2015

A Bayesian Approach to Small Area Estimation of Health Insurance Coverage

Zhengjia Sun
Louisiana State University and Agricultural and Mechanical College, zsun6@lsu.edu

A BAYESIAN APPROACH TO SMALL AREA ESTIMATION OF HEALTH INSURANCE COVERAGE

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Economics

by
Zhengjia Sun
B.S., Dalian University of Technology, 2005
M.S., University of South Carolina, 2008
M.S., Louisiana State University, 2010
May 2015

Acknowledgments

First of all, I would like to express the deepest appreciation to my committee chair, Professor Milton Dek Terrell. Thank you for being a great advisor, devoting so much time and energy into encouraging me as an econometrician. Without his continuous support and advice, this dissertation would not have been possible.

I would express my appreciation to my committee members: Professor Carter R. Hill and Professor Robert J. Newman for their valuable suggestions and help in my academic research. I am very thankful to Professor W. Douglas McMillin, who always been there giving me encouragement and valuable advice when I am in need.

I also appreciate all the other faculties and staffs in the Department of Economics for their patient guidance and helpful comments throughout my graduate study. Moreover, I have received support from my fellows, Quqiong, Xue, Xianliang, Dachao, Bela and Ting. Thank you for being there for me like my family.

Last but not least, I want to thank my husband Jintao Cui for his unconditional love and support in the past ten years of marriage. Thank my dearest daughter, Melissa Sunny Cui, for her beautiful smiles. You make my life more meaningful.

Table of Contents

Abstract

Small area estimation focuses on borrowing strength across area in order to develop a reliable estimator when the auxiliary information is available. The traditional methods for small area estimation borrow strength through linear models that provide links to related areas, which may not be appropriate for some survey data. We examine the empirical best unbiased linear prediction method and hierarchical Bayes method with the Louisiana Health Insurance Survey (LHIS), and a hierarchical Bayes method with probit model to fit the LHIS data by using the single year data in 2013. This approach results in a lower level of posterior standard deviations compared to the other two estimates. Furthermore, we also construct an informative Bayesian prior on the repeated cross-sectional data set 2003-2013, and show a continuous shift from the single year estimates to the pooled estimates. Simulation studies are given to examine the performance of various approaches.

Chapter 1. Introduction

Sample surveys are widely used in providing estimates for both the entire population of interest and for a variety of sub-populations (domains or small areas). Small areas can be defined by geographic areas such as state, county, health service area, or socio-demographic groups such as race, gender or types of industry, in which case they are referred to as domains. The purpose of small area estimation is to produce reliable estimates of characteristics of interest such as means, counts, quantiles for areas or domains for which only small samples or no samples are available. Due to the growing demand for reliable small area statistics for both public and private sectors, small area estimation is becoming important in survey sampling. However, the traditional direct survey estimates for small areas are unlikely to be accepted due to the large standard errors. This makes it necessary to "borrow strength" from related areas to get more accurate estimates for the area with relatively small sample size.

The issue of providing health insurance coverage to children and adults has long been a topic of interest to U.S. policymakers. The pattern of health insurance coverage for adults varies over years. The variations are caused due to the changing of economic environments, insurance policy or people's behavior. For instance, starting from 2001, the number of uninsured Americans increased, primarily because of a decline in employer-sponsored insurance, while the drop in employer coverage was not offset by an increase in public coverage. During the economic recession, health insurance coverage decreased significantly. More recently, on March 23, 2010, the Patient Protection and Affordable Care Act is signed into law. The Congressional Budget Office (2011) has projected that the implementation of health insurance reforms in the Affordable Care Act (ACA) will reduce the number of uninsured Americans by 33 million in 2020, from 56 to 23 million people. Beginning in 2014, most Americans were required to have health insurance coverage meeting certain minimum requirements and would be subject to financial penalties if they did not comply. Estimates of the number of uninsured persons will be a key ingredient in measuring the effectiveness of the Affordable Care Act, particularly if some states deviate from others with regard to some parts of the legislation such as the expansion of Medicaid eligibility.

In Chapter 2, we provide a literature review of the existing approaches to small area estimation. These approaches are usually assumed to be related through some type of linear model. For the linear model, the existing approaches borrow strength by using data from related areas to estimate the interested parameters. Among those linear models, some rely on the direct estimator while the corresponding estimator might become problematic since the direct estimator is either not available or not reliable. Due to the small sample size in some of the sub-populations, it is hard to find a good estimate of the precision of the indirect estimators, as well as the model based estimators. In practice, the assumption of an explicit linking model between variables may not be appropriate for some complicated situations.

In Chapter 3, we introduce a Bayesian approach to small area estimation. We illustrate how to obtain the reliable estimates by applying the hierarchial Bayesian methods, as well as the hierarchical Bayesian methods with a probit model for a binary dependent variable. Given the existing information, we need to apply the Gibbs Sampling and Markov chain Monte Carlo (MCMC) techniques to compute approximately the desired posterior expectations.

In Chapter 4, we employ the Louisiana Health Insurance Survey (LHIS) to estimate the uninsured rates for both adults and children in each of the 64 parishes in Louisiana. The Louisiana Health Insurance Survey starts from 2003, and consists of a series of surveys designed to provide the most accurate and comprehensive assessment of Louisiana's uninsured populations every two years. Compared to existing traditional direct estimates, our methods perform better in terms of posterior standard errors. Starting from a single survey year, we also apply the hierarchial Bayesian method with probit model on cross-sectional data with informative prior.

Chapter 5 contains the simulation procedures and results. To examine the efficiency of our hierarchical Bayesian method with probit model, we generate a data set with different coefficients by employing individuals' information over the past six survey years. Under this circumstance, our estimates show that the flat informative prior impacts the estimates heavily compared to the restricted informative prior. Chapter 6 completes the thesis by summarizing the main findings of my work.

Chapter 2. Small Area Estimation

2.1 History of Small Area Estimation

The history of small area statistics goes back to the eleventh century England. The use of maps to understand the prevalence of a disease for small areas has been used for a long time (Marshall, 1991). The research on small area estimation has received considerable attention in recent years due to growing demand for reliable small area statistics by various federal and local government agencies (such as U.S. Census Bureau, U.S. Bureau of Labor Statistics, Statistics Canada). Over the years, many statisticians have introduced various programs to meet this demand.

A small area usually refers to a subgroup of population from which samples are drawn. The subgroup might be a geographical region such as county or a census division, or a group obtained by cross-classification of demographic factors such as age, race or gender. The importance of reliable small area statistics cannot be over-emphasized as these are needed in regional planning and fund allocation in many federal and local government programs. For example, in both developed and developing countries, governmental policies increasingly demand income and poverty estimates for small areas. In fact, in the U.S. more than $130 billion in federal funds per year are allocated based on these estimates (Jiang and Lahiri, 2006). In addition, states utilize these small area estimates to divide federal funds and their own funds to areas within the state. These funds cover a wide range of community necessities and services including education, public health, and numerous others. Therefore, there is a growing need to improve the methods by which there estimates are made to provide an increased level of accuracy.

Small area estimation attempts to solve the problem of providing reliable estimates of one or several variables of interest in areas where the information available on those variables is not sufficient to provide a valid estimate on its own. The information is usually collected by conducting a survey in some or all areas. The survey may involve the collection of information from the areas themselves or some of the individuals living in those areas, whose data are later used to provide area-based estimates.

Most surveys provide very little information on a particular small area of interest since surveys are generally designed to produce statistics for larger populations. Thus, direct design-based estimators are unreliable since only a few observations are available from the particular small area of interest. The main idea to improve on a design-based survey estimator is to use relevant supplementary information, usually available from various administrative records, in conjunction with the sample survey data.

For the past few decades, sample surveys have taken the place of a complete census as a more cost-effective means of obtaining information on wide-raging topics of interest at frequent intervals over time. Sample survey data can be used to derive reliable estimators of totals and means for large areas. However, the usual direct survey estimators for a small area, based on data only from the sample units in the area, are likely to yield unacceptable large standard errors due to the small size of the sample in the area. Sample sizes for small areas are typically small because the overall sample size in a survey is usually determined to provide specific accuracy at a much higher level of aggregation than that of small areas. The use of survey data in developing reliable small area statistics with the census and administrative data has received more attention recently.

Due to a growing demand for reliable small area statistics from both the public and private sectors, the amount of attention being paid to small area estimation has increased significantly. For example, there may exist geographical subgroups within a given population that are far below the average in certain respects and need a definite upgrade. An identification of such regions is needed, since one would like to have statistical data at the relevant geographical levels. Small area statistics are also needed in the apportionment of government funds, and in regional and city planning. Furthermore, there are demands from the private sector since policy makers for many businesses and industries rely on local socioeconomic conditions. Therefore, the demand for small area statistics can arise from various sources.

2.2 Traditional Methods

Small area estimation methods can be divided broadly into "design-based" and "model-based" methods. The "design-based" methods often use a model for the construction of the estimators, while the bias, variance and other properties of the estimators are evaluated under the randomization distribution. The randomization distribution of an estimator is the distribution over all possible samples that could be selected from the target population of interest under the sampling design used to select the sample, with the population measurements considered as fixed parameters. The "model-based" method normally uses either the frequentist approach or the Bayesian methodology, and in some cases the combination of those two approaches, which is known as "Empirical Bayes" in the literature. Different from the "design-based" method, the "model-based" method is usually conditioned on the selected sample, and the inference is with respect to the underlying model.

2.2.1 Classical Demographic Methods

In this section, I provide a brief review of classical demographic methods for local estimation of population and other characteristics of interest in postcensal years. These methods use current data from administrative registers in conjunction with related data from the last census.

Purcell and Kish (1979) categorize the methods for local estimation of population and other characteristics of interest in postcensal years under the general heading of Symptomatic Accounting Techniques (SAT). Such techniques utilize current data from administrative registers in conjunction with related data from the latest census. The diverse registration data, such as the numbers of births and deaths, existing and new housing units and school enrollments, whose variations are strongly related to changes in population totals.

The vital rate method uses only birth and death data. In a given year t, the annual numbers of births b_t, and deaths d_t are determined for a local area. The crude birth rates r_{1t} and death rates r_{2t} for that local area are estimated by:

$$r_{1t} = r_{10}(\frac{R_{1t}}{R_{10}}), r_{2t} = r_{20}(\frac{R_{2t}}{R_{20}}),$$

where r_{10} and r_{20} denote the crude birth and death rates for the local area in the latest census year ($t = 0$) respectively. R_{1t}, R_{2t} and R_{10}, R_{20} denote the crude birth and death rates in the current and census years for a larger area, which contains the local area, respectively. The population p_t for the local area at year t is estimated by:

$$p_t = \frac{1}{2}(\frac{b_t}{r_{1t}} + \frac{d_t}{r_{2t}}).$$

However, as pointed out by Marker (1983), the success of the vital rates method depends heavily on the validity of the assumption that the ratios r_{1t}/r_{10} and r_{2t}/r_{20} for the local area are approximately equal to the corresponding rates R_{1t}/R_{10} and R_{2t}/R_{20} for the larger area.

The component method is considered as an extension of the vital rates method. The sums are computed independently, particularly by taking census values, adding births, subtracting deaths, and adding an estimate of net migration. Let b_{0t}, d_{0t} and m_{0t} denote the numbers of births, deaths and net migration in the local area during time period $[0, t]$ respectively. Net migration m_{0t} is the sum of immigration i_{0t} minus emigration e_{0t}. Hence, the current population p_t is expressed as:

$$p_t = p_0 + b_{0t} - d_{0t} + m_{0t},$$

where p_0 is the baseline census population.

The estimation methods mentioned above can be identified as special cases of multiple linear regression (Marker, 1983). Regression symptomatic procedures also use multiple linear regression for estimating local area population, utilizing symptomatic variables as independent variables in the regression equation. Two such procedures are the ratio correlation and the difference correlation methods (Rao, 2003).

2.2.2 Ratio Correlation and Difference Correlation Methods

Let 0, 1, and $t(> 1)$ denote two consecutive census years and the current year, respectively. Let p_{ia} and s_{ija} be the population size and the value of jth symptomatic variable ($j = 1, ..., p$) for the ith local area ($i = 1, ..., m$) in the year a $(= 0, 1, t)$. Let p_{ia}/P_a and s_{ija}/S_{ja} be the corresponding proportions, where $P_a = \sum_i p_{ia}$ and $S_{ja} = \sum_i s_{ija}$ are the values for the larger area.

The change in proportional values U_i of the independent variables between census years 0 and 1 for the ith area, are related to the corresponding changes in proportional values z_{ij} of the symptomatic variables for the jth symptomatic variable and the ith area, through multiple linear regression:

$$U_i = \gamma_0 + \gamma_1 z_{i1} + ... + \gamma_p z_{p1} + u_i,$$

where u_i are the random errors assumed to be uncorrelated with zero means and constant variance σ_u^2. In the ratio correlation method ratios are used to measure the changes:

$$U_i = \frac{p_{i1}/P_1}{p_{i0}/P_0}, z_{ij} = \frac{s_{ij1}/S_{j1}}{s_{ij0}/S_{j0}}.$$

The difference correlation method uses differences to measure the changes:

$$U_i = p_{i1}/P_1 - p_{i0}/P_0, z_{ij} = s_{ij1}/S_{j1} - s_{ij0}/S_{j0}.$$

2.2.3 Traditional Synthetic Estimation

Next, I provide a brief discussion of traditional synthetic estimation and related methods under the design based framework. Gonzales (1973) describes synthetic estimates as follows: "An unbiased estimate is obtained from a sample survey for a large area; when this estimate is used to derive estimates for subareas under the assumption that the small areas have the same characteristics as the large area, we identify these estimates as synthetic estimates." The synthetic estimation method is traditionally used for small area estimation, because of its simplicity, applicability to general sampling designs and potential of increased accuracy in estimation by borrowing information from similar small areas (Rao, 2003).

Suppose the population is partitioned into large domains g with reliable estimators $\hat{Y}'_{\cdot g}$ of totals $Y_{\cdot g}$ can be calculated from the survey data. The domain g is divided into several small areas i, so that $Y_{\cdot g} = \sum_i Y_{ig}$, where Y_{ig} is the total for cell (i, g). Assume that auxiliary information in the form of totals X_{ig} is also available. A synthetic estimator of small area total $Y_i = \sum_g Y_{ig}$ is given by:

$$\hat{Y}_i^S = \sum_g (X_{ig}/X_{\cdot g})\hat{Y}'_{\cdot g},$$

where $X_{\cdot g} = \sum_i X_{ig}$ (Ghangurde and Singh, 1977). The above estimator has the desirable consistency property that $\sum_i \hat{Y}_i^S$ equals the reliable direct estimator $\hat{Y}' = \sum_g \hat{Y}'_{\cdot g}$ of the population total Y.

The direct estimator $\hat{Y}'_{\cdot g}$ used in synthetic estimation is typically a ratio estimator of the form

$$\hat{Y}'_{\cdot g} = [(\sum_{l \in s_{\cdot g}} w_l y_l)/(\sum_{l \in s_{\cdot g}} w_l x_l)]X_{\cdot g} = (\hat{Y}_{\cdot g}/\hat{X}_{\cdot g})X_{\cdot g},$$

where $s_{\cdot g}$ denotes the sample in the large domain g and w_l is the sampling weight attached to the lth element. Hence, the synthetic estimator reduces to $\hat{Y}_i^S = \sum_i X_{ig}(\hat{Y}_{\cdot g}/\hat{X}_{\cdot g})$.

If $\hat{Y}'_{\cdot g}$ is approximately design unbiased, the design bias of $\hat{Y}_i^S$ is given by

$$E(\hat{Y}_i^S) - Y_i \doteq \sum_g X_{ig}(Y_{\cdot g}/X_{\cdot g} - Y_{ig}/X_{ig}),$$

which is not zero unless $Y_{ig}/X_{ig} = Y_{\cdot g}/X_{\cdot g}$ for all g. In the special case where the auxiliary information X_{ig} equals the population count N_{ig}, the latter condition is equivalent to assuming that the small area means $\bar{Y}_{ig}$ in each group g equal the overall group mean $\bar{Y}_{\cdot g}$. In fact, synthetic estimators for some small areas can be heavily biased in the design based framework.

A natural way to balance the potential bias of a synthetic estimator against the instability of a direct estimator is to take a weighted average of those two. The composite estimators of the small area totals Y_i could be written as:

$$\hat{Y}_{iC} = \phi_i \hat{Y}_{i1} + (1 - \phi_i)\hat{Y}_{i2},$$

where $\hat{Y}_{i1}$ is a direct estimator, $\hat{Y}_{i2}$ is a synthetic estimator and ϕ_i is a suitable chosen weight $(0 \le \phi_i \le 1)$.

The designed MSE of the composite estimator is given by

$$\text{MSE}_p(\hat{Y}_{iC}) = \phi_i^2 \text{MSE}_p(\hat{Y}_{i1}) + (1-\phi_i)^2 \text{MSE}_p(\hat{Y}_{i2}) + 2\phi_i(1-\phi_i)E_p(\hat{Y}_{i1} - Y_i)(\hat{Y}_{i2} - Y_i).$$

By minimizing the designed MSE with respect to ϕ_i, the optimal weight ϕ_i as follows:

$$\phi_i^* = \frac{\text{MSE}_p(\hat{Y}_{i2}) - E_p(\hat{Y}_{i1} - Y_i)(\hat{Y}_{i2} - Y_i)}{\text{MSE}_p(\hat{Y}_{i1}) + \text{MSE}_p(\hat{Y}_{i2}) - 2E_p(\hat{Y}_{i1} - Y_i)(\hat{Y}_{i2} - Y_i)}$$

$$\approx \text{MSE}_p(\hat{Y}_{i2})/[\text{MSE}_p(\hat{Y}_{i1}) + \text{MSE}_p(\hat{Y}_{i2})],$$

assuming that the covariance term $E_p(\hat{Y}_{i1} - Y_i)(\hat{Y}_{i2} - Y_i)$ is small relative to $\text{MSE}_p(\hat{Y}_{i2})$. The approximate optimal ϕ_i^* is between 0 and 1. Hence, the approximate optimal weight ϕ_i^*

depends only on the ratio of the MSEs. Such that

$$\phi_i^* = 1/(1 + F_i),$$

where $F_i = \text{MSE}_p(\hat{Y}_{i1})/\text{MSE}_p(\hat{Y}_{i2})$.

2.3 An Example in the Small Area Estimation

Consider the following example, in which there are four parishes P_1, P_2, P_3 and P_4. Without loss of generality, assuming parish P_1 has only a small sample available, when the other three parishes have large samples available. Suppose that in the sample we observe the proportion of uninsured rate for parish P_1 is 0.1, while for parishes P_2, P_3 and P_4 are 0.21, 0.22 and 0.24, respectively. Now, we focus on the estimation of the proportion of uninsured rate for parish P_1.

Due to the small sample size, although the direct estimate of uninsured for parish P_1 is an unbiased estimate, it may come with a large variance. An other option to estimate the uninsured rate for parish P_1 could be pooling the data for four parishes. The pooled estimate is the total number of uninsured persons divided by the total number of individuals, who are sampled in all four parishes. Compared with the former estimate, this estimate is much more reliable, because of the large sample size. However, the latter estimate is biased since it is based on the data from other parishes.

Therefore, it is desirable to obtain estimates which are intermediate between the direct estimate and the pooled estimate.

For each parish, the individuals are sampled independently from a distribution particular to that parish, and we also view the means of these parish distributions as coming from an overall distribution on the parish means. In detail, let p_i be the true population uninsured rate for parish i, for $i = 1, ..., N$. Let $\hat{p}_i$ be the estimate of p_i, and let n_i be the sample size for parish i. The model is defined as following:

$$\text{Given}\, p_i, \quad \hat{p}_i \overset{ind}{\sim} \text{Binomial}(p_i, n_i), \;\; i = 1, ..., N, \tag{2.1a}$$

$$\text{Given}\, \mu, \tau, \quad \text{probit}(p_i) \overset{ind}{\sim} TN(\mu, \tau^2), \;\; i = 1, ..., N, \tag{2.1b}$$

where μ and τ are unknown parameters estimated from the data, and $TN(a, b)$ denotes a normal distribution truncated to lie in the region (a, b). Once μ and τ are estimated, we could obtain the estimates of p_i, for $i = 1, ..., N$.

The estimates from the above model are called "composite estimates." In the example, the estimate for parish P_1 will be bigger than the direct estimate 0.10, since the composite estimates apply the information from other parishes, while the other parishes have the higher uninsured rates comparing to Parish 1. This effect is enhanced for the larger parishes P_2, P_3 and P_4, and it would be enhanced even further if there were more of these large parishes. The intuition behind this is that the information on parishes P_2, P_3, and P_4 gives us information on the overall distribution on parish means, which in turn gives us information on parish P_1. However, there are two aspects of weak points regarding the composite estimates, which are accuracy and obtaining confidence intervals. In terms of accuracy, the composite estimate borrows information from other parishes to come up with an estimate for parish P_1, which is more accurate. Compared with the direct estimate, the composite estimate has smaller variance, but it is not unbiased. Second, it is difficult to obtain the confidence intervals for the composite estimate, since the standard approach for obtaining confidence intervals from a point estimate requires the point estimate to be unbiased, or at least nearly so. Some models give accurate point estimates (such as the random effect model); however, it is difficult to derive a formula for the confidence interval.

2.4 Basic Small Area Estimation Models

Traditional methods of indirect estimation are based on implicit models that provide a link to related small areas through supplementary data. In this section, we explore small area models that make specific allowance for area variation.

We assume that unit-specific auxiliary data $x_{ij} = (x_{ij1}, ..., x_{ijp})^T$ are available for each population element j in each small area i. It is often sufficient to assume that only population means $\bar{X}_i$ are known. The variable of interest y_{ij} is assumed to be related to x_{ij} through a one-fold random effect model:

$$y_{ij} = x_{ij}^T \beta + v_i + e_{ij}; \quad i = 1, ..., N_i, j = 1, ..., m. \tag{2.2}$$

The area-specific effects v_i are assumed to be independent and identically distributed random variables satisfying $E_m(v_i) = 0$, $V_m(v_i) = \sigma_v^2(\geq 0)$, where E_m denotes the model expectation and V_m is the model variance. Hence, we denote this assumption as $v_i \sim (0, \sigma_v^2)$. In the model, define $e_{ij} = k_{ij}\tilde{e}_{ij}$ with known constants k_{ij}, and $\tilde{e}_{ij}$ are iid random variables independent of v_i's and $E_m(\tilde{e}_{ij}) = 0$, $V_m(\tilde{e}_{ij}) = \sigma_e^2$. In addition, normality of the v_i's and e_{ij}'s is often assumed. The interested parameters are the small area means $\bar{Y}_i$ or the totals Y_i.

We assume that a sample s_i of size n_i is taken from the N_i units in the i-th area ($i = 1, ..., m$) and that the sample values also obey the assumed model. The latter assumption is satisfied under simple random sampling from each area or more generally for sampling designs that use the auxiliary information x_{ij} in the selection of the sample s_i. Furthermore, we write the model in the matrix form as

$$y_i^P = X_i^P \beta + v_i 1_i^P + e_i^P, \;\; i = 1, ..., m \tag{2.3}$$

where X_i^P is $N_i \times p$, 1_i^P, e_i^P are $N_i \times 1$ vectors and $1_i^P = (1, ..., 1)^T$.

We write the small area mean $\bar{Y}_i$ as

$$\bar{Y}_i = r_i \bar{y}_i + (1 - r_i)\bar{Y}_i^* \tag{2.4}$$

where $r_i = n_i/N_i$ and $\bar{y}_i$ and $\bar{Y}_i^*$ denoting the means of the sampled and non-sampled elements, respectively. It follows from the above equation that estimating the small area mean $\bar{Y}_i$ is equivalent to estimating the realization of the random variable $\bar{Y}_i^*$ given the sample data y_i and auxiliary data X_i^P.

If the population size N_i is large, then we can take the small area means as

$$\bar{Y}_i = \bar{X}_i^T \beta + v_i + \bar{E}_i \tag{2.5}$$

where $\bar{E}_i$ is the mean of the N_i errors e_{ij} ($\bar{E}_i \approx 0$) and $\bar{X}_i$ is the known mean of X_i^P. It follows from the equation that the estimation of $\bar{Y}_i$ is equivalent to the estimation of a linear combination of β and the realization of the random variable v_i.

2.5 Empirical Best Linear Unbiased Prediction Estimates

Small area means or totals can be expressed as linear combinations of fixed and random effects. Best linear unbiased prediction (BLUP) estimators of such parameters can be obtained in the classical frequentist framework, by appealing to general results on BLUP estimation. BLUP estimators minimize the MSE among the class of linear unbiased estimator and do not depend on normality of the random effects. But they depend on the variances and covariances of random effects, which can be estimated by the method of fitting constants or method of moments. Using the estimated components in the BLUP estimator we could obtain a two-stage estimators, which is referred to as the empirical BLUP (EBLUP) estimator (Harville, 1991).

Suppose that the sample data follow the general linear mixed model

$$y = X\beta + Zv + e. \tag{2.6}$$

Here y is the $n \times 1$ vector of sample observations, X and Z are known $n \times p$ and $n \times h$ matrices of full rank, and v and e are independently distributed with mean 0 and covariance matrices G and R depending on some variance parameters $\delta = (\delta_1, ..., \delta_q)^T$. We assume that δ belongs to a specified subset of Euclidean q-space such that $\text{Var}(y) = V(\delta) = R + ZGZ^T$ is nonsingular for all δ belonging to the subset, where $\text{Var}(y)$ denotes the variance-covariance matrix of y.

Next, we list a special case of the above general linear mixed model, which may cover many small area models considered in the literature. For this model

$$\begin{aligned} y =& \text{col}_{1\le i\le m}(y_i) = (y_1^T, ..., y_m^T), \quad X = \text{col}_{1\le i\le m}(X_i), \\ Z =& \text{diag}_{1\le i\le m}(Z_i), \quad v = \text{col}_{1\le i\le m}(v_i), \quad e = \text{col}_{1\le i\le m}(e_i), \end{aligned}$$

where m is the number of small areas, X_i is $n_i \times p$, Z_i is $n_i \times h_i$ and y_i is an $n_i \times 1$ vector with $\sum n_i = n$ and $\sum h_i = h$. Furthermore,

$$\begin{aligned} R =& \text{diag}_{1\le i\le m}(R_i), \\ G =& \text{diag}_{1\le i\le m}(G_i). \end{aligned}$$

Hence, V has a block diagonal structure

$$V = \text{diag}_{1\le i\le m}(V_i) \tag{2.7}$$

with

$$V_i = R_i + Z_i G_i Z_i^T. \tag{2.8}$$

Therefore, the model could be decomposed into m sub-models, such as

$$y_i = X_i\beta + Z_i v_i + e_i, i = 1, ..., m. \tag{2.9}$$

We are interested in estimating linear combinations $\mu_i = 1_i^T\beta + m_i^T v_i, i = 1, ..., m$.

The BLUP estimator of μ_i is:

$$\tilde{\mu}_i^H = t_i(\delta, y_i) = 1_i^t\tilde{\beta} + m_i^T\tilde{v}_i, \tag{2.10}$$

where

$$\tilde{v}_i = G_i Z_i^T V_i^{-1}(y_i - X_i\tilde{\beta}),$$

and

$$\tilde{\beta} = (\sum_i X_i^T V_i^{-1} X_i)^{-1}(\sum_i X_i^T V_i^{-1} y_i).$$

The MSE of the BLUP estimator is:

$$\text{MSE}(\tilde{\mu}_i^H) = g_{1i}(\delta) + g_{2i}(\delta) \tag{2.11}$$

with

$$g_{1i}(\delta) = m_i^T(G_i - G_i Z_i^T V_i^{-1} Z_i G_i)m_i,$$

and

$$g_{2i}(\delta) = d_i^T(\sum_i X_i^T V_i^{-1} X_i)^{-1} d_i,$$

where

$$d_i^T = 1_i^T - b_i^T X_i,$$

with

$$b_i^T = m_i^T G_i Z_i^T V_i^{-1}.$$

Replacing δ by an estimator of $\hat{\delta}$, we get the EBLUP estimator

$$\hat{\mu}_i^H = t_i(\hat{\delta}, y_i) = 1_i^T \hat{\beta} + m_i^T \hat{v}_i. \tag{2.12}$$

In the following part, we consider the basic unit level model and spell out EBLUP estimation, using the general results for the general linear mixed model with block diagonal covariance structure.

Take the i-th small area mean as $\mu_i = \bar{X}_i^T \beta + v_i$, if the population size N_i of the small areas are sufficiently large. In this case, we use the sample part of the model, $y_{ij} = x_{ij}^T \beta + v_i + e_{ij}$, $j = 1, ..., n_i$, $i = 1, ..., m$ which could be written in matrix notation as

$$y_i = X_i \beta + v_i 1_{n_i} + e_i, i = 1, ..., m \tag{2.13}$$

to make inference on $\bar{Y}_i$, by appealing to the general results.

The model (2.13) is a special case of the general model (2.9) with block diagonal covariance structure. We have,

$$y_i = y_i, \;\; X_i = X_i, \;\; Z_i = 1_{n_i},$$
$$v_i = v_i, \;\; e_i = e_i, \quad \beta = (\beta_1, ..., \beta_p)^T,$$

where y_i is the $n_i \times 1$ vector of sample observations y_{ij} from the ith area. Furthermore,

$$G_i = \sigma_v^2, \quad R_i = \sigma_e^2 \text{diag}_{1 \le j \le n_i}(k_{ij}^2),$$

so that

$$V_i = R_i + \sigma_v^2 1_{n_i} 1_{n_i}^T.$$

Also, $\mu_i = \bar{X}_i^T \beta + v_i$ so that $1_i = \bar{X}_i$ and $m_i = 1$. The matrix V_i can be inverted explicitly as

$$V_i^{-1} = \frac{1}{\sigma_e^2}[\text{diag}_j(a_{ij}) - \frac{\gamma_i}{a_{i\cdot}} a_i a_i^T] \tag{2.14}$$

using the following standard result on matrix inversion:

$$(A + uv^T)^{-1} = A^{-1} = A^{-1}uv^T A^{-1}/(1 + v^T A^{-1}u). \tag{2.15}$$

Here we have

$$a_{ij} = k_{ij}^{-2}, a_{i\cdot} = \sum_i a_{ij}, a_i = (a_{i1}, ..., a_{in_i})^T \tag{2.16}$$

and

$$\gamma_i = \sigma_v^2/(\sigma_v^2 + \sigma_e^2/a_{i\cdot}). \tag{2.17}$$

Making the above substitution in the general formula (2.10) and noting that $(\sigma_v^2/\sigma_e^2)(1-\gamma_i) = \gamma_i/a_i$. We get the BLUP estimator of μ_i as

$$\tilde{\mu}_i^{\text{H}} = \bar{X}_i^T \tilde{\beta} + \gamma_i(\bar{y}_{ia} - \bar{x}_{ia}^T \tilde{\beta}), \tag{2.18}$$

where $\bar{y}_{ia}$ and $\bar{x}_{ia}$ are weighted means given by

$$\bar{y}_{ia} = \Sigma_j a_{ij} y_{ij}/a_{i\cdot}, \quad \bar{x}_{ia} = \Sigma_j a_{ij} x_{ij}/a_{i\cdot}, \tag{2.19}$$

and $\tilde{\beta}$ is the BLUE of β:

$$\tilde{\beta} = \left(\Sigma_i X_i^T V_i^{-1} X_i\right)^{-1} \left(\Sigma_i X_i^T V_i^{-1} y_i\right), \tag{2.20}$$

where

$$X_i^T V_i^{-1} X_i = A_i = \sigma_e^{-2}\left(\Sigma_j a_{ij} x_{ij} x_{ij}^T - \gamma_i a_{i\cdot} \bar{x}_{ia} \bar{x}_{ia}^T\right), \tag{2.21}$$

and

$$X_i^T V_i^{-1} y_i = \sigma_e^{-2}\big(\Sigma_j a_{ij} x_{ij} y_{ij} - \gamma_i a_{i\cdot} \bar{x}_{ia} \bar{y}_{ia}\big). \tag{2.22}$$

The BLUP estimator (2.18) can also be expressed as a weighted average of the "survey regression" estimator $\bar{y}_{ia} + (\bar{X}_i - \bar{x}_{ia})^T \tilde{\beta}$ and the regression synthetic estimator $\bar{X}_i^T \tilde{\beta}$:

$$\tilde{\mu}_i^{\mathrm{H}} = \gamma_i[\bar{y}_{ia} + (\bar{X}_i - \bar{x}_{ia})\tilde{\beta}] + (1 + \gamma_i)\bar{X}_i^T \tilde{\beta}. \tag{2.23}$$

The weight $\gamma_i\,(0 \leq \gamma_i \leq 1)$ measures the model variance σ_ν^2, relative to the total variance $\sigma_v^2 + \sigma_e^2/a_{i\cdot}$. If the model variance is relatively small, then γ_i will be small and more weight is attached to the synthetic component. Similarly, more weight is attached to the survey regression estimator as $a_{i\cdot}$ increases. Note that $a_{i\cdot}$ is of order $O(n_i)$ and it reduces to n_i if $k_{ij} = 1$ for all (i, j). Also, in the latter case the survey regression estimator is approximately design-unbiased for μ_i under simple random sampling, provided the total sample size $n = \Sigma_i n_i$ is large.

In the case of general $k_{ij}'s$, it is model-unbiased for μ_i conditional on the realized local effect v_i, provided $\tilde{\beta}$ is conditionally unbiased for β. On the other hand, the BLUP estimator (2.23) is conditionally biased due to the presence of the synthetic component $\bar{X}_i^T \tilde{\beta}$. Under simple random sampling and $k_{ij} = 1$ for all (i, j), the BLUP estimator is design-consistent for $\bar{Y}_i$ as n_i increases because $\gamma_i \to 1$.

The MSE of the BLUP estimator could be obtained either directly or from the general result (2.11) by letting $\delta = (\sigma_v^2, \sigma_e^2)^T$. It is given by

$$\mathrm{MSE}(\tilde{\mu}_i^{\mathrm{H}}) = E(\tilde{\mu}_i^{\mathrm{H}} - \mu_i)^2 = g_{1i}(\sigma_v^2, \sigma_e^2) + g_{2i}(\sigma_v^2, \sigma_e^2), \tag{2.24}$$

where

$$g_{1i}(\sigma_v^2, \sigma_e^2) = r_i(\sigma_e^2/a_{i\cdot}) \tag{2.25}$$

and

$$g_{2i}(\sigma_v^2, \sigma_e^2) = (\bar{X}_i - \gamma_i \bar{x}_{ia})^T \left(\Sigma_i A_i\right)^{-1} (\bar{X}_i - \gamma_i \bar{x}_{ia}) \tag{2.26}$$

with A_i given by (2.21). The first term, $g_{1i}(\sigma_v^2, \sigma_e^2)$, is of order $O(1)$, whereas the second term, $g_{2i}(\sigma_v^2, \sigma_e^2)$, is of order $O(m^{-1})$ for large m, assuming the following regularity conditions:

(i) k_{ij} and n_i are uniformly bounded.

(ii) Elements of X_i are uniformly bounded such that A_i is of order $O(1)$.

