JUE insights: Does mobility explain why slums were hit harder by COVID-19 in Mumbai, India?

Jaymee Sheng [a], Anup Malani [b,*], Ashish Goel [a], Purushotham Botla [c]

[a] Stanford University, United States
[b] University of Chicago, United States
[c] Infinite Analytics, United States

ARTICLE INFO

ABSTRACT

SARS-CoV-2 has had a greater burden, as measured by rate of infection, in poorer communities within cities. For example, 55% of Mumbai slums residents had antibodies to COVID-19, 3.2 times the seroprevalence in non-slum areas of the city according to a sero-survey done in July 2020. One explanation is that government suppression was less severe in poorer communities, either because the poor were more likely to be exempt or unable to comply. Another explanation is that effective suppression itself accelerated the epidemic in poor neighborhoods because households are more crowded and residents share toilet and water facilities. We show there is little evidence for the first hypothesis in the context of Mumbai. Using location data from smart phones, we find that slum residents had nominally but not significantly (economically or statistically) higher mobility than non-slums prior to the sero-survey. We also find little evidence that mobility in non-slums was lower than in slums during lockdown, a subset of the period before the survey.

1. Introduction[27]

While SARS-Cov-2 (or COVID-19) has had a massive global impact on both health and economic activity, the pandemic has had an uneven burden across communities.[1] Urban areas have experienced greater rates of infection. Carozzi et al. (2020) show that COVID-19 hit earlier in US counties with greater population density.[2] Moreover, the attack rate and reproductive rate of COVID-19 increase with city size in the US (Stier et al., 2020) and Brazil (Ribeiro et al., 2020). More relevant for this paper, within urban areas, poorer communities experienced higher rates of infection. For example, poorer zip codes within New York City have experienced significantly higher rates of infection (Whittle and Diaz-Artiles, 2020).[3]

A natural question is why poorer communities in cities bear a greater burden from COVID-19. The poor face a greater burden from many other health conditions for many reasons, including low health human capital, lower income, and poor facilities (Banerjee and Duflo, 2011). However, the rapid pace at which inequalities have emerged during the epidemic suggests that there may be explanations unique to COVID-19 and associated non-pharmaceutical interventions.

One explanation is that restrictions such as shelter-in-place orders had less impact in poorer communities ("higher mobility" hypothesis).[4] There are several potential mechanisms for this rationale. First, individuals from low-income communities may be more likely to be essential workers and exempt from such orders.[5] Second, poorer individuals may have less capacity to work remotely or so little financial buffer that they are unable to survive unless they continue working outside the home. Consistent with these first two mechanisms, work-day mobility was greater during lockdown in poorer neighborhoods of New York City and Santiago Chile[6] and in poorer counties across the US

---

[1] As of this writing, SARS-CoV-2 has infected at least 140 million people and claimed 3 million lives worldwide (Worldometer, 2021).

[2] See also Leatherby (2020).

[3] Consistent with this finding, Reese (2020) finds that poorer neighborhoods of LA County and poorer counties of California had higher COVID-19 infection rates. Emeruwa et al. (2020) finds that pregnant women in poorer neighborhoods in New York City were more likely to test positive for COVID-19. Beyond just cities, Nichols et al. (2020) finds a greater incidence of COVID-19 in poorer zipcodes across the US.

---

[4] Contrary to the bulk of the early literature, Alexander and Karger (2020) finds that the effect of stay-at-home orders on consumer behavior does not vary by income.

[5] Kearney and Munana (2020) show that low-income individuals are relatively more likely to work in essential industries. However, this result is conditional on working. Lower-income individuals may also have faced greater risk of unemployment due to COVID-19 (Parker et al., 2020). Moreover, unemployed individuals may be more likely to be at home during the work day.

[6] Coven and Gupta (2020) use cell phone GPS data to show low-income residents in New York City exhibited greater likelihood to work during the day.

(Huang et al., 2020; Lou and Shen, 2020).[7] A third mechanism is that lower-income individuals may be less compliant with suppression policy for other reasons, including greater optimism about COVID risk (Fan et al., 2020).[8] Evidence for this mechanism comes from Almagro et al. (2020), which finds that mobility outside of work hours was greater in lower income areas.[9] Finally, wealthier households were more likely to flee high-prevalence cities for lower-prevalence areas. Coven and Gupta (2020) provide evidence of this phenomenon among New York City residents at the start of the epidemic.

A different explanation is that stay-at-home orders and the like may, perversely, have accelerated COVID-19 transmission in poor neighborhoods. Such orders may have trapped poor individuals in even more crowded households where infection spread rapidly ("crowded housing" hypothesis). In communities where households lack private water taps and toilets, congregation at public facilities may have spread infection. Consistent with these hypotheses, Almagro et al. (2020) shows that, at the start of the pandemic, crowded housing contributed to the COVID hospitalization rate in New York, though less than mobility did. Once lockdowns led to job losses, however, it finds that the relative importance of household crowding versus mobility doubled.[10]

The two explanations for faster spread of COVID-19 in poor communities are not mutually exclusive. Almagro et al. (2020), for example, finds a role of both mobility and crowded housing. However, the two explanations have divergent implications for policy. The differential mobility explanation recommends greater enforcement of suppression in poor areas. By contrast, the crowded housing explanation does not recommend as strict enforcement of suppression in poor areas. Therefore, it is critical to determine the more important driver of divergent COVID-19 outcomes across neighborhoods.

We consider the greater incidence of COVID-19 in poor communities in the context of India. India has had the second highest number of reported cases in the world. According to Worldometer (2021), India has had over 14 million cases as of this writing. If, as serological studies suggest, the number of actual cases is 40 to 100 times greater than confirmed cases (Malani et al., 2020; Mohanan et al., 2020; Malani et al., 2021), India likely has had the highest number of infections of any country worldwide. Within India, cities and towns have greater infection rates than rural areas. For example, by August 2020, the population prevalence of antibodies (seroprevalence) was 10 percentage points higher in urban areas than rural areas of the state of Karnataka (Mohanan et al., 2020).[11] Moreover, within Indian cities, poor neighborhoods have a sharply higher reproductive rate. As early as mid July 2020, seroprevalence in the slums of Mumbai was 55%, roughly 3.2 times greater in the non-slums (17%) areas of the city (Malani et al., 2020). COVID has spread shockingly fast in slums despite the fact that India has one of the world's strictest lockdowns. According to Our World

in Data, India had a more stringent lockdown than any large, high-income country through July 1, 2020.[12]

This paper tests the higher mobility hypothesis as an explanation for the disproportionate incidence of COVID in Mumbai's slums using data from smart phones.[13] We employ location information from applications (apps) on these phones to pinpoint the location of users' homes and their within-city travel. We first compare the mobility patterns of slum residents versus non-slum residents from the start of the pandemic to the date of the seroprevalence survey by Malani et al. (2020). We employ a difference-in-difference (DD) design that compares the change in mobility from (a) a "baseline" period prior to the pandemic-induced fall in mobility to (b) a "pandemic" period starting after baseline and lasting until the sero-survey was done across the two communities. Second, we test whether India's lockdown was less effective at reducing mobility in slums. Again we employ a DD design that examines the difference between mobility during (a) the pandemic period during the lockdown (24 March–1 June 2020) and (b) a "non-lockdown" period defined as the portion of the pandemic period excluding the lockdown.

We find that slum residents had nominally higher mobility, measured either by the number of unique locations visited or the number of trips made, than non-slum residents in the period before the Malani et al. (2020) sero-study (relative to a baseline period before the pandemic affected mobility). However, the differences are often not statistically significant and typically not economically as well. Interestingly, while the mean level of mobility is higher among slum residents, the median level is not.

We obtain similar findings when we focus on the lockdown, a subset of the period before the sero-study. First, we compare the lockdown period to the remainder of the pandemic period. Second, we change the benchmark and examine relative mobility as between the lockdown period and the baseline period just before the pandemic impacted mobility. In either comparison, slums had nominally higher, but not significantly higher mobility.

Beyond providing evidence on relative mobility during the pandemic and the relative efficacy of suppression in the context of Mumbai, this paper makes two other contributions. First, it highlights the importance of city- and county-specific heterogeneity in the impact of social distancing on different socioeconomic groups. Second, it provides fine-grained measures of home location and mobility using a novel data set of location grabs from smart phones in India.

