Supplemental Appendix For:

Distinguishing Occasional Abstention from Routine Indifference in Models of Vote Choice

In this appendix, we first discuss a series of Monte Carlo experiments that compare the performance of the multinomial logit (MNL) and baseline inflated multinomial logit (BIMNL) models under various conditions. We then describe the operationalizations of several independent and control variables that are used in our main paper's applications, and present the corresponding tables of (BIMNL and MNL) coefficient estimates and standard errors for each application. Finally, we present a table of the 42 multinomial choice articles that we identified as appearing in the *American Political Science Review*, *American Journal of Political Science*, and *Journal of Politics* during the years 2009-2013, along with an indication of the specific multinomial estimator(s) that were used in each article.

1. Monte Carlo Experiments

To evaluate the performance of the MNL and BIMNL models in finite samples, we assess the results from two main Monte Carlo exercises below. The first Monte Carlo exercise ("Experiment 1") compares the MNL and BIMNL models when the baseline choice category of a discrete polytomous dependent variable is generated from two distinct d.g.p's, and thus "inflated". Here, we specifically evaluate the MNL and BIMNL models in cases where the degree of inflation in the baseline category of a (three category) polytomous dependent variable ($\mathbf{y}_1 = 1, 2, 3$) is set to the modest level of approximately 35% (of all observations);[1] which corresponds to 75% inflation among baseline category responses. Following similar simulation studies (e.g., Bagozzi and Mukherjee, 2012) we have chosen to set the inflation level of our dependent variable to this relatively conservative proportion in order to evaluate the boundaries of the BIMNL model under "real world" (i.e. moderate) inflation conditions. For the second Monte Carlo exercise ("Experiment 2"), we repeat the steps described above when using a comparable (three category) polytomous dependent variable ($\mathbf{y}_2 = 1, 2, 3$) that has no inflation in the

[1]That is, 35% on average, for each set of ($sims = 5,000$) simulations conducted.

baseline choice category (i.e., that follows a MNL d.g.p.). The aim of this second Monte Carlo experiment is to evaluate the relative performance of the BIMNL model in situations where it is misapplied.

For our two experiments, we set the number of *sims* $= 5,000$ and then compare the performance of our MNL and BMNL models across finite samples of $N = 5,000$.[2] In each case, we draw our outcome stage covariates $\mathbf{x}$ from $\mathbf{x} = (\mathbf{1}, \mathbf{x_1}, \mathbf{x_2})'$ where $\mathbf{x_1}$ is the natural log of $Uniform[0,100]$ and $\mathbf{x_2} = 1_{Uniform[0,1]>0.5}$. For Experiment 1 our inflation stage covariates $\mathbf{z} = (\mathbf{1}, \mathbf{z_1}, \mathbf{z_2})'$ are similarly drawn from $\mathbf{z_1} = ln(Uniform[0,100])$ and $\mathbf{z_2} = \mathbf{x_2} = 1_{Uniform[0,1]>0.5}$. Note that while these latter covariates do not add inflation to Experiment 2's MNL d.g.p., Experiment 2 nevertheless adds these covariates to the inflation stage specifications of our BIMNL model during estimation in order to best approximate real world instances of BIMNL misapplication. Draws of $\mathbf{x}$ and $\mathbf{z}$ were allowed to vary within our simulations.[3] Parameter values were assigned as $(\beta_{1,c_2}, \beta_{2,c_2}, \beta_{3,c_2})' = (-1.25, 0.25, -0.50)'$ for our second choice outcome (i.e., $\mathbf{y} = 2$) in each experiment; as $(\beta_{1,c_3}, \beta_{2,c_3}, \beta_{3,c_3})' = (-1.00, 1.00, -0.25)'$ for our third choice outcome (i.e., $\mathbf{y} = 3$) in both experiments; and as $(\gamma_1, \gamma_2, \gamma_3)' = (1.75, -0.5, 1.75)'$ for our inflation stage (i.e., $\mathbf{z}$'s). As mentioned above, these specifications produced a three category unordered dependent variable $\mathbf{y} = (1,2,3)$ with either (i) 75% inflation in category $\mathbf{y} = 1$ (Experiments 1)[4] or (ii) 0% inflation (Experiment 2).

1.1. *Experiment 1 Results*

The results for Monte Carlo Experiment 1 appear in Table A.1. Within this table, we report the mean (β and γ) coefficient estimates from our BIMNL and MNL models;[5] the mean

[2]Estimation was undertaken in R using optim() and BFGS.

[3]We thank an anonymous reviewer for this suggestion, and would like to also note that earlier drafts of this Supplemental Appendix reported comparable results to those discussed below under conditions where draws of $\mathbf{x}$ and $\mathbf{z}$ were taken once at each N and then held fixed for each simulation therein.

