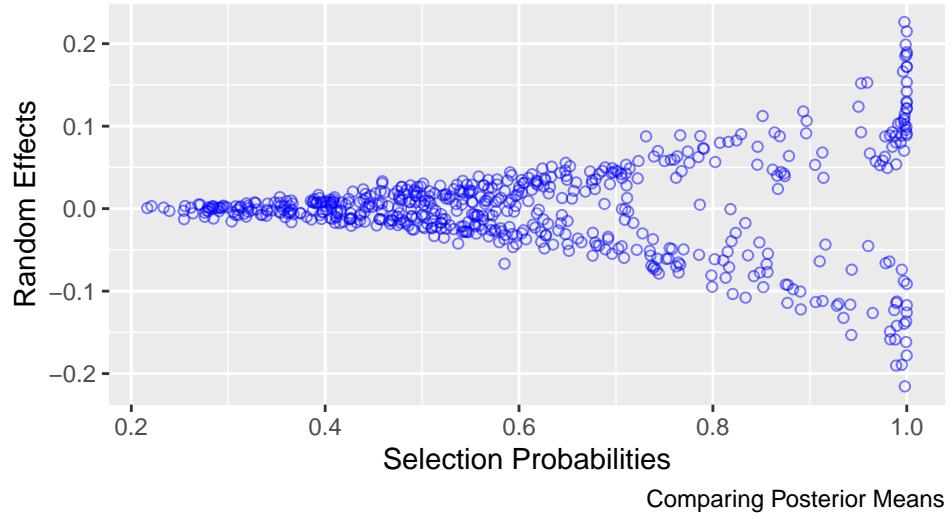


# Additional Analysis on Random Effects

Sho Kawano

2024-12-26

## SSD Model: Selection Probabilities vs Random Effects (S. Atlantic Census Division)



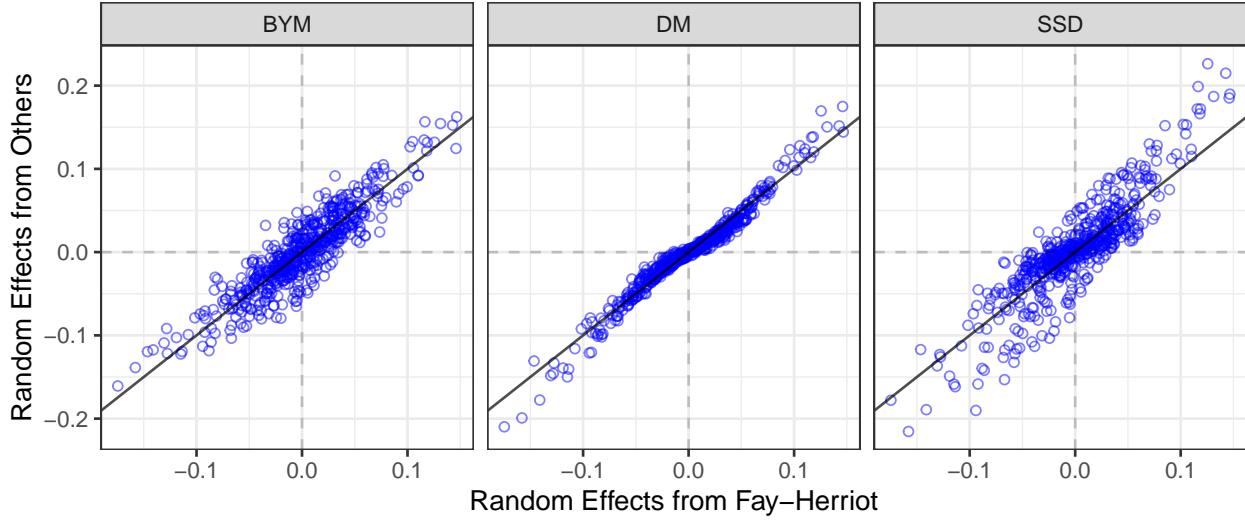
We can see that the areas where the selection probability is high (close to 1) also have larger magnitude random effects (both negative and positive). This makes sense given the posterior selection probability is