Section 2 describes our data, how we identify whether a smart phone user lives in a slum or not, and how we measure mobility over time. Section 3 presents our result on differences in mobility in wealthier and poorer communities. Finally, Section 4 connects our finding to two explanations for unequal burden of COVID-19 risk across socioeconomic groups.

2. Methodology

2.1. Data

We primarily use two types of data. We infer mobility and residential location from smart phone location data. We classify residential community type (slum or non-slum) based on shapefiles of slums.

Smart phone location data. Smart phone or device location is obtained from Infinite Analytics (IA), a marketing analytics company. IA purchases data from data vendors that gather device locations for users that permit location services. GPS coordinates of a device are collected when specific applications (usually in the gaming, news, and utilities categories) are in use. We analyze location data from 1 March 2020 to 31 October 2020 because this covers a period from before India's lock-

---

Carranza et al. (2020) show that mobility fell less in low-income neighborhoods than higher income neighborhoods of Santiago, Chile. Bennett (2020) finds that lockdowns reduced cases more in wealthier sections of Santiago, Chile.

[7] Chiou and Tucker (2020) suggests that the difference in mobility across income groups may be due to lack of access to high-speed internet among the poor. While working from home during COVID-19 is positively correlated with income, that correlation disappears when controlling for access to high-speed internet.

[8] Fan et al. (2020) finds low income households not only believe that the risk from COVID is lower, they also believe that suppression orders are less effective.

[9] However, lower-income households may be less able to afford food delivery at home. Moreover, essential workers may have longer hours that do not overlap the traditional 9-5 work shift (Coven and Gupta, 2020).

[10] Relatedly, Emeruwa et al. (2020) finds that, among pregnant women in New York City, those living in crowded households were more likely to test positive for COVID-19.

[11] A similar pattern was observed in the state of Tamil Nadu (Malani et al., 2021) and more broadly in India (Murhekar et al., 2021).

---

[12] Among the countries with the 10 highest numbers of total cases, only Brazil had a more strict lockdown according to this measure.

[13] We do not directly test the crowded housing hypothesis.

down to a period well past both the release from lockdown and serological surveys in Mumbai. Each observation in these data is an anonymized device ID,[14] timestamp and location; most devices can be seen multiple times on different dates and locations.

To reduce the computation required to analyze location data, we reduce the resolution of that data. Specifically, we convert raw latitude-longitude locations into Uber H3 ("Uber") hexagons.[15] Uber, the ride-share company, has developed a system of partitioning the surface of the globe using hexagons for demarcating location in its ride-share application.[16]

An important choice we face is the size of each hexagonal cell (Fig. A.1). The larger the size of cells (with larger cells having lower H-prefix numbers), the fewer slums we can identify, because many slums are smaller than a given-sized Uber cell and we characterize a cell as a slum if its center is contained in a slum. However, the larger a cell, the more devices for which we can identify a home location because it is more likely we can see the device repeatedly showing up in a location at night. A similar trade-off exists for measuring mobility. The larger the cells, the less we are able to detect shorter trips. However, the larger cells the less likely we are to confuse imperfect triangulation of location as movement outside a home. Note that one need not use the same resolution for identifying residential location and community type as for measuring mobility.

We chose a default Uber H3 resolution of 12 (H12 for short) for determining residential location. This produces hexagons that on average have an edge length of approximately 9.4 m and an area of approximately 307 m squared. We chose a default resolution of 10 for determining mobility. This produces hexagons that have an edge length of approximately 66 m and an area of 15,000 m squared. However, we vary cell size for both residential location and mobility measures in our sensitivity analysis. Our default choice yields results with respect to the differential mobility across slums and non-slums that are close to average across the different cell sizes we consider.

We further process the data by removing duplicate records as well as simultaneous records of multiple Uber H9 locations associated with the same device. 8.0 million devices in the processed data set have more than one observation, i.e., a timestamp-location record; 4.5 million devices are observed on at least 3 days.

Slum location data. It is a challenge to identify the location of slums as these communities are informal and both emerge and disappear with some frequency. Our approach is to use three sources of geospatial data (shapefiles) on slums in Mumbai, both the city and its suburbs. Each data set set may have false positives and false negatives. False positives arise, e.g., because slums were identified at some point in the past and have since been demolished or abandoned and replaced with non-slum developments. False negatives arise because some slums are missed by the screening process used to find slums. On the assumption that disagreement signals greater likelihood of error, we will use the intersection of the three data sets, i.e., classify areas as slums (non-slums) if they are slums (non-slums) in all three data sets.

The first data set is directly from the Mumbai government.[17] The Mumbai government data records slums that are confirmed by visits from government officials, but may not capture all slums. Only slums that have been officially recognized are included; newer or smaller slums, which may include the poorest households, are omitted. The data set is from 2016, and so some of the slums may have been rehabilitated into formal apartments or relocated to formal apartments elsewhere since that time.[18] The second set of data, from Taubenböck and Kraff (2014), are derived from application of an algorithm that attempts to identify slums using satellite imagery. Slums are identified based on building size, density and height. The algorithm may miss slums if slum buildings have common roofs or if there is dense cloud cover. The third source, from Gandhi et al. (2014), takes an earlier (2012) version of the Mumbai government data and supplements it with slums in areas that the Mumbai government did not screen. Whereas the Mumbai government only recognized slums on state or municipal government land, this source also included slums in so-called Special Planning Areas.

To make the data sets comparable, we map each shapefile onto the Uber hexagons. We do this in three steps. First, we convert slum polygons in each shapefile into H14 hexagons (roughly 6.3m^2) whose centers are contained in the polygon. Second, we convert all coordinates in our location data set into H14 hexagons. Third, we characterize each H14 hexagon in our data set as a sure slum (non-slum) if it is a slum (non-slum) in all three shapefiles. Table A.1 compares each shapefile to one other. It reports how many H14's have slums under both shapefiles, under neither shapefile, and under one but not the other. (This table does not report how many false positives and negatives there are in each data set, as we do not know the ground truth.)

2.2. Home location

We associate each device with a home location, defined as the most frequent Uber cell observed for a device between 8pm and 4am over the entire period from 1 March to 31 October 2020. We do this once at resolution H12 and once at resolution H10. To reduce the likelihood of assigning spurious home locations to devices, we require the same home location for each device to be observed on at least 3 different days. Finally, we remove devices with multiple home locations, which constitute only a small proportion (less than 5%) of all devices with valid home locations. This results in 607,124 devices with unique H12 home locations and 693,092 devices with unique H10 home locations satisfying the required minimum frequency.

Once a device is associated with a H10 or H12 cell, we determine whether that location is a slum by merging with slum location data as explained above. For a given Uber resolution used to define residential location, we calculate the fraction of each hexagon that are slum and non-slum H14 hexagons. For example, if we define residential location using a H12 grid, we identify all the H14 hexagons within each H12 hexagon and then determine which fraction of the H14's are sure slums and sure non-slums.[19] Because we define locations as slums or non-slums if all three slums shapefiles agree, we have a number of H14 cells that are undefined. Moreover, we will often analyze a subsample of devices that are on Uber cells with H14 cells that are all slums or all non-slums. These restrictions reduce the sample size of devices for which we can characterize home locations as slum or non-slum. Table 1 reports the total sample size of devices that remain by whether they are in slums or non-slums under these restrictions.

---

[14] While it is conceivable that someone may be able to identify individuals based on inferred home location and mobility patterns, we did not have any other personally identifiable information and made no attempt at de-anonymizing any IDs. Furthermore, we were careful with encrypting any data downloaded locally.

[15] See https://h3geo.org/.

[16] An advantage of a partition that uses hexagons over say triangles or squares is that there is only one distance between the center-point of any two adjacent cells, simplifying distance calculations.

[17] https://www.sra.gov.in//upload/MUMABI_CA_MAP_MUMBAI_GLOBAL_42X72_SLUMCLUSTER.pdf. Geojson file provided by Spatial Data of Municipalities (Maps) Project (http://projects.datameet.org/Municipal_Spatial_Data/) by Data{Meet} made available under the Creative Commons Attribution-ShareAlike 2.5 India.