[4]I.e., 75% inflation as a share of all category 1 responses. This yields 35% (global) inflation as a share of all three outcomes of $\mathbf{y}$ and places approximately 45% of our total observations within the baseline choice category ($\mathbf{y} = 1$).

[5]Our focus on coefficient estimates is consistent with the Monte Carlo simulations conducted by Bagozzi and Mukherjee (2012) for middle-inflated discrete choice estimators. Initial experiments suggest that the main insights discussed below are comparable when one instead evaluates the theoretical and empirical changes in (MNL and BIMNL derived) probabilities, as opposed to the coefficients themselves.

absolute errors[6] (MAE's) of these estimates, the parameter estimates' 95% empirical coverage probabilities (CPs),[7] and a selection of model fit statistics.[8] In addition to also reporting a single mean MAE across all β estimates for a given sample size ($\overline{MAE}_{\beta}$), our aggregate model fit statistics separately report the percentage of times that the correct model (in this case the BIMNL) was identified by (i) a generalized Vuong test statistic, (ii) a superior proportion reduction in error (PRE), (iii) the *AIC*, (iv) the *BIC*, and (v) a likelihood ratio (*LR*) test.[9] Finally, we also report the proportion of non-convergence instances for each model.

Turning to Table A.1, one can first observe that, with $N = 5000$, a majority of our average BIMNL β coefficient estimates are comparable to their true parameters values, whereas the MNL β coefficient estimates often dramatically diverge from these true estimates. For example, our mean BIMNL coefficient estimate for the effect of $\mathbf{x_2}$ on Outcome 2 is 0.287, very close to the true estimate (0.250), whereas the comparable MNL coefficient estimate of -0.413 is nearly twice the size of our true estimate, and in the opposite direction. This case aside, we more generally find the mean MNL coefficient estimates to be biased towards zero, though not the BIMNL estimates. Altogether, these findings are captured quite well in the individual MAE statistics for each coefficient estimate, as well as the global $\overline{MAE}_{\beta}$ statistic for our model as a whole. In the majority of these cases, we find that the MAEs of our individual BIMNL estimates are dramatically lower than those of the MNL model, a trend that is reflected by a BIMNL $\overline{MAE}_{\beta}$ (0.194) that is less than half that of the MNL's $\overline{MAE}_{\beta}$ (0.442). Hence, the BIMNL model exhibits notably superior accuracy in its estimates, relative to the MNL model, when one's d.g.p. is BIMNL. Two modest exhibits include the MAE's for $\mathbf{1}$ and $\mathbf{x_2}$ in Outcome 2, wherein the MNL estimates' MAE's are slightly superior to those of the BIMNL estimates. For this particular Experiment, the 95% empirical CP values favor the BIMNL model in every

[6]Calculated relative to each estimate's true value, over each set of $5,000$ simulations. As reported in an earlier draft of this paper, we obtain comparable results with root mean square errors (RMSE's).

[7]I.e., the average proportion of times—out of $5,000$ simulations—that a true parameter value fell within the 95% confidence intervals of that parameter's estimate.

[8]Taken together, this approach is comparable to that used by Harris and Zhao (2007) and Bagozzi and Mukherjee (2012).

[9]Though note that the *LR* test is not strictly appropriate in this instance, given the non-nested nature of our models. See the main paper for more detailed discussion of each of these test statistics.

instance, with the BIMNL model's empirical CP's consistently falling in the mid 90% range. By comparison, while our MNL estimates occasionally yield empirical CP's in the 90%'s, in most cases they fall to as low as 0%, suggesting that even in those instances where the MNL estimates are relatively accurate, they tend to overstate one's confidence in these estimates. Finally, we can note in Table A.1 that we encountered virtually no convergence problems with either model for Experiment 1.

Experiment 1 also offers a number of insights with respect to model selection. As a whole, the five model fit statistics mentioned above provide mixed results in correctly favoring the BIMNL model over the MNL model for this experiment. The Vuong test statistic correctly selects the BIMNL model approximately 73% of the time, whereas the MNL model is only favored by the Vuong test 19% of the time (with the remainder corresponding to cases where the test was inconclusive). This lends some credibility to the use of Vuong tests for model selection in instances where the d.g.p. is BIMNL and one's sample size is moderate, though we can note that the Vuong test's performance in this case is notably worse than in simulation studies of the ZIOP model (Harris and Zhao, 2007).[10] Regarding the PRE, our BIMNL model exhibited an superior PRE to that of the MNL model only 28% of the time, while the MNL model outperformed the BIMNL model on this metric 72% of the time—indicating that the PRE statistic is an especially poor choice for multinomial model choice under suspected cases of baseline inflation. On the other hand, in virtually every simulation, *AIC*, *BIC*, and *LR* tests correctly favored the BIMNL model over the MNL model. Therefore, when one's d.g.p. is BIMNL, the *AIC*, *BIC*, and *LR* are dramatically superior choices for model selection than are the Vuong test or PRE statistic.

