Revisiting Pediatric Covid-19 cases in Counties With and Without School Mask Requirements—United States, July 1—October 20 2021.

Ambarish Chandra, Ph.D.*
Tracy Beth Høeg, Ph.D.

Abstract
Background: There has been considerable debate around mask requirements in schools in the United States and other countries during the Covid-19 pandemic. To date, there have been no randomized controlled trials of mask requirements in children. All analyses of the effectiveness of school mask mandates have relied on observational studies.

The Centers for Disease Control in the U.S. have released multiple observational studies suggesting that school mask mandates significantly reduce case rates. However, there have also been numerous additional US and international observational studies finding no significant effect of school mask mandates on pediatric cases.

Methods: Our study replicates a highly cited CDC study showing a negative association between school mask mandates and pediatric SARS-CoV-2 cases. We then extend the study using a larger sample of districts and a longer time interval, employing almost six times as much data as the original study. We examine the relationship between mask mandates and per-capita pediatric cases, using multiple regression to control for differences across school districts.

Findings: Replicating the CDC study shows similar results; however, incorporating a larger sample and longer period showed no significant relationship between mask mandates and case rates. These results persisted when using regression methods to control for differences across districts.

Interpretation: School districts that choose to mandate masks are likely to be systematically different from those that do not in multiple, often unobserved, ways. We failed to establish a relationship between school masking and pediatric cases using the same methods but a larger, more nationally diverse population over a longer interval. Our study demonstrates that observational studies of interventions with small to moderate effect sizes are prone to bias caused by selection and omitted variables. Randomized studies can more reliably inform public health policy.

Funding: University of Toronto

* Corresponding Author. University of Toronto, 105 St. George Street, Toronto, ON M5S 3E6, Canada. 647-517-1824. Ambarish.chandra@utoronto.ca

Introduction

There is ongoing scientific debate around mask requirements in schools in the United States and other countries during the Covid-19 pandemic. To date, there have been no randomized controlled trials of mask requirements in children. All analyses of the effectiveness of school mask mandates have relied on observational studies. Though some of these studies report a negative association between mask wearing and case rates, others fail to identify any association.

Studies of this subject face the challenge of controlling for fundamental differences between districts and communities that choose to require or not require masking in schools. For example, a study released by the CDC from Arizona and another U.S. study, found that mask requirements had large negative associations with COVID-19 outbreaks and in-school transmission, respectively.[1,2] In each case, these findings were likely confounded by crucial differences, such as the number of instructional days or contact tracing policy differences, rather than being driven by the effects of masks themselves.

In addition, these studies were inconsistent with more rigorously designed observational studies[3,4] as well as a study using a regression discontinuity design which reported no significant effect.[5] Furthermore, one randomized study in adults found no impact of community cloth masking and only a marginal impact of surgical masking on those over 50.[6] Another randomized study in adults failed to find a 50% reduction in acquisition of SARS-CoV-2 by the wearer of surgical cloth masks in community settings and was not powered to find smaller reductions.[7]

Considering millions of schoolchildren have been required to comply with mask mandates since the start of the pandemic and may be required to do so again in the future, it is incumbent upon us to determine if there is any effectiveness of masking children in schools against COVID-19 or other respiratory illnesses and, if so, that the benefits outweigh the harms.

Our study has two goals: First, we replicate and then extend, a highly cited observational study by the Centers for Disease Control on school mask mandates by Budzyn et al.[8] The second goal is to demonstrate that observational studies are prone to identifying spurious correlations, particularly when limited in population size, diversity, and duration.[9] This is important because studies identifying significant findings are more likely to be published[10] and public health interventions may be quickly adopted despite a potential lack of efficacy.

The recent MMWR publication by Budzyn et al finds that, following school openings in the Fall of 2021, pediatric cases of Covid-19 increased faster in U.S. counties that did not have mask mandates in schools, compared to those that did. While the results do not show a causal relationship, the study has been cited by the CDC in its policy recommendations and in the news media as evidence that mask mandates in schools can lower community spread. Our study replicates that analysis and then extends it, using a larger sample of school districts and a longer study duration.

Methods

We follow the methods from Budzyn et al as closely as possible. Data on pediatric and adult Covid-19 case rates, by county, were obtained from the October 25 release of the CDC's Restricted Case Dataset. Data on school enrollments and mask policies were obtained from the data company MCH, the same source used by Budzyn et al. County level demographic variables and school district to county mappings were obtained from the U.S. Census Bureau.

