Supplementary Materials

Quinn White

2023-08-26

Contents

Specification of the Priors **2**

Details on Each Implementation **2**

Full Set of Simulation Intervals by State **4**

Testing Rate (2021-06-18 - 2021-07-01) **5**

COVID-19 Trends and Impact Survey **6**

Comparing Lag Between Probabilistic Bias Estimates and Covidestim Infections **9**

County Results by Implementation **11**

Variant Data **12**

Specification of the Priors

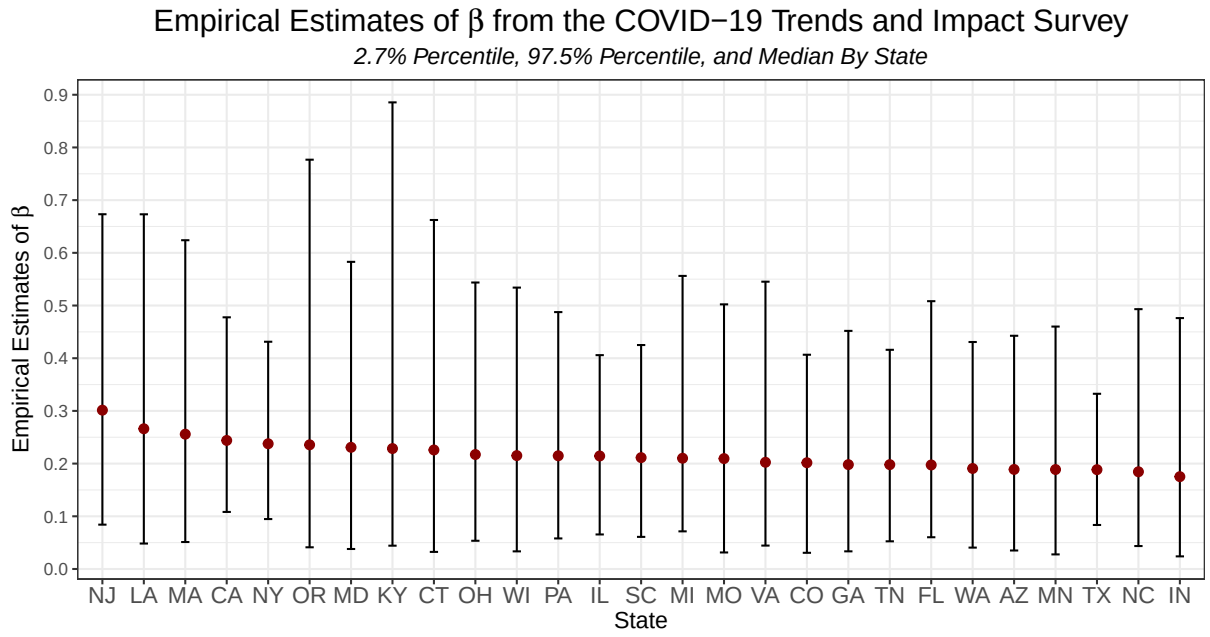

Figure 1: Since β represents the ratio of $\Pr(\text{test}_+|\text{untested}, S_0)$ to $\Pr(\text{test}_+|\text{tested})$, we estimate β from the COVID-19 Trends and Impact Survey as the ratio of the screening test positivity over the overall test positivity. Here, we consider the 2.5% and 97.5% percentiles and mean for each state. Most have a mean near 0.20.

