Appendix – For Online Publication

A Data appendix

A.1 Smartphone visits data

Each observed visit consists of a device, a venue, a timestamp, and an attribution score. PlaceIQ's attribution scores are larger when a device is more likely to have been within a venue, based on the number and density of pings, data source of pings, and proximity of the pings to the polygon defining the venue. We retain all visits with an attribution score greater than a threshold value recommended by PlaceIQ based on their experience correlating their data to a diverse array of truth sets, including consumer spending data and foot-traffic counts. PlaceIQ also reports a lower bound for the visit's duration based on the time between consecutive pings at the same venue.

We also clean the visit data to remove simultaneous visits. For instance, when two venues are in close proximity to one other, a single visit event may have an attribution score for both venues that exceeds the threshold value recommended by PlaceIQ. We retain only the visit to the venue with the highest attribution score. In other cases, the polygons of two different venues overlap.[32] When two polygons overlap, we retain polygons with an identified business category over those lacking a category.

Table A.1 summarizes the smartphone movement data after this cleaning for days between January 20 and March 1, 2020. On the average day, there were 176 million visits produced by 33 million devices visiting 40 million residential and non-residential venues. The average device appears in the data for 25 days between January 20 and March 1, but a notable number appear on only one day.

[32]This could happen, for instance, if the basemap contains one polygon representing a business establishment and a second polygon representing both that building and the accompanying parking lot.

After we apply the device selection criteria we use when computing the LEX and DEX indices (devices that pinged on at least 11 days over any 14-day period from November 1, 2019 through the reporting date), there are 152 million visits from 23 million devices visiting 37 million venues on an average day. The selected devices appear in the data between January 20 and March 1 for 35 days on average.

Table A.1: Summary statistics for cleaned visits and indices samples

	Cleaned visits sample				Indices sample			
	Mean	SD	5th	95th	Mean	SD	5th	95th
Devices	33.43	1.92	31.15	36.58	22.80	0.49	22.05	23.61
Venues	40.46	0.81	39.17	41.51	36.88	0.92	35.35	38.28
Visits	175.85	11.33	154.15	191.12	151.56	11.30	132.59	166.74
Duration	25.81	14.31	1.00	41.00	34.91	9.89	11.00	41.00

NOTES: This table summarizes PlaceIQ data for January 20, 2020 to March 1, 2020 after our cleaning of the visits as described in the text. The counts of devices, venues, and visits are stated in millions per day. Duration is the number of days between a device's first and last appearance in the data (between January 20 and March 1).

A.2 Home assignments

Residential venues are a distinct category in the PlaceIQ data. This allows us to construct a weekly panel of home locations for a subset of devices using the following assignment methodology:

1. For each week, we assign a device to the residential venue where its total weekly visit duration at night (between 5pm and 9am) is longest, conditional on it making at least three nighttime visits to that venue within the week.[33] If a device does not visit any residential location on at least three nights, then on initial assignment that device-week pair has a missing residential location.

[33]Since we only observe minimum duration, there are instances where total duration is 0 across all residential locations. In these cases, we assign the residential venue as the venue a device makes the most nighttime visits.

Table A.2: Venue categories in DEX

Retail	209,274
Restaurants	200,839
Gas Station/Convenience Stores	118,307
Night Clubs/Bars	88,784
Banks	79,150
Shipping	36,745
Hotels	32,303
Home Improvement Stores	27,097
Grocery Stores	25,770
Financial Services	23,238
Pharmacies	22,408
Car Dealerships	20,644
Beauty Stores	15,556
Big Box Stores	11,558
Real Estate Offices	9,732
Gyms	9,289
Car Rental	8,999
Pay Day Loan	6,043
Storage	5,935
Movie Theaters	4,632
Library	1,962
Liquor Stores	1,193

NOTES: This table lists the venue categories that enter the computation of the Device Exposure Index (DEX) and shows the total number of distinct venues on 30 June 2020 in each category. Some venues belong to multiple categories, so the number of distinct venues (about three-quarters of a million) is smaller than the sum of all rows in this table.

2. After this preliminary assignment, we fill in missing weeks and adjust for noisiness in the initial panel using the following interpolation rules:

Rule 1: *Change "$X \cdot X$" to "$X\ X\ X$"*: If the residential assignment for a week is missing and the non-missing residential assignment in the weeks before and after is the same, we replace the missing value with that residential assignment.

