Using Multilevel Regression and Poststratification to Estimate Dynamic Public Opinion

Andrew Gelman *
gelman@stat.columbia.edu

Jeffrey Lax
rjl2124@columbia.edu

Justin Phillips
jhp2121@columbia.edu

Jonah Gabry
jgabry@gmail.com

Robert Trangucci
robert.trangucci@gmail.com

August 28, 2018

Abstract

Multilevel Regression and Poststratification (MRP) has emerged as a widely-used technique for estimating subnational preferences from national polls. This technique, however, has a key limitation—existing MRP technology is best utilized for creating static as opposed to dynamic measures of opinion. We develop an approach for implementing a "dynamic MRP", doing so in the context of changing public support for same-sex marriage. Using a large dataset of survey respondents, we estimate (in a single model) an annual measure of support for same-sex marriage for each state from 1993 through 2004. To evaluate our estimates we examine their face validity and compare them to estimates produced using the standard MRP approach as well as to the estimates produced by actual state-level polls. We also consider the conditions under which dynamic MRP seems to produce more accurate estimates.

*This project is supported by a grant from the National Science Foundation

1 Introduction

Since its emergence in the 1930s, scientific polling has grown into a large industry, fueled by an insatiable demand for information about the American public. Political scientists now have extensive polling on important policy issues as well as nearly-continuous polling on presidential approval and, during campaign season, we are inundated with trial-heat polls. The Roper Center for Public Opinion Research reports that its archive of opinion surveys has grown to over 18,000 datasets and continues to grow by hundreds more each year. Put simply, researchers are now awash in survey data.

While surveys provide researchers with invaluable data about the public's views, preferences, and beliefs, these data are not without limitations. Key among them, is that most polls are conducted by national survey organizations and are only designed to measure opinion at the national level (e.g., 34% of Americans have a favorable opinion of the Affordable Care Act, 53% support same-sex marriage, and so forth). Many important policies, however, are decided by state governments. For that matter, opinions on contentious national issues are themselves typically translated into national policy based on their potential impact on individual congressional or senatorial races or on the electoral college—i.e., "all politics is local." For these reasons summaries of national opinion provide only limited guidance to most lawmakers. They are also of limited value for those social scientists who want to study the ways in which public opinion varies across the geography of American federalism or those who wish to investigate issues of representation and policy responsiveness.

Relying on subnational surveys is not often a reasonable solution. Despite a rich tradition of state-level polling, finding comparable polls across states is nearly impossible. Similar questions are rarely asked in surveys across all or even most states and when they are, differences in timing, question wording, survey techniques, and response categories make comparisons difficult. An alternative and more practical approach is to use national survey data to simulate subnational opinion. Recently, scholars have revived—or more accurately, reinvented—simulation techniques. The first to truly catch on as a widespread tool, is multilevel regression and poststratification (MRP). MRP was developed by Gelman and Little (1997) and extended by Park, Gelman, and Bafumi (2004, 2006), Lax and Phillips (2009a), Pacheco (2011), Warshaw and Rodden (2012), and Kastellec, et al. (2015). It uses individual survey responses from national surveys coupled with advances in Bayesian statistics and multilevel modeling to generate opinion estimates by demographic-geographic subgroups, or "types". The opinion estimates for each demographic-geographic respondent type are then weighted (poststratified) by the percentages of each type in the actual population of each subnational unit of interest. Several research teams have already evaluated and validated MRP (Park, Gelman and Bafumi 2006, Lax and Phillips 2009a, 2013, Warshaw and Rodden 2012, Buttice and Highton 2013). This work suggests that MRP can produce accurate estimates using fairly simple demographic-geographic models of survey response and small amounts of survey data.

However, challenges remain. In particular, the method is not currently well suited for exploring temporal changes in opinion. Existing MRP technology is best utilized for creating static measures of preferences—that is, using national surveys conducted during time *t* (with *t* representing a year or set of years) to create a single opinion estimate for each geographic unit. Though such static measures have already proven invaluable, they do not go far enough. Policymakers, the media, and scholars want to understand how and why public opinion is changing. Researchers, across a range of disciplines, need dynamic measures of the public's preferences in order to better establish causal links between public opinion and outcomes. Political scientists, for example, may want to see whether public policy changes in response to shifting public preferences, while psychologists may want to investigate whether the mental health of lesbian, gay, and bisexual populations improve in places where public where tolerance for homosexuality is rising.

While a few scholars have employed MRP to study opinion change, they have typically created over-time measures of opinion by running separate MRPs on polls from different years or on year subsets and then stringing them together into a time series (see Gelman, Lax and Phillips 2010; Pachecho 2011; Kastellec 2016). Doing so, however, fails to make use of all the available data and employs arbitrary assumptions as to how much change occurs over time. In this paper, we advance MRP such that it can be used to create dynamic measures of public opinion. Advancing MRP in this fashion is not straightforward. To create dynamic measures, it is likely that MRP will need to allow demographic and geographic effects to vary by year; that is, it must allow for models of much higher complexity wherein key predictors are interacted with time. In essence, we aim to transform MRP so that it partially pools survey data not just across *space* both also over *time*. Advancing MRP in this fashion will enable researchers to utilize decades of accumulated survey data to create time-varying measures of public opinion across a variety of subnational units and to potentially generate more accurate estimates of public opinion at any given point in time—by incorporating time trends, researchers will be adding information that should result in more accurate estimates than are currently generated using static MRP.

We explore the potential of "dynamic MRP" in the context of changing public support for same-sex marriage. We have assembled a large dataset consisting of all of the respondents to publicly available opinion surveys conducted from 1993 through 2014 that directly ask about support for same-sex marriage. Our dataset consists of 81,127 respondents from 68 separate polls. Using these data we generate, in a single model, an estimate of support for same-sex marriage for each state in each year. This means that we produce 1,100 separate estimates of state opinion. The models that we develop are estimated in `Rstanarm`, a fully Bayesian approach that allows for the use more complicated models than `lme4`, the package that has traditionaly been used to estimate MRP models. In addition to allowing for added complexity, Rstanarm directly estimates uncertainty as a result of the sampling process, which is important not only for obvious reasons of understanding the limits of our estimates but incorporation of such uncertainty into further analysis (see Kastellec, et al. 2015 as example).

Here we consider a variety of alternative specifications of our dynamic model. To evaluate our estimates we examine their face validity, compare them to estimates produced using the standard MRP approach, and compare them to the results of actual state-level polls. In doing so, we investigate the conditions under which dynamic MRP produces more accurate estimates. We vary model complexity, the extent and details of temporal dynamics (do all parameters or only some vary over time?), the amount of data used to estimate models so as to resemble common real-world data limits, etc.

Ultimately, we succeed in implementing an approach to dynamic MRP (or "MRT"). Our approach produces reasonable estimates of state level support for same-sex marriage. These estimates have face validity and are highly correlated with the results of actual state level polls. We find that with lots of data, even simply MRT model produce accurate estimates of public opinion. As the amount of survey declines however, it becomes increasingly important to include time smoothers in the models. We also find that MRT produces results that are highly correlated with those generated by MRP models, but that MRT does better when it comes to estimating over time opinion in smaller states.

2 Overview of Multilevel Repression and Poststratification

The simulation of subnational public opinion traces back nearly fifty years (Pool, Abelson, and Popkin 1965). Opinion estimates are created for various geographic units according to the demographic distribution of the population within each. The primary flaw in the older versions of this technique is that respondents were generally modeled as differing in their demographic but not their geographic characteristics, so the prediction for any demographic type was unvaried across the subnational unit of interest (Erikson, Wright, and McIver 1993). In other words, using this method, the opinion of white citizens in New England in 1965 regarding civil rights issues could be thought the same as those held by white citizens in the South. Simulated opinion estimates ignoring geographic variation have been shown inferior in less controversial contexts as well. This tool failed to be widely adopted.

Recently, Park, Gelman, and Bafumi (2006) resuscitated the simulation approach, drawing on pioneering work in poststratification by Rod Little (1991,1993) and small-area estimation by Bob Fay (1979). Unlike early simulation approaches, this reincarnation, which is referred to as multilevel regression and poststratification (MRP), takes geography into account. It is also much more sophisticated than earlier simulation techniques in terms of the way it models individual survey responses and demographic variation.

MRP proceeds in two stages. In the first, a multilevel model of individual survey response is estimated. Instead of relying solely on demographic differences like older incarnations of the method, the geographic location of the respondents is used to estimate geographic effects, which themselves can be modeled using additional predictors such as aggregate demographics of the geographic area. Those residents from a particular area yield information as to how much predictions within that area vary from others after controlling for demographics. All individuals in the survey, no matter their location, yield information about demographic patterns which can be applied to all geographic estimates. These demographic-geographic predictors can interact, as well.