Following Budzyn et al, counties that met the following criteria were selected for the analysis: 1) a valid school start date and a known mask requirement for at least one school district in the county, 2) in districts that have made mask policies available, such policies apply uniformly to all students (counties containing districts with conflicting mask policies were discarded), 3) at least three weeks of data were available with seven full days of case data since the start of the 2021—22 school year (the median school start date across districts within a county was used).

Using September 4 as the cut-off, as chosen by Budzyn et al, these criteria resulted in the inclusion of 565 counties, compared to the 520 counties in that study. Using the more recent data release of October 25 resulted in a larger sample of 1832 counties. This study uses the smaller sample to enable a comparison with Budzyn et al, and the larger sample to evaluate robustness.

As in Budzyn et al, we conducted a multiple linear regression of per-capita pediatric case rates for these 1832 counties, including the following controls: median age, race and ethnicity, and population density; pediatric Covid-19 vaccination rates; adult transmission rates in the corresponding county and week; percentage uninsured and percentage living in poverty; social vulnerability index score; Covid-19 community vulnerability index score; and fixed-effects for each week after school opening.

Demographic variables were obtained from the U.S. Census Bureau. The SVI score was obtained from the CDC and the CCVI score from Surgo Ventures. Community transmission is defined as the per-capita rate of adult cases in the county in the corresponding week. The pediatric vaccination rate is defined as the fraction of children in the county ages 12—18 who received two doses of vaccine.

Following the empirical analysis, we use statistical methods to reconcile the difference between our results and those of Budzyn et al. We note various biases that can arise when using observational data on jurisdictions that choose whether to adopt particular public health recommendations. We explore potential cross sectional and temporal biases that can affect observational studies of this nature.

Role of the Funding Source: The funders of this study had no role in study design, data collection, data analysis, data interpretation, the writing of the manuscript or the decision to submit the manuscript for publication.

Results

We follow the methods of Budzyn et al and examine the relationship between school mask requirements and pediatric cases of Covid-19, replicating their analysis as closely as possible. Figure 1 displays results from 565 counties that fit the criteria used in that study and analyzes data from three weeks prior to schools opening to six weeks following opening. After two weeks of schools being open, which was the endpoint for Budzyn et al, the results align with the findings of that study: non-masking counties had, on average, around 30 additional daily pediatric cases per 100,000 children, compared to counties with mask mandates.

Thus, during the Fall of 2021, mask mandates were associated with lower pediatric cases of Sars-CoV-2 in districts where schools started by August 15. However, Figure 1 also shows that cases quickly declined in later weeks and did so faster in counties without mask mandates. In fact, the Budzyn et al study ended at exactly the peak of school case numbers for this sample of counties. By the end of the sixth week after reopening average pediatric case rates in counties without mask mandates were 7.53 per 100,000 while counties with mandates averaged 7.88 per 100,000; the difference was not statistically significant (p=0.9116).

Figure 2 shows a larger sample of 1832 counties, including counties which were not included in Budzyn et al's analysis due to later school district start dates. Note that there are 3142 counties in the United States, but many cannot be classified with regard to school mask policies, due to either insufficient data from MCH or the particular criteria used by Budzyn et al.

The larger sample of counties enables us to employ a much broader geographical representation of districts, in addition to increasing the population size of the study. This is important for full geographic representation as southern states are more likely to have early start dates. The figure shows the lack of a clear relationship between mask mandates and pediatric cases. Counties that required masks in schools saw slightly larger increases in cases in the weeks immediately before and after school opening, but by the second week after the start of school there was no statistical difference. At the end of week 6, cases per 100,000 capita were 22.0 in counties without mask mandates and 21.9 in counties with mandates (p=0.936).

Next, we estimated a multiple regression of pediatric case rates controlling for observable differences across counties. Following Budzyn et al, the regression includes the following county-level covariates: age, race, ethnicity, Covid-19 community transmission, population density, social vulnerability index score, Covid-19 community vulnerability index score, the fraction unemployed, the fraction living in poverty, and the pediatric Covid-19 vaccination rate.

