Supplemental Material (not copyedited or formatted) for: Devin Caughey, Michael C. Dougal, Eric Schickler. 2020. "Policy and Performance in the New Deal Realignment: Evidence from Old Data and New Methods." The Journal of Politics 82(2). DOI: https://doi.org/10.1086/707305.

Supplementary Information for “Policy and Performance in the New Deal Realignment: Evidence from Old Data and New Methods”

Supplemental Material (not copyedited or formatted) for: Devin Caughey, Michael C. Dougal, Eric Schickler. 2020. "Policy and Performance in the New Deal Realignment: Evidence from Old Data and New Methods." The Journal of Politics 82(2). DOI: https://doi.org/10.1086/707305.

Contents

Supplemental Material (not copyedited or formatted) for: Devin Caughey, Michael C. Dougal, Eric Schickler. 2020. "Policy and Performance in the New Deal Realignment: Evidence from Old Data and New Methods." The Journal of Politics 82(2). DOI: https://doi.org/10.1086/707305.

A Group-Level IRT Model

We estimate citizen policy liberalism using the statistical framework of item-response theory. In the two-parameter IRT model, the relationship between responses to question q and the unobserved trait θ_i is governed by the question's threshold κ_{qt}, which captures the base level of support for the question, and its dispersion σ_q, which represents question-specific measurement error. Under this model, respondent i's probability of selecting the liberal response to question q is:

$$\pi_{iq} = \Phi\left(\frac{\theta_i - \kappa_{qt}}{\sigma_q}\right), \tag{2}$$

where the normal CDF Φ maps $(\theta_i - \kappa_{qt})/\sigma_q$ to the $(0, 1)$ interval.[1] The model assumes that greater liberalism (i.e., higher values of θ_i) increases respondents' probability of answering liberally. The strength of this relationship is inversely proportional to σ_q, and the threshold for a liberal response is governed by κ_{qt}.

Accurate estimation of θ_i requires data from many respondents, each of whom answers many items. In the period we consider, however, very few surveys included more than a handful of policy questions, and those questions that were included were rarely asked in consistent form across many years. The fact that each respondent answers no more than a few questions (sometimes only one) prevents us from using an IRT model to estimate the liberalism of individual respondents. Our ultimate interest, however, is not individuals but rather the *average* liberalism of the eligible electorate in each state-year. Fortunately, it is possible to make inferences about the average level of θ_i in each group even when individual-level estimation is not feasible.

Following Caughey and Warshaw (2015), we do this by treating individual citizens as having been randomly sampled from a given subpopulation g defined by demographic and geographic characteristics (e.g., white farmers in Kentucky). As-

[1] A common alternative way of writing the model in Equation (2) is $\Pr(y_{iq} = 1) = \Phi(\beta_q \theta_i - \alpha_q)$, where $\beta_q = 1/\sigma_q$ and $\alpha_q = \kappa_{qt} \times \beta_q$.

suming that $\theta_{i[g]}$ is distributed normally with mean $\bar{\theta}_g$ and standard deviation σ_θ, we can re-write the individual-level IRT model at the group level as

$$p_{gj} = \Phi\left(\frac{\bar{\theta}_g - \kappa_j}{\sqrt{\sigma_j^2 + \sigma_\theta^2}}\right), \tag{3}$$

where p_{gj} is the probability that a randomly sampled citizen from group g will give a liberal answer to item j (Mislevy 1983). We then model group g's total number of liberal answers to item j as

$$s_{gj} \sim \text{Binomial}(n_{gj},\ p_{gj}), \tag{4}$$

where n_{gq} is group g's total number of non-missing responses to question j and s_{gj} is the number of those responses that are liberal.[2] The estimates of $\bar{\theta}_g$ may be of interest in themselves, or they can be poststratified into estimates of, for example, average liberalism in each state (cf. Park, Gelman, and Bafumi 2004).

