





ARTICLE



<https://doi.org/10.1057/s41599-023-02116-9>

OPEN

# Businesses in high-income zip codes often saw sharper visit reductions during the COVID-19 pandemic

Aditya Kulkarni<sup>1</sup>, Minkyong Kim<sup>2</sup>, Jayanta Bhattacharya<sup>3</sup> & Joydeep Bhattacharya <sup>2</sup> 

As the COVID-19 pandemic unfolded, the mobility patterns of people worldwide changed drastically. While travel time, costs, and trip convenience have always influenced mobility, the risk of infection and policy actions such as lockdowns and stay-at-home orders emerged as new factors to consider in the location-visitation calculus. We use SafeGraph mobility data from Minnesota, USA, to demonstrate that businesses (especially those requiring extended indoor visits) located in affluent zip codes witnessed sharper reductions in visits (relative to parallel pre-pandemic times) outside of the lockdown periods than their poorer counterparts. To the extent visits translate into sales, we contend that post-pandemic recovery efforts should prioritize relief funding, keeping the losses relating to diminished visits in mind.

<sup>1</sup>Department of Computer Science, Yale University, New Haven, CT, USA. <sup>2</sup>Department of Economics, Iowa State University, Ames, IA, USA. <sup>3</sup>Department of Medicine, Stanford University, Stanford, CA, USA. email: [joydeep@iastate.edu](mailto:joydeep@iastate.edu)

## Introduction

The urban landscape, a vital engine of social and economic opportunity, is a daily witness to intense face-to-face socioeconomic interactions that exhibit excellent regularity<sup>1</sup>. Without restrictions on spatial mobility, city residents typically have plans for times within the day that they wish to stay indoors vs. outdoors or which businesses or locations to visit. However, the same people are forced to re-optimize once constraints are externally imposed due to a disease-related lockdown<sup>2</sup>. Depending on the timing and nature of the restrictions, some residents (a) cut down most travel outdoors and hunker down at home, (b) reduce mainly non-essential or easily substitutable travel and (c) switch to other businesses and locations different from ones they would patronize pre-lockdown<sup>3,4,5</sup>. Businesses, especially those relying on face-to-face transactions, see the effects of changes in the mobility calculus on the number of visits they see.

In March 2020, the transmission of the Sars-Cov2 virus had reached pandemic proportions, and governments worldwide were scrambling to formulate a non-pharmaceutical policy response. Governments enacted aggressive policies, including “shelter in place” and emergency closures of all non-essential services, with associated economic and social consequences wherever possible.

The first stay-at-home order was imposed in Minnesota, USA, on March 16, 2020. It forced many Minnesotans to remain home most of the time and re-optimize their daily outdoor-indoor mobility calculus. Some had to cut down on restaurant visits, others on trips to movie theaters, doctor offices, pharmacies, or grocery stores<sup>6</sup>. How did the vastly altered mobility landscape affect businesses and point-of-interest locations<sup>7</sup>? Did location matter? Did companies located in high-income zip codes see sharper declines in visits (relative to pre-pandemic levels)? Did timing matter, whether it was a lockdown period or not? This paper documents pandemic-influenced visiting patterns in businesses in Minnesota’s high- and low-income zip codes following the onset of COVID-19.

**Context.** To place these questions in context, consider Fig. 1, which compares the number of visits to restaurants within two cities in Minnesota with vastly different median incomes: Prior Lake has a median household income of \$109,609, more than double Hibbing’s income of \$49,009. The vertical axis measures restaurant visits during 2020–21 as a fraction of the same for 2019. (The data sources are described further below).

As 2020 starts, it is apparent that the yellow (Prior Lake) and the blue (Hibbing) lines are similar through January and February, with the yellow above the blue. However, soon after the first lockdown period started in March 2020, both lines began to show steep declines (the shaded regions represent indoor dining venues’ lockdown periods). Once restaurants reopened in June 2020, differences between the two came into sharp focus: Hibbing’s restaurant visits recovered to near pre-pandemic levels, while those in Prior Lake restaurants only achieved a sluggish recovery and remained at depressed levels for months after. Prior Lake’s restaurants show a significant spike in visits during the second lockdown, possibly due to more curbside pickup visits for Thanksgiving. In contrast, Hibbing’s restaurants slowly lose visitors. Hibbing’s restaurants likely relied more on regular in-person visit traffic since a greater proportion of their visits were of long duration before the second lockdown.

**Objectives and methodology.** The research is framed as follows. Consider visits by potential customers to two otherwise-identical businesses selling the same product, one located in a high and the other in a low-income neighborhood. Suppose, pre-pandemic, the

business located in the high-income area saw a similar number of visits to its counterpart in the low-income area. We want to know, how does the number of visits change post-pandemic? If visits fell, did they fall more (and stay that way longer) in the high-income neighborhood store? What effect did the lockdowns have? In short, the study’s main objective is to compare pre-and post-pandemic visits to similar businesses that differ by the median income of the zip codes they are situated in.

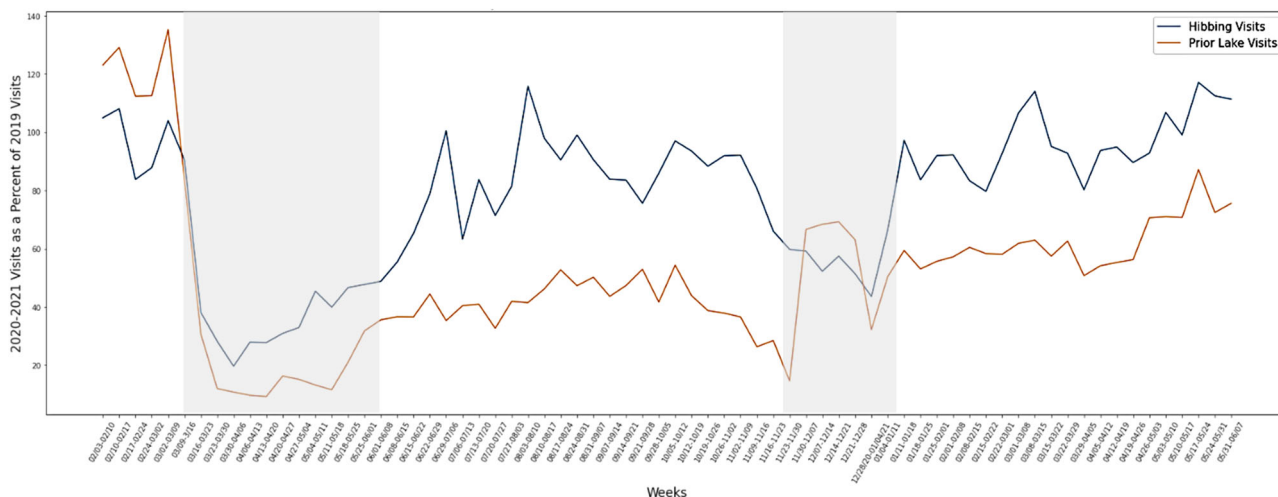
To that end, the paper outlines a pairwise comparison method that compares the reduction in visits to businesses in two groups of zip codes (top and bottom one-third of zip codes by median income). This method accounts for differences in individual zip code population sizes and calibrates it to the usual number of visits to each business location based on past years.

