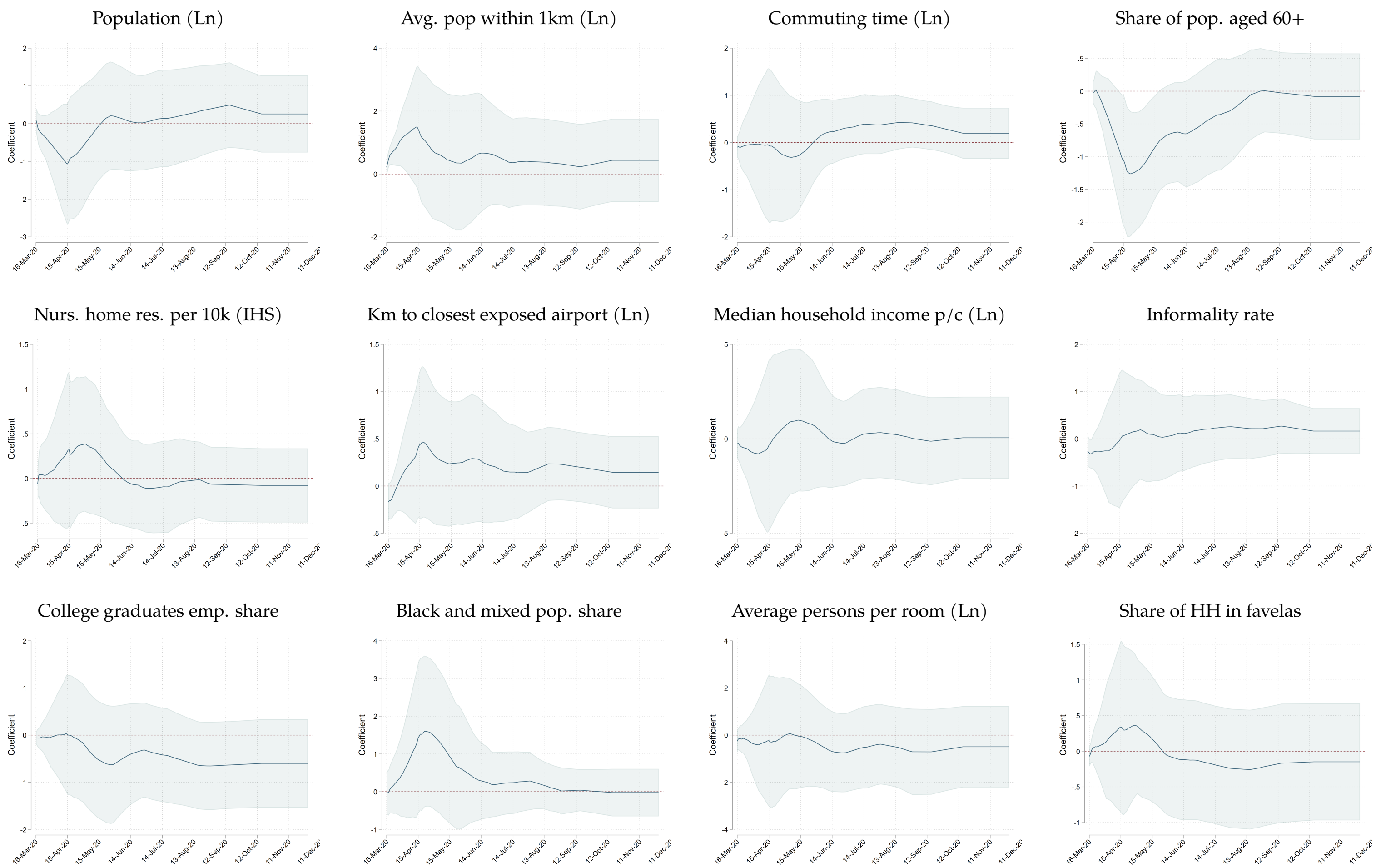
local pre-pandemic characteristics and either COVID-19 or mobility outcomes are not driven by endogenous policy responses.

6 Conclusions

In this paper, I study the correlates of the local impact of COVID-19 in Brazil. I find multiple similarities with equivalent US estimates, starting with the fact that the impact of pre-pandemic characteristics can vary over time, and that population density is a strong predictor of the local toll of the disease. But I also find important differences. While poorer cities in the US experienced more deaths per capita, the opposite is the case for Brazil. Cities with higher income levels in Brazil experienced, on average, more deaths per capita. This finding may be driven by the link between local economic dynamism and the exposure of the local population to the virus, as suggested by the fact that higher-income cities in Brazil also experienced higher mobility during the period of study. However, for other city characteristics, the connection with cases and deaths appears to be largely uncorrelated to mobility. For example, during the first wave the death toll was unusually large in cities with a large presence of favelas, and in cities with high residential crowding – in spite of the higher propensity of their people to stay at home. In addition, pre-pandemic characteristics are not strong predictors of local policy choices over the period of study. This shows that endogenous responses to the crisis can account for only part of the results. Further research is needed to better understand the mechanisms behind the connection between pre-pandemic characteristics and COVID-19 outcomes.

Overall, my work adds to a growing body of evidence showing that cities within the same countries can have large differences in terms of their vulnerability to the pandemic. These differences appear to be structural in nature, and not merely driven by variation in the time of arrival of the disease across locations. This supports the case for geographic prioritization and targeting of containment, vaccination, and recovery efforts.

Figure 4: OLS Estimates of Correlations Between Containment Policy Intensity and Municipality Characteristics



Notes: This figure plots daily OLS estimates from multivariate regressions, for the period between March 16, 2020, and December 11, 2020, with data from 501 Brazilian municipalities. The dependent variable is "containment policy intensity," defined as the number of days, over the immediately prior 30-day window in which the municipality had in place at least one of the following containment policies: workplace restrictions, commerce restrictions, travel restrictions, public transit restrictions, lockdowns, or curfews. All regressors are standardized. All regressions include a constant and state fixed effects. The shaded area shows 95% confidence intervals constructed from robust standard errors clustered at the state level.

References

Ahmad, K., Erqou, S., Shah, N., Nazir, U., Morrison, A. R., Choudhary, G., and Wu, W. C. (2020). Association of poor housing conditions with COVID-19 incidence and mortality across US counties. *PLoS ONE*, 15(11 November):1–13.

Alfaro, L., Becerra, O., and Eslava, M. (2020). EMEs and COVID-19: Shutting down in a world of informal and tiny firms. *NBER Working Paper 27360*.

Allcott, H., Boxell, L., Conway, J. C., Ferguson, B. A., Gentzkow, M., and Goldman, B. (2020). What Explains Temporal and Geographic Variation in the Early US Coronavirus Pandemic? (May):1–54.

Almagro, M. and Orane-Hutchinson, A. (2020). JUE insight: The determinants of the differential exposure to COVID-19 in New York city and their evolution over time. *Journal of Urban Economics*, (October):103293.