Regression results are presented in Table 1. The first column of that Table uses the same specification as Budzyn et al and confirms the lack of association between school mask mandates and case rates: after accounting for covariates, pediatric cases in counties with mask mandates were slightly higher, by 1.27 weekly cases per 100,000, though the effect was not statistically significant (p=0.058).

In column 2 of Table 1 we show the effect of removing the COVID-19 community vulnerability index (CCVI). The CCVI is derived from, and highly correlated with, the social vulnerability index. Omitting this variable results in a significant positive association ($p<0.0001$) between school mask requirements and pediatric case rates.

Discussion

We successfully replicated the main result of the original study by Budzyn et al. However, our extension shows that the results do not hold in a larger sample of school districts, or even in the original sample of districts when extended to a longer time interval. Specifically, when we extend the sample to districts that began later in the Fall, which encompasses a much larger portion of the country, we find no association between mask requirements and pediatric cases. Even in the set of districts studied by Budzyn et al, the initial association between mask requirements and reduced case rates disappeared when we extended the analysis to a longer time interval.

Thus, using the same methods and sample construction criteria as Budzyn et al, but a larger sample size and expanded time frame for analysis, we fail to detect a significant association between school mask mandates and pediatric COVID-19 cases. The discrepancy between these findings and those of Budzyn et al is likely attributable to the inclusion of more counties, a larger geographic area and extension of the study over a longer time period. By ending the analysis on September 4, 2021, Budzyn et al excluded counties with a median school start date later than August 14, 2021. According to the MCH data, this heavily over-samples regions that open schools by mid-August including Florida, Georgia, Kentucky and other southern states. That study would not have incorporated data from New York, Massachusetts, Pennsylvania, and other states that typically start schools in September. While this does not necessarily bias the results, it calls into question whether the results of that study can be representative of the entire country and suggests at least one important geographic confounding variable affects observational studies of school-based mask mandates in the United States.

If mask mandates were assigned at random, as, for example, in Abaluck et al,[10] then the relationship between these mandates and case outcomes may have a causal interpretation, though that study also has important limitations.[11] But mask mandates are not only non-random, but likely to be correlated with unobserved factors in systematic ways, making it inappropriate to infer causality.

School districts that mandate masks are likely to systematically differ from those that do not for many reasons. First, the former are likely to invest in other measures to mitigate transmission and may differ by testing rates and practices. Second, the choices made by school districts reflect the attitudes and behavior of their community. Communities that are concerned about the spread of Covid-19 are also likely to implement other measures, even outside of schools, that may eventually result in lower spread within schools. Finally, the timing of public health interventions is likely to be correlated with that of private behavioral changes. Public health measures are introduced when case counts are high, which is precisely when community members are likely to react to media coverage and change their own behavior.

Our study also uses observational data and does not provide causal estimates either. However, there is an important difference: while the presence of correlation does not imply causality, the *absence* of correlation can suggest causality is highly unlikely, especially if the direction of bias can be reasonably anticipated.[12]

In the case of school mask mandates, the direction of bias can be anticipated quite well. Private behavioral changes are likely to be positively correlated with public health measures to reduce cases, both cross-sectionally and temporally.[13] Therefore, the bias in the estimated coefficient from a naive regression of case outcomes on public health measures will be negative. In other words, an analysis that omits the extent and timing of private behavioral changes, even one that controls for covariates, will tend to overstate the effect of mask mandates, and run the risk of spurious negative correlations between mask mandates and recorded cases.

To see this, consider estimating the following linear regression:

$$y_{it} = b_0 + b_1 x_{it} + b_2 u_{it} + e_{it}$$

where y_{it} denotes case rates in community i at time t, x denotes observed public-health measures such as mask mandates, and u denotes variables that can affect case rates, but are unobserved or difficult to measure, such as private behavioral changes in the population. If we expect that x and u have a correlation $\delta > 0$, but that u is negatively correlated with y, the bias in the estimated coefficient from a naive regression of case outcomes on observed public-health measures is $b_2\delta$ which is negative.[12]

Furthermore, we have demonstrated that removing redundant socioeconomic data from the regression analysis actually resulted in a significant positive association between school mask mandates and COVID-19 cases. No causality can be inferred with the increased cases, but this demonstrates how controlling multiple times for similar variables can alter results and including the CCVI may have strengthened the negative association identified by Budzyn et al.