Details on Each Implementation

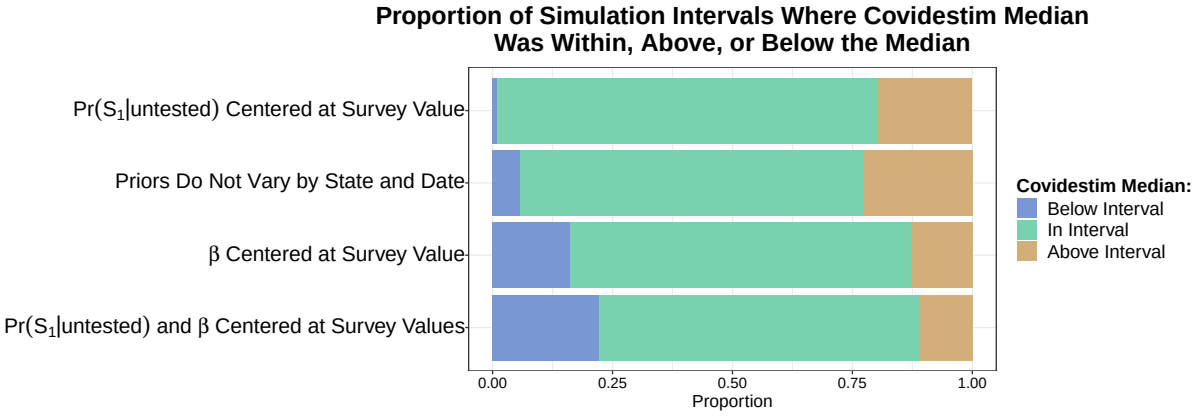

Figure 2: For each implementation, the proportions of simulation intervals where the Covidestim median was below, within, or above the interval are shown, across all states used for the implementation.

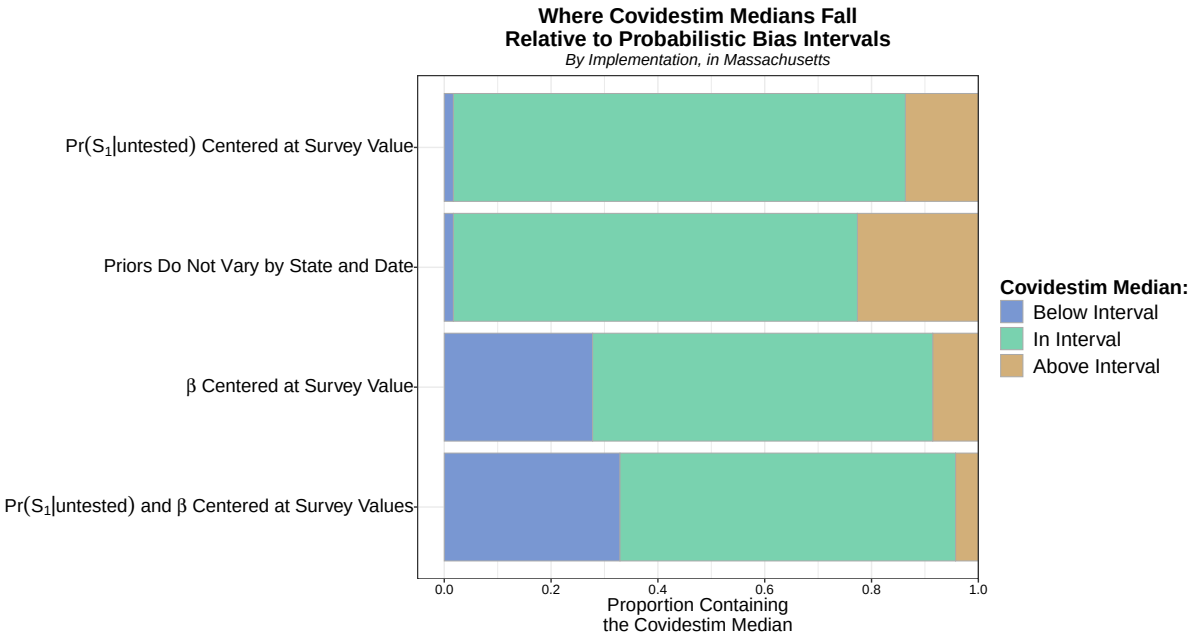

Figure 3: The proportions of simulation intervals where the Covidestim median was below, within, or above the intervals are shown, considering all counties except Dukes/Nantucket, where Covidestim estimates were not produced.

