Small Area Estimation for the Virginia 2020 BRFSS

Submitted to:

Virginia Department of Health
Office of Family Health Services
109 Governor Street
Richmond, Virginia 23219
Attn: Freda Williams, MPH

Submitted by:
Abt Associates Inc.
6130 Executive Boulevard
Rockville, Maryland 20852

January 25, 2022

VDH VIRGINIA DEPARTMENT OF HEALTH

Contents

1. Background

The Virginia Behavioral Risk Factor Surveillance System Survey (BRFSS) is designed to provide stable and accurate estimates on key health outcomes for the household population of adults 18 years of age and older who reside in the Commonwealth of Virginia. The survey is not designed, however, to provide estimates by county. As such, the survey sample sizes are not sufficiently large to provide stable direct estimates of important health outcomes by county (or, in the case of Virginia, by independent city).

Small area estimation (SAE) is often able to provide more reliable estimates for counties than can be obtained from direct survey estimates alone. SAE methods combine direct estimates of outcomes from the survey with estimates modeled using auxiliary or outside data. This auxiliary data can come either from a larger survey such as the U.S Census Bureau's American Community Survey (ACS) or from administrative data such as vital statistics records. Through statistical models of the mathematical relationships between the outcome of interest and area-level characteristics, we can develop estimates for small areas that "borrow strength" from data about other, similar areas. From these models we can get synthetic estimates that represent the expected value of the outcome by county with similar (modeled) characteristics.

There are several important caveats to keep in mind in presenting and using SAEs:

- Whereas direct survey estimates contain sampling error and other sources of survey error, SAEs additionally contain error from the auxiliary data sources as well as the model error. The quality of SAEs thus depends not only on the quality of the underlying survey data but also on the quality of the auxiliary data and the model used to create the estimates. As such, just as survey estimates are presented with information about the survey methodology and measures of uncertainty such as standard errors or confidence intervals, SAEs should be presented with additional information about the auxiliary data and modeling procedures along with information about the quality of the estimates.
- Related to the above, SAEs include modeled or synthetic estimates. As such, an area's SAE may differ from the true value of the outcome when there are key explanatory variables that impact the outcome that are not included in the model. As a hypothetical example, two areas with similar demographic characteristics and administrative data (and thus similar synthetic estimates) may truly differ on fruit and vegetable consumption due to the presence of a farmer's market in one county but not in the other. Users of SAEs should consider the data included in the model along with local conditions or characteristics not included in the model that may impact the accuracy of a given SAE.

The goal of the SAE for the 2020 Virginia BRFSS was to provide estimates of 31 key health outcome measures for the Commonwealth of Virginia's 133 counties and independent cities.

Exhibit 1: County/City Summary Table

County/City	Unweighted N
Accomack County	94
Albemarle County	138
Alleghany County	40
Amelia County	26
Amherst County	56
Appomattox County	50
Arlington County	207
Augusta County	132
Bath County	11
Bedford County	130
Bland County	11
Botetourt County	49
Brunswick County	27
Buchanan County	27
Buckingham County	27
Campbell County	101
Caroline County	26
Carroll County	28
Charles City County	7
Charlotte County	27
Chesterfield County	400
Clarke County	17
Craig County	12
Culpeper County	40
Cumberland County	18

County/City	Unweighted N
Dickenson County	21
Dinwiddie County	56
Essex County	27
Fairfax County	773
Fauquier County	80
Floyd County	29
Fluvanna County	28
Franklin County	63
Frederick County	77
Giles County	19
Gloucester County	84
Goochland County	26
Grayson County	22
Greene County	13
Greensville County	21
Halifax County	69
Hanover County	97
Henrico County	289
Henry County	73
Highland County	11
Isle of Wight County	38
James City County	95
King and Queen County	9
King George County	29
King William County	24
Lancaster County	37

County/City	Unweighted N
Lee County	36
Loudoun County	298
Louisa County	31
Lunenburg County	31
Madison County	22
Mathews County	31
Mecklenburg County	71
Middlesex County	24
Montgomery County	126
Nelson County	12
New Kent County	29
Northampton County	32
Northumberland Count	33
Nottoway County	26
Orange County	41
Page County	24
Patrick County	35
Pittsylvania County	119
Powhatan County	32
Prince Edward County	34
Prince George County	51
Prince William County	366
Pulaski County	51
Rappahannock County	10
Richmond County	76
Roanoke County	184

County/City	Unweighted N
Rockbridge County	45
Rockingham County	126
Russell County	43
Scott County	84
Shenandoah County	56
Smyth County	45
Southampton County	29
Spotsylvania County	118
Stafford County	142
Surry County	10
Sussex County	28
Tazewell County	75
Warren County	41
Washington County	84
Westmoreland County	41
Wise County	95
Wythe County	44
York County	70
Alexandria city	147
Bristol city	7
Buena Vista city	5
Charlottesville city	39
Chesapeake city	269
Colonial Heights city	33
Covington city	3
Danville city	26

County/City	Unweighted N
Emporia city	4
Fairfax city	57
Falls Church city	16
Franklin city	10
Fredericksburg city	36
Galax city	2
Hampton city	169
Harrisonburg city	22
Hopewell city	37
Lexington city	1
Lynchburg city	99
Manassas city	26
Manassas Park city	8
Martinsville city	10
Newport News city	165
Norfolk city	303
Norton city	4
Petersburg city	59
Poquoson city	9
Portsmouth city	167
Radford city	25
Richmond city	222
Roanoke city	155
Salem city	17
Staunton city	14
Suffolk city	104

County/City	Unweighted N
Virginia Beach city	437
Waynesboro city	6
Williamsburg city	13
Winchester city	14
Total	9,552

Please note that the unweighted (raw) count of completed interviews by county represent the total of the 9,552 2020 Virginia BRFSS cases. They do not represent the number of unweighted cases used in the small area estimation (SAE) for each of the thirty-one health indicator analyses. Both unweighted and effective sample sizes are indicator specific – because of the item missing data for all outcomes (due to either nonresponse when the respondent provided a "Don't know" answer or refused to answer, or due to the case coming from a different state, and not filling in the VA-specific modules) and because of subsetting the sample by age, sex or other appropriate characteristics (e.g. shingles vaccination age 50+). Therefore effective sample sizes by county are not presented in this report, because data files of different sizes (number of cases) were used through the SAE process, utilizing all available data.

2. Small Area Estimation Procedures

Small area estimation (SAE) is an area of active growth and research in survey statistics that originated in late 1970s (Fay and Herriot 1978) and has flourished in 1990s and later, driven by growing demands for detailed estimates from data users accompanied with greater availability of computing power. SAE addresses the problem of obtaining reasonable estimates for domains where small sample sizes do not allow direct estimation using survey data only (including domains with zero sample observations), e.g., at the levels of a county or a metro area in national and state samples. An encompassing reference on SAE is Rao and Molina (2015)[1].

The modern approach to SAE involves the use of statistical models to predict the outcome of interest, such as current smoking. Direct estimates from the survey data are combined with synthetic estimators from statistical models to create a composite SAE.

Each of the thirty-one health indicator output files contain results by county. These outputs include the one-sided lower and upper bounds at the 95% confidence interval and the point (best) estimate as well as bell curve charts with the two-tailed 90% confidence interval represented by gray markers.

[1] Rao, J. N. K., and Molina, I. (2015). *Small Area Estimation.* 2nd edition. Wiley Series in Survey Methodology. Hoboken, NJ: Wiley.