An important limitation of our study is that it does not exclude some effectiveness of school mask mandates. We do however show why relying on observational studies can be misleading. Such studies will tend to systematically overstate, rather than understate, the effectiveness of interventions for two reasons: selection bias caused by particular types of jurisdictions choosing to implement interventions; and omitted variable bias, due to such jurisdictions also being likely to implement other, unobserved interventions. Furthermore, publication bias tends to lead to publication of studies with significant findings.[10] Certain journals may also only publish findings that fit their preference, as was the case with our analysis; our expanded version of the original Budzyn et al publication was not accepted for publication by MMWR despite using the same methods, but with an expanded population and time frame. This bias can lead to the published "science" being a self-fulfilling prophecy rather than an unbiased pursuit of truth.

In summary, expanding upon a widely cited CDC study, we fail to find the same evidence that school mask mandates are associated with a reduction in county pediatric COVID-19 cases. We demonstrate how observational studies can be misleading when used to guide public health policy.

Contributors
AC conceptualized the study, collected the data and performed the empirical analysis. AC and TH contributed to the data interpretation. AC and TH contributed to writing the manuscript and AC wrote the first draft. AC and TH verified the data used in the study and approved the final manuscript for submission.

Data sharing
The county case data were obtained from the CDC's Restricted Case Dataset and may not be publicly shared. All other data sources are publicly available.

References

1 Jehn M, Mac McCullough J, Dale AP, Gue M, Eller B, Cullen T, Scott SE. Association between K–12 school mask policies and school-associated COVID-19 outbreaks—Maricopa and Pima Counties, Arizona, July–August 2021. *Morbidity and Mortality Weekly Report*. 2021 Oct 1;70(39):1372.

2 Boutzoukas E, Kanecia O. Zimmerman, Moira Inkelas, M. Alan Brookhart, Daniel K. Benjamin, Sr., Sabrina Butteris, Shawn Koval, Gregory P. DeMuri, Vladimir G. Manuel, Michael J. Smith, Kathleen A. McGann, Ibukunoluwa C. Kalu, David J. Weber, Amy Falk, Andi L. Shane, Jennifer E. Schuster, Jennifer L. Goldman, Jesse Hickerson, Vroselyn Benjamin, Laura Edwards, Tyler R. Erickson, Daniel K. Benjamin, Jr.; School Masking Policies and Secondary SARS-CoV-2 Transmission. *Pediatrics* 2022; 10.1542/peds.2022-056687

3 [preprint] Oster E, Jack R, Halloran C, Schoof J, McLeod D. COVID-19 mitigation practices and COVID-19 rates in schools: Report on data from Florida, New York and Massachusetts. medRxiv. 2021 Jan 1.

4 [preprint] Aapo Juutinen, Emmi Sarvikivi, Päivi Laukkanen-Nevala, Otto Helve. Use of face masks did not impact COVID-19 incidence among 10–12-year-olds in Finland. medRxiv 2022.04.04.22272833.

5 [preprint] Coma, E., M. Catal`a, L. M´endez-Boo, S. Alonso, E. Hermosilla, E. Alvarez-Lacalle, D. Pino, M. Medina-Peralta, L. Asso, A. Gatell, Q. Bassat, A. Mas, A. Soriano-Arandes, F. Fina-Aviles, and C. Prats (2022). Unravelling the role of the mandatory use of face covering masks for the control of SARS-CoV-2 in schools: A Quasi-experimental study nested in a population-based cohort in Catalonia (Spain).

6 Abaluck J, Kwong LH, Styczynski A, Haque A, Kabir MA, Bates-Jefferys E, Crawford E, Benjamin-Chung J, Raihan S, Rahman S, Benhachmi S. Impact of community masking on COVID-19: A cluster-randomized trial in Bangladesh. *Science*. 2021 Dec 2: eabi9069.

7 Bundgaard H, Bundgaard JS, Raaschou-Pedersen DE, von Buchwald C, Todsen T, Norsk JB, Pries-Heje MM, Vissing CR, Nielsen PB, Winsløw UC, Fogh K. Effectiveness of adding a mask recommendation to other public health measures to prevent SARS-CoV-2 infection in Danish mask wearers: a randomized controlled trial. *Annals of internal medicine*. 2021 Mar;174(3):335-43.

