# Estimating California County Ideology Using Multilevel Regression With Poststratification, Item-Response Theory, and Structured Priors

Bertrand Wilden

2024-04-03

## 1 Introduction

Political scientists often measure latent public opinion using item-response theory (IRT) models (Treier and Hillygus 2009; Caughey and Warshaw 2015). Individuals' observed responses to survey questions are used as inputs in these statistical models, which in turn yield estimates of the latent variable underlying these choices. For example, an IRT model might be constructed using survey responses relating to support of gun control, abortion rights, and tax policy to create a measure of public ideology. This assumes that there is some latent unidimensonal Left-Right, or Liberal-Conservative, scale upon which all members of the public can be placed.

Once the latent variable of public opinion has been estimated at the individual-level, it is common to aggregate these values up to some level of geography, such as state or congressional district (Caughey and Warshaw 2015). There may be many ways, however, that a particular survey sample will not match the target population. The most effective way to mitigate this type of bias is to apply multilevel regression with post-stratification (MRP) when performing the aggregation step

(Downes et al. 2018).

I improve upon standard MRP-IRT models by incorporating *structured prior distributions*. Typically the "multilevel" part of the MRP model treats the hierarchical intercepts of the categorical covariates as independently normally distributed. Instead, I explicitly model underlying structure in these covariates. For the ordinal *age* category I use an autoregressive prior component. This results in an increase of information sharing among ordinal age categories that are close together. Intuitively, the ideology of someone in the "55-60 year" group should be informed more by the ideology of those in the "60-65 year" group, rather than by the "18-25 year" group. In addition to the structured prior component for *age*, I add a component to the model to account for *spatial* structure. Due to underlying network effects, geographic units should pool more information with one another if they are more proximate. Gao et al. (2021) use simulation studies to show how adding these structured prior distributions to MRP models helps match the sample estimates to the target population.

In this project I combine the IRT, MRP, and structured prior distributions together into a single model to estimate the ideology of counties in California using survey data from the Cooperative Election Study (CES).[1] California counties offer a good setting for this method because they vary widely in terms of population (9.83 million in Los Angeles county compared to 1,200 in Alpine county). MRP offers a principled method for sharing information between large and small counties, thereby increasing the validity of small county estimates. And the spatial structure com-

---

[1] Kuriwaki, Shiro, 2023, "Cumulative CES Common Content", https://doi.org/10.7910/DVN/II2DB6, Harvard Dataverse, V8

ponent ensures that more of this information sharing comes from neighboring counties which share (unobserved) underlying similarities.

## 2 IRT

The standard latent variable Bayesian IRT model is shown in Equation 1. In the survey context, the outcomes $y_{ij}$ are respondent $i$'s answer to question $j$. For the CES survey this could correspond to a Support/Oppose response to the question "Require criminal background checks on all gun sales". The parameter $\theta_i$ corresponds to the value of the latent variable (ideology in this example) for respondent $i$. Parameters $\gamma_j$ and $\xi_j$ roughly correspond to the ideological component and non-ideological component of the question, respectively. The cumulative density function $\Phi$ transforms the nonlinear parameter terms into a probability value, which is then passed through a Bernoulli distribution to arrive at the observed 1's and 0's—a respondent's answer of support/oppose to a specific survey question.

$$y_{ij} \sim \text{Bernoulli}[\Phi(\gamma_j \theta_i + \xi_j)] \tag{1}$$

As it stands, the model in Equation 1 is not identified. In other words, given specific data $y_{ij}$, there is not a unique set of values for $\theta_i$, $\gamma_j$, or $\xi_j$. Adding (or multiplying) a constant to the term $\gamma_j \theta_i$ can be offset by subtracting (or dividing) a constant from $\xi_j$. Furthermore, there is no way to tell if $\gamma_j$ and $\theta_i$ should be positive or negative. These identifiability issues follow from understanding $\theta$'s nature as a latent variable. Latent variables, such as *ideology*, have no inherent

3

scale (the multiplicative problem), or center (the additive problem). And there is no inherent polarity for a latent variable (the positive/negative problem).

When we estimate an IRT model using Bayesian methods, the multiplicative and additive identification problems can usually be solved using strong priors. Setting a relatively narrow Normal(0, 1) prior distribution on $\theta$, $\gamma$, and $\xi$ constrains the set of likely values for these parameters. This helps avoid the multiplicative and additive problems by setting a default center and range before the model is introduced to data (Bürkner 2020). The positive/negative problem can also be solved using priors if we know the polarity of the responses to the survey questions. Once we re-code all the responses such that Support/1 reflects the conservative position on a question and Oppose/0 reflects the liberal position on a question, the parameter $\gamma$ can be constrained to be strictly positive using priors. This will then produce estimates of $\theta_i$ with negative values corresponding to more liberal ideology and positive values correspond to more conservative ideology.