[18] Mumbai has a Slum Rehabilitation Act that allows private landowners on whose land slums have arisen to enlist the help of the government to remove slums. The government will assist in evacuating slum residents if the landowner allocated sufficient flats in the development that replaces the slum to former slum residents. Slums that are recognized by the government and on public land can also be subject to rehabilitation, though the residents are typically moved to formal government apartment buildings that are not in situ.

[19] Each finer resolution cell has one seventh the area of the coarser resolution cell. Since a hexagon cannot be perfectly subdivided into seven hexagons, the finer cells are only approximately contained within a parent cell.

Table 1
Summary statistics on devices and mobility, by H3 grid used.

Outcome	Data resolution		Non-slums					Slums				
	Home	Mobility	Mean	Std Dev	Devices	Average T	N	Mean	Std Dev	Devices	Average T	N
Locations	H10	H9	1.177	3.935	441,736	31.851	14,069,826	1.413	4.220	50,897	36.394	1,852,369
	H10	H10	1.708	6.199	441,736	31.851	14,069,826	2.145	6.631	50,897	36.394	1,852,369
	H12	H9	1.197	3.952	449,778	31.850	14,325,264	1.430	4.219	52,130	36.185	1,886,327
	H12	H10	1.755	6.231	449,778	31.850	14,325,264	2.183	6.619	52,130	36.185	1,886,327
	H12	H12	3.245	11.881	449,778	31.850	14,325,264	3.965	11.856	52,130	36.185	1,886,327
Trips	H10	H9	1.049	3.807	441,736	31.851	14,069,826	1.183	3.349	50,897	36.394	1,852,369
	H10	H10	1.374	4.958	441,736	31.851	14,069,826	1.608	4.195	50,897	36.394	1,852,369
	H12	H9	1.058	3.830	449,778	31.850	14,325,264	1.193	3.324	52,130	36.185	1,886,327
	H12	H10	1.379	4.963	449,778	31.850	14,325,264	1.601	4.105	52,130	36.185	1,886,327
	H12	H12	1.678	5.791	449,778	31.850	14,325,264	1.966	5.066	52,130	36.185	1,886,327

2.3. Measures of mobility

We use two distinct metrics to assess the mobility of slum and non-slum populations in our dataset: the number of unique Uber cells (other than the home location) that each device travels to and the number of trips a device makes between its home location and other locations. We use 3 resolutions at which to compute mobility: H9, H10 or H12. We compute these metrics on a daily basis for each device. Because each device is not observed each day, we leave missing measurements on mobility for any day a given device is not observed.

2.4. Policy periods

We create a number of binary indicator variables for different control and treatment periods. We define our baseline period as 1 March–14 March 2020. This is before lockdown and the major reduction in mobility due to COVID-19. The Malani et al. (2020) sero-survey was conducted from 29 June–19 July 2020. We define the pandemic period as the period after the baseline but before the end of the sero-survey.

During a subset of the pandemic period, there was a lockdown. The national mandatory lockdown, applicable also to Mumbai, was implemented on 24 March 2020. (There was a voluntary lockdown on Sunday, 22 March 2020.) The lockdown was repeatedly extended in roughly two-week increments from 17 April until 1 June 2020. Then, on 8 June 2020, the government announced a series of "unlocks". Therefore, we define the lockdown as 24 March–1 June 2020.

2.5. Final data set

Our final data set is an unbalanced panel with missing data (Table 1). Our units are devices and our dates are days between 1 March and 31 October 2020. Although each device has a probabilistic home community (slum or non-slum) type with probabilities depending on home location resolution chosen (H10 or H12), we keep only those devices that we assess as having a probability of 1 of being in a slum or of being in a non-slum. Many device-day combinations are missing because we do not observe each device each day. The mobility level depends both on what is measured (locations visited or trips made) and the resolution used to measure mobility (H9, H10 or H12). Each device may be observed for multiple days, but no device is observed every day; the average device is observed on roughly 30 days (Table 1) and Fig. A.2 gives the distribution of number of days per device.

2.6. Regression analysis

We employ regression analysis to estimate the difference between mobility response in slums and non-slums. This analysis attempts to answer two questions.

1. Did slum residents have greater mobility during the pandemic?

2. Did lockdown reduce the mobility of slum residents less than mobility of non-slum residents?

We employ difference-in-difference designs to answer these questions. Table 2 provides an overview of our different designs.

Differential effect of pandemic. To determine if slums had greater mobility during the pandemic we define two periods: a baseline period (1–14 March 2020) and a pandemic period (15 March–19 July 2020). We end the pandemic period on 19 July 2020 because that is the last day of sampling for the Mumbai sero-survey and we are sure that sero-prevalence is higher in slums until the sero-survey.[20] We restrict our sample to the baseline and pandemic period and estimate the following regression:

$$m_{it} = \beta_0 + \beta_1 \text{pandemic}_t + \beta_2 \text{slum}_i + \beta_3 (\text{pandemic}_t \times \text{slum}_i) + e_{it} \qquad (1)$$

where i indexes devices, t indexes days, and pandemic is an indicator for the pandemic period. In all cases, the error is clustered at the Uber cell at which home location is defined because mobility may be serially correlated.

We estimate this regression in three ways. One is ordinary least squares (OLS). Second, we add random effects to reduce the risk that OLS reflects the differential composition of devices observed during the pandemic versus baseline.[21] Third, we estimate a quantile regression. If there is skew in mobility among devices, the mean will give a misleading picture of disease risk: a small number of devices may be at extreme risk while a large number are not.[22]

Differential effect of lockdown. To determine if lockdown is responsible for a decline in mobility and if the effect of lockdown is smaller in slums, we subdivide the pandemic into two periods: a lockdown period (24 March–1 June 2020) and a non-lock down period (15–23 March and 2 June–19 July 2020) after baseline. In our primary analysis, we restrict our sample to these two sub-periods and compare mobility across them. In our sensitivity analysis, we compare the lockdown period to the baseline period defined above. We also try different masures of the lockdown period. In one we account for the fact that it took approximately 1 week for the government to implement the lockdown and in another we extended the lockdown period to 7 June because the formal "Unlock 1.0" policy began June 2020. In all variations, we estimate a

---

[20] Defining the pandemic period to end on the last day of the sero-survey is conservative in the sense that it takes several days after infection to generate antibodies. We try other variants in Table 9.

[21] A fixed effect regression would be better but it is difficult to implement because our hypothesis tests require calculating the average of roughly 50,000 fixed effects for devices in slums.

[22] Our position is not that only the median matters. It is possible that COVID-19 has high k, meaning superspreaders. But if that is not the case, e.g., there are superspreader events rather than persons, then a median model may be more appropriate.

Table 2
Difference-in-difference regression designs.

Question	Specification	Control period	Treatment period	Table
Pandemic	Main	Baseline (1–14 March)	After baseline - end of sero-survey (15 March - 19 July)	3, A2
	Before survey	Baseline (1–14 March)	After baseline - before sero-survey (15 March - 28 June)	A1
Lockdown	Main	After baseline but not lockdown (15–23 March, 2 June–19 July)	Lockdown (24 March–1 June)	4, A3
	Baseline control	Baseline (1–14 March)	Lockdown (24 March–1 June)	5
	Before survey	After baseline but not lockdown (15–23 March, 2 June–28 June)	Lockdown (24 March–1 June)	A1
	Longer lockdown	After baseline but not lockdown (15–23 March, 2 June–19 July)	Lockdown (24 March–8 June)	A1
	Later lockdown	After baseline but not lockdown (15–23 March, 2 June–19 July)	Lockdown (1 April–1 June)	A1

regression of the form:

$$m_{it} = \beta_0 + \beta_1 \text{lockdown}_t + \beta_2 \text{slum}_i + \beta_3(\text{lockdown}_t \times \text{slum}_i) + e_{it} \quad (2)$$

The error is again clustered at the Uber cell at which home location is defined. We estimate this regression in the same three ways we did for (1).

