Leveraging Time-Series Information to Improve Small-Area Estimation

Supplementary Information

Isaac D. Mehlhaff*

Contents

*The University of Chicago; imehlhaff@uchicago.edu.

A Model Estimation

I fit all models in a fully Bayesian framework, which allows me to easily design more complex models than standard frequentist computational tools would allow. It also enables me to produce uncertainty estimates as a direct byproduct of the model-fitting process, which are important for assessing the efficiency of each modeling strategy relative to the otheres. Typically, the default method for fitting Bayesian models like these is to employ Markov chain Monte Carlo (MCMC) techniques to explore the full parameter space. Unfortunately, even the most advanced MCMC methods can be extremely slow to converge when the posterior is complex or the number of data points is large, as are both the case when it comes to the fully dynamic models.

Instead, I use variational inference to fit the models. Variational inference is a method of posterior approximation that is guaranteed to converge, and is easily assessed by convergence criteria. It operates by proposing a family of candidate distributions for the true (analytically intractable) posterior and finds the member of that distribution family most closely resembling the true posterior by minimizing the Kullback-Leibler divergence between the two (Grimmer, 2011). Variational inference has become an indispensable tool in computer science and statistics (e.g. Airoldi et al., 2008; Blei et al., 2003), but its application to political science has been more limited (for examples, see Grimmer, 2013; Imai et al., 2016). Because variational inference is not deterministic and can identify a slightly different posterior distribution each time it is run, I fit each model ten times and save only the model fit with the maximum evidence lower bound (ELBO)—the criterion used to monitor the variational algorithm. When the ELBO is maximized, the Kullback-Leibler divergence is minimized, indicating the estimated posterior is a better match for the true posterior.

Though variational inference can be a powerful tool, it would be helpful to know how well the variational estimates approximate posterior estimates from the more common (and sometimes more reliable) MCMC techniques. The computational resources required to fit each model using MCMC makes it impractical to fit all twenty-nine policy issues; fitting just one model to one time series takes several hours to converge. I therefore focus only on the same-sex marriage issue, a common application in MRP research and one I explore more fully in section D below. I fit the models using MCMC, drawing 1500 samples on each of four chains, with the first 500 samples discarded as warm-up.

Two fit diagnostics, R-hat and effective sample size, suggest the MCMC models successfully converged to the posterior. The distributions of these two diagnostics are included in Figures S1 and S2, respectively. All R-hat statistics are well below the conventional cutoff of 1.1 (Gelman et al., 2013) and all effective sample sizes are sufficiently large, indicating the chains have mixed and likely provide reliable parameter estimates.