In the practice of fitting the combined IRT-MRP-structured priors model, I found that the narrow Normal(0, 1) priors on $\theta$, $\gamma$, and $\xi$ were not enough to fix the multiplicative/additive identification problems. The benefit to fitting Bayesian models using Stan's Hamiltonian Monte Carlo (HMC) sampler is that these identification issues become immediately obvious(Betancourt 2018). HMC runs four (default) parallel chains to explore the log probability density of the posterior distribution. While each chain might start in different regions of this density, ideally each will converge to the same posterior distribution by the time the model finishes sampling. The HMC diagnostic statistic Rhat is a measure of the ratio of total variability among chains to variability within chain. When confronted with a poorly identified model, the HMC chains will have a difficult time agreeing

4

on where the ultimate posterior should be—thus resulting in high Rhat values.

To fix these sampling issues I generate standardized versions of each of the three IRT parameters. This is done by subtracting the mean value from each and dividing by the standard deviation, e.g. $\theta_i^{adj} = \frac{\theta_i - \bar{\theta}}{s_\theta}$. Now the IRT parameters have mean $= 0$ and standard deviation $= 1$. This sets a default center and scale for the model, thereby fixing the additive and multiplicative identification issues. The IRT model used in the final combined MRP-IRT model is Equation 2.

$$y_{ij} \sim \text{Bernoulli}[\Phi(\gamma_j^{adj}\theta_i^{adj} + \xi_j^{adj})] \tag{2}$$

## 3 MRP

Multilevel-regression and poststratification (MRP) is a method for improving the way survey estimates generalize to the target population. MRP can help correct this mismatch between the survey sample and target population in at least two ways. First, the survey may suffer from some form of bias. Among the most common forms of survey bias comes from selective non-response, whereby individuals with certain unobserved characteristics opt out of the sample. If these unobserved characteristics also affect how someone might respond to a survey, the results of the survey will be biased. MRP can help ameliorate survey bias by weighting respondents according to their demographic prevalence in the overall population.

Second, MRP can help improve the reliability for small group-level aggregate estimates. Even a perfectly representative and unbiased survey may only only have a few respondents from sparsely

5

populated geographic units. If inference about these small geographic units is an important research goal, random noise in the sample may yield unreliable results. The hierarchical modeling step in MRP allows small, noisy groups to borrow information from larger groups.

## 3.1 Multilevel Regression

The "multilevel" part of MRP refers to modeling the poststratification covariates (geographic unit, age category, race category, etc) hierarchically. Rather than modeling these variables as "fixed effects", hierarchical models treat each category as drawn from a common distribution. This allows categories to partially pool information among each other—in contrast to the "fixed effects" approach in which no pooling of information between groups occurs. Hierarchical models impose regularization on the model which can greatly improve out of the sample predictions (McElreath 2020).

Building off the IRT model from Equation 2, we can add hierarchical terms to help us both estimate $\theta_i$ and use poststratification in the next step. In Equation 3, the categories in the hierarchical terms $A_c^{county}$, $A_a^{age}$, $A_e^{education}$, $A_r^{race}$, $A_h^{income}$ are each drawn from a common distribution. Continuous variables such as age and income must be discretized into categorical variables in order to be used in the poststratification step. I do not model gender hierarchically because the Census, from which the poststratification table is built, treats this variable as binary.

$$\theta_i \sim \text{Normal}(\mu^\theta + A_{c[i]}^{county} + A_{a[i]}^{age} + A_{e[i]}^{education} + A_{r[i]}^{race} + A_{h[i]}^{income} +$$

$$B^{gender} * \text{Gender}_i, \sigma^\theta)$$

$$A_c^{county} \sim \text{Normal}(0, \sigma^{county})$$

$$A_a^{age} \sim \text{Normal}(0, \sigma^{age})$$

$$A_e^{education} \sim \text{Normal}(0, \sigma^{education})$$

$$A_r^{race} \sim \text{Normal}(0, \sigma^{race})$$

$$A_h^{income} \sim \text{Normal}(0, \sigma^{income})$$

$$B^{gender} \sim \text{Normal}(0, 2)$$

$$\sigma^{county}, \sigma^{age}, \sigma^{education}, \sigma^{race}, \sigma^{income} \sim \text{Normal}_+(0, 1)$$

$$(3)$$

The IRT terms related to survey questions, $\gamma_j$ and $\xi_j$, can also be modeled hierarchically. There are no relevant question covariates to include here, but modeling these terms hierarchically still improves the model by allowing questions to pool information with each other. Note that this is also where we can constrain $\gamma_j$ to be positive by using the $\text{Normal}_+$ distribution in the prior.