To be specific, MRP uses Bayesian statistics and multilevel modeling (Gelman and Little 1997; Park, Gelman, and Bafumi 2006) to improve upon the estimation of the effects of individual and geographic predictors. For data with hierarchical structure (e.g., individuals within states or congressional districts), multilevel modeling is generally an improvement over classical regression. Rather than using "fixed" ("unmodeled") effects, MRP uses "random" ("modeled") effects, for some predictors. These modeled effects (e.g., state or district effects) are related to each other by their grouping structure and thus are partially pooled towards the group mean, with greater pooling when group-level variance is small and for less-populated groups (this is equivalent to assuming errors are correlated within a grouping structure (Gelman and Hill 2007, 244-65)). The degree of pooling within the grouping emerges from the data endogenously. They can be modeled not only in terms of this "shrinkage" (the assumption that they are drawn from some common distribution) but also by including group-level predictors.

The second step is poststratification: the estimates for each demographic-geographic respondent type are weighted (poststratified) by the percentages of each type in actual populations of the relevant geography, so that we can estimate the percentage of respondents within each who have a particular issue position or preference. Poststratification is done using state or congressional district population frequencies obtained from either the Public Use Micro Data Samples supplied by the Census Bureau (and available going back to early 20th century) or similar data. Compared to previous simulation methods and classical methods, multilevel modeling now makes possible the use of many more respondent types. This too greatly improves accuracy.

Importantly, poststratification corrects for differences between survey samples and the actual population. National surveys, while representative at the national level, are often flawed in terms of representativeness or geographic coverage at the state or congressional district level, due to clustering and other survey techniques utilized by polling firms (Norrander 2007, 154). Indeed, with the increasing popularity of internet survey techniques and cell phones, it is becoming increasingly difficult to even find data with a random sample of the national population, let alone random samples of subsets of interest such as demographic slices or residents of particular states. MRP addresses these concerns. Moreover, it can be difficult to combine different surveys using traditional methods, if they use different question wording, sampling techniques, demographic questions, etc. MRP, on the other hand, can bridge surveys by modeling and accounting for such differences.

The value of using MRP to estimate subnational public opinion has been confirmed in several articles that have appeared in top peer-reviewed journals. Importantly, these articles have been authored by separate research teams. Park, Gelman, and Bafumi (2006) compare MRP estimates of opinion to those of alternative techniques that also model individual survey responses. They show that MRP substantially outperforms the older style simulation approaches as well as simulation approaches that do not partially pool information across respondents. Lax and Phillips (2009a) show that MRP notably outperforms its main competitor (disaggregation), yielding smaller errors, higher correlations, and more reliable estimates. They also establish the face and external validity of MRP estimates by comparing them to actual state polls and election results, demonstrating that a single national poll (approximately 1,400) and a very simple demographic-geographic model can, in some contexts, suffice for MRP to produce highly accurate state-level opinion estimates. In a parallel analysis, Warshaw and Rodden (2012) demonstrate that MRP can produce accurate estimates of opinion by congressional districts (using sample sizes of just 2,500 survey respondents) and state senate districts (with a sample of 5,000). Most recently, scholars have worked to provide sets of guidelines and cautions to MRP users (Buttice and Highton 2013; Lax and Phillips 2013; Toshkov 2015).

Collectively, these efforts demonstrate that MRP can produce accurate estimates of public opinion across a variety of subnational units using fairly simple models of survey preferences and modest sample sizes. Relatively small national samples can be used to produce accurate measures of subnational opinion because the multilevel models used in MRP borrow strength by partially pooling respondent types across space (i.e., the subnational units of interest) to an extent determined (endogenously) by the data. As a result, MRP is "emerging as a widely used gold standard for estimating preferences from national surveys" (Selb and Munzert 2011, 456). Indeed, research employing MRP has already appeared in the top political science journals, and MRP has been employed to study to myriad substantive questions, including the responsiveness of state governments (Lax and Phillips 2009b, 2012), state supreme court abortion decisions (Caldarone, Canes-Wrone, and Clark 2009), roll call voting in Congress (Kastellec, Lax, and Phillips 2010, Kastellec et al. 2015; Krimmel, Lax, ad Phillips 2016), and the diffusion of public policy (Pacheco 2012). This substantive work has also further developed the method, devising techniques for postratifiying by non-census variables and for estimating uncertainty around MRP estimates and then incorporating this uncertainty into subsequent empirical analyses (Kastellec et al. 2015).

3 Developing a Dynamic MRP

While the methodological and substantive potential of MRP has clearly been established, more work needs to be done before the full potential of national surveys can be unlocked. In particular, existing MRP technology seems to be best utilized for creating static measures of preferences—that is, using national surveys conducted during time t (with t representing a year or set of years) to create a single opinion estimate for each geographic unit of interest. We do know that there are better and worse ways to do MRP (Lax and Phillips 2013). For example, for the most typical use of creating state estimates from national data, it is important to include a state-level variable such as presidential vote or ideology to help smooth across states and enable a proper degree of pooling across states. But how do we explore time trends and combine data across years instead of geographic divisions?

One simple way is just to do MRP for each year, standing alone, to create a state estimate for each year possible. This doesn't make use of the "secret sauce" of MRP for the state problem—whereby partial pooling enables large efficiency gains. So we want to allow for some type of pooling of information over time...but combining data across years raises many concerns. How do we help MRP capture trends across time? How do we enable the "right" degree of pooling across time, without over smoothing and hiding opinion change or opinion differences across states? How do we avoid chasing noise in the guise of opinion swings? Do we just need an over-time state smoother such as presidential vote? Should we model time trends more explicitly? How do we know if we are discovered trends or creating them through modeling?

Some researchers have developed various "patches" for these problems, but these do not necessarily deal with said problems and, perhaps worse, we cannot tell if they do so. For example, consider the creation of over-time measures of opinion by running separate MRPs on polls from different years or on year subsets and then stringing them together into a time series (see Gelman, Lax and Phillips 2010; Pachecho 2011). Doing so, however, fails to make use of all the available data and employs arbitrary assumptions as to how much change occurs over time. Pooling is complete within subset and barred completely across subsets. At least subsetting does some, if oddly constructed, degree of pooling. The most careful of this type of work so far is Pachecho (2011), which created yearly estimates of state-level voter ideology (from 1977 through 2007) by pooling surveys over three or five year time-periods. So, to get an estimate of opinion in year t she estimates an MRP using responses from polls conducted in years t *-1*, t, and t + *1*. Doing so assumes away potential short-term changes in opinion. It also means (according to our back-of-the-envelope calculations) that for each year's opinion estimate, 90% of the survey data—just over 292,000 responses—are ignored so that we learn nothing about pooling parameters, etc., from them. This suggests there are gains to be made. We believe that MRP can be improved so that researchers do not need to make these unpalatable decisions.

In particular, we suspect that by incorporating much more data (than static MRP) we should be able to improve opinion estimation for any given year. But we need to know how to model time and opinion change in a flexible-enough way that still yields the desired efficiency gains, without obscuring time trends or creating spurious ones. We need to compare different methods of so doing. For example, Pacheco (2011) presented evidence that three-year windows beat five-year windows for partisanship and ideology data. We seek general evidence across a wider array of possibilities for doing MRP over Time...or MRT.

