Data Descriptor for Nature Scientific Data

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

[1]Geographic Data Science Lab, Department of Geography and Planning, University of Liverpool, Liverpool, UK
[2]Centre for Demographic Urban and Environmental Studies, El Colegio de Mxico, Ciudad de Mxico, Mxico
[*]c.cabrera-arnau@liverpool.ac.uk

ABSTRACT

COVID-19 triggered a reduction in the frequency and extent of people's movement. Existing evidence suggests that while the impact of the pandemic on mobility was widespread, the extent of this impact was unequally felt across socioeconomic groups in the early stages of the pandemic. Here, we find that the most deprived locations have experienced a more accelerated recovery towards pre-pandemic levels of mobility in the long term. Furthermore, the socioeconomic disparities in the patterns of mobility triggered by the first outbreak of COVID-19, have persisted as of April 2023. These findings are based on the analysis of time-series mobility data corresponding to X urban areas from Latin American countries collected from Meta-Facebook users upon their consent. Our research highlights the importance of timely mobility data with high spatiotemporal resolution to understand the long-term effects of the pandemic and to inform equitable policy responses that address societal challenges in urban areas.

Main

Spatial human mobility is key to creating sustainable, livable and inclusive cities. At the societal level, spatial mobility enables the transfer of knowledge, skills and labour to places they are needed (Ackers 2005). Spatial mobility also shapes service and transport demand across urban spaces (Chen et al. 2016), and enables the monitoring and control of transmissible diseases (Belik, Geisel, and Brockmann 2011). At the individual level, mobility enables people to access and achieve opportunities and aspirations in space (Klugman 2009). Understanding spatial human mobility is thus important to supporting appropriate policy responses to address societal challenges relating to carbon emissions, urban planning, service delivery, public health, disaster management and transport (Barbosa et al. 2018; Chinazzi et al. 2020).

The COVID-19 pandemic resulted in a notable decrease in mobility, particularly in cities (Nouvellet et al. 2021). Coupled with fears of contagion in crowded public spaces, nonpharmaceutical interventions to contain the spread of COVID19 prompted this decrease in the overall levels of mobility (Nouvellet et al. 2021; Rowe, González-Leonardo, and Champion 2023). Especially during lockdowns, mobility recorded reductions in the frequency, distance and time of trips in cities across the globe (Abdullah et al. 2020; Bonaccorsi et al. 2020; Abu-Rayash and Dincer 2020; Lee et al. 2023). Higher engagement with remote working, online schooling and shopping activity reduced the need to travel for work, education, shopping and leisure, hence giving rise to more geographically localised mobility patterns (Borkowski, Jadewska-Gutta, and Szmelter-Jarosz 2021).

However, reductions in mobility levels were highly unequal reflecting existing socioeconomic inequalities in our societies (Chang et al. 2020). In most countries, affluent individuals tended to record the greatest drops in mobility levels as they are predominantly employed in knowledge-intensive jobs which can be done fully or partly remotely (Fraiberger et al. 2020; Bonaccorsi et al. 2020; Weill et al. 2020; Dueñas, Campi, and Olmos 2021; Santana et al. 2023). During the COVID-19 pandemic, the adoption of remote work reduced the need of commuting for knowledge-intensive, non-public facing jobs (Florida, Rodríguez-Pose, and Storper 2021).At the same time, individuals from less privileged socioeconomic backgrounds displayed less pronounced declines mirroring the nature of their work requiring public-facing, face-to-face interaction, and thus a requirement for daily work commutes (Dueñas, Campi, and Olmos 2021; Santana et al. 2023).

Thus, while a growing body of empirical evidence has contributed to advancing our understanding of the impacts of the COVID-19 pandemic on spatial mobility within cities, existing research has focused on more developed countries and the immediate effects of the pandemic during 2020. Less is known about the longer term patterns of resilience in urban mobility in less developed countries extending beyond this period (Rowe et al. 2023). Urban spaces have changed considerably since then, from going through waves of high COVID-19 fatality, infections, school and business closures to the removal of all COVID-19 restrictions as the UN World Health Organization (WHO) declared an end to the pandemic as a public health emergency; yet, different configurations of hybrid working have remained in the norm across many sector of the economy (Barrero, Bloom, and Davis 2021; Aksoy et al. 2022). Thus, assessing the extent to which the level of mobility has returned back to the pre-pandemic baseline level across socioeconomic groups is important to understand the potentially unequal long-term impacts of hybrid working.

A key barrier to monitor changes in geographic mobility patterns in less developed countries during and post the COVID-19 pandemic has been the lack of suitable data (Rowe et al. 2023). Traditionally census and survey data have been employed to analyse human mobility patterns in these countries (Green, Pollock, and Rowe 2021). Yet, these data streams are not frequently updated and suffer from slow releases, with census data for example being collected over intervals of ten years in most countries (Bell et al. 2014). Traditional data streams thus lack the temporal granularity to analyse population movements over short-time periods and to offer an up-to-date representation of the urban mobility system (Rowe 2023b). Data resulting from social interactions on digital platforms have emerged as an unique source of information to deliver this representation and capture human population movement in less developed countries at scale (Rowe 2023b). Particularly location data from mobile phone applications have become a prominent source to sense patterns of human mobility at higher geographical and temporal resolution in real time (Calafiore et al. 2023).