3. CDC BRFSS SAE Methodology

3.1 BRFSS SAE Estimation

The Centers for Disease Control and Prevention (CDC) has developed a SAE system for a portion of the BRFSS data referred to as SMART (Selected Metropolitan/Micropolitan Area Risk Trends) for counties that have sample sizes of at least 500. Pierannunzi et. al. 2016[2] outline the CDC method to model health outcomes at the county level using this procedure (referred to as SMART-SAE), which is described below. *Our comments regarding the implementation of these procedures and potential improvements, are provided as sub-items and are italicized.*

1. BRFSS data are re-raked to state-level demographic variables (age by gender, race/ethnicity, education, marital status, (housing) tenure, gender by race-ethnicity, age by race-ethnicity, region by age, gender, and race-ethnicity), as well as county-level targets (county by sex, age, and race) using Nielsen Claritas data. Weights are then rescaled to the nominal sample size for the county.

    a. *The Nielsen Claritas data set was apparently chosen for historic reasons, as it was used in prior work by the CDC for similar purposes. The American Community Survey (ACS) based margins appear to be more reliable.*

    b. *There is evidence (Pfeffermann et. al. 1998) that scaling by the effective sample size, rather than by the nominal sample size, works better in reducing small sample biases of variance parameters in mixed models.*

2. BRFSS outcomes are imputed using a hot-deck procedure.

    a. *The hot-deck procedure is somewhat restrictive in that it has a certain low-dimensional structure in mind. Namely, that the missingness is conditionally independent of the outcomes within the imputation cells. We believe that the regression model for the outcome that will be proposed at a later stage does a similar or a better job incorporating the missing data in the outcomes.*

    b. *In addition, imputation introduces a source of variation in the data that needs to be accounted for in the standard errors of the SAEs down the line. To incorporate the imputation variation correctly, multiple imputation procedures and Rubin (1978) rules should be used. It is unclear whether that was done in SMART-SAE by the CDC.*

    c. *Given this, for our work with the Virginia BRFSS SAE, we chose to forego the imputation step, as the demographic predictors were used in the model for the outcome, which would have been done in an imputation model, anyway, as well.*

[2] Pierannunzi C, Xu F, Wallace RC, Garvin W, Greenlund KJ, Bartoli W, et al. A Methodological Approach to Small Area Estimation for the Behavioral Risk Factor Surveillance System. *Prevention of Chronic Diseases* 2016;13:150480. DOI: http://dx.doi.org/10.5888/pcd13.150480.

3. ACS public use microdata series (PUMS) data are obtained.

    a. *Pierannunzi et. al. (2016) describe this step as creating a single-year data set from multiple years of ACS data. We believe it is better to use the data set created by the Census Bureau specifically for the purposes of providing sufficient sample sizes at low levels of geography, namely the 5-year data set. For 2020 Virginia BRFSS SAE we used the ACS 5-year data set for 2015–2019. The 2020 ACS data set is released as "experimental" since ACS failed to reach the target response rate because of the impact of COVID-19 pandemic on the field, and hence the data cannot be fully endorsed as official. Delivery of the data sets was delayed, as well, and the 5-year data for 2016–2020 was not available at the time of completion of this project.*

4. ACS data are approximately subset to county levels. The finest level of geography provided in ACS PUMS data is that of the public use microdata area (PUMA), a contiguous geographic area with a total population of about 100,000. Large counties can be split into several PUMAs, and smaller counties are aggregated into a single PUMA, so there is no 1:1 relation. Population fraction weights (% of the PUMA population found in a county) are used to distribute the total population of a PUMA to its component counties if needed. SMART-SAE used the data from Missouri Census Data Center (http://mcdc.missouri.edu/websas/geocorr.html) to obtain these fractions.

    a. *Alternatively, accurate county variables are available in the protected ACS data available to researchers through research data centers. We did not have access to these, and utilized the MCDC conversion factors, as did Pierannunzi et. al. (2016).*

5. ACS data are re-raked to Nielsen Claritas data.

    a. *The utility of this step is unclear to us. ACS is a better-quality data set than any commercial data set. The only reasonable justification is to align the totals to the same ones used in BRFSS raking. Since we do not have the Nielsen Claritas data set we cannot perform the SMART-SAE raking steps 1 and 5. However, we can rely on using the low level ACS data (as detailed in 3.a. above).*

6. ACS and BRFSS data are stacked together to prepare for modeling.

7. A logistic random effects model with county as a random effect, and race, gender, and age as the main effects, is fit to the data.

    a. *The model is known as the unit-level model (Rao and Molina 2015, Sec. 4.2), where modeling happens at the level of an individual. The use of the model for SAE assumes that the values of the explanatory variables are known for all units in the population – (see 5a in Section 3.2 Challenges and Solutions below). No data source exists for the U.S. population that can act as a complete population register with the required race, gender, and age information.*

    b. *An alternative is a model in which the response variable is the direct estimate for an area, and all explanatory variables are at the area level. This approach is known as an area-level model in SAE literature (Rao and Molina 2015, Sec. 4.3). Area-level*

models do not require the knowledge of the values of explanatory variables for all units in the population, and thus are easier to use in the U.S.

c. *The SMART-SAE model used by the CDC can be improved by adding county-level contextual variables. It is unclear to us why Pierannunzi et. al. (2016) did not incorporate county-level predictors. In our experience detailed below, such predictors improve performance of the model.*

8. Predicted probabilities are obtained for the age-gender-race cells within a county.

a. *This was not spelled out as a separate step by Pierannunzi et. al. 2016, arguably because its implementation is through the single SAS PROC GLIMMIX procedure call utilized in step 7 listed above. However, there are a variety of ways to create predictions. In Stata, which we use for modeling, the model fitting steps and generating prediction steps are separated in syntax.*

b. *Pierannunzi et. al. (2016) use the language of "best linear unbiased prediction". However, these do not exist for the nonlinear models like logistic regression used here. The appropriate concept for the binary data are empirical best predictions (Jiang and Lahiri 2001[3]).*

c. *Aggregation of predictions to the county level is not described in the Pierannunzi et. al. paper in sufficient detail. Several implementations are possible.*

i. *The predicted probabilities (incorporating the random effect of the county) could be obtained for the units in the ACS PUMS using the regression coefficients obtained on BRFSS data. These predicted probabilities can then be added up with their appropriate ACS weights (split between counties as needed, as explained above in item 4c) to form county-level SAE.*

ii. *The predicted probabilities for all age-gender-race groups within a county can be obtained (incorporating the random effect of the county), and then these predicted probabilities can be added up using the estimated proportions of these population groups using aggregate ACS data available at http://data.census.gov (CEDSCI system). This is the essence of the multilevel regression with poststratification (MRP) approach. This approach may produce more precise estimates, as only a fraction of ACS data is released as the public use data, while the CEDSCI tables are based on the complete ACS data set. However, computation of the standard errors for these predictions through aggregation of cell-level standard errors is difficult.*

d. *In a strict sense of SAE methodology, SMART-SAE appears to be synthetic estimates, in that they only use the estimated model. Better estimates can be obtained by combining the synthetic estimates with the direct survey estimates into composite estimates. The latter can increase the effective sample size used in estimation, and*

[3] Jiang, J., and Lahiri, P. (2001) Empirical Best Prediction for Small Area Inference with Binary Data. *Annals of the Institute of Statistical Mathematics*, **53**, 217–243.

better protect against possible model violations, such as important covariates not included in the model.