The models that we employ here are estimated using an R package called RStanarm. RStanarm is a full Bayesian counterpart to lme4 (a more approximate but simpler method, the usual way to implement multilevel model estimation for the purposes of MRP).[1] One advantage of this Stan approach is that uncertainty around estimates is directly output as part of the estimation process rather than requiring a type of post-estimation bootstrapping from model parameters (as done in the previous MRP paper incorporating uncertainty, see Kastellec et al 2015). More importantly perhaps, it is a fully Bayesian approach incorporating uncertainty (including uncertainty in the hierarchical variance parameters of the models). While this makes for negligible differences for simple multi-level models and normal MRP, the increased complexity of MRT warrants allowing for as many sources of uncertainty as possible. At this point in time, Stan estimation is feasible for these purposes.[2]

[1]It encapsulates several popular model classes, such as linear mixed effect models (estimated using stan-lmer) and generalized linear mixed models (estimated using stan-glmer). These functions accept lme4 model formula syntax but ultimately use Stan to do Bayesian inference on the models. Specifically, Stan uses a flavor of Hamiltonian Monte Carlo, a Markov Chain Monte Carlo algorithm, called the No-U-Turn-Sampler, which adaptively tunes the hyperparameters required by HMC. HMC is powerful because it takes the geometry of the posterior distribution into account when generating proposals for the next step in a Markov Chain. In doing so, the sampler can take sequential steps that are far apart, yielding draws from the posterior that are near independent. This differs from the maximum marginal likelihood (MML) algorithm that undergirds lme4 in several ways. When viewed through the lens of Bayesian inference, lme4 generates maximum a posteriori (MAP) estimates for random effects and hierarchical variance parameters with uniform priors on the half interval (0, Inf). Gelman 2006 writes that this leads the model to overstate the expectation of true variance, which can lead to less-than-optimal shrinkage of random intercepts towards the common mean. Additionally, MML cannot be applied to generalized mixed models in closed form and thus is estimating a normal approximation to the posterior (a Laplace approximation). To summarize, MML's point estimates for hierarchical variance parameters are positively biased, but MML understates posterior uncertainty because it uses a normal approximation to the posterior. Stan doesn't suffer from any of these shortcomings. We can put proper priors on hierarchical variance parameters and then generate random draws from the true posterior. We can diagnose when the sampler encounters problems, but MML cannot offer any theoretical tools to assess the statistical validity of its estimates aside from the convergence of its log-likelihood function.

[2]Stan is flexible, so it allows you to add terms to your model and estimate the thing in the same framework (i.e. run Stan, extract samples, take means, variances and quantiles, etc), rather than requiring you to change packages and figure out how to shoehorn your model into another estimation procedure.

3.1 Models Estimated

In order to demonstrate the robustness of MRT and to make recommendations about the sorts of specifications that researchers should utilize, we estimate 18 models of varying complexity. Table 1 notes all of the variables that we employ, though the specific variables utilized differs across models.

Table 1: **Variables Used in Dynamic MRP Models**

Variable	**Description**
`favor`	Number of respondents favoring gay marriage in cell i
`oppose`	Number of respondents opposing gay marriage in cell i
`year_std`	Year running from 1993 to 2015, standardized
`year_sq_std`	Square of year, running from 1993 to 2015, standardized
`state`	Categorical state variable, 50 levels, unordered
`age`	Categorical age variable, 4 levels
`edu`	Categorical education variable, 4 levels
`sex_race`	Categorical sex and race variable, 8 levels
`year`	Categorical year variable, 19 levels
`Ideology`	Measure of state-level ideology from Erikson, Wright and McIver (1993)

We begin with a baseline/simple model of demographic smoothing which we label as $\{D\}$. This model includes all of our demographic predictors plus state and year random effects. This statistical model can be written as:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

All subsequent models add complexity to the baseline specification in some way. Specifically, the additional models add one or more of the following:

- S: state-level smoother, helping to model state random effect intercept shifts, in this case state-level Ideology
- S_t: state level smoother, allowed the coeffiecient on ideology to vary by year

- T: a continuous time smoothing variable to help model year intercept shifts, composed of year and year squared
- T_s: time smoothing, the coefficients on year and year squared are allowed to vary by state, so that trends can vary by state
- D_t: demographic coefficients (such as coefficients on age category) vary by time

Every model will have demographic smoothing, and given that these are MRT models each will have some form of time smoothing. That leaves the set of all models as:

$$D \bigcup s \mid s \in \mathcal{P}(S), , , S : S, , S, , S_t, , T, , T, , T_s, , D_t$$

For a more complete explication of all 18 models see the Appendix. In addition to these, we also estimate MRP models that reflect the more "patchy" approaches that scholars have used to generate time varying estimates of preferences (the basic MRP model we estimate here includes the age, education, and sex_race demographic predictors, a state random effect, and uses stat ideology as a state level smoother). The first of these is to run a separate MRP model for each year using only survey data from that year (this labeled as MRP-t), the second is to estimate an MRP using a rolling window 3-year window of data (this is labeled as MRP-3t-S), the third is to again estimate an MRP model on a 3 year window of data but this time without included the a state level smoother (this is labeled as MRP-3t and replicates the approach of Pacheco 2011).

The saturated MRT model, with all variable sub-types included is akin to running a separate MRP model for each year but with the expected benefits of borrowing strength across years, pooling coefficients over time (e.g., age coefficients get pooled over time, etc.). A less than fully saturate MRT approach constrains some things to not vary as they would be allowed to if we ran fully separate models by year. This might be useful if we are asking too much of the data, risking adding more noise than signal from the response model complexity. Or it might be useful for computational reasons, if we wish to keep the modeling simpler for time concerns or ease of presentation. Then the question is when a simpler model suffices, not only whether the more complex model is strictly better or worse. MRT models without an explicit time element do partially pool over time, but do not take advantage of the actual time series structure (that is, all years are pooled together without any assumption that years closer in time are related to each other).

When we do MRP and have data for most or nearly all years, and then compare to measures of truth in those years, we are not pooling over time. But, when we soon turn to more sparse data situations, and we need to make a prediction relative to an actual state poll for a year in which we do not have an actual MRP result, we do linearly interpolate between estimates, which is a form of smoothing over time itself.

3.2 Our Data

We estimate our dynamic models of MRP using a large database of survey respondents. The database includes respondents from all national publicly available polls from 1993 through 2014 that directly asked about support for same-sex marriage.[3] Respondents were excluded if we do not have data about their demographic characteristics or geographic location (not all polls report respondents state of residence). We located polls using iPoll which is housed at the Roper Center for Public Opinion Research. For each respondent, we code race, age, gender, level of education, and state of residence. In total our dataset includes a total of 81,127 useable respondents from 68 unique polls. Our initial efforts maintain all data, to be trimmed with simulated smaller samples later.

We have also collected the results of 347 state-level polls that ask about support for same sex marriage. These polls were conducted during our time period of interest and were found using news archives and interest group websites. While such surveys are conducted unevenly across states, they serve as benchmarks to which we can compare and evaluate our MRT estimates (of course, surveys themselves are imperfect estimates of preferences).

[3]Polls that ask about support for both same-sex marriage and civil unions are excluded as well as those that ask about support for a constitutional amendment banning same-sex marriage.

3.3 Results

We compare the estimates of state-level opinion generated by each of our MRT and MRP models to corresponding state polls at the relevant point in time, computing the root mean squared error between them. At this point, we treat models for which the RMSE is smaller as being more accurate.

Figure 1 presents the RMSE of the median state-year opinion estimate for each model, considering the results for all states and then separately for the largest 10 states (in terms of population size) and the smallest 10 states. The right hand legend shows the number of state polls that are being used to capture "true" state opinion for the state sample in question. Importantly, the MRT models all seem to perform reasonably well with very small differences in RMSE across model specifications or across state size. Even the MRP methods do fine for large or all states, though they do notably worse than a full MRT approach when it comes to producing measures of support for same-sex marriage among the smallest states.

Figure 2 looks at similar patterns but this time we use fewer polls to estimate our MRT and MRP models. We do this to in an attempt to simulate "real world" data limitations—rarely do researchers have such a wealth of survey data as do those focused on same-sex marriage opinion. To simulate more restrictive data environments, we estimate our models using every second poll, every fourth poll, and every eighth poll, after the first. Here we present RMSE resulting from models that only use every fourth poll (this means we have approximately 20,000 survey respondents spread out over our time period of interest). We now see greater separation among MRT approaches. First, the MRT models that do well are those that use a time smoother (T). Models that include only demographic predictors, state random effects, and year random effects do not perform well (even if the coefficients on these are allowed to vary by year). Such models perform particularly poorly among small states.

Second, the MRP models for "every fourth" perform similarly to the MRT models that utilize a time smoother. This is not that surprising since the MRP models have sufficient data to perform well, by basic MRP benchmark standards. The gains from borrowing strength in MRT across years of data is small enought to be ignored perhaps. Once we move to "every eighth" poll then we do see that MRP approaches are clearly worse for small states, as shown in Figure 3. In other words, there are gains to be had by partial pooling over time, which can improve estimation, similar to how MRP's partial pooling over states improves estimation.

Figure 4 and Figure 5 reorder these results to enable a closer look at how having fewer polls affects MRT. These figures compare RMSEs within a modeling approach using every poll, every second poll, every fourth poll, and every eighth poll. Figure 4 shows data for all states and Figure 5 for only the smallest ten states. The dangers of having fewer polls are particularly bad for those approaches that omit any direct time modeling in the form of the T component. In this sense, having much more poll data and modeling time are substitutes when it comes to accurate estimation. This is particular clear for smaller states. The same problem occurs in doing MRP without any pooling of data across years to gather sufficient data (MRP_t as compared to the MRP_3t models).