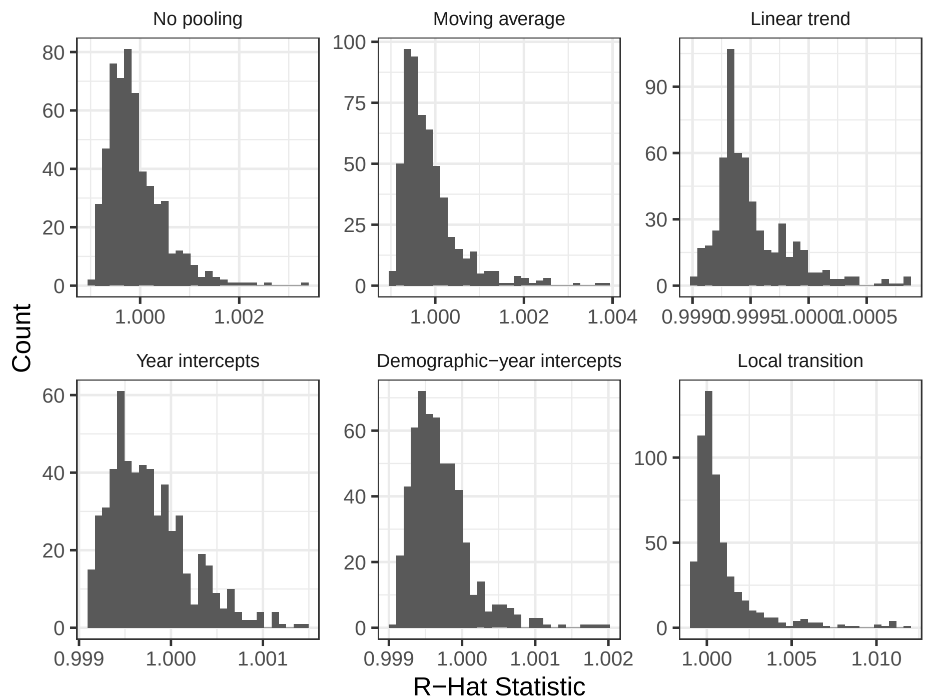

Figure S1: R-Hat Statistics for Poststratification Estimates by Model

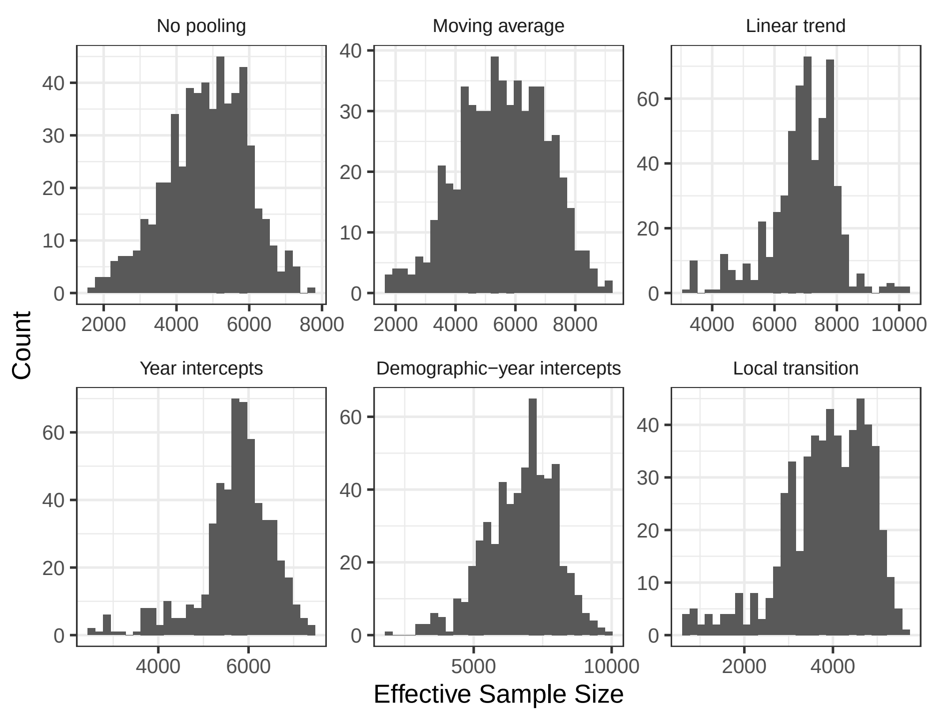

Figure S2: Effective Sample Sizes for Poststratification Estimates by Model

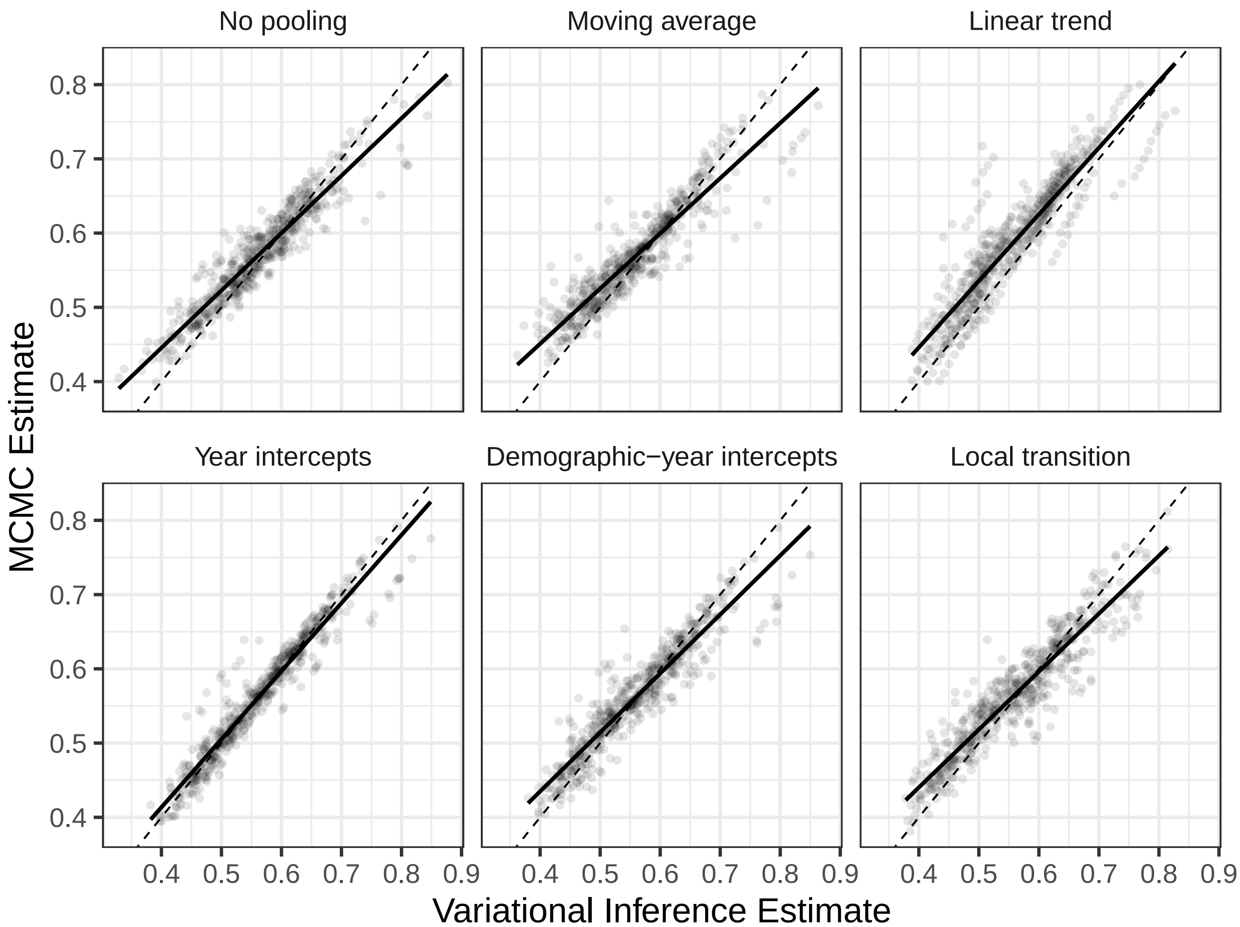

Figure S3: Correlation Between Model Estimates using Variational Inference and Markov Chain Monte Carlo. Dashed lines show $y = x$ line, solid lines show line of best fit.

Figure S3 displays the relationship between poststratification estimates when using MCMC and variational inference. For all six models, estimates are very closely related regardless of which estimation strategy is used. Correlations range from 0.91 for the linear trend model to 0.99 for the model with demographic-year intercepts. Using variational inference to approximate the posterior appears to produce estimates that are nearly identical to those produced with the more common MCMC approach.

B CES Analysis

B.1 Survey Items

I use twenty-nine survey items from the Cooperative Election Studies (CES), which repeatedly asked a nationally representative sample of Americans a wide range of policy issue items over multiple years. There were two criteria for inclusion: A policy issue must have been asked in at least four consecutive years and it must have been either a binary survey item or an ordinal survey item that could be converted to binary. For example, an item that gave the response options "strongly agree," "agree," "disagree," and "strongly disagree" was eligible for inclusion, but items that included a response option like "neither agree nor disagree" or that were asked on a scale from 0-100 were ineligible. Table S1 provides survey item wording for each of the twenty-nine policy items, as well as for the four demographic items I include as individual-level predictors in the MRP models. Note that some of the policy items may have slight wording changes over time that are not reflected here.

Table S1: Survey Item Wording

Variable	Question	Response Options	Years
	Demographics		
Gender	What is your gender?	Male, Female	2006-2021
Race	What racial or ethnic group best describes you?	White, Black, Hispanic, Asian, Native American, Middle Eastern, Mixed, Other	2006-2021
Age	In what year were you born?		2006-2021
Education	What is the highest level of education you have completed?	No HS, High school graduate, Some college, 2-year, 4-year, Post-grad	2006-2021
	Policy Issues		
Abortion access	There has been some discussion about abortion during recent years. Which one of the opinions on this page best agrees with your view on this issue? (2006-2013); Always allow a woman to obtain an as a matter of choice (2014-2021)	By law, abortion should never be permitted, The law should permit abortion only in case of rape, incest or when the woman's life is in danger, The law should permit abortion for reasons other than rape, incest, or danger to the woman's life, but only after the need for the abortion has been clearly established, By law, a woman should always be able to obtain an abortion as a matter of personal choice (2006-2013); Support, Oppose (2014-2021)	2006-2021
Employer covers abortion	Allow employers to decline coverage of abortions in insurance plans	Support, Oppose	2014-2018
Prohibit funding abortion	Prohibit the expenditure of funds authorized or appropriated by federal law for any abortion	Support, Oppose	2014-2019
Regulate carbon	Give Environmental Protection Agency power to regulate Carbon Dioxide emissions	Support, Oppose	2014-2021
Raise fuel efficiency	Raise required fuel efficiency for the average automobile from 25 mpg to 35 mpg	Support, Oppose	2014-2021

