TÉLÉCOM PARIS

MODS207 - Project in Applied Economics

The Effect of Telecommunication Infrastructure on COVID-19 in Brazil

Professor Ulrich Laitenberger

José Lucas Barretto
Lucas Miguel Celinga Agrizzi

Paris, June 20th, 2021

Summary

1. Introduction and Background

At the end of 2019, the world watched the insurgence of the coronavirus and the start of a global pandemic. This event changed the dynamic of countries all over the globe, becoming one of the biggest challenges that humanity has faced in this century. Worldwide measures were taken to slow down the spread of the virus, the main one consisting of limiting social interaction between humans, through lockdowns and social distancing measures. However, these measures directly impact the economy, as the number of interactions is lower and this discourages all commercial activities. For this reason, most countries have faced a controversial choice between maintaining economic activities or enforcing strict public health measures. Besides governmental restrictions, other regional factors are correlated to the spread of the virus - and hence, to the number of COVID cases and deaths in a given region - such as access to information, access to basic sanitation and health facilities, and the types of jobs that inhabitants perform.

In Brazil, this scenario is not different. In a country with continental proportions and with varying levels of infrastructure over its area, the impact of access to information and job qualification can have a high influence on how the numbers of cases and deaths behave. Some structural problems can be highlighted during the pandemic as shown in [Peterson K. Ozili, 2020]. For instance, the migration of on-site work to other modalities such as 'work from home' and 'hybrid', is not a choice available to every employee - blue-collar workers in essential sectors had to remain working on-site even during lockdowns. The possibility of this migration for remote jobs can be correlated with the worker's level of specialty and access to technological infrastructure [*Krenz, Astrid 2020*] [*Xiaoqun Zhang 2020*]. This last aspect is also tightly correlated to the population's access to information - which may help to combat the spread of the virus since many helpful resources such as WHO sanitary measures, government recommendations, and guidance on the disease are available on most information sources such as television and internet.

This project aims to measure the impact of access to information, which is interpreted as the population's access to telecommunications infrastructure, on the number of COVID-19 deaths per capita in different Brazilian municipalities. This infrastructure is measured through the number of per capita accesses to fixed and mobile broadband, cable television, and fixed telephony.

2. Data Collection

This stage of the project is responsible for collecting all the data used throughout the project. We divided this section into three stages: ingestion and processing.

2.1 Ingestion

Our main goal was to collect data related to certain aspects that explain the influence of Covid-19 data at a city level. To do this, we divided our data search into different categories:

- City-level demographic data - *single value per city*
- Telecommunication accesses - *time series per city*
- Social Isolation index - *time series per city*
- Covid data - *time series per city*
- Lockdown data - *time series some cities*

To collect this data we utilized many different resources.

2.1.1 Characteristics of the City

This data is available in a web API offered by the government called [SIDRA](). SIDRA offers all the official data available from the Government related to national demographic research. In this platform, you can set the data that you want, the year, the level (city, state, country), and you can pre-visualize how the table is going to be shown in the end. After that, you can download it as a *.csv*.

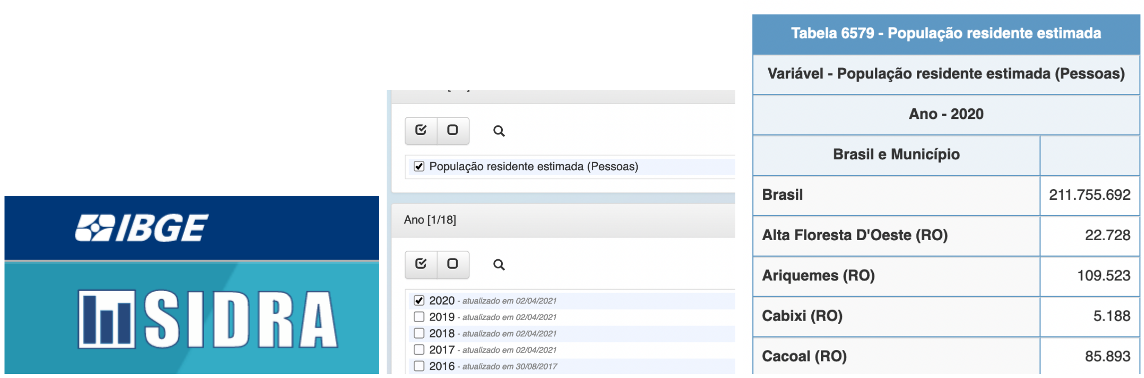


Figure 1. Sidra platform.

2.1.2 Telecommunications Accesses

This data is available in the regulatory agency of Brazilian telecommunications, Anatel. They offer a platform where you can access the data for different telecommunication technologies: cable TV, mobile and fixed broadband, fixed telephony, etc.

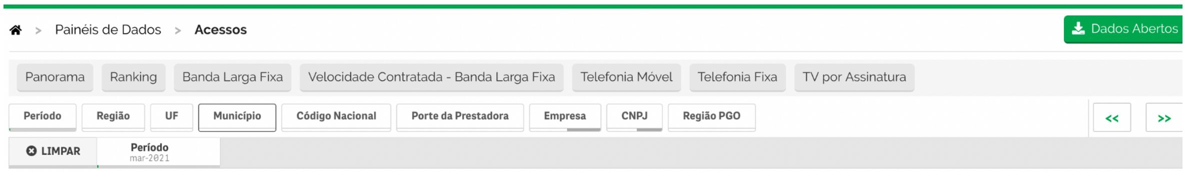


Figure 2. Anatel platform.

2.1.3 Social Isolation Index

This data was a bit complicated to find because it is not well organized at the municipality level by a single entity. Each state in Brazil has the autonomy of controlling its own pandemic-related measures, and almost all the states use the Social Isolation Index as a measure of the accuracy of their policies. The main problem is that not all the states have transparency of this data.

Analyzing all the 27 states, we only found 2 states that have consistent data at a city level: Bahia and São Paulo. Because of that, this data was the main limitation on the cities we used, but they offer a large variety of data since these are states from two different regions, with big and small cities.

The state of São Paulo offers the time series data in a Tableau, and it is available to **download** as a .csv. The main issue is that they do not show all the cities at the state level, but we have 140 of 570 cities during 364 days.

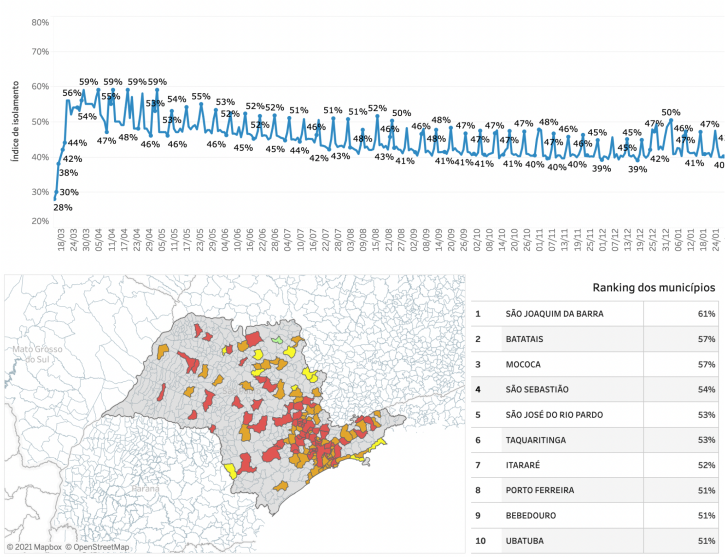


Figure 3. São Paulo state social isolation index per city.