Drawing on a dataset of 213 million observations from Meta-Facebook users' mobile location data, we aim to assess socioeconomic differences in the extent and persistence of decline in urban mobility in Argentina, Chile, Colombia and Mexico during and after the COVID-19 pandemic from March 2020 to March 2023. We use Meta-Facebook data to measure origin-destination flows from March 2020 to May 2022, and Meta Prophet time-series forecasting machine learning algorithm (Taylor and Letham 2017) to predict origin-destination flows from June 2022 to March 2023. We use Functional Urban Areas (FUAs) boundaries from the Global Human Settlement Layer (GHSL), developed by the European Comission's Joint Research Centre (Schiavina M. 2019) to define urban areas; and the Global Gridded Relative Deprivation Index (GRDI) developed by NASA's Socioeconomic Data and Applications Centre (Columbia University 2022) from sociodemographic and satellite data inputs. Building on existing evidence (e.g. Rowe et al. 2022; Wang et al. 2022), we hypothesised that (1) urban mobility has recovered returning to the pre-pandemic baseline level of movement as nonpharmaceuthical restrictions were lifted; and, that (2) socioeconomic differences in urban mobility have endured the pandemic reflecting deep societal inequalities as knowledge-intensive businesses adopt hybrid working.

Latin America provides an ideal test-bed for testing these hypotheses because of its exceptionally high levels of inequalities (De Ferranti 2004; Carranza, De Rosa, and Flores 2023) and urbanisation (United Nations and Affairs. 2023). Half of the 20 most unequal countries in the planet are in this region. The average income Gini index of the region is 4 percentage points higher than that of Africa and 11 higher than China (Milanovic 2016), and cities display some of the starkest inequalities (Habitat 2022). Currently, over 80% of the population in Latin America live in urban areas. By 2050, this share is predicted to reach 89%, with the largest share concentrating in a few megacities (Habitat 2022). Developing an understanding of human mobility in Latin America is thus important to support sustainable and inclusive spaces (Habitat 2022).

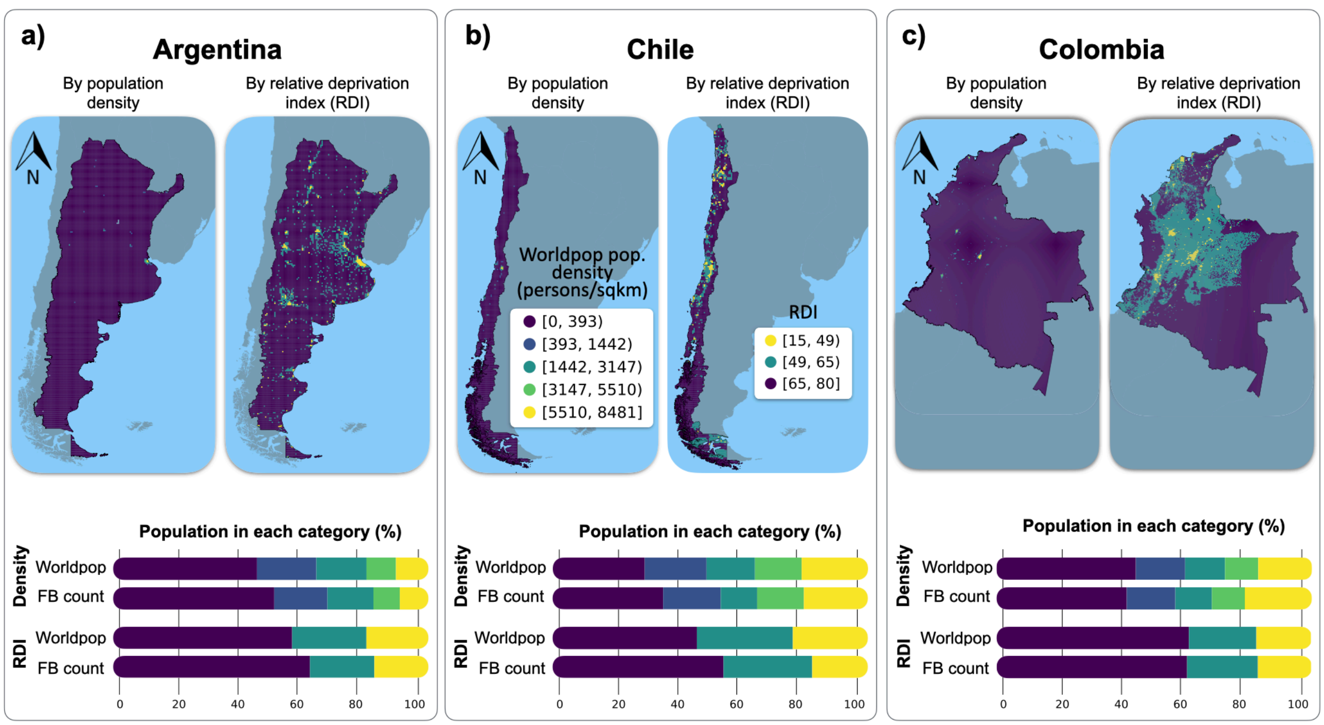
Results

Analysing the variability in the impact of COVID-19 on mobility across population groups

