Supplementary Information for

Policy and weather influences on mobility during the early US COVID-19 pandemic

Yihan Wu,[1*] Todd A. Mooring,[2] and Marianna Linz[1,2]

[1]*School of Engineering and Applied Sciences, Harvard University, Cambridge MA 02138;*
[2]*Department of Earth and Planetary Sciences, Harvard University, Cambridge MA 02138*

* Corresponding Author
Email: yihanwu@college.harvard.edu

This PDF file includes:

Supplementary Text
Figures S1 to S5
Table S1
SI References

Supplementary Information Text

As an additional check to ensure that our results are not an artifact of multiple testing, we used the false discovery rate-controlling procedure of Benjamini and Hochberg (1995).[1-3] In our context, the false discovery rate (FDR) is the expected fraction of states incorrectly detected as having a significant temperature-mobility correlation when such a correlation does not in fact exist. The procedure[4] requires as input the p-values of the family of tests to be analyzed, which we compute from the state-specific null distributions as described in the Methods section. We chose to analyze the temperature-park visitation and temperature-potential encounters correlations as separate families of 51 tests each (50 states plus D.C.) and to control the false discovery rate at 0.05 in each family.

After applying this procedure using the (2020) p-values and our chosen false discovery rate, we identified 11 states as having significant temperature-park visitation correlations and zero states with significant temperature-potential encounters correlations. This is compared to 17 states and one state, respectively, identified as having significant correlations before applying this multiple testing correction. Although fewer states are identified as having significant temperature-park visitation correlations than in our primary analysis, our main focus is not on the significance of correlations in individual states but on whether a temperature effect on human behavior is apparent in the dataset as a whole (i.e. nationwide). Thus, at the national scale we again observe statistically significant correlations between temperature and park visitation and a lack of such correlations between temperature and potential encounter rates, further supporting our conclusions. The results of the main analysis and the FDR analysis are shown in Table S1.

We also investigated an alternative method for evaluating the significance of the individual state-level temperature and mobility correlations, by estimating whether the correlations are significantly different from zero via a bootstrapping technique. By resampling days at random with replacement from our March 23 to May 1, 2020 time series only (i.e., without using previous years of weather data), we generate 4000 additional synthetic datasets of temperature and mobility per state and then use these synthetic datasets to recompute 4000 synthetic correlation values per state. These distributions of synthetic correlation values form the basis of the correlation significance information presented in Fig. S5. Note that the observed correlations shown in this figure are necessarily independent of this resampling technique and are therefore reproduced exactly from Figure 3.

It is important to note that these results are comparable to those obtained from our primary analysis using the state-specific null distributions of correlation values. In comparing Fig. 3 with Fig. S5, we observe similar relationships between the patterns of significant temperature-park visitations correlations and significant temperature-encounters correlations. While this bootstrapping technique identifies correlations in a few more states as significant, states identified as having significant correlations in our primary analysis are also found to have significant correlations in this tertiary analysis. We have not applied any form of multiple testing correction to our bootstrapping-based results.

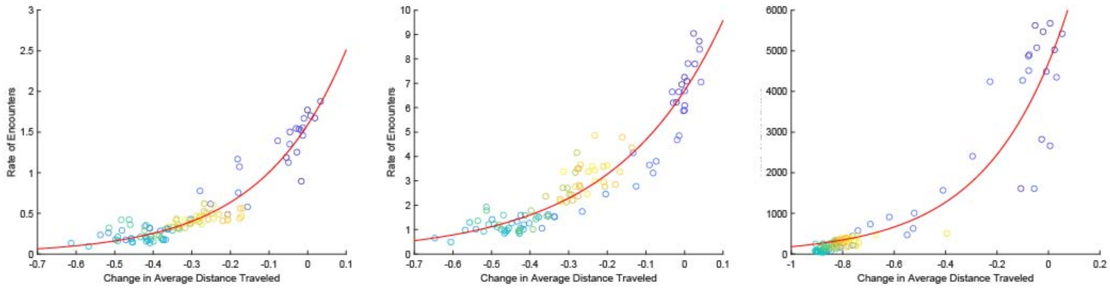

Fig. S1. Relationship between change in average distance traveled and change in rate of encounters in California (left), Florida (middle), and New York County (Manhattan), New York (right), colored by date (dark blue is February 24, light yellow is June 6). Data are expressed relative to pre-pandemic baselines. Exponential model fits are plotted in red.