**Theory.** The literature, heretofore, has convincingly shown that, during various stages of the pandemic, visitations to most POI locations and businesses fell, *in the aggregate*, relative to pre-pandemic levels. In a sense, that is to be expected since (a) many people lost family members and employment and livelihoods were destroyed, (b) people’s innate aversion to infection risk kept them indoors voluntarily, and (c) government “stay-at-home” orders mandated little mobility. We take the literature forward by further disciplining the pre- and post-pandemic comparison to get at the disaggregated behaviors. We focus on businesses selling the same product but in two locations, a high and a low-income neighborhood. The underlying factors causing these changes are income and preferences, and they act in concert to generate behavioral responses. But as Weill et al. (2020) point out, “unpacking the mechanisms through which income is associated with behavioral responses to social distancing policies is a long-run challenge...[involving] access to information, mapping of information into subjective probabilities of outcomes and risk preferences, and constraints affecting capacity or ability to respond”.

To unpack the underlying mechanisms, assume innate preferences for various goods and services at both locations do not change because of the pandemic. Also, assume most residents in affluent areas did not experience a pandemic-induced, big negative income shock. Suppose we detect a sharp decline in visits to rich-neighborhood businesses post-pandemic. In that case, this would, *prima facie*, suggest that the underlying causal factor is likely a higher aversion to infection risk in such areas since their incomes and preferences for the products had not changed. However, if there is a sharp decline in visits to poor-neighborhood businesses, the reason could be income shock or disease-risk preference.

If, *additionally*, we discover that the visits to poor-neighborhood businesses fell *more sharply* relative to similar ones in wealthier neighborhoods, a causal explanation could be the bigger income (hence, demand) shock hitting poorer areas. It is also possible that the nature of businesses visited changed – more online deliveries, fewer visits to movie theaters, less dining-in – in the wealthier areas. Our data limitations do not allow a clear identification of income vs. preference mechanisms.

Nor is a difference-in-difference regression setup possible if low and high-income zip codes respond differently to government policies. The assumption underlying difference-in-differences is that, in the absence of the “stay-at-home” policy, both high- and low-income zip codes would respond similarly. However, as noted, unobserved differences in risk aversion between the income groups could violate this assumption. Moreover, the presence of multiple overlapping policies and intricate behavioral feedback further undermines the suitability of



**Fig. 1** Comparison of visits to full-time restaurants for Hibbing and Prior Lake Cities.

difference-in-differences as an evaluation method (Weill et al. 2021). This is why we simulate the entire potential distribution of differences by comparing *all possible* pairs of low and high-income zip codes. This simulation-based approach permits a more comprehensive analysis.

**Main findings.** The main table (Table 2) shows twenty-one different North American Industry Classification System (NAICS) business categories where businesses in wealthier zip codes saw a more significant decrease in visits than businesses in poorer zip codes. These metrics are calculated for periods in different stages of the pandemic.

- The blue cells (indicating a higher reduction in the poorer zip codes) are concentrated in the lockdown period.
- In comparison, the red (higher reduction in the rich zip codes) cells are concentrated in businesses that encourage extended indoor visits (sit-down restaurants, religious organizations, and movie theaters) *throughout the pandemic* (not just during the lockdown).
- In low-income zip codes, the drop in the number of visits to all kinds of businesses during the lockdown was more pronounced than in high-income zip codes.
- This paper also showcases kernel density estimation plots of all possible pairwise comparisons of visit changes for select business categories to show the dispersion of comparisons during different periods.

**Policy relevance.** The paper studies how non-pharmaceutical interventions designed to contain disease spread affected businesses based on their location in high or low-income areas. This focus on the location of the business in a high or low-income locality allows the research to offer limited guidance on post-pandemic governmental relief allocation.

This research is not intended to shed light on the measures required to contain the spread of the disease (Hsiang et al. 2020; Kraemer et al. 2020); it takes the disease, its progression, and the policies as given and attempts to understand their effect on commerce. Our answers could be valuable to policymakers trying to design non-pharmaceutical responses to epidemics under twin objectives: (a) contain or reduce disease spread and (b) generate minimal disruption to people’s lives and business operations. If the singular aim is to contain the epidemic, then indiscriminate lockdowns are probably the best policy response. This work

suggests that if objective (b) is taken seriously, the knowledge that businesses in rich and poor urban areas withstand mobility shocks differently can help decide which companies and locations to place under lockdown and which to spare.

Ideally, we would like to know how the changed visitation calculus generally affected sales or GDP. Before proceeding further, it bears emphasis that we do not have access to expenditure data, i.e., we do not know the identity of the visitors nor how much was spent or earned due to a single visit to any location. For us, visits are an intermediate marker of economic activity. While visits to a store in regular times closely track sales, that correlation may have weakened during the pandemic, especially if sales moved online, businesses started delivery programs, or an UberEats driver executed multiple restaurant pickup orders in a single visit<sup>8</sup>.

In a broader context, the present work documents one aspect of the significant changes in urban interactions due to social distancing measures, even as cities embrace new strategies to address inequities caused by such measures<sup>9</sup>. Limited by a lack of spending/revenue data, this research offers tentative guidance on where scarce pandemic relief dollars should go. For example, suppose movie theaters are visited mainly by the affluent, who sharply curtail theater visits (but not pharmacy visits) post-lockdown. In that case, it could be argued relief dollars should go to theater owners before going to pharmacy owners. Bear in mind businesses witnessing the sharpest reductions in visits are often places that hire many low-wage workers (Koren and Pető, 2020).

**Literature review and research gap.** This paper’s research question and the pairwise-comparison methodology are novel. While we recognize the sharp decline in visitations relative to pre-pandemic levels, our interest lies in documenting how the visitation reductions differed across the same type of business in rich vs. poor towns. To that end, unlike much of this literature, we do not simply compute changes in the volume of visitations – we hold the type of location (say, a dine-in restaurant) fixed and ask how such locations fared depending on whether they were established in a high or low-income area.

Several studies have analyzed the changes in human mobility after the onset of the pandemic (Gao et al. 2020; Huang et al. 2020; Lee et al. 2020; Sevtsuk et al. 2021; Weill et al. 2020). However, many employed a relatively short observation period and did not account for seasonality, as their pre-pandemic benchmark was set in early 2020, just before the pandemic. For example, Kim and Kwan (2021) examine the change in people’s

travel distance over seven months using anonymized mobile phone data. They discovered a substantial decline in mobility during the first lockdown, but also that it quickly recovered. However, their baseline was March 2020, which failed to account for seasonality. Many studies use February as a reference point for pre-pandemic levels. Arguably, February is not a representative baseline because it fails to capture seasonal changes, such as variations in weather or holidays. Instead, our study employs the 2019 levels at a parallel time of the year as a counterfactual for pre-pandemic levels.