$$\hat{p}_i = \frac{p \cdot \phi(y_i | x_i^\top \beta + v_{1i} + v_{2i}, d_i)}{p \cdot \phi(y_i | x_i^\top \beta + v_{1i} + v_{2i}, d_i) + (1-p) \cdot \phi(y_i | x_i^\top \beta, d_i)}$$

for area  $i = 1, \dots, n$ . Thus, if the random effect is large for area  $i$ , the difference between  $\phi(y_i | x_i^\top \beta + v_{1i} + v_{2i}, d_i)$  and  $\phi(y_i | x_i^\top \beta, d_i)$  will be larger, impacting the selection probability.

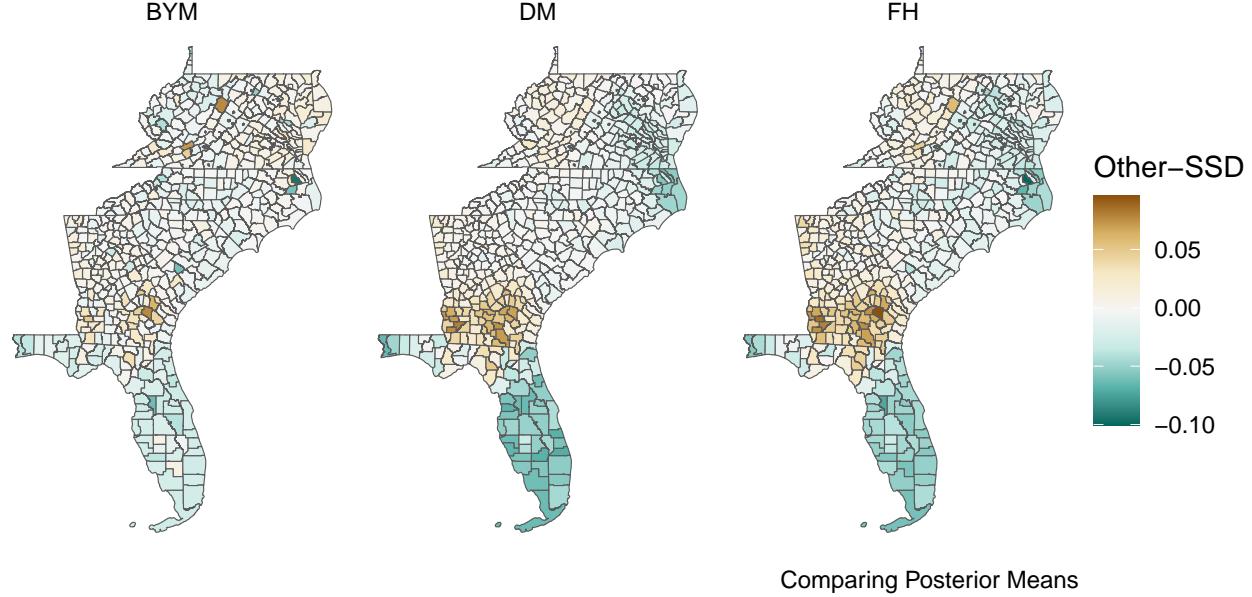
## Comparing Random Effects from Different Models

### Scatterplot Comparisons (Posterior Means)



We can see that the DM model most closely resembles the Fay-Herriot random effects, as they both assume independent and identically distributed (IID) random effects. The SSD model is somewhere between the BYM and the DM model, as expected. Of the models, SSD model allows for the biggest range in the random effect values.