The state of Bahia offers a similar [Tableau](), and it is available to **download**, and we could collect all the 417 cities data for 364 days.

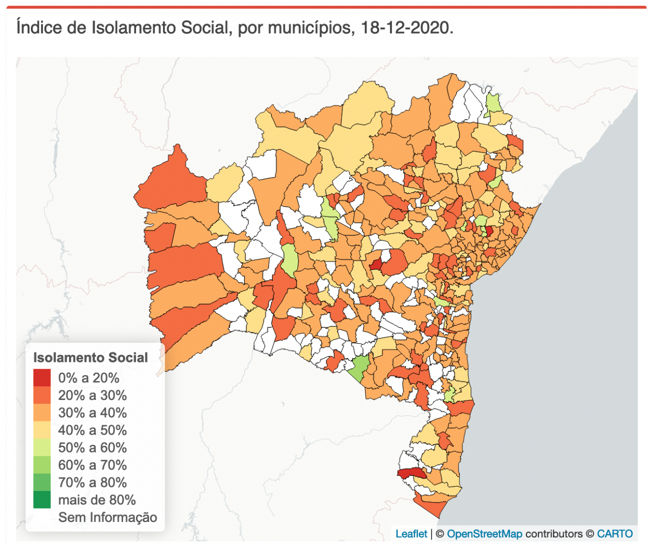


Figure 4. Bahia state social isolation index per city.

In a state-level analysis, we were forced to web scrape the data, because the data providers do not allow it to be publicly downloaded. To do this, we used a library that scraped the data from Tableau panels and used it to collect the time series from each state from Brazil.

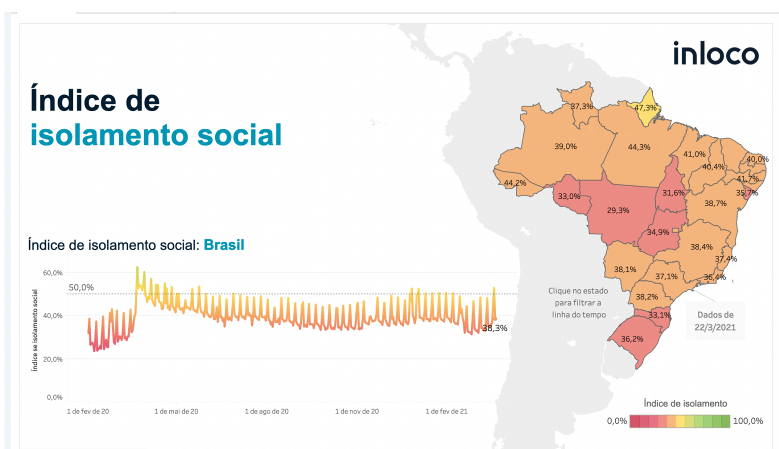
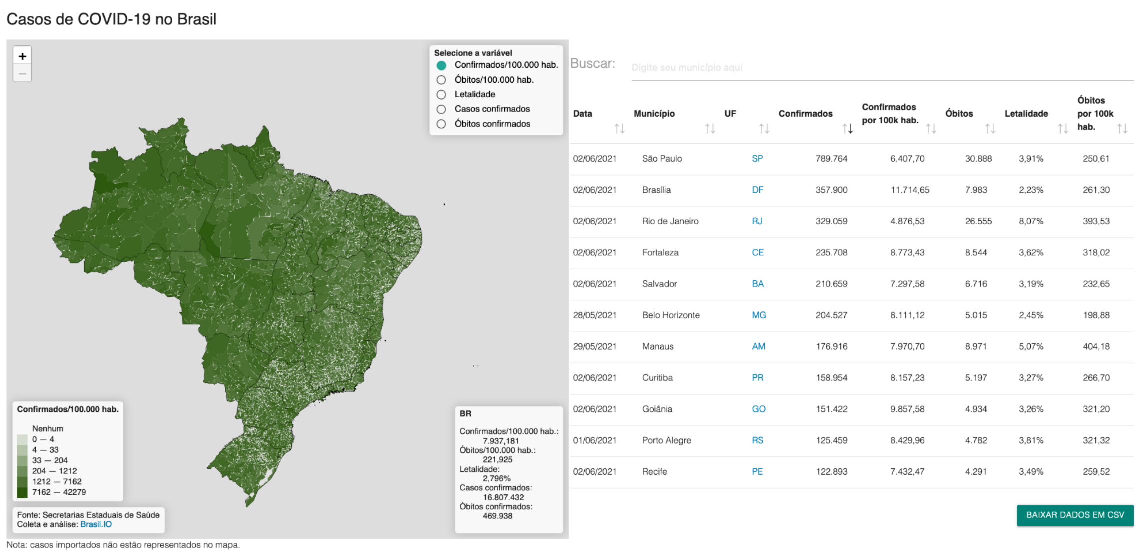


Figure 5. Brazilian states' social isolation index.

2.1.4 Covid Data

All the data was available in an official group of researchers in COVID-19 studies from Brazil and available to **download** as a .csv in this [Tableau](#).



Data	Município	UF	Confirmados	Confirmados por 100k hab.	Óbitos	Letalidade	Óbitos por 100k hab.
02/06/2021	São Paulo	SP	789.764	6.407,70	30.888	3,91%	250,61
02/06/2021	Brasília	DF	357.900	11.714,65	7.983	2,23%	261,30
02/06/2021	Rio de Janeiro	RJ	329.059	4.876,53	26.555	8,07%	393,53
02/06/2021	Fortaleza	CE	235.708	8.773,43	8.544	3,62%	318,02
02/06/2021	Salvador	BA	210.659	7.297,58	6.716	3,19%	232,65
28/05/2021	Belo Horizonte	MG	204.527	8.111,12	5.015	2,45%	198,88
29/05/2021	Manaus	AM	176.916	7.970,70	8.971	5,07%	404,18
02/06/2021	Curitiba	PR	158.954	8.157,23	5.197	3,27%	266,70
02/06/2021	Goiânia	GO	151.422	9.857,58	4.934	3,26%	321,20
01/06/2021	Porto Alegre	RS	125.459	8.429,96	4.782	3,81%	321,32
02/06/2021	Recife	PE	122.893	7.432,47	4.291	3,49%	259,52

Figure 6. COVID-19 cases and deaths per city in Brazil.

2.1.5 Lockdown Data

This data was manually collected because we did not find it in any centralized place. Also, most of them were in news, web pages, and reports built by official entities. As the data was sparse, we built a table and exported it in a .csv. The two main places that we found were a [Wikipedia page](#) and the official [portal of São Paulo](#) with some maps of regions with the colors indicating the measures.