This paper aligns with recent literature analyzing variation in mobility during the Covid-19 pandemic in the U.S., finding that high-income neighborhoods increased days at home substantially more than low-income neighborhoods (Jay et al. 2020; Ruiz-Euler et al. 2020). Some of this is explained by the enhanced capacity of the rich to work from home (Dingel and Neiman, 2020). Another angle that may explain the difference in stay-at-home days is risk perception (Bundorf et al. 2021; Erchick et al. 2021). Indeed, the literature has primarily attributed the perceived risk of catching COVID-19 as a theoretical framework explaining reduced business visitations. For example, while some studied how mobility restrictions led to a sharp decline in dine-in restaurant demand, others wondered which kind of mobility restrictions had the biggest effect on disease spread apparently, shutting bars and restaurants and closing non-essential businesses are key (Wellenius et al. 2021).

Numerous studies take a broader view and investigate human mobility patterns utilizing mobile devices. González et al. (2008) discovered that human trajectories display a high degree of regularity, while Lu et al. (2021) demonstrated that residents' income could be distinguished based on their activity patterns. As socioeconomic characteristics are correlated with mobility patterns, an event like Covid-19 that restricts mobility is an excellent natural experiment to comprehend said disparities (Li et al. 2022).

Bonaccorsi et al. (2020) find that Italian municipalities richer in terms of social indicators and fiscal capacity are more affected by the loss in mobility efficiency in the aftermath of the lockdown. Indeed, they report two seemingly opposite patterns: Individual indicators (average income) show that the poorest are more exposed to the economic consequences of the lockdown; conversely, aggregate indicators (municipality level), deprivation, and fiscal capacity, reveal that wealthier municipalities are those more severely hit by mobility contraction induced by the lockdown. They do not, however, study the effect on businesses.

Iio et al. (2021) found that high-income groups experienced a more significant decrease in mobility, with a particular emphasis on changes in travel distance compared to right before the pandemic outbreak. However, their study was limited in terms of its observation time frame and did not investigate where the reduction in mobility occurred. The initial step toward understanding the effects of mobility inflexibility among low-income populations during the pandemic would be to quantify the reduction in various types of businesses. Likewise, Lee et al. (2020) discovered that higher income groups began to spend more time at home following the COVID-19 outbreak, diverging from the uniform trend of staying at home across different sociodemographic groups. Yabe et al. (2023) found income diversity of urban encounters has substantially decreased from the pandemic through 2021 using mobile phone user data.

**Data**

**Mobility pattern.** We use SafeGraph COVID-19 Data Consortium data to observe human mobility patterns. SafeGraph collects GPS information from about forty-five million

anonymized smartphone devices (10% of devices) and 3.6 million point-of-interest (POI) locations in the U.S.<sup>10,11</sup> The GPS data comes from mobile applications where users have consented to location tracking. Within the SafeGraph COVID-19 Data Consortium, there are three primary datasets, of which we use two: Weekly Patterns and Core Places. From the Weekly Patterns dataset, we use data on the number of visits each week to each POI in the area we are analyzing. SafeGraph counts a visit to the POI by checking if the GPS location matches the inner boundary of a POI location<sup>12</sup>. We use the Core Places dataset to bring additional information on each POI, such as the North American Industry Classification System (NAICS) code (U.S. Census Bureau 2022), street address, city, region, and zip code. We merge both datasets by matching the unique POI identifiers “safe-graph\_id” (to separate POIs of the same name but different locations) to create a single dataset with millions of records. Each record contains the POI name, the two dates marking the beginning and end of a specific week, the number of visits in that week, the POI’s city, the POI’s zip code, and the POI’s NAICS code, and the POI’s “safegraph\_id”.

**Income and population.** Income and population data comes from the 2015–2019 American Community Survey 5-year Estimates (U.S. Census Bureau 2020). The zip code is the geographic unit in this study. To divide rich and poor areas, we rank zip codes by the median income of residents; the top one-third are classified as high-income, and the bottom one-third as low-income.

**Methodology**

We chose to analyze the difference in mobility patterns in Minnesota’s businesses because of the range of clearly articulated policies employed during various pandemic stages.

The first stage is the pre-pandemic baseline, from February 2, 2020, to March 16, 2020; the second is the first lockdown phase, from March 17, 2020, to June 1, 2020; the third is the interim when businesses reopened, from June 2, 2020, to November 23, 2020. The next phase is the second lockdown, from November 24, 2020, to January 11, 2021, and the last is the reopening period, from January 11, 2021, to May 31, 2021. These dates line up with the initial plans for reopening Minnesota.

The geographical unit is the zip code, denoted by *j*. Let *i* represent the business category (e.g., full-time restaurants), *k*, the number of businesses under category *i*, and *t* the period mentioned above. Let *v* denote visits, the measure of mobility. First, we compute the number of visits to a given business in a zip code,  $v_{ijkt}$ . Next, we compute  $V_{ijt} \equiv \sum_k v_{ijkt}$ , the sum of visits to every business in category *i* in zip code *j*. Then, we compare  $V_{ijt}$  against the number of visits in the corresponding weeks of the year 2019, denoted by  $V_{ijtpre}^{13}$ . Comparing the number of visits in 2019, we can account for normal variations such as holidays or seasonality. In addition, it removes the tracking bias consistent across 2019–2021 in the same area.

Table 1 shows the corresponding period dates from 2020–2021 to 2019.

	<b>Year 2019</b>	<b>Year 2020–2021</b>
Normal	2/04/19–3/18/19	2/03/20–3/16/20
Lockdown 1	3/18/19–6/03/19	3/16/20–6/01/20
Reopening 1	6/03/19–11/25/19	6/01/20–11/23/20
Lockdown 2	11/25/19–1/14/20	11/23/20–1/11/21
Reopening 2	1/13/20–6/01/20	1/11/21–5/31/21



