Estimation of mask effectiveness perception for small domains using multiple data sources

Aditi Sen and Partha Lahiri *

December 14, 2020

Abstract

All pandemic are local; so learning about the impacts of pandemic on public health and related societal issues at granular level is of great interest. COVID-19 is affecting everyone in the globe and mask wearing is one of the few precautions against it. To quantify people's perception of mask effectiveness to prevent the spread of COVID-19 for small areas, i.e. for 50 US states and District of Columbia, we use Understanding America Study's survey data on COVID-19. To this end the direct estimate derived from wave data are not reliable: for very small sample size the estimates can have high variability. Synthetic estimates on the other hand provide robust results. We will estimate proportions of mask effectiveness using synthetic estimator based on logistic models with features of respondents as well as states. We establish its reliability over direct estimates in terms of stability and standard errors through Cross Validation, Benchmarking and Jackknifing techniques. This proposed modelling approach gives us a statistical tool to produce more reliable estimates which we can apply the developed methodology to other binary variables related derived from the UAS survey, for e.g. proportion of people suffering from mental health issues or high fear of job loss.

Keywords: Adjusted maximum likelihood method; empirical Bayes; empirical best linear unbiased prediction; linear mixed model

*Aditi Sen is Student Researcher. Partha Lahiri is Professor, The Joint Program in Survey Methodology & Department of Mathematics, University of Maryland, College Park, Maryland, US. The second author's' research was partially supported by the U.S. National Science Foundation Grant SES-1758808.

1 Introduction

Having originated in China at the end of 2019 a novel coronavirus has created worldwide crisis in the year 2020. Cascella et al. (2020) in Features, Evaluation, and Treatment of Coronavirus state that published literature can trace the beginning of symptomatic individuals back to the beginning of December 2019 and as they were unable to identify the causative agent, these first cases were classified as ”pneumonia of unknown etiology.” This new virus is very contagious and has quickly spread globally. The etiology of this illness was consequently attributed to a novel virus belonging to the coronavirus (CoV) family. In a meeting on January 30, 2020, the outbreak was declared by the World Health Organization a Public Health Emergency of International Concern as it had spread to 18 countries with four countries reporting human-to-human transmission. The disease the virus coronavirus causes was named Covid-19 by the World Health Organization on 11th February 2020 as depicted in WHO situation reports and activity logs, papers and articles such as Organization (2020b). The first case of Covid-19 in the USA emerged on 20th January; Holshue et al. (2020) discuss that the infected was a 35-year-old man who returned to Washington State on January 15 after traveling to visit family in Wuhan, and as of November end.

The symptoms of Covid-19 range from uncomplicated (mild) Illness to moderate pneumonia to severe pneumonia. Cascella et al. (2020) discuss that for mild cases patients usually present with symptoms of an upper respiratory tract viral infection, including mild fever, cough (dry), sore throat, nasal congestion, malaise, headache, muscle pain, or malaise. New loss of taste and/or smell, diarrhoea, and vomiting are usually observed. Moderate cases show respiratory symptoms such as cough and shortness of breath without signs of severe pneumonia whereas severe pneumonia depicts fever associated with severe dyspnea, respiratory distress, tachypnea, and hypoxia, however, the fever symptom must be interpreted carefully as even in severe forms of the disease, it can be moderate or even absent. The authors further state that based on data from the first cases in Wuhan and investigations conducted by the China CDC and local CDCs, the incubation time could be generally within 3 to 7 days and up to 2 weeks as the longest time from infection to symptoms was 12.5. This data also showed that this novel epidemic doubled about every seven days, whereas the basic reproduction number is 2.2 i.e. on average, each patient transmits the infection to an additional 2.2 individuals.

Since early months of 2020 every country has experienced the impact of the pandemic in various ways, affecting physical as well as mental and economic health. Pfefferbaum and North (2020) discuss that the public health emergencies may affect the health, safety, and well-being of both individuals and communities. These effects may translate into a range of emotional reactions and extensive research in disaster mental health has established that emotional distress is ubiquitous in affected populations. Due to the pandemic life and livelihood has changed manifold with people losing jobs or living with fear of losing income, alienated from society with months of staying at home and as a result of these there is increase in crimes. Referencing Pfefferbaum and North (2020), from April 2020, we get to learn that a review of psychological sequelae in samples of quarantined people and of health care providers may be revealed numerous emotional outcomes, including stress, depression, irritability, insomnia, fear, confusion, anger, frustration, boredom, and stigma associated with quarantine, some of which persisted after the quarantine was lifted. Specific stressors included greater duration of confinement, having inadequate supplies, difficulty securing medical care and medications, and resulting financial losses.

As of November 30th, there have been more than 62 million cases of which close to 18 million are active. In terms of the magnitude of infection and deaths the USA remains the country with the most confirmed cases of Covid-19, as of November end the total infection count has surpassed 13 million , obtained from worldometer information: Worldometer (2020). In mid April the death toll for USA reached highest in the world, surpassing Italy; on 11th April, total death count became greater 20 thousand in USA, as depicted by various reports like BBC (2020). Through time, testing rates have improved and the Covid Tracking report provides this through charts, Cov (2020), and daily numbers on hospitalization, recovery, death rates along with testing for various states of the US. USA being one of the most affected countries, is in discussion and an article from the Pew Research Center Social Demographic Trends from March 2020, Pew (2020), note that nearly nine-in-ten U.S. adults say their life has changed at least a little as a result of the COVID-19 outbreak, including 44% who say their life has changed in a major way. It is further noted that about nine-in-ten U.S. adults (91%) say that, given the current situation, they would feel uncomfortable attending a crowded party. Roughly three-quarters (77%) would not want to eat out at a restaurant. In the midst of a presidential election year, about two-thirds (66%) say they wouldn't feel comfortable going to a polling place to vote. And smaller but still substantial shares express discomfort even with going to the grocery store (42%) or visiting with a close friend or family member in their home (38%).

Response to the pandemic in the US varied from state to state and was famously characterized by an explosion of cases in the state of New York before lockdown was imposed. Gershman (2020) discuss that most U.S. states have imposed lockdown measures restricting gathering and social contact, disrupting the lives of hundreds of millions of people and the operations of thousands of businesses. Some states, however, have announced or instituted plans to relax restrictions and several states did not impose lockdown and internal travel was generally unrestricted. Along with quarantine policies the other preventive measures adopted to fight the virus are sanitization and regular hand-washing, wearing masks or face covering while going out in public in order to reduce potential spread of the virus without causing any decelerating impact on the economy like that of lockdown. Guner et al. (2020) emphasize that with increased testing capacity, detecting more positive patients in the community will also enable the reduction of secondary cases with stricter quarantine rules, but in COVID-19, which has no approved treatment, it is very important to prevent the spread in the society and the main points in preventing the spread in society are hand hygiene, social distancing and quarantine. Earlier due to lack of clarity on the severity of Covid-19 (viz. how fast it spreads or how it can be asymptomatic) some public health officials had suggested it wasn't mandatory to wear masks or face coverings, but with further development of cases, organizations like the W.H.O. and CDC have suggested it as an effective measure for both people who are affected to stop spreading of the virus and those around not getting affected from contact with affected people: WHO Interim Guidance in January 2020, Organization (2020a), state that "Wearing a medical mask is one of the prevention measures to limit spread of certain respiratory diseases, including 2019-nCoV, in affected areas. However, the use of a mask alone is insufficient to provide the adequate level of protection and other equally relevant measures should be adopted. If masks are to be used, this measure must be combined with hand hygiene and other IPC measures to prevent the human-tohuman transmission of 2019-nCov".