Full Set of Simulation Intervals by State

Testing Rate (2021-06-18 - 2021-07-01)

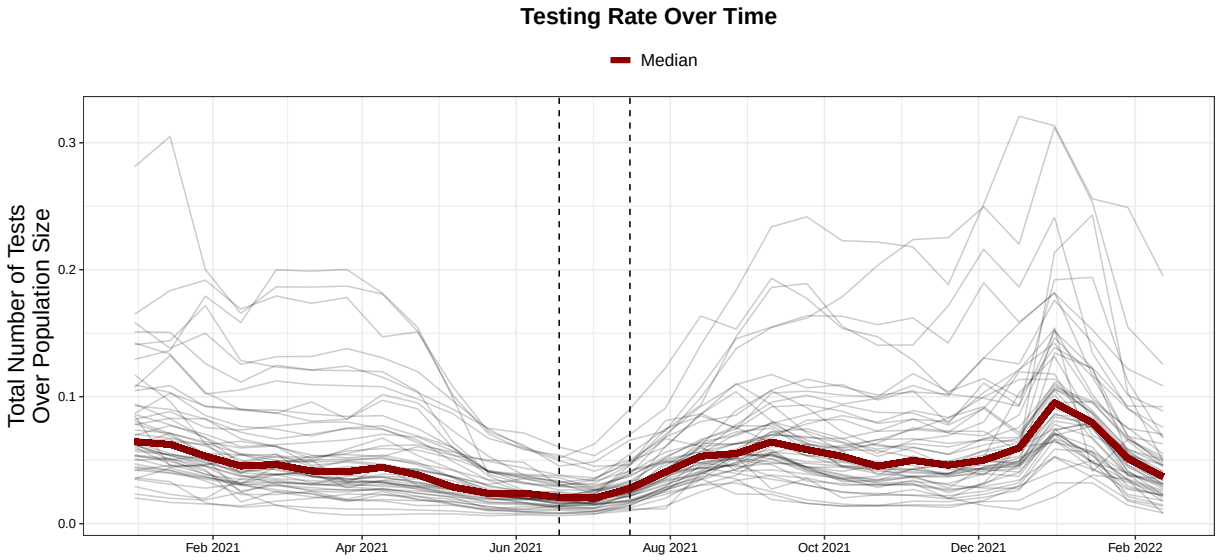

Figure 4: Testing rate, calculated as the total number of tests for a 2-week interval in a state over the census population in that state, over time. The median across states for each two-week interval is shown in red. In relation to Figure 1 in the main text, where we see that the period with the highest ratio of estimated to observed infections is June 18, 2021 through July 1, 2021, we see here that testing rates across states are at a minimum during this time period.