The leading term of the MSE of the BLUP estimator is given by $g_{1i}(\sigma_v^2, \sigma_e^2) = \gamma_i(\sigma_e^2/a_{i\cdot})$. Comparing this term to $\sigma_e^2/a_{i\cdot}$, the leading term of the MSE of the sample regression estimator, it is clear that the BLUP estimator provides considerable gain in efficiency over the sample regression estimator γ_i is small. Therefore, models with smaller γ_i, which is the sample variance relative to the total variance, should be preferred, provided they provide an adequate fit in terms of residual analysis and other model diagnostics (Rao, 2003).

The BLUE $\tilde{\beta}$ and its covariance matrix $(\Sigma_i X_i^T V_i^{-1} X_i)^{-1}$ can be calculated using ordinary least squares (OLS) by first transforming the model (2.13) with correlated errors $u_{ij} = v_i + e_{ij}$ to a model with uncorrelated errors u_{ij}^*. The transformed model is given by

$$k_{ij}^{-1}(y_{ij} - \tau_i \bar{y}_{ia}) = k_{ij}^{-1}(x_{ij} - \tau_i \bar{x}_{ia})^T \beta + u_{ij}^*, \tag{2.27}$$

where $\tau_i = 1 - (1 - \gamma_i)^{1/2}$ and the u_{ij}^*'s have mean zero and constant variance σ_e^2 (Stukel and Rao, 1997). If $k_{ij} = 1$ for all (i, j), equation (2.27) reduces to the transformed model of Fuller and Battese (1973). In practice, τ_i is estimated from the data.

The BLUP estimator (2.23) depends on the variance ratio σ_v^2/σ_e^2, which is unknown in practice. Replacing σ_v^2 and σ_e^2 by estimators $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$, we obtain and EBLUP estimator

$$\hat{\mu}_i^{\mathrm{H}} = \hat{\gamma}_i |[\bar{y}_{ia} + (\bar{X}_i - \bar{x}_{ia})^T \hat{\beta}] + (1 - \hat{\gamma}_i) \bar{X}_i^T \hat{\beta}, \tag{2.28}$$

where $\hat{\gamma}_i$ and $\hat{\beta}$ are the values of γ_i and $\tilde{\beta}$ when (σ_v^2, σ_e^2) is replaced by $(\hat{\sigma}_v^2, \hat{\sigma}_e^2)$.

Chapter 3. Bayesian Analysis

3.1 Basic Theory

Bayesian econometrics is based on simple rules of probability, which is one of the chief advantages of the Bayesian approach. Bayesian methods can be used in estimating the parameters of a model, comparing different models, or obtaining predictions from a model. Hence, the researchers can use Bayesian methods to learn about a phenomenon by using data.

To motivate the simplicity of the Bayesian approach, let us consider two random variables, A and B. The rules of probability imply:

$$p(A, B) = p(A|B)p(B)$$

where $p(A, B)$ is the joint probability of A and B occurring, $p(A|B)$ is the probability of A occurring conditional on B having occurred, and $p(B)$ is the marginal probability of B. Alternatively, we can reverse the roles of A and B and get an expression for the joint probability of A and B:

$$p(A, B) = p(B|A)p(A).$$

Equating these two expressions for $p(A, B)$ and rearranging provides us with Bayes's Rule, which lies as the core theory in Bayesian Econometrics:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}. \tag{3.1}$$

In economics, we work with models which depend upon parameters. For the regression model, the researchers are interested in estimating the coefficients. In this case, the coefficients are the parameters under study. Let y be a vector of matrix of data and θ be a vector or matrix which contains the parameters for a model that seeks to explain y. We are interested in learning about θ based on the data y. We could rewrite the core theory of Bayesian as:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}, \tag{3.2}$$

where $p(\theta)$ is the assumed prior distribution of the unknown parameters θ, $p(y|\theta) = l(\theta|y)$ is the likelihood function. Bayesians treat $p(\theta|y)$ as being of fundamental interest, which is the posterior distribution given the prior of the unknown parameters $p(\theta)$ and the data y. Intuitively, the Bayesian approach addresses the question, Given the data, what do we know about θ? After establishing the $p(\theta|y)$ as the fundamental interest for the econometrician interested in using data to learn about parameters in a model, we return to equation (3.2). Since we are only interested in learning about θ, the term $p(y)$ is essentially a constant with respect to θ. We can write the posterior distribution as:

$$p(\theta|y) \propto p(y|\theta)p(\theta), \tag{3.3}$$

where the symbol $\propto$ signifies that the posterior distribution is "proportional" to the likelihood augmented with the prior.

3.2 Empirical Bayes Methods

In the previous chapter, we discuss the Empirical Best Linear Unbiased Prediction, which is applicable to linear models. However, the linear mixed models are designed for continuous variables, while they are not suitable for handling binary or count data. Empirical Bayes and hierarchical Bayes methods are applicable in handling binary and count data, which will be discussed in the following two sections.

Morris (1983) lists an excellent account of the Empirical Bayes approach, which could be summarized as follows:

Step 1. Obtain the posterior density, $f(\mu|y,\lambda)$ of the small area parameters of interest μ, given the data y, using the conditional density $f(y|\mu,\lambda_1)$ of y given μ and the density $f(y|\mu,\lambda_2)$ of μ, where $\lambda = (\lambda_1^T, \lambda_2^T)^T$ denotes the vector of model parameters.

Step 2. Estimate the model parameters, λ, from the marginal density, $f(y|\lambda)$.

Step 3. Use the estimated posterior density, $f(\mu|y,\hat{\lambda})$, for making inferences about μ, where $\hat{\lambda}$ is an estimator of λ.

Assuming normality, the linear mixed model with block diagonal covariance structure may be expressed as

$$y_i|v_i \overset{ind}{\sim} N(X_i\beta + Z_i v_i, R_i) \tag{3.4}$$

$$v_i \overset{ind}{\sim} N(0, G_i), i = 1, ..., m, \tag{3.5}$$

where G_i and R_i depend on variance parameters δ. The Bayes estimator of realized $\mu_i = 1_i^T\beta + m_i^T v_i$ is given by the conditional expectation of μ_i given y_i, β and δ:

$$\hat{\mu_i}^B(\beta,\delta) = E(\mu_i|y_i,\beta,\delta) = 1_i^T\beta + m_i^T\hat{v_i}^B \tag{3.6}$$

where

$$\hat{v_i}^B = E(v_i|y_i,\beta,\delta) = G_i Z_i^T V_i^{-1}(y_i - X_i\beta)$$

and $V_i = R_i + Z_i G_i Z_i^T$. The results (3.6) follow from the posterior distribution of μ_i given y_i:

$$\mu_i|y_i,\beta,\delta \overset{ind}{\sim} N(\hat{\mu_i}^B, g_{1i}(\delta)), \tag{3.7}$$

where $g_{1i}(\delta)$ is given by equation (2.11).

The estimator $\hat{\mu}^B$ depends on the model parameters β and δ which are estimated from the marginal distribution

$$y_i \overset{ind}{\sim} N(X_i\beta, V_i), i = 1, ..., m. \tag{3.8}$$

Denoting the estimators as $\hat{\beta}$ and $\hat{\delta}$, we obtain the empirical Bayes estimator of μ_i from μ_i^B for B and $\hat{\delta}$ for δ:

$$\hat{\mu_i}^H B(\hat{\beta}, \hat{\delta}) = 1_i^T \hat{\beta} + m_i^T \hat{v_i}^B(\hat{\beta}, \hat{\delta}). \tag{3.9}$$

Therefore, the EB estimator $\hat{\mu_i}^{EB}$ is identical to the EBLUP estimator (2.12).

3.3 Hierarchical Bayes Methods

In the hierarchical Bayes (HB) approach, a subjective prior distribution $f(\lambda)$ on the model parameters λ is specified. Moreover, given the data y, the posterior distribution $f(\mu|y)$ of the small area parameters of interest μ is obtained. Using Bayes theorem, the two-stage model, $f(y|\mu, \lambda_1)$ and $f(y|\mu, \lambda_2)$, is combined with the subjective prior on $\lambda = (\lambda_1^T, \lambda_2^T)^T$ to arrive at the posterior $f(\mu|y)$ by using the Bayes theorem. Here inferences are based on $f(\mu|y)$. In particular, a parameter of interest, say $\phi = h(\mu)$, is estimated by its posterior mean $\hat{\phi}^{HB} = E[h(\mu)|y]$. The posterior variance $V[h(\mu)|y]$ is used as a measure of precision of the estimator, provided they are finite.

The Hierarchical Bayes approach is straightforward, and its inferences are "exact." But the inferences require the specification of a subjective prior $f(\lambda)$ on the model parameters λ. Priors on λ might be informative, when based on substantial prior information, such as previous studies judged relevant to the current data set y.

On the other hand, diffuse (or noninformative) priors are designed to reflect lack of information about λ. One may take different choices for a diffuse prior, and some diffuse improper priors could lead to improper posteriors. Moreover, under the frequentist framework, it is desirable to select a diffuse prior that leads to well-calibrated inferences for the sake of validity. In practice, both the frequentist bias $E(\hat{\phi}^{HB} - \phi)$ of the HB estimator $\hat{\phi}^{HB}$ and the relative frequentist bias of the posterior variance as an estimator of MSE $(\hat{\phi}^{HB})$ should be small (Browne and Draper, 2006).

Datta, Fay and Ghosh (1991) applied the hierarchical Bayes approach to the estimation of small area means $\bar{Y}$'s, under general mixed linear models. In the HB approach, a prior distribution on the model parameters is specified, which is equivalent to assuming β has a uniform distribution. Then, the posterior distribution of the parameters of interest is obtained, and the interested parameter is estimated by its posterior mean and its precision is measured by its posterior variance.

Applying Bayes theorem, we have

$$f(\mu, \lambda|y) = \frac{f(y, \mu|\lambda)f(\lambda)}{f_1(y)}, \tag{3.10}$$

where $f_1(y)$ is the marginal density of y:

$$f_1(y) = \int f(y, \mu|\lambda)f(\lambda)d\mu d\lambda. \tag{3.11}$$

The desired posterior density $f(\mu|y)$ is obtained from (3.10) as

$$f(\mu|y) = \int f(\mu, \lambda|y)d\lambda \tag{3.12}$$

$$= \int f(\mu|y, \lambda)f(\lambda|y)d\lambda. \tag{3.13}$$

It follows from (3.13) that $f(\mu|y)$ is a mixture of conditional densities $f(\mu|y, \lambda)$. Here $f(\mu|y, \lambda)$ is used for EB inferences.

It is clear from (3.10) and (3.12) that the evaluation of $f(\mu|y)$ and associated posterior quantities, such as $E[h(\mu)|y]$, involves multi-dimensional integrations. However, it is possible to simply analytically perform integration with respect to some of the components of μ and λ. If the reduced problem involves only one- or two-dimensional integration, it can use direct numerical integration to calculate the desired posterior quantities. For complex problems, Markov Chain Monte Carlo (MCMC) methods are broadly used to evaluate high dimensional integrals, which will be discussed in the following section. The required regularity conditions will be also be discussed later. The MCMC methods have the desired properties that overcome the computational difficulties to a large extent, while these methods generate samples from the posterior distribution.

3.4 Markov Chain Monte Carlo Methods

Obtaining the posterior distribution function is difficult and computationally intensive, requiring the calculation of high dimensional integrals or sampling from unknown distribution. Before the 1990s, the evaluation of the posterior distribution represented the major issue in the empirical application of Bayesian analysis. The development and implementation of Markov Chain Monte Carlo (MCMC) methods have overcome the computational difficulties to a large extent.

3.4.1 Markov Chain

Let $\eta = (\mu^T, \lambda^T)^T$ be the vector of small area parameters μ and model parameters λ. It is in general not feasible to draw independent samples from the joint posterior $f(\eta|y)$ because of the intractable denominator $f_1(y)$. MCMC methods avoid this difficulty by constructing a Markov chain $\{\eta^{(k)}, k = 0, 1, 2, \cdots\}$ such that the distribution of $\eta^{(k)}$ converges to a unique stationary distribution equal to $f(\eta|y)$, denoted by $\pi(\eta)$. Therefore, after a sufficiently large "burn-in," say d, we can regard $\eta^{(d+1)}, \cdots, \eta^{(d+D)}$ as D dependent samples from the target distribution $f(\eta|y)$, regardless of the starting point $\eta^{(0)}$.

To construct a Markov Chain, we need to specify a one-step transition probability $P(\eta^{(k+1)}|\eta^{(k)})$ which depends only on the current "state" $\eta^{(k)}$ of the chain, which means that the conditional distribution of $\eta^{(k+1)}$ given $\eta^{(0)}, \cdots \eta^{(k)}$ does not depend on the previous $\{\eta^{(0)}, \cdots \eta^{(k-1)}\}$. Meanwhile, the transition kernel must satisfy the stationarity condition:

$$\int \pi(\eta^{(k)}) P(\eta^{(k+1)}|\eta^{(k)}) d\eta^{(k)} = \pi(\eta^{(k+1)}). \tag{3.14}$$

Equation (3.14) shows that if $\eta^{(k)}$ is from $\pi(\cdot)$, then $\eta^{(k+1)}$ will also be from $\pi(\cdot)$. Stationarity is satisfied if the chain is "reversible":

$$\pi(\eta^{(k)})P(\eta^{(k+1)}|\eta^{(k)}) = \pi(\eta^{(k+1)})P(\eta^{(k)}|\eta^{(k+1)}). \tag{3.15}$$

It follows from (3.15) that the stationary distribution of the chain generated by $P(\cdot|\cdot)$ is $\pi(\cdot)$.

It is also necessary to make sure that $P(k)(\eta^{(k)}|\eta^{(0)})$, which denotes the distribution of $\eta^{(k)}$ given $\eta^{(0)}$, converges to $\pi(\eta^{(k)})$ regardless of $\eta^{(0)}$. Thus the chain needs to be "irreducible" and "aperiodic" (Rao, 2003). Irreducibility means that from all starting points $\eta^{(0)}$ the chain will eventually reach any nonempty set in the state space with positive probability. Aperiodicity means that the chain is not permitted to oscillate between different sets in a periodic manner. For an irreducible and aperiodic chain, the following theorem holds:

$$\bar{h}_D = \frac{1}{D}\sum_{k=d+1}^{d+D} h(\eta^{(k)}) \rightarrow_p E[h(\eta)|y], \tag{3.16}$$

as $D \rightarrow \infty$, where $\rightarrow_p$ denotes convergence in probability. Therefore, for sufficiently large D, we are able to obtain an estimator $\bar{h}_D$, of $E[h(\eta)|y]$ with adequate precision.

3.4.2 Gibbs Sampler

The Gibbs Sampler, also called alternating conditional sampling, is another core of the Markov Chain algorithm. In order to generate the samples $\eta^{(k)}$, following Rao (2003), we partition η into suitable blocks $\eta_1, \cdots, \eta_r$. Some of the blocks may contain only single elements, while others contain more than one element. For instance, consider the basic unit level model with $\mu = (\theta_1, \cdots, \theta_m)^T = \theta$ and $\lambda = (\beta^T, \sigma_\nu^2)^T$. In this case η may be partitioned as $\eta_1 = \beta$, $\eta_2 = \theta_1, \cdots, \eta_{m+1} = \theta_m, \eta_{m+2} = \sigma_\nu^2$, hence $r = m + 2$. The following set of Gibbs conditional distributions is needed: $f(\eta_1|\eta_2, \cdots, \eta_r, y)$, $f(\eta_2|\eta_1, \eta_3, \cdots, \eta_r, y), \cdots, f(\eta_r|\eta_1, \cdots, \eta_{r-1}, y)$. The Gibbs sampler uses these conditional distributions to construct a transition kernel, $P(\cdot|\cdot)$, such that the stationary distribution of the resulting Markov Chain is $\pi(\eta) = f(\eta|y)$. This result follows from the fact that $f(\eta|y)$ is uniquely determined by the set of Gibbs conditionals.

For the standard conditional distribution, such as normal inverse-gamma, samples can be generated directly from the conditional distribution. Otherwise, Metropolis-Hastings (M-H) rejection sampling, can be used to generate samples from the conditional distribution. Therefore, the Gibbs sampler represents a special case of M-H algorithm.

The Gibbs sampling algorithm involves the following steps:

Step 0. Choose a starting point $\eta^{(0)}$ with components $\eta_1^{(0)}, \cdots, \eta_r^{(0)}$; set $k = 0$. For example, we could use REML or moment estimates of model parameters λ and EB estimates of μ as starting values.

Step 1. Generate $\eta^{(k+1)} = (\eta_1^{(k+1)}, \cdots, \eta_r^{(k+1)})$ as follows:
draw $\eta_1^{(k+1)}$ from $f(\eta_1|\eta_2^{(k)}, \cdots, \eta_r^{(k)}, y)$;
draw $\eta_2^{(k+1)}$ from $f(\eta_2|\eta_1^{(k+1)}, \eta_3^{(k)}, \cdots, \eta_r^{(k)}, y)$;
$\cdots$;
draw $\eta_r^{(k+1)}$ from $f(\eta_r|\eta_1^{(k+1)}, \cdots, \eta_{r-1}^{(k+1)}, y)$.

Step 2. Set $k = k + 1$ and go to Step 1.

Steps 1 and 2 constitute one cycle for each k. The sequence $\{\eta^{(k)}\}$ generated by the Gibbs sampler is a Markov chain with stationary distribution $\pi(\eta) = f(\eta|y)$ (Gelfand and Smith, 1990).

3.4.3 Choice of a Prior

Diffuse priors $f(\lambda)$, reflecting a lack of information about the model parameters λ, are commonly used in the HB approach to small area estimation. For informative data, the posterior distribution is robust over a wide range of priors. On the other hand, for non-informative data, the characteristics of the prior used, such as location (mean) and precision (the inverse of the variance), become especially meaningful for the posterior distribution. For instance, if the diffuse is improper, such that $\int f(\lambda)d\lambda = \infty$, then the Gibbs sampler could lead to seemingly reasonable inferences about a nonexistent posterior $f(\mu, \lambda|y)$. As pointed out by Natarajan and McCulloch (1995), Hobert and Casella (1996), this happens when the posterior is improper and yet all the Gibbs conditionals are proper.

Consider the simple nested error model without covariates:

$$y_{ij} = \mu + v_i + e_{ij},$$

where $v_i \overset{iid}{\sim} N(0, \sigma_v^2)$ and $e_{ij} \overset{iid}{\sim} N(0, \sigma_e^2)$. Hill (1965) points out that, if we choose an improper prior of the form $f(\mu, \sigma_v^2, \sigma_e^2) = f(\mu)f(\sigma_v^2)f(\sigma_e^2)$ with $f(\mu) \propto 1$, $f(\sigma_v^2) \propto \sigma_v^{-2}$ and $f(\sigma_e^2) \propto \sigma_e^{-2}$, then the joint posterior of μ, $v = (v_1, \cdots, v_m)^T$, σ_v^2 and σ_e^2 is improper. On the other hand, all the Gibbs conditionals are proper for this choice of prior. Particularly, σ_v^{-2} conditional on all others follows a gamma distribution, while σ_e^{-2} conditional on all others also follows a gamma distribution. Meanwhile, μ conditional on others follows normal distributions, as well as v_i.

Following Gilks et al. (1995), we use diffuse proper priors of the form $\mu\ N(0, \sigma_0^2)$, $\sigma_v^{-2}\ G(a_0, a_0)$ and $\sigma_e^{-2}\ G(a_0, a_0)$ as default priors, where σ_0^2 is chosen very large (say 10,000) and a_0 very small (say 0.001) to reflect lack of prior information on μ, σ_v^2 and σ_e^2[1]. The posterior resulting from the above prior remains proper as $\sigma_0^2 \to \infty$, but it becomes improper as $a_0 \to 0$. Therefore, the posterior is nearly improper for very small a_0, and this feature can affect the convergence of the Gibbs sampler (Rao, 2003).

[1] Note that $G(a, b)$ denotes a gamma distribution with shape parameter a and scale parameter b and that the variance of $G(a_0, a_0)$ is $1/a_0$ which becomes very large as $a_0 \to 0$.

3.5 Basic Unit Level Model

In this section, following Rao's (2003) study, we apply the HB approach to the basic unit level model (2.13) with equal error variances (that is, $k_{ij} = 1$), assuming a prior distribution on the model parameters $(\beta, \sigma_v^2, \sigma_e^2)$.

We first consider the case of known σ_ν^2 and σ_e^2, and assume a "flat" prior on β: $f(\beta) \propto 1$. We rewrite (2.13) as a HB model:

(i) $y_{ij}|\beta, v_i, \sigma_e^2 \overset{ind}{\sim} N(x_{ij}^T\beta + v_i, \sigma_e^2),\ j = 1, \ldots, n_i;\ i = 1, \ldots, m$

(ii) $v_i|\sigma_v^2 \overset{iid}{\sim} N(0, \sigma_v^2),\ i = 1, \ldots, m$

(iii) $f(\beta) \propto 1$.

We then extend the results to the case of unknown σ_v^2 and σ_e^2 in the above HB model, with replacing the condition (iii) by

$$f(\beta, \sigma_v^2, \sigma_e^2) = f(\beta)f(\sigma_v^2)f(\sigma_e^2) \propto f(\sigma_v^2)f(\sigma_e^2). \tag{3.17}$$

where $f(\sigma_v^2)$ and $f(\sigma_e^2)$ are the priors on σ_v^2 and σ_e^2. For simplicity, we take $\mu_i = \bar{X}_i^T\beta + v_i$ as the i-th small area mean, assuming the population size, N_i, is large.

3.5.1 Known σ_v^2 and σ_e^2

Ideally, assume σ_v^2 and σ_e^2 are known, and a flat prior of β, the HB and BLUP approaches under normality leading to identical point estimates and measures of variability. This result is valid for a general linear mixed model with known variance parameters. Hence, the HB estimator of μ_i is given by:

$$\tilde{\mu}_i^{\text{HB}}(\sigma_v^2, \sigma_e^2) = E(\mu_i|y, \sigma_v^2, \sigma_e^2) = \tilde{\mu}_i^{\text{H}}, \tag{3.18}$$

where y is the vector of sample observations and $\tilde{\mu}_i^{\mathrm{H}}$ is the BLUP estimator given by (2.18). Similarly, the posterior variance of μ_i is

$$V(\mu_i|\sigma_v^2, \sigma_e^2, y) = M_{1i}(\sigma_v^2, \sigma_e^2) = \mathrm{MSE}(\tilde{\mu}_i^H), \tag{3.19}$$

where $M_{1i}(\sigma_v^2, \sigma_e^2)$ is given by (2.24).

3.5.2 Unknown σ_v^2 and σ_e^2

Practically, σ_v^2 and σ_e^2 are unknown and it is necessary to take account of the uncertainty about σ_v^2 and σ_e^2 by assuming a prior on σ_v^2 and σ_e^2. The HB model is given by (i) and (ii) and the equation given by (3.17). We obtain the HB estimator of μ_i and the posterior variance of μ_i as

$$\hat{\mu}_i^{\mathrm{HB}} = E(\mu_i|y) = E_{\sigma_v^2,\sigma_e^2}[\tilde{\mu}_i^{\mathrm{HB}}(\sigma_v^2, \sigma_e^2)] \tag{3.20}$$

and

$$V(\mu_i|y) = E_{\sigma_v^2,\sigma_e^2}[M_{1i}(\sigma_v^2, \sigma_e^2)] + V_{\sigma_v^2,\sigma_e^2}[\tilde{\mu}_i^{\mathrm{HB}}(\sigma_v^2, \sigma_e^2)], \tag{3.21}$$

where $E_{\sigma_v^2,\sigma_e^2}$ and $V_{\sigma_v^2,\sigma_e^2}$, denote the expectation and variance with respect to the posterior distribution $f(\sigma_v^2, \sigma_e^2|y)$, respectively.

For the basic unit level model, the posterior $f(\sigma_v^2, \sigma_e^2|y)$ could be obtained from the restricted likelihood function $L_R(\sigma_v^2, \sigma_e^2)$ as

$$f(\sigma_v^2, \sigma_e^2|y) \propto L_R(\sigma_v^2, \sigma_e^2) f(\sigma_v^2) f(\sigma_e^2). \tag{3.22}$$

Under flat priors $f(\sigma_v^2) \propto 1$ and $f(\sigma_e^2) \propto 1$, the posterior $f(\sigma_v^2, \sigma_e^2|y)$ is proper and proportional to $L_R(\sigma_v^2, \sigma_e^2)$. Evaluation of the posterior mean (3.20) and the posterior variance (3.21), using $f(\sigma_v^2, \sigma_e^2|y) \propto L_R(\sigma_v^2, \sigma_e^2)$, involves two-dimensional integration.

If we assume a diffuse gamma prior, $G(a_e, b_e)$ with $a_e \geq 0$ and $b_e > 0$, then it is possible to integrate out σ_e^2 with respect to $f(\sigma_v^2, \sigma_e^2|y)$, where $\tau_v = \sigma_v^2/\sigma_e^2$. The evaluation of (3.20) and (3.21) is now reduced to single-dimensional integration with respect to the posterior of τ_v, $f(\tau_v|y)$. Datta and Ghosh (1991) expressed $f(\tau_v)$ as $f(\tau_v|y) \propto h(\tau_v)$ and obtained an explicit expression for $h(\tau_v)$, assuming a gamma prior on τ_v^{-1}: $G(a_v, b_v)$ with $a_e \geq 0$ and $b_e \geq 0$; note that a_v is the shape parameter and b_v is the scale parameter.

Next we apply Gibbs sampling to the basic unit level model, assuming the prior (3.17) on $\beta, \sigma_v^2, \sigma_e^2$ with $\sigma_v^{-2} \sim G(a_v, b_v)$, $a_v \geq 0$, $b_v > 0$ and $\sigma_e^{-2} \sim G(a_e, b_e)$, $a_e \geq 0$, $b_e > 0$.

The precision parameter of each of the variance components is assumed to follow an inverse gamma distribution with different parameters, $\sigma_e^2 \sim IG(\lambda_1, \tau_1)$ and $\sigma_v^2 \sim IG(\lambda_2, \tau_2)$. The joint posterior distribution function is as follows,

$$f(\beta, \sigma_v^2, \sigma_e^2 | y_{ij}, 1 \leq j \leq n, 1 \leq i \leq m) =$$

$$\prod_{i=1}^{m}[\prod_{j=1}^{n_i}(\frac{1}{\sigma_e^2})^{\frac{1}{2}} e^{-\frac{1}{2\sigma_e^2}(y_{ij}-x_{ij}^T\beta-v_i)^2}(\frac{1}{\sigma_v^2})^{\frac{1}{2}} e^{-\frac{1}{2\sigma_v^2}v_i^2}]$$

$$\times[\prod_{l=1}^{p}(\frac{1}{h_l^2})^{\frac{1}{2}} e^{-\frac{1}{2h_l^2}\beta_l^2}](\frac{1}{\sigma_e^2})^{\lambda_1+1} e^{-\frac{\tau_1}{\sigma_e^2}}(\frac{1}{\sigma_v^2})^{\lambda_2+1} e^{-\frac{\tau_2}{\sigma_v^2}} \tag{3.23}$$

Solving for the marginal posterior distribution from equation (3.23) gives the following conditions.

$$\beta | y_{ij}, v_i, \sigma_v^2, \sigma_e^2 \sim N_p(\Lambda \sigma_e^{-2} \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - v_i)x_{ij}, \Lambda) \tag{3.24}$$

$$v_i | y_{ij}, \beta, \sigma_v^2, \sigma_e^2 \sim N((n_i + \frac{\sigma_e^2}{\sigma_v^2})^{-1}\sum_{j=1}^{n_i}(y_{ij} - x_{ij}^T\beta), (\frac{n_i}{\sigma_e^2} + \frac{1}{\sigma_v^2})^{-1}) \tag{3.25}$$

$$\sigma_e^2 | y_{ij}, \beta, v_i, \sigma_v^2 \sim G(\lambda_1 + \frac{1}{2}\sum_{i=1}^{m} n_i, \tau_1 + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - x_{ij}^T\beta - v_i)^2) \tag{3.26}$$

$$\sigma_v^2 | y_{ij}, \beta, v_i, \sigma_e^2 \sim G(\lambda_2 + \frac{m}{2}, \tau_2 + \frac{1}{2}\sum_{i=1}^{m} v_i^2) \tag{3.27}$$

where $\Lambda = (\sigma_e^{-2}\sum_{i=1}^{m}\sum_{j=1}^{n_i} x_{ij}x_{ij}^T + H^{-1})^{-1}$.

In the first stage of our estimation, we use equation (3.24) to (3.27) in Gibbs sampling (Gelfard and Smith, 1990) to simulate the marginal posterior distributions of σ_e^2, σ_u^2. In particular, we set $\lambda_1 = 1$, $\tau_1 = 0.002$, $\lambda_2 = 1$, and $\tau_2 = 0.002$.

The Gibbs sampling is based on these conditions:

(1) Set $y^{*[0]} = y$ and apply the starting values for $\sigma_e^{2[0]}$ and $\sigma_v^{2[0]}$;

(2) Draw $\beta^{[1]}|y^{*[0]}, \sigma_e^{2[0]}, \sigma_v^{2[0]}$ from equation (3.24);

(3) Using the drawn $\beta^{[1]}$ and the initial values for $\sigma_e^{2[0]}$ and $\sigma_v^{2[0]}$, draw and update $v_i^{[1]}$ with equation (3.25);

(4) Draw and update $\sigma_e^{2[1]}$ conditional on initial $\sigma_v^{2[0]}$ and the updated values of $\beta^{[1]}, v_i^{[1]}$;

(5) Draw and update $\sigma_v^{2[1]}$ given new values of $\beta^{[1]}, v_i^{[1]}$ and $\sigma_e^{2[1]}$.

The process is repeated 25,000 times to product 25,000 draws for each conditional marginal posterior, and the first 5,000 draws were burnt.

Next, the Markov Chain Monte Carlo (MCMC) methods are used to generate samples from the posterior distribution, and then used in the simulated samples to approximate the desired posterior quantities.

Denote the MCMC samples from a single large run by $\{\beta^{(k)}, v^{(k)}, \sigma_v^{2(k)}, \sigma_e^{2(k)},\ k = d + 1, \ldots, d + D\}$. The marginal MCMC samples[2] $\{\beta^{(k)}, v^{(k)}\}$ can be used directly to estimate the posterior mean of μ_i as

$$\hat{\mu}_i^{\mathrm{HB}} = \frac{1}{D} \sum_{k=d+1}^{d+D} \mu_i^{(k)}, \tag{3.28}$$

where $\mu_i^{(k)} = \bar{X}_i^T \beta^{(k)} + v_i^{(k)}$. Similarly, the posterior variance of $\mu_i$is estimated as

$$V(\mu_i|y) = \frac{1}{D-1} \sum_{k=d+1}^{d+D} (\mu_i^{(k)} - \hat{\mu}_i^{\mathrm{HB}})^2. \tag{3.29}$$

[2]The sequence $\{\beta^{(k)}, v^{(k)}\}$ generated by the Gibbs sampling is a Markov Chain with stationary distribution, see Gelfand and Smith (1990), Rao (2003).

Alternatively, Rao-Blackwell estimators of the posterior mean and the posterior variance of μ_i may be used to obtain:

$$\mu_i^{\mathrm{HB}} = \frac{1}{D}\sum_{k=d+1}^{d+D} \tilde{\mu}_i^{\mathrm{HB}}(\sigma_v^{2(k)}, \sigma_e^{2(k)}) = \tilde{\mu}_i^{(\mathrm{HB})}(\cdot,\cdot), \tag{3.30}$$

and

$$\begin{aligned} V(\mu_i|y) = &\frac{1}{D}\sum_{k=d+1}^{d+D} [g_{1i}(\sigma_v^{2(k)}, \sigma_e^{2(k)}) + g_{2i}(\sigma_v^{2(k)}, \sigma_e^{2(k)})] \\ &+ \frac{1}{D-1}\sum_{k=d+1}^{d+D} [\tilde{\mu}_i^{\mathrm{HB}}(\sigma_v^{2(k)}, \sigma_e^{2(k)}) - \tilde{\mu}_i^{\mathrm{HB}}(\cdot,\cdot)]^2. \end{aligned} \tag{3.31}$$

3.6 Hierarchical Bayes Method with Probit Model

The probit model is commonly used when the dependent variable is a qualitative one indicating an outcome in one of two categories. It is usually motivated as arising when an individual is making a choice (Koop, 2003).

The individual i gets some utility from alternative 0 and another level of utility from alternative 1. The utility depends on a variety of individual specific characteristics, hence, the key result is the difference between the individuals utility from alternative 1 and utility from alternative 0.

Let $y^* = (y_1^*, ..., y_N^*)^T$ denote the dependent variable. The model is written as

$$y_i^* = x_i^T\beta + e_i \tag{3.32}$$

where $x_i = (1, x_{i2}, ..., x_{ik})^T$. The difference in utility y_i^* depends on observed individual characteristics x_i, a vector of unobserved parameters β, and the random error component e_i. However, y_i^* cannot be observed. Instead, we observe the individual's choice y_i, set equal to one if alternative 1 is chosen and 0 if alternative 0 is picked.

For the probit model, the relationship between y and y^* takes the form

$$y_i = \begin{cases} 1 & \text{if}\, y_i^* > 0 \\ \\ 0 & \text{if}\, y_i^* \le 0 \end{cases} . \tag{3.33}$$

Choosing a distribution for the error e_i completes the basic model.

To complete the statistical model, specify the prior of β as $\beta \sim N(\beta_p, \Sigma_p)$, where β_p is a $k \times 1$ vector and Σ_p is the $k \times k$ covariance matrix. Let $H_p = \Sigma_p^{-1}$ denote the prior precision. The conditional density of β given y^* is:

$$\beta | y^*, h \sim N(\bar{\beta}, \bar{H}^{-1}) \tag{3.34}$$

where $\bar{H} = H_p + h(\mathbf{x}'\mathbf{x})$, $\bar{\beta} = \bar{H}^{-1}(H_p\beta_p + h\mathbf{x}'y^*) = \bar{H}^{-1}(H_p\beta_p + h(\mathbf{x}'\mathbf{x})\hat{\beta})$, and $\hat{\beta} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'y^*$.

The conditional density for y^* is a truncated normal where the truncation depends on y:

$$y_i^* | y_i, \beta, h \sim \begin{cases} TN_{(0,\infty)}(x_i'\beta, h^{-1}) & \text{if } y_i = 1 \\ \\ TN_{(-\infty,0)}(x_i'\beta, h^{-1}) & \text{if}\, y_i = 0 \end{cases} \tag{3.35}$$

where $TN_{(a,b)}$ denotes a normal distribution truncated to lie in the region (a, b).

In addition to parameter estimates, it is useful to present information about the choice probabilities. These can be derived from the posterior of the parameters by noting that, for any particular values of the parameters,

$$\begin{aligned} Pr(y_i = 1 | \beta, h) &= Pr(y^* > 0 | \beta, h) \\ &= Pr(x_i'\beta + e_i > 0 | \beta, h) \\ &= Pr(\sqrt{h} e_i > -\sqrt{h} x_i'\beta | \beta, h). \end{aligned} \tag{3.36}$$

Since the errors are assumed to be normally distributed, the last term in (3.36) is simply one minus the cumulative distribution function of the standard normal (i.e. $\sqrt{h}e_i$ is $N(0,1)$). If we define $\Phi(a)$ as the cumulative distribution function of the standard normal distribution, then the probability of choosing alternative 1 is $1 - \Phi(-\sqrt{h}x_i'\beta)$.