Variable	Question	Response Options	Years
Require renewable fuel	Require that each state use a minimum amount of renewable fuels (wind, solar, and hydroelectric) in the generation of electricity even if electricity prices increase a little	Support, Oppose	2014-2021
Clean Air/Water Acts	Strengthen the Environmental Protection Agency enforcement of the Clean Air Act and Clean Water Act even if costs US jobs	Support, Oppose	2014-2021
Background checks	Background checks for all sales, including at gun shows and over the Internet	Support, Oppose	2013-2021
Reveal gun owners	Prohibit state and local governments from publishing the names and addresses of all gun owners	Support, Oppose	2013-2017
Ban assault rifles	Ban assault rifles	Support, Oppose	2013-2021
Concealed-carry	Make it easier for people to obtain concealed-carry permit	Support, Oppose	2013-2021
Repeal ACA	Repeal Affordable Care Act	Support, Oppose	2012-2021
Medicare for all	Expand Medicare to a single comprehensive public health care coverage program that would cover all Americans	Support, Oppose	2018-2021
Legal status	Grant legal status to all illegal immigrants who have held jobs and paid taxes for at least 3 years, and not been convicted of any felony crimes	Support, Oppose	2010-2021
Border security	Increase the number of border patrols on the US-Mexican border	Support, Oppose	2010-2021
Police question undocumented	Allow police to question anyone they think may be in the country illegally	Support, Oppose	2010-2017
Sanction undocumented hiring	Fine U.S. businesses that hire illegal immigrants	Support, Oppose	2012-2017

Variable	Question	Response Options	Years
Deport undocumented	Identify and deport illegal immigrants	Support, Oppose	2014-2017
Police report undocumented	Withhold federal funds from any local police department that does not report to the federal government anyone they identify as an illegal immigrant	Support, Oppose	2017-2021
Ensure oil supply	Would you approve of the use of U.S. military troops in order to ensure the supply of oil?	Support, Oppose	2010-2016
Destroy terrorist camp	Would you approve of the use of U.S. military troops in order to destroy a terrorist camp?	Support, Oppose	2010-2016
Stop genocide	Would you approve of the use of U.S. military troops in order to intervene in a region where there is genocide or a civil war?	Support, Oppose	2010-2016
Spread democracy	Would you approve of the use of U.S. military troops in order to assist the spread of democracy?	Support, Oppose	2010-2016
Protect allies	Would you approve of the use of U.S. military troops in order to protect American allies under attack by foreign nations?	Support, Oppose	2010-2016
Help UN	Would you approve of the use of U.S. military troops in order to help the United Nations uphold international law?	Support, Oppose	2010-2016
China tariffs	Tariffs on 200 billion dollars worth of goods imported from China	Support, Oppose	2018-2021
Canada/Mexico tariffs	25 percent tariffs on all imported steel and 10 percent on imported aluminum, INCLUDING from Canada and Mexico	Support, Oppose	2018-2021

Variable	Question	Response Options	Years
Same-sex marriage	President Bush recently spoke out in favor of a Constitutional Amendment defining marriage as strictly between a man and a woman. Do you support or oppose a Constitutional amendment banning gay marriage? (2006-2007); Do you support a Constitutional Amendment banning Gay Marriage? (2008-2011); Do you favor or oppose allowing gays and lesbians to marry legally? (2012-2016)	Strongly support, Somewhat support, Somewhat oppose, Strongly oppose (2006-2007); Support, Oppose (2008-2016)	2006-2016
Affirmative action	Affirmative action programs give preference to racial minorities and to women in employment and college admissions in order to correct for discrimination. Do you support or oppose affirmative action?	Strongly support, Somewhat support, Somewhat oppose, Strongly oppose	2008-2014

B.2 Results with Unweighted Benchmark Values

Several scholars have noted the thoughful debate on the appropriateness of using survey weights in conjunction with MRP to correct for non-representativity (Bisbee, 2019; Gelman, 2013; Lax & Phillips, 2009b). Results in the main text use weights to correct population-level benchmarks, but sample data fed to the MRP models remains unweighted. By design, the poststratification stage in MRP should correct for the same sorts of variation. I make this decision on the grounds that survey weights tend to bring statistical estimates closer in line with the "ground truth," but because I am comparing MRP models to each and not to disaggregation or another method of small-area estimation, whether or not I weight benchmarks matters relatively little for the main purpose of the analysis—assessing performance among MRP models.

Nevertheless, I can also calculate error metrics using unweighted data. Figure S4 displays these metrics, replicating Figure 1 in the main text. Overall, substantive takeaways remain mostly unchanged. There is still a great deal of variation among models within issues and among issues within models. Models with linear trends appear to perform slightly better when compared to this unweighted benchmark—it is less likely to produce severely biased estimates relative to other