Since it seems clear that some version of this MRT process can work well enough, and we see hints above of how to best do MRT, we move to direct recommendations for better MRT practices. The first question is how to best set up the MRT model. How should "time" come into the model and what should be time-varying? And, should we do MRP or MRT in the situations we study?

We start by noting that standard advice for MRP to include a state-level substantive variable to help determine the proper degree of pooling in a state might not be necessary in the MRT context, if one has enough data. Indeed, even using one eighth of the polls we have, we have over 10,000 respondents (at that level, even disaggregation can suffice for producing very good estimates, as shown in Lax and Phillips 2009)

Is S needed? We again compare our estimates to actual state polls, looking at the RMSE difference for including ideology in the model. If one has sufficient samples by state, as we do even when using a subset of our data such as every eighth poll, then the state-level variable (adding an S term, in our parlance) contributes little. Here, we find the value of adding an S term to be negligibly small when it comes to comparisons to actual state polls. The RMSE changes on the order of one-tenth of a percentage point, no matter what the other features of the MRT model are. That said, since this is standard in MRP, it would seem appropriate to leave such a term in.

The deeper question is whether to allow the slope with respect to this S variable to vary over time, or instead assume a constant slope over time. This too matters little for the RMSE to state polls, even for our every 8 polls, smallest states sample. Thus, one might as well omit this complication. (Later, more complicated simulations of constructed data will explore this further.)

Next, when doing MRT, do we need to let the effects of the standard MRP demographic variables change over time? If we simply run an MRP model for each year, then we are in effect allowing these demographic coefficients to vary from model to model, so perhaps allowing variation within MRT is important. When doing so in MRT, as compared to separate MRPs, we get the gains of pooling these demographic coefficients over time, borrowing strength. But does allowing this variation improve RMSE to actual state polls? There are six different simpler models to which one could add Dt, allowing demographic intercepts to be time-variant. It could matter what the rest of the MRT model structure is, so that the value of the Dt term depends on which other terms are included. That is not the case. When using all polls, the value (in smaller RMSE) of the Dt term is nearly zero across the board. It remains a tenth of a percentage point or smaller even for the every eighth poll analysis. We see no need to include this complication at this point. This finding is parallel to that in previous MRP work showing that rather simple demographic structures suffice (Lax and Phillips 2009, Lax and Phillips 2013).

Finally, even in the most basic MRT model, there are year random effects (varying intercepts). As noted above, MRT is like running an MRP for each year but allowing partial pooling. But one can bring time in more explicitly, allowing for a time-smoother variable akin to the substantive state-smoother variable. Does adding our continuous time smoother help? (Recall this means including year and squared year.) There are six models that could have T added. The RMSE improvement for all data, all states is around a tenth of a percentage point across the board, but as polls become more sparse, the RMSE improvement increases. For every fourth poll, for example, the reduction to RMSE is around 3.3 percentage points regardless of other model components. For every fourth poll, smallest states only, the benefits are roughly 4.5 percentage points regardless of model components. We conclude that including a continuous smoothing measure of time is important.

Does one benefit further from allowing the time slope to vary by state, adding the Ts term? In the all state comparisons, it often made RMSE worse, suggesting it just induced noise. This was true regardless of other model components or state sample (all or small), suggesting this complication be avoided. Future work could explore how much this is needed when state trends vary more than they do in the gay marriage situation.

Finally, we present estimated support for same sex marriage over time for 12 states alongside the national average. These states (though not a random sample) contains a mix of large, medium, and small population states as well as states from across the ideological spectrum. Figure 6 does so using what we would probably think of the simplest normal MRT model (labeled as TDS above). This model employs the demographic predictors D, the time smoother T, and the state smoother S. Figure 7 uses our most saturated model (labeled as $TDSStTsDt$ above). This model utilizes the D and S smoothers but allows their effects to vary by year. It also utilizes the T smoother, but allows the effect to vary by state.

Both figures provide face validity for our estimates. The results (unsurprisingly) are consistent with what one would anticipate. Support for same-sex marriage rises over time, and does so quite dramatically after 2005. The states that are most supporting of same-sex marriage in our sample are those in the northeast (Massachusetts, New Hampshire, and New York), plus California. The states with the lowest levels of support are located in the south—Tennessee and Alabama. As one can see, the estimates across states clearly move together, though the extent to which they do depends somewhat upon the model. Obviously, the more saturated model builds in this type of flexibility. Indeed, while both models produce very similar RMSEs (even with smallish sample sizes and in small states), researchers may prefer the more saturated model because of this added flexibility. Or, they may fear that this shows simply noise. Again, further work is needed.

4 Further Questions/Next Steps

Whereas our previous attempts at handling MRT (not shown) ran into problems of clear over-pooling over time, at this point we have demonstrated that MRT can successfully be implemented. We were able to generate (in a single model) annual opinion estimates for all 50 states over a 22-year period. These estimates are comparable to estimates obtained through traditional MRP or MRP using a three-year sliding window of survey data (similar to Pacheco 2011). The estimates we produced using MRT are also strongly correlated to actual state-level polls. This is especially true when variables are included to capture time trends. At this point, we feel that we can say that MRT is advantageous over the alternatives approaches we consider when working in environments with smaller amount of data—estimating opinion for small states or estimating models with fewer survey respondents.

References

Buttice, Matthew K and Benjamin Highton. 2013. "How Does Multilevel Regression and Post-stratification Perform with Conventional National Surveys?" *Political Analysis* 21: 449-67.

Caughey, Devin and Christopher Warshaw. 2015. "Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model," *Political Analysis* 23(2): 191211.

Caldarone, Richard P., Brandice Canes-Wrone, and Tom S. Clark. 2009. "Partisan Labels and Democratic Accountability: An Analysis of State Supreme Court Abortion Decisions." *The Journal of Politics* 71(2):560-73.

Chipman, H. A., George, E. I., and McCulloch, R. E. 2008. "BART: Bayesian additive regression trees. Technical report, Department of Mathematics and Statistics," Acadia University, Nova Scotia, Canada.

Erikson, Robert S., Gerald C. Wright, and John P. McIver. 1993. *Statehouse Democracy: Public Opinion and Policy in the American States*. Cambridge: Cambridge University Press.

Fay, Robert E., and Herriot, Roger A. 1979. "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association* 74: 269-77.

Feller, Avi, Andrew Gelman and Boris Shor. 2012. "Red State/Blue State Divisions in the 2012 Presidential Election." *The Forum* 10(4): 127-31.

Gelman, Andrew. 2004. "Parameterization and Bayesian Modeling." *Journal of the American Statistical Association* 99(466):537-45.

Gelman, Andrew. 2005. "Comment: Anova as a Tool for Structuring and Understanding Hierarchical Models." *Chance* 18(3):33.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel-Hierarchical Models*. Cambridge: Cambridge University Press.

Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau, and Yu-Sung Su. 2008. "A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models. Annals Applied Statistic. 2(4): 1360-83.

Gelman, Andrew, and Thomas C. Little. 1997. "Poststratification into Many Categories Using Hierarchical Logistic Regression." *Survey Methodology* 23(2): 127-35.

Gelman, Andrew and Yair Ghitza. 2013. "Deep Interactions with MRP: Election Turnout and Voting Patterns Among Small Electoral Subgroups." *American Journal of Political Science* 57(3): 762-76.

Gelman, Andrew, Jeffrey Lax, and Justin Phillips. 2010. "Over Time, a Gay Marriage Groundswell. *The New York Times*, page WK3.

Gelman, Andrew, David Park, Boris Shor, and Jeronimo Cortina. 2009. *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*, second edition. Princeton University Press.

Gelman, Andrew, Boris Shor, Joseph Bafumi, and David Park. 2007. "Rich State, Poor State, Red State, Blue State: Whats the Matter with Connecticut?" *Quarterly Journal of Political Science* 2:34567.

Gelman, Andrew, Nate Silver, and Daniel Lee. 2009. "The Senates Health Care Calculations." *New York Times*, 19 November, A35.

Gelman, Andrew, David van Dyk, Zaiying Huang, and John W. Boscardin. 2007. "Transformed and Parameter-Expanded Gibbs Samplers for Multilevel Linear and Generalized Linear Models." *Journal of Computational and Graphical Statistics*.

Kastellec, Jonathan. 2018. "Judicial Federalism and Representation." *Journal of Law and Courts* 6(1): 51-92.

