Appendix to the Paper "Using Mobile Device Data to Understand the Effect of Stay-at-Home Orders on Residents' Mobility"

Guihua Wang
Jindal School of Management, University of Texas at Dallas
guihua.wang@utdallas.edu

Appendix A: Related Literature on Stay-at-Home Orders

Table A1 Summary of Related Literature on Stay-at-Home Orders

Paper	Methodology	Mobility Data	Study Period
Abouk et al. 2020	difference-in-differences	Google	02/15/20 - 04/25/20
Alexander et al. 2020	difference-in-differences	Unacast	03/01/20 - 04/17/20
Andersen 2020	difference-in-differences	SafeGraph	01/01/19 - 06/14/20
Askitas et al. 2020	multi-events model	Google	01/03/20 - 02/06/20
Barrios et al. 2020	difference-in-differences	Unacast	02/24/20 - 03/31/20
Bonaccorsi et al. 2020	linear regression	Facebook	02/23/20 - 04/04/20
Brodeur et al. 2020	difference-in-differences	Google, Unacast	03/03/20 - 04/24/20
Bryant et al. 2020	Bayesian model	Google	02/05/20 - 03/29/20
Brzezinski et al. 2020a	difference-in-differences	SafeGraph	03/01/20 - 04/19/20
Brzezinski et al. 2020b	difference-in-differences	SafeGraph	01/01/20 - 04/23/20
Brzezinski et al. 2020c	difference-in-differences	SafeGraph	02/01/20 - 03/31/20
Bushman et al. 2020	Granger causality	SafeGraph	01/21/20 - 07/03/20
Charoenwong et al. 2020	difference-in-differences	SafeGraph	02/01/20 - 03/30/20
Chinazzi et al. 2020	transmission model	Baidu	12/31/19 - 02/11/20
Chiou et al. 2020	difference-in-differences	SafeGraph	02/01/20 - 03/21/20
Cook et al. 2020	difference-in-differences	Google, SafeGraph	02/15/20 - 05/27/20
Cronin et al. 2002	difference-in-differences	SafeGraph	01/01/20 - 05/15/20
Dasgupta et al. 2020	negative binomial	Descartes Labs, Google	04/15/20 - 04/17/20
Dave et al. 2020	difference-in-differences	SafeGraph	03/08/20 - 04/17/20
Ding et al. 2020	difference-in-differences	SafeGraph	01/22/20 - 04/30/20
Engle et al. 2020	difference-in-differences	Unacast	02/24/20 - 03/25/20
Fang et al. 2020	difference-in-differences	Baidu	01/01/20 - 02/29/20
Fowler et al. 2020	difference-in-differences	Not Applicable	03/01/20 - 05/07/20
Gao et al. 2020	linear regression	Descartes Labs, SafeGraph	03/11/20 - 04/10/20
Gatto et al. 2020	transmission model	Not Applicable	02/24/20 - 03/25/20
Grossman et al. 2020	difference-in-differences	SafeGraph	03/01/20 - 03/31/20
Guo 2020	synthetic control method	SafeGraph	02/01/20 - 04/14/20
Gupta et al. 2020	difference-in-differences	PlaceIQ, SafeGraph	01/01/20 - 03/31/20
Huang 2020	difference-in-differences	Google, PlaceIQ, SafeGraph	02/20/20 - 04/20/20
Jacobsen et al. 2020	summary statistics	Apple, Google	02/15/20 - 03/29/20
Jay et al. 2020	linear regression	SafeGraph	01/06/20 - 05/01/20
Klein et al. 2020	summary statistics	Cuebiq	01/01/20 - 03/25/20
Kraemer et al. 2020	generalized linear model	Baidu	01/01/20 - 02/04/20
Lin et al. 2020	difference-in-differences	Google	03/01/20 - 04/25/20
Luther 2020	difference-in-differences	Google	02/15/20 - 04/26/20
Lyu et al. 2020	event study	Not Applicable	03/21/20 - 05/15/20
McGrail et al. 2020	linear regression	Google	02/15/20 - 06/05/20
Painter et al. 2020	difference-in-differences	SafeGraph	01/01/20 - 04/23/20
Wang 2020	difference-in-differences	Google	03/07/20 - 04/03/20
Weill et al. 2020	difference-in-differences	Google, SafeGraph	01/01/20 - 04/21/20
Wellenius et al. 2020	regression discontinuity	Google	01/03/20 - 03/29/20
Wright et al. 2020	difference-in-differences	Unacast	02/23/20 - 05/01/20
Yang et al. 2020	transmission model	SafeGraph	03/01/20 - 06/07/20
Yilmazkuday 2020	difference-in-differences	Google	02/15/20 - 05/20/20

Note: This table summarizes related literature on stay-at-home orders. To provide a complete picture of the current literature, we include papers written both before and after the completion of this paper. However, it is worth noting that some of the papers included in this literature review are incomplete working papers with preliminary results that are subject to change.

