# BIOS 8372 Final Project: Bayesian spatial models with Asthma ED visits in California counties

Ya-Chen Lin

# Introduction

Asthma is a persevering and chronic disease. In California, approximately 13.8% of adults reported that they had ever been diagnosed with asthma (lifetime asthma) and 8.1% said they still have asthma (current asthma). The prevalence of the Asthma for both groups has not changed significantly since 2001. The Asthma patients usually followed the asthma management plan following Global Initiative for Asthma (GINA) report. However, it is reported that only 38% of children with asthma have a recent asthma management plan. This will most likely lead to Asthma exacerbation. One measurement to measure asthma exacerbation is the ED visit for Asthma patients. For this study, we are interested in learning the Asthma patients ED visits rate in California counties between 2012 and 2017. To further investigate our question of interest, we conducted county-level and zipcode-level analysis. In the county-level analysis, we used the total ED visits in each county as outcome. In this level, we compared different models with and without considering spatial correlation. We assessed the convergence and conducted model selection. We then applied the selected model with the zipcode level analysis. In this level, we will consider the ED visits at the zipcode level within each county. We want to compare the results between county and zipcode-level analysis and further compared the result with frequentist approach where the incorporation of spatial information is currently not feasible for generalized linear model.

# Method

## Data source

The dataset was published by California Health and Human Services containing zipcodes level Asthma population information in each California county for 2012 and 2017. The number of Asthma ED visits, stratified by age (Adult and Children groups) is reported at the zipcode level. There's additional column mapping each zipcode to the county. There are 55 counties and 1030 zipcodes. In this study, we focused on the Adult population.

## Proposed models

Since the ED visits are count-based data, Poisson model was fitted for both level analyses. For the spatial correlation, we implemented the power exponential covariance structure.

### County level

The outcome of interest in the level analysis will be the total number of ED counts in each county. Here are some notations:

- Denote outcome as $y_k$ where k = 1,..., $N$ counties
- $b1_k$ indicates the log counts for each county at baseline.
- $b3_k$ indicates the additional log counts at 2017 for each county comparing to 2012. The exponentiated values will be the rate ratio between 2017 and 2012 for county k and this is our variable of interest.

The outcome y follows Poisson distribution. We proposed to compare the models with and without spatial correlation. All the candidate models share the same likelihood:

$$y_{k,t} \sim \text{Pois}(\lambda_{k,t})$$

$$\log(\lambda_{k,t}) = b1_k \text{I}(\text{county} = \text{k}) + b3_k \text{I}(\text{t} = 2017)$$

The candidate models contains different prior assumptions on $b1_k$ and $b3_k$ and can be categorized into models with and without spatial correlation.

No spatial correlation model (No spatial):

$b1_k \sim N(\theta, \tau_{1,k})$
$b3_k \sim N(\theta_2, \tau_{2,k})$
$\tau_{1,k} = 1/\sigma^2_{\theta_{1,k}}$ ; $\tau_{2,k} = 1/\sigma^2_{\theta_{2,k}}$
$\theta \sim N(0, 0.000001)$
$\theta_2 \sim N(0, 0.000001)$
$\sigma^2_{\theta_{1,k}} \sim$ dunif(0,1000)
$\sigma^2_{\theta_{2,k}} \sim$ dunif(0,1000)

Spatial correlation model 1 (Spatial 1):

$b1_k \sim N(\theta, \Sigma_{\phi_1,\tau_1,\kappa_1})$
$\Sigma_{\phi_1,\tau_1} = 1/\tau_1$Exp(-$\phi_1\kappa_1$ dist)
$b3_k \sim N(\theta_2, \Sigma_{\phi_2,\tau_2,\kappa_2})$
$\Sigma_{\phi_2,\tau_2} = 1/\tau_2$Exp(-$\phi_2\kappa_2$ dist)
$\theta \sim N(0, 0.000001)$
$\theta_2 \sim N(0, 0.000001)$
$\phi_1 \sim$ dunif(0.05,50)
$\phi_2 \sim$ dunif(0.05,50)
$\tau_1 \sim$ Gamma(0.001,0.001)
$\tau_2 \sim$ Gamma(0.001,0.001)
$\kappa_1 \sim$ dunif(0.05,1.95)
$\kappa_2 \sim$ dunif(0.05,1.95)

Spatial correlation model 2 (Spatial 2):

$b1_k \sim N(\theta, \Sigma_{\phi_1,\tau,\kappa_1})$
$\Sigma_{\phi_1,\tau_1} = 1/\tau$Exp(-$\phi_1\kappa_1$ dist)
$b3_k \sim N(\theta_2, \tau_{2,k})$
$\theta \sim N(0, 0.000001)$
$\theta_2 \sim N(0, 0.000001)$
$\tau_{2,k} = 1/\sigma^2_{\theta_{2.k}}$
$\phi_1 \sim$ dunif(0.05,50)
$\tau \sim$ Gamma(0.001,0.001)
$\kappa_1 \sim$ dunif(0.05,1.95)
$\sigma^2_{\theta_{2,k}} \sim$ dunif(0,1000)

**Zipcode level**

The outcome of interest in the level analysis will be the ED count in each zipcode. Each zipcode corresponds to 1 county.

- Denote county as k where k = 1,...,N county
- Denote j as zipcode in each county where j = 1,....,$N_k$. $N_k$ is the number of zipcodes in each county
- $y_{k,j,t}$ denotes the ED county at $j^{th}$ zipcode in county k at year t
- $b1_k$ indicates the log counts for each county at baseline.
- $b3_k$ indicates the additional log counts at 2017 for each county comparing to 2012. The exponentiated values will be the rate ratio between 2017 and 2012 for county k and this is our variable of interest.

$$y_{k,j,t} \sim \text{Pois}(\lambda_{j,k,t})$$

$$\log(\lambda_{j,k,t}) = b1_k \text{I(county = k)} + b3_k \text{I(t = 2017)}$$

The priors and hyper-priors are same as Spatial 2.

Originally, the goal was to fit b1 and b3 for each zipcode and assign higher level hierarchical with region effects (the complete original model plan is summarized in the appendix.) However, The model fails to converge with more than 2000 parameters. This simplified model treats the zipcodes in each county as repeated measurements.

## Analysis

### County level

The priors in all models were selected to be non-informative first. All models were run with 3 chains, each with 5000 iterations (first 2500 discarded) in openbugs. The initial values are described as the following:

- $\theta$ is a random draw from N(2,2)
- $\theta_2$ is a random draw from N(0,2)
- All $\sigma^2$ are random draws from uniform(0, 10)
- All $\phi = \kappa = \tau = 1$

**Model convergence and selection**

The model selection process was conducted in this level between No spatial, Spatial 1 and Spatial 2 model. The DIC criteria will be used for model selection. $b3_k$ is of interest to obtain the rate ratio between year 2012 and 2017. The convergence plots for $b3_k$ of the selected model are presented while the convergence for other 2 models were described in the main text. Our main goal is to see if implementing spatial correlation will improve the model. The posterior predictive checks for Spatial 1 and Spatial 2 were conducted.

**Priors sensitivity analysis**

Some sensitivity analyses of priors were carried out on $\theta$, $\theta_2$ and $\tau$. For $\tau$ distribution, different priors including uniform, gamma (0.001,0.001) and gamma (1,1) were implemented to compare the influence on $\tau$ distribution. For $\theta$ distribution, we will compare the original non-informative priors with informative priors. The informative priors were selected from GLM (Generalized Linear Model) estimates (mean and variance). For $\theta_2$, the mean and variance of the estimates from GLM were used to approximate and construct the informative prior. For $\theta$, the informative priors are obtained as the following: estimates from GLM were considered as random effects.

- Denote $\gamma$ as the GLM estimates which $\gamma$ is a N x 1 vector.
- $\gamma$ can be re-written as $\beta_0$ + b where b is a N x 1 vector residual
- Estimate b with empirical variogram and conducted iterative weighted least squares method to estimate $\beta_0$
- $\beta_0$ and its variance will be incorporated into as mean and variance for $\theta$ priors

**Frequentist approach**

GLM was run with glm function in R 3.5.2.

### Zipcode level

The computation, priors and initial values are identical to county level analysis.

# Result

## Exploratory Analysis

Table 1 summarizes the number of zipcodes, baseline total counts of ED visits and the rate ratio of ED visits between 2012 and 2017 in each county. The information can also be visualized with Figure 1. From Figure 1, left and Table 1, Los Angeles has the highest number of zipcodes within the county and the most total ED counts while Mono has only 1 zipcode within the county and 14 ED visits at 2012. In Figure 1, right, We see that Mono has the highest rate ratio of 2.071 while the rate ratio for Los Angeles is around 1. The lowest rate ratio was observed in Colusa county where the total ED count is only 27 in 2012. Most of the extreme rate ratios happen in the county where the baseline total ED visits count is small. However, if we look deeper into zipcode level data within Los Angeles, the ED visit counts range from 7 to 703 and the rate ratio ranges from 0.05 to 31.8. In addition, we noticed that Northern part of California has fewer zipcodes in each county than Southern part. Therefore, the rate ratios are more extreme. This observation could potentially be modeled with spatial correlation.

## County level

### Model convergence and selection

The criteria for model selection was based on DIC. DIC for No spatial model was 1111. The DIC for Spatial 1 and 2 were 1009 and 1001, respectively. The DIC values for the models with some spatial correlation were lower comparing to model without any spatial correlation. In addition, the posterior mean for $\phi_2$ in Spatial 1 is around 30 with wide range. It seems to be a uniform-like distribution from 10 to 50, similar to our prior distribution. (Figure 3) The maximum distance in the dataset is around 15, indicating that $\phi_2$ was largely driven by the priors. On the other hand, $\phi_1$ is a sharp peak at around 0.8 despite our wide range uniform prior, indicating the present of correlation in b1 parameters and low correlation between the b3 parameters. Therefore, Spatial 2 was selected as our main model to proceed for priors sensitivity analysis. Since the $\phi_2$ for spatial indicates that there's no correlation between b3 parameters, we still assumed they share same variance in Spatial 1. Therefore, we will also include Spatial 1 inference to explore the same variance assumption. All models converged well with Rhat values for all parameters less than 1.1. The convergence plots for Spatial 2 b3 parameters are shown in Figure 2. In addition, the posterior predictive checks plots for Spatial 2 are included in the appendix Figure 22 and 23 while the plots for Spatial 1 are included in the appendix Figure 24 and 25. The predictive counts are located in the center of each of the predictive distributions. Therefore, we can conclude that both models fit well. It is worth noting that for counties with few sample sizes (ex: Inyo, Mono or Colusa), the observed ED visits values are more towards the right or left of the distribution comparing to Spatial 2 where the observed visits are centered.

### Prior sensitivity

We assessed the priors for Spatial 2 on $\theta$, $\theta_2$, $\phi$ and $\tau$. As mentioned in the Methods section, the prior distribution for $\theta$ and $\theta_2$ were obtained from the glm estimates. The informative prior for $\theta$ is N(5.984782, 0.01571683) and informative prior for $\theta_2$ is t(0.2050, 14.34982,7). The t distribution was chosen based on the wide range of posterior distribution in Figure 4. The results for informative priors are shown in Figure 5. We do not see much shift in the center of the posterior distribution for both parameters, indicating that the posterior distribution was robust with different priors settings. However, more smoothness was observed for $\theta_2$ with informative prior where t distribution was able to capture the variation over the original normal prior. Different priors for the precision $\tau$ give us similar results for the posterior. In Figure 6, the posterior distribution of $\sigma^2$ shows a peak around 5. The uniform prior gives the longest tail range to around 65 while the gamma(1,1) prior gives a tail to around 30. The posterior distributions of $\phi$ with different priors are shown in Figure 7. We observe the same sharp peak as in Figure 3. The difference in prior only results in the range of the posterior density plots. However, we do see the shape and center are similar. The density for values larger than 5 are very close to 0 regardless of the priors used. The convergence plots for b3, $\phi$, $\tau$, are presented in Appendix Figure 14. All Rhat values are less than 1.1. The b3 estimates with informative priors are highly similar to non-informative priors. We will proceed the following with results from informative priors model.

### Comparison with frequentist model

The summary statistics for the Spatial 2 was summarized in Table 2. The line plots for 95% credible intervals were shown in Figure 8. The summary statistics for the Spatial 1 was summarized in Table 3. The line plots for 95% credible intervals were shown in Figure 9. The summary statistics for glm results was summarized in Table 4. The convergence and density plots for b1 were shown in Appendix Figure 16 and 17.

The visual results for all three models were summarized in Figure 10. The posterior density plots for b3 from Spatial 2 and 1 were summarized in Appendix Figure X and X. In Figure 8 and 9, we see that the counties with wide credible intervals were those with few baseline visits. These counties were usually the ones with more extreme rate values. We noticed that in general, the posterior mean estimates are higher than the glm for Spatial 2. The increase is more obvious for counties with wider intervals. For example, the log rate estimate is -0.708 for Colusa with Bayesian model and -0.81 with glm. In contrast, Spatial 1 tends to shrink the estimates toward 0. We can also observe this trend in Figure 10 where comparing to glm, Spatial 2 has larger upper and lower bound while Spatial 1 has lower upper and lower bound, indicating the shrinking toward rate 1. Therefore, we could expect that more counties will not be significant due to the shrinkage. Mono and Inyo are the the obvious counties that were not significant as the estimates and intervals were largely shrunk towards the null.

To further investigate the larger rates observation, we looked into b1 estimates which estimated the baseline ED visits count. The results were listed in Table 5. We can observe that for Spatial 2, the estimated baseline counts are smaller than estimates from the glm. In contrast, for Spatial 1, the estimates are usually larger for counties with positive b3 and lower for negative b3 parameters. When calculated rates, this difference could affect counties with small baseline counts, for example, Colusa, Inyo or Mono.

## Zipcode level

For the zipcode level analysis, we directly adapt the settings in county level but treated the zipcode data as repeated measurements. The glm results for zipcode level is the same as county level data as shown in Table 4. The summary statistics for Spatial 2 and Spatial 1 were summarized in Table 6 and Table 7. The line plots for 95% credible intervals for Spatial 2 and Spatial 1 were shown in Figure 11 and Figure 12. The visual results for all models were summarized in Figure 13. The posterior convergence and density plots for b3 from the proposed model was summarized in Appendix Figure 18 and 19. The convergence and density plots for b1 were shown in Appendix Figure 20 and 21.

From Table 6, 7 and Figure 11, 12, we can notice that the posterior mean for each county becomes larger comparing to county level data for both Spatial 1 and 2 models. Counties with less average sample/zipcode are more sensitive to the change. Similar to County level analysis, Spatial 1 still shrinks the estimates towards the null comparing to Spatial 2 and glm, the estimates are larger for those counties that are sensitive to change. For example, the estimates for Inyo and Mono were 0.2558 and 0.1899 for Spatial 1 in county level while the estimates were 0.4372 and 0.3435 for zipcode level. The estimates for Inyo and Mono were 0.7537 and 0.8346 for Spatial 1 in county level while the estimates were 0.9889 and 1.1869 for zipcode level. In figure 13, we also notice that the upper bound for Spatial 2 becomes more than 3 and more than 1.5 for Spatial 1.

Similar to county level analysis, the baseline ED visits estimates were similar for counties with larger visits (Table 8). For counties like Colusa, we notice the decrease for both Spatial 2 and 1 comparing to county level data (23.6, 18.93 vs 15.34, 13.13). For Mono, the estimates for Spatial 2 was 12 and Spatial 1 was 18 comparing to 14 from glm for county level analysis. The estimates for Spatial 2 is 8 and Spatial was 14 comparing to 14 from glm for zipcode level analysis. The baseline estimates from zipcode level are smaller than from county level and the decrease is more obvious in counties with smaller sample size.

# Discussion

The goal of the study was to evaluate the implementation of spatial correlation structure when modeling the ED visits in the California counties between the year of 2012 and 2017. We included the results for two types of models where spatial correlation was incorporated differently. The convergence of the models and priors sensitivity analyses were assessed as well. The results were then compared with frequentist method where the incorporation of spatial correlation is not feasible in GLM. The DIC selection indicates that the incorporation of spatial correlation is needed with the difference of around 100 with DIC criteria. Among the two spatial models proposed, we found that parameter $\phi$ and $\phi_2$ behaves differently in Spatial 1 where $\phi$ is sharp around 3 while $\phi_2$ is uniform-like distribution with large numbers. This indicates that $b_3$ parameters are mostly uncorrelated. Then, the difference between Spatial 1 and 2 will be the assumption of parameters $b_3$ sharing same variance or not. For priors sensitivity analysis, the different priors do not affect most of the parameters in the models. The original Normal flat priors of $\theta_2$ might not be capturing the correct variation as seen in Figure 4. Therefore, distribution that can have heavier tails, like the t-distribution could better capture the variation. Finally, different types of priors for $\tau$ results in same shape of the posterior distribution but with different ranges. However, the posterior densities become 0 after the values of 15 for all different priors used.

The behavior of the parameters of interest, $b_{3k}$ were heavily influenced by the models we assumed when the sample size is small. When sample size is large, Spatial 1, 2 and GLM behaves similarly. In counties like Colusa, Mono or Inyo, with sample size less than 50, the posterior means of the log rates shifted more towards positive monotonically in Spatial 2 while in Spatial 1, the estimates shifted toward 0 comparing to GLM. The difference between these two models could lead to difference inference drawn. We can infer that by assuming both parameters sharing same variance, shrinkage effects toward null was observed. In these case, the wide credible intervals for counties with few sample sizes will most likely not be significant. The behavior observed in Spatial 2 is more obvious with zipcode-level analysis where the interval shifts more towards positive comparing to county-level analysis. For Spatial 1, the shrinkage effect towards null still exists while the interval shifts toward positive monotonically comparing to county-level analysis. The GLM results remain the same for both level analysis.