**Table 2 Percent of Difference in Differences less than 0 for each period.**

Main Category	Name	2/03/20 - 3/16/20	3/16/20 - 6/01/20	6/01/20 - 11/23/20	11/23/20 - 1/11/21	1/11/21 - 5/31/21
Percentage of pairwise comparison: $(V_{iht} - V_{iht_{pre}}) - (V_{ilt} - V_{ilt_{pre}}) < 0$ , for $\forall h \in H$ and $\forall l \in L$						
Food Services	Full-Service Restaurants	44.7%	74.5%	66.7%	68.3%	58.9%
	Limited-Service Restaurants	56.1%	55.7%	60.5%	49.3%	55.1%
	Meat, Seafood, & Fruit/Vegetable Markets	47.6%	34.1%	50.8%	50.7%	49.2%
Essential Goods/Services	Supermarkets & Grocery Stores	42.3%	48.9%	60.4%	58.5%	57.8%
	Gasoline Stations	52.7%	57.5%	56.1%	59.0%	43.7%
	Plumbing, Heating, & A.C. Contractors	43.5%	41.1%	38.4%	37.8%	55.0%
	Offices of Physicians	46.0%	56.8%	59.4%	54.6%	54.6%
	General Medical & Surgical Hospitals	43.7%	54.6%	55.3%	48.2%	46.3%
	General, Surgical, Psychiatric, & Specialty Hospitals	49.5%	41.5%	48.2%	43.0%	44.3%
	Pharmacies & Drug Stores	58.4%	42.9%	51.4%	43.0%	49.7%
Medical Locations	Veterinary Services	49.6%	39.1%	51.6%	38.1%	54.8%
	Nursing Care & Child Day Care	50.4%	47.3%	58.8%	49.2%	63.8%
	Funeral Homes & Services	45.3%	41.5%	45.9%	38.6%	39.8%
	Car and Other Vehicle Rental	66.4%	33.9%	50.6%	43.3%	37.3%
Travel Locations	Hotels & Motels	39.4%	58.0%	64.1%	54.8%	54.2%
	Correctional Institutions	64.8%	46.1%	42.9%	40.9%	46.6%
Judicial Services	Police Stations	40.9%	36.4%	45.5%	27.3%	63.6%
	Libraries & Archives	55.2%	38.8%	50.7%	39.5%	44.2%
	Religious Organizations	46.2%	63.2%	62.3%	60.1%	57.3%
Miscellaneous	Barber Shops, Beauty & Nail Salons	50.3%	46.3%	44.4%	38.7%	43.3%
	Movie Theaters	60.5%	58.8%	66.0%	56.1%	64.9%
	Fitness & Recreational Centers (Gyms)	55.0%	48.7%	50.4%	52.7%	51.1%

The identifying assumption is, had the pandemic not happened,  $V_{ijt}$  would be very similar to the same period in 2019,  $V_{ijt_{pre}}$ . This assumption allows us to ascribe changes in visits to the lockdown, the fear of the virus, or both. We divide the number of visits by the zip code population to make changes comparable across zip codes.

In our sample,  $j = \{1, 2, \dots, 574\}$ ,  $j \in H \cup L$  where 287 zip codes are classified as low-income (denoted  $L$ ), and the rest are high-income (denoted  $H$ ). Next, we compare all possible (82,369 combinations) pairwise differences in the change in the number of visits for high- and low-income zip codes. That is, we take one high-income zip code and compare its population-deflated  $\Delta_{ij} = (V_{ijt} - V_{ijt_{pre}})_{j \in H}$  with the same for a low-income zip code  $\Delta_{ij} = (V_{ijt} - V_{ijt_{pre}})_{j \in L}$ . We do this for all possible pairwise combinations for all  $j$ .

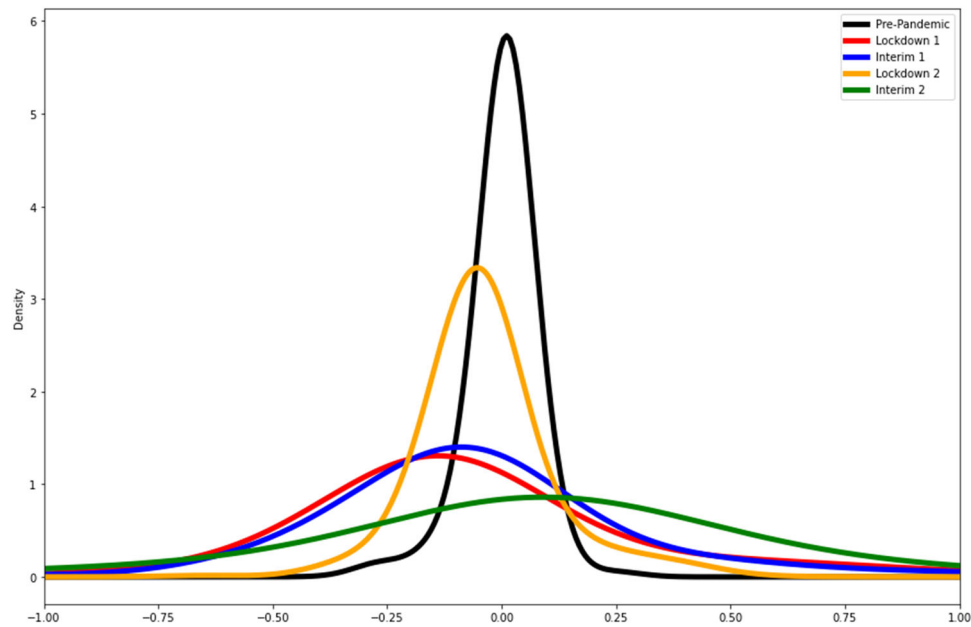
$$(V_{iht} - V_{iht_{pre}}) - (V_{ilt} - V_{ilt_{pre}}) \text{ for } \forall h \in H \text{ and } \forall l \in L$$

If this number is negative, it follows that visits to business  $i$  during  $t$  in high-income zip codes saw a sharper reduction compared to pre-pandemic levels in low-income zip codes.

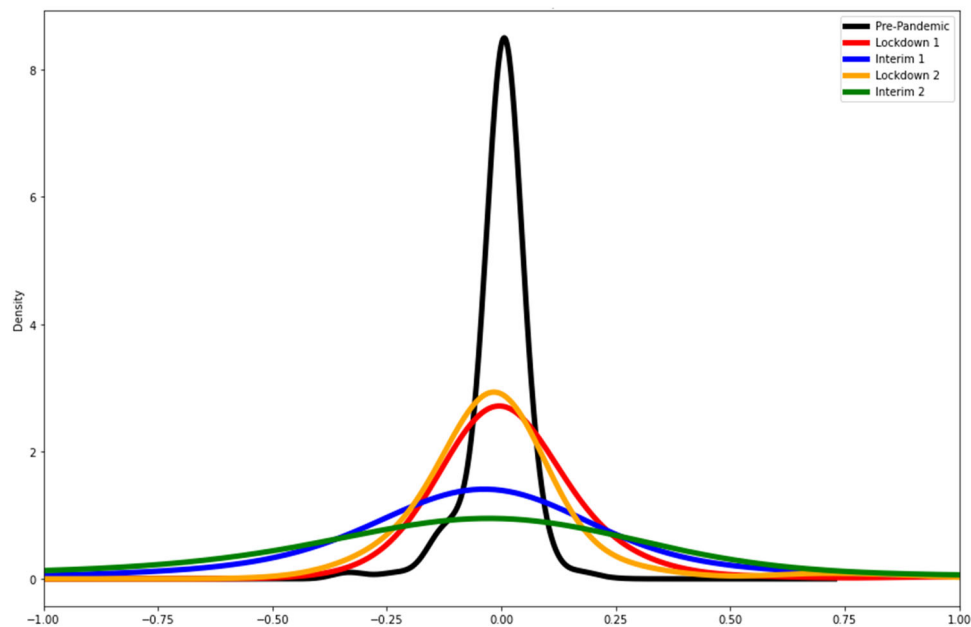
Finally, we construct the distribution of all differences across all possible combinations. We produced such distributions for  $i = \{1, 2, \dots, 21\}$  different business categories and  $t = \{1, \dots, 5\}$  different periods. The results are summarized in Table 2<sup>14</sup>.

