

APPENDIX

This appendix consists of three parts:

- The appendix to the final published version of the article
- The Fast Track Ahead Of Print version of the article, which was published online on May 14, 2020
- The appendix published online with that article, consisting of the tables and figure that now appear in the final published version of the article (*Health Affairs*, vol. 39, no. 7, July 2020)



Appendix

A. Adjustments for Likely Data Errors

The original county-by-day cases data contained some likely reporting errors. First, the COVID-19 case counts on Sunday, March 15, 2020 were identical to those from March 14, 2020 in all large counties that had cases at the time, even though substantial growth was reported on March 16, 2020. This is almost certainly due to lack of reporting on that particular Sunday rather than a true lack of new cases. We therefore imputed the cases for March 15, 2020 so that growth from March 14, 2020 to March 16, 2020 was evenly split between the two days.

Additionally, 964 of the 182,004 county-by-day observations reported a reduction in cases from the previous day, which is implausible since case counts were cumulative. Approximately half of these were cases where the count increased by one on one day and then decreased by one the next. In such cases, the original increase was likely an error (e.g. presumed case that was later found to be negative, duplicate, or person who lived in a different county/state), so we removed the initial increase so that the case count remained constant over the three days. We set remaining instances of negative growth rates to zero. As shown later in this appendix, our results were virtually unchanged if we simply dropped observations with case reductions as well as all observations on March 15, 2020.

We also corrected some apparent errors in the social distancing policy dates. Many Kansas counties had reported dates for gathering bans that were after the date for the statewide ban, so we changed those dates to match the statewide ban.¹ We removed the reported SIPO in Wyoming, which was listed as March 28, 2020, as sources such as the Kaiser Family Foundation² as well as various news articles³ listed the state as not having any SIPO in place after the date listed. Some school closing dates also appeared incorrect. School closing dates were checked against those listed at edweek.org⁴. If there was more than a day mismatch, then a third source was found to decide the date used, leading to changes for Alabama⁵, Alaska⁶, Iowa⁷, and Wyoming⁸. We changed the SIPO start date for MS⁹ and gathering ban date for KY¹⁰ as they

¹ The counties were Allen, Anderson, Atchison, Barber, Barton, Bourbon, Brown, Butler, Chase, Chautauqua, Cherokee, Cheyenne, Clark, Clay, Cloud, Coffey, Comanche, Cowley, Crawford, Decatur, Dickinson, and Doniphan.

² See: <https://www.kff.org/health-costs/issue-brief/state-data-and-policy-actions-to-address-coronavirus/>

³ See: <https://www.usnews.com/news/best-states/wyoming/articles/2020-03-30/wyoming-governor-still-no-plan-for-state-stay-at-home-order>, <https://abcnews.go.com/Health/9-states-issue-formal-stay-home-orders-amid/story?id=69959039>

⁴ See: <https://www.edweek.org/ew/section/multimedia/map-coronavirus-and-school-closures.html>

⁵ See: <https://www.al.com/news/2020/03/alabama-closes-all-k-12-schools.html>

⁶ See: <https://www.adn.com/alaska-news/education/2020/03/16/alaskas-statewide-school-closure-is-about-to-begin-no-one-knows-quite-how-it-will-work/>

⁷ See: <https://www.usnews.com/news/best-states/iowa/articles/2020-03-16/iowa-schools-closing-taking-coronavirus-cue-from-governor>

⁸ See: https://trib.com/news/local/education/school-districts-across-wyoming-announce-closures-following-governors-recommendation/article_333c5100-c6a6-57f1-b04b-948546b59ffd.html

⁹ See: https://www.sos.ms.gov/Content/documents/ed_pubs/Exec%20Orders/1466.pdf

¹⁰ See: https://governor.ky.gov/attachments/20200319_Order_Mass-Gatherings.pdf

were incorrect in the source data. The SIPOs for IN, MN, NH, OH, and TN had an effective start time of 11:59 pm. We changed the start day for these states to the following day as they were in effect for only one minute on their actual start date. Below we list the state and county (if different from state) policy dates that we ultimately used.

Appendix Exhibit 1 – Effective Dates of State Social Distancing Measures

State	Shelter in Place	No Gatherings of 500+	Public Schools Closed	Rest./ Ent./Gym Order
AL	4/4	3/13	3/16	3/19
AK	3/28	3/24	3/19	3/17
AZ	3/31	3/17	3/16	3/17
AR		3/26	3/17	3/19
CA	3/19	3/19	3/19	3/15
CO	3/26	3/13	3/23	3/17
CT	3/23	3/23	3/17	3/16
DE	3/24	3/24	3/16	3/12
DC	4/1	3/25	3/16	3/16
FL	4/3	3/17	3/16	3/20
GA	4/3	3/17	3/18	3/20
HI	3/25	3/25	3/23	3/20
ID	3/25	3/25	3/23	3/23
IL	3/21	3/18	3/17	3/18
IN	3/25	3/13	3/19	3/17
IA		3/17	4/2	3/17
KS	3/30	3/17	3/18	3/24
KY	3/26	3/20	3/16	3/16
LA	3/23	3/17	3/16	3/17
ME	4/2	3/18	3/15	3/18
MD	3/30	3/13	3/16	3/16
MA	3/24	3/17	3/17	3/17
MI	3/24	3/24	3/16	3/16
MN	3/28	3/27	3/18	3/17
MS	4/3	3/25	3/20	3/25
MO	4/6	3/21	3/19	3/21
MT	3/28	3/28	3/16	3/20
NE		3/19	3/19	4/4
NV	4/1	3/24	3/16	3/20
NH	3/28	3/24	3/16	3/16
NJ	3/21	3/16	3/18	3/16
NM	3/24	3/18	3/16	3/18
NY	3/22	3/13	3/18	3/16
NC	3/30	3/12	3/16	3/17
ND			3/16	3/19
OH	3/24	3/15	3/17	3/15

OK		3/29	3/17	4/1
OR	3/23	3/16	3/16	3/17
PA	4/1	3/23	3/16	3/17
RI	3/28	3/16	3/16	3/23
SC	4/7	3/18	3/16	3/18
SD			3/16	3/23
TN	4/1	3/25	3/20	3/23
TX	4/2	3/14	3/23	3/19
UT		3/12	3/16	3/18
VT	3/25	3/25	3/18	3/17
VA	3/30	3/24	3/16	3/16
WA	3/23	3/11	3/17	3/17
WV	3/24	3/24	3/16	3/24
WI	3/25	3/17	3/18	3/17
WY		3/20	3/20	3/19

Appendix Exhibit 2 – Effective Dates of County Social Distancing Measures

State	County	Date Issued	State	County	Date Issued
AL	Jefferson	3/26	MO	Randolph	3/25
AK	Anchorage	3/22	MO	Ray	3/25
CO	Adams	3/24	MO	St. Louis	3/23
CO	Arapahoe	3/24	PA	Allegheny	3/23
CO	Archuleta	3/24	PA	Beaver	3/28
CO	Boulder	3/24	PA	Berks	3/27
CO	Denver	3/24	PA	Bucks	3/23
CO	Douglas	3/24	PA	Butler	3/27
CO	Eagle	3/24	PA	Centre	3/28
CO	Grand	3/24	PA	Chester	3/23
CO	Jefferson	3/24	PA	Delaware	3/23
CO	La Plata	3/24	PA	Erie	3/24
CO	Pitkin	3/24	PA	Lackawanna	3/27
CO	San Miguel	3/24	PA	Lancaster	3/27
FL	Alachua	3/25	PA	Lehigh	3/25
FL	Broward	3/27	PA	Luzerne	3/27
FL	Duval	3/25	PA	Monroe	3/23
FL	Hillsborough	3/27	PA	Montgomery	3/23
FL	Leon	3/25	PA	Northampton	3/25
FL	Miami-Dade	3/25	PA	Philadelphia	3/23
FL	Monroe	3/30	PA	Pike	3/27
FL	Orange	3/26	PA	Washington	3/28
FL	Osceola	3/26	PA	Wayne	3/27
FL	Palm Beach	3/30	PA	Westmoreland	3/27
FL	Pinellas	3/26	PA	York	3/27
GA	Chatham	3/24	TN	Davidson	3/25
GA	Dougherty	3/21	TN	Franklin	3/25
GA	Fulton	3/24	TX	Austin	3/24
KS	Johnson	3/24	TX	Bell	3/24
KS	Leavenworth	3/24	TX	Bexar	3/24
KS	Sedgwick	3/25	TX	Brazos	3/24
KS	Wyandotte	3/24	TX	Collin	3/24
MS	Lafayette	3/22	TX	Dallas	3/23
MS	Lauderdale	3/31	TX	El Paso	3/24
MO	Boone	3/25	TX	Fort bend	3/24
MO	Cass	3/24	TX	Galveston	3/24
MO	Clay	3/24	TX	Harris	3/24
MO	Cole	3/23	TX	McLennan	3/24
MO	Greene	3/26	TX	Tarrant	3/24
MO	Jackson	3/24	UT	Salt Lake	3/30
MO	Jefferson	3/24	UT	Summit	3/27
MO	Platte	3/24			

B. Event-Study Econometric Model

$$\begin{aligned}
& \ln(cases_{cd}) - \ln(cases_{c,d-1}) \\
&= \beta_0 + \sum_{i=-2}^{-1} \beta_{1i} SIPO_{cdi} + \sum_{i=1}^5 \gamma_{1i} SIPO_{cdi} + \sum_{i=-2}^{-1} \beta_{2i} schools_{cdi} \\
&+ \sum_{i=1}^5 \gamma_{2i} schools_{cdi} + \sum_{i=-2}^{-1} \beta_{3i} events_{cdi} + \sum_{i=1}^5 \gamma_{3i} events_{cdi} \\
&+ \sum_{i=-2}^{-1} \beta_{4i} entertainment_{cdi} + \sum_{i=1}^5 \gamma_{4i} entertainment_{cdi} \\
&+ \sum_{i=0}^1 \alpha_i (\ln(tests_{c,d-i}) - \ln(tests_{c,d-1-i})) + county_c + day_by_cd_{cd} + \varepsilon_{cd}
\end{aligned}$$

$cases_{cd}$ = cumulative confirmed COVID-19 cases in county c on day d (with 1 added to prevent the variable from being undefined when the county does not yet have a case).

$SIPO_{cdi}$ = indicator for whether a shelter-in-place order was enacted i time periods before day d in county c , where the time periods are defined as

$$i = \left\{ \begin{array}{l} -2 \equiv 10 \text{ or more days before } d \\ -1 \equiv 5 \text{ to } 9 \text{ days before } d \\ 1 \equiv 1 \text{ to } 5 \text{ days after } d \\ 2 \equiv 6 \text{ to } 10 \text{ days after } d \\ 3 \equiv 11 \text{ to } 15 \text{ days after } d \\ 4 \equiv 16 \text{ to } 20 \text{ days after } d \\ 5 \equiv 21 \text{ or more days after } d \end{array} \right\}, \text{ leaving } 0 \text{ to } 4 \text{ days before } d \text{ as the omitted}$$

reference range; i.e. the coefficients reflect differences between the predicted growth rate under a policy enacted i periods ago versus one that will be enacted within the next four days.

$schools_{cdi}$ = indicator for whether schools were closed starting i periods before day d in county c .

$events_{cdi}$ = indicator for whether large events were banned (at least as stringent as 500 people) starting i periods before day d in county c .

$entertainment_{cdi}$ = indicator for whether restaurants or gyms/entertainment centers were closed starting i periods before day d in county c .

$tests_{cd}$ = cumulative COVID-19 tests recorded in the state containing county c on day d (with 1 added to prevent the variable from being undefined when the state has not yet run a test); the summation term before the log-difference in tests reflects the fact that the lag of the log difference is also included.

$county_c$ = fixed effect for county c ; this captures unobserved county-level variables that do not vary during our sample period, which is why we do not control for county characteristics, such as population density, demographics, or voting behavior, that might influence case growth rates.

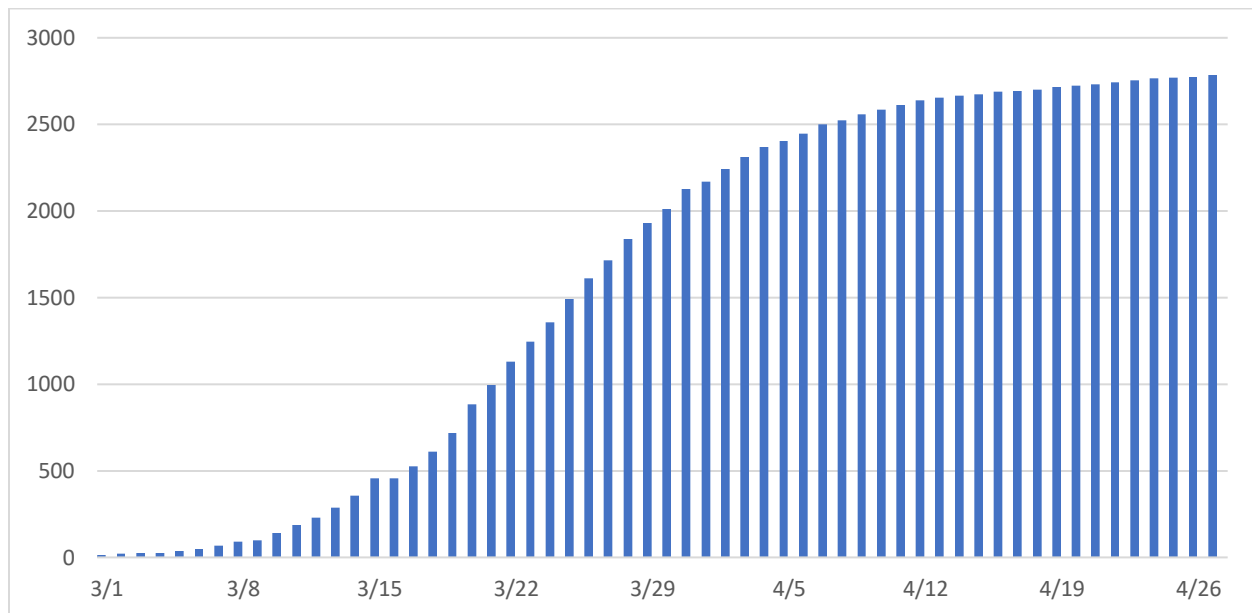
$day_by_cd_{cd}$ = fixed effect for day d (ranging from March 1 to April 27, 2020) in the Census Division containing county c ; this captures unobserved variables that trend over time, and allows these time trends to vary by Census Division.

ε_{cd} = error term for county c on day d

C. Diffusion of COVID-19 during Sample Period

The figure below shows the growth in the number of counties with confirmed COVID-19 cases during our sample period of March 1 through April 27, 2020. These are simple counts, not adjusted for population. 0.4% of counties had a COVID-19 case on 3/1, but 11% of the US population lived in those counties. Similarly, more than half of US residents lived in counties with a case by March 13, even though just 9% of counties had a case. By April 1, 97% of residents lived in counties with a case, compared to 71% of the unadjusted number of counties.

Appendix Exhibit 3 – Number of Counties with Confirmed COVID-19 Cases on Each Day



Note: All cases for Bronx, Kings, Queens, and Richmond counties in NY were lumped into New York county in the source data.

D. Robustness Checks

Multicollinearity

Our event-study specification included 24 different treatment variables (6 for each of the 4 policies) as well as Census-Division-by-day fixed effects, raising the question of whether there was sufficient identifying variation to precisely identify the impacts of the different policies. In particular, this could conceivably have explained the null results for public school closures and large event bans. Our first two robustness checks therefore utilized a simpler specification that included only a single variable for each policy type: whether the policy was in effect 10 days ago in the first regression and 20 days ago in the second. (The lag allowed for incubation periods, testing delays, etc.) Also, we included day fixed effects rather than their interactions with each of the nine Census Divisions. This means that policy differences across Census Divisions were allowed to contribute to the identifying variation, whereas the baseline model relied only on within-Census-Division variation.

The results are below. We observed the same pattern of results as we did in our main model: SIPOs had the largest statistically significant effect, restaurant/entertainment closures had the second largest, and large event bans and school closures had no discernable impacts.