$$\gamma_j \sim \text{Normal}_+(0, \sigma^\gamma)$$

$$\xi_j \sim \text{Normal}(0, \sigma^\xi)$$

$$\sigma^\gamma, \sigma^\xi \sim \text{Normal}_+(0, 1)$$

The final component in the hierarchical IRT model is a sub-model for counties. In Equation 3 $A_c^{county}$ is drawn from $\text{Normal}(0, \sigma^{county})$, but the model can be improved by adding additional covariates. Counties in California are grouped into five regions: North-Central Valley, Central and Southern, Bay Area, Southern, and Northern Mountain. Just as the respondent-level model for

7

$\theta_i$ benefits from taking into account which county the respondent lives in, the county-level sub-model benefits from taking into account which region the county belongs to. Lastly, I include the county-level 2020 general election Republican vote share.[2]

$$
\begin{aligned}
A_c^{county} &\sim \text{Normal}(A_{n[c]}^{region} + B^{RepVote} * \text{RepVote}_c, \sigma^{county}) \\
A_n^{region} &\sim \text{Normal}(0, \sigma^{region}) \\
B^{RepVote} &\sim \text{Normal}(0, 2)
\end{aligned}
\tag{4}
$$

## 3.2 Poststratification

Now that we have a regression model for estimating each respondent's ideal point, $\theta_i^{adj}$ in the survey data, we can proceed with the poststratification step of MRP. During poststratification we are weighting each value of $\theta_i^{adj}$ in a county by the number of individuals in a county (based on the Census) who match the same demographic characteristics. This gives us an estimate of ideology per county, $\theta_c^{MRP}$ from Equation 5.

$$
\theta_c^{MRP} = \frac{\sum_{i \in c} N_i \theta_i^{adj}}{\sum_{i \in c} N_i}
\tag{5}
$$

In Equation 5, $N_i$ corresponds to the number of people in a particular cell of the poststratification table. These cells are created by counting the number of individuals in the Census who have a set of particular characteristics. For example, as we see in the first row of Figure 1 there are 43 white male individuals between the ages of 18 and 22 with no high school education earning between \$30,000

---

[2]MIT Election Data and Science Lab, 2018, "County Presidential Election Returns 2000-2020", https://doi.org/10.7910/DVN/VOQCHQ, Harvard Dataverse, V12.

and $100,000 who live in Alameda county.

Figure 1: Example Poststratification Table

| County | Age | Gender | Education | Race | Income | n |
|--------|-----|--------|-----------|------|--------|---|
| Alameda | 18-22 | Male | No HS | White | $30k-100k | 43 |
| Alameda | 18-22 | Male | No HS | White | $100k+ | 66 |
| Alameda | 18-22 | Male | No HS | Black | $0-30k | 72 |
| Alameda | 18-22 | Male | No HS | Black | $30k-100k | 29 |
| Alameda | 18-22 | Male | No HS | Black | $100k+ | 51 |
| Alameda | 18-22 | Male | No HS | Hispanic | $0-30k | 40 |
| Alameda | 18-22 | Male | No HS | Hispanic | $30k-100k | 326 |

Constructing a poststratification table involves a few practical considerations. Continuous variables such as age and income must be broken up into discrete categories in order to have reliable counts for individuals in each cell. Increasing the number of categories increases the precision of the regression model. But because the poststratification table is itself constructed using a survey—in this case the Census's 5-year 2017-2022 American Community Survey (ACS)—it is possible to have unreliable estimates if there are too few individuals in each cell. In this project I treat the ACS population counts as if they were a complete Census, but future work could incorporate the uncertainty of these measurements into the larger model.

The second downside to including too many categories for continuous variables in the poststratification table is computational. Each cell in the poststratification table corresponds to a unique combination of variable categories, so increasing these categories can cause the table to explode in size quickly. Calculating $\theta_c^{MRP}$ can then become computationally difficult. For this project, I landed on constructing a poststratification table with the 58 California counties $\times$ 28 age categories[3]

---

[3]The age variable was split into relatively many categories to help the structured prior component.

$\times$ 2 gender categories $\times$ 5 education categories $\times$ 7 race/ethnicity categories $\times$ 3 income categories for a total of 341,040 cells.[4]

# 4 Structured Priors

While IRT models have been combined with MRP before (Caughey and Warshaw 2015), the primary contribution of this project is including structured prior distributions to the hierarchical ideal point model. In the context of MRP, these structured priors have been shown to dramatically reduce non-response bias (Gao et al. 2021). They also impose additional regularization to the IRT MRP model, thereby helping with out-of-sample prediction, in a much more interpretable and intuitive way than traditional machine learning approaches (Bisbee 2019).

