Methods Appendix for Analysis of Medicaid Expansion Preference Using Multilevel Regression and Poststratification

January 29, 2018

1 The Regression Model

In 2014 the CCES asked respondents if they were in favor of their state rejecting medicaid expansion. Respondents are divided into cells based on demographic characteristics and the state that they live in. Letting i index cells, the distribution of yes responses (meaning the respondent thinks their state should reject expansion) for the respondents in cell i is modeled with a binomial distribution

$$y_i \sim Binomial\left(n_i, p_i\right) \tag{1}$$

Where n_i is the number of respondents in cell i. To infer the probabilities for each cell p_i, we use the logistic inverse link function to map from probability to a linear predictor η_i

$$p_i = \frac{1}{1 + \exp \eta_i} = logistic\left(\eta_i\right) \tag{2}$$

The linear predictors are then modeled with the following regression equation

$$\eta_i = \alpha^0 + \alpha^{female}_{g(i)} + \alpha^{college}_{e(i)} + \alpha^{race}_{r(i)} + \alpha^{age}_{a(i)} + \alpha^{state}_{s(i)} + \alpha^{race:college}_{r(i),e(i)} + \alpha^{race:gender}_{r(i),e(i)} + \alpha^{race:state}_{r(i),s(i)} \tag{3}$$

The model includes:

- binary gender, indexed by g(i)
- binary education, 1 for college and 0 for non college, indexed by e(i)
- 5 race categories white, black, white Hispanic, Latino, and other race, indexed by r(i)
- 5 age categories: 18-29, 30-44, 45-54, 55-64, and 65+, indexed by a(i)
- 51 state categories (DC is counted as a state) indexed by s(i)
- 10 categories for race - education interaction
- 10 categories for race - gender interaction
- 255 categories for race - state interaction

There are a total of 5100 cells representing all combinations of these categories. Since our regression model does not account for every variable that may affect selection probability, we use the design effect correction method of [1] to adjust the size and number of yes responses in each cell.

To facilitate partial pooling between cells, we use a hierarchical regression for the state intercepts which are drawn from a normal distribution

$$\alpha^{state}_s \sim Normal\left(\beta^{region}_{p(s)} + \beta^{Trump} x^{Trump}_s + \beta^{Evang} x^{Evang}_s + \beta^{Income} x^{Income}_s + \beta^{Expansion} x^{Expansion}_s, \sigma^2_{state}\right) \tag{4}$$

We use the eight regions from the Bureau of Economic Analysis. Four state level predictors are also included, 2016 Trump vote share x^{Trump}_s as a measure of state partisanship, the share of the population that is Evangelical x^{Evang}_s, state median household income x^{Income}_s, and a binary indicator for whether or not the state expanded medicaid $x_s^{Expansion}$. Similarly, hierarchical normal priors are placed on α_r^{race}, α_a^{age}, $\alpha_{r,e}^{race:college}$, $\alpha_{r,g}^{race:gender}$, $\alpha_{r,s}^{race:state}$ i.e.

$$\alpha^k \sim Normal\left(0, \sigma_k^2\right) \tag{5}$$

Following Si et al [2], we place a structured, informative prior on the group scale parameters σ_k to stabilize estimates of interaction parameters. Each scale parameter is taken as the product of a global scale σ and a local scale λ_k so that

$$\sigma_k = \lambda_k \sigma \tag{6}$$

For main effects, the local scales are drawn from a half normal distribution

$$\lambda_k^{(1)} \sim Normal_+ (0, 1) \tag{7}$$

For 2nd order interactions, the local scales are the product of a relative magnitude parameter $\delta^{(2)} \sim Normal_+ (0, 1)$ and the local scales of the main effect. The local scale for race - state interactions becomes

$$\lambda_{race:state}^{(2)} = \delta^{(2)} \lambda_{race}^{(1)} \lambda_{state}^{(1)} \tag{8}$$

with other interactions defined similarly. To complete the specification of the structured prior, the prior for the global scale is a half Cauchy $\sigma \sim Cauchy_+ (0, 1)$. All hierarchical normal models use non-centered parameterizations for more efficient sampling. The inference is done with the No U-Turn Sampler, a variant of the Hamiltonian Monte Carlo method, which has been implemented in the open source package PyMC3.

2 Poststratification

Poststratified estimates for each of the 5100 cells are obtained by simulation as follows. For each realization, we draw binomial probabilities p_i^{post} from the posterior distribution of the regression model, and then draw simulated responses for each cell from

$$y_i^{post} \sim Binomial\left(n_i^{post}, p_i^{post}\right) \tag{9}$$

The poststratification counts for each cell n_i^{post} are constructed from the census microdata from the 5 year estimates from the 2015 American Community Survey, and only includes voting age citizens. After drawing y_i^{post} we can aggregate results from different cells together to get state and national level estimates for each response.

