Evaluating the effects of shelter-in-place policies during the COVID-19 pandemic

Christopher R. Berry, Anthony Fowler, Tamara Glazer, Samantha Handel-Meyer, and Alec MacMillen

*Harris School of Public Policy, University of Chicago, Chicago, IL 60637

We estimate the effects of shelter-in-place (SIP) orders during the first wave of the COVID-19 pandemic. We do not find detectable effects of these policies on disease spread or deaths. We find small but measurable effects on mobility that dissipate over time. And we find small, delayed effects on unemployment. We conduct additional analyses that separately assess the effects of expanding versus withdrawing SIP orders and test whether there are spillover effects in other states. Our results are consistent with prior studies showing that SIP orders have accounted for a relatively small share of the mobility trends and economic disruptions associated with the pandemic. We reanalyze two prior studies purporting to show that SIP orders caused large reductions in disease prevalence, and show that those results are not reliable. Our results do not imply that social distancing behavior by individuals, as distinct from SIP policy, is ineffective.

COVID-19 | shelter-in-place policies | mobility | disease spread | government policy

The rapid onset and unprecedented scope of the COVID-19 pandemic has prompted dramatic changes in individual behavior and public policy. On March 13, 2020, the White House declared a state of national emergency in the United States. Shortly thereafter, many states and counties began enacting shelter-in-place (SIP) orders intended to reduce human interaction and, in turn, the likelihood of disease transmission. The timing, severity, and enforcement of state orders, however, have significantly varied, with some governors holding off on imposing SIP orders altogether. SIP policies have been controversial, and beliefs about their effectiveness vary widely.

A better understanding of the impacts of SIP policies will inform political leaders and public health officials as they continue to navigate the pandemic and consider trade-offs between the potential health benefits and the potential economic and personal costs of restrictive policies. Beyond the current crisis, our study also offers lessons about the extent to which citizens comply with government orders and the extent to which public policy can significantly alter behavior over a short period of time. Our results suggest that county- and state-level policies did not have large effects on behavior and health. Perhaps this is because many citizens were already altering their behavior voluntarily in the absence of government policies, possibly in response to messaging from public health experts and government officials. And perhaps this is because the pandemic has been highly politicized and some citizens are unwilling to change their behavior even in the presence of government orders. Our results differ from those of at least two prior studies claiming that SIP policies caused large reductions in disease prevalence (1, 2). We reanalyze those studies and show that their conclusions are not robust to reasonable changes in model specification. In particular, the results of Hsiang et al. (1) are not robust to the inclusion of day fixed effects, while those of Dave et al. (2) depend on weighting by population and excluding New York and New Jersey.

Estimating the effects of SIP orders is complicated by preexisting trends in our outcomes of interest, as well as complex epidemiological dynamics in the spread of the disease. As such, we do not hang our hats on any one particular estimate but rather explore a variety of reasonable statistical modeling choices. None of the models indicates that SIP orders significantly reduced COVID-19 cases or deaths.

Overall, our results suggest that SIP orders in the first wave of the pandemic did not produce large health benefits but also accounted for a small share of pandemic-related economic disruptions. To be clear, our findings do not mean that sheltering in place and social distancing behaviors had no effect on the disease. Indeed, the health benefits of SIP orders were likely limited because many people were already social distancing before the introduction of SIP orders, and others failed to comply with SIP orders in a highly politicized pandemic. Our results also do not mean that other government actions, such as emergency declarations or public health advisories, had no effect, nor do they mean that future SIP orders could not be more effective.

Related Literature

The COVID-19 pandemic has attracted immense scholarly interest and spawned a rapidly growing empirical and theoretical literature. We focus specifically on empirical studies of the effects of state and local policies in the United States. Prior studies have focused on the effects of SIP orders on three related types of outcomes: social distancing, economic outcomes, and disease prevalence.