Alon, T. M., Kim, M., Lagakos, D., and Van Vuren, M. (2020). How should policy responses to the COVID-19 pandemic differ in the developing world? *NBER Working Paper 27273*.

Avery, C., Bossert, W., Clark, A., Ellison, G., and Ellison, S. F. (2020). Policy implications of models of the spread of coronavirus: Perspectives and opportunities for economists. *NBER Working Paper 27007*.

Baqui, P., Bica, I., Marra, V., Ercole, A., and van der Schaar, M. (2020). Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. *The Lancet Global Health*, 8(8):e1018–e1026.

Bellemare, M. F. and Wichman, C. J. (2020). Elasticities and the Inverse Hyperbolic Sine Transformation. *Oxford Bulletin of Economics and Statistics*, 82(1):50–61.

Benitez, J. A., Courtemanche, C. J., and Yelowitz, A. (2020). Racial and ethnic disparities in COVID-19: Evidence from six large cities. *NBER Working Paper 27592*.

Brotherhood, L., Cavalcanti, T., Da Mata, D., and Santos, C. (2020). Slums and Pandemics. *Mimeo*.

Brown, C., Ravallion, M., and Van de Walle, D. (2020). Can the world's poor protect themselves from the new coronavirus? *NBER Working Paper 27200*.

Brown, C. S. and Ravallion, M. (2020). Inequality and the coronavirus: Socioeconomic covariates of behavioral responses and viral outcomes across US counties. *NBER Working Paper 27549*.

Bryan, G., Glaeser, E., and Tsivanidis, N. (2020). Cities in the developing world. *Annual Review of Economics*, 12:273–297.

Buss, L. F., Prete, C. A., Abrahim, C. M., Mendrone, A., Salomon, T., De Almeida-Neto, C., França, R. F., Belotti, M. C., Carvalho, M. P., Costa, A. G., Crispim, M. A., Ferreira, S. C., Fraiji, N. A., Gurzenda, S., Whittaker, C., Kamaura, L. T., Takecian, P. L., Da Silva Peixoto, P., Oikawa, M. K., Nishiya, A. S., Rocha, V., Salles, N. A., De Souza Santos, A. A., Da Silva, M. A., Custer, B., Parag, K. V., Barral-Netto, M., Kraemer, M. U., Pereira, R. H., Pybus, O. G., Busch, M. P., Castro, M. C., Dye, C., Nascimento, V. H., Faria, N. R., and Sabino, E. C. (2021). Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. *Science*, 371(6526):288–292.

Busso, M., Camacho, J., Messina, J., and Montenegro, G. (2020). The challenge of protecting informal households during the COVID-19 pandemic: Evidence from Latin America. *IADB Discussion Paper 780*, (June).

Candido, D. S., Claro, I. M., de Jesus, J. G., Souza, W. M., Moreira, F. R., Dellicour, S., Mellan, T. A., du Plessis, L., Pereira, R. H., Sales, F. C., Manuli, E. R., Thézé, J., Almeida, L., Menezes, M. T., Voloch, C. M., Fumagalli, M. J., Coletti, T. M., da Silva, C. A., Ramundo, M. S., Amorim, M. R., Hoeltgebaum, H. H., Mishra, S., Gill, M. S., Carvalho, L. M., Buss, L. F., Prete, C. A., Ashworth, J., Nakaya, H. I., Peixoto, P. S., Brady, O. J., Nicholls, S. M., Tanuri, A., Rossi, Á. D., Braga, C. K., Gerber, A. L., de C Guimarães, A. P., Gaburo, N., Alencar, C. S., Ferreira, A. C., Lima, C. X., Levi, J. E., Granato, C., Ferreira, G. M., Francisco, R. S., Granja, F., Garcia, M. T., Moretti, M. L., Perroud, M. W., Castiñeiras, T. M., Lazari, C. S., Hill, S. C., de Souza Santos, A. A., Simeoni, C. L., Forato, J., Sposito, A. C., Schreiber, A. Z., Santos, M. N., de Sá, C. Z., Souza, R. P., Resende-Moreira, L. C., Teixeira, M. M., Hubner, J., Leme, P. A., Moreira, R. G., Nogueira, M. L., Ferguson, N. M., Costa, S. F., Proenca-Modena, J. L., Vasconcelos, A. T. R., Bhatt, S., Lemey, P., Wu, C. H., Rambaut, A., Loman, N. J., Aguiar, R. S., Pybus, O. G., Sabino, E. C., and Faria, N. R. (2020). Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science*, 369(6508):1255–1260.

Carozzi, F., Provenzano, S., and Roth, S. (2020). Urban Density and COVID-19. *IZA Discussion Paper No. 13440*, (13440).

Chauvin, J. P. (2021). Cities and Public Health in Latin America. *Unpublished manuscript, Inter-American Development Bank*.

Chauvin, J. P., Fowler, A., and L, N. H. (2020). The younger age profile of COVID-19 deaths in developing countries . *IADB Working Paper*.

Chauvin, J. P., Glaeser, E., Ma, Y., and Tobio, K. (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics*, 98:17–49.

Chauvin, J. P., Glaeser, E. L., and Kestelman, S. (2021). Impact of Subnational Responses to a National Crisis: Evidence from Brazil. *Mimeo*.

Cho, S.-W. (2020). Quantifying the impact of nonpharmaceutical interventions during the COVID-19 outbreak: The case of Sweden. *Econometrics Journal*, 23(3):323–344.

Cintra, H. P. and Fontinele, F. N. (2020). Estimative of real number of infections by COVID-19 in Brazil and possible scenarios. *Infectious Disease Modelling*, 5:720–736.

Clark, A., Jit, M., Warren-Gash, C., Guthrie, B., Wang, H. H., Mercer, S. W., Sanderson, C., McKee, M., Troeger, C., Ong, K. L., Checchi, F., Perel, P., Joseph, S., Gibbs, H. P., Banerjee, A., and Eggo, R. M. (2020). Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *The Lancet. Global health*, (20):1–15.

DATASUS, M. d. S. (2021). Sistema de Informacao Hospitalar/SUS. *Available at http://tabnet.datasus.gov.br*.

De la Roca, J. and Puga, D. (2017). Learning by Working in Big Cities. *Review of Economic Studies*, 84(1):106–142.

Desmet, K. and Wacziarg, R. (2021). JUE insight: Understanding Spatial Variation in COVID-19 across the United States. *Journal of Urban Economics*.

Fang, H., Wang, L., and Yang, Y. (2020). Human mobility restrictions and the spread of the Novel Coronavirus (2019-nCoV) in China. *Journal of Public Economics*, 191:104272.

Glaeser, E. L. (2009). The death and life of cities. *Making Cities Work: Prospects and Policies for Urban America*, pages 22–62.