Consider the following simple illustration: imagine two rich areas,  $a$  and  $b$ , and two poor areas,  $x$  and  $y$ . For each business type, there are four possible pairwise comparisons. For business-type  $i$ , the possible comparisons are  $\Delta_{ia} - \Delta_{ix}$ ,  $\Delta_{ia} - \Delta_{iy}$ ,  $\Delta_{ib} - \Delta_{ix}$ ,  $\Delta_{ib} - \Delta_{iy}$ . For example,  $\Delta_{ia} - \Delta_{ix} < 0$  means visits to business  $i$  in area  $a$  decreased more than in area  $x$  compared to visits in 2019. Similarly,  $\Delta_{ia} - \Delta_{ix} = 0$  means the change in visits is identical in areas  $a$  and  $x$ . We obtain the distribution by computing this pairwise comparison for all possible cases. A distribution centered at zero means that the average change in visits in low- and high-income areas is identical.

**Theory and background.** In economic analysis, the demand for a good or service depends on various variables, such as income and prices. In pandemic times, factors such as infection risk aversion, stay-at-home orders, and the ability to work from home also come into play. We do not have data on demand; we proxy demand by the number of visits. A regression relation can summarize the statistical relationship between the above-mentioned variables. Causal identification of, say, the impact of lockdown policies could, in principle, be done using a difference-in-difference (DID) approach. Such an approach compares the difference in the number of visits between low-income and high-income areas before and after the implementation of lockdowns. However, for the DID approach to be valid, it is necessary for the parallel trend assumption to be held, meaning that in the absence of the lockdown policy, the



**Fig. 2** Relative changes in the number of visits to full-time restaurants between low- and high-income zip codes.



**Fig. 3** Relative changes in the number of visits to groceries and supermarkets between low- and high-income zip codes.

trends in visits for both income groups would have followed similar patterns.

Unfortunately, the pandemic and the lockdown policies introduced severe threats to the parallel trend assumption. For instance, the type of occupation of a household significantly influences its income and ability to work from home. Further, as mentioned earlier, income may also affect individuals' response to the infection risk, and therefore, the unobserved risk-aversion response of rich and poor zip codes may differ.

To address this issue, we simulate the potential distribution of differences by comparing all possible pairs of low and high-income zip codes. This simulation-based approach allows for a

more comprehensive analysis, accounting for the varying responses and potential disparities between income groups. There remains a potential threat regarding the impact of Covid-19 on income. To address this, we categorized zip codes into high- and low-income groups based on pre-pandemic estimates of median income for each zip code and assumed that the relative rankings of zip codes by median income stayed the same during the pandemic.

#### Results: kernel density plots

Figure 2 and Fig. 3 show the plot of the kernel densities (KDE) derived from a smooth, continuous density estimation of all

possible pairwise comparisons of changes in visits to full-time restaurants and groceries between low- and high-income zip codes. Each color represents the distribution during the indicated period.

The black distribution represents the change in visits between low- and high-income zip codes during the normal pre-pandemic period. It is centered at zero implies that the change in visits to full-time restaurants or groceries was relatively similar in low- and high-income areas. This suggests a parallel trend in visits between the two income groups prior to the pandemic. Additionally, the small variance in the distribution indicates a low level of heterogeneity within both high- and low-income zip codes. This implies that the changes in visitation patterns were relatively consistent within each income group, further supporting the notion of a parallel trend. The same trend is observable for almost all categories, *i*.

However, as the pandemic unfolded, the situation changed. The red line in Fig. 2 represents the distribution during the first lockdown in March 2020. In this period, the distribution is centered to the left of zero, indicating that, on average, restaurants in affluent zip codes experienced a more significant decrease in visitors compared to the same period in 2019, compared to businesses in low-income zip codes. This pattern persists until the last reopening period, starting in 2021.<sup>15</sup>

On the other hand, the pattern of visits to groceries and supermarkets in Fig. 3 exhibits similarity between low- and high-income zip codes across all periods. The distribution is centered around zero, indicating both the rich and the poor zip codes reduced visits similarly. One way to explain this finding is to classify visits to groceries and supermarkets as essential-oriented and where income differences have minimal influence. People in both zip codes demand fresh produce, milk, and other perishables. Although the frequency of visits may have been affected by the pandemic, the overall difference in the number of visits between low- and high-income zip codes remained the same. Under this classification, visits to full-time restaurants are not essential-oriented.

### Results: pairwise comparisons frequencies, Table 2

Table 2 shows how the pairwise comparisons are distributed using  $(V_{iht} - V_{iht_{pre}}) - (V_{ilt} - V_{ilt_{pre}})$ . For example, consider a city with three high-income zip codes (A, B, and C) and three low-income zip codes (1, 2 and 3). Not knowing whether A is a representative zip code, we take every zip code and construct nine pairwise comparisons: (A, 1), (A, 2), (A, 3), (B, 1), (B, 2), (B, 3), (C, 1), (C, 2), (C, 3). Next, we compare the change in visits (population deflated) to locations in low-income zip codes to high-income ones. It stands to reason the number of visits will be different across zip codes. To make the comparison fair, we compare changes to their pre-pandemic levels. Therefore, each pairwise comparison represents a *double difference*; (1) difference in visits pre- and post-pandemic in each zip code, (2) difference across low- and high-income zip codes.

In Table 2, each cell represents a percentage indicating the share of pairwise comparisons demonstrating negative changes. For example, 66.6% means the change is negative in six out of the entire set (nine in this case) of pairwise comparisons. Specifically, in six of nine pairwise matches, visits to locations in high-income zip codes experienced a greater decline (relative to pre-pandemic levels) than in low-income zip codes. The entire distribution is centered at a negative number in this specific example. If the decline in visits had been similar across low- and high-income zip codes, the distribution (for large samples) would be centered around zero: roughly 50% of pairwise

matches would demonstrate a negative number, while the remaining 50% would display a positive number. A number *greater* than 50% indicates that the distribution is centered to the left of zero. This means that compared to pre-pandemic times, locations in high-income zip codes experienced a more substantial decline in visitors than those in low-income zip codes. We colored cells dark red if the percentage exceeds 70%, light red for greater than 60% and less than 70%, light blue for greater than 30%, less than 40%, and dark blue for less than 30%. Thus, blue indicates that poorer zip codes saw a sharper reduction in visits to that type of service than rich zip codes, and the color red, and vice versa.

Table 2 shows that people in more prosperous zip codes reduced visits to restaurants, religious organizations, and movie theaters. On the other hand, poorer zip code people reduced visits to plumbing, heating & A.C. contractors, veterinary services, funeral homes, police stations, and libraries. An important observation from Table 2 is that several blue cells appear in the lockdown period, indicating that individuals in lower-income zip codes considerably curtailed their movements during the lockdowns.