Appendix B: Probit Analysis

Table B1 Results from the Probit Analysis

Variable	Coefficient	Standard Error
RaceBlack	0.073	0.054
RaceOthers	0.016	0.028
HispanicYes	0.017	0.040
HighSchoolYes	−0.087	0.158
InsuranceYes	0.073	0.107
StayAtHome	0.243	0.167
Number of States	50	
Adjusted R-Squared	0.277	

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. This table summarizes the results from the probit analysis, where the dependent variable is whether a state implemented a stay-at-home order during our study period and independent variables are state features and pre-treatment outcomes.

Appendix C: Robustness Checks

Table C1 Effect of Stay-at-Home Orders on Staying at Home (Leave-State-Out Analysis)

Left-Out State	Coef.	Std. Err.	Left-Out State	Coef.	Std. Err.
Alabama	2.804 ***	0.329	Missouri	2.908 ***	0.325
Alaska	2.840 ***	0.328	Montana	2.852 ***	0.328
Arizona	2.830 ***	0.328	Nebraska	2.785 ***	0.341
Arkansas	2.892 ***	0.346	Nevada	2.835 ***	0.327
California	2.773 ***	0.339	New Hampshire	2.834 ***	0.327
Colorado	2.828 ***	0.331	New Jersey	2.710 ***	0.307
Connecticut	2.814 ***	0.327	New Mexico	2.866 ***	0.328
Delaware	2.828 ***	0.327	New York	2.740 ***	0.316
DC	2.829 ***	0.326	North Carolina	2.859 ***	0.333
Florida	2.811 ***	0.330	North Dakota	2.825 ***	0.335
Georgia	2.779 ***	0.330	Ohio	2.824 ***	0.335
Hawaii	2.821 ***	0.327	Oklahoma	2.984 ***	0.316
Idaho	2.863 ***	0.326	Oregon	2.850 ***	0.329
Illinois	2.803 ***	0.339	Pennsylvania	2.821 ***	0.335
Indiana	2.842 ***	0.332	South Carolina	2.816 ***	0.330
Iowa	2.833 ***	0.346	South Dakota	2.834 ***	0.338
Kansas	2.852 ***	0.331	Tennessee	2.856 ***	0.329
Kentucky	2.922 ***	0.324	Texas	2.695 ***	0.312
Louisiana	2.856 ***	0.333	Utah	2.958 ***	0.318
Maine	2.851 ***	0.327	Vermont	2.831 ***	0.327
Maryland	2.810 ***	0.328	Virginia	2.831 ***	0.341
Massachusetts	2.799 ***	0.326	Washington	2.838 ***	0.330
Michigan	2.781 ***	0.331	West Virginia	2.854 ***	0.330
Minnesota	2.858 ***	0.332	Wisconsin	2.860 ***	0.331
Mississippi	2.821 ***	0.331	Wyoming	2.811 ***	0.329

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. This robustness check addresses a potential concern that our result in the main analysis is driven by a particular state, by leaving one state out each time and re-estimating the treatment effect.

Table C2 Effect of Stay-at-Home Orders on Staying at Home (Leave-Days-Out Analysis)

Left-Out	Coef.	Std. Err.	Left-Out	Coef.	Std. Err.
Days of Week			A Month		
Mondays	2.879 ***	0.322	January	2.721 ***	0.342
Tuesdays	2.871 ***	0.333	February	2.862 ***	0.353
Wednesdays	2.776 ***	0.338	March	2.574 ***	0.365
Thursdays	2.836 ***	0.335	April	3.082 ***	0.407
Fridays	2.823 ***	0.325	May	2.973 ***	0.354
Saturdays	2.846 ***	0.321	June	2.841 ***	0.317
Sundays	2.832 ***	0.326	July	2.767 ***	0.296
Weekdays/Weekends			August	2.812 ***	0.307
Weekdays	2.846 ***	0.321	September	2.842 ***	0.313
Weekends	2.746 ***	0.439	October	2.891 ***	0.324

Note: *** p < 0.01, ** p <0.05, * p < 0.1. This robustness check addresses a potential concern that our result in the main analysis is driven by a particular period of time, by performing three different leave-days-out analyses in which we leave out (1) the same day (e.g., Mondays) of a week, (2) weekdays or weekends, and (3) a month.