COVID-19 Trends and Impact Survey

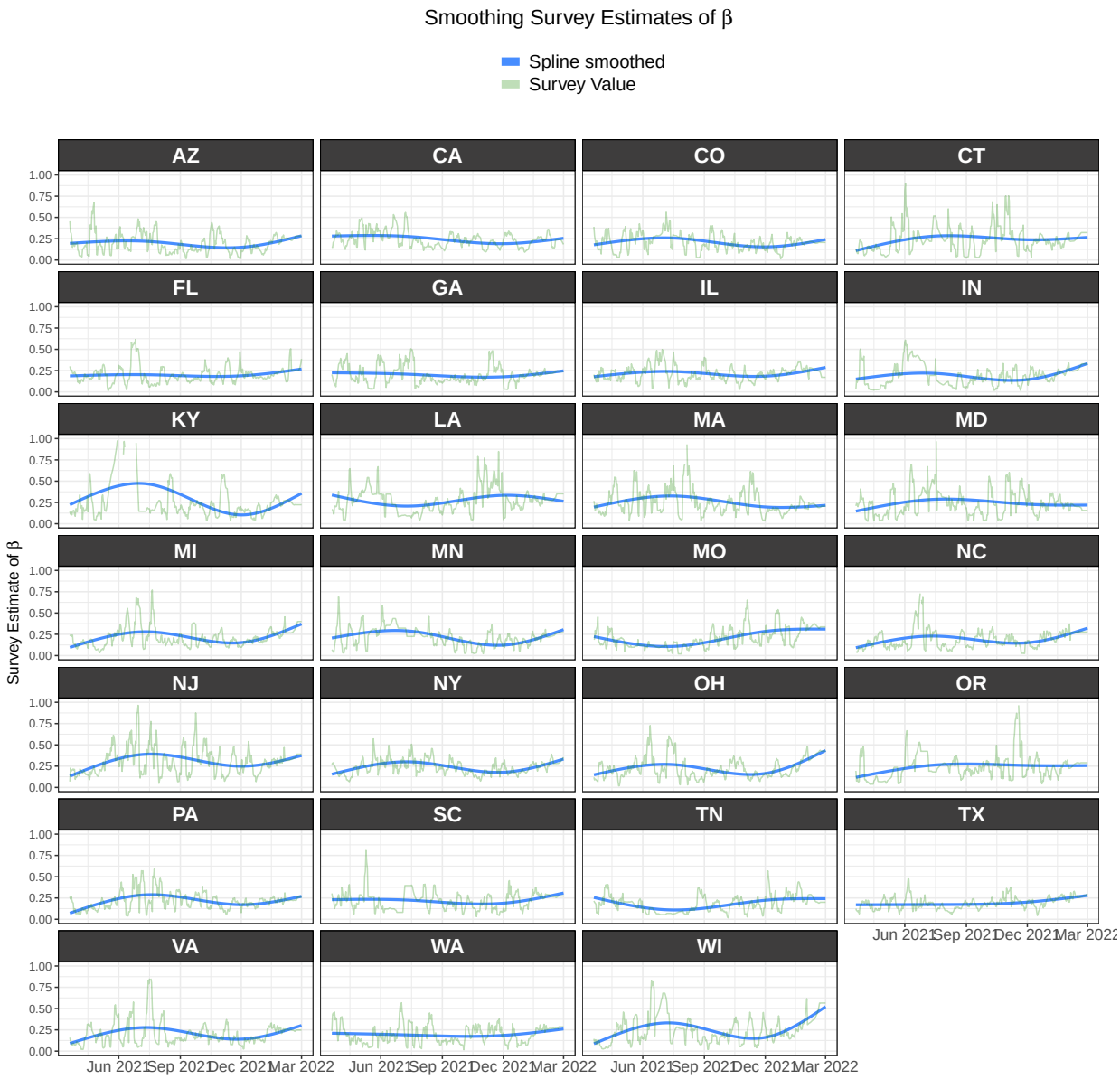

Figure 5: The ratio of the screening test positivity over the overall test positivity from the COVID-19 Trends and Impact Survey is taken to be the estimate of β. Because of the noise present inthe data, we use the spline smoothed values to inform the priors. More specifically, we use cubic spline smoothing with 2 knots (July 15th, 2021 and December 1st, 2021). While there are estimates of $\Pr(S_1|\text{untested})$ for all states (seen in the following figure), there are 27 states with estimates of β.