To filter these measures, we only counted for cities with official measures and considerate strict measures, such as lockdowns and curfews. In the SP state, these occurred more often because they have an automatic policy based on the number of infections. The strict measures can be seen in the red regions of the image.

In the end, we collected **47 cities** with strict measures over the year. There were some problems detecting cities in the borders of the regions and since the data was mostly built through manual work, it is only a high-level indicator for strict measures and not a precise measurement of government policies.

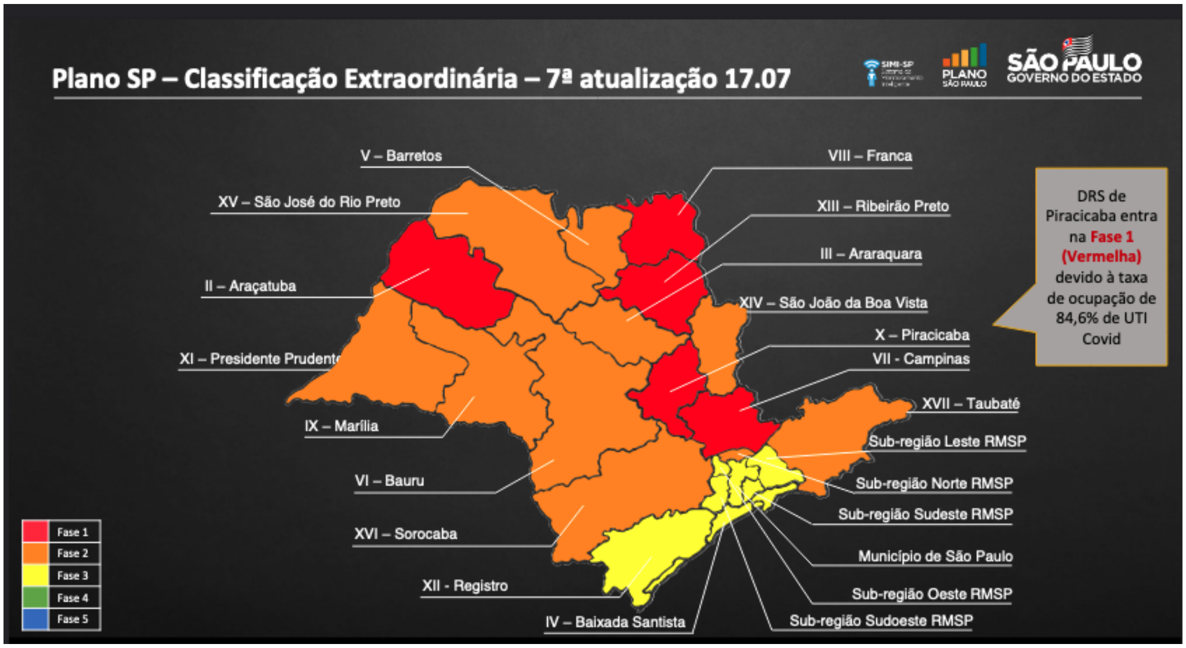


Figure 7. SP state measures per region.

2.2 Processing

Our data came from different sources and was not standardized. The goal at this stage was to create tables that we could use to make our analysis easier. For example, the Social Isolation index from each state came with different structures, as displayed in the next image.

	Município1	Código Município IBGE	População estimada (2020)	UF1	26/02/20	27/02/20	28/02/20	29/02/20
0	AMERICANA	3501608	242018	SP	36%	34%	34%	35%
1	AMPARO	3501905	72677	SP	33%	33%	33%	35%
2	ANDRADINA	3502101	57202	SP	37%	36%	36%	36%
3	ARAÇATUBA	3502804	198129	SP	32%	NaN	NaN	NaN
4	ARARAQUARA	3503208	238339	SP	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...
135	VARGEM GRANDE PAULISTA	3556453	53468	SP	38%	35%	34%	38%
136	VÁRZEA PAULISTA	3556503	123071	SP	31%	NaN	NaN	33%
137	VINHEDO	3556701	80111	SP	37%	34%	34%	38%
138	VOTORANTIM	3557006	123599	SP	32%	30%	NaN	33%
139	VOTUPORANGA	3557105	95338	SP	35%	37%	37%	37%

	City Name	Dt	Isolated_city
0	Abaíra	07/11/2020	63,6%
1	Abaíra	06/11/2020	78,3%
2	Abaíra	05/11/2020	72,4%
3	Abaíra	04/11/2020	57,7%
4	Abaíra	03/11/2020	65,4%
...	...	...	...
149084	Xique-Xique	05/02/2020	28,5%
149085	Xique-Xique	04/02/2020	26,9%
149086	Xique-Xique	03/02/2020	24,0%
149087	Xique-Xique	02/02/2020	34,8%
149088	Xique-Xique	01/02/2020	28,8%

Figure 8. Social Isolation index in different formats.

We also had many different work categories classified as white-collar or blue-collar work to simplify the data. In the end, we finished with simpler tables.

city_name	city_code	date	state	withdrawal	cable_tv_accesses	fixed_bandwidth_accesses	fixed_telephony_accesses	2G_mobile_accesses	3G_m
ABAÍRA	2900108	2020-04-13	BA	0.571	979	220	735.0	580.0	
ABAÍRA	2900108	2020-04-14	BA	0.556	979	220	735.0	580.0	
ABAÍRA	2900108	2020-04-15	BA	0.596	979	220	735.0	580.0	
ABAÍRA	2900108	2020-04-16	BA	0.588	979	220	735.0	580.0	
ABAÍRA	2900108	2020-04-17	BA	0.500	979	220	735.0	580.0	

	date	state	city_name	confirmed_acc	deaths_acc	city_code	death_rate	confirmed_day	deaths_day
0	2020-04-13	BA	ABAÍRA	1	0	2900108	0.0000	0.0	0.0
1	2020-04-14	BA	ABAÍRA	1	0	2900108	0.0000	0.0	0.0
2	2020-04-15	BA	ABAÍRA	1	0	2900108	0.0000	0.0	0.0
3	2020-04-16	BA	ABAÍRA	1	0	2900108	0.0000	0.0	0.0
4	2020-04-17	BA	ABAÍRA	1	0	2900108	0.0000	0.0	0.0

Figure 9. Final tables.

3. Data Summary

Once the data collection process is finished, we can analyze what we managed to gather. In the beginning, we had a set of data, but in the modeling process, we saw that some items had an unexpected noise, and it made the model less accurate. To solve this, we applied a filter and only kept cities where the population is more than 100,000 people.:

Characteristic	Elements before filter	Elements before filter
Number of Cities	556 (412 from BA + 144 from SP)	97 (17 from BA + 80 from SP)
Number of days	364	333
First day	2020-02-26	2020-03-19
Last day	2021-02-26	2021-02-26

Table 1. Instances.

We can see the distribution of the data in the following map and how the average social isolation rate in each city behaves before and after the filter.