Rule 2: *"$a\ X\ Y\ X\ b$" to "$a\ X\ X\ X\ b$" where $a \neq Y$ and $b \neq Y$*: If a device has a residential assignment Y that does not match the assignment X in

the week before or after, we replace Y with X as long as Y was not the residential assignment two weeks before or two weeks after.[34]

3. After step 2's interpolation, for any spells of at least four consecutive weeks where a device is assigned the same residential venue, we assign that venue as a device's "home" for those weeks. Spells of less than four weeks are set to missing.

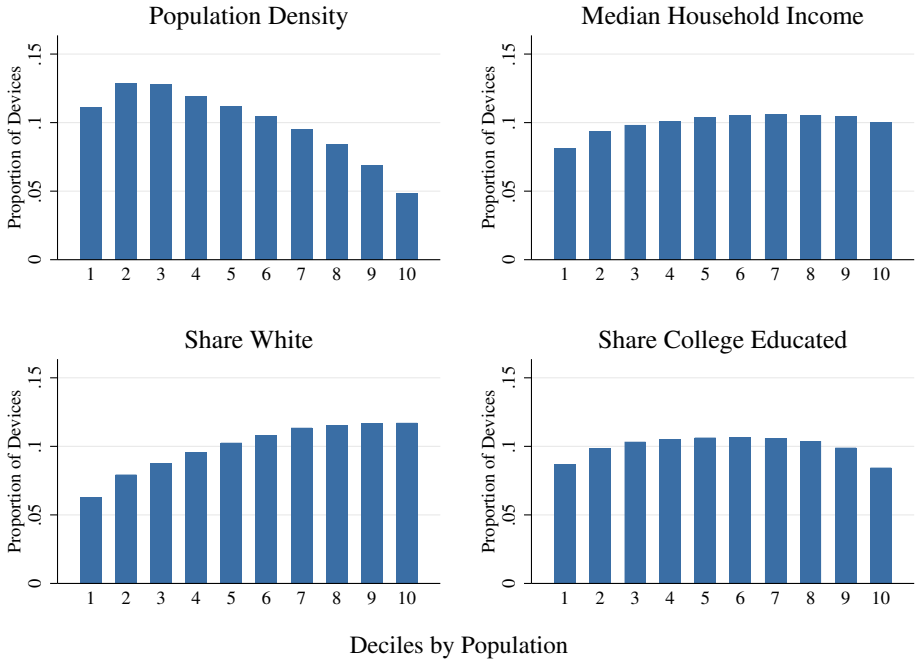
4. If a device has more than one home assignment and the pairwise distance between them is less than 0.1 kilometers, we keep the home that appears for the most weeks.

5. If a device has the same home assignment in two non-consecutive periods and no other home assignments in between, then we assign all weeks in between to that home assignment.

[34]For cases where a device's residential location is bouncing between two places ("Y X Y X X") we are not able to ascertain whether Y or X is more likely to be a device's residence in a given week