Even with our large-scale dataset of survey respondents, many group cells are likely to be small or empty in a given year. To address this sparseness, we use a dynamic linear model to smooth the estimated group means across both time and states. Letting ξ_t be an intercept common to all groups and $\boldsymbol{x}_{g\boldsymbol{\cdot}}$ a vector of hierarchical predictors (*Race*, *Urban*, and *State*), we model the group means in each year as

$$\bar{\theta}_{gt} \sim N(\delta_t \bar{\theta}_{g,t-1} + \xi_t + \mathbf{x}'_{g\cdot}\gamma_t,\ \sigma^2_{\bar{\theta}t}), \tag{5}$$

That is, the prior expected value for $\bar{\theta}_{gt}$ is a weighted combination of its lagged value

[2]Following Ghitza and Gelman (2013), we adjust the raw values of s_{gq} and n_{gq} to account for survey weights and for respondents who answer multiple questions. The latter is particularly important in this application because of the way that we deal with ordinal questions, which is to break each such question into a set of dichotomous questions, each of which indicates whether the response is above a given response level. For example, a question with three ordinal response choices, (1) "disagree", (2) "neutral", and (3) "agree," would be converted into two dichotomous variables respectively indicating whether the response is above "disagree" and above "neutral."

and predictions based on demographically similar groups, with the variance of the prior determined by $\sigma^2_{\bar{\theta}}$. If there are no survey responses from group g in year t, (5) acts as an imputation model for the missing data.

We estimate this model on a dataset of 273 different polls fielded between December 1936 and July 1952. We analyze responses to 258 distinct policy questions related to social welfare, economic regulation, and other New Deal issues. We coded the responses to each question according to their relative liberalism and dichotomized questions with more than two response options at an appropriate midpoint. We dropped questions with a presidential or candidate cue as well as ones that explicitly reference the policy status quo.

Over 340,000 unique survey respondents are represented in this dataset, with the sample size in each year ranging from a minimum of 2,301 in 1951 to 56,411 in 1937. We classify respondents into one of 288 demographic groups, defined by the cross-classification of *State*, *Black/Non-Black*, and *Urban/Rural/Farm*. Within groups, we weight respondents to match the marginal population distributions of gender, phone ownership, and professional status in that group. We estimate the model in the Bayesian simulation program Stan, as implemented by the R package `dgo` (Dunham, Caughey, and Warshaw 2016). We ran four chains of the sampler, each with 2,000 iterations, discarding the first 1,500 and thinning every 2. The estimation produced 1,000 samples from the posterior distribution of average liberalism in 288 demographic groups in each of 17 years (1936–52). To estimate average liberalism in the voting-eligible public, we weighted each group estimate in proportion to the group's proportion in the over-21 population (for non-Southern states) or the over-21 white population (for Southern states). We based the poststratification on population targets estimated by Caughey et al. (2016). We conducted the poststratification separately for each Monte Carlo sample, resulting in in 1,000 draws from the distribution of liberalism estimates.

Supplemental Material (not copyedited or formatted) for: Devin Caughey, Michael C. Dougal, Eric Schickler. 2020. "Policy and Performance in the New Deal Realignment: Evidence from Old Data and New Methods." The Journal of Politics 82(2). DOI: https://doi.org/10.1086/707305.

B Propagating Uncertainty with the Method of Composition

The key independent variable in this study—mass support for New Deal liberalism—is a latent quantity whose values must be inferred rather than directly observed. As such, it is subject to measurement error. Even if it is independent and mean-zero, measurement error can bias point estimates and standard errors. However, since we can estimate the values' measurement error from their Monte Carlo sampling distributions, we can account for measurement error using a technique known as the "method of composition" (MOC) (Tanner 1996, 52; Treier and Jackman 2008, 215–6; Kastellec et al. 2015, 791–2; Caughey and Warshaw 2017, Supplementary Appendix D).