We focus on capturing how mobility patterns for different population groups have been impacted by COVID-19 in Latin America. We consider populations groups by classifying the spatial units of analysis into categories according to their population density and their socioeconomic level. More details for the classification method are provided in section X. The two criteria for classification are chosen due to their relevance for … We use five population density categories and three socioeconomic categories defined according to the relative deprivation index. The geographic distribution of categories for each of the countries is displayed in Figures (**Fig1-ARG?**) a), (**Fig1-CHL?**) a) and (**Fig1-COL?**) a) for Argentina, Chile and Colombia respectively.

A key challenge in the in the analysis of population counts and movements with Facebook user data is the absence of small count records, which is a result of privacy-protection techniques applied to ensure that the location of individuals or small groups cannot be identified. As a consequence, the countries of the analysis have locations with null values for Facebook population counts and flows. These missing data are not distributed randomly. For example, the missing values for FB population counts display high spatial autocorrelation as shown in Supplementary Figure SF1. Therefore, simply removing the missing records from the analysis could lead to geographically biased results (Afghari et al., 2019). To address this, we designed a data processing method to estimate the missing values, both for Facebook population counts and flows. Furthermore, this method also applies a correction factor to eliminate the fluctuations in the daily number of observations, assuming that the representativeness of Facebook data across spatial units remains stable during the study period. The data processing method is described in section X.

Even after applying the imputation method to estimate missing values, we acknowledge that Facebook data may still overrepresent certain population groups while underrepresenting others. For example, Figures (**Fig1-ARG?**) b), (**Fig1-CHL?**) b) and (**Fig1-COL?**) b) show, for each country, the proportion of the population across the various analysis categories. These comparisons use WorldPop population estimates and Facebook population counts, with the latter reflecting active users on an average day during a pre-pandemic baseline period (see Section X of the methodology for details). In all three countries, notable discrepancies appear between the population distributions according to WorldPop and Facebook data. For instance, in Argentina and Chile, WorldPop data indicates a higher proportion of people living in low-density areas, suggesting that Facebook data underrepresents populations in these regions. Similarly, in these countries, Facebook data shows an overrepresentation of the most affluent socioeconomic group. While addressing this kind of representativity bias is beyond the scope of this paper, we recognise its significance and the importance of addressing it in future analyses. This issue is revisited in the Discussion section.



The analysis of mobility is therefore done using Facebook movement data that has been pre-processed according to our methodology. Figures (**Fig1-ARG?**) c), (**Fig1-CHL?**) c) and (**Fig1-COL?**) c) show, for Argentina, Chile and Colombia respectively, changes in the intensity of movement relative to pre-pandemic baseline levels, at two points during the period of analysis, May 2020 and March 2022. The patterns show remarkable consistency across countries for both classification criteria.

COVID-19 unevenly impacted human mobility across population categories in Latin America

The evolution of the percentage change in the number of movements is measured with respect to a baseline period prior to the pandemic as described in Methods. For the purposes of the analysis, we aggregate the raw movement data temporally into months and spatially into administrative units at various levels according to GADM, the Database of Global Administrative Areas (GADM 2024). The analysis focuses on administrative units that are within the boundaries of functional urban areas as specified by the Global Human Settlement Layer (GHSL).

For each administrative unit, we compute the Relative Deprivation Index (RDI) based on data from NASA's Socioeconomic Data and Applications Centre (SEDAC). Figure 1 displays the administrative units included in the study, coloured according to their average RDI. Predictions about the evolution of the percentage change in the number of movements are made using the Prophet forecasting procedure. Further details are provided in Methods.

We analyse the evolution of the percentage change in the number of movements with respect to a baseline period prior to the pandemic. Specifically, we focus the analysis on short-distance movements in urban areas to represent local and routine mobility (Owen and Green 1992), so only movements covering a distance of at most 70 km are considered. For a movement to be classified as urban it needs to start or end within a functional urban area from Argentina, Chile and Colombia. The observed data is available for a two-year period starting in April 2020, just after the first wave of COVID-19 pandemic cases, and ending in March 2022. After 2022 no observations are available, however, we generate a 12-month forecast up to March 2023 in order to gain a better understanding of the recovery trends.

Figure 2 displays the patterns of recovery for the mobility levels in the administrative units belonging to functional urban areas in the countries of interest. The three lines in each panel represent the mean levels of mobility for administrative units grouped into one of three terciles, according to their average RDI.