Studies of social distancing have relied heavily on cell phone mobility data. While all studies document a large reduction in mobility associated with the spread of COVID-19, most find the trends were occurring even in states without SIP policies, suggesting that the behavioral changes were largely voluntary. Some have found that state and local policies have been more effective
in reducing mobility in wealthier places (3) and more Democratic places (4). On the whole, however, the estimated effects of state and local policies on social distancing appear to be small and to dissipate quickly (5, 6).

Studies of the effects of SIP policies on economic outcomes have also found that, while there have been large economic disruptions associated with the pandemic, SIP policies account for a relatively small share of observed declines in economic

---

Fig. 1. Mobility by state over time. Day indicates days after February 24. Mobility is the proportional change relative to the beginning of the time series. The dashed lines indicate the start of a state SIP policy, and the dotted lines indicate the end. There are clear nationwide trends, and mobility was already declining in most states before a SIP policy.
activity, again suggesting that behavioral changes were largely voluntary (6–9).

Despite the reportedly small effects of SIP policies on mobility and social distancing and economic activity, several studies have found that SIP orders lead to large reductions in COVID-19 cases (1, 2, 10). These findings pose a puzzle for the literature. How can SIP policies be so effective in reducing the spread of COVID-19 if they don’t meaningfully reduce social distancing? One possibility is that we do not have the right measures of social distancing. Perhaps SIP orders do not meaningfully reduce mobility (as measured by the GPS locations of smartphones), but they do reduce the extent to which people come into close contact with one another. Another possibility is that previous studies have overestimated the effects of SIP orders on COVID-19 cases, and these policies have not dramatically mitigated the spread of the disease. Our reanalysis of the prior studies shows that the latter possibility is indeed the most likely one.

Materials and Methods

We study two outcomes that we take to be the main target of SIP policies: the number of COVID cases and the number of COVID-related deaths. We normalize each of these variables by state population to measure new cases and deaths per million residents.* Our period of study runs from February 24, 2020 through May 30, 2020, allowing us to examine trends well before the introduction of SIP policies and well after their removal.

In addition, we study the primary mechanism through which SIP orders are thought to influence cases and deaths: social distancing. No single metric entirely captures this behavior. Acknowledging this limitation, we utilize anonymized, aggregated smartphone data from the Norwegian location data company UNACAST that measures individual movement at the county level. Specifically, we use a daily measure of the change in average distance traveled compared to a pre–COVID-19 baseline period as a proxy for social distancing. The baseline is represented as the average distance traveled for the same county on the same weekday day over the 4 wk prior to March 8, 2020. The data assign an individual to a county based on where a device is present for the longest amount of time each day, and the data follow 15 million to 17 million devices. This study uses data collected from February through May 2020.

Our final outcome is unemployment, which is meant to capture some of the unintended disruptive impacts of SIP policies. We utilize data from the Department of Labor on the insured unemployment rate in each state, which is computed from unemployment insurance claims. This measure surely understates total unemployment, but it has the advantage, for our purposes, of being updated weekly at the state level.

Utilizing data on the timing of state and county SIP orders, we compute the proportion of state residents who are under a SIP order on a particular day. In cases where the state has implemented a SIP order, this variable takes a value of one, and it can take values between zero and one if there is no state order but one or more counties have implemented SIP orders. SIP orders include “stay-at-home,” “safer-at-home,” and related orders, as well as policies ordering the closure of nonessential businesses. Additional details regarding data sources are provided in SI Appendix.

While estimating policy effects is not easy even in the best of times, there are additional challenges in this setting. First, there were very large changes in behavior taking place both before SIP policies were adopted in states that ultimately adopted them and, at the same time, in states that never adopted SIP policies during our study period. Fig. 1 illustrates this point by showing trends in mobility for every state during our study period. Dashed lines indicate the days on which various states implemented SIP orders, and the dotted lines indicate days on which states withdrew SIP orders. We see, in the figure, that there were dramatic nationwide trends in mobility that are unrelated to state SIP policies, and, in many states, mobility was already declining before the policies were implemented. Analogous figures for our other outcomes are shown in SI Appendix, Figs. S1–S3. These figures show that there were preexisting trends in all outcomes prior to the adoption of SIP policies, meaning that it will be especially important to control for pre-treatment trends when attempting to isolate changes in outcomes due to policy.