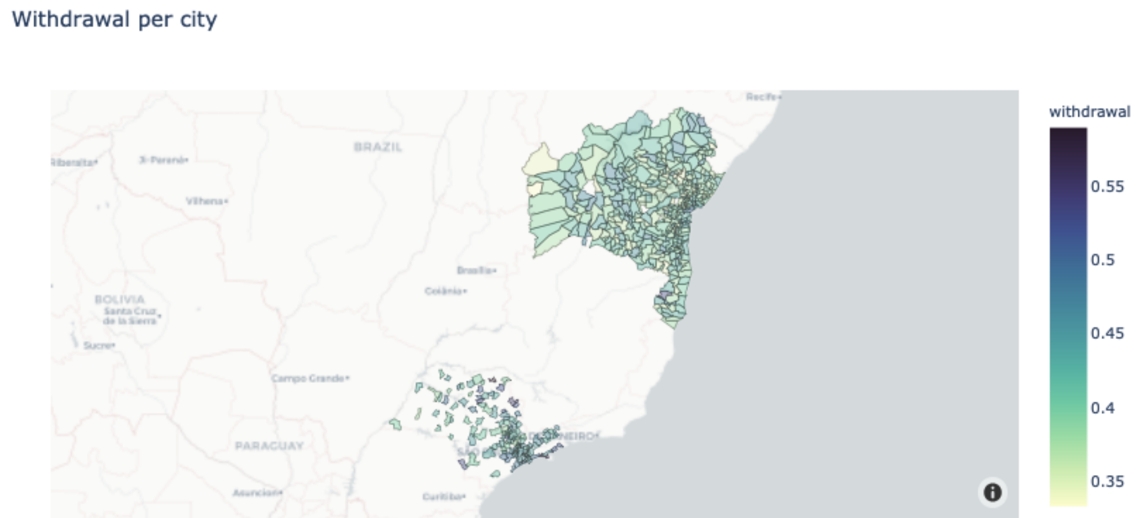


Figure 10. Social Isolation and city distribution before filtering.

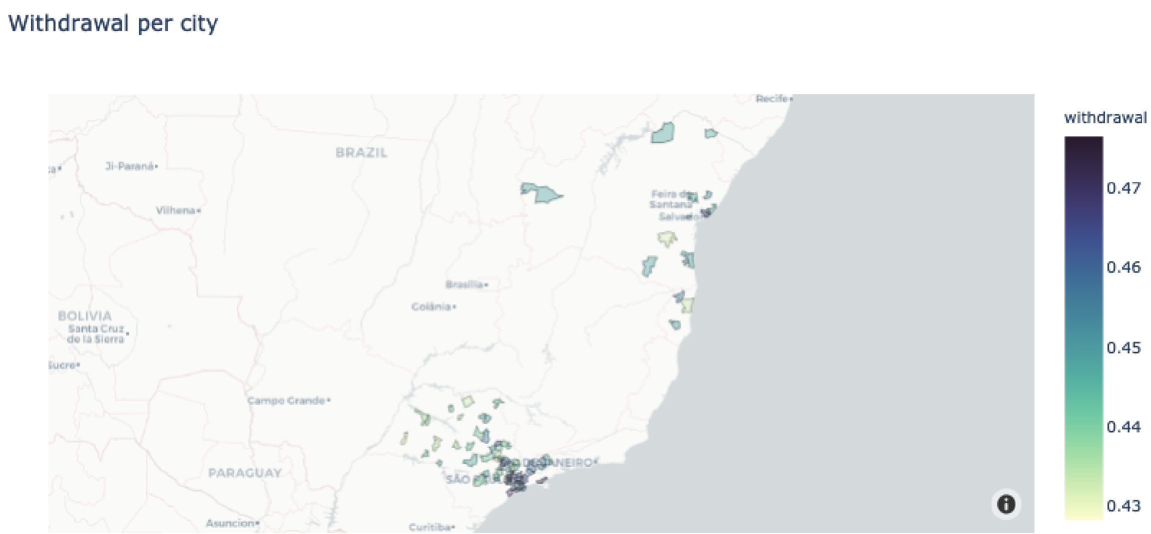


Figure 11. Social Isolation and city distribution after the filter.

	Feature (per 100000)	**elements**	**Mean**	**Std**	**Min**	**Max**
COVID	daily cases	12790	13.6	17.2	0.07	363.4
	daily deaths	12790	0.4	0.7	0.0	17.4
	cable_tv_accesses	12790	10884.0	5807.4	834.1	28899.1
	fixed_broadband_accesses	12790	22424.7	8252.5	3618.2	42785.8

Accesses	fixed_telephony_accesses	12790	38159.6	17026.8	5133.9	114025.1
	2G_mobile_accesses	12790	14732.6	53686.3	2495.1	938882.6
	3G_mobile_accesses	12790	21767.1	37857.6	3489.2	462844.9
	4G_mobile_accesses	12790	90752.7	28028.2	37405.4	333226.3
Social Isolation	Withdrawal (social isolation)	12790	2.2e-06	1.2e-06	3.3e-08	6.7e-06

Table 2. Descriptive of the time series data.

4. Data Treatment

4.1 Filtering

To define our filtering method, we first want to see how the data is distributed. One way to see this is by plotting graphs of the number of COVID-19 cases (deaths and cases) and the Social Isolation Index to understand how they are related.

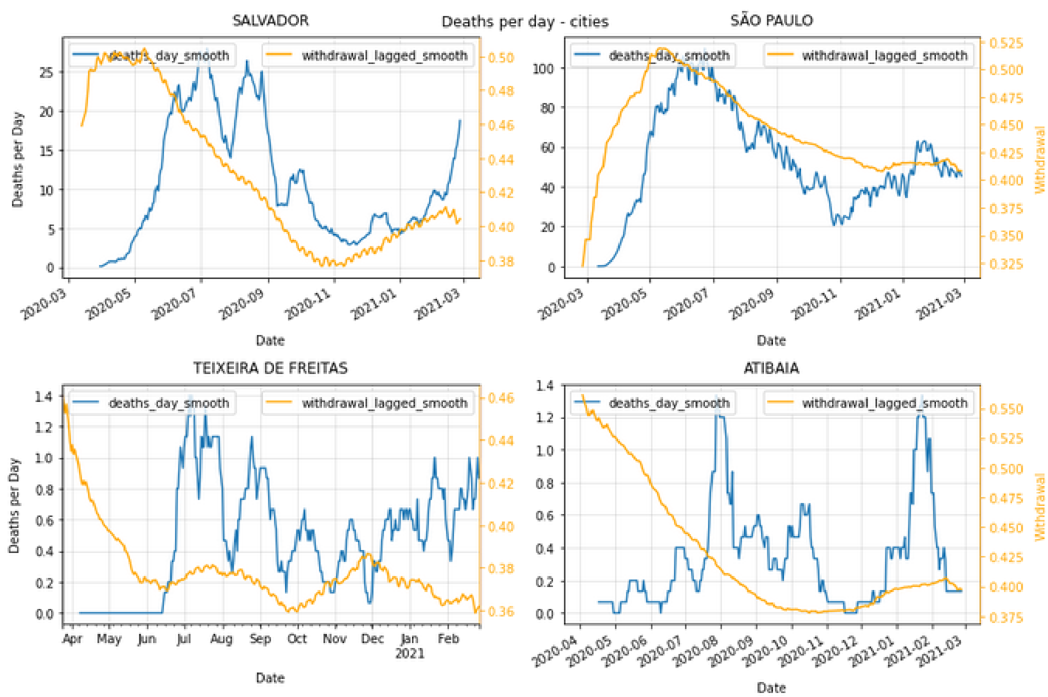


Figure 12. Social Isolation and COVID deaths.

