# Appendix A. Notation

* Census tracts are denoted by
* Observations are denoted by for the
* - outcome ( smoking) for observation
* - is the census tract for observation
* - is the raked sampling weight for observation
* - zip code for observation
* -True census tract level population prevalence

# Appendix B. Imputation of Census Tract

The Department of Housing and Urban Development (HUD) provides the estimated proportion of a zip code's total residential addresses, which fall within each census tract (CT). For example, if zip code overlaps with 5 census tracts, we would have and where

and . We assume that the probability of being from CT given you are in zip code is .

For observations with missing census tracts, we impute a census tract based on a multinomial distribution with the HUD probabilities (). For a single imputation each observation is assigned to census creating a set of indices . The imputation is repeated times, resulting in complete observation data sets (no missing census tracts).

# Appendix C. Hierarchical Bayesian Model

For each complete data set our approach is to summarize the data in census tract via the asymptotic distribution of the estimator of , which we denote , the Hajek estimator (1) of , with corresponding variance estimator var. In this way the design is acknowledged in both the estimator and the variance. We define the area-level data summary as as the empirical logistic transform of . This approach constrains the probability to lie in (0,1). The likelihood is then taken as the asymptotic distribution

We employ three-stage models with the first stage given by above, which was shown to perform well in a small area estimation context and has been applied to annual zipcode-level BRFSS data (2). When the we have which is problematic in the first stage of our model. In these cases we use the method of moments based on a beta-binomial model described in the supplementary materials of (3) to provide and that are adjusted to be non-zero.

At the second stage of the model we introduce the spatial random effects terms, corresponding to the convolution model of (4), and denote the area-specific parameters as

where is the overall risk level, is an independent census tract random effect, and following an intrinsic conditional autoregressive (ICAR) model (5). The ICAR model is a non-parametric, stochastic smoothing model with

where indexes the set of neighbors of area , is the number of such neighbors and is the mean of the neighbors.

For the third stage we use assign Gamma priors on the spatial conditional precision and the precision parameter . The rate and shape parameters, 0.5 and 0.008, respectively, were selected such that the 95% range is on the interpretable and scale of (0.056,4.04). We use an improper flat prior on .

# Appendix D. Combining Estimates

Our goal is to describe the posterior distribution of given the observed data. We assume are the set of observed census tracts and are the missing census tracts

where . In our current implementation is the multinomial distribution based on HUD data, but in principle could involve other covariates.

Given our set of smoothed estimates of each we find the mean posterior estimate

where is the estimated from the th complete data set.

Similarly we find a variance

where (the posterior variance of based on the th complete dataset). This variance estimate has contributions from within and between the sets of estimates (1).

Finally, estimates for census tract are derived from expit() and 95% credible intervals are generated using

# Appendix E. Scatter Plots Comparing Direct Estimates with Smoothed Estimates By Sample Size

| Point Estimates | CI Width |
| --- | --- |
| Direct to Complete Case Smoothed  C:\Users\songl\Dropbox\PHSKC_SAE\Results\CompareHT_Smooth.jpeg | Direct to Complete Case Smoothed  C:\Users\songl\Dropbox\PHSKC_SAE\Results\IntervalWidth.jpeg |
| Direct to MI Smoothed  C:\Users\songl\Dropbox\PHSKC_SAE\Results\CompareHT_MI_Smooth.jpeg | Direct to MI Smoothed  C:\Users\songl\Dropbox\PHSKC_SAE\Results\IntervalWidth_Direct_MI_Smooth.jpeg |

**REFERENCES**

1. Hajek, J. (1971). Comments on a paper by D. Basu. In: *Foundations of Statistical*
2. *Inference* (eds: V.P. Godambe and D.A. Sprott). Toronto: Holt, Rinehart and Winston.
3. Mercer LD, Wakefield J, Pantazis A, Lutambi AA, Masanja H, and Clark SJ. “Space-Time Smoothing of Complex Survey Data: Small Area Estimation for Child Mortality”. *Annals of Applied Statistics.* 9(4) *(2015*): 1889-1905.
4. Chen, Cici, Jon Wakefield, and Thomas Lumely. "The use of sampling weights in Bayesian hierarchical models for small area estimation." *Spatial and spatio-temporal epidemiology* 11 (2014): 33-43.
5. Besag, Julian, Jeremy York, and Annie Mollié. "Bayesian image restoration, with two applications in spatial statistics." *Annals of the institute of statistical mathematics* 43.1 (1991): 1-20.
6. Rue, Havard, and Leonhard Held. *Gaussian Markov random fields: theory and applications*. CRC Press, 2005.