Evaluating Policies Early in a Pandemic: Bounding Policy Effects with Nonrandomly Missing Data

Supplementary Appendix

Brantly Callaway* Tong Li†

July 5, 2022

This supplementary appendix contains additional details for (i) alternative estimation strategies to the doubly robust estimator used in the main text, (ii) some alternative strategies involving extra assumptions that can be used to deliver tighter bounds than the main bounds in the text (up to point identification), and (iii) additional results for the application on Tennessee's expanded testing policy.

S-A Alternative Estimation Strategies

In this section, we discuss an alternative estimation strategy based on matching. In particular, for each county in Tennessee, we find a "match county" from among its six surrounding states based on pre-policy county characteristics, Z_{lt^*-1}. Then, we compute an estimate of the ATT by taking the average outcome experienced by counties in Tennessee relative to average outcomes among their matches.

In practice, to construct the matched dataset, we match based on the same covariates used in the main text: the seven day lags of the cumulative number of tests, confirmed cases, and deaths per 1000 people in the county as well as the log of county population. We pick matches from comparison counties without replacement. To actually construct the matches, for county l in Tennessee, we construct its match by choosing the county in comparison states that minimizes the Mahalanobis distance (Rosenbaum and Rubin (1985))

$$d(Z_{lt^*-1}, Z_{jt^*-1}) = (Z_{lt^*-1} - Z_{jt^*-1})' \hat{S}^{-1} (Z_{lt^*-1} - Z_{jt^*-1})$$

where j indexes counties in comparison states and $\hat{S}$ is the estimated covariance matrix of Z_{jt^*-1} for the entire sample. Finally, we drop Shelby County (where Memphis is) due to not having a close match. This results in a matched dataset that includes 91 counties in Tennessee. For inference, we

*Department of Economics. University of Georgia. brantly.callaway@uga.edu
†Department of Economics. Vanderbilt University. tong.li@vanderbilt.edu

cluster standard errors by match along the lines of Abadie and Spiess (2022). Results using matching are provided in Figure S-4 below.

S-B Tighter Bounds under Additional Conditions

In this section, we develop tighter bounds on ATT_C relative to the ones in the main text under some additional assumptions. For this section, we maintain Assumptions 1 and 3 to 5 as well as Extra Conditions (i)-(iii), the primitive conditions to rationalize Assumption 5 in Appendix B. We take as a starting point the expression for $ATT_C(Z_{lt^*-1})$ in Equations (7) to (9). These conditions immediately imply that $\phi_1(Z_{lt^*-1}) \leq \phi_0(Z_{lt^*-1}) \leq \gamma_0(Z_{lt^*-1})/\tau_0(Z_{lt^*-1})$ which we use below. The first inequality means that the probability of having had Covid-19 among untested individuals in treated locations is less than or equal to the probability of having had Covid-19 among untested individuals in untreated locations conditional on locations having similar pre-treatment characteristics, and this holds by the same sort of arguments as in Appendix B. The second inequality holds by Assumption 1. The same conditions further imply that $\tau_1(Z_{lt^*-1}) \geq \tau_0(Z_{lt^*-1})$.

Next, we introduce a first assumption that can lead to tighter bounds. Along the lines of Remark 1, we consider an assumption that the fraction of untested individuals who have had Covid-19 is not "too different" between treated and untreated locations with similar pre-treatment characteristics.

Assumption S-B.1. *The difference between Covid-19 infection rates among untested individuals conditional on pre-treatment characteristics for untreated locations relative to treated locations is uniformly bounded with the bound given by*

$$\phi_0(Z_{lt^*-1}) - \phi_1(Z_{lt^*-1}) \leq \alpha$$

for some $\alpha \geq 0$.

Assumption S-B.1 does not affect the upper bound from Proposition 4, but it does increase the lower bound. In particular, the lower bound is now given by $\gamma_1(Z_{lt^*-1}) - \gamma_0(Z_{lt^*-1}) - \alpha(1-\tau_1(Z_{lt^*-1})) - \frac{\gamma_0(Z_{lt^*-1})}{\tau_0(Z_{lt^*-1})}(\tau_1(Z_{lt^*-1}) - \tau_0(Z_{lt^*-1}))$; this bound comes essentially immediately from the same types of arguments as in Proposition 4. This bound is generally tighter than the one in Proposition 4. And it is interesting to consider the limiting cases. When $\alpha = 0$, then the difference in the bounds arises because the difference between $\tau_1(Z_{lt^*-1})$ and $\tau_0(Z_{lt^*-1})$ is less than $1 - \tau_0(Z_{lt^*-1})$. On the other hand (and slightly abusing notation), if α were equal to $\gamma_0(Z_{lt^*-1})/\tau_0(Z_{lt^*-1})$ (which is its maximum possible value under the conditions in the main text), then these bounds reduce to the same ones in Proposition 4.

In practice, we note that it is likely to end up being quite *ad hoc* how to set a plausible value of α. On the one hand, as noted in the text, our largest bounds on policy effects occur in cases where the policy eliminates all Covid-19 cases among the untested, but the rate of actual Covid-19 cases would be the same for the untested as for the tested without the policy. In practice, this case seems unlikely

and choosing some smaller value of α can rule out this case. That being said, it also seems unlikely that $\alpha = 0$. In particular, α would tend to be be positive in cases where (i) there are some Covid-19 cases among the untested and either (ii) the policy expands testing and there is self-selection into the test (as is very likely to be the case in our application on Tennessee) or (iii) the policy actually reduces Covid-19 cases (as seems to be the case for our application on Tennessee and might be the case in other policy evaluation studies for early pandemic policies). Besides the two limiting cases, though, it seems difficult to interpret different values of α.

Finally, we consider an even stronger assumption, but one that leads to point identification.

Assumption S-B.2. *The probability of having had Covid-19 among the untested in treated locations is the same as the probability of having had Covid-19 among the untested in untreated locations conditional on pre-treatment characteristics. Moreover, this probability is constant across pre-treatment characteristics, and this probability is known. That is,*

$$\phi_1(Z_{lt^*-1}) = \phi_0(Z_{lt^*-1}) = \beta$$

for some β that is known to the researcher.