Table C3 Effect of Stay-at-Home Orders on Staying at Home (Other Robustness Checks)

Robustness Check	Coef.	Std. Err.	Robustness Check	Coef.	Std. Err.
Other Statewide Policies			Lagged StayAtHome		
Workplace Closures	2.373 ***	0.322	1 day before	0.988 ***	0.106
Mask Mandates	2.822 ***	0.319	2 days before	1.158 ***	0.132
Restaurant Restrictions	2.642 ***	0.314	3 days before	1.194 ***	0.134
School Closures	2.837 ***	0.330	4 days before	1.238 ***	0.140
All Measures	2.340 ***	0.319	5 days before	1.290 ***	0.150
County Fixed Effects	2.959 ***	0.316	6 days before	1.310 ***	0.155
Potential Endogeneity Issues	1.399 ***	0.302	7 days before	1.261 ***	0.154
Alternative Treatment Group	1.664 ***	0.473	$ln(StayAtHome)$	0.113 ***	0.014
Baseline StayAtHome	2.877 ***	0.320	Alternative Data	2.103 ***	0.011

Note: *** p < 0.01, ** p <0.05, * p < 0.1. This table summarizes the results from seven robustness checks in which we (1) control for other statewide policies such as workplace closures, mask mandates, restaurant restrictions, and school closures, (2) include county fixed effects to control for systematic differences across counties, (3) address potential endogeneity issues by using counties with early implementation dates as a control group and those with late implementation dates as a treatment group, (4) use nearby states as an alternative treatment group, (5) include the baseline percentage of residents staying at home or a lagged version of the dependent variable as an additional independent variable, (6) use the log-transformed percentage of residents staying at home, $ln(StayAtHome)$, as an alternative dependent variable, or (7) use alternative data from SafeGraph.

Table C4 Heterogeneous Effects of the Orders on Staying at Home (Alternative Data)

	(1) Uninsured		(2) NoHighSchool	
Variable	Coef.	Std. Err.	Coef.	Std. Err.
Order Implementation	2.408***	0.017	2.465***	0.016
× Target Feature	−0.043***	0.001	−0.030***	0.001
Order Expiration	Included		Included	
× Target Feature	Included		Included	
Number of Observations	9,006,184		9,006,184	
Adjusted R-Squared	0.350		0.350	

Note: *** p < 0.01, ** p <0.05, * p < 0.1. This robustness check uses SafeGraph's Social Distancing Metrics as alternative data. In columns (1)–(2), the target features are *Uninsured* and *NoHighSchool*, respectively.

Appendix D: Effect of Stay-at-Home Orders on Work and Non-Work Trips

Figure D1 compares the number of work and non-work trips in the control and treatment groups over different days. To examine the treatment effect heterogeneity across insurance, we divide counties into two categories of equal size based on the percentage of residents without health insurance. We informally designate these two categories as "Counties with Higher Percentages of Residents without Health Insurance" and "Counties with Lower Percentages of Residents without Health Insurance". For each category, we compare the number of work (Figure D2) and non-work (Figure D4) trips. Similarly, to examine the treatment effect heterogeneity across education, we divide counties into two categories of equal size based on the percentage of residents who did not attend high school. We informally designate these two categories as "Counties with Higher Percentages of Residents without High School Diploma" and "Counties with Lower Percentages of Residents without High School Diploma". For each category, we compare the number of work (Figure D3) and non-work (Figure D5) trips. In all figures, the two vertical dashed lines indicate the first (March 19, 2020) and last (April 7, 2020) orders implemented different states.