Appendix Exhibit 4 – Estimated Effect of Social Distancing Policies on the Growth Rate of COVID-19 Cases in Models with a Single Variable for Each Policy

	10-Day Lag	20-Day Lag
Shelter-in-Place Order in Effect	-3.997 (0.705)*	-4.106 (0.631)*
Restaurant/Gym/Entertainment Closures	-2.417 (0.726)*	-1.951 (0.797)*
No Large Events	-0.486 (1.020)	-0.953 (1.136)
Public Schools Closed	1.705 (1.062)	0.168 (0.903)

Notes: Sample size = 182,004 county-by-day observations. * indicates statistically significant at the 5% level; + 10% level. Standard errors, in parentheses, are heteroskedasticity-robust and clustered by state. Observations are weighted by county population. Regressions include the growth rate in cumulative tests in each of the past two days as well day and county fixed effects.

Note also that, though we do not show the results here, the estimated treatment effects from our baseline event-study model were not meaningfully sensitive to the use of region-by-day or simply day fixed effects instead of Census Division-by-day fixed effects. However, the pre-trend for SIPOs was worse using the sparser specifications.

Testing

Here we conduct robustness checks aimed to ensure that the results were not driven by the way in which we controlled for testing. Specifically, we were concerned about two issues. First, the baseline model's functional form for tests was the daily log difference in cumulative tests (as well as its first lag), which mirrored our specification for the outcome variable, effectively assuming that a given percentage increase in testing corresponded to a given percentage increase in cases. However, it is possible that the relationship between testing and cases is better characterized by a different functional form. Second, number of tests performed depends not only on availability, but also the severity of the outbreak. The coefficient estimates for the testing variables were therefore likely biased, and bias could have spilled over to the coefficient estimates for the policy variables.

Accordingly, we run four checks. The first three use different functional forms for the testing variables: 1) $\ln(\text{new tests})$ – i.e. the log of the difference rather than the difference in the logs, 2) new tests (not logged) per 1,000 residents, and 3) cumulative tests run to date. The fourth check simply excludes testing variables completely. The results, presented below alongside the corresponding estimates from the baseline model, show that testing was indeed positively associated with cases, but our coefficient estimates of interest were remarkably stable across the different specifications. While none of the testing measures are perfect, this stability increased our confidence that differential trends in access to testing across counties did not drive our conclusions.

Appendix Exhibit 5 – Robustness Checks Related to Different Ways of Controlling for Testing

	Baseline	ln(New Tests)	New Tests per 1,000	ln(Cumulative Tests)	No Control for Tests
<u><i>Shelter-in-Place Order</i></u>					
1-5 Days Ago	-1.514	-1.643	-1.641	-1.639	-1.633
6-10 Days Ago	-3.033*	-3.229*	-3.098*	-3.165*	-3.131*
11-15 Days Ago	-4.482*	-4.671*	-4.530*	-4.658*	-4.630*
16-20 Days Ago	-5.950*	-6.155*	-5.977*	-6.132*	-6.104*
21 or More Days Ago	-8.600*	-8.905*	-8.564*	-8.821*	-8.781*
<u><i>Restaurant/Gym/Entertainment Closures</i></u>					
1-5 Days Ago	-4.372*	-4.415*	-4.349*	-4.500*	-4.335*
6-10 Days Ago	-4.710*	-4.805*	-4.751*	-4.878*	-4.713*
11-15 Days Ago	-6.125*	-6.151*	-6.088*	-6.293*	-6.031*
16-20 Days Ago	-5.594*	-5.595*	-5.588*	-5.861*	-5.539*
21 or More Days Ago	-5.177*	-5.149*	-5.156*	-5.500*	-5.135*
<u><i>No Large Events</i></u>					
1-5 Days Ago	0.172	0.264	0.258	0.260	0.227
6-10 Days Ago	-0.061	0.033	0.130	-0.010	0.125
11-15 Days Ago	0.013	0.055	0.144	-0.053	0.109
16 or More Days Ago	-1.041	-1.039	-0.888	-1.140	-0.933
21 or More Days Ago	-1.272	-1.269	-1.127	-1.404	-1.164
<u><i>Public Schools Closed</i></u>					
1-5 Days Ago	0.304	0.170	0.002	0.298	0.109
6-10 Days Ago	1.004	0.928	0.683	1.149	0.864
11-15 Days Ago	2.443	2.406	2.157	2.751	2.432
16 or More Days Ago	3.465	3.475	3.179	3.868	3.505
21 or More Days Ago	3.095	3.173	2.812	3.591	3.232
Testing Variable for Day t	0.010 ⁺	0.291*	0.005*	1.235*	--
Testing Variable for Day t-1	0.012*	-0.064	0.005*	-0.676	--

See notes for Appendix Exhibit 4.

Timing

Our next series of robustness checks related to the subjective nature of the start date of the sample. For our baseline results, we began the sample period on March 1. We excluded prior days because there were no new reported cases in the U.S. on most days in January in February. We did not start later in March because using the entire month allowed for ten days prior to the implementation of any policy with which we could examine pre-trends. However, arguments could be made for either earlier or later start dates. On one hand, an earlier start date would utilize more data and allow for a longer pre-treatment period. On the other hand, the earlier the start date, the more counties started with sample with no cases, exacerbating any potential issues created by adding one to the log of cases to prevent the outcome from being undefined.

Our first two checks simply started the sample a half month earlier and later: Feb. 15 and March 15, respectively. The third check used all available data, meaning a Jan. 22 start date. The fourth check used this longer period but avoided adding one to the log of cases by running two separate regressions: 1) one with the full sample but an indicator for any cases as the outcome (linear probability model due to the bias inherent in nonlinear fixed effects models), and 2) one with exponential growth rate as the outcome, but without adding one to the log of cases when creating the variable, thereby restricting the sample to counties with cases on both the current and previous days.¹¹ Next, we return to the original March 1 start date and avoid adding one to the log of cases a different way: by using the inverse hyperbolic sine transformation, which is interpretable in the same manner as natural logs but defined at zero.¹²

The table below displays the results. Across all start dates, we continued to find that SIPOs and restaurant/entertainment center closures – but not the other two policies – reduced the growth rate of cases. Interestingly, the results from the two-part specification suggested that SIPOs reduced COVID-19 spread along both the extensive margin (delaying the arrival of the first case in the county) and the intensive margin (slower spread once the first case was recorded), whereas restaurant/entertainment center closures did so only along the intensive margin. The results from the regression using the inverse hyperbolic sine specification showed this more directly, as those results were virtually identical to those from the baseline model.

¹¹ Note that this is not exactly equivalent to the conventional two-part model, which would use a dichotomized version of the continuous outcome as the dependent variable in the first part. In our setting, that would mean using the original outcome variable (with one added to the log of cases), and conditioning the sample for the second part on having a non-zero growth rate, meaning counties with cases but no growth in the previous day would also be dropped. Given these distinctions, it is not obvious to us how one might compute combined marginal effects that encompass both parts of the model, so we do not do so.

¹² For further detail, see John B. Burbidge, Lonnie Magee and A. Leslie Robb. “Alternative Transformations to Handle Extreme Values of the Dependent Variable.” *Journal of the American Statistical Association* Vol. 83, No. 401 (Mar., 1988), pp. 123-127.

Appendix Exhibit 6 – Robustness Checks Related to Sample Start Dates

	Start on 2/15	Start on 3/15	Start on 1/22	Start on 1/22; Any Cases as Outcome	Start on 1/22; Counties with Cases Only	Start on 3/1; Inverse Hyperbolic Sine
<u><i>Shelter-in-Place Order</i></u>						
1-5 Days Ago	-1.144	-0.783	-0.994	-0.024*	-1.045	-1.844
6-10 Days Ago	-2.332*	-1.736*	-2.028*	-0.059*	-2.117*	-3.456*
11-15 Days Ago	-3.412*	-2.652*	-2.929*	-0.090*	-2.898*	-4.794*
16-20 Days Ago	-4.517*	-3.626*	-3.847*	-0.112*	-3.666*	-6.052*
21 or More Days Ago	-6.541*	-5.480*	-5.536*	-0.143*	-5.239*	-8.727*
<u><i>Restaurant/Gym/Entertainment Closures</i></u>						
1-5 Days Ago	-4.060*	-5.779*	-3.936 ⁺	0.011	-6.778*	-4.821*
6-10 Days Ago	-4.164 ⁺	-6.406*	-3.952 ⁺	0.002	-6.242*	-5.123*
11-15 Days Ago	-5.372*	-8.016*	-5.081*	0.010	-8.085*	-6.556*
16-20 Days Ago	-4.648 ⁺	-7.825*	-4.302 ⁺	0.027	-7.694*	-5.997*
21 or More Days Ago	-4.067 ⁺	-7.849*	-3.683	0.046	-7.549*	-5.556*
<u><i>No Large Events</i></u>						
1-5 Days Ago	0.176	0.528	0.213	0.028	-1.713	0.007
6-10 Days Ago	-0.082	0.185	-0.011	0.036	-2.107	-0.509
11-15 Days Ago	-0.012	0.553	0.089	0.035	-2.300	-0.302
16-20 Days Ago	-1.077	-0.317	-0.941	0.033	-3.802	-1.414
21 or More Days Ago	-1.349	-0.300	-1.177	0.030	-4.260	-1.568
<u><i>Public Schools Closed</i></u>						
1-5 Days Ago	-0.070	0.260	-0.188	-0.062	1.909	1.033
6-10 Days Ago	0.231	0.334	-0.0004	-0.085	2.656	2.153
11-15 Days Ago	1.335	1.137	1.025	-0.095	3.390	3.956
16-20 Days Ago	2.090	1.786	1.732	-0.107	3.851	5.160
21 or More Days Ago	1.495	1.038	1.116	-0.115	2.774	5.032
Sample Size	229,074	138,072	298,110	298,110	93,245	182,004

See notes for Appendix Exhibit 4.

Dropping States

Our next group of checks asked whether results were robust to dropping particular, unique locations: New York, the hardest-hit state; Washington, the first state to be hit; and California, which had the nation's largest population and earliest shelter-in-place order. Appendix Exhibit 7 contains the results, which show that dropping any of these states has little impact on the estimates.

Appendix Exhibit 7 – Robustness Checks Dropping States

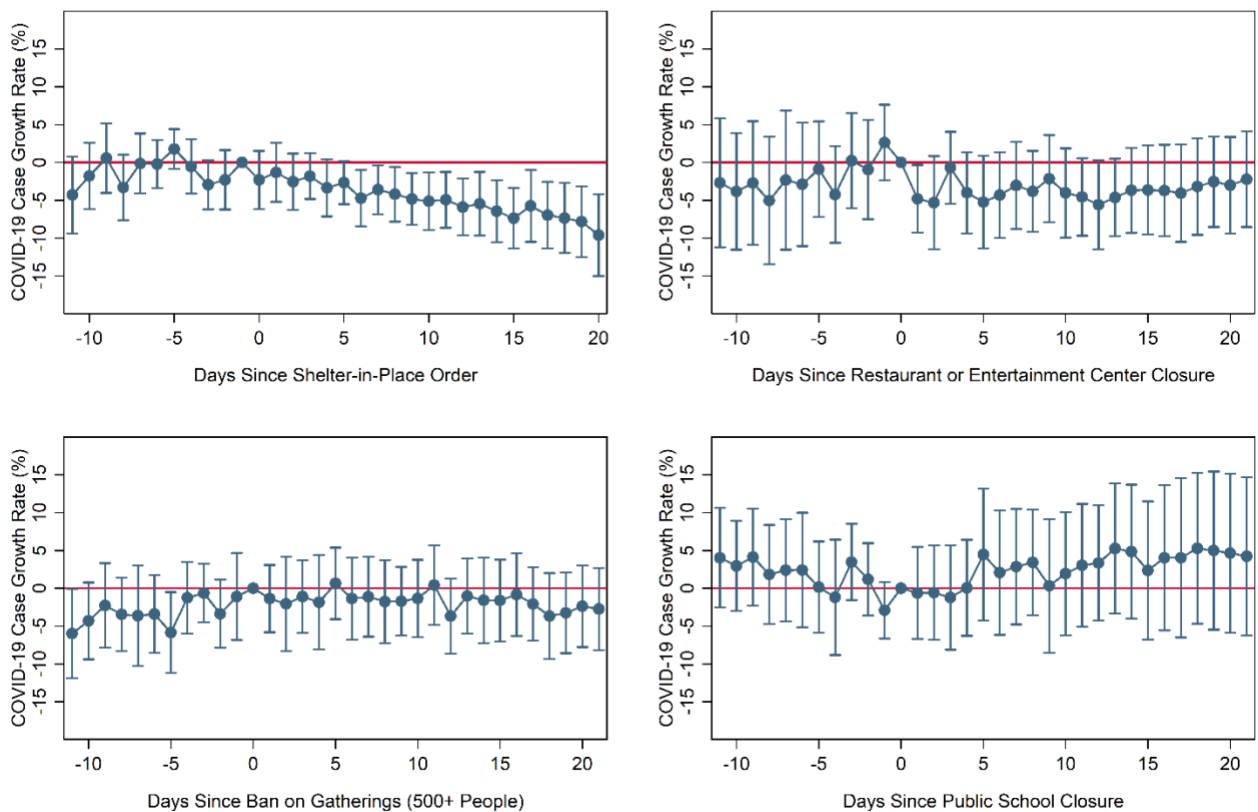
	Drop NY	Drop WA	Drop CA
<u><i>Shelter-in-Place Order</i></u>			
1-5 Days Ago	-1.320	-1.564	-1.475
6-10 Days Ago	-2.853*	-2.987*	-2.770*
11-15 Days Ago	-4.248*	-4.616*	-4.531*
16-20 Days Ago	-5.681*	-6.069*	-5.940*
21 or More Days Ago	-8.336*	-8.776*	-8.523*
<u><i>Restaurant/Gym/Entertainment Closures</i></u>			
1-5 Days Ago	-4.691*	-4.991*	-5.068*
6-10 Days Ago	-5.189*	-5.293*	-5.220*
11-15 Days Ago	-6.645*	-6.621*	-6.430*
16-20 Days Ago	-6.313*	-6.271*	-6.111*
21 or More Days Ago	-5.947*	-5.875*	-5.603*
<u><i>No Large Events</i></u>			
1-5 Days Ago	-0.051	0.483	0.609
6-10 Days Ago	-0.414	0.723	0.719
11-15 Days Ago	0.354	0.850	0.700
16 or More Days Ago	-0.388	-0.004	-0.239
21 or More Days Ago	-0.230	-0.050	-0.356
<u><i>Public Schools Closed</i></u>			
1-5 Days Ago	0.923	0.644	0.521
6-10 Days Ago	1.285	2.084	2.368
11-15 Days Ago	2.500	3.614	3.445
16 or More Days Ago	3.442	4.851	4.697
21 or More Days Ago	2.808	4.665	4.398
Sample Size	178,640	179,742	178,640

See notes for Appendix Exhibit 4.

Event-Study Model with One-Day Bins

Next, we estimate a more fine-grained event study model that included indicators for each day from 10 days before a policy took effect to 20 days after, rather than grouping them into 5-day increments. Appendix Exhibit 8 below displays the results. Not surprisingly, the estimates were less precise than those obtained using 5-day groupings, and the wider confidence intervals generally made the results less informative. In particular, we no longer observed significant effects for restaurant/entertainment center closures, despite most of the point estimates remaining roughly similar to those obtained using the 5-day intervals. The confidence intervals included reductions as large as 8 or more percentage points, making the lack of statistical significance uninformative. In contrast, the confidence intervals for SIPOs remain smaller, and most of those estimates remain statistically significant. One useful result from this check is that the apparent upward pre-treatment trend for gathering bans was concentrated among the first half of the pre-treatment window and largely disappeared four days before treatment. This increased our confidence that the failure to find a causal effect of these bans was not due to confounding from pre-trends.

Appendix Exhibit 8 – Results from Event-Study Model Including Day Indicators



Other Miscellaneous Checks

Next, two robustness checks considered alternate constructions of two of our policy variables. The first of these used bans on 50 or more, rather than 500 or more, people for the “event” variables. The second used closures of restaurants and gyms/entertainment centers, rather than one or the other, modeled by including both measures separately and then summing their coefficients. (Since the two measures were so strongly correlated, parsing out their separate effects with meaningful precision is not possible.)

We then explored sensitivity to dropping the observations with questionable case numbers, rather than imputing them. As described above, these included the entire day of Sunday, March 15, 2020 as well as all instances of negative reported case growth.