The reason for the behaviors observed in the two spatial models is still not clear. One possible hypothesis is that with the implementation of spatial correlation, the estimates shifts toward null. For example, in Spatial 2, the spatial correlation was implemented on only $b_1$. Therefore, the baseline counts estimates are smaller than GLM. No spatial correlation was implemented on $b_3$. Therefore, the counts estimates for year of 2017 are similar to GLM, resulting in monotone increase in $b_3$. In contrast, spatial correlation was implemented on both $b_1$ and $b_3$ in Spatial 1 leading to shrinkage for counts estimates for both baseline and year of 2017 and further results in the shrinkage behavior towards null.

There are some limitations in the study. We assumed the spatial covariance structure followed the power exponential distribution. In reality, this assumption might not be correct. However, the DIC and posterior distribution of $b_1$ indicates that the neighboring areas are correlated and the current assumption of correlation might be better than no correlation assumption. For the ED visit counts distribution, we assumed the counts followed Poisson distribution and did not take into account of the dispersion. We have tried to incorporate the dispersion parameters but the estimation was not valid. Therefore, it is possible that other distribution like negative binomial could be used to compare the result with Poisson distribution. In summary, we have observed different behaviors by incorporating spatial correlation differently. Moreover, these different behaviors affect the counties with few visits more profoundly than counties with larger number of visits.

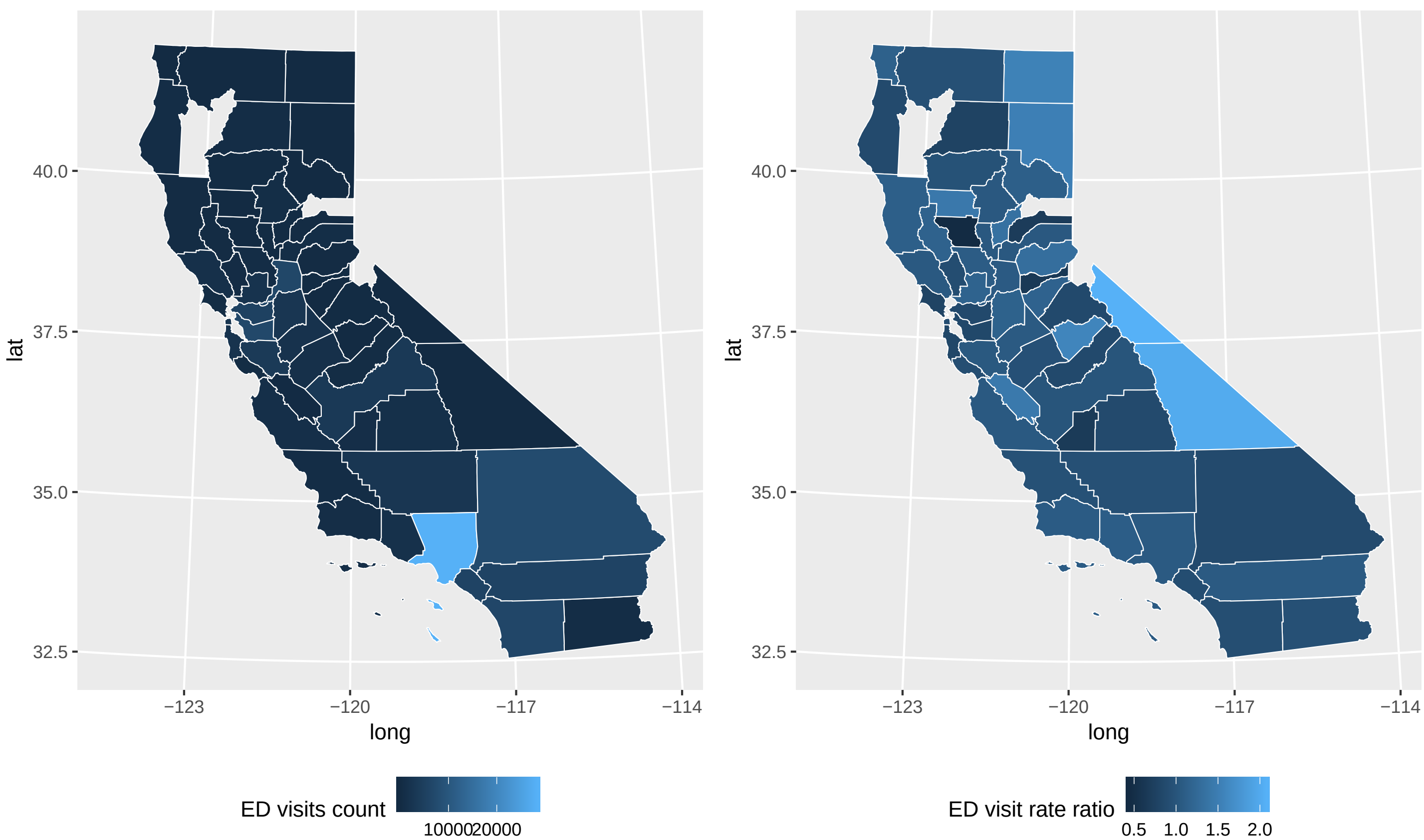


Figure 1: The left plot shows the number of ED visits in each county in 2012. The lighter colors indicate larger number of ED visits. The right plots show the ratio of ED visit between 2012 and 2017 for each county in California. The lighter colors indicate the ratio between 2017 and 2012 is larger.

Table 1: Number of zipcodes, total ED visits at 2012 and the ED visit rate ratio between 2012 and 2017 in each county

| | Number of zipcodes | Total.vistis | Rate.ratio |
|---|---|---|---|
| Alameda | 41 | 6788 | 0.828 |
| Amador | 4 | 172 | 0.640 |
| Butte | 8 | 646 | 1.037 |
| Calaveras | 2 | 67 | 1.164 |
| Colusa | 1 | 27 | 0.444 |
| Contra Costa | 32 | 5209 | 0.849 |
| Del Norte | 1 | 78 | 1.141 |
| El Dorado | 5 | 273 | 1.300 |
| Fresno | 34 | 3309 | 0.996 |
| Glenn | 2 | 61 | 1.426 |
| Humboldt | 6 | 588 | 0.862 |
| Imperial | 7 | 601 | 0.935 |
| Inyo | 1 | 27 | 2.000 |
| Kern | 24 | 2670 | 0.937 |
| Kings | 5 | 842 | 0.662 |
| Lake | 7 | 281 | 1.160 |
| Lassen | 1 | 47 | 1.511 |
| Los Angeles | 249 | 28265 | 1.059 |
| Madera | 4 | 434 | 0.848 |
| Marin | 7 | 391 | 0.777 |
| Mariposa | 1 | 26 | 1.577 |
| Mendocino | 3 | 266 | 1.120 |
| Merced | 12 | 1214 | 0.942 |
| Modoc | 1 | 28 | 1.536 |
| Mono | 1 | 14 | 2.071 |
| Monterey | 13 | 1011 | 1.046 |
| Napa | 3 | 369 | 0.894 |
| Nevada | 4 | 270 | 0.681 |
| Orange | 74 | 5662 | 0.903 |
| Placer | 10 | 700 | 1.037 |
| Plumas | 2 | 32 | 1.125 |
| Riverside | 59 | 5679 | 1.056 |
| Sacramento | 42 | 6428 | 1.063 |
| San Benito | 1 | 123 | 1.439 |
| San Bernardino | 51 | 7376 | 0.859 |
| San Diego | 70 | 6342 | 0.909 |
| San Francisco | 21 | 2348 | 0.897 |
| San Joaquin | 23 | 2682 | 1.153 |
| San Luis Obispo | 11 | 478 | 0.948 |
| San Mateo | 18 | 1887 | 0.814 |

| | Number of zipcodes | Total.vistis | Rate.ratio |
|---|---|---|---|
| Santa Barbara | 13 | 742 | 1.078 |
| Santa Clara | 44 | 3575 | 1.044 |
| Santa Cruz | 8 | 441 | 0.907 |
| Shasta | 6 | 581 | 0.764 |
| Siskiyou | 2 | 64 | 0.938 |
| Solano | 10 | 1982 | 1.173 |
| Sonoma | 16 | 1232 | 1.050 |
| Stanislaus | 18 | 1786 | 1.058 |
| Sutter | 3 | 167 | 1.042 |
| Tehama | 3 | 198 | 0.955 |
| Tulare | 15 | 1179 | 0.858 |
| Tuolumne | 2 | 141 | 0.851 |
| Ventura | 20 | 1504 | 1.104 |
| Yolo | 7 | 506 | 1.087 |
| Yuba | 2 | 163 | 1.337 |

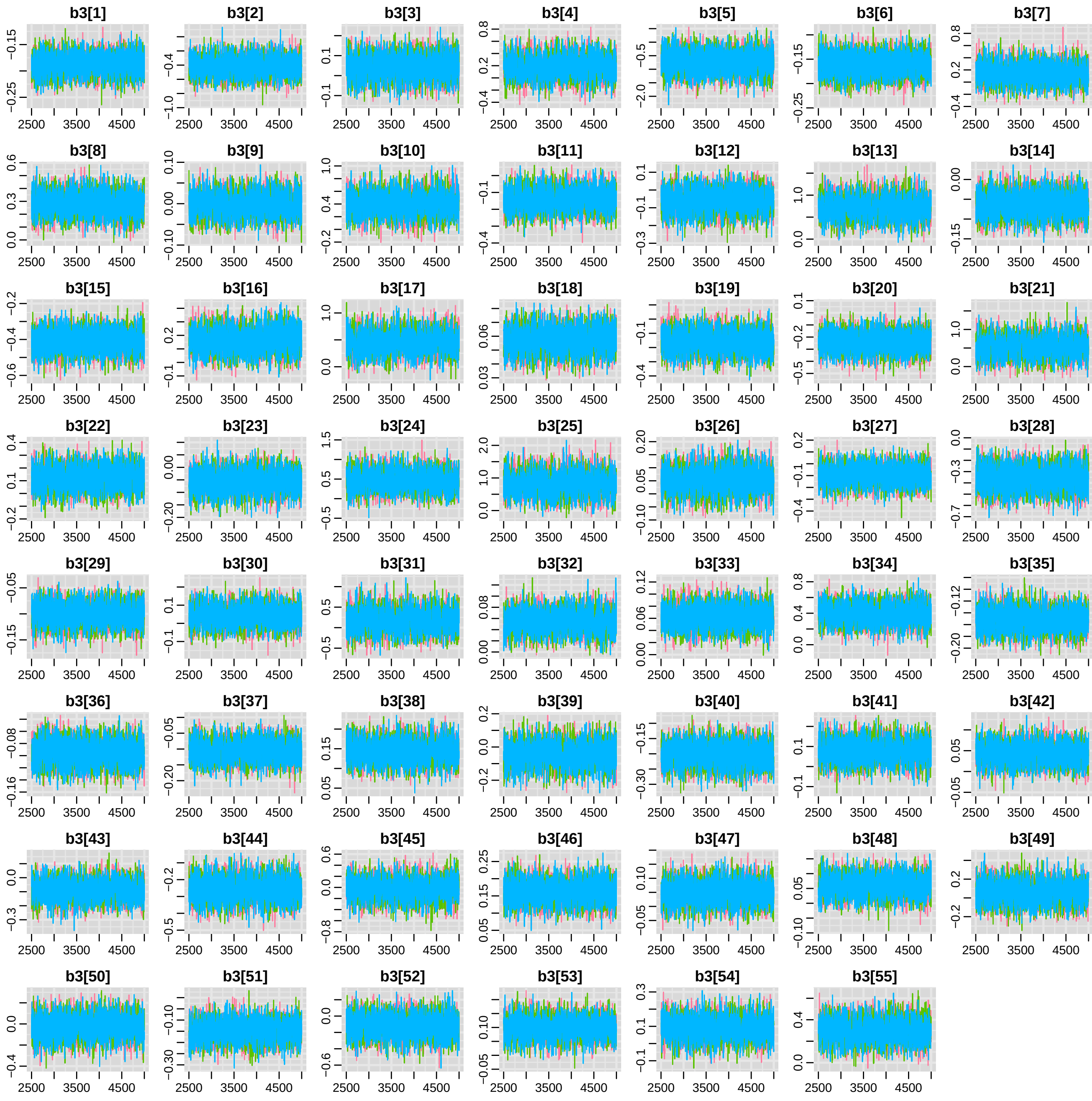

Figure 2: Convergence plots for b3 for county level analysis with non-informative priors. The numbers in the bracket indicate the order of the counties. All counties were ordered alphabetically as in Table 1