Figure D1 Compare Control and Treatment Counties (Number of Work and Non-Work Trips)

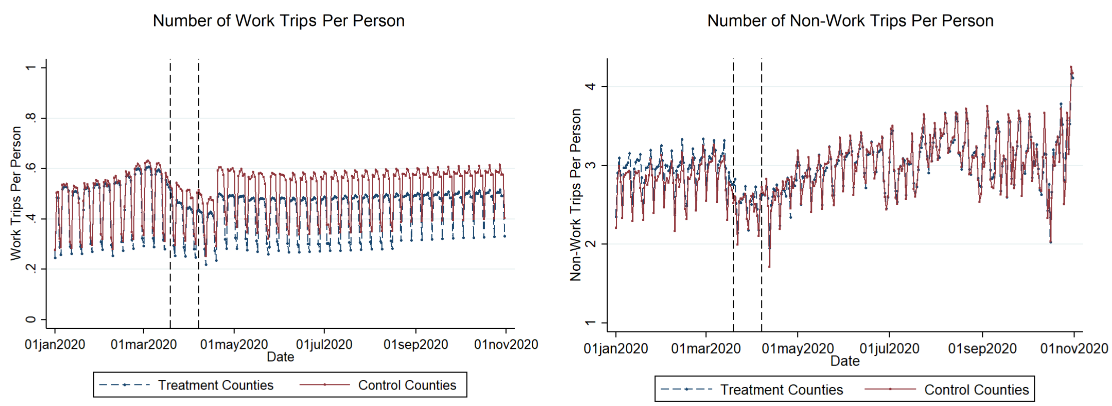


Figure D2 Comparison by Percentage of Residents without Health Insurance (Work Trips)

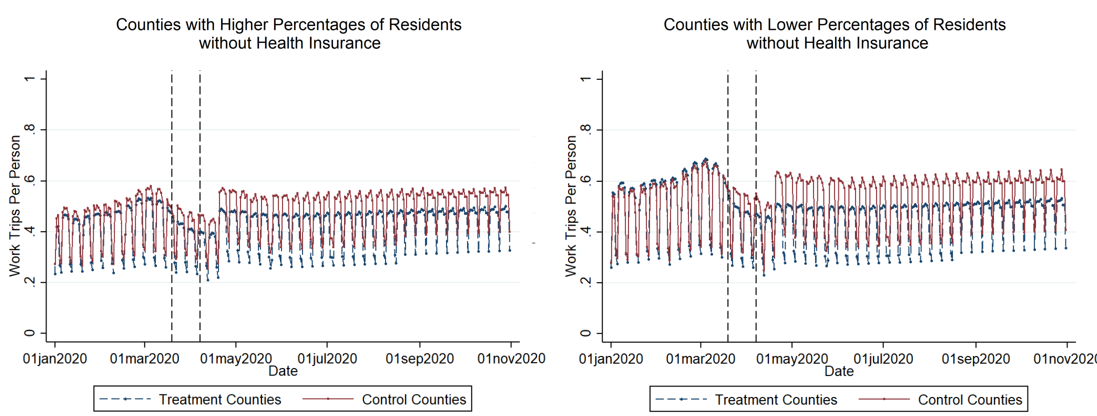


Figure D3 **Comparison by Percentage of Residents without High School Diploma (Work Trips)**

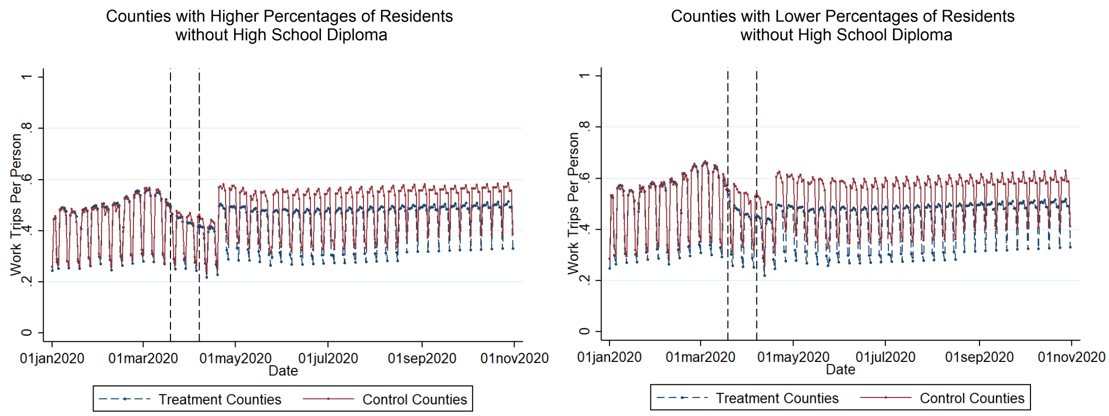


Figure D4 **Comparison by Percentage of Residents without Health Insurance (Non-Work Trips)**

