Supporting Information for "On Polarization and Partisan Attachments: Greater Consistency but not Greater Strength"

Table of Contents:

A Information for panel switching estimates generating Figure 2

Table A.1. Comparing Rates of American's Major-Party Switching Across Panel Studies

Study Years	Four-year Switching Rate Presidential Support		Party Identification	
(Party ID n = Observed/Imputed)	Observed	Imputed	Observed	Imputed
ANES 1956-60	25.1	26.8	6.77	6.77
(n = 1212/1237)	(1.5)	(1.6)	(.69)	(.68)
ANES 1972-76	25.2#	30.2	7.04	7.03
(n = 1306/1320)	(1.6)	(1.6)	(.74)	(.71)
ANES 1992-96	13.6	19.3	10.71	10.30
(n = 593/1005)	(1.9)	(1.6)	(1.43)	(1.25)
ANES 2000-04	11.2	13.3	8.78	8.87
(n = 835/1807)	(1.9)	(1.1)	(1.58)	(1.04)
ISCAP 2012-16	14.8#	14.2	5.34	8.16
(n = 1068/2606)	(1.1)	(1.3)	(.71)	(.80)
TAPS 2012-16	9.9	14.5	6.81	10.46
(n = 852/1761)	(1.9)	(1.9)	(1.04)	(1.44)
GSS/ANES 2016-20	6.2+	13.3	8.60	8.09
(n = 797/2085)	(1.7)	(1.6)	(1.03)	(1.03)

Sample average percentage point estimates adjusted by post-stratification weights (when available), with standard errors in parentheses. Switching rates for presidential support are based on respondents supporting major-party candidates; these are imputed for respondents supporting third-party candidates. Studies of comparison in the American Panel Survey (TAPS) are between December 2012 and November 2016 for presidential support and between October 2012 and December 2016 for party identification. GSS estimates are for the 2016 sample of respondents only.

- Excludes non-voters, who were not asked which candidate they would have supported.

+ - Spring 2020 recollection of 2016 presidential support.

Imputation approach

A general approach to imputation was followed consistent with Smidt (2017). Chained imputations were created by modeling missing responses within a panel wave based on non-missing responses within that wave or a prior wave. All survey estimates represent the complete set of initial panel wave respondents and using the initial panel's sampling or post-stratification weight if available. The estimates are based on a set of 20 imputations (with 10 burn-in iterations) for the 1956-60 and 1972-76 panels with minimal attrition, and a set of 40 imputations (with 20 burn-in iterations) for the other panels with 50% attrition rates or higher.

B Graded Response Model Illustration and Identification

The graded response IRT modeling framework is advantageous for this analysis for the following reasons:

A. It allows for nested model fit testing within a maximum likelihood framework and a test of the null hypothesis that the measurement performance of these items is constant. It also has a defined set of tools or evaluating these possibilities; see the differential item functioning discussion in Appendix C.