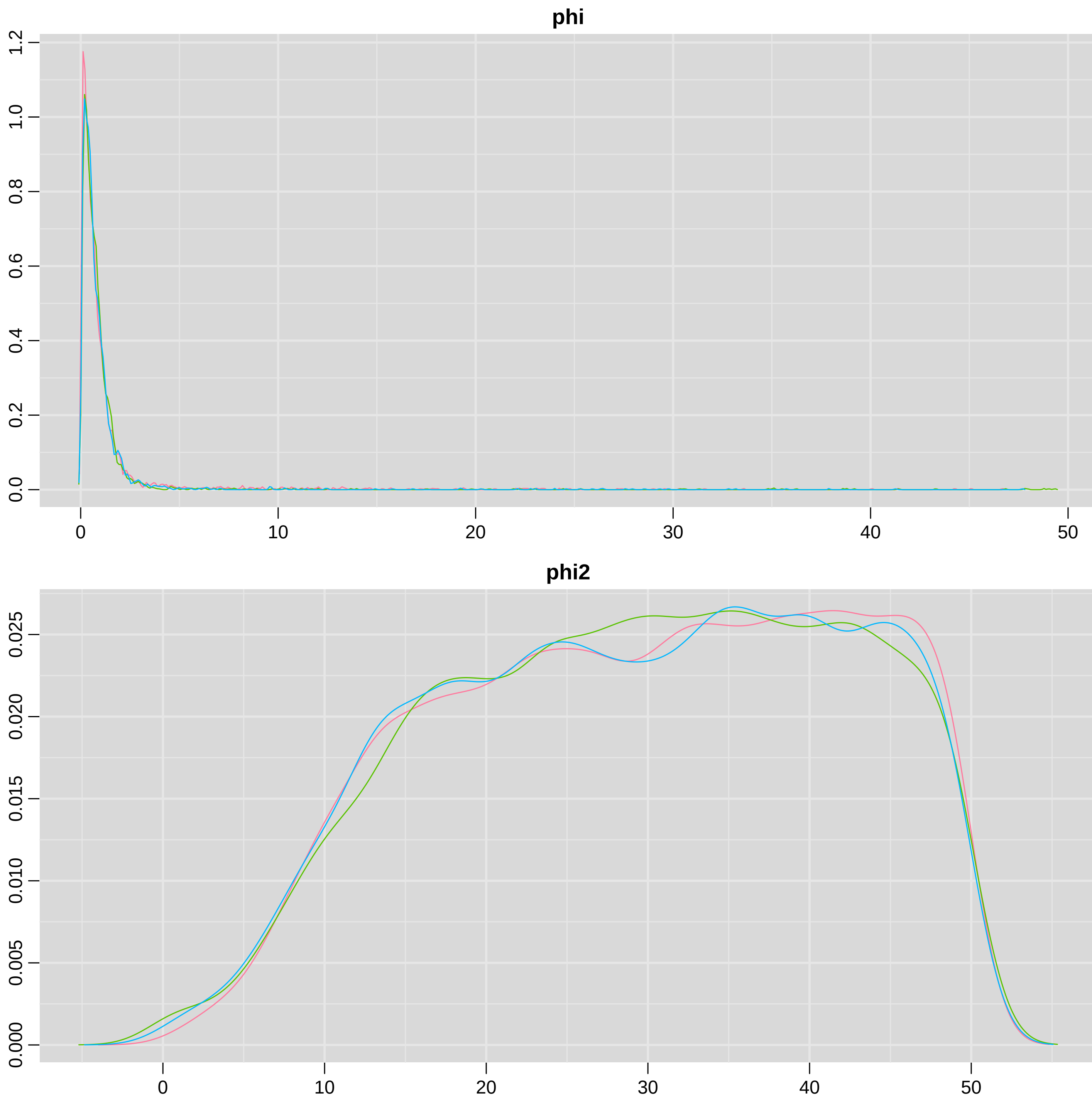


Figure 3: Density plots for both decaying parameters in Spatial 1

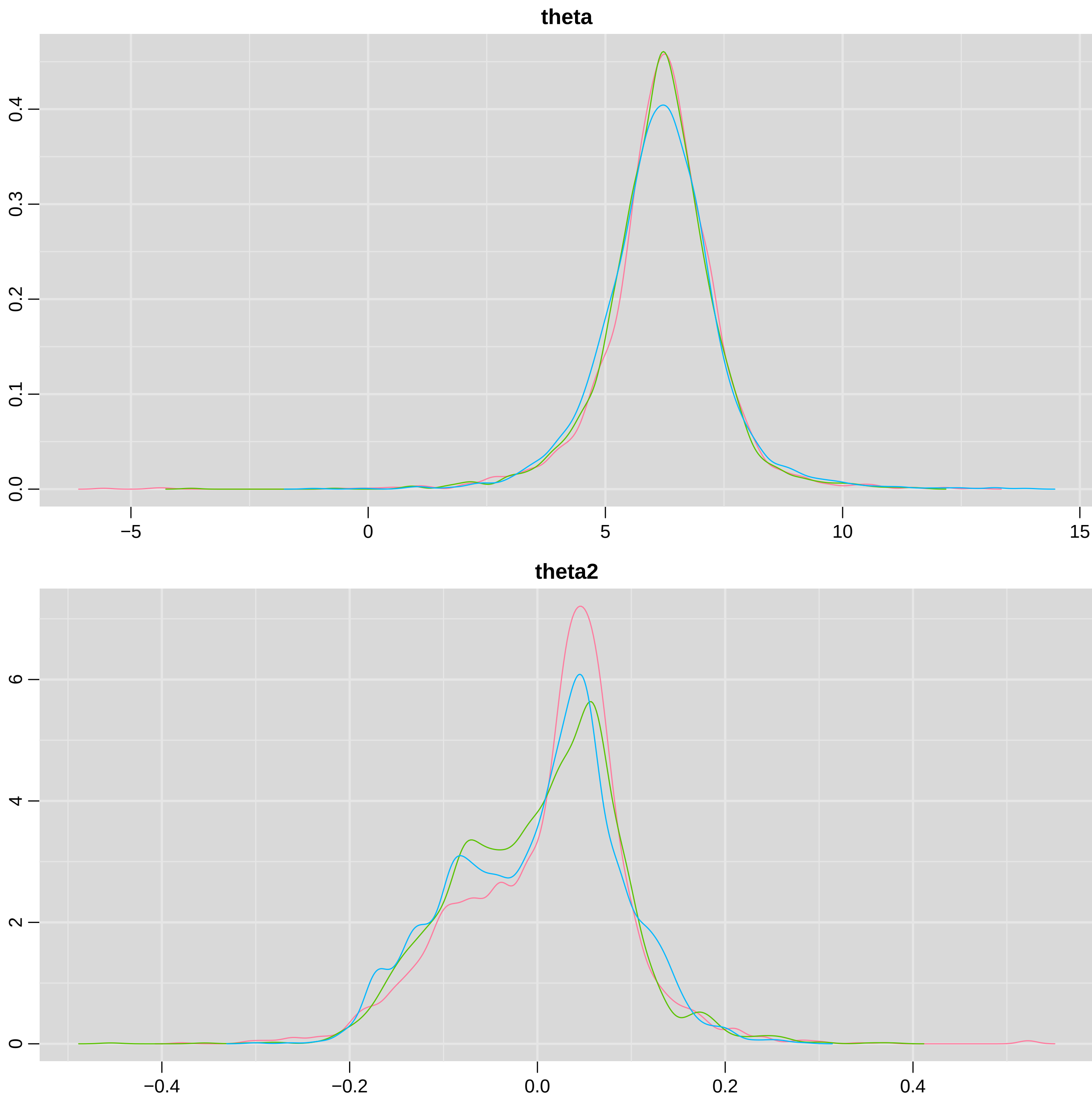


Figure 4: Density plots for theta distribution with non-informative priors in Spatial 2

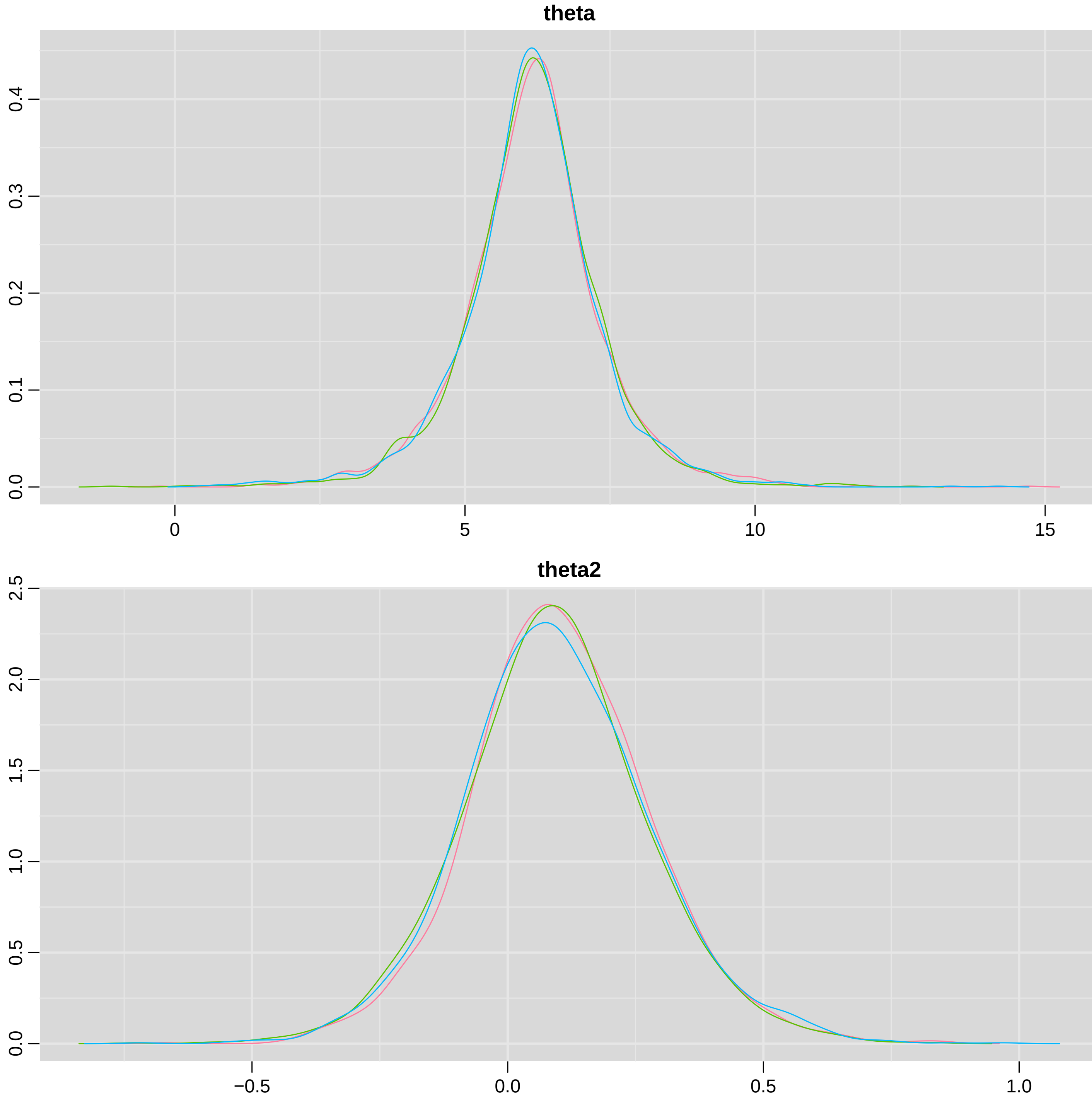


Figure 5: Density plots for theta distribution with informative priors in Spatial 2

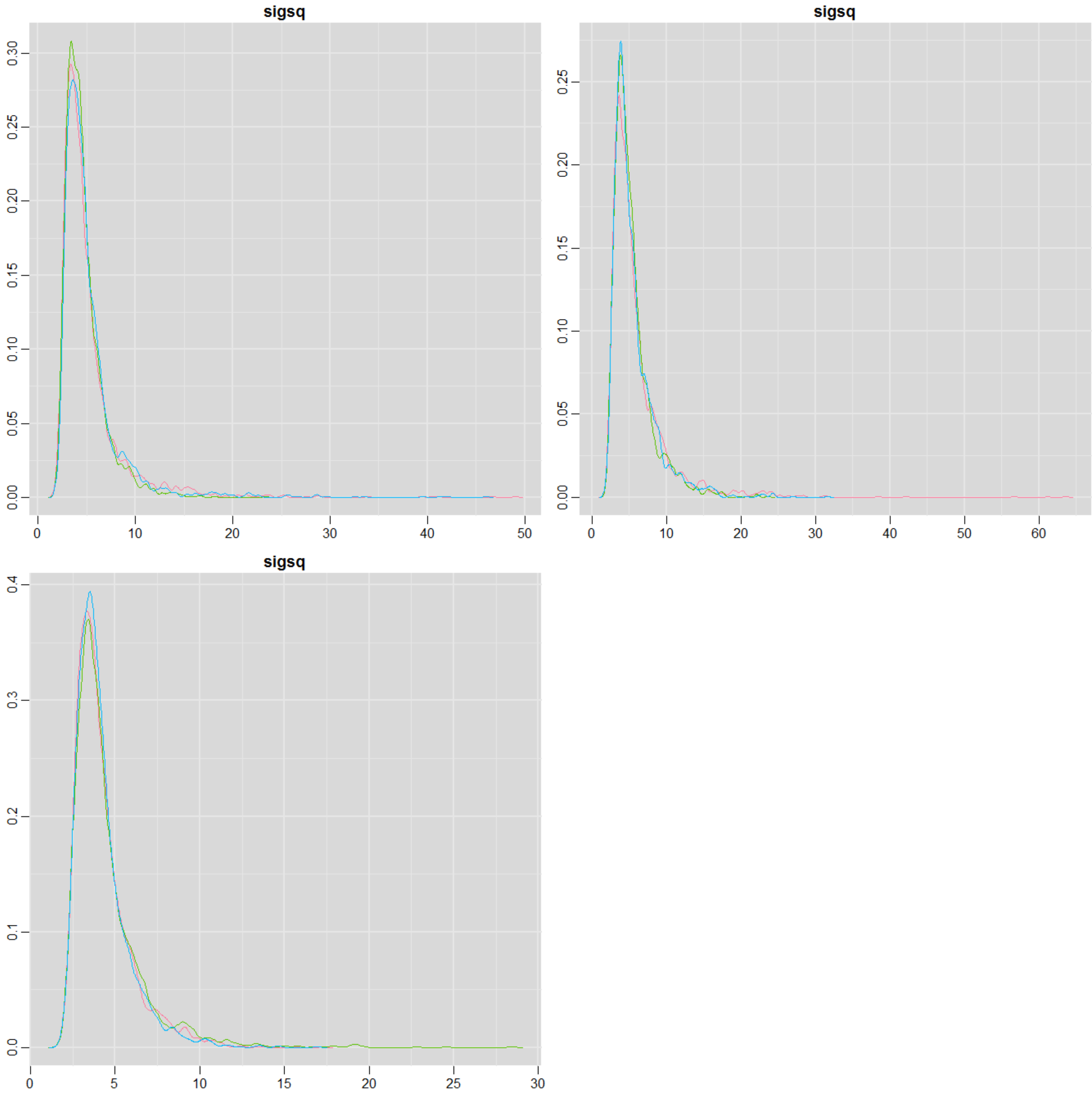


Figure 6: Posterior density plots for sigma2 (1/tau). The top left figure corresponds to tau prior with gamma(0.001,0.001). The top right figure corresponds to tau prior with uniform(0,1000). The bottom left figures correspnds to tau prior with gamma(1,1).

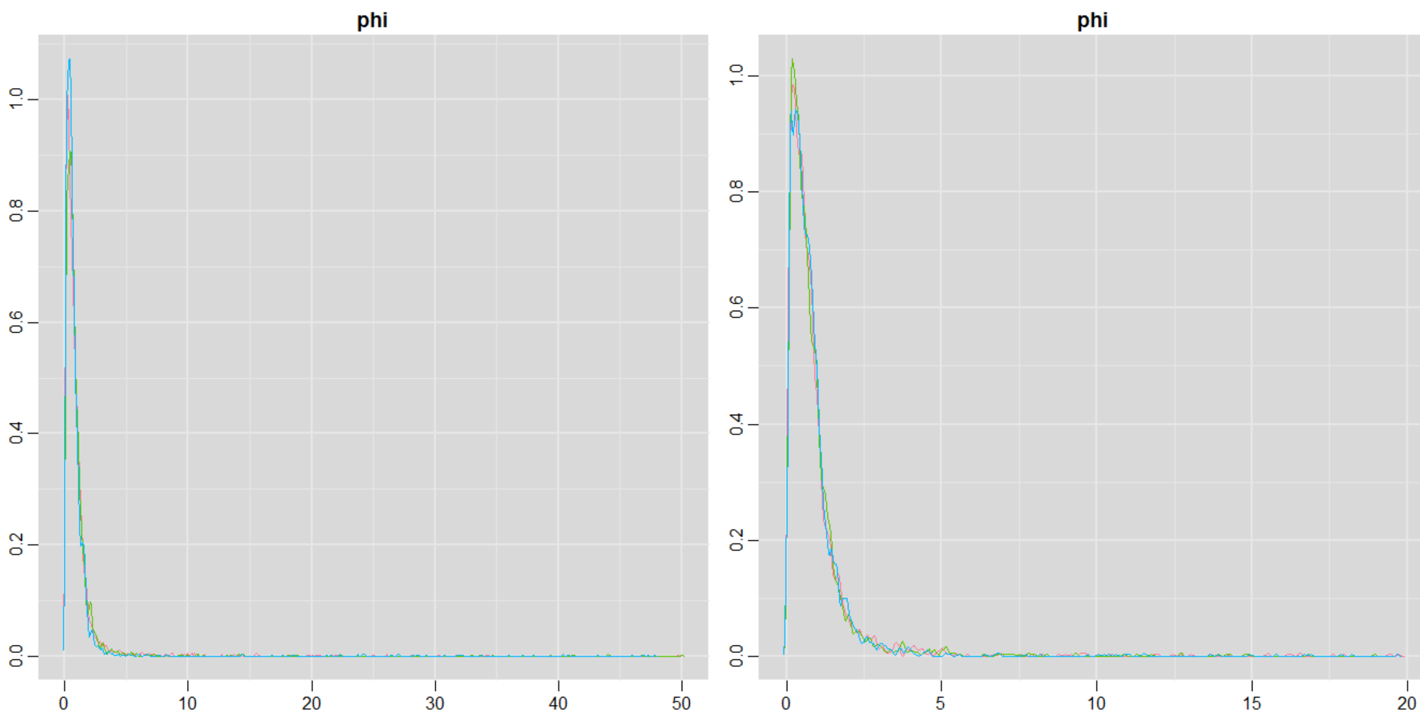


Figure 7: Posterior density plots for phi. The left figure corresponds to wider uniform prior (0.05,50). The right figure corresponds to narrower uniform prior (0.05,20).

Table 2: Summary statistics table for Spatial 2 with informative priors including posterior mean, 95 percent credible intervals and if b3 (log rate) intervals contain 0 for county level analysis.

| | Mean | 2.5% | 97.5% | Contain.zero |
|---|---|---|---|---|
| Alameda | −0.1879 | −0.2237 | −0.1524 | TRUE |
| Amador | −0.4162 | −0.6566 | −0.1756 | TRUE |
| Butte | 0.0427 | −0.0689 | 0.1521 | FALSE |
| Calaveras | 0.1896 | −0.1406 | 0.5311 | FALSE |
| Colusa | −0.7083 | −1.4290 | −0.0122 | TRUE |
| Contra Costa | −0.1629 | −0.2032 | −0.1228 | TRUE |
| Del Norte | 0.1551 | −0.1489 | 0.4681 | FALSE |
| El Dorado | 0.2836 | 0.1245 | 0.4431 | TRUE |
| Fresno | −0.0032 | −0.0508 | 0.0443 | FALSE |
| Glenn | 0.4111 | 0.0880 | 0.7387 | TRUE |
| Humboldt | −0.1429 | −0.2607 | −0.0264 | TRUE |
| Imperial | −0.0657 | −0.1795 | 0.0487 | FALSE |
| Inyo | 0.7537 | 0.2848 | 1.2420 | TRUE |
| Kern | −0.0637 | −0.1185 | −0.0088 | TRUE |
| Kings | −0.4091 | −0.5168 | −0.3021 | TRUE |
| Lake | 0.1664 | −0.0001 | 0.3273 | FALSE |
| Lassen | 0.4683 | 0.0945 | 0.8481 | TRUE |
| Los Angeles | 0.0579 | 0.0413 | 0.0743 | TRUE |
| Madera | −0.1563 | −0.2979 | −0.0153 | TRUE |
| Marin | −0.2439 | −0.3963 | −0.0986 | TRUE |
| Mariposa | 0.5687 | 0.0817 | 1.0820 | TRUE |
| Mendocino | 0.1205 | −0.0432 | 0.2810 | FALSE |
| Merced | −0.0574 | −0.1386 | 0.0233 | FALSE |
| Modoc | 0.4734 | 0.0153 | 0.9643 | TRUE |
| Mono | 0.8346 | 0.2131 | 1.4980 | TRUE |
| Monterey | 0.0519 | −0.0354 | 0.1381 | FALSE |
| Napa | −0.0992 | −0.2493 | 0.0532 | FALSE |
| Nevada | −0.3655 | −0.5563 | −0.1751 | TRUE |
| Orange | −0.1021 | −0.1397 | −0.0643 | TRUE |
| Placer | 0.0433 | −0.0583 | 0.1441 | FALSE |
| Plumas | 0.1997 | −0.2783 | 0.6888 | FALSE |
| Riverside | 0.0550 | 0.0189 | 0.0910 | TRUE |
| Sacramento | 0.0619 | 0.0284 | 0.0960 | TRUE |
| San Benito | 0.3951 | 0.1580 | 0.6305 | TRUE |
| San Bernardino | −0.1513 | −0.1846 | −0.1189 | TRUE |
| San Diego | −0.0956 | −0.1306 | −0.0600 | TRUE |
| San Francisco | −0.1070 | −0.1640 | −0.0467 | TRUE |
| San Joaquin | 0.1442 | 0.0923 | 0.1960 | TRUE |
| San Luis Obispo | −0.0497 | −0.1809 | 0.0815 | FALSE |
| San Mateo | −0.2039 | −0.2725 | −0.1361 | TRUE |

| | Mean | 2.5% | 97.5% | Contain.zero |
|---|---|---|---|---|
| Santa Barbara | 0.0799 | −0.0209 | 0.1790 | FALSE |
| Santa Clara | 0.0440 | −0.0010 | 0.0897 | FALSE |
| Santa Cruz | −0.0927 | −0.2279 | 0.0450 | FALSE |
| Shasta | −0.2603 | −0.3838 | −0.1390 | TRUE |
| Siskiyou | −0.0289 | −0.3833 | 0.3185 | FALSE |
| Solano | 0.1614 | 0.1019 | 0.2223 | TRUE |
| Sonoma | 0.0524 | −0.0263 | 0.1304 | FALSE |
| Stanislaus | 0.0583 | −0.0059 | 0.1230 | FALSE |
| Sutter | 0.0632 | −0.1482 | 0.2793 | FALSE |
| Tehama | −0.0248 | −0.2234 | 0.1734 | FALSE |
| Tulare | −0.1499 | −0.2356 | −0.0638 | TRUE |
| Tuolumne | −0.1239 | −0.3738 | 0.1204 | FALSE |
| Ventura | 0.0999 | 0.0311 | 0.1710 | TRUE |
| Yolo | 0.0923 | −0.0256 | 0.2122 | FALSE |
| Yuba | 0.3106 | 0.1079 | 0.5168 | TRUE |