Next, recall that one of the motivations for interacting the day fixed effects with Census Division indicators was to account for different parts of the country starting their outbreaks at different times. As an alternative way to account for this, we replaced these day-by-Census Division fixed effects with fixed effects for days since (or until) the first case in the state. Since there was a case in every state prior to any distancing restrictions being implemented, the timing of this first case was not endogenous to our variables of interest.

Finally, we returned to the model with Census Division-by-day fixed effects but controlled for extrapolated county-specific pre-treatment trends in case growth rates. These were computed by restricting the sample to March 1-11, 2020 before any policies were enacted, running county-specific linear regressions of our outcome variable on day, and using the coefficient estimate to project changes after March 11.

The table below shows that the results for these checks were similar to those from our baseline model.

Appendix Exhibit 9 – Miscellaneous Robustness Checks

	Bans on 50 Instead of 500	Restau- rants <u>and</u> Gyms, not <u>or</u>	Drop Ques- tionable Data	Fixed effects for days since first case	Control for Pre- Trend
<u><i>Shelter-in-Place Order</i></u>					
1-5 Days Ago	-1.733	-1.393	-1.402	-2.361*	-1.525
6-10 Days Ago	-3.189*	-2.800*	-2.912*	-4.176*	-3.038*
11-15 Days Ago	-4.732*	-4.302*	-4.333*	-5.867*	-4.461*
16-20 Days Ago	-6.262*	-5.780*	-5.775*	-7.47*	-5.888*
21 or More Days Ago	-8.871*	-8.351*	-8.383*	-10.414*	-8.438*
<u><i>Restaurant/Gym/Entertainment Closures</i></u>					
1-5 Days Ago	-4.581*	-3.797*	-4.144*	-3.218 ⁺	-4.407*
6-10 Days Ago	-4.905*	-3.068	-4.427 ⁺	-5.351*	-4.762*
11-15 Days Ago	-6.612*	-4.642*	-5.740*	-6.575*	-6.186*
16-20 Days Ago	-6.330*	-3.988	-5.159 ⁺	-7.109*	-5.656*
21 or More Days Ago	-5.970*	-3.357	-4.666 ⁺	-7.092*	-5.221*
<u><i>No Large Events</i></u>					
1-5 Days Ago	1.971	-0.046	0.188	0.208	0.172
6-10 Days Ago	1.333	-0.787	-0.073	0.488	-0.055
11-15 Days Ago	1.451	-0.684	-0.008	0.643	0.030
16 or More Days Ago	1.085	-1.602	-1.064	-0.358	-1.012
21 or More Days Ago	0.280	-1.800	-1.316	-0.191	-1.218
<u><i>Public Schools Closed</i></u>					
1-5 Days Ago	0.649	0.080	0.315	3.207	0.356
6-10 Days Ago	1.921	0.736	0.933	1.717	1.092
11-15 Days Ago	3.521	2.151	2.367	1.396	2.544
16 or More Days Ago	4.882	3.244	3.335	0.641	3.567
21 or More Days Ago	4.798	2.881	2.933	-0.030	3.176
Sample Size	182,004	182,004	182,004	182,004	182,004

See notes for Appendix Exhibit 4.

E. Creation of Counterfactuals in Exhibits 4 and 5

In Exhibit 4, we plot the daily growth rate of COVID-19 cases in the U.S. under three scenarios. The first is actual case growth rates, which reflect the presence of all four social distancing measures when and where they were actually implemented. The second is counterfactual growth rates when we assume no jurisdictions ever implemented a SIPO but did implement the other measures the same way they did in reality. Formally, this is done by subtracting out the effect of SIPOs from the actual growth rate so that

$$counterfactual1_{cd} = \ln(cases_{cd}) - \ln(cases_{c,d-1}) - \hat{\gamma}_1 \sum_{i=1}^5 SIPO_{cdi}.$$

The third is the counterfactual when we assume no jurisdictions ever implemented any social distancing restriction:

$$\begin{aligned} counterfactual2_{cd} &= \ln(cases_{cd}) - \ln(cases_{c,d-1}) - \hat{\gamma}_1 \sum_{i=1}^5 SIPO_{cdi} - \hat{\gamma}_2 \sum_{i=1}^5 schools_{cdi} \\ &\quad - \hat{\gamma}_3 \sum_{i=1}^5 events_{cdi} - \hat{\gamma}_4 \sum_{i=1}^5 entertainment_{cdi} \end{aligned}$$

Note that we do *not* also subtract out the “placebo” effects of the future policy implementation variables, since those are intended to capture unobserved confounders rather than part of the causal effect of the policies. After obtaining the counterfactuals, we aggregate all three growth rate variables to the nation-by-day level by weighting each county-by-day observation by county population.

Exhibit 5 builds on Exhibit 4 to predict the number of cases under each counterfactual scenario. Note that $\ln(cases_{cd}) - \ln(cases_{c,d-1}) = \ln\left(\frac{cases_{cd}}{cases_{c,d-1}}\right)$ implies that

$$cases_{cd} = cases_{c,d-1} \exp\left[\ln\left(\frac{cases_{cd}}{cases_{c,d-1}}\right)\right].$$

With a predicted growth rate, $\ln\left(\frac{\widehat{cases}_{cd}}{cases_{c,d-1}}\right)$, this becomes

$$\widehat{cases}_{cd} = \widehat{cases}_{c,d-1} \exp\left[\ln\left(\frac{\widehat{cases}_{cd}}{cases_{c,d-1}}\right)\right] \exp(\widehat{\varepsilon}_{cd}), \quad (A1)$$

where $\exp(\widehat{\varepsilon}_{cd})$ is the average of $\exp(\widehat{\varepsilon}_{cd})$ in the estimation sample.¹³

Therefore, predicting number of cases under each scenario requires the estimated residual from our regression and predicted values of $\ln\left(\frac{cases_{cd}}{cases_{c,d-1}}\right)$ under the counterfactuals discussed above. Each of these estimates is then plugged into equation (A1) to predict counterfactual cases in county/day observations following the start date of a relevant policy. Prior to a policy beginning in a given county, the “counterfactual” cases are no different from the observed cases. Once the number of cases are predicted in each period under a given counterfactual, we sum the observed and predicted cases by day to create nation-by-day totals for observed and counterfactual cases. When these sums are presented on a logarithmic scale, as they are in Exhibit 5, the natural log must be calculated again after summing each variable.

This calculation is subject to several technical caveats that together mean the results should be interpreted as rough approximations rather than exact numbers. There is always uncertainty with calculations based on underlying parameters that are estimated, and it is difficult to accurately quantify the amount of this uncertainty. In this particular calculation, errors can compound over time, as \widehat{cases}_{cd-1} in equation (A1) is itself an estimate. Even when \widehat{cases}_{cd} is predicted from observed policies and using observed cases in $d - 1$, the predicted value is larger than $cases_{cd}$ on average in our sample. Additionally, the average difference is always positive later in the sample. To err on the side of caution and avoid having these errors accumulate over time, we discount our counterfactual predictions based on the average ratio of observed to predicted cases whenever the predicted cases (based on observed policies) are greater than observed cases on average.

We also err on the side of caution in one other way. When calculating the predicted case counts for each counterfactual, we set the coefficient on the variables indicating a policy start date 21 or more days in the past ($i = 5$ in the summation above) equal to the corresponding coefficient for 16-20 days. As discussed above and in the manuscript, our sample period includes up to 20 days after each policy. This means there is geographic balance for the variables reflecting up to 20 days, enabling apples-to-apples comparisons. However, the comparability of the estimates for the 21-or-more-day indicators is less clear, as earlier SIPOs remain in that category longer than more recent SIPOs. It is therefore not clear that the additional effect of SIPOs observed after 21+ days relative to 16-20 days is a “true” increase as opposed to simply the artifact of the compositional changes. We therefore make the conservative assumption that the true effect peaked and remained constant after 16 days. Without this adjustment, the estimated cases with no social distancing policies rises from ~35 million to ~50 million.

¹³ Don M. Miller (1984) Reducing Transformation Bias in Curve Fitting, The American Statistician, 38:2, 124-126, DOI: [10.1080/00031305.1984.10483180](https://doi.org/10.1080/00031305.1984.10483180)

An assumption that pushes in the direction of overstatement is that the pool of potentially infectable individuals is fixed over time. This ignores immunity from fighting off the virus. By the end of the sample period, with nearly a tenth of the US population predicted to have been sick enough to reach the threshold for a confirmed case, this assumption may not be benign. However, debate continues over the extent to which exposure to the virus conveys immunity, so the exact implications of this assumption are not clear.

By Charles Courtemanche, Joseph Garuccio, Anh Le, Joshua Pinkston, and Aaron Yelowitz

Strong Social Distancing Measures In The United States Reduced The COVID-19 Growth Rate

DOI: 10.1377/hlthaff.2020.00608
HEALTH AFFAIRS 39, NO. 7 (2020): –
©2020 Project HOPE—The People-to-People Health Foundation, Inc.

ABSTRACT State and local governments imposed social distancing measures in March and April of 2020 to contain the spread of novel coronavirus disease 2019 (COVID-19). These included large event bans, school closures, closures of entertainment venues, gyms, bars, and restaurant dining areas, and shelter-in-place orders (SIPOs). We evaluated the impact of these measures on the growth rate of confirmed COVID-19 cases across US counties between March 1, 2020 and April 27, 2020. An event-study design allowed each policy's impact on COVID-19 case growth to evolve over time. Adoption of government-imposed social distancing measures reduced the daily growth rate by 5.4 percentage points after 1–5 days, 6.8 after 6–10 days, 8.2 after 11–15 days, and 9.1 after 16–20 days. Holding the amount of voluntary social distancing constant, these results imply 10 times greater spread by April 27 without SIPOs (10 million cases) and more than 35 times greater spread without any of the four measures (35 million). Our paper illustrates the potential danger of exponential spread in the absence of interventions, providing relevant information to strategies for restarting economic activity. [Editor's Note: This Fast Track Ahead Of Print article is the accepted version of the peer-reviewed manuscript. The final edited version will appear in an upcoming issue of Health Affairs.]

Charles Courtemanche (courtemanche@uky.edu) is an associate professor of economics at the University of Kentucky, in Lexington, Kentucky.

Joseph Garuccio is a PhD student in economics at Georgia State University, in Atlanta, Georgia.

Anh Le is a PhD student in economics at the University of Kentucky.

Joshua Pinkston is an associate professor of economics at the University of Louisville, in Louisville, Kentucky.

Aaron Yelowitz is a professor of economics at the University of Kentucky.

A critical question during the COVID-19 pandemic is the effectiveness of the social distancing policies adopted by US states and localities in bending the curve. Although these policies take a variety of forms—such as imposing shelter-in-place orders (SIPOs); restricting dine-in at restaurants; closing other non-essential business such as bars, entertainment venues, and gyms; banning large social gatherings; and closing public schools—their effectiveness depends critically on the cooperation of the public. For example, although California's first-in-the-nation SIPO carries threats of fines and incarceration, its effectiveness fundamentally relies on social pressure.¹ Compliance with so-

cial distancing orders appears to be related to local income, partisanship, and political beliefs in the US; and compliance with self-quarantines is related to potential losses in income in Israel.^{2–4}

Some epidemiological models forecast the eventual number of COVID-19 cases and fatalities based on untested assumptions about the impact of social distancing policies in contemporary society. The widely cited Imperial College model assumes contact outside the home, school or workplace declines by 75 percent, school contact rates are unchanged, workplace contact rates fall by 25 percent, and household contact rates rise by 25 percent.⁵ Another study assumes social distancing measures will reduce the aver-

age contact rate by 38 percent, based on evidence from the 1918 influenza pandemic.⁶

In the US the literature on models of social distancing during the COVID-19 pandemic is evolving rapidly, and at the time of our writing, we were aware of several working papers that examined the consequences of social distancing policies. Recent work found significant effects of stronger measures (like SIPOs) on movement using difference-in-differences methods and state-level data from Google.⁷ Similar findings have been obtained in a study with SafeGraph mobility data,⁸ although a different study using PlaceIQ and SafeGraph data found strong measures were not important.⁹ Another paper used synthetic control methods to show that California's SIPO significantly reduced COVID-19 cases.¹ A study of SIPOs across the U.S. also found a reduction in cases, as well as higher rates of staying home full-time.¹⁰ Other authors used interrupted time-series methods and found that early statewide social distancing measures were associated with decreases in states' COVID-19 growth rates, but later SIPOs did not lead to further reductions.¹¹

At issue is not whether isolation works to limit the spread of disease, but rather whether the particular government restrictions designed to encourage social distancing in the US reduced spread relative to simply providing information and recommendations. Individuals may voluntarily engage in avoidance behavior, such as hand washing or wearing masks, once they fully perceive the risks of contagion.^{12,13} Critics of more stringent government measures highlight Sweden's less intrusive response to COVID-19, although Sweden's strategy is increasingly questioned.¹⁴ Rigorous empirical research is needed to determine the impacts of the various aspects of state and local governments' responses in the U.S.

Our work—which leveraged both state and county policy variation and used a flexible event-study method that allowed for effects to vary across measures and over time—estimated the impacts of four types of social distancing measures on confirmed COVID-19 case growth rates through April 27, 2020. The reduced-form approach captures any potential pathways driven by these mandates, including complementary avoidance behaviors that the public may engage in if these orders provide an informational shock in addition to increasing social distancing.

Study Data And Methods

STUDY DATA The unit of observation was daily US county/county equivalents. Although there are 3,142 counties in the US, official COVID-19 rec-

ords report New York City as a whole rather than dividing it into five counties, reducing this number to 3,138. Our dataset tracked counties over 58 days from March 1, 2020 to April 27, 2020, leading to a sample size of 182,004. We chose March 1 as the start date because no new cases were reported in the entire U.S. on most days in January and February. The April 27 end date was chosen to coincide with the first removal of one of four types of restrictions we analyzed (the re-opening of restaurants and other entertainment facilities in Georgia).¹⁵ Each county observation was weighted by population using 2018 estimates from the United States Department of Agriculture's Economic Research Service.¹⁶

OUTCOME OF INTEREST We examined the daily growth rate in confirmed COVID-19 cases at the county level, which originated from the 2019 Novel Coronavirus COVID-19 Data Repository provided by the Johns Hopkins Center for Systems Science and Engineering. This repository collected data on COVID-19 cases worldwide from a range of sources including government and independent health institutions.¹⁷

The daily exponential growth rate was calculated as the natural log of cumulative daily COVID-19 cases minus the log of cumulative daily COVID-19 cases on the prior day. We chose this functional form because epidemiological models predict exponential growth in the absence of intervention. Percent growth in cases is identical to percent growth in cases per capita since reported county populations did not vary during the sample period. The growth rate was multiplied by 100 and can be read as percentage point changes. In computing the growth rate, we followed a recent COVID-19 study and added one to the case counts to avoid dropping counties that started with zero cases.¹⁸

COVARIATES The data on the timing of state and local government social distancing interventions was gathered from a host of sources and made available by Johns Hopkins University.¹⁹ Part A of the online appendix explains a few corrections we made to the dates and provides a list of state- and county-level policies used in the analysis.²⁰

We focused on four government-imposed interventions: SIPOs, public school closures, bans on large social gatherings, and closures of entertainment-related businesses. For large gatherings, we used the date of the first prohibition that was at least as restrictive as 500 people. Most of the bans were much more restrictive: 95 percent of the time (in our population-weighted sample) the prohibition extended to 50 people. For entertainment-related businesses, we used the date of the first closure of either restaurant dining areas (including bars) or gyms/

entertainment centers. 96 percent of the time, if one such prohibition was in place, the other was in place as well.

We included control variables related to the availability of COVID-19 tests. The same Data Repository that provides cases also includes daily counts of positive, negative, and pending tests in each state on each day, which we added together.¹⁷ To mirror our measure of cases, we converted this testing variable to the exponential daily growth rate of cumulative tests performed. Since COVID-19 test results are generally not available immediately, we also included the one-day lag of this growth rate. Further lags (out to 10 days) were considered but always statistically insignificant, so we did not include them. Most states did not report any pending tests, meaning that they did not officially record tests until the results were obtained. This likely explains the lack of a longer lag between testing growth and case growth.

