RESEARCH REPORT SERIES
(Statistics #2011-01)

A Random Effect Approach to Protection Against Model Error in Logistic Models of Census Coverage

Donald J. Malec
Julie H. Tsay
Elizabeth T. Huang

Center for Statistical Research & Methodology
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: March 2, 2011

A Random Effect Approach to Protection against Model Error in Logistic Models of Census Coverage

Donald J. Malec, Julie H. Tsay and Elizabeth T. Huang[1]

1 Introduction.

The application of logistic models and variable selection procedures to modeling Census coverage error has enabled the use of an expanded list of independent variables and has provided a way to eliminate higher order interactions in models; interactions in models which were implicitly included in the post-stratum used in the 2000 Census coverage (Mule et al., 2008). In the past, predictive categories of coverage such as a person's Race/ethnicity, age, sex and whether they lived in a rented home needed to be included in a model with all of their resulting cross classifications. In the logistic model, only the level of interactions that are important need to be included in the model. In addition, continuous variables, such as single year age splines, Census mail return rates and other local rates can be included. Although the logistic modeling approach has the potential to improve dual-system estimates of coverage, there is still the potential for bias due to erroneously leaving out independent variables that are tested as non-significant (type II error) such as a subset of the aforementioned higher-order interactions. It may be difficult to add specific interactions terms into the 2010 model after viewing the 2010 data due to the need to guard against perceptions of manipulating the model.

Griffin and Olson (2010) evaluated estimates of the 2000 Census correct enumeration rates using a variety of link functions, including the logistic. In that comparison, a pattern was observed in which the low predicted rates were consistently below their corresponding design-based estimates. As suggested in that report, "...control at the low end, where there is often not enough data to perform statistical testing...demonstrates a problem." In other words, appropriate interactions may be difficult to include because their significance tests are based on only a small sample size. The typical bias versus variance trade-off in model selection is present here. Leaving out an independent variable that should have been included (type II error) results in extra bias and putting in an independent variable whose parameter is really zero (type I error) results in extra variance.

The following presents an alternative to using only significance testing to select variables into a model. Variables that may be important predictors but fail a significance test due to inadequate sample size may still be included in the model as random effects. As will be seen, these terms will not have the full effect as if they were entered in the model as fixed effects but can still provide predictive power. Unlike the variable selection approach where an interaction is either zero or in the model as a fixed effect, random effect interactions are always in the model but are weighted towards zero when their precision is low, exhibiting the "shrinkage effect" (e.g. Carter and Rolph, 1974). Adding random effects could be a way to include additional terms in the model without over-fitting the model with a large number of fixed effects.

Section 2 provides a simple illustration of the mean squared error reducing properties of using a random effect model in place of variable selection. Section 3 provides an illustration of the potential of random effects inclusion in a small model for match probability using the 2000 Accuracy and Coverage Evaluation Survey (ACE) revision data. Section 4 uses a model which is similar in character to the type of "production models" that may be used for 2010 coverage estimates. There, using a model of match rates, estimation

[1]This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The authors would like to thank Martin Klein and Douglas Olson for their review and many helpful comments. The authors would also like to thank members of the Census Coverage Measurement (CCM) group for their useful comments during a presentation of these results and Lynn Weidman for useful discussions at the beginning of this project.

issues are detailed. Section 5 provides some discussion of issues that would need to be addressed before this technique could be implemented. An important part of this project included determining software in which the implementation of the random effect model approach to estimation is as straightforward as possible. This section includes a brief summary of those issues. Lastly, section 6 provides a final summary.

2 Background and Illustration.

Mosteller (1948) showed that the use of a preliminary test of significance to build a model will result in a biased estimate. The bias is a function of sample size and the actual effect size. Using a model with normally distributed errors, Mosteller provided an example that illustrated the relative gains and losses in MSE when using a preliminary test to fit a model. He also introduced the use of substituting a random effect into the model instead of including the effect in the model testing with a significance test.

The following illustrates this idea of using random effects in a very simple binary outcome model (since Census coverage data is binary). For illustration, we will use the absolutely simplest binary outcome model, that of estimating the probability that a coin comes up heads. In terms of a logistic model, one can specify the probability of a heads as:

$$p = \frac{e^{\alpha}}{1 + e^{\alpha}}$$

2.1 "Significance testing" approach to estimating p

Typically, p, and hence, α, are estimated by flipping the coin n times, counting the number of heads, m, and estimating p by m/n (or estimating α by ln(m/(n-m))).

To illustrate the variable selection approach, a significance test that $\alpha = 0$ is performed. If significant, α is estimated by ln(m/(n-m)), if not significant, $\alpha = 0$ is used. Although this is a very nonstandard way to estimate the probability of a heads, it is an exact analogy to how interactions currently are included in the logistic models for Census coverage measurement. For this example, suppose an exact, two-sided, significance test is used at the .05 level[2].

Given that the coin will be flipped n times, figure 1 plots two different estimates of p for each possible outcome. One type of estimate (labeled "sample mean") is just the sample mean. The other type of estimate (labeled "selection") uses the sample mean if it is significantly different from $\frac{1}{2}$, otherwise the value $\frac{1}{2}$ is used as the estimate. As can be seen, as the sample size gets larger, inference is more precise and the estimator based on the significance test uses the standard estimated value of p, rather than the value of $\frac{1}{2}$, more often.

A reason why one might want to use the "significance test" estimator instead of the standard approach is provided in the mean squared error[3] (MSE) plots in figure 2. Specifically, this figure provides the mean squared error of the "significance test" estimator relative to the MSE of the "sample mean" estimator given the true probability of heads, p, is known. As can be seen when the true p is near $\frac{1}{2}$ (i.e., $\alpha = 0$), the MSE of the significance test estimator is smaller than the MSE of the standard estimator. However, the significance test estimator can perform poorly relatively nearby, especially for small sample sizes.

2.2 Random effect approach to estimation

The random effect approach is to specify that the effect, α, is random from a distribution with a mean equal to the mean specified by the null hypothesis and with a variance to be estimated or specified. To facilitate a more transparent understanding of the procedure, the random effect for α in this example will be specified

[2]The cutoff is determined as $max_c Pr(m \leq c|p = .5) + Pr(m \geq n - c|p = .5) \leq .05$.

