Review of: "A Multi-factor Model of COVID-19 Epidemic in California"

Rand Ressler[1]

1 Georgia Southern University

Potential competing interests: No potential competing interests to declare.

Referee Report on

“A Multi-factor Model of COVID-19 Epidemic in California”

Other than the Kermack-McKendrick SEIR model, there is no review of the literature other than the assertion that “many models have been developed….” Later we are told “More elaborate models incorporated additional factors….” Without a reasonably thorough review of the literature, this paper is incomplete.

I think the results are being driven by one variable: population This makes sense with the number of COVID-19 cases in the observed county as the dependent variable. A better approach is to use the (number of COVID cases)/(1,000 county residents) as the dependent variable.

Income is a good explanatory variable since it is highly correlated with both the level of education and access to healthcare. As such, I would expect Income to be inversely related to per capita COVID cases.

Population Density is a well-known catalyst for higher contraction rates and an accelerated spread of infections. I would expect this variable to have a positive impact on cases per capita.

The Gini Coefficient is an interesting (and unusual) choice for an independent variable. More typically, the poverty rate is used instead, since income is also an explanatory variable. It might be informative to try an estimation with the poverty rate and compare it to the estimates generated by using the Gini Coefficient.

I do not understand why Land Area is included since the model includes population density. Other than it being easy to include, what is the motivation here?