<!- –Patterns of recovery for urban mobility in administrative units from Argentina, Chile, Colombia and Mexico. —!>

Generally, there was a drop in the levels of mobility with respect to the baseline period in all four countries. This drop was especially large for Argentina, Chile and Colombia, with Mexico displaying a smaller decrease in the percentage change in the number of movements with respect to the baseline. Following the initial drop in mobility, all four countries evolve towards the recovery of baseline levels of mobility, with a generally increasing trend. There are, however, fluctuations from the general trend which manifest differently for each country. These fluctuations mirror each other in the case of Argentina and Colombia, where urban mobility sharply bounces back closer to pre-pandemic levels around July of 2020. Chile and Mexico display more progressive patterns of recovery, although Chile never reaches baseline levels. These fluctuations are unique to each country and can be attributed to local factors such as the effects of seasonality or the different stringency measures imposed by the national governments during the pandemic.

From Figure 2, we observe that there is a consistent tendency in how administrative units with varying levels of deprivation were affected by the pandemic. For all four countries, we observe that the administrative units in the most deprived tercile are the ones that experienced the smallest loss in levels of mobility at the beginning of the pandemic. Differences in the levels of mobility across relative deprivation terciles diminish with time. Argentina and Chile stand out as the countries with the largest differences in mobility levels for different relative deprivation terciles.

Recovery trajectories towards baseline mobility levels are heterogeneous across population categories in Latin America

In this section we explore further the role of socioeconomic deprivation in the evolution of the levels of urban mobility. For a given point in time (i.e. a month), we start by considering the relationship between the percentage change in the number of movements relative to the pre-pandemic baseline period and the average RDI, at the administrative unit level. We assume that this relationship is linear and we use a linear regression to estimate the slope and intercept characterising the line of best fit. This is shown for April 2020 and March 2022 in the right-hand side panels of Figure 3. After obtaining the slope and intercept for every month, we are able to plot the evolution of these parameters for both the observed and forecasted data, as displayed on the left-hand-side panels of the same Figure.

We find patterns in the evolution of the estimated parameters that characterise the relationship between the levels of urban mobility and RDI. In Argentina, Colombia and Mexico, we observe that the slope of this relationship evolves to become smaller over time. The tendency is not apparent in Chile, where the slope of the relationship remains approximately the same over time despite the temporary fluctuations. The slope captures the extent of differences in the level of urban mobility across administrative units with varying levels of socioeconomic deprivation. It can therefore be regarded as a measure of inequality in mobility patterns across socioeconomic groups. A slope equal to zero would mean that all administrative units display the same level of mobility regardless of their average RDI. Given the patterns observed in Argentina, Colombia and Mexico, we find that at the beginning of the pandemic there were notable inequalities between socioeconomic groups in terms of the levels of urban mobility. While it has taken more than two years for Argentina and Mexico to close the gap (their slope is close to zero from spring 2022), inequalities persist in Chile and Colombia as of March 2023.

The intercept of the relationship displays stronger patterns, which are consistent across the four countries. The intercept estimates the urban mobility levels that would be observed in administrative units where the RDI is zero. The intercept was below the baseline level at the early stages of the pandemic. As observed in Figure 3, while there are some differences between countries in the evolution of the intercept, the general tendency is for the intercept to increase. While Argentina and Mexico reach values that are closer to the baseline towards the end of the forecast period, the intercept for Chile and Colombia remains lower. Therefore, if there were areas with no socioeconomic deprivation, we would have seen a recovery in the levels of mobility, especially in Argentina and Mexico

Discussion

Using location data from Meta-Facebook users, our study aimed to examine the evolution of patterns of mobility across socioeconomic groups in functional urban areas from Argentina, Chile, Colombia and Mexico from April 2020 to March 2023, following the COVID-19 pandemic. We found a systematic drop in the number of population movements in April 2020, with the largest reductions observed in the most affluent administrative units within functional urban areas (FUAs) from Argentina, Chile and Colombia. While mobility rebounded closer to pre-pandemic levels approximately two years later, when COVID-19 restrictions eased, the number of movements remained below pre-pandemic in Chile. Furthermore, we found that at the beginning of the pandemic there were inequalities between socioeconomic groups in terms of the levels of urban mobility. While it has taken more than two years for Argentina and Mexico to gradually reduce gap, inequalities persist as of March 2023, especially in Chile and Colombia according to our estimated data.

We focused the analysis on short-distance movements in urban areas, especifically those covering 70 km or less. These journeys are typically considered to represent local and routine mobility (Owen and Green 1992) . However, due to the characteristics of the Meta-Facebook movement data, we are unable to distinguish the purpose of these short-distance movements. Hence, some of our data could be capturing journeys that involve a permanent change of place of residence. Our work therefore motivates the need to answer questions regarding the validity of digital footprint data for the analysis of human mobility. Further research should focus on inferring more specific information about the nature of the journeys, following similar approaches to those proposed by Cabrera-Arnau et al. (2023), and quantifying the extent to which the digital footprint data mirrors the true mobility patterns.

Conducting research on urban mobility using digital footprint data is not straightforward, due to the challenges in accessing and working with unstructured data sets which are often subject to biases and statistical representation issues. These biases often arise from inequalities in access and usage of digital technologies across demographic groups (Rowe 2023a). Despite these challenges, the data and analysis that we used for this work provide evidence for non-trivial patterns that are consistent across four countries in Latin America and with other parts of the world. Our findings highlight the dynamic interplay between socioeconomic status and urban mobility, and shall be used to motivate and inform the public debate regarding the deep societal consequences of urban mobility disparities on the wider socioeconomic landscape of Latin American countries.