Another noteworthy observation is that affluent zip codes witnessed a greater decrease in visits to places considered more susceptible to infection. This suggests that individuals in wealthier areas had a higher ability to adapt to the pandemic by finding alternative options. They were more likely to modify their behavior and avoid locations with a higher risk of infection, reflecting their greater capacity to respond to the challenges posed by the pandemic. For example, fewer restaurant visits could be because more prosperous people started using more delivery services. The change in visits to essential services such as supermarkets, gasoline stations, and medical services is similar for rich and poor areas. The number of rich and poor zip codes with at least one location of a particular business category are listed in the last two columns within the Appendix, Table A1. These two columns shed light on the distribution of a particular business type between rich and poor zip codes (e.g., at least one childcare business location in 149 of the rich zip codes compared to only 73 of the poor zip codes).

A notable fact from Figs. 2 and 3 is that dispersion in the number of visits increased during the pandemic. Compared to the pre-pandemic level, the variance in the change in visits across rich and poor zip codes increased regardless of the type of business. This is observable even in industries (e.g., groceries) where the distribution of the difference post-lockdown is centered at zero. This means that, on average, the reduction of visits to groceries is similar between the rich and poor areas. However, within similar income groups, differences in mobility patterns emerged after the pandemic. For example, some affluent areas sharply reduced outdoor trips compared to others from similar income areas. Some poor areas maintained a level equivalent to the pre-pandemic level, while others declined to go outside. This implies that factors other than median income may influence the rich and poor areas to behave differently. These factors may include income inequality, the number of jobs that can be done remotely, the age composition of the population in the zip code, the perception of the epidemic, and other relevant factors.

In summary, prior to the pandemic, visits to various businesses were relatively similar across different income groups, as indicated by the pre-pandemic distribution centered at zero. The low variance of the pre-pandemic distribution suggests a homogeneous mobility pattern within each income group. However, rich and poor areas diverged into different visitation patterns after the pandemic, as evidenced by the shift in the distribution of different types of services. Furthermore, higher pairwise distribution variance during the pandemic period suggests that

people diverge from the typical income group pattern after the pandemic. This divergence may be explained by factors other than income.

### Discussion and conclusion

As shown above, the COVID-19 pandemic and subsequent restrictions impacted different types of businesses in zip codes with high and low median incomes. We show that the effects differed across various business categories in distinct periods. Comparing the number of visits during the pandemic to data from 2019, more than a year before the COVID-19 pandemic hit the U.S., and adjusting visit totals relative to current population counts in each zip code were factors we used to normalize data into a similar format. Regardless of how businesses were hit, there was a considerable divergence in mobility between richer and poorer zip codes in most business categories. In addition, within each income group, the mobility of people diverged from the pre-Covid pattern of the group.

**Policy.** The results can contribute to the design of post-pandemic recovery programs because a different pattern of visit changes across high and low-median-income zip codes was demonstrated. Specifically, the fact that a higher reduction in visits for the poorer zip codes is concentrated in the lockdown periods suggests some pandemic dollars may be earmarked for income support during those periods. Further, since businesses in rich neighborhoods that encourage extended indoor visits lost visits significantly more than in poorer areas, and throughout the pandemic, not just the lockdown periods, scarce pandemic dollars should be targeted accordingly.

This research can also assist policymakers in implementing lockdown policies and business-related economic disruptions to a minimum. Policymakers try to design non-pharmaceutical responses to epidemics under twin objectives, (a) contain or reduce disease spread and (b) generate minimal disruption to people's lives and business operations. If all they want to do is contain the epidemic, then indiscriminate lockdowns are probably fine. Our work suggests that if objective (b) is considered, the knowledge that businesses in rich and poor urban areas withstand mobility shocks differently can help decide which companies and locations to place under lockdown and which to spare.

While this paper does not explicitly investigate variations in mobility among individuals with different income levels, it can serve as a valuable foundation for comprehending inequities. Suppose the frequency of business visits within a specific zip code indicates its residents' mobility. In that case, this research suggests that affluent individuals were able to decrease visits that involved prolonged indoor stays throughout the pandemic. Considering indoor visits increase the risk of contracting COVID-19, such behavior could be considered a health-conscious measure or a defensive investment. In such instances, policies should be formulated to address this inequity.

**Limitations.** Below, we record several limitations of our study. First, the analysis is based on changes in the number of visits, not spending – we do not have granular data on spending at each visit. The lack of data on expenditures is a significant limitation. For example, residents in low-income neighborhoods may be unable to buy certain items in bulk (hence, have to make more visits) compared to their counterparts in more affluent areas. Similarly, we do not have access to data on the modes of transportation used: some visits could be on foot, some on bicycles or private vehicles, even as public modes of transportation such as public buses and trains had stopped operating.

Second, externally defined job classifications, such as the “essential worker” category, may have contributed to the observed visit patterns. Third, when it comes to restaurant and grocery-store visits, it bears mention that residents of the more affluent neighborhoods are more likely to have used delivery options such as Uber Eats. By the same token, an Uber Eats driver could have executed multiple orders to a restaurant in one visit.

**Future research.** Future research could explore how mobility changed based on various metrics, such as the most common jobs held by people in different zip codes. It could also look at changes in people's spending or modes of transportation, spending patterns, business revenue, and associated job losses to assess the economic impact of the policy. Although this paper is not designed to study changes in mobility across high and low-income individuals, it could be a good starting point for understanding the inequity. Lastly, there is potential for different mobility patterns between richer and poorer zip codes in states other than Minnesota due to various states' unique approaches to mitigating COVID-19 exposure.

### Data availability

The data supporting this study's findings are available from SafeGraph ([www.safegraph.com](http://www.safegraph.com)) upon academic request.

### Code availability

No custom code was used to generate the results. Any code used will be made available upon e-mail request to any authors.

Received: 17 February 2023; Accepted: 7 September 2023;

Published online: 17 October 2023

### Notes

- 1 There is a lot of consistency in urban mobility at different hours, days, and weeks (Nemeškal et al. 2020; Sevtsuk and Ratti 2010).
- 2 Sometimes, it is not up to the individual: during the pandemic, many low-income workers were classified as doing “critical” work, and as such, were not allowed to work from home.
- 3 Bundorf et al. (2021) present evidence of large activity reductions in the presence of lockdowns in the U.S.: “40 percent of people reported reducing their activity by a lot for grocery shopping, while 79 percent of people reported reducing their activity by a lot for restaurant visits, consistent with grocery shopping being a more essential activity.” Galeazzi et al. (2021) find for France, Italy, and UK, lockdowns create “smallworldness —i.e., a substantial reduction of long-range connections in favor of local paths.” Cronin and Evans (2020) study foot traffic data for the US and find self-regulating behavior on the part of customers resulting from the changed calculus explains more than three-quarters of the decline in foot traffic in most industries; restrictive regulation explains half the decline.
- 4 McKinsey reports that until March 2020, consumers roughly spent even amounts between retail outlets (such as grocery stores and supermarkets) and food-service companies (such as restaurants, hospitals, and schools). After the pandemic started and mobility restrictions were put in place, retail outlets saw a 29% boost in sales while sales declined by 27% at restaurants, fast-food locations, coffee venues, and casual dining locations (Felix et al. 2021). Williams (2020) at the Center for Research on the Wisconsin Economy studies the buying patterns of Wisconsin consumers and documents a shift toward on-line activity during the pandemic. As the pandemic spread, in-store sales were down 30% in March while online sales rose 25%. Specifically, in-store transactions in Wisconsin were down over 40% for the last three weeks of March, “roughly consistent with the measures of activity from foot traffic”.
- 5 Shin et al. (2021) report significant changes in foot-traffic behavior among Seoul residents. They find that reduced visitations are driven by temporary business closures rather than citizens' risk-avoidance behavior.
- 6 Sevtsuk et al. (2021) study mobility patterns for a single town, Somerville, Massachusetts, and find that the most heavily impacted individual business types included furniture stores (NAICS 442), clothing stores (NAICS 448), and hotels (NAICS 721), which lost 71–78% of their visits compared to the same period the year before. Hardware and home improvement stores (NAICS 444) and grocery stores