We find that our national level predicted responses are consistent with the estimates produced using the survey weights from the CCES. Below we have plotted the MRP predicted distribution and mean along with the estimates using the CCES weights. The two methodologies lead to comparable national results, and are different from one another by about half a percentage point.

The MRP simulations can also be used to asses congruence in the states that refused expansion. For each simulation, we count the number of states who refused expansion where support for refusing expansion was greater than support for accepting it. The resulting distribution is shown in figure 2, where we see that we have congruence in no more than five states in any simulation, but it is much more likely that only 2 or 3 states are congruent.

Wyoming and Oklahoma are the rejection states with the highest likelihood of congruence, where we estimate a bare majority of support for rejecting expansion.

References

[1] Devin Caughey and Christopher Warshaw. Dynamic estimation of latent opinion using hierarchical group-level irt model. *Political Analysis*, 23:197–211, 2015.

[2] Yajuan Si, Rob Trangucci, Jonah Sol Gabry, and Andrew Gelman. Bayesian hierarchical weighting adjustment and survey inference. *Arxiv.org*, 2017.

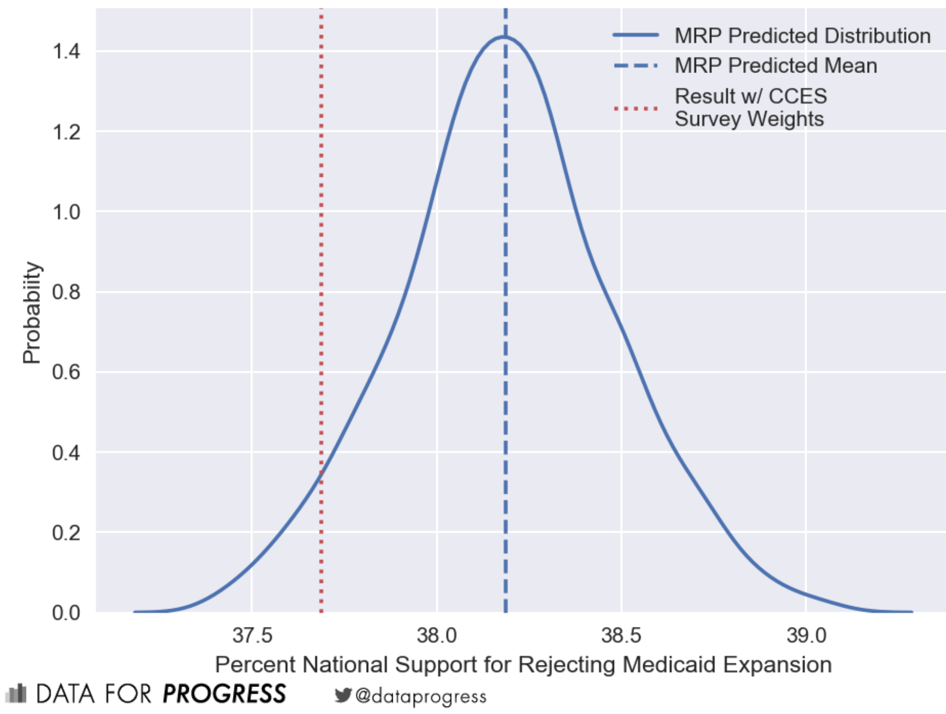

Figure 1: Comparison of National Level Estimates Using MRP and the CCES Survey Weights