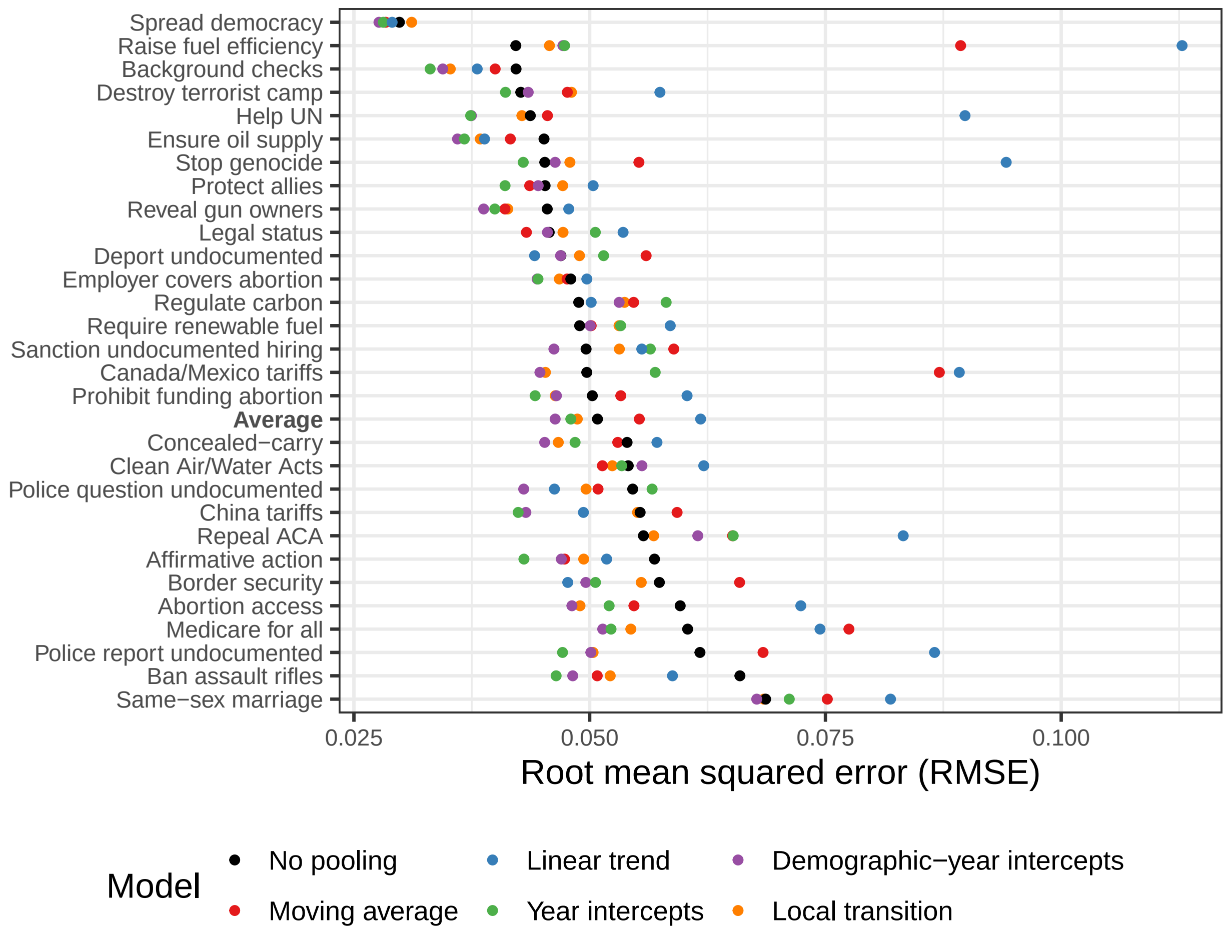

Figure S4: RMSE of CES Time Series Estimates using Unweighted Benchmarks. Performance metrics reflect model estimates for all state-years within each policy issue. Policy issues ordered on y-axis according to no-pooling RMSE.

models of the same issue—but it is still a poorly performing model in general. From this analysis, little conclusive evidence exists to suggest the clear superiority of one model over the others.

B.3 Alternative Error Metrics

In the main text, I used root mean squared error (RMSE) to assess model performance on the twenty-nine CES time series. This section presents alternative error metrics: mean squared error (MSE) in Figure S5, mean absolute error (Bisbee, 2019) in Figure S6, standardized bias (Buttice & Highton, 2013) in Figure S7, and correlation in Figure S8. All metrics tell a similar story to the ones presented in the main text: Model performance varies across issues, there is no one model that frequently performs best, and the gap between the best- and worst-performing model on any given issue can be wide, though correlation metrics tend to be slightly less variable.

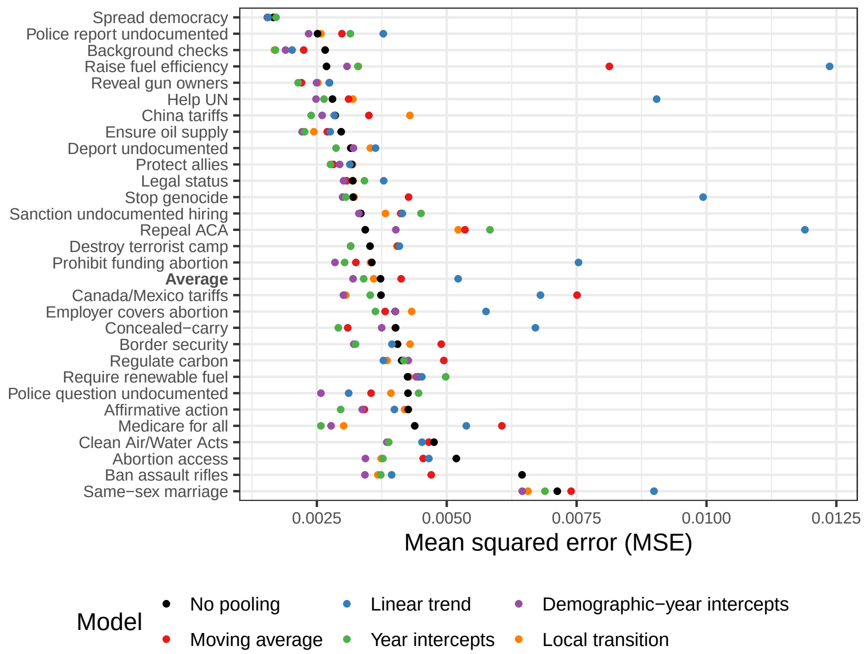

Figure S5: MSE of CES Time Series Estimates. Performance metrics reflect model estimates for all state-years within each policy. Policies ordered on y-axis according to no-pooling MSE.

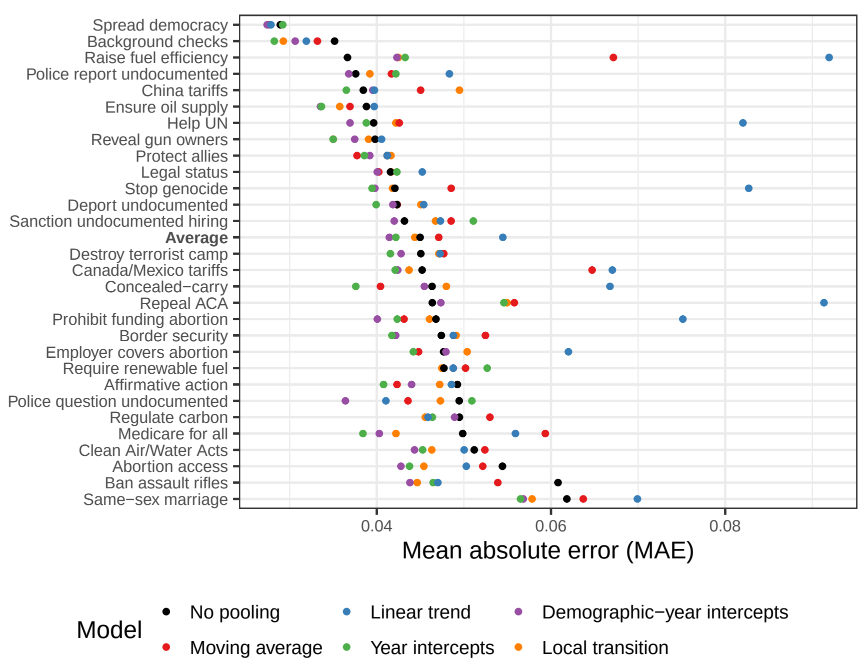

Figure S6: MAE of CES Time Series Estimates. Performance metrics reflect model estimates for all state-years within each policy. Policies ordered on y-axis according to no-pooling MAE.