In both regression analyses, we measure treatment effects, i.e., differential mobility across slums and non-slums, in two ways. First, we examine the level of mobility in slums versus non-slums during the "treatment" period. Looking at levels is the right approach if a slum resident and a non-slum resident with the same level of mobility produce the same number of "pings" on their smartphone. Second, we examine the percent change in mobility among slum versus non-slum residents relative to the baseline period. This is the right approach if slum and non-slum residents with the same level of mobility do not produce the same number of pings. This is possible if, for example, for a given level of mobility slum residents are less likely to turn on location services on apps or to use apps with location services turned on.

3. Results

Whether we look at the raw data via graphical analysis or the regression results, we see that slums have nominally higher average mobility during the pandemic and even during lockdown, but the differences are typically statistically insignificant or economically insignificant. Even these nominal differences go away if we account for the fact that it took some time for the lockdown to be enforced or we focus on mobility before the sero-survey started rather than ended. Indeed, if we examine the mobility of the median device, we see no difference across the two communities regardless of how we define treatment and control periods.

3.1. Graphical analysis

Average mobility. Visualization of the average number of unique locations visited provides modest evidence that mobility was higher in slums prior to the sero-survey, but less evidence on the differential effect of the lockdown (Fig. 1). Non-slum residents and to some extent slum residents reduced the daily number of locations visited prior to the announcement of the first nationwide lockdown on 24 March 2020. The average number of locations per day by slum and non-slum dwellers decreased from around 3.5 and 4.5, respectively, to around 1 by the time the lockdown went into effect and stayed at that level until early June when some of the lockdown restrictions began to be lifted.

Although the number of locations visited during lockdown is similar across both communities, non-slum areas have a somewhat lower level of mobility during the first week of lockdown. Moreover, the decline in locations visited during the bulk of the lockdown is a larger percentage reduction for non-slum residents.

After release of lockdown, locations visited rose for both slum residents and non-slum residents. The level of mobility after release is larger for slums; the same is true of the percentage return relative to baseline: slums return to nearly their baseline levels, while non-slums claw back only 50% of their baseline mobility. That said, the big post-lockdown

gaps between slums and non-slums show up during the sero-survey and are unlikely to be the cause of the sero-survey results because it takes > 1 week for an infection to trigger detectable antibody levels in respondents (Long et al., 2020; Okba et al., 2020; Zhao et al., 2020). These patterns are repeated regardless of how we define home location and measure mobility.

With the exception of a short period just before lockdown, the average number of trips (as opposed to locations visited) taken suggests that, if anything, non-slums had a greater level of mobility in the relevant periods. Non-slums reduced the average number of trips taken from roughly 2 to 1 at the end of baseline through the first week of lockdown; but trips recovered to nearly baseline levels through lockdown and until the start of the sero-survey period. During the study, the number of trips increased above baseline levels, but too late to affect sero-survey results. These patterns are repeated regardless of how we define home location and measure mobility.

Median mobility. It has been noted that a small number of people may be responsible for a large percentage of the spread of COVID (Laxminarayan et al., 2020). In the short run this can increase the rate of spread. However, the higher the skew of contact rates in the population, the faster the reproductive rate may decline (Lloyd-Smith et al., 2005). This suggests that one may value medians as well as means as measure of population-level infection. The difference between slums and non-slums largely disappears, however, when we examine median locations visited or trips made (Fig. 2).[23] These patterns are repeated regardless of how we define home location and measure mobility.

Taken together, data on number of locations visited and trips suggests that both communities adjusted more on the margin of where they went as opposed to how often they went out. For example, perhaps they shopped for food or meals at fewer locations rather than less often. The data also suggests that there was significant skew in both measures of mobility.

3.2. Statistical analysis

Mobility during the pandemic. Regression analysis suggests that mobility was moderately higher in slums during the pandemic, defined as the period after baseline and before the end of the sero-survey. Table 3 presents our results for our default grid: identifying home location with an H12 grid and measuring mobility on an H10 grid. We found that the average number of locations visited were actually higher in slums (columns 1–2, bottom panel). When measured as average percent change from baseline, non-slums have bigger declines than slums. However, the magnitude of the difference in levels, roughly 0.32 locations in column 2, is roughly 10% of overall mobility during baseline. Moreover, if one examines medians, there is no difference between slums and non-slums (column 3).

We also found that the number of trips was higher in slums during the pandemic (column 4, $p = 0.001$). Although trips increased in both

---

[23] The 95th percentile graph, shown in Fig. A.3, looks a bit more like mean mobility.
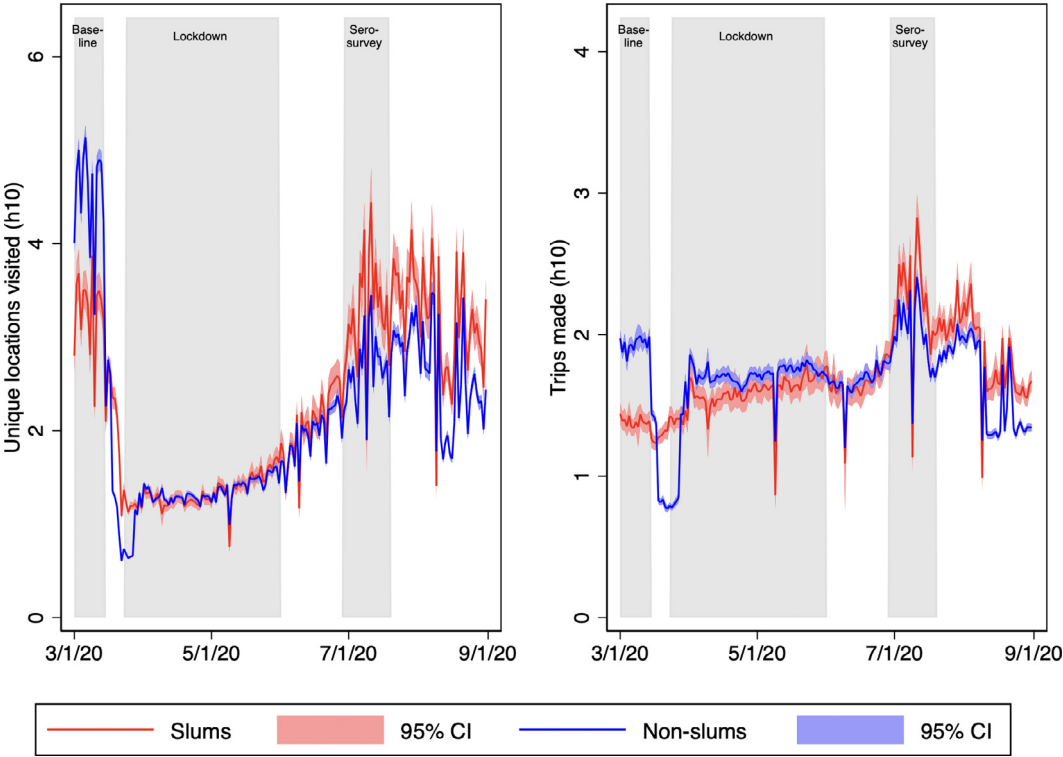
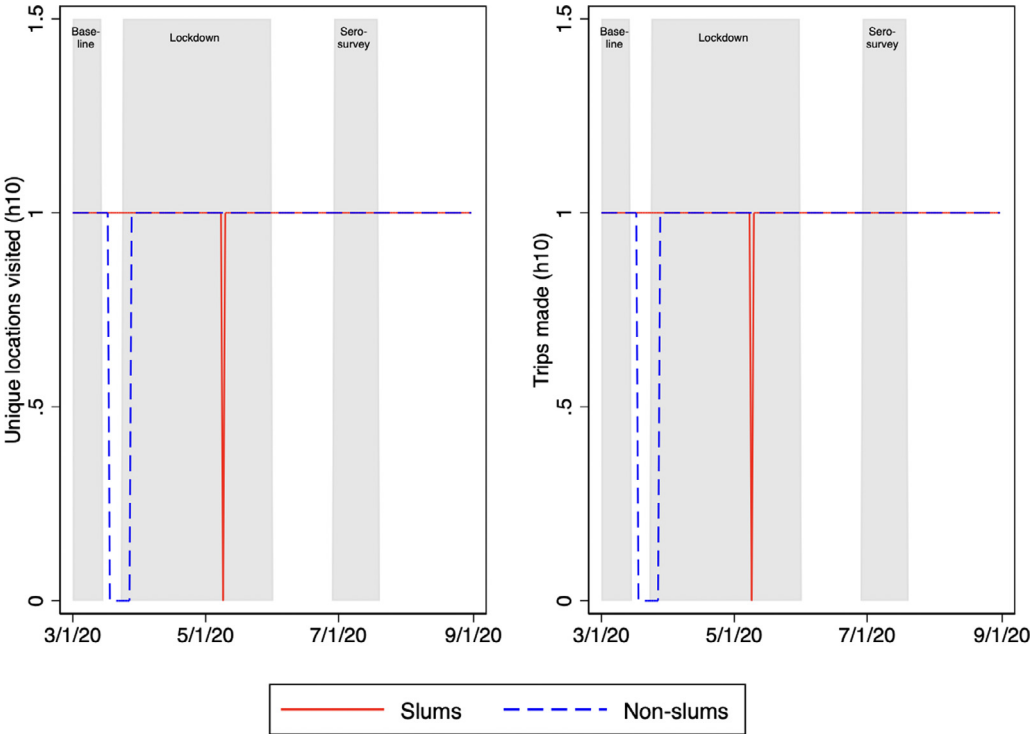**Fig. 1.** Average mobility among devices from slums and non-slums.