The idea behind MOC is that we wish to estimate the marginal distribution (i.e., the posterior or, given flat priors, likelihood) of a parameter vector $\boldsymbol{\beta}$, integrating over the measurement error in variables $\boldsymbol{X}$:

$$p(\boldsymbol{\beta}|\boldsymbol{W}) = \int_X p(\boldsymbol{\beta}|\boldsymbol{W}, \boldsymbol{X}) p(\boldsymbol{X}|\boldsymbol{Z}) d\boldsymbol{X}, \tag{6}$$

where $\boldsymbol{W}$ is a matrix of variables measured without error and $\boldsymbol{X}$ is a matrix of variables estimated with error conditional on data $\boldsymbol{Z}$ (and a measurement model). That is, we wish to integrate the joint density of $\boldsymbol{\beta}$ and $\boldsymbol{X}$ over the distribution of $\boldsymbol{X}$. As Treier and Jackman (2008, 215) explain, this can be done via the following iterative procedure: At each iteration s,

1. Sample $\boldsymbol{X}^{(s)}$ from the distribution $p(\boldsymbol{X}|\boldsymbol{Z})$.

2. Sample $\tilde{\boldsymbol{\beta}}^{(s)}$ from $p(\boldsymbol{\beta}|\boldsymbol{W}, \boldsymbol{X}^{(s)})$ in two steps:

   (a) Conditional on $\boldsymbol{W}$ and $\boldsymbol{X}^{(s)}$, estimate the parameter vector $\hat{\boldsymbol{\beta}}^{(s)}$ and its variance-covariance matrix $\hat{\boldsymbol{V}}^{(s)}$.

Supplemental Material (not copyedited or formatted) for: Devin Caughey, Michael C. Dougal, Eric Schickler. 2020. "Policy and Performance in the New Deal Realignment: Evidence from Old Data and New Methods." The Journal of Politics 82(2). DOI: https://doi.org/10.1086/707305.

```r
run_moc <- function (data, its, model, vcov = plmHC, prog.int = 100,
                     rsq = TRUE) {
  ## Method of Composition (as described by Treier and Jackman)
  tildeB <-
    as.data.frame(matrix(nrow = length(its), ncol = length(coef(model))))
  R2 <- data.frame(rsq = rep(NA, length(its)), adjrsq = rep(NA, length(its)))
  for (s in seq_along(its)) {
    if (!s %% prog.int) print(s)
    it <- its[s]
    ## (1) Sample from p(x)
    data_s <- subset(data, It == it)
    ## (2) Sample from p(B|x,y):
    ##     (a) Estimate B_s and Cov(B_s) conditional on x_s.
    mod_s <- update(model, data = data_s) # Re-estimate model on sample
    hatB_s <- coef(mod_s)
    hatV_s <- vcov(mod_s)
    ##     (b) Sample \tilde{B_s} from MV(\hat{B_s}, \hat{Cov(B_s)}).
    tildeB[s, ] <- MASS::mvrnorm(n = 1, mu = hatB_s, Sigma = hatV_s)
    if (rsq) {
      R2$rsq[s] <- summary(mod_s)$r.squared["rsq"]
      R2$adjrsq[s] <- summary(mod_s)$r.squared["adjrsq"]
    }
  }
  names(tildeB) <- names(coef(model))
  if (rsq) {
    tildeB$rsq <- R2$rsq
    tildeB$adjrsq <- R2$adjrsq
  }
  return(tildeB)
}
```

Listing 1: R Function for the Method of Composition

(b) Draw one sample $\tilde{\boldsymbol{\beta}}^{(s)}$ from a multivariate normal distribution with mean vector $\hat{\boldsymbol{\beta}}^{(s)}$ and variance-covariance matrix $\hat{\boldsymbol{V}}^{(s)}$.

Each draw $\tilde{\boldsymbol{\beta}}^{(s)}$ is a sample from (the normal approximation of) the marginal distribution $p(\boldsymbol{\beta}|\boldsymbol{W})$, with the iterative algorithm performing the intergration over $p(\boldsymbol{X}|\boldsymbol{Z})$.[3] The R function we used to implement the MOC algorithm is reproduced in Figure 1.