Furthermore, equation (3.36) illustrates an identification problem which is said to occur if multiple values for the model parameters give rise to the same value for the likelihood function. In the probit model, there are an infinite number of values for β and h which yields exactly the same model. The standard solution is to set $h = 1$.

After the adjustment of the latent variable, we apply the Gibbs sampling and MCMC method described in section 3.4.

Chapter 4. Application: Insurance Coverage for Louisiana Parishes

4.1 Introduction

The issue of providing insurance coverage to children and adults has long been a topic of interest to U.S. policymakers. Estimates of the number of uninsured persons will be a key ingredient in measuring the effectiveness of the Affordable Care Act (ACA), particularly if some states deviate from others with regard to the some parts of the legislation such as the expansion of Medicaid eligibility[1].

The Congressional Budget Office (2011) has projected that the implementation of health insurance reforms in the Affordable Care Act (ACA) will reduce the number of uninsured Americans by 33 million in 2020, from 56 to 23 million people. Although this still falls short of universal coverage, the number of uninsured people will be reduced by more than half. Most of the coverage gains will come from expending Medicaid to everyone below 133% of the poverty line (138% with income disregards) and from creating health insurance exchanges. Principally, knowing about who will remain uninsured will assist safety net providers and programs, organizations, and support systems to determine further needs for uninsured access and also the optimal structures for achieving those needs. Beginning in 2014, most Americans will be required to have health insurance coverage meeting certain minimum requirements and will be subject to financial penalties if they do not comply. For those people who cannot afford insurance, or some other specialized circumstances, such as people who are Native Americans, prisoners or have religious objections, exemptions will be granted. Medicaid eligibility will expand greatly for adults in many states; however, only

[1] The Supreme Court ruling on the Affordable Care Act allowed states to opt out of the law's Medicaid expansion. Until January 27, 2015, there were 28 states that accepted Medicaid expansion; however, the state of Louisiana is not expanding Medicaid at this time.

small or zero increases will be seen for children. Due to the Children's Health Insurance Program (CHIP), children's eligibility levels for public coverage are already much higher than for adults (Buettgens et al., 2011).

In this section, we estimate the uninsured rates for children and adults by using three methodologies in small area estimation as mentioned in the previous sections. The purpose of small area estimation is obtaining reliable estimates from subpopulations (such as district, county, state, sex, race, sex-race combination, etc.) when the data has few observations in some of the subpopulation (Datta and Ghosh, 1991; Datta et al., 1996, 2000, 2002; Rao 2003). Starting from direct estimates obtained from survey data, we describe a range of Bayesian hierarchical models that incorporate different types of random effects and show that these give improved estimates. Although implementation of complex Bayesian models requires computationally intensive Markov Chain Monte Carlo simulation algorithms (Gilks et al., 1995), there are still a number of potential benefits of the Bayesian approach for small area estimation. The Bayesian approach can handle different types of target variables (such as continuous, dichotomous, categorical), different random effects structures (such as independent, spatially correlated), areas with no direct survey information, models to smooth the survey sample variance estimates and so on (Gomez-Rubio et al., 2008).

4.2 Data

The Louisiana Health Insurance Survey (LHIS), which starts from 2003, is a series of surveys designed to provide the most accurate and comprehensive assessment of Louisiana's uninsured populations every two years. Each round of the LHIS has been based on more than 10,000 Louisiana households (roughly 27,000 Louisiana residents), which allows researchers to estimate the uninsured populations for each parish[2], the Department of Health and Hospitals' nine regions, and also for specific subpopulations (e.g. children under 200% of federal

[2]The state of Louisiana is divided into 64 parishes (French: paroisses) in the same way the 48 other states of the United States are divided into counties.

poverty). Each round of the LHIS has also incorporated methodological improvements to ensure that the survey results reflect the best understanding of how to estimate uninsured populations. For example, the 2007 LHIS incorporated an innovative methodological tool to adjust uninsured estimates for the Medicaid undercount at the individual level. The 2009 LHIS incorporated uninsured estimates from a cell phone sample to improve coverage of cell-only households. This improvement helps researchers to estimate the uninsured rate more accurately, since national surveys estimate there are 31.6% households that are cell-only. The prior research also indicates that cell-only households are more likely to be younger, poorer, ethnic minorities, and uninsured. Therefore, the 2011 LHIS expands coverage of cell-only households by increasing the cell phone sample from 500 to 2,000 completed interviews, which is an improvement.

The LHIS survey gauges uninsured status through a household-level approach in which individual respondents are asked to report on the health insurance status of each member of the household. To assure reporting is as accurate as possible, initial respondents are screened to make sure they are the most knowledgeable person in the household about family health care and health insurance. Once the most knowledgeable person in the household has been selected, respondents are asked to identify all members of the household covered by particular types of insurance including employer sponsored insurance, privately purchased insurance, Medicaid or LaCHIP, Medicare, or military insurance. Respondents are asked to verify uninsured status for any individual in the household not identified as having some form of insurance coverage. Only household members who are identified as not having any form of insurance coverage and who are verified as uninsured are included in the final estimate of the uninsured population. Moreover, the probability of being selected into the final sample was dependent on the parish in which the respondent resided. To account for this, the results were weighted to adjust for sampling differences across parishes. Specifically, the sampling weight was constructed as the parish population according to the 2010 Census divided by the number of individuals sampled in the parish. Because differences in response rates among different segments of the population may also result in biased estimates of uninsured rates, the data were weighted to match demographic characteristics as estimated by the most recently available U.S. census data.

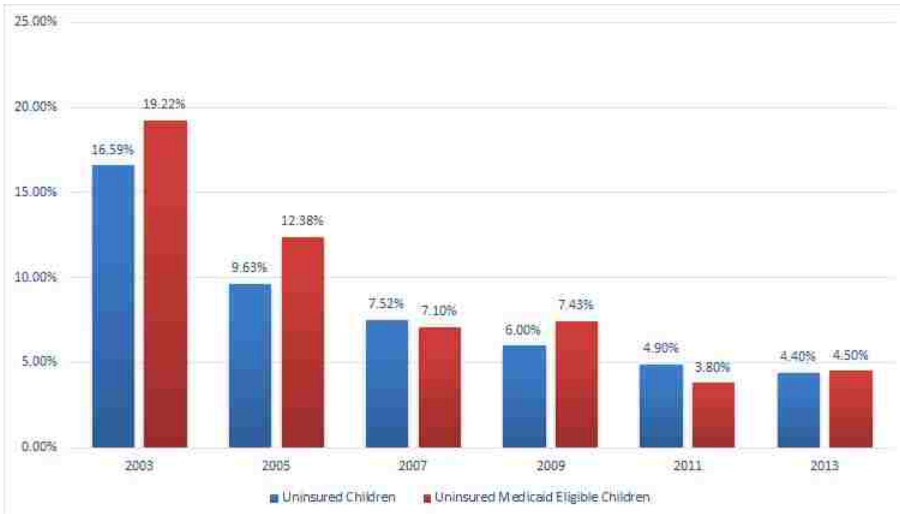

Figure 4.1: Uninsured Children 2003 - 2011

Figure 4.1 shows the percent of uninsured children in the state of Louisiana over the past 10 years. Both the uninsured children and uninsured medicaid eligible children have been declining over the years. The uninsured rates for children who are eligible for Medicaid are slightly higher than over all uninsured rates in each survey year except year 2007. From 2003-2011, the percent of uninsured children declined from 11.1% to 3.5% translating into 101,162 fewer uninsured children in Louisiana[3]. There is a similar decline pattern in the percent of uninsured Medicaid eligible children from 12.9% in 2003 to 3.8% in 2011. Recently, the percent of uninsured children decreased slightly to 4.4%, 5,542 fewer children uninsured than in 2011. For children who are eligible for Medicaid, the uninsured rate increased to 4.5%, an overall increase of 6,592 children since 2011.

Figure 4.2 is the map for the Department of Health & Hospitals (DHH) regions in Louisiana. The state of Louisiana is divided into 9 DHH regions geographically. In terms of

[3]The calculation is based on the estimates of population, which is provided by United States Census Bureau. Source: Louisiana's Uninsured Population: A Report from the 2013 Louisiana Health Insurance Survey (Barnes et al., 2013).

Figure 4.2: Department of Health & Hospitals (DHH) regions in Louisiana

median household income, the wealth levels are similar within each regions. For instance, East Carroll Parish in the Northeast Region (Region 8) has the lowest median household income, roughly \$25,321, while the median household income is \$44,874 over the entire state[4]. On the other hand, St. Tammany Parish in the Northshore Region (Region 9) has the highest median household income at \$60,799. Meanwhile, East Baton Rouge Parish, which is the location of the state capital, has a median household income as \$48,506.

Figure 4.3 lists the regional variation in uninsured rates for children over the past five years. Comparing survey year 2009 and 2011, the uninsured rates went down in every region with the notable exception of the Northshore region (Region 9), a region that already had the lowest uninsured rate in the state, where the shift in uninsured rates was barely perceptible. This slightly increasing uninsured rate may lead to a small decrease in our estimate of the number of uninsured children. In other regions, uninsured rates declined

[4]Source: United States Census Bureau, the median household income from 2009 to 2013.

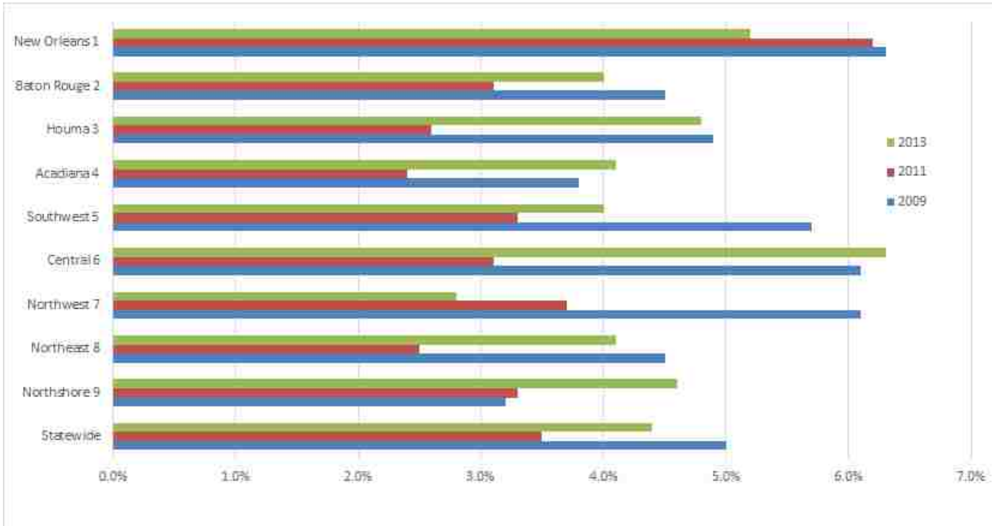

Figure 4.3: Regional Variation in Uninsured Rates for Children

over years, especially in the Central region (Region 6) where the uninsured rate for children dropped from 6.1% in 2009 to 3.1% in 2011. The smallest decline occurred in the Baton Rouge region (Region 2) where the uninsured rates dropped from 4.5% in 2009 to 3.1% in 2011. After that, the uninsured rates for children increased to the level of 2009. Overall, in the past five years, the uninsured rates have slightly decreased. The statewide uninsured rate decreased about 0.5%. In terms of uninsured population, the number of uninsured children has decreased by 5,924. As we mentioned earlier, Children's Health Insurance Program (CHIP) provides health insurance to families with incomes that are modest but too high to qualify for Medicaid. Due to the expansion of the CHIP and Medicaid, the uninsured rate for children has dramatically decreased since 2003, and stays fairly stable after 2009.

Figure 4.4 shows the percent of uninsured non-elderly adults (19-64) in Louisiana over the past decade. Different from children, the number of uninsured adults has no clear decreasing pattern. In the recent surveys, the health insurance coverage for adults is slightly decreased by 1.6% (in particularly, the uninsured adults decrease from 633,943 to 622,033 since 2011). The trend among uninsured adults under 200% of federal poverty is quite similar. The proportion of uninsured adults under 200% of the federal poverty level is around one third

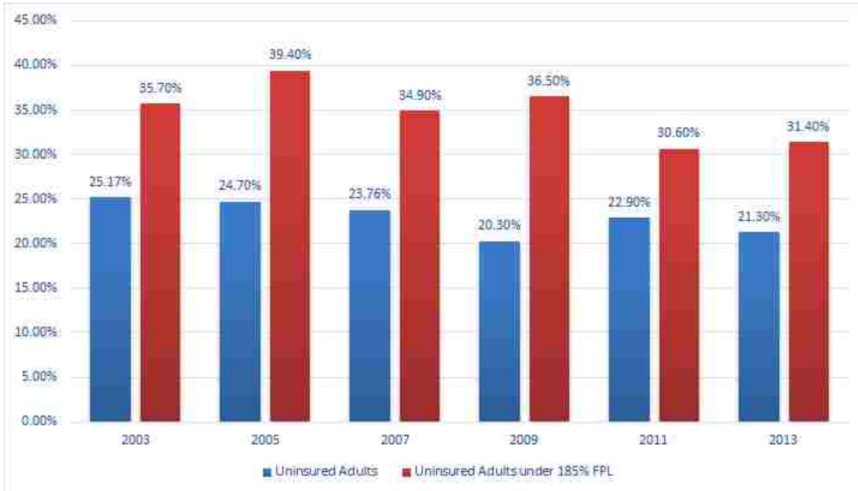

Figure 4.4: Uninsured non-elderly adults 2003 - 2011

over years. Because adults do not have the same social safety net as children, their uninsured rates show a much stronger relationship to economic conditions in Louisiana, which include a lower unemployment rate in 2013 than in 2011.

The health insurance status is highly correlated with personal characteristics, such as race, income, poverty, education, age, etc. For instance, the health insurance coverage is relatively low for African American, poorer, less educated and younger adults. In the following, we will present the differences in health insurance coverage status across gender, race, income, poverty, age, and education. First of all, there are only minor differences in insurance status depending on gender, while male adults and female children are slightly more likely to be uninsured. However, the gender-based differences in health insurance coverage are small for both adults and children. Next, we discuss the different behaviors of health insurance coverage by race. About 32% of African-American non-elderly adults are uninsured compared to 17.7% of Caucasians. Notably, uninsured rates for African-Americans have increased more dramatically than for Caucasians. In 2013, 29.6% of African-Americans were uninsured compared to 16.8% of Caucasians. The differences are small among children: 4.0% of African American children and 3.0% of Caucasian children are uninsured. Uninsured rates for African American and Caucasian children have steadily declined since 2005 when 7.9%

of African American children and 6.4% of Caucasian children were reported as uninsured. In 2013, 5.2% of African American and 4.0% of Caucasian children had no health insurance coverage. The uninsured patterns are similar for children when we consider the different level of income and poverty. It turns out that the level of income and poverty do not affect children as much as adults. As we mentioned earlier, this may be caused by the availability of Medicaid/LaCHIP programs.

For adults, income is also an important predictor of uninsured status either when measured as household income or in relation to federal poverty guidelines. But there are less clear relationships between income and insurance status for children. From the LHIS survey, the highest uninsured rates for children occur in income ranges between $65,000 and $74,999, which is high enough to be ineligible for public assistance but perhaps still low enough that budget constrained families are less likely to purchase insurance. About 6% of children in this income range are uninsured. For adults, being uninsured is strongly related to income. Forty-six percent of adults earning between $10,000 and $14,999 are uninsured compared to 6.3% of adults earning $95,000 or more. When we examine the uninsured status relative to federal poverty, it shows a similar pattern, which accounts for family size in determining the sufficiency of available financial resources. For adults, being uninsured is strongly correlated with poverty. For example, nearly 47.2% of adults between 50-100% of FPL are uninsured. For children, the greatest risk for being uninsured is to fall outside the range of Medicaid eligibility: 5.8% of children between 200-300% of federal poverty are uninsured compared to just 1.7% between 150-200% of FPL.

Besides the income and poverty level, education is also strongly associated with uninsured rates for adults, such that less educated respondents are considerably more likely to be uninsured. There exists a steady decline in uninsured rates with education increases. Forty-four percent of respondents with less than a high school education were uninsured, 29.1% with a high school education, 19.6% with some college, while the uninsured rates are 10.0% and 7.3% for those with a college degree and with a graduate degree, respectively.

Furthermore, age is also associated with uninsured status as young children are least likely to be uninsured. Younger adults (19-29) are most likely to be uninsured. Overall, uninsured rates for adults decrease as age increases. Relative to 2011, uninsured rates have increased significantly for adults age 30 and older while they have remained relatively stable for younger adults (19-29). This stability may reflect the impact of the Patient Protection and Affordable Care Act as younger adults are able to remain covered through a parent's employer-sponsored insurance policy through age 26. For children, because of Medicaid/LaCHIP programs, young children (0-5) are least likely to be uninsured. Only 2.4% of children 0-5 have no health insurance coverage as are 3.2% of children between 6-13 and 4.6% of children between 14-18.

4.3 Variable Definitions

As we mentioned earlier, the health insurance coverage is highly related to several personal characteristics. Most of the variables are categorical ones. The primary variables of our interest are listed below:

1. Black. A binary variable indicating the race of the adults/children. If the race is Africa-American, "Black" takes the value of 1. Otherwise, "Black" takes the value of 0.

2. Female. A binary variable indicating the gender of the adults/children. If the person is female, "Female" takes the value of 1. Otherwise, "Female" takes the value of 0.

3. Working percent. A continuous variable ranging between 0 and 1, indicates the percent of working age adults in the family who are employed.

4. Income. A continuous variable indicating the household income[5].

[5] In order to match other relatively small values variables, we adjust the income variable as the Household income/10,000.

5. Poverty. A binary variable indicating the poverty level of adults'/children's family. If the adult/child lives in a family below 185% of the federal poverty line, "Poverty" takes the value of 1. Otherwise, "Poverty" takes the value of 0.

6. Age. A numerical variable which reported the age of the adults being interviewed in the survey.

7. Age Group(i). $i = 1, 2, 3$. A binary variable indicating the age group for children. If the child's age is between 5 and 9, "Age group (1)" takes the value of 1. If the child's age is between 10 and 14, "Age group (2)" takes the value of 1. If the child's age is between 14 and 18, "Age group (3)" takes the value of 1. Otherwise, "Age group (i)" takes the value of 0.

8. Parish. A factor identifying the parish which the resident belongs to. The 64 parishes start from "Acadia Parish", "Allen Parish", ..., "East Baton Rouge Parish",...,"Winn Parish".

4.4 Model Setup

In this section, we specify three estimators of uninsured rate for 64 parishes in the state of Louisiana, which are the best linear unbiased prediction estimators (EBLUP), hierarchical Bayes estimators (HB), and hierarchical Bayes method with probit model, as we described in Chapter 3. We use a nested error linear regression model with cross sectional data at parish level. The model develop is based on the basic unit level nested error regression model by Battese et al. (1988) and extensions by Prasad and Rao (1999) and You and Rao (2003).

Suppose that the $i-th$ parish or small area population size N_i for $i = 1, 2, ..., 64$ is known to us. We use μ_i to denote the percentage of health insurance coverage of adults/children for parish i. We are interested in estimating μ_i for each parish.

The top parts of Table 4.1 and Table 4.2 list the sample summary statistics of 64 parishes

Table 4.1: Summary statistics by parish for adults 2013

Survey Sample Variable	Obs	Mean	Std. Dev.	Min	Max
% Black	64	0.2652	0.1503	0.0000	0.6706
Household income	64	66036	18678	24484	106522
% Female	64	0.5497	0.0315	0.4902	0.6706
% P185	64	0.3998	0.1106	0.1845	0.6404
Working Percent	64	0.6056	0.0686	0.3820	0.7255
Age	64	46.15	1.97	42.72	50.84
n_i	64	230	170	85	896
Population Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
% Black	64	0.3207	0.1454	0.0150	0.6930
Household income	64	41196	9062	25267	66173
% Female	64	0.4918	0.0406	0.2870	0.5240
% P185	64	0.7098	0.1245	0.4610	0.9500
Working Percent	64	0.6571	0.0849	0.4120	0.8070
Age	64	44.07	1.09	41.80	47.00
N_i	64	45165	62199	2809	290720

in Louisiana for adults and children in the survey year 2013, respectively, while the bottom parts list the population summary statistics which come from the sources U.S. Department of Labor: Bureau of Labor Statistics and U.S. Census Bureau. The number of observations n_i in each parish range between 85 (in East Carroll Parish) to 896 (in East Baton Rouge Parish) for adults, and range from 21 (in East Carroll Parish) to 290 (in East Baton Rouge Parish) for children.

Comparing the top and bottom parts in Table 4.1, there are some variations between survey sample means $\bar{x}$ and population means $\bar{X}$. For instance, the average of black adults rates is 26.52% in the survey sample, while the rate is roughly 32% from the U.S. Census Bureau. As we have seen from the bottom part of Table 1, the rates of black adults range from 1.5% (Cameron Parish) to 69.3% (East Carroll Parish) within 64 parishes, while the rates of black adults from the survey sample range from 0 (Cameron Parish) to 67.06%

Table 4.2: Summary statistics by parish for children 2013

Survey Sample Variable	Obs	Mean	Std. Dev.	Min	Max
% Black	64	0.3172	0.1851	0.0000	0.7667
Household income	64	67001	20792	19014	106761
% Age 5-8	64	0.2637	0.0576	0.0714	0.4054
% Age 9-13	64	0.2904	0.0687	0.1053	0.5000
% Age 14-18	64	0.2595	0.0490	0.1515	0.3667
% Female	64	0.4848	0.0629	0.3030	0.6271
Poverty	64	0.4729	0.1493	0.2093	0.8571
Working percent	64	0.6496	0.0861	0.4405	0.8426
n_i	64	84	63	21	290
Population Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
% Black	64	0.3662	0.1630	0.0330	0.7380
Household income	64	41196	9062	25267	66173
% Age 5-8	64	0.2690	0.0112	0.2190	0.2940
% Age 9-13	64	0.2676	0.0137	0.2160	0.2890
% Age 14-18	64	0.2046	0.0212	0.1750	0.3340
% Female	64	0.4868	0.0072	0.4600	0.4980
Poverty	64	0.6357	0.1015	0.3910	0.9500
Working percent	64	0.6571	0.0849	0.4120	0.8070
N_i	64	18515	23824	1281	113177

(East Carroll Parish). For the variable household income, the average of survey sample means is about 66,036 over all parishes, while the average of household income from the U.S. Department of labor is only 41,196. The summary statistics are based on the means of each parish, that may have caused the variations. The details of sample means for each parish are available in the Appendix.

Table 4.3 provides the estimates of β in equation (2.2) for three methodologies, as well as the marginal effects for the hierarchical Bayes method with probit model. The first column is the estimates from ordinary least squares regression. The estimates show that holding other variables constant, the African-American people are more likely to be uninsured by 5.4%. Notably, uninsured rates for African-Americans have increased more dramatically than for Caucasians. In 2009, 27.6% of African-Americans were uninsured compared to 15.8% of Caucasians. The uninsured rates for African-Americans is 29% compared to 18.2% of Caucasians in survey year 2013. The gender based differences in uninsured rates are relatively small: 0.3% lower for female adults. Considering the percentage of working adults, the higher the portion of working adults, the lower the probability of uninsured. Overall, uninsured rates for adults decrease as age increases. Relative to 2011, uninsured rates have increased significantly for adults at age 30 and older while they have remained relatively stable for younger adults (19-29). This stability may reflect the impact of the Patient Protection and Affordable Care Act as younger adults are able to remain covered through a parent's employer-sponsored insurance policy through age 26. From the estimates, with a one-year increase of adult's age, the probability of being uninsured decreases by 0.34%. For adults, income is also an important predictor of uninsured status either when measured as household income or in relation to federal poverty guidelines. For instance, 46% of adults earning between \$10,000 and \$14,999 are uninsured compared to 6.3% of adults earning \$95,000 or more. When we examine the uninsured status relative to federal poverty, it shows a similar pattern, which accounts for family size in determining the sufficiency of available financial resources. For adults, being uninsured is strongly correlated with poverty. For instance, nearly 47.2% of adults between 50-100% of FPL are uninsured.

Table 4.3: Estimates of β for adults

Adults 2013	OLS Estimate	HB Estimate	HB_Probit	Marginal Effect
Constant	0.2771***	0.2768***	-0.5311	-0.1688
	(0.0169)	(0.0166)	(0.0616)	
Black	0.0538***	0.0537***	0.1811	0.0567
	(0.0083)	(0.0077)	(0.0289)	
Family income	-0.0025***	-0.0024	-0.0192	-0.0061
	(0.0004)	(0.0036)	(0.0023)	
Female	-0.0034	-0.0035	-0.0153	-0.004
	(0.0061)	(0.0063)	(0.0256)	
Poverty	0.2188***	0.2188***	0.7454	0.236
	(0.0094)	(0.0069)	(0.0338)	
Working Percent	-0.0060	-0.0061	-0.028	-0.0089
	(0.0105)	(0.0069)	(0.0395)	
Age	-0.0034***	-0.0034***	-0.0137	-0.0044
	(0.0002)	(0.0001)	(0.0009)	

Note: Numbers in parentheses are standard errors and posterior standard deviations.

* Statistically significantly different from zero at the 10% level.

** Statistically significantly different from zero at the 5% level.

*** Statistically significantly different from zero at the 1% level.

The estimates in Table 4.3 shows that, if the adult is living in a family below 185% of the federal poverty line, he or she is more likely to be uninsured by 21.9%.

Table 4.4: Estimates of β for children

Children 2013	OLS Estimate	HB Estimate	HB_Probit	Marginal Effect
Constant	0.0445***	0.0445***	-1.6474	-0.1697
	(0.0107)	(0.0149)	(0.1333)	
Black	0.0104	0.0104	0.1447	0.0149
	(0.0072)	(0.0073)	(0.0762)	
Family income	-0.0008**	-0.0008	-0.0306	-0.032
	(0.0004)	(0.0053)	(0.0091)	
Age 5-8	0.0024	0.0024	0.0341	0.0035
	(0.0079)	(0.0098)	(0.0988)	
Age 9-13	0.0104	0.0103	0.1177	0.0121
	(0.0080)	(0.0071)	(0.0966)	
Age 14-18	0.0162*	0.0162**	0.1529	0.0157
	(0.0085)	(0.0066)	(0.0984)	
Female	-0.0055	-0.0054	-0.0483	-0.0050
	(0.0056)	(0.0061)	(0.0621)	
Poverty	0.0194**	0.0194***	0.0653	0.0067
	(0.0081)	(0.0050)	(0.0980)	
Working Percent	-0.0155	-0.0154***	-0.1146	-0.0118
	(0.0094)	(0.0037)	(0.0936)	

Note: Numbers in parentheses are standard errors and posterior standard deviations.

* Statistically significantly different from zero at the 10% level.

** Statistically significantly different from zero at the 5% level.

*** Statistically significantly different from zero at the 1% level.

The pattern is quite similar for children. Uninsured rates for African American and Caucasian children have steadily declined since 2005 when 7.9% of African American children and 6.4% of Caucasian children were reported as uninsured. In 2013, 5.0% of African-American and 3.6% of Caucasian children were uninsured. The African-American child is more likely to be uninsured by 1%, while a female child is less likely to be uninsured, although both of them are insignificant. Different from adults, the uninsured probabilities have no trend over age categories for children. Furthermore, the more adults working in the household, the less likely to have a child uninsured.

We could observe similar patterns of uninsured rates for children as adults when we consider the different level of income and poverty. It turns out that the level of income and poverty do not affect children as much as adults. It may because of the availability of Medicaid/LaCHIP programs[6]. Again, the household income has a negative impact on a child's uninsured probabilities. From the LHIS survey, the highest uninsured rates for children occur in income ranges between \$65,000 and \$74,999, which is high enough to be ineligible for public assistance but perhaps still low enough that budget constrained families are less likely to purchase insurance. About 6% of children in this income range are uninsured.

The second columns in Tables 4.3 and 4.4 list the estimate of β for hierarchial Bayes methods for adults and children, respectively. The third columns are the estimates of hierarchical Bayes method with probit model. Different from the linear regression model, coefficients of a probit model rarely have any direct interpretation. In our study, the coefficient signs of hierarchical Bayes methods with a probit model are the same as the OLS estimates and the hierarchical Bayes method.

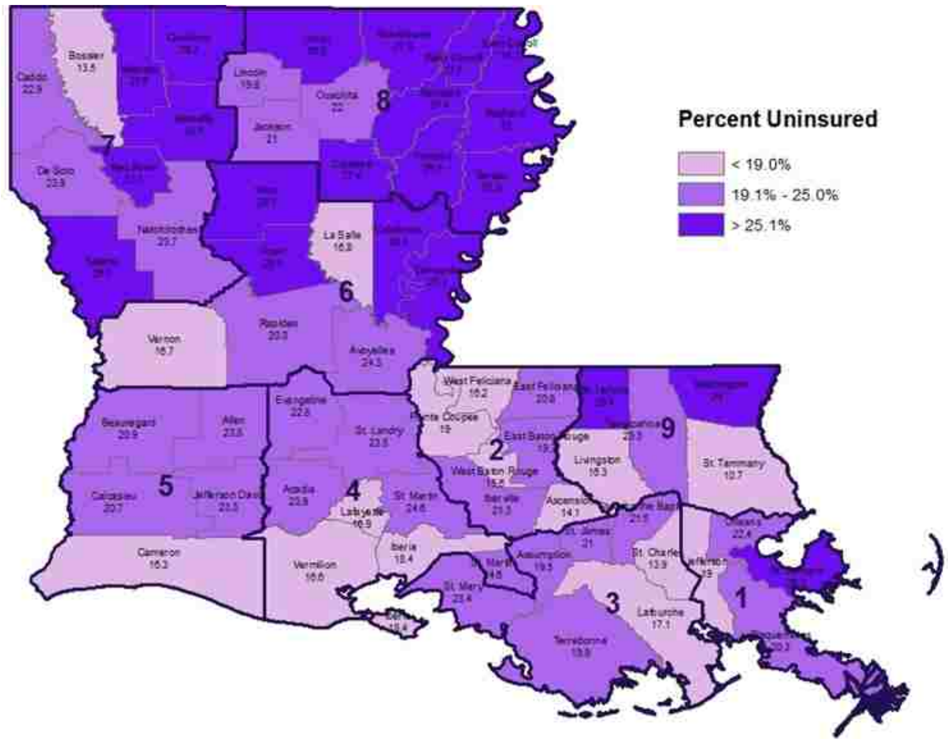

Figure 4.5: Uninsured Rates for Adults age 19-65 (2013)

4.5 Results

Figure 4.5 shows a map of the estimated uninsured rates for adults in survey year 2013. The uninsured rates rage from an estimated 14.5% (as in Cameron Parish) to an estimated 36.0% (as in West Carroll Parish). Figure 4.6 shows the mean household income by parish. These two maps show a similar pattern. For instance, Figure 4.5 shows higher uninsured rates in the northeastern DHH region, high poverty parishes of Louisiana, and lower uninsured rates throughout the wealthier I-10 corridor. In terms of parish, unsurprisingly, St. Tammany Parish, the parish with the highest household income, has the lowest estimated uninsured rate for adults over the entire state. On the other hand, Madison Parish, the parish with the lowest household income, has one of the highest uninsured rate for adults.

[6] The Louisiana ChildrenŠs Health Insurance Program (LaCHIP) provides health coverage to uninsured children up to age 19. It is a no-cost health program that pays for hospital care, doctor visits, prescription drugs, shots and more.

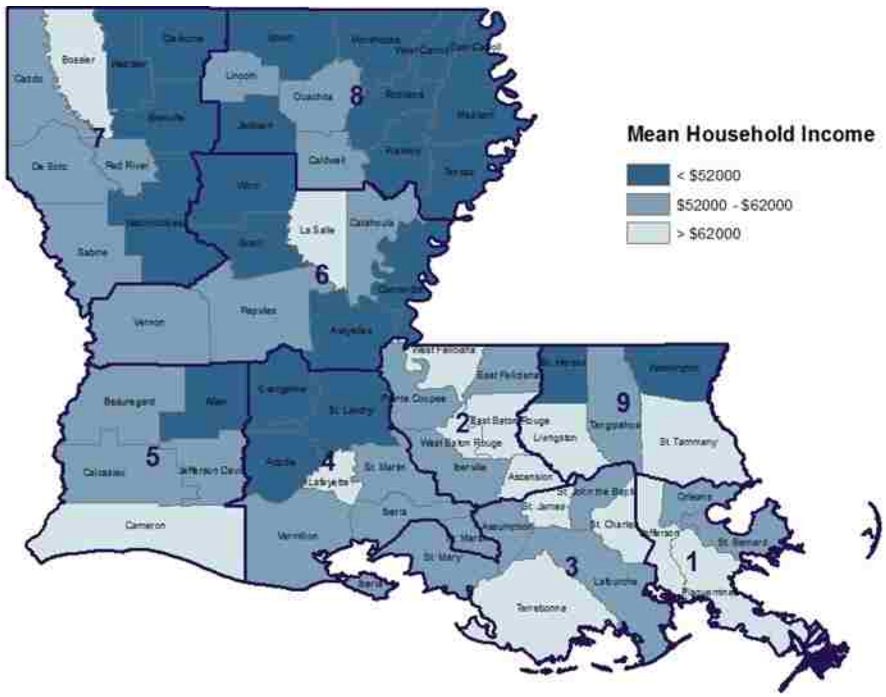

Figure 4.6: Mean Household Income by Parish (2013)

The estimated uninsured rates for children have a different pattern than the adults. Figure 4.7 shows the uninsured rates for children under 19 by parish in survey year 2013. It shows that there is no systematic geographic pattern of high uninsured rates among children. As mentioned earlier, while poverty tends to be present in geographic clusters, Medicaid and LaCHIP enrollments offset the pattern of low employer and private insurance coverage in poor parishes.

Table 4.5 and Table 4.6 list the three estimates as well as the mean square error (posterior standard deviation) of uninsured adults and children for 64 parishes in Louisiana, respectively. We also convert the estimated uninsured rate into the estimates of the number of uninsured for adults and children based on the information of the population[7]. Consider the EBLUP estimates, the probability of adults' uninsured rate ranges from 12.79% (as in Cameron Parish) to 44.15% (as in West Carroll Parish). In terms of an uninsured person, the range is between 601 (as in Cameron Parish) to 82,996 (as in East Baton Rouge Parish).

[7]We calculate the uninsured adults and children based on the estimates of uninsured rates from Hierarchical Bayes method with probit Model.