Table 3: Summary statistics table for Spatial 1 with informative priors including posterior mean, 95 percent credible intervals and if b3 (log rate) intervals contain 0 for county level analysis.

| | Mean | 2.5% | 97.5% | Contain.zero |
|---|---|---|---|---|
| Alameda | −0.1863 | −0.2214 | −0.1518 | TRUE |
| Amador | −0.2466 | −0.4484 | −0.0410 | TRUE |
| Butte | 0.0387 | −0.0656 | 0.1421 | FALSE |
| Calaveras | 0.0966 | −0.1354 | 0.3326 | FALSE |
| Colusa | −0.1229 | −0.4361 | 0.1715 | FALSE |
| Contra Costa | −0.1606 | −0.2008 | −0.1203 | TRUE |
| Del Norte | 0.0836 | −0.1433 | 0.3203 | FALSE |
| El Dorado | 0.2170 | 0.0720 | 0.3630 | TRUE |
| Fresno | −0.0025 | −0.0495 | 0.0442 | FALSE |
| Glenn | 0.2033 | −0.0330 | 0.4516 | FALSE |
| Humboldt | −0.1265 | −0.2419 | −0.0144 | TRUE |
| Imperial | −0.0581 | −0.1696 | 0.0515 | FALSE |
| Inyo | 0.2558 | −0.0214 | 0.5641 | FALSE |
| Kern | −0.0626 | −0.1179 | −0.0077 | TRUE |
| Kings | −0.3709 | −0.4739 | −0.2673 | TRUE |
| Lake | 0.1349 | −0.0101 | 0.2836 | FALSE |
| Lassen | 0.2131 | −0.0428 | 0.4815 | FALSE |
| Los Angeles | 0.0576 | 0.0412 | 0.0742 | TRUE |
| Madera | −0.1303 | −0.2623 | −0.0032 | TRUE |
| Marin | −0.2022 | −0.3426 | −0.0660 | TRUE |
| Mariposa | 0.1728 | −0.1014 | 0.4691 | FALSE |
| Mendocino | 0.1001 | −0.0480 | 0.2495 | FALSE |
| Merced | −0.0558 | −0.1358 | 0.0228 | FALSE |
| Modoc | 0.1605 | −0.1206 | 0.4408 | FALSE |
| Mono | 0.1899 | −0.1071 | 0.5243 | FALSE |
| Monterey | 0.0532 | −0.0302 | 0.1374 | FALSE |
| Napa | −0.0835 | −0.2215 | 0.0536 | FALSE |
| Nevada | −0.2705 | −0.4441 | −0.1070 | TRUE |
| Orange | −0.1011 | −0.1392 | −0.0625 | TRUE |
| Placer | 0.0385 | −0.0595 | 0.1357 | FALSE |
| Plumas | 0.0688 | −0.2166 | 0.3528 | FALSE |
| Riverside | 0.0541 | 0.0167 | 0.0906 | TRUE |
| Sacramento | 0.0609 | 0.0273 | 0.0954 | TRUE |
| San Benito | 0.2613 | 0.0707 | 0.4573 | TRUE |
| San Bernardino | −0.1504 | −0.1842 | −0.1161 | TRUE |
| San Diego | −0.0946 | −0.1303 | −0.0588 | TRUE |
| San Francisco | −0.1052 | −0.1642 | −0.0464 | TRUE |
| San Joaquin | 0.1406 | 0.0898 | 0.1929 | TRUE |
| San Luis Obispo | −0.0444 | −0.1629 | 0.0745 | FALSE |
| San Mateo | −0.1959 | −0.2612 | −0.1298 | TRUE |

| | Mean | 2.5% | 97.5% | Contain.zero |
|---|---|---|---|---|
| Santa Barbara | 0.0725 | −0.0210 | 0.1697 | FALSE |
| Santa Clara | 0.0433 | −0.0019 | 0.0889 | FALSE |
| Santa Cruz | −0.0808 | −0.2057 | 0.0464 | FALSE |
| Shasta | −0.2276 | −0.3459 | −0.1144 | TRUE |
| Siskiyou | −0.0147 | −0.2595 | 0.2273 | FALSE |
| Solano | 0.1556 | 0.0950 | 0.2160 | TRUE |
| Sonoma | 0.0481 | −0.0289 | 0.1236 | FALSE |
| Stanislaus | 0.0577 | −0.0053 | 0.1199 | FALSE |
| Sutter | 0.0470 | −0.1349 | 0.2236 | FALSE |
| Tehama | −0.0262 | −0.1962 | 0.1527 | FALSE |
| Tulare | −0.1405 | −0.2221 | −0.0567 | TRUE |
| Tuolumne | −0.0694 | −0.2720 | 0.1304 | FALSE |
| Ventura | 0.0969 | 0.0291 | 0.1672 | TRUE |
| Yolo | 0.0809 | −0.0318 | 0.1934 | FALSE |
| Yuba | 0.2295 | 0.0541 | 0.4156 | TRUE |

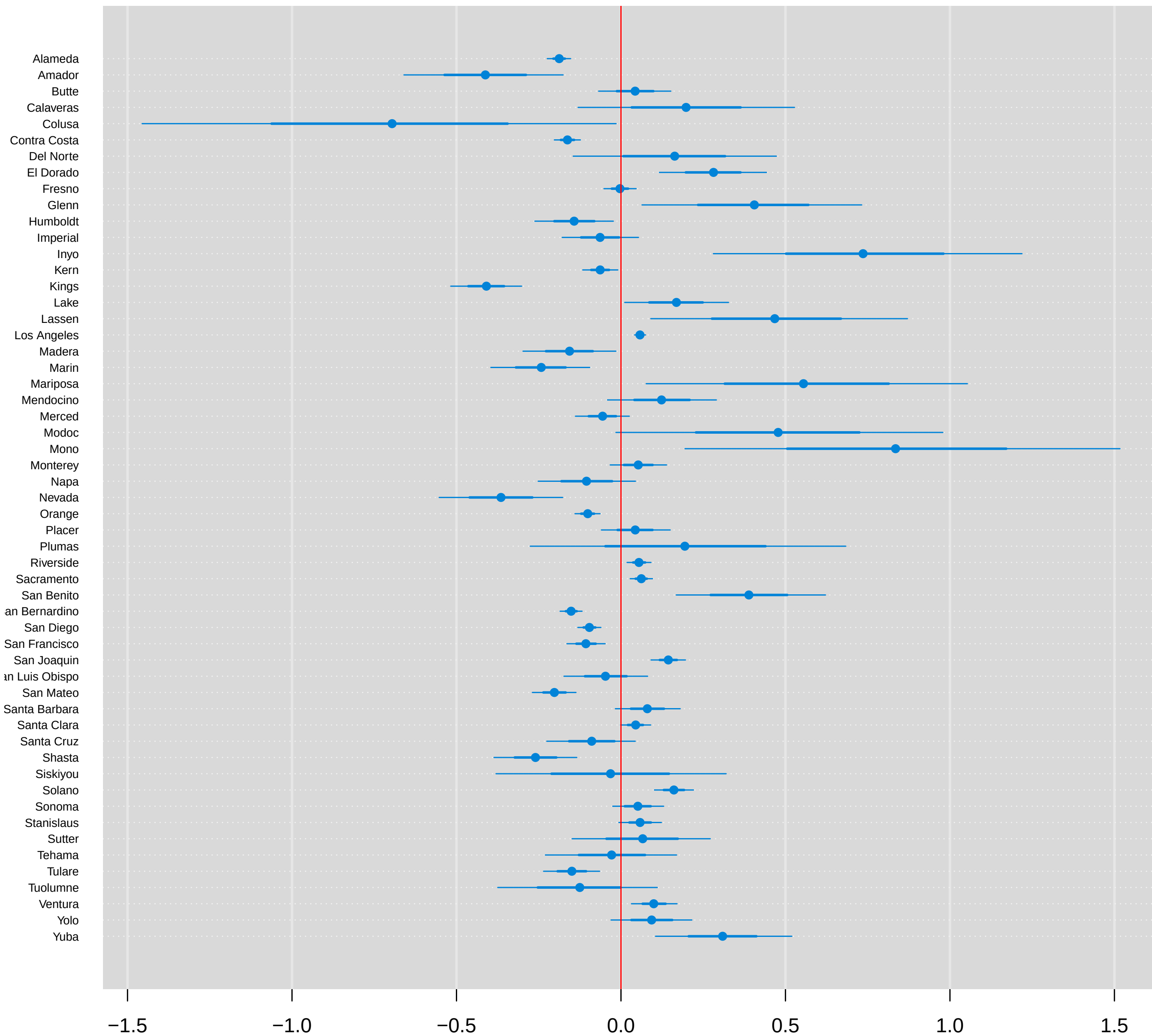


Figure 8: Line plots for Spatial 2 model with informative priors, 95 percent credible intervals in each county for county level analysis

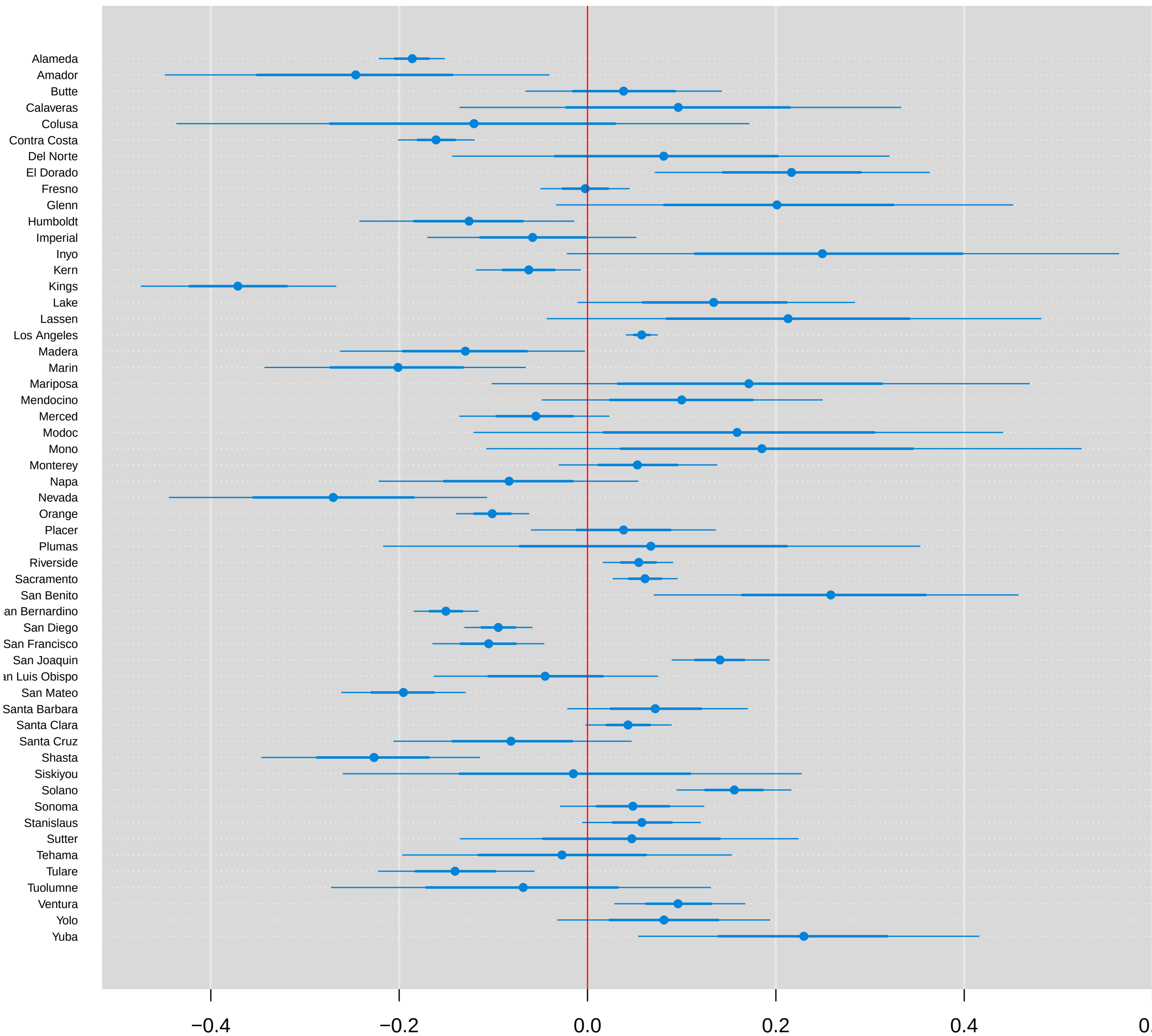


Figure 9: Line plots for Spatial 1 model with informative priors, 95 percent credible intervals in each county for county level analysis

Table 4: Summary statistics table for b3 with glm result. The table includes the glm estimate, lower and upper bound of 95 percent confidence intervals and if the analysis survive FDR correction at 0.1. The analysis used county level data

| | Estimate | lowerCI | upperCI | sig |
|---|---|---|---|---|
| Alameda | −0.1886 | −0.2240 | −0.1533 | TRUE |
| Amador | −0.4470 | −0.6889 | −0.2051 | TRUE |
| Butte | 0.0365 | −0.0772 | 0.1502 | FALSE |
| Calaveras | 0.1520 | −0.1764 | 0.4804 | FALSE |
| Colusa | −0.8109 | −1.4919 | −0.1300 | TRUE |
| Contra Costa | −0.1638 | −0.2172 | −0.1104 | TRUE |
| Del Norte | 0.1319 | −0.1741 | 0.4380 | FALSE |
| El Dorado | 0.2626 | 0.1010 | 0.4243 | TRUE |
| Fresno | −0.0039 | −0.0637 | 0.0559 | FALSE |
| Glenn | 0.3550 | 0.0258 | 0.6842 | TRUE |
| Humboldt | −0.1482 | −0.2722 | −0.0243 | TRUE |
| Imperial | −0.0671 | −0.1874 | 0.0532 | FALSE |
| Inyo | 0.6931 | 0.2298 | 1.1565 | TRUE |
| Kern | −0.0650 | −0.1300 | 0.0000 | TRUE |
| Kings | −0.4132 | −0.5259 | −0.3005 | TRUE |
| Lake | 0.1485 | −0.0149 | 0.3120 | FALSE |
| Lassen | 0.4125 | 0.0423 | 0.7828 | TRUE |
| Los Angeles | 0.0574 | 0.0185 | 0.0963 | TRUE |
| Madera | −0.1650 | −0.3083 | −0.0216 | TRUE |
| Marin | −0.2517 | −0.4057 | −0.0977 | TRUE |
| Mariposa | 0.4555 | −0.0372 | 0.9481 | FALSE |
| Mendocino | 0.1136 | −0.0555 | 0.2827 | FALSE |
| Merced | −0.0603 | −0.1484 | 0.0279 | FALSE |
| Modoc | 0.4290 | −0.0483 | 0.9063 | FALSE |
| Mono | 0.7282 | 0.0894 | 1.3671 | TRUE |
| Monterey | 0.0454 | −0.0477 | 0.1386 | FALSE |
| Napa | −0.1117 | −0.2644 | 0.0409 | FALSE |
| Nevada | −0.3835 | −0.5742 | −0.1928 | TRUE |
| Orange | −0.1022 | −0.1539 | −0.0504 | TRUE |
| Placer | 0.0365 | −0.0732 | 0.1461 | FALSE |
| Plumas | 0.1178 | −0.3597 | 0.5953 | FALSE |
| Riverside | 0.0548 | 0.0042 | 0.1055 | TRUE |
| Sacramento | 0.0611 | 0.0120 | 0.1102 | TRUE |
| San Benito | 0.3640 | 0.1312 | 0.5967 | TRUE |
| San Bernardino | −0.1518 | −0.2006 | −0.1031 | TRUE |
| San Diego | −0.0959 | −0.1461 | −0.0457 | TRUE |
| San Francisco | −0.1088 | −0.1774 | −0.0401 | TRUE |
| San Joaquin | 0.1426 | 0.0799 | 0.2052 | TRUE |
| San Luis Obispo | −0.0537 | −0.1870 | 0.0796 | FALSE |
| San Mateo | −0.2058 | −0.2819 | −0.1297 | TRUE |

| | Estimate | lowerCI | upperCI | sig |
|---|---|---|---|---|
| Santa Barbara | 0.0753 | −0.0307 | 0.1812 | FALSE |
| Santa Clara | 0.0427 | −0.0152 | 0.1006 | FALSE |
| Santa Cruz | −0.0976 | −0.2375 | 0.0423 | FALSE |
| Shasta | −0.2689 | −0.3974 | −0.1404 | TRUE |
| Siskiyou | −0.0645 | −0.4185 | 0.2894 | FALSE |
| Solano | 0.1592 | 0.0896 | 0.2288 | TRUE |
| Sonoma | 0.0483 | −0.0373 | 0.1340 | FALSE |
| Stanislaus | 0.0566 | −0.0171 | 0.1303 | FALSE |
| Sutter | 0.0411 | −0.1742 | 0.2563 | FALSE |
| Tehama | −0.0465 | −0.2489 | 0.1559 | FALSE |
| Tulare | −0.1527 | −0.2439 | −0.0616 | TRUE |
| Tuolumne | −0.1613 | −0.4073 | 0.0847 | FALSE |
| Ventura | 0.0987 | 0.0205 | 0.1769 | TRUE |
| Yolo | 0.0834 | −0.0424 | 0.2092 | FALSE |
| Yuba | 0.2907 | 0.0847 | 0.4968 | TRUE |

