42 Using Google's Mobility Data to understand park visitation during the COVID-19 pandemic: A note of caution

William Rice[1], **Bing Pan**[2], [1]University of Montana, USA. [2]Pennsylvania State University, USA

The COVID-19 pandemic has dramatically impacted park visitation around the globe. In an effort to understand the factors influencing these changes, numerous attempts have been made to use big data to monitor changes in park use (e.g., Venter et al., 2020). Google's Community Mobility Reports represent a dataset with significant potential in this regard. Released in April 2020, these reports were generated on the hypothesis that "aggregated, anonymized data could be helpful [to] make critical decisions to combat COVID-19" (Fitzpatrick & DeSalvo, 2020, para. 1). The heading on the reports' website asks browsers to "see how your community is moving around differently due to COVID-19" (Google 2020b). The data released through the reports are generated from "aggregated, anonymized sets of data from [Google] users who have turned on the Location History setting, which is off by default" (Google 2020b).

Methods

To understand drivers of changes in park visitation during the first wave of the COVID-19 pandemic in the western United States, we gathered data from Google's COVID-19 Community Mobility Reports (Google, 2020c). This data contains daily mobility trends—calculated as difference from the "baseline" period of January 3rd to February 6th, 2020—for areas such as national parks, public beaches, marinas, dog parks, plazas, and public gardens" (Google, 2020a, p. 1). We generated a coefficient for 97 U.S. counties representing the average daily change in park mobility from April through June 2020. These 97 counties were selected based on data availability, representing a continuous swath of counties having park mobility data available for at least half of the days of the study period.

Using a spatial lag model, we assessed a number of independent variables with relation to their influence on changes in park visitation (see Table 1). Each of these variables were selected either because of previous demonstration of their influence on park visitation or their use by policymakers in controlling and adapting to the COVID-19 pandemic.

Results

The results of the spatial lag model are listed in Table 1. On average, among 97 counties examined in this study, there was a 20.2% increase in park visitation compared to the baseline period. Concerning the spatial lag model, according to the R^2, the independent variables' variance account for 62% of the variance within the dependent variable. Just two variables were found to be statistically significantly predictive of change in park visitation: elevation and latitude. Duration of safer-at-home orders and median age showed borderline significance (below 95%) in their prediction of change in park visitation at a 92.5% confidence interval.

Discussion and Conclusion

All else being equal, the overall increase in park visitation from the baseline period indicates that individuals living in the study area were able to visit parks despite the limitations of the pandemic. However, our results indicate that the use of January and February 2020 park visitation levels as a baseline for calculating changes in park visitation is troublesome. This contention is based on the finding that only elevation and latitude—not any of the variables directly related to the pandemic—were predictive of changes in park visitation during the first wave of the COVID-19 pandemic in the western United States. This suggests that much of the change in park visitation depicted in Google's Community Mobility Reports is the function of seasonality rather than the pandemic. Climate, influenced by elevation and latitude, is a noted driver of seasonal changes in park visitation (Smith, 1993). We therefore posit that Google's park mobility data are misleading, biased by geography. Researchers must be very careful when using big data to assess visitor use trends in parks, as the curation of the data may be less-than-transparent.

Table 1: Model Summary

	Definition	Source	Min.	Max.	Mean	Coefficient	Std. Error	p-value
Dependent Variable								
Park Visitation	Average percent change in daily park use among county residents during the study period (April 1st – June 30th, 2020) from the baseline period. Baseline use is calculated are the median values, for the corresponding day of the week, during the 5-week period January 3rd to February 6th, 2020.	Google (2020b)	-59.0	101.8	20.2			
Independent Variables								
Population density	Population per square mile based on 2018 census data	United States Census Bureau (2018)	1.8	18384.2	514.5	-0.0004	0.0013	0.7450
Median age	Median age of county residents based on 2018 census data	United States Census Bureau (2018)	29.6	53.9	39.1	-0.7954^{+}	0.4123	0.0537
Duration of Safer-at-home order	Number of days throughout the study area where county-level safer-at-home order was in place	Killeen et al. (2020)	38	72	46.7	-0.4693^{+}	0.2613	0.0726
Confirmed COVID-19 Cases within county	Total confirmed cases within county as of June 30th, 2020	Centers for Disease Control and Prevention (2020)	5	103,529	3,737.2	-0.0002	0.0002	0.2645
Latitude	Centroid latitude of county	ESRI (2020)	37.0	62.5	49.3	3.33057***	0.5757	< 0.0001
Elevation	Average elevation (meters) of county	ESRI (2020)	1	2,118	356.1	0.0135**	0.0049	0.0063
Population within ½ mile of park	Portion of population within a buffer of ½ mile radius of a park	Centers for Disease Control (2019)	0.12	0.99	0.59	-4.5577	14.4969	0.7532
Model Specs								
Spatial lag effect						0.2090^{+}	0.1140	0.0667
Constant						-96.0302***	27.4610	0.0005

$^{+}$p < .075, *p < .05, **p < .01, ***p < .001 Breusch-Pagan test: 13.26, p = 0.066 $R^2 = 0.623$
Multicollinearity condition number = 27.12 Likelihood Ratio Test: 2.92, p = 0.088 AIC = 895.699; BIC = 918.871

References

Centers for Disease Control and Prevention. 2020. https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/ ESRI. 2020. ArcGis Pro [Software]. Google. 2020a. https://www.gstatic.com/covid19/mobility/2020-0502_US_Iowa_Mobility_Report_en.pdf. Google. 2020b. https://www.google.com/covid19/mobility/. Google. 2020c. https://www.google.com/covid19/mobility/data_documentation.html?hl=en. Killeen, B. D., Wu, J. Y., Shah, K., Zapaishchykova, A., Nikutta, P., Tamhane, A., Chakraborty, S., Wei, J., Gao, T., Thies, M., & Unberath, M. 2020. https://arxiv.org/abs/2004.00756. Smith, K. 1993. https://doi.org/10.1002/j.1477-8696.1993.tb05828.x. United States Census Bureau. 2018. https://www.arcgis.com/home/item.html?id=45ede6d6ff7e4cbbbffa60d34227e462. Venter, Z. S., Barton, D. N., Gundersen, V., Figari, H., & Nowell, M. 2020. https://doi.org/10.1088/1748-9326/abb396.