[3]For any estimator, $\tilde{p}(m, n)$, MSE[$\tilde{p}(m, n)$]= $\sum_{m=0}^{n}(\tilde{p}(m, n) - p)^2 pr(m|n, p)$, where $pr(m|n, p)$ is the binomial probability

Figure 1: Values of sample mean and “selection” estimators for sample sizes, n=10, 20, 50, 100, 1000

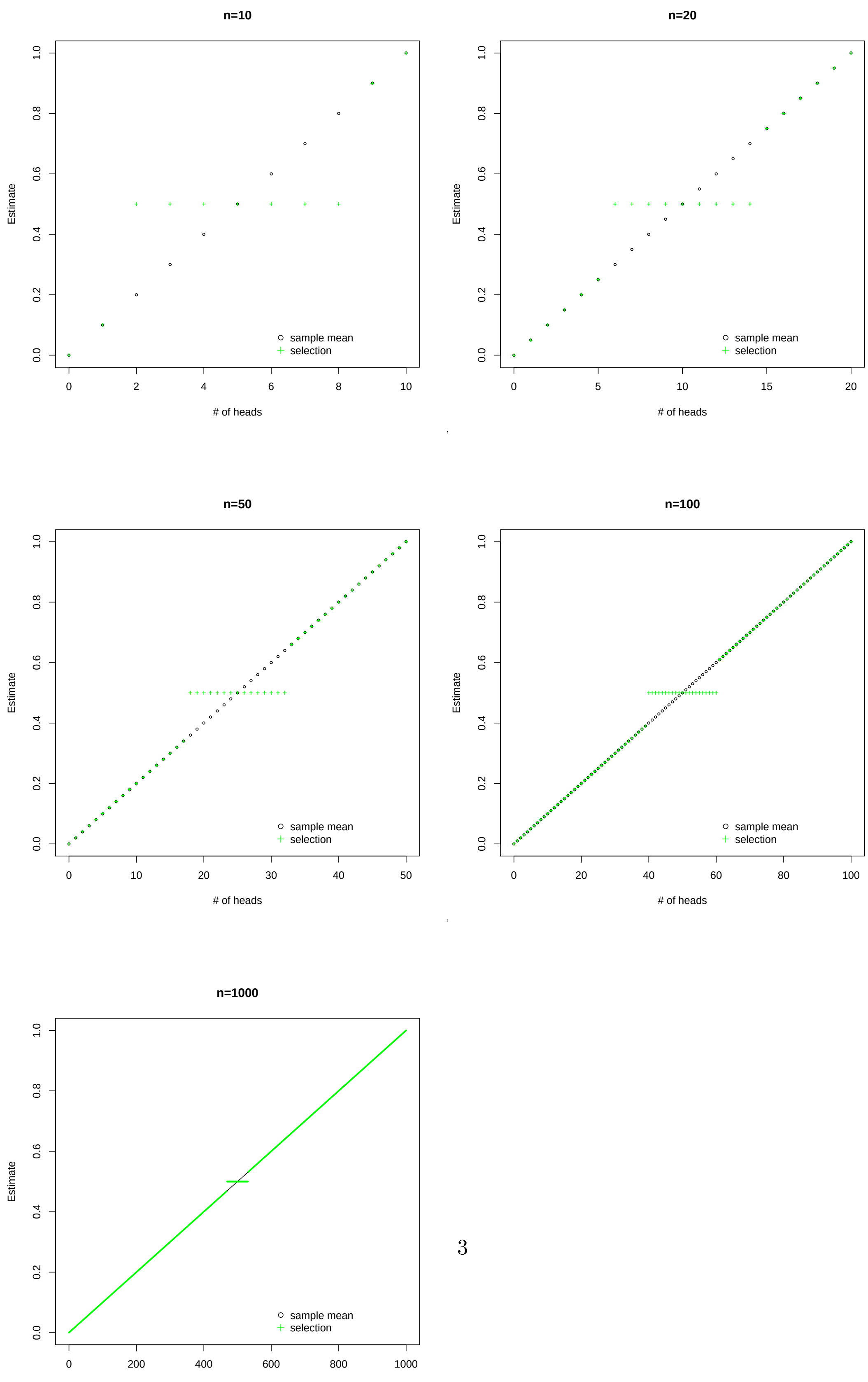
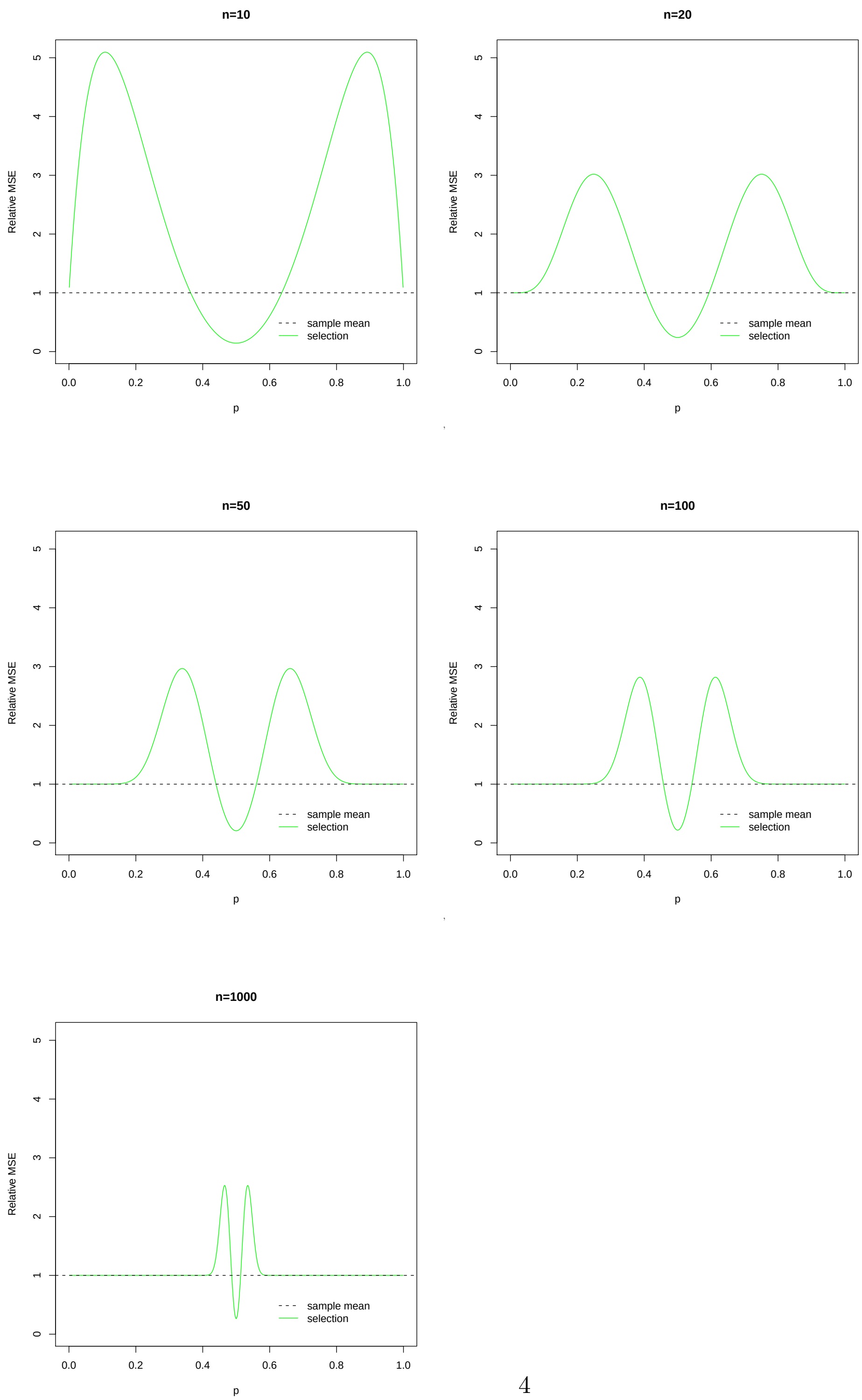