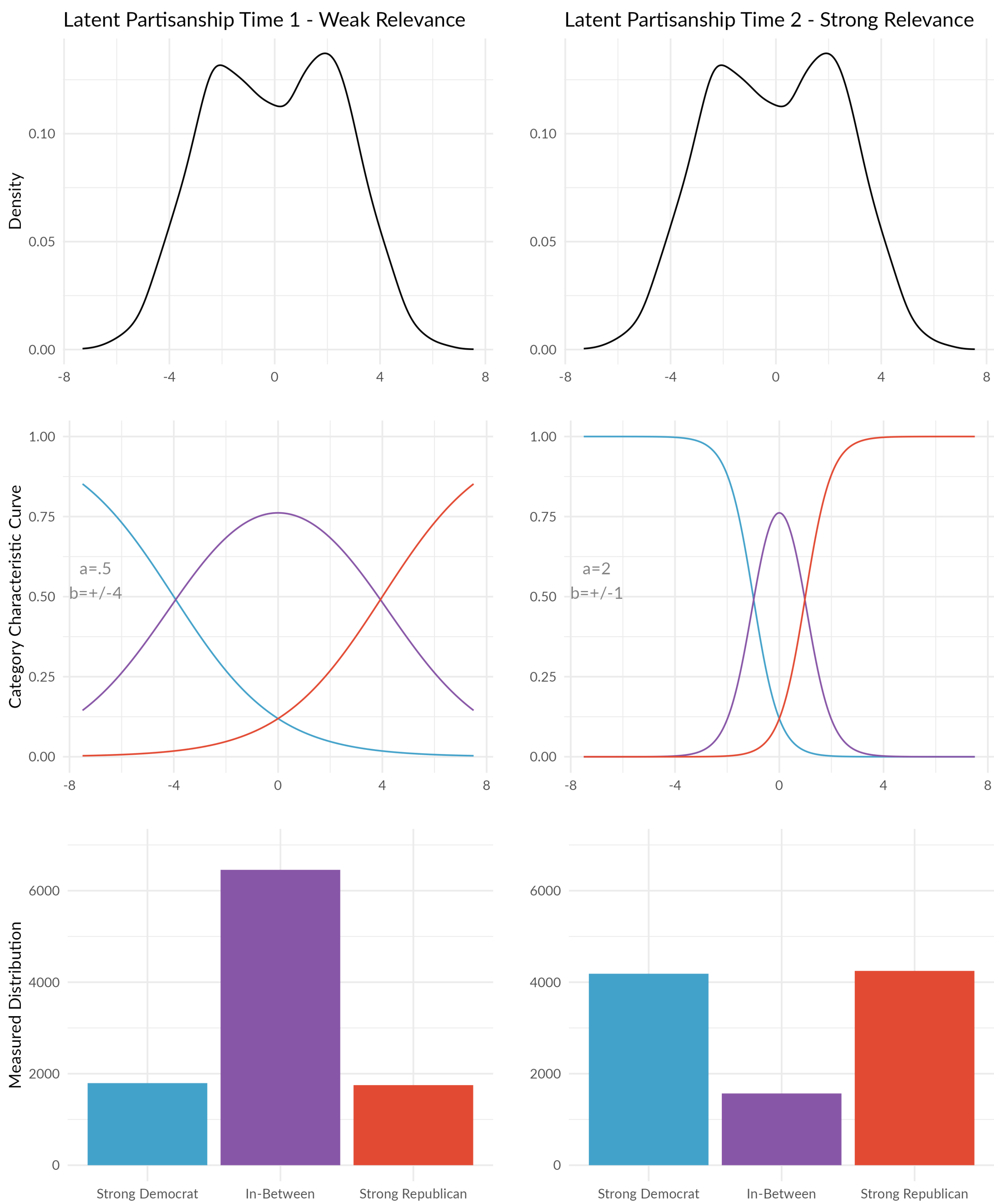
B. These models are available in Stata 16.1 with the appropriate accommodations for survey weights to generate accurate estimates of the United States adult population and accurate standard errors.

C. These models also allow the researcher to specify alternative fixed parameterizations that enable meaningful and comparable metrics across time: the balancing of the difficulty parameters defines teh zero-point as no partisan attachment; and the scale of attachment is made meaningful by anchoring the discrepancy of the party identification item to be 1 across all years, with a constant difficulty parameter.

Illustration
To illustrate the dynamics discussed in the manuscript, Figure B.1 graphs out how a three-category measure can appear more polarized across time simply because it has greater relevance to partisanship. The latent distribution of partisanship is the same across both time periods. In the first time period, the item was a less relevant measure of partisanship, such that only very strong partisans had a high probability of providing a party-conforming response, as shown in middle row's category characteristic curve. But the item becomes more relevant to partisanship by the second time period. Like partisan ideological proximity (Bafumi and Shapiro 2009), this item is a more salient feature to partisanship and, because of greater elite uniformity, it is an easier item for partisans of all levels of sophistication to respond consistently (Zingher 2022). Its discrimination parameter strengthens and its location parameters shrink toward zero. Consequently, as shown in the bottom row of tabulations, individuals with any party leanings have a greater probability of responding like strong partisans by this time, but that simply may be because the item is more relevant to partisanship rather than an indicator of strengthening partisan attachments.
Identification Testing
To prevent group-comparison tests that were biased towards the proposed model, identification tests were performed on the fixed relevance model. Identification requires that the latent partisanship dimension has an identified scaling and an identified zero point. The default assumption for identifying IRT models is either to assume the latent trait has a mean of 0 and a variance of 1, either for the population or one baseline group, or to have a mean of 0 and scale the discrimination parameter to unity for one of the indicators.

The models presented exchange the mean of zero assumption for an assumption that gives the zero-point the substantive interpretation of pure independence. This is done by constraining the discrimination parameter of the partisan categorization item to be unity and by setting the sum of the two cutpoints for the party identification categorization item to be equal to zero. This make the two cutpoints an equal distance on either side of zero, such that a score of latent partisanship at

Figure B.1. The effects that changes in partisanship's relevance can have on observed measures of partisan attitudes



zero represents someone with the highest probability of choosing the party neutral response and an equal probability of choosing a Republican or Democratic favorable response.

This assumption does not change model fit. When estimating the model with the default identification constraint of setting the mean of latent partisanship at zero, the log-likelihood is -331148.6; when estimating the model by fixing the party identification the log-likelihood is the same, -331148.6.