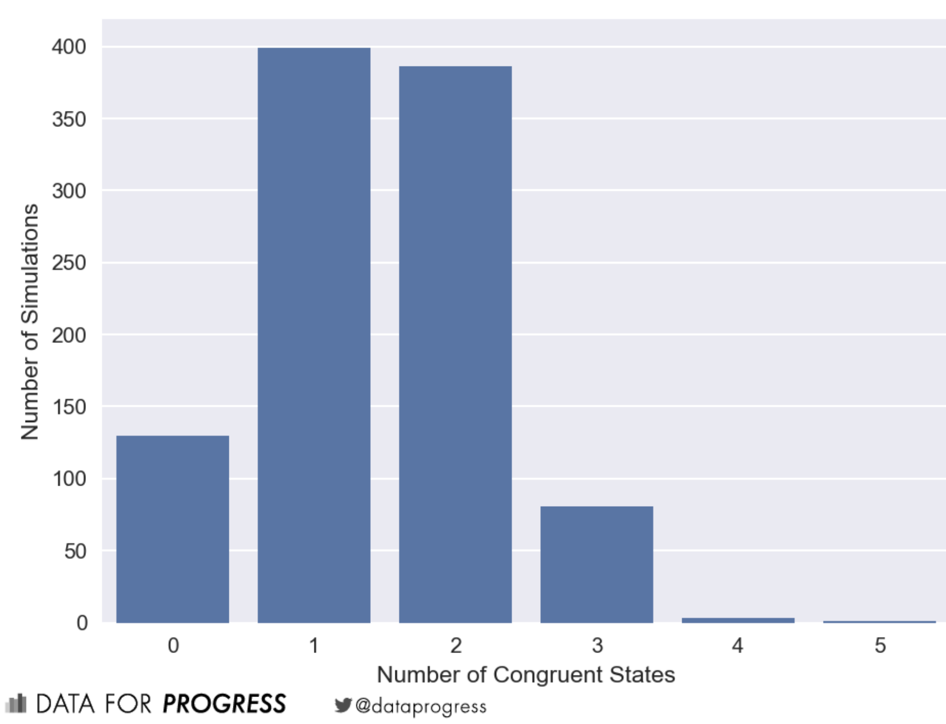

Figure 2: Distribution of the Number of Congruent Rejection States in the MRP Simulations

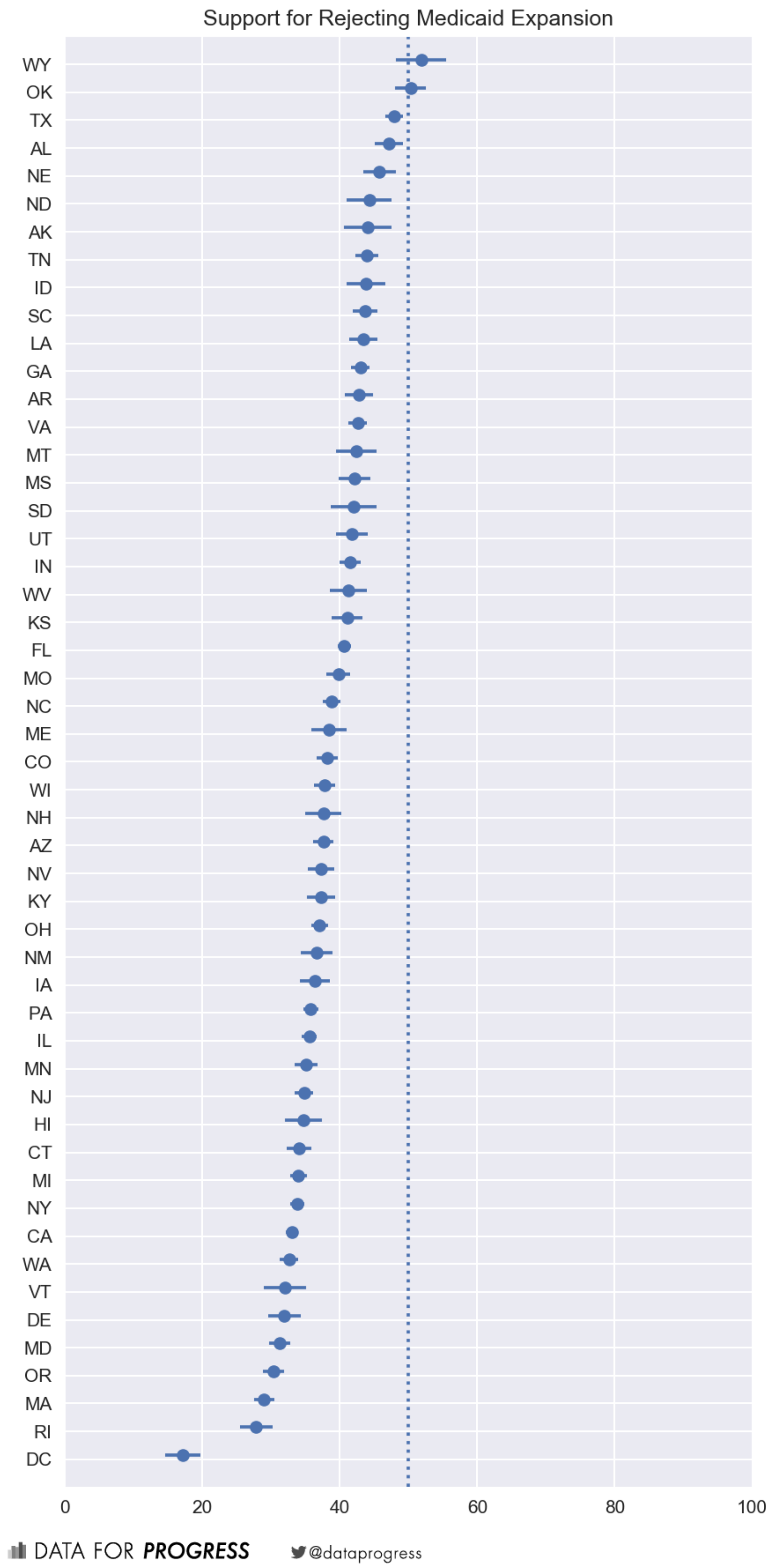

Figure 3: MRP Results by State