Figure 2: Relative Mean Squared Error for the “selection” estimator for known p for sample sizes, n=10,20,50,100,1000



through a distribution of $p = e^{\alpha}/(1+e^{\alpha})$. A tractable random effect model to work with, in this case, is to specify that p has a beta distribution with mean equal to the null hypothesis: $\frac{1}{2}$ (i.e. $p \sim beta(\tau\frac{1}{2}, \tau\frac{1}{2})$ and variance estimated by $(\frac{1}{2}\text{-}\hat{p})^2$, where $\hat{p}$ is the sample mean.

The best predictor of p under squared error loss is the conditional expectation, $E(p|\hat{p},\tau)$, Rao (2003). In this case, $E(p|\hat{p},\tau) = \hat{p}\times(n/(n+\tau)) + \frac{1}{2}\times(\tau/(n+\tau))$. Choosing an empirical best predictor as an estimator by plugging in an estimate of τ obtained by equating $(\frac{1}{2}\text{-}\hat{p})^2$ to $Var(p)$, the random effect estimator takes the form: $(1-w)\hat{p} + w\frac{1}{2}$, where w=$\hat{\tau}/(n+\hat{\tau})$ and $\hat{\tau}$=$\frac{(\frac{1}{2})^2}{(\frac{1}{2}-\hat{p})^2} - 1$.

Figure 3 compares the random effect estimator with the significance test estimator. As can be seen, the random effect estimator does not abruptly jump to the hypothesized value but instead is a weighted average.

In figure 4, the MSE of the random effect eliminates some of the inefficiencies in the selection model estimator while still improving on the sample mean. However, each of the three estimators, the simple mean, the "significance test" estimator, and the random effect estimator, is optimal for selected values of p.

3 A Simple Illustration Using 2000 ACE

As stated by U.S. Census Bureau (2004), "The U.S. Census Bureau conducted the Accuracy and Coverage Evaluation (A.C.E.) survey to measure the coverage of the population in Census 2000. The A.C.E. was designed to serve two purposes: (1) to measure the net coverage of the population, both in total and for major subgroups, and (2) to provide data that could serve as the basis for correcting the census counts for such uses as Congressional redistricting, state and local redistricting, funds allocation, and governmental program administration." In order to measure net coverage, statistical models are developed to estimate both the census erroneous enumeration rate and the rate in which people are captured in the Census.

This section and the next look closer at the modeling of the Census capture rate (often referred to as the match rate, since it is obtained as the proportion in the coverage survey that match to the correct Census enumerations) in order to demonstrate the potential use of including random effects as a substitute for the fixed effect interaction terms left out of the model. This section looks at a simple fixed effect model using only main effects while the next section illustrates the method using a more realistic fixed effects model. Here, a logistic model for census capture is based on only a main effect model with the following categories:

ROAST[4] main effects used in fixed effects model:

- Indicators of Race/Ethnicity Domains
  - American Indian on Reservation
  - American Indian off Reservation
  - Hispanic
  - Black
  - Native Hawaiian or Pacific Islander
  - Asian
  - White and other race
- Indicators of Age/sex
  - ages 0-9
  - ages 10-17

[4] "ROAST" is an acronym for Race/Origin Age Sex and Tenure

Figure 3: Values of sample mean, "selection" and random effects estimator for sample sizes, n=10, 20, 50, 100, 1000