Previous work conducted by the CDC[4] compared a number of possible approaches to SAE, including:

1. Unweighted logistic random effects model;
2. Weighted logistic random effects model;
3. Multi-level model;
4. Aggregation over time (7-year window);
5. Empirical best linear unbiased prediction, weighted and benchmarked;
6. Constrained and benchmarked model (essentially, a highly detailed raking procedure); and
7. Bayesian SAE.

The weighted logistic regression performed best, and was chosen as the backbone of the SMART-SAE procedure outlined above.

3.2 Challenges and Solutions

From the total survey error perspective, aside from measurement errors, the errors in representation and modeling in the proposed SAE estimation plan are detailed below.

Challenges:

1. Coverage error for the non-telephone population. Extrapolation to the non-telephone population is implicitly performed by using the ACS areal frame data. However, the limited use of the person-level demographic variables effectively assumes that phone coverage is sufficiently well explained by these variables only. In formal missing data language, individuals who do not have phones and are not covered by the BRFSS sampling design are assumed to be missing at random conditional on the demographics used in BRFSS weighting and/or in our SAE modeling (but not missing completely at random).
2. BRFSS nonresponse error (partially compensated by weighting).
3. For Steps 1 and 5 of the CDC's SMART-SAE procedure as detailed in *Section 3.1*, the county-level population estimates used as targets for raking are likely to have nonzero levels of uncertainty. The resulting BRFSS estimates thus partially inherit sampling error from the ACS which can be nontrivial for small counties where ACS sample sizes may be in the low hundreds. Sampling errors in targets can be quantified for the ACS data;

[4] Gotway-Crawford, C., D. Ford and C. Pierannunzi (2014) "Comparison of Small Area Estimation Methods for use by the BRFSS". Presentation at the annual AAPOR conference.

however, properly accounting for them downstream becomes a complicated exercise (Dever and Valliant 2010[5]).

4. Overall, the utility of step 5 of the CDC's SMART-SAE procedure as detailed in *Section 3.1* is unclear, as the direct estimates for which the re-calibrated weights may arguably provide improved inference with lesser biases are never obtained.

Solutions:

1. If county-level variables are to be used in modeling the outcomes, a very broad range of possibilities opens up for the data sources. We identified the following data sets that can be used as sources for the county-level explanatory variables:

   a. **ACS Tables** (the possibilities are really endless, as CEDSCI contains several thousand tables at the county level)**.** It can be reasonably expected that some of the demographic variables may have predictive power. Also, socioeconomic status variables (poverty rates and the use of government assistance programs) and limited health-related variables (health insurance status) may turn out to be helpful.

   b. **Census Planning Database.** While this database mostly reuses the ACS data, it also contains additional variables such as the Census mail return rates. It is somewhat unclear whether these field process metadata variables can be useful in modeling health outcomes, but they may serve as proxies for the willingness of the population to cooperate with the government (which in turn may be associated with the health outcomes in programs that have explicit health policies, such as vaccinations or smoking).

   c. **National Center for Education Statistics Data** for school districts that, in case of Virginia, align with counties. Variables such as percent of students eligible for free or reduced price lunch are proxies for the county-level socio-economic status, while teacher-to-student ratio, availability of the local government funding.

   d. **U.S. Environmental Protection Agency Environmental Quality Index (EQI).** This index includes the measured concentrations of certain known pollutants, aggregated to the air quality, water quality and land quality indices, sociodemographic domain and built environment domain quality indices, and the overall environmental equality index.

   e. **County-level health estimates from the 2000 SAE project of the National Cancer Institute** (http://sae.cancer.gov; Raghunathan et. al. 2007). Four estimates are available: current smoking, ever smoked, Pap smear, and mammography.

[5] Dever, J. A. and R. Valliant (2010). A comparison of variance estimators for poststratification to estimated control totals. Survey Methodology 36 (1), 45-56.

f. **County Health Rankings,** a collaboration between the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute (http://www.countyhealthrankings.org/app/virginia/2018/overview).

Additionally, the Virginia Department of Health provided the following data sets which were used in the modeling:

a. Age-Adjusted Malignant Cancer Incidence Rates and Counts for Selected Cancer Sites by sex, VA 2013.

b. Mortality rates by source and age, VA 2010–2014.

c. Hospitalization rates, by source, VA 2005–2013.

Finally, we used the SAE estimates from the 2014 and 2017 rounds.

Challenges:

5. Given the multitude of the possible area-level variables available, the issue of selecting the best ones will arise. There is a number of helpful developments in the SAE literature. Pfeffermann (2013[6]) proposed to use conditional Akaike information criterion to find the best fitting model given the eventual focus on prediction for the existing set of areas (rather than a generalization to a fictional universe of all possible areas implied by the Gaussian distribution of the random effects). Application of this proposal would require fitting a broad number of random effects models, which does not appear to be computationally feasible (especially given the interest of the current project in 36 health indicators, each of which will likely require its own model). Lahiri and Suntornchost (2015)[7] modified the estimate of the residual MSE in the Fay-Herriot model for the contribution of the design variances, and suggested plugging the modified estimator into standard diagnostic measures such as AIC, BIC, Mallows' C_p for i.i.d. regression models. An entirely different set of model selection approaches is available in Bayesian paradigm in the form of posterior predictive checks (for applications to SAE, see Moura and Migon 2002[8]); since we do not use Bayesian approaches in this project, these alternatives are not pursued.

   a. A faster search can be conducted with more aggregated data that would not require mixed modeling.

[6] Pfeffermann, D. (2013). New important developments in small area estimation. *Statistical Science* 28 (1), 40-68.

[7] Lahiri, P., and Suntornchost, J., (2015). Variable selection for linear mixed models with applications in small area estimation. Sankhya B , 77(2), 312–320.

[8] Moura, F. A. S., and Migon, H. S. (2002). Bayesian spatial models for small area estimation of proportions. *Statistical Modelling*, 2, 183–201.

i. Fay-Herriot (1978[9]) area-level regression model is formulated, where the dependent variable is given by the direct estimates of the outcome at the area level (transformed to stabilize the error variance), and explanatory variables are observed at the area level, as well.

ii. Least absolute sum shrinkage and selection operator (lasso; Efron and Hastie 2016[10]) can be fit quicky to an extensive set of predictors to identify the relevant ones.

b. To account for the potential model selection error, a limited number of the best fitting models can be retained for each outcome. Variation between the estimates based on the different models serve as a measure of the error due to the model uncertainty.

6. The sampling error in the regression coefficients (step 4 in the SMART-SAE procedure) needs to be accounted for.

7. The sampling error in the ACS aggregation (step 5 in the SMART-SAE procedure) needs to be accounted for. We utilized a faster Taylor series linearization using the strata and cluster variables provided with the IPUMS data.

8. Self-contained analytical expressions for the Mean Squared Error (MSE) of the SAEs (Rao and Molina 2015) account for the sampling and the model fit error when a fixed model is fit to the (area-level or unit-level) outcomes. In other words, these expressions would be appropriate if a fixed model was fit to the BRFSS data, and predictions were obtained directly from that model. As far as we can see, additional steps are required:

   a. Extrapolation to the ACS data (step 7 of the SMART-SAE procedure);

   b. Model selection (issue #5 in the current list).

[9] Fay, R. E. and R. A. Herriot (1978). Estimates of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association* 74 (366), 269-277.

[10] Efron, B., and Hastie, T. (2016). *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge University Press.