But the issue of wearing mask or face coverings has created a lot of controversy due to difference in opinion and has been highly politicized too, especially in the USA. Chughtai et al. (2020) discuss the effectiveness of cloth masks to protect the wearer from respiratory infections and note that the use of cloth masks during the coronavirus disease (COVID-19) pandemic is under debate. The filtration effectiveness of cloth masks is generally lower than that of medical masks and respirators; however, cloth masks may provide some protection if well designed and used correctly and hence conclude that cloth masks are a more suitable option for community use when medical masks are unavailable. Mask usage and effectiveness has also been studied through surverys on COVID-19. Knotek II et al. (2020) comment that variation is seen in perception of mask effectiveness due to factors like age saying while most respondents indicated that they were extremely likely to wear a mask if required by public authorities, the reported likelihood is strongly dependent on age and perceived mask efficacy i.e. young aged people not considering masks to be that effective as older people. There is also denial from a section of general public as they find it a breach to their freedom if compelled to mask up. Due to mixed messages even from highest authority there has been difference in approach for different states as to whether mask wearing is so effective as to make it a rule to wear them when social distancing is not as much effective. In the vast nation that is the US the approach of states to this issue has varied with some states having made it a mandate to wear masks or face coverings, like that of California by requirements of the state's Department of Public Health released in Angell and Newsom (2020), which state that "people who are infected but are asymptomatic or presymptomatic play an important part in community spread. The use of face coverings by everyone can limit the release of infected droplets when talking, coughing, and/or sneezing, as well as reinforce physical distancing".

Understanding America Study is a panel of households at the University of Southern California (USC) of approximately 9,000 respondents representing the entire United States. Their survey Understanding Coronavirus in America, conducted by members of Center for Economic and Social Research (CESR), part of the USC Dornsife College of Letters, Arts and Sciences, is a survey on the coronavirus pandemic in the United States, where respondents answer surveys on a computer, tablet, or smart phone, wherever they are and whenever they wish to participate. This survey has been live from March and till now (mid November) there have been 16 waves. The survey asks pertinent questions related to a variety of topics from COVID-19 Symptoms, Testing, and Medical Care, COVID-19 knowledge, expectations and behaviors, COVID-19 risk perceptions, mental Health and substance abuse, discrimination and stigma, economic and food Security, social safety nets, housing and debt, crime and safety. While for the whole country of USA, national estimates can be effectively derived by weighted means or proportions from respondent level data using relevant variables, but to draw conclusion on small areas such as the different state, for which populations vary a great deal and hence sample size from states vary as well, direct methods of estimation as inappropriate as well as misleading with very low or high estimates and highly variable standard errors. Methods of synthetic estimation using explicit models or otherwise are hereby relevant and depicts very well performance in terms of sensible estimate values and robust standard errors. Along with UAS data, we combine census data, Covid Tracking Report data to create features for models and draw inference on the small areas. The small states of USA like Rhode Island, Wyoming hence can be well represented by synthetics estimates as can be the large states like California, New York.

It is of interest to find estimates of proportion of people considering mask to be highly effective at state level along with standard errors for USA. In our analysis of mask effectiveness we have five main sections viz. UAS Data, Supplementary Data, Synthetic Method, Data Analysis, Conclusion and Acknowledgements. Thus we first explain the different sources of data, with particular emphasis on UAS first, it's data structure, weighting procedure etc., then the supplementary data like Covid Tracking and US Census Bureau. Then we derive mathematically the synthetic estimate using logistic regression at respondent level followed by data analysis section. In the data analysis section we first define 6 logistic models with various combination of features, which we fit on multiple waves, 10 waves from the recent most, and note the significant variables from them using p-value criteria. In the model selection section we use cross validation technique for selecting the best performing model and thus calculate the synthetic estimate and benchmarked synthetic estimate with benchmarking ratio method for small areas (states of USA) using the estimated model coefficients from each wave. Next we estimate the variance of the synthetic estimator with Jackknifing technique. Finally we evaluate the estimates derived from this synthetic method by comparative analogy of plotting with direct estimates for a handful of states, some small like District of Columbia, Rhode Island, North Dakota and large states like New York, California, Florida. We conclude the paper with the effectiveness of the methods described in the paper and how they can be extended to any other binary, categorical or continuous variable from this survey or any other with little adjustments or modifications. We extend our thanks to UAS in the acknowledgements section and provide references of articles, research papers and journals at the end of the paper.

2 UAS Data

The Understanding America Study (UAS) is a nationally representative panel of U.S. households, where a household is broadly defined as anyone living together with the initial person who signed up to become a participant in the UAS. The Coronavirus of UAS survey was launched on March 10, wherein a total of 9063 UAS panel members were invited to participate including 8,547 panel members who were eligible to be included in the weighted sample. Each panel member was randomized to respond on a pre-assigned day of the week, distributed so that a full sample is invited to participate over a 14-day period. Respondents have 14 days to complete the survey but receive an extra monetary incentive for completing the survey on the day they are invited to participate. Data for the full sample is thus final after a 28 day period. Since respondents have two weeks to answer the survey, the total field period is 4 weeks, so that responses during the last two weeks of a field period of one survey overlap with responses in the first two weeks of the subsequent survey. The wave details like name of the wave, time period, sample size are described in Table 1.

Sampling batch and frame: The UAS is sampled in batches through address based sampling and as of November 2020 there are 21 batches (latest being added in August which is "21 MSG 2020/08 Nat. Rep. Batch 11") as described in 2. Most batches use a two-stage sample design, in which zip codes are drawn first, and then households are randomly drawn from the sampled zip codes (except for two small sub-groups which are simple random samples from lists). The National batches draw zip codes without replacement, but the Los Angeles County batches draw with replacement and do sometimes contain the same zip code in different batches. Batches cannot be viewed as strata for sampling design as typically, strata are different sub populations, while UAS batches are drawn from the same population (or in the case of Native Americans/California/LA County, sub populations of the other batches' population). The UAS draws from multiple frames, but each batch draws from only one frame. Thus sample frame is of 4 types; Nationally Representative Sample, Native Americans, LA County, California and their relation with batches are shown in Table 3.

Weighting procedure: A two step weighing procedure is followed by UAS where first base weights are computed, using a logit model to estimate the probability that a zip code is sampled, which correct for the unequal probabilities of sampling UAS members, and

Table 1: UAS Wave Details

Wave Number	Wave Name	Time period	Sample size (n)
1	UAS 230	March 10,2020 - March 31,2020	6,932
2	UAS 235	April 1, 2020 - April 28, 2020	5,478
3	UAS 240	April 15, 2020 - May 12, 2020	6,287
4	UAS 242	April 29, 2020 - May 26, 2020	6,403
5	UAS 244	May 13, 2020 - June 9, 2020	6,407
6	UAS 246	May 27, 2020 - June 23, 2020	6,408
7	UAS 248	June 10, 2020 - July 8, 2020	6,346
8	UAS 250	June 24, 2020 - July 22, 2020	6,077
9	UAS 252	July 8, 2020 - Aug 5, 2020	6,289
10	UAS 254	July 22, 2020 - Aug 19, 2020	6,371
11	UAS 256	Aug 5, 2020 - Sep 2, 2020	6,238
12	UAS 258	Aug 19, 2020 - Sep 16, 2020	6,284
13	UAS 260	Sep 2, 2020 - Sep 30, 2020	6,284
14	UAS 262	Sep 16, 2020 - Oct 14, 2020	6,129
15	UAS 264	Sep 30, 2020 - Oct 27, 2020	6,181
16	UAS 266	Oct 14, 2020 - Nov 11, 2020	6,181

Table 2: UAS List of Batches

1 ASDE 2014/01 Nat.Rep.	11 MSG 2016/04 Nat.Rep. Batch 7
2 ASDE 2014/01 Native Am.	12 MSG 2016/05 Nat.Rep. Batch 8
3 ASDE 2014/11 Native Am.	13 MSG 2016/08 LA County Batch 2
4 LA County 2015/05 List Sample	14 MSG 2017/03 LA County Batch 3
5 MSG 2015/07 Nat.Rep. Batch 1	15 MSG 2017/11 California Batch 1
6 MSG 2016/01 Nat.Rep. Batch 2	16 MSG 2018/02 California Batch 2
7 MSG 2016/01 Nat.Rep. Batch 3	17 MSG 2018/08 Nat.Rep. Batch 9
8 MSG 2016/01 Nat.Rep. Batch 4	18 MSG 2019/04 LA County Batch 4
9 MSG 2016/02 Nat.Rep. Batch 5	19 MSG 2019/05 LA County Batch 5
10 MSG 2016/03 Nat.Rep. Batch 6	20 MSG 2019/11 Nat. Rep. Batch 10
	21 MSG 2020/08 Nat. Rep. Batch 11

Table 3: UAS Batch and Frame Relationship

Batch	Reference Population
1	U.S.
2,3	Native American
4	Los Angeles County young mothers
5 to 12	U.S.
13,14,18,19	Los Angeles County
15,16	California
17,20,21	U.S.

then post-stratification or final weights are generated using raking algorithm which align the sample of each study to the reference population along certain socio-economic dimensions such as gender, age, race-ethnicity, household income among others. The benchmark distributions against which UAS surveys are weighted are derived from the Basic Monthly Current Population Survey (CPS).Weights are provided for all batches, except batch 4, which comprises of Los Angeles County young mothers, and non-Native American households in batches 2 and 3.

