APPENDICES TO “Protest Movements and Citizen Discontent”

Appendix A: Question Wordings

IDEOLOGY: “How would you describe your views on most political matters? Generally do you think of yourself as liberal, moderate, or conservative?”

EDUCATION “What was the last grade in school you completed?” Possible responses: Not a High School grad, High School grad, Some college (trade or business), College grad, Post-grad work or degree (Masters, MBD, JD, MD, PhD).

WRONG TRACK: “Do you feel things in this country are generally going in the right direction or do you feel things have pretty seriously gotten off on the wrong track?”

TEA PARTY: “Do you consider yourself to be a supporter of the Tea Party movement, or not?”

OCCUPY: “Do you consider yourself to be a supporter of the Occupy Wall Street movement, or not?”

GRIDLOCK: “These days, who do you blame more for the difficulties in reaching agreements and passing legislation in Congress--the Republicans in Congress or Barack Obama and the Democrats in Congress?” Response of “both” counted at ‘1’.

Appendix B: Fits for IRT Model

We here discuss the general evidence in favor for an item-response theoretic model for these data in general, and then, for the more specific (constrained) model that is in accord with our theoretical arguments. We begin by attempting to see whether, in each of the six categories formed by the cross-classification of ideology and education, the item response theory model (the Rasch model) is an adequate characterization of the data.

Table B-1 presents fit statistics for the model, most importantly, the R_{1c} statistic (Glas 1988), which is a test of unidimensionality. It will be noted that while we would accept the Rasch model as an adequate characterization for the data for liberals, and for high education conservatives, it does not fit well for moderates, or for low education conservatives. The reason for the moderates is not hard to find: as we saw in Table 1, the odd-ratio for moderates is near zero, and while the magnitude of the zero-order odds ratio is not itself a strong indication as to whether the Rasch model explains some data (Duncan 1984), in this case, there is little traction for the model in other associations as well. However, we note that in the four item scale, we do not reject the Rasch model for the moderates; since the estimates of the item parameters and the trait value are nearly identical, we take this to mean that the Rasch model still provides a useful description of the data. Some of the badness of fit results from change between the two waves, which we take into account in the models presented in Table 3.

<<Table B-1 about here>>

Having demonstrated that the model is a good one for all our education/ideology groups separately, we go on to present models for the combined data. This pooling of ideological and educational categories allows us to make theoretically derived constraints, in this case, to require certain parameters (but not others) to be identical across groups that share an ideology, but differ education. Second, it allows us to disaggregate the two waves of data (such a disaggregation would be implausible given the sample sizes within each subgroup formed by the cross-classification of education and ideology). This can prevent us from being confused by pooling data from two different samples, which might affect our capacity to test theoretically informed models. We continue to analyze both the three-item and four-item scales, as the former allows us to examine change.

Finally, there is a third reason for pooling: previous research has demonstrated interesting patterns of geographical variation in supported for at least the Tea Party (see Cho, Gimpel and Shaw [2012: 116], McVeigh, Beyerlein, Vann, and Trivedi [2014]; Ulbig and Macha [2012: 105]). Our sample sizes were too small to make possible incorporating regional differences when the models were conducted independently. With the pooled data, we were able to include random disturbances at the state level in the trait. However, we were not able to discern significant variation at the state-level; models tended to have difficulty converging and parameter estimates were very close to those that we estimate when we omit the state-level random effects. We hence conclude that state level variation is negligible,[1] and stick with the more reliable two-level (response nested within person) models instead.

Having pooled the data, we can recast the individual Rasch tests from Table B-1 as a single model. In other words, it is an unconstrained model that allows all parameters to vary across our six subgroups; this appears below as a base model (M_1), for the two waves separately. With this pooled version of the data, we go on to impose the theoretically derived constraints of Table 2 as model M_2. This model is nested in M_1, but adds constraints (equating item difficulties

[1]. This is not altogether surprising as prior results occurred at either much more fine-grained (e.g. county-level) or coarser (e.g. region) geographies.

across categories). For the reason, we expect that the fit will be worse than M_1. However, we can get a sense of whether the theoretical arguments are on the right track by comparing to the results when we fit a third model, in which we *reverse* the theoretical predictions, including the interactions with education for liberals when we predict Occupy support, and those with education for conservatives when we predict Tea Party support (M_3). In addition, for the four item scale, we also examine a model (M_4) that constrains there to be no interactions between education and difficulty for the "both parties bad" item (one which is inherently non-partisan), but does not impose constraints on the Tea Party and Occupy difficulties, and therefore sits between both M_2 and M_3 on the one hand and on M_1 the other.