METHODS We estimated the relationship between social distancing policies and the exponential growth rate of confirmed COVID-19 cases using an event-study regression with multiple treatments. Statistical analysis was conducted using Stata MP (version 15). This approach is akin to difference-in-differences but more flexible, as it interacts the treatment variables with multiple indicators of time since implementation, thereby tracing out the evolution of the treatment effects over time.²¹

For each of the four policies, we include seven variables: whether it was implemented 1–5, 6–10, 11–15, 16–20, or more than 20 days ago; and whether it will be implemented 5–9 or 10 or more days later. Implementation on the current day through four days from now was, therefore, the reference group. If a county never adopted the policy, each of these variables was set to 0 throughout the sample period.

An event study model is particularly useful to study the impact of social distancing policies on COVID-19 cases for two reasons. First, after accounting for the incubation period and time between onset of first symptoms and positive test result, such policies likely only affect official cases after a considerable lag.²² Additionally, the inclusion of variables reflecting future implementation allows for an analysis of pretreatment trends. Since it is not plausible for policies that have not yet been implemented to causally affect current cases, finding such associations could suggest misspecification. For instance, one might expect counties with rapidly growing case counts to be the most likely to enact these measures, leading to a reverse-causal relationship between current cases and future policies that would be detected by our model.

Each policy was implemented at least 10 days after the start of the sample period and at least 20 days prior to the end. Therefore, each policy contributes to the identifying variation for all coefficients except those for more than 20 days ago and 10 or more days from now. Since the estimated policy effects at those two “catch-all” time periods could partially reflect compositional changes, they should therefore be interpreted with more caution than the estimates for the other time intervals.

In addition to the testing controls discussed above, the model also included fixed effects for geography and time. County fixed effects accounted for the likelihood that, even aside from differences in policies, case growth rates may have varied due to a number of county characteristics. These characteristics include population density and residents’ education, political orientation, and age.^{3,4} Fixed effects for each day in each of the nine U.S. Census Divisions (522 fixed effects in total) allowed for flexible underlying trends in growth rates that could vary in different parts of the country, helping to account for the staggered nature of the outbreak across locations.²³ We report 95% confidence intervals, with standard errors robust to heteroskedasticity and clustered by state, the level of most of the policy variation. Part B of the appendix provides the formal notation for the event-study model.²⁰

LIMITATIONS There are several limitations to our analysis. Official COVID-19 case counts are known to understate the true prevalence of the disease, as they do not include asymptomatic carriers, those who are not ill enough to seek medical care, and those who are unable to obtain a test due to supply constraints.¹ Nonetheless, confirmed case counts are crucial to the Trump administration’s “Opening Up America Again” plan, which proposes either a “downward trajectory of documented cases within a 14-day period” or “downward trajectory of positive tests as a percent of total tests within a 14-day period (flat or increasing volume of tests)” as criteria to loosening social distancing measures.²⁴ Moreover, to the extent that testing shortages led to only the sickest individuals receiving them, official case counts can loosely be interpreted as the prevalence of moderate-to-severe illnesses, a relevant metric for policy purposes.

A related caveat is that, ideally, we would like to be able to control more precisely for access to testing. Available data only allowed us to control for number of tests performed at the state, rather than county, level. However, most of our policy variation is at the state level, so state-level testing should go a long way towards alleviating bias. Additionally, number of tests performed is not an ideal measure of the ease of obtaining a test

because it also reflects the level of illness in the community.

Also, we might ideally want to estimate a richer econometric model. It would be interesting to trace out the timing of impacts more exactly and study the policies' interactions with each other or county characteristics. Future work should also examine the impacts of other social distancing policies such as closing public parks and beaches, the requirement to wear masks in public, restrictions on visitors in nursing homes, state announcements of first cases or fatalities, and federal government actions such as prohibiting international travel.⁹ However, it is difficult to include numerous correlated treatment variables without reducing precision to the point where statistical inference is uninformative.

Finally, as is typical of observational data analyses, we cannot rule out all possible threats to causal inference. Numerous possible confounders could vary across time and space, including the other policies mentioned above, informal encouragement by government officials to wear masks or improve hygiene, changing business practices, and social norms regarding distancing. That said, including Census-Division-by-day and county fixed effects in our model and examining pretreatment trends helped us to push in the direction of causality.

Study Results

DESCRIPTIVE INFORMATION Confirmed COVID-19 cases grew rapidly during the sample period, from just 30 on March 1 to 978,047 on April 27. Part C of the appendix shows the number of counties with any COVID-19 cases on each day.²⁰ On March 1, the vast majority of counties had zero cases, and across all days, 49 percent of unweighted county-by-day observations were zero. However, counties with zero cases tended to have low populations, so our population weights limited the influence of these counties on the results.

Supplemental exhibit 1 illustrates the reach of social distancing policies on the US population over time.²⁰ The SIPO was generally the last policy to be implemented, and adoption was uniformly lower than the other policies. On March 1, no jurisdiction had implemented all four measures. By March 22, nearly 25 percent of the US population was covered by all the measures, growing to approximately 65 percent by March 29 and 95 percent by April 7, when the last SIPO took effect.

IMPACT OF SOCIAL DISTANCING POLICIES Supplemental exhibit 2 illustrates the coefficients (and confidence intervals) for SIPOs and bans on large gatherings derived from the event-study

model.²⁰ Relative to the reference category of 0–4 days before implementation, SIPOs lead to statistically significant ($p < 0.01$) reductions in the COVID-19 case growth rate of 3.0 percentage points after 6–10 days, 4.5 after 11–15 days, 5.9 after 16–20 days, and 8.6 from day 21 onward. Because the model held constant the other types of policies, these estimates should be interpreted as the additional effect of SIPOs *beyond* shutting down schools, large gatherings, and entertainment-related businesses. This additional effect may come from either the requirement/strong advisement to shelter-in-place aside from “essential” activities or the accompanying closure of any “non-essential” businesses that remained open. We did not observe any statistically significant “placebo” effects of SIPOs in the periods prior to implementation, giving credence to a causal interpretation of our main results. If anything, the pre-trend appears to point upward, which would make our estimates in the post-treatment period conservative.

We found no evidence that bans on large social gatherings influenced the growth rate. The point estimates for banning gatherings were statistically insignificant ($p > 0.56$ in all cases). However, the 95% confidence intervals included reductions of up to 3–6 percentage points, so the lack of evidence of an effect should not be misinterpreted as clear evidence of no effect. Also, the lack of a statistically significant reduction in the post-treatment period could potentially be due to an upward (though not statistically significant) pre-treatment trend. However, results from the aforementioned event study with separate variables for each day showed that the pre-trend disappeared four days prior to implementation.

Supplemental exhibit 3 shows estimates for the restaurant-and-entertainment-related businesses and school closings.²⁰ Closing restaurant dining rooms/bars and/or entertainment centers/gyms led to statistically significant reductions in the growth rate of COVID-19 cases in all time periods after implementation ($p < 0.05$). The estimated effect was 4.4 percentage points after 1–5 days, 4.7 after 6–10 days, 6.1 after 11–15 days, 5.6 after 16–20 days, and 5.2 after 21 or more. Prior to implementation, policies related to businesses showed no effect on the growth rate, again passing the “placebo” test.

In contrast, we found no evidence that school closures influenced the growth rate. The point estimates were never close to statistically significant ($p > 0.37$ in all cases), but the 95% confidence intervals meant that we could not rule out reductions of up to 4–5 percentage points.

Adding the coefficient estimates for each policy gives the combined effect of implementing all

four social distancing policies. In days 1–5 after implementation, the bundle of restrictions reduced the growth rate of COVID-19 cases by 5.4 percentage points. In days 6–10 after implementation, the growth rate fell by 6.8 percentage points. This reduction grew to 6.8 percentage points after 6–10 days, 8.2 percentage points after 11–15, 9.1 after 16–20, and 12.0 after 21 or more. As discussed previously, the estimate for 21+ days should be viewed with caution, as it did not utilize the same geographic balance of treatments as the estimates for the other time intervals. A conservative interpretation of these results would therefore be that the impact reached 9.1 percentage points after 16–20 days and appeared to remain at least as high after that.

ROBUSTNESS CHECKS Part D of the appendix presents and discusses results from a number of robustness checks designed to address possible concerns with our model.²⁰ These checks begin with regressions with just one variable per policy to help rule out the null results for gathering bans and school closures being due to multicollinearity (appendix exhibit 4).²⁰ We then evaluated robustness to using different functional forms for the testing controls or omitting them (appendix exhibit 5).²⁰ Next, appendix exhibit 6 varied the sample start date and the approach used to dealing with counties with no cases.²⁰ Appendix exhibit 7 shows results from dropping the uniquely affected states of NY, WA, and CA.²⁰ Appendix exhibit 8 displays results from a more fine-grained event-study model with separate variables for each day from 10 days before treatment to 20 days after.²⁰ Finally, appendix exhibit 9 presents results from other checks related to data and measurement issues as well as controlling for county-specific pre-treatment trends.²⁰ The general pattern of results was robust to these different specifications.

COUNTERFACTUAL SIMULATIONS Supplemental exhibit 4²⁰ uses the results from the baseline model to compare the observed growth rate of COVID-19 cases to two counterfactuals: 1) none of the four social distancing measures ever being imposed and 2) no SIPO ever being imposed. The process for creating these counterfactuals is described in the appendix.²⁰ The mean exponential growth rate without any interventions was 16.2 percent over the full time period. The observed and both counterfactual growth rates peaked on March 19, 2020 at 26–28 percent but started to diverge afterward, eight days after the earliest restriction. Without any social distancing policies, the model predicts the case growth rate would have stayed similarly high for another week before gradually falling to 14 percent by April 27, 2020. Without SIPOs—but keeping the other restrictions—the growth

rate would have fallen to 11 percent. The actual growth rate, which reflects all implemented distancing policies including SIPOs, fell to 3 percent by that date.

Supplemental exhibit 5 compares the reported number of COVID-19 cases over time to the number of cases predicted by our event-study regression under these same two counterfactual scenarios.²⁰ Part E of the appendix describes the technical details of these simulations along with the required assumptions.²⁰ The graph uses the natural logarithm of nationwide cases (or predicted cases) for the y-axis scale, but with corresponding numbers labeled on the y-axis instead of logs.

In all three scenarios, cases increased roughly linearly on the log scale, as expected under exponential growth, until the last week of March—approximately two weeks after the first restrictions and one week after the first SIPO. The actual curve then began to flatten substantially, eventually leading to 978,047 cases by April 27. In contrast, the two counterfactual curves only flattened slightly. By the end of the sample period, the model predicts that cases would have been 10 times higher without SIPOs (10,224,598) and 35 times higher (35,257,098) without any social distancing restrictions. Interestingly, the closures of restaurants/entertainment facilities accounted a larger share of the reduction in cases than SIPOs, despite SIPOs having larger coefficient estimates. This is because restaurant/entertainment facilities were implemented earlier and in more places than SIPOs.

Discussion

Nuance is required when interpreting the results in supplemental exhibit 5.²⁰ We view the simulation as providing an illustration of the power of exponential growth and the effectiveness of social distancing restrictions at “flattening the curve,” even when their impacts are not immediately visible. As late as April 6, nearly a month after the earliest interventions, the number of cases would still have been under 1,000,000 even without any restrictions—just 2.4 times the actual number of cases. The explosion in cases without social distancing measures happens later, and by time it is happening, the lagged effects of these measures mean it is too late to stop it.

At the same time, we urge caution about taking the specific numbers of cases averted too literally. Simulations that use estimated parameters to predict outside the range of observed policy variation are inherently subject to a high level of uncertainty that is difficult to quantify. Moreover, had policymakers not taken action and

COVID-19 had continued to spread throughout April in the manner depicted by our simulations, voluntary social distancing by individuals and businesses would have likely increased as panic over the rising death toll and hospital overcrowding across the country mounted. In technical terms, the Census-Division-by-day fixed effects would have evolved differently than what we observed. This would have likely offset at least some of the additional predicted cases—though, because of the lag to impact, it is unclear how much of this offsetting could have occurred before the end of our sample period.

Relatedly, testing shortages would likely have prevented official case counts from reaching the numbers presented in our simulations. However, this is largely a semantic distinction, as these infections would still be severe enough to warrant testing in the absence of a shortage. If anything, not being confirmed as a COVID-19 case could lead to inadequate treatment.

As striking as our counterfactual estimates are, they still are not worst-case scenarios because they account for at least some voluntary social distancing. Even without any government restrictions, supplemental exhibit 4²⁰ illustrated a 14.3 percentage point drop from the peak growth rate to the end of the sample period. The most plausible explanation is responses of individuals and businesses to information about the severity of the pandemic and federal guidelines.

While our results suggest both SIPO and non-SIPO measures can be effective at averting COVID-19 cases, the lack of evidence of effects of school closures or bans on large social gatherings is noteworthy. We cannot rule out the possibility that these null results are due to statistical imprecision, but it is also possible that both policies may displace social interaction rather than reducing it. For example, school closures may have led families to continue social interactions outside of the school setting, such as at day care centers or parks. Google mobility data through April 5, 2020 show increases of 10 percent or more in visits to parks in 28 states.²⁵ A new study finds that schools are only slightly more dangerous than parks and playgrounds for COVID-19 transmission, supporting this explanation.²⁶ Alternatively, school closures primarily affect children and the vast majority of children experience mild symptoms and therefore may not be included in confirmed cases.²⁷ While asymptomatic children can pass the virus to adults who become more severely ill, our results imply that the extent to which this led to confirmed cases did not change when schools were closed.

Similarly, official group events may have sim-

ply been replaced by informal gatherings. Alternatively, official prohibitions may have been largely redundant since the largest events (such as college and professional sports) were already being cancelled due to CDC guidance or other information.

Also note that school closures and large event bans occurred prior to the implementation of SIPOs, meaning substitute types of social gatherings were still allowed. Our results, therefore, should not be interpreted as a forecast about what would happen if schools were reopened or certain large gatherings were allowed while other aspects of SIPOs remained in place.

Conclusion

We estimated the separate and combined impact of four widely adopted social distancing policies. Both SIPOs and closures of restaurants/bars/entertainment-related businesses substantially slowed the spread of COVID-19. We did not find evidence that bans on large events and closures of public schools also did, though the confidence intervals cannot rule out moderately sized effects. Interestingly, two recent papers on the effect of social distancing restrictions on mobility found the same pattern as we did in terms of which restrictions mattered and which ones did not, suggesting that null effects of gathering bans and school closures on case growth are at least plausible.^{7,8}

Our contribution was to provide credible empirical evidence on whether US social distancing measures worked as intended in flattening the curve. Estimating other important benefits and costs from social distancing, including the total lives saved and economic harm, was beyond the scope of our study. Other work has attempted to estimate job losses, simulate effects on the overall economy and economic growth, or estimate distributional consequences from current and past pandemics.^{1,6,28–31}

Nonetheless, we provide important information about benefits of social distancing for policymakers to consider as they decide on strategies for restarting economic activity. For instance, our results argue against returning to partial measures such as school closures and restrictions on large gatherings, while removing the restrictions that prevent the redirection of social activity to other settings. At issue moving forward is whether cases averted simply turn into cases delayed, and a premature return to light measures would make this more likely. At the same time, our results are not informative about the effectiveness of intermediate measures, such as lifting a SIPO but requiring masks in public or opening restaurants at reduced capacity. Further

research is needed as gradual, untested steps toward reopening are taken across the country. ■

Charles Courtemanche has received funding in the past year from the National Institutes of Health, United

States Department of Agriculture, and the American Beverage Institute.
[Published online May 14, 2020.]