Glaeser, E. L., Gorback, C., and Redding, S. J. (2020). JUE insight: How much does COVID-19 increase with mobility? Evidence from New York and four other U.S. cities. *Journal of Urban Economics*, (October).

Goldberg, P. K. and Reed, T. (2020). The Effects of the Coronavirus Pandemic in Emerging Market and Developing Economies: An Optimistic Preliminary Account. *Brookings Papers on Economic Activity*, 2020(2):161–235.

Hallal, P. C., Hartwig, F. P., Horta, B. L., Silveira, M. F., Struchiner, C. J., Vidaletti, L. P., Neumann, N. A., Pellanda, L. C., Dellagostin, O. A., Burattini, M. N., Victora, G. D., Menezes, A. M., Barros, F. C., Barros, A. J., and Victora, C. G. (2020). SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. *The Lancet Global Health*, 8(11):e1390–e1398.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology*, 34(3):623–642.

Harris, J. E. (2020). The subways seeded the massive coronavirus epidemic in New York city. *NBER Working Paper 27021*.

Hausmann, R. and Schetter, U. (2020). Horrible trade-offs in a pandemic: Lockdowns, transfers, fiscal space, and compliance. *CID Faculty Working Paper No. 382, Harvard University*.

Henderson, V. (2002). Urbanization in developing countries. *World Bank Research Observer*, 17(1):89–112.

IBGE (2012). Censo Demografico 2010: Microdados da Amostra. *Available at https://www.ibge.gov.br/estatisticas/sociais/saude/9662-censo-demografico-2010.html*.

IBGE (2016). *Arranjos Populacionais e Concentrações Urbanas do Brasil*. IBGE, Rio de Janeiro, second edition.

IBGE (2020). Aglomerados subnormais 2019: classificação preliminar e informações de saúde para o enfrentamento à COVID-19. *Available at https://www.ibge.gov.br/geociencias/organizacao-do-territorio/tipologias-do-territorio/15788-aglomerados-subnormais.html*.

InLoco (2021). Índice de Isolamento Social, obtained at https://inloco.com.br.

Justen, Á. (2021). Brasil.io: Painel COVID-19 (accessed on July 13, 2020).

Knittel, C. R. and Ozaltun, B. (2020). What does and does not correlate with COVID-19 death rates. *NBER Working Paper 27391*.

Levin, A. T., Cochran, K. B., and Walsh, S. P. (2020). Assessing the age specificity of infection fatality rates for COVID-19: Meta-analysis & public policy implications. *NBER Working Paper 27597*.

Marinho, F., Torrens, A., Teixeira, R., França, E., Nogales, A. M., Xavier, D., and Fujiwara, T. (2020). Excess Mortality in Brazil: Detailed Description of Trends in Mortality During the COVID-19 Pandemic. *Vital Strategies, Conselho Nacional de Secretários de Saúde (CONASS), and Associação Nacional dos Registradores de Pessoas Naturais (Arpen-Brasil), available at: https://www.conass.org.br/indicadores-de-obitos-por-causas-naturais/*.

Maugeri, A., Barchitta, M., Battiato, S., and Agodi, A. (2020). Estimation of Unreported Novel Coronavirus (SARS-CoV-2) Infections from Reported Deaths: A Susceptible–Exposed–Infectious–Recovered–Dead Model. *Journal of Clinical Medicine*, 9(5):1350.

McLaren, J. (2020). Racial disparity in COVID-19 deaths: Seeking economic roots with census data. *NBER Working Paper 27407*.

Metro de Sao Paulo (2017). Pesquisa Origem e Destino da Regiao Metropolitana de Sao Paulo. *Data*.

O'Driscoll, M., Ribeiro Dos Santos, G., Wang, L., Cummings, D. A., Azman, A. S., Paireau, J., Fontanet, A., Cauchemez, S., and Salje, H. (2021). Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature*, 590(7844):140–145.

Roser, M., Ritchie, H., Ortiz-Ospina, E., and Hasell, J. (2021). Coronavirus Disease (COVID-19). *Our World in Data. Available at https://ourworldindata.org/coronavirus*.

Schiavina, M., Freire, S., and MacManus, K. (2019). GHS population grid multitemporal (1975, 1990, 2000, 2015) R2019A. *European Commission, Joint Research Centre (JRC)*.

Sheng, J., Malani, A., Goel, A., and Botla, P. (2021). JUE Insight: Does Mobility Explain Why Slums Were Hit Harder by COVID-19 in Mumbai, India? *Journal of Urban Economics*, page 103357.

Wiemers, E. E., Abrahams, S., AlFakhri, M., Hotz, V. J., Schoeni, R. F., and Seltzer, J. A. (2020). Disparities in vulnerability to severe complications from COVID-19 in the United States. *NBER Working Paper 27294*.

A Data Appendix

Table A1: Outcome Variables Description

Variable	Sources	Description / comments
COVID-19 deaths	Brasil.io	Number of confirmed COVID-19 deaths reported by state health secretaries at the municipality level. In cities that comprise more than one municipality the death counts of all municipalities in the city are added.
COVID-19 cases	Brasil.io	Number of confirmed COVID-19 cases reported by state health secretaries at the municipality level. In cities that comprise more than one municipality the case counts of all municipalities in the city are added.
Share staying at home	InLoco	InLoco's Social Isolation Index, computed at the municipality-date level, and defined as the share of phone users in the municipality that stayed at home on that date. Staying at home is defined as remaining within 450 meters of the location identified as the residence. The original data source is anonymized location data from over 60 million cell phones. In cities that comprise more than one municipality I use the population-weighted average of the index across all municipalities in the city. The series is smoothed using a seven-days rolling average.
Containment policy intensity	Chauvin et al. (2021)	Number of days, over the immediately prior 30-days window, in which the municipality had in place at least one of the following containment policies: workplace restrictions, commerce restrictions, travel restrictions, public transit restrictions, lockdowns, and curfews. The original sources are local legislation documents obtained from public online records, from which the text of individual articles was extracted and analyzed to attribute policies to municipalities and dates.

Notes: The Brasil.io data can be obtained from https://brasil.io/dataset/covid19/caso_full/. The InLoco data are not publicly available. Information on gaining access can be obtained at https://inloco.com.br. The containment policy intensity data from Chauvin et al. (2021) will be made publicly available upon publication.