To summarize the above findings more concisely, we can conclude from Experiment 1 that when a discrete (three category) unordered dependent variable becomes contaminated by moderate levels of inflation in its baseline category, the parameter estimates recovered from BIMNL models are generally superior to those derived from a comparable set of MNL models. This

[10]Our inflation proportion is lower than that used by Harris and Zhao (2007). Hence, the poorer performance of the Vuong test in our case is not particularly surprising.

holds for both accuracy (MAE) and empirical coverage, and for a majority of our specific parameter estimates. In these contexts, our global BIMNL $RMSE_{\beta}$ was found to be less than half to that of the MNL model, suggesting that estimating an MNL model on an even modestly inflated polytomous dependent variable can lead to severe inaccuracy in one's primary coefficient estimates of interest. Looking more closely at our BIMNL model estimates exclusively, we generally find that the BIMNL model's β estimates exhibit higher accuracy than do the BIMNL model's γ estimates. However, the empirical CP's for both sets of BIMNL estimates are comparable. BIMNL non-convergence was approximately 0%, which is comparable to the ZIOP and MIOP model convergence levels observed in similar simulations, as reported by Harris and Zhao (2007) and Bagozzi and Mukherjee (2012), respectively, and is much better than the ZIOP convergence levels reported by Bagozzi et al. (2014) in simulations examining moderately lower inflation proportions (i.e., levels of inflation as low as 10%). As a whole, a majority of the model fit statistics in Experiment 1 correctly distinguished between the BIMNL and MNL models at a commensurate rate although two (the PRE and Vuong) did not.

1.2. *Experiment 2*

Experiment 2 evaluates the performance of the BIMNL and MNL models when the d.g.p. is explicitly MNL. The summary results for this Experiment are reported in Table A.2. Here we find that when the d.g.p. is MNL, the MNL model marginally outperforms the BIMNL model on most metrics. For instance, in comparing the the MNL and BIMNL models directly via their β parameter estimates in Table A.2, we find that both the MNL and BIMNL models recover the true parameter estimates relatively well. To this end, one can first observe that there appear to be only modest discrepancies between the mean β parameter estimates recovered by each model, as well as between each set of estimates and the true β parameter estimates. Our MAE findings reinforce this point. The individual MAEs for each BIMNL and MNL estimate are virtually identical, although the MNL MAEs are consistently lower than (and thus slightly superior to) the BIMNL MAEs. Turning to the $\overline{MAE}_{\beta}$, we find similar effects for this aggregate MAE summary statistic: the MNL model $\overline{MAE}_{\beta}$ suggests superior accuracy to that

Table A.1: Experiment 1: Marginal Effects For BIMNL DGP with N=5000

		Outcome 2 Estimates					Outcome 3 Estimates		
		True	MNL	BIMNL			True	MNL	BIMNL
x_1	*mean*	-1.250	-1.239	-1.233	x_1		-1.000	-0.436	-0.977
	MAE		(0.168)	(0.250)				(0.564)	(0.228)
	CP		0.946	0.965				0.002	0.965
		True	MNL	BIMNL			True	MNL	BIMNL
x_2	*mean*	0.250	-0.413	0.287	x_2		1.000	0.168	1.038
	MAE		(0.663)	(0.133)				(0.832)	(0.138)
	CP		0.000	0.950				0.000	0.951
		True	MNL	BIMNL			True	MNL	BIMNL
x_3	*mean*	-0.500	-0.316	-0.528	x_3		-0.250	-0.041	-0.278
	MAE		(0.212)	(0.228)				(0.209)	(0.190)
	CP		0.801	0.962				0.041	0.962
		Inflation Estimates					Model Fit		
		True	MNL	BIMNL				MNL	BIMNL
z_1	*mean*	1.750	.	1.170	$\overline{MAE}_\beta$			0.442	0.194
	MAE		.	(0.650)					
	CP		.	0.710	Vuong			0.193	0.734
		True	MNL	BIMNL	PRE			0.724	0.276
z_2	*mean*	-0.250	.	-0.333					
	MAE		.	(0.175)	AIC			0	1
	CP		.	0.625					
					BIC			0	1
		True	MNL	BIMNL					
z_3	*mean*	1.750	.	1.176	LR			0	1
	MAE		.	(0.574)					
	CP		.	0.017	Non Con.			0.00	0.00

of the BIMNL model, although both model exhibit fairly similar $\overline{MAE's}_\beta$. The 95% empirical coverage probabilities appear comparable for each model, falling in every case fall around the 95%-level, and in most cases slightly favoring the MNL model. Hence, when one's d.g.p. is MNL, it appears that the BIMNL and MNL models report comparable levels of uncertainty. Examining the inflation (i.e., γ) estimates for our BIMNL model under this, we observe in Table A.2 that the estimate for the intercept (γ_1), as well as, to a lesser extent, our two primary covariate estimates (γ_2-γ_3), are in each case large and positive. Given that the "dependent variable" for this stage is the probability of non-inflation in category $\mathbf{y} = 1$, this implies that our BIMNL model is effectively assigning each observation in our data set to have a probability of being non-inflated that is close to one. Lastly, while we again find non-convergence to be virtually non-existent for the MNL model, our BIMNL model under Exhibit 2 exhibits notable convergence issues, failing to converge approximately 42% of the time.