Under Assumption S-B.2, it follows immediately from Equations (7) to (9) that ATT_C is point identified, and it is given by

$$ATT_C = \mathbb{E}[\gamma_1(Z_{lt^*-1}) - \gamma_0(Z_{lt^*-1})|D_l = 1] + \beta\mathbb{E}[\tau_0(Z_{lt^*-1}) - \tau_1(Z_{lt^*-1})|D_l = 1]$$

Relative to Assumption S-B.1, this assumption implies that $\alpha = 0$ and that the probability of having Covid-19 among the untested is somehow known. This is likely to be a very strong assumption in practice, but in early-pandemic applications, it is likely to effectively be required in order to point identify causal effects of the policy on actual Covid-19 cases. One way that this sort of assumption can be operationalized in practice is the following. There exist estimates of the actual number of Covid-19 cases (as distinct from confirmed cases) for some locations at different points in time for the early part of the pandemic (e.g., Figure 1 in the main text). In the context of our application, for example, the IHME provides daily estimates of the actual number of cases at the state-level. Given this estimate (and given that it is constant across counties), one can back out the probability of having Covid-19 among the untested. In particular, along the lines of Equation (1), we have that

$$\begin{aligned}\mathrm{P}(C_{ilt^*} = 1|D_l = 1) = \mathrm{P}(C_{ilt^*} = 1|T_{ilt^*} = 1, D_l = 1)\mathrm{P}(T_{ilt^*} = 1|D_l = 1) \\ + \mathrm{P}(C_{ilt^*} = 1|T_{ilt} = 0, D_l = 1)\mathrm{P}(T_{ilt^*} = 0|D_l = 1)\end{aligned}$$

and every term on the right hand side of this equation is identified except for $\mathrm{P}(C_{ilt^*} = 1|T_{ilt} = 0, D_l = 1)$. However, if we replace the left hand side, $\mathrm{P}(C_{ilt^*} = 1|D_l = 1)$, with an existing estimate of the fraction of individuals that have had Covid-19, then we can recover $\mathrm{P}(C_{ilt^*} = 1|T_{ilt^*} = 0, D_l = 1)$

and set β equal to this value. Thus, under Assumption S-B.2, we can point identify the effect of the policy.

Next, we use this approach in the context of our application on Tennessee's expanded testing policy. On May 9, the IHME estimates that there had been 50,100 cumulative cases in Tennessee (for comparison, there had been 14,528 confirmed cases on the same date) which corresponds to 0.73% of Tennessee's population. As an aside, these estimates come from information about confirmed cases, hospitalizations, deaths, and seroprevalence studies;[1] see Barber et al. (2022) for a detailed discussion. This estimate is towards the lower end of the state-level bounds on per capita actual Covid-19 cases presented in Figure 7. If the IHME estimate is taken as the truth, then we can back out the the fraction of untested individuals who have had Covid-19; doing this, we calculate that 0.47% of untested individuals in Tennessee had had Covid-19 by May 9 (relative to 5.9% of tested individuals). Plugging this in for ϕ_1 and ϕ_0 in from Equations (7) to (9), we therefore estimate that the policy reduced the number of actual Covid-19 cases by 1.29 per 1000 people. Continuing to take the IHME estimate of the overall infection rate as being true, this suggests that the policy reduced the number of Covid-19 cases in Tennessee by about 15% relative to what they would have been without the policy.

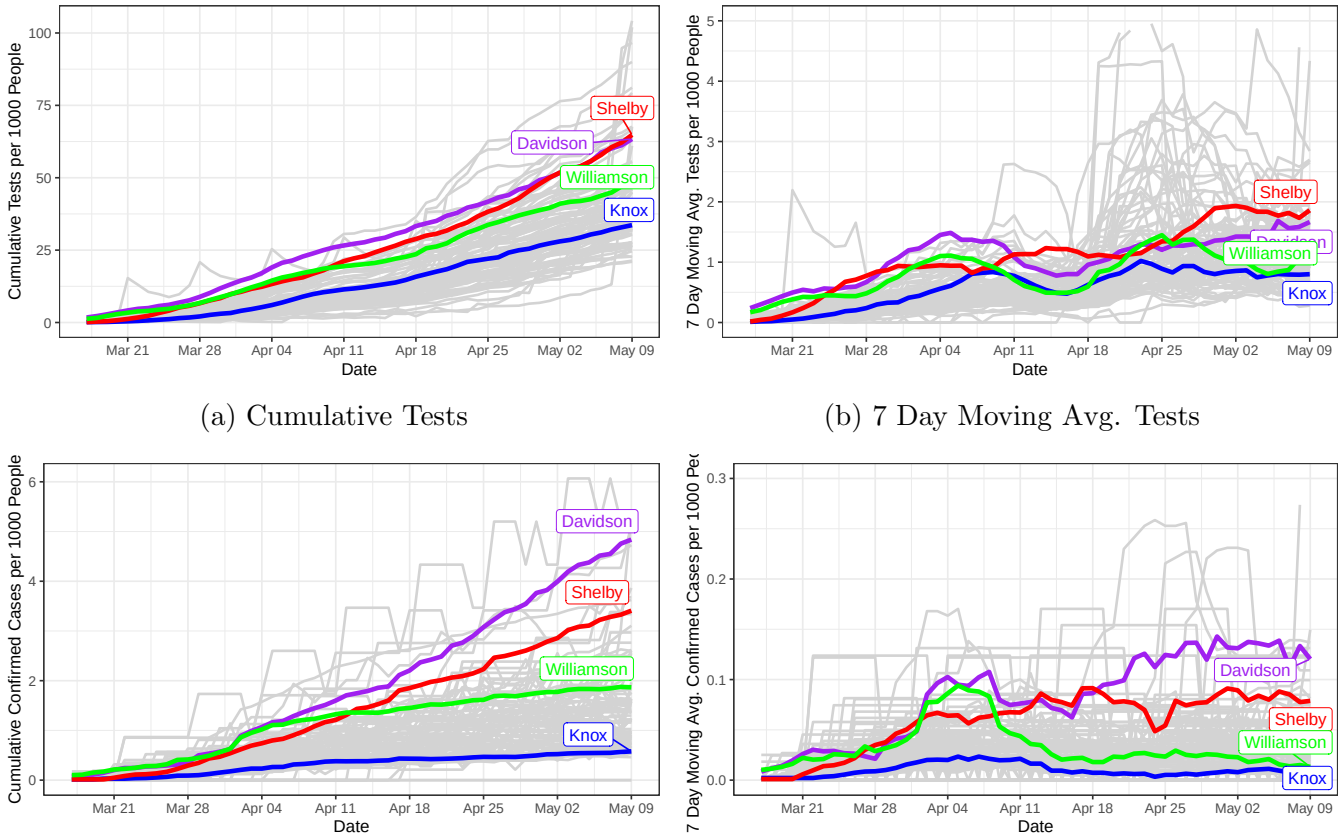
To conclude this section, we emphasize that the discussion above requires a number of extra assumptions that were not required for our bounding approach in the main text. In particular, as discussed above, it is unlikely that $\alpha = 0$ (which is required here). Even it were, it is additionally unlikely that β is the same across counties with different characteristics (in particular, it seems unlikely that β would be the same across counties with different values of Z_{lt^*-1} such as having had a much different number of Covid-19 cases and/or tests or for locations with much different populations). Moreover, the underlying data used to estimate the overall probability of having had Covid-19 is not likely to be very precise at the frequency needed to study state-level policies (due to pooling information across time and locations). The main advantage of this approach is that it delivers a point estimate of the effect of the policy on the number of actual Covid-19 cases. And, among different possible approaches that could deliver point identification, this one seems likely to be the most reasonable choice. However, if this were the only type of result reported on the effects of a policy early in the pandemic, it would not be clear whether the results were due to the actual effects of the policy or whether they were due to the extra assumptions from above. Thus, while our approach in the main text does not result in point identification, finding evidence that a policy does indeed reduce the number of actual Covid-19 cases without requiring these extra sort of conditions is likely to be substantially more credible.