8 Budzyn SE, Panaggio MJ, Parks SE, Papazian M, Magid J, Eng M, Barrios LC. Pediatric COVID-19 cases in counties with and without school mask requirements—United States, July 1–September 4, 2021. *Morbidity and Mortality Weekly Report*. 2021 Oct 1;70(39):1377.

9 Lawlor DA, Smith GD, Bruckdorfer KR, Kundu D, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence?. *The Lancet*. 2004 May 22;363(9422):1724-7.

10 Ioannidis JP. Why most published research findings are false. *PLoS medicine*. 2005 Aug 30;2(8):e124.

11 [preprint] Chikina, M., W. Pegden, and B. Recht (2021). A note on sampling biases in the Bangladesh mask trial. arXiv preprint arXiv:2112.01296 .

12 Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

13 Wada K, Oka-Ezoe K, Smith DR. Wearing face masks in public during the influenza season may reflect other positive hygiene practices in Japan. *BMC Public Health*. 2012 Dec;12(1):1-6.

Research in context

Evidence before this study

During the Covid-19 pandemic research on the effectiveness of school mask mandates has been mostly limited to observational studies, which are prone to various kinds of statistical bias. School districts and counties with mask mandates are generally fundamentally different than those without. Isolating the effect of mask mandates alone on pediatric COVID-19 case rates has been challenging without randomized data, though one regression discontinuity design study and two other large studies identified no effect of school mask mandates on child case rates. Randomized studies of adults have not found any impact of cloth mask mandates on COVID-19 case counts and only a marginal benefit of surgical masks in those over the age of 50. One highly publicized study in the United States found a negative correlation between school mask mandates and county pediatric case counts but was limited by size, geographic location and duration.

Added value of this study

We replicated and then expanded upon a highly cited study published by the CDC which reported a negative correlation between school mask mandates. Our expanded analysis, which included increased population size, geographic area and duration, failed to find a correlation between school-based mask mandates and county-wide pediatric covid-19 case counts.

Implications of all the available evidence

Our study calls into question the results of the CDC study which was used, among other studies, to justify national school mask mandate recommendations. Our expanded analysis using a longer time period and a larger, more diverse population had findings consistent with the existing randomized literature on mask mandates in adults and also with the highest quality observational studies in children. Our re-analysis, and subsequent statistical critique, demonstrates how observational studies, especially those that are smaller or of short duration, are prone to identifying spurious or confounded results.

Figure 1: Mean daily pediatric Covid-19 rates per 100,000 children, 565 counties, July 1--October 8, 2021