In conclusion, we argue that this work goes beyond the analysis of specific patterns by demonstrating the potential of digital footprint data for policy-relevant research on human mobility at an unprecedented level of spatiotemporal granularity. While we have seen a rise in initiatives to improve data services and methodological frameworks to facilitate the use of digital footprint data for social good, progress is still limited, especially in some parts of the world including Latin America. It is in the hands of governments and public organisations to prioritise the maximisation the societal benefits that digital footprint data has to offer. This includes engaging in activities such as building strategic partnerships with commercial data-holders and academic institutions to establish a unified framework for the use of digital footprint data in policy and research. In particular, we call for the creation of resources like those developed by the European Commission Joint Research Centre (Commission et al. 2022) and the UN Statistics Division (Division 2019), which identify sources of non-traditional data and set methodological protocols for incorporating mobile phone data into official mobility statistics. While current resources tend to have a global reach, we advocate for more tailored local initiatives that acknowledge disparities in regional data availability and adoption of digital technologies.

Data and methods

Meta-Facebook data

To capture population movements, we used anonymised aggregate mobile phone location data from Meta users for Argentina, Colombia, Chile and Mexico, covering a 24-month period from April 2020 to March 2022. We used the dataset Facebook Movements created by Meta and accessed through their Data for Good Initiative (https://dataforgood.facebook.com). The data are built from Facebook app users who have the location services setting turned on on their smartphone. Prior to releasing the datasets, Meta ensures privacy and anonymity by removing personal information and applying privacy-preserving techniques (Maas et al. 2019). Small-count dropping is one of these techniques. A data entry is removed if the population or movement count for an area is lower than 10. The removal of small counts may mean that population counts in small sparsely populated areas are not captured. A second technique 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. Third, spatial smoothing using inverse distance-weighted averaging is also applied is applied to produce a smooth population count surface. The Facebook Movements dataset offers information on the total number of Facebook users moving between and within spatial units in the form of origin-destination matrices. The data is temporally aggregated into three daily 8-hour time windows (i.e. 00:00-08:00, 08:00-16:00 and 16:00-00:00). The dataset includes a baseline capturing the number of movements before COVID-19 based on a 45-day period ending on March 10th 2020. The baseline is computed using an average for the same time of the day and day of the week in the period preceding March 10th. For instance, the baseline for Monday 00:00-08:00 time window is obtained by averaging over data collected on Mondays from 00:00 to 8:00 for the 45-day period. Details about the baseline can be found in (Maas et al. 2019). The Bing Maps Tile System developed by Microsoft (Microsoft) is used a spatial framework to organise the data. The Tile System is a geospatial indexing system that partitions the world into tile cells in a hierarchical way, comprising 23 different levels of detail (Microsoft). At the lowest level of detail (Level 1), the world is divided into four tiles with a coarse spatial resolution. At each successive level, the resolution increases by a factor of two. The data that we used are spatially aggregated into Bing tile levels 13. That is about 4.9 x 4.9km at the Equator (Maas et al. 2019).

WorldPpop population data

We used data from WorldPop (Tatem 2017) to classify the spatial units of analysis according to their level of urbanisation, and to estimate missing baseline values in the Facebook population data. WorldPop offers open access gridded population estimates at a resolution as small as 3 arc-seconds approximately 100m and 1km at the Equator, respectively. WorldPop produces these gridded datasets using top-down (i.e. dissagregating administrative area counts into smaller grid cells) or bottom-up (i.e. interpolating data from counts from sample locations into grid cells) approaches. For the purposes of this work, we use gridded population data at a resolution of 1km^2 in raster format. We perform a spatial join of the Facebook spatial units (Bing tiles) with the gridded population data and compute the sum of Worldpop populations corresponding to each of the Facebook spatial units.

Socioeconomic deprivation data

We use the Global Gridded Relative Deprivation Index (GRDI), Version 1 (GRDIv1) data set as a measure of socioeocnomic deprivation. The GRDI data is made available via NASA's Socioeconomic Data and Applications Centre (SEDAC), at a spatial resolution of 30 arc-seconds, or 1 km^2 approximately. The index quantifies the relative levels of multidimensional deprivation and poverty, where a value of 100 represents the highest level of deprivation and a value of 0 the lowest. We perform a spatial join of the Facebook spatial units and the gridded relative deprivation data and compute the average RDI corresponding to each of the Facebook spatial units.

Classification of tiles according to level of urbanisation and socioeconomic deprivation

Processing Facebook data

To ensure the privacy and anonymity of the users' data, Meta removes information corresponding to data entries where the population or movement count is less than 10 for a specific time or day, retaining only information about the percentage change in the number of counts with respect to the baseline period is retained (Maas et al. 2019). Consequently, we observe many tiles where the population or movement count for either the baseline or crisis period are blank.