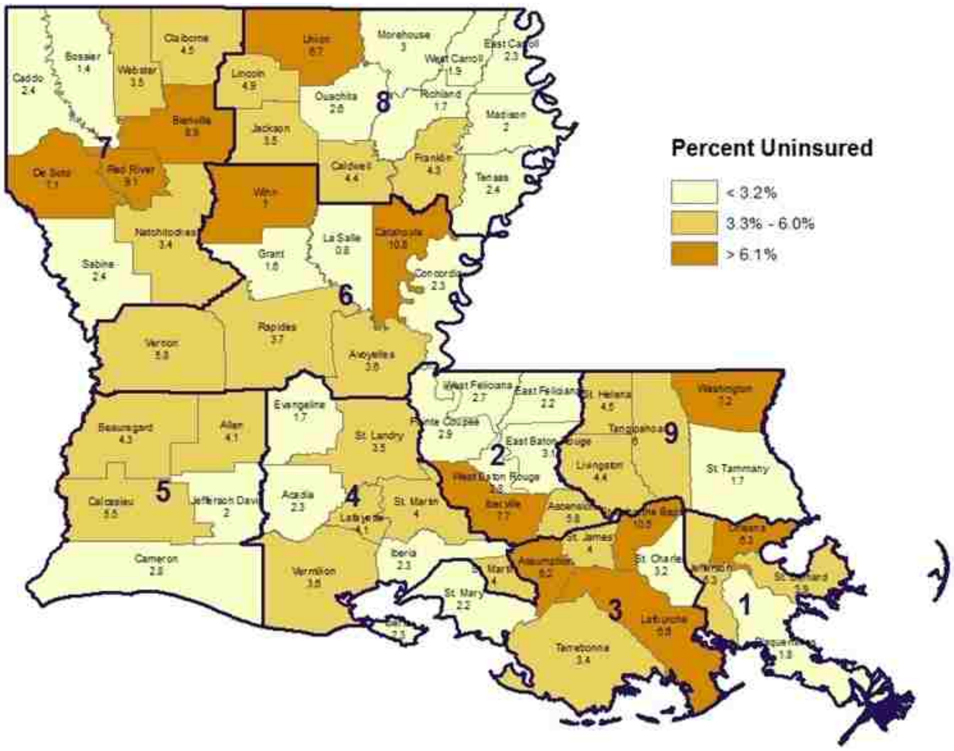

Figure 4.7: Uninsured Rates for Children under Age 19 (2013)

Comparing the estimates between the hierarchical Bayes estimates and hierarchical Bayes method with probit model, the uninsured rates are quite similar for adults. The uninsured rate from the hierarchical Bayes method with probit model ranges from 14.51% to 36.01%, which is narrower than the EBLUP estimates. The mean square error (posterior standard deviation) is a measure of the desirability of efficiency. For adults, the mean square error (posterior standard deviation) is of a similar magnitude for each parish among the three methodologies.

As can be seen in the first column in Table 4.6, the EBLUP method provides some negative estimates, which is the disadvantage of the EBLUP method. For instance, the EBLUP method estimates the uninsured rate for children in East Carroll Parish is -1.07%, while the estimates are -0.25% in LaSalle Parish, and -0.3% in Madison Parish. The negative uninsured rates are unrealistic. This is the reason that we explore a better method to estimate the uninsured probabilities for children. The second and third column of Table 8 provides the hierarchical Bayes estimates and hierarchical Bayes method with probit model of uninsured rates for children. In terms of posterior standard deviation, the latter perform better than the regular hierarchical Bayes method. Therefore, we calculate the uninsured children based on the hierarchical Bayes method with probit model. Overall, in the survey year 2013, the estimates of uninsured rates for children range between 1.6% (as in LaSalle Parish) and 8.77% (as in Bienville Parish), in terms of uninsured children, the number of uninsured children rage from 58 (as in Cameron Parish and Tensas Parish) to 5,379 (as in Orleans Parish).

Next, we discuss the behavior of coefficients in the hierarchical Bayes method. As we discussed in Section 3.4, the Markov Chain Monte Carlo methods construct a stationary distribution. Hence, after a sufficiently large "burn-in," we can regard the small area parameters as dependent samples from the target distribution, regardless of the starting point. Figure 4.8 shows the β's distributions after 500 warmups. We observe that all the βs converge to a stable values after 5,000 iterations.

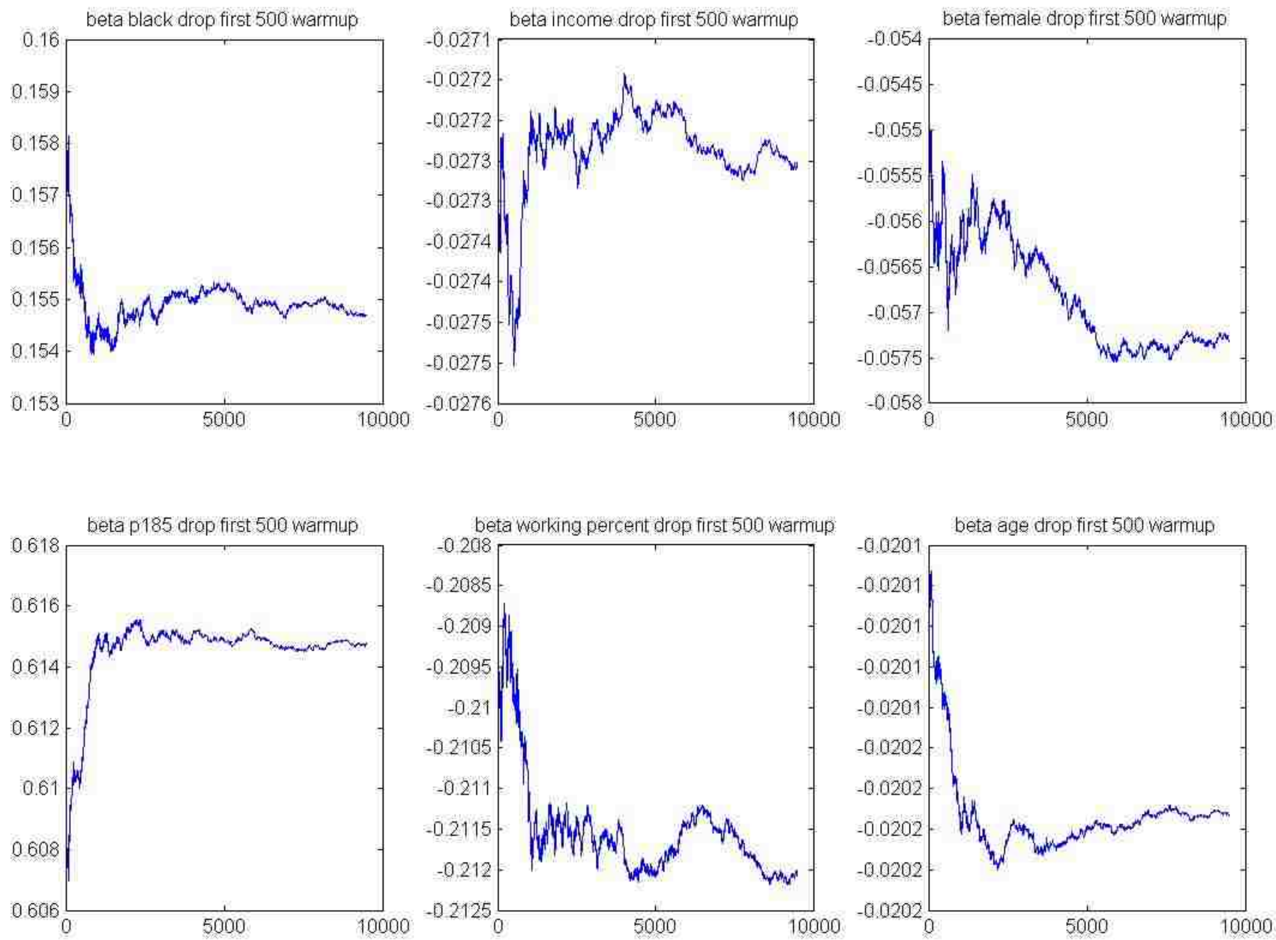

Figure 4.8: β's distribution after 500 warmups

Table 4.5: Estimates of Uninsured Adults in 2013

parish_name	Sample mean	EBLUP	MSE	HB	Posterior std.dev.	HB_Probit	Posterior std.dev.	Sample Size	19-64 population	Uninsured 19-64
Acadia	0.2430	0.3234	0.0254	0.3115	0.0235	0.2813	0.0263	214	38633	10866
Allen	0.1975	0.2704	0.0296	0.2738	0.0249	0.2494	0.0265	157	16473	4109
Ascension	0.1495	0.2222	0.0214	0.2227	0.0193	0.1889	0.0197	301	70066	13236
Assumption	0.1679	0.2207	0.0324	0.2433	0.0274	0.2247	0.0259	131	14276	3208
Avoyelles	0.1970	0.3040	0.0265	0.3046	0.0229	0.2877	0.0270	198	25354	7294
Beauregard	0.1856	0.2427	0.0266	0.2444	0.0230	0.2157	0.0232	194	22023	4750
Bienville	0.2458	0.2721	0.0342	0.2753	0.0272	0.2553	0.0277	118	8220	2099
Bossier	0.1590	0.2107	0.0199	0.2145	0.0199	0.1894	0.0185	346	76251	14440
Caddo	0.2285	0.2990	0.0161	0.2933	0.0168	0.2692	0.0186	534	158369	42639
Calcasieu	0.1799	0.2655	0.0151	0.2595	0.0166	0.2336	0.0169	617	120197	28079
Caldwell	0.2824	0.3324	0.0324	0.3092	0.0270	0.2768	0.0299	131	6152	1703
Cameron	0.0882	0.1279	0.0367	0.1647	0.0297	0.1451	0.0226	102	4142	601
Catahoula	0.1964	0.3191	0.0351	0.3134	0.0281	0.2936	0.0311	112	6402	1880
Claiborne	0.3060	0.3586	0.0320	0.3492	0.0268	0.3322	0.0306	134	10854	3606
Concordia	0.2047	0.2949	0.0329	0.3010	0.0272	0.2853	0.0293	127	12280	3503
DeSoto	0.2640	0.2721	0.0331	0.2758	0.0268	0.2516	0.0270	125	16259	4091
East Baton Rouge	0.1596	0.3124	0.0130	0.3061	0.0190	0.2855	0.0172	896	290720	82996
East Carroll	0.2353	0.3169	0.0403	0.3418	0.0307	0.3366	0.0341	85	4681	1576
East Feliciana	0.2544	0.3444	0.0348	0.3375	0.0281	0.3100	0.0314	114	13025	4038

Continued on next page

Table 4.5 – *Continued from previous page*

parish_name	Sample mean	EBLUP	MSE	HB	Posterior std.dev.	HB_Probit	Posterior std.dev.	Sample Size	19-64 population	Uninsured 19-64
Evangeline	0.2671	0.3461	0.0307	0.3313	0.0260	0.3087	0.0298	146	20091	6201
Franklin	0.2991	0.3931	0.0359	0.3652	0.0297	0.3308	0.0342	107	11843	3918
Grant	0.2701	0.3266	0.0317	0.3074	0.0267	0.2747	0.0294	137	14278	3922
Iberia	0.1803	0.2817	0.0245	0.2776	0.0220	0.2555	0.0245	233	44621	11402
Iberville	0.1511	0.3081	0.0317	0.3127	0.0260	0.2952	0.0300	139	21598	6376
Jackson	0.1926	0.2573	0.0319	0.2634	0.0261	0.2377	0.0268	135	9681	2301
Jefferson	0.1409	0.2623	0.0150	0.2584	0.0163	0.2259	0.0177	646	274951	62111
Jefferson Davis	0.1437	0.2220	0.0288	0.2278	0.0245	0.2043	0.0239	167	18545	3788
LaSalle	0.1469	0.2277	0.0175	0.2235	0.0184	0.1973	0.0177	463	9195	1814
Lafayette	0.1449	0.2484	0.0170	0.2430	0.0191	0.2112	0.0183	490	147813	31220
Lafourche	0.1766	0.2734	0.0196	0.2636	0.0199	0.2349	0.0206	368	60934	14312
Lincoln	0.1931	0.2623	0.0261	0.2677	0.0258	0.2398	0.0243	202	32163	7713
Livingston	0.2164	0.2805	0.0194	0.2700	0.0182	0.2388	0.0214	365	82068	19594
Madison	0.2913	0.3583	0.0365	0.3569	0.0289	0.3389	0.0325	103	7730	2620
Morehouse	0.2207	0.3301	0.0309	0.3296	0.0255	0.3128	0.0299	145	16315	5104
Natchitoches	0.2500	0.3235	0.0334	0.3157	0.0267	0.2954	0.0297	124	24450	7223
Orleans	0.2103	0.2795	0.0160	0.2830	0.0188	0.2644	0.0180	542	248136	65601
Ouachita	0.2048	0.2683	0.0174	0.2647	0.0188	0.2393	0.0189	454	95082	22755
Plaquemines	0.1443	0.2198	0.0377	0.2389	0.0291	0.2115	0.0274	97	14640	3097
Pointe Coupee	0.1885	0.3161	0.0338	0.3135	0.0274	0.2880	0.0306	122	13431	3869

Continued on next page

Table 4.5 – *Continued from previous page*

parish_name	Sample mean	EBLUP	MSE	HB	Posterior std.dev.	HB_Probit	Posterior std.dev.	Sample Size	19-64 population	Uninsured 19-64
Rapides	0.2179	0.3075	0.0169	0.2987	0.0183	0.2735	0.0199	491	79689	21796
Red River	0.2294	0.3458	0.0357	0.3389	0.0280	0.3191	0.0325	109	5300	1691
Richland	0.2358	0.3342	0.0335	0.3278	0.0272	0.3047	0.0309	123	12532	3819
Sabine	0.3404	0.3922	0.0312	0.3532	0.0280	0.3118	0.0336	141	14060	4384
St. Bernard	0.2613	0.3257	0.0353	0.3137	0.0284	0.2879	0.0302	111	26896	7745
St. Charles	0.1214	0.2127	0.0259	0.2202	0.0227	0.1906	0.0223	206	33347	6355
St. Helena	0.2247	0.3016	0.0394	0.3152	0.0305	0.3025	0.0322	89	6643	2010
St. James	0.1343	0.2114	0.0321	0.2386	0.0275	0.2188	0.0276	134	13272	2904
St. John Baptist	0.1838	0.3277	0.0320	0.3317	0.0268	0.3143	0.0315	136	28063	8820
St. Landry	0.1535	0.2713	0.0234	0.2755	0.0218	0.2553	0.0247	254	48942	12493
St. Martin	0.1545	0.2784	0.0252	0.2789	0.0227	0.2604	0.0252	220	32637	8499
St. Mary	0.1881	0.2732	0.0262	0.2811	0.0246	0.2664	0.0256	202	32916	8768
St. Tammany	0.1279	0.1964	0.0152	0.1962	0.0171	0.1649	0.0149	602	146545	24169
Tangipahoa	0.2553	0.3407	0.0205	0.3310	0.0190	0.3100	0.0239	329	77521	24028
Tensas	0.2900	0.3795	0.0371	0.3694	0.0293	0.3540	0.0333	100	2809	994
Terrebonne	0.1467	0.2595	0.0216	0.2563	0.0213	0.2267	0.0222	300	69821	15827
Union	0.1769	0.2837	0.0326	0.2859	0.0268	0.2621	0.0287	130	13384	3508
Vermilion	0.0905	0.1789	0.0263	0.1953	0.0239	0.1748	0.0227	199	35410	6191
Vernon	0.1538	0.2153	0.0236	0.2264	0.0211	0.2049	0.0215	247	34099	6986
Washington	0.2746	0.3469	0.0311	0.3443	0.0258	0.3304	0.0302	142	27955	9237

Continued on next page

Table 4.5 – *Continued from previous page*

parish_name	Sample mean	EBLUP	MSE	HB	Posterior std.dev.	HB_Probit	Posterior std.dev.	Sample Size	19-64 population	Uninsured 19-64
Webster	0.2609	0.3040	0.0292	0.2937	0.0247	0.2650	0.0268	161	24236	6423
West Baton Rouge	0.1268	0.2228	0.0312	0.2421	0.0260	0.2161	0.0268	142	15380	3324
West Carroll	0.3790	0.4415	0.0333	0.3910	0.0304	0.3601	0.0368	124	6711	2417
West Feliciana	0.1649	0.2325	0.0377	0.2548	0.0295	0.2233	0.0281	97	11076	2474
Winn	0.2391	0.3118	0.0316	0.3026	0.0261	0.2787	0.0284	138	9390	2617

Table 4.6: Estimates of Uninsured Children in 2013

parish_name	Sample mean	EBLUP	MSE	HB	Posterior std.dev.	HB_Probit	Posterior std.dev.	Sample Size	U19 population	Uninsured Children
Acadia	0.0253	0.0228	0.0230	0.0498	0.0475	0.0318	0.0131	79	17527	557
Allen	0.0417	0.0469	0.0294	0.0718	0.0561	0.0477	0.0196	48	6077	290
Ascension	0.0672	0.0701	0.0188	0.1201	0.0747	0.0526	0.0167	119	32273	1698
Assumption	0.1277	0.1204	0.0296	0.1350	0.0525	0.0722	0.0259	47	5837	422
Avoyelles	0.0345	0.0332	0.0220	0.0593	0.0432	0.0379	0.0144	87	10564	401
Beauregard	0.0588	0.0617	0.0222	0.0869	0.0584	0.0472	0.0166	85	9744	460
Bienville	0.1282	0.1228	0.0324	0.1226	0.0479	0.0877	0.0326	39	3468	304
Bossier	0.0079	0.0078	0.0183	0.0544	0.0626	0.0203	0.0091	126	31876	648
Caddo	0.0505	0.0499	0.0140	0.0918	0.0516	0.0421	0.0114	218	66674	2808
Calcasieu	0.0616	0.0627	0.0143	0.0953	0.0548	0.0516	0.0131	211	51134	2639
Caldwell	0.0303	0.0391	0.0351	0.0676	0.0574	0.0408	0.0185	33	2454	100
Cameron	0.0333	0.0328	0.0366	0.0718	0.0643	0.0335	0.0163	30	1730	58
Catahoula	0.1176	0.1140	0.0287	0.1288	0.0516	0.0769	0.0268	51	2476	190
Claiborne	0.0588	0.0494	0.0345	0.0826	0.0437	0.0490	0.0202	34	3390	166
Concordia	0.0161	0.0284	0.0261	0.0576	0.0449	0.0382	0.0162	62	5426	207
DeSoto	0.0943	0.0874	0.0279	0.1102	0.0479	0.0611	0.0216	53	6925	423
East Baton Rouge	0.0172	0.0249	0.0122	0.0774	0.0644	0.0236	0.008	290	113177	2675
East Carroll	0.0000	-0.0107	0.0463	0.0540	0.0429	0.0437	0.0205	21	2045	89
East Feliciana	0.0213	0.0241	0.0297	0.0696	0.0504	0.0351	0.0154	47	4400	155

Continued on next page

Table 4.6 – *Continued from previous page*

parish_name	Sample mean	EBLUP	MSE	HB	Posterior std.dev.	HB_Probit	Posterior std.dev.	Sample Size	U19 population	Uninsured Children
Evangeline	0.0000	0.0056	0.0314	0.0407	0.0457	0.0314	0.0149	42	9563	301
Franklin	0.0714	0.0464	0.0421	0.0853	0.0466	0.0499	0.0211	28	5538	276
Grant	0.0000	0.0054	0.0309	0.0365	0.0490	0.0286	0.014	43	5250	150
Iberia	0.0278	0.0353	0.0199	0.0703	0.0548	0.0310	0.0119	108	20777	645
Iberville	0.0862	0.0955	0.0269	0.1294	0.0535	0.0661	0.0234	58	7838	518
Jackson	0.0526	0.0476	0.0336	0.0744	0.0463	0.0449	0.0189	38	3888	175
Jefferson	0.0519	0.0557	0.0136	0.0963	0.0602	0.0453	0.0119	231	101097	4578
Jefferson Davis	0.0169	0.0221	0.0265	0.0497	0.0539	0.0307	0.0138	59	8757	268
LaSalle	0.0000	-0.0025	0.0196	0.0356	0.0636	0.0159	0.008	173	3713	59
Lafayette	0.0464	0.0511	0.0168	0.0946	0.0673	0.0417	0.013	151	57866	2415
Lafourche	0.0847	0.0883	0.0190	0.1165	0.0650	0.0664	0.0197	118	24886	1653
Lincoln	0.0435	0.0492	0.0249	0.0814	0.0511	0.0457	0.0172	69	12135	554
Livingston	0.0442	0.0464	0.0153	0.0799	0.0609	0.0392	0.0117	181	36726	1439
Madison	0.0000	-0.0031	0.0406	0.0475	0.0404	0.0410	0.0193	25	3127	128
Morehouse	0.0244	0.0230	0.0317	0.0593	0.0400	0.0424	0.0184	41	7173	304
Natchitoches	0.0217	0.0285	0.0301	0.0630	0.0465	0.0377	0.0164	46	10835	409
Orleans	0.0817	0.0814	0.0143	0.1263	0.0491	0.0644	0.0148	208	83488	5379
Ouachita	0.0363	0.0407	0.0149	0.0770	0.0565	0.0372	0.0115	193	42998	1601
Plaquemines	0.0000	0.0075	0.0332	0.0634	0.0652	0.0274	0.0137	37	6700	183
Pointe Coupee	0.0000	0.0106	0.0321	0.0561	0.0541	0.0318	0.0153	40	5696	181

Continued on next page

Table 4.6 – *Continued from previous page*

parish_name	Sample mean	EBLUP	MSE	HB	Posterior std.dev.	HB_Probit	Posterior std.dev.	Sample Size	U19 population	Uninsured Children
Rapides	0.0482	0.0536	0.0160	0.0856	0.0548	0.0455	0.0134	166	35638	1622
Red River	0.1000	0.0887	0.0303	0.1151	0.0475	0.0667	0.0236	50	2387	159
Richland	0.0000	0.0057	0.0325	0.0455	0.0467	0.0317	0.0152	39	5590	177
Sabine	0.0000	0.0097	0.0321	0.0447	0.0506	0.0299	0.0146	40	6225	186
St. Bernard	0.0625	0.0651	0.0356	0.0866	0.0515	0.0486	0.0209	32	10700	520
St. Charles	0.0317	0.0359	0.0258	0.0863	0.0634	0.0349	0.0146	63	14464	505
St. Helena	0.0526	0.0605	0.0329	0.0892	0.0428	0.0529	0.0212	38	2863	151
St. James	0.0612	0.0617	0.0291	0.1086	0.0577	0.0459	0.0186	49	5784	266
St. John Baptist	0.1538	0.1444	0.0253	0.1698	0.0501	0.0873	0.0265	65	12547	1095
St. Landry	0.0354	0.0389	0.0195	0.0679	0.0463	0.0386	0.0137	113	23741	917
St. Martin	0.0235	0.0348	0.0223	0.0686	0.0547	0.0354	0.0145	85	14424	511
St. Mary	0.0253	0.0300	0.0231	0.0633	0.0505	0.0337	0.0137	79	14340	483
St. Tammany	0.0175	0.0178	0.0137	0.0710	0.0768	0.0219	0.0078	228	62108	1362
Tangipahoa	0.1043	0.1030	0.0192	0.1260	0.0488	0.0740	0.0205	115	33018	2442
Tensas	0.0333	0.0268	0.0366	0.0661	0.0409	0.0453	0.0195	30	1281	58
Terrebonne	0.0328	0.0387	0.0188	0.0747	0.0642	0.0341	0.0124	122	30163	1029
Union	0.1034	0.0987	0.0373	0.1050	0.0504	0.0590	0.0249	29	5427	320
Vermilion	0.0361	0.0452	0.0225	0.0749	0.0599	0.0419	0.0162	83	16100	675
Vernon	0.0602	0.0644	0.0225	0.0863	0.0531	0.0513	0.0179	83	15027	771
Washington	0.1429	0.1314	0.0313	0.1284	0.0426	0.0790	0.0276	42	12283	970

Continued on next page

Table 4.6 – *Continued from previous page*

parish_name	Sample mean	EBLUP	MSE	HB	Posterior std.dev.	HB_Probit	Posterior std.dev.	Sample Size	U19 population	Uninsured Children
Webster	0.0303	0.0273	0.0252	0.0562	0.0438	0.0371	0.0148	66	10126	376
West Baton Rouge	0.0233	0.0362	0.0310	0.0847	0.0640	0.0374	0.0169	43	6256	234
West Carroll	0.0189	0.0214	0.0280	0.0474	0.0442	0.0332	0.0149	53	2985	99
West Feliciana	0.0222	0.0267	0.0303	0.0833	0.0655	0.0342	0.0156	45	2737	94
Winn	0.0667	0.0699	0.0265	0.0823	0.0428	0.0533	0.0191	60	3524	188

4.6 Model Specification for Cross-sectional Data

In the previous section, we estimate the health insurance coverage for 64 parishes in Louisiana by employing the single survey year 2013. As we mentioned in the previous section, the LHIS data is a biannual survey that was conducted each year since 2003, which allows us to pool data from every year together in order to get a significant increase of sample size. However, the hypothesis test rejects the null hypothesis that parishes have equal coefficients over all years. Next, we use an informative prior which allows for a continuous shift from single year estimates to pooled year estimates.

The regression model is written as follows,

$$\begin{aligned} y_{ij} =& \sum_{k=1}^{6} \beta_{ik} x_{ik} + \sum_{k=7}^{11} \beta_{ij} * D_{year} + \sum_{k=12}^{17} \beta_{ik} x_{ik} * D_{2003} \\ &+ \sum_{k=18}^{23} \beta_{ik} x_{ik} * D_{2005} + \sum_{k=24}^{29} \beta_{ik} x_{ik} * D_{2007} \\ &+ \sum_{k=30}^{35} \beta_{ik} x_{ik} * D_{2009} + \sum_{k=36}^{41} \beta_{ik} x_{ik} * D_{2011} + v_i + e_{ij} \end{aligned} \tag{4.1}$$

where $i = 1, ..., 64$, $j = 1, ...m_i$, $D_{2003}, ..., D_{2011}$ are dummy variables that takes the value one if the individual is collected in that particular year.

We use survey year 2013 as the reference year. This means that if we are interested in the parameters for year 2013, we focus on β_1 through β_6. For all other years, the parameters β are obtained by adding to the reference year's β. For instance, the values of the corresponding β parameters for variables in survey year 2003 are the β values $\beta_{1,2003} = \beta_1 + \beta_7 + \beta_{12}$, which is the estimate coefficient for variable "black" in survey year 2003.

We define the following matrices and vectors,

$$Z = \begin{bmatrix} X_{1,2003} & 0 & 0 & 0 & \cdots & 0 & 0 \\ X_{1,2005} & X_{1,2005} & 0 & 0 & \cdots & 0 & 0 \\ X_{1,2007} & 0 & X_{1,2007} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ X_{64,2011} & 0 & 0 & 0 & \cdots & 0 & X_{64,2011} \end{bmatrix} \tag{4.2}$$

where $X_{1,2003}$, $X_{1,2005}$, $\cdots$, $X_{1,2011}$, $\cdots$, $X_{64,2011}$ are the independent variables matrices from all years and all parishes, with $N = N_{1,2003} + \cdots + N_{1,2013} + N_{2,2003} + \cdots + N_{64,2013}$, the number of individuals obtained by adding up the number of individuals for each parish.

Let $K_i = \sum_{n=1}^{m_i} K_j$ and $K_0 = 0$. The rows from 1 to K_1 are the N_1, $N_1 = N_{1,2003}+, \cdots, N_{64,2003}$, stacked observations for the survey year 2003. The rows from K_1+1 to K_2 are the N_2 stacked observations for the survey year 2005, and so on. Therefore, we have

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{41} \end{pmatrix}, \; y = \begin{pmatrix} y_{1,2003} \\ y_{1,2005} \\ \vdots \\ y_{64,2011} \end{pmatrix}, \; \Sigma = \begin{pmatrix} \Sigma_{1,2003} & 0 & 0 & \cdots & 0 \\ 0 & \Sigma_{1,2005} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Sigma_{64,2011} \end{pmatrix} \tag{4.3}$$

where Σ is the data variance-covariance $N * N$ diagonal matrix in which the variance corresponding to the n's individual is inserted in row n. The covariance terms are equal to zero. The error term $e_{ij} \sim N(0, \sigma_e^2)$ are still normally distributed.

We define an informative prior for $\beta \sim N(\beta_p, H_p^{-1})$, where β_p is a $41*1$ vector of constants, and H_p is a 41*41 positive definite matrix of constants. We specify $\beta_p = (0, 0, \cdots, 0)^T$ as the null vector and construct the prior precision matrix as a diagonal matrix. We also define a scale factor S, which is a constant used to multiply rows 7 to 41 of the prior precision matrix. Next, we run the Gibbs sampler for different values of the scale factor S, S=0.0001; 0.001; 0.01; 0.1; 1; 10; 100; 1,000; 10,000; 100,000; 1,000,000; 100,000,000; and 1,000,000,000. The larger the value of scale factor S, the stronger the prior. Intuitively, the magnitudes of the scale factor indicate the level of the mixture of the model. For instance, the small value of S, $S = 1$ or smaller, will generate the basic model as we described earlier. With the increasing of the scale factor S, the models are pushed to the pooled estimates over all years.

Therefore, the four conditional distributions are derived based on the same references as we discussed in section 3.4.

The process is repeated 25,000 times to product 25,000 Markov Chains Monte Carlo iterations with 5,000 "burn-in" draws.

4.7 Empirical Results for Cross-sectional Data

In this section, we present the results in graphs. We illustrate the impacts of the informative priors on the estimates.

Figure 4.9 shows the estimates of the coefficients for the cross-term of the independent variable "female" and each survey year. The horizontal axis lists the logarithm of the scale factor which ranges from -5 to 9. We can observe that under the informative prior, with the increasing the strength of the prior, the estimated coefficients converge to zero eventually. For the variable "female," the convergence starts from scale factor $S = 0$, and roughly close to zero when the scale factor $S = 1,000$.

Figure 4.10 lists the convergence of the cross term of the independent variable "poverty" and different survey years. Similar to the variable "female," the convergence starts from the scale factor $S = 0$, and close to zero once the scale factor increases to $S = 1,000$. We also provide the 90% highest density regions in the graph. The graph also points out that the posterior highest density regions of the parameters are generally wider for weak scale factors, while the regions are much smaller for stronger scale factors.

Figure 4.11 shows the variation of the cross-term of independent variable "income" and each survey years. Different from the previous two variables, the cross-terms of income and year convergence to zero with both positive and negative starting values. For instance, with the base as survey year 2013, the starting points are all negative except year 2007. Given the estimated coefficients of the variable "income" is negative, the individual's uninsured probabilities decrease with the increasing of the household income. For the survey year 2007, holding other variables constant, the impact of household income on an individual's uninsured probabilities are smaller than the impact of the year 2013, while for other years, the impact of household income are slightly larger on individual's uninsured probabilities.

Figure 4.12 shows the convergence of the most different variable "age." In the model, variable "age" has the maximum magnitudes relative to other variables[8]. As with other variable, the convergence starts from scale factor S=0, which is equivalent to no strength on the prior. After a large fluctuation, the variable converges to zero until the scale factor S=1 billion. Under this circumstance, estimates of uninsured rates of the population may be impacted heavily due to the convergence speed of the "age" variable. We also illustrate the impact of variable "age" on the estimates of uninsured rates.

Figure 4.13 shows the estimates of selected variables' convergence over the informative prior. We observe that when the prior's strength is small (S=0.00001), the estimates are close to the cross-sectional estimates (survey year 2013), which reflect the column L1 in Table 4.7. With the increasing strength of the prior, the estimates converge to pooled estimates over the year 2003-2013, which is listed in the last column (L2) in Table 4.7.

[8]As mentioned earlier, the variable household income has been adjusted as household income/10,000.