Second, SIP policies likely have both direct and indirect effects. Direct effects arise, for example, from business and school closures, because of which individuals no longer have reason to travel to those locations. Indirect effects arise from the information conveyed to the public by the enactment of the policy. The announcement of a SIP order might be an informative signal about the severity of disease conditions that could, for instance, cause some people to work from home even in the absence of a workplace closure. Such indirect effects could spill over across state borders, if one state’s policy announcement induces voluntary changes in behavior in other states. SI Appendix presents several additional analyses intended to address concerns about spillovers across states.

To account for differences across states and nationwide trends, we implement differences-in-differences designs with state and day fixed effects. State fixed effects account for any unchanging differences between states, and day fixed effects account for secular trends, allowing us to estimate the effect of SIP policies by testing for differential trends when those policies change. In addition, we control for lagged values of SIP policies precisely as COVID-19 is trending poorly in that state (relative to the rest of the country), we also control for lagged values of the dependent variable in the prior 14 d. This approach flexibly controls for potentially nonlinear prior trends in the outcome of interest. We show, in SI Appendix, that the inclusion or exclusion of the lagged dependent variable does not meaningfully affect any of our estimates but, in some cases, affects the magnitude, and we explain why we put more faith in the estimates that account for these lags.

This specification is depicted by the following equation:

$$DV_t = ISP_{it} + \gamma_i + \delta_t + \sum_{k=1}^{14} \theta_k DV_{t-k} + \epsilon_t,$$

where $DV_t$ is the outcome of interest in state $i$ on day $t$, $ISP_{it}$ is our measure of the share of a state covered by SIP policies in state $i$ on day $t$. $\gamma_i$ represents state fixed effects, $\delta_t$ represents day fixed effects, and the quantity of interest is $\beta$, a local average treatment effect of SIP policies. To complement this analysis, we also test for lagged or decaying effects of SIP orders over time by including lagged values of the treatment variable from the prior 14 d. This allows us to estimate and visualize how the effects of the policy change over time,

$$DV_t = \sum_{k=0}^{14} \rho_k ISP_{i,t-k} + \gamma_i + \delta_t + \sum_{k=1}^{14} \theta_k DV_{t-k} + \epsilon_t.$$

In Eq. 2, $\rho_k$ corresponds to the effect of SIP policies $k$ days ago. So, to estimate the effect of having SIP policies in place for two continuous days, for example, we would add $\rho_0$, $\rho_1$, and $\rho_2$. In SI Appendix, we show more detailed results where we include or exclude lagged values of the treatment, lagged values of the outcome, and leading values of the outcome.

Results

Table 1 presents our results for our four main outcomes based on Eq. 1. We look at COVID-19 cases, COVID-19 deaths, mobility, and unemployment. Cases and deaths are coded as new cases or new deaths per million residents. Mobility is measured as the proportional change in distance traveled. Unemployment is the insured unemployment rate.

We find no evidence that SIP policies led to reductions in new COVID cases or deaths; indeed, the point estimates for both outcomes are positive but insignificant. We find that SIP policies did decrease mobility. The estimate implies that SIP orders decreased mobility, on average, by 0.7%. This effect is very small relative to the nationwide trend. Between late February and mid-April, mobility nationwide decreased by about 50%, so SIP orders explain only a tiny fraction of this general trend. The estimated effects of SIP orders on unemployment is positive but not statistically significant (although it is significant in some alternative specifications—see SI Appendix).