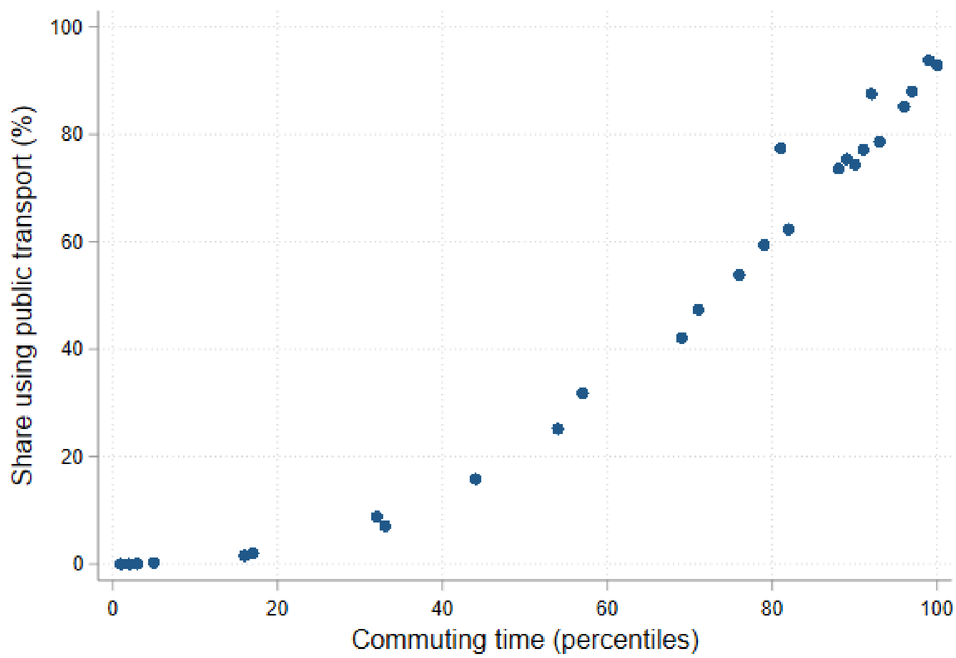
Table A2: Pre-pandemic Characteristics – Variables Description

Variable	Sources	Description / comments
Population	Census 2010 (IBGE)	Projected number of persons living in the city in 2019.
Population density	Global Human Settlement Layer	Total population living within 1 km of the average inhabitant of the city. It is calculated using 250m x 250m cells. First, each cell is attributed a neighborhood population, defined as the sum of the population of the cell and of all the cells within a 1km radius. Then the city-level value is computed taking the weighted average of the neighborhood populations across all cells within the city boundaries, where the weights are the cell populations.
Commuting time	Census 2010 (IBGE)	Average commuting time of workers, estimated based on midpoints of the time intervals reported in the microdata.
Share of 60+ in population	Census 2010 (IBGE)	Share of individuals aged 60 or older in the projeted 2019 population.
Share of nursing home residents per 10k pop	Census 2010 (IBGE)	Share of population living in nursing homes which is aged 65 or more per 10,000 population
Km to closest airport connecting to hot spots	National Civil Aviation Agency (ANAC)	Distance from city centroid to the nearest airport having at least a flight from USA, UK, FR, SP, IT.
Median income per capita	Census 2010 (IBGE)	City-level median of the household income per capita, calculated dividing total household income from all sources by the number of people in the household.
Informality rate	Census 2010 (IBGE)	An informal worker is someone who during the period of reference worked without a signed work card, or was self-employed. The informality rate is the share of informal workers in the labor force.
Share of college in employment	Census 2010 (IBGE)	Share of workers with at least college degree in the employed population.
Share of black and mixed in population	Census 2010 (IBGE)	Share of self-identified black and mulatto individuals in total population.
Average persons per room	Census 2010 (IBGE)	Cross-households average of the number of persons per room.
Share of households in favelas	IBGE (2020)	2019 projection of the share of household located in "abnormal agglomerations" by the census definition.
Patients with at least one precondition in the population	Health ministry (DATASUS)	Number of patients with at least one of the morbidities associated with severe COVID-19 complications identified in Clark et al. (2020).
Number of doctors	Health ministry (DATASUS)	Number of doctors in local hospitals in February 2020.
ICU beds	Health ministry (DATASUS)	Number of beds in Intensive Care Units in local hospitals in February 2020.
Ventilators	Health ministry (DATASUS)	Number of ventilators in local hospitals in February 2020.
Maximum yearly temperature	University of East Anglia Climatic Research Unit (CRU)	Highest registered temperature between 1900-2019, interpolated to the city level.
Average yearly precipitation	University of East Anglia Climatic Research Unit (CRU)	Average yearly precipitation 1900-2019, interpolated to the city level.
Distance to Sao Paulo	Census 2010 (IBGE)	Distance (in Km) of the shortest path between the city's centroid and Sao Paulo's centroid.

Notes: All city characteristics data are publicly available. The 2010 census data is available at https://downloads.ibge.gov.br.The DATASUS microdata can be obtained from http://www2.datasus.gov.br/DATASUS/index.php. The CRU data is available at http://www.cru.uea.ac.uk/data. The Global Human Settlement Layer is available at https://ghsl.jrc.ec.europa.eu/download.php. Lastly, the IBGE estimates of the 2019 share of households in favelas is available at https://www.ibge.gov.br/geociencias.