This assumption is sufficient for identification, but additional identification assumptions allow sharper estimates of latent partisanship and a more meaningful substantive interpretation. The measurement model's fit may be slightly worse, but provide more anchors to the scale of latent partisanship to enhance its meaning and comparability across time, which is the focus of Figures 3 - 5.

Three additional sets of constraints were added to the published model. The zero points of pure independence was also assumed to also assumed to be the location where a respondent has the highest probability of net favorability toward either party in the party thermometer scores and the party affect scores and to have an equal and lowest probability of reporting a strong identification toward either party. For each of these three items, the cutpoint parameters differentiating the neutral option between the closest positive or negative valence categories are constrained to sum to zero, such that they are equidistant and reflective about zero. See Figure C.2 for an example.

The loss of fit with these constraints is relatively small. The log-likelihood decreases from -331148.6 in the minimal constraint model to -331171.2 in the published model (a likelihood ration test statistic of 45.2; $\chi^2_{\alpha=.05}$ = 7.815, with 3 degrees of freedom). The BIC changes from 662708.6 to 662720.5. For comparison, an increase of less than 12 in the BIC for a model with 67,495 observations is relatively small compared to the increase in the BIC of 2851.2 between the fixed relevance model in Table 2 to the varying relevance model in Table 2.

Table B.1 replicates the presentation in Table 2 with the minimally-constrained variants of each model. The nature and scale of all substantive conclusions are the same. The worse-fitting fixed relevance model finds latent partisan strength has greatly increased, whereas the varying relevance model estimates latent partisan strength to be essentially the same, if not slightly weaker. The estimates of average latent partisan strength are equal to the first decimal place. Of the small differences, the biggest is the estimate of mean of latent partisanship (μ_θ). By excluding additional non-party identification item constraints, the minimal constraint model finds Americans slightly more Democratic in their leaning (a difference of -.88 vs. -.82 and -.69 vs. -.62 for the respective pre-polarization and polarization eras). Moreover, the estimates for both these estimates are less precise, as the standard errors increase by about 5%.

The main consequence of these additional constraints is to re-scale the location of empirical Bayes partisan strength estimates between 1952-1998 in the ANES, since some of these years have no party thermometer scores. The correlation between the latent estimates of the two models is 0.9998, but Americans are estimated to have a smaller lean in favor of Democrats from 1952-1998 (-1.11 vs. -1.01). The difference in estimates of partisan strength, however, are not nearly as large across 1952-1998 (3.75 vs. 3.77), and is essentially zero when looking at 2000-2020 (3.70 vs. 3.70).

In summary, the additional constraints used for identification purposes enhance the Empirical Bayes latent partisanship score with greater meaning and greater precision without any clear differences that distort or modify any conclusions and inferences presented in the main text.

Support for Unidimensionality

Previous studies justify the study of partisanship as a unidimension polar constract (Green 1988). Figure B.2 presents the estimated Eigenvalue decomposition from a factor analysis of the pre-2000

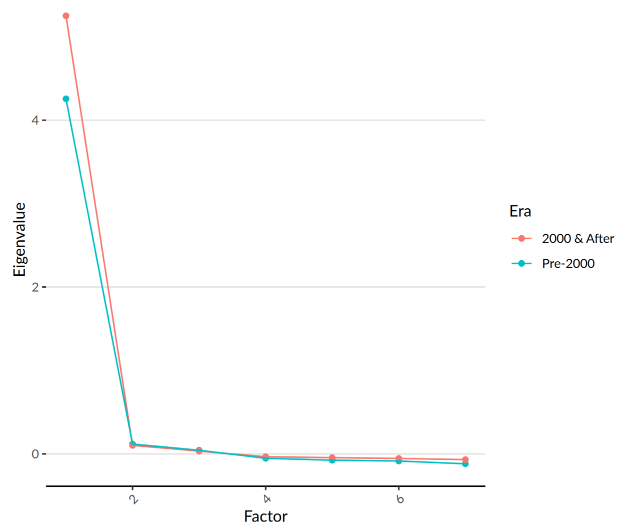
Table B.1. Latent Partisan Strength Estimates and Tests - Minimal Constraints (1952-2020)

	Fixed Relevance Model		Varying Relevance Model	
Latent Partisanship	Pre-2000	2000 & After	Pre-2000	2000 & After
Average Strength				
$\lvert\hat{\theta}_i^{EB} - 0\rvert$	3.20	4.15	3.75	3.70
	(.01)	(.02)	(.01)	(.02)
μ_θ	−0.876	−0.687	−1.116	−0.440
	(.028)	(.048)	(.034)	(.053)
σ_θ^2	16.367	26.143	22.169	21.207
	(.453)	(.781)	(.726)	(.824)
Log-Likelihood	−331148.6		-329530.2^{+}	
n	67,495		67,495	