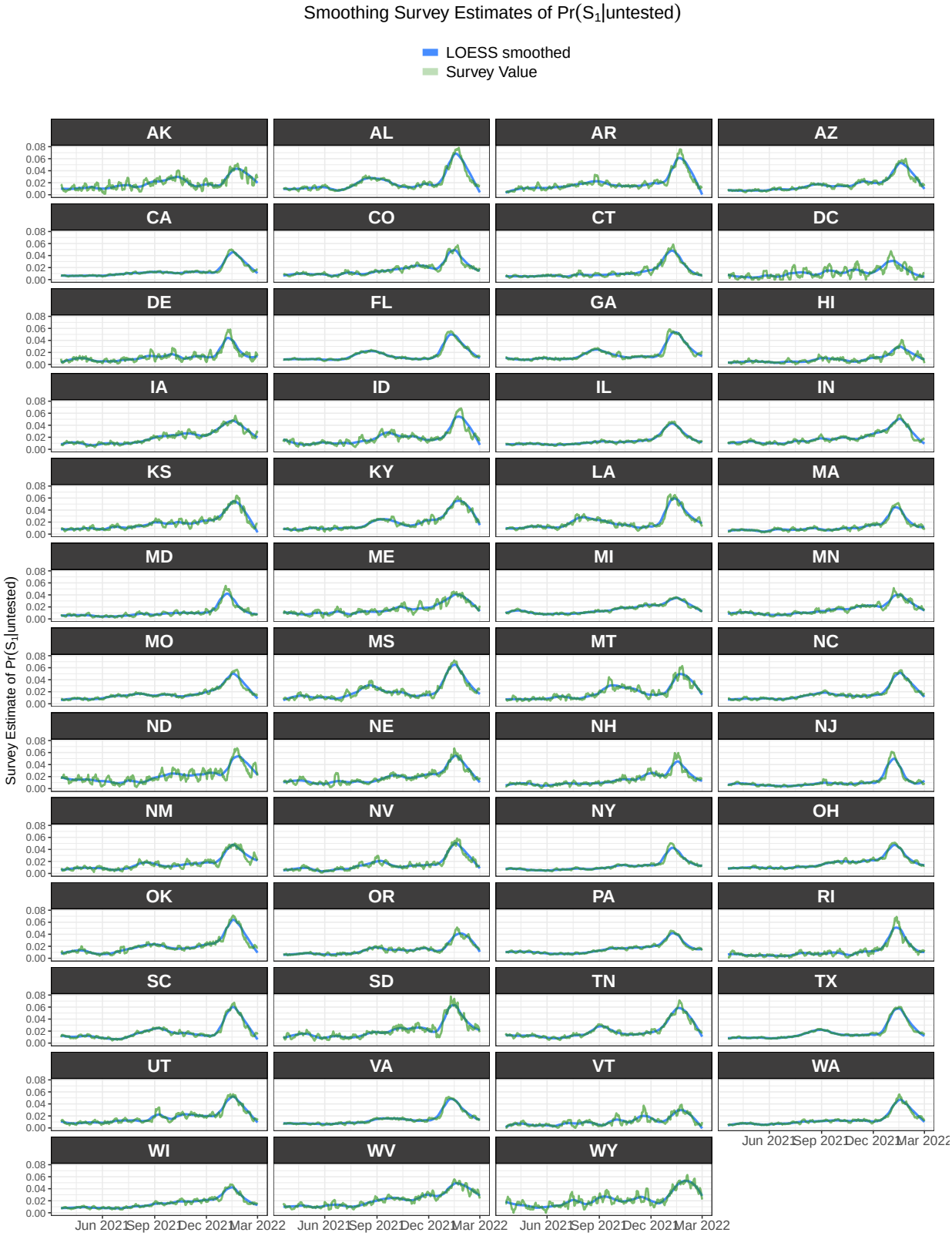

Figure 6: The percentage of the population experiencing COVID-19-like illness is taken to be the estimate of $\Pr(S_1|\text{untested})$. The LOESS smoothed estimate with a span of 0.2 is shown in red. Survey data on this variable is available for all states.

