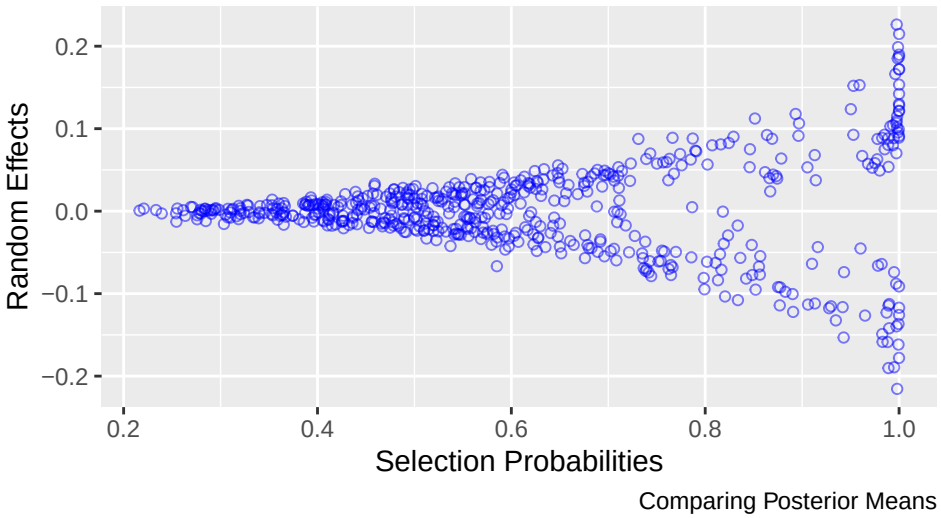
Additional Analysis on Random Effects

Sho Kawano

2024-12-26

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



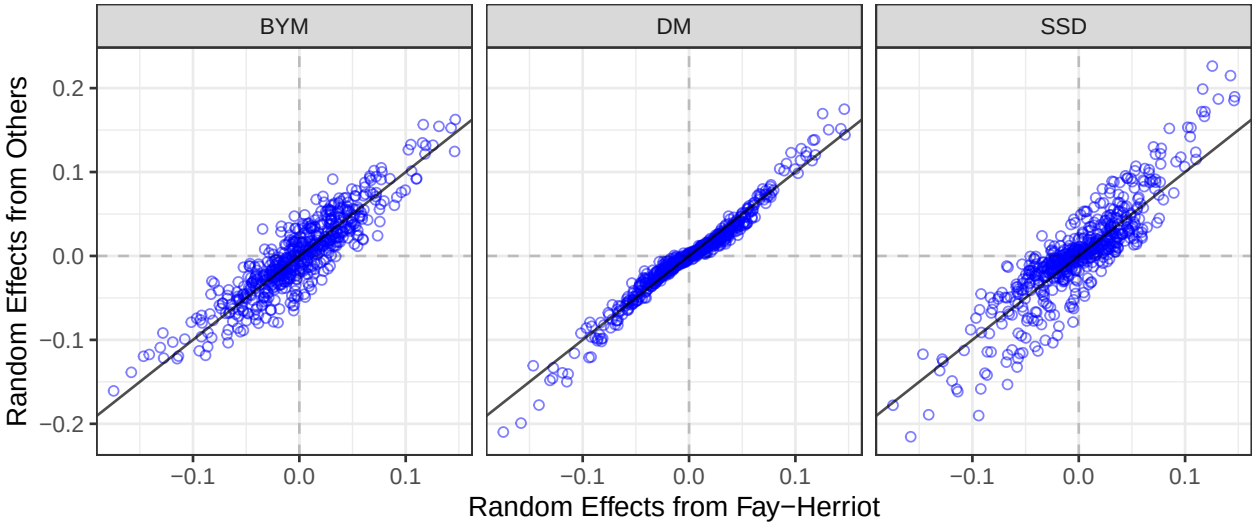
We can see that the areas where the selection probability is high (close to 1) also have larger magnitude random effects (both negative and positive). This makes sense given the posterior selection probability is

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

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

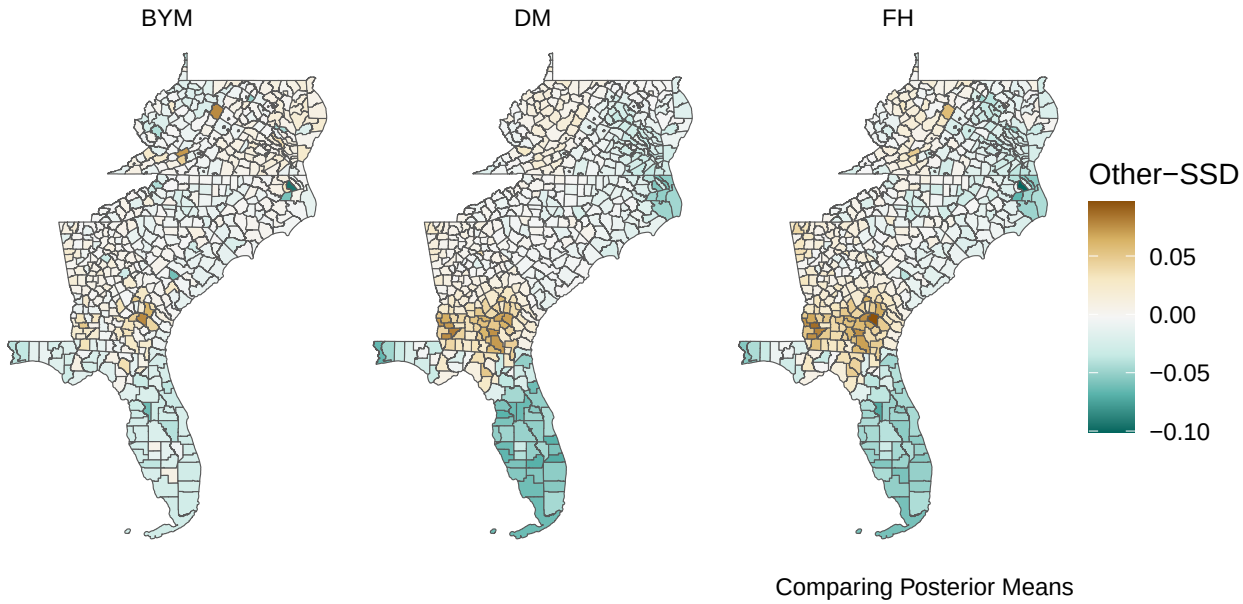
Comparing Random Effects from Different Models