Turning to the model fit statistics for Experiment 2, we find that—similar to Experiment 1—several of our primary model fit statistics do not decisively favor either model. Nevertheless, for Experiment 2, the model fit statistics do marginally tend to favor the MNL model over the BIMNL model in all cases in Table A.2. Beginning first with the Vuong test statistics, we can note that Vuong tests favor the MNL models over the BIMNL models in roughly 64% of our simulations, and only favor the BIMNL model 18% of the time (with the remaining cases corresponding to instances where the Vuong test favored neither model). These proportions, while suboptimal from a model selection standpoint, are nevertheless a dramatic improvement over the Vuong test results reported by Harris and Zhao (2007) in their Monte Carlo comparisons of OP and ZIOP models under an OP d.g.p.[11] Similar to Experiment 1, the PRE often indicates that our MNL and BIMNL models predict **y** equally well. However, for the current experiment, the MNL model now predicts the remaining cases better than the BIMNL. *AIC*, *BIC*, and *LR* tests each favor the MNL the majority of the time in Table A.2, with the *BIC* (100%) performing better than either the *LR* test (94%) or the *AIC* (89%). The most striking difference between Experiments 1 and 2 arises with the BIMNL convergence levels for Experiment 2: we find here that an average of roughly 42% of the simulations saw the BIMNL model fail to converge.[12]

In sum, Experiment 2 suggests that for polytomous dependent variables with no inflation, the parameter estimates recovered from an MNL model will be slightly more accurate than will estimates from a BIMNL model. In this respect, we found that the global MNL $\overline{MAE}_\beta$ was marginally better than that of the BIMNL model, although the two $\overline{MAE}_\beta$ statistics were fairly comparable. The BIMNL model, in addition to exhibiting slightly lower levels of accuracy, also exhibits significant convergence problems when applied to dependent variables that follow an

[11] Which had indicated that, under these conditions, the Vuong test never correctly selected the OP model (Harris and Zhao, 2007).

[12] As noted immediately below, it is likely that these higher (MNL d.g.p.) convergence levels are arising from our inclusion of variables in the inflation stage that in this case are wholly uncorrelated with the true (MNL) d.g.p.

MNL d.g.p: across all sample sizes, the BIMNL model failed to converge approximately 42% of the time in Experiment 2. It is worth noting, however, that these high convergence problems may not have been as severe had we not included inflation stage covariates in the BIMNL model that were wholly uncorrelated with the true d.g.p., and we plan to examine this potential further in future extensions of this project. At present, the convergence problems identified in Experiment 2 are similar to those reported for the ZIOP(C) models when applied to OP data by Bagozzi et al. (2014), though more severe than those reported by Harris and Zhao (2007). On average, we found mixed results for our model fit statistics' abilities to correctly select the MNL model over the BIMNL model with the d.g.p. was MNL. Harris and Zhao (2007) report similar results for their OP/ZIOP comparisons when the d.g.p. is OP. In our case, the Vuong and PRE tests performed most poorly in distinguishing BIMNL and MNL models under an MNL d.g.p, and only identified the correct (MNL) model 64% and 19% of the time, respectively. On the other hand, the *AIC*, *BIC*, and *LR* tests each identified the MNL model as the correct model at commensurate rates, ranging from 89% of simulations to 100% of simulations.

1.3. *Discussion*

To summarize, we have found above that, when one's d.g.p. is BIMNL (Experiment 1), the BIMNL outperforms the MNL substantially in both accuracy and empirical coverage provided that one's sample size approaches $N = 5,000$. In this regard, a number of our mean MNL coefficient estimates exhibited entirely reversed signs to our true coefficients of interest. Even in those few cases where our MNL estimates were comparable to the BIMNL model in magnitude and MAE, we further observed notably worse empirical coverage probabilities. This latter finding suggests that not only are MNL estimates biased when misapplied to BIMNL data, but they will also often overstate one's certainty about one's biased estimated effects. By contrast, the MNL only marginally outperforms the BIMNL model in accuracy, and to a lesser extent in empirical coverage, when the d.g.p. is MNL (Experiment 2), wherein we find that our mean coefficient estimates are virtually identical among converged models. Overall, the BIMNL model's convergence problems were nonexistent when the true d.g.p. was BIMNL, but were