When analyzing big cities, we can see that size plays a role, as in São Paulo and Salvador, there seems to exist some correlation between the social distance and the number of deaths. On the other hand, this correlation is not clear in small cities, which would require a more precise classification of the cities to improve the model.

Due to this fact, we will only consider cities with more than 100,000 population for the following analyses, since they convey the cities that were mostly affected by the COVID-19 pandemic. Additionally, we will only consider data entries that registered more than one case of COVID-19 per day and with a social isolation index higher than 30%.

The data of the social isolation index has a lot of variation during the week, which we also tried to avoid. To correct this, we applied a gaussian moving average to smooth out the high noise and delayed it from 21 days with a standard deviation of 7 days, based on some literature [Weier Wang MB 2020] and organization websites 'Center for Disease Control and Prevention' and 'wordometers'. Doing this, we can see the real influence of isolation on daily contamination and death numbers.

In the first row of the following Image, we can see the capitals from the states of SP and BA: more populated cities and with strong economic power and policies, while in the second row, we see the opposite. In this graph, we can see that the smaller cities have a lower social isolation rate on average.

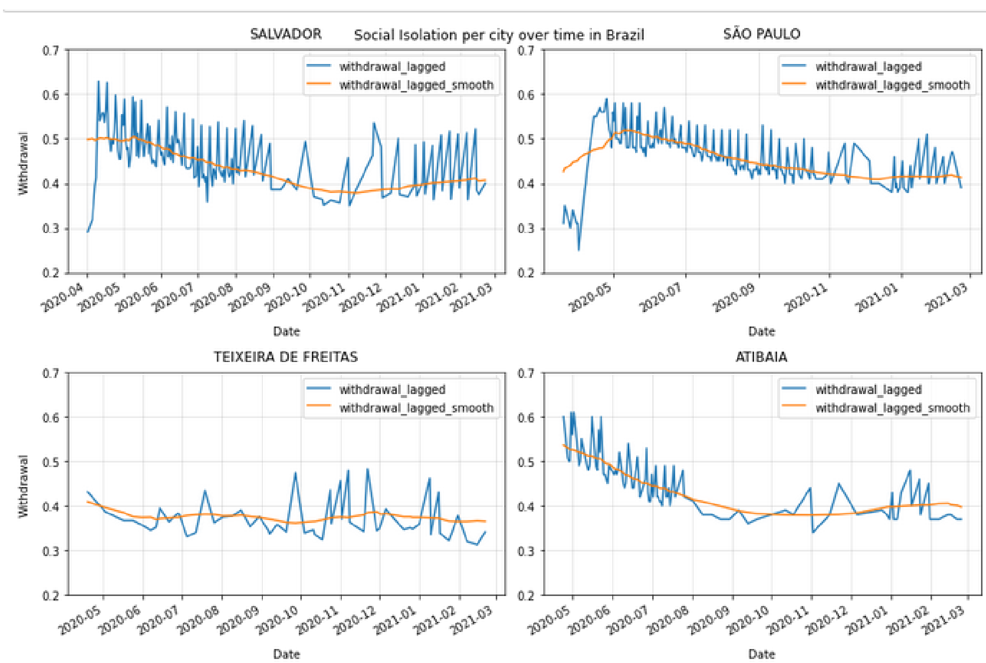


Figure 13. Noisy and smoothed social isolation index.

5. Data Analysis

5.1 Analyzing Social Isolation

In this study, we use the social isolation index as a proxy for how the population is respecting social distance, and as a result of governmental policies. We can see in the image below that for social isolation ratings higher than 40%, both the number of new daily cases and the number of deaths are inversely correlated with isolation. This shows that social distance is indeed positive for containing the spread of the virus.

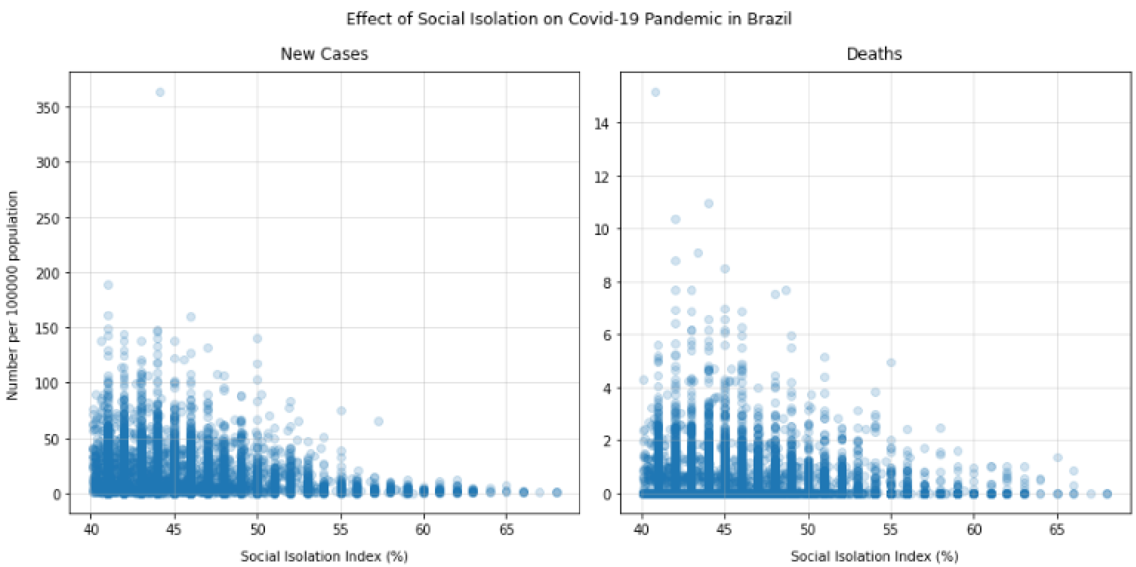


Figure 14. Social Isolation and COVID metrics.

However, if we look at the rates of death and new cases at the lower end of the isolation index, we notice a positive correlation. This may happen because when the number of cases is low, people start to go out and the government loosens restrictions, leading to an increase in the number of cases.

5.2 Correlation Between Variables

Before running regressions, we want to see the correlation between the variables that we will use in each model to see the influence in the model. As we will use a two-stage model, we can see the correlation between the variables in each part. The first one is the average correlation between the telecommunication variables for every city.