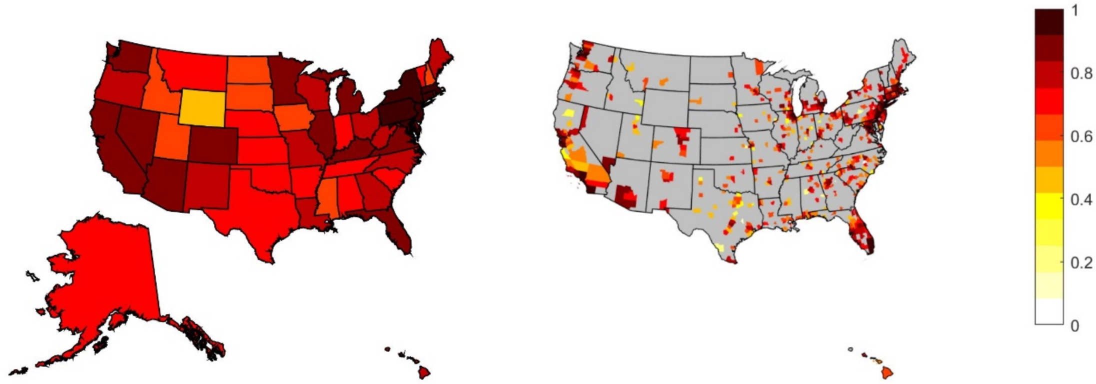

Fig. S2. Adjusted R^2 values of exponential fits on the state (left) and county (right) levels. Alaska is not shown on the county map as Anchorage Municipality is the only region available (R^2=0.68). Regions with missing data are colored in gray.

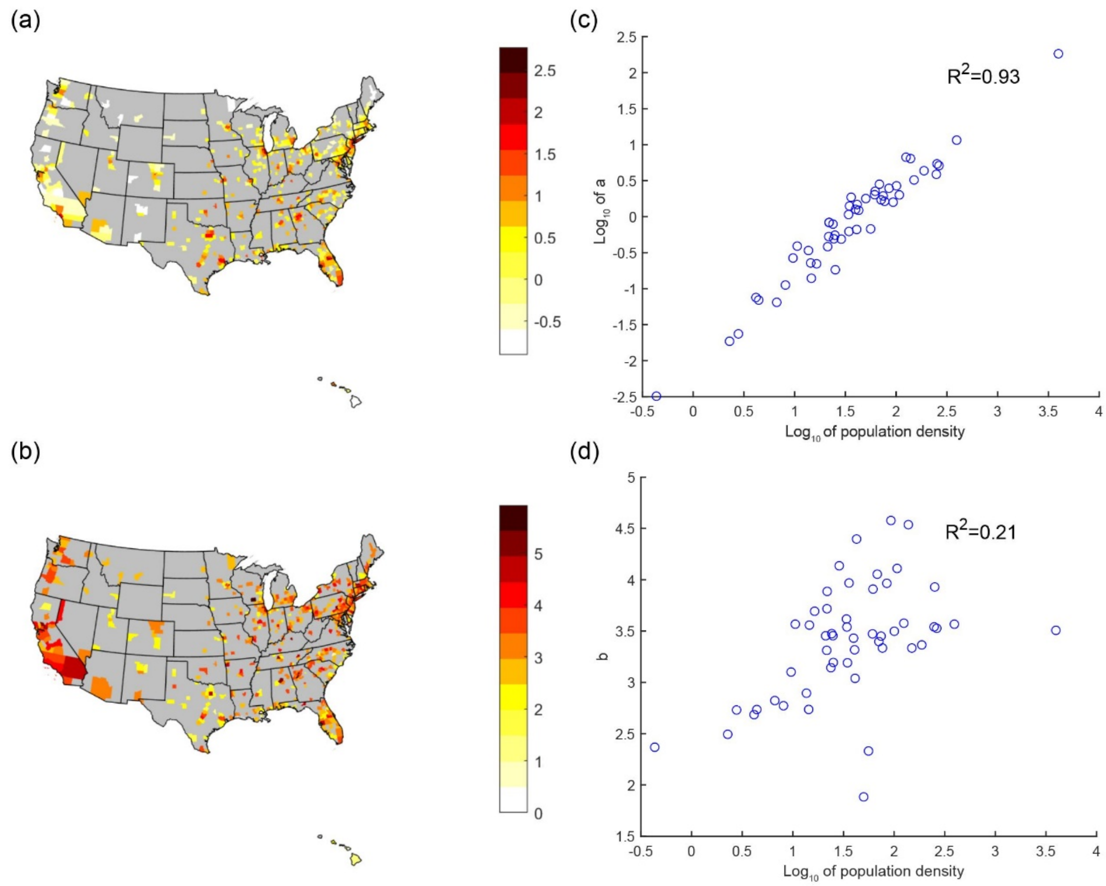

Fig. S3. Map of $\log_{10}$ of leading coefficient *a* (a) and map of growth rate coefficient *b* (b) on the county level. In (a), the value for New York County (Manhattan), New York ($\log_{10} a$ =3.60) is not shown. In (a) and (b), Alaska is not shown as Anchorage Municipality is the only region available ($\log_{10} a$=-0.213, *b*=2.28). Regions with missing data are colored in gray. Scatter plots of coefficients at the state level are shown against the $\log_{10}$ of population density for $\log_{10} a$ in (c) and for *b* in (d).