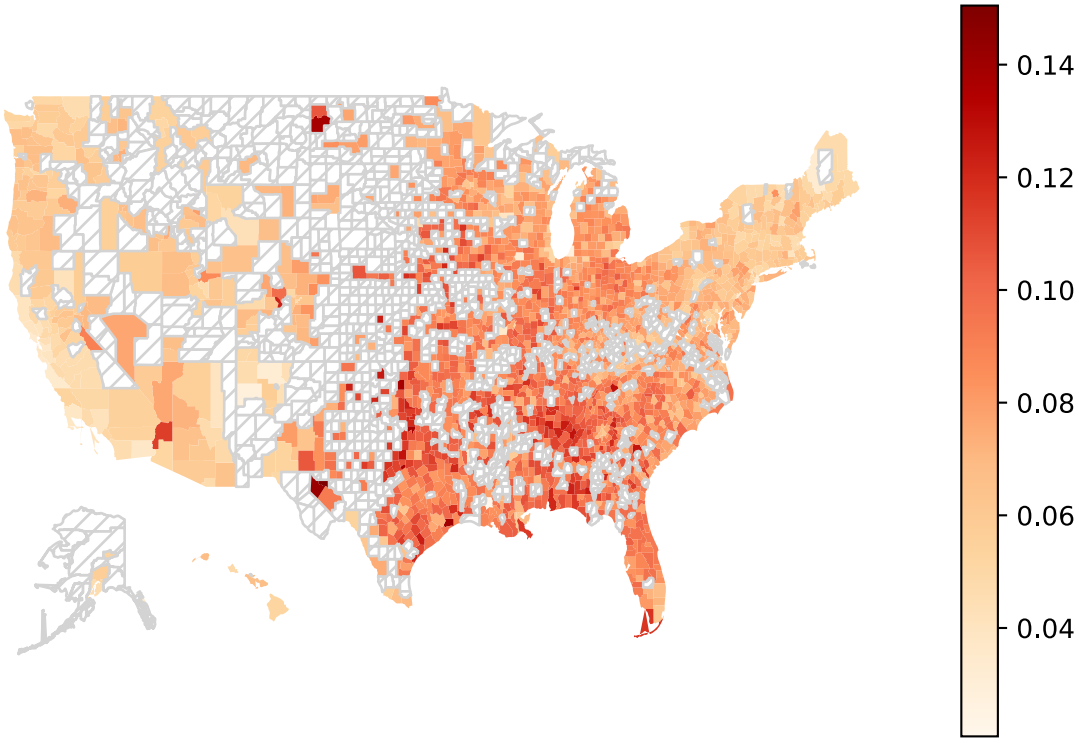
B Figures appendix

Figure B.1: Balance of devices' residences across block groups by national demographic deciles



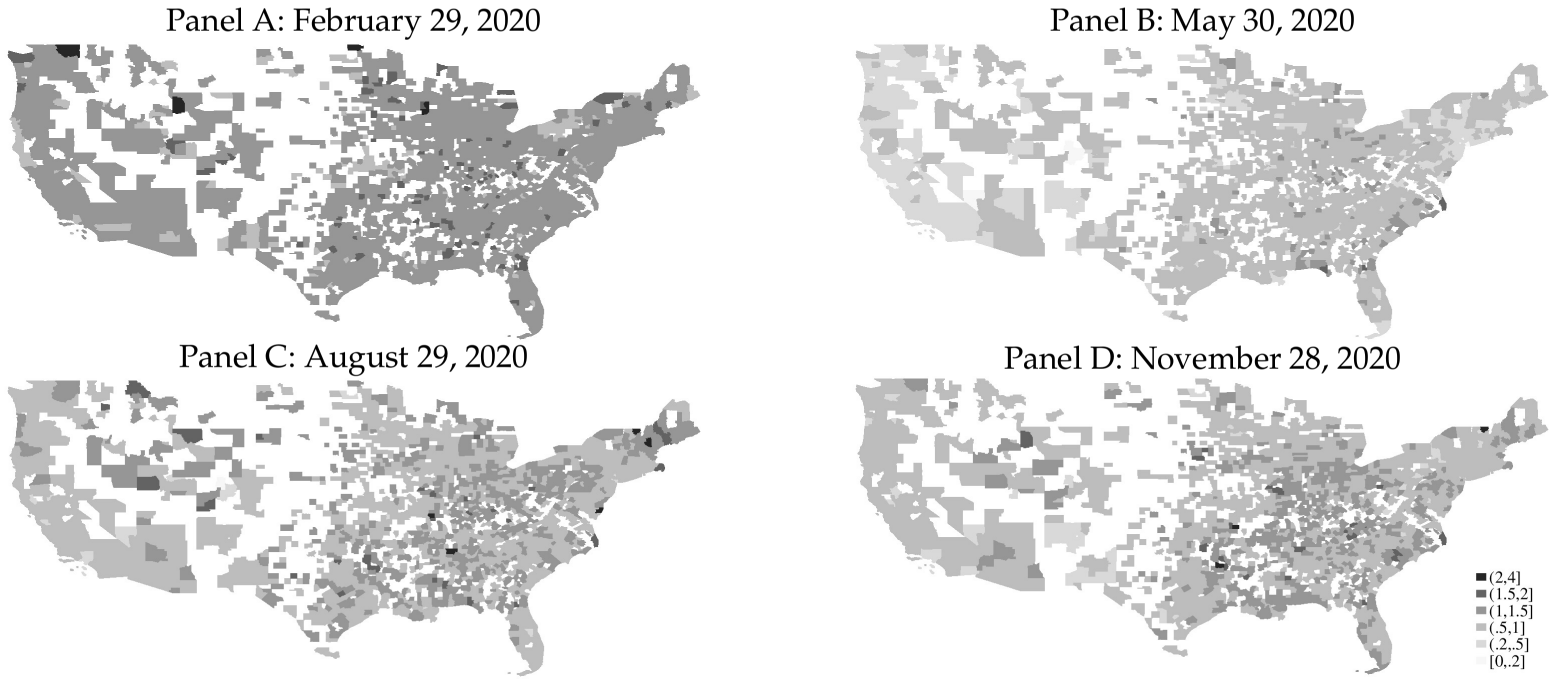
NOTES: This figure shows the total share of devices living in census block groups corresponding to the national deciles for each of the four demographic categories.