## 4.1 Directed Structured Priors

The baseline hierarchical ideal point model in Equation 3 treats the varying intercept for each age category, $A_a^{AgeCat.}$ as independently drawn from a mean-zero Normal$(0, \sigma^{AgeCat.})$ distribution. We can improve upon this by using a *first order autoregressive* specification for age (Gao et al. 2021). Instead of each age category intercept being partially pooled towards the global regression center, we increase the information shared between the each category and its previous ordinal category. In other words, we expect proximate age groups to be more informative regarding the ideology of each respondent. The autoregressive structured prior in Equation 6 accomplish this without imposing

---

[4]This poststratification table was extremely sparse, with a majority of cells with zero individuals. Because the weight of these cells is zero, they do not contribute to the calculation of $\theta_c^{MRP}$. These cells, along with those for counties not found in the survey data (Alpine and Trinity), were removed from the final poststratification table to avoid wasting computational time. The final number of cells was 79616

any type of linear assumption on age. To ensure identification of these parameters we also constrain their sum to equal one.

$$A_a^{AgeCat.} \sim \text{Normal}(A_{a-1}^{AgeCat.}, \sigma^{AgeCat})$$
$$\sigma^{AgeCat} \sim \text{Normal}(0, 1) \tag{6}$$
$$\sum_{a=1}^{a} A_a^{AgeCat.} = 1$$

In theory, a similar structured prior could be applied to the two other ordinal variables in the hierarchical ideal point model: education and income level. Gao et al. (2021) find that the benefit to these types of priors only starts to kick in after splitting the variable into 12 or more categories. Given the computational issues discussed previously with adding too many cells to the poststratification table, I opt to restrict education and income to only a few categories each and use standard mean-zero hierarchical priors instead. Subsequent research could investigate the benefits associated with using multiple directed structured priors in MRP models.

## 4.2 Undirected Spatial Priors

In addition to the autoregressive structured prior for age in the model, I add a component which helps account for spatial structure in the data. This comes from a class of priors known as Conditional Autoregressive (CAR) models. In the context of our IRT-MRP model, the intuition behind this prior is that the ideology of each county in California should be informed more by its geographic neighbors than simply by the global regression average.

The first step to adding a CAR component to our model is constructing a $N \times N$ adjacency

11

matrix of the $N$ counties in California. For each binary relationship, $c \sim k$, the values in the adjacency matrix are 1 if $n_c$ and $n_k$ are neighbors and are 0 otherwise. Relationships are not reflexive, so $c \neq k$.

Next we define the spatial random variable $\phi_c$ as:

$$\phi_c \sim \text{Normal}(0, [D-W]^{-1})$$

Where $D$ is the $N \times N$ adjacency matrix described above, and $W$ is a $N \times N$ matrix where the diagonal elements are the number of neighbors the $c$th county has and all off-diagonal elements are 0. Morris et al. (2019) show that the prior for $\phi_c$ can be rewritten as Equation 7, which is the parameterization I use in the final model.

$$p(\phi_c) = -\frac{1}{2}\left(\sum_{c \sim k}(\phi_c - \phi_k)^2\right) \tag{7}$$

The CAR model in Equation 7 imposes a strict assumption on the model that hierarchical county intercepts should only be informed by their neighbors. We can relax this by using a mixing parameter, $\rho \in [0,1]$ which controls how much weight to give the spatial structure versus the standard, unstructured hierarchical county effects. The Besag, York, and Mollié model (BYM2) in Equation 8 is a frequently used method to accomplish this [Besag, York, and Molli (1991)].

$$A_c^{county} = \sqrt{1-\rho} \times \alpha_c + \sqrt{\rho/s} \times \phi_c$$
$$p(\phi_c) = -\frac{1}{2} \left( \sum_{c \sim k} (\phi_c - \phi_k)^2 \right)$$
$$\sum_{c=1}^{58} \phi_c = 0 \tag{8}$$
$$\alpha_c \sim \text{Normal}(0, I)$$
$$\rho \sim \text{Beta}(0.5, 0.5)$$

In Equation 8 the hierarchical county intercepts used in the final IRT-MRP model, $A_c^{county}$ are now a convolution between the unstructured effects, $\alpha_c$ and the spatial effects $\phi_c$. The constant, $s$ is a scaling factor which enforces $Var(\alpha_c) \approx Var(\phi_c) \approx 1$. As in the autoregressive prior for the age hierarchical effects, $\phi_c$ is not identified so we constrain it to sum to 0. The mixing parameter $\rho$ lies between 0 and 1 so a Beta$(0.5, 0.5)$ prior is a natural choice.