Weighted estimates by using ANES post-stratification weights and adjusting them to make each study an equivalent sample size so that discrimination and location estimates are averages of the two eras. $\hat{\theta}^{EB}$ indicates Empirical Bayes estimates of latent partisanship. $^{+}$ indicates likelihood-ratio test rejects null hypothesis that the unconstrained, varying relevance model provides no improvement in the log-likelihood relative to the nested, fixed relevance model.

and 2000 & after era basic model measures. It is clear that the first dimension explains nearly all of the shared variance and the unidimensional framework is equally supported across the two eras.

Figure B.2. Screeplot of Basic Model Items Eigenvalue Decomposition across the Two Eras



C Basic and Expanded Model Estimates and Supporting Materials for Table 2

Table C.1 presents the constituent parameter estimates that were used to generate the estimates and findings presented in Table 2 and Figures 3 and 4.

Table C.1. Coefficient Estimates from the Basic Graded Response IRT Model (1952-2020)

Variable	Component	Pre-2000	2000 & After
Latent Partisanship	Mean	-1.01 (.03)	-0.474 (.05)
	Variance	22.64 (.74)	21.15 (.82)
		Discrimination Estimates	
Partisanship	Categorization	1 (Constrained)	
	Strength	0.703 (.014)	0.743 (.019)
Party Affect		0.436 (.009)	0.549 (.013)
Thermometer Differences	Party	0.765 (.016)	1.118 (.029)
	President	0.534 (.012)	0.800 (.021)
Party Support	Pres. Support	0.602 (.016)	0.978 (.034)
	House Vote	0.389 (.009)	0.463 (.014)

Coefficient estimates with standard errors in parentheses. Higher values of latent partisanship indicate stronger Republican attachments, such that Democratic thermometer items with higher scores have a negative relationship.

Differential Item Functioning Tests

IRT models generally refer to parameter instability across groups as differential item functioning (DIF). Uniform DIF refers to the lack of interaction between the groupings and the latent partisanship score; there is a difference in scoring across groups, but it is constant across latent partisanship. Nonuniform DIF describes situations where there is an interaction between latent partisan strength and groupings, such that the probability of a partisan response differs by groups across the latent partisanship score. The varying relevance model is an example of accounting for nonuniform differential item function across the two eras.

There are a variety of alternative ways to test for differential item functioning that go beyond the tests in Table 2. One popular method is to use the predicted latent score from the IRT estimate within regression models that assume uniform DIF and nonuniform DIF across a group of interest (Swaminathan and Rogers 1990). The approach to the ordinal case used in this paper by regressing each item on the estimated latent partisanship score from the fixed relevance model within ordered logistic regression. A nested comparison is made using the likelihood ratio test, comparing the nested uniform model, where there is no difference in relationship between the latent score and the item by group, to the nonuniform model that includes an interaction term between the group and the latent score to test for nonuniform differences in how the scores perform by that group.

Table C.2. Nonuniform DIF Tests for each item within the Basic Fixed Relevance Model

Item	χ^2	p-value
Partisan Categorization	1937.21	0.00
Partisan Strength	2086.24	0.00
Party Affect	926.86	0.00
Party Therm. Diff.	13.13	0.00
Pres. Therm. Diff.	1173.44	0.00
Pres. Support	348.73	0.00
House Vote	172.66	0.00

Likelihood-ratio tests comparing the log-likelihood of the nonuniform DIF model to its nested variant which assumes no differences by era (1952-1998 and 2000-2020). Estimates are based on an ordered logit regression for each item using the latent partisanship score from the fixed relevance model presented in Table 2.

Table C.2 presents the nonuniform DIF test results. All the items indicate significant, nonuniform differential item function across the two eras. This adds further justification for the use of a varying relevance model to estimate partisan strength.

Additional Results from Table 2

Figure C.1 plots out the change in each item's discrimination parameter across eras based on the estimates from Table C.1. Figure C.2 compares the item-specific cutpoint location parameters across eras for both the fixed and varying relevance models from Table 2. Table C.4 reproduces Table 2 when using the estimates of the expanded model that includes ideological and issue items. And Table C.4 presents the constituent parameter estimates that were used to generate the estimates and findings presented in Table C.3.