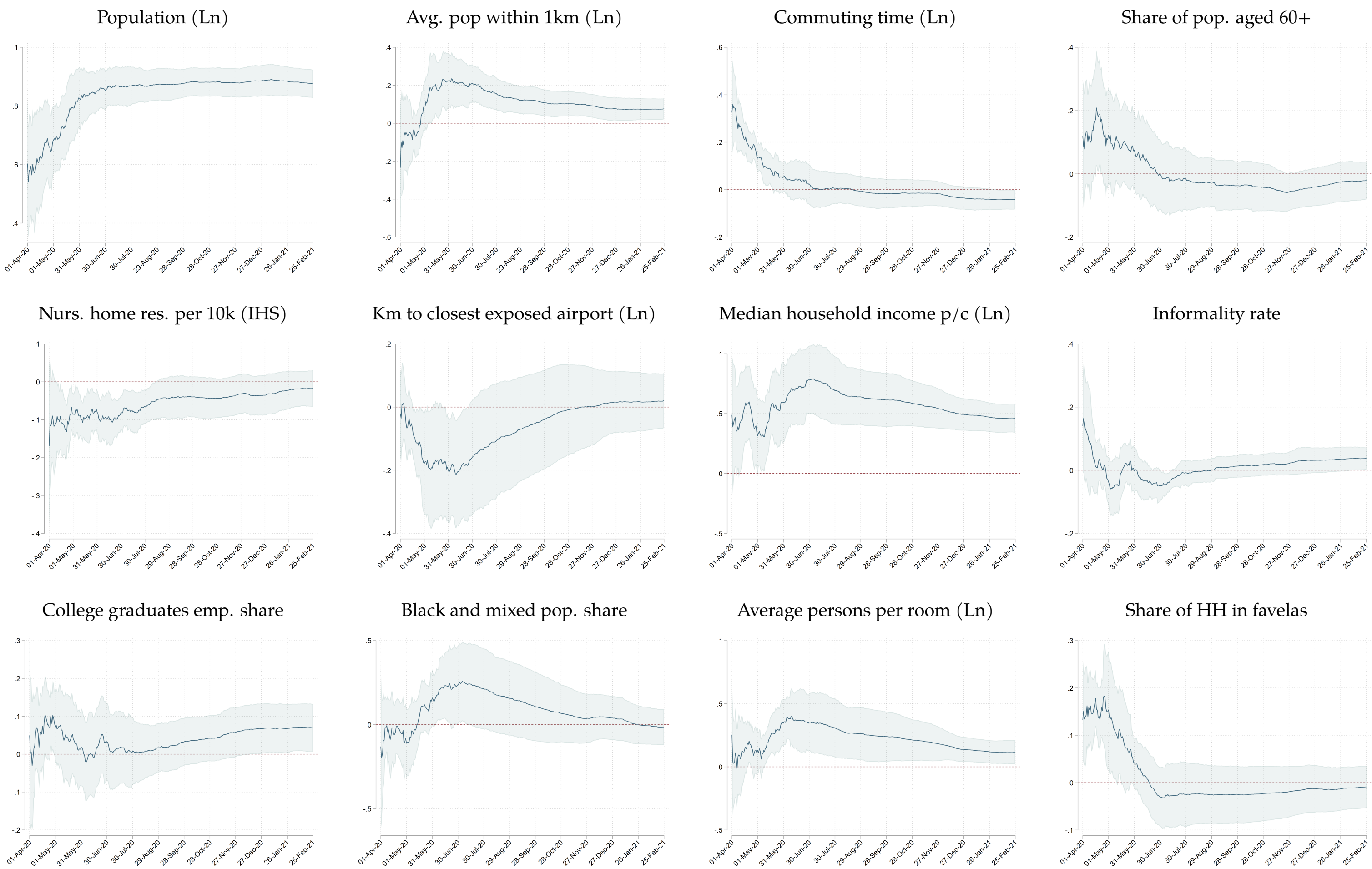
B Additional Figures and Tables

Figure B1: Length of Commute vs. Share of Commuters Using Public Transport, São Paulo, 2017



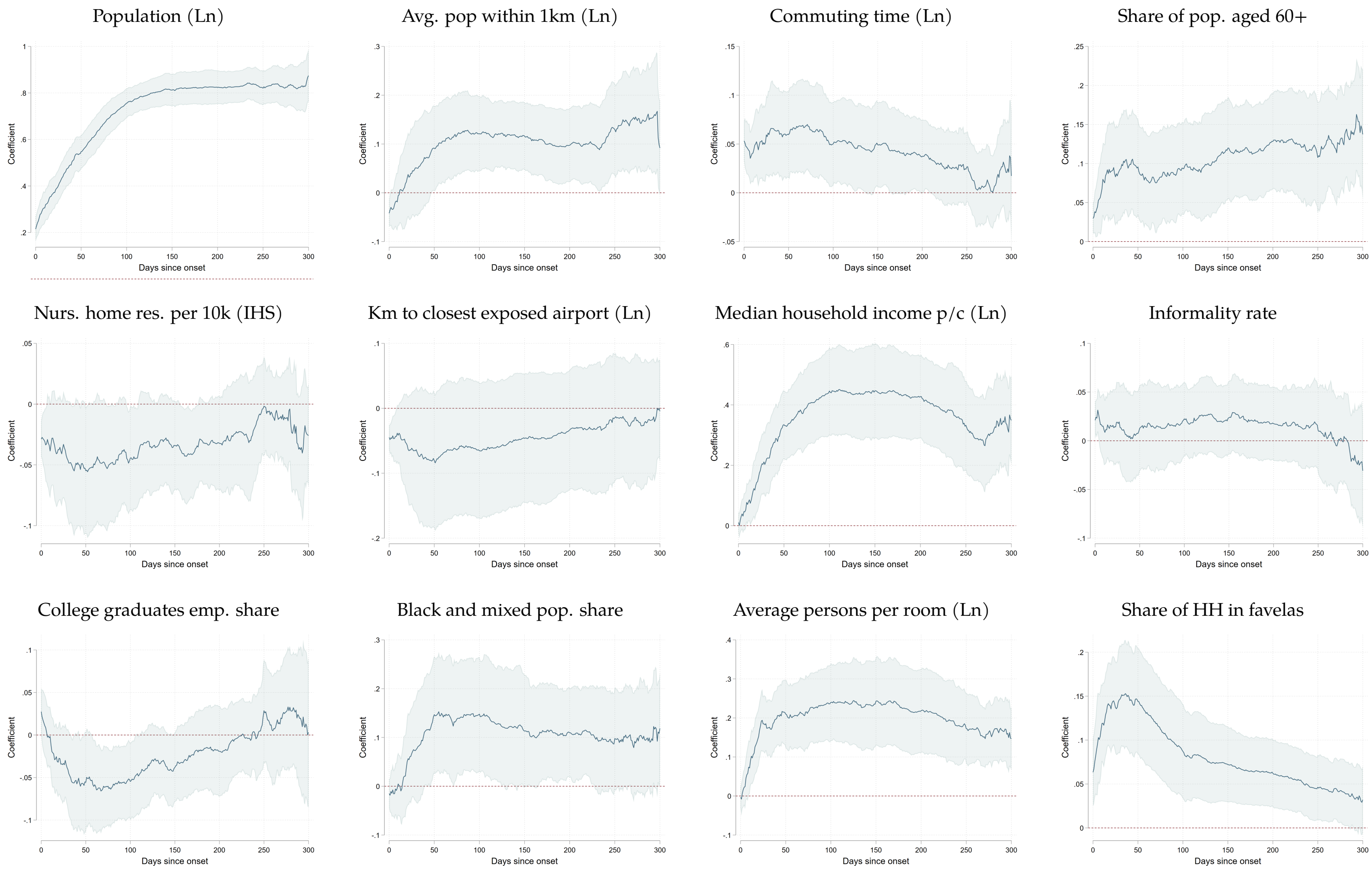
Notes: Author's calculations with data from the 2017 Origin-Destination survey of São Paulo.