Figure B.2: Ratio of devices to residents by county



Notes: This figure depicts ratio of smartphone devices to residential population for each county in the contiguous United States. The number of devices is the average number of devices contributing to the county-level DEX during January 20 through February 28, 2020. We report our indices for 2,018 counties that account for about 97 percent of the national population.

Figure B.3: County-Level Device Exposure (DEX) Normalized to Level on February 1, 2020



Notes: This figure shows the county-level average device exposure (DEX) on a given date (February 29 in Panel A, May 30 in Panel B, August 29 in Panel C, and November 28 in Panel D) as a percentage of its level on February 1, 2020.

B.1 DEX values by block-group demographics

Figure B.8 reveals variation in the reduction in activity across educational attainment and race. Panel A depicts each DEX-education quartile relative to the aggregate DEX on March 7 using a 7-day moving average. Prior to the onset of COVID-19 in the U.S., residents of block groups with more college graduates were more exposed to other devices than average.[35] In March, exposure fell for residents of all block groups, but residents of block groups with more college graduates exhibited a proportionately greater decline. As a result, by the end of March 2020, there was little discernible difference in device exposure across neighborhoods with different shares of college graduates. Exposure for all education quartiles rose during late spring and plateaued in the summer. Initial differences across quartiles reemerged, though greatly attenuated.

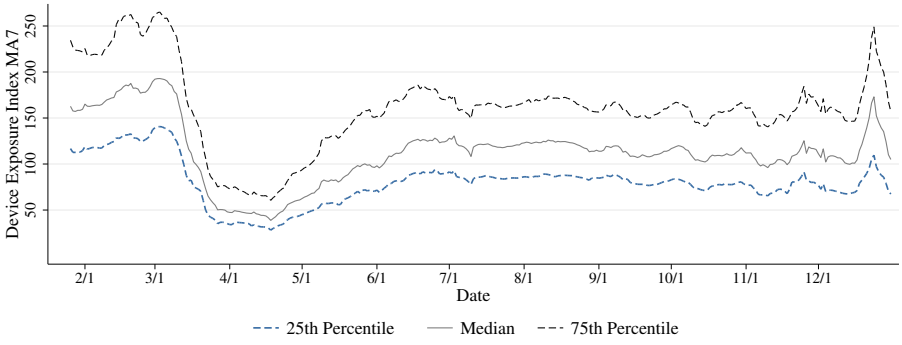
Panel B of Figure B.8 depicts device exposure by racial/ethnic demographics using a 7-day moving average. Prior to the pandemic, devices living in block groups with more Black, Hispanic, and White residents had similar levels of exposure, while devices living in block groups with more Asian residents had higher DEX values. From mid-March onwards, all four demographic groups exhibit fairly similar exposure levels.

The limited variation in device exposure across different demographic groups after March 15 may imply a limited role for heterogeneous exposure rates in explaining differences in these demographic groups' infection and mortality rates during the pandemic. Researchers investigating these questions could combine these local measures of social contact by demographic traits with other observed demographic differences that may explain disparate outcomes.

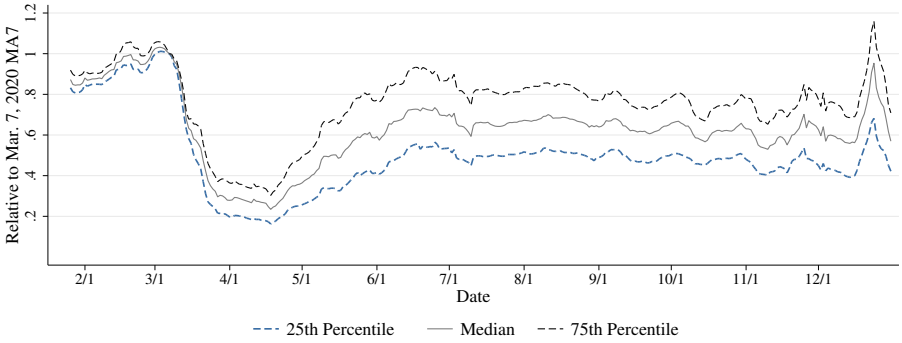
[35] This is consistent with the finding that devices from higher-income neighborhoods visit more places (Chen and Pope, 2020).