NOTES

- 1 Friedson AI, McNichols D, Sabia JJ, Dave D. Did California's Shelter in Place Order Work? Early Evidence on Coronavirus-Related Health Benefits [Internet]. Cambridge (MA): National Bureau of Economic Research; 2020 Apr [cited 2020 May 11]. (NBER Working Paper No. 26992). Available from: <https://www.nber.org/papers/w26992>
- 2 Bodas M, Peleg K. Self-Isolation Compliance In The COVID-19 Era Influenced By Compensation: Findings From A Recent Survey In Israel. Health Aff (Millwood). 2020 Apr 9. DOI: 10.1377/hlthaff.2020.00382. [Epub ahead of print].
- 3 Painter M, Qiu T. Political Beliefs affect Compliance with COVID-19 Social Distancing Orders [Internet]. St. Louis (MO): Saint Louis University; 2020 Apr 30 [cited 2020 May 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3569098
- 4 Wright AL, Sonin K, Driscoll J, Wilson J. Poverty and Economic Dislocation Reduce Compliance with COVID-19 Shelter-in-Place Protocols [Internet]. Chicago (IL): University of Chicago, Becker Friedman Institute for Economics; 2020 Apr 29 [cited 2020 May 11]. (Working Paper No. 2020-40). Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3573637
- 5 Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand [Internet]. London: Imperial College London; 2020 Mar 16 [cited 2020 May 11]. Available at: <https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf>
- 6 Thunstrom L, Newbold S, Finnoff D, Ashworth M, Shogren JF. The Benefits and Costs of Using Social Distancing to Flatten the Curve for COVID-19 [Internet]. Laramie (WY): University of Wyoming; 2020 Mar 27 [cited 2020 May 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3561934
- 7 Abouk R, Heydari B. The Immediate Effect of COVID-19 Policies on Social Distancing Behavior in the United States [Internet]. Wayne (NJ): William Paterson University; 2020 Apr 8 [cited 2020 May 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3571421
- 8 Andersen M. Early Evidence on Social Distancing in Response to COVID-19 in the United States [Internet]. Greensboro (NC): University of North Carolina Greensboro; 2020 Apr 6 [cited 2020 May 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3569368
- 9 Gupta S, Nguyen TD, Rojas FL, Raman S, Lee B, Bento A, et al. Tracking Public and Private Responses to the COVID-19 Epidemic: Evidence from State and Local Government Actions [Internet]. Cambridge (MA): National Bureau of Economic Research; 2020 Apr [cited 2020 May 11]. (NBER Working Paper No. 27027). Available from: <https://www.nber.org/papers/w27027>
- 10 Dave DM, Friedson AI, Matsuzawa K, Sabia JJ. When Do Shelter-in-Place Orders Fight COVID-19 Best? Policy Heterogeneity Across States and Adoption Time [Internet]. Cambridge (MA): National Bureau of Economic Research; 2020 May [cited 2020 May 11]. (NBER Working Paper No. 27091). Available from: <https://www.nber.org/papers/w27091>
- 11 Siedner MJ, Harling G, Reynolds Z, Gilbert RF, Venkataramani A, Tsai AC. Social distancing to slow the U.S. COVID-19 epidemic: an interrupted time-series analysis [Internet]. Boston (MA): Massachusetts General Hospital; 2020 [cited 2020 May 11]. Available for download from: https://www.researchgate.net/publication/340518052_Social_distancing_to_slow_the_US_COVID-19_epidemic_an_interrupted_time-series_analysis
- 12 Abaluck J, Chevalier JA, Christakis NA, Forman HP, Kaplan EH, Ko A, et al. The Case for Universal Cloth Mask Adoption and Policies to Increase Supply of Medical Masks for Health Workers [Internet]. New Haven (CT): Yale University; 2020 Apr 1 [cited 2020 May 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3567438
- 13 Harris JE. The Coronavirus Epidemic Curve is Already Flattening in New York City [Internet]. Cambridge (MA): National Bureau of Economic Research; 2020 Apr [cited 2020 May 11]. (NBER Working Paper No. 26917). Available from: <https://www.nber.org/papers/w26917>
- 14 Ahlander J, O'Connor P. Sweden's liberal pandemic strategy questioned as Stockholm death toll mounts. Reuters [serial on the Internet]. 2020 Apr 3 [cited 2020 May 7]. Available from: <https://www.reuters.com/article/us-health-coronavirus-sweden/swedens-liberal-pandemic-strategy-questioned-as-stockholm-death-toll-mounts-idUSKBN21L23R>
- 15 The State of Georgia. Providing guidance for reviving a healthy Georgia in response to COVID-19 [Internet]. Atlanta (GA): Governor's Office; 2020 Apr 23 [cited 2020 May 11]. (Executive Order). Available for download from: <https://gov.georgia.gov/executive-action/executive-orders/2020-executive-orders>
- 16 US Department of Agriculture, Economic Research Service. Population estimates for the U.S., States, and counties, 2010–18 [Internet]. Washington (DC): USDA; [updated 2019; cited 2020 May 11]. Available from: <https://data.ers.usda.gov/reports.aspx?ID=17827>
- 17 Johns Hopkins University and Medicine, Coronavirus Resource Center. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) [Internet]. Baltimore (MD): Johns Hopkins University; 2020 [cited 2020 May 11]. Available from: <https://coronavirus.jhu.edu/map.html>
- 18 Bursztnyn L, Rao A, Roth C, Yanagizawa-Drott D. Misinformation During a Pandemic [Internet]. Chicago (IL): University of Chicago, Becker Friedman Institute for Economics; 2020 Apr [cited 2020 May 11]. (Working Paper No. 2020-44). Available from: https://bfi.uchicago.edu/wp-content/uploads/BFI_WP_202044.pdf [accessed 2020 7 May].
- 19 Killeen BD, Wu JY, Shah K, Zapaishchykova A, Nikutta P, Tamhane A, et al. A County-level Dataset for Informing the United States' Response to COVID-19 [Internet]. Baltimore (MD): Johns Hopkins University; 2020 [cited 2020 May 11]. Available from: <https://arxiv.org/pdf/2004.00756v1.pdf>
- 20 To access the appendix, click on the Details tab of the article online.
- 21 Saloner B, Maclean JC. Specialty

Editor's Note: The appendix referred to in note 20 is no longer online. It is now part of this document and follows the article.

- Substance Use Disorder Treatment Admissions Steadily Increased In The Four Years After Medicaid Expansion. *Health Aff (Millwood)*. 2020;39(3):453–61.
- 22 Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med*. 2020 Mar 10. [Epub ahead of print].
 - 23 US Census Bureau. Census Regions and Divisions of the United States [Internet]. Washington (DC): Census Bureau; [cited 2020 May 11]. Available from: https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf
 - 24 White House. Opening up America again. Washington (DC): White House; 2020.
 - 25 Google. COVID-19 Community Mobility Report [Internet]. [Updated 2020 Apr 5; cited 2020 May 11]. Available from: https://www.gstatic.com/covid19/mobility/2020-04-05_US_Mobility_Report_en.pdf
 - 26 Benzell S, Collis A, Nicolaides C. Rationing Social Contact During the COVID-19 Pandemic: Transmission Risk and Social Benefits of US Locations [Internet]. Cambridge (MA): Massachusetts Institute of Technology; 2020 Apr 18 [cited 2020 May 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3579678
 - 27 Editorial. Pandemic School Closures: Risks and Opportunities. *The Lancet Child & Adolescent Health*. 2020; 4(5):341.
 - 28 Scherbina AD. Determining the Optimal Duration of the COVID-19 Suppression Policy: A Cost-Benefit Analysis [Internet]. Boston (MA): Brandeis University; 2020 Mar 24 [cited 2020 May 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3562053
 - 29 Hall RE, Jones CI, Klenow PJ. Trading Off Consumption and COVID-19 Deaths [Internet]. Stanford (CA): Stanford University; 2020 Apr 24 [cited 2020 May 11]. (Working Paper). Available from: https://web.stanford.edu/~chadj/Consumption_v_Covid.pdf
 - 30 Greenstone M, Vishan N. Does Social Distancing Matter? [Internet]. Chicago (IL): University of Chicago, Becker Friedman Institute for Economics; 2020 Mar 30 [cited 2020 May 11]. (Working Paper No. 2020-26). Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3561244
 - 31 Correia S, Luck S, Verner E. Pandemics Depress the Economy, Public Health Interventions Do Not: Evidence from the 1918 Flu [Internet]. Washington (DC): Board of Governors of the Federal Reserve System; 2020 Mar 26 [cited 2020 May 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3561560

Appendix

Contents:

A. Adjustments for Likely Data Errors	2
Appendix Exhibit 1	4
Appendix Exhibit 2	6
B. Event Study Econometric Model	7
C. Diffusion of COVID-19 During Sample Period	8
Appendix Exhibit 3	9
D. Robustness Checks	10
Appendix Exhibit 4	10
Appendix Exhibit 5	12
Appendix Exhibit 6	14
Appendix Exhibit 7	15
Appendix Exhibit 8	16
Appendix Exhibit 9	18
E. Creation of Counterfactuals in Supplemental Exhibits 4 and 5	19
SUPPLEMENTAL EXHIBITS	
Supplemental Exhibit 1	22
Supplemental Exhibit 2	23
Supplemental Exhibit 3	24
Supplemental Exhibit 4	25
Supplemental Exhibit 5	26

A. Adjustments for Likely Data Errors

The original county-by-day cases data contained some likely reporting errors. First, the COVID-19 case counts on Sunday, March 15, 2020 were identical to those from March 14, 2020 in all large counties that had cases at the time, even though substantial growth was reported on March 16, 2020. This is almost certainly due to lack of reporting on that particular Sunday rather than a true lack of new cases. We therefore imputed the cases for March 15, 2020 so that growth from March 14, 2020 to March 16, 2020 was evenly split between the two days.

Additionally, 964 of the 182,004 county-by-day observations reported a reduction in cases from the previous day, which is implausible since case counts were cumulative. Approximately half of these were cases where the count increased by one on one day and then decreased by one the next. In such cases, the original increase was likely an error (e.g. presumed case that was later found to be negative, duplicate, or person who lived in a different county/state), so we removed the initial increase so that the case count remained constant over the three days. We set remaining instances of negative growth rates to zero. As shown later in this appendix, our results were virtually unchanged if we simply dropped observations with case reductions as well as all observations on March 15, 2020.

We also corrected some apparent errors in the social distancing policy dates. Many Kansas counties had reported dates for gathering bans that were after the date for the statewide ban, so we changed those dates to match the statewide ban.¹ We removed the reported SIPO in Wyoming, which was listed as March 28, 2020, as sources such as the Kaiser Family Foundation² as well as various news articles³ listed the state as not having any SIPO in place after the date listed. Some school closing dates also appeared incorrect. School closing dates were checked against those listed at edweek.org⁴. If there was more than a day mismatch, then a third source was found to decide the date used, leading to changes for Alabama⁵, Alaska⁶, Iowa⁷, and Wyoming⁸. We changed the SIPO start date for MS⁹ and gathering ban date for KY¹⁰ as they

¹ The counties were Allen, Anderson, Atchison, Barber, Barton, Bourbon, Brown, Butler, Chase, Chautauqua, Cherokee, Cheyenne, Clark, Clay, Cloud, Coffey, Comanche, Cowley, Crawford, Decatur, Dickinson, and Doniphan.

² See: <https://www.kff.org/health-costs/issue-brief/state-data-and-policy-actions-to-address-coronavirus/>

³ See: <https://www.usnews.com/news/best-states/wyoming/articles/2020-03-30/wyoming-governor-still-no-plan-for-state-stay-at-home-order>, <https://abcnews.go.com/Health/9-states-issue-formal-stay-home-orders-amid/story?id=69959039>

⁴ See: <https://www.edweek.org/ew/section/multimedia/map-coronavirus-and-school-closures.html>

⁵ See: <https://www.al.com/news/2020/03/alabama-closes-all-k-12-schools.html>

⁶ See: <https://www.adn.com/alaska-news/education/2020/03/16/alaskas-statewide-school-closure-is-about-to-begin-no-one-knows-quite-how-it-will-work/>

⁷ See: <https://www.usnews.com/news/best-states/iowa/articles/2020-03-16/iowa-schools-closing-taking-coronavirus-cue-from-governor>

⁸ See: https://trib.com/news/local/education/school-districts-across-wyoming-announce-closures-following-governors-recommendation/article_333c5100-c6a6-57f1-b04b-948546b59ffd.html

⁹ See: https://www.sos.ms.gov/Content/documents/ed_pubs/Exec%20Orders/1466.pdf

¹⁰ See: https://governor.ky.gov/attachments/20200319_Order_Mass-Gatherings.pdf

were incorrect in the source data. The SIPOs for IN, MN, NH, OH, and TN had an effective start time of 11:59 pm. We changed the start day for these states to the following day as they were in effect for only one minute on their actual start date. Below we list the state and county (if different from state) policy dates that we ultimately used.

Appendix Exhibit 1 – Effective Dates of State Social Distancing Measures

State Shelter in Place No Gatherings of 500+ Public Schools Closed Rest./ Ent./Gym Order

AL	4/4	3/13	3/16	3/19
AK	3/28	3/24	3/19	3/17
AZ	3/31	3/17	3/16	3/17
AR		3/26	3/17	3/19
CA	3/19	3/19	3/19	3/15
CO	3/26	3/13	3/23	3/17
CT	3/23	3/23	3/17	3/16
DE	3/24	3/24	3/16	3/12
DC	4/1	3/25	3/16	3/16
FL	4/3	3/17	3/16	3/20
GA	4/3	3/17	3/18	3/20
HI	3/25	3/25	3/23	3/20
ID	3/25	3/25	3/23	3/23
IL	3/21	3/18	3/17	3/18
IN	3/25	3/13	3/19	3/17
IA		3/17	4/2	3/17
KS	3/30	3/17	3/18	3/24
KY	3/26	3/20	3/16	3/16
LA	3/23	3/17	3/16	3/17
ME	4/2	3/18	3/15	3/18
MD	3/30	3/13	3/16	3/16
MA	3/24	3/17	3/17	3/17
MI	3/24	3/24	3/16	3/16
MN	3/28	3/27	3/18	3/17
MS	4/3	3/25	3/20	3/25
MO	4/6	3/21	3/19	3/21
MT	3/28	3/28	3/16	3/20
NE		3/19	3/19	4/4
NV	4/1	3/24	3/16	3/20
NH	3/28	3/24	3/16	3/16
NJ	3/21	3/16	3/18	3/16
NM	3/24	3/18	3/16	3/18
NY	3/22	3/13	3/18	3/16
NC	3/30	3/12	3/16	3/17
ND			3/16	3/19
OH	3/24	3/15	3/17	3/15
OK		3/29	3/17	4/1
OR	3/23	3/16	3/16	3/17
PA	4/1	3/23	3/16	3/17
RI	3/28	3/16	3/16	3/23
SC	4/7	3/18	3/16	3/18
SD			3/16	3/23

TN	4/1	3/25	3/20	3/23
TX	4/2	3/14	3/23	3/19
UT		3/12	3/16	3/18
VT	3/25	3/25	3/18	3/17
VA	3/30	3/24	3/16	3/16
WA	3/23	3/11	3/17	3/17
WV	3/24	3/24	3/16	3/24
WI	3/25	3/17	3/18	3/17
WY		3/20	3/20	3/19

Appendix Exhibit 2 – Effective Dates of County Social Distancing Measures

State	County	Date Issued	State	County	Date Issued
AL	Jefferson	3/26	MO	Randolph	3/25
AK	Anchorage	3/22	MO	Ray	3/25
CO	Adams	3/24	MO	St. Louis	3/23
CO	Arapahoe	3/24	PA	Allegheny	3/23
CO	Archuleta	3/24	PA	Beaver	3/28
CO	Boulder	3/24	PA	Berks	3/27
CO	Denver	3/24	PA	Bucks	3/23
CO	Douglas	3/24	PA	Butler	3/27
CO	Eagle	3/24	PA	Centre	3/28
CO	Grand	3/24	PA	Chester	3/23
CO	Jefferson	3/24	PA	Delaware	3/23
CO	La Plata	3/24	PA	Erie	3/24
CO	Pitkin	3/24	PA	Lackawanna	3/27
CO	San Miguel	3/24	PA	Lancaster	3/27
FL	Alachua	3/25	PA	Lehigh	3/25
FL	Broward	3/27	PA	Luzerne	3/27
FL	Duval	3/25	PA	Monroe	3/23
FL	Hillsborough	3/27	PA	Montgomery	3/23
FL	Leon	3/25	PA	Northampton	3/25
FL	Miami-Dade	3/25	PA	Philadelphia	3/23
FL	Monroe	3/30	PA	Pike	3/27
FL	Orange	3/26	PA	Washington	3/28
FL	Osceola	3/26	PA	Wayne	3/27
FL	Palm Beach	3/30	PA	Westmoreland	3/27
FL	Pinellas	3/26	PA	York	3/27
GA	Chatham	3/24	TN	Davidson	3/25
GA	Dougherty	3/21	TN	Franklin	3/25
GA	Fulton	3/24	TX	Austin	3/24
KS	Johnson	3/24	TX	Bell	3/24
KS	Leavenworth	3/24	TX	Bexar	3/24
KS	Sedgwick	3/25	TX	Brazos	3/24
KS	Wyandotte	3/24	TX	Collin	3/24
MS	Lafayette	3/22	TX	Dallas	3/23
MS	Lauderdale	3/31	TX	El Paso	3/24
MO	Boone	3/25	TX	Fort bend	3/24
MO	Cass	3/24	TX	Galveston	3/24
MO	Clay	3/24	TX	Harris	3/24
MO	Cole	3/23	TX	McLennan	3/24
MO	Greene	3/26	TX	Tarrant	3/24
MO	Jackson	3/24	UT	Salt Lake	3/30
MO	Jefferson	3/24	UT	Summit	3/27
MO	Platte	3/24			

B. Event-Study Econometric Model

$$\begin{aligned}
& \ln(cases_{cd}) - \ln(cases_{c,d-1}) \\
&= \beta_0 + \sum_{i=-2}^{-1} \beta_{1i} SIPO_{cdi} + \sum_{i=1}^5 \gamma_{1i} SIPO_{cdi} + \sum_{i=-2}^{-1} \beta_{2i} schools_{cdi} \\
&+ \sum_{i=1}^5 \gamma_{2i} schools_{cdi} + \sum_{i=-2}^{-1} \beta_{3i} events_{cdi} + \sum_{i=1}^5 \gamma_{3i} events_{cdi} \\
&+ \sum_{i=-2}^{-1} \beta_{4i} entertainment_{cdi} + \sum_{i=1}^5 \gamma_{4i} entertainment_{cdi} \\
&+ \sum_{i=0}^1 \alpha_i (\ln(tests_{c,d-i}) - \ln(tests_{c,d-1-i})) + county_c + day_by_cd_{cd} + \varepsilon_{cd}
\end{aligned}$$

$cases_{cd}$ = cumulative confirmed COVID-19 cases in county c on day d (with 1 added to prevent the variable from being undefined when the county does not yet have a case).