Figure B2: OLS Estimates of Correlations between Cumulative COVID-19 Cases and City Characteristics



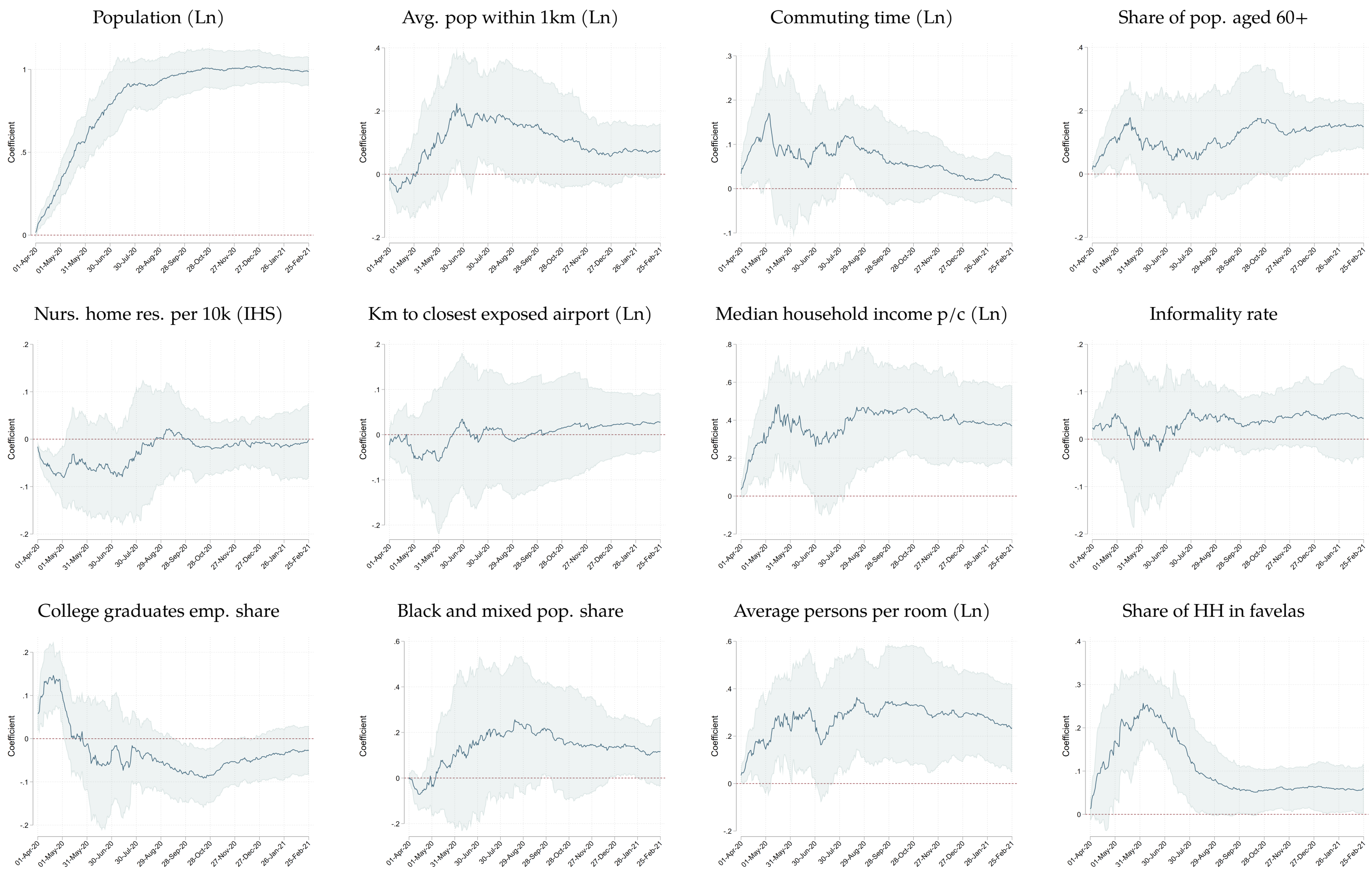
Notes: This figure plots daily OLS estimates from multivariate regressions, for the period between April 1, 2020 and February 25, 2021, with data of 2,509 Brazilian cities. All regressors are standardized. All estimations include a constant and state fixed effects. The shaded area shows 95% confidence intervals constructed from robust standard errors clustered at the state level.

Figure B3: OLS Estimates of Correlations between Cumulative COVID-19 Deaths and City Characteristics, by Days since the Local Onset of the Epidemic



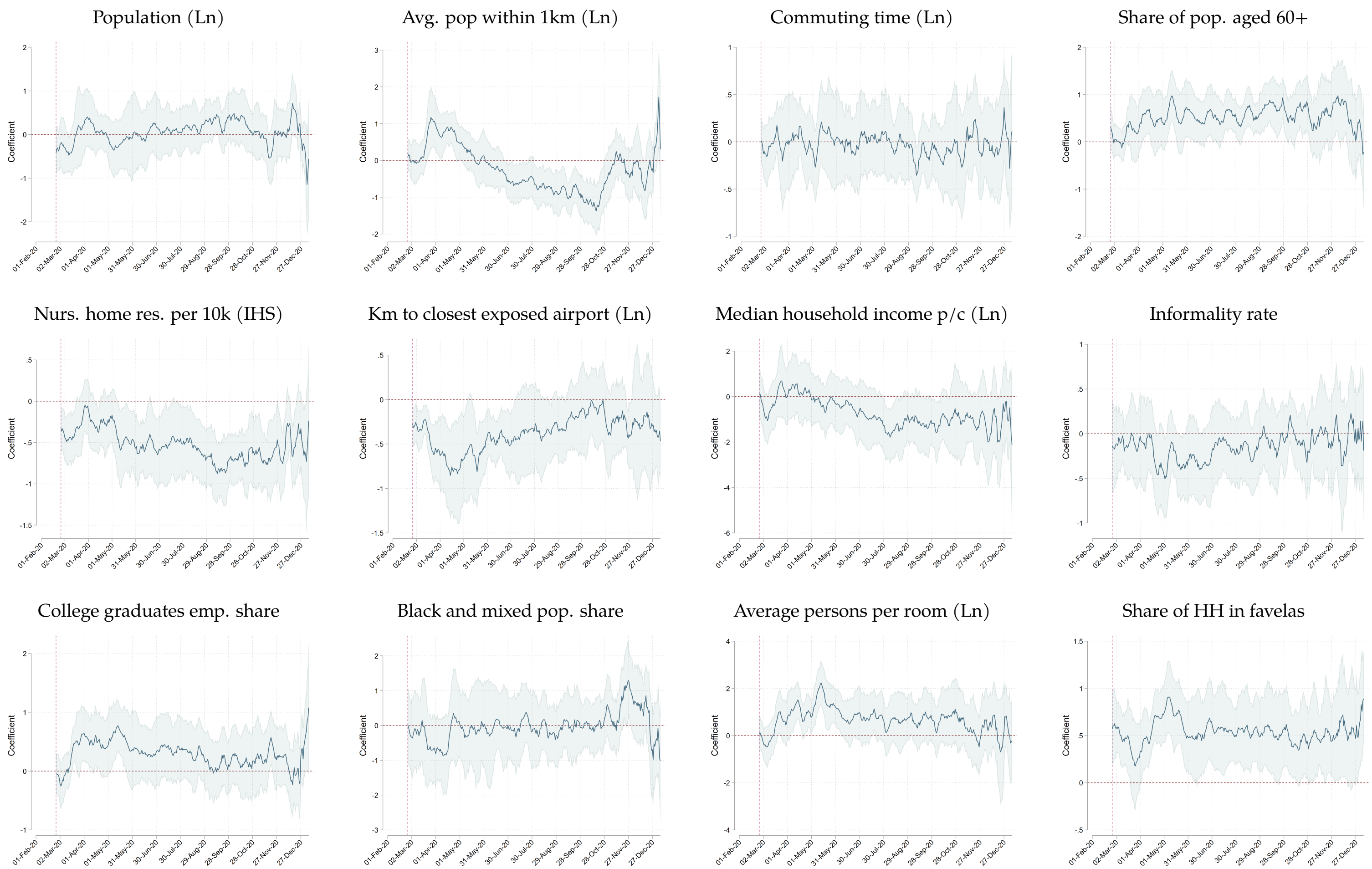
Notes: This figure plots OLS estimates from multivariate regressions, with data of 2,509 Brazilian cities. All regressors are standardized. For each city, onset is defined as the day at which the number of deaths per 100,000 population is equal or greater than 0.5. The dependent variable is the logarithm of cumulative deaths. All estimations include a constant and state fixed effects. The shaded area shows 95% confidence intervals constructed from robust standard errors clustered at the state level.