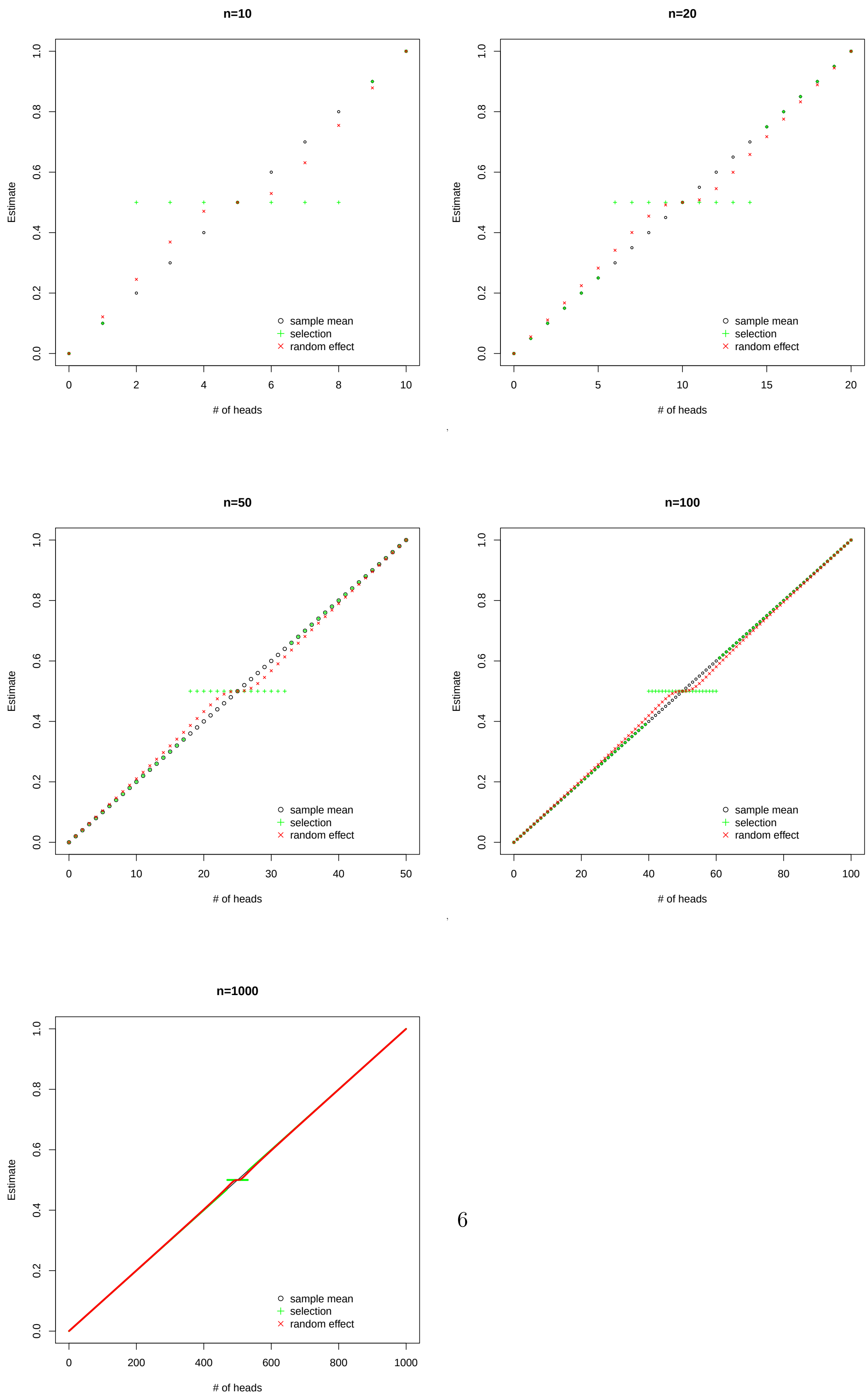
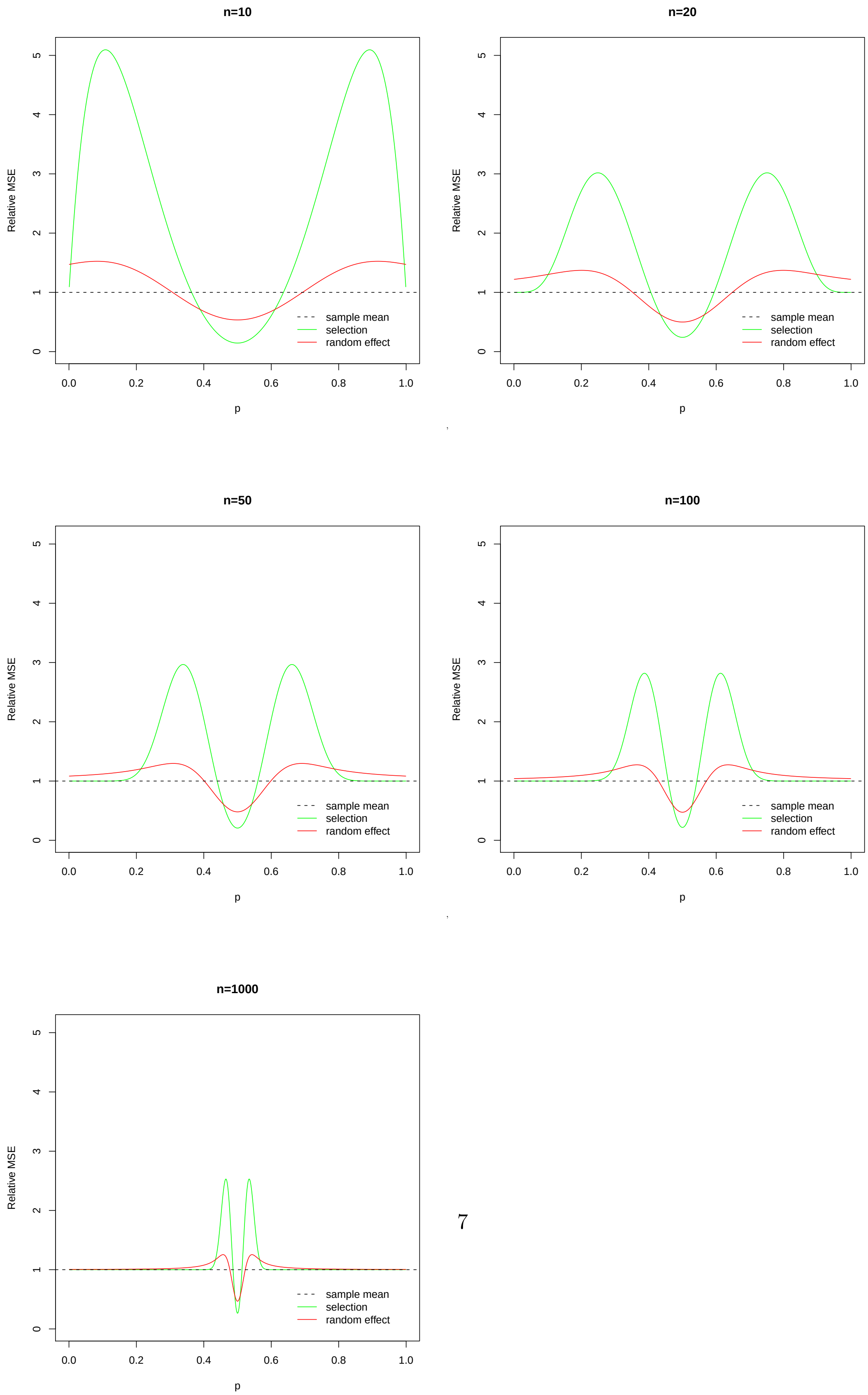


Figure 4: Relative MSE for “selection” and random effects estimator for sample sizes, n=10, 20, 50, 100, 1000



- ages 18-29, male
- ages 18-29, female
- ages 30-49, male
- ages 30-49, female
- ages 50+, male
- ages 50+, female

- Indicator that housing unit is owned by a household member

Suppose that the 112 ROAST domains, defined as the complete cross-classification of the ROAST main effects, are indentified as being of possible interest.

3.1 Models and Estimation

All models start with an indicator: $\delta_i = 0/1$ when a person in the follow-up sample, i, was not/was in the census.

Estimates are made by determining the parameters that maximize a pseudo likelihood - assuming independence between observations with each observation weighed by its adjusted weight. Estimates based on two models of match probability are made. The two models are below.

Fixed, main effect model:

$$p_i = prob(\delta_i = 1|\underline{b}) = \frac{e^{\underline{x}_i\underline{b}}}{1+e^{\underline{x}_i\underline{b}}} \tag{1}$$

Added random effect model:

$$p_i = prob(\delta_i = 1|\underline{b},\underline{\alpha}) = \frac{e^{\underline{x}_i\underline{b}+\underline{u}_i\underline{\alpha}}}{1+e^{\underline{x}_i\underline{b}+\underline{u}_i\underline{\alpha}}}, \qquad \alpha_d \sim N(0,\sigma^2), independent \tag{2}$$

The vector, $\underline{x}_i$, holds the fixed effect terms pertaining to observation i. The vector, $\underline{u}_i$, contains all zeros except for a single value of indicating the ROAST interaction cell that observation i is in.