## 5    Results and Discussion

We can use the structured priors IRT-MRP model to estimate the ideology of California counties: $\theta_c^{MRP}$. As the results from this model demonstrate, California's consistent statewide Democratic support belies significant ideological diversity at the sub-state level. The data I use come from 7,998 respondents in the combined 2019, 2020, and 2021 CES samples. Thirteen questions on abortion (2), the environment (3), gun control (3), healthcare (2), and immigration (3) are used in the IRT portion of the model. Fitting the full structured priors IRT-MRP model in Stan resulted in only 1% divergent transitions, and all parameters with $Rhat < 1.05$, suggesting that there were no major issues during computation.

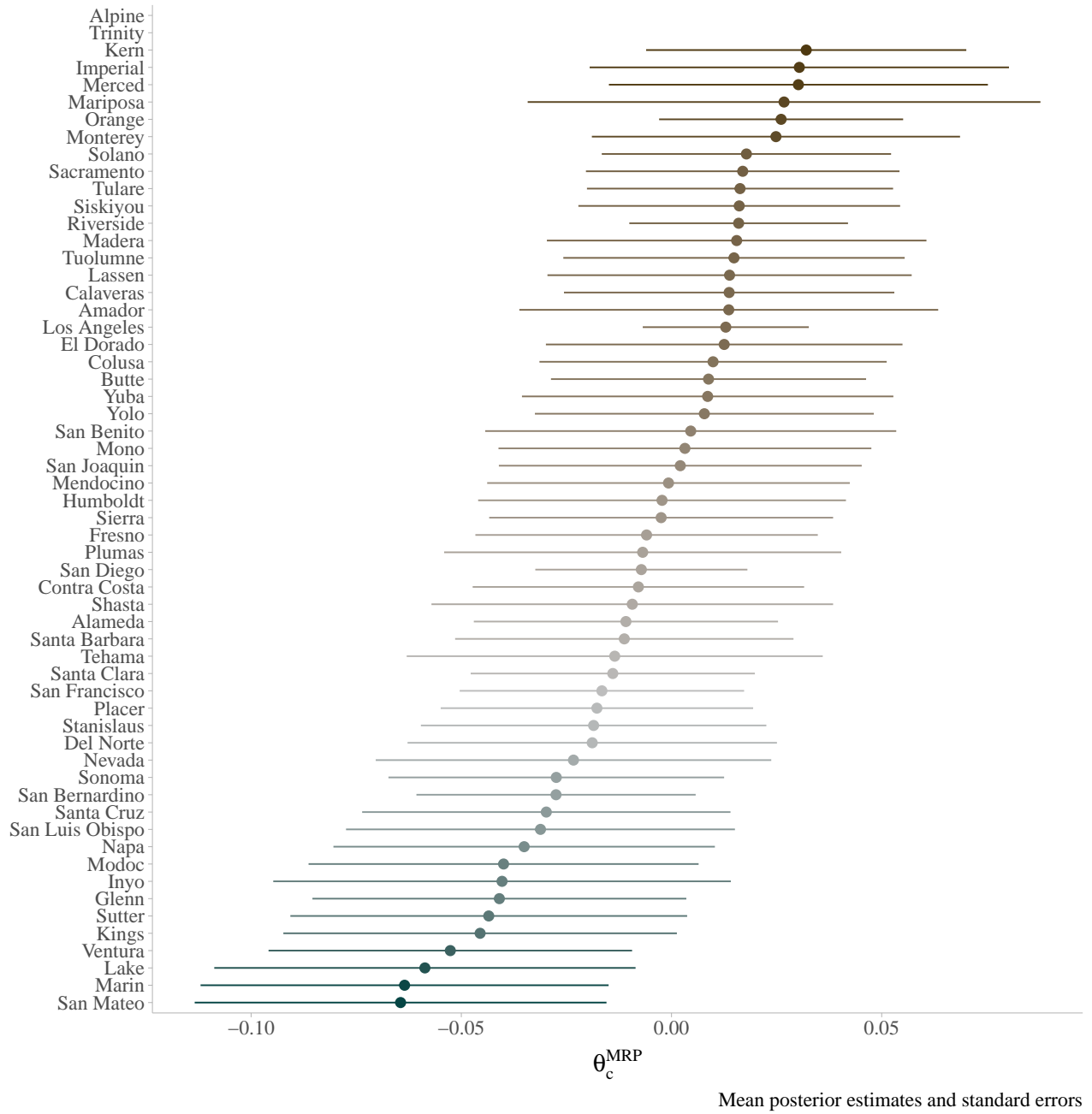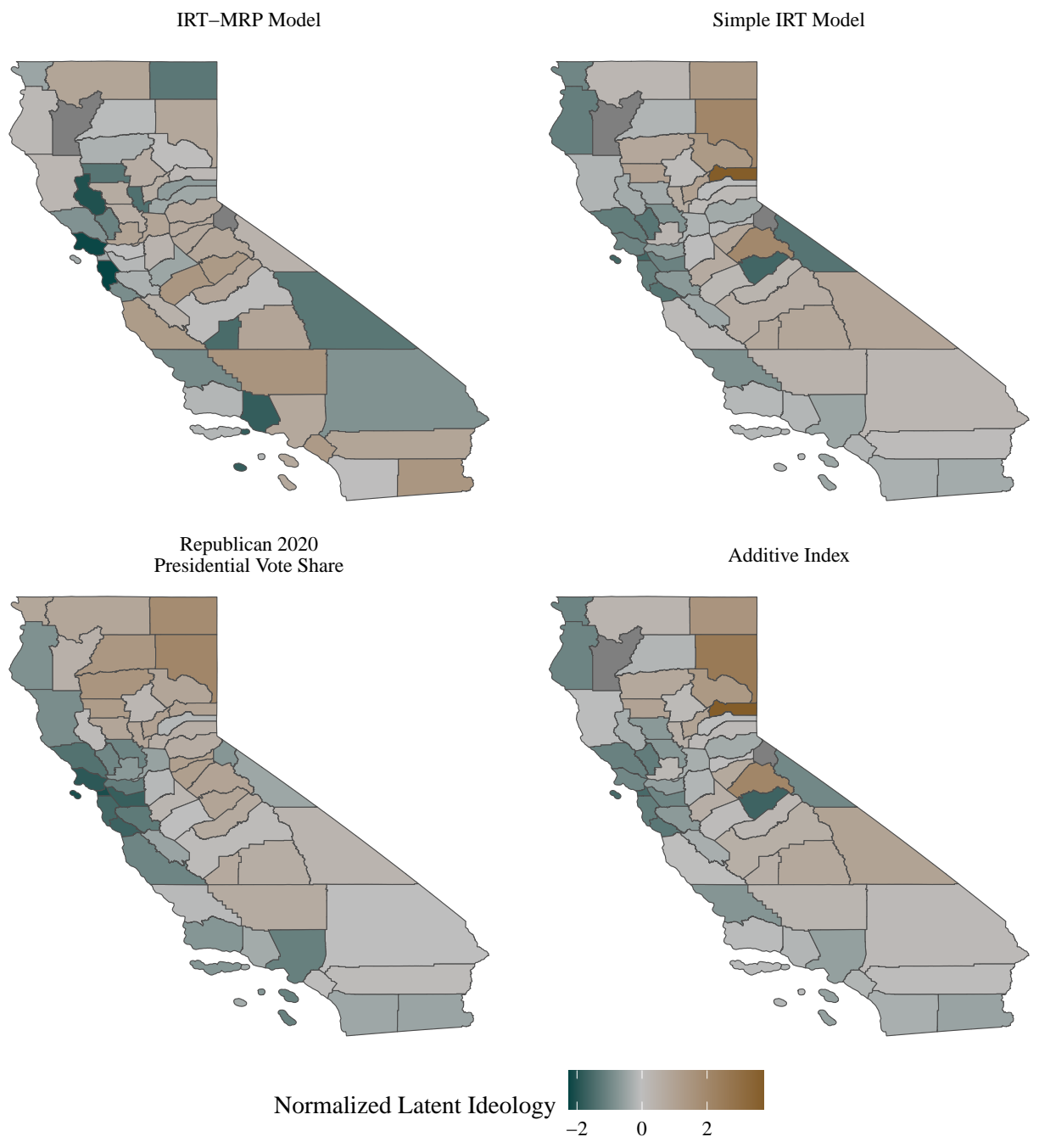Figure 2: California County Ideology Estimates