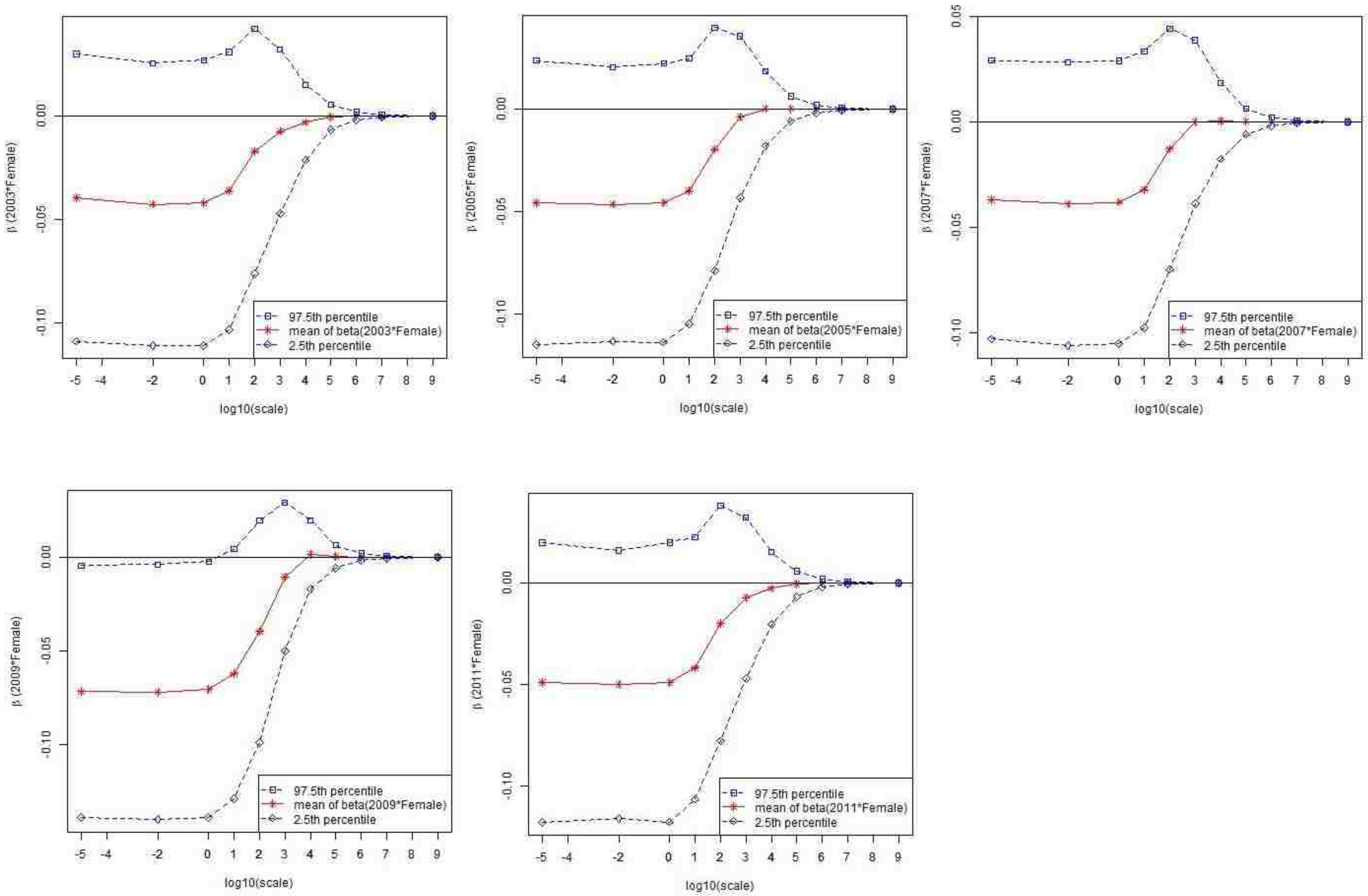

Figure 4.9: Estimates of β_{Female} in different survey years for adults

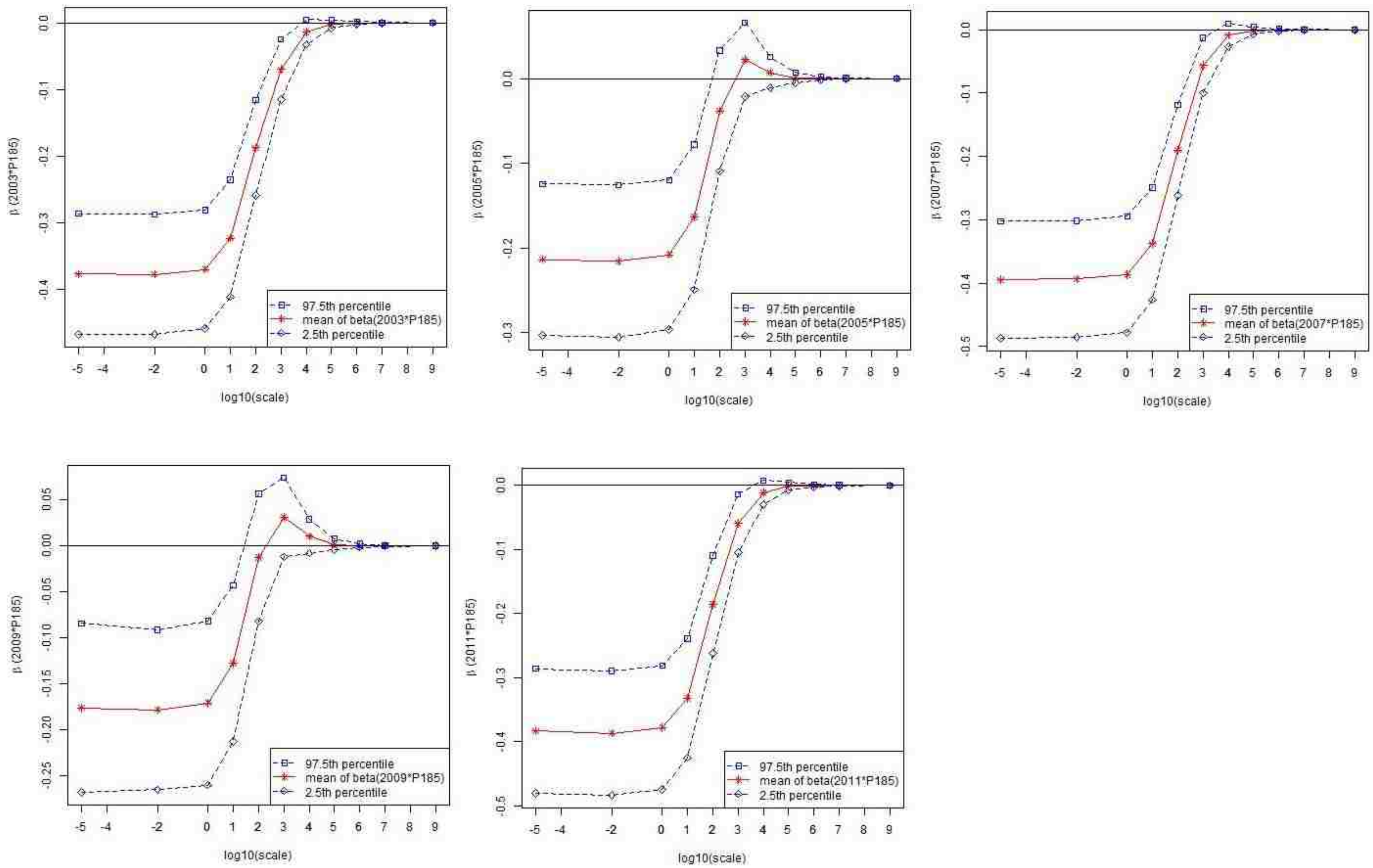

Figure 4.10: Estimates of $\beta_{Poverty}$ in different survey years for adults

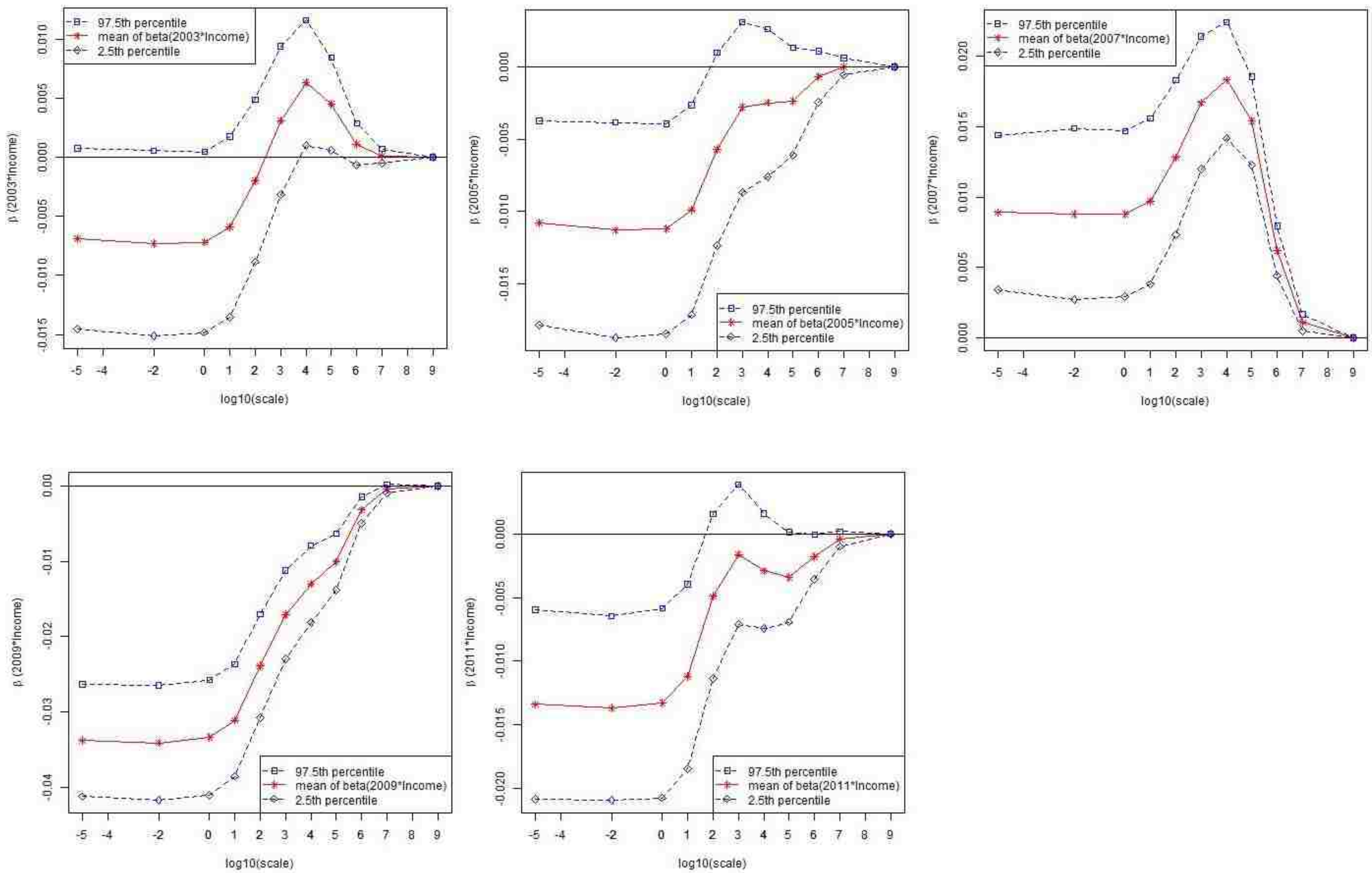

Figure 4.11: Estimates of β_{Income} in different survey years for adults

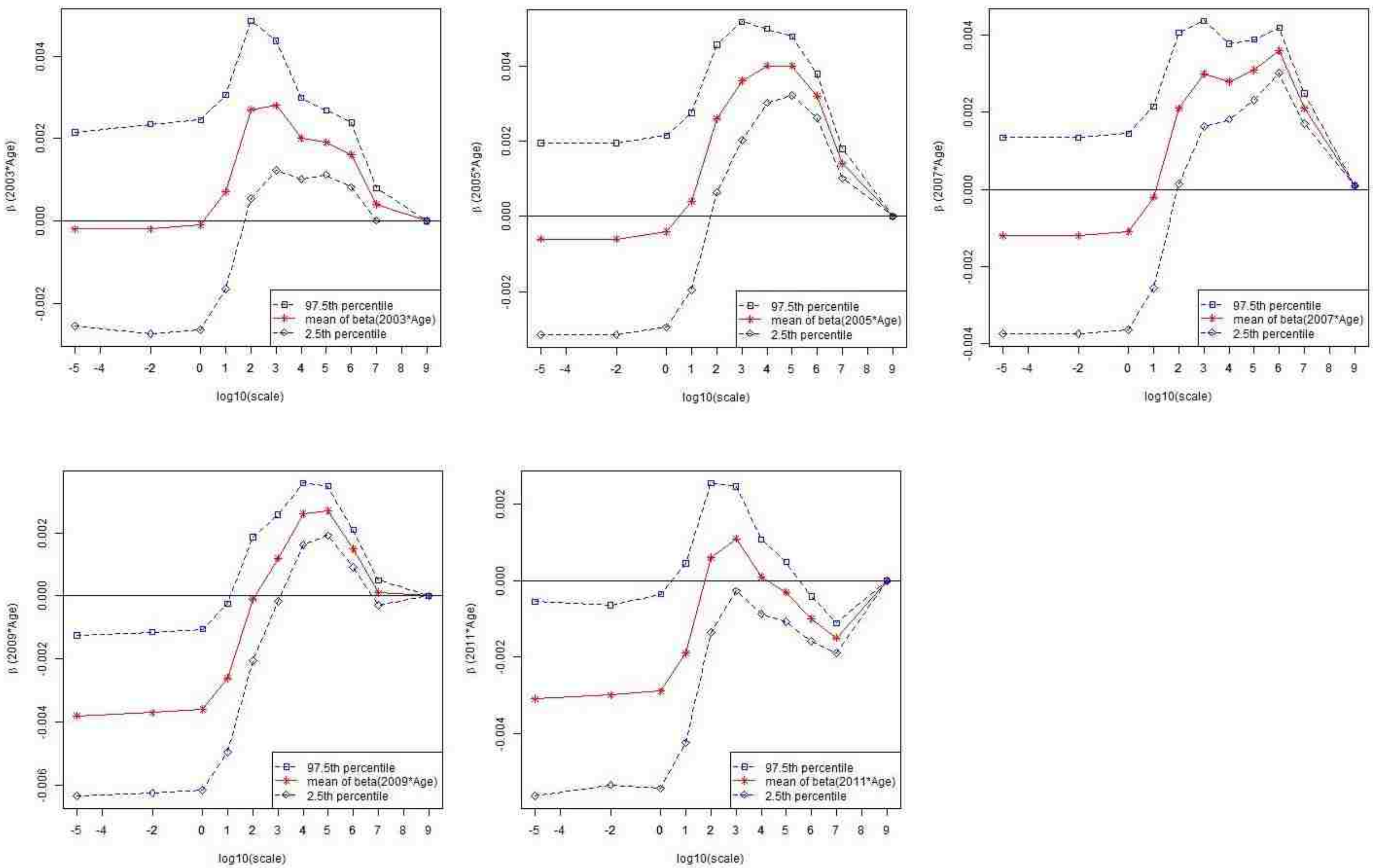

Figure 4.12: Estimates of β_{Age} in different survey years for adults

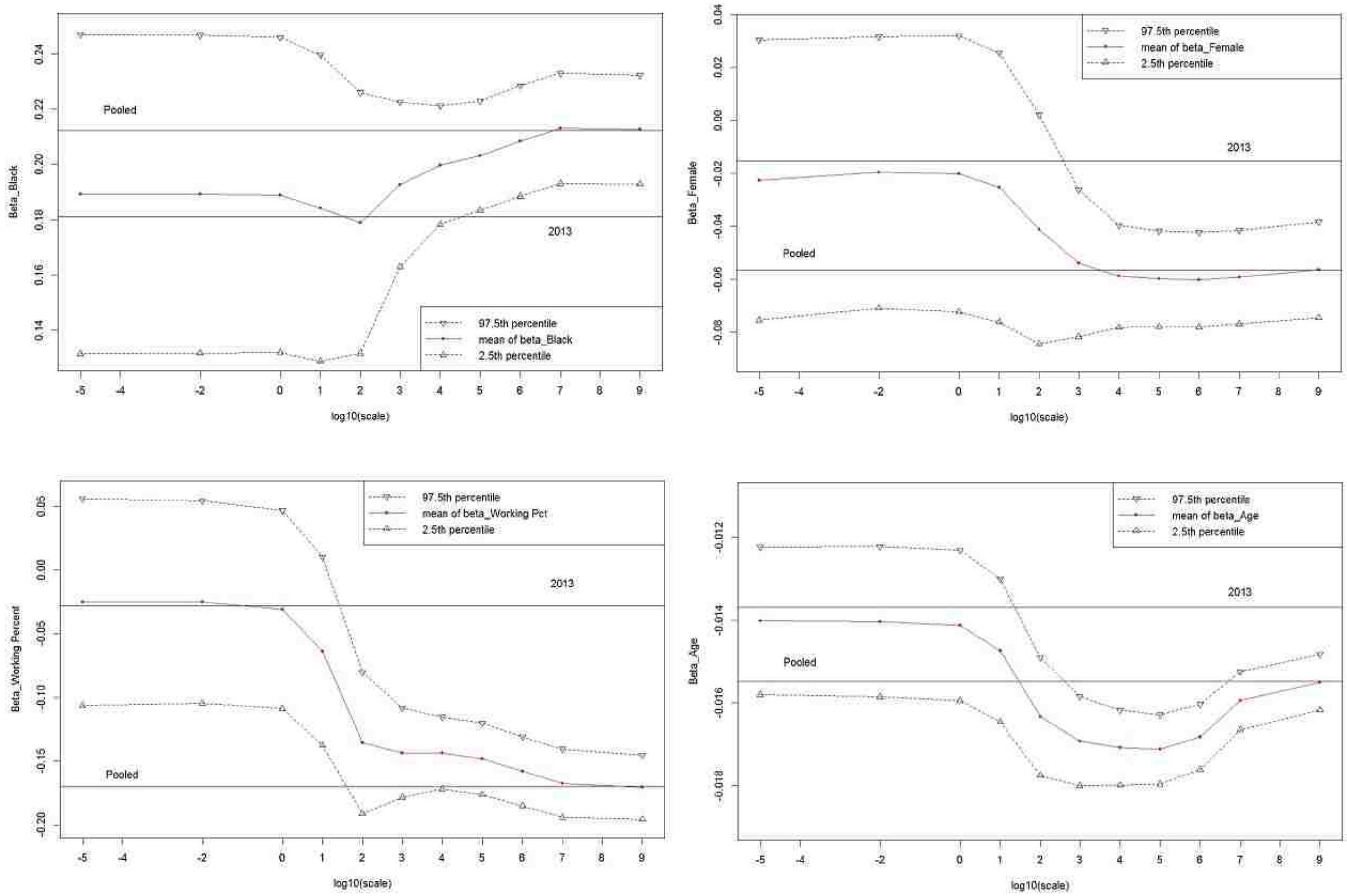

Figure 4.13: Parameters with 95% highest density region for adults

Table 4.7 lists the posterior means and standard deviations together with the 95% percent highest density regions[9] for the parameters in the survey year 2013. Column L1 lists the results for the single year estimates from the earlier section, while the last column L2 indicates the pooled estimates over years 2003 to 2013. We also plot the graphs for each variable to illustrate the continuous shift from single year estimates to pooled year estimates.

Figure 4.14 to Figure 4.21 lists the estimates for the uninsured rates for each parish under the sequence of informative prior S. First of all, the uninsured rates increased with the increased strength of the prior. As we mentioned earlier, the varying strength of the informative prior could be considered as the shrinking of the pooled estimator towards cross-sectional results. Therefore, the convergence for each parish shows that the uninsured rates are higher for the pooled estimator. Secondly, the convergence of the uninsured rates for each parish has certain fluctuation. For instance, most of them experience a local maximum at scale factor S=100 and a local minimum at scale factor S=100,000. As we mentioned earlier, this may be caused by the relatively large magnitudes of variable "age." As well as the adult's graph, the estimates of uninsured rates for children (Figure 4.22 to Figure 4.29) also converge to the pooled estimates with the increasing of the prior's strength. However, since there is no large magnitudes variable for children (like the variable "age" for adults), the convergence is smooth over the increasing of the prior's strength.

[9]Posterior moments are computed based on 25,000 points generated from the Gibbs sampling algorithm with the first 5,000 as the burn-in samples. The end points of the 95% confidence region are the 2.5^{th} and the 97.5^{th} percentiles of the posterior marginal densities.

Table 4.7: Posterior Means, Standard Deviation, and 95% Highest Density Region (Adults)

Variables	L1	S=0.0001	S=0.01	S=1	S=10	S=100	S=1,000	S=100,000	S=10,000,000	S=1,000,000,000	L2
Constant	-0.5311	-0.5544	-0.5541	-0.5404	-0.4506	-0.2167	-0.1162	-0.1019	-0.0832	-0.0552	-0.0549
Post. S. D	(0.0616)	(0.0648)	(0.0647)	(0.0661)	(0.0577)	(0.0389)	(0.0253)	(0.0219)	(0.0217)	(0.0208)	(0.0207)
[H.D.R.]	[-0.6518,-0.4104]	[-0.6813,-0.4274]	[-0.6808,-0.4274]	[-0.6699,-0.4108]	[-0.5637,-0.3375]	[-0.2929,-0.1405]	[-0.1658,-0.0666]	[-0.1448,-0.0589]	[-0.1258,-0.0406]	[-0.096,-0.0143]	[-0.0954,-0.0144]
Black	0.1811	0.1892	0.1892	0.1889	0.1841	0.1789	0.1927	0.2032	0.213	0.2127	0.2124
Post. S. D	(0.0289)	(0.0295)	(0.0294)	(0.0291)	(0.0283)	(0.0241)	(0.0152)	(0.0101)	(0.0102)	(0.01)	(0.01)
[H.D.R.]	[0.1245,0.2377]	[0.1314,0.247]	[0.1316,0.2468]	[0.1319,0.246]	[0.1286,0.2396]	[0.1316,0.2262]	[0.1629,0.2225]	[0.1834,0.223]	[0.193,0.233]	[0.193,0.2323]	[0.1929,0.2319]
Income	-0.0192	-0.0237	-0.0235	-0.0235	-0.0239	-0.0256	-0.0285	-0.0298	-0.0291	-0.0292	-0.0292
Post. S. D	(0.0023)	(0.0023)	(0.0024)	(0.0024)	(0.0024)	(0.0022)	(0.002)	(0.0014)	(0.001)	(0.001)	(0.001)
[H.D.R.]	[-0.0237,-0.0147]	[-0.0283,-0.0191]	[-0.0282,-0.0187]	[-0.0282,-0.0189]	[-0.0285,-0.0192]	[-0.0298,-0.0213]	[-0.0323,-0.0246]	[-0.0325,-0.027]	[-0.0311,-0.0271]	[-0.0311,-0.0273]	[-0.0312,-0.0272]
Female	-0.0153	-0.0226	-0.0196	-0.0203	-0.0253	-0.0412	-0.054	-0.0599	-0.0592	-0.0565	-0.0565
Post. S. D	(0.0256)	(0.027)	(0.0261)	(0.0266)	(0.0259)	(0.0221)	(0.0142)	(0.0092)	(0.009)	(0.0092)	(0.0091)
[H.D.R.]	[-0.0655,0.0349]	[-0.0755,0.0303]	[-0.0709,0.0316]	[-0.0723,0.0318]	[-0.0761,0.0254]	[-0.0844,0.0021]	[-0.0817,-0.0262]	[-0.0779,-0.0418]	[-0.0768,-0.0416]	[-0.0746,-0.0384]	[-0.0744,-0.0386]
P185	0.7454	0.7796	0.7786	0.7719	0.7293	0.6075	0.522	0.5011	0.489	0.4566	0.4555
Post. S. D	(0.0338)	(0.0354)	(0.0346)	(0.0346)	(0.0332)	(0.0257)	(0.0169)	(0.0125)	(0.0123)	(0.0117)	(0.0117)
[H.D.R.]	[0.6792,0.8116]	[0.7102,0.8491]	[0.7109,0.8464]	[0.704,0.8398]	[0.6643,0.7943]	[0.5571,0.6579]	[0.4888,0.5552]	[0.4766,0.5255]	[0.4649,0.5131]	[0.4337,0.4795]	[0.4325,0.4785]
Working Percent	-0.028	-0.0251	-0.0251	-0.0309	-0.0637	-0.1357	-0.1434	-0.1481	-0.1673	-0.1705	-0.1695
Post. S. D	(0.0395)	(0.0414)	(0.0405)	(0.0397)	(0.0377)	(0.0284)	(0.0179)	(0.0143)	(0.0136)	(0.0128)	(0.0126)
[H.D.R.]	[-0.1054,0.0494]	[-0.1063,0.056]	[-0.1045,0.0542]	[-0.1088,0.047]	[-0.1376,0.0102]	[-0.1913,-0.0801]	[-0.1785,-0.1083]	[-0.1762,-0.1201]	[-0.194,-0.1406]	[-0.1956,-0.1453]	[-0.1943,-0.1448]
Age	-0.0137	-0.014	-0.014	-0.0141	-0.0147	-0.0163	-0.0169	-0.0171	-0.0159	-0.0155	-0.0155
Post. S. D	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0007)	(0.0006)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
[H.D.R.]	[-0.0155,-0.0119]	[-0.0158,-0.0122]	[-0.0159,-0.0122]	[-0.0159,-0.0123]	[-0.0165,-0.013]	[-0.0178,-0.0149]	[-0.018,-0.0158]	[-0.018,-0.0163]	[-0.0167,-0.0152]	[-0.0162,-0.0148]	[-0.0162,-0.0148]

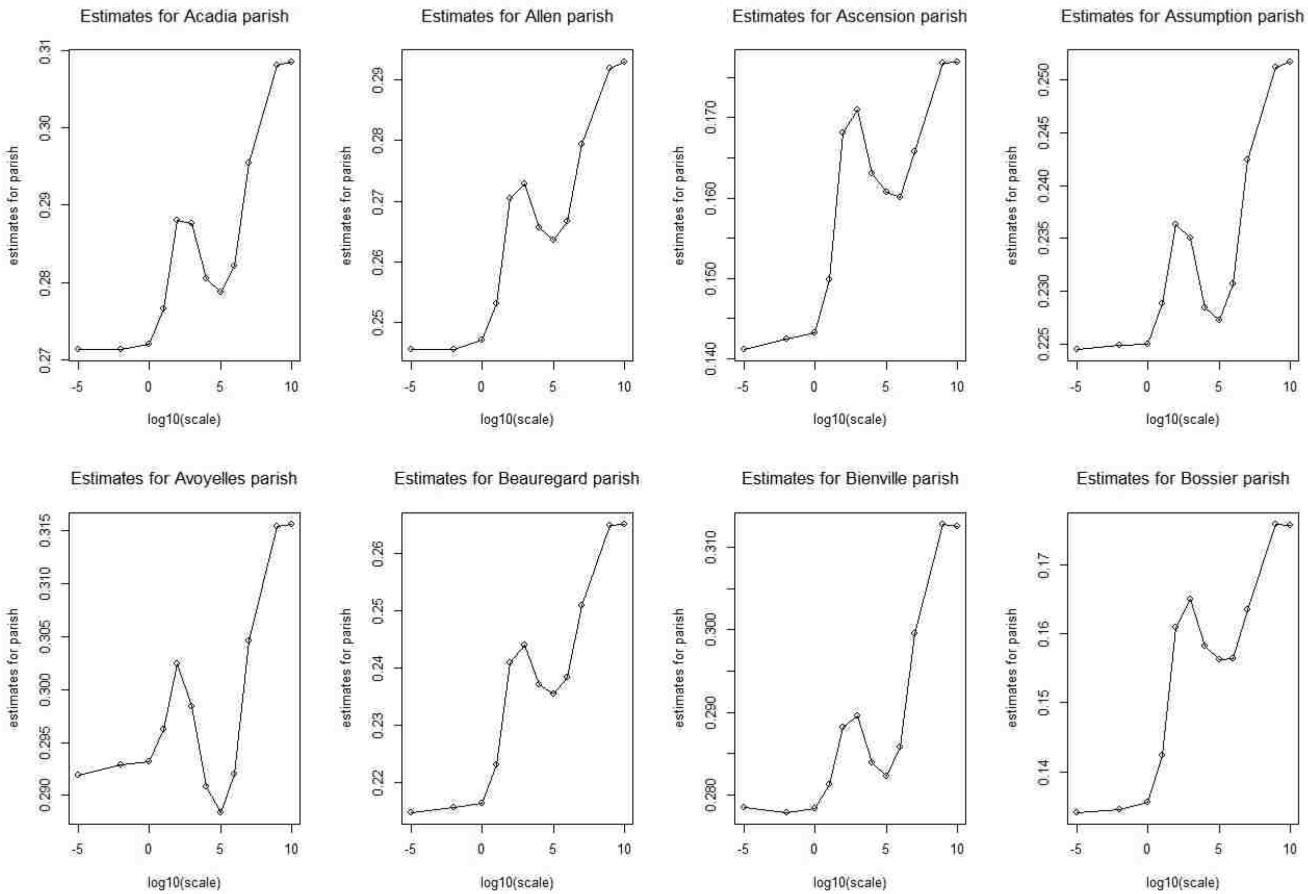

Figure 4.14: Estimates for parishes based on informative priors (adults)

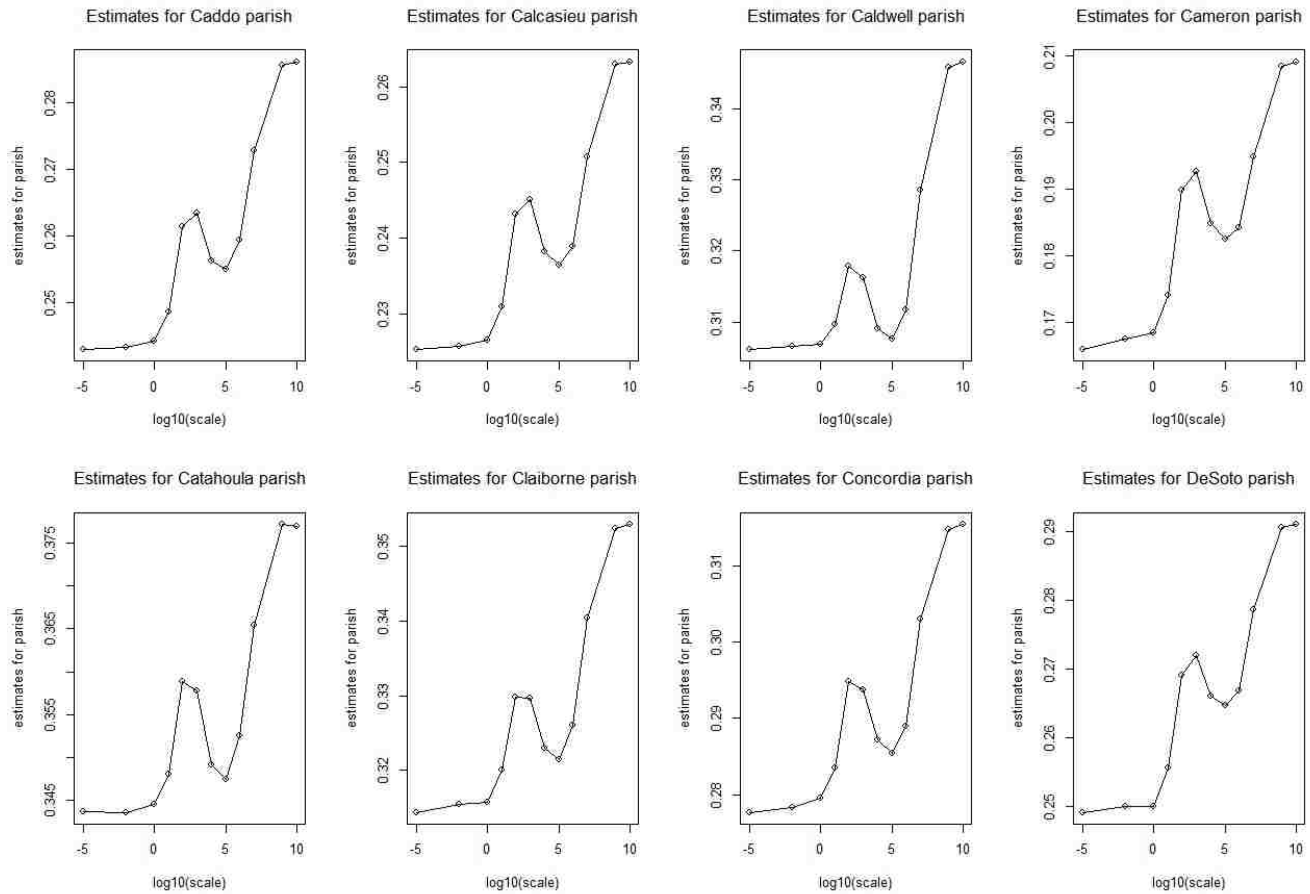

Figure 4.15: Estimates for parishes based on informative priors (adults)

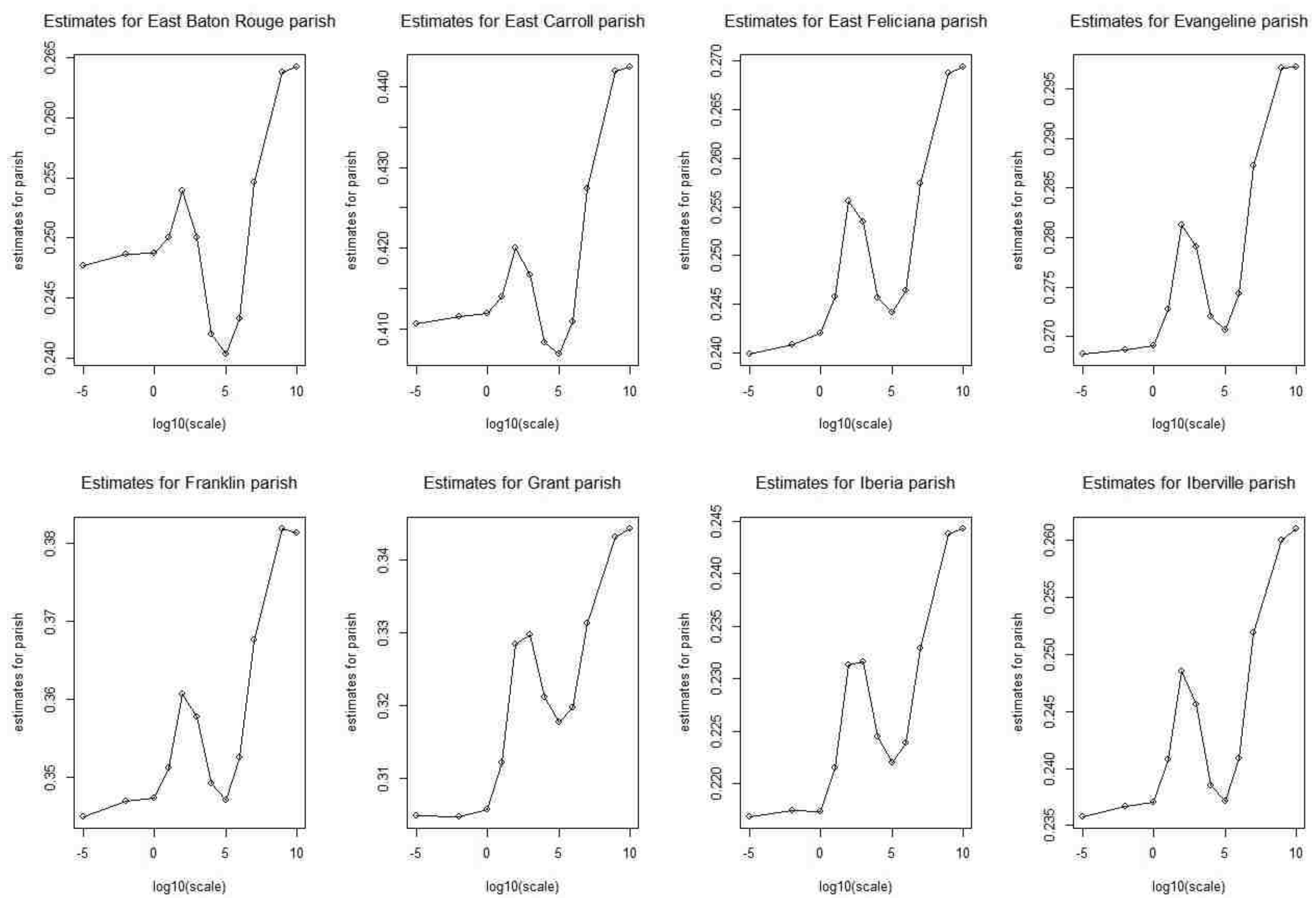

Figure 4.16: Estimates for parishes based on informative priors (adults)