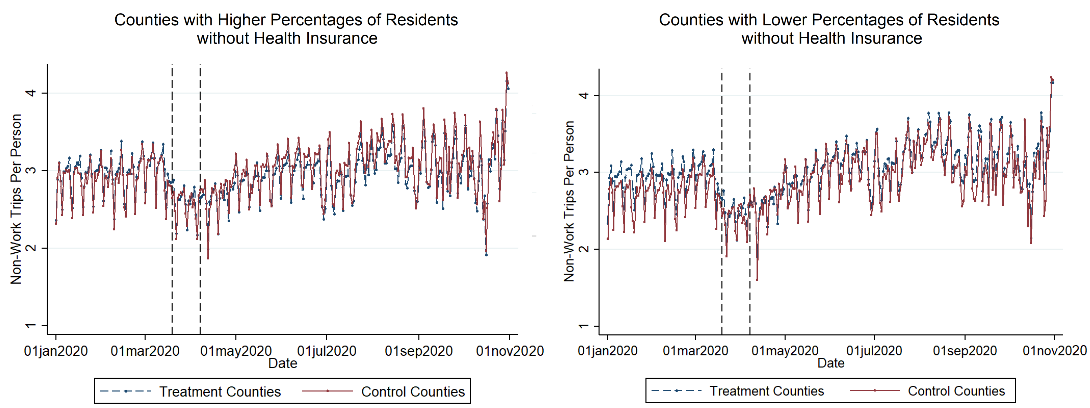
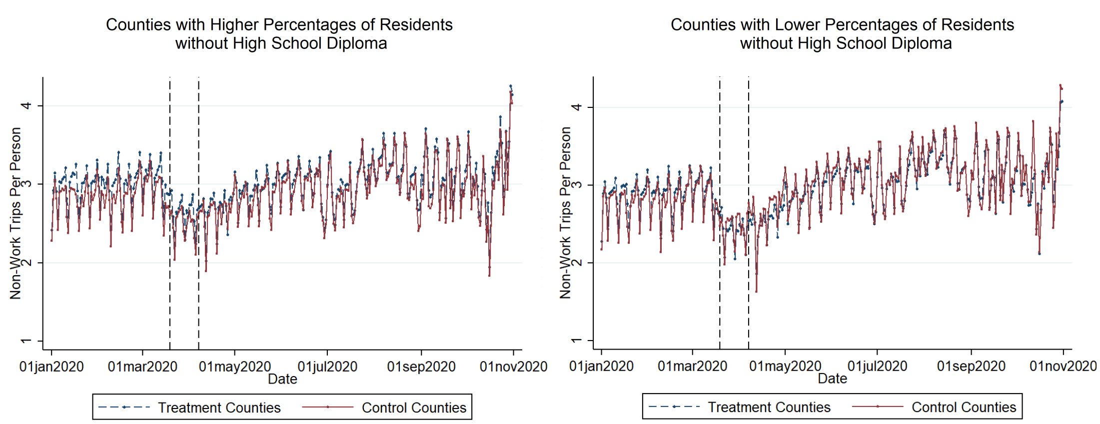


Figure D5 **Comparison by Percentage of Residents without High School Diploma (Non-Work Trips)**



Appendix E: Impact of Mobility on Infections

In this section, we first describe our empirical strategy for estimating the impact of mobility on infections. We then describe the two-stage least square regressions (2SLS) model to address potential endogeneity issues. Finally, we discuss the results from the 2SLS model and perform robustness checks to analyze the sensitivity of our result.

E.1. Empirical Strategy

Analyzing the impact of mobility on infections presents two major challenges. The first challenge is endogeneity issue. There might be unobservable factors that affect both residents' mobility and infections. For example, wearing a mask might increase residents' mobility and, at the same time, reduce infections. The second challenge is reverse causality. While an increase in residents' mobility might increase infections, an increase in infections might reduce residents' mobility. Therefore, simply regressing infections on residents' mobility might lead to biased results.