<<Table B-2 about here>>

The results of our tests are found in Table B-2.[2] We show results for the three-item scales in both waves and results from the four item version in the second. The top panel condenses the results of our model comparison, which are found in the bottom panel. Because both M_2 and M_3 (as well as M_4 for the four-item scale) are nested within M_1, we can use a conventional chi-square test. This test shows that the difference between M_1 and M_2 is not statistically significant at .05 level for the first wave three item scale, although it is for the four item scale and is near significant (p=.06) for the second wave three item scale. We also examine model selection criteria AIC and BIC; the latter is known to be more conservative, in that it prefers more parsimonious models than the former. We find that M_2 is consistently preferred by

---

[2]. Note that in non-linear hierarchical linear models, we can only approximate the likelihood; however, the degree of imprecision in the approximation is generally of orders of magnitude less than that of the test we conduct. While one would use a test of restricted maximum likelihoods for two models who differ in their random effects, for comparisons of models differing in their fixed parts, as do these, we want to use the maximum likelihood estimates of the final likelihoods (Snijders and Boskers 2000: 57, 89).

BIC, while in two of the three cases, AIC prefers the full model (though one will note that even here, M_2 is quite close to M_1). Significantly, the reversed-model, M_3, is *never* preferred, and consistently fits worse. Thus, there is evidence that our restriction of the effects of sophistication (here tapped by education) to increasing the difficulty of supporting the ideologically opposed protest movement, while somewhat over-parsimonious, is in keeping with the main patterns in the data.

Appendix C: Multilevel Realization of the IRT Model

We treat each observation Y_{ij} as the j^{th} individual confronting the i^{th} item, and producing a response coded '1' if the answer is in a positive direction, and 0 otherwise.[3] Let π_j represent individual j's degree of the trait (disgruntlement). Each of the items has a difficulty, which we here convert to an "easiness" for consistency with standard notation for hierarchical linear models. Because the easiness of items varies by persons (at least, within certain groups), we denote the easiness of the i^{th} item for person j as α_{ij}. We thus treat our data as item responses (level 1) nested within persons (level 2), and write a probability of a positive response as a level 1 equation:

$$log\left[\frac{Prob\left(Y_{ij}=1\right)}{1-Prob\left(Y_{ij}=1\right)}\right]=\pi_j+\alpha_{ij} \qquad (1)$$

Thus level-1 comprises a measurement model for the response, while level-2 consists of multivariate "explanatory" model (de Boeck and Wilson 2004). Here individual j's latent disgruntlement is a function of a set of education and ideology fixed parameters and a person-specific random parameter, $u_j \sim N(0, \tau)$:

$$\pi_j=\gamma_1 CONS_j+\gamma_2 LIB_j+\gamma_3 COLLEGE_j+\gamma_4 CONS*COLLEGE_j+\gamma_5 LIB*COLLEGE_j+u_j \quad (2)$$

In omitting an intercept for π_j, we simply scale the latent trait such that 0 is the discontent of the average non-college educated moderate[4]. The fixed parameters ($\gamma_1,\ldots,\gamma_5$) are interpreted just as

---

3. Because we nest measurements within individuals, our own i is actually particular item-response combination for any individual, but for ease of exposition, we ignore this here.

4. There is no substantive difference that results from the decision to omit. It is simply a more convenient specification for attaining estimates of item difficulties for all items. Lacking the constraint, we would have to omit an indicator variable for one item.

one would interpret parameters in an ordinary regression equation. They represent the additive effects of college education, ideology, and their interactions.

Differential item functioning across groups is accounted for in the level-2 equations predicting the level-1 parameters α_{1j}, …, α_{4j}. To avoid overidentifying the model we must, by necessity, constrain the difficulty of at least one item (in this case “wrong track”) to be constant across individuals conditional on latent discontent. This is our “anchor item.” The constant difficulty of the wrong track item is given by δ_{11}. The difficulties of the other items for non-college educated moderates are represented by δ_{i1}, where i = 2, 3, 4. We do not include person-specific random effects for the difficulties of any item. Thus,

$$\alpha_{1j} = \delta_{11} \tag{3}$$

and

$$\alpha_{ij} = \delta_{i1} + \delta_{i1}CONS_j + \delta_{i2}LIB_j + \delta_{i3}COLLEGE_j + \delta_{i4}CONS*COLLEGE_j + \delta_{i5}LIB*COLLEGE_j \tag{4}$$

for i = 2, 3, and 4. If the Tea Party is our i = 2, and Occupy i = 3, our model M_2 can then be understood as the restrictions $\gamma_{03} = \delta_{23} = \delta_{24} = \delta_{33} = \delta_{35} = 0$, and for the four trait model, the additional restriction that $\delta_{43} = \delta_{44} = \delta_{45} = 0$.

Typically, mixed effects DIF models specify separate variances for the latent trait in each group (e.g. we allow the possibility that there is greater variation in disgruntlement among conservatives without a college degree than among college-educated conservatives). We do not do so in the models reported here. We have attempted such models; however, we have found that the groups differ so little in their internal heterogeneity that such models cannot converge within any reasonable level of tolerance. That is, between-group differences in trait variance are so small that they cannot be estimated within the boundary of the parameter space.