Comparing Lag Between Probabilistic Bias Estimates and Covidestim Infections

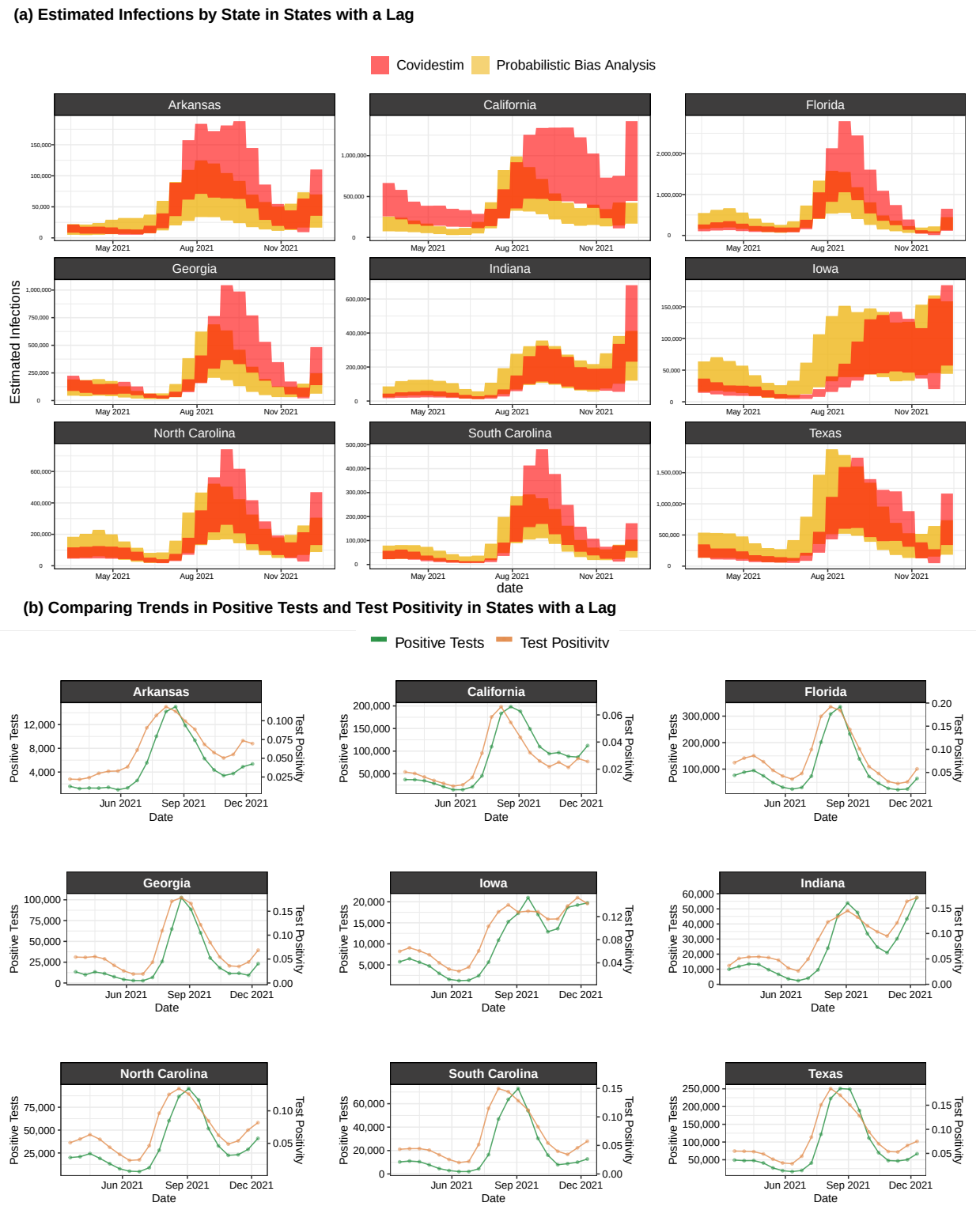

Figure 7: States where the Covidestim infection estimates are lagged relative to the probabilistic bias estimates, that is, they are most highly correlated when the Covidestim infections are lagged by one or more biweeks. (a) Simulation intervals demonstrating the lag seen around the delta peak in summer 2021. (b) Test positivity and positive tests for each state. Increases in test positivity tend to precede increases in positive tests, and these increases in test positivity result in increases in the estimates probduced by the probabilistic bias analysis seen in part (a).