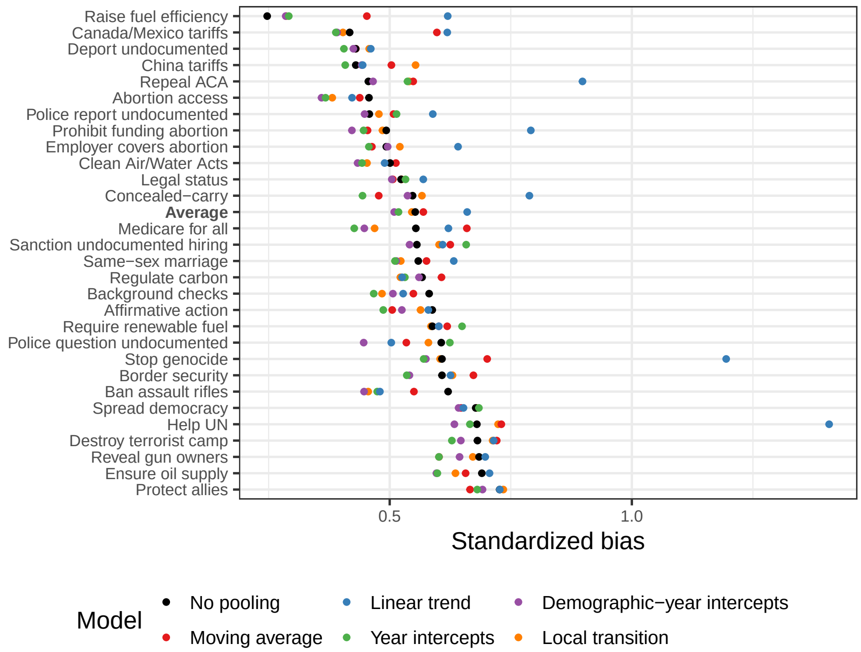

Figure S7: Standardized Bias of CES Time Series Estimates. Performance metrics reflect model estimates for all state-years within each policy. Policies ordered on y-axis according to no-pooling standardized bias.

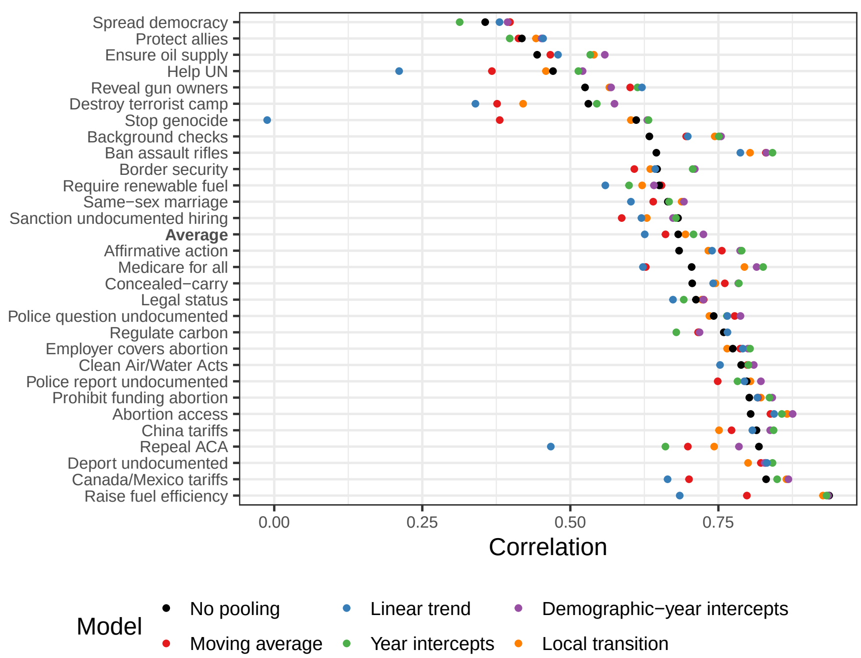

Figure S8: Correlation of CES Time Series Estimates with Benchmarks. Performance metrics reflect model estimates for all state-years within each policy. Policies ordered on y-axis according to no-pooling correlation.

B.4 Time Series Characteristics

The main text describes how model performance can vary depending on how volatile the time series is and how many time periods are available to the model. The left plot of Figure S9 displays the standard deviation of opinion across years, within states to give a sense of how much the population-level opinion is changing over time for each issue. The right plot shows how many years of data are available for each issue. I do not consider survey items with less than three consecutive years of data.

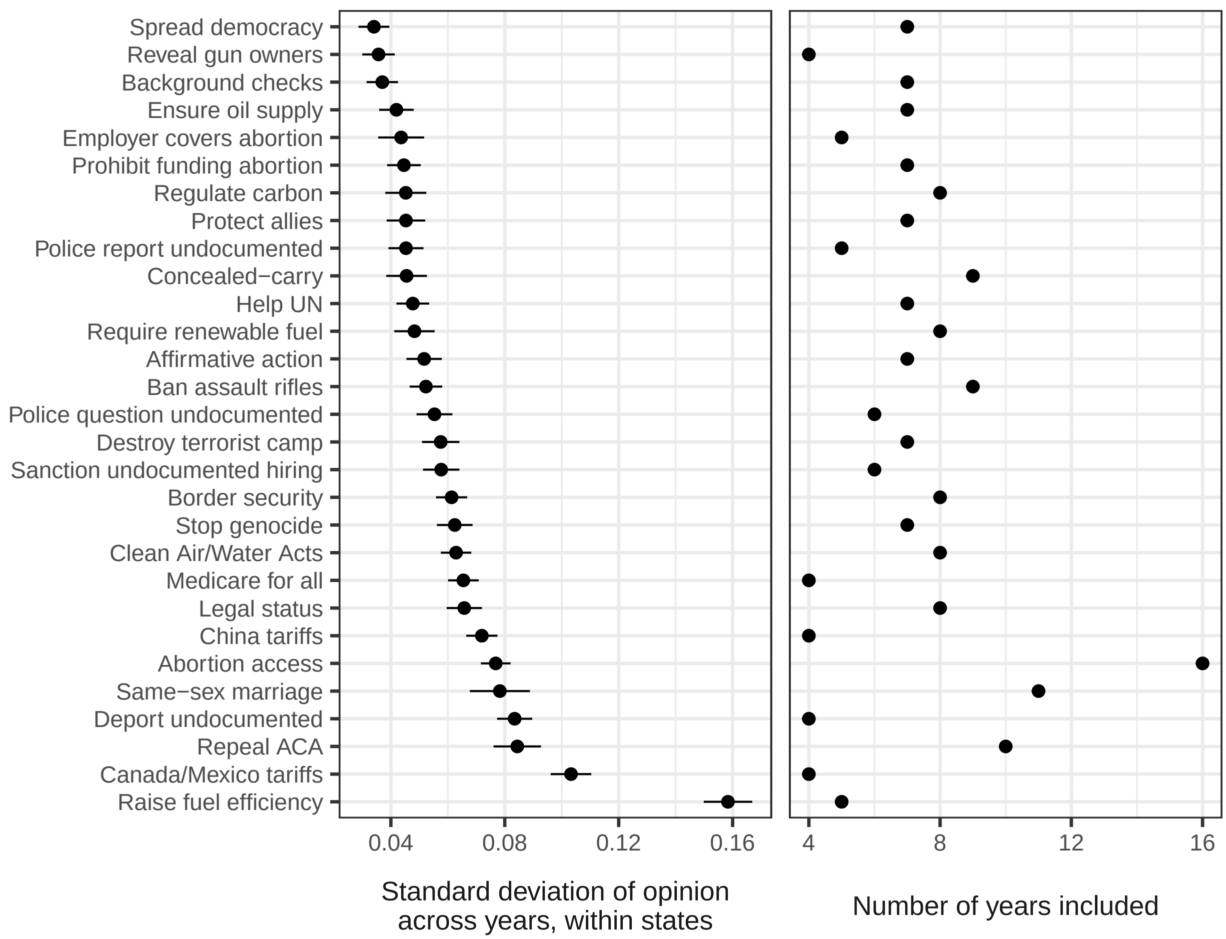

Figure S9: Over-Time Change and Length of CES Items. Policy issues ordered on y-axis according to standard deviation across years. Error bars give 95% confidence intervals.