Figure C.1. Item Discrimination Estimates, Basic Model

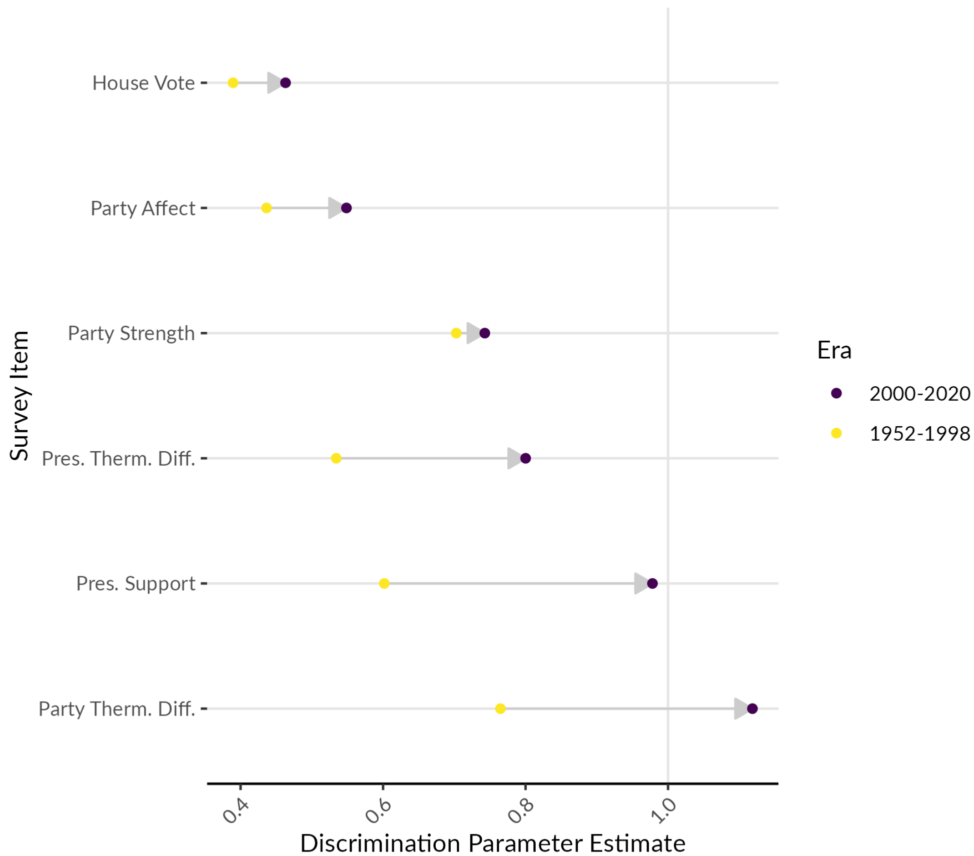


Figure C.2. Item Location Estimates, Basic Model

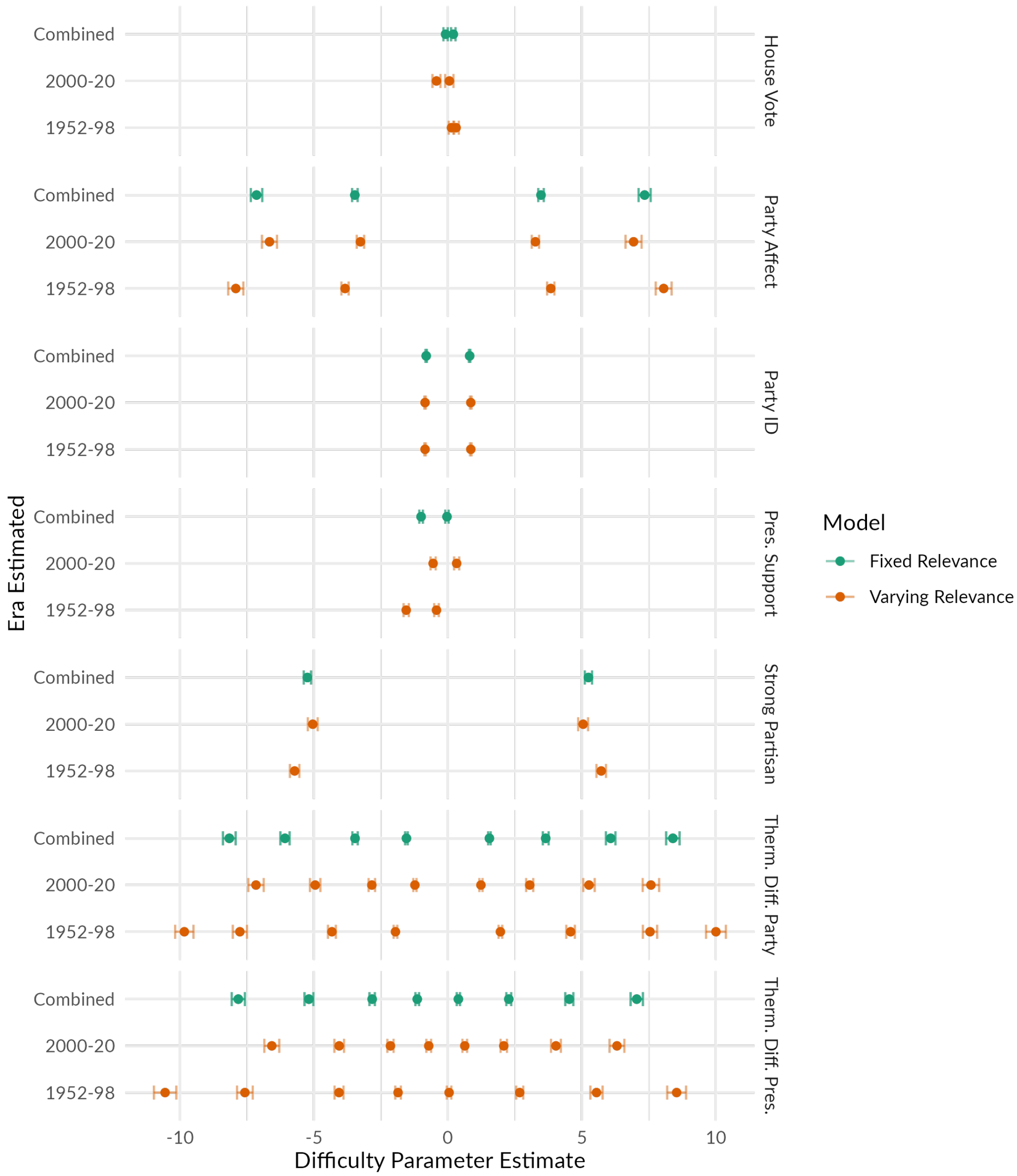


Table C.3. Expanded Model Latent Partisan Strength Estimates and Tests (1952-2020)

	Fixed Relevance Model		Varying Relevance Model	
Latent Partisanship	Pre-2000	2000 & After	Pre-2000	2000 & After
Average Strength				
$\lvert\hat{\theta}_i^{EB} - 0\rvert$	3.10	3.91	3.58	3.48
	(.01)	(.02)	(.01)	(.02)
μ_θ	−0.781	−0.602	−0.952	−0.456
	(.025)	(.044)	(.030)	(.042)
σ_θ^2	15.465	23.120	20.349	18.441
	(.392)	(.635)	(.609)	(.664)
Log-Likelihood	−412961.8		-411253.8^{+}	
n	67,502		67,502	

Weighted estimates by using ANES post-stratification weights and adjusting them to make each study an equivalent sample size so that discrimination and location estimates are averages of the two eras. $\hat{\theta}^{EB}$ indicates Empirical Bayes estimates of latent partisanship. $^{+}$ indicates likelihood-ratio test rejects null hypothesis that the unconstrained, varying relevance model provides no improvement in the log-likelihood relative to the nested, fixed relevance model.

Table C.4. Coefficient Estimates from the Expanded Graded Response IRT Model (1952-2020)

Variable	Component	Pre-2000	2000 & After
Latent Partisanship	Mean	-.952 (.03)	-0.456 (.04)
	Variance	20.35 (.61)	18.44 (.66)
		Discrimination Estimates	
Partisanship	Categorization	1 (Constrained)	
	Strength	0.724 (.014)	0.749 (.018)
Party Affect		0.462 (.009)	0.570 (.013)
Thermometer Differences	Party	0.812 (.015)	1.092 (.025)
	President	0.583 (.012)	0.861 (.021)
Party Support	Pres. Support	0.647 (.016)	1.036 (.034)
	House Vote	0.407 (.009)	0.487 (.014)
Ideology	Proximity	0.359 (.008)	0.464 (.012)
Issues	Abortion	0.228 (.011)	0.272 (.008)
	Gov. Service	0.346 (.011)	0.449 (.012)
	Aid to Blacks	0.285 (.009)	0.392 (.011)

Coefficient estimates with standard errors in parentheses. Higher values of latent partisanship indicate stronger Republican attachments, such that Democratic thermometer items with higher scores have a negative relationship.