4. Virginia BRFSS Small Area Estimation

Exhibit 2: Indicator Summary Table

Indicator #	Indicator Name
1	Binge Drinking- 2014, 2017, 2020
2	Colorectal Cancer Screening- 2014, 2017, 2020
3	Overweight or Obese – 2014, 2017, 2020
4	No Physical Activity in the Past Month- 2014, 2017, 2020
5	Poor Mental Health- 2014, 2017, 2020
6	Has a Regular Healthcare Provider- 2014, 2017, 2020
7	Ever Diagnosed with Diabetes- 2014, 2017, 2020
8	Flu vaccination past year – 2017, 2020
9	Ever Had Arthritis- 2014, 2017, 2020
10	Current Smoker- 2014, 2017, 2020
11	Ever Had Asthma- 2014, 2017, 2020
12	Dental Visit in the Past Year- 2014, 2017, 2020
13	Poor Health: Physical or Mental Health – 2014, 2017, 2020
14	Tetanus vaccination in the past 10 years – 2017, 2020
15	Depressive Disorder- 2014, 2017, 2020
16	Ever Had Stroke- 2014, 2017, 2020
17	Ever Had Heart Attack- 2014, 2017, 2020
18	HIV Test Ever- 2014, 2017, 2020
19	Diabetes Test Past Year- 2014, 2017
20	Current E-Cigarette Use (among tobacco users)- 2017, 2020
21	Ever Had COPD- 2014, 2017, 2020
22	Pneumonia Shot Ever- 2014, 2017, 2020
23	Adverse Childhood experiences, percentage who reported three or more ACE experiences – 2017, 2020
24	Ever Had Coronary Heart Disease, 2014, 2017, 2020
25	Pre-diabetes, 2014, 2017, 2020
26	No Health coverage, 2014, 2017, 2020
27	No doctor due to cost, 2014, 2017, 2020
28	Any tobacco use – 2014, 2017, 2020
29	Routine Checkup last year – 2017, 2020
30	Tetanus, diphtheria or TDAP 2014, 2017, 2020
31	Shingles Vaccine ages 50+ - 2014, 2017, 2020

4.1 Health Outcomes of Interest

Thirty-one outcomes of interest were analyzed in this project, listed in Exhibit 2. (Note that in other years, the internal indicator numbers may have been different.) Some of the outcomes analyzed with the 2017 data could not be analyzed as the Virginia state module was not administered in 2020, and only the core module data were available.

4.2 Data

The data set used for the SAE was the 2020 Virginia BRFSS, of which there were 9,552 cases. In the demographic section of the survey respondents are asked what county they live in. For the 2020 Virginia BRFSS, 8,589 (89.9% unweighted) respondents provided a response of a county within Virginia (`ctycode2` variable); for other cases, an open text entry was encoded by CDC (`cpcounty` variable). According to the BRFSS SMART procedure, a respondent's missing county was imputed by CDC according to the following process:

- For landline numbers, the frame county (based on the most prevalent county in a given 100-block) was used.
- For cell phone numbers, the county was coded based on:
    a. An open-ended self-reported location;
    b. Self-reported ZIP code; or
    c. For the records lacking the above, the largest county population by age and race/ethnicity was used.

The accuracy of the latter imputation step is likely to be low, and it may bias estimates for the counties affected (mostly, the counties into which the imputation was made).

4.3 Virginia BRFSS SAE Procedures

Retaining the main steps and methods of the CDC SMART-SAE procedure outlined above in *Section 3. CDC BRFSS SAE Methodology*, we used the following SAE algorithm for the 2020 Virginia BRFSS.

1. Produce direct estimates of the health outcomes at the county level.
2. Perform lasso selection of a concise predictive Fay-Herriot type area-level model.
    a. The first pass of lasso selection used the main effects, relied on 10-fold validation and lasso BIC to select the best main effects model. Out of this model, up to five variables (those that first entered the lasso path) were interacted with other main effect variables.
    b. The second pass of lasso was performed on the set that included the full list of variables from the first pass, and the interactions, and the model with the lowest cross validated MSE minus one s.e. rule was selected as the final model.
    c. A set of up to five models between that selected by the one standard deviation rule and the one selected by the best cross-validated MSE was retained for each

outcome where the selected models would have lasso BIC within 3 units of the best BIC and would have positive cross-validated R^2_{adj}.

3. Fit a weighted logistic regression model to the BRFSS data with outcomes from the survey data, age, race, and gender from the survey data, and the county-level predictors selected in step 2
4. Fit a mixed logistic model with both area level and unit level predictors. (For most outcomes, the variances of random effects were estimated to be zero, and the model reduced to the generalized linear model (GLM) weighted logistic regression model with these predictors.)
5. Obtain the (unit-level) predicted values from the final model for the matched ACS PUMS data.
6. Obtain the (area-level) point estimates as weighted averages across areas of the predictions obtained in the previous step using the ACS PUMS weights.
7. Obtain the (area-level) variance component due to the regression model parameters being estimated using the delta method and the variance-covariance matrix of regression parameter estimates.
8. Obtain the (area-level) variance component due to the posterior empirical Bayes distribution of area effects, if applicable (estimated variance > 0).
9. Obtain the (area-level) ACS sampling standard errors using the Taylor series linearization method.
10. Assuming that the model selection error is independent of both the ACS sampling error and the SAE MSE, calculate the between-model error as a variance within a small subset of the best-fitting candidate models from step 2.
11. Add the variances obtained in steps 7–10 above (i.e., assuming independence between the sources of error), and use the square root of thus estimated MSE to report as the standard error of the SAEs obtained in step 6.

This procedure was repeated for all 31 indicators of interest for all 133 counties and independent cities.

The project consists of about 1,900 files, including 32 human-created Stata do-files, 31 automatically created do-files for model search, 31 automatically created do-files for mixed model estimation, 275 intermediate data files and 564 graphic files.

4.4 Modeling results

The big picture of modeling results is given in Exhibit 3. Our initial expectations were:

1. Most outcomes would have several area-level predictors associated with them.

Exhibit 3: Model summaries.

Outcome	Area covariates used?	F-H cross-validated R2	Var[area effect] >0?	Prior years SAE among predictors?	Models used	SAE outside direct CI	CIs partial / don't overlap
1: Binge Drinking	Yes	0.285	Not sig / No	No	6	11	18 / 0
2: Colon Cancer Screening 45+ past year	Yes	<0	No / No	Yes; not sig	5	4	24 / 0
3: Overweight or Obese	Yes	0.239	Yes / Not sig	Yes; not sig	3	10	21 / 0
4: Physical Activity in the Past Month	Yes	0.346	Yes / No	No	6	9	26 / 0
5: Any Days Mental Health Not Good	No	<0	No / No	No	6	13	28 / 0
6: Has personal doctor	Yes	0.218	Not sig / Not sig	No	6	15	36 / 0
7: Ever Diagnosed with Diabetes	Yes	0.275	Not sig / No	Yes; not sig	6	5	18 / 0
8: Flu Vaccine Past Year	Yes	0.222	Yes / No	No	6	9	14 / 0
9: Ever had arthritis, gout, lupus	Yes	0.589	Yes / No	Yes; not sig	6	22	34 / 0
10: Current Smoker	Yes	0.411	Yes / No	Yes; not sig	6	5	11 / 0
11: Ever had asthma	No	<0	Not sig / Not sig	No	2	1	5 / 0
12: Dental visit in the past year	Yes	0.276	Yes / Not sig	Yes; not sig	6	9	24 / 0
13: Poor physical or mental health days	Yes	0.118	No / No	No	6	2	6 / 0
14: Tetanus Vaccination in the Past 10 Years	Yes	0.196	Not sig/No	Yes; not sig	5	5	17 / 0
15: Ever Diagnosed with Depression	Yes	0.224	Not sig/No	No	6	12	18 / 0
16: Ever had a stroke	Yes	0.258	Not sig/No	No	6	2	12 / 0
17: Ever had a heart attack	Yes	0.206	No / No	Yes; not sig	6	1	6 / 0
18: Ever Tested for HIV	Yes	0.383	Yes / No	Yes; not sig	6	9	22 / 0
19: Diabetes test past 3 years	No	<0	Not sig/Not sig	No	5	15	24 / 0
20: Current E-Cigarette Use Among Tobacco Users	Yes	0.072	No / No	No	1	6	21 / 0
21: Ever had COPD	Yes	0.335	Yes / Not sig	Tobacco use	6	1	8 / 0
22: Pneumonia Vaccination	Yes	0.263	Not sig/No	No	4	16	32 / 0
23: 3+ adverse childhood experiences	Yes	0.167	Not sig/No	No	6	7	21 / 0