Facebook population data

To imput Facebook population baseline values, we first identify all the baseline values that are available for each spatial unit and weekday (of the three available time windows, we only consider one per day). We then estimate the missing baseline values for each weekday, based on a linear model for the relationship between the Worldpop population and the Facebook population counts, which are fitted using ordinary least squares. This is illustrated in Supplementary Figure **??**.

We then use the complete baseline Facebook population values to compute missing Facebook population counts during the crisis period. This is possible because, as mentioned above, Meta reports the percentage change in the number of counts with respect to the baseline, even if the counts are not reported due to low value.

Facebook movement data

The imputation of Facebook movement baseline values is done according to a spatial interaction model (see e.g. (Rowe, Lovelace, and Dennett 2022)). We considered the population flow between an origin and a destination tile, and model this variable as a function of the Facebook population count at the origin tile on the same weekday, the Facebook population count at the destination on the same weekday and the distance between origin and destination; we also included indicator variables to capture the day of the week and the pair of population density classes corresponding to the origin and destination tiles. Mathematically, this model can be expressed as

$$ {}_{ijw} = {}_0 + {}_1 pop_{iw} + {}_2 pop_{jw} + {}_3 d_{ij} + {}_4 W + {}_5 C + \tag{1}$$

where ${}_{ijw}$ is the expectation of the flow of people from tile i to tile j on the weekday w; ${}_0$ is an intercept pop_{iw} and pop_{jw} are the Facebook population counts at the origin and destination on weekday w during the baseline period, d_{ij} is the distance between origin and destination, W is a series of indicator variables capturing the day of the week and similarly C is a series of indicator variables reflecting the pair of population density classes for the origin and destination tiles, resulting in X pairs (10 classesŒ10 classes minus one so it is not collinear with ${}_0$), ${}_{0,1,2,3,4,5}$ are model parameters to be estimated from the observed data. The error term is denoted by .

To estimate the model parameters, we used a count data regression model. Specifically, we used a negative binomial regression which is a generalised linear model (GLM) where overdispersion of the error term is assumed, i.e. the variance exceeds the mean.

We compute missing Facebook movement counts during the crisis period by considering the complete Facebook movements baseline and the percentage change in the number of counts with respect to the baseline, which is reported in the Facebook Movements datasets even when the count is not reported due to its low value.

Correction factors??? For representativeness

Data analysis

Code availability

For all studies using custom code in the generation or processing of datasets, a statement must be included under the heading "Code availability", indicating whether and how the code can be accessed, including any restrictions to access. This section should also include information on the versions of any software used, if relevant, and any specific variables or parameters used to generate, test, or process the current dataset.

References

Abdullah, Muhammad, Charitha Dias, Deepti Muley, and Md. Shahin. 2020. "Exploring the Impacts of COVID-19 on Travel Behavior and Mode Preferences." *Transportation Research Interdisciplinary Perspectives* 8 (November): 100255. https://doi.org/10.1016/j.trip.2020.100255.

Abu-Rayash, Azzam, and Ibrahim Dincer. 2020. "Analysis of Mobility Trends During the COVID-19 Coronavirus Pandemic: Exploring the Impacts on Global Aviation and Travel in Selected Cities." *Energy Research & Social Science* 68 (October): 101693. https://doi.org/10.1016/j.erss.2020.101693.

Ackers, Louise. 2005. "Moving People and Knowledge: Scientific Mobility in the European Union1." *International Migration* 43 (5): 99–131. https://doi.org/10.1111/j.1468-2435.2005.00343.x.

Aksoy, Cevat Giray, Jose Maria Barrero, Nicholas Bloom, Steven Davis, Mathias Dolls, and Pablo Zarate. 2022. "Working from Home Around the World." https://doi.org/10.3386/w30446.

Alexander, Lauren, Shan Jiang, Mikel Murga, and Marta C. González. 2015. "Origin–Destination Trips by Purpose and Time of Day Inferred from Mobile Phone Data." *Transportation Research Part C: Emerging Technologies* 58: 240–50. https://doi.org/https://doi.org/10.1016/j.trc.2015.02.018.

Barbosa, Hugo, Marc Barthelemy, Gourab Ghoshal, Charlotte R. James, Maxime Lenormand, Thomas Louail, Ronaldo Menezes, José J. Ramasco, Filippo Simini, and Marcello Tomasini. 2018. "Human Mobility: Models and Applications." *Physics Reports* 734 (March): 1–74. https://doi.org/10.1016/j.physrep.2018.01.001.

Barrero, Jose Maria, Nicholas Bloom, and Steven Davis. 2021. "Why Working from Home Will Stick." https://doi.org/10.3386/w28731.

Belik, Vitaly, Theo Geisel, and Dirk Brockmann. 2011. "Natural Human Mobility Patterns and Spatial Spread of Infectious Diseases." *Physical Review X* 1 (1). https://doi.org/10.1103/physrevx.1.011001.

Bell, Martin, Elin Charles-Edwards, Dorota Kupiszewska, Marek Kupiszewski, John Stillwell, and Yu Zhu. 2014. "Internal Migration Data Around the World: Assessing Contemporary Practice." *Population, Space and Place* 21 (1): 1–17. https://doi.org/10.1002/psp.1848.