Table A.2: Experiment 2: Marginal Effects For MNL DGP with N=5000

		Outcome 2 Estimates				Outcome 3 Estimates		
		True	MNL	BIMNL		True	MNL	BIMNL
x_1	*mean*	-1.250	-1.261	-1.146	x_1	-1.000	-1.004	-0.921
	MAE		(0.164)	(0.227)			(0.121)	(0.179)
	CP		0.953	0.943			0.953	0.949
		True	MNL	BIMNL		True	MNL	BIMNL
x_2	*mean*	0.250	0.252	0.236	x_2	1.000	1.002	0.995
	MAE		(0.048)	(0.060)			(0.035)	(0.046)
	CP		0.952	0.943			0.957	0.958
		True	MNL	BIMNL		True	MNL	BIMNL
x_3	*mean*	-0.500	-0.503	-0.522	x_3	-0.250	-0.251	-0.271
	MAE		(0.130)	(0.146)			(0.083)	(0.101)
	CP		0.947	0.952			0.947	0.950
		Inflation Estimates				Model Fit		
		True	MNL	BIMNL			MNL	BIMNL
z_1	*mean*	0.000	.	11.804	$\overline{MAE}_\beta$		0.097	0.126
	MAE		.	(11.821)				
	CP		.	0.976	Vuong		0.642	0.178
		True	MNL	BIMNL	PRE		0.19	0.056
z_2	*mean*	0.000	.	5.274				
	MAE		.	(6.018)	AIC		0.886	0.114
	CP		.	0.800				
					BIC		1	0
		True	MNL	BIMNL				
z_3	*mean*	0.000	.	9.427	LR		0.944	0.056
	MAE		.	(13.429)				
	CP		.	0.998	Non Con.		0.00	0.42

noticeable when the d.g.p. was MNL. In both cases, BIMNL non-convergence levels were similar to ZIOP(C) convergence problems reported under comparable levels of inflation when one's d.g.p. was either ZIOP or OP (Harris and Zhao, 2007; Bagozzi et al., 2014), and for the MNL d.g.p. are at least partially attributable to the fact that we include a set of wholly irrelevant covariates in the inflation stage of our BIMNL model. In sum, when one suspects even a low to moderate level of baseline category inflation, the BIMNL model generally provides more accurate estimates of one's outcome stage covariates relative to the MNL model. However, convergence problems may suggest that one's BIMNL model is in fact being misapplied to an MNL-generated dependent variable.

Given a polytomous dependent variable with an ambiguous degree of baseline category inflation, and the potential choice of using a MNL or BIMNL model, which model fit statistics best inform us of the correct choice? Our Monte Carlo experiments help to shed light on this question. To this end, Experiments 1 and 2 indicate that standard information-based model selection criteria, such as *BIC* and *AIC*, correctly choose between the BIMNL and MNL models (under each d.g.p.) nearly 100% of the time. *LR* tests perform comparably, correctly choosing the BIMNL and MNL models in 94-100% of our simulations. These three results are very consistent with the OP/ZIOP simulation results reported in Harris and Zhao (2007). Also in line with these OP/ZIOP findings, the generalized Vuong test statistic described above was highly accurate in choosing the BIMNL model when the d.g.p. was BIMNL, but performed poorly in selecting the MNL model when the d.g.p. was MNL. The PRE fared even worse in both experiments and in most simulations favored either the MNL model, or neither model over the other. Hence, researchers interested in using model fit statistics to supplement their (BIMNL versus MNL) model selection decisions should employ a combination of the test statistics mentioned above, and should place more weight in the *BIC*, *LR* test, and *AIC*, relative to the Vuong test and the PRE.

Lastly, it is worth emphasizing that the Monte Carlo findings and insights discussed above are dependent upon our specific choices of sample size, inflation proportion, and parameter values. While some additional experimentation on our part, as well as past simulation studies of similar models, suggests that higher inflation proportions (than those examined above) will favor the BIMNL model over the MNL model even more dramatically than found in Experiment 1—and will reduce the convergence problems highlighted earlier—this remains an active area of research.

2. **Applications**

This section presents the variable operationalizations and full tables of coefficient estimates for our two replication studies. We begin by presenting the materials corresponding to our replication of Arceneaux and Kolodny (2009). Descriptions of the variables used in this replication are presented immediately below, followed by the full table of BIMNL and MNL estimates.