Figure B4: OLS estimates of correlations between cumulative COVID-19 deaths and municipality characteristics



Notes: This figure plots daily OLS estimates from multivariate regressions, for the period between April 1, 2020 and February 25, 2021, with data from 501 Brazilian municipalities. All regressors are standardized. All estimations include a constant and state fixed effects. The shaded area shows 95% confidence intervals constructed from robust standard errors clustered at the state level.

Figure B5: OLS estimates of correlations between share staying at home and municipality characteristics



Notes: This figure plots daily OLS estimates from multivariate regressions, for the period between April 1, 2020 and February 25, 2021, with data of 501 Brazilian municipalities. All regressors are standardized. All estimations include a constant and state fixed effects. The shaded area shows 95% confidence intervals constructed from robust standard errors clustered at the state level.

Table B1: Descriptive Statistics of Pre-Pandemic City Characteristics Employed in the Analysis

	Mean	Median	SD	Min	Max
Panel A: Main characteristics					
Population (Ln)	10.218	9.978	0.908	9.211	16.892
Avg. population within 10km (Ln)	4.863	4.828	1.28	0.18	9.833
Commuting time (Ln)	3.142	3.132	0.173	2.38	3.911
Share of people aged 60+	0.128	0.128	0.038	0.033	0.27
Nursing home residents per 10k pop (IHS)	1.317	0	1.644	0	5.177
Km to closest airport connecting to hot spots (Ln)	6.064	6.165	0.832	0	8.044
Median household income p/c (Ln)	5.792	5.756	0.473	4.069	6.833
Informality rate	0.16	0.155	0.051	0.043	0.444
College graduates employment share	0.077	0.072	0.036	0.005	0.219
Black and mixed population share	0.56	0.616	0.222	0.041	0.933
Average persons per room (Ln)	0.529	0.501	0.123	0.36	1.552
Share of households located in favelas	0.011	0	0.044	0	0.55
Panel B: Public health					
Share of patients with at least 1 precondition	0.14	0.134	0.054	0.027	0.669
Number of doctors (IHS)	7.775	7.744	0.331	5.769	9.546
ICU beds (IHS)	1.258	0	1.733	0	5.982
Ventilators (IHS)	1.957	2.398	1.732	0	5.882
Panel C: Weather and geography					
Maximum yearly temperature (F)	85.174	85.935	4.824	71.503	93.608
Average yearly precipitation	3.756	3.752	1.495	0.844	9.931
Distance to Sao Paulo	6.932	7.124	0.822	0	8.167

Notes: Sample restricted to cities with projected populations of at least 10,000 in 2019.

Table B2: Descriptive Statistics of Outcome Variables

	Mean	Median	SD	Min	Max
Panel A: City-level outcome variables					
IHS cases	5.303	5.489	2.097	0	14.353
IHS deaths	1.78	1.444	1.575	0	11.067
Log cases	4.632	4.812	2.084	0	13.66
Log deaths	1.674	1.386	1.404	0	10.374
Share staying at home	39.015	38.989	6.047	4.55	81.82
Panel B: Municipality-level outcome variables					
Number of days with any containment policy in the last 30 days	25.396	30	10.032	0	30
IHS deaths	2.612	2.776	1.922	0	9.108
Share staying at home	39.224	38.799	4.916	18.717	80.95

Notes: Panel A uses data restricted to cities with projected populations of at least 10,000 in 2019. Descriptive statistics correspond to the city-day level data. Panel B uses data from 501 representative municipalities at the municipality-day level.

Table B3: City Characteristics and COVID-19 Toll as of February 25, 2021. Logarithms Specification.

	(1) Log cases	(2) Log cases	(3) Log deaths	(4) Log deaths
Panel A: Baseline regressors				
Population (Ln)	0.90***	0.87***	0.88***	0.85***
	(0.02)	(0.02)	(0.02)	(0.02)
Avg. population within 10km (Ln)	0.07**	0.08***	0.13***	0.14***
	(0.03)	(0.03)	(0.03)	(0.03)
Commuting time (Ln)	-0.05**	-0.04**	0.05**	0.03*
	(0.02)	(0.02)	(0.02)	(0.02)
Share of people aged 60+	-0.03	-0.02	0.08***	0.14***
	(0.03)	(0.03)	(0.02)	(0.03)
Nursing home residents per 10k pop (IHS)	-0.01	-0.02	-0.01	-0.01
	(0.02)	(0.02)	(0.01)	(0.01)
Km to closest airport connecting to hot spots (Ln)	0.03	0.02	-0.04	-0.04
	(0.04)	(0.04)	(0.04)	(0.04)
Median household income p/c (Ln)	0.42***	0.46***	0.25***	0.40***
	(0.04)	(0.06)	(0.06)	(0.07)
Panel B: Full specification				
Informality rate		0.04**		0.02
		(0.02)		(0.02)
College graduates employment share		0.07**		0.01
		(0.03)		(0.03)
Black and mixed population share		-0.01		0.09
		(0.05)		(0.06)
Average persons per room (Ln)		0.12**		0.21***
		(0.05)		(0.05)
Share of households located in favelas		-0.01		0.03*
		(0.02)		(0.02)
Observations	2,509	2,509	2,509	2,509
R^2	0.80	0.80	0.80	0.81
R^2 of the per-capita specification	0.38	0.39	0.31	0.33

Notes: OLS regressions at the city level. Sample restricted to cities with projected populations of at least 10,000 in 2019. All regressors are standardized. The R^2 of the per-capita specification comes from regressions in which the outcomes are directly expressed in per capita terms, and population is excluded from the regressors set. All regression include a constant and state fixed effects. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B4: City Characteristics and COVID-19 Toll as of February 25, 2021: Specification without State Fixed Effects.