[1]There are several state-level seroprevalence studies; probably most relevant is Bajema et al. (2021) though, to our knowledge, this same sort of information is not available at the county-level, and the sample sizes tend to be small for particular states (e.g., Bajema et al. (2021) provides monthly repeated cross sections data from July to September 2020 with close to 1000 observations collected in Tennessee in each cross section). See also Arora et al. (2021) for a detailed database of seroprevalence studies.

S-C Additional Results on Tennessee's Expanded Testing Policy

This section provides some additional descriptive figures regarding the data used in the application as well as a number of additional results for the application on Tennessee's expanded Covid-19 testing policy.

Figure S-1: Tests and Confirmed Cases by County in Tennessee



(a) Cumulative Tests

(b) 7 Day Moving Avg. Tests

(c) Cumulative Confirmed Cases

(d) 7 Day Moving Avg. Confirmed Cases

Notes: The figure provides cumulative and seven day moving averages of tests and confirmed cases for all counties in Tennessee excluding Bledsoe, Trousdale, and Lake. The highlighted counties are Davidson (county of Nashville), Shelby (county of Memphis), Knox (county of Knoxville), and Williamson (suburban county of Nashville).

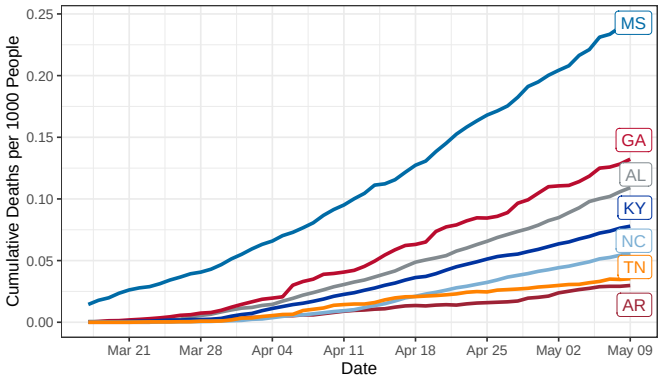
Sources: CDC COVID Data Tracker, COVID-19 Integrated County View

Figure S-1 provides county-level data about tests and confirmed cases for counties in Tennessee. This is analogous to the state-level plot of similar data in Figure 2 in the main text.

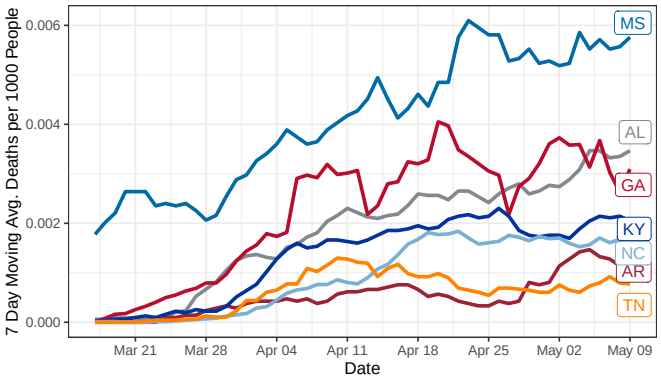
Next, Figure S-2 provides state-level data on Covid-19 deaths and hospitalizations.

Figure S-3 provides information about county-level deaths for counties in Tennessee.