Independent and Control Variable Operationalizations for Arceneaux and Kolodny (2009)

- *Canvassing Treatment*: Binary indicator denoting survey respondents that received door-to-door canvassing treatment from liberal interest group
- *Phone Treatment*: Binary indicator denoting survey respondents that received telephone endorsement treatment from liberal interest group
- *Democrat*: Binary indicator of whether survey respondent considered themselves to be a Democrat
- *Republican*: Binary indicator of whether survey respondent considered themselves to be a Republican
- *CanvassingXDem*: Interaction of *Canvassing Treatment* and *Democrat*
- *CanvassingXRep*: Interaction of *Canvassing Treatment* and *Republican*
- *PhoneXDem*: Interaction of *Phone Treatment* and *Democrat*
- *PhoneXRep*: Interaction of *Phone Treatment* and *Republican*
- *Vote 2004*: Binary indicator of whether survey respondent reported having voted in 2004 election
- *Age*: Respondent's age in years
- *Female*: Binary indicator of whether survey respondent was female or male
- *House Size*: Count of the number of members in respondent's household
- *District 156*: Binary indicator of whether survey respondent lived in District 156 or District 161

Table A.3: BIMNL and MNL Models of Vote Choice (Pooled Sample)

	MNL			BIMNL			
	R/O vs. Abstain	D vs. Abstain	NR vs. Abstain	R/O vs. Abstain	D vs. Abstain	NR vs. Abstain	Inflation Stage
Canvassing Treat	1.222***	0.335	0.000	1.566***	0.680	0.382	.
	(0.455)	(0.387)	(0.466)	(0.558)	(0.514)	(0.567)	
Phone Treat	0.763*	0.464	-0.095	1.315**	1.045**	0.464	.
	(0.438)	(0.353)	(0.427)	(0.555)	(0.504)	(0.545)	
Democrat	0.129	1.160***	-1.612*	0.663	1.741***	-1.057	.
	(0.581)	(0.042)	(0.831)	(0.702)	(0.588)	(0.920)	
Republican	1.598***	-0.309	-1.375*	2.003***	0.102	-0.939	.
	(0.472)	(0.460)	(0.717)	(0.600)	(0.601)	(0.805)	
CanvassingXDem	-1.443**	-0.330	0.217	-1.462*	-0.346	0.164	.
	(0.674)	(0.496)	(0.975)	(0.828)	(0.702)	(1.085)	
CanvassingXRep	-1.153**	-0.942*	0.336	-1.529**	-1.348*	-0.064	.
	(0.535)	(0.572)	(0.838)	(0.694)	(0.735)	(0.946)	
PhoneXDem	-0.651	-0.524	0.817	-1.090	1.006	0.374	.
	(0.639)	(0.458)	(0.891)	(0.773)	(0.655)	(0.997)	
PhoneXRep	-0.989*	-1.064**	0.368	-1.551**	-1.671**	-0.194	.
	(0.509)	(0.513)	(0.779)	(0.676)	(0.678)	(0.692)	
Vote 2004	2.068***	1.732***	3.004***	1.195*	0.873	2.141**	2.410***
	(0.245)	(0.222)	(0.721)	(0.628)	(0.619)	(0.918)	(0.389)
Age	0.025***	0.030***	0.019***	0.083***	0.093***	0.077***	-0.036**
	(0.004)	(0.004)	(0.006)	(0.015)	(0.016)	(0.016)	(0.016)
Female	-0.154	-0.400***	0.469**	-0.434**	-0.779***	0.136	0.203
	(0.128)	(0.131)	(0.229)	(0.202)	(0.213)	(0.277)	(0.238)
House Size	0.142**	0.121*	0.038	0.208*	0.179	0.095	0.081
	(0.066)	(0.066)	(0.125)	(0.114)	(0.115)	(0.155)	(0.122)
District 156	0.208*	0.280**	-0.078	-0.203	-0.087	-0.473*	0.641***
	(0.123)	(0.125)	(0.203)	(0.221)	(0.223)	(0.271)	(0.216)
Constant	-4.677***	-3.796***	-5.154***	-6.004***	-5.361***	-6.435***	1.055
	(0.534)	(0.462)	(0.912)	(0.777)	(0.742)	(1.071)	(1.340)

Note: $N = 1,998$. D = vote for Democratic candidate, $R/0$ = vote for Republican or other candidate, NR = nonresponse, *Abstain* = did not vote. Standard error in parentheses *** indicates $p < .01$; ** indicates $p < .05$; * indicates $p < .10$

We next present the materials corresponding to our replication of Campbell and Monson (2008). Variable descriptions are presented immediately below, again followed by the full table of BIMNL and MNL estimates.

Independent and Control Variable Operationalizations from Campbell and Monson (2008)

- *Ideology*: Responses to the question: do you consider yourself generally liberal, moderate, or conservative? 1 = Liberal, 2 = Moderate, 3 = Conservative
- *Female*: Binary indicator of whether a respondent was female or male
- *African American*: Binary indicator of whether a respondent described themselves as African American