	(1) IHS cases	(2) IHS cases	(3) IHS deaths	(4) IHS deaths
Panel A: Baseline regressors				
Population (Ln)	0.96***	0.91***	0.97***	0.92***
	(0.03)	(0.03)	(0.05)	(0.05)
Avg. population within 10km (Ln)	0.07	0.07*	0.12*	0.13**
	(0.05)	(0.04)	(0.07)	(0.06)
Commuting time (Ln)	-0.12***	-0.12***	-0.04	-0.06
	(0.03)	(0.03)	(0.04)	(0.04)
Share of people aged 60+	-0.12***	-0.04	-0.04	0.07*
	(0.04)	(0.03)	(0.04)	(0.03)
Nursing home residents per 10k pop (IHS)	-0.07*	-0.08**	-0.05	-0.06**
	(0.04)	(0.03)	(0.03)	(0.03)
Km to closest airport connecting to hot spots (Ln)	0.05	0.04	-0.02	-0.04
	(0.04)	(0.04)	(0.05)	(0.05)
Median household income p/c (Ln)	0.26***	0.32***	0.16**	0.29***
	(0.05)	(0.07)	(0.07)	(0.09)
Panel B: Full specification				
Informality rate		0.05**		0.04
		(0.02)		(0.03)
College graduates employment share		0.07		0.06
		(0.04)		(0.05)
Black and mixed population share		0.02		0.10
		(0.05)		(0.07)
Average persons per room (Ln)		0.17***		0.19***
		(0.06)		(0.06)
Share of households located in favelas		0.06**		0.09***
		(0.03)		(0.03)
Observations	2,509	2,509	2,509	2,509
R^2	0.75	0.76	0.71	0.73
R^2 of the per-capita specification	0.22	0.26	0.13	0.32

Notes: OLS regressions at the city level. Sample restricted to cities with projected populations of at least 10,000 in 2019. All regressors are standardized. The R^2 of the per-capita specification comes from regressions in which the outcomes are directly expressed in per capita terms, and population is excluded from the regressors set. All regression include a constant. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B5: City Characteristics and COVID-19 Toll as of February 25, 2021. Specifications With Additional Controls.

	(1) IHS cases	(2) IHS cases	(3) IHS cases	(4) IHS deaths	(5) IHS deaths	(6) IHS deaths
Panel A: Main specification						
Population (Ln)	0.88***	0.88***	0.90***	0.89***	0.89***	0.93***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Avg. population within 10km (Ln)	0.08***	0.07**	0.06**	0.16***	0.16***	0.14***
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Commuting time (Ln)	-0.04**	-0.04**	-0.04**	0.03	0.03	0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Share of people aged 60+	-0.02	-0.01	-0.01	0.15***	0.15***	0.15***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Nursing home residents per 10k pop (IHS)	-0.02	-0.02	-0.02	-0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Km to closest airport connecting to hot spots (Ln)	0.02	0.01	-0.00	-0.03	-0.04	-0.03
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)
Median household income p/c (Ln)	0.46***	0.43***	0.44***	0.45***	0.43***	0.37***
	(0.06)	(0.05)	(0.05)	(0.08)	(0.08)	(0.04)
Informality rate	0.04**	0.04**	0.04**	0.02	0.02	0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
College graduates employment share	0.07**	0.04	0.04	-0.00	-0.01	-0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Black and mixed population share	-0.01	-0.01	-0.04	0.10	0.10	0.07
	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.07)
Average persons per room (Ln)	0.12**	0.12**	0.11**	0.23***	0.24***	0.18***
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.04)
Share of households located in favelas	-0.01	-0.01	-0.02	0.02	0.02	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Panel B: Public health controls						
Share of patients with at least 1 precondition		-0.01	-0.01		0.04	0.03
		(0.02)	(0.02)		(0.02)	(0.02)
Number of doctors		0.09***	0.09***		0.05*	0.04**
		(0.02)	(0.02)		(0.02)	(0.02)
ICU beds		-0.05***	-0.05***		-0.02	-0.03
		(0.02)	(0.02)		(0.02)	(0.02)
Ventilators		0.03***	0.03**		0.00	0.00
		(0.01)	(0.01)		(0.01)	(0.01)
Panel C: Weather and geography controls						
Maximum yearly temperature			0.02			0.06
			(0.06)			(0.06)
Average yearly precipitation			0.03			0.20**
			(0.08)			(0.09)
Distance to Sao Paulo			0.08			0.05
			(0.08)			(0.06)
Observations	2,509	2,509	2,458	2,509	2,509	2,458
R^2	0.81	0.81	0.81	0.78	0.78	0.79
R^2 of the per-capita specification	0.39	0.41	0.42	0.32	0.32	0.36

Notes: OLS regressions at the city level. Sample restricted to cities with projected populations of at least 10,000 in 2019. All regressors are standardized. The R^2 of the per capita specification comes from regressions in which the outcomes are directly expressed in per capita terms, and population is excluded from the regressors set. All regression include a constant and state fixed effects. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B6: City Characteristics and COVID-19 Toll as of February 25, 2021 for Cities of Population 100k and Larger.

	(1) IHS cases	(2) IHS cases	(3) IHS deaths	(4) IHS deaths
Panel A: Baseline regressors				
Population (Ln)	0.96***	0.96***	1.13***	1.14***
	(0.06)	(0.06)	(0.11)	(0.11)
Avg. population within 10km (Ln)	0.01	-0.01	0.11	0.14
	(0.15)	(0.14)	(0.23)	(0.19)
Commuting time (Ln)	-0.12**	-0.15***	-0.22***	-0.23***
	(0.05)	(0.04)	(0.07)	(0.06)
Share of people aged 60+	0.11*	0.18***	0.22**	0.34***
	(0.06)	(0.06)	(0.09)	(0.11)
Nursing home residents per 10k pop (IHS)	-0.29**	-0.28**	-0.16*	-0.19***
	(0.11)	(0.12)	(0.08)	(0.07)
Km to closest airport connecting to hot spots (Ln)	0.00	0.01	-0.06	0.02
	(0.03)	(0.03)	(0.05)	(0.05)
Median household income p/c (Ln)	0.31**	0.51**	-0.10	0.31
	(0.14)	(0.20)	(0.25)	(0.24)
Panel B: Full specification				
Informality rate		-0.12		0.05
		(0.17)		(0.09)
College graduates employment share		-0.05		-0.13*
		(0.05)		(0.07)
Black and mixed population share		0.16		-0.30**
		(0.20)		(0.12)
Average persons per room (Ln)		0.19		0.44***
		(0.12)		(0.11)
Share of households located in favelas		0.01		0.07**
		(0.02)		(0.03)
Observations	115	115	115	115
R^2	0.96	0.96	0.96	0.97
R^2 of the per-capita specification	0.65	0.67	0.50	0.61

Notes: OLS regressions at the city level. Sample restricted to cities with projected populations of at least 100,000 in 2019. All regressors are standardized. The R^2 of the per capita specification comes from regressions in which the outcomes are directly expressed in per capita terms, and population is excluded from the regressors set. All regression include a constant and state fixed effects. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.