The influence of socioeconomic deprivation on the recovery of urban mobility after COVID-19 in Latin America

Carmen Cabrera-Arnau *[1], Miguel González-Leonardo[2], Andrea Nasuto[1], Ruth Neville[1], and Francisco Rowe[1]

[1]Geographic Data Science Lab, Department of Geography and Planning, University of Liverpool, Liverpool, United Kingdom
[2]Center for Demographic, Urban and Environmental Studies, El Colegio de México (COLMEX), Mexico City, Mexico

Abstract

*Corresponding author: c.cabrera-arnau@liverpool.ac.uk

1. Introduction

1.1. General context

1.2. New forms of digital data to track human mobility

1.3. Contribution

2. Data

The aim of this study is to uncover the role of socioeconomic deprivation in determining the pace of post-pandemic recovery of urban mobility in Latin America. To do this, we leverage high-resolution data from different sources.

2.1. Meta-Facebook data

The multinational technology conglomerate Meta offers a range of data products aimed for social good through their Data for Good programme, which is open to trusted partners including universities, non-profit organisations, and international institutions. The data is available at small spatial and temporal scales and has the potential to improve how we respond to real-world crises or unusual events, such as earthquakes, hurricanes, floods or pandemics (Maas et al. 2019). In particular, the Data for Good programme offers location data, gathered from Facebook app users who have the Location Services setting turned on on their smartphone.

We analyse human mobility during the COVID-19 pandemic using the datasets Coronavirus Disease Prevention Map of Facebook Population During Crisis (Tile Level) and Coronavirus Disease Prevention Map of Facebook Movements During Crisis (Tile Level), hereinafter Facebook Population and Facebook Movements datasets. Specifically, we analyse data from four Latin American countries: Argentina, Chile, Colombia and Mexico. The datasets include data corresponding to a two-year period, starting on the 10th March 2020 and ending in mid-March 2022. Data are temporally aggregated into three 8-hour daily time windows (00:00-08:00, 08:00-16:00 and 16:00-00:00).The data is spatially aggregated into units called Tiles, according to the Bing Maps Tile System developed by Microsoft (Microsoft). This widely-used system offers a variety of world partitions, where the spatial units are square cells at various levels of resolution. The Data for Good datasets are typically generated using Bing tile levels 13 through 16, where level 13 results in tiles that are about 4.9 x 4.9 km at the Equator (Maas et al. 2019). The datasets also include data for baseline levels before COVID-19 based on a 45-day period ending on the 10th of March of 2020. The baseline data is computed using an average for the same time of the day and day of the week in the period preceding the crisis (e.g., average over all data collected on a Monday from 00:00 to 8:00 or average over all data collected on a Wednesday from 16:00 to 00:00). For more details on how the baseline values are computed, see (Maas et al. 2019).

The Facebook Population data provides the number of mobile app users who have the Location Services setting turned on, aggregated by tile. The location of each user in a given 8-hour time window is determined by the tile where they spent most of the time within that window. The Facebook Movement data captures the number of mobile app users who have the the Location Services setting turned on moving between a pairs of tile. The origin and destination of a movement are defined as the locations where a user spent most time between two subsequent time windows (e.g., 00:00-8:00 and 8:00-16:00). In addition to the count data, both the Facebook Population and the Facebook Movement datasets include the percentage difference between the number of counts during the crisis period and the corresponding baseline level for each entry.

Prior to releasing the datasets, information on personal characteristics of users is removed and several techniques are applied to ensure privacy and anonymity. Small-count dropping is one of these techniques, whereby a data entry is removed from the data set if population or movement counts are lower than 10 during the crisis period, the baseline period or both. While this technique makes it harder to identify individual users based on their movement patterns, the removal of data entries containing locations with small counts may lead to an underrepresentation of the population in these places. Another of these techniques consists in adding a small undisclosed amount of random noise to ensure that it is not possible to ascertain precise, true counts for sparsely populated locations. In order to produce a smooth population count surface, spatial smoothing using inverse distance-weighted averaging is also applied (see (Maas et al. 2019) for details).