Fig. 2. Median mobility among devices from slums and non-slums.

Table 3
Mobility prior to sero-survey (relative to baseline).

	1	2	3	4	5	6
Estimation method	OLS	OLS with rand. effects	Quantile	OLS	OLS with rand. effects	Quantile
Dep. Variable	Unique locations	Unique locations	Unique locations	Trips	Trips	Trips
Pre-survey	-3.001***	-2.297***	0.000	-0.360**	0.189***	0.000
	0.082	0.046	0.008	0.141	0.063	0.007
Slum	-1.293***	-0.679***	0.000	-0.549***	0.067	0.000
	0.127	0.238	0.001	0.048	0.144	0.002
Pre-survey	1.460***	1.009***	0.000	0.606***	0.154**	0.000
x slum	0.141	0.075	0.009	0.178	0.074	0.008
Constant	4.547***	3.381***	1.000***	1.935***	0.940***	1.000***
	0.100	0.217	0.001	0.040	0.134	0.002
Observations	9,550,499	9,550,499	9,550,499	9,550,499	9,550,499	9,550,499
No. devices		405,194			405,194	
Level						
Slum	1.712	1.414	1.000	1.632	1.349	1.000
Non-slum	1.545	1.084	1.000	1.575	1.128	1.000
Slum - non-slum (p)	0.393	0.000	1.000	0.001	0.290	1.000
Decline						
Slum	0.474	0.477	0.000	-0.177	-0.340	0.000
Non-slum	0.660	0.679	0.000	0.186	-0.201	0.000
Slum - non-slum (p)	0.000	0.107	1.000	0.784	0.010	1.000

Notes. Home location defined on H12 and mobility on H10 grid. Cells in top panel contain coefficient and standard errors clustered on home location cells. ***/**/* indicates $p < 0.01/0.05/0.1$. Middle panel reports number of device fixed effects. Bottom panel contains measures of mobility x community type, and p values for differences across communities.

Table 4
Effect of lockdown (relative to remainder of pre-survey period) on mobility.

	1	2	3	4	5	6
Estimation method	OLS	OLS with rand. effects	Quantile	OLS	OLS with rand. effects	Quantile
Dep. Variable	Unique locations	Unique locations	Unique locations	Trips	Trips	Trips
Lockdown	-0.671***	-0.509***	0.000	0.074	0.175***	0.000
	0.185	0.116	0.004	0.070	0.015	0.003
Slum	0.434	0.477	0.000	0.182	0.273	0.000
	0.329	0.300	0.012	0.243	0.205	0.011
Lockdown	-0.367*	-0.198	0.000	-0.203***	-0.0979***	0.000
x slum	0.217	0.139	0.004	0.075	0.026	0.003
Constant	1.938***	1.392***	1.000***	1.532***	1.008***	1.000***
	0.270	0.270	0.011	0.209	0.191	0.010
Observations	8,977,712	8,977,712	8,977,712	8,977,712	8,977,712	8,977,712
No. devices		401,062			401,062	
Level						
Slum	1.334	1.161	1.000	1.585	1.359	1.000
Non-slum	1.267	0.882	1.000	1.606	1.184	1.000
Slum - non-slum (p)	0.571	0.092	1.000	0.909	0.011	1.000
Decline						
Slum	0.437	0.379	0.000	0.075	-0.060	0.000
Non-slum	0.346	0.366	0.000	-0.048	-0.174	0.000
Slum - non-slum (p)	0.068	0.562	1.000	0.020	0.388	1.000

Notes. Home location defined on H12 and mobility on H10 grid. Cells in top panel contain coefficient and standard errors clustered on home location cells. ***/**/* indicates $p < 0.01/0.05/0.1$. Middle panel reports number of device fixed effects. Bottom panel contains measures of mobility x community type, and p values for differences across communities.

slums and non-slums once you adjust for composition effects with random effects, it increased more in slums (column 5, $p = 0.01$). However, when adjusting for composition effects, the difference in levels is not significant ($p = 0.29$) and difference in change-from-baseline is only about 15% of lockdown level mobility in non-slums. Moreover, there are no differences in medians (column 6).

Differential Mobility during the lockdown. Regression analysis suggests that, relative to other periods after baseline and before the sero-study concluded, lockdowns had a significant negative effect on mobility, but that the effect was not much larger in slums. Results for our default grid are in Table 4. The effect of lockdown on average levels of or on percentage change in locations visited was not statistically sig-

nificant (columns 1 and 2, bottom panel). When using random effects, the effect of lockdown on the number of trips (column 5) was significant ($p=0.011$) but the effect on the decline in trips was not ($p=0.388$). However the differential effect is measured, lockdown has no effect on median locations or trips (columns 3 and 6).

If we compare lockdown to baseline rather than other post-baseline periods (Table A.2), we find that slums on average experience a lower reduction in visit (column 2), but the effect is not significant ($p = 0.136$). They also experience a larger increase in trips (column 5), an effect that is significant ($p = 0.017$). However, there is again no difference in median locations or trips (columns 3 and 6) even in this design. Moreover, it is unclear whether baseline should be the control for lockdown.

Sensitivity analysis. We test the effect of varying the definition of the pandemic period and lockdown period in appendix Table A.3. The table reports results from OLS with random effects.

The first two columns look at the impact of the pandemic on relative mobility. We find that redefining the pandemic period to end before the sero-survey (to reflect the fact that it takes over one week for antibodies to develop) leads us to conclude that, although mobility levels were not significantly higher in slums, mobility declines were. That said, even the declines were not economically significant (roughly 5% difference).

The last 6 columns examine the effect of lockdown on relative mobility. Redefining the pandemic period (i.e., the control period) as above causes the effect of lockdown to not significantly differ across slums and non-slums. Likewise, lengthening the lockdown period (i.e., the treatment period), delaying the start of that period, or starting that lockdown period earlier[24] does not lead to substantially different estimated effects of lockdown across communities.

To test whether the results above are a byproduct of the partition we use to locate homes and measure mobility, we consider 5 combinations of this partition. Tables A.4 and A.5 report the results of our test of whether slums have greater levels of mobility or smaller declines in mobility during the pandemic period (relative to the baseline period) and during lockdown (relative to the rest of the pandemic period), respectively, for different partitions. We find the results are roughly consistent with the results when home location is defined on an H12 grid and mobility measured on an H10 grid.

4. Conclusion and limitations

We do not find significant evidence that higher mobility can explain higher seroprevalence among Mumbai slum residents. We find that slum residents likely had greater mobility prior to the end of the sero-survey in which they were found to have 3.2 times greater seroprevalence than non-slum residents. This was true whether one examined level of mobility or the change in mobility relative to baseline. However, these differences were often not significant statistically; even when they were, they were not always significant economically. Moreover, the difference was zero when we considered the median device in each community.

It is likewise unclear that lockdowns had differential effects on slum residents' mobility. We find that slum residents had higher mobility during the lockdown. However, the effect is often not statistically and economically significant. Again, the differential effect is roughly zero when examining medians.