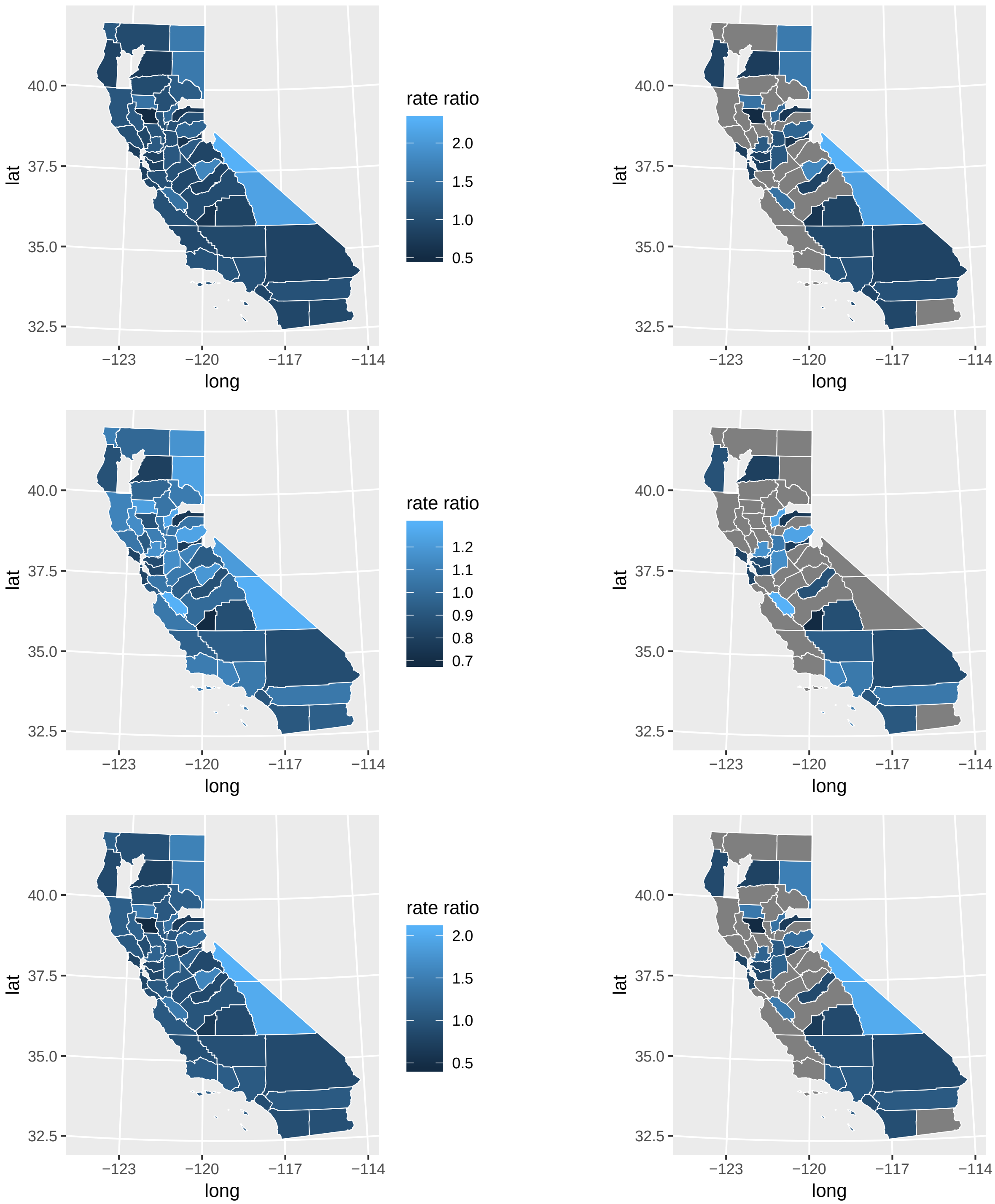


Figure 10: The plots in the first, second and third rows corresponding to Spatial 2, Spatial 1 and GLM. The left column denote the posterior mean or glm estimates of rate (exp(b3)). The right column denotes the significance of the estimated rates. The gray shaded areas were not significant. The results are shown for county level analysis

Table 5: Estimated baseline ED visits with GLM, Spatial 2 and Spatial 1 models in county level analysis.

| | GLM | Spatial.2 | Spatial.1 |
|---|---|---|---|
| Alameda | 6788 | 6783.2445 | 6777.2254 |
| Amador | 172 | 166.3100 | 154.5984 |
| Butte | 646 | 641.1760 | 642.8571 |
| Calaveras | 67 | 63.9486 | 67.1125 |
| Colusa | 27 | 23.5591 | 18.9316 |
| Contra Costa | 5209 | 5203.3844 | 5197.9840 |
| Del Norte | 78 | 75.6149 | 78.5514 |
| El Dorado | 273 | 267.0410 | 277.0292 |
| Fresno | 3309 | 3304.5540 | 3303.2774 |
| Glenn | 61 | 57.3898 | 64.8496 |
| Humboldt | 588 | 584.7102 | 580.0870 |
| Imperial | 601 | 599.2747 | 597.2179 |
| Inyo | 27 | 25.1766 | 34.3873 |
| Kern | 2670 | 2666.6706 | 2664.8682 |
| Kings | 842 | 838.1612 | 825.7262 |
| Lake | 281 | 275.4661 | 280.0228 |
| Lassen | 47 | 44.1345 | 51.1524 |
| Los Angeles | 28265 | 28255.8558 | 28258.2294 |
| Madera | 434 | 429.6604 | 424.2506 |
| Marin | 391 | 387.3610 | 380.2269 |
| Mariposa | 26 | 22.8949 | 28.8791 |
| Mendocino | 266 | 263.5740 | 266.3529 |
| Merced | 1214 | 1210.0413 | 1209.1471 |
| Modoc | 28 | 26.4108 | 31.5938 |
| Mono | 14 | 12.2595 | 18.4850 |
| Monterey | 1011 | 1004.2092 | 1002.8524 |
| Napa | 369 | 364.3103 | 361.5351 |
| Nevada | 270 | 264.7842 | 254.2583 |
| Orange | 5662 | 5660.2945 | 5656.7032 |
| Placer | 700 | 694.6489 | 696.4275 |
| Plumas | 32 | 29.0938 | 31.2650 |
| Riverside | 5679 | 5676.4077 | 5679.6078 |
| Sacramento | 6428 | 6422.0261 | 6426.0287 |
| San Benito | 123 | 118.8550 | 128.2800 |
| San Bernardino | 7376 | 7371.8355 | 7369.7009 |
| San Diego | 6342 | 6338.8037 | 6336.6193 |
| San Francisco | 2348 | 2343.4547 | 2341.1639 |
| San Joaquin | 2682 | 2676.9856 | 2682.7480 |
| San Luis Obispo | 478 | 475.5451 | 474.2697 |
| San Mateo | 1887 | 1882.7749 | 1875.1959 |

| | GLM | Spatial.2 | Spatial.1 |
|---|---|---|---|
| Santa Barbara | 742 | 738.1798 | 740.9617 |
| Santa Clara | 3575 | 3569.9893 | 3570.9042 |
| Santa Cruz | 441 | 438.1232 | 435.5890 |
| Shasta | 581 | 575.5496 | 566.9544 |
| Siskiyou | 64 | 61.3901 | 60.9119 |
| Solano | 1982 | 1977.1688 | 1982.9426 |
| Sonoma | 1232 | 1227.0229 | 1229.5288 |
| Stanislaus | 1786 | 1782.2276 | 1781.8170 |
| Sutter | 167 | 162.8440 | 164.3944 |
| Tehama | 198 | 193.3671 | 193.3971 |
| Tulare | 1179 | 1175.1975 | 1169.4648 |
| Tuolumne | 141 | 135.3925 | 131.8087 |
| Ventura | 1504 | 1501.4349 | 1503.3632 |
| Yolo | 506 | 501.2066 | 504.3098 |
| Yuba | 163 | 159.2520 | 166.5654 |

Table 6: Summary statistics table for Spatial 2 including posterior mean, 95 percent credible intervals and if b3 (log rate) intervals contain 0 for zipcode level analysis.

| | Mean | 2.5% | 97.5% | Contain.zero |
|---|---|---|---|---|
| Alameda | −0.1858 | −0.2215 | −0.1507 | TRUE |
| Amador | −0.3165 | −0.5793 | −0.0310 | TRUE |
| Butte | 0.0549 | −0.0531 | 0.1661 | FALSE |
| Calaveras | 0.3162 | −0.0266 | 0.6652 | FALSE |
| Colusa | −0.2794 | −1.0535 | 0.4875 | FALSE |
| Contra Costa | −0.1610 | −0.2015 | −0.1214 | TRUE |
| Del Norte | 0.2623 | −0.0494 | 0.5752 | FALSE |
| El Dorado | 0.3404 | 0.1718 | 0.5132 | TRUE |
| Fresno | 0.0001 | −0.0489 | 0.0476 | FALSE |
| Glenn | 0.5174 | 0.1890 | 0.8554 | TRUE |
| Humboldt | −0.1327 | −0.2536 | −0.0138 | TRUE |
| Imperial | −0.0547 | −0.1668 | 0.0583 | FALSE |
| Inyo | 0.9889 | 0.5067 | 1.5055 | TRUE |
| Kern | −0.0607 | −0.1160 | −0.0051 | TRUE |
| Kings | −0.4003 | −0.5101 | −0.2912 | TRUE |
| Lake | 0.2025 | 0.0416 | 0.3673 | TRUE |
| Lassen | 0.6358 | 0.2400 | 1.0360 | TRUE |
| Los Angeles | 0.0577 | 0.0413 | 0.0740 | TRUE |
| Madera | −0.1305 | −0.2712 | 0.0106 | FALSE |
| Marin | −0.2267 | −0.3720 | −0.0765 | TRUE |
| Mariposa | 0.9038 | 0.3632 | 1.4620 | TRUE |
| Mendocino | 0.1512 | −0.0176 | 0.3243 | FALSE |
| Merced | −0.0507 | −0.1322 | 0.0301 | FALSE |
| Modoc | 0.6790 | 0.1767 | 1.1925 | TRUE |
| Mono | 1.1869 | 0.5278 | 1.9040 | TRUE |
| Monterey | 0.0672 | −0.0231 | 0.1554 | FALSE |
| Napa | −0.0746 | −0.2246 | 0.0773 | FALSE |
| Nevada | −0.3271 | −0.5222 | −0.1316 | TRUE |
| Orange | −0.1006 | −0.1388 | −0.0620 | TRUE |
| Placer | 0.0560 | −0.0497 | 0.1617 | FALSE |
| Plumas | 0.3597 | −0.1413 | 0.8624 | FALSE |
| Riverside | 0.0566 | 0.0213 | 0.0925 | TRUE |
| Sacramento | 0.0630 | 0.0287 | 0.0963 | TRUE |
| San Benito | 0.5809 | 0.3042 | 0.8949 | TRUE |
| San Bernardino | −0.1501 | −0.1839 | −0.1170 | TRUE |
| San Diego | −0.0949 | −0.1312 | −0.0592 | TRUE |
| San Francisco | −0.1022 | −0.1619 | −0.0431 | TRUE |
| San Joaquin | 0.1473 | 0.0944 | 0.1999 | TRUE |
| San Luis Obispo | −0.0363 | −0.1620 | 0.0894 | FALSE |
| San Mateo | −0.1997 | −0.2681 | −0.1336 | TRUE |

| | Mean | 2.5% | 97.5% | Contain.zero |
|---|---|---|---|---|
| Santa Barbara | 0.0878 | −0.0114 | 0.1894 | FALSE |
| Santa Clara | 0.0462 | 0.0001 | 0.0927 | TRUE |
| Santa Cruz | −0.0728 | −0.2074 | 0.0632 | FALSE |
| Shasta | −0.2396 | −0.3645 | −0.1190 | TRUE |
| Siskiyou | 0.0575 | −0.3027 | 0.4178 | FALSE |
| Solano | 0.1674 | 0.1065 | 0.2290 | TRUE |
| Sonoma | 0.0563 | −0.0222 | 0.1353 | FALSE |
| Stanislaus | 0.0632 | −0.0012 | 0.1296 | FALSE |
| Sutter | 0.1110 | −0.1052 | 0.3275 | FALSE |
| Tehama | 0.0309 | −0.1728 | 0.2309 | FALSE |
| Tulare | −0.1426 | −0.2271 | −0.0598 | TRUE |
| Tuolumne | −0.0336 | −0.2913 | 0.2301 | FALSE |
| Ventura | 0.1034 | 0.0355 | 0.1736 | TRUE |
| Yolo | 0.1064 | −0.0130 | 0.2255 | FALSE |
| Yuba | 0.3777 | 0.1703 | 0.5837 | TRUE |

Table 7: Summary statistics table for Spatial 1 including posterior mean, 95 percent credible intervals and if b3 (log rate) intervals contain 0 for zipcode level analysis.

| | Mean | 2.5% | 97.5% | Contain.zero |
|---|---|---|---|---|
| Alameda | −0.1850 | −0.2205 | −0.1497 | TRUE |
| Amador | −0.1841 | −0.4236 | 0.0739 | FALSE |
| Butte | 0.0558 | −0.0478 | 0.1617 | FALSE |
| Calaveras | 0.2083 | −0.0563 | 0.4829 | FALSE |
| Colusa | −0.0289 | −0.4039 | 0.3424 | FALSE |
| Contra Costa | −0.1587 | −0.1985 | −0.1187 | TRUE |
| Del Norte | 0.1801 | −0.0854 | 0.4351 | FALSE |
| El Dorado | 0.3049 | 0.1415 | 0.4813 | TRUE |
| Fresno | 0.0009 | −0.0462 | 0.0480 | FALSE |
| Glenn | 0.3201 | 0.0552 | 0.5955 | TRUE |
| Humboldt | −0.1180 | −0.2361 | −0.0044 | TRUE |
| Imperial | −0.0469 | −0.1597 | 0.0662 | FALSE |
| Inyo | 0.4372 | 0.1161 | 0.7947 | TRUE |
| Kern | −0.0595 | −0.1133 | −0.0062 | TRUE |
| Kings | −0.3702 | −0.4715 | −0.2656 | TRUE |
| Lake | 0.1828 | 0.0293 | 0.3392 | TRUE |
| Lassen | 0.3639 | 0.0751 | 0.6719 | TRUE |
| Los Angeles | 0.0576 | 0.0416 | 0.0738 | TRUE |
| Madera | −0.1100 | −0.2404 | 0.0234 | FALSE |
| Marin | −0.1961 | −0.3346 | −0.0558 | TRUE |
| Mariposa | 0.3716 | 0.0308 | 0.7344 | TRUE |
| Mendocino | 0.1392 | −0.0160 | 0.2959 | FALSE |
| Merced | −0.0481 | −0.1285 | 0.0307 | FALSE |
| Modoc | 0.3043 | −0.0224 | 0.6475 | FALSE |
| Mono | 0.3435 | −0.0140 | 0.7417 | FALSE |
| Monterey | 0.0729 | −0.0125 | 0.1584 | FALSE |
| Napa | −0.0610 | −0.2013 | 0.0778 | FALSE |
| Nevada | −0.2509 | −0.4284 | −0.0740 | TRUE |
| Orange | −0.0992 | −0.1374 | −0.0615 | TRUE |
| Placer | 0.0578 | −0.0433 | 0.1591 | FALSE |
| Plumas | 0.1682 | −0.1505 | 0.4947 | FALSE |
| Riverside | 0.0568 | 0.0209 | 0.0933 | TRUE |
| Sacramento | 0.0634 | 0.0294 | 0.0979 | TRUE |
| San Benito | 0.4475 | 0.1986 | 0.7354 | TRUE |
| San Bernardino | −0.1495 | −0.1831 | −0.1163 | TRUE |
| San Diego | −0.0934 | −0.1290 | −0.0562 | TRUE |
| San Francisco | −0.0997 | −0.1588 | −0.0407 | TRUE |
| San Joaquin | 0.1458 | 0.0946 | 0.1987 | TRUE |
| San Luis Obispo | −0.0299 | −0.1512 | 0.0939 | FALSE |
| San Mateo | −0.1930 | −0.2579 | −0.1260 | TRUE |

| | Mean | 2.5% | 97.5% | Contain.zero |
|---|---|---|---|---|
| Santa Barbara | 0.0830 | −0.0156 | 0.1827 | FALSE |
| Santa Clara | 0.0461 | 0.0011 | 0.0912 | TRUE |
| Santa Cruz | −0.0631 | −0.1929 | 0.0662 | FALSE |
| Shasta | −0.2125 | −0.3309 | −0.0922 | TRUE |
| Siskiyou | 0.0420 | −0.2362 | 0.3177 | FALSE |
| Solano | 0.1648 | 0.1043 | 0.2256 | TRUE |
| Sonoma | 0.0552 | −0.0212 | 0.1301 | FALSE |
| Stanislaus | 0.0629 | 0.0000 | 0.1260 | TRUE |
| Sutter | 0.1037 | −0.0904 | 0.2959 | FALSE |
| Tehama | 0.0281 | −0.1529 | 0.2107 | FALSE |
| Tulare | −0.1348 | −0.2176 | −0.0507 | TRUE |
| Tuolumne | 0.0052 | −0.2235 | 0.2349 | FALSE |
| Ventura | 0.1017 | 0.0337 | 0.1686 | TRUE |
| Yolo | 0.1049 | −0.0117 | 0.2230 | FALSE |
| Yuba | 0.3045 | 0.1183 | 0.4964 | TRUE |