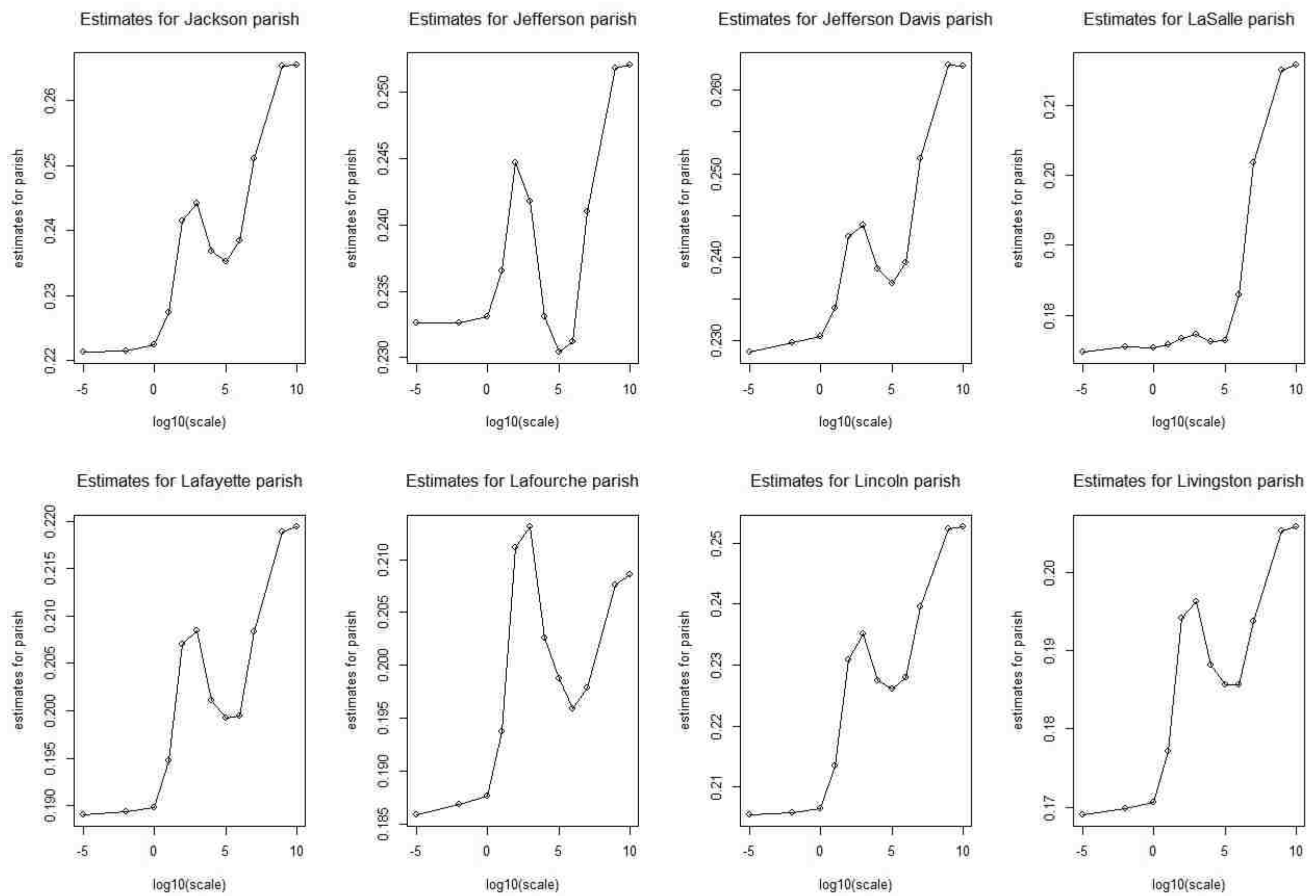

Figure 4.17: Estimates for parishes based on informative priors (adults)

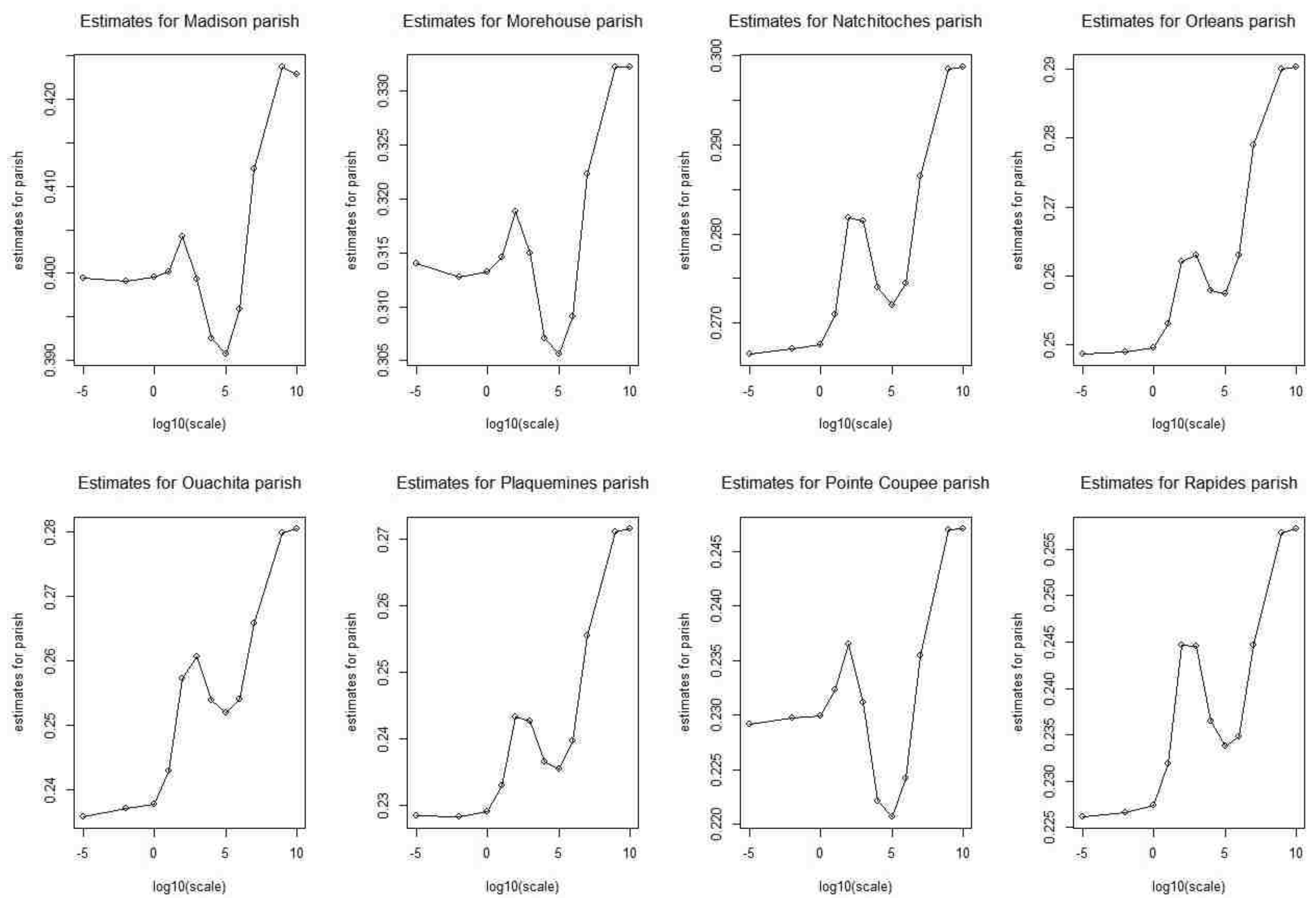

Figure 4.18: Estimates for parishes based on informative priors (adults)

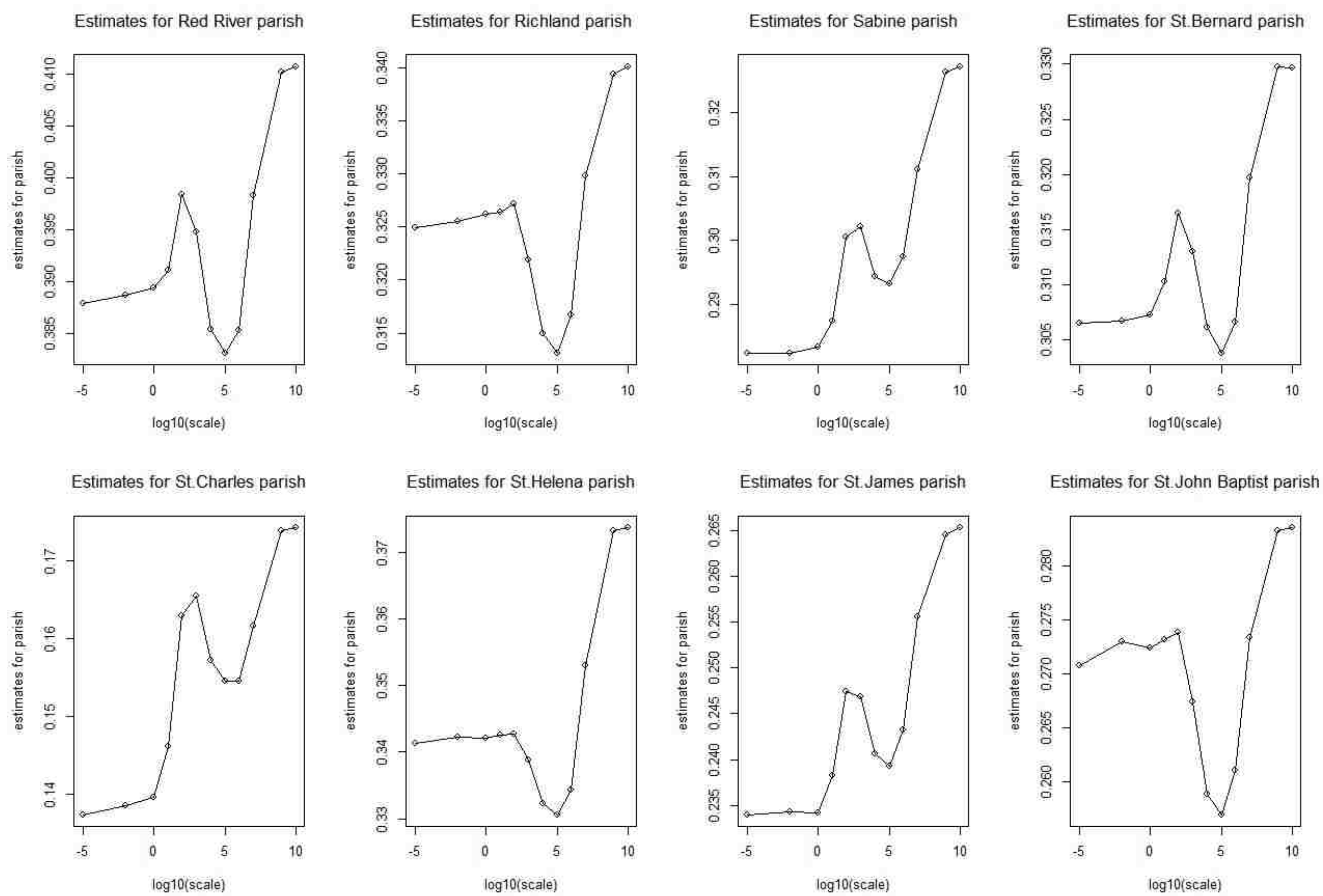

Figure 4.19: Estimates for parishes based on informative priors (adults)

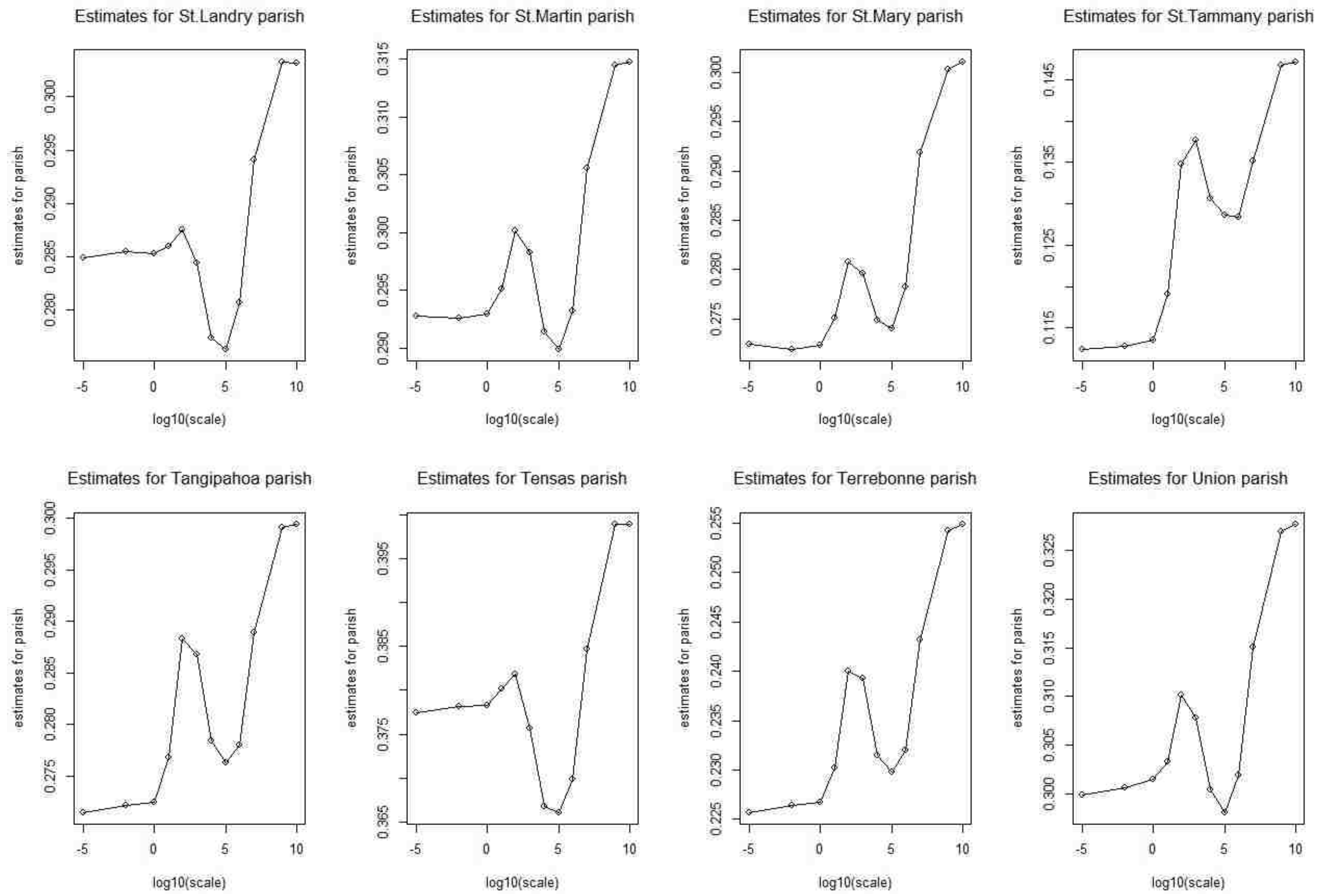

Figure 4.20: Estimates for parishes based on informative priors (adults)

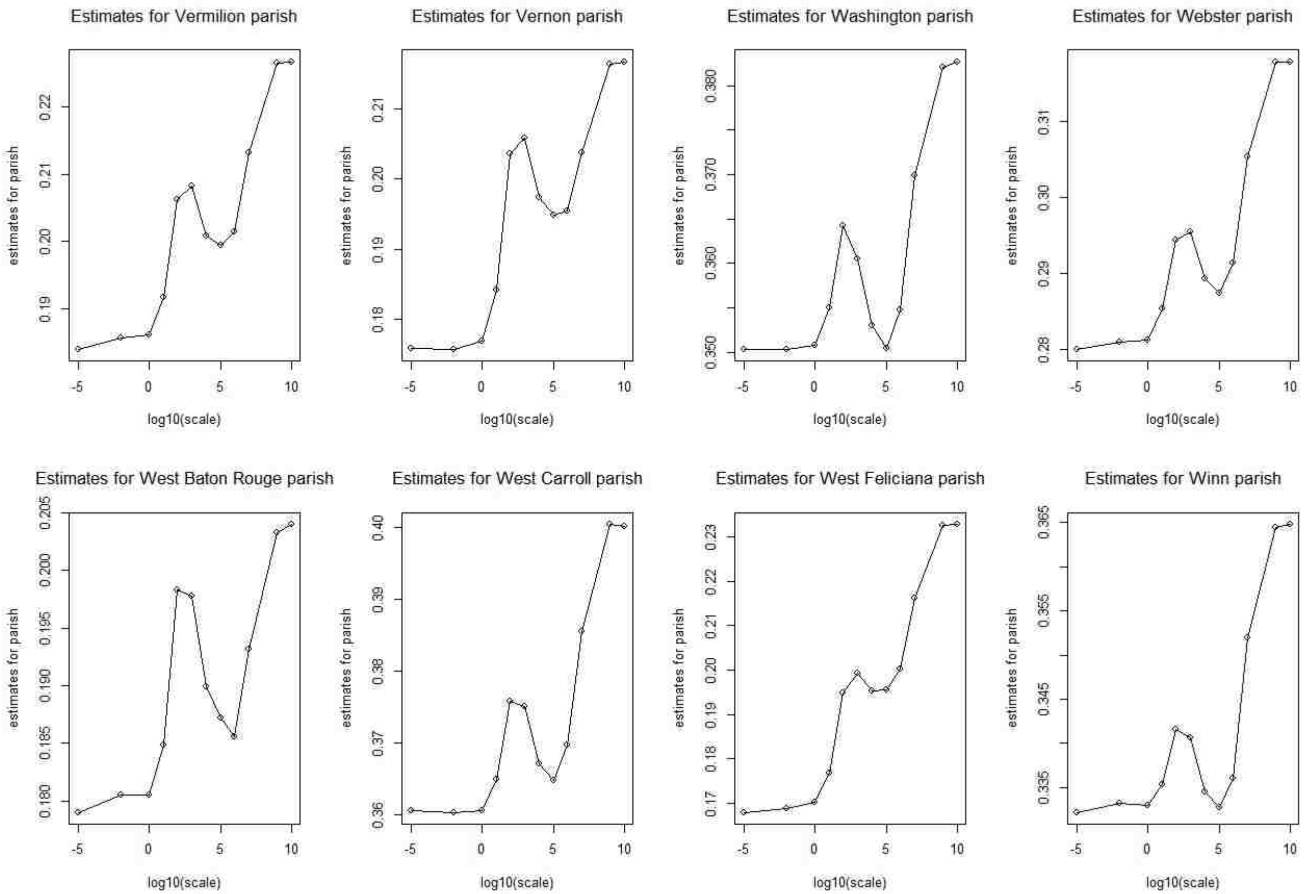

Figure 4.21: Estimates for parishes based on informative priors (adults)

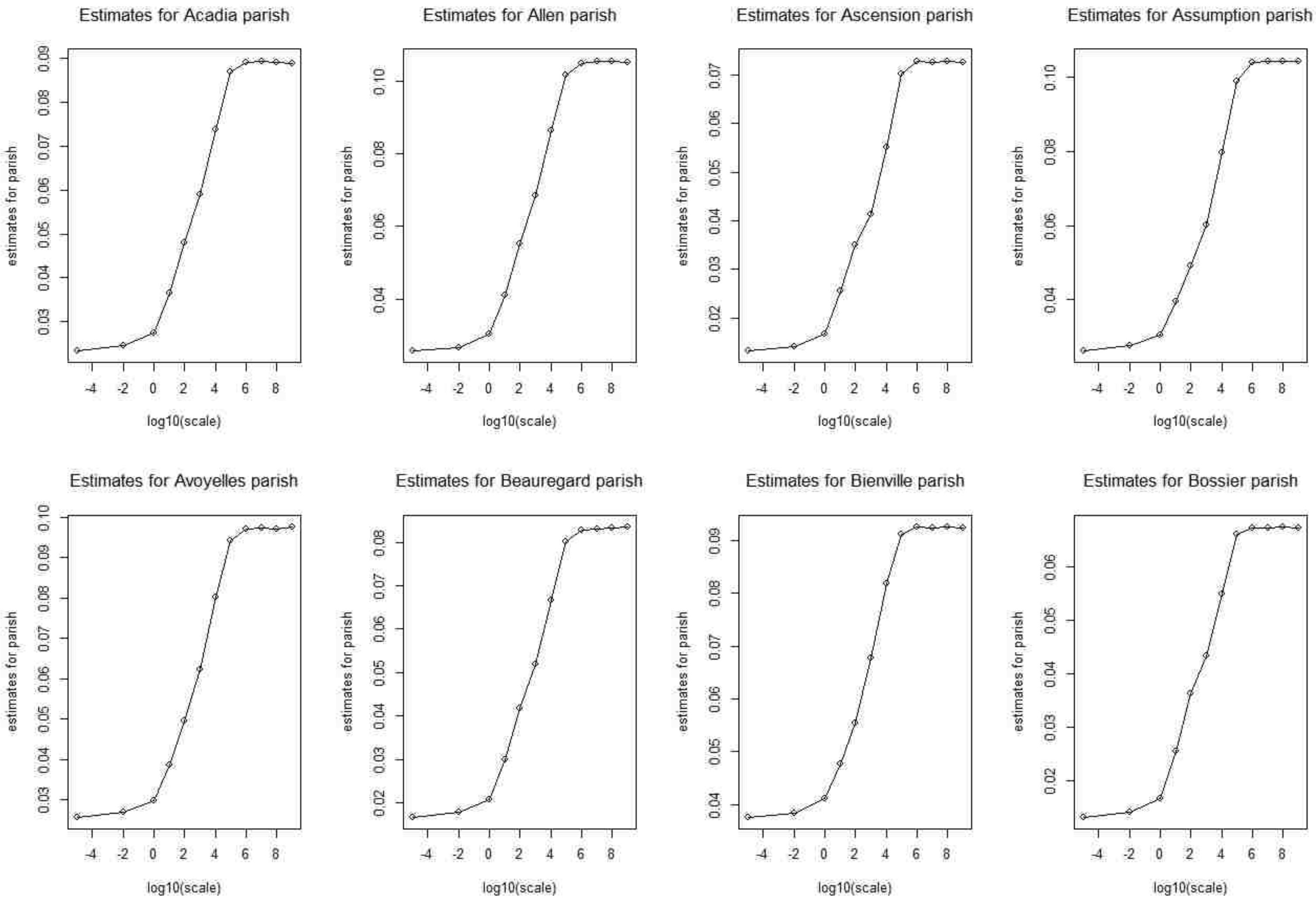

Figure 4.22: Estimates for parishes based on informative priors (children)

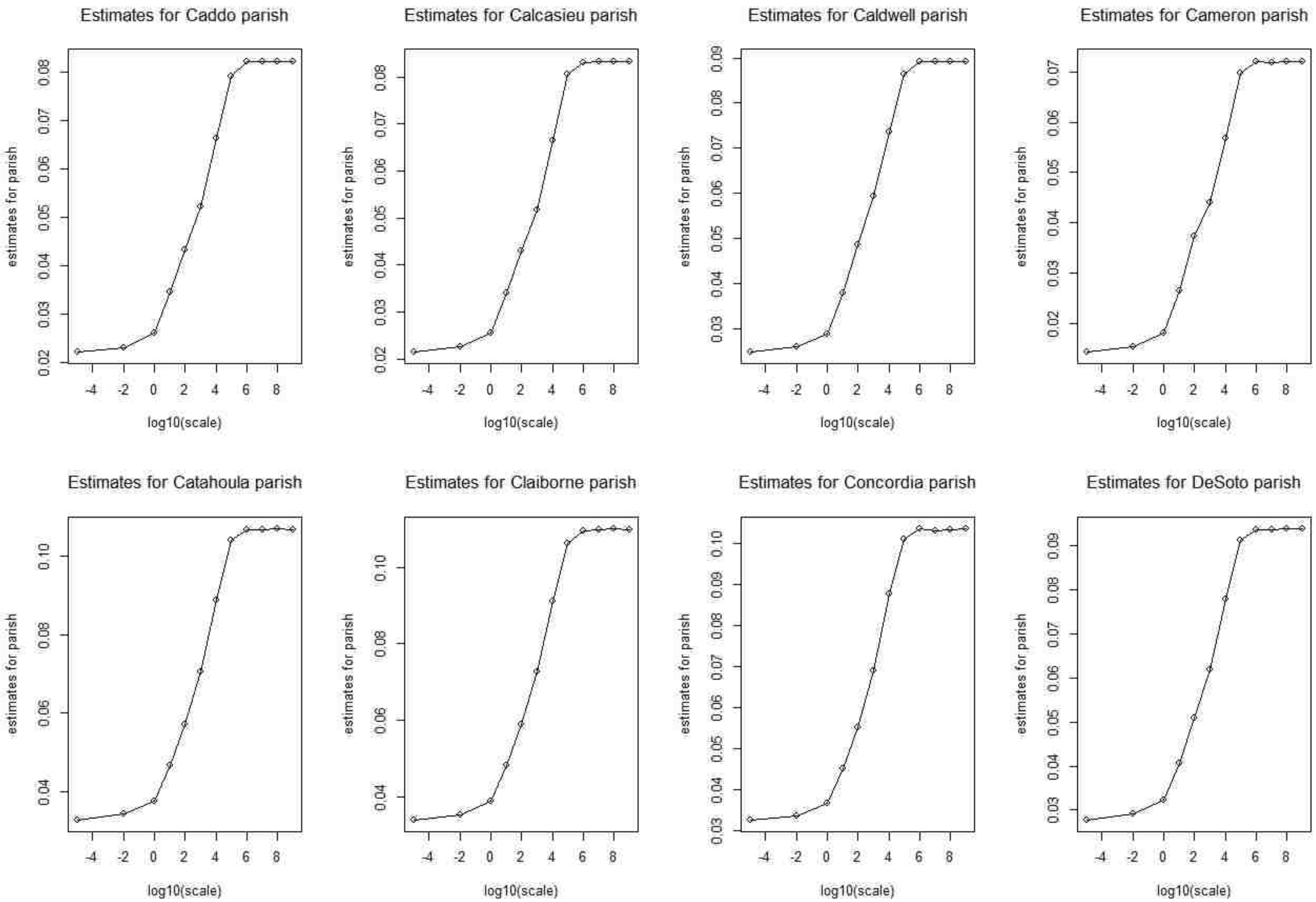

Figure 4.23: Estimates for parishes based on informative priors (children)

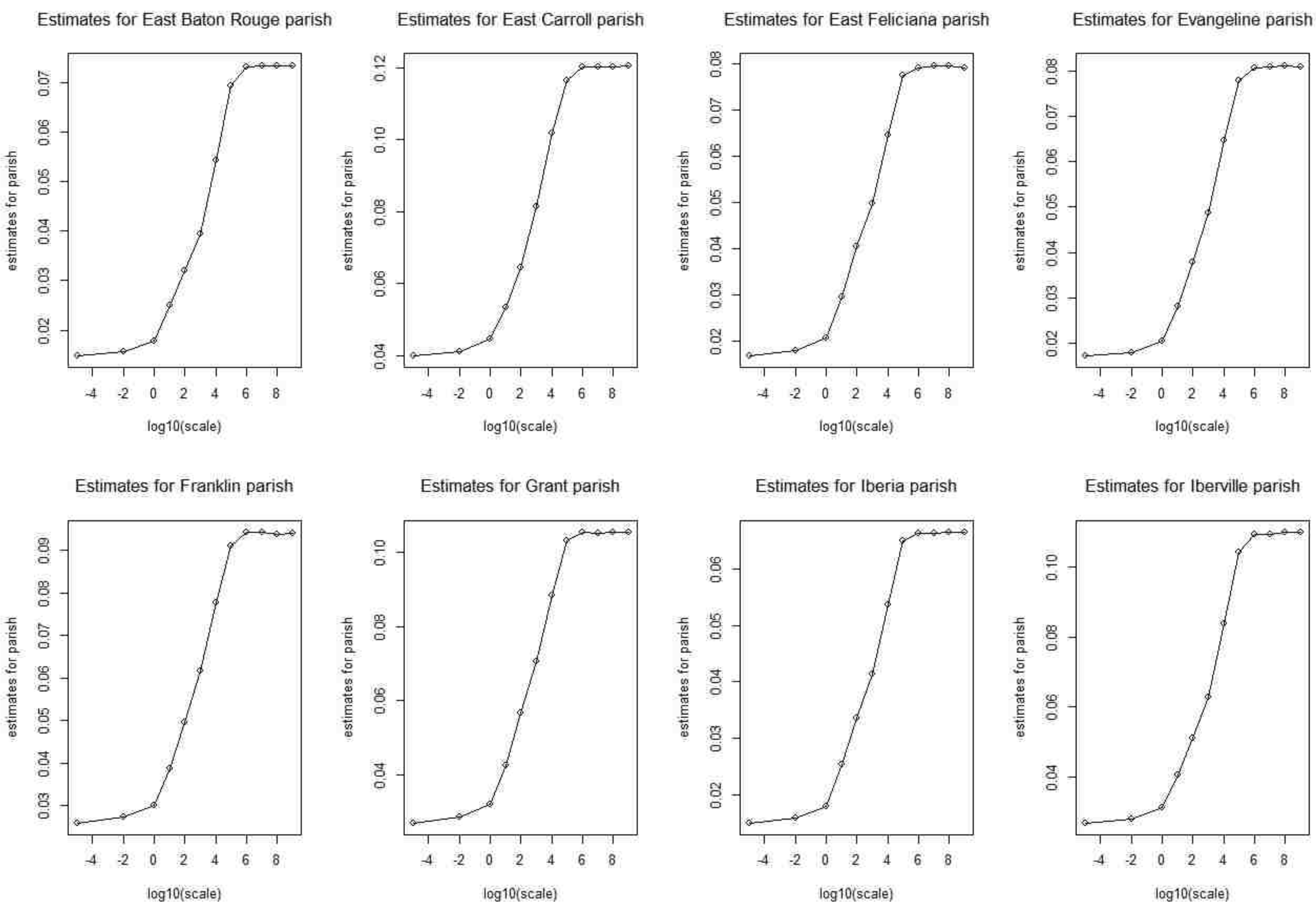

Figure 4.24: Estimates for parishes based on informative priors (children)

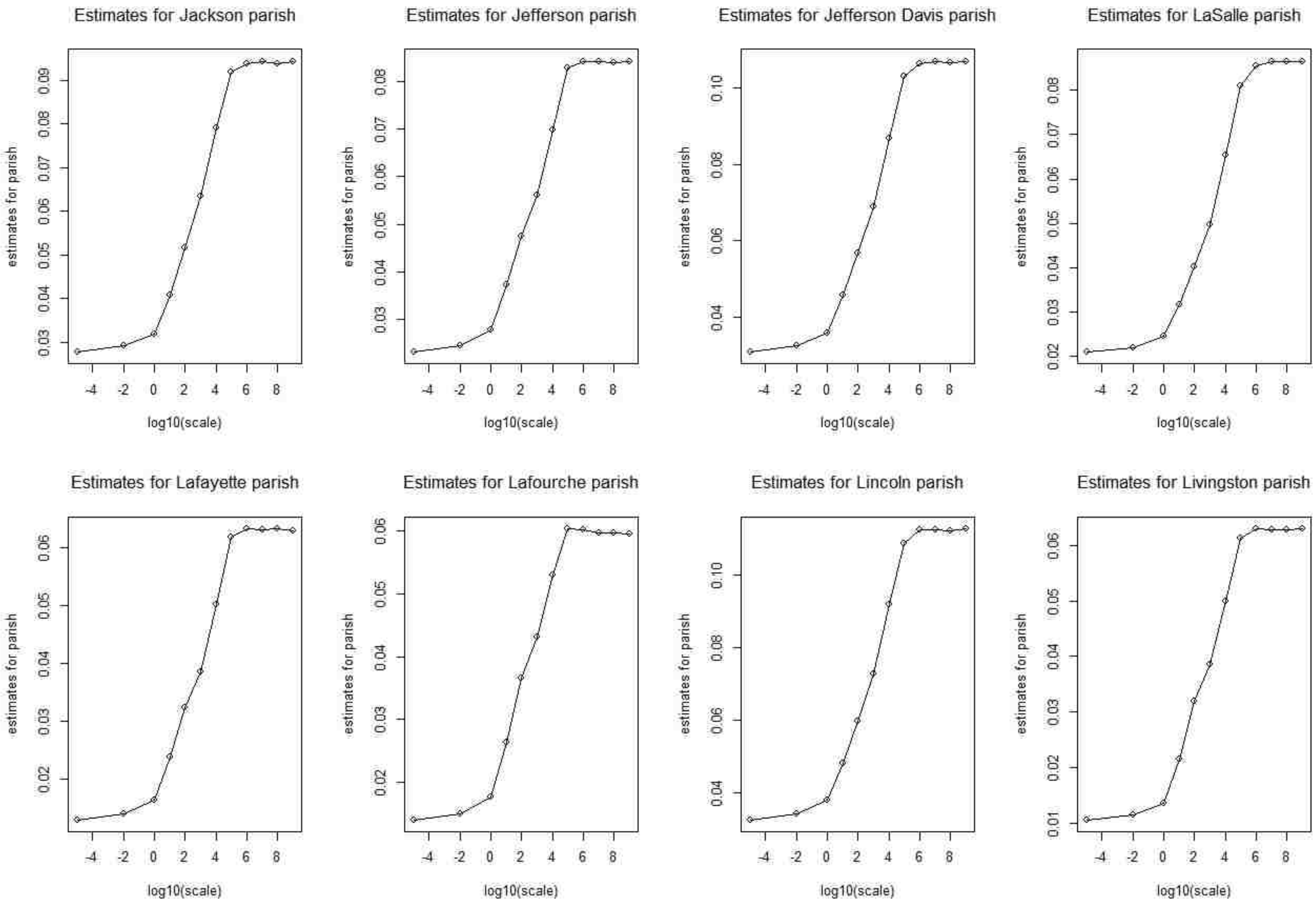

Figure 4.25: Estimates for parishes based on informative priors (children)

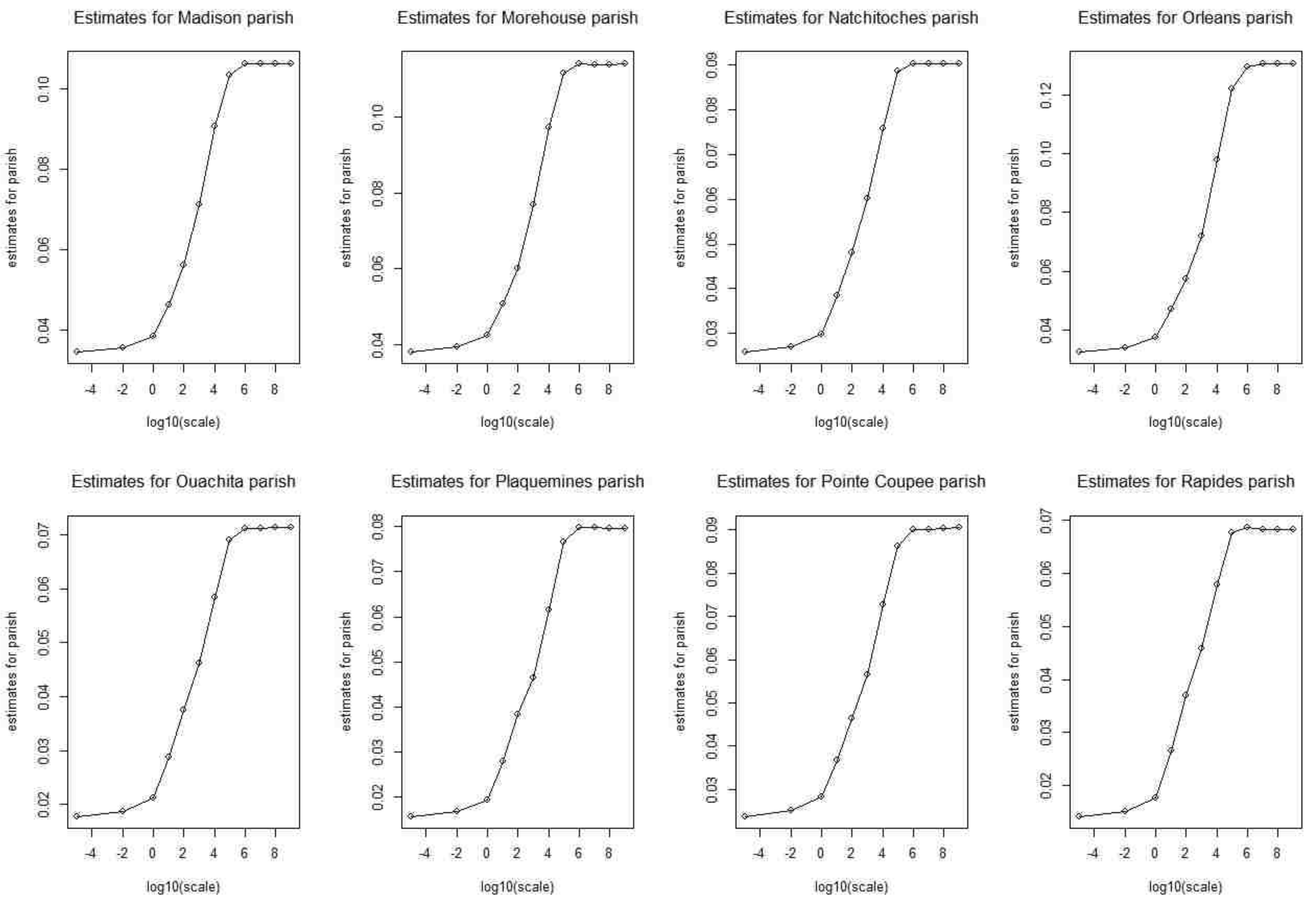

Figure 4.26: Estimates for parishes based on informative priors (children)