Mean posterior estimates and standard errors

Figure 2 shows the mean and standard deviation of the posterior estimates for $\theta_c^{MRP}$. Tracking general expectations, the most liberal counties are in the Bay Area (San Mateo and Marin), whereas the most conservative are in the Central Valley (Kern and Merced). Orange county, a Republican stronghold through most of the past half-century, is also among the most conservative. Perhaps contrary to expectations, the counties of San Francisco and Los Angeles are more ideologically centrist than their electoral history and reputation might suggest.

There is a lot of uncertainty in the model results from Figure 2. The extent of the overlap among the posterior standard deviations makes it difficult to make strong claims about the ideological ordering of California counties. But model uncertainty would likely be even worse without the MRP component and structured priors. Small counties such as Sierra (population 3,200) have had their ideology estimates partially pooled towards both the global average and their neighboring counties. This ultimately results in more reliable estimates.

There is no clear way to definitively validate these results. Due to its latent nature, no measure of ideology has a "ground truth" against which to compare. Simulation studies can be conducted in order to inform our understanding about when a measurement model performs better or worse (Gao et al. 2021), but there is no guarantee that simulations reflect important aspects of reality. Instead of running a simulation study, we can illustrate some of the advantages of the structured prior IRT-MRP model by comparing it to other plausible measures of ideology.

Figure 3 shows four measures of ideology across California's counties. The top left has the main structured prior IRT-MRP model estimates from Figure 2. Marin and San Mateo counties, the two most liberal, stand out in dark green on the left side of the state. In the top right of Figure 3

15

Figure 3: California County Ideology Map

IRT–MRP Model

Simple IRT Model

Republican 2020
Presidential Vote Share

Additive Index

Normalized Latent Ideology

−2    0    2

are estimates from a simple IRT model (from Equation 1) using the CES data and aggregating individual ideology, $\theta_i$ by each county. These estimates differ from those on the top right by having none of the structured priors or MRP components in the model. Below the simple IRT model, in the bottom right, are the results from a simple additive index of ideology aggregated by county. Here a CES respondent's ideology is the sum of all conservative answers they gave on the thirteen policy questions. This "model" is equivalent to the simple IRT model from Equation 1, but with the question parameters $\gamma_j$ and $\xi_j$ removed. The implicit assumption behind additive index measures is that each policy question has equal relevance to the underlying latent variable being measured.

Lastly, in the bottom right of Figure 3 are the results of the Republican party vote-share in the 2020 presidential election by county. While acknowledging that partisanship and ideology are not the same thing (Green, Palmquist, and Schickler 2002), presidential election vote share can serve as a valuable gut-check comparison. All four of the variables in Figure 3 have different natural scales, so they have been normalized with mean $= 0$ and standard deviation $= 1$ for comparison purposes. The polarity of each variable has been set such that lower values correspond to ideologically liberal and higher values correspond to ideologically conservative.

The results from both the simple IRT model and additive index in Figure 3 are very similar. This suggests that each policy question in the CES contributes a similar amount of information towards a respondent's underlying ideology. Where both these models fall short, however, is in overfitting to the CES sample. Take Sierra county for example. This extreme conservative outlier (see the narrow horizontal county in the north-east) only had one respondent in the CES. That person took the conservative position on all but one of the policy questions, thereby receiving a normalized ideology

17

score of 3.74 in the simple IRT model and 3.73 from the additive index. The 2020 Republican presidential vote share was relatively high (for California) in Sierra county at 0.59, but this hardly suggests that the county should be almost four standard deviations more ideologically conservative than the average California county.

The structured prior IRT-MRP model developed in this project ameliorates overfitting problems like this in two ways. First, the multilevel/hierarchical aspect shrinks extreme estimates towards the global regression average (and towards a county's neighbors due to the spatial structured prior). We are right to be skeptical of extreme outliers arising out of an idiosyncratic sample—and Bayesian hierarchical modeling provides a principled method for handling such data. The second benefit from using the more complex IRT-MRP model comes from the poststratification step. The ideology of this single respondent from Sierra county should be weighted based on how demographically representative they are compared to the rest of the county. If white men between the ages of 71 and 75 with a two-year college degree earning between $30,000 and $100,000 per year are relatively common in Sierra county, this respondent's ideology estimate will count for more when assigning a final ideology estimate to the whole county, and vice versa. The structured prior IRT-MRP model gives Sierra county a moderately-conservative normalized ideology score of 0.18.

It is important to note that the CES includes weights for each respondent in the survey. These weights can be used to make survey estimates align more closely with nation-wide estimates. But they are less useful for producing county-level estimates since CES weights are constructed with poststratification at the state-level. Also, these types of survey weights are not helpful if we are interested in quantifying uncertainty for county-level estimates. Despite its singular respondent,

18

the uncertainty for Sierra county's ideology is on par with most other county estimates in Figure 2. This is due to the partial pooling of information in the hierarchical component of the IRT-MRP model. Uncertainty for a counties with only one, or a few, respondents would be unquantifiable or explode without this key model component.

One more discrepancy to note between the structured prior IRT-MRP model and the simpler IRT/additive index models is in the central foothills county of Mariposa.[5] This county voted 58% for Donald Trump in 2020, yet the survey-only models score it as the most liberal in California (-1.61 from the simple IRT model, and -1.67 from the additive index). Although there are eight total respondents this time, as opposed to only one, the same overfitting issue which impacted Sierra county's estimate is taking place in Mariposa. Despite the CES's high quality sampling procedures, sampling error is inevitable when the target population is so small (Mariposa county has the sixth lowest population in California with 16,919 residents). The structured prior IRT-MRP estimate for Mariposa is 1.32, which is more reasonable given the fact that the county has voted for the Republican candidate in every presidential election since 1992.[6]

---

[5]Home to Yosemite Valley.