For the fixed effect model, the parameter estimates are obtained by maximizing the pseudo log-likelihood:

$$\sum_i w_i^{'}(\delta_i ln(p_i) + (1-\delta_i)ln(1-p_i))$$

where i is summed over all sampled persons and $w_i^{'}$ is obtained be adjusting the original sample weights so that $\sum_i w_i^{'}$= the sample size.

For the random effects model, an integrated pseudo log-likelihood is maximized:

$$\int\cdots\int\prod_i p_i^{\delta_i w_i^{'}}(1-p_i)^{(1-\delta_i)w_i^{'}}d\alpha_1\cdots d\alpha_D$$

3.2 Comparison of Estimates

Figure 5 plots the observed domain means (i.e., the sample-weighted domain means) for each of the 112 ROAST domains against its corresponding predicted rate from the main effect only, fixed effect model. If the main effect only, fixed-effects model had been a perfect fit, the domain means (the dots) would have fallen on the 45° line. The four points with the largest distance from their predicted value will be examined in more detail. These are the two left-most points, the left-most being American Indians on reservation, male, age 18-29, renters, and the next point representing Native Hawaiians, male, age 18-29, renters. The other two points, closest to the bottom of the plot, represent Native Hawaiians, male, age 18-29, owners (the bottom-most) and Native Hawaiians, female, age 50+, renters (the next bottom-most).

There may be questions about the responses of these points and, hence, the authenticity of the estimates. However, if the responses have passed all edits, they will be regarded as legitimate here and their effect taken into account.

Next, in addition to estimation from the main effects-only, fixed effect logistic model, estimates from the random effect model specified in equation (2) are also made. The random effect model includes the same independent variable as the main effects-only, fixed effect logistic model with the addition of one random effect for each of the 112 ROAST domains. Figure 6 includes the resulting predicted rates from this model along with the previous fixed effect model results. The characteristic "compromise" of the random effects, being somewhere between the fixed-effect estimates and the observed values, is apparent. Of interest is the amount of smoothing the random effect model provides for the aforementioned four outliers. All show characteristic smoothing. Since this is a logistic model, an explicit formula which shows how the estimator is smoothed is not available. However, the sample size of the domain and the estimated rate both play a role in the sampling variance and, hence, in the amount of smoothing. The four points in order of description have an unweighted sample size of 43, 35, 32 and 38, respectively. For reference, the domain with the largest weight-adjusted sample size, 372, is white and other race, female aged 50+ owner.

4 A Realistic Example of the Benefits of Including Random effects in a Fixed Effect Model of Correct Enumeration

As a more realistic illustration, one of the early models developed by the Census Coverage Measurement (CCM) Estimation Modeling sub-team (see, Griffin and Olson 2010 for details) is used to evaluate the addition of random effects. Using ACE Revision II data, the match rate was modeled using the following types of variables:

- Renter
- Sex
- Three categories of enumeration area (TEA): 1=mailout/mailback; 2=update/leave; other
- Age Splines: Quadratic 0-17; Linear 17-20; Quadratic 20-50; Linear 50-80
- Age Offsets: Indicators for age=0, age= 80, and for age in (20,60) divisible by 5
- Larger Domains: Black, Hispanic and Asian can be interacted with other characteristics (with the smaller domains collapsing with White)

These variables were combined to yield a final fixed effect model with 87 terms.

The same 112 ROAST domains, defined in section 3, will be evaluated here.

Figure 5: ROAST estimates from fixed effect model using only ROAST main effects

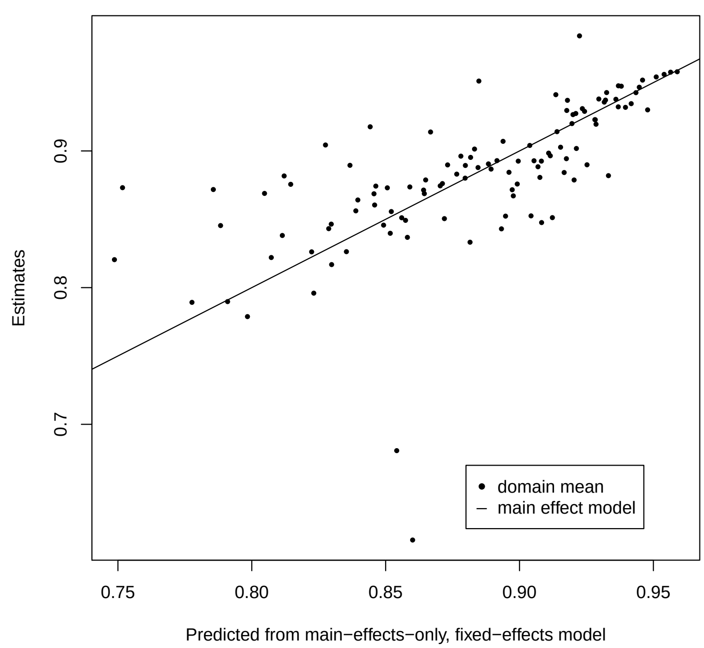
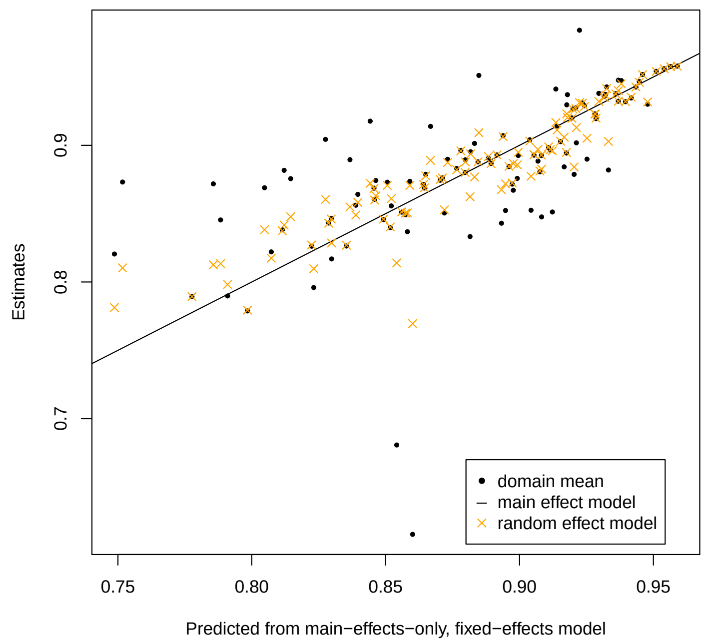


Figure 6: ROAST estimates from fixed effect model using only ROAST main effects and Random effects Model



The two types of models outlined in section 3.1 will be evaluated using the larger set of covariates. As part of the evaluation, one other model, a fixed effect model that includes both the original 87 parameter fixed effects plus the 112 domains as fixed effects will be included. For completeness, all three models to be evaluated are listed below.