C Computational Efficiency

Scholars must often choose between model performance and computational efficiency; less accurate models can be fit rather quickly while more accurate models may require more consider-able resources, especially when the number of estimated parameters or the number of observations are large. Simulation evidence above showed that more complex models—particularly those with demographic-year random intercepts and local-level transitions—tend to outperform other dynamic MRP models across use cases. But do these models also come with a higher computational price tag and, if so, is the gain in performance worth the extra cost to fit the model?

To assess this tradeoff, I run each model on one CPU and record the peak memory (RAM) and elapsed time required to fit the model. I begin with one time period and add time periods one by one, calculating computational efficiency each time. I repeat this process ten times and average over these ten iterations.

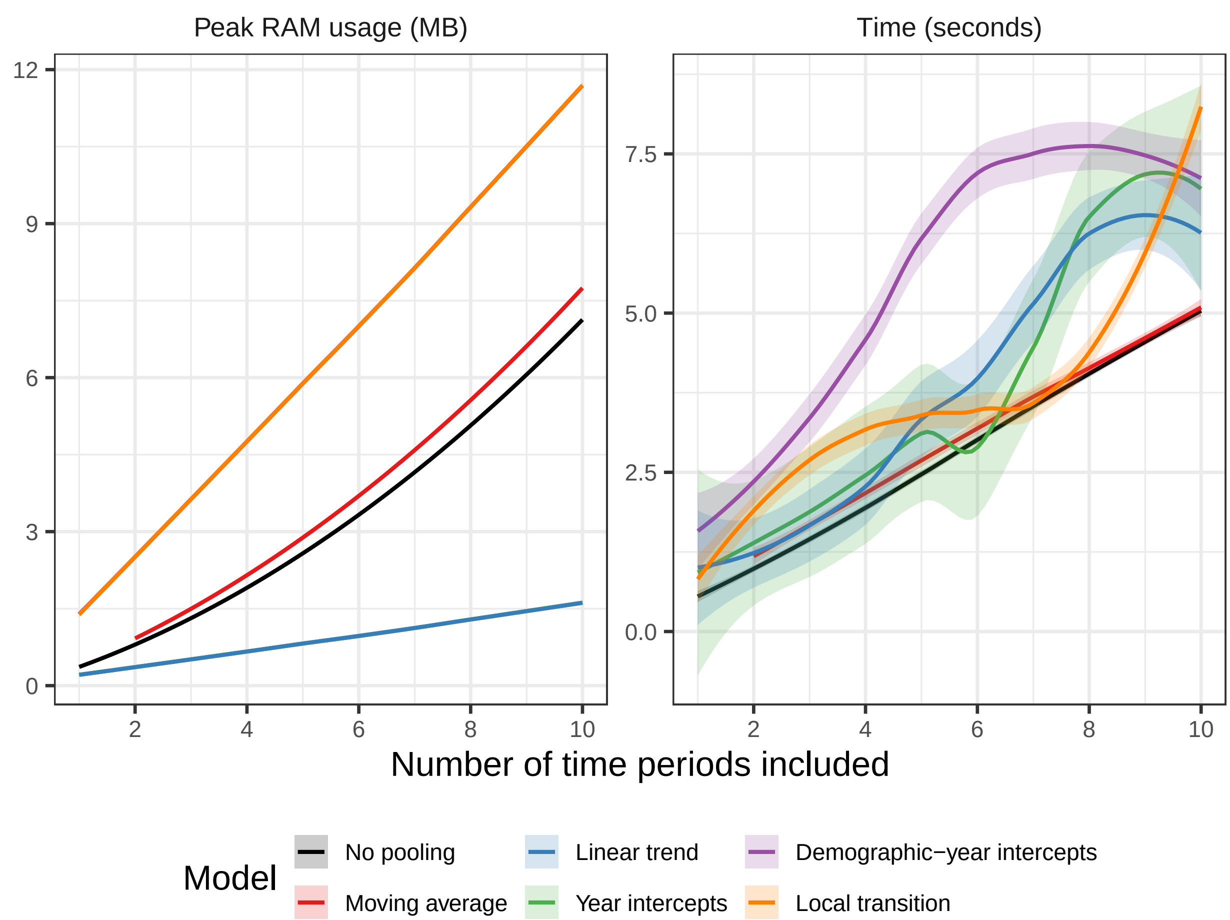

Figure S10: Computational Efficiency as T Increases. Peak RAM usage in megabytes and elapsed real time in seconds with one CPU. Trend lines show 95% confidence intervals across ten iterations.

Figure S10 displays the RAM, in megabytes (MB), and time, in seconds, required to fit each model as the number of time periods increases. The moving average model is identical to a no-pooling model when $T = 1$, so I only estimate the moving average model for all $T \geq 2$. When there are few time periods, the difference in memory usage among all six models is relatively minimal. However, models with demographic-year random intercepts and local-level transitions—owing to their larger inventory of parameters—consistently require more memory than all other models, and the rate of memory usage increase is greater for these two models than for the other dynamic models. Because the no-pooling and moving average models must be fit once for each time period, their memory requirements increase exponentially while the requirements to fit the dynamic models—which need only be fit once—increase linearly.

The dynamic models are also slightly more computationally expensive when it comes to the time required to fit the model. When there are only a handful of time periods, all six models can be fit in approximately the same amount of time. However, each additional time period after five or six begins to increase computation time for dynamic models more than it does for no-pooling and moving average models. Models with demographic-year random intercepts are especially time-intensive, with an appreciable difference beginning to show up after just three or four time periods. However, time requirements for this model also level off at higher numbers of time periods, at which point computational requirements for the local-level transition model begin to increase rapidly. Computational efficiency metrics will vary for each unique dataset, but this analysis suggests that when scholars need to produce small-area estimates for many time periods, they can expect to pay for enhanced model performance with their time. When the number of time periods is few, the more performant models take nearly the same amount of time as the simpler ones and require comparable amounts of memory.

D Application: Same-Sex Marriage Policy Responsiveness

One of the most prominent applications of MRP estimates is to questions of policy responsiveness and substantive representation (Kastellec et al., 2010; Pacheco, 2013; Tausanovitch & Warshaw, 2014). Scholars especially made significant gains in understanding the link between public support for same-sex marriage and its gradual, state-by-state legalization throughout the the late 2000s and early 2010s (Lax & Phillips, 2009a; Lewis & Jacobsmeier, 2017; Warshaw, 2016). To align dynamic MRP with this research agenda, I focus on the yearly estimates of same-sex marriage support generated in the CES analysis in the main text and in section B. State-level estimates run from 2006 to 2015, when the Supreme Court recognized same-sex marriage at the federal level in *Obergefell v. Hodges*. I combine these subnational opinion estimates with information about whether a state allowed same-sex marriage in a given year to produce an analysis similar to the one conducted by Caughey and Warshaw (2019).