D Measures of Strength's Performance in Identifying Future Switchers

Table D.1. Predicting Partisan Switching (AUC) by Measure of Partisan Strength and ANES Panel Study

Study Years	Measure of Attachment Strength		
(*n*)	Subj. Strength	Therm. Diff.	Latent Strength
Observed Sample			
ANES 1956-60	.655		.716
(1032)	(.028)		(.029)
ANES 1972-76	.629	.540	.689
(1135)	(.029)	(.030)	(.029)
ANES 1992-96	.652	.646	.742
(525)	(.031)	(.037)	(.032)
ANES 2000-04	.700	.632	.748
(752)	(.033)	(.039)	(.033)
ANES Reinterview 2016-20	.680	.731	.842
(2475)	(.015)	(.019)	(.014)
Imputed Sample			
ANES 1956-60	.657		.717
(1050)			
ANES 1972-76	.628	.541	.688
(1145)			
ANES 1992-96	.680	.646	.743
(1005)			
ANES 2000-04	.745	.683	.775
(1807)			

Nonparametric estimates of the area under the receiver operator characteristic curve (AUC) with standard errors in parentheses are presented for the observed data estimates. Imputed sample estimates present average AUC estimates from logistic regression model in-sample estimates across imputation samples.

Comparing Latent Strength's Correlation with other Indicators of Strong Partisan Behavior

The latent strength measure also outperforms other observed measures of strength in differentiating other types of strong partisan behavior that are not direct functions of the input variables. Strong partisans are typically observed to be more likely to vote straight ticket and more likely to participate in campaigns. Table D.2 reports the correlation coefficient estimates of the latent partisan strength measure with these two outcomes and compares it to two alternative, observed measures of partisan strength. Across both eras and for both types of behavior, the latent measure of partisan strength has a strong linear association than either two observed survey measures.

Table D.2. Correlation of Additional Partisan Behaviors by Measure of Partisan Strength

	Measure of Attachment Strength		
Era	Subj. Strength	Therm. Diff.	Latent Strength
Pre-2000 Sample			
Straight Ticket Voting ($n = 8776$)	.162	.197	.358
Campaign Participation ($n = 26,715$)	.155	.149	.199
2000 & After Sample			
Straight Ticket Voting ($n = 13,478$)	.152	.223	.320
Campaign Participation ($n = 13,513$)	.208	.247	.289

Pearson correlation coefficients estimates from unweighted era samples. Straight ticket voting is calculated from VCF0709 and campaign participation is calculated from VCF0723 in the ANES cumulative data file.

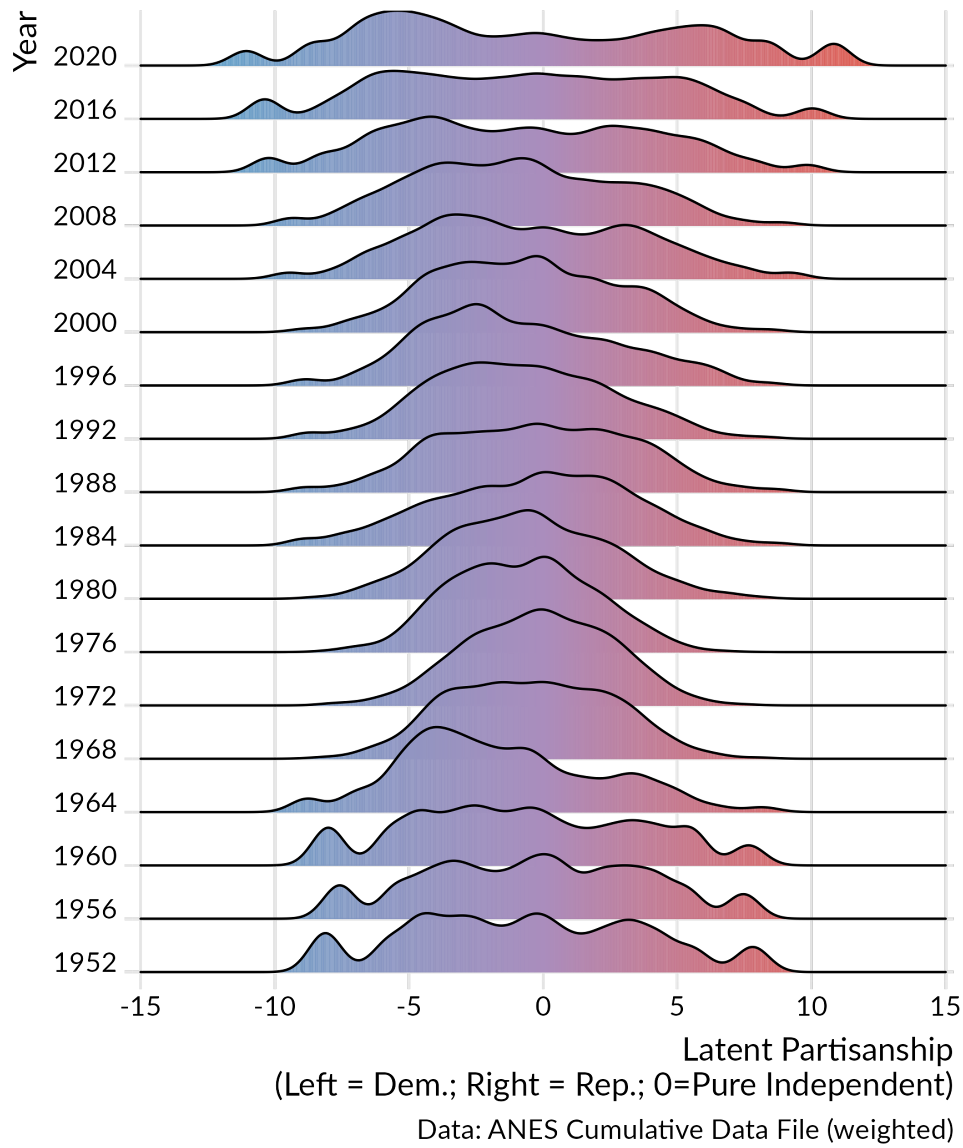
E Supporting information for the year-specific model estimates