Outcome	Area covariates used?	F-H cross-validated R2	Var[area effect] >0?	Prior years SAE among predictors?	Models used	SAE outside direct CI	CIs partial / don't overlap
24: coronary heart disease	Yes	0.132	No / No	Heart disease mortality	6	4	9 / 0
25: pre-diabetes or borderline diabetes	No	<0	No / No	No	2	5	8 / 0
26: Health Care Coverage - Ages 18-65	Yes	<0	Yes / No	No	6	13	36 / 0
27: Doctor Inaccessible Due to Cost	Yes	0.029	Yes / Not sig	No	6	5	18 / 0
28: Current Tobacco Use	Yes	0.450	Yes / No	Yes; sig	6	4	14 / 0
29: Routine Checkup Past Year	Yes	0.257	Not sig/No	No	6	11	27 / 0
30: Tetanus Diphtheria Tdap shot	Yes	0.136	Not sig/No	No	6	6	17 / 0
31: Shingles Vaccination - Ages 50+	Yes	<0	Not sig/No	Yes; not sig	5	8	29 / 0

Notes:

1. Outcome sequential number and short label.
2. Are covariates used? Yes: the final model includes (fixed) person-level and area-level covariates; No: the final model only includes the fixed person-level covariates because the model selected by lasso was intercept only.
3. Fay-Herriot model cross-validated R^2_{adj} statistic.
4. Is the mixed model estimate of random effect variance positive? No: the estimated variance is computationally zero (e.g. reported as 0.000) regardless of declared significance; Yes: significant at level 5% or below; Not sig[nificant]: a positive estimate is reported but the estimate is not significant at 5% level (algebraically equivalent to fewer than 8 d.f.s obtained via chi-square approximation of the variance component).
5. Prior years SAE among predictors: were the estimates of the same indicator selected by lasso in the Fay-Herriot regression? ("Yes" is an indicator of the overall internal validity of the SAE exercise, as being performed with consistent methodology over time.)
6. Models used: number of models used to estimate the model uncertainty component (capped at main + five closest alternative models that had reasonably close performance; for some outcomes, there were fewer models)
7. SAE outside direct CI: number of areas for which the SAE point estimates are outside the direct confidence interval, expected to happen about 5% of the time ~ 6.65 areas on average.
8. CIs partial / don't overlap: Number of areas for which the SAE and direct confidence intervals overlap partially, or not at all (the latter never happened). The complement, and the most typical result, is that the SAE CI is entirely contained in the direct CI. Frequent discord between CIs can be understood as evidence of poor model performance although the judgment is qualitative.

2. The area-level models would not predict the outcome perfectly, so there would be remaining error variances associated with the models.
3. If there is inertia in the health outcome trajectories at the population level, SAEs from prior years would be correlated with the current prevalence of an outcome and appear in the corresponding regression models.
4. The R^2 of the regression models would be in the range typical for aggregated behavioral data, i.e., from 10% to 50% (in the person-level models, R^2 is typically lower, and does not often exceed 10%, but for aggregate data, the trends may be more pronounced).
5. There would be agreement between the SAEs and direct estimates:
    a. Direct estimates are assumed unbiased, but they have wide confidence intervals.
    b. SAEs are more efficient, their CIs are narrower, and would be expected to lie entirely inside the direct CIs or have substantial overlap with the direct CIs.
    c. In the ideal situation of SAEs nailing the true population prevalence's exactly, the number of times these estimates would be outside the 95% direct estimate confidence interval would follow binomial distribution with n=133, the number of areas, and p=0.05, the significance level.

Let us address the degree to which these expectations held with the SAE modeling exercise.

In all but seven of the final models with both unit-level demographic variables and area-level aggregated variables, the area-level random effect variance was estimated to be zero. Exhibit 4 provides the estimates. Rao and Molina (2015, Sec. 9.3) discuss this as an undesirable artefact of mixed models, as in this case, compositing of the model-based and direct-estimates is limited, and the estimates for the areas with larger effective sample sizes do not benefit from greater precision of direct estimates. With zero area variances, model-based estimates should be interpreted as interpolation/extrapolation of the area-level effects using the selected covariates. If such model approximation is subject to specification error, interpolation and especially extrapolation errors may result. The issue is partially controlled for by incorporating the variance component over the range of several best performing models. The number of such models is reported under the model table, and ranges from 1 to 6. To the extent that a given area may be outlying on its area-level covariates, and thus at risk of extrapolation biases, it will also be likely to have a greater variability of estimates obtained from the competing models. As the latter component is incorporated into the SAE standard errors, we believe that risks of extrapolation errors are mitigated. Outcomes that had non-trivial variances of random effects were:

3. Overweight or obese
6. Has personal doctor
11. Ever had asthma
12. Dental visit past year
19. Diabetes test
21. Ever had COPD
27. Doctor inaccessible due to cost

Note that in the 2014 VA BRFSS SAE project, all models had zero final random effect variances, while in the 2017 VA BRFSS SAE project, the outcomes with non-zero final random effect variances were different. We believe the current approach based on lasso is somewhat more conservative in model selection, as it results with models retaining non-zero estimates of area variances more often (which, as pointed out above, is desirable).

Estimated Intraclass Correlation Coefficient (ICC) for counties varies from 0 to 7.0% (insurance coverage without area predictors) in the models with demographic variables only. In the models with area predictors, the only notable ICCs are 1.2% for asthma and 1.3% for the doctor inaccessibility due to cost. Here, ICC is estimated as

$$\mathrm{ICC} = \frac{\sigma_u^2}{\sigma_u^2 + \pi^2/3}$$

where $\pi^2/3$ is the variance of logistic distribution. In other words, the reported estimated standard deviation of the area variance should be compared to the standard deviation of the logistic distribution, $\pi/\sqrt{3} = 1.814$, to gauge the relative importance of the area effects, on top of the demographic effects, unaccounted for in that model. Note that some important person-level covariates that have not been used neither in Pierannunzi et al (2016) nor here was socio-economic status (which can be proxied by income and education) or interactions of demographic variables (e.g., age by gender). Kolenikov and Pitblado (2015)[11] provided a summary of ICCs for descriptive statistics across health research and found that these range from 0 to high single digit percentage points. For statistical models like those used in this SAE modeling effort, ICCs should be lower as some of the variability would be explained by the predictors.