Survey questionnaire, wave and variables: The survey includes a national bi-weekly long-form questionnaire and a weekly Los Angeles County short-form questionnaire administered in each bi-weekly wave. As of November 11, there are 16 waves, as described in Table 1 with their time periods, and for our study we have considered some of the latest waves till wave 16 (UAS 266). Each wave data consists of, on an average, six thousand observations, with variables comprising of some default variables like household identifiers and demographic variables from My Household (a quarterly administered survey which inquires UAS respondents of their age, ethnicity, education, marital status, work status, state of residence and family structure among other matters), weight variables comprising of base weight, final weight of every respondent and survey specific variables related to COVID-19.

Problem formulation For the mask effectiveness problem we focus on the following question from survey questionnaire:

How effective is wearing a face mask such as the one shown here for keeping you safe from coronavirus?

- *Extremely Ineffective*
- *Somewhat Ineffective*
- *Somewhat Effective*
- *Extremely Effective*
- *Unsure*

The answer choices of respondents have been used to create a binary variable where 1 is taken if mask if considered to be Extremely effective by respondent and 0 otherwise. Using this binary variable the direct estimate works really well at overall national level with low SE. At national level first to understand the broader question on the identification of demographic factors influencing such effectiveness perceptions certain domains or groups are created based on race-ethnicity x age. These four groups are Non Hispanic White Age 18-44, Non Hispanic White Age 45+, Other race Age 18-44 and Other race Age 44+ and sensible variation among those groups is observed across multiple waves with all standard errors (SE) from direct estimates around 2%, after which it is chosen for further estimation study. The direct estimates (i.e.Horvitz Thompson estimates using weighted mean of respondents in the survey sample) at national level as well as domain level from waves 1 to 9 are provided in Table **??** along with the standard errors in parenthesis. From such numbers we see that the overall national estimate and the domain NH White Age(45+) behave similarly (for example 50% and 49% at wave 10). The Other Race Age (45+) domain has the highest perception of mask effectiveness (for example 66% at wave 10), whereas the domain NH White Age (18-44) has the least value of such estimate (for example 38% at wave 10). Thus this break down of the population into domain can be used further into the analysis during modelling. We have used R survey package to compute such estimates with the weights of

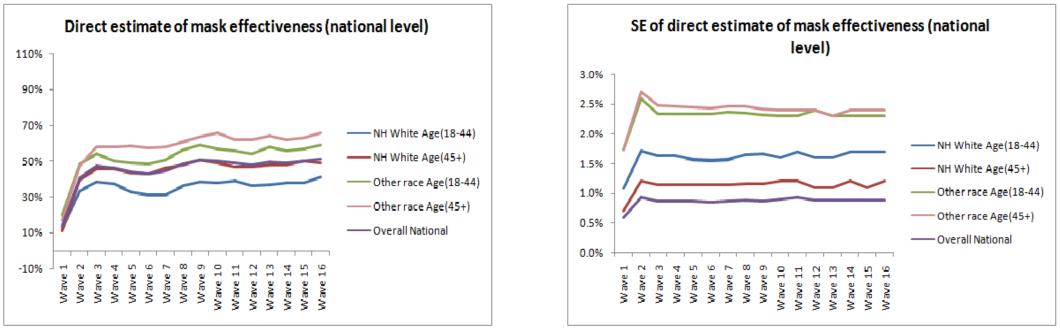


Figure 1: National direct wave estimates of mask effectiveness and associated standard error direct estimates; overall estimates as well as estimates for four groups are provided.

respondents as provided in the wave data. In each wave the total weight is equal to the total number of observations in the sample.

Motivation of synthetic estimation The survey data contains respondents residing in 50 states and DC but naturally they are not evenly distributed i.e. for larger states like California or Florida there is sizable volume in the sample of even as high as 2000 respondents and for smaller states like Delaware or Wyoming there is very little representation of even 3 or 4 respondents. In such scenarios by using direct survey wave based estimator there are serious issues of making wrong estimation, for example we see for the first three waves 0% of people in Wyoming think mask is extremely effective which happens because all the respondents in the sample take the value 0 for binary response variable mask effectiveness. Hence this is not a good method to draw conclusion for the whole population of the states. In addition there arises the issue of extremely variable Standard Error (SE) or margin of error (ME). Estimated SE, or equivalently, estimated ME for a state depends on sample size and value of estimated proportion. For states with small sample sizes, say less than 12, SE is either 0 or very high. From computations of direct estimates or weighted mean i.e. Horvitz Thompson estimator, from multiple waves we see that for Rhode Island, a state whose contribution in the wave is small with 2 or 3 respondents, estimated SE in the first few waves (1 and 2) is 0%. The reason for 0 SE can be either a sample size of only 1 respondent or all binary observations taking the same value (either 0 or 1). In this case of Rhode Island the cause is latter. But as soon as we have a mix of 0s and 1s, SE becomes very high, as high as even 30% from wave 5 on wards to wave 9 for Rhode Island. We have showed this erratic behavior of direct estimates and standard errors as comparative view of four states with varying population sizes (as estimated from the Census Bureau's PEP data)- one with high population (**California** - estimated adult population of 30 million from PEP), one with medium population (**New York** - estimated adult population of 15 million from PEP), one with small population (**Maryland** - estimated adult population of 4 million from PEP) and one with very small population (**Rhode Island** - estimated adult population of 800k from PEP), in Figure **??**. The curves for Rhode Island are the most variable (both state level direct estimate and SE), next is Maryland which vary but not too much and those of New York and California which are quite stable. These SE estimates are thus surely very unstable or unreliable and typically, in public opinion polls like the famous Gallup polls, margin of errors (2x se) is targeted at a low level such as 3% or 4%.

The Figure 2 for Rhode Island demonstrates high variability in state estimates for smaller states. Along with high variability a demonstration of high bias in the direct state estimates can also be viewed. Since we do not know the truth for mask effectiveness, we cannot demonstrate bias properties for mask effectiveness. But we can say if we consider another outcome variable for which "truth" is known from the PEP data, we can at least partially justify our claim. Using Figure 3 we show that UAS estimates of proportions of people falling in the four demographic groups or domains we considered do not match up with PEP data for states, but they more or less match at the national level. For large states

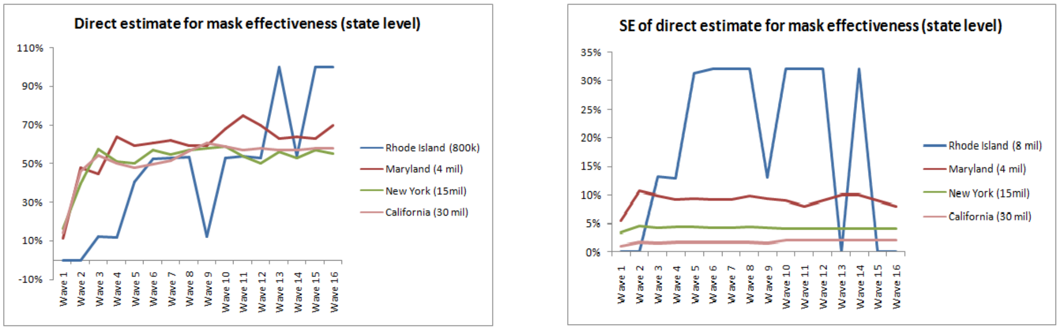


Figure 2: Direct estimates of mask effectiveness and associated standard errors for four selected states.

like California, the difference between PEP estimate of percentage of adult population and UAS direct wave estimate is little, similar it is for medium sized states like Maryland and New York, but for small states like Rhode Island and North Dakota, referring to Figure 3, we see the percentages vary a great with even 0% or no contribution in some domains.