Table E.1. Year-Specific Latent Partisanship Estimates (1952-2020)

Parameter	Year	Fixed Relevance Model		Varying Relevance Model	
μ_θ	1952	−0.665	(.12)	−1.269	(.18)
	1956	−0.279	(.11)	−0.722	(.14)
	1960	−0.684	(.17)	−0.752	(.19)
	1964	−1.568	(.10)	−1.994	(.19)
	1968	−0.454	(.09)	−0.792	(.13)
	1972	−0.144	(.06)	−0.645	(.08)
	1976	−0.777	(.07)	−0.831	(.08)
	1980	−0.595	(.08)	−0.781	(.11)
	1984	−0.334	(.08)	−0.381	(.10)
	1988	−0.292	(.09)	−0.296	(.10)
	1992	−0.845	(.07)	−0.806	(.09)
	1996	−0.841	(.10)	−0.840	(.14)
	2000	−0.668	(.10)	−0.801	(.13)
	2004	−0.422	(.14)	−0.405	(.17)
	2008	−0.950	(.11)	−0.945	(.15)
	2012	−0.765	(.09)	−0.426	(.08)
	2016	−0.679	(.10)	−0.348	(.07)
	2020	−0.359	(.09)	−0.110	(.08)
σ^2_θ	1952	22.40	(1.26)	33.29	(4.98)
	1956	19.25	(1.05)	25.82	(2.78)
	1960	21.24	(1.48)	23.39	(3.80)
	1964	14.76	(.73)	26.69	(3.76)
	1968	9.90	(.46)	15.49	(1.80)
	1972	8.44	(.36)	11.26	(1.08)
	1976	8.04	(.35)	10.41	(.95)
	1980	10.21	(.50)	13.09	(1.36)
	1984	13.79	(.59)	17.08	(1.60)
	1988	13.44	(.58)	16.09	(1.53)
	1992	11.76	(.50)	15.38	(1.32)
	1996	13.94	(.68)	21.69	(2.74)
	2000	12.58	(.59)	19.55	(2.16)
	2004	18.04	(.91)	24.65	(3.84)
	2008	16.02	(.74)	26.92	(2.96)
	2012	24.30	(.89)	17.60	(1.16)
	2016	26.47	(1.01)	10.57	(.75)
	2020	36.40	(1.26)	20.99	(1.18)
Log-Likelihood		−280895.26		-274444.05^{+}	
Constraints		489		104	
n		46,816		46,816	

Weighted estimates with standard errors in parentheses. $^{+}$ indicates likelihood-ratio test rejects null hypothesis that the unconstrained, varying relevance model provides no improvement in the log-likelihood relative to the nested, fixed relevance model.

Figure E.1. Distribution of Latent Partisanship ($\hat{\theta}_i^{EB}$) by Presidential Election (Fixed Relevance Model)



References

Bafumi, Joseph, and Robert Y. Shapiro. 2009. "A New Partisan Voter." *The Journal of Politics* 71 (1): 1–24.

Green, Donald Philip. 1988. "On the Dimensionality of Public Sentiment toward Partisan and Ideological Groups." *American Journal of Political Science* 32 (3): 758–780.

Smidt, Corwin D. 2017. "Polarization and the Decline of the American Floating Voter." *American Journal of Political Science* 61 (2): 365–381.

Swaminathan, Hariharan, and H Jane Rogers. 1990. "Detecting differential item functioning using logistic regression procedures." *Journal of Educational Measurement* 27 (4): 361–370.

Zingher, Joshua N. 2022. *Political Choice in a Polarized America: How Elite Polarization Shapes Mass Behavior.* New York: Oxford University Press.