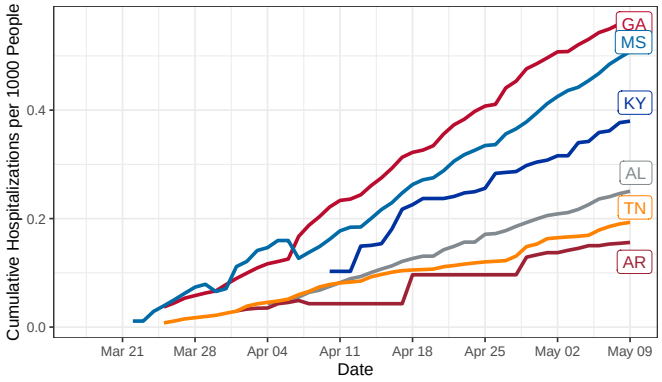
Figure S-2: Deaths and Hospitalizations by State



(a) Cumulative Deaths



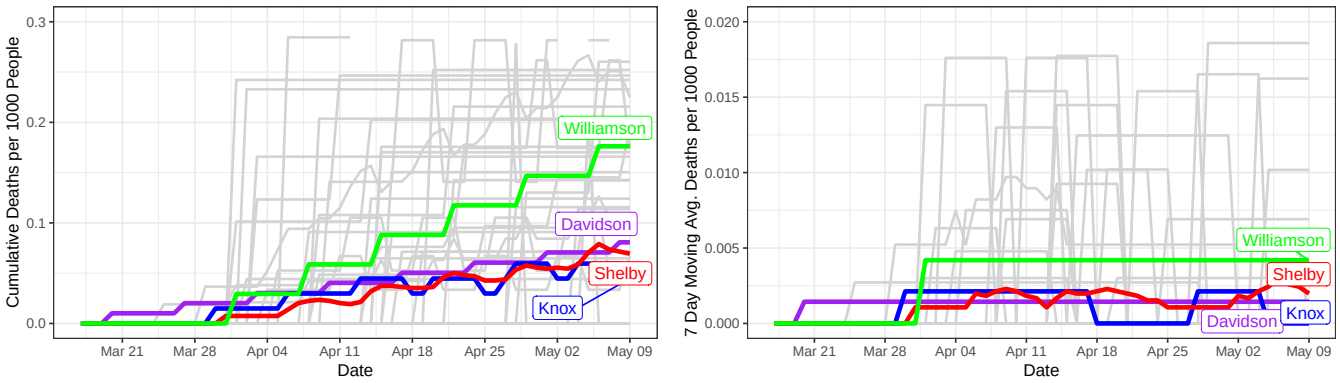
(b) 7 Day Moving Average Deaths



(c) Cumulative Hospitalizations

Notes: The figure provides (i) the cumulative and seven day moving average of deaths and (ii) cumulative hospitalizations for Tennessee, Alabama, Arkansas, Georgia, Kentucky, Mississippi, and North Carolina (hospitalizations are not available for North Carolina) from March 18 to May 9, 2020; not all dates are available for hospitalizations.
Sources: For deaths: CDC COVID Data Tracker, https://covid.cdc.gov/covid-data-tracker/#trends_totaldeaths. For hospitalizations, The COVID Tracking Project, https://covidtracking.com/

Figure S-3: Deaths by County in Tennessee



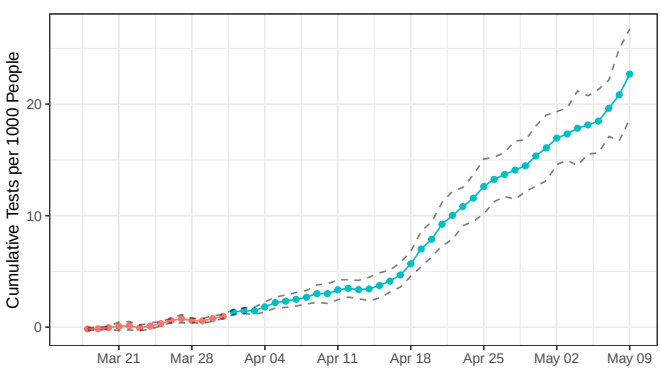
(a) Cumulative Deaths

(b) 7 Day Moving Average Deaths

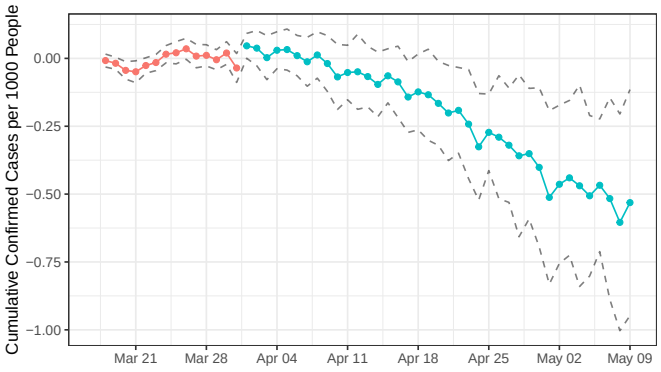
Notes: The figure provides cumulative and seven day moving averages of deaths for all counties in Tennessee excluding Bledsoe, Trousdale, and Lake. The highlighted counties are Davidson (county of Nashville), Shelby (county of Memphis), Knox (county of Knoxville), and Williamson (suburban county of Nashville).

Sources: CDC COVID Data Tracker, COVID-19 Integrated County View

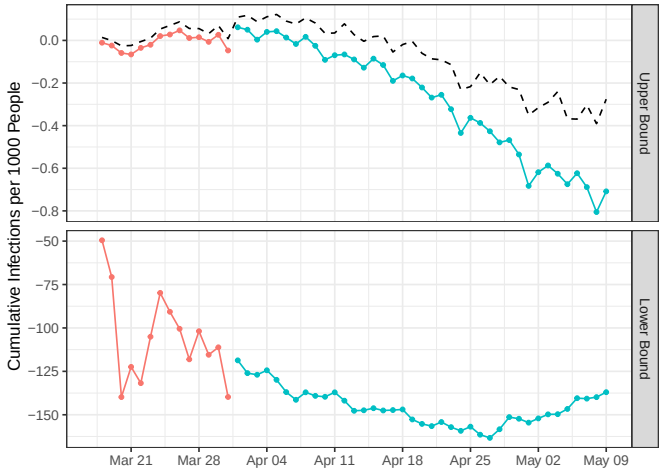
Figure S-4: Tests, Confirmed Cases, and Bounds on Actual Cases using Matching



(a) Cumulative Tests



(b) Cumulative Confirmed Cases



(c) Bounds on Cumulative Actual Cases

Notes: The figure provides estimates of the effects of Tennessee's expanded testing policy on cumulative tests per 1000 people, cumulative confirmed cases per 1000 people, and bounds on actual cases per 1000 people using the matching estimator described in Appendix S-A. The red points in the figure are estimates before April 1 while the blue points are for after April 1. The dashed line provides a 90% confidence interval.