Kastellec, Jonathan, Jeffrey Lax, and Michael Malecki, and Justin Phillips. 2015. “Distorting the Electoral Connection? Partisan Representation in Supreme Court Confirmation Politics.” *The Journal of Politics*.

Kastellec, Jonathan, Jeffrey R. Lax, and Justin H. Phillips. 2010. “Public Opinion and Senate Confirmation of Supreme Court Nominees.” *The Journal of Politics* 72(3): 767-84.

Krimmel, Kate, Jeffrey Lax, and Justin Phillips. 2016. “Gay Rights in Congress: Public Opinion and (Mis)Representation.” *Public Opinion Quarterly*.

Lax, Jeffrey R., and Justin H. Phillips. 2009a. “How Should We Estimate Public Opinion in the States?” *American Journal of Political Science* 53(1): 107-21.

Lax, Jeffrey R., and Justin H. Phillips. 2009b. “Gay Rights in the States: Public Opinion and Policy Responsiveness,” *American Political Science Review* 103(3): 367385.

Lax, Jeffrey and Justin H. Phillips. 2012. “The Democratic Deficit in the States,” *American Journal of Political Science* 56(1), 148166.

Lax, Jeffrey R., and Justin H. Phillips. 2013. “How Should We Estimate Sub-National Opinion Using MRP? Preliminary Findings and Recommendations” Presented at the Annual Meeting of the American Political Science Association, Chicago, IL.

Little, Roderick J. A. 1991. “Inference with survey weights.” *Journal of Official Statistics* 7: 405-424.

Little, Roderick J. A. 1993. “Post-stratification: a modelers perspective.” *Journal of the American Statistical Association* 88: 1001-12.

Norrander, Barbara. 2001. “Measuring State Public Opinion with the Senate National Election Study.” *State Politics and Policy Quarterly* 1(1): 111-25.

Norrander, Barbara. 2007. “Choosing Among Indicators of State Public Opinion.” *State Politics and Policy Quarterly* 7(2): 111.

Pacheco, Julianna. 2011. "Using National Surveys to Measure Dynamic U.S. State Public Opinion: A Guideline for Scholars and an Application." *State Politics and Policy Quarterly* 11:41539.

Pacheco, Julianna. 2012. "The Social Contagion Model: Exploring The Role of Public Opinion on the Diffusion of Anti-Smoking Legislation across the American States." *The Journal of Politics* 74(1): 187-202.

Park, David K., Andrew Gelman, and Joseph Bafumi. 2004. "Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls." *Political Analysis* 12 (4):375-85.

Park, David K., Andrew Gelman, and Joseph Bafumi. 2006. "State Level Opinions from National Surveys: Poststratification using Multilevel Logistic Regression," in *Public Opinion in State Politics* (ed. Jeffrey E. Cohen). Stanford, CA: Stanford University Press.

Pool, Ithiel de Sola, Robert Abelson, and Samuel L. Popkin. 1965. *Candidates, Issues, and Strategies*. Cambridge, MA: M.I.T. Press.

Salganik, Matthew, Tian Zheng, and Tyler H. McCormick. 2010. "How Many People Do You Know? Efficiently Estimating Network Size." *Journal of the American Statistical Association* 105(489): 59-70.

Selb, Peter and Simon Munzert. 2011. "Estimating Constituency Preferences from Sparse Survey Data Using Auxiliary Geographic Information." *Political Analysis* 19(3): 455-70.

Shirley, Kenneth E. and Andrew Gelman. 2014. "Hierarchical Models for Estimating State and Demographic Trends in US Death Penalty Public Opinion." *Journal of the Royal Statistical Society A*

Toshkov, Dimiter. 2015. "Exploring the Performance of Multilevel Modeling and Poststratification with Eurobarometer Data," *Political Analysis* 23(3):455-60.

Warshaw, Christopher and Jonathan Rodden. 2012. "How Should We Measure District-Level Public Opinion on Individual Issues?" *The Journal of Politics* 74(1): 203-19.

Zheng, Tian, Matthew J. Salganik, and Andrew Gelman. 2006 "How Many People Do You Know in Prison?: Using Overdispersion in Count Data to Estimate Social Structure in Networks." *Journal of the American Statistical Association* 101(474): 409-23.