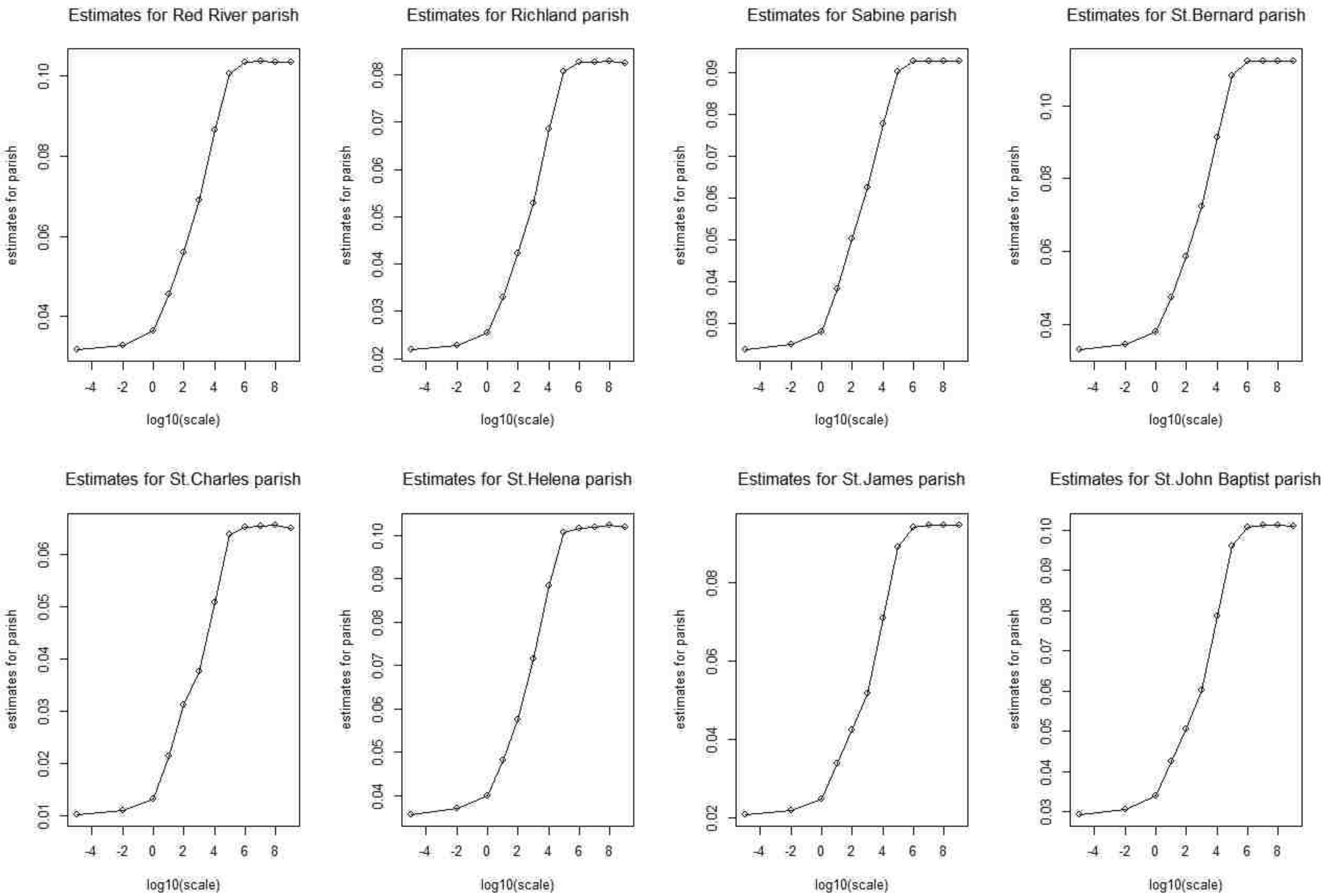

Figure 4.27: Estimates for parishes based on informative priors (children)

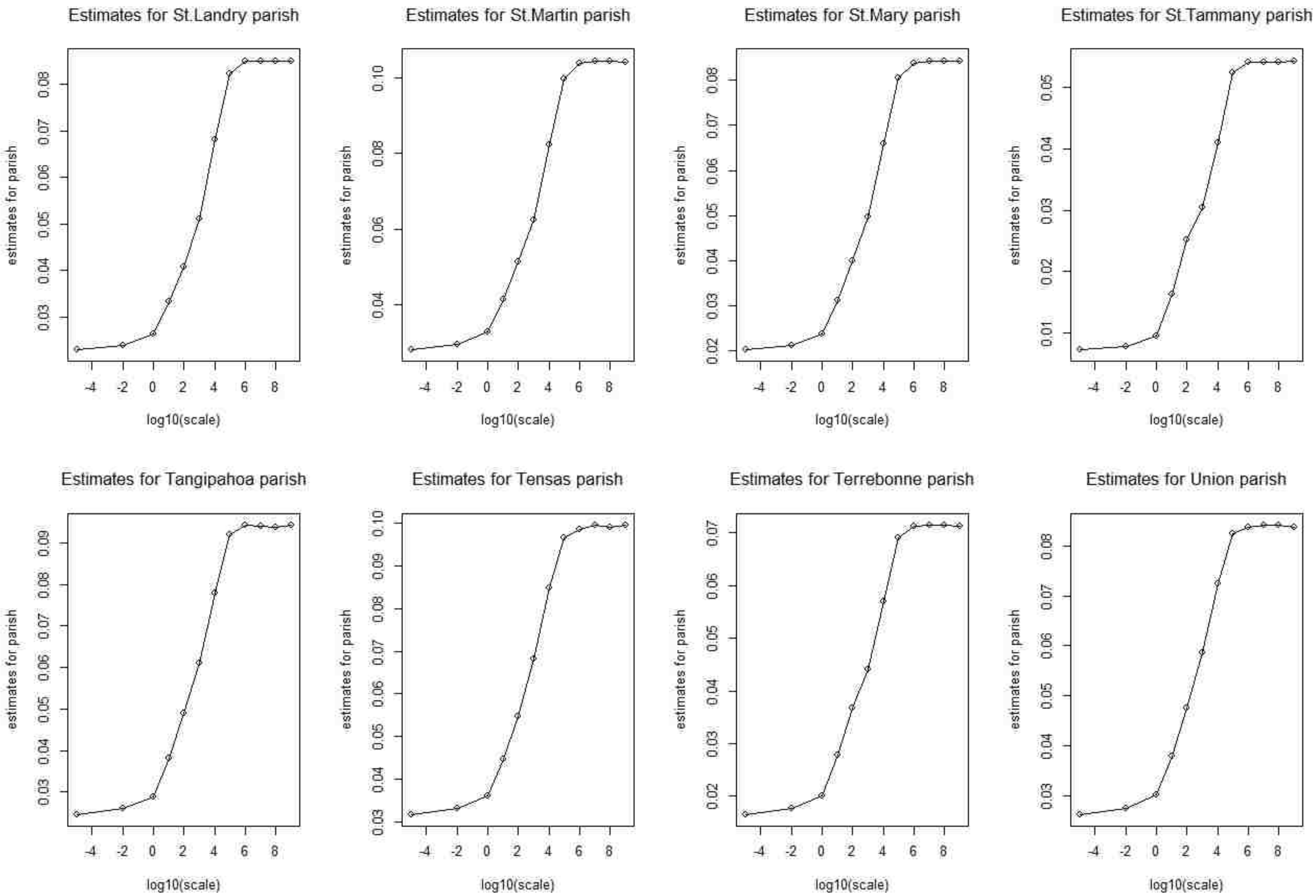

Figure 4.28: Estimates for parishes based on informative priors (children)

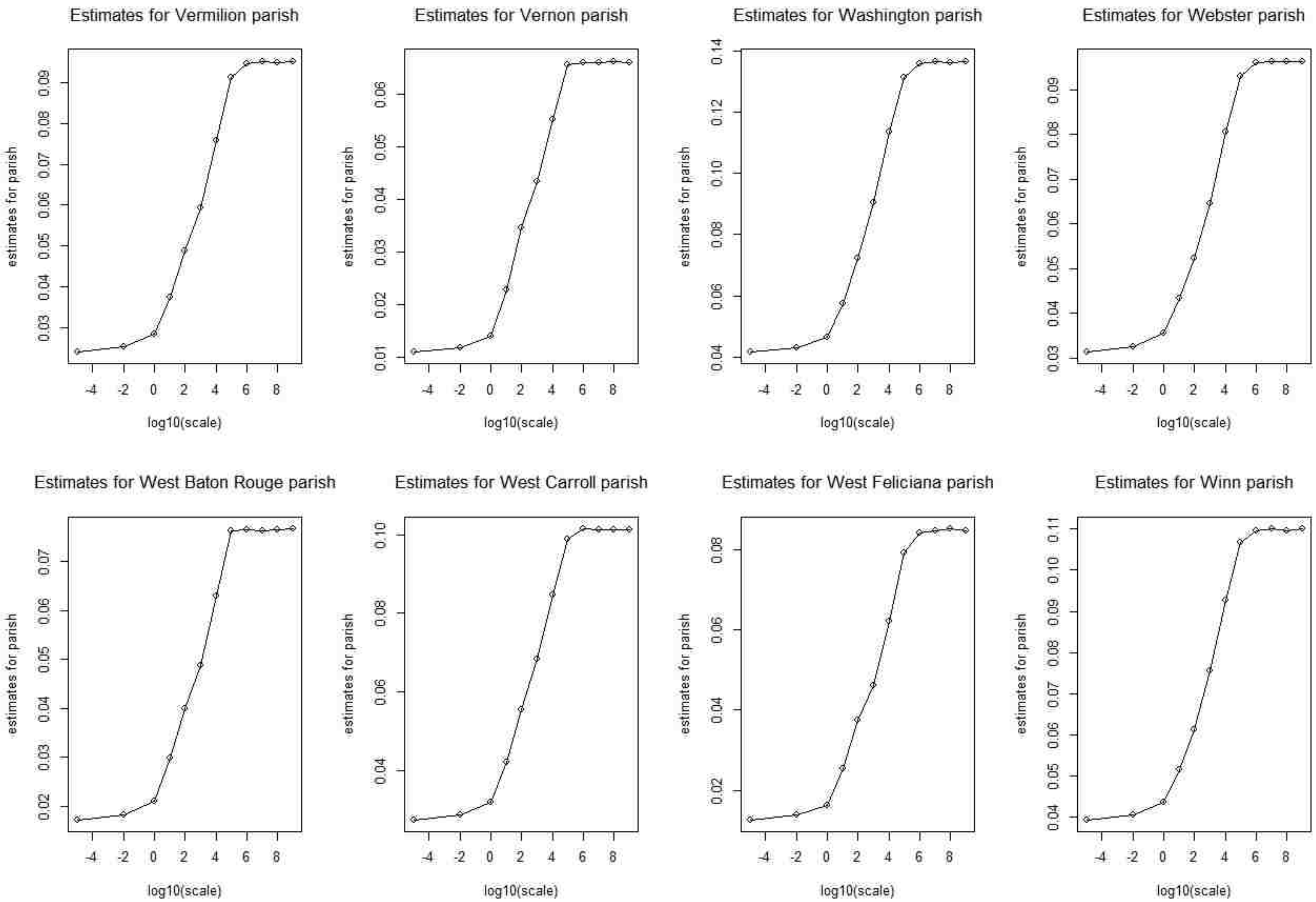

Figure 4.29: Estimates for parishes based on informative priors (children)

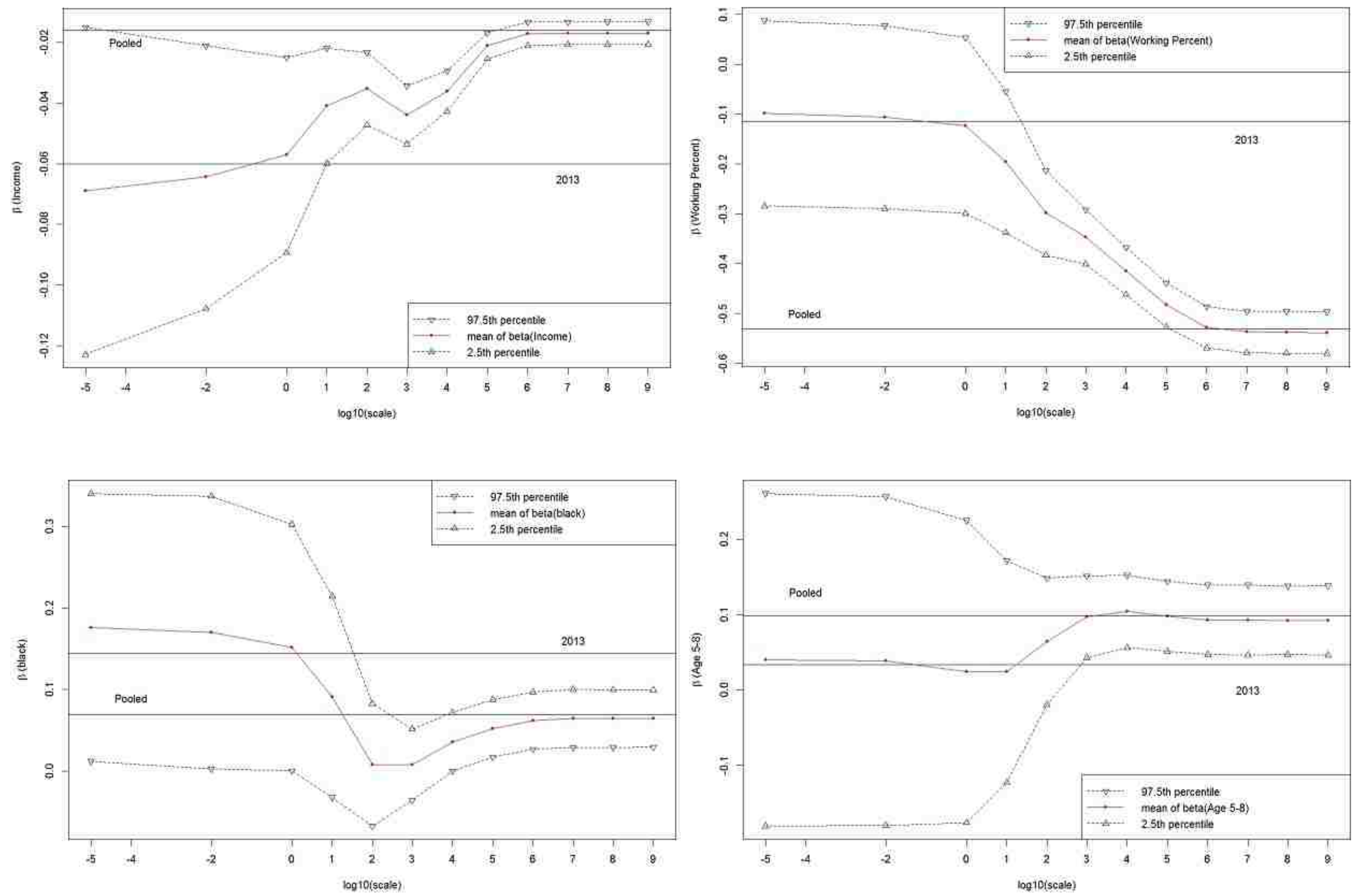

Figure 4.30: Parameters with 95% highest density region for children

Table 4.8: Posterior Means, Standard Deviation, and 95% Highest Density Region (Children)

Variables	L1	S=0.0001	S=0.01	S=1	S=10	S=1,000	S=100,000	S=10,000,000	S=1,000,000,000	L2
Constant	-1.6474	-1.8015	-1.7846	-1.7298	-1.5945	-1.2696	-1.1016	-1.056	-1.0541	-1.0171
Post. S. D	(0.1333)	(0.1832)	(0.1874)	(0.1579)	(0.0946)	(0.0328)	(0.0286)	(0.0284)	(0.0287)	(0.0314)
[H.D.R.]	[-1.9087,-1.3861]	[-2.1606,-1.4424]	[-2.1519,-1.4173]	[-2.0393,-1.4203]	[-1.7799,-1.4091]	[-1.3339,-1.2053]	[-1.1577,-1.0455]	[-1.1117,-1.0003]	[-1.1104,-0.9978]	[-1.0786,-0.9556]
Black	0.1447	0.1763	0.1699	0.1516	0.0912	0.0081	0.0523	0.0645	0.0645	0.0692
Post. S. D	(0.0762)	(0.0837)	(0.0852)	(0.077)	(0.0628)	(0.0223)	(0.018)	(0.0182)	(0.0178)	(0.0224)
[H.D.R.]	[-0.0047,0.2941]	[0.0122,0.3404]	[0.0029,0.3369]	[0.0007,0.3025]	[-0.0319,0.2143]	[-0.0356,0.0518]	[0.017,0.0876]	[0.0288,0.1002]	[0.0296,0.0994]	[0.0253,0.1131]
Income	-0.0306	-0.069	-0.0644	-0.0571	-0.0409	-0.0439	-0.0211	-0.017	-0.0169	-0.0159
Post. S. D	(0.0091)	(0.0275)	(0.0221)	(0.0164)	(0.0097)	(0.0049)	(0.0022)	(0.0019)	(0.0019)	(0.0024)
[H.D.R.]	[-0.0484,-0.0128]	[-0.1229,-0.0151]	[-0.1077,-0.0211]	[-0.0892,-0.025]	[-0.0599,-0.0219]	[-0.0535,-0.0343]	[-0.0254,-0.0168]	[-0.0207,-0.0133]	[-0.0206,-0.0132]	[-0.0206,-0.0112]
Age 0-4	0.0341	0.0402	0.0385	0.0247	0.0247	0.0973	0.098	0.093	0.0927	0.0991
Post. S. D	(0.0988)	(0.1128)	(0.1115)	(0.1025)	(0.0751)	(0.0277)	(0.0237)	(0.0238)	(0.0236)	(0.0427)
[H.D.R.]	[-0.1595,0.2277]	[-0.1809,0.2613]	[-0.18,0.257]	[-0.1762,0.2256]	[-0.1225,0.1719]	[0.043,0.1516]	[0.0515,0.1445]	[0.0464,0.1396]	[0.0464,0.139]	[0.0154,0.1828]
Age 5-9	0.1177	0.1656	0.1624	0.1413	0.1127	0.0777	0.0866	0.0856	0.0855	0.0902
Post. S. D	(0.0966)	(0.1165)	(0.1103)	(0.1016)	(0.0732)	(0.0268)	(0.0225)	(0.0222)	(0.0226)	(0.0278)
[H.D.R.]	[-0.0716,0.307]	[-0.0627,0.3939]	[-0.0538,0.3786]	[-0.0578,0.3404]	[-0.0308,0.2562]	[0.0252,0.1302]	[0.0425,0.1307]	[0.0421,0.1291]	[0.0412,0.1298]	[0.0357,0.1447]
Age 10-14	0.1529	0.2357	0.23	0.2137	0.1889	0.2297	0.2084	0.1991	0.1982	0.2115
Post. S. D	(0.0984)	(0.114)	(0.1137)	(0.0993)	(0.0737)	(0.0278)	(0.0239)	(0.0235)	(0.0238)	(0.0262)
[H.D.R.]	[-0.04,0.3458]	[0.0123,0.4591]	[0.0071,0.4529]	[0.0191,0.4083]	[0.0444,0.3334]	[0.1752,0.2842]	[0.1616,0.2552]	[0.153,0.2452]	[0.1516,0.2448]	[0.1601,0.2629]
Female	-0.0483	-0.07	-0.0668	-0.0666	-0.0726	-0.0517	-0.023	-0.0235	-0.0236	-0.0258
Post. S. D	(0.0621)	(0.068)	(0.0701)	(0.0664)	(0.0577)	(0.0203)	(0.0159)	(0.0157)	(0.0158)	(0.0271)
[H.D.R.]	[-0.17,0.0734]	[-0.2033,0.0633]	[-0.2042,0.0706]	[-0.1967,0.0635]	[-0.1857,0.0405]	[-0.0915,-0.0119]	[-0.0542,0.0082]	[-0.0543,0.0073]	[-0.0546,0.0074]	[-0.0789,0.0273]
P185	0.0653	0.173	0.1624	0.1473	0.123	0.1314	0.1096	0.0829	0.082	0.0146
Post. S. D	(0.098)	(0.1122)	(0.116)	(0.1038)	(0.0741)	(0.0258)	(0.022)	(0.0219)	(0.0219)	(0.0183)
[H.D.R.]	[-0.1268,0.2574]	[-0.0469,0.3929]	[-0.065,0.3898]	[-0.0561,0.3507]	[-0.0222,0.2682]	[0.0808,0.182]	[0.0665,0.1527]	[0.04,0.1258]	[0.0391,0.1249]	[-0.0213,0.0505]
Working Pct	-0.1146	-0.0984	-0.106	-0.1231	-0.196	-0.3466	-0.4833	-0.5371	-0.5384	-0.531
Post. S. D	(0.0936)	(0.0949)	(0.0937)	(0.09)	(0.0723)	(0.0278)	(0.0226)	(0.0212)	(0.0214)	(0.0249)
[H.D.R.]	[-0.2981,0.0689]	[-0.2844,0.0876]	[-0.2897,0.0777]	[-0.2995,0.0533]	[-0.3377,-0.0543]	[-0.4011,-0.2921]	[-0.5276,-0.439]	[-0.5787,-0.4955]	[-0.5803,-0.4965]	[-0.5798,-0.4822]

Chapter 5. Simulation and Results

In this chapter, we construct a series of simulations. Starting from a basic simulation, we explore the performances of the three methodologies as we discussed earlier in small area estimation (such as the empirical best linear unbiased predictions, hierarchical Bayes method and hierarchical Bayes method with a probit model). The simulation on the informative priors derive the impacts of varying coefficients in the cross-sectional situation.

5.1 Simulation Model Setup

In the first simulation, we consider five small areas with two independent variables. N_i denotes the population size for each small area. We generate N_i from a uniform distribution.

$$N_i \sim \text{Uniform}[50, 200], i = 1, ..., 5.$$

Define $(y_{ij}, x_{ij1}, x_{ij2})$ is the j^{th} observation in the i^{th} small areas. Particularly, we assume $x_1 \sim N(\mu = 1, \sigma = 0.5)$ and $x_2 \sim N(0.5, 0.32)$. The interested parameter y_{ij}'s are generated through the following model:

$$\textbf{Model 1}: \quad y_{ij} = 0.5x_{ij1} + x_{ij2} + v_i + e_{ij} \tag{5.1}$$

where $v_i \sim N(0, 1)$ and $e_{ij} \sim N(0, 1)$. Hence, the population mean of dependent variable $\mu_i \sim N(\mu = 1, \sigma = 0.15), i = 1, ..., 5$.

We list the direct estimates, EBLUP estimates and HB estimates in the following table, as well as the true values.

The first column in Table 5.1 represents the label of the small area; the second column

Table 5.1: Results of simulation for Model 1

Small Area	N_i	Method	True Value	Point Estimate	Standard Deviation	95% credible intervals	
1	64	Direct	2.1063	1.4719	1.3016	-1.0792	4.0231
2	72		1.4104	0.7516	1.1704	-1.5424	3.0457
3	87		1.1913	0.9572	1.2780	-1.5476	3.4620
4	108		0.8731	0.4357	1.1498	-1.8179	2.6892
5	196		1.8975	1.2516	1.1510	-1.0044	3.5075
1	64	EBLUP	2.1063	2.1001	0.1411	1.8235	2.3767
2	72		1.4104	1.1504	0.1276	0.9003	1.4004
3	87		1.1913	1.2408	0.1214	1.0029	1.4787
4	108		0.8731	0.9087	0.1122	0.6887	1.1287
5	196		1.8975	1.6881	0.0869	1.5178	1.8584
1	64	HB	2.1063	2.0201	0.1346	1.7563	2.2839
2	72		1.4104	1.2307	0.1216	0.9924	1.4690
3	87		1.1913	1.2177	0.115	0.9923	1.4431
4	108		0.8731	0.879	0.108	0.6673	1.0907
5	196		1.8975	1.7399	0.0848	1.5737	1.9061

represents the sample size of each small area; the third column indicates the estimation methods; the fourth column lists the true values of the population. The following two columns list the point estimates and standard deviation (posterior standard deviation). The last two columns list the 95% credible intervals for each estimation. Comparing with the direct estimates, the EBLUP and HB estimations provide the narrower credible intervals. However, two out of five credible intervals in the EBLUP estimations did not contains the true values. Therefore, the HB estimation performs the best among three methodologies.

Next, in consistency with our data set, we consider the binary dependent variable. Hence, we rewrite the first model as the probit model.

$$\textbf{Model 2}: \quad y^*_{ij} = 0.5 * x_{ij1} + x_{ij2} + v_i + e_{ij} \tag{5.2}$$

$$y_{ij} = \begin{cases} 1 & \text{if } y_{ij}^* > 0 \\ \\ 0 & \text{if } y_{ij}^* \leq 0 \end{cases} . \tag{5.3}$$

Again, we list the information and results for model 2 in Table 5.2. Similar to the results for model 1, the EBLUP and HB estimations provide much smaller credible intervals relative to direct estimates. Three out of five credible intervals in the EBLUP estimation did not contains the true values. Moreover, the point estimates for the first small area is greater than one, which is unrealistic. Therefore, the HB estimation performs the best among three methodologies.

Table 5.2: Results of simulation for Model 2

Small Area	N_i	Method	True Value	Point Estimate	Standard Deviation	95% credible intervals	
1	64	Direct	0.9115	0.8750	0.3333	0.2217	1.5283
2	72		0.817	0.7361	0.4438	-0.1338	1.6060
3	87		0.7775	0.8391	0.3696	0.1147	1.5635
4	108		0.7122	0.6389	0.4826	-0.3069	1.5847
5	196		0.888	0.8622	0.3463	0.1836	1.5409
1	64	EBLUP	0.9115	1.0152	0.0507	0.9158	1.1147
2	72		0.817	0.8201	0.0459	0.7301	0.9102
3	87		0.7775	0.8974	0.0436	0.8120	0.9828
4	108		0.7122	0.7622	0.0405	0.6829	0.8416
5	196		0.888	0.9796	0.0313	0.9184	1.0409
1	64	HB	0.9115	0.9151	0.028	0.8602	0.9700
2	72		0.817	0.8004	0.0275	0.7465	0.8543
3	87		0.7775	0.8239	0.0267	0.7716	0.8762
4	108		0.7122	0.7679	0.0369	0.6954	0.8404
5	196		0.888	0.9111	0.0207	0.8705	0.9517

5.2 Model setup for LHIS data set

In this section, we employed the modified the Louisiana Health Insurance Survey (LHIS) data set. As we mentioned earlier, the LHIS data set is a biannual data set which starts from 2003 and provides the most accurate and comprehensive assessment of Louisiana's uninsured populations every two years. The economic environment has changed over the past decade. In particular, in the state of Louisiana, the effects of Hurricane Katrina[1], in August 2005, were catastrophic and widespread. The LHIS data set was collected during the summer of 2005, hence, the LHIS data in survey year 2005 reflect the insurance status before the Hurricane Katrina. Therefore, we could consider that public health insurance status are quite similar in the survey year 2003 and 2005. After that, the economy in Louisiana, as well as the health insurance states are in recovery from the disaster.

Later, the United States housing bubble affected many parts of the U.S. housing market in over half of American states. The credit crisis resulting from the bursting of the housing bubble is the primary cause of the 2007-2009 recession in the United States. During the recession, job loss was more pronounced, and it was often paired with a loss of health insurance coverage. After the long and deep recession, the economic recovery began in mid-2009.

As we mentioned in the previous chapter, we are planing to combine different data sets together in order to reach an increasing sample set. In our simulation, we combine the years 2003 and 2005 together, leave year 2007 and year 2009 separate, while year 2011 and 2013 are combined as one group. In particular, except survey year 2011, we specify a unique coefficient for the independent variable "Poverty," which indicates whether the adult lives in a family below 185% of the federal poverty line[2].

Following equation (2.9), the model is set up as follows:

$$y_i^* = X_i\beta + v_i + u_i, i = 1, ..., m. \tag{5.4}$$

[1] Hurricane Katrina made landfall in Louisiana on August 29, 2005, as a Category 3 hurricane. The storm was large and had an effect on several different areas, for instance, all counties in Mississippi and Louisiana, 22 counties in Western Alabama and 11 counties in Florida.

[2] The health insurance coverage of children is different from that of adults. Due to the Children's Health Insurance Program (CHIP), the children's health insurance coverage is less sensitive to the economic environment. Therefore, in this Chapter, we only apply the simulation on adults.

where X_i is the matrix containing the individual's information, β is the specified coefficients, v_i is the area-specific effect, and $u_i \sim N(0,1)$. For the cross-sectional data, we follows the regression model (69) to specify the coefficients.

$$\begin{aligned} y_{ij} =& \sum_{k=1}^{6} \beta_{ik} x_{ik} + \sum_{k=7}^{11} \beta_{ij} * D_{year} + \sum_{k=12}^{17} \beta_{ik} x_{ik} * D_{2003} \\ &+ \sum_{k=18}^{23} \beta_{ik} x_{ik} * D_{2005} + \sum_{k=24}^{29} \beta_{ik} x_{ik} * D_{2007} \\ &+ \sum_{k=30}^{35} \beta_{ik} x_{ik} * D_{2009} + \sum_{k=36}^{41} \beta_{ik} x_{ik} * D_{2011} + v_i + e_{ij} \end{aligned} \tag{5.5}$$

where $i = 1, ..., 64,\ j = 1, ...m_i,\ D_{2003}, ..., D_{2011}$ are dummy variables that take the value one if the individual is collected in that particular year. Hence, the following table lists the values of each β_{ik}.

Table 5.3: Simulation Coefficients

Variables	2003	2005	2007	2009	2011	2013
Constant	0.4551	0.4978	0.5718	0.4926	0	-0.5544
Black	0	0	0.0615	0.0337	0	0.1892
Income	0	0	0.0089	-0.0338	0	-0.0237
Female	0	0	0.0370	-0.0717	0	-0.0226
Poverty	-0.4	-0.4	-0.4	-0.4	0	0.7796
Working Percent	0	0	-0.2230	0.1257	0	-0.0251
Age	0	0	-0.0012	-0.0038	0	-0.0140

Furthermore, for the binary dependent variable, we need to specify the observed dependent variable y_{ij} as:

$$y_{ij} = \begin{cases} 1 & \text{if } y_{ij}^* > 0 \\ 0 & \text{if } y_{ij}^* \leq 0 \end{cases} . \tag{5.6}$$

5.3 Simulation Results

In this section, we present the results from the simulation. The Figures 5.1 - 5.4 show the convergence of the estimates for variables Age, Working percent, Income and Poverty, respectively. All the estimates converge to zero with the increasing of the strength of the scale factor S, roughly after scale factor becomes 100,000. Furthermore, similar to the raw data set, the Bayesian informative prior shrink the pooled estimates to the cross-sectional estimates. As listed in Figure 5.5, the selected estimate plots for parishes show that the trend is not smooth around scale level S=100; 1,000; and 10,000. Based on the principles of the scale factor, the larger the scale factor S, the stronger the prior. The non-smooth trend may be caused by the relatively large value of variables "Age" and "Income"[3].

Practically, we could specify a particular value for scale factor S, in order to get the estimates for the health insurance coverage for each parish. Furthermore, due to the different performances of the strength of the scale factor S, we could specify a set of scale factors. For instance, we construct a more strength scale factor on variable "Age", and a strength scale factor on variable "Income", while a plain scale factor on others.

[3]During the regression, we use the adjusted household income as household income/10,000. But it is still quite large relative to other variables.