From the aforementioned observations, it is clear the direct estimates are not stable even at the state level. We need improved estimator like the synthetic estimator that is smoother than the direct estimator. These synthetic estimators essentially would borrow strength from other states through implicit or explicit models and combine information from the sample survey, various administrative/census records, or previous surveys. Synthetic estimators are highly effective and appealing in estimating any response for small areas; in this case for US states the proportion of people considering mask to be highly effective against the spread of coronavirus. Referring to synthetic estimation methods explained in Lahiri and Pramanik (2019) we have used unit level logistic model with respondent level features like the age x race-ethinicity group the respondent belongs to along with state level features like which region of the US the respondent's state of residence belongs to (Northeast, Midwest, South or West) or what is the party affiliation of the state (Democratic

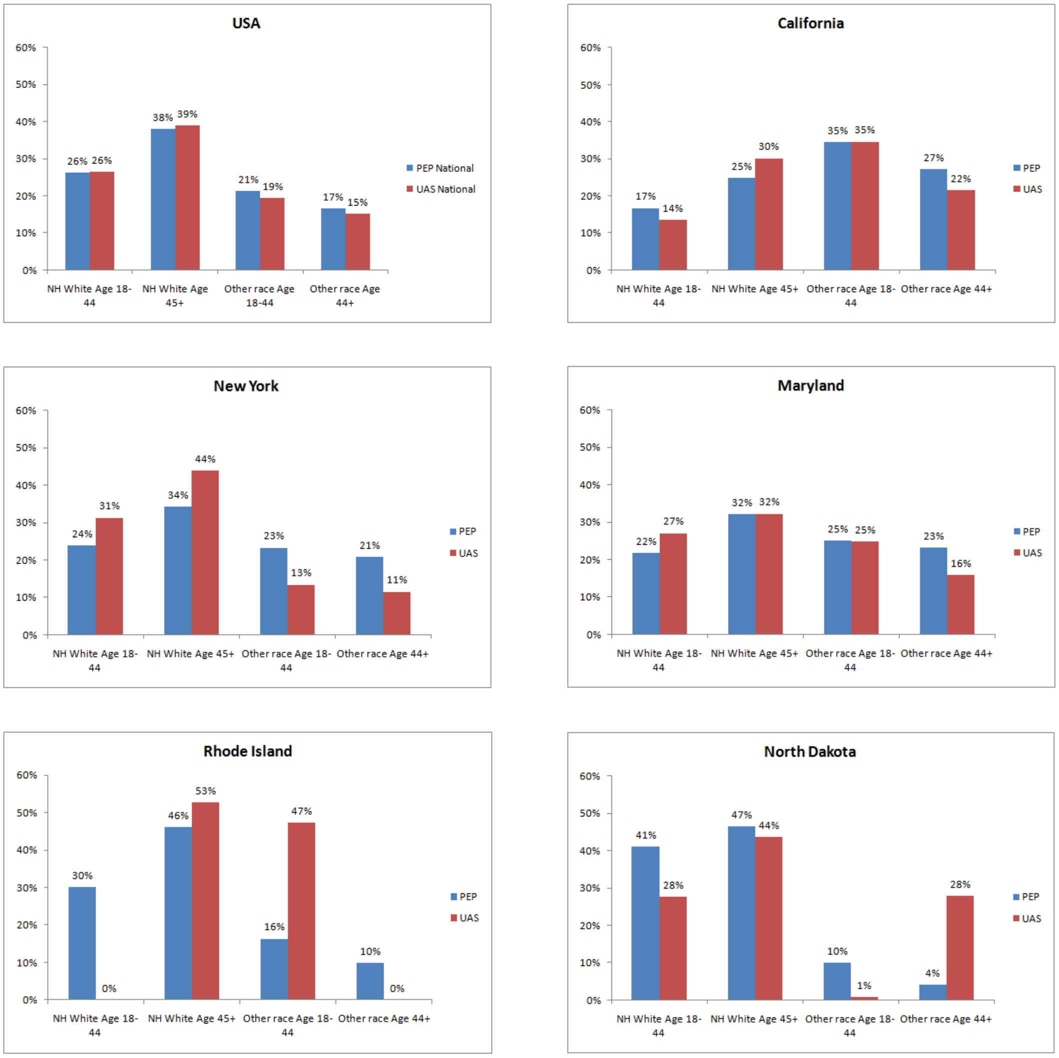


Figure 3: PEP vs. UAS estimates of 4 domains

or Republic) and even what is the testing rate or positivity rate of Covid-19 in the state. Thus we have combined the data in UAS coronavirus survey in conjunction with US Census Bureau data and Covid Tracking Project data to derive state level synthetic estimator of population means and totals for the variable of interest. We will talk about the data in details in the next section.

3 Supplementary Data

Covid Tracking Project: Along with UAS data another data used to facilitate the estimation process is the COVID Tracking Project which collects and publishes testing data daily for US as a whole and also for states and territories. From this data we get to know that for 50 states and DC combined the daily test count has been increasing fast with more than 1 million in April to more than 50 million in August. There are numerous variables on the state level data relating to topics like COVID-19 daily total testing, total test results, positive/negative, confirmed, death, recovery count (as obtained from Johns Hopkins data on coronavirus), hospitalization, ventilation etc. of which we have used the following variables:

1. totalTestResults: gives total number of tests with positive or negative results

2. positive: gives total number of positive tests

Hence we calculated the following for 50 states and DC to be used in the regression model to predict mask effectiveness after checking the correlation with the binary variable in question:

1. Testing rate: Total tests with positive or negative results/Total population of state

2. Positivity rate: Total positive tests/ Total tests with positive or negative results

Census Bureau: To estimate mask effectiveness using synthetic method we need information from reliable larger datasets and hence we use US Census Bureau's Population Estimates Program (PEP) data. Annual estimates of population counts for adult population (18+) is required at overall state level and also by demographics for the domains or groups we have discussed in previous section i.e. NH White Age(18-45), NH White Age(45+) and so on. The Census Bureau essentially obtains these estimates using 2010 decennial census as the base and updates by births, deaths, migration etc. available from the administrative records and others obtained from the ACS survey. Population density estimates for US states for the year 2010 are also obtained from US Census Bureau and we create a categorical feature from it with three levels indicating low (for eg. North Dakota, Wyoming, Alaska etc.), medium(for eg. Georgia, Michigan, Virginia) and high population density(for eg. New York, California, DC) for all states of the USA and DC. We have used two data sources as follows:

1. SCPRC-EST2019-18+POP-RES: Estimates of the Resident Population Age 18 Years and Older for the US states from July 1, 2019 (released on Dec 2019) which can be directly used.

2. SC-EST2019-ALLDATA5: Estimates of population by "Age, Sex, Race, and Hispanic Origin - 5 race groups (5 race alone or in combination groups). This data need to be adjusted by filtering out 18+ population (with "AGE") for the above-mentioned domains (using variables "RACE" for white and rest as other race and "ORIGIN" for Hispanic or Non-Hispanic). Sex is not used, although present in the data and hence set to value 0 for all. The domain wise populations are then adjusted with a factor (i.e. multiplying with domain wise population/total state population) so that the

sum of all the domains equalize with the total state level estimate mentioned before.