[6]MIT Election Data and Science Lab, 2018, "County Presidential Election Returns 2000-2020", https://doi.org/10.7910/DVN/VOQCHQ, Harvard Dataverse, V12.

# 6 Conclusion

# References

Besag, Julian, Jeremy York, and Annie Molli . 1991. "Bayesian Image Restoration, with Two Applications in Spatial Statistics." *Annals of the Institute of Statistical Mathematics* 43 (1): 1–20. https://doi.org/10.1007/BF00116466.

Betancourt, Michael. 2018. "A Conceptual Introduction to Hamiltonian Monte Carlo." *arXiv*, 60.

Bisbee, James. 2019. "BARP: Improving Mister P Using Bayesian Additive Regression Trees." *American Political Science Review* 113 (4): 1060–65. https://doi.org/10.1017/S0003055419000480.

Bürkner, Paul-Christian. 2020. "Bayesian Item Response Modeling in R with Brms and Stan." *arXiv:1905.09501 [Stat]*, February. http://arxiv.org/abs/1905.09501.

———. 2023. *Brms: Bayesian Regression Models Using Stan.* https://github.com/paul-buerkner/brms.

Caughey, Devin, and Christopher Warshaw. 2015. "Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model." *Political Analysis* 23 (2): 197–211. https://doi.org/10.1093/pan/mpu021.

Donegan, Connor. 2023. *Geostan: Bayesian Spatial Analysis.* https://connordonegan.github.io/geostan/.

Downes, Marnie, Lyle C Gurrin, Dallas R English, Jane Pirkis, Dianne Currier, Matthew J Spittal,

and John B Carlin. 2018. "Multilevel Regression and Poststratification: A Modeling Approach to Estimating Population Quantities From Highly Selected Survey Samples." *American Journal of Epidemiology* 187 (8): 1780–90. https://doi.org/10.1093/aje/kwy070.

Gao, Yuxiang, Lauren Kennedy, Daniel Simpson, and Andrew Gelman. 2021. "Improving Multilevel Regression and Poststratification with Structured Priors." *Bayesian Analysis* 16 (3). https://doi.org/10.1214/20-BA1223.

Green, Donald P, Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters.* New Haven, CT: Yale University Press.

Greg Freedman Ellis, Derek Burk, and Finn Roberts. 2023. *Ipumsr: An r Interface for Downloading, Reading, and Handling IPUMS Data.* https://tech.popdata.org/ipumsr/.

Kay, Matthew. 2023. *Ggdist: Visualizations of Distributions and Uncertainty.* https://mjskay.github.io/ggdist/.

Landau, William Michael. 2024a. *Stantargets: Targets for Stan Workflows.* https://docs.ropensci.org/stantargets/.

———. 2024b. *Targets: Dynamic Function-Oriented Make-Like Declarative Pipelines.* https://docs.ropensci.org/targets/.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan.* 2nd ed. CRC Texts in Statistical Science. Boca Raton: Taylor; Francis, CRC Press.

Morris, Mitzi, Katherine Wheeler-Martin, Dan Simpson, Stephen J. Mooney, Andrew Gelman, and Charles DiMaggio. 2019. "Bayesian Hierarchical Spatial Models: Implementing the Besag York Mollié Model in Stan." *Spatial and Spatio-Temporal Epidemiology* 31 (November): 100301.

https://doi.org/10.1016/j.sste.2019.100301.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files.* https://here.r-lib.org/.

Treier, Shawn, and D. Sunshine Hillygus. 2009. "The Nature of Political Ideology in the Contemporary Electorate." *Public Opinion Quarterly* 73 (4): 679–703. https://doi.org/10.1093/poq/nfp067.

Wickham, Hadley. 2023. *Stringr: Simple, Consistent Wrappers for Common String Operations.* https://stringr.tidyverse.org.

Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, Dewey Dunnington, and Teun van den Brand. 2024. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics.* https://ggplot2.tidyverse. org.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation.* https://dplyr.tidyverse.org.

Wickham, Hadley, and Lionel Henry. 2023. *Purrr: Functional Programming Tools.* https://purrr. tidyverse.org/.

Wickham, Hadley, Davis Vaughan, and Maximilian Girlich. 2024. *Tidyr: Tidy Messy Data.* https: //tidyr.tidyverse.org.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with Kable and Pipe Syntax.* http:// haozhu233.github.io/kableExtra/.