- (NAICS 445) were some of the least impacted types of establishments losing 29% and 46% of visits, on average, respectively.
- 7 Apedo-Amah et al. (2020) document the severe, widespread, and persistent negative impact on sales across firms worldwide.
  - 8 Fernández-Villaverde and Jones (2020) use Google Mobility data and argue it has several key advantages over sales or GDP-related measures: available at a daily frequency rather than quarterly or monthly; reported with a very short lag; and available at a very disaggregated geographic level. Kim et al. (2022) use individual-level monthly panel data of Singaporeans and find households with above-median net worth (who presumably reside in high-income neighborhoods) reduced their spending more than households with below-median net worth. Alexander and Karger (2021) found county-level measures of mobility declined 6–7% rapidly after a stay-at-home order went into effect and such orders caused “large reductions in spending in sectors associated with mobility: small businesses and large retail chains.” Weill et al. (2020) performed a panel regression analysis on four aggregated mobility metrics by census tract and county across the U.S. and concluded high-income areas had larger reductions in mobility than low-income areas.
  - 9 Target 11.5 of the SDG 11 asks governments to “substantially decrease the direct economic losses relative to global gross domestic product caused by disasters, including water-related disasters, with a focus on protecting the poor and people in vulnerable situations”.
  - 10 While it is not possible to gather device-level demographic data, SafeGraph provides estimates regarding the census block group to which the device owner’s home belongs. The data is well-sampled across household income averages, educational attainment, and demographic categories, according to the aggregate summary of SafeGraph data and the characteristics of a Census block group (Squire 2019).
  - 11 Strictly speaking, SafeGraph’s sample is not truly random and is subject to sampling errors. The company has provided evidence that the sample correlates highly with U.S. census data.
  - 12 The mode of travel to a POI is not recorded in SafeGraph data, only the number of visits to it. We do not know if a specific person visited a particular POI, A, in 2019 and did or did not visit A in 2020. We also have no data on spending at a POI, only the amount of time spent. Also, our unit of analysis is the zip code.
  - 13 We look at the number of visits to restaurants by duration and find that visits of less than five minutes, which include pick-up and delivery orders, did not increase significantly during the lockdown. Unfortunately, we are unable to distinguish between delivery orders and pick-up orders, as a single driver can pick up multiple delivery orders. The change is plotted in the appendix, Figures A2–A3.
  - 14 We also conducted a sensitivity analysis by changing the thresholds for high- and low-income zip codes from one-third to one-fourth. In that case, only 216 zip codes were reclassified as high-income and 216 as low-income. We also changed the threshold from one-third to two-fifths, so there were 344 high-income and 344 low-income zip codes. After constructing distributions based on the reclassifications, we found the results were similar to the baseline one-third threshold, the one we chose to report.
  - 15 For the first re-opening period, restaurants, bars, theaters, gyms, amusement parks, and personal care services faced an extension of temporary closure (starting from June 2020). But restaurants were able to provide outdoor service provided they met some requirements. Order 20–63 allowed bars and restaurants to serve food outdoors beginning on June 1, 2020. So there were periods when indoor dining was prohibited, but outdoor dining was allowed. However, regardless of the median income of the zip code, everyone in Minnesota was subject to the same order. To our knowledge, no restrictions differed by location (e.g., streets).
- ## References
- Alexander D, Karger E (2021) Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior. *Rev Econ Stat*, 1–25. [https://doi.org/10.1162/rest\\_a\\_01108](https://doi.org/10.1162/rest_a_01108)
- Apedo-Amah MC, Avdiu B, Cirera X, Cruz M, Davies E, Grover A, Iacovone L, Kilinc U, Medvedev D, Maduko FO, Poupakis S, Torres J, Tran TT (2020) Unmasking the impact of COVID-19 on businesses: firm level evidence from across the world. <https://doi.org/10.1596/1813-9450-9434>
- Bonaccorsi G, Pierri F, Cinelli M, Flori A, Galeazzi A, Porcelli F, Lucia A, Michele C, Scala A, Quattrocchi W (2020) Economic and social consequences of human mobility restrictions under COVID-19. *Proc Natl Acad Sci* 117(27):15530–15535. <https://doi.org/10.1073/pnas.2007658117>
- Bundorf MK, DeMatteis J, Miller G, Polyakova M, Streeter J, Wivagg J (2021). Risk perceptions and protective behaviors: evidence from COVID-19 pandemic. In *Ssrn*. <https://doi.org/10.3386/w28741>
- Cronin CJ, Evans WN. (2020). Private precaution and public restrictions: what drives social distancing and industry foot traffic in the COVID-19 era? <https://doi.org/10.3386/W27531>
- Dingel JI, Neiman B (2020) How many jobs can be done at home? *J Publ Econ* 189:104235. <https://doi.org/10.1016/j.jpubeco.2020.104235>
- Erchick DJ, Zapf AJ, Baral P, Edwards J, Mehta SH, Solomon SS, Gibson DG, Agarwal S, Labrique AB (2021) COVID-19 risk perceptions of social interaction and essential activities and inequity in the United States: results from a nationally representative survey. *MedRxiv*, 2021.01.30.21250705. <https://doi.org/10.1101/2021.01.30.21250705>
- Felix I, Martin A, Mehta V, Mueller C (2021) US food supply chain: disruption and implication from Covid-19. *McKinsey & Company*, July, 1–13. <https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/us-food-supply-chain-disruptions-and-implications-from-covid-19>
- Fernández-Villaverde J, Jones CI (2020) Macroeconomic outcomes and COVID-19: A progress report. <https://doi.org/10.3386/W28004>
- Galeazzi A, Cinelli M, Bonaccorsi G, Pierri F, Schmidt AL, Scala A, Pammolli F, Quattrocchi W (2021) Human mobility in response to COVID-19 in France, Italy and UK. *Sci Rep* 11(1):1–10. <https://doi.org/10.1038/s41598-021-92399-2>
- Gao S, Rao J, Kang Y, Liang Y, Kruse J, Dopfer D, Sethi AK, Mandujano Reyes JF, Yandell BS, Patz JA (2020) Association of mobile phone location data indications of travel and stay-at-home mandates with COVID-19 infection rates in the US. *JAMA Network Open* 3(9):e2020485. <https://doi.org/10.1001/jamanetworkopen.2020.20485>
- González MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. *Nature* 453(7196):779–782. <https://doi.org/10.1038/nature06958>
- Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, Druckenmiller H, Huang LY, Hultgren A, Krasovich E, Lau P, Lee J, Rolf E, Tseng J, Wu T(2020)The effect of large-scale anti-contagion policies on the COVID-19 pandemic.*Nature* 584(7820):262–267. <https://doi.org/10.1038/s41586-020-2404-8>
- Huang X, Li Z, Jiang Y, Li X, Porter D (2020) Twitter reveals human mobility dynamics during the COVID-19 pandemic. *PLoS ONE*, 15(Nov). <https://doi.org/10.1371/journal.pone.0241957>
- Iio K, Guo X, Kong X, Rees K, Bruce Wang X (2021) COVID-19 and social distancing: disparities in mobility adaptation between income groups. *Transp Res Interdiscip Perspect*, 10. <https://doi.org/10.1016/j.trip.2021.100333>
- Jay J, Bor J, Nsoesie EO, Lipson SK, Jones DK, Galea S, Raifman J (2020) Neighbourhood income and physical distancing during the COVID-19 pandemic in the United States. *Nat Hum Behav*, 4. <https://doi.org/10.1038/s41562-020-00998-2>
- Kim J, Kwan MP (2021) The impact of the COVID-19 pandemic on people’s mobility: a longitudinal study of the U.S. from March to September of 2020. *J Transp Geogr*, 93. <https://doi.org/10.1016/j.jtrangeo.2021.103039>
- Kim S, Koh K, Zhang X (2022) Short-term impact of COVID-19 on consumption spending and its underlying mechanisms: evidence from Singapore. *Can J Econ/Revue Canadienne d’économie* 55(S1):115–134. <https://doi.org/10.1111/CAJE.12538>
- Koren M, Petó R (2020) Business disruptions from social distancing. *PLOS ONE* 15(9):e0239113. <https://doi.org/10.1371/JOURNAL.PONE.0239113>
- Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, du Plessis L, Faria NR, Li R, Hanage WP, Brownstein JS, Layan M, Vespignani A, Tian H, Dye C, Pybus OG, Scarpino SV (2020) The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368(6490):493–497. [https://doi.org/10.1126/SCIENCE.ABB4218/SUPPL\\_FILE/PAPV2.PDF](https://doi.org/10.1126/SCIENCE.ABB4218/SUPPL_FILE/PAPV2.PDF)
- Lee M, Zhao J, Sun Q, Pan Y, Zhou W, Xiong C, Zhang L (2020) Human mobility trends during the early stage of the COVID-19 pandemic in the United States. *PLOS ONE* 15(11):e0241468. <https://doi.org/10.1371/JOURNAL.PONE.0241468>
- Li W, Wang Q, Liu Y, Small ML, Gao J (2022) A spatiotemporal decay model of human mobility when facing large-scale crises. *Proc Natl Acad Sci USA* 119(33):e2203042119. [https://doi.org/10.1073/PNAS.2203042119/SUPPL\\_FILE/PNAS.2203042119.SAPP.PDF](https://doi.org/10.1073/PNAS.2203042119/SUPPL_FILE/PNAS.2203042119.SAPP.PDF)
- Lu J, Zhou S, Liu L, Li Q (2021) You are where you go: inferring residents’ income level through daily activity and geographic exposure. *Cities* 111:102984. <https://doi.org/10.1016/j.cities.2020.102984>
- Nemeškal J, Ouředníček M, Pospíšilová L (2020) Temporality of urban space: daily rhythms of a typical week day in the Prague metropolitan area. *J Maps* 16(1):30–39. [https://doi.org/10.1080/17445647.2019.1709577/SUPPL\\_FILE/TJOM\\_A\\_1709577\\_SM3437.PDF](https://doi.org/10.1080/17445647.2019.1709577/SUPPL_FILE/TJOM_A_1709577_SM3437.PDF)
- Ruiz-Euler A, Privitera F, Giuffrida D, Zara I (2020) Mobility patterns and income distribution in times of crisis. In *SSRN Electronic Journal*. <https://ssrn.com/abstract=3572324>
- Sevtsuk A, Hudson A, Halpern D, Basu R, Ng K, de Jong J (2021) The impact of COVID-19 on trips to urban amenities: examining travel behavior changes in Somerville, MA. *PLOS ONE* 16(9):e0252794. <https://doi.org/10.1371/JOURNAL.PONE.0252794>
- Sevtsuk A, Ratti C (2010) Does urban mobility have a daily routine? Learning from the aggregate data of mobile networks. 17(1), 41–60. <https://doi.org/10.1080/10630731003597322>
- Shin J, Kim S, Koh K (2021) Economic impact of targeted government responses to COVID-19: evidence from the large-scale clusters in Seoul. *J Econ Behav Org* 192:199–221. <https://doi.org/10.1016/j.jebo.2021.10.013>