To address the endogeneity issue, we follow Glaeser et al. (2020) by using residential teleworkable shares as an instrumental variable for residents' mobility. As described in Glaeser et al. (2020), the share of teleworking workers is a valid instrument, because it correlates residents' mobility (i.e., the relevance condition) but does not directly impact infections (i.e., the exclusion condition). To address the second challenge, we use a lagged version of residents' mobility as an independent variable. The intuition behind this approach is that the lagged mobility impacts infections but not vice versa. This approach is consistent with existing medical literature (see e.g., Li et al. 2020) that there is a lag between exposure and infection.

E.2. Econometric Model

To describe the econometric model, we let $Infection_{it}$ denote per capita infection rate in count i on day t.[1] We follow existing public health literature (see e.g., Silva 2020) to calculate per capita infection rate by dividing the number of infections by the number of residents in a county. We use per capita infection rate instead of the number of infections as the dependent variable, because correcting for population size allows us to compare counties of different sizes. As robustness check, we use the number of infections as an alternative dependent variable to check the sensitivity of our results.

Our main independent variable of interest is the percentage of residents staying at home in county i on day t (denoted by $StayAtHome_{it}$). We include county features (denoted by $Feature_i$) to control for the differences across counties, a vector of state dummies (denoted by $State_i$) to control for the systematic differences across these counties' states, and a vector of day dummies (denoted by Day_t) to control for time trends of infection rate. To describe the instrumental variable, we denote by $Teleworking_i$ the share of teleworking workers in county i. We estimate the effect of staying at home on infection rate using the 2SLS model:

$$StayAtHome_{it} = \alpha_0 + \alpha_1 Teleworking_i + \alpha_2 State_{j(i)} + \alpha_3 Day_t + \alpha_4 Feature_i + \varepsilon_{it}, \tag{1}$$

$$Infection_{i,t+k} = \beta_0 + \beta_1 \widehat{StayAtHome}_{it} + \beta_2 State_{j(i)} + \beta_3 Day_t + \beta_4 Feature_i + \zeta_{it}, \tag{2}$$

where k indicates the lag between staying at home and infection, $\widehat{StayAtHome}_{it}$ is the predicted value from equation (3), β_1 is the bias-corrected 2SLS estimator of the impact of staying at home on infection rate, and ε_{it} and ζ_{it} are idiosyncratic errors.

E.3. Results and Discussion

The results from the first-stage equation are summarized in the first row of Table E1. We see the coefficient of the instrumental variable is significantly different from zero at the 1% significance level, which suggests the instrumental variable has a strong first stage. The positive coefficient suggests a positive correlation between the share of teleworking workers and the percentage of residents staying at home.

Table E1 Impact of Mobility on Infections

	IV Regression		Non-IV Regression	
Variable	Coef.	Std. Err.	Coef.	Std. Err.
First Stage: IV	0.202 ***	0.002		
StayAtHome	−4.844 ***	0.728	1.303 ***	0.079
County Features	Included		Included	
Day Fixed Effects	Included		Included	
State Fixed Effects	Included		Included	
Number of Observations	924,396		924,396	

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The results from the second-stage equation are summarized in the second row of Table E1.[2] We see the coefficient of $StayAtHome$ is negative and significantly different from zero at the 1% significance level, which suggests infection rate decreases as the percentage of residents staying at home increases. More specifically, a coefficient of -4.844 means a one percentage point increase in the percentage of residents staying at home reduces infection rate by 4.844.

As a comparison, the right side of Table E1 summarizes the results from a multivariate regression model without an instrumental variable. We see the coefficient of $StayAtHome$ is positive and significantly different from zero at the 1% significance level, which suggests infection rate increases as the percentage of residents staying at home increases. This comparison suggests ignoring potential endogeneity issues leads to a biased estimate of the impact of staying at home on infection rate.

E.4. Robustness Checks

We perform five robustness checks to analyze the sensitivity of our results. The first robustness check performs leave-state-out analyses by leaving out one state at a time. The second robustness check performs leave-days-out analyses by leaving the same day of a week, weekdays or weekends, or a specific month at a time. The third robustness check uses the independent variable with a lag of 0, 1, ..., 6, or 14 days. The fourth robustness check includes mask mandates as an additional independent variable. The fifth robustness check uses the log-transformed infection rate and the number of infections as alternative dependent variables.