4 Synthetic Method

Despite certain limitations of synthetic estimation, there is a widespreed use of synthetic estimation in small area estimation; see, e.g., Rao and Molina (2015),Ghosh (2020),Citro and Kalton (2000),Elbers et al. (2003),Ericksen (1974),Gonzalez and Hoza (1978),Gonzalez and Waksberg (Gonzalez and Waksberg),Hansen et al. (1953),Marker (1995),Marker (1999),Mule (2010),for Health Statistics (1968),Nicholls (1977),Purcell and Kish (1980),A.L. and Zaslavsky (1997),E. et al. (1991), and others. For the synthetic method of estimation of mask effectiveness for small areas i.e. at state level we first define the following mathematical notations and then derive the formula for the estimator from logistic regression model, as our response variable for this study is binary. We can also use this method for categorical or continuous variable. Let y_k denote the value of outcome (or dependent) variable for the kth respondent ($k = 1, \cdots, n$), where n denotes the number of persons in a given wave (say, wave 2 covering April 1-April 28) of the UAS survey. The outcome variable is binary as defined by $y_k = 1$ if person k considers mask wearing highly effective. Let $x_k = (x_{k1}, \cdots, x_{kp})'$ denote the value of a vector of auxiliary variables (same as independent variables or predictor variables or covariates) for respondent k. We have focused on the following two criteria for selecting the auxiliary variables for the unit level logistic regression model:

(1) The auxiliary variables should have good explanatory power in explaining the outcome variable of interest, which we have checked from correlation values or p-value criteria of significance (as shown in **??** for multiple waves).

(2) The total or mean of these auxiliary variables for the population should be available from a big data such as a bigger survey, administrative records or decennial census. For example the domain level populations (NH white age 18-44 etc.) are all available from the Census Bureau's PEP data.

Taking the above into consideration we have created the following variables:

- Y: Mask wearing (0, 1)
- X_1: Non Hispanic White Age(18-44) (0,1)
- X_2: Non Hispanic White Age(44+) (0,1)
- X_3: Other Race Age(18-44) (0,1)
- X_4: Testing Rate
- X_5: Positivity Rate
- X_6: Population density (1,2,3 for low, medium, high)
- X_7: party affiliation Democratic (0,1)
- X_8: region Northeast (0,1)
- X_9: region Midwest (0,1)
- X_10: region South (0,1)

To avoid multicollinearity for the indicator variables, one is dropped from each group. Let N_i be the population size of the ith state, as we are considering adult population N_i should be interpreted as the adult 18+ population and let N_{gi} be the population size of the gth cell in state i. As discussed previously in the data section the N_{gi} and N_i

values are obtained from US Census Bureau. From the aforementioned list of variables, x_{1k} is a binary dummy variable indicating if the $k$the person is young white (1 if the person is white aged 44 or less and 0 otherwise), x_{2k}: older white etc. for G=4 groups. Let y_{gik} be the value of the outcome variable for kth person in state i for the gth group ($g = 1, \cdots, G;\ i = 1, \cdots, m;\ k = 1, \cdots, N_{ig}$). Here $m = 51$ (50 states and DC) are the small areas. Let z_i be a state specific auxiliary variable and here we have considered a total seven of such state specific variables (as observed from the previous list). For the estimation of mask-effectiveness variable for the 50 states and the District of Columbia, we write the population model as:

$$\begin{aligned} \text{Level 1:} \qquad & y_{gik}|\theta_{gi} \overset{ind}{\sim} f(\theta_{gi}) \\ \text{Level 1:} \qquad & h(\theta_{gi}) = x_g'\beta + z_i'\gamma, \end{aligned}$$

for $k = 1, \cdots, N_{gi},\ \ i = 1, \cdots, m;\ \ g = 1, \cdots, G$, where $f(\theta_{gi})$ is suitable distribution with parameter θ_{gik} (here for binary variable this is a Bernoulli distribution with success probability θ_{gik}), $g(\theta_{gik})$ is a suitable known link function (here for binary variable, we take logit link); β and γ are unknown parameter to be estimated using UAS micro data i.e. at respondent or unit level using survey weights.

We estimate population mean for state i by:

$$\hat{\bar{Y}}_i^{syn} = \sum_{g=1}^{G} \frac{N_{gi}}{N_i}\hat{\theta}_{gi} = \sum_{g=1}^{G} \frac{N_{gi}}{N_i} h^{-1}(x_g'\hat{\beta} + z_i'\hat{\gamma}),$$

where h^{-1} is the inverse function of h; $\hat{\beta}$ and $\hat{\gamma}$ are the survey-weighted estimator of β and γ, respectively.

If $h(\cdot)$ is a logit function, we have

$$\hat{\bar{Y}}_i^{syn} = \sum_{g=1}^{G} \frac{N_{gi}}{N_i}\hat{\theta}_{gi} = \sum_{g=1}^{G} \frac{N_{gi}}{N_i} \frac{\exp(x_g'\hat{\beta} + z_i'\hat{\gamma})}{1 + \exp(x_g'\hat{\beta} + z_i'\hat{\gamma})}.$$

For continuous data, we can assume the following model:

$$y_{gik} = \theta_{gi} + \epsilon_{gik} = x_g'\beta + z_i'\gamma + \epsilon_{gik},$$

where ϵ_{gik} could be iid errors. In this case, we estimate the population mean for state i by: $\left(\sum_{g=1}^{G} \frac{N_{gi}}{N_i} x_g\right)' \beta + z_i'\gamma$, where $N_i = \sum_{g=1}^{G} N_{gi}$. The regression coefficients β and γ are estimated from the UAS survey using nationwide data.

5 Data Analysis

5.1 Model definition

For the data analysis we have used total of 11 recent most wave data from wave 6 to wave 16. We have evaluated four logistic regression models for the indicator response variable mask effectiveness with the different combination of explanatory variables as explained in **??**. We will compare the results from these with one model with mean effect i.e. no covariate. In every case we use R survey package to run weighted logistic regression with quasi-binomial family, where weights are the final post-stratification weights as provided by UAS and design is defined with such weights and no strata or cluster. Summary of one such model is given in **??** (which is Model 1 on wave 16 data). Significant variables are noted from the p-value criterion from the "svyglm" object summary, where p-value < 0.05. The variables for which p-value < 0.05 vary for different waves for a fixed model. A list of variables observed to be significant from such criteria for all models from multiple waves (wave 16 to 16) are noted in **??**. Variables for the groups 1 and 2 are always found to be significant and variables like population density are never found to be significant from such criteria.

Table 4: Significant covariates in Model 1 for different waves

Wave	intercept	NH White Age(18-44) (indicator)	NH White Age(44+) (indicator)	Other race Age(18-44) (indicator)	testing rate	positivity rate	population density (categorical)	region Northeast (indicator)	region Midwest (indicator)	region South (indicator)	Democratic party (indicator)
1	***	*	***						•	**	
2		***	*		•						
3		***	***				***				
4		***	***	*			***			**	
5	*	***	***	*		**	**	•			
6		***	***	*			***				
7		***	***	*			**				
8		***	***				***		*		
9		***	***		•		***				
10		***	***	*	*		**				
11		***	***	•			***		•	•	
12		***	***	*			***				
13		***	***	•	*		**			**	
14		***	***	•			***				
15		***	***	•					•		
16		***	***	*			***				

Table 5: A list of competing models

Wave	intercept	NH White Age(18-44) (indicator)	NH White Age(44+) (indicator)	Other race Age(18-44) (indicator)	testing rate	positivity rate	population density (categorical)	region Northeast (indicator)	region Midwest (indicator)	region South (indicator)	Democratic party (indicator)
M1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
M7	✓	✓	✓	✓			✓				
M8	✓	✓	✓	✓	✓		✓				
M9	✓	✓	✓	✓			✓			✓	
M10	✓	✓	✓	✓	✓		✓			✓	
M11	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
M12	✓	✓	✓	✓	✓		✓	✓	✓	✓	