### Mapping Random Effects from the Different Models

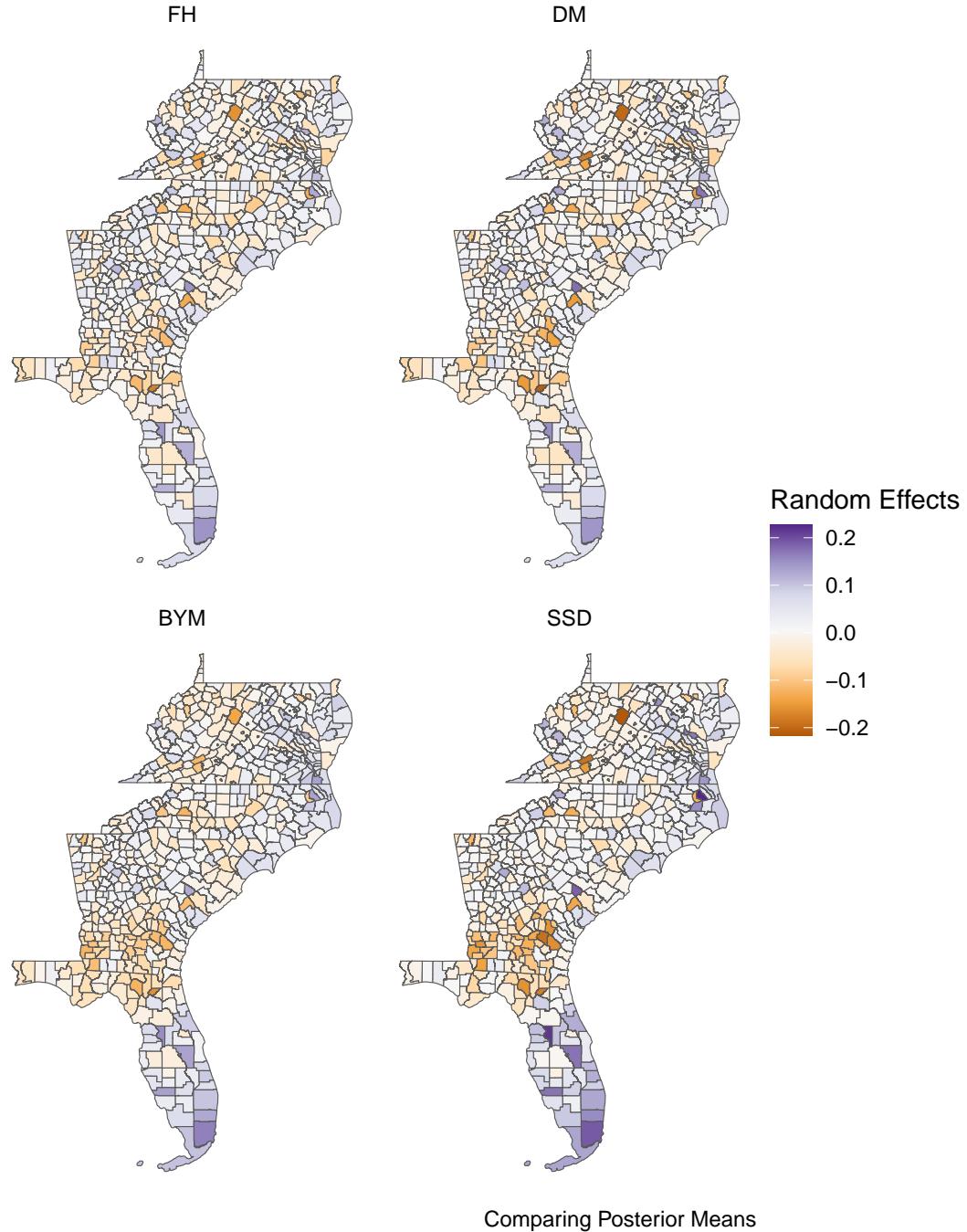


Comparing Posterior Means

Here is a comparison of the differences between the SSD random effects and the other models. As expected, difference compared to the IID models (DM & FH) have a strong spatial pattern. The difference between the BYM and the SSD is marked by a few spots of large differences. It is interesting, however, that the difference in the random effects between these two models has still has a spatial pattern down by Florida. We can also see that compared to the non selection models (FH & BYM), there are a few counties where SSD random effects are much larger in magnitude.

In fact, we can see from the map below that the both models with selection (DM & SSD) has larger magnitude

effects. This is due to these models not having a common variance assumption. This can be seen in other places in the census division.



### Comparing Posterior Means

Overall, we can see that all of the models have fairly similar spatial patterns. A Monte-Carlo simulation of a Geary's C test was performed on the random effects for the non-spatial models (FH & DM). Here are the p-values: 0.0004999 and 0.0009998.