- *Hispanic*: Binary indicator of whether respondent described themselves as Hispanic
- *Party Identification*: 1 = Strong Republican, 2 = Weak Republican, 3 = Independent-leaning Republican, 4 = Pure Independent, 5 = Independent-leaning Democrat, 6 = Weak Democrat, 7 = Strong Democrat
- *Shared Values W/Bush*: Responses to the question: "Think of George W. Bush. In your opinion, does the phrase 'shares my values' describe George W. Bush extremely well, quite well, not too well, or not well at all?"
- *Presidential Battleground State*: Binary indicator of whether or not respondent lives in a presidential battleground state
- *South*: Binary indicator of whether or not a respondent lives in a Southern state
- *GMB*: A binary indicator of whether or not a state had a gay marriage ban on the ballot in 2004
- *White Evangelical* : Binary indicator of whether respondent indicated that they considered themselves to be a born again Christian or an evangelical Christian, and racially identified as white.
- *White Evangelical X GMB*: Interaction of *White Evangelical* and *GMB*
- *Catholic*: Binary indicator of whether respondent selected that they considered themselves to be a Catholic
- *Catholic X GMB*: Interaction of *Catholic* and *GMB*
- *Other Religion*: Binary indicator of whether respondent selected that they considered themselves to be "other" from multiple choice religious preference question
- *Secular*: Binary indicator of whether respondent selected that they considered themselves to be "No preference/no religious affiliation" in response to multiple choice religious preference question
- *Secular X GMB*: Interaction of *Secular* and *GMB*
- *Education*: Ordinal indicator of highest level of education completed, where 1 = elementary school only, 2 = some high school, 3 = completed high school, 4 = some college, 5 = two-year college degree, 6 = four-year college degree, 7 = some graduate work, 8 = completed masters or professional degree, 9 = advanced graduate work or Ph.D.
- *Age*: Ordinal indicator of age cohort, where 1 = 18-29, 2 = 30-54, 3 = 55 and over
- *Mobilization Index*: Additive index of responses to questions of, whether during the previous campaign the respondent (1) Received a letter or mail piece from a campaign, (2) Received a request to donate money to a campaign, (3) Had a face-to-face conservation or contact with someone from a campaign, (5) Received an email from a campaign, or (6) Heard a radio ad from a campaign.

Table A.4: 2004 Presidential Vote Choice

	MNL		BIMNL		
	Kerry vs. Abstain	Bush vs. Abstain	Kerry vs. Abstain	Bush vs. Abstain	Inflation Stage
Ideology	-0.271	0.067	-0.349	0.004	.
	(0.186)	(0.196)	(0.233)	(0.249)	
Female	-0.614***	-0.311	-0.667**	-0.378	.
	(0.257)	(0.267)	(0.313)	(0.331)	
African American	0.540	-1.542	0.506	-1.398	.
	(0.591)	(0.847)	(0.786)	(1.103)	
Hispanic	-0.956	-1.298	-1.274*	-1.715*	.
	(0.721)	(0.868)	(0.741)	(0.908)	
Party Identification	0.344***	-0.386***	0.344***	-0.390***	.
	(0.075)	(0.076)	(0.089)	(0.092)	
Shared Values W/Bush	-0.514***	1.139***	-0.391***	1.343***	.
	(0.155)	(0.170)	(0.194)	(0.216)	
Presidential Battleground State	-0.092	-0.335	-0.126	-0.455	.
	(0.276)	(0.286)	(0.335)	(0.356)	
South	0.133	0.432	0.590	1.008**	.
	(0.328)	(0.289)	(0.480)	(0.500)	
GMB	0.499	0.372	0.653	0.609	.
	(0.410)	(0.422)	(0.548)	(0.584)	
White Evangelical	-0.899	-0.651	-0.556	-0.321	.
	(0.621)	(.608)	(0.911)	(0.886)	
White Evangelical X GMB	0.048	0.545	-0.671	-0.230	.
	(0.991)	(0.967)	(1.324)	(1.282)	
Catholic	-0.487	-0.455	-0.594	-0.527	.
	(0.533)	(0.554)	(0.626)	(.659)	
Catholic X GMB	-0.398	0.052	0.369	0.829	.
	(0.800)	(0.801)	(1.210)	(1.245)	
Other Religion	-1.329***	1.431***	-1.442***	-1.536***	.
	(0.800)	(0.476)	(0.536)	(0.570)	
Secular	-1.113**	-0.666	-0.670	-0.095	-1.295*
	(0.553)	(0.612)	(0.786)	(0.895)	(0.789)
Secular X GMB	-0.694	-1.468*	-1.181	-2.121	.
	(0.689)	(0.796)	(0.937)	(1.125)	
Education	0.292***	0.208***	0.076	-0.033	1.507***
	(0.074)	(0.078)	(0.100)	(0.107)	(0.528)
Age	0.888***	0.572***	0.850***	0.649***	0.743*
	(0.621)	(0.202)	(0.233)	(0.251)	(0.437)
Mobilization Index	0.785***	0.776***	0.666***	0.656***	1.139***
	(0.118)	(0.123)	(0.149)	(0.162)	(0.313)
Constant	-2.189**	-3.018***	-0.724	-1.632	-4.938**
	(0.996)	(1.061)	(1.253)	(1.322)	(1.985)