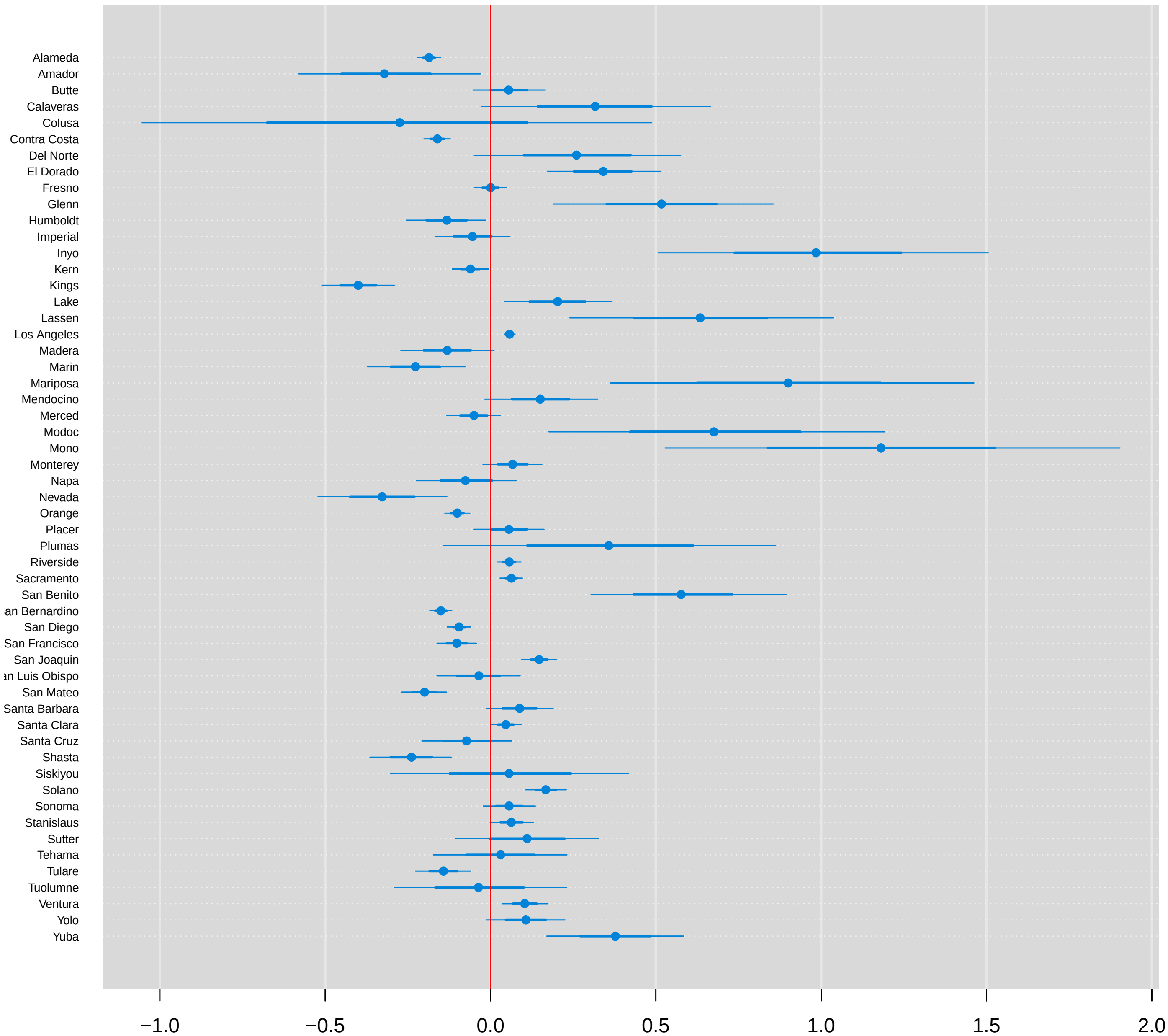


Figure 11: Line plots for Spatial 2 model, 95 percent credible intervals in each county for zipcode level analysis

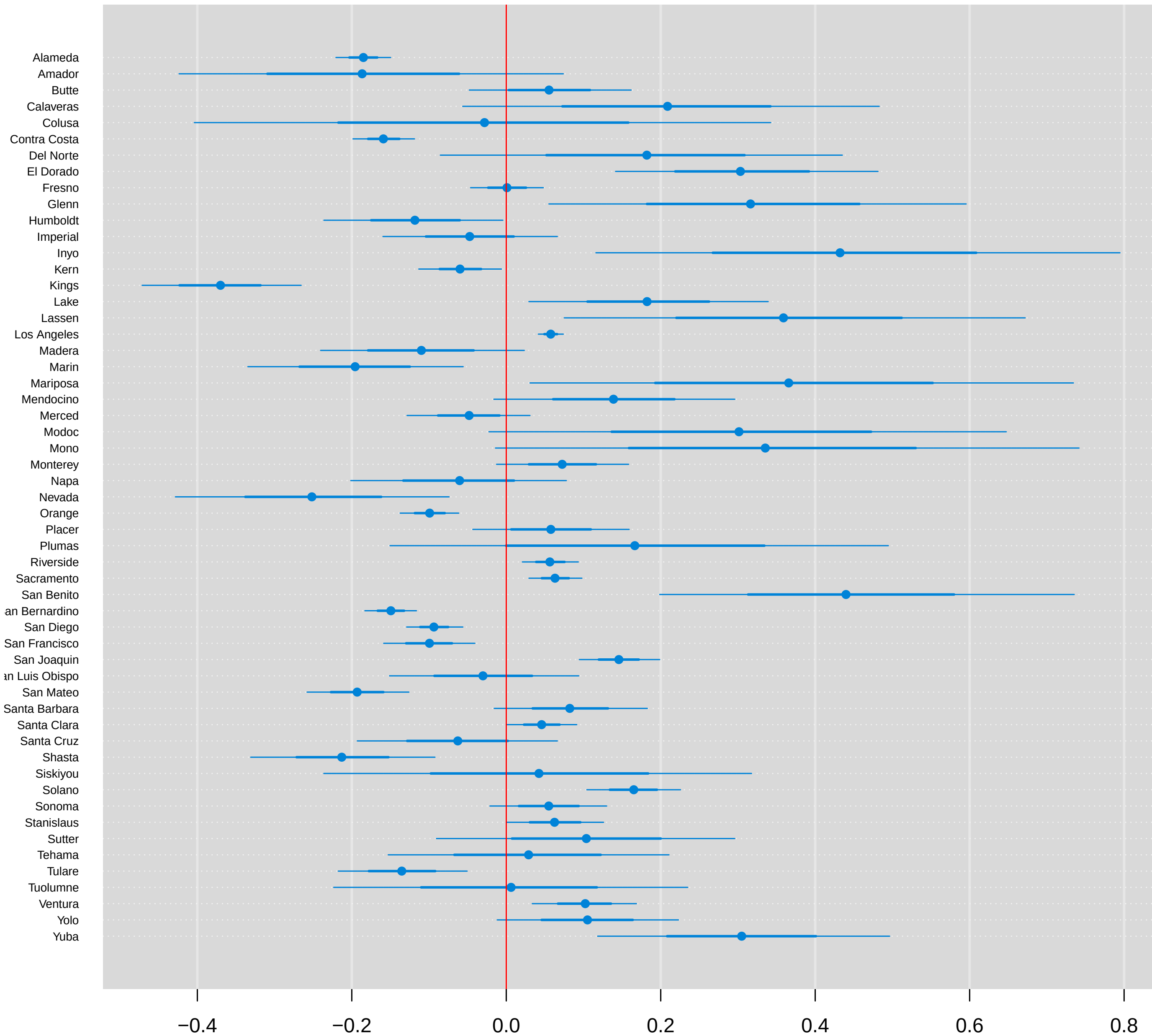


Figure 12: Line plots for Spatial 2 model, 95 percent credible intervals in each county for zipcode level analysis

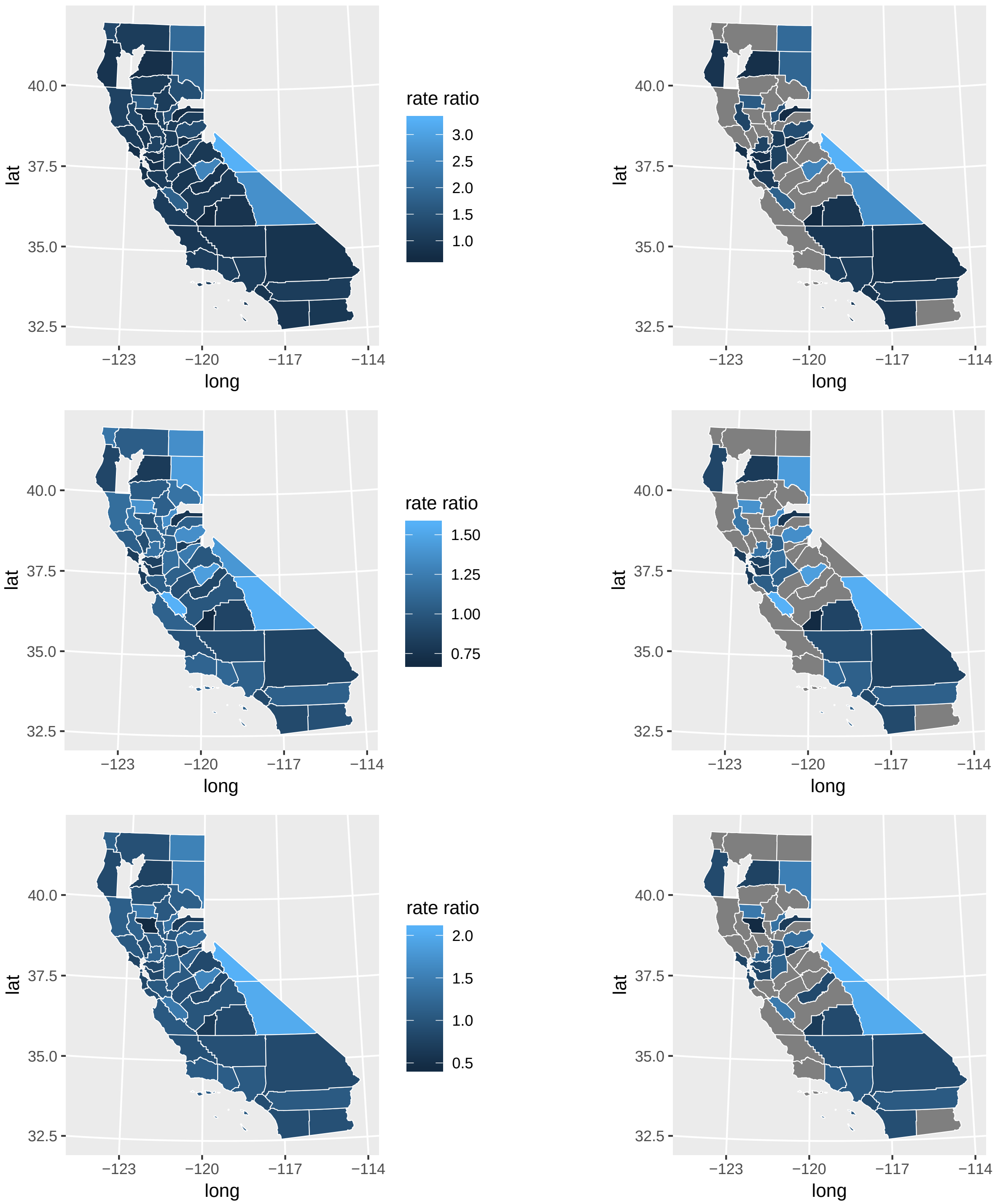


Figure 13: The plots in the first, second and third rows corresponding to Spatial 2, Spatial 1 and GLM. The left column denote the posterior mean or glm estimates of rate (exp(b3)). The right column denotes the significance of the estimated rates. The gray shaded areas were not significant. The results are shown for zipcode level analysis

Table 8: Estimated baseline ED visits with GLM, Spatial 2 and Spatial 1 models in zipcode level analysis

| | GLM | Spatial.2 | Spatial.1 |
|---|---|---|---|
| Alameda | 165.5610 | 165.1140 | 165.0296 |
| Amador | 43.0000 | 37.5679 | 34.8907 |
| Butte | 80.7500 | 79.2250 | 79.0344 |
| Calaveras | 33.5000 | 28.1742 | 29.5002 |
| Colusa | 27.0000 | 15.3461 | 13.1366 |
| Contra Costa | 162.7813 | 162.2823 | 162.0873 |
| Del Norte | 78.0000 | 67.9821 | 70.3823 |
| El Dorado | 54.6000 | 50.3922 | 50.8811 |
| Fresno | 97.3235 | 96.9086 | 96.8410 |
| Glenn | 30.5000 | 25.7270 | 28.5744 |
| Humboldt | 98.0000 | 96.3812 | 95.5408 |
| Imperial | 85.8571 | 84.6476 | 84.1370 |
| Inyo | 27.0000 | 19.8024 | 27.7977 |
| Kern | 111.2500 | 110.7742 | 110.6720 |
| Kings | 168.4000 | 166.0890 | 163.7971 |
| Lake | 40.1429 | 37.9801 | 38.1905 |
| Lassen | 47.0000 | 37.2311 | 43.3830 |
| Los Angeles | 113.5141 | 113.4812 | 113.4743 |
| Madera | 108.5000 | 104.6843 | 103.3096 |
| Marin | 55.8571 | 54.3346 | 53.4282 |
| Mariposa | 26.0000 | 16.3304 | 21.9019 |
| Mendocino | 88.6667 | 85.2430 | 85.4122 |
| Merced | 101.1667 | 100.1759 | 99.9741 |
| Modoc | 28.0000 | 21.4687 | 26.4190 |
| Mono | 14.0000 | 8.6808 | 14.7243 |
| Monterey | 77.7692 | 76.0842 | 75.6917 |
| Napa | 123.0000 | 118.3820 | 117.1909 |
| Nevada | 67.5000 | 63.6787 | 61.1578 |
| Orange | 76.5135 | 76.3848 | 76.3104 |
| Placer | 70.0000 | 68.5063 | 68.3217 |
| Plumas | 16.0000 | 12.4173 | 13.5770 |
| Riverside | 96.2542 | 96.0779 | 96.0467 |
| Sacramento | 153.0476 | 152.7338 | 152.6548 |
| San Benito | 123.0000 | 98.6943 | 105.1177 |
| San Bernardino | 144.6275 | 144.3952 | 144.3516 |
| San Diego | 90.6000 | 90.4814 | 90.4022 |
| San Francisco | 111.8095 | 111.0602 | 110.8751 |
| San Joaquin | 116.6087 | 116.0465 | 116.0964 |
| San Luis Obispo | 43.4545 | 42.6602 | 42.4246 |
| San Mateo | 104.8333 | 104.1615 | 103.7327 |

| | GLM | Spatial.2 | Spatial.1 |
|---|---|---|---|
| Santa Barbara | 57.0769 | 56.3315 | 56.3767 |
| Santa Clara | 81.2500 | 80.9686 | 80.9396 |
| Santa Cruz | 55.1250 | 53.7319 | 53.3580 |
| Shasta | 96.8333 | 93.9623 | 92.5660 |
| Siskiyou | 32.0000 | 28.1064 | 28.0214 |
| Solano | 198.2000 | 196.5272 | 196.6831 |
| Sonoma | 77.0000 | 76.3212 | 76.3129 |
| Stanislaus | 99.2222 | 98.5459 | 98.4954 |
| Sutter | 55.6667 | 51.7188 | 51.5292 |
| Tehama | 66.0000 | 60.9535 | 60.6089 |
| Tulare | 78.6000 | 77.7985 | 77.4156 |
| Tuolumne | 70.5000 | 61.6727 | 59.9248 |
| Ventura | 75.2000 | 74.8119 | 74.8252 |
| Yolo | 72.2857 | 70.5530 | 70.3850 |
| Yuba | 81.5000 | 74.5691 | 77.4863 |