Figure 1: RMSE Compared to State Polls

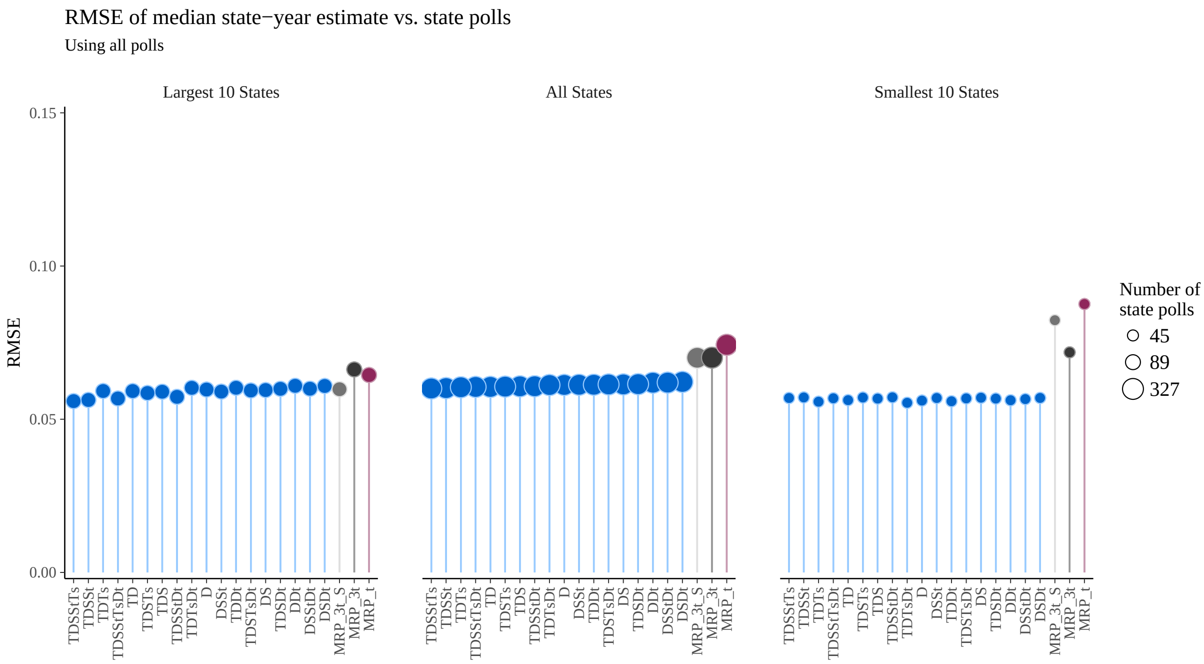


Figure 2: RMSE Compared to State Polls, Using Every 4th Poll

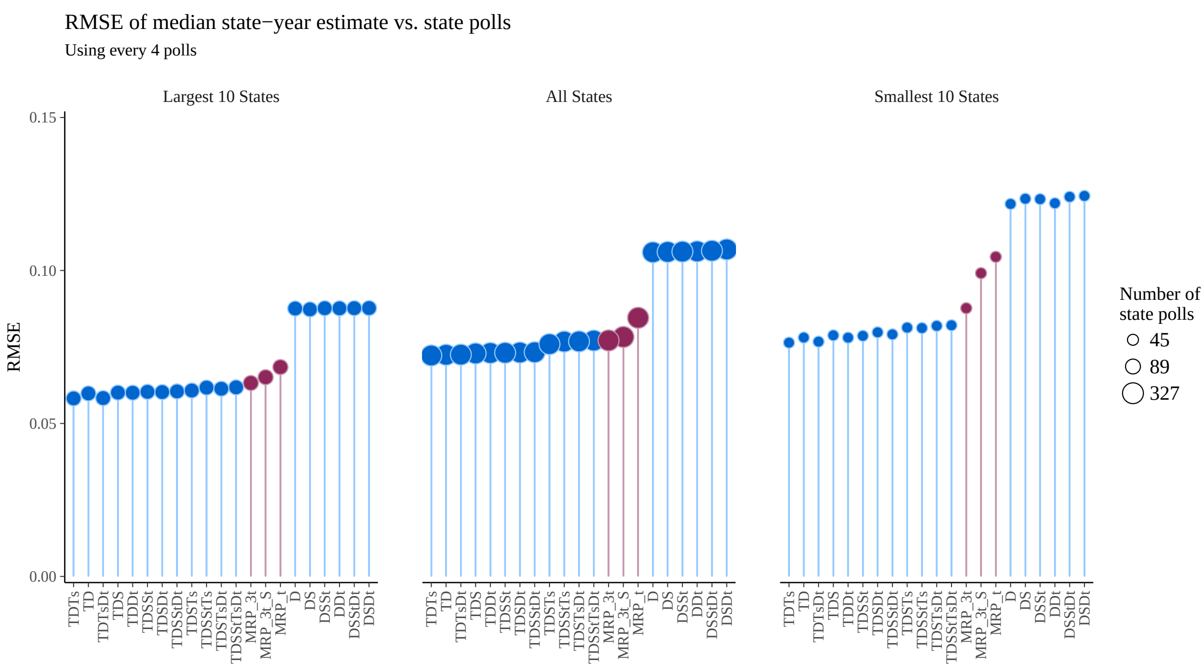


Figure 3: RMSE Compared to State Polls, Using Every 8th Poll

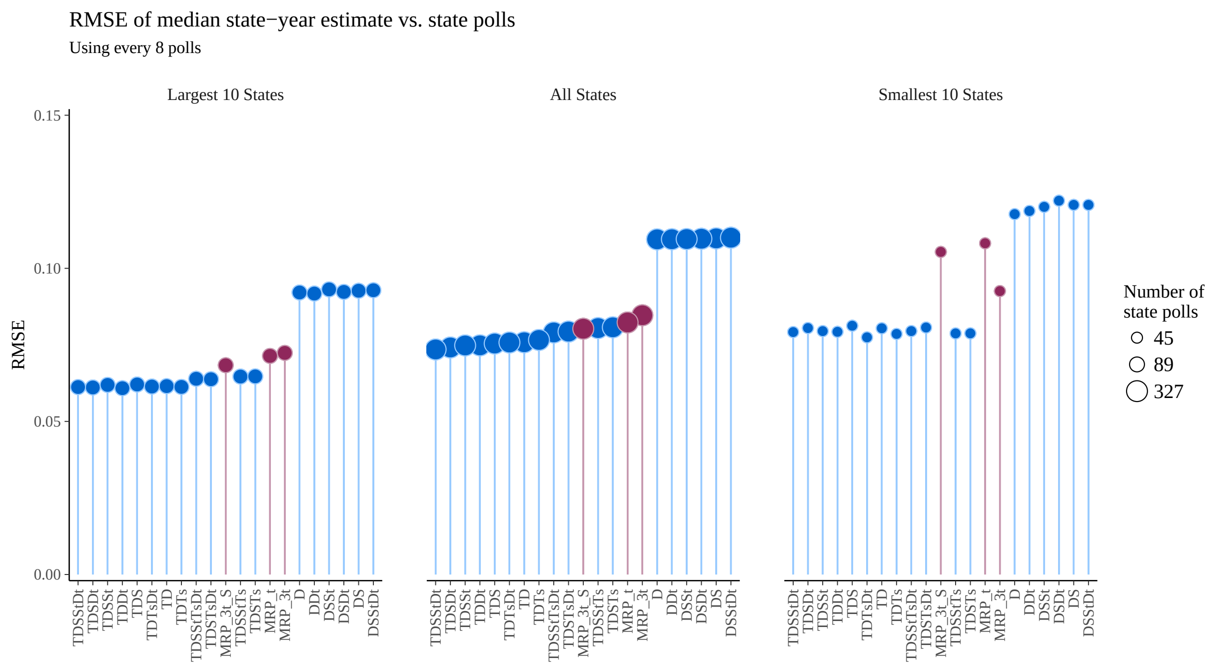


Figure 4: RMSE Compared to State Polls: All States with Every kth Poll Used

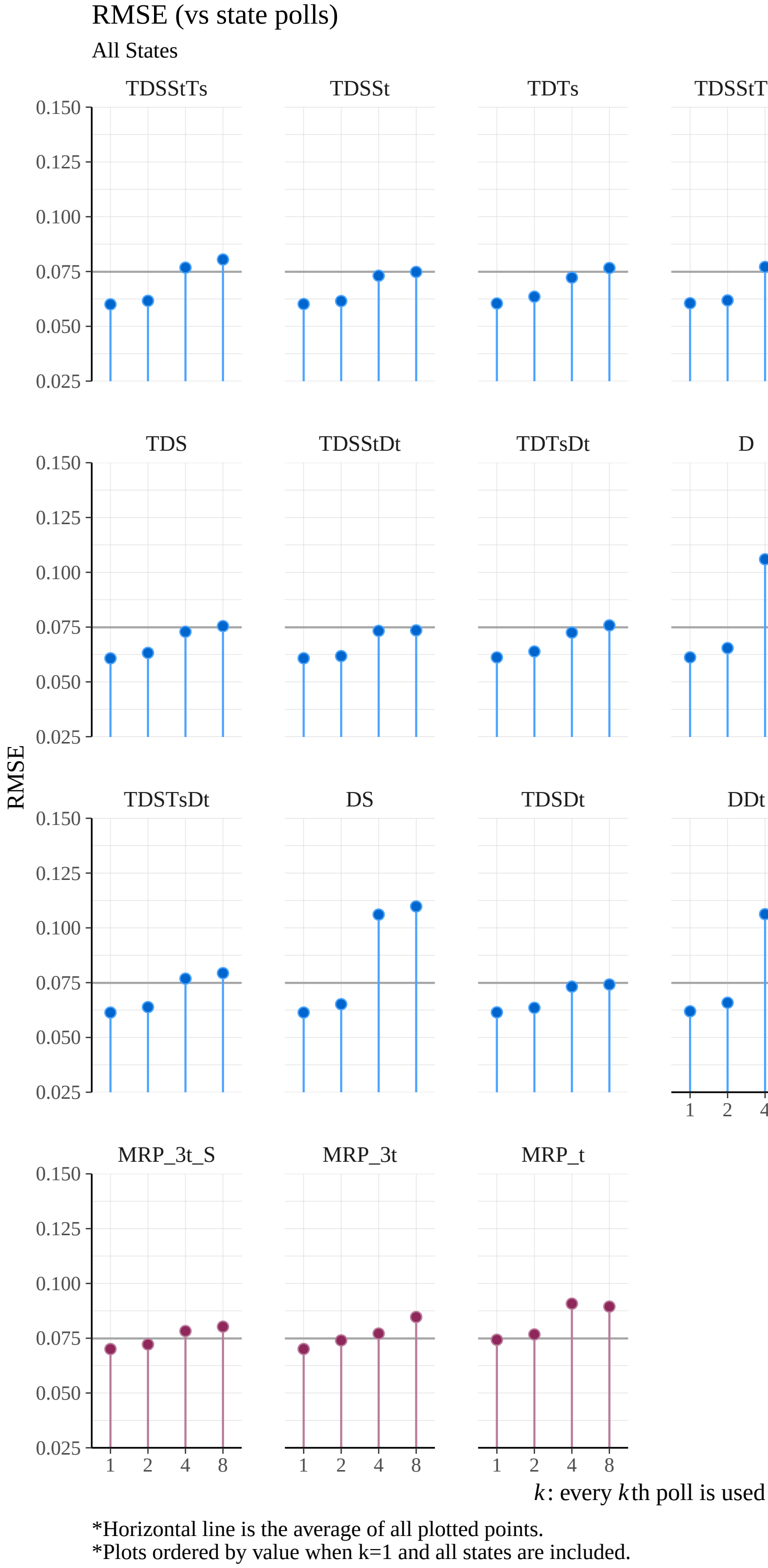
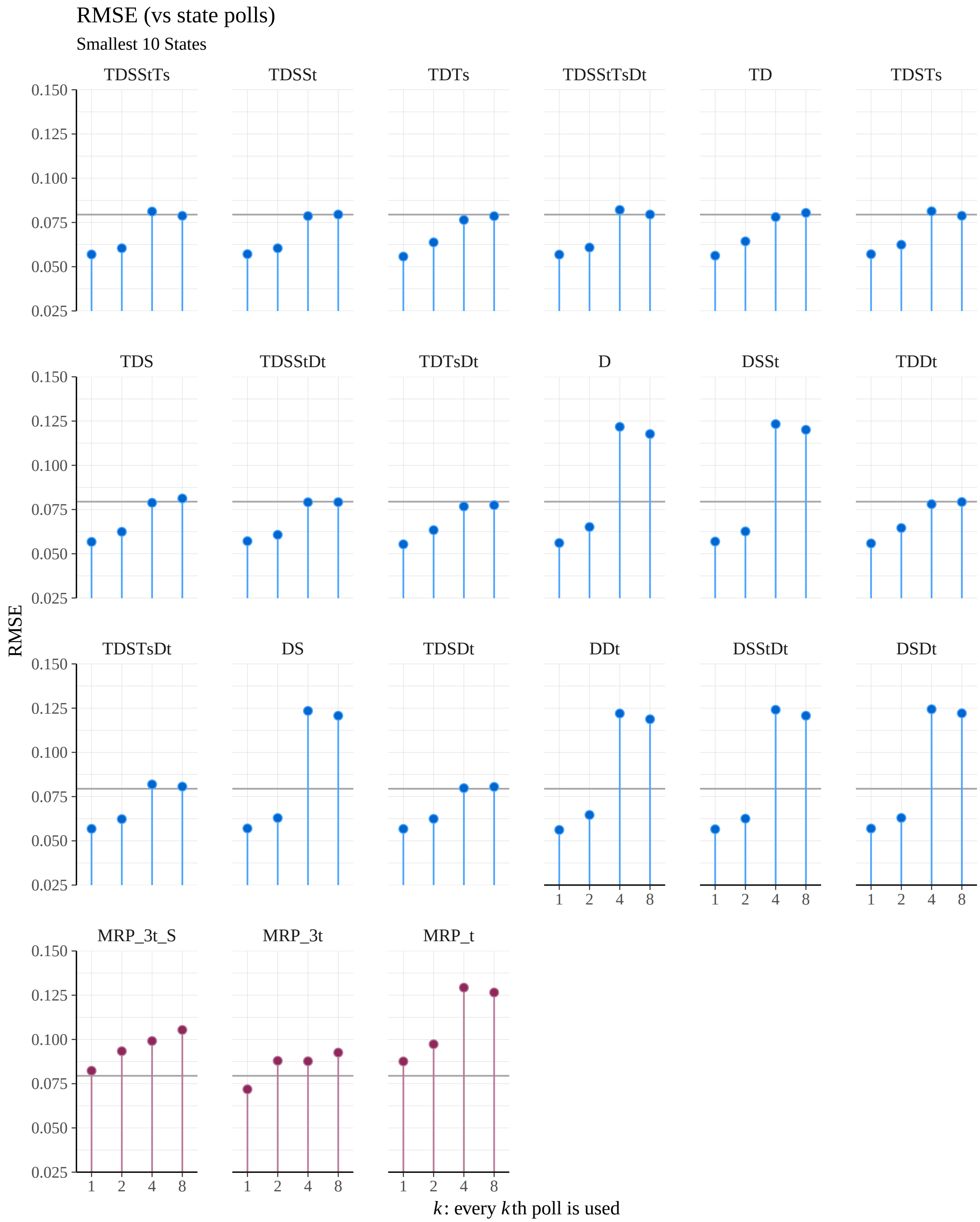