This paper has a number of important limitations. First, data on smart phone locations may not capture mobility in slums. While India had 439 million smartphone users by 2020, users are disproportionately the wealthy. While a significant fraction of Mumbai slum households have one smartphone,[25] it is often shared. Moreover, they may be less likely to use it, have location services turned on, or use apps with location services turned on. To some extent these concerns are addressed by looking at percentage changes in mobility within such

Table A.1
Overlap between three slum shapefiles.

		Taubenbock Shapefile		
		Slum H14	Non-slum H14	Total H14
Government Shapefile	Slum H14	3,325,507	607,471	3,932,978
	Not slum H14	1,614,019	20,302,232	21,916,251
	Total H14	4,939,526	20,909,703	
		Gandhi Shapefile		
		Slum H14	Non-slum H14	Total H14
Government Shapefile	Slum H14	3,343,658	589,320	3,932,978
	Not slum H14	987,089	20,929,162	21,916,251
	Total H14	4,330,747	21,518,482	
		Gandhi Shapefile		
		Slum H14	Non-slum H14	Total H14
Taubenbock Shapefile	Slum H14	3,478,653	1,460,873	4,939,526
	Not slum H14	852,094	20,057,609	20,909,703
	Total H14	4,330,747	21,518,482	

communities. When that is done, we still find small differences in mobility.

Second, perhaps the composition of devices that we examine varies depending on sample period. For example, even among device users in slums, perhaps those who are more mobile before the pandemic are more (less) likely to go out during the pandemic. This composition effect would cause us to overestimate (underestimate) mobility in slums during the pandemic. To mitigate this problem we use random effects, which causes us to focus on devices observed during both treatment and control periods. Even this, however, is inadequate as it is short of fixed effect estimation. (We do not run fixed effects because our hypothesis tests are difficult to run with ≈500,000 device fixed effects.)

Third, the mean or median level of mobility may not be the appropriate moment or statistic to examine to understand COVID risk from differential mobility. Perhaps there are superspreaders and we should, instead, examine the 75th or 95th percentile of devices in terms of mobility. However, dispersion in contact rates has mixed effects. While it can hasten spread of an epidemic to start, it can also hasten the decline of the reproductive rate (Lloyd-Smith et al., 2005).

Fourth, this paper does not consider and cannot provide support for, nor rule out, the crowded-housing explanation for higher seroprevalence in slums. Nor does it test for any interaction between the two explanations. As such the paper only indirectly discriminates between the first and second explanation of why the infection rate is higher in poor neighborhoods. Moreover, the study cannot show that some third factor associated with poverty (but not crowded housing) is responsible for higher seroprevalence. Nor can it rule out that voluntary social distancing rather than lockdown was responsible for keeping people in crowded homes in slums.[26]

Finally, the paper only casts light on conditions in Mumbai. It is possible that, in other Indian cities, and in cities outside India, suppression had less effect on mobility in poorer communities.

Appendix

---

[24] This test addresses two issues. One is the sharp decline in mobility in non-slums just before the start of the 24 March national lockdown. The differential decline in non-slums is sufficiently short that it does not change the topline results. The other is that Maharashtra announced smaller lockdown steps before the national lockdown. This included a ban on outdoor gatherings after 14 March, closing of shops on alternate days in Mumbai starting 18 March, and closing all workplaces on 20 March 2020. Starting the lockdown earlier does not imply that the higher mobility in slums was economically significant.

[25] Based on ethnographic work in Mumbai slums by one of the authors, while per-capita smartphone possession rates may be low, per-household possession is high. The vast majority of households in notified slums have 1 member, usually a teenager or young adult, with a smartphone.

---

[26] However, while this would exonerate lockdowns, it would be consistent with crowded homes rather than mobility being the mechanism for disparate impact of social distancing.

Table A.2
Effect of lockdown (relative to baseline) on mobility.

	1	2	3	4	5	6
Estimation method	OLS	OLS with rand. effects	Quantile	OLS	OLS with rand. effects	Quantile
Dep. Variable	Unique locations	Unique locations	Unique locations	Trips	Trips	Trips
Lockdown	-3.279***	-2.860***	0.000	-0.329***	0.189**	0.000
	0.050	0.063	0.007	0.115	0.076	0.007
Slum	-1.293***	-0.970***	0.000	-0.549***	-0.050	0.000
	0.127	0.182	0.001	0.048	0.107	0.002
Lockdown	1.360***	1.152***	0.000	0.528***	0.161*	0.000
x slum	0.102	0.093	0.010	0.155	0.089	0.009
Constant	4.547***	3.747***	1.000***	1.935***	1.030***	1.000***
	0.100	0.161	0.001	0.040	0.098	0.002
Observations	5,888,788	5,888,788	5,888,788	5,888,788	5,888,788	5,888,788
No. devices		365,042			365,042	
Level						
Slum	1.334	1.069	1.000	1.585	1.329	1.000
Non-slum	1.267	0.887	1.000	1.606	1.218	1.000
Slum - non-slum (p)	0.571	0.000	1.000	0.001	0.551	1.000
Decline						
Slum	0.590	0.615	0.000	-0.143	-0.357	0.000
Non-slum	0.721	0.763	0.000	0.170	-0.183	0.000
Slum - non-slum (p)	0.000	0.136	1.000	0.909	0.017	1.000

Notes. Home location defined on H12 and mobility on H10 grid. Cells in top panel contain coefficient and standard errors clustered on home location cells. ***/**/* indicates $p < 0.01/0.05/0.1$. Middle panel reports number of device fixed effects. Bottom panel contains measures of mobility x community type, and p values for differences across communities.

Table A.3
Sensitivity to different definitions of treatment periods.

Effect of:	Pandemic		Lockdown							
Sensitivity to:	Pandemic period ends before sero-survey		Pandemic period ends before sero-survey		Lockdown ends later (7 June)		Lockdown starts later (1 April)		Lockdown starts earlier (15 March)	
Outcome:	Locations	Trips	Locations	Trips	Locations	Trips	Locations	Trips	Locations	Trips
Difference in level	0.267	0.200	0.249	0.108	0.277	0.176	0.290	0.116	0.277	0.214
	0.160	0.345	0.081	0.497	0.099	0.375	0.061	0.480	0.147	0.337
Difference in decline	-0.184	-0.123	0.015	0.185	0.027	0.122	0.018	0.192	-0.012	0.029
	0.000	0.030	0.821	0.265	0.285	0.028	0.689	0.152	0.817	0.688

Notes. Estimation method is OLS with random effects for devices. Cell contains differences in level or decline and p-value of difference.

Table A.4
Relative mobility before sero-survey, as measured with different Uber cell sizes.

			Outcome					
	Data set		Location			Trips		
Test	Home grid	Location grid	OLS	OLS with rand. effects	Quantile regression	OLS	OLS with rand. effects	Quantile regression
Level	H10	H9	0.049	0.143	-1.000	0.020	0.124	-1.000
			0.617	0.203	0.000	0.878	0.370	0.000
	H10	H10	0.184	0.343	0.000	0.079	0.236	0.000
			0.330	0.075	1.000	0.702	0.247	1.000
	H12	H9	0.038	0.131	-1.000	0.008	0.114	-1.000
			0.701	0.259	0.000	0.950	0.423	0.000
	H12	H10	0.167	0.330	0.000	0.057	0.221	0.000
			0.393	0.107	1.000	0.784	0.290	1.000
	H12	H12	0.333	0.738	0.000	0.127	0.321	0.000
			0.482	0.103	1.000	0.645	0.243	1.000
Decline	H10	H9	-0.154	-0.163	1.000	-0.349	-0.117	1.000
			0.000	0.000	0.000	0.001	0.047	0.000
	H10	H10	-0.192	-0.210	0.000	-0.388	-0.146	0.000
			0.000	0.000	1.000	0.000	0.009	1.000
	H12	H9	-0.150	-0.157	1.000	-0.337	-0.117	1.000
			0.000	0.000	0.000	0.001	0.039	0.000
	H12	H10	-0.186	-0.203	0.000	-0.363	-0.140	0.000
			0.000	0.000	1.000	0.001	0.010	1.000
	H12	H12	-0.255	-0.281	0.000	-0.376	-0.020	0.000
			0.000	0.000	1.000	0.003	0.707	1.000