Bonaccorsi, Giovanni, Francesco Pierri, Matteo Cinelli, Andrea Flori, Alessandro Galeazzi, Francesco Porcelli, Ana Lucia Schmidt, et al. 2020. "Economic and Social Consequences of Human Mobility Restrictions Under COVID-19." *Proceedings of the National Academy of Sciences* 117 (27): 15530–35. https://doi.org/10.1073/pnas.2007658117.

Borkowski, Przemysaw, Magdalena Jadewska-Gutta, and Agnieszka Szmelter-Jarosz. 2021. "Lockdowned: Everyday Mobility Changes in Response to COVID-19." *Journal of Transport Geography* 90 (January): 102906. https://doi.org/10.1016/j.jtrangeo.2020.102906.

Cabrera-Arnau, Carmen, Chen Zhong, Michael Batty, Ricardo Silva, and Soong Moon Kang. 2023. "Inferring Urban Polycentricity from the Variability in Human Mobility Patterns." *Scientific Reports* 13 (1): 5751. https://doi.org/10.1038/s41598-023-33003-7.

Calafiore, Alessia, Krasen Samardzhiev, Francisco Rowe, Martin Fleischmann, and Daniel Arribas-Bel. 2023. "Inequalities in Experiencing Urban Functions. An Exploration of Human Digital (Geo-)footprints." *Environment and Planning B: Urban Analytics and City Science*, November. https://doi.org/10.1177/23998083231208507.

Carranza, Rafael, Mauricio De Rosa, and Ignacio Flores. 2023. "Wealth Inequality in Latin America."

Chang, Serina, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. 2020. "Mobility Network Models of COVID-19 Explain Inequities and Inform Reopening." *Nature* 589 (7840): 82–87. https://doi.org/10.1038/s41586-020-2923-3.

Chen, Cynthia, Jingtao Ma, Yusak Susilo, Yu Liu, and Menglin Wang. 2016. "The Promises of Big Data and Small Data for Travel Behavior (Aka Human Mobility) Analysis." *Transportation Research Part C: Emerging Technologies* 68 (July): 285–99. https://doi.org/10.1016/j.trc.2016.04.005.

Chinazzi, Matteo, Jessica T. Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, et al. 2020. "The Effect of Travel Restrictions on the Spread of the 2019 Novel Coronavirus (COVID-19) Outbreak." *Science* 368 (6489): 395–400. https://doi.org/10.1126/science.aba9757.

Columbia University, Center for International Earth Science Information Network. CIESIN. 2022. "Global Gridded Relative Deprivation Index (GRDI), Version 1." Palisades, New York: NASA Socioeconomic Data; Applications Center (SEDAC). https://doi.org/10.7927/3xxe-ap97.

Commission, European, Joint Research Centre, C Bosco, S Grubanov-Boskovic, S Iacus, U Minora, F Sermi, and S Spyratos. 2022. *Data Innovation in Demography, Migration and Human Mobility*. Publications Office of the European Union. https://doi.org/doi/10.2760/027157.

De Ferranti, David M. 2004. *Inequality in Latin America: Breaking with History?* World Bank publications.

Division, United Nations Statistical. 2019. "Handbook on the Use of Mobile Phone Data for Official Statistics."

Dueñas, Marco, Mercedes Campi, and Luis E. Olmos. 2021. "Changes in Mobility and Socioeconomic Conditions During the COVID-19 Outbreak." *Humanities and Social Sciences Communications* 8 (1). https://doi.org/10.1057/s41599-021-00775-0.

Florida, Richard, Andrés Rodríguez-Pose, and Michael Storper. 2021. "Critical Commentary: Cities in a Post-COVID World." *Urban Studies* 60 (8): 1509–31. https://doi.org/10.1177/00420980211018072.

Fraiberger, Samuel P., Pablo Astudillo, Lorenzo Candeago, Alex Chunet, Nicholas K. W. Jones, Maham Faisal Khan, Bruno Lepri, et al. 2020. "Uncovering Socioeconomic Gaps in Mobility Reduction During the COVID-19 Pandemic Using Location Data." https://doi.org/10.48550/ARXIV.2006.15195.

GADM. 2024. "GADM Maps and Data."

Green, Mark, Frances Darlington Pollock, and Francisco Rowe. 2021. "New Forms of Data and New Forms of Opportunities to Monitor and Tackle a Pandemic." In, 423–29. Springer International Publishing. https://doi.org/10.1007/978-3-030-70179-6_56.

Habitat, UN. 2022. "World Cities Report 2022: Envisaging the Future of Cities." *United Nations Human Settlements Programme: Nairobi, Kenya*, 41–44.

Klugman, Jeni. 2009. "Human Development Report 2009. Overcoming Barriers: Human Mobility and Development." *Overcoming Barriers: Human Mobility and Development (October 5, 2009). UNDP-HDRO Human Development Reports.*

Lee, Sujin, Eunjeong Ko, Kitae Jang, and Suji Kim. 2023. "Understanding Individual-Level Travel Behavior Changes Due to COVID-19: Trip Frequency, Trip Regularity, and Trip Distance." *Cities* 135 (April): 104223. https://doi.org/10.1016/j.cities.2023.104223.