$SIPO_{cdi}$ = indicator for whether a shelter-in-place order was enacted i time periods before day d in county c , where the time periods are defined as

$$i = \left\{ \begin{array}{l} -2 \equiv 10 \text{ or more days before } d \\ -1 \equiv 5 \text{ to } 9 \text{ days before } d \\ 1 \equiv 1 \text{ to } 5 \text{ days after } d \\ 2 \equiv 6 \text{ to } 10 \text{ days after } d \\ 3 \equiv 11 \text{ to } 15 \text{ days after } d \\ 4 \equiv 16 \text{ to } 20 \text{ days after } d \\ 5 \equiv 21 \text{ or more days after } d \end{array} \right\}, \text{ leaving } 0 \text{ to } 4 \text{ days before } d \text{ as the omitted}$$

reference range; i.e. the coefficients reflect differences between the predicted growth rate under a policy enacted i periods ago versus one that will be enacted within the next four days.

$schools_{cdi}$ = indicator for whether schools were closed starting i periods before day d in county c .

$events_{cdi}$ = indicator for whether large events were banned (at least as stringent as 500 people) starting i periods before day d in county c .

$entertainment_{cdi}$ = indicator for whether restaurants or gyms/entertainment centers were closed starting i periods before day d in county c .

$tests_{cd}$ = cumulative COVID-19 tests recorded in the state containing county c on day d (with 1 added to prevent the variable from being undefined when the state has not yet run a test); the summation term before the log-difference in tests reflects the fact that the lag of the log difference is also included.

$county_c$ = fixed effect for county c ; this captures unobserved county-level variables that do not vary during our sample period, which is why we do not control for county characteristics, such as population density, demographics, or voting behavior, that might influence case growth rates.

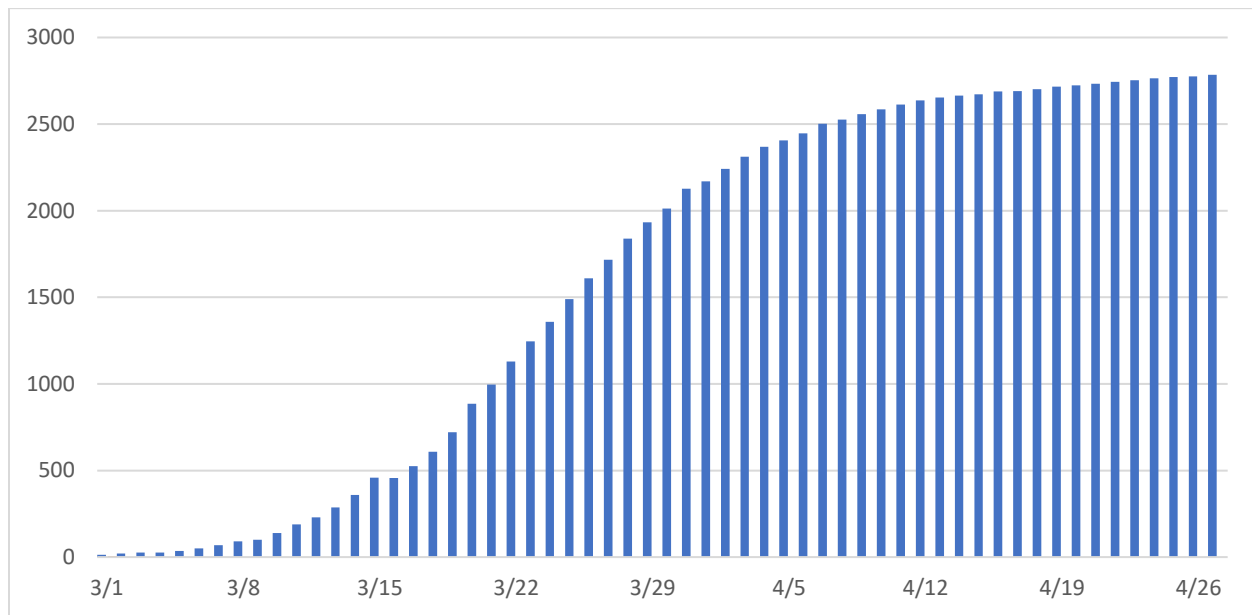
$day_by_cd_{cd}$ = fixed effect for day d (ranging from March 1 to April 27, 2020) in the Census Division containing county c ; this captures unobserved variables that trend over time, and allows these time trends to vary by Census Division.

ε_{cd} = error term for county c on day d

C. Diffusion of COVID-19 during Sample Period

The figure below shows the growth in the number of counties with confirmed COVID-19 cases during our sample period of March 1 through April 27, 2020. These are simple counts, not adjusted for population. 0.4% of counties had a COVID-19 case on 3/1, but 11% of the US population lived in those counties. Similarly, more than half of US residents lived in counties with a case by March 13, even though just 9% of counties had a case. By April 1, 97% of residents lived in counties with a case, compared to 71% of the unadjusted number of counties.

Appendix Exhibit 3 – Number of Counties with Confirmed COVID-19 Cases on Each Day



Note: All cases for Bronx, Kings, Queens, and Richmond counties in NY were lumped into New York county in the source data.

D. Robustness Checks

Multicollinearity

Our event-study specification included 24 different treatment variables (6 for each of the 4 policies) as well as Census-Division-by-day fixed effects, raising the question of whether there was sufficient identifying variation to precisely identify the impacts of the different policies. In particular, this could conceivably have explained the null results for public school closures and large event bans. Our first two robustness checks therefore utilized a simpler specification that included only a single variable for each policy type: whether the policy was in effect 10 days ago in the first regression and 20 days ago in the second. (The lag allowed for incubation periods, testing delays, etc.) Also, we included day fixed effects rather than their interactions with each of the nine Census Divisions. This means that policy differences across Census Divisions were allowed to contribute to the identifying variation, whereas the baseline model relied only on within-Census-Division variation.

The results are below. We observed the same pattern of results as we did in our main model: SIPOs had the largest statistically significant effect, restaurant/entertainment closures had the second largest, and large event bans and school closures had no discernable impacts.

Appendix Exhibit 4 – Estimated Effect of Social Distancing Policies on the Growth Rate of COVID-19 Cases in Models with a Single Variable for Each Policy

	10-Day Lag	20-Day Lag
Shelter-in-Place Order in Effect	-3.997 (0.705)*	-4.106 (0.631)*
Restaurant/Gym/Entertainment Closures	-2.417 (0.726)*	-1.951 (0.797)*
No Large Events	-0.486 (1.020)	-0.953 (1.136)
Public Schools Closed	1.705 (1.062)	0.168 (0.903)

Notes: Sample size = 182,004 county-by-day observations. * indicates statistically significant at the 5% level; * 10% level. Standard errors, in parentheses, are heteroskedasticity-robust and clustered by state. Observations are weighted by county population. Regressions include the growth rate in cumulative tests in each of the past two days as well day and county fixed effects.

Note also that, though we do not show the results here, the estimated treatment effects from our baseline event-study model were not meaningfully sensitive to the use of region-by-day or simply day fixed effects instead of Census Division-by-day fixed effects. However, the pre-trend for SIPOs was worse using the sparser specifications.

Testing

Here we conduct robustness checks aimed to ensure that the results were not driven by the way in which we controlled for testing. Specifically, we were concerned about two issues. First,

the baseline model's functional form for tests was the daily log difference in cumulative tests (as well as its first lag), which mirrored our specification for the outcome variable, effectively assuming that a given percentage increase in testing corresponded to a given percentage increase in cases. However, it is possible that the relationship between testing and cases is better characterized by a different functional form. Second, number of tests performed depends not only on availability, but also the severity of the outbreak. The coefficient estimates for the testing variables were therefore likely biased, and bias could have spilled over to the coefficient estimates for the policy variables.

Accordingly, we run four checks. The first three use different functional forms for the testing variables: 1) $\ln(\text{new tests})$ – i.e. the log of the difference rather than the difference in the logs, 2) new tests (not logged) per 1,000 residents, and 3) cumulative tests run to date. The fourth check simply excludes testing variables completely. The results, presented below alongside the corresponding estimates from the baseline model, show that testing was indeed positively associated with cases, but our coefficient estimates of interest were remarkably stable across the different specifications. While none of the testing measures are perfect, this stability increased our confidence that differential trends in access to testing across counties did not drive our conclusions.

Appendix Exhibit 5 – Robustness Checks Related to Different Ways of Controlling for Testing

	Baseline	ln(New Tests)	New Tests per 1,000	ln(Cumulative Tests)	No Control for Tests
<u><i>Shelter-in-Place Order</i></u>					
1-5 Days Ago	-1.514	-1.643	-1.641	-1.639	-1.633
6-10 Days Ago	-3.033*	-3.229*	-3.098*	-3.165*	-3.131*
11-15 Days Ago	-4.482*	-4.671*	-4.530*	-4.658*	-4.630*
16-20 Days Ago	-5.950*	-6.155*	-5.977*	-6.132*	-6.104*
21 or More Days Ago	-8.600*	-8.905*	-8.564*	-8.821*	-8.781*
<u><i>Restaurant/Gym/Entertainment Closures</i></u>					
1-5 Days Ago	-4.372*	-4.415*	-4.349*	-4.500*	-4.335*
6-10 Days Ago	-4.710*	-4.805*	-4.751*	-4.878*	-4.713*
11-15 Days Ago	-6.125*	-6.151*	-6.088*	-6.293*	-6.031*
16-20 Days Ago	-5.594*	-5.595*	-5.588*	-5.861*	-5.539*
21 or More Days Ago	-5.177*	-5.149*	-5.156*	-5.500*	-5.135*
<u><i>No Large Events</i></u>					
1-5 Days Ago	0.172	0.264	0.258	0.260	0.227
6-10 Days Ago	-0.061	0.033	0.130	-0.010	0.125
11-15 Days Ago	0.013	0.055	0.144	-0.053	0.109
16 or More Days Ago	-1.041	-1.039	-0.888	-1.140	-0.933
21 or More Days Ago	-1.272	-1.269	-1.127	-1.404	-1.164
<u><i>Public Schools Closed</i></u>					
1-5 Days Ago	0.304	0.170	0.002	0.298	0.109
6-10 Days Ago	1.004	0.928	0.683	1.149	0.864
11-15 Days Ago	2.443	2.406	2.157	2.751	2.432
16 or More Days Ago	3.465	3.475	3.179	3.868	3.505
21 or More Days Ago	3.095	3.173	2.812	3.591	3.232
Testing Variable for Day t	0.010 ⁺	0.291*	0.005*	1.235*	--
Testing Variable for Day t-1	0.012*	-0.064	0.005*	-0.676	--

See notes for Appendix Exhibit 4.

Timing

Our next series of robustness checks related to the subjective nature of the start date of the sample. For our baseline results, we began the sample period on March 1. We excluded prior days because there were no new reported cases in the U.S. on most days in January in February. We did not start later in March because using the entire month allowed for ten days prior to the implementation of any policy with which we could examine pre-trends. However, arguments could be made for either earlier or later start dates. On one hand, an earlier start date would utilize more data and allow for a longer pre-treatment period. On the other hand, the earlier the start date, the more counties started with sample with no cases, exacerbating any potential issues created by adding one to the log of cases to prevent the outcome from being undefined.

Our first two checks simply started the sample a half month earlier and later: Feb. 15 and March 15, respectively. The third check used all available data, meaning a Jan. 22 start date. The fourth check used this longer period but avoided adding one to the log of cases by running two separate regressions: 1) one with the full sample but an indicator for any cases as the outcome (linear probability model due to the bias inherent in nonlinear fixed effects models), and 2) one with exponential growth rate as the outcome, but without adding one to the log of cases when creating the variable, thereby restricting the sample to counties with cases on both the current and previous days.¹¹ Next, we return to the original March 1 start date and avoid adding one to the log of cases a different way: by using the inverse hyperbolic sine transformation, which is interpretable in the same manner as natural logs but defined at zero.¹²

The table below displays the results. Across all start dates, we continued to find that SIPOs and restaurant/entertainment center closures – but not the other two policies – reduced the growth rate of cases. Interestingly, the results from the two-part specification suggested that SIPOs reduced COVID-19 spread along both the extensive margin (delaying the arrival of the first case in the county) and the intensive margin (slower spread once the first case was recorded), whereas restaurant/entertainment center closures did so only along the intensive margin. The results from the regression using the inverse hyperbolic sine specification showed this more directly, as those results were virtually identical to those from the baseline model.

¹¹ Note that this is not exactly equivalent to the conventional two-part model, which would use a dichotomized version of the continuous outcome as the dependent variable in the first part. In our setting, that would mean using the original outcome variable (with one added to the log of cases), and conditioning the sample for the second part on having a non-zero growth rate, meaning counties with cases but no growth in the previous day would also be dropped. Given these distinctions, it is not obvious to us how one might compute combined marginal effects that encompass both parts of the model, so we do not do so.

¹² For further detail, see John B. Burbidge, Lonnie Magee and A. Leslie Robb. “Alternative Transformations to Handle Extreme Values of the Dependent Variable.” *Journal of the American Statistical Association* Vol. 83, No. 401 (Mar., 1988), pp. 123-127.

Appendix Exhibit 6 – Robustness Checks Related to Sample Start Dates

	Start on 2/15	Start on 3/15	Start on 1/22	Start on 1/22; Any Cases as Outcome	Start on 1/22; Counties with Cases Only	Start on 3/1; Inverse Hyperbolic Sine
<u><i>Shelter-in-Place Order</i></u>						
1-5 Days Ago	-1.144	-0.783	-0.994	-0.024*	-1.045	-1.844
6-10 Days Ago	-2.332*	-1.736*	-2.028*	-0.059*	-2.117*	-3.456*
11-15 Days Ago	-3.412*	-2.652*	-2.929*	-0.090*	-2.898*	-4.794*
16-20 Days Ago	-4.517*	-3.626*	-3.847*	-0.112*	-3.666*	-6.052*
21 or More Days Ago	-6.541*	-5.480*	-5.536*	-0.143*	-5.239*	-8.727*
<u><i>Restaurant/Gym/Entertainment Closures</i></u>						
1-5 Days Ago	-4.060*	-5.779*	-3.936 ⁺	0.011	-6.778*	-4.821*
6-10 Days Ago	-4.164 ⁺	-6.406*	-3.952 ⁺	0.002	-6.242*	-5.123*
11-15 Days Ago	-5.372*	-8.016*	-5.081*	0.010	-8.085*	-6.556*
16-20 Days Ago	-4.648 ⁺	-7.825*	-4.302 ⁺	0.027	-7.694*	-5.997*
21 or More Days Ago	-4.067 ⁺	-7.849*	-3.683	0.046	-7.549*	-5.556*
<u><i>No Large Events</i></u>						
1-5 Days Ago	0.176	0.528	0.213	0.028	-1.713	0.007
6-10 Days Ago	-0.082	0.185	-0.011	0.036	-2.107	-0.509
11-15 Days Ago	-0.012	0.553	0.089	0.035	-2.300	-0.302
16-20 Days Ago	-1.077	-0.317	-0.941	0.033	-3.802	-1.414
21 or More Days Ago	-1.349	-0.300	-1.177	0.030	-4.260	-1.568
<u><i>Public Schools Closed</i></u>						
1-5 Days Ago	-0.070	0.260	-0.188	-0.062	1.909	1.033
6-10 Days Ago	0.231	0.334	-0.0004	-0.085	2.656	2.153
11-15 Days Ago	1.335	1.137	1.025	-0.095	3.390	3.956
16-20 Days Ago	2.090	1.786	1.732	-0.107	3.851	5.160
21 or More Days Ago	1.495	1.038	1.116	-0.115	2.774	5.032
Sample Size	229,074	138,072	298,110	298,110	93,245	182,004

See notes for Appendix Exhibit 4.