County Results by Implementation

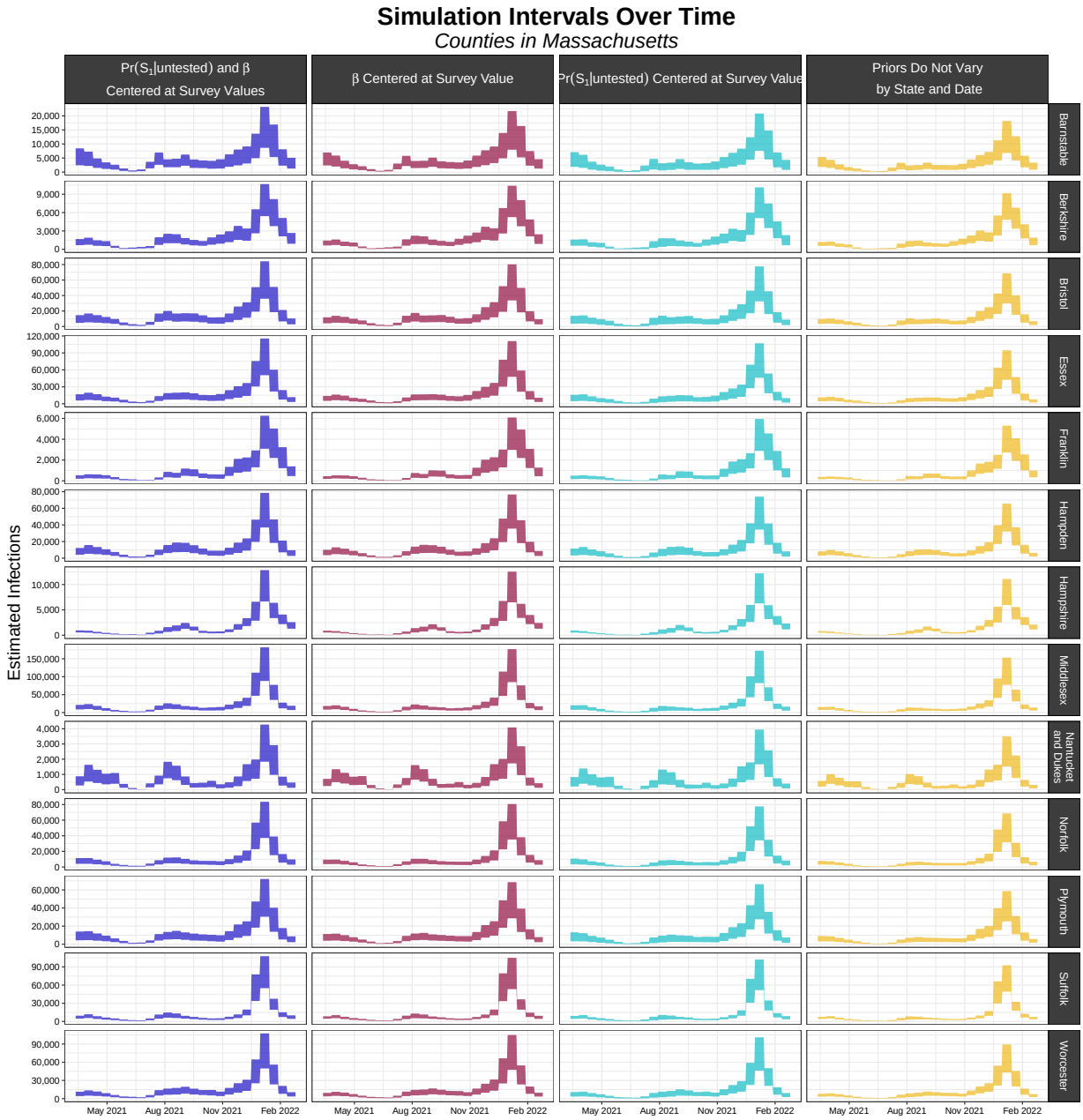

Figure 8: Simulation intervals for counties in Massachusetts. Each interval corresponds to a 95% simulation interval for the total number of estimated infections for that county in that two-week time interval. The columns represent different implementations of the probabilistic bias analysis: in the first through the third columns, the priors vary by state and date, while for the fourth, they are the same for all states and time intervals considered. For the implementation in the first column, we center the distribution of β at the ratio of the screening test positivity to the overall test positivity from the survey, and we center the distribution of $\Pr(S_1|\text{untested})$ at the percentage of the population experiencing COVID-19-like illness from the survey for each two-week interval. The second column centers only β at the aformentioned value, and the third column only centers $\Pr(S_1|\text{untested})$ at the aforementioned value. The fourth column corresponds to the implementation