Results using the matching estimator described in Appendix S-A are provided in Figure S-4. These results are very similar to the results using the doubly robust estimator described in the main text. In particular, the estimated effects of the policy on tests are virtually identical. The estimated effects on the number of confirmed cases is somewhat smaller here. For example, in the main text, we estimated that on May 9 there had been about 0.9 fewer confirmed cases per 1000 people under the policy than there would have been without the policy. Here, we estimate about 0.5 fewer cumulative cases per 1000 people though these estimates are not statistically different from each other. In both cases, the standard errors are somewhat larger using the matching estimator. This is not surprising as it effectively uses fewer observations from the untreated group of counties. The bounds are also broadly similar to the main text; relative to the main text, the bounds are somewhat tighter here though the qualitative results are the same as in the main text.

Next, consider the case where the policy's implementation date is set to be March 25 rather than April 1. Recall that, in the main text, we argued that it was not immediately obvious where to set the date of the policy due to tradeoffs regarding capturing the full effect of Tennessee's expanded testing policy, using valid comparison counties, and not leaving out useful comparison counties. The results in Figure S-5 are very similar to the ones reported in the main text. If anything, we estimate slightly larger effects on testing and virtually the same effects on confirmed cases and actual cases due to the policy in this case. This is in line with our arguments from the main text where setting the policy date earlier would tend to additionally capture the effects of the earlier (and smaller) part of Tennessee's policy.

Figure S-6 provides similar results but for the case where the policy's implementation date is set to be April 18 (the date where the open-testing policy was implemented) rather than April 1. For these results, our estimates of the effects of the policy, both on cumulative tests and confirmed cases, are smaller than when we set the implementation date earlier. This is in line with our expectations given the related discussion in the main text. For example, notice that there are some statistically significant effects of the policy in the pre-policy periods in this case. These effects are not included in the cumulative effects for dates after April 18. There are also somewhat tighter bounds in this case as well. That said, all of the results in this case are still qualitatively the same as above and in the main text: the policy is estimated to have increased the number of tests while decreasing the number of confirmed cases and actual cases.

Next, Figure S-7 provides a robustness check on our results from the main text by including the change in the number of confirmed cases in the prior week as an additional covariate. As discussed in the main text, our motivation for the particular covariates that we include comes from models of pandemics from the epidemiology literature. Some variations of these models, particularly ones that include location-specific unobserved heterogeneity, can suggest including the pre-treatment change in the number of cases as an additional covariate. These results are very similar to the ones reported in the main text.

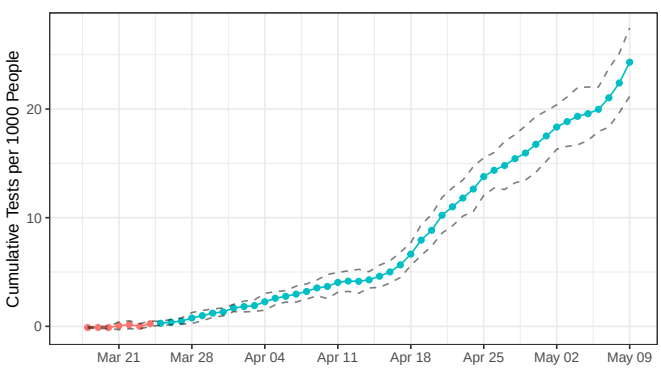
Figure S-8 provides an additional robustness check where we replace the log of county population with the county's population density as a covariate. The results are quite similar to the ones in the main text. Here, we estimate essentially the same effect of the policy on testing and marginally larger effects on the number of confirmed and actual Covid-19 cases than in the main text.

Next, we provide results that exclude counties from Arkansas from being in the comparison group in Figure S-9. Recall that, unlike all the other states that we consider, Arkansas did not implement a shelter-in-place order. The estimates that exclude Arkansas are very similar to the ones in the main text — the estimated effect of the policy on tests is virtually identical to that presented in the main text, and the effect on confirmed cases is slightly larger (i.e., more negative) here than in the main text. Likewise, the upper bound on actual cases is somewhat more negative (indicating a slightly larger estimate of the effect of the policy).

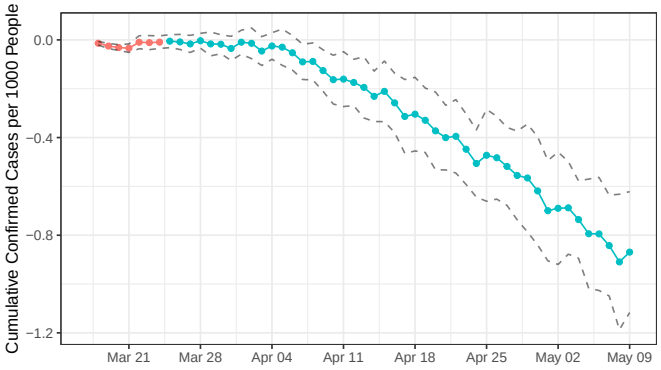
Finally, we provide estimates of the effects of the policy on deaths Figure S-10. We estimate large

effects of the policy on deaths: by May 9, we estimate that the policy had reduced the number of Covid-19 deaths by about 0.11 per 1000 people. These are quite large effects. For example, with a 1% infection fatality rate, it would indicate that the policy had reduced the number of actual cases by about 11 per 1000 people which is notably larger than our estimates above. We tend to favor the estimates using confirmed cases that we have emphasized in the main text though. The main reason is that the number of deaths per county is quite small this early in the pandemic (e.g., there had been 241 total deaths in Tennessee by May 9, and there are 95 counties in Tennessee). This same issue also suggests that these results are likely to be more sensitive to our needing to impute "suppressed" numbers of deaths at the county level (recall that 42% of the county-level data on deaths is suppressed in our data which indicates that the seven-day moving average of deaths was positive in that county at that point in time but small enough to not be included for privacy reasons). That being said, at least qualitatively, these results are in line with those from the main text and suggest that Tennessee's policy decreased the number of Covid-19 cases in Tennessee.