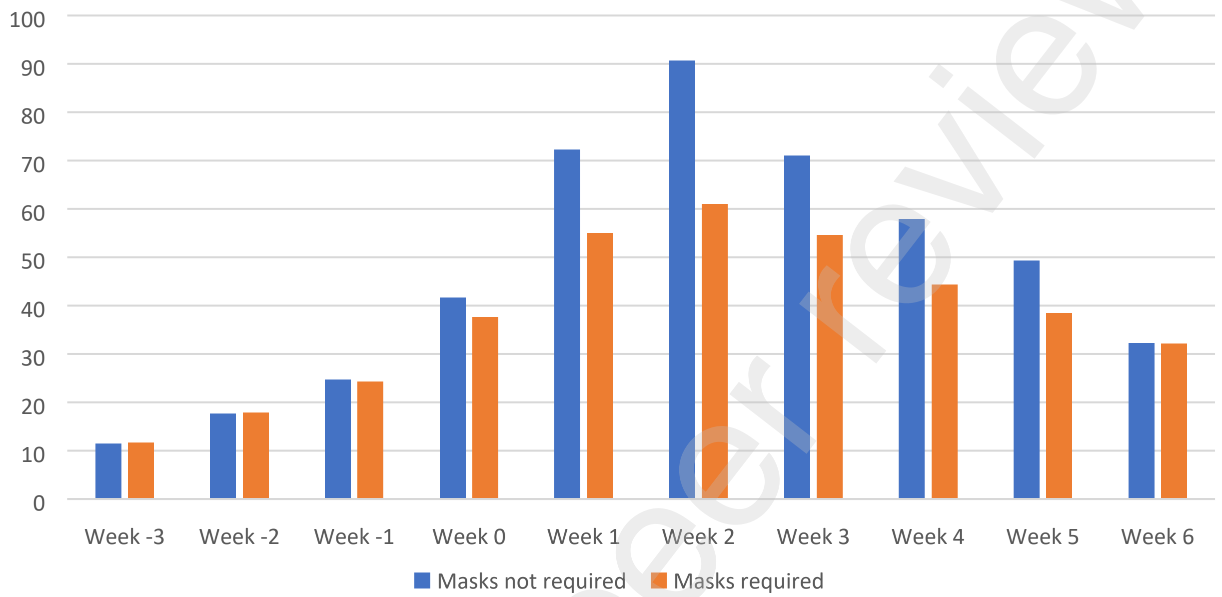

Figure 2: Mean daily pediatric Covid-19 rates per 100,000 children, 1832 counties, July 1--October 20, 2021

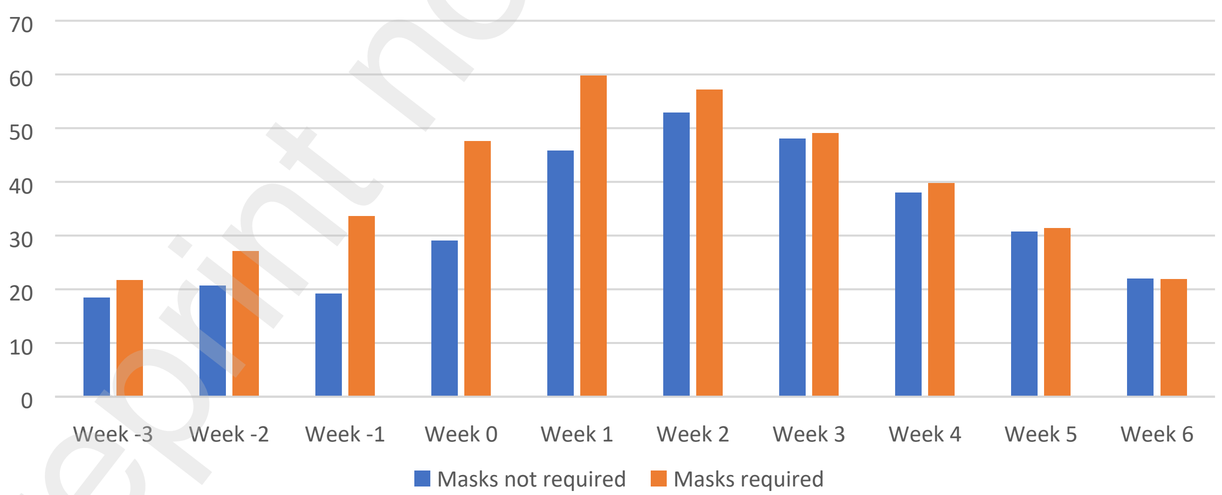

Table 1: Regression of pediatric Covid-19 case rates per 100K

	From Budzyn et al (1)	From Budzyn et al minus CCVI (2)
School Mask Requirement	1.279	3.507
	(0.058)	(0.000)
Adult Cases per 100K	1.194	1.196
	(0.000)	(0.000)
Percent Uninsured	-0.558	-0.353
	(0.000)	(0.000)
Percent in Poverty	0.531	0.312
	(0.000)	(0.000)
Population Density	-0.001	0.001
	(0.162)	(0.329)
Social Vulnerability Index	-13.558	8.911
	(0.000)	(0.000)
Community Vulnerability Index	27.154	
	(0.000)	
Percent Non-Hispanic White	-0.892	-1.836
	(0.667)	(0.380)
Median Age	0.768	0.738
	(0.000)	(0.000)
Pediatric Vaccination Rate	1.320	-1.352
	(0.546)	(0.540)
Intercept	-40.113	-35.973
	(0.000)	(0.000)
R^2	0.673	0.666
Observations	12824	12824

Notes: Values in the table are coefficient estimates obtained from a regression of county level reported per-capita pediatric cases of Sars-Cov-2 on the control variables listed in the table. *p*-values are in parentheses. Regressions also include week fixed effects, that are not reported. The specification in Column 1 follows Budzyn et al. The specification in Column2 omits the CCVI index due to its high correlation with the Social Vulnerability Index. See text for sample construction and other details.