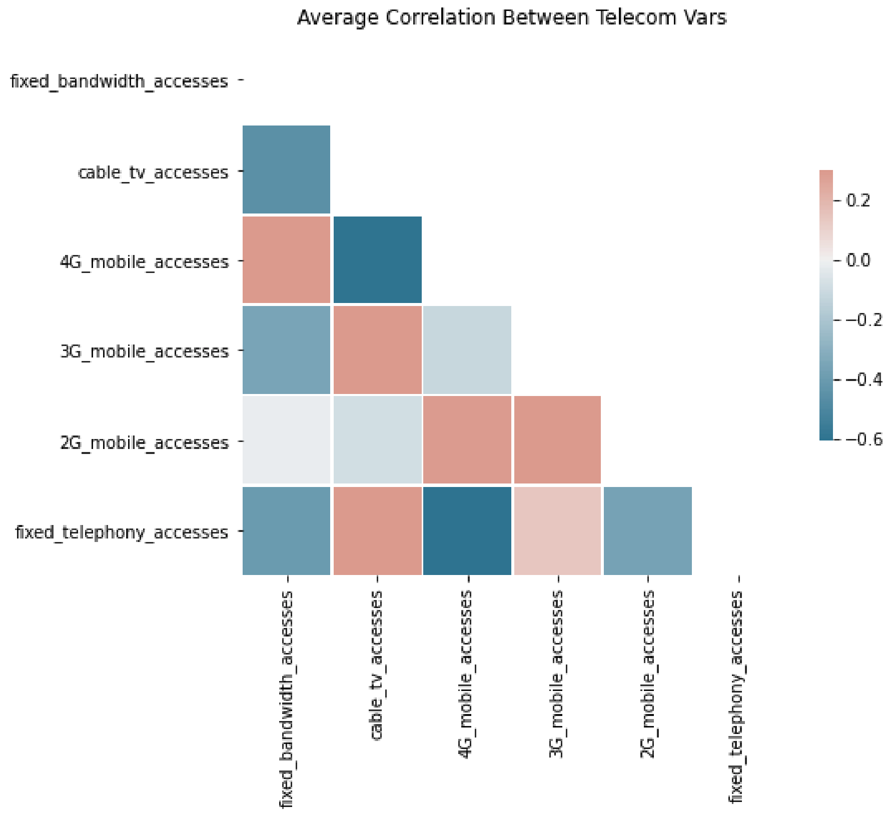


Figure 15. Heatmap telecommunications variables.

With this heatmap, we can see that the correlation is not that big for the majority of the variables. The biggest influence is the negative relationship between 4G x fixed telephony and 4G x cable TV, showing that higher usage of modern technologies tends to decrease the usage of less modern ones.

The other part of the model is to see the correlation between the demographic variables. And we can see that is not a big issue, the biggest one is the negative correlation between urbanization and blue-collar type of work. This says that the more urbanized a city is, the fewer blue-collar worker it has. Also, it is interesting to see that urbanization is not highly correlated with GDP per capita, showing that rural zones can also have a high GDP per capita.

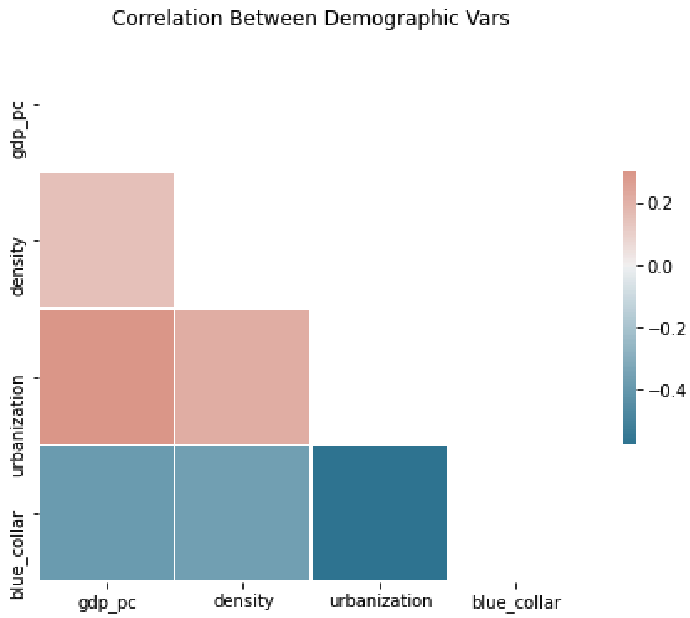


Figure 16. Heatmap demographic variables.

5.3 Panel Data Analysis Regression Models

For the OLS regression analysis, we decided to follow a two-stage model approach which can help to remove the correlation between the explained variable and the explanatory variable [Andrew Gelman 2005]. To do this we try to use not only variables as lockdown measures, geographic and infrastructure as auxiliary variables. In [David H. Glass, 2020], an SEIR two-stage model is used with differential equations to associate the lockdown to the number of deaths. In [Feinhandler, I., 2020], a simple two-stage model is applied to associate the political opinion with the number of deaths too. In this sense, our approach was to use the two-stage model to associate the telecommunications infrastructure with the number of deaths by COVID-19.

The first stage is responsible for estimating the social isolation index, and the second is responsible for estimating the number of COVID-19 deaths per capita. In this sense, we used the interaction between telecommunications infrastructure of cities and lockdown measures as instrumental variables to estimate the number of deaths per capita. This is ultimately based on the assumption that the access to information (which is proxied by the number of per capita accesses to telecommunication services) does not influence the number of COVID-19 deaths directly, but it does so indirectly by affecting the population's respect to social isolation during a lockdown, which, in turn, affects the number of deaths. Another reason for choosing the two-stage approach is due to the simultaneity between the number of COVID-19 deaths and the social isolation. When the number of cases starts to rise, restrictive isolation measures are enforced by the government, driving up the social isolation index. This implies that there is reverse causality between these two variables, which further justifies the use of the instrumental variable approach.

Regarding the model choice, we used a Fixed-Effects (FE) model in both stages. The main reason behind this choice lies in the assumption that the unobserved effects are correlated with our explanatory variables. Since our models are very simplistic and only account for a very limited number of city-level characteristics, such as GDP per capita, urbanization rate, and percentage of blue-collar workers, it is reasonable to admit that other significant unobserved effects are also correlated with the included variables. Besides, the FE models allow us to include city-level fixed effects and time-effects, which account for city-level access to health services (fixed effects), and improvements in medical treatment of COVID-19 cases (time-effects).

For the first stage, which estimates the social isolation index, the model includes time-varying data, such as the interaction between telecommunication accesses per capita and lockdown; and time-invariant demographic data, such as GDP per capita, urbanization rate, population. We also included an interaction term between the percentage of blue-collar workers and lockdown measures, which proxies for work conditions of the population (blue-collar workers are less able to work from home, for instance), further isolating the effect of access to information contained in the telecommunications infrastructure data.

For the second model, our main explanatory variable is the estimated social isolation rate, but we also included municipality demographic data.