1. fixed effects model: $prob(\delta_i = 1|\underline{b}) = \frac{e^{\underline{x}_i \underline{b}}}{1+e^{\underline{x}_i \underline{b}}}$
2. random effects model: $prob(\delta_i = 1|\underline{b}, \underline{\alpha}) = \frac{e^{\underline{x}_i \underline{b} + \underline{u}_i \underline{\alpha}}}{1+e^{\underline{x}_i \underline{b} + \underline{u}_i \underline{\alpha}}}$
   $\alpha_d \sim N(0, \sigma^2), independent$
3. fixed effects model with additional domain fixed effects (if estimable): $prob(\delta_i = 1|\underline{b}, \underline{a}) = \frac{e^{\underline{x}_i \underline{b} + \underline{u}_i \underline{a}}}{1+e^{\underline{x}_i \underline{b} + \underline{u}_i \underline{a}}}$

In all three models, $\underline{x}_i$ represents the independent variables in the 87-characteristic model while $\underline{u}_i$ consists of 0/1 variables indicating which one of the 112 domains the observation is in.

4.1 Comparison of predicted domain rates

Figures 7 through 9 display estimates of the 112 domains from the three models described in section 4, along with the sample mean of each ROAST interaction. For each model, the estimate for domain, k, is constructed by taking the model-based predicted probability of correct enumeration of each Census data-defined enumeration that falls in domain, k, and determining the weighted average of these cases. For each model, these predicted domains are plotted against the estimates from model 1, i.e., the original 87-term fixed effects model.

Figure 7 compares the sample estimates against the fixed effect model. Note that when the 112 ROAST domains are added to the fixed effects model, the predictions (using the estimable parameters to avoid multi-collinearity) are identical to the sample means. Figure 8 labels some of the largest deviations in figure 7 (label: race/ethnicity, age group, sex then tenure status). The extreme differences are all in either the "Native Hawaiian and Pacific Islander" group ("NH") or the "American Indian or Alaskan Native on Reservation" group ("AIO"); all groups with a relatively small sample size.

Figure 9 includes the estimates from the random effects model obtained by plugging the MLE for $\underline{b}$ and the best predictors for $\underline{\alpha}$ (using the Restricted MLE of the variance component to determine the variance component) into the logistic function of model 2. Here, the estimates based on the random effects model are not identical to the fixed effect model, but they are very close, indicating that the variance component of the random effect has been estimated to be near zero.

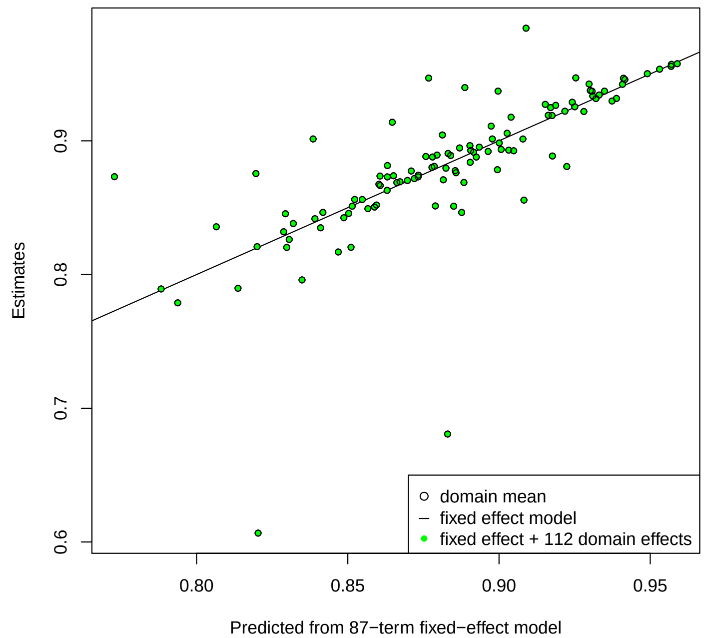
4.2 An Alternative Random Effect Model

One explanation why the estimates from the random effect model and the fixed effect model are similar may be due to small sample size, resulting in large errors. Another explanation could be that all 112 random effects do not come from the same distribution and do not all share a common variance component; most could have a very small variance component while only a few have a large variance component. To assess this second possibility, the 112 effects are formed into two groups. The group that has a relatively small p-value (when being tested as a fixed effect using the basic chi-squared test in R) is modeled as having random effects from the same distribution. The remaining effects are simply set equal to zero.

Table 1 lists the corresponding domains whose fixed effect estimate had a p-value $< .1$. These 16 domains were modeled as one random effect group.

The resulting domain estimates from this model (labeled the "pre-test" model), compared to the previous models, are listed in figure 10. Several of the new estimates have shifted noticeably towards their sample

Figure 7: ROAST estimates from fixed effect models with and without ROAST effects added



Race/Ethnicity	age group	sex	tenure	p-value
American Indian or Alaskan Native (off reservation)	30-49	female	owner	0.0015
Native Hawaiian or Pacific Islander	50+	female	renter	0.0021
American Indian or Alaskan Native (off reservation)	18-29	female	renter	0.0083
White or other race	50+	female	owner	0.0108
Hispanic	30-49	male	owner	0.0114
White or other race	10-17		renter	0.0215
Native Hawaiian or Pacific Islander	18-29	male	owner	0.0235
Hispanic	50+	male	owner	0.0262
American Indian or Alaskan Native (off reservation)	18-29	female	owner	0.0383
White or other race	30-49	female	renter	0.0402
White or other race	<10		renter	0.0422
American Indian or Alaskan Native (off reservation)	30-49	female	renter	0.0525
American Indian or Alaskan Native (off reservation)	18-29	male	owner	0.0586
Black	18-29	male	owner	0.0757
American Indian or Alaskan Native (off reservation)	50+	female	renter	0.0776
White or other race	30-49	female	owner	0.0818