Figure B.4: Interquartile range of DEX over time

Panel A: Raw Device Exposure Index



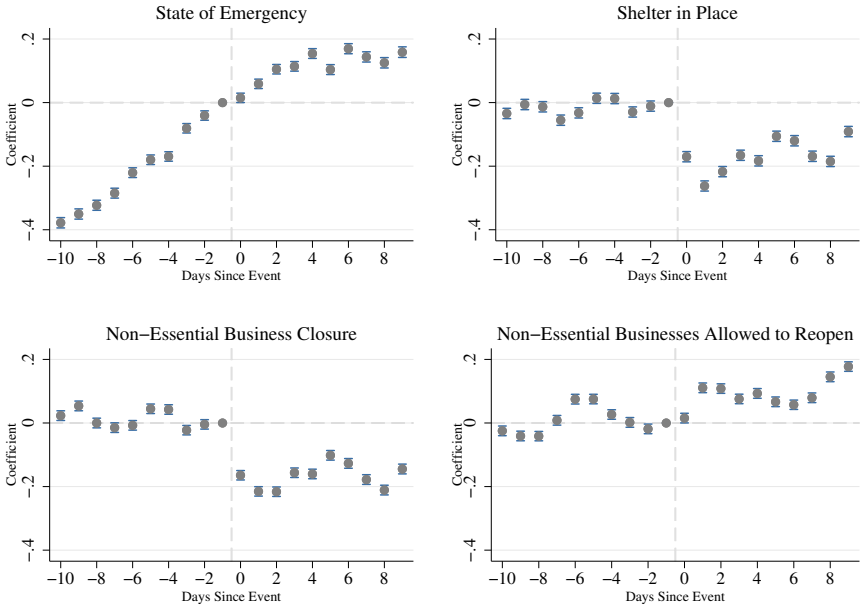
Panel B: Normalized to March 7, 2020



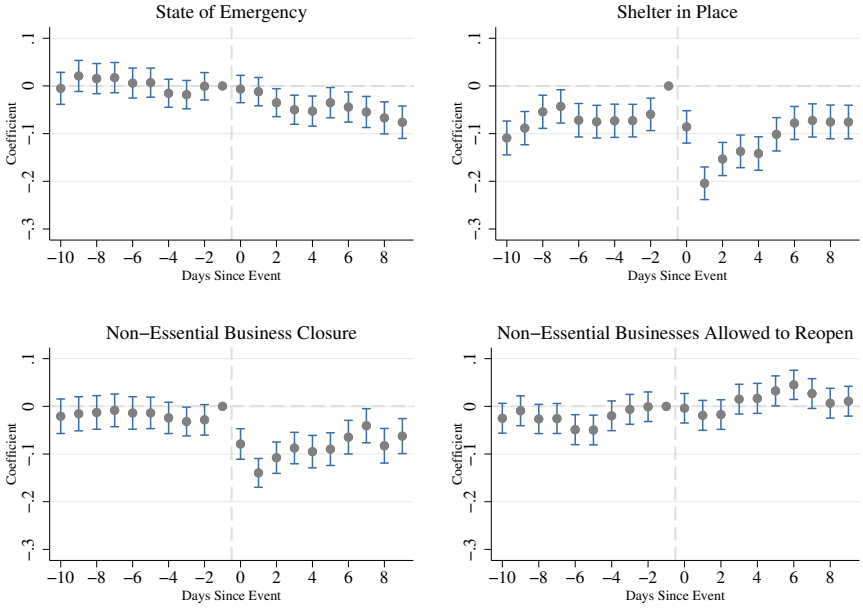
NOTES: This figure shows the population-weighted median and interquartile range of the device exposure index over time.