Notes. Cells give levels or difference in percent changes and p-values. Negative level means slums have less mobility. Positive decline means slums have more decline from baseline.
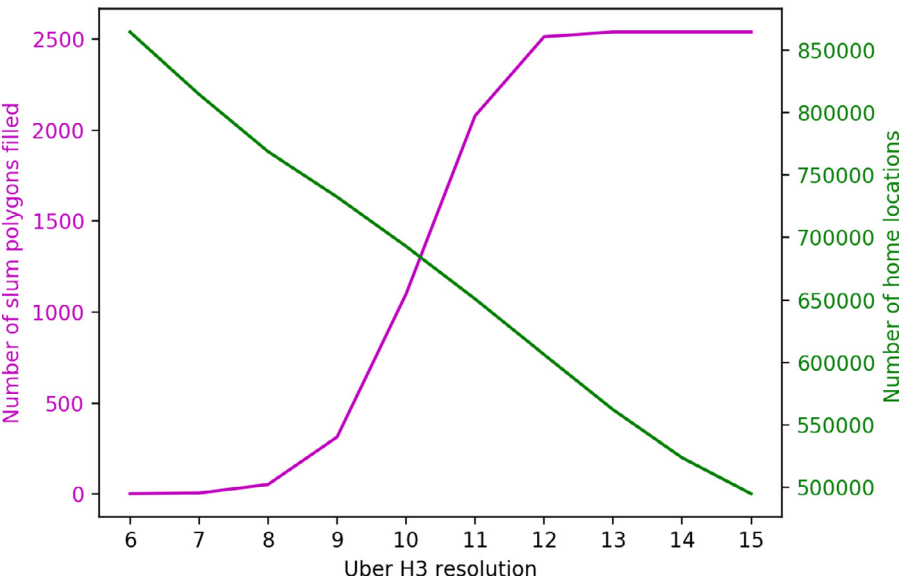
Fig. A.1. Trade-off between number of slums and home locations identified.
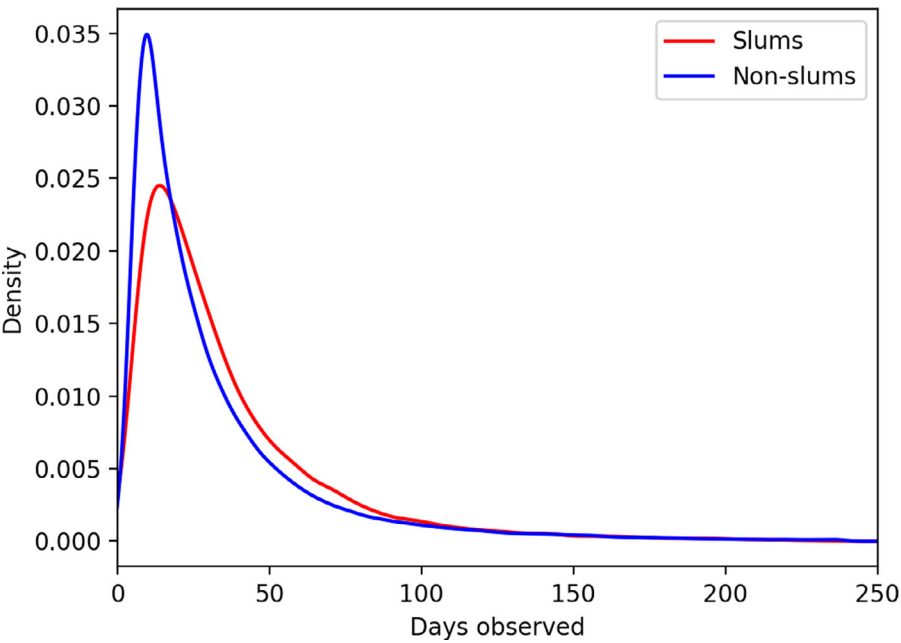

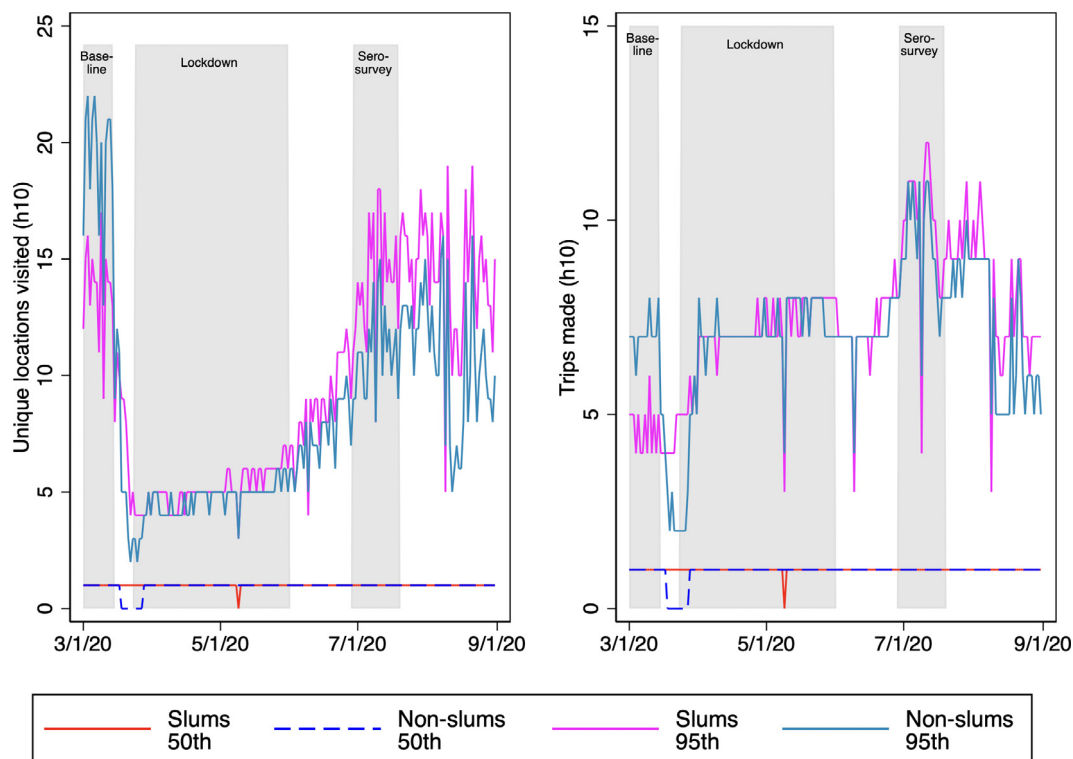**Fig. A.2.** Distribution of the number of days observed for devices in slums and non-slums based on H12 home location.

Fig. A.3. 95th percentile and median mobility among devices from slums and non-slums.

Table A.5
Relative effect of lockdown (relative to remainder of pre-survey period), as measured with different Uber cell sizes.

| | Data set | | Outcome | | | | | |
| | | | Location | | | Trips | | |
Test	Home grid	Location grid	OLS	OLS with rand. effects	Quantile regression	OLS	OLS with rand. effects	Quantile regression
Level	H10	H9	0.049	0.143	-1.000	0.020	0.124	-1.000
			0.617	0.203	0.000	0.878	0.370	0.000
	H10	H10	0.184	0.343	0.000	0.079	0.236	0.000
			0.330	0.075	1.000	0.702	0.247	1.000
	H12	H9	0.038	0.131	-1.000	0.008	0.114	-1.000
			0.701	0.259	0.000	0.950	0.423	0.000
	H12	H10	0.167	0.330	0.000	0.057	0.221	0.000
			0.393	0.107	1.000	0.784	0.290	1.000
	H12	H12	0.333	0.738	0.000	0.127	0.321	0.000
			0.482	0.103	1.000	0.645	0.243	1.000
Decline	H10	H9	-0.154	-0.163	1.000	-0.349	-0.117	1.000
			0.000	0.000	0.000	0.001	0.047	0.000
	H10	H10	-0.192	-0.210	0.000	-0.388	-0.146	0.000
			0.000	0.000	1.000	0.000	0.009	1.000
	H12	H9	-0.150	-0.157	1.000	-0.337	-0.117	1.000
			0.000	0.000	0.000	0.001	0.039	0.000
	H12	H10	-0.186	-0.203	0.000	-0.363	-0.140	0.000
			0.000	0.000	1.000	0.001	0.010	1.000
	H12	H12	-0.255	-0.281	0.000	-0.376	-0.020	0.000
			0.000	0.000	1.000	0.003	0.707	1.000

Notes. Cells give levels or difference in percent changes and *p*-values. Negative level means slums have less mobility. Positive decline means slums have more decline from baseline.