Table 1: Domains with p-values < .1

Figure 8: Fixed effect estimates with identifying large differences

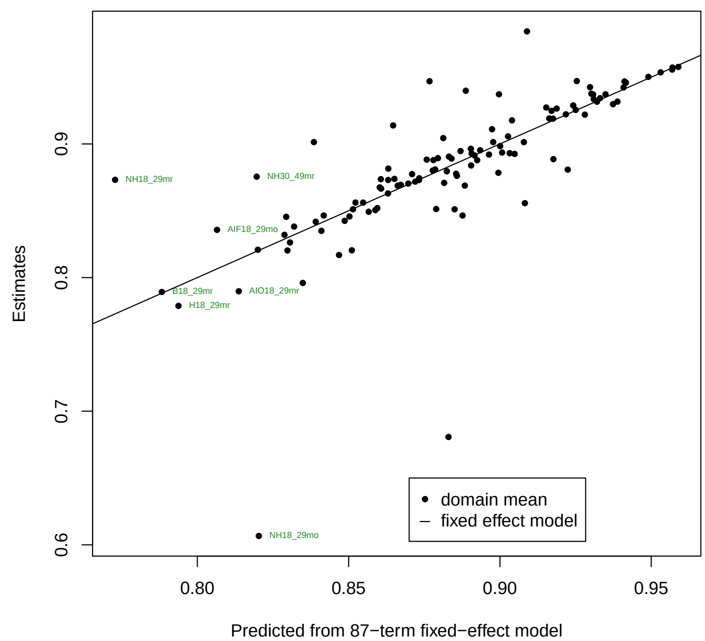


Figure 9: Estimates of 112 domains including the random effects model

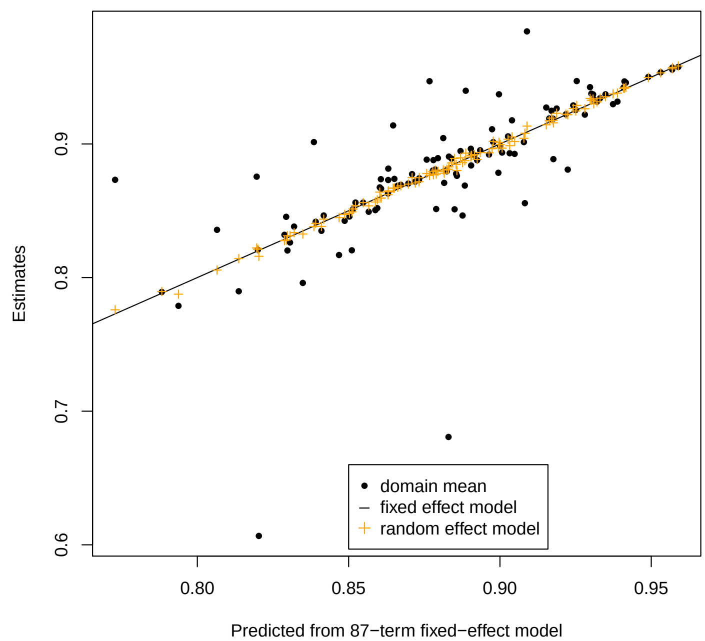
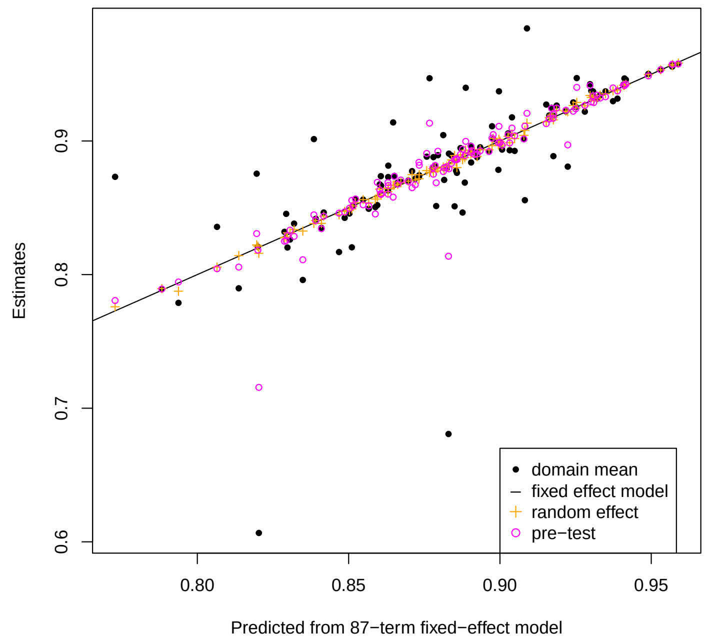


Figure 10: Estimates of 112 domains including the "pre-test" model



estimates. This indicates that specification of the variance component model can have a noticeable effect on the resulting estimates. Figure 11 pinpoints the domains that were in the random effects models and those that were not. As can be seen from this figure, many of the sample mean estimates with large deviations were not part of the random effects model.

5 Discussion

Although random effects are used, here, as a way to guard against type II error, the same type of model can be motivated as a model for small domains. A similar type of model applied to small areas (local Census offices) was used by Malec and Maples (2008) and given a full Bayes treatment.

This method is even closer to the Census coverage smoothing work reported in Isaki et al. (1991,2000) because in both cases, small domains are smoothed using a random effects model. Isaki et al. model the direct estimates of Census coverage factors while this work models the components that make up the Census coverage factor [5]. Also, while the Isaki et al. work is framed as a smoothing of adjustment factor and the work here proposes a way to "roughen" the estimates from a fixed effects model, the intent is the same; to make coverage estimates that include more interaction effects combined with a fixed effect model to reduce the overall mean squared error. Table 2 summarizes some of the similarities and differences between the two approaches.

[5]Only the match-rate component is modeled here, for illustration. The other two components: the data-defined rate and the correct enumeration rate were not modeled.

Figure 11: Identification of random effects group in "pre-test" model

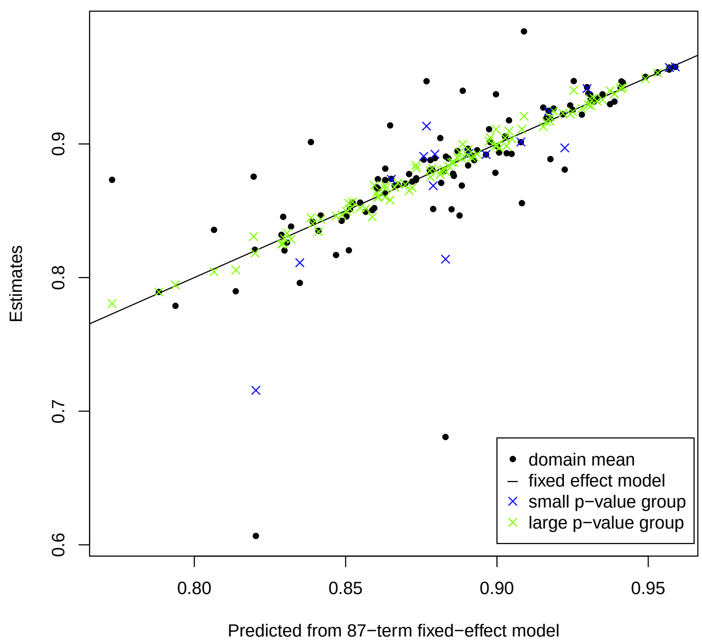


Table 2: Summary of features between post-stratum smoothing of Isaki et al. and the logistic random effect model