5.2 Model selection

A method of model selection based on cross validation, Leave One State Out method, is followed to arrive at the best performing model from the total of six models. Then we use the Benchmarking technique to get the benchmarked synthetic estimate at national as well as state level. We talk about these in details as follows:

Cross Validation Leave one state out method: We leave out the entire UAS survey data on the outcome variable y_i (e.g., mask effectiveness) for state i and predict the vector of outcome variables for all sampled units of the leave out state using x_g for the sampled unit and z_i for the leave out state. Let $f(y_i|y_{-i})$ denote the joint density of y_i, all the observations in state i, conditional on the data from the rest of the states, say $y_{-i)}$. For the Bernoulli distribution of y_i for state i, using independence, we have for known model parameters β and γ:

$$\begin{aligned}
\log f(y_i|y_{-i};\beta,\gamma) &= \sum_{g=1}^{G}\sum_{k=1}^{n_{gi}} \left[y_{gik}\log\theta_{gi} + (n_{gi}-y_{gik})\log(1-\theta_{gi})\right] \\
&= \sum_{g=1}^{G}\sum_{k=1}^{n_{gi}} \left[y_{gik}\log\left(\frac{\theta_{gi}}{1-\theta_{gi}}\right) + n_{gi}\log(1-\theta_{gi})\right] \\
&= \sum_{g=1}^{G}\sum_{k=1}^{n_{gi}} \left[y_{gik}(x_g'\beta + z_i'\gamma) - n_{gi}\log\left(1+\exp(x_g'\beta+z_i'\gamma)\right)\right].
\end{aligned}$$

Using data from the rest of states, i.e. $y_{(-i)}$ we get the survey-weighted estimates β and γ and plug in the above expression. Let these estimates be $\hat{\beta}_{w,(-i)}$ and $\hat{\gamma}_{w;(-i)}$. We then define out model selection criterion as:

$$C = \sum_{i=1}\sum_{g=1}^{G}\sum_{k=1}^{n_{gi}} w_{gik}\left[y_{gik}(x_g'\hat{\beta}_{w,(-i)} + z_i'\hat{\gamma}_{w;(-i)}) - n_{gi}\log\left(1+\exp(x_g'\hat{\beta}_{w,(-i)} + z_i'\hat{\gamma}_{w;(-i)})\right)\right].$$

Table 6: Cross validation leave one state out statistic for all models

Model	Minimum	Median	Mean	Maximum
M1	-89.02	-79.84	-79.04	-62.90
M7	-91.95	-83.66	-84.60	-74.75

We compute C for different models and compare the six models with a baseline model with no auxiliary variable, that is $x_g'\beta + z_i'\gamma = \mu$ and choose the one with largest C as the final model. Table 8 shows the C values, which are obtained using the aforementioned formula and further dividing each value by the sample size to scale down the numbers for ease of comparison. The C values are all negative, as probability densities are values between 0 and 1 and logarithm of fractions make these out to be negative. For every state, iteratively regressions are run and regression estimates are obtained which are used in the formula. We observe from Table 6 that the average C value over multiple waves is maximum for M1, although for all models such values are really close. Hence we choose M1 as the best performing model. The estimates of coefficients of auxiliary variables, p-values and other regression summaries for M1 (from wave 16) are noted in **??**.

Bench-marking Ratio method We define Benchmark Ratio (BR) as the ratio of the overall direct national estimate to the synthetic estimate (aggregated at the national level). The synthetic estimates which are obtained at state level by is aggregated by multiplying by the ratio of the adult state population to the overall US adult population estimate and then adding up. The closer the value of BR is to 1 the better is the model. We see from Table 7 that BR is close to 1 for all waves, using which we compute the Benchmarked or BR synthetic estimate .

Table 7: Benchmarking ratios and national synthetic and benchmarked synthetic estimates for last five waves; synthetic estimates are based on Model 1

Model	Wave 12	Wave 13	Wave 14	Wave 15	Wave 16
Benchmarking Ratio	0.98	0.98	0.99	0.99	0.99
Synthetic estimate	48.39	50.30	49.89	50.77	51.67
BR Synthetic estimate	47.69	49.47	49.18	50.22	51.22

5.3 Model diagnostics

For the chosen model M1 we create a state level comparative diagram of benchmarked synthetic estimate with direct estimate in Figure 4 with data from wave 16. As the synthetic estimate and benchmarked synthetic estimate at state level are really close, we have not plotted synthetic estimates for ease if viewing. We observe that the synthetic estimate obtained using regression method is a much more stable one than direct. The states arranged in increasing order of total population show that the issue of highly variable state level direct estimate in the smaller states has been removed in the synthetic one. For largely populated states as well as for small ones the benchmarked synthetic estimate is doing a good job of estimating the proportion of the response variable. We next check the robustness of the synthetic estimator in terms of variance through Jackknife method.

5.4 Estimation of variance of synthetic estimator

We have used resampling methods such as the jackknife to evaluate the variance of the synthetic estimator. We obtain ith jackknife resample by deleting all survey observations in batch j. Thus we have $m = 20$ jackknife resamples from wave 14 onwards because there are 20 batches in total, whereas earlier for waves 1 to 13 there were in total 19 batches in each wave data, the latest addition being "21 MSG 2020/08 Nat. Rep. Batch 11" in August and LA County Young mothers is not present in any of the waves. For each jackknife resample, we recompute replicate synthetic estimate given using (1). We will get m such replicate estimates, say, $\hat{\hat{Y}}_{i(-j)}^{syn}$ $(j = 1, \cdots, m)$. We can then estimate the variance of $\hat{\hat{Y}}_i^{syn}$ by

$$v(\hat{\hat{Y}}_i^{syn}) = \frac{m-1}{m} \sum_{j=1}^{m} \left(\hat{\hat{Y}}_{i(-j)}^{syn} - \frac{1}{m} \sum_{j=1}^{m} \hat{\hat{Y}}_{i(-j)}^{syn} \right)^2 . \tag{1}$$

We have run M1 on wave 16 and obtained at state level the Jackknife estimates of variance and thus the standard deviations and provided a comparative view with the SE from direct estimates at state level. In the two plots in Figure 5, which are from the wave 16 data, states are arranged in increasing order of sample sizes and y-axis is ratio of direct estimate (survey-weighted) and synthetic estimate in the first graph, whereas in the second graph the y-axis shows the ratio of STD/SE i.e. SE of direct estimate coming right from UAS (treating states as domains) and STD of benchmarked synthetic which is the jackknife STD described in the last section. For states with small sample sizes (e.g. Rhode Island, Wyoming), we see a lot of differences between the survey-weighted direct estimates and the synthetic estimates. For states with large sample sizes (e.g., California), the ratio is approaching to 1 (as plotted by the straight line) as the auxiliary variables used to construct the synthetic estimator are reasonable. We observe all the Jackknife estimates are much lesser than direct estimates and we conclude that the model is a fair one at estimating the mask effectiveness at state level.