Using this approach, we accounted for measurement error in mass liberalism as well as party identification. Independently for each measure, we drew 1,000 samples from the joint posterior distribution of state-year values. We combined each two-variable sample with a copy of the other variables in our dataset, which were presumed to

[3]Treier and Jackman (2008, 216) note that this procedure relies on two conditional independence assumptions. The first is that the data used to estimate the latent variables $\boldsymbol{X}$ do not supply information about the model parameters $\boldsymbol{\beta}$ except through $\boldsymbol{X}$. That is, $p(\boldsymbol{\beta}|\boldsymbol{W},\boldsymbol{X}) = p(\boldsymbol{\beta}|\boldsymbol{W},\boldsymbol{X},\boldsymbol{Z})$. The second assumption is that inferences about each latent variable $\boldsymbol{x}_j$ are not informed by the other latent variables $\boldsymbol{X}_{-j}$ or by the manifest variables $\boldsymbol{W}$. That is, $p(\boldsymbol{x}_j|\boldsymbol{Z}) = p(\boldsymbol{x}_j|\boldsymbol{Z},\boldsymbol{X}_{-j},\boldsymbol{W})$. Together, these assumptions separate the estimation of the measurement model from the estimation of the structural model.

be measured without error. Then, following the algorithm above, we re-estimated each model on the 1,000 versions of the dataset and drew 1,000 samples of the model parameters. We then used these samples to calculate point estimates and standard errors, on which we based our inferences.

C Congressional Vote

In this section we replicate the regression analyses in our paper with congressional instead of presidential vote as the dependent variable.