Figure 5: RMSE Compared to State Polls: Smallest 10 States with Every *k*th Poll Used



*Horizontal line is the average of all plotted points.
*Plots ordered by value when k=1 and all states are included.

Figure 6: Support for Same-Sex Marriage, using model TDS

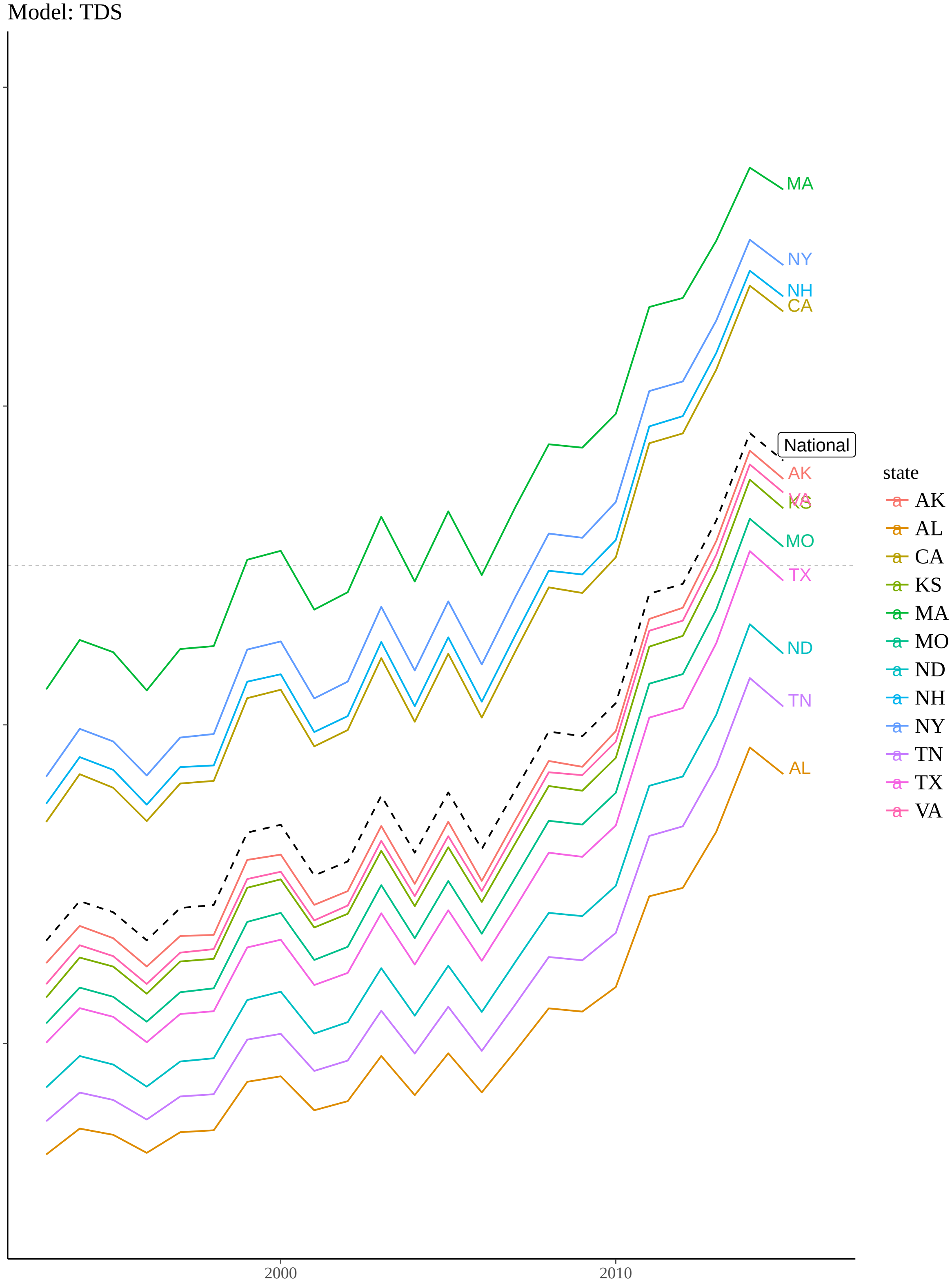
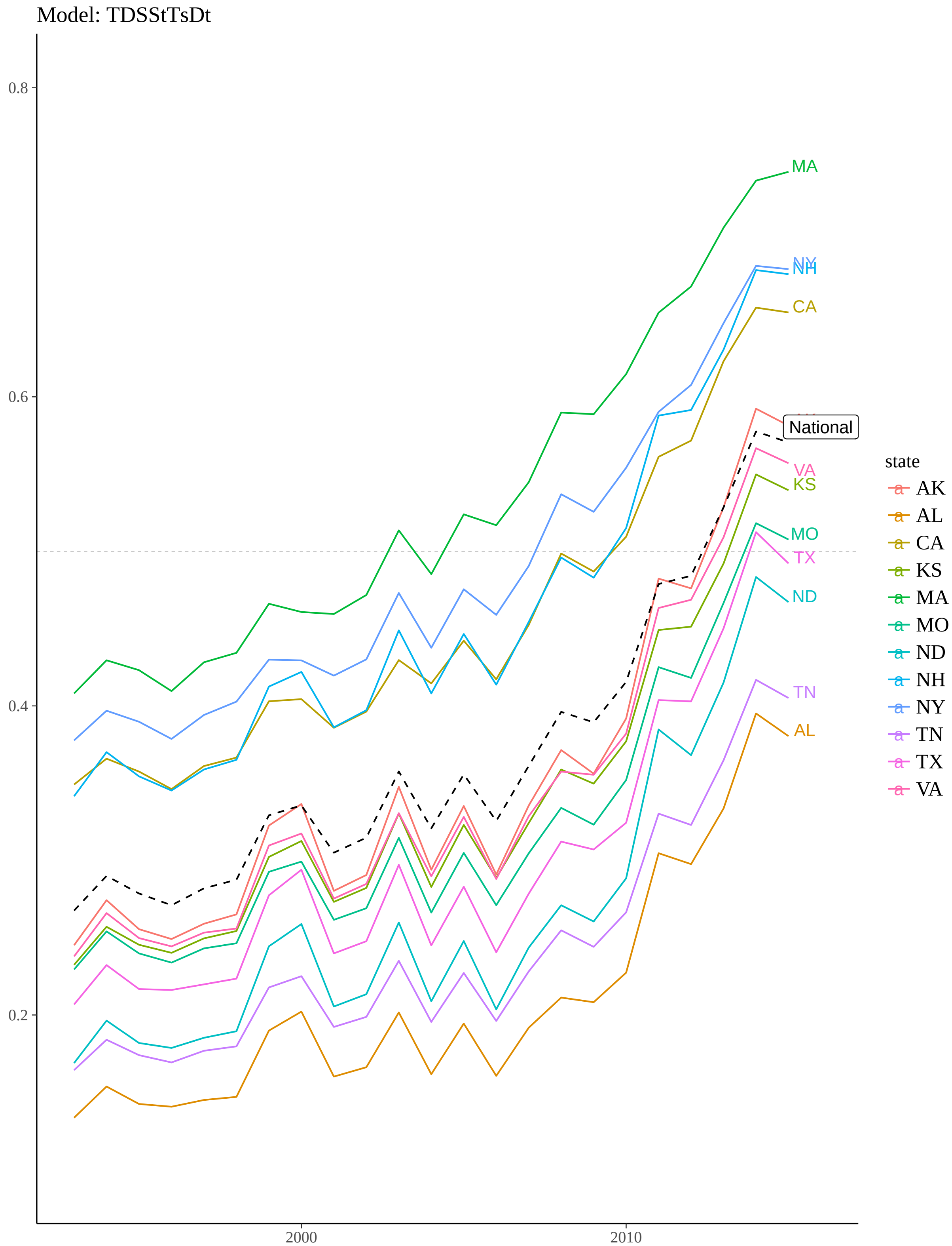