The models exhibit complicated patterns of variability between areas. Some of that variability is captured by the area effect variances, and other components may be captured by the explanatory variables. E.g., for outcomes "20. E-cigarette use" and "24. Coronary Heart Disease", the model without the area covariates had zero estimated area variances, which would typically be interpreted as area-level information not being able to contribute to the outcome explanation, yet the full model had significant area-level predictors. Also notably, for outcomes "5. Nonzero days of mental health issues", "11. Ever diagnosed with asthma" and "25. Diagnosed with prediabetes", the final models contained no covariates, and had estimated zero area-level variances. In other words, for these models, all of the explanation and information borrowing across areas comes from the person-level demographic characteristics.

For both direct estimates and composite estimates, the effective sample size is calculated by reversing the variance of an i.i.d. statistic formulae, namely:

$$n^* = \frac{\hat{p}(1-\hat{p})}{(s.e.[\hat{p}])^2}$$

where the standard error is either the design-corrected standard error for the direct estimate, or the composite standard error for the composite estimate (see Step 11 of the SAE algorithm above).

[11] Kolenikov, S., and Pitblado, J. (2015). Analysis of Complex Health Survey Data. Ch. 29 in: Johnson, T. P. (2015), *Handbook of Health Survey Methods*. Wiley Handbooks in Survey Methodology: Wiley, Hoboken, NJ.

Exhibit 4: Estimated Standard Deviations of County-Level Random Effects.

	Demographic variables only			Demographic + area variables		
Outcome	**st. dev. of area effects**	**(std. error.)**	**ICC**	**st. dev. of area effects**	**(std. error.)**	**ICC**
outc1_binge_drink	0.1398	(0.1018)	0.0059	0.0000	(0.0467)	0.0000
outc2_colon_test	0.0000	(0.0000)	0.0000	0.0001	(0.0000)	0.0000
outc3_oweight_obese	0.2752	(0.0449)	0.0225	0.0829	(0.0900)	0.0021
outc4_phys_activity	0.3827	(0.0651)	0.0426	0.0000	(0.0218)	0.0000
outc5_ng_ment_hlth	0.0000	(0.0000)	0.0000	0.0000	(749.7023)	0.0000
outc6_doctor_person	0.1948	(0.0760)	0.0114	0.1702	(0.0887)	0.0087
outc7_diabetes_diag	0.2186	(0.0614)	0.0143	0.0000	(0.0001)	0.0000
outc8_flu_vacc	0.2686	(0.0509)	0.0215	0.0000	(1.0155)	0.0000
outc9_arthritis	0.3048	(0.0581)	0.0275	0.0000	(0.3311)	0.0000
outc10_curr_smoker	0.4866	(0.0649)	0.0671	0.0000	(0.0000)	0.0000
outc11_asthma	0.1998	(0.0708)	0.0120	0.1998	(0.0708)	0.0120
outc12_dentist	0.3514	(0.0767)	0.0362	0.0284	(0.6929)	0.0002
outc13_poorhlthdays	0.0000	(0.0000)	0.0000	0.0000	(1.1605)	0.0000
outc14_tetanus_vacc	0.1757	(0.0765)	0.0093	0.0000	(0.0000)	0.0000
outc15_depression	0.2382	(0.0812)	0.0170	0.0000	(0.0000)	0.0000
outc16_stroke	0.3707	(0.1615)	0.0401	0.0000	(1.3796)	0.0000
outc17_heart_attack	0.0000	(0.0000)	0.0000	0.0000	(0.0047)	0.0000
outc18_hiv_test	0.2128	(0.0506)	0.0136	0.0000	(0.0000)	0.0000
outc19_diabetes_test	0.0490	(0.2697)	0.0007	0.0490	(0.2697)	0.0007
outc20_ecig	0.0000	(0.0000)	0.0000	0.0000	(8.7838)	0.0000
outc21_copd	0.4124	(0.0715)	0.0492	0.1852	(0.1358)	0.0103
outc22_pneum_vacc	0.0603	(0.1686)	0.0011	0.0000	(0.1063)	0.0000
outc23_ace	0.2251	(0.0668)	0.0152	0.0000	(0.0000)	0.0000
outc24_heart_disease	0.0000	(0.0000)	0.0000	0.0000	(0.9213)	0.0000
outc25_prediab_diag	0.0000	(0.0000)	0.0000	0.0000	(3.7776)	0.0000
outc26_health_cover	0.4990	(0.0995)	0.0704	0.0000	(0.0000)	0.0000
outc27_doctor_cost	0.3506	(0.0743)	0.0360	0.2086	(0.0915)	0.0131
outc28_tobacco_user	0.4686	(0.0505)	0.0626	0.0000	(0.0000)	0.0000
outc29_rt_checkup	0.1130	(0.0444)	0.0039	0.0000	(0.1245)	0.0000
outc30_tdap_vacc	0.2094	(0.0615)	0.0132	0.0000	(0.0000)	0.0000
outc31_shingles_vacc	0.1440	(0.0669)	0.0063	0.0001	(0.0000)	0.0000

There was a considerable increase in the effective sample sizes associated with the SAE modeling as shown in Exhibit 5.

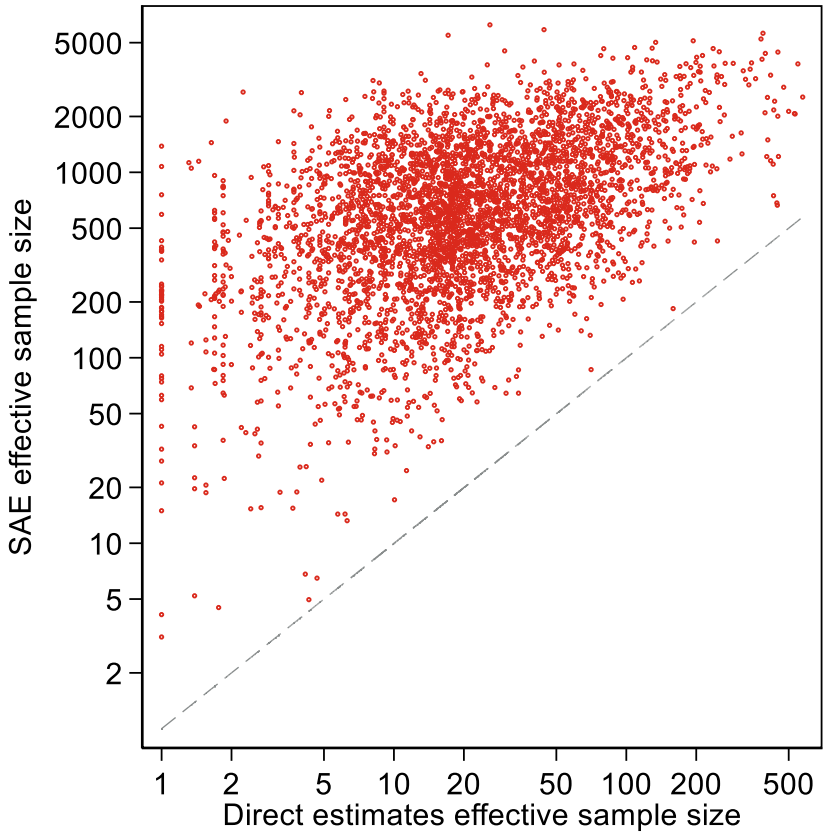


Exhibit 5: Effective sample size gain.