To assess the degree to which changes in public support for same-sex marriage are associated with subsequent moves to legalize same-sex marriage in the states, I fit a series of two-way fixed effects models using the legal status of same-sex marriage in each state-year as the dependent variable and the subnational time-series data generated by each of the six models from the main text as the only explanatory variable. I do not intend to produce a dispositive test of policy responsiveness—many other scholars have addressed that question. I simply aim to demonstrate that model accuracy and efficiency can have meaningful effects on downstream analyses by deploying a model that takes advantage of the dynamic nature of the data.

Table S2: Effect of Public Support on Same-Sex Marriage Legalization

	No-pooling	Moving average	Linear trend	Year intercepts	Demographic-year intercepts	Local transition
Measurement error excluded	0.6^* (0.093)	0.99^* (0.184)	-7.113 (8.881)	5.401 (8.768)	0.877 (0.771)	0.383 (0.286)
Measurement error included	0.283 (0.262)	0.387 (0.333)	-6.986^* (3.047)	5.108 (3.181)	0.173 (0.563)	0.177 (0.338)

Note: $^*p < 0.05$. Coefficients on state-year public support from two-way fixed effects models. Values in parentheses give standard errors of coefficient estimates.

The top row of Table S2 displays the coefficient estimates from the two-way fixed effects models using state-level opinion estimates from each of the six MRP models. These models use the opinion estimates directly in the model, without accounting for any measurement error that might be generated in the MRP process. Most models uncover the expected, positive association, but coefficient estimates vary widely. Local-level transition estimates produce a coefficient of only 0.383, while year random intercepts estimates produce a coefficient of 5.401. Only no-pooling and moving average models—the ones that do not take time into consideration—are statistically significant at $p < 0.05$. The year-intercepts model produces the largest relationship of all—more than fourteen times larger than that of the local-level transition model—but the standard error is also very large, perhaps reflecting the uncertainty with which the year-intercepts MRP models tended to estimate opinion in the main text simulations. The linear trend model is the odd one out, producing coefficients that are in the opposite direction than what many scholars have previously found.

Model efficiency is important, as it indicates the model is generating more precise estimates that suffer from less measurement error. Several authors have emphasized the importance of incorporating this measurement error into downstream analyses when using learned proxies as explanatory or dependent variables (Knox et al., 2022; Tai et al., forthcoming). To assess whether the efficiency of each dynamic MRP model carries consequences for this analysis of policy responsiveness, I use the method of composition to propagate the measurement error from the MRP estimates through to the coefficient estimates in the two-way fixed effects models. On the method of composition, see Caughey and Warshaw (2018), Tanner (1993), and Treier and Jackman (2008).

As one would expect, incorporating measurement error causes coefficients to decrease in magnitude. Deploying the method of composition on these data attenuates effect estimates by about forty percent, similar to what Caughey and Warshaw (2018) find. Both coefficients that were previously statistically significant are no longer, although the coefficient deriving from the linear trend data gains significance, owing to the smaller standard error. In sum, the choice of MRP model matters for the inferences one can make. In applications where the point estimate of the effect is important—such as in policy analysis—different MRP models may return very different substantive conclusions. Even in cases where the researcher is comfortable making a more heuristic assessment of effect direction and statistical significance, using a model that makes inefficient estimates may result in different conclusions.

E Replication: Racial Resentment

All CES time series and simulations I analyze in the main text and in previous sections of the SI use binary dependent variables. This is a common practice in survey research and especially in MRP, where authors almost exclusively formulate their multilevel models with a logit link function. However, small-area estimation comprises many more applications than can be expressed as Bernoulli draws, and survey research commonly produces data more accurately described as ordinal or even—in some cases—continuous. The models explicated in the main text can be applied to any data format that can be modeled with a link function in the exponential family, making them applicable to a wide variety of data structures.

To demonstrate, I replicate the analysis in Smith et al. (2020), who use a model with random intercepts by year to estimate racial resentment in each US state over nine years. To create individual-level estimates of racial resentment, they use a four-item racial resentment battery that was repeatedly asked on the American National Election Studies. Each item is asked on a five-point scale, meaning the total additive index varies on the interval $[4,20]$, with higher values indicating greater racial resentment. To model these data, I replace the inverse logit link function in main text equation (1) with a normal distribution, making the first-stage model an ordinary least squares (OLS) regression. I also need to add a prior on the variance parameter for that normal distribution, σ^2, which I do by specifying $\sigma^2 \sim \mathrm{N}^+(0,1)$. Following Smith et al., I use gender, race, age, and education as individual-level covariates and ideology as a state-level covariate.

Figure S11 shows the over-time results using each of the six dynamic MRP models as well as the original estimates from Smith et al. in eight exemplar states, chosen by Smith et al. to