# Appendix

## Original Zipcode model

$$y_{k,j,t} \sim \text{Pois}(\lambda_{j,k,t})$$

$$\log(\lambda_{j,k,t}) = b1_{j,k}\text{I(county = k)} + b3_{j,k}\text{I(t = 2017)}$$

$b1_{j,k} \sim N(\theta_{1,k}, \tau_{1,k})$
$b3_{j,k} \sim N(\theta_{2,k}, \tau_{2,k})$
$\theta_{1,k} \sim N(\theta_{o,1}, \Sigma_{\phi,\tau,\kappa})$
$\theta_{2,k} \sim N(\theta_{o,2}, \tau_3)$
$\tau_{1,k} = 1/\sigma^2_{\theta_{1,k}}$
$\sigma^2_{\theta_{1,k}} \sim$ dunif(0,1000)
$\tau_{2,k} = 1/\sigma^2_{\theta_{2,k}}$
$\sigma^2_{\theta_{2,k}} \sim$ dunif(0,1000)
$\tau_3 = 1/\sigma^2_{\theta_3}$
$\sigma^2_{\theta_3} \sim$ dunif(0,1000)
$\phi \sim$ dunif(0.05,50)
$\tau \sim$ Gamma(0.001,0.001)
$\kappa \sim$ dunif(0.05,1.95)

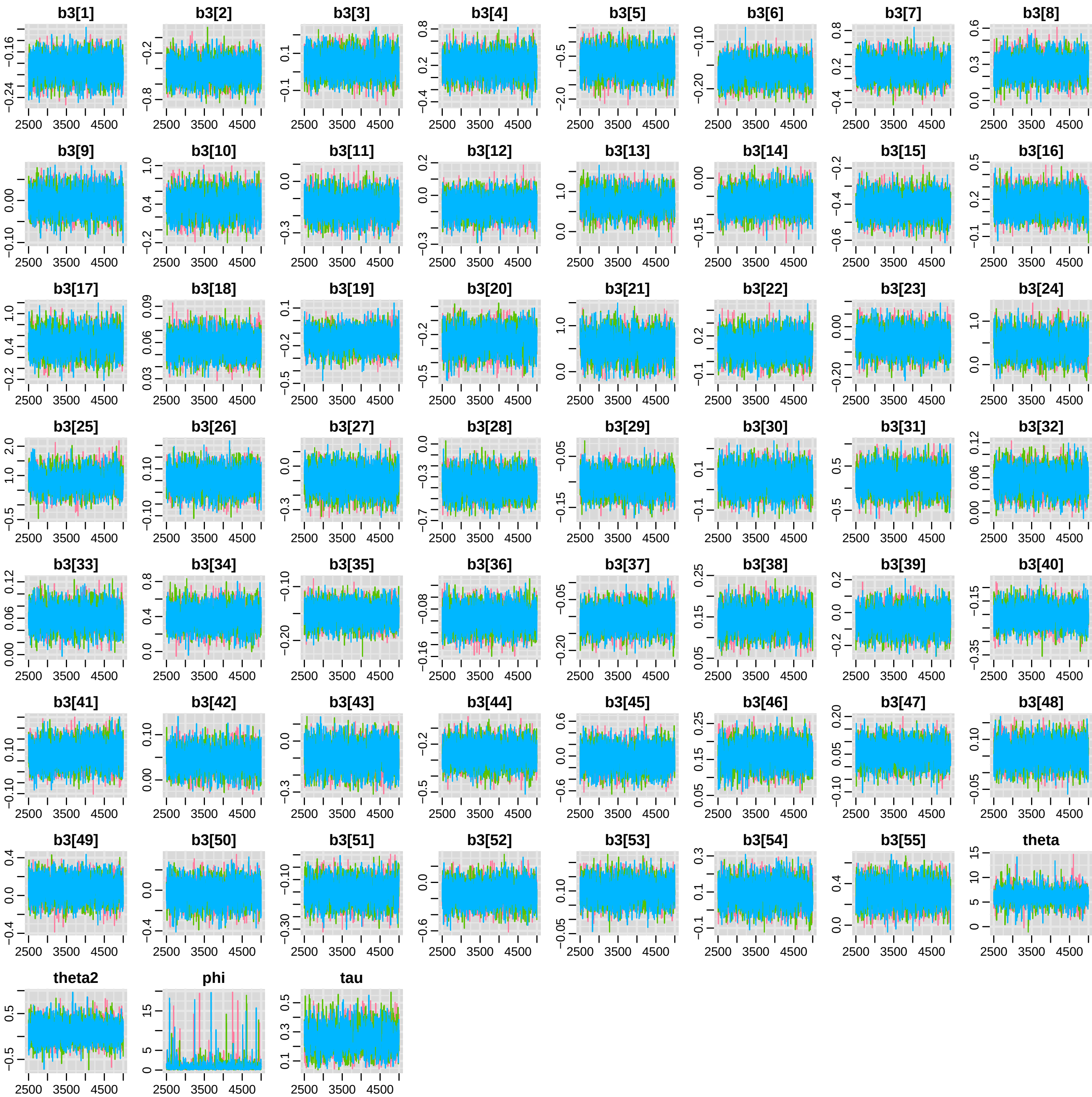

Figure 14: Convergence plots for b3, theta, theta2, phi and tau for Spatial 2 with informative priors at county level analysis

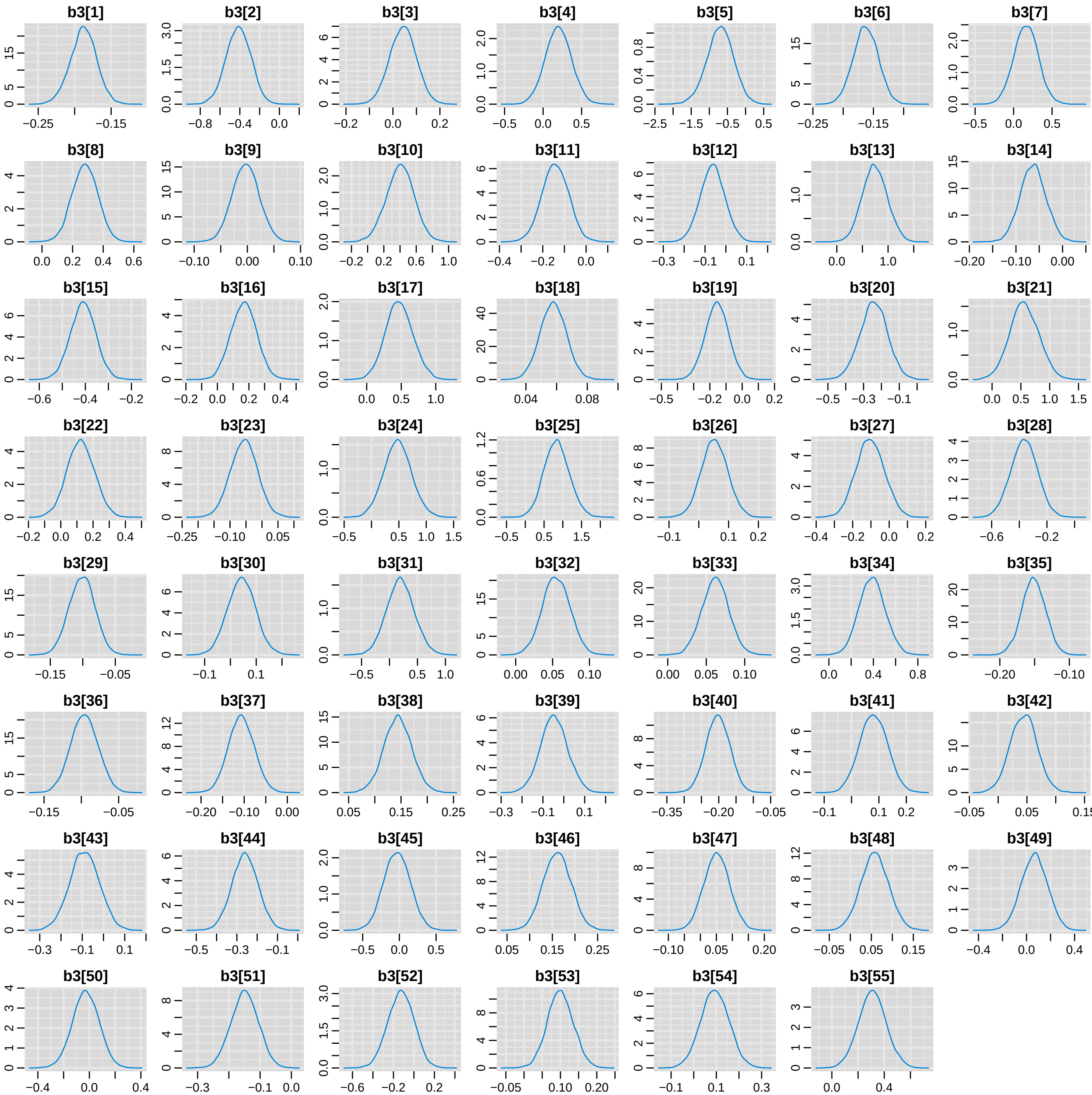

Figure 15: Density plots for Spatial 2, b3 in county level analysis with informative priors

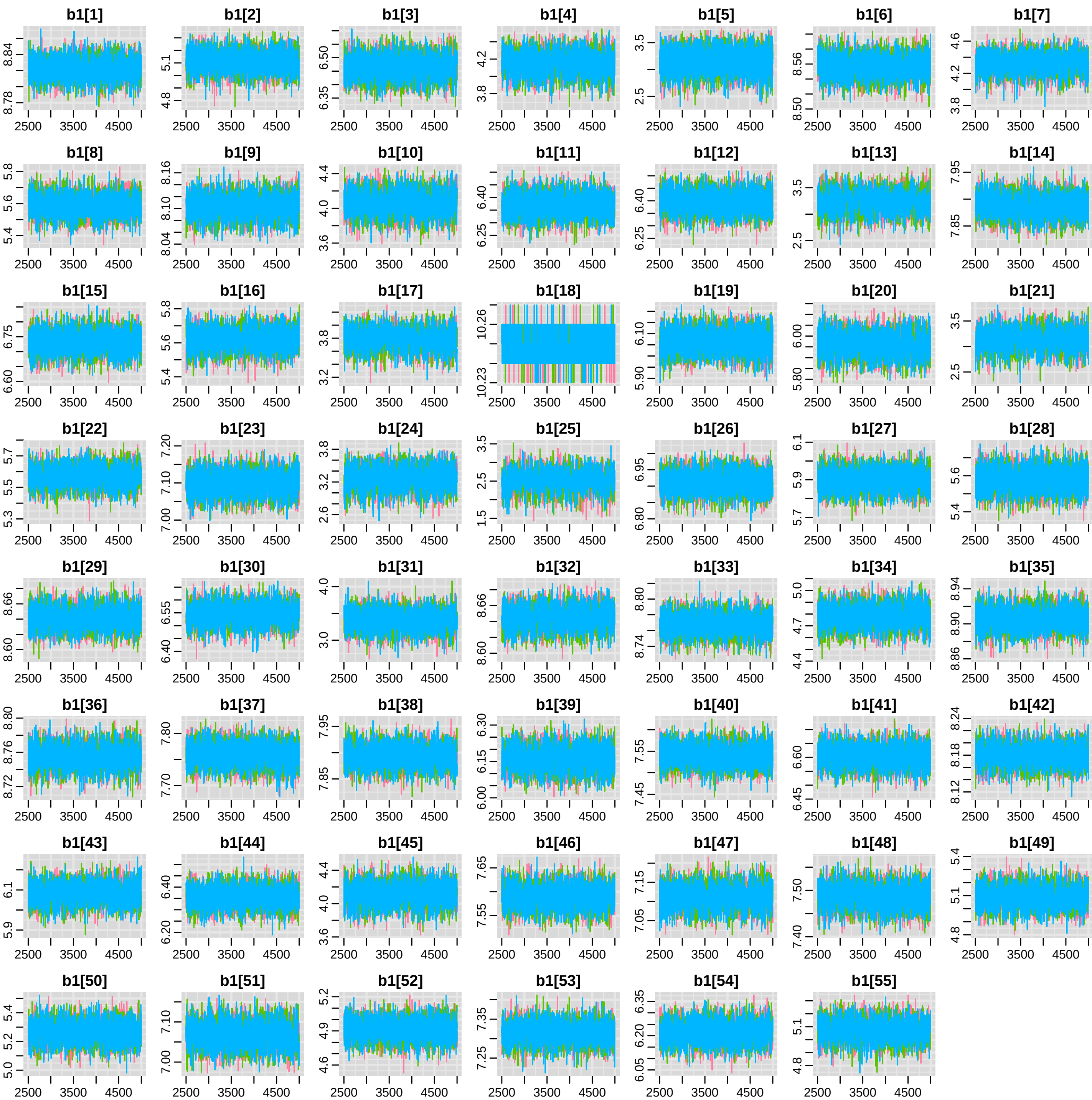

Figure 16: Convergence plots for b1 for Spatial 2 with informative priors at county level analysis

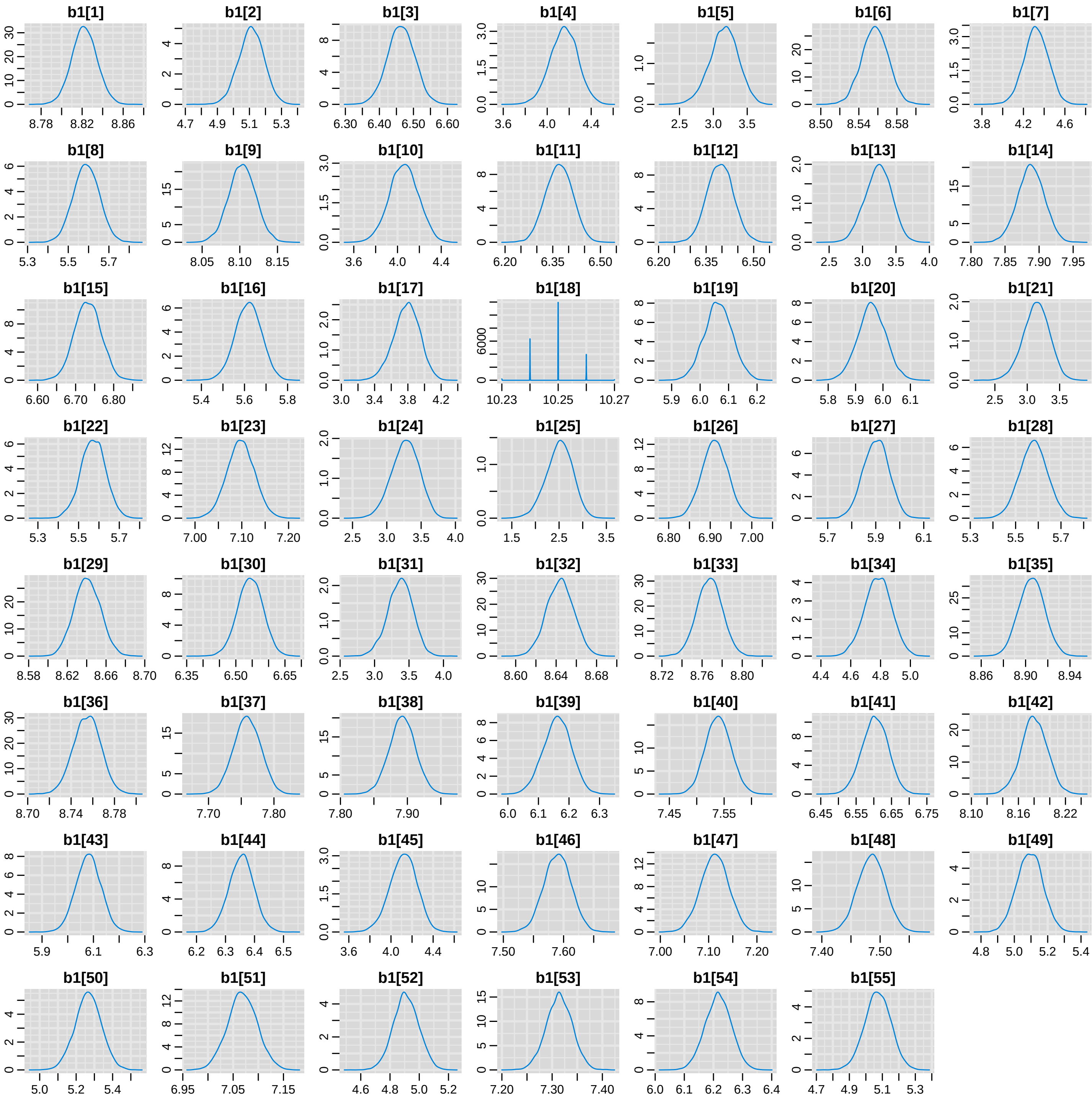

Figure 17: Density plots for Spatial 2, b1 in county level analysis with informative priors