Because of the lack of convergence for maximum likelihood approximation (via Laplace transform) in Wave 1, for consistency, we present unit-specific penalized quasi-likelihood (PQL) results for the models in Table 3. For both Wave 2 models, the PQL estimates are consistent with those obtained under the Laplace approximation of the maximum likelihood, the basis for the model fit comparisons just discussed. Given the low number of well-educated ideologues who support the contra-indicated movement, it is not surprising that some parameter estimates are labile to the inclusion of respondents with incomplete data.[5] Here we have analyzed only data from respondents with complete data. Finally, we reiterate that we display the item parameters in terms of *easiness*, not difficulty, to facilitate a conventional interpretation of coefficients from a hierarchical linear model. The effect on difficulty is found simply by multiplying any coefficient by -1.

---

[5]. In particular, the seemingly anomalously small coefficient for the three item scale, wave 1, for the college education effect among conservatives for the *Occupy* difficulty, becomes large and negative. We therefore do not believe that this indicates change over time.

Table B-1: Rasch Model Fit Statistics

Table B-1a: Fits for Three Item Scale

Ideology	R_{1c}	df	$p(R_{1c})$	N_{obs}	N_{perf}	CLL	LL
Liberals							
Low Ed	1.178	2	.555	148	42	-78.37	-166.30
High Ed	3.083	2	.214	191	50	-67.95	-163.20
Moderates							
*Low Ed**	12.903	2	.002	314	77	-165.50	-349.79
*High Ed**	16.466	2	.000	281	77	-126.52	-280.87
Conservatives							
*Low Ed**	34.800	2	.000	319	50	-124.36	-287.15
High Ed	0.258	2	.879	210	16	-32.77	-100.56

* Rasch model rejected by unidimensionality test.

Table B-1b: Fits for Four Item Scale

Ideology	R_{1c}	df	$p(R_{1c})$	N_{obs}	N_{perf}	CLL	LL
Liberals							
*Low Ed**	17.062	6	.009	95	10	-105.04	-180.37
*High Ed**	14.048	6	.029	123	16	-110.44	-199.38
Moderates							
Low Ed	6.897	6	.330	209	36	-190.65	-342.22
High Ed	9.721	6	.137	212	30	-178.95	-328.14
Conservatives							
*Low Ed**	25.222	6	<.001	204	25	-172.74	-315.89
High Ed	3.726	6	.714	148	9	-72.51	-152.37

Columns are, in order, chi-square, degrees of freedom, and probability for R1c test; number of non-missing observations; number of observations that are all 1s or all 0s; conditional log likelihood and log likelihood.

Note: *N* includes perfect scores, but not those with any missing values on any item.

Table B-2: Results of Model Comparison: Laplace estimates of marginal maximum likelihood

Table B-2a: Summarized

Scale/Sample	*Deviance*	*Df*	*P-value* M_2 *vs.* M_1	*AIC*	*BIC*
4 items	17.44	7	.015	M_1	M_2
3 items, First Wave	4.90	4	.300	M_2	M_2
3 items, Second Wave	9.02	4	.060	M_1	M_2

Table B-2b: Further Information

Four Items	Deviance	Parameters	AIC	BIC
M_1 (original)	11638.61	25	11688.61	11845.84
M_2 (theoretically predicted)	11656.05	18	11691.05	11805.25
M_3 (anti- M_2)	11670.51	18	11706.51	11819.71
M_4 (in between M_1 and M_2)	11647.53	22	11691.53	11829.89

Three Items, W1[6]	Deviance	Parameters	AIC	BIC
M_1 (original)	1363.76	19	1401.76	1501.62
M_2 (theoretically predicted)	1368.68	15	1398.68	1477.52
M_3 (anti- M_2)	1374.97	15	1404.97	1483.80

Three Items, W2	Deviance	Parameters	AIC	BIC
M_1 (original)	8467.82	19	8505.82	8619.85
M_2 (theoretically predicted)	8476.84	15	8506.84	8596.86
M_3 (anti- M_2)	8496.18	15	8526.18	8616.20

[6]. Results for this subtable came from using the lme4 library (Bates et al. 2015) in R; results for the others from HLM7 (Raudenbush et al 2011). HLM7 did not converge for all models in this subtable; replication of the models where HLM7 did converge demonstrated that the deviances are related to one another by a constant proportional to *N*.

ADDITIONAL REFERENCES FOR APPENDICES

(OTHER REFERENCES IN MAIN PAPER)

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67(1); doi: 10.18637/jss.v067.i0.

Glas, C. A. W. 1988. The Derivation of Some Tests for the Rasch Model from the Multinomial Distribution." *Psychometrika* 53: 525–546.

Snijders, Tom A. B. and Roel J. Bosker. 2000. *Multilevel Analysis*. London: Sage.