Face validity of the system of estimates is partially supported by presence of past estimates as predictors. In 11 models, such past estimates were selected by lasso (although in most cases, were sometimes rendered insignificant once person-level predictors were incorporated). For the "24. Coronary Heart disease" outcome, past heart disease mortality, a closely related concept, was among predictors. For "21. COPD", earlier tobacco use SAEs were among predictors, which is also related to lung disease prevalence.

As a part of internal verification check, synthetic simulated outcomes were created (i.i.d. with success proportions 10%, 25% and 50%; synthetic outcome correlated with gender only; synthetic outcome with moderate ICC of 4.5% for counties and no other correlates). For all of these simulated outcomes except the latter, the model selection correctly identified the empty area-level Fay-Herriot model no predictors and with negative cross-validated R^2. For the simulated outcome with geographic patterns, two area-level predictors were picked that resulted in cross-validated R^2 of 0.163.

In the ideal situation of SAEs nailing the true population prevalence's exactly, the number of times these estimates would be outside the 95% direct estimate confidence interval would follow binomial distribution with n=133, the number of areas, and p=0.05, the significance level. In reality, the SAE estimates will outside the direct CIs more often as they have their own error, likely more so when the

models perform poorly. The mean frequency of SAE outside direct CI is 7.90, slightly higher than the expected frequency 6.65 (=133*0.05), but not significantly so (p=0.179). The Pearson test rejects the null of the binomial distribution (test statistic = 199.19 for 15 degrees of freedom) with the unit bins, but the expected frequencies are all except one smaller than 5, so the validity of the test is suspect. The Q-Q plot (Exhibit 6) shows divergence of the distribution from the target, in the form of overdispersion: too many small values less than 3 (1st expected quantile), and too many large values above 12 (99th expected quantile). There is some evidence that the outlying values are associated with poorly performing models: of 7 such models in Exhibit 3, four have outlying values of the SAE occurrences outside the direct CIs.

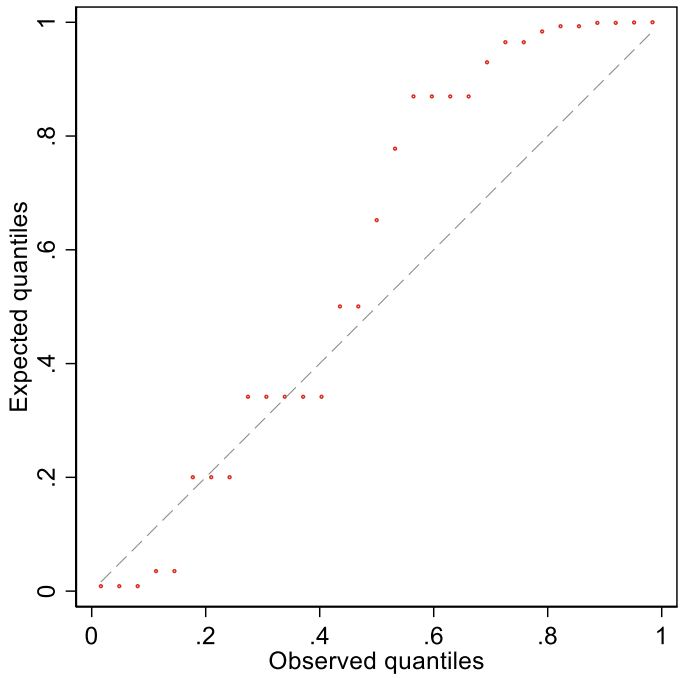


Exhibit 6: Q-Q plot of frequencies of SAEs outside direct CIs.

5. Reporting Small Area Estimation Results

The small area estimation output for each of the 31 outcomes is provided in four subsections: 1) definition of the variable and coding; 2) small area regression models; 3) estimates and confidence intervals; and 4) graphical representation of the results. Each of these is detailed below.

5.1 Definition of the Variable and Coding

The first subsection of the results gives the definition of the variable and how it is coded into a binary 0/1 variable. Note that some outcomes are negatively worded. For example, the poor health days variable has a value of 1 for those who report nonzero number of days when poor health prevented the respondent from daily activities. Another example is for the health insurance coverage variable, which has a value of 1 for those who do not have insurance.

5.2 Small Area Regression Models

The second subsection provides the selected SAE models that resulted from the model search. All models include demographic covariates following Pierannunzi et al (2016). For each outcome, three models are reported.

1. The first model is a weighted mixed logistic regression model with demographic variables only, and area effects included as random effects (i.e., no area level covariates are being used). This is the model that would have resulted from the SMART-SAE approach of Pierannunzi et al (2016).
2. The second model is a logistic regression model without area effects.
3. The third model is the mixed model with both demographic unit-level variables and aggregated area-level covariates in the best fitting Fay-Herriot model. This third model "Mixed with area covariates" is the final model used for the SAE.

5.3 Estimates and Confidence Intervals

The third subsection provides the numeric summaries for each county and independent city, grouped by health districts on each page. These summaries include the direct estimate (i.e., the estimate obtained from the survey data alone using BRFSS weights), its confidence intervals (CI), and the composite estimate with its confidence interval. The confidence intervals are Wilson confidence intervals. Dean and Pagano (2015[12]) found those to be among the most accurately and robustly performing ones in their comparison of seven different methods for proportion confidence intervals in complex surveys. For the proportion estimate $\hat{p}$ and effective sample size n^*, the Wilson confidence is found by solving for p the coverage equation

$$(\hat{p} - p)^2 \leq z^2_{1-\frac{\alpha}{2}} \frac{p(1-p)}{n^*}$$

which produces the confidence interval of the form.

[12] Dean, N. and M. Pagano (2015). Evaluating Confidence Interval Methods for Binomial Proportions in Clustered Surveys. J Surv Stat Methodology, 3 (4): 484-503. doi:10.1093/jssam/smv024

$$\frac{\hat{p} + \frac{z^2_{1-\frac{\alpha}{2}}}{2n^*} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n^*} + \frac{z^2_{1-\frac{\alpha}{2}}}{n^*}}}{1 + \frac{\left(z^2_{1-\frac{\alpha}{2}}\right)}{n^*}}$$

Wilson confidence intervals are always contained between 0 and 1 (unlike Wald confidence intervals $\hat{p} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n^*}}$), and exhibit more and more pronounced asymmetry as estimates get closer to zero or one.

5.4 Graphical Representation of the Results

In the fourth subsection for each outcome, SAE results are visualized using inchworm plots (Rhoda, 2016; see https://github.com/BiostatGlobalConsulting/inchworm-plots-stata). The plots are designed to visually convey some of the properties of the confidence intervals for proportions that are not necessarily obvious in the tabular representation of the estimates, their standard errors, and confidence intervals. First, confidence intervals are based on a density, with the highest, most likely values near the estimate, and the least likely values further away from it. Second, asymmetry of confidence intervals for estimates near zero or one and low values of $\hat{p}n^*, (1-\hat{p})n^*$ is, likewise, not immediately obvious from the numeric summaries. Inchworm plots make these features more prominent.

On each plot, the red curve and the shaded area beneath it represent an approximation to the posterior distribution of the estimate for the given area-and-outcome-specific value of the effective sample size n^*. The grey bookend markers represent Wilson confidence intervals. Both the estimates and the CIs are additionally reported in the numeric output panel on the right[13].