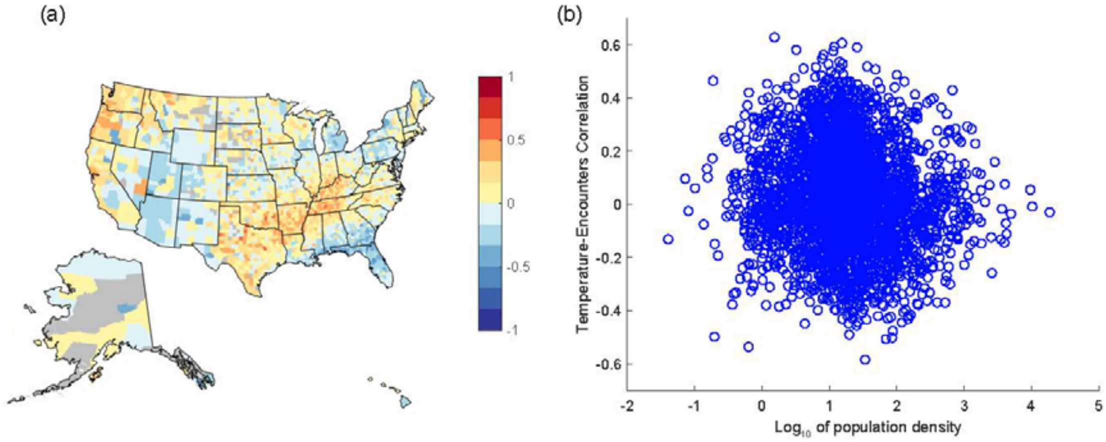

Fig. S4. (a) Observed county-level correlation coefficient of temperature and potential encounter rate. (b) Relationship between county-level correlation coefficient of temperature and potential encounter rate and $\log_{10}$ of population density.

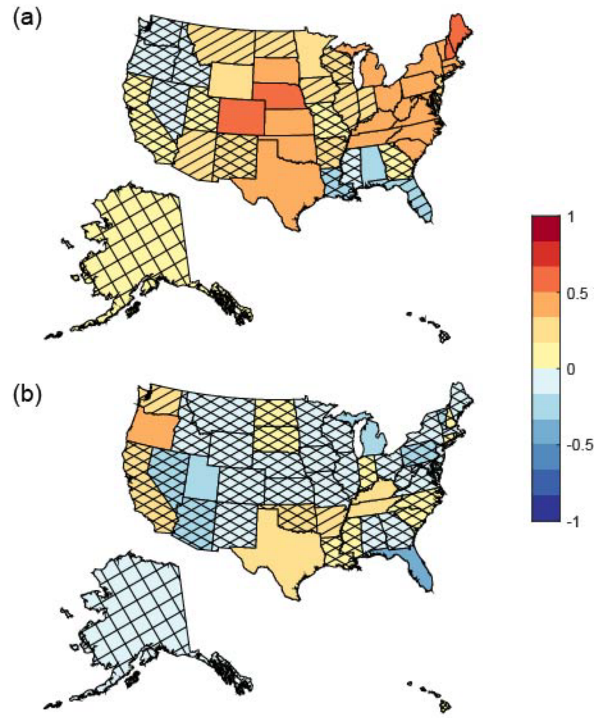

Fig. S5. State-level correlation coefficient of temperature and park visitation (a) or correlation coefficient of temperature and encounter rate (b), both reproduced from Fig. 3. Significance of the correlations was estimated using a bootstrapping technique (see Supplementary Information Text for details). Single hatching marks areas where zero is within 1.96 standard deviations of the observed correlation (i.e., significant at ~95% or less), while cross hatching shows where zero is within 1.64 standard deviations of the observed correlation (significant at ~90% or less). A lack of hatching indicates where zero is at least 1.96 standard deviations from the observed correlation (significant at ~95% or more). Rainy days have been excluded from this analysis and all time series were high-pass filtered prior to computing the correlations (see Methods for details).