Figure B.5: Changes in DEX relative to lockdown policies

Panel A: Using All Variation

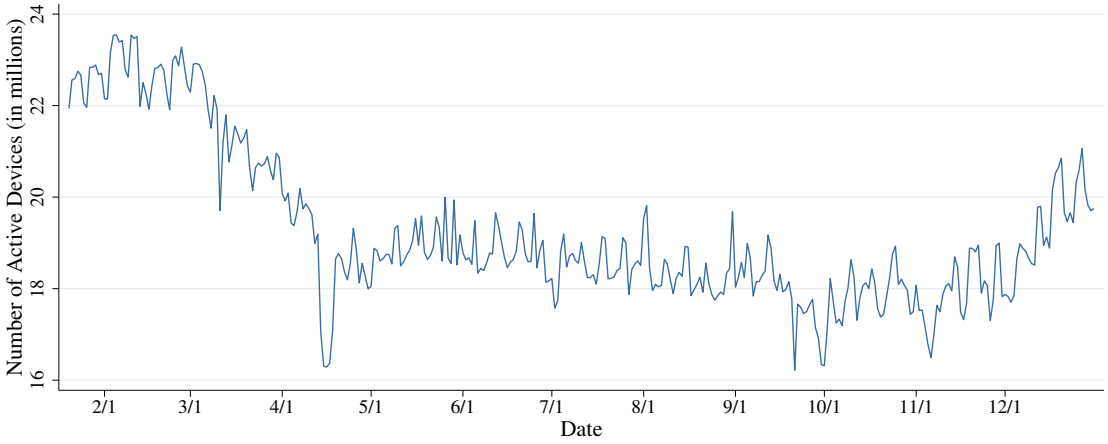


Panel B: Only Using Cross-State Variation within Commuting Zones



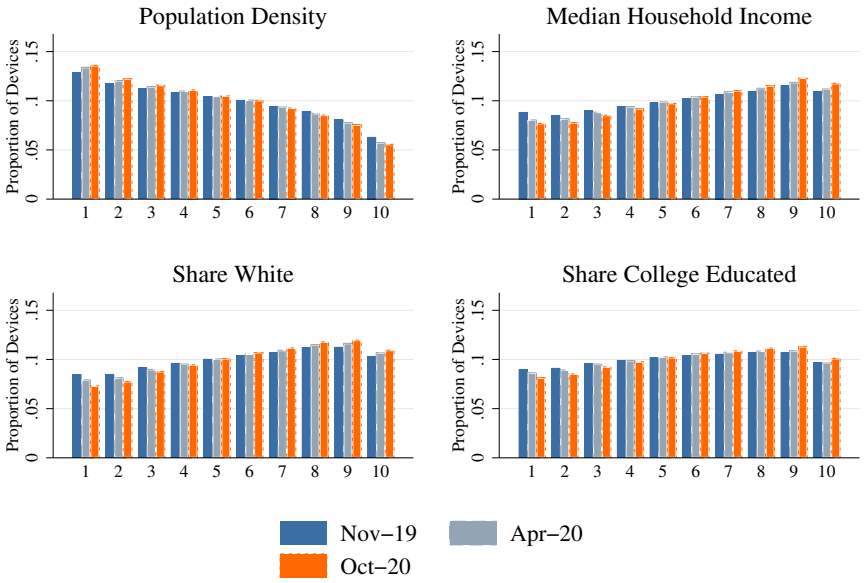
NOTES: Each plot in this figure presents the coefficients estimated in a regression of the county-level device exposure index on dummies for the time since a given policy change. In Panel A, these regressions also include county and date fixed effects. In Panel B, the regressions include county and commuting zone-by-date fixed effects. Each plot presents the results for a different state-wide policy, each drawn from Raifman, Nocka, Jones, Bor, Lipson, Jay, and Chan (2020). Each point represents the coefficient on the dummy for a given number of days since the policy was instituted, with the bands reflecting 95% confidence bounds on those estimates.