Dropping States

Our next group of checks asked whether results were robust to dropping particular, unique locations: New York, the hardest-hit state; Washington, the first state to be hit; and California, which had the nation's largest population and earliest shelter-in-place order. Appendix Exhibit 7 contains the results, which show that dropping any of these states has little impact on the estimates.

Appendix Exhibit 7 – Robustness Checks Dropping States

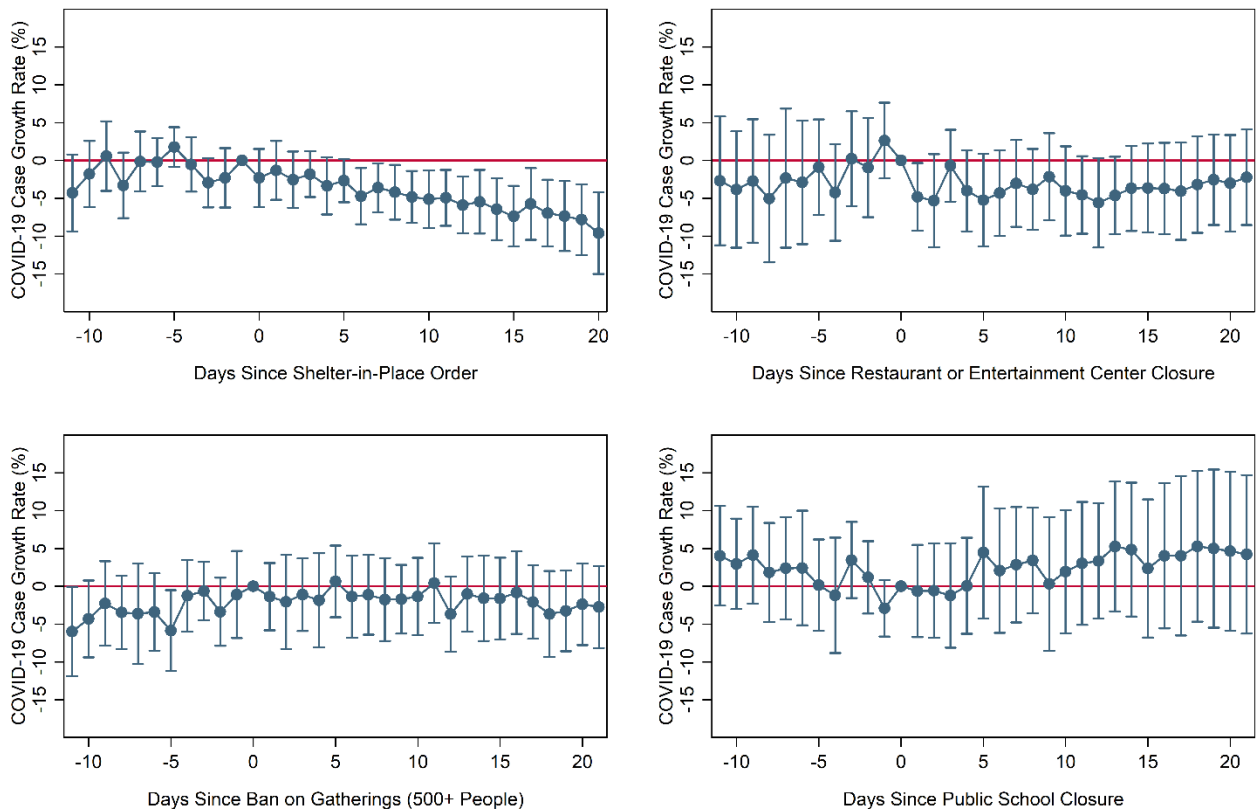
	Drop NY	Drop WA	Drop CA
<u><i>Shelter-in-Place Order</i></u>			
1-5 Days Ago	-1.320	-1.564	-1.475
6-10 Days Ago	-2.853*	-2.987*	-2.770*
11-15 Days Ago	-4.248*	-4.616*	-4.531*
16-20 Days Ago	-5.681*	-6.069*	-5.940*
21 or More Days Ago	-8.336*	-8.776*	-8.523*
<u><i>Restaurant/Gym/Entertainment Closures</i></u>			
1-5 Days Ago	-4.691*	-4.991*	-5.068*
6-10 Days Ago	-5.189*	-5.293*	-5.220*
11-15 Days Ago	-6.645*	-6.621*	-6.430*
16-20 Days Ago	-6.313*	-6.271*	-6.111*
21 or More Days Ago	-5.947*	-5.875*	-5.603*
<u><i>No Large Events</i></u>			
1-5 Days Ago	-0.051	0.483	0.609
6-10 Days Ago	-0.414	0.723	0.719
11-15 Days Ago	0.354	0.850	0.700
16 or More Days Ago	-0.388	-0.004	-0.239
21 or More Days Ago	-0.230	-0.050	-0.356
<u><i>Public Schools Closed</i></u>			
1-5 Days Ago	0.923	0.644	0.521
6-10 Days Ago	1.285	2.084	2.368
11-15 Days Ago	2.500	3.614	3.445
16 or More Days Ago	3.442	4.851	4.697
21 or More Days Ago	2.808	4.665	4.398
Sample Size	178,640	179,742	178,640

See notes for Appendix Exhibit 4.

Event-Study Model with One-Day Bins

Next, we estimate a more fine-grained event study model that included indicators for each day from 10 days before a policy took effect to 20 days after, rather than grouping them into 5-day increments. Appendix Exhibit 8 below displays the results. Not surprisingly, the estimates were less precise than those obtained using 5-day groupings, and the wider confidence intervals generally made the results less informative. In particular, we no longer observed significant effects for restaurant/entertainment center closures, despite most of the point estimates remaining roughly similar to those obtained using the 5-day intervals. The confidence intervals included reductions as large as 8 or more percentage points, making the lack of statistical significance uninformative. In contrast, the confidence intervals for SIPOs remain smaller, and most of those estimates remain statistically significant. One useful result from this check is that the apparent upward pre-treatment trend for gathering bans was concentrated among the first half of the pre-treatment window and largely disappeared four days before treatment. This increased our confidence that the failure to find a causal effect of these bans was not due to confounding from pre-trends.

Appendix Exhibit 8 – Results from Event-Study Model Including Day Indicators



Other Miscellaneous Checks

Next, two robustness checks considered alternate constructions of two of our policy variables. The first of these used bans on 50 or more, rather than 500 or more, people for the “event” variables. The second used closures of restaurants and gyms/entertainment centers, rather than one or the other, modeled by including both measures separately and then summing their coefficients. (Since the two measures were so strongly correlated, parsing out their separate effects with meaningful precision is not possible.)

We then explored sensitivity to dropping the observations with questionable case numbers, rather than imputing them. As described above, these included the entire day of Sunday, March 15, 2020 as well as all instances of negative reported case growth.

Next, recall that one of the motivations for interacting the day fixed effects with Census Division indicators was to account for different parts of the country starting their outbreaks at different times. As an alternative way to account for this, we replaced these day-by-Census Division fixed effects with fixed effects for days since (or until) the first case in the state. Since there was a case in every state prior to any distancing restrictions being implemented, the timing of this first case was not endogenous to our variables of interest.

Finally, we returned to the model with Census Division-by-day fixed effects but controlled for extrapolated county-specific pre-treatment trends in case growth rates. These were computed by restricting the sample to March 1-11, 2020 before any policies were enacted, running county-specific linear regressions of our outcome variable on day, and using the coefficient estimate to project changes after March 11.

The table below shows that the results for these checks were similar to those from our baseline model.

Appendix Exhibit 9 – Miscellaneous Robustness Checks

	Bans on 50 Instead of 500	Restau- rants <u>and</u> Gyms, not <u>or</u>	Drop Ques- tionable Data	Fixed effects for days since first case	Control for Pre- Trend
<u><i>Shelter-in-Place Order</i></u>					
1-5 Days Ago	-1.733	-1.393	-1.402	-2.361*	-1.525
6-10 Days Ago	-3.189*	-2.800*	-2.912*	-4.176*	-3.038*
11-15 Days Ago	-4.732*	-4.302*	-4.333*	-5.867*	-4.461*
16-20 Days Ago	-6.262*	-5.780*	-5.775*	-7.47*	-5.888*
21 or More Days Ago	-8.871*	-8.351*	-8.383*	-10.414*	-8.438*
<u><i>Restaurant/Gym/Entertainment Closures</i></u>					
1-5 Days Ago	-4.581*	-3.797*	-4.144*	-3.218 ⁺	-4.407*
6-10 Days Ago	-4.905*	-3.068	-4.427 ⁺	-5.351*	-4.762*
11-15 Days Ago	-6.612*	-4.642*	-5.740*	-6.575*	-6.186*
16-20 Days Ago	-6.330*	-3.988	-5.159 ⁺	-7.109*	-5.656*
21 or More Days Ago	-5.970*	-3.357	-4.666 ⁺	-7.092*	-5.221*
<u><i>No Large Events</i></u>					
1-5 Days Ago	1.971	-0.046	0.188	0.208	0.172
6-10 Days Ago	1.333	-0.787	-0.073	0.488	-0.055
11-15 Days Ago	1.451	-0.684	-0.008	0.643	0.030
16 or More Days Ago	1.085	-1.602	-1.064	-0.358	-1.012
21 or More Days Ago	0.280	-1.800	-1.316	-0.191	-1.218
<u><i>Public Schools Closed</i></u>					
1-5 Days Ago	0.649	0.080	0.315	3.207	0.356
6-10 Days Ago	1.921	0.736	0.933	1.717	1.092
11-15 Days Ago	3.521	2.151	2.367	1.396	2.544
16 or More Days Ago	4.882	3.244	3.335	0.641	3.567
21 or More Days Ago	4.798	2.881	2.933	-0.030	3.176
Sample Size	182,004	182,004	182,004	182,004	182,004

See notes for Appendix Exhibit 4.

E. Creation of Counterfactuals in Supplemental Exhibits 4 and 5

In Supplemental Exhibit 4, we plot the daily growth rate of COVID-19 cases in the U.S. under three scenarios. The first is actual case growth rates, which reflect the presence of all four social distancing measures when and where they were actually implemented. The second is counterfactual growth rates when we assume no jurisdictions ever implemented a SIPO but did implement the other measures the same way they did in reality. Formally, this is done by subtracting out the effect of SIPOs from the actual growth rate so that

$$counterfactual1_{cd} = \ln(cases_{cd}) - \ln(cases_{c,d-1}) - \hat{\gamma}_1 \sum_{i=1}^5 SIPO_{cdi}.$$

The third is the counterfactual when we assume no jurisdictions ever implemented any social distancing restriction:

$$\begin{aligned} counterfactual2_{cd} &= \ln(cases_{cd}) - \ln(cases_{c,d-1}) - \hat{\gamma}_1 \sum_{i=1}^5 SIPO_{cdi} - \hat{\gamma}_2 \sum_{i=1}^5 schools_{cdi} \\ &\quad - \hat{\gamma}_3 \sum_{i=1}^5 events_{cdi} - \hat{\gamma}_4 \sum_{i=1}^5 entertainment_{cdi} \end{aligned}$$

Note that we do *not* also subtract out the “placebo” effects of the future policy implementation variables, since those are intended to capture unobserved confounders rather than part of the causal effect of the policies. After obtaining the counterfactuals, we aggregate all three growth rate variables to the nation-by-day level by weighting each county-by-day observation by county population.

Supplemental Exhibit 5 builds on Supplemental Exhibit 4 to predict the number of cases under each counterfactual scenario. Note that $\ln(cases_{cd}) - \ln(cases_{c,d-1}) = \ln(\frac{cases_{cd}}{cases_{c,d-1}})$ implies that

$$cases_{cd} = cases_{c,d-1} \exp \left[\ln \left(\frac{cases_{cd}}{cases_{c,d-1}} \right) \right].$$

With a predicted growth rate, $\ln(\frac{\widehat{cases}_{cd}}{cases_{c,d-1}})$, this becomes

$$\widehat{cases}_{cd} = \widehat{cases}_{c,d-1} \exp \left[\ln \left(\frac{\widehat{cases}_{cd}}{cases_{c,d-1}} \right) \right] \exp(\varepsilon_{cd}), \quad (A1)$$

where $\widehat{\exp(\varepsilon_{cd})}$ is the average of $\exp(\widehat{\varepsilon_{cd}})$ in the estimation sample.¹³

Therefore, predicting number of cases under each scenario requires the estimated residual from our regression and predicted values of $\ln\left(\frac{cases_{cd}}{cases_{c,d-1}}\right)$ under the counterfactuals discussed above. Each of these estimates is then plugged into equation (A1) to predict counterfactual cases in county/day observations following the start date of a relevant policy. Prior to a policy beginning in a given county, the “counterfactual” cases are no different from the observed cases. Once the number of cases are predicted in each period under a given counterfactual, we sum the observed and predicted cases by day to create nation-by-day totals for observed and counterfactual cases. When these sums are presented on a logarithmic scale, as they are in Supplemental Exhibit 5, the natural log must be calculated again after summing each variable.

This calculation is subject to several technical caveats that together mean the results should be interpreted as rough approximations rather than exact numbers. There is always uncertainty with calculations based on underlying parameters that are estimated, and it is difficult to accurately quantify the amount of this uncertainty. In this particular calculation, errors can compound over time, as \widehat{cases}_{cd-1} in equation (A1) is itself an estimate. Even when \widehat{cases}_{cd} is predicted from observed policies and using observed cases in $d - 1$, the predicted value is larger than $cases_{cd}$ on average in our sample. Additionally, the average difference is always positive later in the sample. To err on the side of caution and avoid having these errors accumulate over time, we discount our counterfactual predictions based on the average ratio of observed to predicted cases whenever the predicted cases (based on observed policies) are greater than observed cases on average.