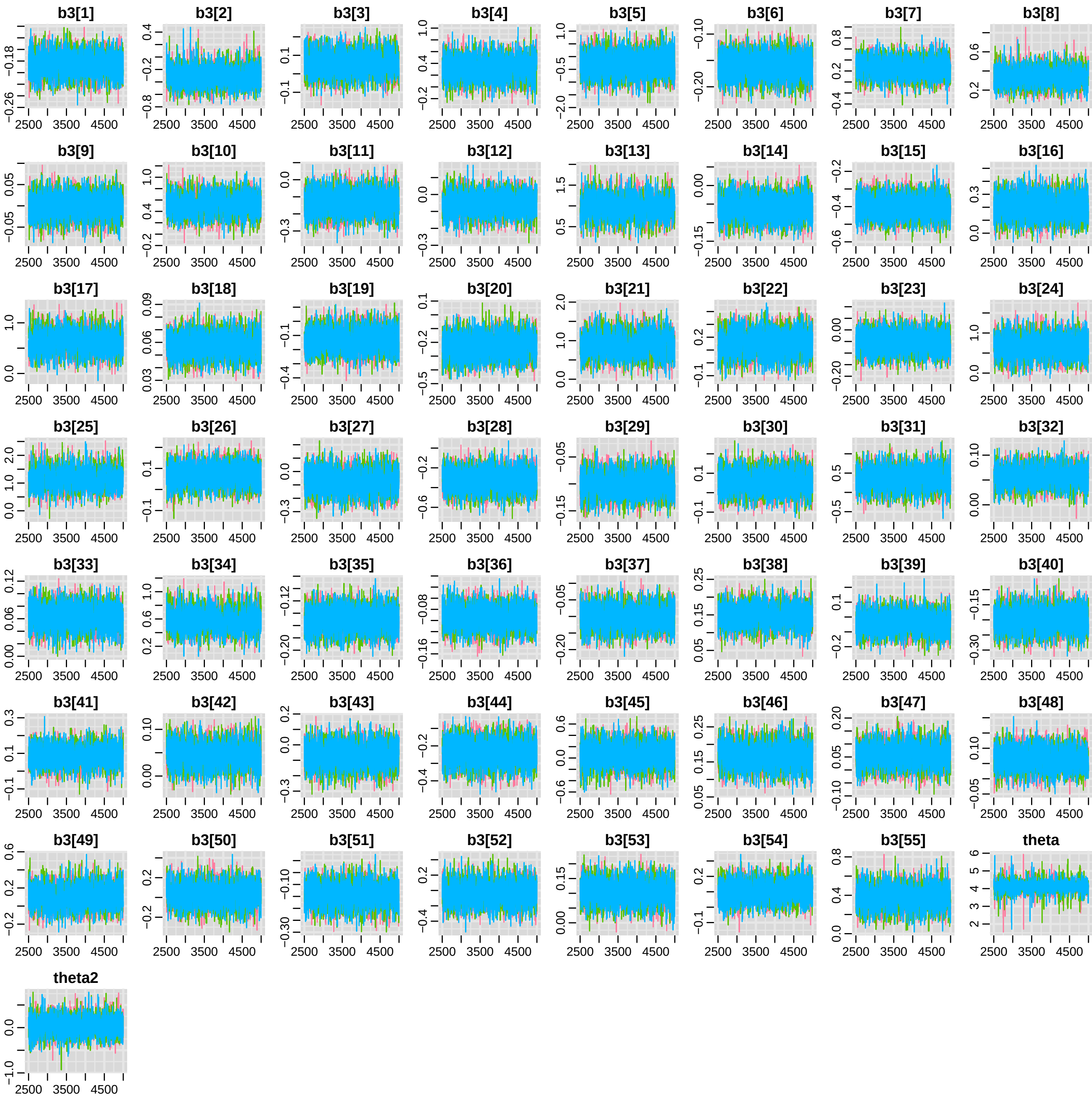


Figure 18: Convergence plots for b3, theta, theta2 and phi for zipcode level Spatial 2 with informative priors at zipcode level analysis

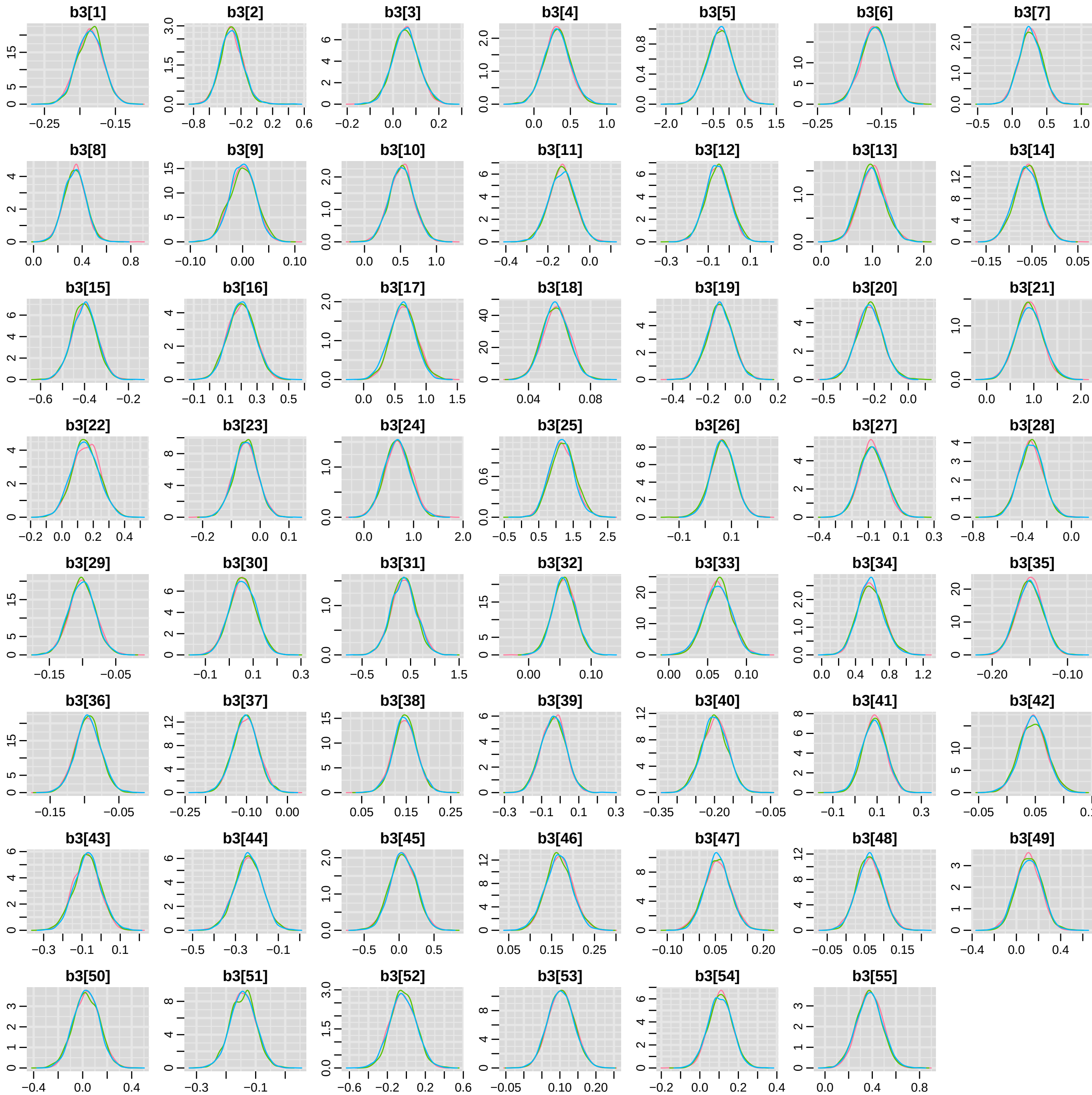

Figure 19: Density plots for Spatial 2, b3 in zipcode level analysis with informative priors

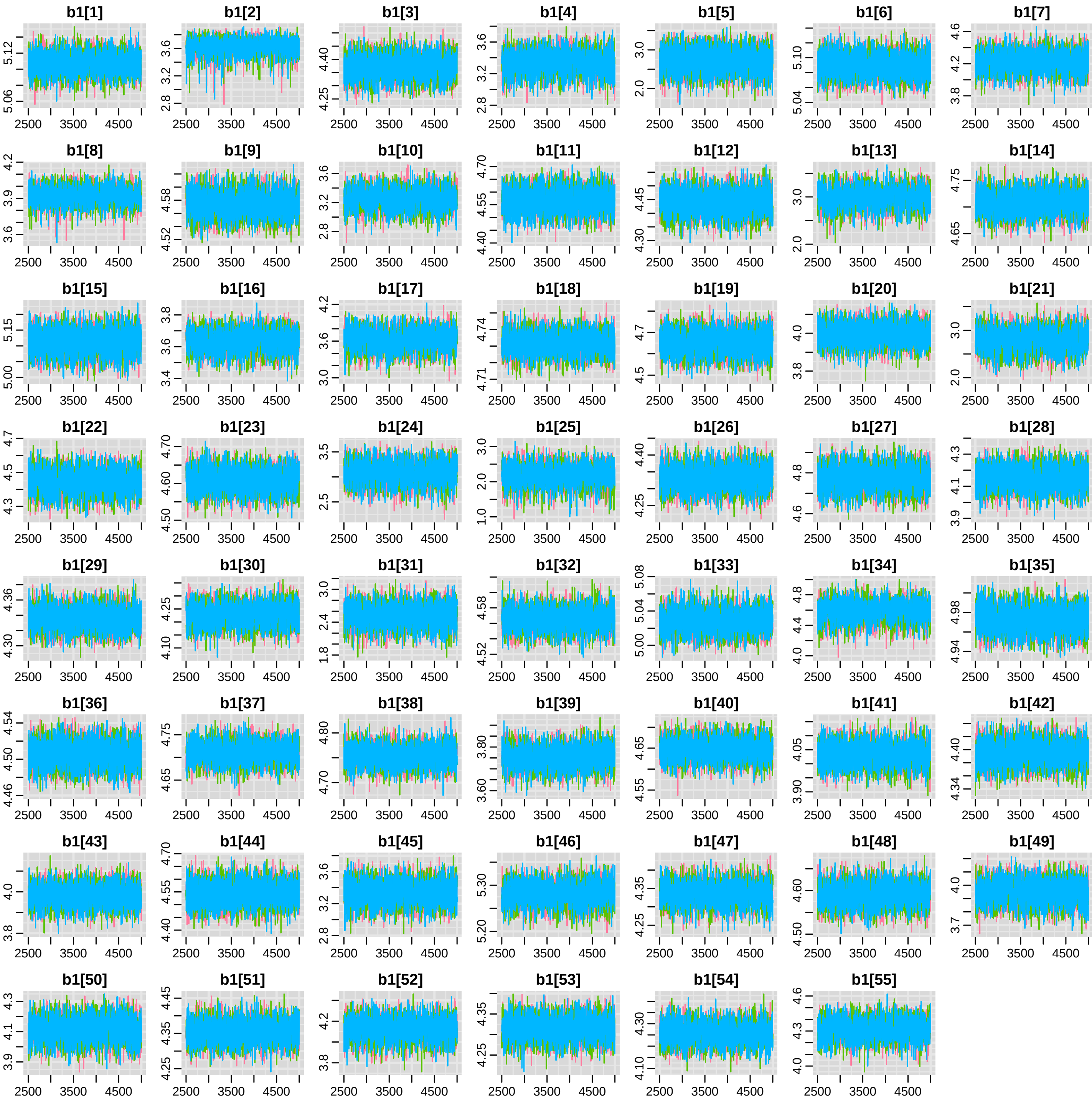


Figure 20: Convergence plots for b1 for zipcode level Spatial 2 with informative priors at zipcode level analysis

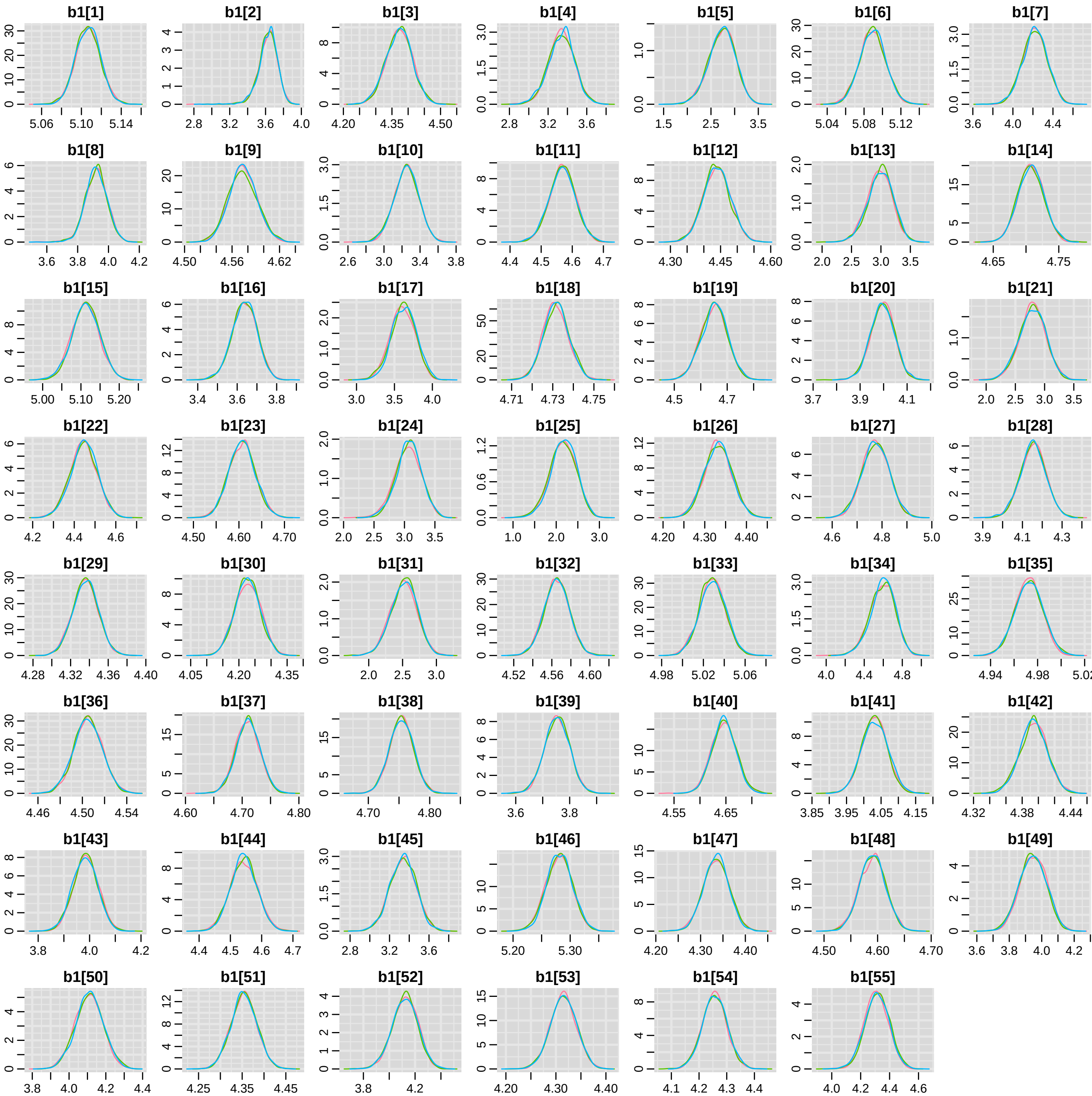


Figure 21: Density plots for Spatial 2, b1 in county level analysis with informative priors


Figure 22: Distribution of the predicted number of ED visits at baseline for Spatial 2, county level analysis were plotted along with each of the observed values (pink vertical line) and the corresponding posterior p-value. Since the observed number of ED visit counts are located in the center of each of the predictive distributions (larger posterior p-values), we can conclude that the model fits these data quite well.


Figure 23: Distribution of the predicted number of ED visits at year 2017 for Spatial 2, county level analysis were plotted along with each of the observed values (pink vertical line) and the corresponding posterior p-value. Since the observed number of ED visit counts are located in the center of each of the predictive distributions (larger posterior p-values), we can conclude that the model fits these data quite well.


Figure 24: Distribution of the predicted number of ED visits at baseline for Spatial 1, county level analysis were plotted along with each of the observed values (pink vertical line) and the corresponding posterior p-value. Since the observed number of ED visit counts are located in the center of each of the predictive distributions (larger posterior p-values), we can conclude that the model fits these data quite well.


Figure 25: Distribution of the predicted number of ED visits at year 2017 for Spatial 1, county level analysis were plotted along with each of the observed values (pink vertical line) and the corresponding posterior p-value. Since the observed number of ED visit counts are located in the center of each of the predictive distributions (larger posterior p-values), we can conclude that the model fits these data quite well.

## Supplemental documents

The R scripts and bug model scripts are saved in the Bayesian_final_report folder. There are 4 sub folders under the main folder. In the Datasets folder, the original dataset is asthma-ed-visit-rates-by-zip-code.csv. Other files were the cleaned dataset obtained with Data_cleaning.R under R_script folder. The R script folder contains the R scripts to run the bugs model in R2openbugs. The R scripts were separated into county-level and zipcode level. Each folder was then further separated into No spatial and Spatial analysis. The Model script folder contains the .txt files corresponding to each bugs R scripts in the R script folder. Lastly, the Models_saved folder contains the bugs models that were saved for convenience in reports generating.

## Reference

1. Asthma Prevalence in California, A Surveillance Report, January 2017

2. Asthma: Findings from the 2011ąV2012 California Health Interview Survey, Ulfat Shaikh and Robert S. Byrd, Population Health Management, 19, 2, 145-151, 2016

3. Spatial Data Analysis, Sudipto Banerjee, Annual Review of Public Health 2016 37:1, 47-60