Figure 7: Support for Same-Sex Marriage, using model TDSStTsDt



Ridiculously Long Bonus Appendix

5 MRT Models Estimated in Paper

5.1 $\{D\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$
$$\alpha_i = \mu + \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state) +
        (1 | year)
```

5.2 $\{D, D_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$
$$+ \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state) +
        (1 | year) +
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

5.3 $\{D, S\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta\,\texttt{ideology}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state) +
        (1 | year)
```

5.4 $\{D, S, D_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta\, \texttt{ideology}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

$$+ \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state) +
        (1 | year) +
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

5.5 $\{D, S, S_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta_{\text{year[i]}} \, \texttt{ideology}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology
        (1 + ideology | year) +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state)
```

5.6 $\{D, S, S_t, D_t\}$

Statistical model:

$$
\begin{aligned}
n_i^{\text{favor}} &\sim \text{BinomialLogit}(N_i, \alpha_i) \\
\alpha_i &= \mu + \delta_{\text{year}[i]}\,\texttt{ideology} \\
&\quad + \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]} \\
&\quad + \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}
\end{aligned}
$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology
        (1 + ideology | year) +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state) +
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

5.7 $\{T, D\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \gamma_1 \, \texttt{year\_std} + \gamma_2 \, \texttt{year\_sq\_std}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        year_std + year_sq_std +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state)
        (1 | year)
```

5.8 $\{T, D, D_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \gamma_1 \, \texttt{year\_std} + \gamma_2 \, \texttt{year\_sq\_std}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

$$+ \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        year_std + year_sq_std +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state)
        (1 | year)
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

5.9 $\{T, D, S\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta\,\texttt{ideology} + \gamma_1\,\texttt{year\_std} + \gamma_2\,\texttt{year\_sq\_std}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        year_std + year_sq_std +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state)
        (1 | year)
```

5.10 $\{T, D, T_s\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta\,\texttt{ideology} + \gamma_{1,\text{state}[i]}\,\texttt{year\_std} + \gamma_{2,\text{state}[i]}\,\texttt{year\_sq\_std}$$
$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        year_std + year_sq_std +
        (1 + year_std + year_sq_std | state) +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | year)
```

5.11 $\{T, D, S, D_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta\,\texttt{ideology} + \gamma_1\,\texttt{year\_std} + \gamma_2\,\texttt{year\_sq\_std}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

$$+ \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        year_std + year_sq_std +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state) +
        (1 | year) +
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

5.12 $\{T, D, S, S_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta_{\text{year}[i]} \, \texttt{ideology} + \gamma_1 \, \texttt{year\_std} + \gamma_2 \, \texttt{year\_sq\_std}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        (1 + ideology | year) +
        year_std + year_sq_std +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state)
```

5.13 $\{T, D, S, T_s\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta\,\texttt{ideology} + \gamma_{1,\text{year}[i]}\,\texttt{year\_std} + \gamma_{2,\text{year}[i]}\,\texttt{year\_sq\_std}$$
$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        year_std + year_sq_std +
        (1 + year_std + year_sq_std | state) +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | year)
```

5.14 $\{T, D, T_s, D_t\}$

Statistical model:

$$
\begin{aligned}
n_i^{\text{favor}} &\sim \text{BinomialLogit}(N_i, \alpha_i) \\
\alpha_i &= \mu + \gamma_{1,\text{year}[i]}\, \texttt{year\_std} + \gamma_{2,\text{year}[i]}\, \texttt{year\_sq\_std} \\
&\quad + \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]} \\
&\quad + \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}
\end{aligned}
$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        year_std + year_sq_std +
        (1 + year_std + year_sq_std | state) +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | year)
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

5.15 $\{T, D, S, S_t, D_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta_{\text{year}[i]}\,\texttt{ideology} + \gamma_1\,\texttt{year\_std} + \gamma_2\,\texttt{year\_sq\_std}$$
$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$
$$+ \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        (1 + ideology | year) +
        year_std + year_sq_std +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state)
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

5.16 $\{T, D, S, S_t, T_s\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta_{\text{year}[i]}\,\texttt{ideology} + \gamma_{1,\text{year}[i]}\,\texttt{year\_std} + \gamma_{2,\text{year}[i]}\,\texttt{year\_sq\_std}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        (1 + ideology | year) +
        year_std + year_sq_std +
        (1 + year_std + year_sq_std | state) +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race)
```

5.17 $\{T, D, S, T_s, D_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta\, \texttt{ideology} + \gamma_{1,\text{year}[i]}\, \texttt{year\_std} + \gamma_{2,\text{year}[i]}\, \texttt{year\_sq\_std}$$
$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$
$$+ \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        year_std + year_sq_std +
        (1 + year_std + year_sq_std | state) +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | year)
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

5.18 $\{T, D, S, S_t, T_s, D_t\}$

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta_{\text{year}[i]}\, \texttt{ideology} + \gamma_{1,\text{year}[i]}\, \texttt{year\_std} + \gamma_{2,\text{year}[i]}\, \texttt{year\_sq\_std}$$

$$+ \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]} + \beta_{\text{year}[i]}$$

$$+ \beta_{\text{age}[i],\text{year}[i]} + \beta_{\text{edu}[i],\text{year}[i]} + \beta_{\text{sex\_race}[i],\text{year}[i]} + \beta_{\text{state}[i],\text{year}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        (1 + ideology | year) +
        year_std + year_sq_std +
        (1 + year_std + year_sq_std | state) +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | age:year) +
        (1 | edu:year) +
        (1 | sex_race:year) +
        (1 | state:year)
```

6 MRP Model Estimated in Paper

Model 0 (MRP-t)

Statistical model:

$$n_i^{\text{favor}} \sim \text{BinomialLogit}(N_i, \alpha_i)$$

$$\alpha_i = \mu + \delta\,\texttt{ideology} + \beta_{\text{age}[i]} + \beta_{\text{edu}[i]} + \beta_{\text{sex\_race}[i]} + \beta_{\text{state}[i]}$$

`lmer` syntax:

```
cbind(favor, oppose) ~
        ideology +
        (1 | age) +
        (1 | edu) +
        (1 | sex_race) +
        (1 | state)
```