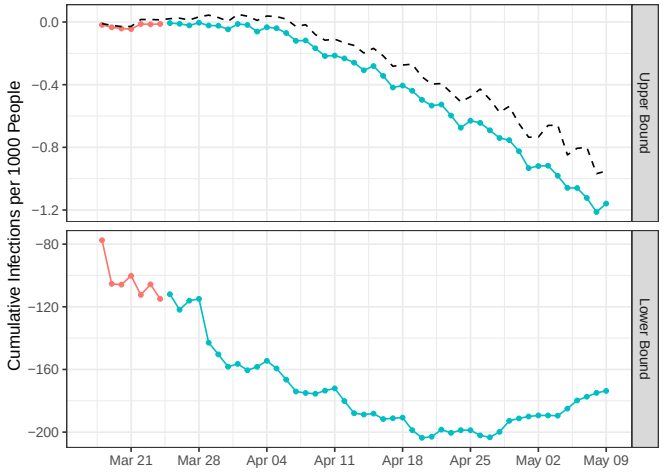
Figure S-5: Tests, Confirmed Cases, and Bounds on Actual Cases using March 25 Policy Date



(a) Cumulative Tests



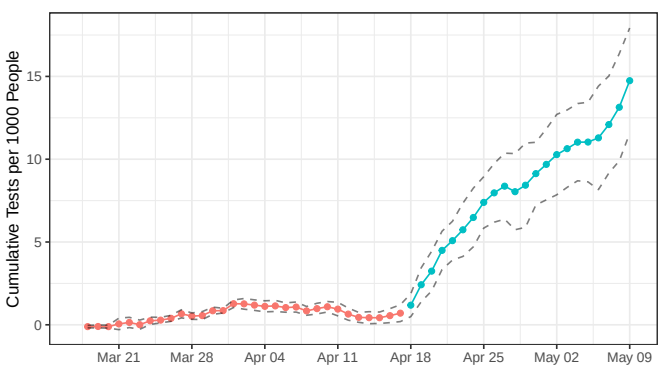
(b) Cumulative Confirmed Cases



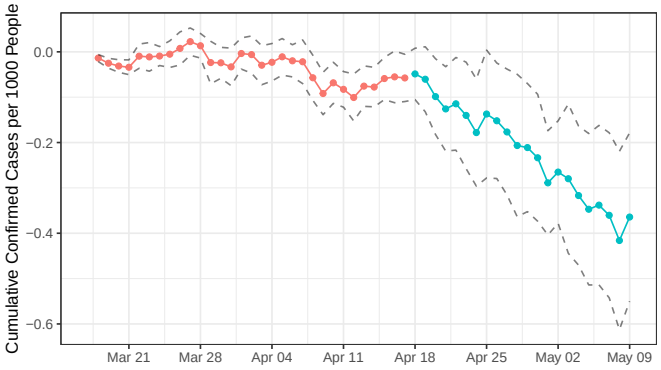
(c) Bounds on Cumulative Actual Cases

Notes: The figure provides estimates of the effects of Tennessee's expanded testing policy on cumulative tests per 1000 people, cumulative confirmed cases per 1000 people, and bounds on actual cases per 1000 people using the approach described in the main text while setting the policy date to be March 25 rather than April 1. The red points in the figure are estimates before March 25 while the blue points are for after March 25. The dashed line provides a 90% confidence interval.

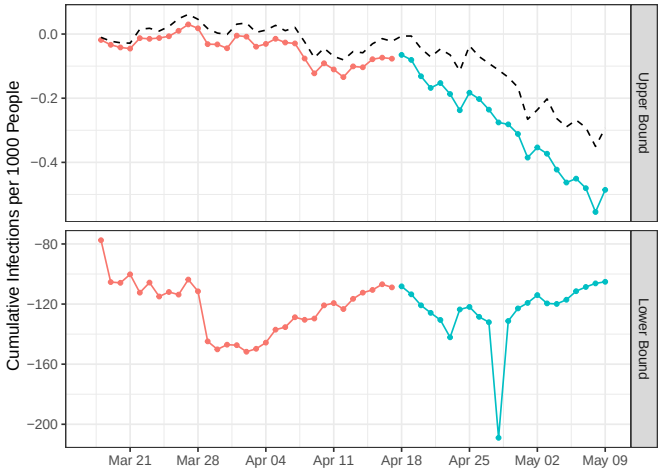
Figure S-6: Tests, Confirmed Cases, and Bounds on Actual Cases using April 18 Policy Date



(a) Cumulative Tests



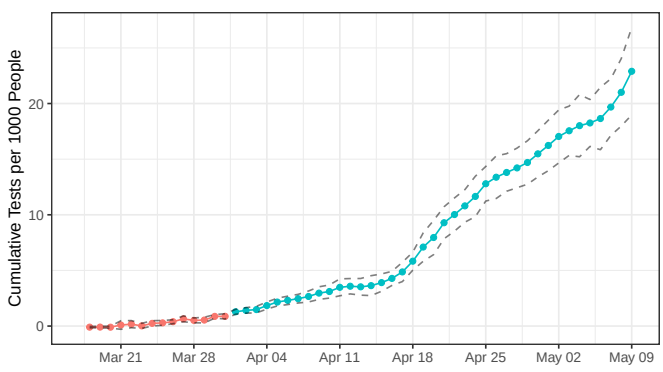
(b) Cumulative Confirmed Cases



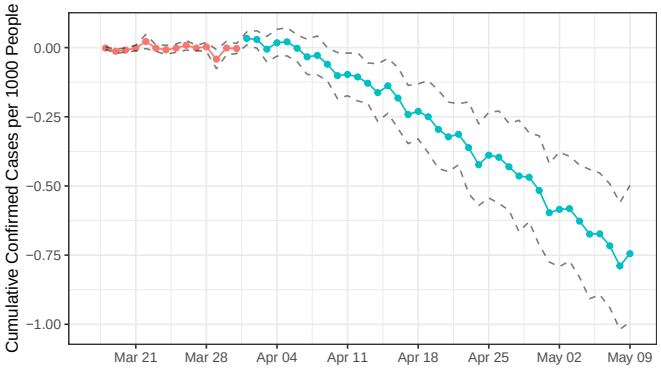
(c) Bounds on Cumulative Actual Cases

Notes: The figure provides estimates of the effects of Tennessee's expanded testing policy on cumulative tests per 1000 people, cumulative confirmed cases per 1000 people, and bounds on actual cases per 1000 people using the approach described in the main text while setting the policy date to be April 18 rather than April 1. The red points in the figure are estimates before April 18 while the blue points are for after April 18. The dashed line provides a 90% confidence interval.