CRediT authorship contribution statement

Jaymee Sheng: Data curation, Formal analysis, Software, Validation, Visualization, Writing - original draft. **Anup Malani:** Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Visualization, Writing - original draft, Writing - review & editing. **Ashish Goel:** Conceptualization, Methodology, Project administration, Supervision, Writing - review & editing. **Purushotham Botla:** Data curation, Supervision.

References

Alexander, D., Karger, E., 2020. Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior.

Almagro, M., Coven, J., Gupta, A., Orane-Hutchinson, A., 2020. Racial disparities in frontline workers and housing crowding during covid-19: evidence from geolocation data. Available at SSRN 3695249.

Banerjee, A., Duflo, E., 2011. Poor Economics: a Radical Rethinking of the Way to Fight Global Poverty. Public Affairs.

Bennett, M., 2020. All things equal? Heterogeneity in policy effectiveness against Covid-19 spread in Chile.

Carozzi, F., Provenzano, S., Roth, S., 2020. Urban Density and COVID-19. IZA Discussion Papers. Bonn

Carranza, A., Goic, M., Lara, E., Olivares, M., Weintraub, G.Y., Covarrubia, J., Escobedo, C., Jara, N., Basso, L.J., 2020. The Social Divide of Social Distancing: Lockdowns in Santiago during the COVID-19 Pandemic. Report.

Chiou, L., Tucker, C., 2020. Social Distancing, Internet Access and Inequality. National Bureau of Economic Research Working Paper Series No. 26982 doi:10.3386/w26982.

Coven, J., Gupta, A., 2020. Disparities in Mobility Responses to Covid-19. Report. NYU Stern Working Paper.

Emeruwa, U.N., Ona, S., Shaman, J.L., Turitz, A., Wright, J.D., Gyamfi-Bannerman, C., Melamed, A., 2020. Associations between built environment, neighborhood socioeconomic status, and SARS-COV-2 infection among pregnant women in new york city. JAMA 324 (4), 390–392. doi:10.1001/jama.2020.11370.

Fan, Y., Orhun, A.Y., Turjeman, D., 2020. Heterogeneous Actions, Beliefs, Constraints and Risk Tolerance During the Covid-19 Pandemic. National Bureau of Economic Research Working Paper Series No. 27211 doi:10.3386/w27211.

Gandhi, S., Tandel, V., Pethe, A., Phatak, V., Risbud, S., 2014. Real estate prices in mumbai: does the metro rail have an impact? Econ. Polit. Wkly. 49 (9), 55–61.

Huang, X., Li, Z., Jiang, Y., Ye, X., Deng, C., Zhang, J., Li, X., 2020. The characteristics of multi-source mobility datasets and how they reveal the luxury nature of social distancing in the u.s. during the covid-19 pandemic. medRxiv, 2020.07.31.20143016. 10.1101/2020.07.31.20143016

Kearney, A., Munana, C., 2020. Taking Stock of Essential Workers. Report. Kaiser Family Foundation.

Laxminarayan, R., Wahl, B., Dudala, S.R., Gopal, K., Mohan B, C., Neelima, S., Jawahar Reddy, K.S., Radhakrishnan, J., Lewnard, J.A., 2020. Epidemiology and transmission dynamics of Covid-19 in two indian states. Science 370 (6517), 691. doi:10.1126/science.abd7672.

Leatherby, L., 2020. The Worst Virus Outbreaks in the U.S. are Now in Rural Areas. New York Times.

Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. Nature 438 (7066), 355–359. doi:10.1038/nature04153.

Long, Q.-X., et al., 2020. Antibody responses to SARS-COV-2 in patients with Covid-19. Nat. Med. 26 (6), 845–848. doi:10.1038/s41591-020-0897-1.

Lou, J., Shen, X., 2020. Are Stay-at-Home Orders More Difficult to Follow for Low-Income Groups? Report. Working Paper.

Malani, A., Ramachandran, S., Tandel, V., Parasa, R., Sudharshini, S., Prakash, V., Yogananth, Y., Raju, S., Selvavinayagam, T. S., 2021. Sars-cov-2 seroprevalence in Tamil Nadu in October-November 2020. medRxiv, 2021.02.03.21250949. 10.1101/2021.02.03.21250949

Malani, A., Shah, D., Kang, G., Lobo, G.N., Shastri, J., Mohanan, M., Jain, R., Agrawal, S., Juneja, S., Imad, S., Kolthur-Seetharam, U., 2020. Seroprevalence of SARS-COV-2 in slums versus non-slums in Mumbai, India. Lancet Glob. Health doi:10.1016/S2214-109X(20)30467-8.

Mohanan, M., Malani, A., Krishnan, K., Acharya, A., 2020. Prevalence of Covid-19 in rural versus urban areas in a low-income country: findings from a state-wide study in Karnataka, India. medRxiv, 2020.11.02.20224782. 10.1101/2020.11.02.20224782

Murhekar, M.V., et al., 2021. SARS-COV-2 antibody seroprevalence in India, August, September, 2020: findings from the second nationwide household serosurvey. Lancet Glob. Health doi:10.1016/S2214-109X(20)30544-1.

Nichols, M., Thorson, M., Procell, C., 2020. We Looked at Coronavirus in 8500 Zip Codes Across America. Here's What we Found. USA Today.

Okba, N.M.A., Müller, M., Li, W., Wang, C., GeurtsvanKessel, C., Corman, V., Lamers, M., Sikkema, R., de Bruin, E., Chandler, F., Yazdanpanah, Y., Le Hingrat, Q., Descamps, D., Houhou-Fidouh, N., Reusken, C.B.E.M., Bosch, B.-J., Drosten, C., Koopmans, M.P.G., Haagmans, B., 2020. Severe acute respiratory syndrome coronavirus 2 specific antibody responses in coronavirus disease patients. Emerg. Infect. Dis. J. 26 (7), 1478. doi:10.3201/eid2607.200841.

Parker, K., Horowitz, J.M., Brown, A., 2020. About Half of Lower-Income Americans Report Household Job or Wage Loss Due to COVID-19. Report. Pew Research Center.

Reese, P., 2020. High-Poverty Neighborhoods Bear the Brunt of Covid's Scourge. KHN.

Ribeiro, H.V., Sunahara, A.S., Sutton, J., Perc, M., Hanley, Q., 2020. City size and the spreading of Covid-19 in Brazil. PLoS One 15.

Stier, A., Berman, M., Bettencourt, L., 2020. Covid-19 Attack Rate Increases With City Size. Mansueto Institute for Urban Innovation Research Paper. Forthcoming

Taubenböck, H., Kraff, N.J., 2014. The physical face of slums: a structural comparison of slums in Mumbai, India, based on remotely sensed data. J. Hous. Built Environ. 29 (1), 15–38. doi:10.1007/s10901-013-9333-x.

Whittle, R.S., Diaz-Artiles, A., 2020. An ecological study of socioeconomic predictors in detection of covid-19 cases across neighborhoods in New York city. BMC Med. 18 (1), 271. doi:10.1186/s12916-020-01731-6.

Worldometer, 2021. COVID-19 Coronavirus Pandemic. Report.

Zhao, J., Yuan, Q., Wang, H., Liu, W., Liao, X., Su, Y., Wang, X., Yuan, J., Li, T., Li, J., Qian, S., Hong, C., Wang, F., Liu, Y., Wang, Z., He, Q., Li, Z., He, B., Zhang, T., Fu, Y., Ge, S., Liu, L., Zhang, J., Xia, N., Zhang, Z., 2020. Antibody responses to SARS-COV-2 in patients with novel coronavirus disease 2019. Clin. Infect. Dis. 71 (16), 2027–2034. doi:10.1093/cid/ciaa344.