What can smartphone location data tell us about the pandemic?

November 5 2020, by Dylan Walsh

Credit: NASA Earth Observatory/Wikimedia

Kevin Williams, an associate professor of economics at Yale SOM, was in the middle of a study of consumer retail behavior using real-time cell phone location data when COVID-19 hit. Shelter-in-place orders rolled across the country. Foot traffic at stores froze.

"If you thought the retail apocalypse looked bad a few years ago, that is likely small potatoes compared to right now," he says.

But as the prospects for one line of research closed down, another avenue opened. Could the same location data be useful in studying COVID-19, he wondered? Partnering with several other researchers,

Williams gathered—and continues to gather—real-time information on the movement of millions of individuals across the country as they adapt to the constraints of a pandemic. He hopes that this public repository of information will spur other researchers to investigate patterns and changes in the movement of individuals and how it affects transmission of the disease.

"Typically, in economics, you obtain data, conduct analysis, write a paper and submit it for publication—and the paper is the product," Williams says. "This work was a bit different: we put the data first. In light of COVID-19, our idea was to create a public good for investigating how devices move across geographies as well as how devices potentially interact at retail establishments."

By June, the data set contained 53 million devices, each of which had reported location data at least 11 of every 14 days starting in November 2019. The researchers connected each to a demographic profile by making an educated guess at the owner's home location—based on where the device generally spends the night—and then matching it against census-reported block groups.

The data is organized into two indexes. First, what the researchers refer to as the location exposure index, or LEX. This captures the broad, geographic movement of devices on a rolling two-week basis. How many move across county lines? How many across state lines? If a device was in New York two weeks ago, where has it traveled since? Second is the device exposure index, or DEX. At the county level, this examines the potential for one device to encounter another at any given establishment.

"If one day I go to the grocery store, and then to a big box retailer, DEX looks at the number of other unique devices that visited these establishments on the same day," Williams says. "So we're keeping track of these interactions at the device level and then reporting the average

size of the exposure set for each county."

Williams and his colleagues ran a few preliminary analyses to demonstrate how other researchers might use the indexes. They found, for instance, that travel between states fell sharply in March and April before beginning to rise again in late April. In the hard-hit New York City area, travel from Manhattan to surrounding counties dropped in the spring but there didn't appear to be a similar change in the Houston area. Another analysis found that exposure levels measured by DEX were uniformly low across races, as inferred from Census data, suggesting that varying levels of exposure outside of the home cannot explain dramatic differences in infection and mortality rates.

As they continue to update both the LEX and DEX data sets, the researchers are looking into the development of a third "venue exposure index," or VEX. This index would be used to assess the potential of exposure based on the size of a given venue and its internal traffic. "What does traffic look like for a venue relative to its size? What are changes in retail patterns based on where and how big a venue is?" Williams says. "Questions like these—and there are many more—are obviously super important when it comes to reopening the economy."

But finding answers to those kinds of questions, Williams notes, is beyond the scope of the current project. "We set out to build and publicize this dataset, and we did that," he says. "We're all just hoping people will be able to take advantage of it."

  More information: Measuring movement and social contact with smartphone data: a real-time application to COVID-19: github.com/COVIDExposureIndice … lob/master/CDGHW.pdf