---

*Other studies have examined the natural logarithm of the cumulative number of cases or the change in the natural logarithm of cumulative cases. We discuss and replicate these studies in SI Appendix. Our results are unchanged if we utilize those measures, but, in our view, a per capita measure better captures what policy makers should care about—the health of the population. A substantively insignificant increase in cases could correspond to a large change in log(cases) if a state was starting at a low level, and a substantively important increase in cases could correspond to a small change in log(cases) if a state already had a lot of previous cases. Furthermore, although there could be an analog of a log-linear specification for studying the proportional change of new cases or deaths, this approach is not feasible because, as discussed in SI Appendix, there are observations for which the reported number of new cases or deaths is negative.
Table 1. Effects of SIP policies on COVID-19, mobility, and unemployment

<table>
<thead>
<tr>
<th>DV</th>
<th>Deaths</th>
<th>Mobility</th>
<th>Unemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shelter in Place</td>
<td>3.804</td>
<td>0.328</td>
<td>0.007**</td>
</tr>
<tr>
<td>Day FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>State FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Controls for lagged DV</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Observations</td>
<td>4,150</td>
<td>4,150</td>
<td>4,200</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.676</td>
<td>0.696</td>
<td>0.953</td>
</tr>
</tbody>
</table>

State-clustered SEs are in parentheses; **P < 0.01. Shelter in place is measured as the proportion of a state’s residents under a SIP order on a given day. Cases and deaths are coded as new cases or new deaths per million residents. Mobility is measured as the proportional change in distance traveled. Unemployment is the insured unemployment rate. For cases, deaths, and mobility, we include 14 controls for the lagged dependent variable for each of the preceding 14 d, and, for unemployment, which is measured weekly, we include two controls for the lagged dependent variable from 7 and 14 d prior. DV, dependent variable; FE, fixed effects. X indicates they were included in the regression.

Our null results on the effects of SIP orders on COVID-19 cases are not easily attributable to imprecision. The lower bound of our 95% CIs from columns 1 and 2 in Table 1 suggest that we can statistically reject the possibility that SIP policies prevented 1.8 new cases per million residents per day or that they prevented 0.02 new deaths per million residents per day. If we multiply these lower-bound estimates with the population of each state and the number of days the policy was in effect in each state, we correspond with 32,112 cases and 372 deaths prevented nationwide during our period of study. Therefore, we can statistically reject the estimates from previous studies concluding that SIP policies saved many thousands of lives.

Fig. 2 graphically represents the results of our event study analyses based on the model in Eq. 2. For both cases and deaths, estimated effects are generally positive but insignificant. While we would not expect a SIP order to affect either of these outcomes instantaneously, the results indicate that there are no declines even after the policy has been in place for 14 or more days.

Fig. 2 does show statistically significant evidence of small, negative effects of SIP orders on mobility in the days immediately after the policy is enacted. By the end of the first week, however, changes in mobility appear to have returned to trend. Finally, Fig. 2 provides some evidence that SIP orders lead to increases in unemployment when they have been in place for 10 or more continuous days.

Substantively, Fig. 2 suggests that having two continuous weeks of a SIP policy increases the insured unemployment rate by one percentage point. The average insured unemployment rate was 1.3% at the beginning of our study period in late February and early March, so a one percentage point effect represents a notable increase over that baseline. However, by early May, the average insured unemployment rate reached 13.9%, meaning the one percentage point effect we estimate represents only a small share of the nationwide increase in unemployment during the pandemic, consistent with other studies showing that SIP policies accounted for a relatively small share of COVID-related economic declines (6, 7).

SI Appendix, Tables S1–S4 show more detailed results for each of our four outcomes. Some specifications include leading values of the treatment in order to test whether SIP orders tended to be implemented at times when the outcomes were already trending differently. These coefficients are generally close to zero and statistically insignificant, and their inclusion does not meaningfully change the other coefficients, suggesting that preexisting trends do not meaningfully bias our estimates. Some specifications include or exclude lagged values of the outcome, which also do not meaningfully affect the other estimates. And some specifications include lagged values of the outcome variable, allowing us to test how the effects of SIP policies vary as more days pass. The detailed results bolster the findings in Table 1 and Fig. 2. SIP orders have no detectable effect on COVID-19 cases or deaths; a small,

Fig. 2. Effects of SIP policies over time. The figure shows the estimated effects of having a SIP policy in place for 0 d to 14+ d in a row. Estimates that are statistically distinguishable from zero (P < 0.05) are red. We regressed the outcome of interest on state fixed effects, day fixed effects, lags of the dependent variable, our SIP policy variable, and 14 lags of the SIP policy variable. To estimate the effect of 2 d of SIP policies, for example, we add the coefficients associated with the SIP policy variable, the 1-d lag, and the 2-d lag, and we conduct an F test of the null hypothesis that this sum is equal to zero. Complete results for these regressions are shown in column 7 of SI Appendix, Tables S1–S4.
negative, short-lived effect on mobility; and a positive lagged effect on unemployment.