2.2. Geographic delimitations

Our analysis focuses on intraurban mobility. The analysis is done at two spatial scales: for functional urban areas (FUAs) and for smaller spatial units, communes.

The rationale for this is: FUAs give us an indication of commuter regions, so they are the natural spatial unit to track intra-urban mobility. However, there is a lot of socioeconomic inequalities in Latin America, so we want to know what happens at smaller spatial scales, such as the commune level. For this reason, we also use data from W, X, Y and Z, which give us the geographical delimitations for FUAs in the four countries analysed here respectively, as well as delimitations for the communes forming the different FUAs.

2.3. COVID-19 stringency data

To understand the patterns of mobility in the context of the COVID-19 pandemic, we use the stringency index as a measure of the level of nonpharmaceutical interventions to COVID-19, such as social distancing and lockdowns. The stringency index ranges from 0 to 100, with 100 being the value corresponding to the most strict scenario. The values for the stringency index were retrieved from the COVID-19 government response tracker (https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker). For more information, see Hale et al. (Hale et al. 2021).

2.4. Relative deprivation data

2.5. Worldpop population data

An additional data set from WorldPop was used to capture the spatial distribution of actual population in the different countries analysed here. The WorlPop poulation dataset is in raster format and contains gridded population data at 1 sqkm resolution. We processed the WorlPop population data so it is spatially aggregated according to the Bing Maps Tile System. The level of detail was chosen so it coincides with that used in the Facebook Population and the Facebook Movements data for each of the countries analysed in this study.

3. Methods

3.1. Determining recovery

Explain here the baseline level of mobility and how just after the pandemic, there was a generalised drop in mobility across the urban-rural continuum. As the pandemic evolved and some of the stringency measures were relaxed, the levels of mobility bounced back to levels closer to those observed during the baseline period. Show Figure illustrating this evolution.

However, for the subsequent parts of the analysis, we need a formal definition of “recovery” since we want to quantitatively assess the association between relative deprivation and the pace of return to normal levels of urban mobility.

For each spatial unit (FUA or commune) the recovery time t_R^* is defined as the time at which we can reject the hypothesis ($p < 0.05$) that the level of mobility $m(t)$ is below the baseline level. The level of mobility m could be the percentage change relative to the baseline of the number of trips per user in a given spatial unit.

To do this, we need to apply the following method:

- Model the time evolution of m taking into account exogenous regressors such as the specific month of the year (to account for seasonality?), the average population density and, importantly, the stringency index. The model needs to be predictive and needs to give the probability distribution of the outcome variable.
- For the model, each sample will be formed by a spatial unit. Each sample has as many time steps as months in the movement dataset. Each sample also has information that should be incorporated as exogenous regressors.
- Compute t_R^* such that $t_R = min\{t_R : Pr(m(t_R) < 0) < 0.05\}$.

3.2. Analysis

For each spatial unit, we will be left with a value of t_R*. We can also compute the average relative deprivation of each spatial unit, denoted by d. Then, we can check whether there is an association between these two variables. Do we observe a trend wherer cities that reach recovery first are those where deprivaton is lower? If so, which are those cities?

4. Results

5. Discussion

Hale, Thomas, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, et al. 2021. "A Global Panel Database of Pandemic Policies (Oxford COVID-19 Government Response Tracker)." *Nature Human Behaviour* 5 (4): 529–38. https://doi.org/10.1038/s41562-021-01079-8.

Maas, P., S. Iyer, A. Gros, W. Park, L. McGorman, C. Nayak, and P. A. Dow. 2019. "Facebook Disaster Maps: Aggregate Insights for Crisis Response and Recovery." In *16th International Conference on Information Systems for Crisis Response and Management*, 836–47.

Microsoft. "Bing Maps Tile System - Bing Maps." https://learn.microsoft.com/en-us/bingmaps/articles/bing-maps-tile-system.