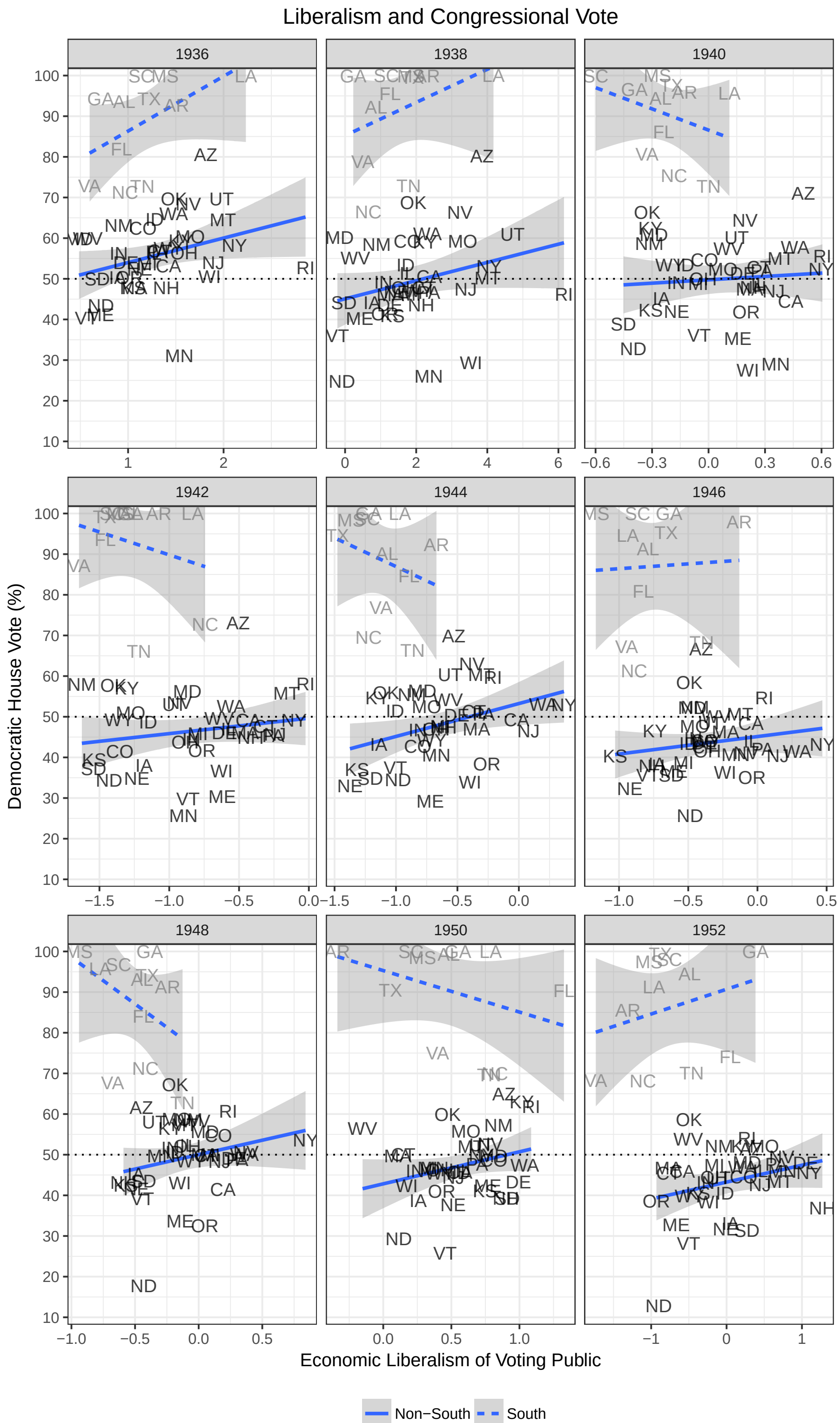

Figure A1: Economic Liberalism and Democratic Congressional Vote, 1936–52.

	DV: Democratic Percent of Two-Party Congressional Vote								
	1936	1938	1940	1942	1944	1946	1948	1950	1952
Income Growth$_t$	**.19**	.14	.26	−.01	**−.19**	−.21	**.25**	−.17	**.43**
	(.09)	(.15)	(.26)	(.06)	(.11)	(.18)	(.12)	(.18)	(.18)
Mass Liberalism$_t$	−.93	−.65	.30	2.61	1.82	−2.00	**5.38**	1.08	1.08
	(1.73)	(.67)	(2.56)	(2.03)	(1.48)	(2.42)	(3.02)	(2.63)	(1.21)
Dem Cong$_{t-2}$	**.99**	**1.10**	**.84**	**.85**	**.88**	**.68**	**1.00**	**.74**	**.82**
	(.13)	(.09)	(.05)	(.09)	(.09)	(.11)	(.16)	(.10)	(.09)
South	16.38	21.32	14.99	−20.11	**−16.93**	**−19.35**	7.43	10.67	4.29
	(13.02)	(14.70)	(9.53)	(15.80)	(9.98)	(10.52)	(12.93)	(9.25)	(14.30)
Inc Growth$_t$ × South	−.29	.60	−.16	−.11	**.59**	−.09	**−.52**	−.40	**−4.14**
	(.39)	(.57)	(.42)	(.33)	(.32)	(.38)	(.23)	(.44)	(2.14)
Mass Lib$_t$ × South	4.36	−.48	−3.18	−5.74	−7.41	**9.01**	**−13.12**	−.91	−1.14
	(4.27)	(1.91)	(8.15)	(5.28)	(5.91)	(5.37)	(6.89)	(3.85)	(4.17)
Dem Cong$_{t-2}$ × South	−.22	−.12	−.10	**.34**	.04	**.45**	−.14	.09	.31
	(.18)	(.16)	(.10)	(.17)	(.13)	(.15)	(.20)	(.12)	(.19)
Constant	−.16	**−9.80**	**6.64**	6.53	**9.30**	**10.62**	4.02	**11.61**	2.71
	(7.42)	(5.15)	(3.60)	(4.79)	(4.98)	(5.58)	(7.37)	(6.36)	(4.67)
Observations	48	48	48	48	48	48	48	48	48
R^2	.92	.95	.97	.97	.97	.96	.95	.96	.96

Table A1: Predictors of congressional election results, by year. Bold coefficients are significant at the 10% level. Standard errors are robust to heteroskedasticity. Estimates are corrected for measurement error in mass liberalism.

Supplemental Material (not copyedited or formatted) for: Devin Caughey, Michael C. Dougal, Eric Schickler. 2020. "Policy and Performance in the New Deal Realignment: Evidence from Old Data and New Methods." The Journal of Politics 82(2). DOI: https://doi.org/10.1086/707305.