In SI Appendix, we also test additional hypotheses and explore the robustness of our results in various ways. We separately investigate the effects of expanding and withdrawing SIP orders, by focusing on different time periods. Interestingly, the introduction of SIP orders has bigger effects on mobility and unemployment than their withdrawal.

One concern with our empirical approach is that the effects of a SIP policy in one state could spill over to other states, biasing our estimates. Suppose, for example, that California’s SIP order affected behavior in Oregon. This could make the effect of California’s policy appear smaller than it really is, because, although we treat Oregon as a control state with respect to California, it was indirectly treated. In SI Appendix, we present several additional tests designed to address this concern. First, we implement our test only for smaller states, assuming that, to the extent that the effects of policies spill over to other states, this phenomenon will be more relevant for the policies of large states. And, to the extent that large states affected smaller states, this will be accounted for by the day fixed effects. Second, we implement our test only for states that implemented SIP orders later. Here, the idea is similarly that, to the extent there were spillover effects, they were probably more relevant for the early-acting states. Lastly, we explicitly estimate the extent to which having a SIP order in at least one neighboring state (as defined by a land border) may have affected behavior in that state over and above the state’s own policy. Although we do find some evidence that SIP policies affected mobility in neighboring states, the estimated effects on COVID-19 cases, mobility, and unemployment in that state are largely unchanged when we account for the possibility of spillovers between states.

SI Appendix also shows county-level estimates. Most of the variation in SIP orders comes from state policies, although some counties implemented their own orders before their states did so, and we obtain slightly more statistical precision by estimating effects at the county level. The estimates are extremely similar to those from our state-level analysis, although we do not examine unemployment, since weekly unemployment claims are only available at the state level.

We also account for changes in the extent of testing in various states. Specifically, we utilize information from Hsiang et al. (1) on the dates on which various states changed their testing regime, and, instead of including state fixed effects, we include state-by-testing-regime fixed effects. When we do this, the estimates are virtually unchanged.

SI Appendix also contains replications and extensions of two related studies attempting to estimate the effects of government policies on COVID-19 cases (1, 2). In both cases, we show that the estimated benefits of SIP policies and other similar government policies are highly sensitive to specification. Reassessing those designs and specifications leads us to conclude that previous studies have overestimated the effects of these policies. This helps to resolve an apparent contradiction in the literature. Since studies have shown that the effects of SIP policies on social distancing and mobility are small (5, 6), it would be surprising to find large effects on COVID-19 cases, and our analyses suggest that these effects are indeed small.

Concerns and Limitations. Our analysis produces no evidence that SIP orders led to substantial reductions in mobility, COVID-19 cases, or COVID-19–related deaths. We emphasize, however, that estimating the effects of SIP orders is challenging, and no estimation method is unassailable. The most notable challenge is that there are pretreatment trends in all the outcomes we study, as evident in Fig. 1 and in SI Appendix, Figs. S1 and S2. This means that the parallel trends assumption of a standard difference-in-differences design—that is, that the treated and untreated states would have followed the same pattern but for the imposition of the treatment—is suspect.

We might worry that policy makers implemented SIP orders precisely when they otherwise expected COVID-19 cases and deaths to increase in their state or county relative to the nationwide trends. While we cannot entirely rule out those concerns, we have compared multiple approaches to account for preexisting trends. We implement many different specifications with and without leads and lags of the treatment variable and with and without lagged values of the outcome (SI Appendix, Tables S1–S4). We separately estimate the effects of imposing and withdrawing SIP orders, and we account for spillover effects across states. In SI Appendix, we also present results that control for epidemiological projections prior to the implementation of SIP orders (12)—the same kinds of projections that likely informed the decisions of policy makers. We do not view any of these analyses as definitive in isolation. Rather, we emphasize that none of the modeling alternatives provides compelling evidence that SIP policies were effective.