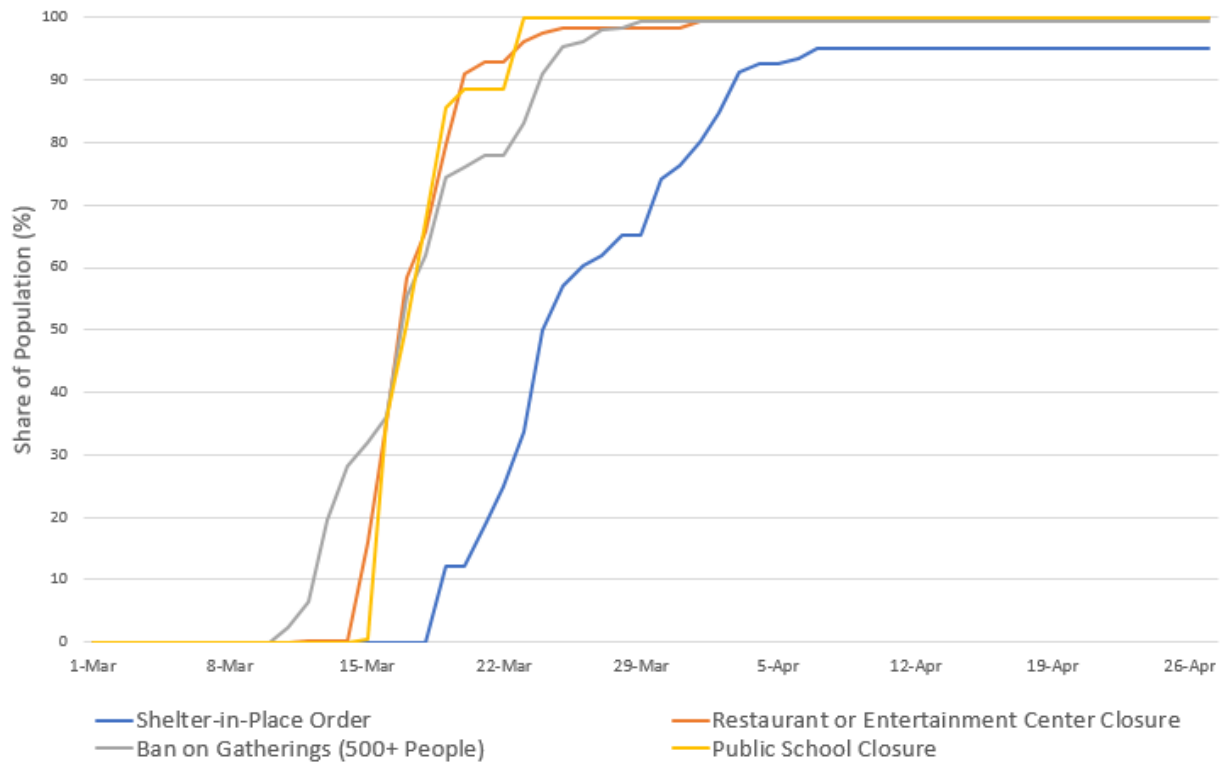
We also err on the side of caution in one other way. When calculating the predicted case counts for each counterfactual, we set the coefficient on the variables indicating a policy start date 21 or more days in the past ($i = 5$ in the summation above) equal to the corresponding coefficient for 16-20 days. As discussed above and in the manuscript, our sample period includes up to 20 days after each policy. This means there is geographic balance for the variables reflecting up to 20 days, enabling apples-to-apples comparisons. However, the comparability of the estimates for the 21-or-more-day indicators is less clear, as earlier SIPOs remain in that category longer than more recent SIPOs. It is therefore not clear that the additional effect of SIPOs observed after 21+ days relative to 16-20 days is a “true” increase as opposed to simply the artifact of the compositional changes. We therefore make the conservative assumption that the true effect peaked and remained constant after 16 days. Without this adjustment, the estimated cases with no social distancing policies rises from ~35 million to ~50 million.

¹³ Don M. Miller (1984) Reducing Transformation Bias in Curve Fitting, *The American Statistician*, 38:2, 124-126, DOI: [10.1080/00031305.1984.10483180](https://doi.org/10.1080/00031305.1984.10483180)

An assumption that pushes in the direction of overstatement is that the pool of potentially infectable individuals is fixed over time. This ignores immunity from fighting off the virus. By the end of the sample period, with nearly a tenth of the US population predicted to have been sick enough to reach the threshold for a confirmed case, this assumption may not be benign. However, debate continues over the extent to which exposure to the virus conveys immunity, so the exact implications of this assumption are not clear.

SUPPLEMENTAL EXHIBIT 1

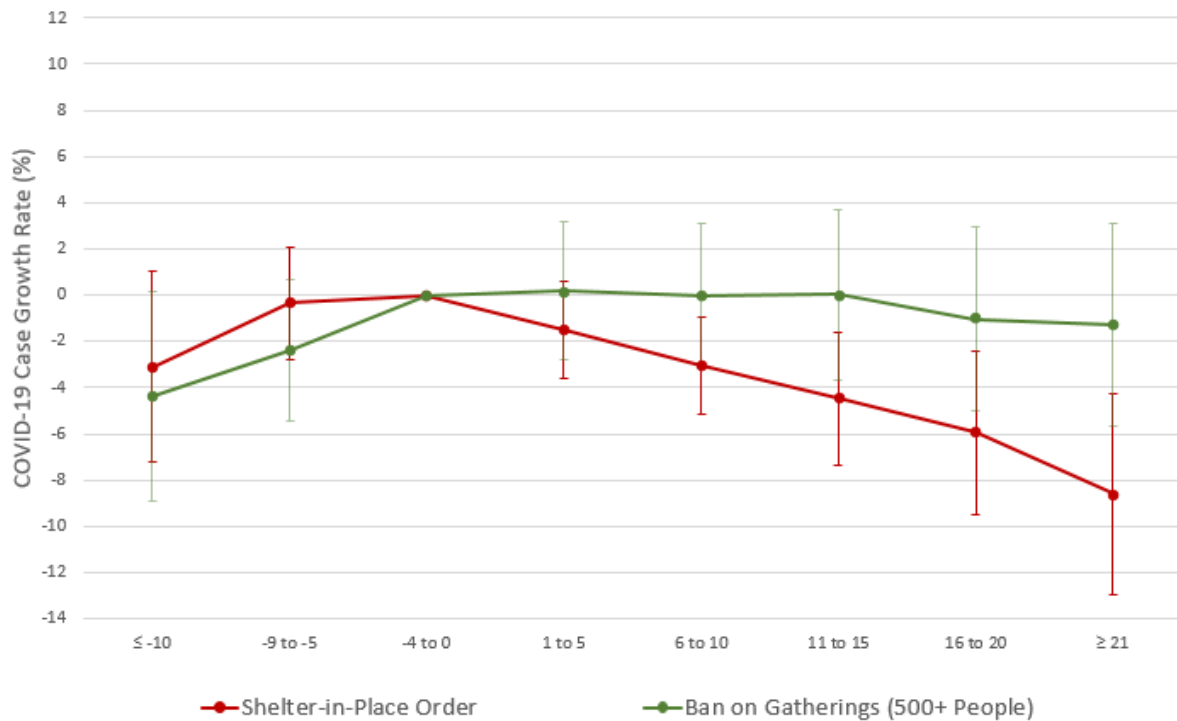
Fraction of US Population Covered by Social Distancing Measures



SOURCES: Authors' calculations from COVID-19 Data by the Johns Hopkins Center for Systems Science and Engineering. NOTES: Estimates are weighted by county population.

SUPPLEMENTAL EXHIBIT 2

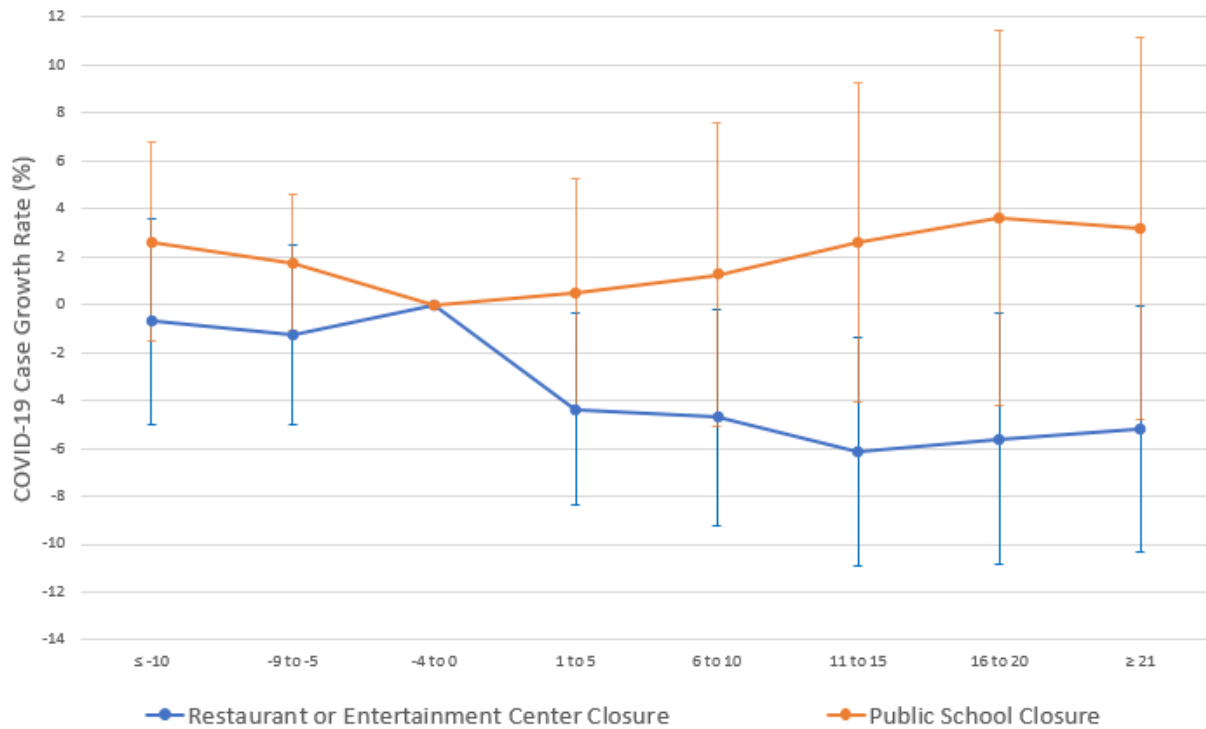
Coefficient Estimates from Event-Study Model – SIPO and Bans on Large Gatherings



SOURCE: Authors' analysis of county-level COVID-19 case data. NOTES: Day, county, and Census-Division-by-day fixed effects and testing growth controls were included. Standard errors were heteroskedasticity-robust and clustered by state.

SUPPLEMENTAL EXHIBIT 3

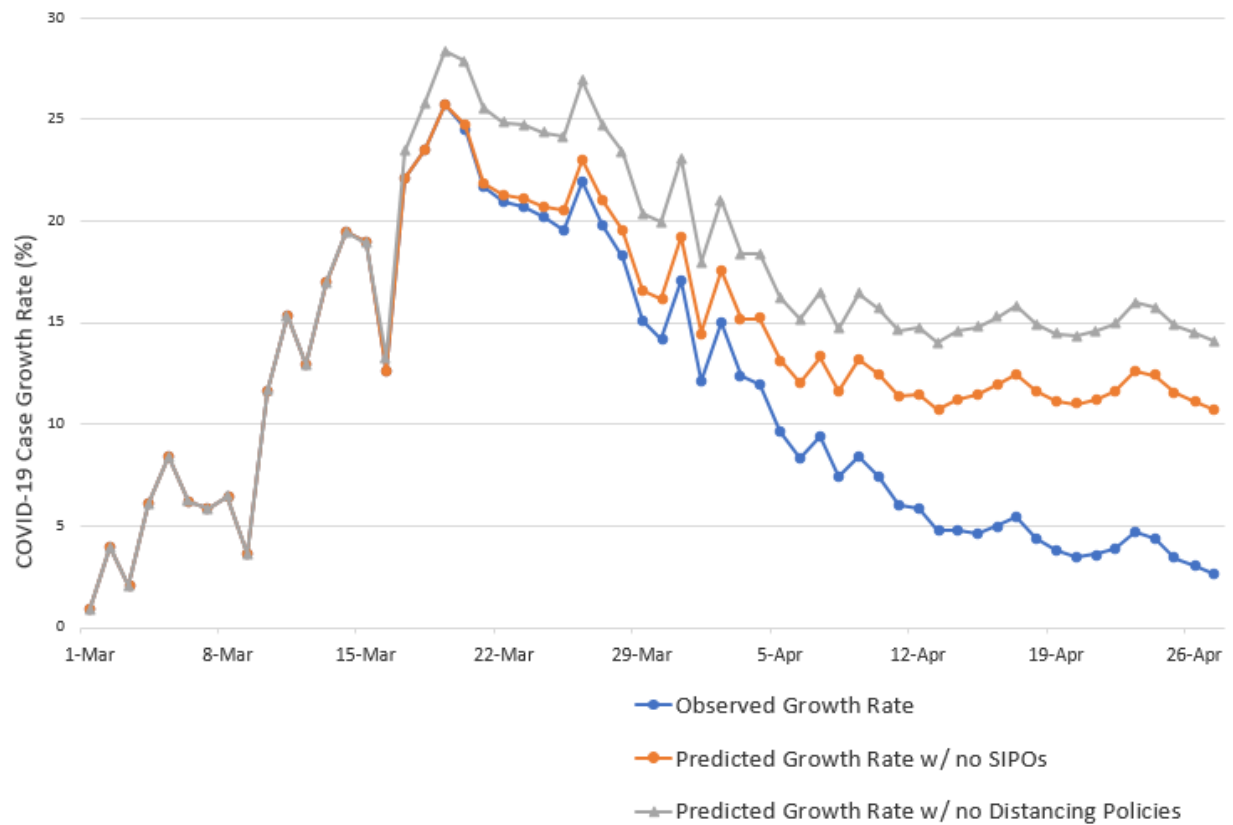
Coefficient Estimates from Event-Study Model – Public School Closure and Bans on Restaurants/ Entertainment Centers



SOURCE: Authors' analysis of county-level COVID-19 case data. NOTES: Day, county, and Census-Division-by-day fixed effects and testing growth controls were included. Standard errors were heteroskedasticity-robust and clustered by state.

SUPPLEMENTAL EXHIBIT 4

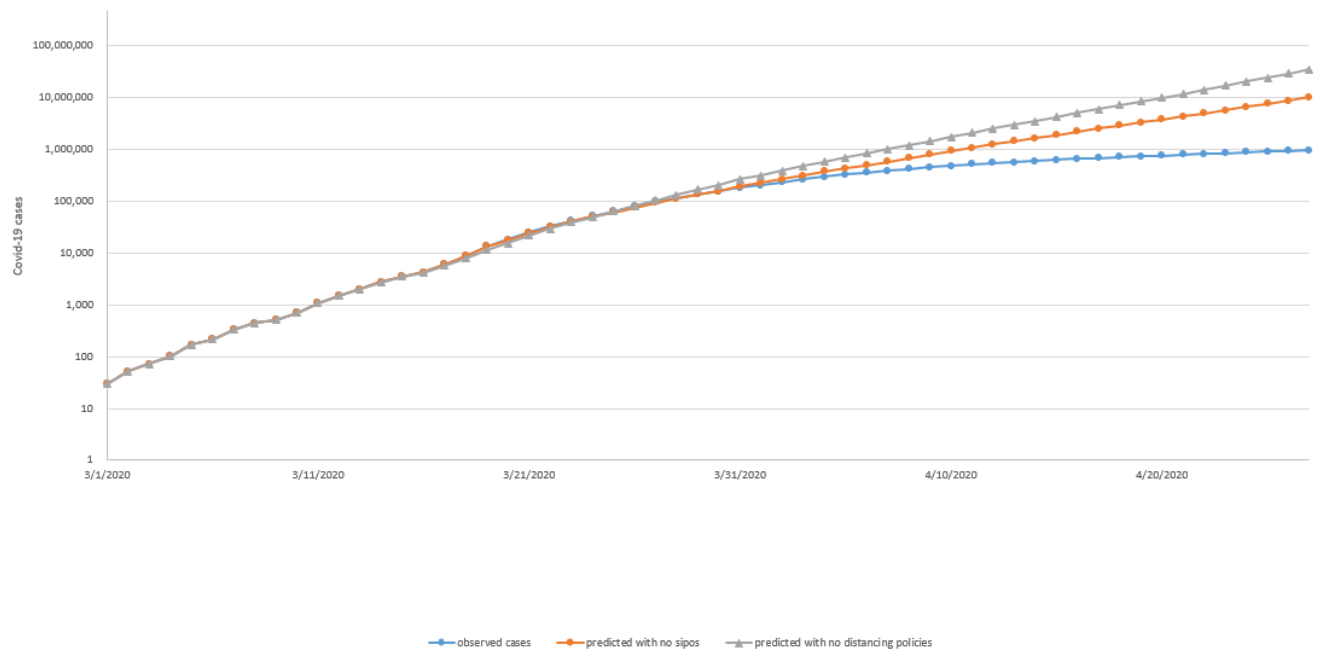
Comparison of Observed COVID-19 Growth Rate Compared to Estimated Growth Rates without Social Distancing Measures



SOURCE: Authors' calculation based on observed data and coefficient estimates from the event-study model. NOTES: The counterfactual growth rates set the SIPO or all four social distancing policies to zero in the event study window for all US counties.

SUPPLEMENTAL EXHIBIT 5

Strong Social Distancing Policies Flatten the Curve of Confirmed COVID-19 Cases (Natural Log Scale)



SOURCE: Estimates derived from authors' event-study model.