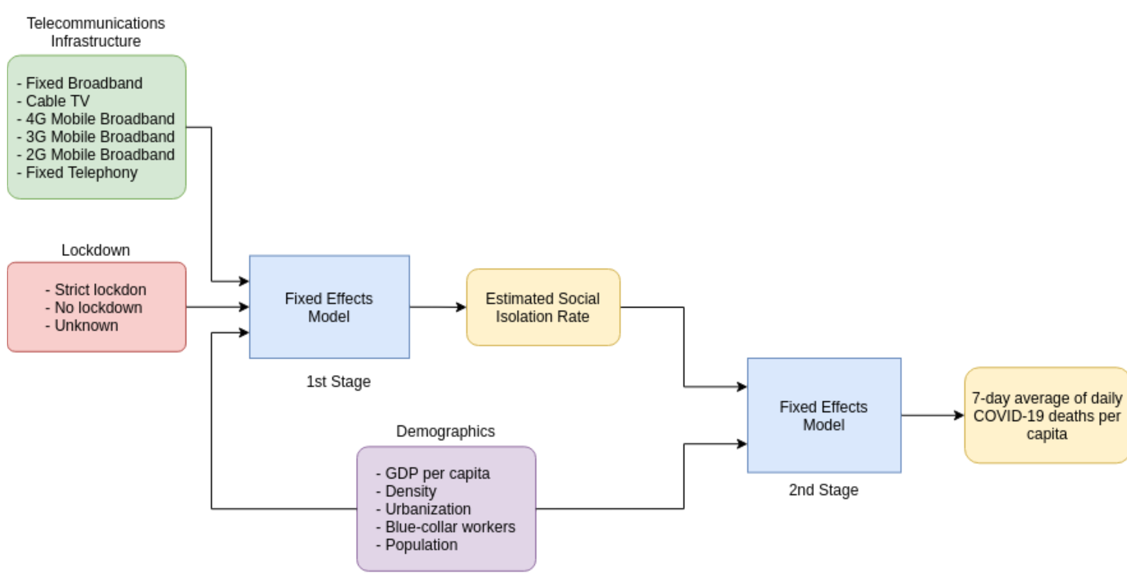


Figure 17. Two-stage panel data model scheme

The model can be described through the following equations:

$$social\ isolation\ index_{it} = \alpha_0 + \alpha_1(telecommunications_{it} * lockdown_{it}) + \alpha_2 demographics_i$$

$$log(7 - day\ average\ daily\ deaths)_{it} = \alpha_0 + \alpha_1\ estimated\ social\ isolation\ index_{it} + \alpha_2 demographics_i$$

5.4 Results and Interpretations

After defining the model, we performed several regressions, whose results are reported and analyzed in this section.

Regression Analysis of the Effect of Telecommunications Infrastructure on the Population's Respect to Social Isolation			
Stage 1 - Dependent variable: city-level social isolation index (%)			
Regressor	**(1)**	**(2)**	**(3)**
fixed broadband accesses per capita * strict_lockdown	-0.0608* (0.0330)	0.1672*** (0.0323)	-0.0334 (0.0351)
cable TV accesses per capita * strict_lockdown	0.1419*** (0.0529)	-0.1297*** (0.0489)	0.1389*** (0.0539)
4G mobile accesses * strict_lockdown	0.0065 (0.0048)	-0.0035 (0.0055)	0.0087* (0.0051)
3G mobile accesses per capita * strict_lockdown	-0.0050 (0.0046)	0.0162*** (0.0036)	-0.0134** (0.0062)
2G mobile accesses per capita * strict_lockdown	-0.0015 (0.0022)	-0.0061*** (0.0019)	0.0025 (0.0027)
fixed telephony accesses per capita * strict_lockdown	-0.0133 (0.0140)	-0.0317** (0.0134)	-0.0031 (0.0143)
pct. of blue-collar workers * strict_lockdown	-	-	-0.0003*** (8e-5)
GDP per capita	-	-	-9.9e-5*** (6.2e-6)
population density	-	-	2e-7*** (7.1e-8)
city urbanization pct.	-	-	0.0010*** (8.5e-5)
log(population)	-	-	-0.0012***

			(0.0003)
Fixed city effects	No	Yes	Yes
Fixed state effects	No	No	Yes
Time effects	Yes	Yes	Yes
R^2	0.0016	0.0274	0.1594
The individual coefficient is statistically significant at the *10%, **5%, or ***1% significance level.			

Table 3. 1st stage of Panel Data OLS Regression

Here, we can see that three telecommunications variables are statistically significant: cable TV and 4G mobile broadband, which are positively correlated with social isolation, and 3G mobile broadband, which is negatively correlated with social isolation. Out of all these variables, cable TV has by far the largest coefficient, and it is the only one that is statistically significant at the 1% confidence level. This may indicate that, during a lockdown, access to cable TV is the largest source of information that prevents people from not practicing social distancing. The results also point out that increasing infrastructure for cable television, 4G and 2G mobile broadband may lead to a relevant increase in access to information on important social distancing and government measures, which motivates people to stay isolated.

We can also see that the interaction term between the percentage of blue-collar workers and lockdown has a statistically significant negative coefficient. This result makes sense because blue-collar workers are usually not able to work from home, even during lockdowns. For this reason, social isolation is smaller in cities with higher percentages of blue-collar workers.

The only outlier in the analysis was the 3G mobile data, whose results point out that it has a statistically significant negative correlation with social isolation. While this variable is not highly correlated with any other one, it may be that its effects were absorbed by other services which are many times used together with 3G, such as 4G and 2G mobile broadband, which explains the unexpected result.

Regression Analysis of the Effect of Social Isolation on the 7-day Average of COVID-19 Deaths per Capita			
Stage 2 - Dependent variable: Log-transformed 7-day average of daily COVID-19 deaths per capita			
Regressor	**(1)**	**(2)**	**(3)**
estimated social isolation index	6.4e-6*** (6.6e-7)	-1.7e-5*** (3.6e-6)	-1e-5** (4.6e-6)
GDP per capita	-	-	-9.7e-10* (5.2e-10)
population density	-	-	2.2e-11*** (2.2e-12)
city urbanization pct.	-	-	3.4e-8*** (5.3e-9)
pct. of blue-collar workers	-	-	-2.6e-8*** (1.5e-9)
log(population)	-	-	-8.5e-7*** (2.2e-8)
Fixed city effects	No	Yes	Yes
Fixed state effects	No	No	Yes
Time effects	Yes	Yes	Yes
R^2	0.0098	0.0011	0.125
The individual coefficient is statistically significant at the *10%, **5%, or ***1% significance level.			

Table 4. 2nd stage of Panel Data OLS Regression

As for the second stage model, we see that social isolation is negatively correlated with the 7-day average of daily COVID-19 deaths per capita, and its effect is statistically significant at the 2% confidence level. We can see the formulas below:

$$\alpha_1 * Social\ Isolation_{predicted} + \ldots = log_{10}(deaths_{per\ day} + 1) / city\ population$$

$$deaths_{per\ day} = 10^\wedge(\alpha_1 * Social\ Isolation_{predicted} * city\ population + \ldots) - 1$$

Given α_1= -1e-5 found in the regression, supposing a variation of 1% in the Social Isolation Index and an average city with 500,000 habitants, we have:

$$\Delta deaths_{per\ day} = -0.1$$

The results indicate that a 1% increase in the social isolation rate results in a reduction of 0.1 in the average number of daily deaths in a city with a 500,000 population if all other factors remain constant. So if you count this in a year, you saved 36.5 lives.