Table S1. State-level temperature-mobility correlation results. Before applying the FDR procedure, 17 states exhibit significant temperature-park visitation correlations (bolded), and one state exhibits a significant temperature-potential encounters correlation (bolded). After applying the FDR procedure, 11 states exhibit significant temperature-park visitation correlations, and zero states exhibit significant temperature-potential encounters correlations as shown in the "FDR Significance" columns.

	Temperature-park visitations			Temperature-potential encounters		
State	Correlation Coefficient (2020)	p-value	FDR Significance	Correlation Coefficient (2020)	p-value	FDR Significance
AL	-0.3697	0.1467	N	-0.2041	0.5529	N
AK	0.1384	0.6961	N	-0.0080	0.9531	N
AZ	0.2753	0.1441	N	-0.2382	0.2523	N
AR	0.3099	0.3165	N	0.2333	0.2309	N
CA	0.0350	0.9412	N	0.1711	0.3305	N
CO	**0.6566**	**0.0018**	**Y**	0.0143	0.8723	N
CT	**0.3602**	**0.0250**	**N**	0.0247	0.8042	N
DE	0.0361	0.7895	N	-0.2572	0.3868	N
D.C.	**0.4150**	**0.0008**	**Y**	-0.1480	0.5412	N
FL	-0.2275	0.2240	N	**-0.4658**	**0.0245**	**N**
GA	0.1836	0.7998	N	-0.0874	0.8996	N
HI	-0.0846	0.6826	N	0.0935	0.7617	N
ID	0.0086	0.7714	N	-0.0245	0.8937	N
IL	0.2936	0.1084	N	-0.1944	0.5761	N
IN	0.3226	0.1282	N	-0.1092	0.9557	N
IA	0.2564	0.1040	N	-0.1758	0.5143	N
KS	**0.3781**	**0.0250**	**N**	-0.1247	0.5608	N
KY	0.3983	0.0823	N	0.2476	0.1824	N
LA	-0.2614	0.2690	N	0.0207	0.6278	N
ME	**0.5565**	**0.0001**	**Y**	-0.0056	0.8258	N
MD	**0.3479**	**0.0035**	**Y**	-0.0437	0.9147	N
MA	**0.3753**	**0.0068**	**Y**	-0.0112	0.7955	N
MI	**0.4656**	**0.0086**	**Y**	-0.3627	0.1405	N
MN	0.3110	0.0588	N	-0.0743	0.6386	N
MS	-0.1197	0.5993	N	0.0284	0.7167	N
MO	0.1851	0.4224	N	-0.1634	0.6676	N
MT	0.2379	0.3380	N	-0.0149	0.9648	N
NE	**0.5483**	**0.0008**	**Y**	-0.1312	0.5367	N
NV	-0.0226	0.8801	N	-0.2293	0.4560	N
NH	**0.4758**	**0.0007**	**Y**	0.0528	0.6252	N
NJ	0.2168	0.3225	N	-0.0569	0.7510	N
NM	0.1904	0.4584	N	-0.1154	0.6378	N
NY	0.3718	0.0665	N	-0.1245	0.6180	N
NC	0.3226	0.1095	N	0.0741	0.5831	N
ND	0.2566	0.2452	N	0.0343	0.9551	N
OH	**0.4970**	**0.0446**	**N**	-0.1913	0.9180	N
OK	0.4512	0.0736	N	0.1674	0.3496	N
OR	-0.0488	0.8803	N	0.3382	0.0539	N
PA	**0.4957**	**0.0107**	**Y**	-0.2853	0.3118	N
RI	**0.4587**	**0.0029**	**Y**	0.0948	0.6825	N
SC	0.3095	0.1091	N	-0.0495	0.9818	N

SD	**0.3993**	**0.0197**	**N**	0.0069	0.9309	N
TN	0.3463	0.0579	N	0.2405	0.2049	N
TX	0.3803	0.1613	N	0.1748	0.1069	N
UT	0.0574	0.9163	N	-0.3123	0.1269	N
VT	**0.5007**	**0.0022**	**Y**	-0.1210	0.3614	N
VA	**0.2930**	**0.0217**	**N**	-0.0892	0.8975	N
WA	0.0421	0.8459	N	0.3485	0.2398	N
WV	**0.4344**	**0.0350**	**N**	-0.2317	0.5929	N
WI	0.2321	0.1646	N	-0.1793	0.5066	N
WY	0.3178	0.1596	N	-0.0076	0.8520	N

Supplementary Information References

1. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289–300 (1995).
2. Y. Benjamini, D. Yekutieli, The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics* **29**, 1165–1188 (2001).
3. Y. Benjamini, Discovering the false discovery rate. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **72**, 405–416 (2010).
4. D. Groppe, fdr_bh. *MATLAB Central File Exchange* (February 27, 2021).