Figure B.6: Number of devices active per day



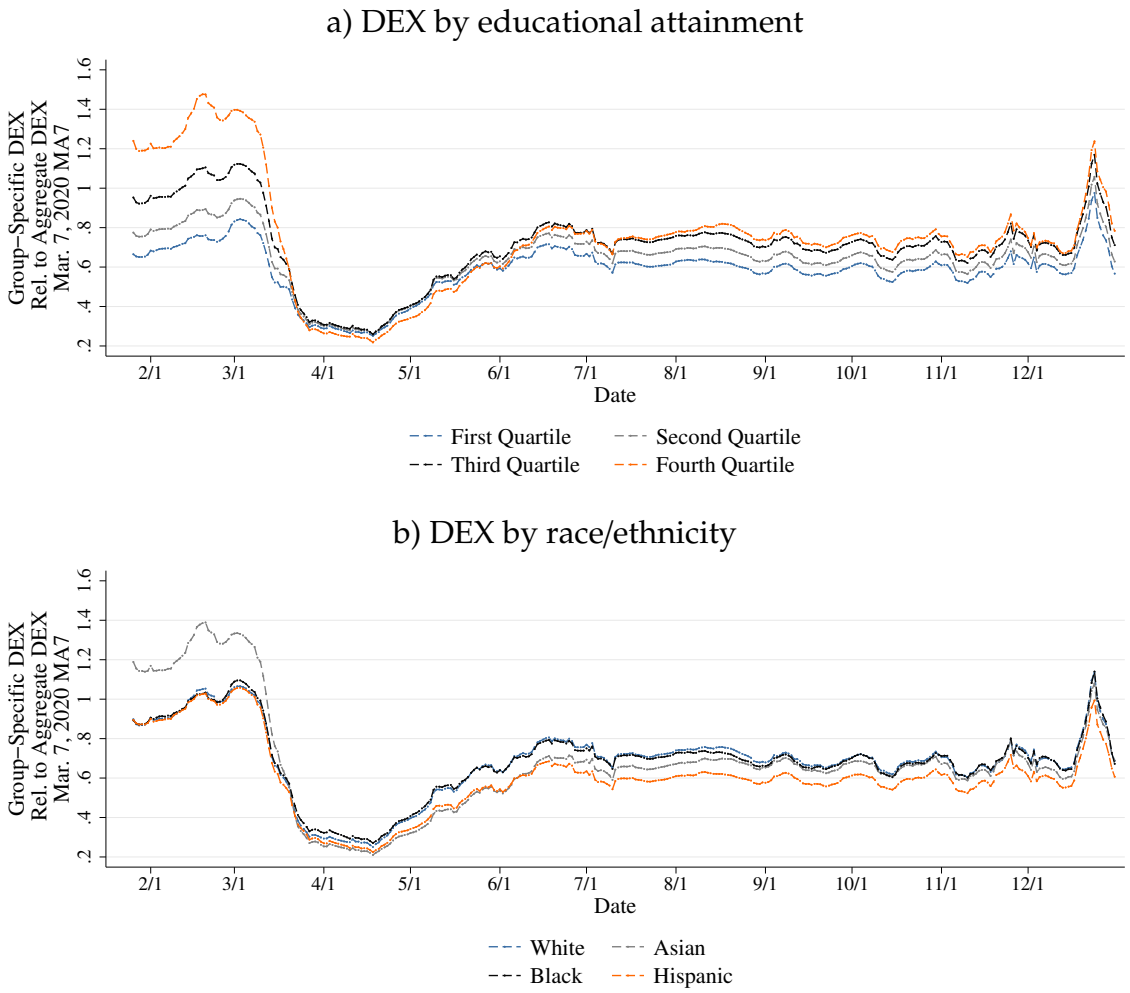
NOTES: This figure depicts the number of devices meeting the sample selection criteria defined in Section 2.2 that pinged at least once on each day from January 20 to December 31, 2020. See footnote 9 for a discussion of the April 14-18, 2020 anomaly.

Figure B.7: Evolving share of devices by block-group demographic decile



NOTES: This figure depicts the share of devices residing in block groups in each county-specific decile of population density, median household income, share of white residents, and share of residents over 25 years with a bachelor's degree or higher. These block group characteristics are from the 2014-2018 American Community Survey. Device shares are determined by the share of devices living in block group j in month t conditional on a device having a block group of residence assigned in month t.

Figure B.8: DEX values by block-group demographics



Notes: These plots depict the state-level DEX by demographic groups on each day from January 20 to December 31, 2020. For each state, the demographic DEX time series is smoothed using a 7-day moving average and then divided by the level of the aggregate DEX on March 7, 2020. The depicted series is a device-weighted average over all states. Panel A depicts this series for DEX by education and Panel B depicts this series for DEX by race/ethnicity as defined in Section 3.