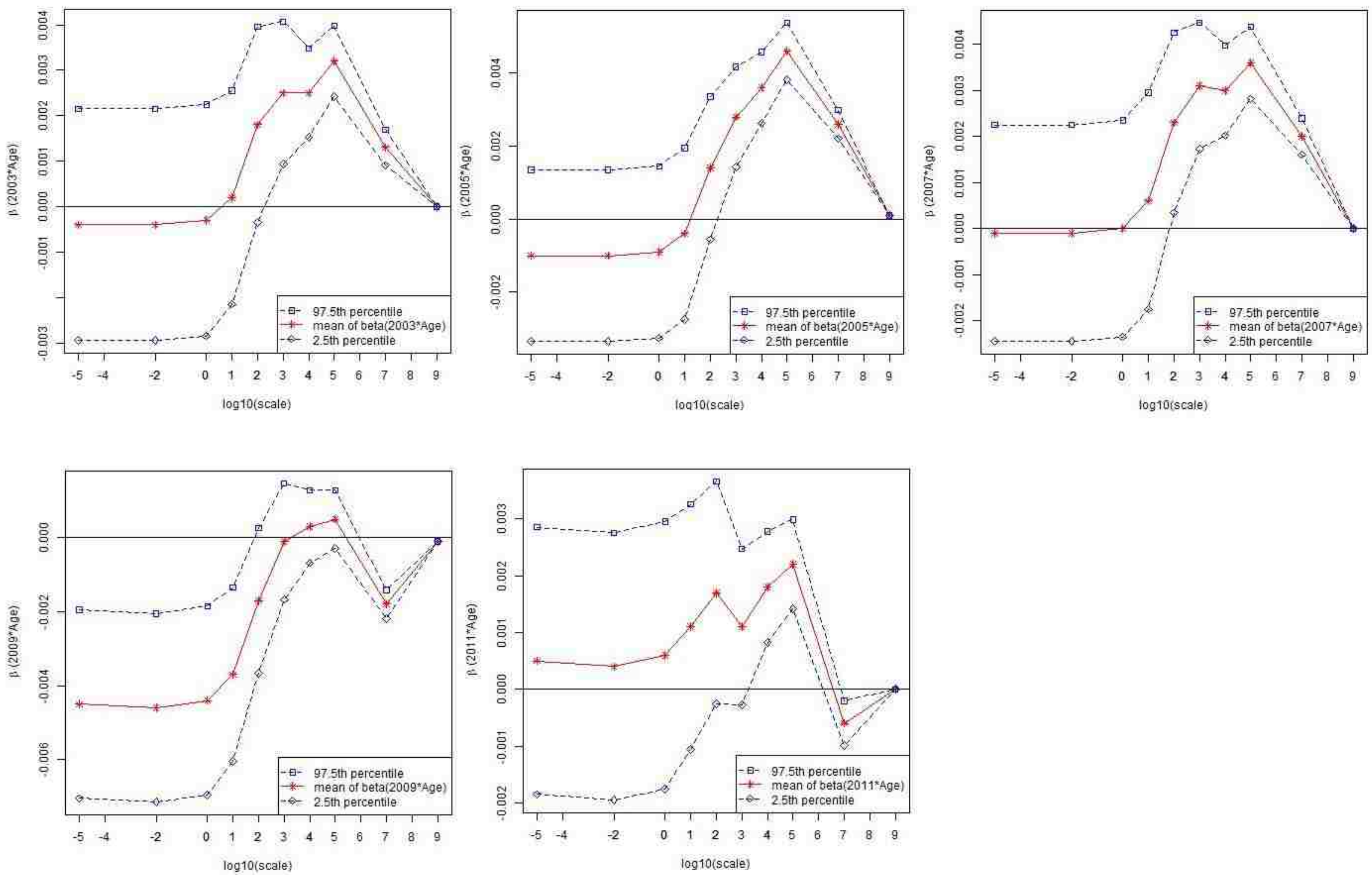

Figure 5.1: Simulation Results for β_{Age}

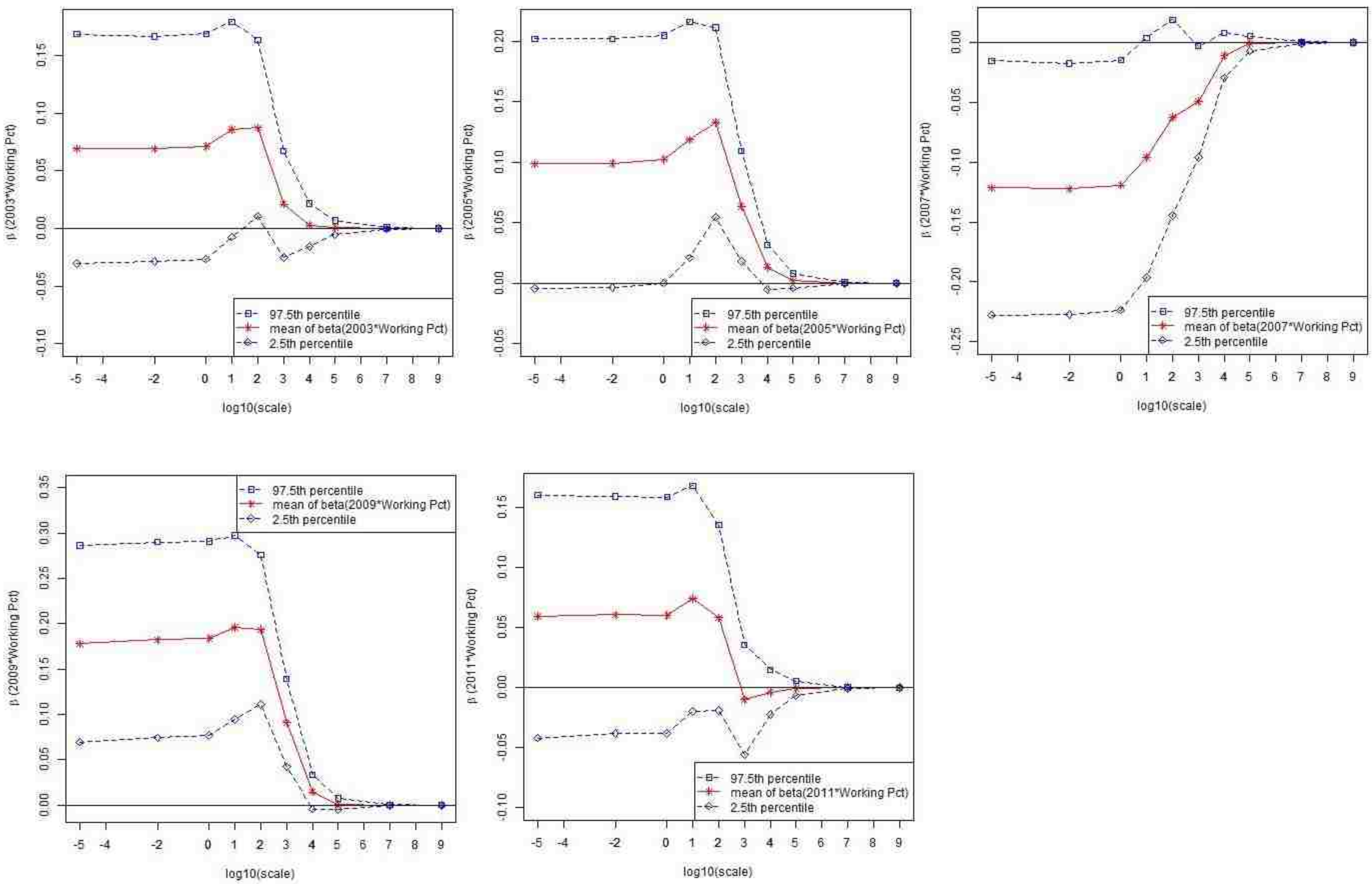

Figure 5.2: Simulation Results for $\beta_{WorkingPercent}$

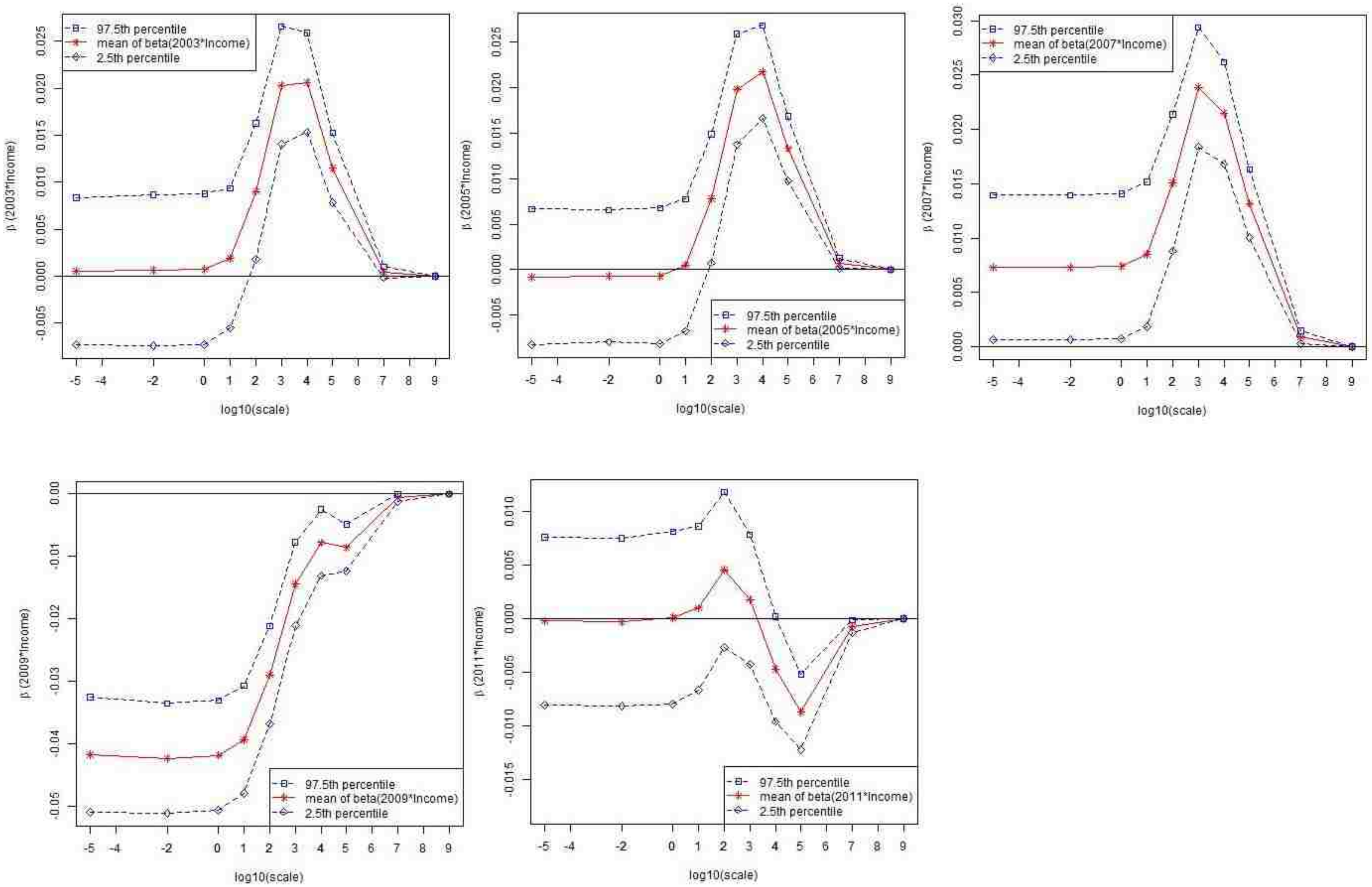

Figure 5.3: Simulation Results for β_{Income}

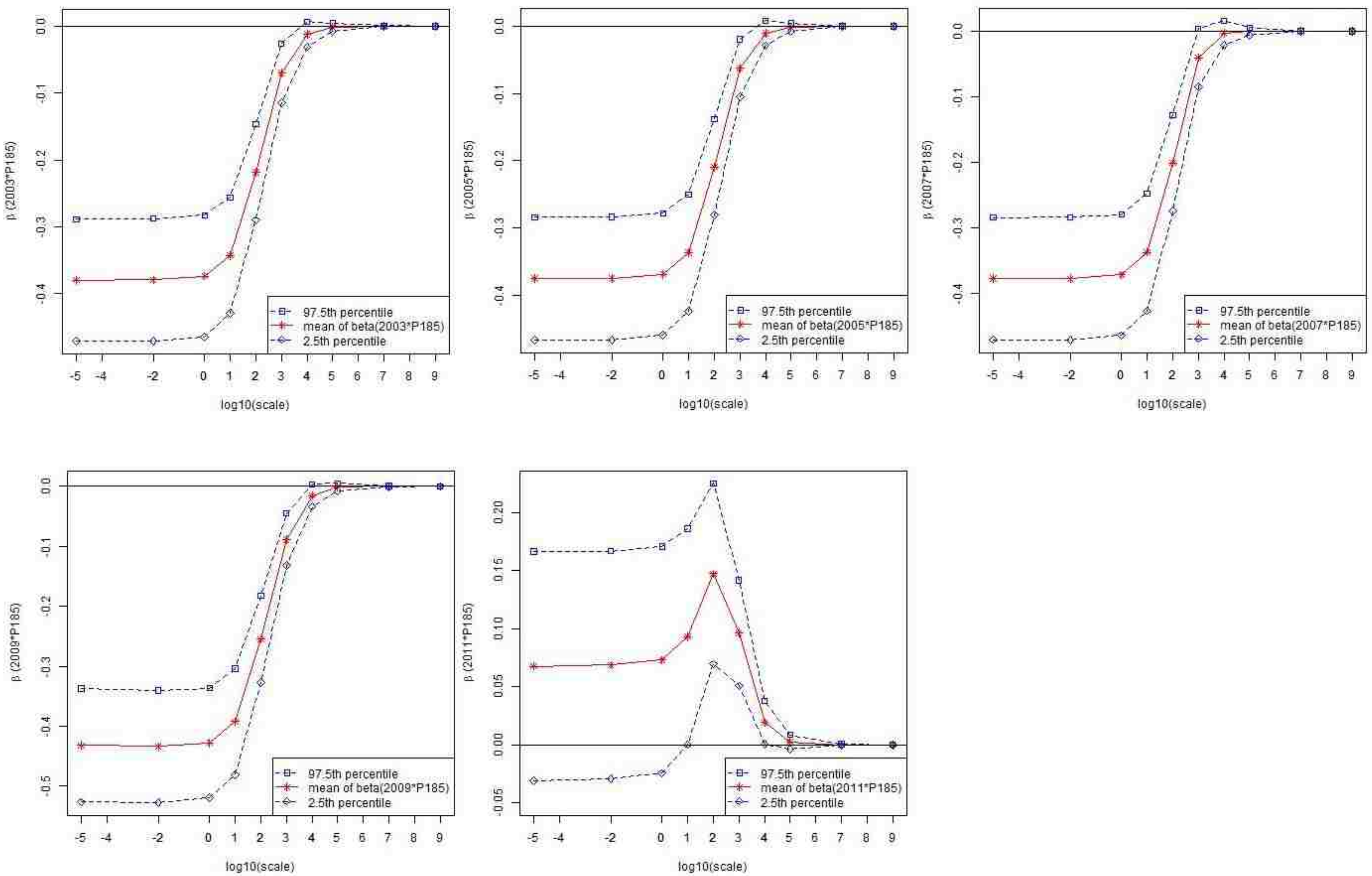

Figure 5.4: Simulation Results for $\beta_{Poverty}$

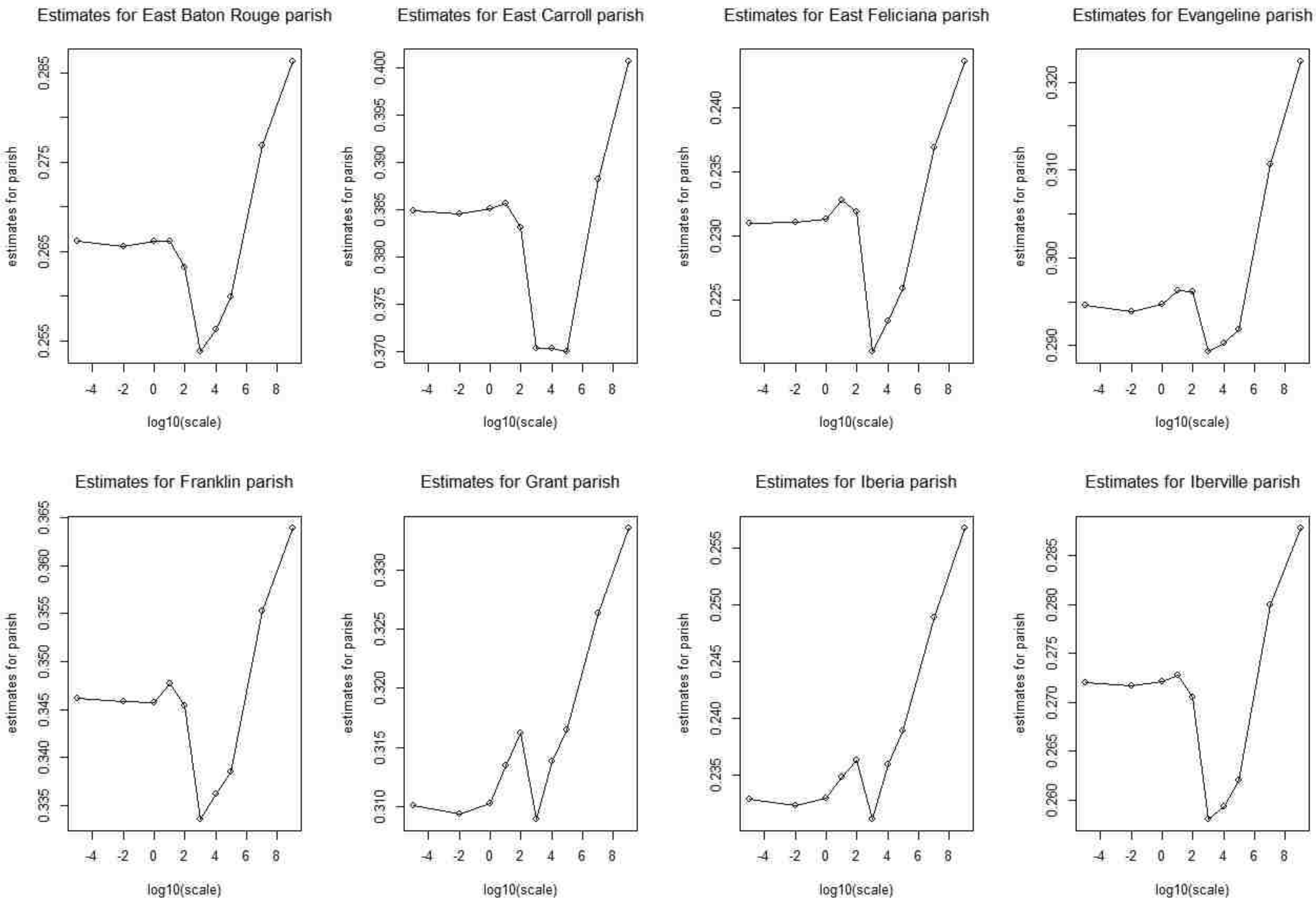

Figure 5.5: Estimates for selected parishes

Chapter 6. Summary of Conclusions

In most practical applications, sample sizes are not large enough to allow direct estimation, while sample surveys provide a cost effective way of obtaining estimates for characteristics of interest at both population level and subpopulation level. When direct estimates are not possible, one has to rely upon alternative methods that depend on the availability of population level information. Small area estimation provides the possibilities of the estimation.

In this dissertation, we explore three methodologies, such as the empirical best linear unbiased predictions, hierarchical Bayes method and hierarchical Bayes method with a probit model in small area estimation. We apply these methodologies to the Louisiana Health Insurance Survey (LHIS), and estimate the health insurance coverage for adults and children for the 64 parishes since 2003 in Louisiana. Among the three methods, the estimates are similar for adults. On the other hand, the estimates of health insurance coverage for children show that the hierarchical Bayes estimation with a probit model performs better for the binary dependent variable. The simulation results also show that direct estimators and traditional estimators will become problematic when the direct estimator is either unavailable or unreliable.

Furthermore, we also propose a Bayesian informative approach for cross-sectional data. The results show that the informative prior in essence shrinks the pooled estimation towards the cross-sectional estimation significantly and improves the performance of the estimation. The simulation results also indicate that adults' health insurance coverage has changed due to the changes in the economic environment, such as economic recession, and nature disasters, such as hurricanes, while almost no impact was seen in children's health insurance coverage due to the Louisiana Children's Health and Insurance Program.

References

[1] Baltagi, B. H., Griffn, J. M, and Xiong, W. (2008). To pool or not to pool, *The Econometrics of Panel Data*, Chapter 16, Springer.

[2] Barnes, S., Goidel, K., Terrell, D. and Virgets, S. (2013). Louisiana's Uninsured Population: A Report from the 2013 Louisiana Health Insurance Survey.

[3] Battese, G. E., Harter, R. M. and Fuller, W. A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 80:28-36.

[4] Blumberg, S., Luke, J., Ganesh, N., Davern, M., and Boudreaux, M. (2012). Wireless substitution: State-level estimates from the national health interview survey. *National Health Statistics Reports* 16, 2010-201.

[5] Brackstone, G. J. (1987). Small area data: policy issues and technical challenges. In R. Platek, J. N. K. Rao, C. E. Sarndal, and M. P. Singh, eds., *Small Area Statistics*, pp. 3-20. John Wiley & Sons, New York.

[6] Browne, W. J. and Draper, D. (2006). A comparison of Bayesian and likelhood-based methods for fitting multilevel models. *Bayesian Analysis*, 3: 473-514.

[7] Buettgens, M. and Hall, M. (2011). Who Will Be Uninsured after Health Insurance Reform? *Princeton, NJ: Robert Wood Johnson Foundation.*

[8] Bundorf, M. K. and Pauly, M. V. (2006). Is Health Insurance Affordable for the Uninsured? *Journal of Health Economics*, 25(4), 650673.

[9] Cohen, R. A. and Martinez, M. E.(2012). Health Insurance Coverage: Early Release of Estimates from the National Health Interview Survey. *Center for Health Statistics.*

[10] Datta, G. S. and Ghosh, M. (1991). Bayesian prediction in linear models: Applications to small area estimation. *The Annals of Statistics*, 19:1748-1770.

[11] Datta, G. S., Ghosh, M., Nangia, N. and Natarajan, K. (1996). Estimation of median income of four-person families: A Bayesian approach. *Bayesian Analysis in Statistics and Econometrics, 129-140.* Wiley. eds. D. A. Berry and K. M. Chaloner and J. K. Geweke.

[12] Datta, G. S., Ghosh, M., and Waller, L. (2000). Hierarchical and empirical Bayes methods for environmental risk assessment, *Handbook of Statistics, Bioenvironmental and Public Health Statistics (18 ed.), 223- 245. North-Holland. eds. P. K. Sen and C. R. Rao.*

[13] Datta, G. S., Lahiri, P. and Maiti, T. (2002). Empirical bayes estimation of median income of four-person families by state using time series and cross-sectional data. *Journal of Statistical Planning and Inference* 102, 83-97.

[14] Efron, B. and Morris, C. N. (1975). Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association*, 70:311-319.

[15] Ericksen, E. P. (1974). A Regression Method for Estimation Population Changes for Local Areas. *Journal of the American Statistical Association*, 69:867-875.

[16] Fay, R. E. and Herriot, R. A. (1979). Estimates of income for small places: an application of James-Stein procedure to census data. *Journal of the American Statistical Association*, 74: 269-277.

[17] Fronstin, P. (2011). Sources of Health Insurance and Characteristics of the Uninsured: Analysis of the March 2012 Current Population Survey. *EBRI Issue Brief*, 376.

[18] Fuller, W. A. and Harter, R. M. (1987). The Multivariate Components of Variance Model for Small Area Estimation. *Small Area Statistics*, An International Symposium. pp. 103-123. John Wiley & Sons, New York.

[19] Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85: 398-409.

[20] Ghosh, M. and Lahiri, P. (1992). A hierarchical Bayes approach to small area estimation with auxiliary information (with discussion). *Bayesian Analysis in Statistics and Econometrics*, pp. 107-125. Springer-Verlag, Berlin.

[21] Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1995). Markov Chain Monte Carlo in Practice. *Chapman & Hall/CRC, Boca Raton, Florida.*

[22] Ghangurde, D. D. and Singh, M. P. (1977). Synthetic estimators in periodic housholds surves. *Survey Methodology 3* 152-181.

[23] Goidel, K., Barnes, S. and Terrell, D. (2012). Louisiana's Uninsured Population: A Report from the 2011 Louisiana Health Insurance Survey.

[24] Gomez-Rubio, V., Best, N., Richardson, S. and Li, G. (2008). Bayesian statistics for small area estimation. *United Kingdom: Office for National Statistics.*

[25] Gonzales, M. E. (1973). Use and evaluation of synthetic estimates. *Procedings of the Social Statistics Section*, American Statistical Association, 33-36.

[26] Gonzalez, M. and Waksberg, J. (1973). Estimation of the Error of Synthetic Estimates. Paper presented at the first meeting of the International Association of Survey Statisticians, Vienna, Austria.

[27] Harville, D. A. (1991). Commment on, "That BLUP is a good thing: The estimation of random effects," by Robinson, G. K. *Statistical Science*, 6: 35-39.

[28] Hill, B. M. (1965). Inference about Variance Components in the One-Way Model. *Journal of the American Statistical Association*, 60: 806-825.

[29] Hobert, J. P. and Casella, G. (1996). The Effect of Improper Priors on Gibbs Sampling in Hierarchical Linear Mixed Models. *Journal of the American Statistical Association*, 91: 1461-1473.

[30] Holahan, J. (2010). The 2007-09 recession and health insurance coverage. *Health Affairs*, 30(1).

[31] Holt, D., Smith, T.M.F. and Tomberlin, T.J. (1979). A Model-Based Approach to Estimation for Small Subgroups of a Population. *Journal of the American Statistical Association*, 74: 405-410.

[32] Jiang, J. (2006b). Linear and generalized linear mixed models and their applications. Springer-Verlag, working monograph.

[33] Jiang, J., and Lahiri, P. (2006). Mixed model prediction and small area estimation (with discussions), *Test*, 15:1-96.

[34] Koop, G. (2003). *Bayesian Econometrics*, John Wiley & Sons, New York.

[35] Marker, D. A. (1983). Organization of small area estimators. *Journal of the American Statistical Association*, 409-414.

[36] Marker, D. A. (1999). Organization of small area estimators using a generalized linear regression framework. *Journal of Official Statistics*, 15:1-24.

[37] Laake, P. (1979). A Prediction Approach to Subdomain Estimation in Finite Populations. *Journal of the American Statistical Association*, 74: 355-358.

[38] Marshall, R. J. (1991). A review of methods for the statistical analysis of spatial patterns of disease. *Journal of the Royal Statistical Society*. Series A, 154:421-441.

[39] McCullagh, P. and Zidek, J.(1987). Regression methods and performance criteria for small area population estimation. In R. Platek, J. N. K. Rao, C. E. Sarndal, and M. P. Singh, eds., *Small Area Statistics*, pp. 62-74. John Wiley & Sons, New York.

[40] Morris, C. (1983). Parametric empirical Bayes inference: Theory and applications (with discussions). *Journal of the American Statistical Association*, 78: 47-65.

[41] Morrison, P. (1971). Demographic Information for Cities: A Manual for Estimating and Projecting Local Population Characteristics. Report R-618, Santa Monica: The Rand Corporation.

[42] Mukhopadhyay, P. (1998). Small Area Estimation in Survey Sampling. Narosa Publishing House, New Delhi.

[43] Natarajan, R. and McCulloch, C. E. (1995). A Note on the Existence of the Pisterior Distribution for a Class of Mixed Models for Binomial Responses. *Biometrika*, 82: 639-643. Narosa Publishing House, New Delhi.

[44] National Center for Health Statistics (1968). Synthetic State Estimates of Disability (*PHS Publication No.1759*). Washington, D.C.:U.S. Goverment Printing Office.

[45] National Institute on Drug Abuse (1979). Synthetic Estimates for Small Areas. (*NIDA Research Monograph 24*). Washington, D.C.:U.S. Goverment Printing Office.

[46] National Research Council (1980). Panel on Small Area Estimates of Population and Income. *Estimation Population and Income of Small Areas*. National Academy Press, Washington, DC.

[47] Pfeffermann, D. (2002). Small area estimation – new developments and directions. *International Statistical Review*, 70:125-C143.

[48] Prasad, N. G. N. and Rao, J. N. K. (1990). The estimation of the mean squared error of small area estimators. *Journal of the American Statistical Association*, 85:163-171.

[49] Prasad, N. G. N. and Rao, J. N. K. (1999). On Robust Small Area Estimation Using a Simple Random Effects Model. *Survey Methodology*, 25: 67-72.

[50] Purcell, N. J. and Kish, L. (1979). Estimation of small domain. *Biometrics*, 35:365-384.

[51] Rao, J. N. K. (2003). Small area estimation. Hoboken, NJ: John Wiley & Sons, Inc.

[52] Sarndal, C. E. (1984). Design-Constent versus Model-Dependent Estimation for Small Domains. *Journal of American Statistical Association*, 79:624-631.

[53] Shaible, W. L. (1979). A Composite Estimator for Small Area Statistics, in Synthetic Estimates for Small Areas. (NIDA Research Monograph 24), Washington, D.C.:U.S. Goverment Printing Office.

[54] Smith, T. M. K. (1983). On the Validity of Inferences from Non-Random Samples, *Journal of the Royal Statistical Society*, A(146):394-403.

[55] Suciu, G., Hoshaw-Woodard, S., Elliott, M. and Doss H. (2001). Uninsured Estimates by County: A Review of Options and Issues. *working paper.*

[56] Woodruff, R.S. (1966). Use of a Regression Techinique to Produce Area Breackdowns of the monthly National Estimates of Retal Trade. *Journal of American Statistical Association*, 79:496-504.

[57] You, Y. and Rao, J. N. K. (2003). Pseudo hierarchical Bayes small area estimation combining unit level models and survey weights. *Journal of Statistical Planning and Inference* 111, 197-208.

[58] Zidek, J. V. (1982). A review of methods for estimating the populations of local areas. Technique Report 82-4, University of British Columbia, Vancouver.

Appendix

In the appendix, we include the number of individuals from the Louisiana Health Insurance Survey for reference.

Table A1: Number of individuals for each survey year (Children)

parish_name	Year						Overall
	2003	2005	2007	2009	2011	2013	
Acadia	46	152	130	131	125	79	663
Allen	69	99	68	98	114	48	496
Ascension	149	53	113	138	155	119	727
Assumption	216	69	51	92	69	47	544
Avoyelles	143	110	90	121	90	87	641
Beauregard	133	120	129	90	141	85	698
Bienville	13	46	62	79	54	39	293
Bossier	11	119	119	178	188	126	741
Caddo	234	233	331	310	276	218	1602
Calcasieu	495	354	301	456	357	211	2174
Caldwell	9	37	117	67	43	33	306
Cameron	2	56	36	65	41	30	230
Catahoula	54	45	47	42	47	51	286
Claiborne	28	44	52	56	71	34	285
Concordia	95	65	88	92	55	62	457
DeSoto	31	53	54	76	52	53	319
East Baton Rouge	976	629	568	504	498	290	3465
East Carroll	30	29	126	25	71	21	302
East Feliciana	67	14	116	76	77	47	397
Evangeline	240	124	97	99	68	42	670
Franklin	15	64	74	48	44	28	273

Continued on next page

Table A1 – *Continued from previous page*

parish_name	Year						Overall
	2003	2005	2007	2009	2011	2013	
Grant	34	78	37	70	86	43	348
Iberia	164	164	112	169	153	108	870
Iberville	158	39	76	78	71	58	480
Jackson	6	70	60	70	48	38	292
Jefferson	71	353	328	503	343	231	1829
Jefferson Davis	187	91	93	93	102	59	625
LaSalle	22	92	70	93	72	173	522
Lafayette	173	284	260	338	297	151	1503
Lafourche	157	187	141	173	210	118	986
Lincoln	24	80	73	76	87	69	409
Livingston	166	211	222	250	281	181	1311
Madison	39	54	50	44	39	25	251
Morehouse	123	40	76	80	84	41	444
Natchitoches	90	90	83	115	99	46	523
Orleans	1418	375	516	315	271	208	3103
Ouachita	166	199	208	168	200	193	1134
Plaquemines	43	53	154	73	46	37	406
Pointe Coupee	51	22	81	48	56	40	298
Rapides	78	261	181	271	322	166	1279
Red River	31	43	74	52	74	50	324
Richland	140	62	73	69	51	39	434
Sabine	13	87	62	55	89	40	346
St. Bernard	113	75	60	42	48	32	370
St. Charles	88	87	92	132	148	63	610
St. Helena	13	17	157	34	41	38	300
St. James	167	13	56	28	50	49	363
St. John Baptist	152	75	143	92	75	65	602
St. Landry	523	152	133	149	133	113	1203
St. Martin	286	142	74	125	113	85	825

Continued on next page

Table A1 – *Continued from previous page*

parish_name	Year						Overall
	2003	2005	2007	2009	2011	2013	
St. Mary	357	84	150	124	89	79	883
St. Tammany	259	304	345	450	452	228	2038
Tangipahoa	252	192	166	277	204	115	1206
Tensas	1	10	64	16	37	30	158
Terrebonne	217	216	193	216	178	122	1142
Union	14	53	50	62	47	29	255
Vermilion	97	153	128	110	115	83	686
Vernon	40	149	76	131	218	83	697
Washington	120	117	136	134	89	42	638
Webster	100	111	83	110	82	66	552
West Baton Rouge	0	76	35	101	80	43	335
West Carroll	10	50	67	33	47	53	260
West Feliciana	34	6	168	53	44	45	350
Winn	91	45	87	56	54	60	393
Total	9344	7577	8262	8521	8061	5387	47152

Table A2: Number of individuals for each survey year (Adults)

parish_name	Year						Overall
	2003	2005	2007	2009	2011	2013	
Acadia	59	385	248	301	306	214	1513
Allen	131	188	180	204	217	157	1077
Ascension	235	152	205	220	434	301	1547
Assumption	342	178	118	216	164	131	1149
Avoyelles	309	207	208	269	264	198	1455
Beauregard	213	288	250	255	321	194	1521
Bienville	23	106	156	166	117	118	686
Bossier	28	282	297	335	465	346	1753

Continued on next page

Table A2 – *Continued from previous page*

parish_name	Year						Overall
	2003	2005	2007	2009	2011	2013	
Caddo	399	671	669	643	754	534	3670
Calcasieu	922	869	717	944	967	617	5036
Caldwell	14	131	270	183	142	131	871
Cameron	4	125	112	120	131	102	594
Catahoula	97	89	86	141	143	112	668
Claiborne	53	107	163	168	158	134	783
Concordia	174	123	185	185	173	127	967
DeSoto	50	139	116	177	148	125	755
East Baton Rouge	1752	1465	1178	1095	1276	896	7662
East Carroll	42	42	188	76	91	85	524
East Feliciana	155	38	240	185	181	114	913
Evangeline	465	226	143	214	202	146	1396
Franklin	23	147	172	154	166	107	769
Grant	52	127	100	175	196	137	787
Iberia	225	405	195	357	374	233	1789
Iberville	281	128	145	205	180	139	1078
Jackson	24	166	169	145	154	135	793
Jefferson	134	906	694	1308	1004	646	4692
Jefferson Davis	373	230	189	234	225	167	1418
LaSalle	38	154	161	237	189	463	1242
Lafayette	318	621	516	645	745	490	3335
Lafourche	318	395	301	421	524	368	2327
Lincoln	70	206	183	193	235	202	1089
Livingston	211	420	379	482	609	365	2466
Madison	84	119	126	129	99	103	660
Morehouse	209	130	165	153	165	145	967
Natchitoches	133	209	192	237	232	124	1127
Orleans	2227	1037	1355	840	687	542	6688
Ouachita	250	399	459	304	490	454	2356

Continued on next page

Table A2 – *Continued from previous page*

parish_name	Year						Overall
	2003	2005	2007	2009	2011	2013	
Plaquemines	69	115	296	137	124	97	838
Pointe Coupee	79	49	198	127	138	122	713
Rapides	139	603	356	582	634	491	2805
Red River	58	122	163	130	179	109	761
Richland	188	124	154	193	126	123	908
Sabine	45	190	153	145	182	141	856
St. Bernard	195	199	107	122	137	111	871
St. Charles	149	176	243	305	390	206	1469
St. Helena	12	47	419	95	125	89	787
St. James	246	47	131	89	134	134	781
St. John Baptist	307	151	329	197	227	136	1347
St. Landry	848	350	280	317	302	254	2351
St. Martin	540	319	180	301	284	220	1844
St. Mary	638	199	246	250	215	202	1750
St. Tammany	374	757	735	948	1051	602	4467
Tangipahoa	336	511	304	533	564	329	2577
Tensas	10	47	154	66	83	100	460
Terrebonne	380	474	372	494	492	300	2512
Union	29	119	143	153	153	130	727
Vermilion	109	324	263	270	316	199	1481
Vernon	43	340	193	232	396	247	1451
Washington	181	255	261	305	235	142	1379
Webster	170	289	175	263	232	161	1290
West Baton Rouge	0	164	81	219	205	142	811
West Carroll	18	110	172	112	125	124	661
West Feliciana	69	29	380	122	122	97	819
Winn	178	123	238	169	150	138	996
Total	15847	17843	17956	19192	20249	14748	105835

Vita

Zhengjia Sun was born in Changchun, Jilin Province, China. She earned her Bachelor of Science in Applied Mathematics from Dalian University of Technology in China in 2005 and her Master of Science in Statistics from University of South Carolina in the U.S. in 2008. She joined Louisiana State University in 2008 and earned her Master of Science in Economics in 2010.