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.

Meng, Chuishi, Yu Cui, Qing He, Lu Su, and Jing Gao. 2017. "Travel Purpose Inference with GPS Trajectories, POIs, and Geo-Tagged Social Media Data." In *2017 IEEE International Conference on Big Data (Big Data)*, 1319–24. https://doi.org/10.1109/BigData.2017.8258062.

Milanovic, Branko. 2016. *Global Inequality: A New Approach for the Age of Globalization.* Harvard University Press.

Nilufer Sari Aslam, Tao Cheng, Mohamed R. Ibrahim, and Yang Zhang. 2021. "ActivityNET: Neural Networks to Predict Public Transport Trip Purposes from Individual Smart Card Data and POIs." *Geo-Spatial Information Science* 24 (4): 711–21. https://doi.org/10.1080/10095020.2021.1985943.

Nouvellet, Pierre, Sangeeta Bhatia, Anne Cori, Kylie E. C. Ainslie, Marc Baguelin, Samir Bhatt, Adhiratha Boonyasiri, et al. 2021. "Reduction in Mobility and COVID-19 Transmission." *Nature Communications* 12 (1). https://doi.org/10.1038/s41467-021-21358-2.

Owen, D., and A. Green. 1992. "Migration Patterns and Trends." In *Migration Processes and Patterns: Research Progress and Prospects*, edited by T. Champion and T. Fielding. Belhaven Press.

Rowe, Francisco. 2023a. "9.: Big Data." In *Concise Encyclopedia of Human Geography*, 42–47. Cheltenham, UK: Edward Elgar Publishing. https://doi.org/10.4337/9781800883499.ch09.

———. 2023b. "Big Data." In *Concise Encyclopedia of Human Geography*, 42–47. Edward Elgar Publishing.

Rowe, Francisco, Carmen Cabrera-Arnau, Miguel González-Leonardo, Andrea Nasuto, and Ruth Neville. 2023. "Reduced Mobility? Urban Exodus? Medium-Term Impacts of the COVID-19 Pandemic on Internal Population Movements in Latin American Countries." https://doi.org/10.48550/ARXIV.2311.01464.

Rowe, Francisco, Alessia Calafiore, Daniel Arribas-Bel, Krasen Samardzhiev, and Martin Fleischmann. 2022. "Urban Exodus? Understanding Human Mobility in Britain During the COVID-19 Pandemic Using Facebook Data." https://doi.org/10.48550/ARXIV.2206.03272.

Rowe, Francisco, Miguel González-Leonardo, and Tony Champion. 2023. "Virtual Special Issue: Internal Migration in Times of COVID-19." *Population, Space and Place*, March. https://doi.org/10.1002/psp.2652.

Rowe, Francisco, Robin Lovelace, and Adam Dennett. 2022. "Spatial Interaction Modelling: A Manifesto." http://dx.doi.org/10.31219/osf.io/xcdms.

Santana, Clodomir, Federico Botta, Hugo Barbosa, Filippo Privitera, Ronaldo Menezes, and Riccardo Di Clemente. 2023. "COVID-19 Is Linked to Changes in the Time–space Dimension of Human Mobility." *Nature Human Behaviour*, July. https://doi.org/10.1038/s41562-023-01660-3.

Schiavina M., Maffenini L., Moreno-Monroy A. 2019. "GHS-FUA R2019A - GHS Functional Urban Areas, Derived from GHS-UCDB R2019A, (2015), R2019A." European Commission, Joint Research Centre (JRC).

Tatem, Andrew J. 2017. "WorldPop, Open Data for Spatial Demography." *Scientific Data* 4 (1). https://doi.org/10.1038/sdata.2017.4.

Taylor, Sean J, and Benjamin Letham. 2017. "Forecasting at Scale." *PeerJ Preprints* 5 (September): e3190v2. https://doi.org/10.7287/peerj.preprints.3190v2.

United Nations, Department for Economic, and Social Affairs. 2023. *World Population Prospects 2022: Summary of Results.* UN.

Wang, Yikang, Chen Zhong, Qili Gao, and Carmen Cabrera-Arnau. 2022. "Understanding Internal Migration in the UK Before and During the COVID-19 Pandemic Using Twitter Data." *Urban Informatics* 1 (1). https://doi.org/10.1007/s44212-022-00018-w.

Weill, Joakim A., Matthieu Stigler, Olivier Deschenes, and Michael R. Springborn. 2020. "Social Distancing Responses to COVID-19 Emergency Declarations Strongly Differentiated by Income." *Proceedings of the National Academy of Sciences* 117 (33): 19658–60. https://doi.org/10.1073/pnas.2009412117.

Acknowledgements

The work was supported by Research England. . . , project no. . . . , 'RECAST'.

Author information

Authors and affiliations

Contributions

A.A. conceived the experiment(s), A.A. and B.A. conducted the experiment(s), C.A. and D.A. analysed the results. All authors reviewed the manuscript.

Ethics declarations

Competing interests

The authors declare no competing interests.

Supplementary information

Rights and permissions