	DV: Democratic Percent of Two-Party Congressional Vote					
	1936–52	1936–52	1938–52	1938–52	1938–52	1938–52
	(1)	(2)	(3)	(4)	(5)	(6)
Income Growth$_t$	**.17**	.04	.004	.01	.003	.003
	(.09)	(.04)	(.05)	(.05)	(.05)	(.05)
Mass Lib$_t$	**2.82**	.40				
	(1.38)	(.51)				
Mass Lib$_{t-1}$			.42		.37	.16
			(.50)		(.57)	(.65)
Mass Lib$_{t-2}$				.41		.25
				(.47)		(.55)
Dem PID$_{t-1}$					**.13**	**.13**
					(.06)	(.06)
Dem Cong$_{t-2}$		**.87**	**.86**	**.86**	**.76**	**.76**
		(.02)	(.02)	(.02)	(.06)	(.06)
South	**31.69**	−1.67	**6.95**	**8.08**	−9.54	−8.66
	(5.86)	(3.92)	(3.69)	(3.60)	(10.50)	(10.08)
Inc$_t$ × South	.15	−.10	−.02	−.05	−.02	−.05
	(.32)	(.15)	(.15)	(.15)	(.14)	(.14)
Lib$_t$ × South	−1.71	−.17				
	(3.58)	(1.03)				
Lib$_{t-1}$ × South			−.04		.82	1.66
			(1.31)		(1.49)	(1.63)
Lib$_{t-2}$ × South				−1.57		−1.56
				(1.04)		(1.26)
PID$_{t-1}$ × South					.33	.31
					(.22)	(.21)
Cong$_{t-2}$ × South		.05	**.08**	**.08**	−.08	−.07
		(.03)	(.03)	(.03)	(.13)	(.13)
Year × South FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	432	432	384	384	384	384
R^2	.77	.95	.95	.95	.96	.96

Table A2: Predictors of congressional election results, pooling years. Bold coefficients are significant at the 10% level. Standard errors are robust to heteroskedasticity and serial correlation. Estimates are corrected for measurement error in mass liberalism and partisanship.

References for Appendix

Caughey, Devin, Adam J. Berinsky, Sara Chatfield, Erin Hartman, Eric Schickler, and Jasjeet J. Sekhon. 2016. "Population Estimation and Calibration Weighting for Nonresponse and Sampling Bias: An Application to Quota-Sampled Opinion Polls, 1936–52." Unpublished manuscript, June 17. https://www.dropbox.com/s/qnol9sffrb2bci1/CaugheyEtAl-Weights.pdf?dl=0.

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

———. 2017. "Policy Preferences and Policy Change: Dynamic Responsiveness in the American States, 1936–2014." *American Political Science Review.* Pre-published. doi:10.1017/S0003055417000533.

Dunham, James, Devin Caughey, and Christopher Warshaw. 2016. *dgo: Dynamic Estimation of Group-Level Opinion. R package version 0.2.3.* https://jamesdunham.github.io/dgo/.

Ghitza, Yair, and Andrew Gelman. 2013. "Deep Interactions with MRP: Election Turnout and Voting Patterns Among Small Electoral Subgroups." *American Journal of Political Science* 57 (3): 762–776.

Kastellec, Jonathan P., Jeffrey R. Lax, Michael Malecki, and Justin H. Phillips. 2015. "Polarizing the Electoral Connection: Partisan Representation in Supreme Court Confirmation Politics." *Journal of Politics* 77 (3): 787–804.

Mislevy, Robert J. 1983. "Item Response Models for Grouped Data." *Journal of Educational Statistics* 8 (4): 271–288.

Park, David K., Andrew Gelman, and Joseph Bafumi. 2004. “Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls.” *Political Analysis* 12 (4): 375–385.

Tanner, Martin A. 1996. *Tools for Statistical Inference Methods for the Exploration of Posterior Distributions and Likelihood Functions.* 3rd ed. New York: Springer.

Treier, Shawn, and Simon Jackman. 2008. “Democracy as a Latent Variable.” *American Journal of Political Science* 52 (1): 201–217.