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.

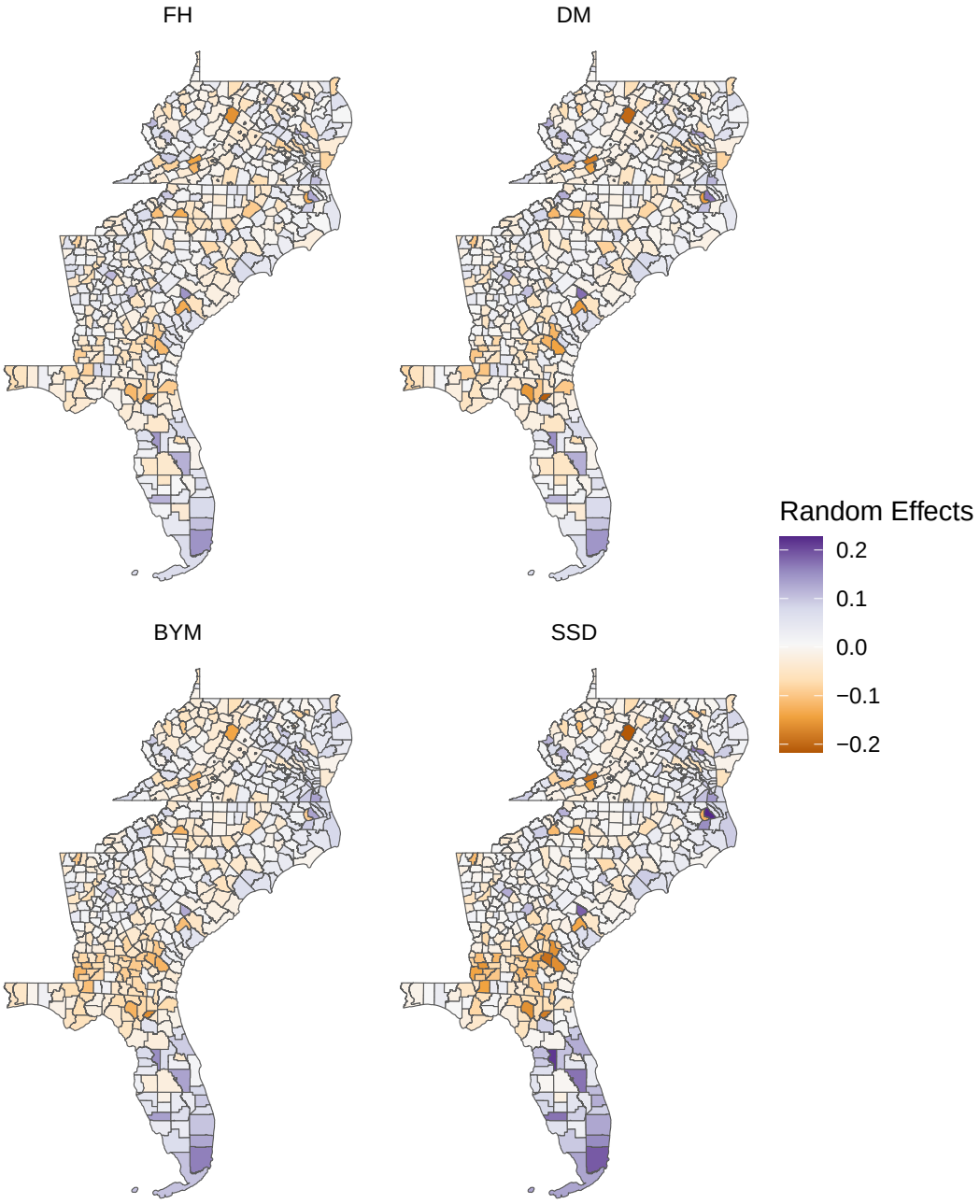
Mapping Random Effects from the Different Models



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 a spatial pattern down by Florida.

It seems that the BYM effects in Florida are drawn upwards by the presence of one county with very large effects in south Florida. This can be seen in the maps below. We can see that the models with selection

can allow for counties with very large magnitude effects without unnecessarily impacting the surrounding counties that may not need random effects. This can be see in other places in the census division.



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 Fay-Herriot and Datta-Mandal. Here are the p-values: 0.0004999 and 0.0009998.