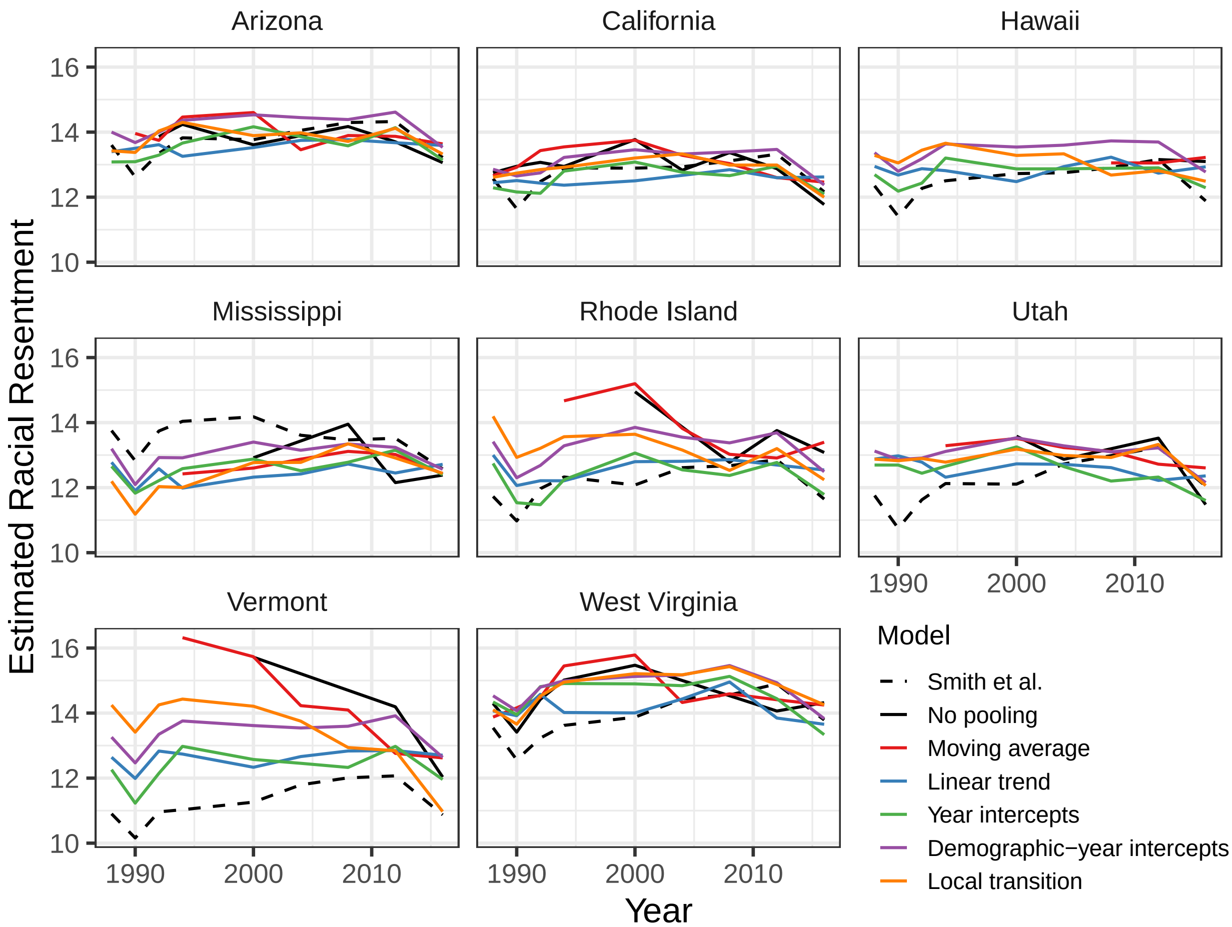

Figure S11: Estimates of State-Level Racial Resentment over Time. Replicates Figure 2 in Smith et al. (2020).

represent a wide range of contexts.[1] The dashed lines indicate the original estimates from Smith et al., while the solid lines show results from each of the six models. Because some states were not administered the racial resentment battery until later in the time series, the no-pooling and moving average model estimates are left-truncated in these states. All other models take the full time series into consideration and therefore can produce estimates even in years where a particular state was not surveyed.

Like the same-sex marriage application above, this replication exercise demonstrates that model selection can matter. Trends estimated by the six dynamic MRP models in this paper often mirror the original estimates closely, especially in states like Arizona and California. In others, however, estimated racial resentment differs quite dramatically, both in terms of trend direction and absolute

[1]This analysis replicates Figure 2 in Smith et al. (2020).

level.[2] Smith et al. point out high racial resentment in Mississippi, but all six dynamic MRP models suggest slightly lower values from 1988 through 2004. The opposite is true in Vermont, which Smith et al. estimate as having very low levels of racial resentment. By contrast, all dynamic MRP models uncover higher index values in that state, and the no-pooling and moving average models especially suggest—likely inaccurately, given the results in the main text—that Vermont is among the most racially resentful among the states examined. In sum, dynamic MRP models can be applied to a wide variety of data types, and careful model selection remains important regardless of the link function used to connect data to covariates.

References

Airoldi, E. M., Blei, D. M., Fienberg, S. E., & Xing, E. P. (2008). Mixed Membership Stochastic Blockmodels. *Journal of Machine Learning Research*, *9*, 1981–2014.

Bisbee, J. (2019). BARP: Improving Mister P Using Bayesian Additive Regression Trees. *American Political Science Review*, *113*(4), 1060–1065.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Buttice, M. K., & Highton, B. (2013). How Does Multilevel Regression and Poststratification Perform with Conventional National Surveys? *Political Analysis*, *21*(4), 449–467.

Caughey, D., & Warshaw, C. (2018). Policy Preferences and Policy Change: Dynamic Responsiveness in the American States, 1936–2014. *American Political Science Review*, *112*(2), 249–266.

Caughey, D., & Warshaw, C. (2019). Public Opinion in Subnational Politics. *The Journal of Politics*, *81*(1), 352–363.

Gelman, A. (2013). *Last Word on Mister P (For Now)*. https://statmodeling.stat.columbia.edu/2013/10/15/last-word-on-mister-p-for-now/

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). CRC Press.

Grimmer, J. (2011). An Introduction to Bayesian Inference via Variational Approximations. *Political Analysis*, *19*(1), 32–47.

Grimmer, J. (2013). Appropriators not Position Takers: The Distorting Effects of Electoral Incentives on Congressional Representation. *American Journal of Political Science*, *57*(3), 624–642.

[2] The original analysis contains a coding error which causes the model to discard information about some racial categories in some years. I have left the original estimates unaltered, but corrected this error when preparing the data for the six dynamic MRP models in this paper. Early years in the time series may therefore appear to depart from my estimates more than they otherwise would.

Imai, K., Lo, J., & Olmsted, J. (2016). Fast Estimation of Ideal Points with Massive Data. *American Political Science Review*, *110*(4), 631–656.

Kastellec, J. P., Lax, J. R., & Phillips, J. H. (2010). Public Opinion and Senate Confirmation of Supreme Court Nominees. *The Journal of Politics*, *72*(3), 767–784.

Knox, D., Lucas, C., & Cho, W. K. T. (2022). Testing Causal Theories with Learned Proxies. *Annual Review of Political Science*, *25*, 419–441.

Lax, J. R., & Phillips, J. H. (2009a). Gay Rights in the States: Public Opinion and Policy Responsiveness. *American Political Science Review*, *103*(3), 367–386.

Lax, J. R., & Phillips, J. H. (2009b). How Should We Estimate Public Opinion in the States? *American Journal of Political Science*, *53*(1), 107–121.

Lewis, D. C., & Jacobsmeier, M. L. (2017). Evaluating Policy Representation with Dynamic MRP Estimates: Direct Democracy and Same-Sex Relationship Policies in the United States. *State Politics & Policy Quarterly*, *17*(4), 441–464.

Pacheco, J. (2013). The Thermostatic Model of Responsiveness in the American States. *State Politics & Policy Quarterly*, *13*(3), 306–332.

Smith, C. W., Kreitzer, R. J., & Suo, F. (2020). The Dynamics of Racial Resentment across the 50 US States. *Perspectives on Politics*, *18*(2), 527–538.

Tai, Y., Hu, Y., & Solt, F. (forthcoming). Democracy, Public Support, and Measurement Uncertainty. *American Political Science Review*.

Tanner, M. A. (1993). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions* (2nd ed.). Springer-Verlag.

Tausanovitch, C., & Warshaw, C. (2014). Representation in Municipal Government. *American Political Science Review*, *108*(3), 605–641.

Treier, S., & Jackman, S. (2008). Democracy as a Latent Variable. *American Journal of Political Science*, *52*(1), 201–217.

Warshaw, C. (2016). The Application of Big Data in Surveys to the Study of Elections, Public Opinion, and Representation. In R. M. Alvarez (Ed.), *Computational Social Science: Discovery and Prediction* (pp. 27–50). Cambridge University Press.