- Squire RF (2019) What about bias in the SafeGraph dataset? <https://www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset>
- U.S. Census Bureau (2020) American Community Survey 2015–2019 5-Year Data
- U.S. Census Bureau (2022) North American Industry Classification System (NAICS). <https://www.census.gov/naics/>
- Weill JA, Stigler M, Deschenes O, Springborn MR (2020) Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proc Natl Acad Sci USA* 117(33):19658–19660. <https://doi.org/10.1073/pnas.2009412117>
- Weill J, Stigler M, Deschenes O, Springborn M (2021) COVID-19 Mobility policies impacts: how credible are difference-in-differences estimates? *SSRN Electron J*. <https://doi.org/10.2139/ssrn.3896512>
- Wellenius GA, Vispute S, Espinosa V, Fabrikant A, Tsai TC, Hennessy J, Dai A, Williams B, Gadepalli K, Boulanger A, Pearce A, Kamath C, Schlosberg A, Bendebury C, Mandayam C, Stanton C, Bavadekar S, Pluntke C, Desfontaines D, Gabrilovich E(2021) Impacts of social distancing policies on mobility and COVID-19 case growth in the US. *Nat Commun* 12(1):1–7. <https://doi.org/10.1038/s41467-021-23404-5>
- Williams N (2020) The Wisconsin Economy During COVID-19: Lockdown and Reopening
- Yabe T, García B, Bueno B, Dong X, Pentland A, Moro E (2023) Behavioral changes during the COVID-19 pandemic decreased income diversity of urban encounters. *Nat Commun* 14:1–10. <https://doi.org/10.1038/s41467-023-37913-y>

### Author contributions

The paper was jointly conceived, and all four authors who contributed to the writing analyzed the data.

### Competing interests

The authors declare no competing interests.

### Ethical approval

This article does not contain any studies with human participants performed by the authors.

### Informed consent

This article does not contain any studies with human participants performed by any authors; therefore, no informed consent is needed.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1057/s41599-023-02116-9>.

**Correspondence** and requests for materials should be addressed to Joydeep Bhattacharya.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023