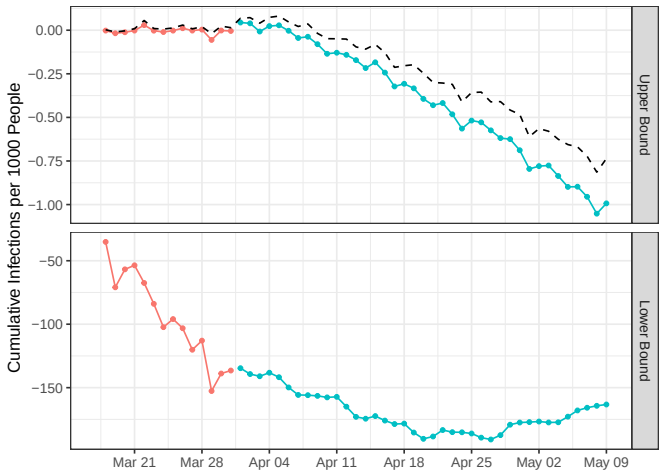
Figure S-7: Tests, Confirmed Cases, and Bounds on Actual Cases including Pre-Treatment Change in Confirmed Cases



(a) Cumulative Tests



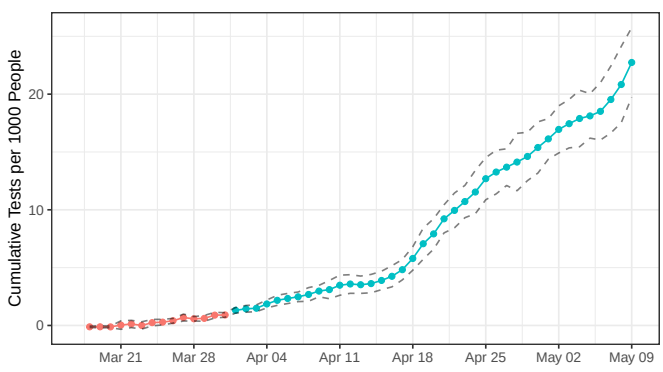
(b) Cumulative Confirmed Cases



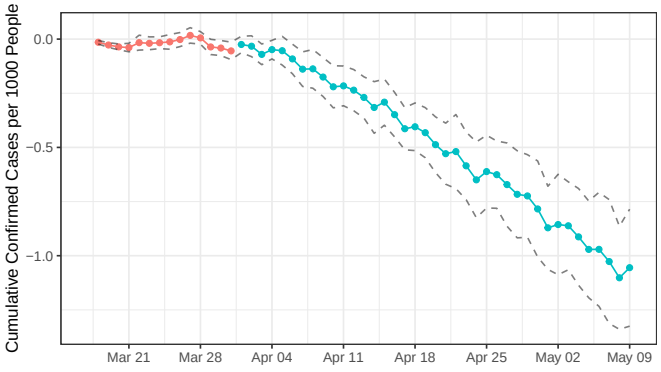
(c) Bounds on Cumulative Actual Cases

Notes: The figure provides estimates of the effects of Tennessee's expanded testing policy on cumulative tests per 1000 people, cumulative confirmed cases per 1000 people, and bounds on actual cases per 1000 people using the approach described in the main text except that it additionally includes the change in confirmed cases in the prior week as an additional covariate. The red points in the figure are estimates before April 1 while the blue points are for after April 1. The dashed line provides a 90% confidence interval.

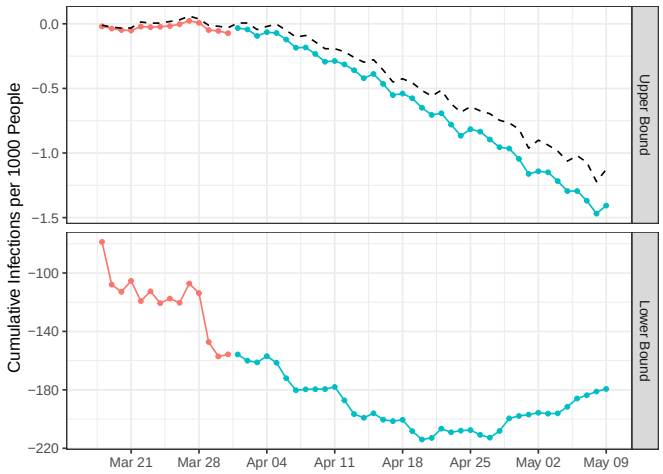
Figure S-8: Tests, Confirmed Cases, and Bounds on Actual Cases using Population Density rather than Log Population