We can try to combine the models to arrive at a final result.

$$\beta_1 * Cable\ TV_{per\ capita} + \beta_2 * 4G_{per\ capita} + \ldots = Social\ Isolation$$

$$Social\ Isolation \in [0, 1]$$

Combining both results, we can calculate the indirect effect of access to telecommunications on the average number of daily COVID-19 deaths, given that all other factors remain constant. These effects are shown in the table below.

In a city of 500,000, a reduction of COVID-19 deaths over one year by adding 5% more accesses per capita of a given telecommunication service.	
Service	Reduction on death over one year
Cable TV	28.0
4G mobile broadband	1.82
2G mobile broadband	0.52

Table 5. Effects of adding new telecommunications services on the number of average COVID-19 daily deaths

These statistically significant results indicate that the indirect effect of increasing the access to telecommunication services, and thus, to information, can lead to a reduction in the number of COVID-19 deaths. Therefore, we can conclude with this study, that the growth of telecommunications accesses influences the number of deaths, but they do not have a high impact on the number of deaths.

6. Conclusion

As mentioned in other works [David H. Glass, 2020] [Feinhandler, I., 2020], the two-stage regression approach is a good model when it comes to the explainability of variables. However, this is not verified for all the variables - it is a simple linear approximation, with a relatively low number of variables, which in many cases is not enough to account for the complexity of real-world problems. Even with more complex models, it is difficult to estimate the impact of specific measures since many assumptions are considered during the modeling stage, and the model results may turn out to be different from the theoretical expectation. As for the data that we collected, we also had to make some assumptions, and since they contain significant noise, we could not achieve high significance results or very assertive statements. For the results we found, we saw that the coefficient values vary a lot with small model changes, which shows that this approach is not the best to determine precise values or to be used to create policies. Instead, it is a relevant method to understand overall trends and correlations. Through this study, it is possible to achieve a comparative understanding of the different telecommunications services' impact on the number of COVID-19 deaths in Brazil.

The results allow us to quantify how access to Cable TV, 4G, and 2G mobile services can reduce the number of deaths by COVID-19. This is due to multiple effects, such as access to information sources, the option to work from distance, and the possibility to stay more at home. Our models estimate that for a city with a 500,000 population, an increase of 5% to 4G accesses can save 1.8 lives during one year. At the same time, in a city with this population, a 1% increase in the social isolation rate would save 36.5 lives a year, showing the importance of social isolation in the combat of the virus.

Moreover, the project went through the whole data pipeline, starting from the collection of the data, the conception of the model, and analysis of results. The study provides new ideas of policies that can help us combat current and future events of pandemics. Even if the results are not assertive, the project was a source of personal and academic development for the members in the areas of applied economics, econometrics, and public policies. Besides, this initial study also opens a lot of possibilities for future projects regarding telecommunications infrastructure and COVID-19 containment policies in Brazil. Some ideas are including the state-level isolation index and expanding the analyses for all cities that do not have a city-level measure, collecting government measures for more cities, and controlling for more demographic factors, such as daily commute hours, education levels, or city-level access to health services.

7. References

[1] ***Krenz, Astrid, and Holger Strulik***. “The Benefits of Remoteness – Digital Mobility Data, Regional Road Infrastructure, and COVID-19 Infections.” *German Economic Review*, 2021, doi:10.1515/ger-2020-0068.

[2] ***Ozili, Peterson K.*** “COVID-19 Pandemic and Economic Crisis: the Nigerian Experience and Structural Causes.” *Journal of Economic and Administrative Sciences*, ahead-of-print, no. ahead-of-print, 2020, doi:10.1108/jeas-05-2020-0074.

[3] ***Wang, Weier,*** et al. “Updated Understanding of the Outbreak of 2019 Novel Coronavirus (2019‐NCoV) in Wuhan, China.” *Journal of Medical Virology*, vol. 92, no. 4, 2020, pp. 441–447., doi:10.1002/jmv.25689.

[4] ***Wang, Weier,*** et al. “Updated Understanding of the Outbreak of 2019 Novel Coronavirus (2019‐NCoV) in Wuhan, China.” *Journal of Medical Virology*, vol. 92, no. 4, 2020, pp. 441–447., doi:10.1002/jmv.25689.

[5] ***Zhang, Xiaoqun***. “Broadband and Economic Growth in China: An Empirical Study during the COVID-19 Pandemic Period.” *Telematics and Informatics*, vol. 58, 2021, p. 101533., doi:10.1016/j.tele.2020.101533.

[6] ***Feinhandler, Ian, et al***. “Predictors of Death Rate during the COVID-19 Pandemic.” *Healthcare*, vol. 8, no. 3, 2020, p. 339., doi:10.3390/healthcare8030339.

[7] ***Gelman, Andrew.*** “Two-Stage Regression and Multilevel Modeling: A Commentary.” *Political Analysis*, vol. 13, no. 4, 2005, pp. 459–461., doi:10.1093/pan/mpi032.

[8] ***Glass, David H.*** “European and US Lockdowns and Second Waves during the COVID-19 Pandemic.” *Mathematical Biosciences*, vol. 330, 2020, p. 108472., doi:10.1016/j.mbs.2020.108472.

Appendices

Data Sources:

- **COVID cases and depths:** https://qsprod.saude.gov.br/extensions/covid-19_html/covid-19_html.html (city-level, 2020/2021)
- **Income per capita**: https://sidra.ibge.gov.br/tabela/5938 (city-level GDP 2018)
- **Access to public sanitation:** https://sidra.ibge.gov.br/tabela/1238 (city-level access to sanitation info, 2008/2010) - PNAD contínua
- **Urbanization:** https://sidra.ibge.gov.br/tabela/185 (city-level urbanization rate, 2010)
- **Population:** https://sidra.ibge.gov.br/tabela/6579 (city-level population, 2020) - Francisco - corrected 2018 data. OK
- **Density:** Francisco, city-level area. OK
- **Workers level of study per state:** https://sidra.ibge.gov.br/tabela/3586 (city-level pct of worker type, 2010) - Francisco, updated 2018 city-level data.
- **Workers sector, city-level, 2018**
- **Social withdrawal Index:**
    - SP https://www.saopaulo.sp.gov.br/coronavirus/isolamento/ (city-level, 2020/2021)
    - BA ftp://ftp.sei.ba.gov.br/covid19 (city-level, 2020/2021)
    - State-level Data from Inloco
- **Telecommunications infrastructure:**
    - Broadband fixed access (2020) - city level
    - Fixed phone services (2020) - city level
    - cable TV (2020) - city level
    - Broadband mobile (2020) - city level
- **Lockdown data**
    - Some cities Wikipedia: https://pt.wikipedia.org/wiki/Lockdown_no_Brasil_em_2021
    - Cities SP: https://www.saopaulo.sp.gov.br/planosp