C Indices appendix

C.1 Density DEX including large venues

The baseline DEX excludes visits to large venues. To appropriately measure device exposure in both large and small venues, we need to adjust for venue size. To that end, we have calculated a version of the DEX that includes visits to all identified business establishments, including large venues, scaling exposure by the land area of the venue.

First, we calculate the density-weighted exposure of device i on date d as the number of distinct devices that visited the same commercial venues, each weighted by the inverse area of the densest establishment that both devices visited:

$$\text{EXP}_{i,d}^{\text{Density}} = \frac{1}{\sum_{j \in \mathcal{J}_{i,d}} a_j^{-1}} \sum_{i'} \max_{j: \mathcal{J}_{i,d} \cap \mathcal{J}_{i',d}} a_j^{-1},$$

where a_j is the land area of establishment j, $\mathcal{J}_{i,d} \cap \mathcal{J}_{i',d}$ is the set of venues visited by both i and i' on date d, and the maximum of an empty set is defined to be zero.

The density-weighted DEX is then defined as the average density-weighted exposure for devices that reside in geographic unit g on date d:

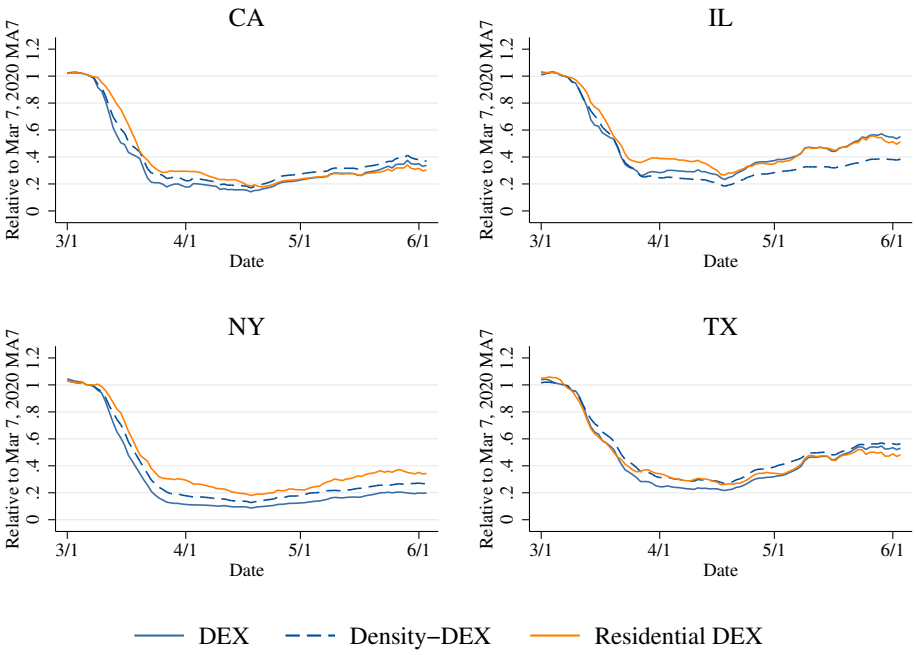
$$\text{DEX}_{g,d}^{\text{Density}} \equiv \frac{1}{|\mathcal{G}_{g,d}|} \sum_{i \in \mathcal{G}_{g,d}} \text{EXP}_{i,d}^{\text{Density}}.$$

Figure C.1 shows that the inclusion of large venues, adjusted for density, does not meaningfully alter time-series patterns in the DEX.

C.2 Residential DEX

We complement the DEX with a "residential" DEX that tracks exposure at residential venues. This index is calculated using the same formula as the baseline DEX

Figure C.1: Comparison of device exposure indices



Notes: This figure depicts three different versions of the DEX for California, Illinois, New York, and Texas. The solid blue line depicts the baseline DEX as defined in the main text. The Density DEX depicts the density-weighted DEX defined in Appendix C.1. The residential DEX depicts the DEX for residential visits defined in Appendix C.2. Each series depicts a 7-day moving average relative to its value on March 7, 2020.

but tracks overlapping visits to residences, rather than commercial venues. We exclude visits to residences within 0.1km of a device's own residence, since we are unable to rule out that these "visits" may simply be time spent at home.

Figure C.1 shows that exposure in residential venues did not dip by as much during the lockdown period of late March and early April but recovered to a similar level by late May.