(a) Cumulative Tests



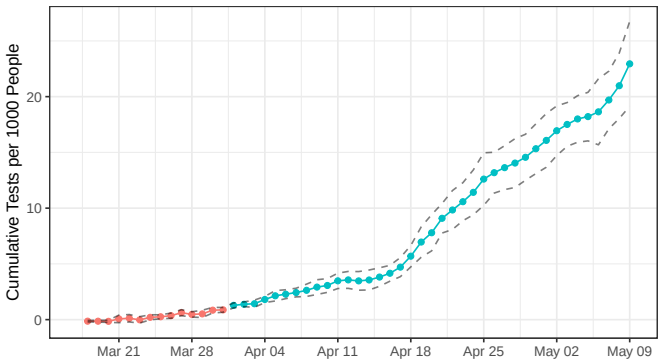
(b) Cumulative Confirmed Cases



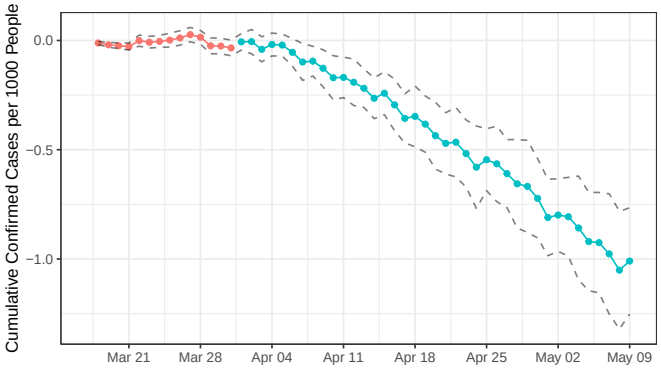
(c) Bounds on Cumulative Actual Cases

Notes: The figure provides estimates of the effects of Tennessee's expanded testing policy on cumulative tests per 1000 people, cumulative confirmed cases per 1000 people, and bounds on actual cases per 1000 people using the approach described in the main text except that it uses population density as a covariate rather than the log of county population. The red points in the figure are estimates before April 1 while the blue points are for after April 1. The dashed line provides a 90% confidence interval.

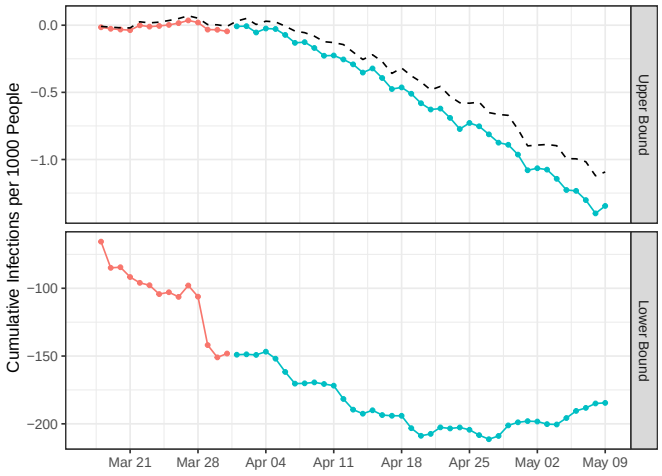
Figure S-9: Tests, Confirmed Cases, and Bounds on Actual Cases without Arkansas



(a) Cumulative Tests



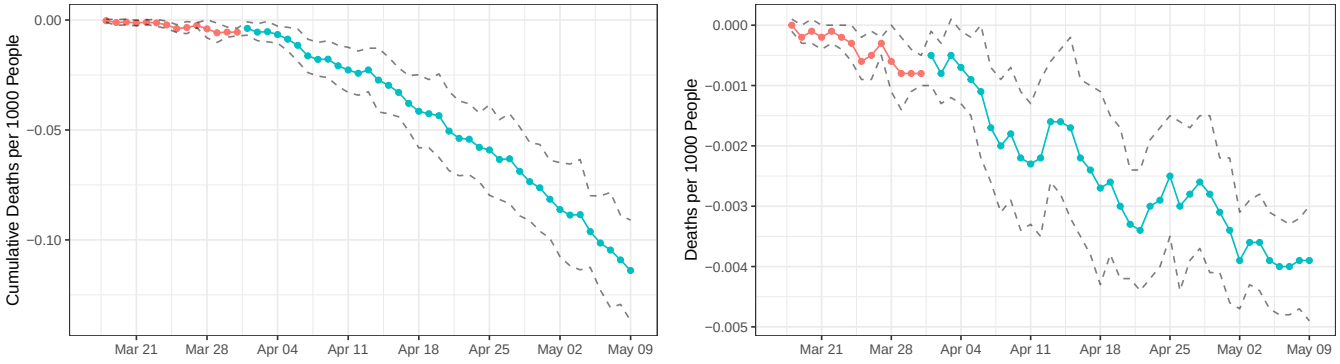
(b) Cumulative Confirmed Cases



(c) Bounds on Cumulative Actual Cases

Notes: The figure provides estimates of the effects of Tennessee's expanded testing policy on cumulative tests per 1000 people, cumulative confirmed cases per 1000 people, and bounds on actual cases per 1000 people using the approach described in the main text except that the results in this figure exclude counties in Arkansas from being in the comparison group. The red points in the figure are estimates before April 1 while the blue points are for after April 1. The dashed line provides a 90% confidence interval.

Figure S-10: Policy Effects on Deaths



(a) Cumulative Deaths

(b) 7 Day Moving Avg. Deaths

Notes: The figure provides estimates of the effects of Tennessee's expanded testing policy on cumulative deaths per 1000 people and the seven day moving average of deaths per 1000 people using the approach described in the text. The red points in the figure are estimates before April 1 while the blue points are for after April 1. The dashed line provides a 90% confidence interval.

References

[1] Abadie, Alberto and Jann Spiess. "Robust post-matching inference". *Journal of the American Statistical Association* 117.538 (2022), pp. 983–995.

[2] Arora, Rahul K, Abel Joseph, Jordan Van Wyk, Simona Rocco, Austin Atmaja, Ewan May, Tingting Yan, Niklas Bobrovitz, Jonathan Chevrier, Matthew P Cheng, et al. "SeroTracker: A global SARS-CoV-2 seroprevalence dashboard". *The Lancet Infectious Diseases* 21.4 (2021), e75–e76.

[3] Bajema, Kristina L, Ryan E Wiegand, Kendra Cuffe, Sadhna V Patel, Ronaldo Iachan, Travis Lim, Adam Lee, Davia Moyse, Fiona P Havers, Lee Harding, et al. "Estimated SARS-CoV-2 seroprevalence in the US as of September 2020". *JAMA Internal Medicine* 181.4 (2021), pp. 450–460.

[4] Barber, Ryan M, Reed JD Sorensen, David M Pigott, Catherine Bisignano, Austin Carter, Joanne O Amlag, James K Collins, Cristiana Abbafati, Christopher Adolph, Adrien Allorant, et al. "Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: A statistical analysis". *The Lancet* (2022).

[5] Rosenbaum, Paul R and Donald B Rubin. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score". *The American Statistician* 39.1 (1985), pp. 33–38.