We also emphasize that our estimates across different outcomes are logically consistent and reinforcing. In particular, we find that SIP policies had a small but short-lived effect on mobility, a finding that is consistent with other high-quality studies (e.g., ref. 6). If SIP orders did not have large effects on behavior, it is hard to imagine how they could have had large effects on COVID-19 cases and deaths. Furthermore, if unobserved differences do bias our estimates, we would expect, if anything, that our mobility estimates overstate the effect of SIP policies on mobility, but the fact that we detect small mobility effects suggests that the extent of this bias is likely small.

Finally, when we reexamine prior studies reporting large effects of SIP orders on COVID-19 cases and deaths (1, 2), we find that their results are not robust to reasonable alterations in modeling choices. For the reasons just enumerated, we conclude that the weight of the evidence suggests that SIP policies have not had large effects. Of course, we do not rule out the possibility that alternative approaches or additional information may produce different results. But, precisely because the policy implications are so important, null findings should be taken seriously. At the very least, our results suggest that policy makers should not begin with the presumption that SIP policies are known to be effective.

Conclusion

Although we estimate modest effects of SIP policies, our results should not be taken to imply that the actions of government officials had little effect on the pandemic. There may have been other policies that better mitigated the spread of COVID-19, although SIP orders have been arguably the most drastic and controversial policy. Furthermore, we observe nationwide trends in all outcomes, and these trends may have been highly responsive to the public health recommendations, emergency declarations, and the behaviors of high-profile politicians. Our results also do not mean that sheltering in place per se is an ineffective way to mitigate the spread of COVID-19. If SIP policies did not meaningfully increase the extent to which people actually sheltered in place or socially distanced, our results have nothing to say about the health and societal benefits of staying at home and reducing physical contact with others.

†Although state policy was surely influenced by beliefs about future trends in COVID-19 cases, many additional factors likely influenced these decisions, including the different priorities and values of government officials. For example, Bacini and Brodeur (11) find that Democratic governors were much more likely to implement SIP policies in the early period of the pandemic. Using our own data, we find that the median state with a Democratic governor implemented a statewide SIP order when there were 1.2 cases per 10,000 residents and 1.5 deaths per million residents, while the median state with a Republican governor waited until there were 4.1 cases per 10,000 residents and 8.0 deaths per million residents.
The explanation for our null findings is likely nuanced and multifaceted. One part of the explanation is that many people were already staying at home and social distancing voluntarily even in the absence of SIP policies. Another part of the explanation is, perhaps, that few people who weren’t already changing their behavior complied with the policies. After all, SIP orders appeared to cause less than a 1% decrease in mobility. There was, however, approximately a 50% decrease in mobility nationwide between February and April. The nationwide reaction to COVID-19 almost surely decreased the spread of the disease. SIP orders likely would have been more effective in slowing the spread had more people complied with them, and future SIP orders would likely be more effective if they are coupled with greater enforcement. But we find little evidence that SIP orders, as implemented, had much effect over and above all the other public messaging and voluntary behavior changes occurring nationwide. Although we find no detectable health benefits of SIP orders, we also find that they accounted for a small share of economic costs associated with the pandemic, consistent with other studies (6–9).

Our study is certainly not the last word on this topic. Assessing the effects of SIP orders is difficult, and more information and better designs may become available in the future that enable more precise or more credible estimates. Furthermore, our study focuses on the early months of the pandemic, and the effectiveness of SIP orders could change over time. However, the previously presented evidence on the effectiveness of SIP orders appears to be misleading, and there is currently no compelling evidence to suggest that SIP policies saved a large number of lives or significantly mitigated the spread of COVID-19. However, this does not mean that voluntary social distancing—SIP practice as distinct from policy—was ineffective.

Data Availability. Data have been deposited in Dataverse (DOI: 10.7910/DVN/JKSG8C).