where we specify priors that are the same for all dates. The implementation that centers both $\Pr(S_1|\text{untested})$ and β at the survey values is consistently the highest among the implementations, followed by the implementation that centers only β at the survey value, followed then by the implementation only centering $\Pr(S_1|\text{untested})$ at the survey value, and then the lowest among the implementations is that where the priors do not vary by date.

Variant Data

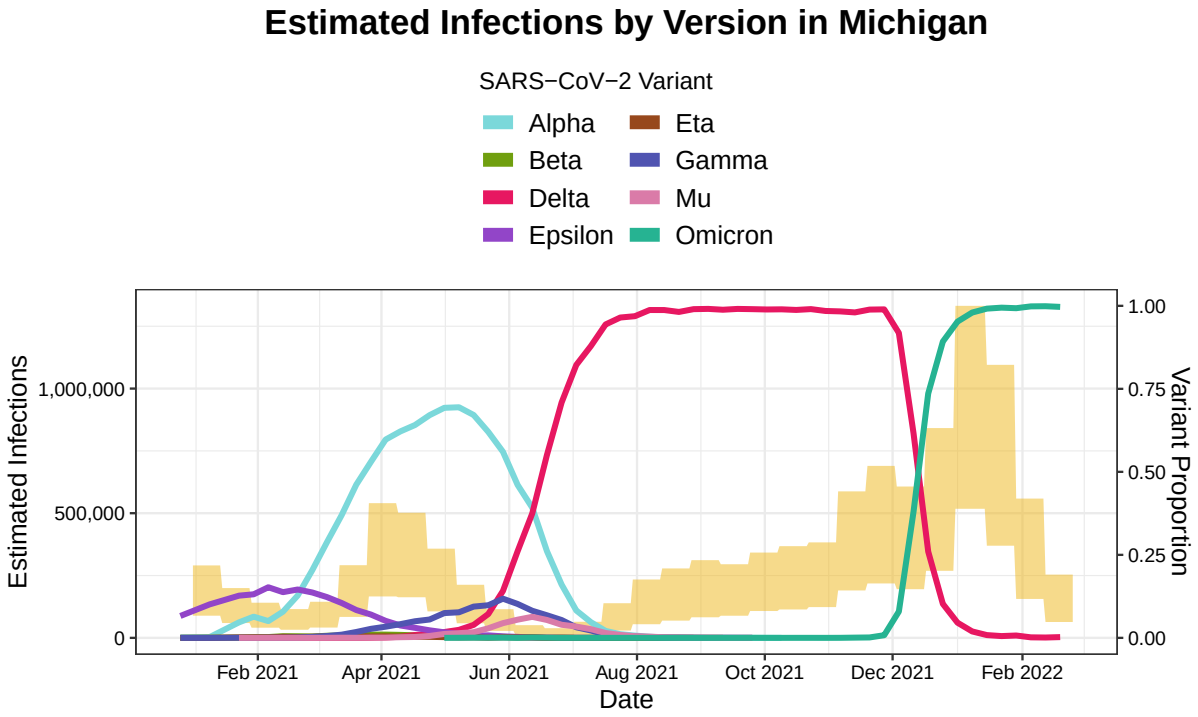

Figure 9: Simulation intervals for Michigan along with the variant proportions for variants designated as *Variant of Interest* or *Variant of Concern* by the World Health Organization. These variant proportions are not specific to Michigan; the estimates are for the United States. The rise in Omicron corresponds to the dramatic increase in both the probabilistic bias intervals and Covidestim estimates in December of 2021 through January of 2022. We also see the increase in the Delta variant that precedes the Delta wave in the summer of 2021.