Tables E2, E3, and E4 summarize the results from these robustness checks. We see all coefficients of $StayAtHome$ are negative and significantly different from zero at the 1% significance level, which suggests an increase in the percentage of residents staying at home reduces infection rate. Note that the results from the last two robustness checks are different from that in the main analysis because they use different dependent variables. These robustness checks suggest our results are not driven by a particular state or period, the value of a lag, other statewide policies such as mask mandates, or the form of the dependent variable.

Table E2 Results from the Instrumental Variable Model (Leave-State-Out Analysis)

Left-Out State	Coef.	Std. Err.	Left-Out State	Coef.	Std. Err.
Alabama	-4.741 ***	0.728	Missouri	-4.536 ***	0.754
Alaska	-4.858 ***	0.745	Montana	-4.693 ***	0.716
Arizona	-4.826 ***	0.725	Nebraska	-4.196 ***	0.699
Arkansas	-4.646 ***	0.732	Nevada	-4.886 ***	0.731
California	-5.089 ***	0.790	New Hampshire	-4.815 ***	0.726
Colorado	-4.272 ***	0.697	New Jersey	-4.891 ***	0.733
Connecticut	-4.859 ***	0.729	New Mexico	-4.849 ***	0.757
Delaware	-4.856 ***	0.729	New York	-5.389 ***	0.749
DC	-4.844 ***	0.728	North Carolina	-4.989 ***	0.733
Florida	-4.028 ***	0.725	North Dakota	-5.252 ***	0.745
Georgia	-4.955 ***	0.714	Ohio	-4.862 ***	0.773
Hawaii	-4.844 ***	0.727	Oklahoma	-4.653 ***	0.732
Idaho	-4.975 ***	0.729	Oregon	-4.689 ***	0.732
Illinois	-4.930 ***	0.736	Pennsylvania	-5.243 ***	0.735
Indiana	-5.196 ***	0.753	South Carolina	-5.096 ***	0.737
Iowa	-4.358 ***	0.728	South Dakota	-5.035 ***	0.729
Kansas	-5.302 ***	0.738	Tennessee	-5.259 ***	0.674
Kentucky	-5.501 ***	0.924	Texas	-2.921 ***	0.744
Louisiana	-5.001 ***	0.718	Utah	-4.938 ***	0.710
Maine	-4.840 ***	0.727	Vermont	-4.857 ***	0.727
Maryland	-4.890 ***	0.734	Virginia	-4.655 ***	0.718
Massachusetts	-5.015 ***	0.747	Washington	-5.049 ***	0.724
Michigan	-4.832 ***	0.738	West Virginia	-4.964 ***	0.760
Minnesota	-4.751 ***	0.734	Wisconsin	-4.484 ***	0.739
Mississippi	-4.987 ***	0.717	Wyoming	-5.333 ***	0.758

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table E3 Results from the Instrumental Variable Model (Leave-Days-Out Analysis)

Left-Out	Coef.	Std. Err.	Left-Out	Coef.	Std. Err.
Days of Week			A Month		
Mondays	-4.785 ***	0.801	January	-5.205 ***	0.782
Tuesdays	-4.875 ***	0.818	February	-5.279 ***	0.792
Wednesdays	-4.888 ***	0.814	March	-5.556 ***	0.809
Thursdays	-4.252 ***	0.802	April	-4.843 ***	0.831
Fridays	-5.063 ***	0.677	May	-3.503 ***	0.732
Saturdays	-4.879 ***	0.775	June	-4.029 ***	0.783
Sundays	-5.152 ***	0.807	July	-4.170 ***	0.771
Weekdays/Weekends			August	-3.798 ***	0.694
Weekdays	-3.653 ***	1.273	September	-6.046 ***	0.745
Weekends	-5.246 ***	0.872	October	-5.951 ***	0.721

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table E4 Results from the Instrumental Variable Model (Other Robustness Checks)

Robustness Check	Coefficient	Standard Error
Include Lagged StayAtHome		
same day	$-4.753***$	0.714
1 day before	$-4.752***$	0.714
2 days before	$-4.767***$	0.716
3 days before	$-4.785***$	0.719
4 days before	$-4.799***$	0.721
5 days before	$-4.815***$	0.724
6 days before	$-4.834***$	0.726
14 days before	$-4.953***$	0.743
Control for Mask Mandates	$-4.844***$	0.727
$ln(InfectionRate)$ as Dependent Variable	$-0.049***$	0.004
Number of Infections as Dependent Variable	$-0.414***$	0.124

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.