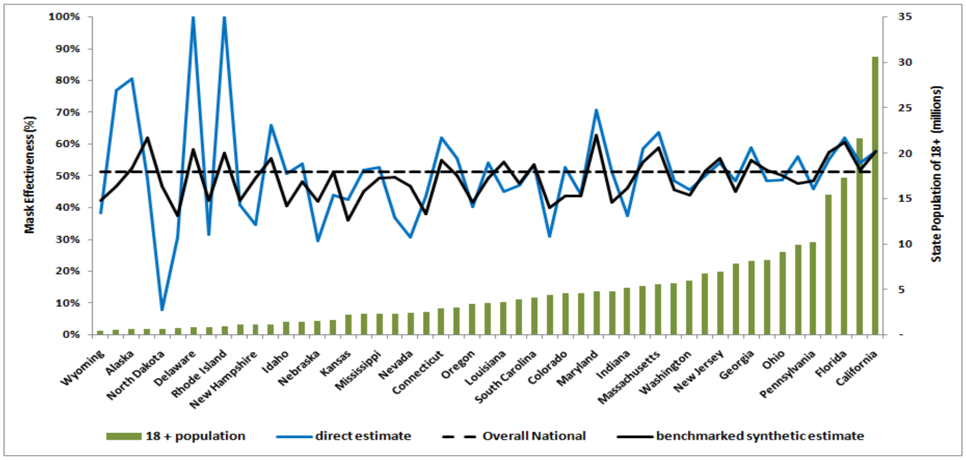
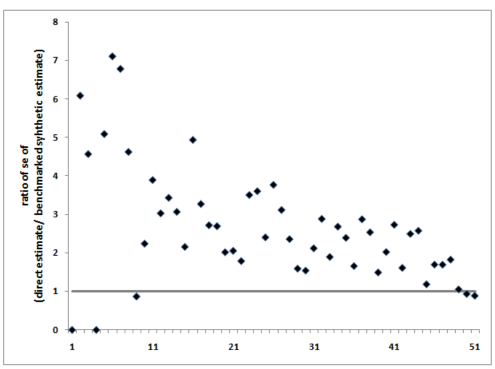


Figure 4: State level comparison of direct estimator and synthetic estimator from M1 on wave 16



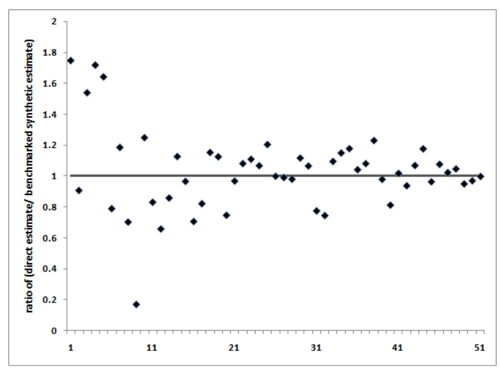


Figure 5: Comparison of direct estimator with benchmarked synthetic estimator through ratio of SE/Jackknife STD and ratio of estimates from M1 on wave 16; domains arranged in increasing order of sample size.

5.5 Evaluation

We will now compare our synthetic estimates with the corresponding direct sample survey estimates (i.e., weighted proportion of people from UAS who believe mask is highly effective) for the 50 states and district of Columbia. This gives us an idea about the magnitude of bias in the synthetic estimates because direct estimates, though unreliable, are unbiased or approximately so.

In problem formulation of UAS section through Figure 2, we pointed out issues with direct estimates and associated se estimates for four states (add DC and North Dakota because of the two questions you asked on them – very high pop density (DC) and testing rates exceeding 1). Now we compare them with the benchmarked synthetic estimates in Figure 6 and 7 in consecutive pages. In page 1 of graphs, we have 6 plots corresponding to the same 6 states (3 with small population - District of Columbia, North Dakota, Rhode Island, 3 with large population - California, New York, Florida) of point estimates (direct and benchmarked synthetic) vs waves, which gives a time series trend from wave 1 to wave 16. In page 2 of graph same kind of plots for SE of direct wave estimated and STD from Jacknife method on benchmarked synthetic estimate. The plots of estimates of mask effectiveness % show great improvement for small states, for example for DC the value was 0% in waves 1 and 2, which has improved to 18% and 39% using benchmarked synthetic estimate and it is line with the national estimates for such waves. Similar issues of 100% mask effectiveness proportion for waves 13 and 15 for Rhode Island have also been eliminated. If we focus on the error graphs the values from direct estimates get as high as 32% for small states (i.e. one with low contribution to overall sample size). Using benchmarked synthetic estimates at state level, error has reduced to almost 6 times with as low as 6% STD from Jackknife method. For larger states like New York and Florida the errors reduce using benchmarked synthetic estimate, although not to a great extent and for the state contributing most to the sample size, California, the STD and SE from Jackknife and direct estimate are more or less similar.

6 Conclusion

The issue of mask effective is a critical one and insight into this behavioral aspect of people helps in understanding the future impacts or spread of the disease at state level. There is impact of political affiliation and views on such behavioural aspect. Needless to say, it will be effective to monitor the this proportion in the coming time period after mid-November, which can be insightful as to the change in administration. From September we have seen rise in cases in the US and in November we have seen total case count to have surpassed 11 million. There have been signs of downfall in the infection curve but then there comes a new wave and certainly this pandemic is still a serious health risk. Wearing masks is undoubtedly one of the few and most effective precautionary measures and people's awareness of such could be tracked from an analysis such as this one, which uses response of the public in surveys from a wide range of time period from March till now (November). Variation in perception through time is observed as the estimates show for overall US in March 14% of people viewed mask wearing to be extremely effective, whereas in November this number is at around 51%. Although this analysis has been done at wave level we can definitely look into estimating such proportions at day level, which would require deep dive into sampling frames and further development of methodology.

The method of estimating population means or totals for the states of USA explained in the paper is a robust one which provides sensible and numerically sound estimates and the model selection and evaluation methods provide satisfactory results with all standard

Figure 6: Time series trend of direct and benchmarked synthetic estimate for 6 sample states (3 small, 3 large)

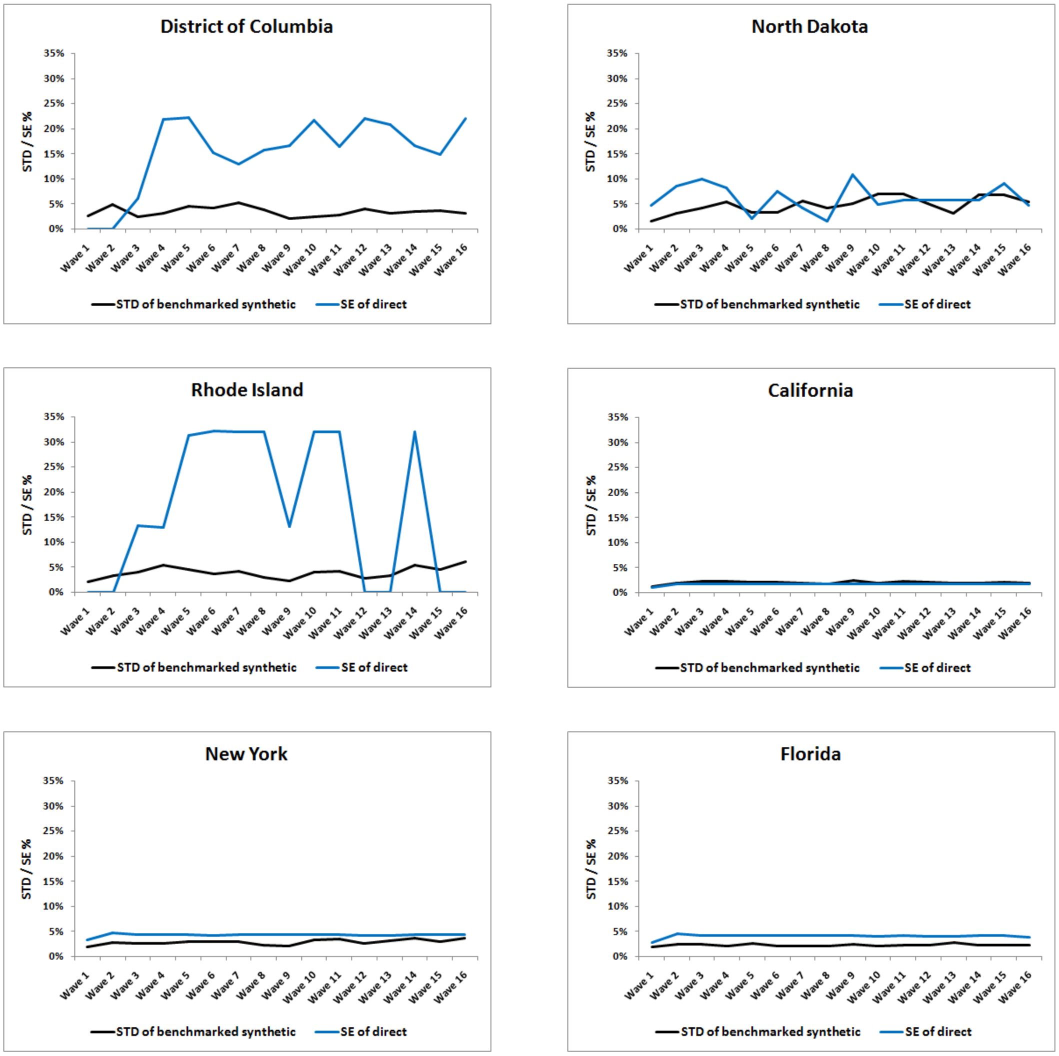


Figure 7: Time series trend of SE of direct and benchmarked synthetic estimate for 6 sample states (3 small, 3 large)