Consider the results for the first outcome, binge drinking, in Central Shenandoah and Lord Fairfax health districts (Exhibit 5). Most results are found to be in the 5% to 20% range. Several results are of note. First, one can note that low estimates are associated with asymmetric confidence intervals, such as those for Bath County and Highland County. Most other confidence intervals, however, are (approximately) symmetric, as sufficiently large (effective) sample sizes are available for these other counties and cities, and estimates are further away from zero. The second important observation is that the lack of information is reflected in very wide confidence intervals for Highland County and especially Buena Vista City. Also, since each curve represents an approximation to the distribution of the estimate with an area of 1, the wider confidence intervals for those two areas are associated with lower heights of the curves at peak/point estimate.

5.5 Additional Diagnostics

Some additional diagnostics regarding the performance of SAE models in comparison to direct estimates are collected in an additional diagnostic Excel file. Along with the county name,

[13] There are small discrepancies in the width and endpoints of the confidence intervals between those reported on the plots and those reported in the tables. We have been able to identify the differences as having to do with the different definitions of the effective sample sizes used in our code and in inchworm plots code.

County FIPS, Direct estimate, Direct 95% CI, SAE composite, and SAE 95% CI (same as reported in the Word file), the following information is included:

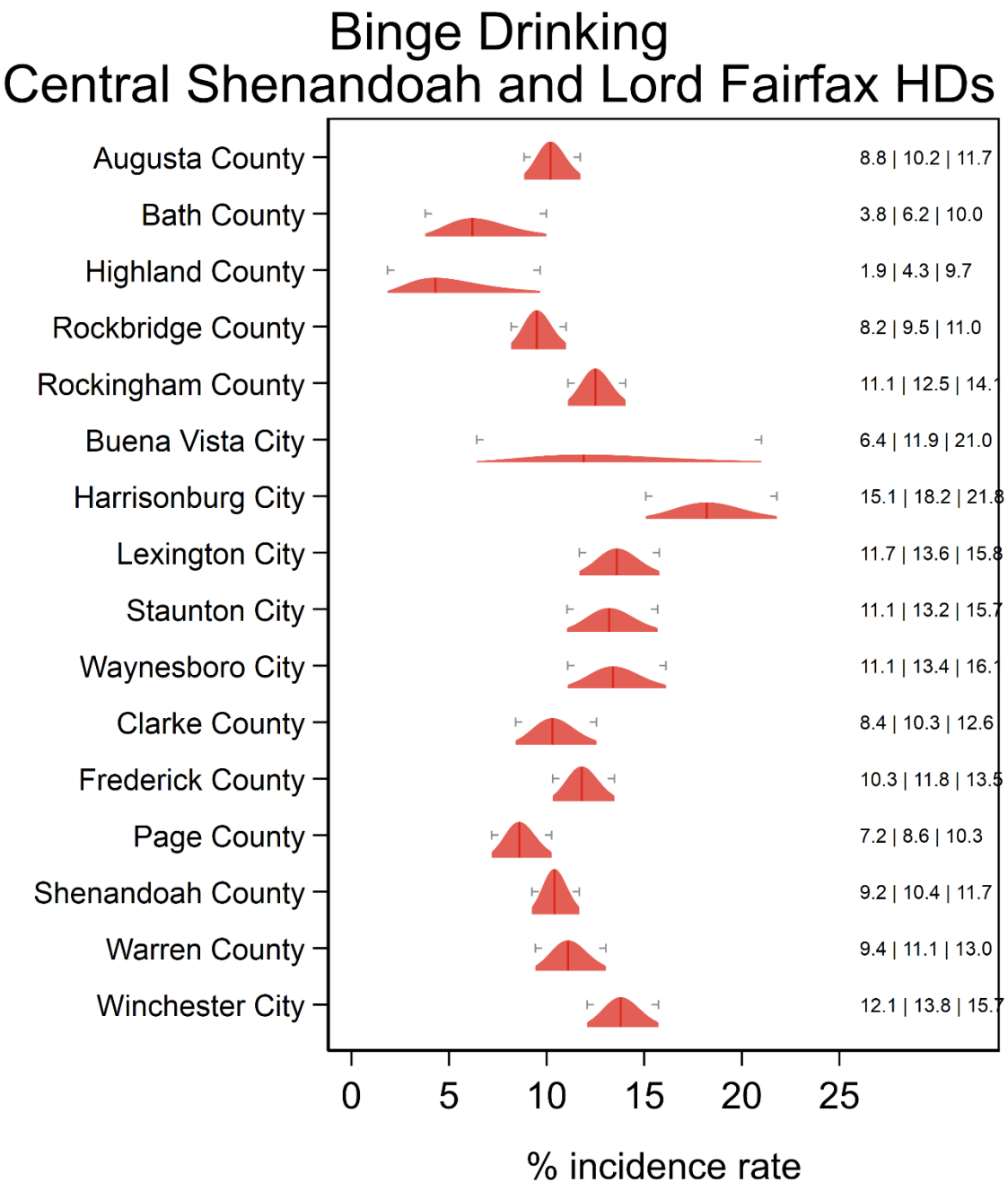
- **Direct Point vs. SAE CI**
  Results: "Check" if the direct estimate is inside the SAE CI; "Fail" if the direct estimate is outside of the SAE CI.
  Rationale: It is reasonable to expect that the direct estimates and composite SAE estimates will be close. However, since the SAE and direct estimates are not independent, it is difficult to say whether this event should happen. For most outcomes, the direct estimates are outside the SAE CI about half of the time.

- **SAE Point vs. Direct CI**
  Results: "Check" if the SAE estimate is inside the direct CI; "Fail" if the SAE estimate is outside of the direct CI.
  Rationale: It is reasonable to expect that the SAE estimate will be within the direct interval. We should expect this to happen quite often, as direct CIs are quite wide, and this diagnostic check is a very low bar to pass. If the model is successful in producing highly accurate estimates with low bias and low variance, then direct estimates, being unbiased, should contain the SAE estimates approximately 95% of the time.

- **SAE CI vs. Direct CI**
  Results: "Fully within" if the SAE CI is fully contained within the direct CI; "Some overlap" if SAE and direct CIs overlap, but each has portions not covered by the other; and "No overlap" if SAE and direct CIs do not overlap at all.
  Rationale: This is probably the most reliable check on the relation between the SAE and the direct estimates. When the SAE CI is fully within the direct CI, we can say that SAE helped zoom in on the range where the true value is likely to be. Lack of overlap should be particularly troubling.

- **SAE S.E. Less Than Direct S.E.**
  Results: "Check" if SAE standard error (s.e.) is less than or equal to that of the direct standard error; "Fail" if SAE standard error is greater that the (nonzero) direct standard error; "Zero direct s.e." if the direct standard error is zero (so the SAE is bound to be greater than the direct standard error).
  Rationale: We should reasonably expect the standard errors to go down in SAEs as we incorporate the regression model into the estimates.

Exhibit 5: Example of the inchworm plots, Binge Drinking, Central Shenandoah and Lord Fairfax Health Districts



Text at right: 2-sided 95% LCB | Point Estimate | 2-sided 95% UCB
Gray markers: 95% confidence interval

- Percent variance due to ACS sampling error.

- Percent variance due to the BRFSS model coefficients sampling error (this is the dominant component for most areas and outcomes).
- Percent variance due to random effect variance (unmodeled area effects)
- Percent variance due to model uncertainty.

A summary of these results is also provided in the fifth section of the outcome-specific reports.