- **assumed sampling distribution**
  - Isaki et al.: assumes direct estimates of Census coverage factors are normally distributed
  - logistic: the three parts of the coverage factors are based on Bernoulli random variables
- **model of domain variation**
  - Isaki et al.: linear mixed model
  - logistic: logistic mixed model
- **Hierarchical Random Effects?**
  - Isaki et al.: yes
  - logistic: no
- **interactions of interest**
  - Isaki et al.: all post-strata
  - logistic: 112 ROAST interactions
- **types of covariates allowed in model**
  - Isaki et al.: discrete variables
  - logistic: continuous and discrete variables
- **assumptions on design effect**
  - Isaki et al.: area-level model, smoothed fixed-effect design covariance model
  - logistic: unit level model, weighted pseudo-likelihood with independent components
    (Note: Used for illustration, further design adjustment may be needed)
- **assumptions about model based variance**
  - Isaki et al.: normally distributed variation, separate variance for 4 regions and, also, by demographic groups
    (note: see 2000 paper)
  - logistic: normally distributed variation, one variance component,
    recommend using more that one variance component
- **effect on estimates**
  - Isaki et al.: published results shows smoothing and reduction in estimated variance
    (no comparison available to estimates from fixed effects model only)
  - logistic: Shows extreme smoothing to fixed effect model,
    smoothing to a lesser extend using more than one variance component

Both Isaki et al. (2000) and this work recognize the possible benefits of using multiple variance components. Isaki et al. let variance components differ by both region demographic groups. The work here attempts to cluster random effects into different groups using p-values. Isaki et al. (2000) address problems in specifying the variances associated with the estimates of the interaction terms. Sensitivity to the specification of the pseudo-likelihood was not considered to be in-scope of the initial work, here, but would certainly be needed for further development.

A parallel purpose of this project was to determine what computing software could be used make empirical Bayes estimates in a production setting. Empirical Bayes estimates were selected for this project in order to avoid the specialized software and computer intensive methods of full Bayes methods. In short, the pseudo-likelihood was treated as if it was an actual likelihood. MLEs were determined for all fixed effect parameters and the variance components. An estimate of the best predicted value of a random effect was substituted into the logistic model and the resulting probability was used to make an estimate of the corresponding rate. Originally, the possibility of using replication methods to estimate the design-based variance of these estimates was planned but was never begun.

The evaluation and development of software to make the empirical Bayes estimates was a major part of this project. SAS has two relevant procedures: PROC GLIMMIX and PROC NLMIXED. Both procedures seemed to work well for small models but had convergence issues for large models. This is not a criticism of the numerical aspects of these procedures. However, it is a criticism that they could not be implemented quickly for this project while other methods could. Each of the SAS procedures have a wide variety of user-defined settings and the procedures appear to developed for a very wide class of mixed models. Given that PROC GLIMMIX had not been fully implemented and the initial problems with convergence using PROC NLMIXED, a more transparent approach using PROC IML was developed. Using PROC IML, variance component estimation using the fourth-order approximation detailed in Breslow and Lin (1995) was used. A linear approximation conditional on the variance component was used to make estimates of the parameters of the fixed effect linear components of the model. These results were compared to corresponding results obtained from the “glmer” function in the lme4 library of the R-project. This library is also not fully implemented but produced comparable answers to the large model-results from PROC IML.

6 Summary

An approach that adds coefficients as random effects into a model, as opposed to testing whether they should be included/excluded as fixed effects via a significance test, was described and evaluated. The benefit of adding random effects to a clearly under-parameterized model was clear; however, little benefit was realized when a larger, more realistic, fixed effect model was used in conjunction with a single random effect distribution. Grouping the random effects into more then one distribution was more successful.

Software for these types of models (GLMM) is available but not mature. SAS (Proc NLMIXED and GLIMMIX) were difficult to use and settings to achieve convergence uncertain. Implementation of Breslow and Lin’s approximation in SAS IML appeared successful, however. The “glmer” function in the lme4 library of R worked well, however, all the features such as the use of initial values, the maximum number of iterations, etc. did not appear to be operational, yet.

More work is needed before a production version of this procedure could be achieved. Of importance is a more comprehensive incorporation of the sample design into the logistic likelihood, the determination of how many different random effect distributions are needed to represent the model error and, aslo, specification for variance estimation. The procedure, as outlined above could still be used, as is, as a “sensitivity analysis” to assess the effects of missing terms in the model. In addition, a quick screening of interactions that are not in the model, but of possible subject matter importance, for significance or marginal insignificance may also useful.

7 References

Breslow, Norman E. and Xihong Lin. (1995). "Bias Correction in Generalized Linear Mixed Models with a Single Component of Dispersion," *Biometrika*, 82, 81-91.

Carter, Grace M. and Rolph, John E. (1974). "Empirical Bayes methods applied to estimating fire alarm probabilities," *Journal of the American Statistical Association*, 69, 880-885.

Griffin, Richard A. and Douglas Olson (2010). "Research of Alternative Linking Functions," DSSD 2010 Census Coverage Measurement Memorandum Series: 2010-E-23.

Isaki, Cary T., Huang, Elizabeth T. and Tsay, Julie H. (1991). "Smoothing adjustment factors from the 1990 Post Enumeration Survey," ASA Proceedings of the Social Statistics Section, 338-343, American Statistical Association (Alexandria, VA).

Isaki, C.T., Tsay, Julie H. and Fuller, Wayne A. (2000). "Estimation of Census adjustment factors," *Survey Methodology*, 26, 31-42.

Malec, Donald and Maples, Jerry. (2008). "Small area random effects models for capture/recapture methods with applications to estimating coverage error in the U.S. Decennial Census," *Statistics in Medicine*, 27, 4038-4056.

Mosteller, Frederick. (1948). "On Pooling Data," *Journal of the American Statistical Association*, 43, 231-242.

Mule, Vincent T., Malec, Donald, Maples, Jerry and Schellhamer, Teresa. (2008). "Using continuous variables as modeling covariates for net coverage estimation," ASA Proceedings of the Joint Statistical Meetings, 1941-1948 American Statistical Association (Alexandria, VA).

Rao, J.N.K. (2003). Small Area Estimation, New York: Wiley.

U.S. Census Bureau, (2004) "Accuracy and Coverage Evaluation of Census 2000: Design and Methodology". In www.census.gov/prod/2004pubs/dssd03-dm.pdf, November 19, 2010.