error of estimates within 2% as per the standards. We noticed high variability of synthetic estimates in state level estimation. We further note that while direct UAS estimates are designed to produce approximately unbiased estimates at the national level, they are subject to biases for the state level estimation. Biases in the direct proportion estimates at the state level may arise from the fact that they are essentially ratio estimates since the state sample sizes are random and expected sample sizes are small for most states. Moreover, the UAS weights are not calibrated at the state level.

From our investigation, we found that synthetic estimates improve on UAS direct estimates in terms of variance reduction, especially for the small states. But since synthetic estimates are derived using a working model, they are subject to biases when working model is not reasonable. However, we observe that the benchmarking ratios for all waves are consistently around 1 showing lack of evidence for bias. Our benchmarked synthetic estimates are close to the synthetic estimates because the benchmarking ratios are close to 1. None-the-less by benchmarking synthetic estimates we achieve data consistency and it is reasonable to expect to reduce biases as well. We add that it is possible to biases at the state level by benchmarking the synthetic estimates to the UAS direct estimates for a goup of states (e.g., benchmarking with a division). This may be need for other synthetic estimation problems.

This method can be replicated or tried out for any binary variable as has been done for mask effectiveness and even for categorical or continuous ones. There are numerous interesting areas in the UAS survey which can be studied to find state level estimates. To name a few areas:

- **Physical health**: proportion of people experiencing three or more flu-like symptoms which can be looked at as supposed indication of outbreak (viz. Fever or chills cr001a, Runny or stuffy nose cr001b, Cough cr001d, Sore throat, cr001e etc.)

- **Mental health**: proportion of people seeking help because of mental health issues
- **Economic/financial anxiety**: proportion of people fearing job loss due to Covid-19
- **Other state level Covid-19 estimates**: proportion of people tested, Covid positive, proportions among pre-existing conditions (like cancer, diabetes, high blood pressure, asthma, heart disease etc.)

7 Acknowledgements

The project described in this paper relies on data from survey(s) administered by the Understanding America Study, which is maintained by the Center for Economic and Social Research (CESR) at the University of Southern California. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of USC or UAS. The collection of the UAS COVID-19 tracking data is supported in part by the Bill Melinda Gates Foundation and by grant U01AG054580 from the National Institute on Aging.

References

(2020). Coronavirus: Us death toll overtakes italy as world's highest. *BBC News: US and Canada*.

(2020). Most americans say coronavirus outbreak has impacted their lives. *Pew Research Center: Social Demographic Trends*.

(2020). Us daily tests. *The Covid Tracking Report*.

A.L., S. and A. Zaslavsky (1997). Reweighting households to develop microsimulation estimates for states. *American Statistical Association, Proceedings of the Survey Research Methods Section*.

Angell, S. C. D. o. P. H. and G. G. Newsom (2020). Guidance for the use of face coverings. *State of California—Health and Human Services Agency - California Department of Public Health*.

Cascella, M., M. Rajnik, A. Cuomo, S. Dulebohn, and R. Napoli (2020). Features, evaluation, and treatment of coronavirus. *National Center for Biotechnology Information*.

Chughtai, A., H. Seale, and C. Macintyre (2020). Effectiveness of cloth masks for protection against severe acute respiratory syndrome coronavirus 2. *Centers for Disease Control and Prevention*.

Citro, C. and G. e. Kalton (2000). Small-area income and poverty estimates: Priorities for 2000 and beyond. *National Academy Press, Washington, D.C.*.

E., S., P. Goel, and D. Rumsey (1991). County estimates of wheat production. *Survey Methodology 17*(2), 211–225.

Elbers, C., J. O. Lanjouw, and P. Lanjouw (2003). Micro–level estimation of poverty and inequality. *Econometrica 71*(1), 355–364.

Ericksen, E. (1974). A regression method for estimating population changes of local areas. *Journal of the American Statistical Association 69*, 867–875.

for Health Statistics, N. C. (1968). Synthetic state estimates of disability. *PHS Publication*.

Gershman, J. (2020). A guide to state coronavirus reopenings and lockdowns. *The Wall Street Journal*.

Ghosh, M. (2020). Small area estimation: Its evolution in five decades. *Statistics in Transition New Series, Special Issue on Statistical Data Integration*, 1–67.

Gonzalez, M. and C. Hoza (1978). Small-area estimation with application to unemployment and housing estimates. *Journal of the American Statistical Association 73*(361), 7–15.

Gonzalez, M. and J. Waksberg. Estimation of the error of synthetic estimates. In *First meeting of the international Association of Survey Statisticians, Vienna, Austria*, Volume 18.

Guner, R., I. Hasanoglu, and F. Aktas (2020). Covid-19: Prevention and control measures in community. *National Center for Biotechnology Information*.

Hansen, M., W. Hurwitz, and W. Madow (1953). *Sample Survey Methods and Theory, Vol. 1*. Wiley-Interscience.

Holshue, M., C. DeBolt, S. Lindquist, K. Lofy, J. Wiesman, C. Bruce, H. Spitters, K. Ericson, S. Wilkerson, A. Tural, G. Diaz, and e. a. f. t. W. S. .-n. C. I. T. Cohn, A. (2020). First case of 2019 novel coronavirus in the united states. *The New England Journal of Medicine*.

Knotek II, E., R. Schoenle, A. Dietrich, G. Müller, M. K., and M. M. (2020). Consumers and covid-19: Survey results on mask-wearing behaviors and beliefs. *Federal Reserve Bank of Cleveland*.

Lahiri, P. and S. Pramanik (2019). Estimation of average design-based mean squared error of synthetic small area estimators. *Austrian Journal of Statistics 48*, 43–57.

Marker, D. (1995). *Small Area Estimation: A Bayesian Perspective*. Phd thesis, University of Michigan, Ann Arbor, MI.

Marker, D. (1999). Organization of small area estimators using a generalized linear regression framework. *Journal of Official Statistics 15*, 1–24.

Mule, T. (2010). U.s. census coverage measurement survey plans. In *Proceedings of the Section on Survey Research Methods, Alexandria, VA: American Statistical Association*.

Nicholls, A. (1977). A regression approach to small area estimation. *Australian Bureau of Statistics, Canberra*.

Organization, W. H. (2020a). Advice on the use of masks in the community, during home care and in health care settings in the context of the novel coronavirus (2019-ncov) outbreak. *World Health Organization Interim Guidance*.

Organization, W. H. (2020b). Timeline of who's response to covid-19.

Pfefferbaum, B. and C. North (2020). Mental health and the covid-19 pandemic. *The New England Journal of Medicine*.

Purcell, N. and L. Kish (1980). Postcensal estimates for local areas (or domains). *International Statistical Review/Revue Internationale de Statistique*, 3–18.

Rao, J. N. K. and I. Molina (2015). *Small Area Estimation, 2nd Edition*. Wiley.

Worldometer (2020). Coronavirus u.s.