Note: $N = 1,341$. Standard error in parentheses *** indicates $p < .01$; ** indicates $p < .05$; * indicates $p < .10$

Table A.5: Multinomial Choice Estimator Usage, 2009-2013

Citation	MNL Used	CL Used	MNP Used
Arceneaux and Kolodny (2009)	✓		
Wright (2009)	✓		
Fukumoto (2009)	✓		
Feddersen et al. (2009)	✓		✓
Brown and Mobarak (2009)	✓		
Prior (2009)	✓		
Scott and Bornstein (2009)	✓		
Conrad and Moore (2010)	✓		
Eifert et al. (2010)	✓		
Stasavage (2010)	✓		
Kalyvas and Balcells (2010)	✓		
Duch et al. (2010)		✓	
Bertelli and John (2010)	✓		
Gent and Shannon (2010)	✓		
Kam and Simas (2010)	✓		
Tir (2010)	✓		
Gehlbach and Malesky (2010)	✓		
Greene (2011)	✓		
Campbell et al. (2011)	✓		
Huth et al. (2011)	✓		
Croco (2011)	✓		
Liu (2011)	✓		
Benson (2011)	✓		
Thomson (2011)	✓		
Panagopoulos (2011)	✓		
Kam and Kinder (2012)	✓		
Kayser and Peress (2012)		✓	
Sinclair et al. (2012)	✓		
Benmelech et al. (2012)	✓		
Noel (2012)	✓		
Curini and Hino (2012)		✓	
Gerber et al. (2013)	✓		
Kardesheva (2013)	✓		
Blinder et al. (2013)	✓		
Lupu (2013)	✓		
Dancey and Sheagley (2013)	✓		✓
Hart (2013)	✓		
Ansell and johannes Lindvall (2013)	✓		
Panagopoulos (2013)	✓		
Stanton (2013)	✓		
Utych and Kam (2013)	✓		
Potter and Baum (2013)	✓		

Note: Sample includes articles appearing in the *APSR*, *AJPS*, and *JOP*. Usage applies to main or robustness analysis. MNL = "Multinomial Logit", CL = "Conditional Logit", MNP = "Multinomial Probit".

REFERENCES

Ansell, B. and johannes Lindvall (2013). The political origins of primary education systems: Ideology, institutions, and interdenominational conflict in an era of nation building. *American Political Science Review 107*(3), 505–522.

Arceneaux, K. and R. Kolodny (2009). Educating the least informed: Group endorsements in a grassroots campaign. *American Journal of Political Science 53*(4), 755–770.

Bagozzi, B. E., D. W. J. Hill, W. H. Moore, and B. Mukherjee (2014). Modeling two types of peace: The zero-inflated ordered probit (ziop) model in conflict research. *Journal of Conflict Resolution*.

Bagozzi, B. E. and B. Mukherjee (2012). A mixture model for middle category inflation in ordered survey responses. *Political Analysis 20*(3), 369–386.

Benmelech, E., C. Berrebi, and E. F. Klor (2012). Economic conditions and the quality of suicide terrorism. *Journal of Politics 74*(1), 113–128.

Benson, B. V. (2011). Unpacking alliances: Deterrent and compellent alliances and their relationship with conflict, 1816-2000. *Journal of Politics 73*(4), 1111–1127.

Bertelli, A. M. and P. John (2010). Government checking government: How performance measures expand distributive politics. *Journal of Politics 72*(2), 545–558.

Blinder, S., R. Ford, and E. Ivarsflaten (2013). The better angels of our nature: How the antiprejudice norm affects policy and party preferences in great britain and germany. *American Journal of Political Science 57*(4), 841–857.

Brown, D. S. and A. M. Mobarak (2009). The transforming power of democracy: Regime type and the distribution of electricity. *American Political Science Review 103*(2), 193–213.

Campbell, D. E., J. C. Green, and G. C. Layman (2011). The party faithful: Partisan images, candidate religion and the electoral impact of party identification. *American Journal of Political Science 55*(1), 42–58.

Campbell, D. E. and J. Q. Monson (2008). The religion card: Gay marriage and the 2004 presidential election. *Public Opinion Quarterly 72*(3), 399–419.

Conrad, C. R. and W. H. Moore (2010). What stops the torture? *American Journal of Political Science 54*(2), 459–467.

Croco, S. E. (2011). The decider's dilemma: Leader culpability, war outcomes, and domestic punishment. *American Political Science Review 105*(3), 457–477.

Curini, L. and A. Hino (2012). Missing links in party-system polarization: How institutions and voters matter. *Journal of Politics 74*(2), 460–473.

Dancey, L. and G. Sheagley (2013). Heuristics behaving badly: Party cues and voter knowledge. *American Journal of Political Science 57*(2), 312–325.

Duch, R. M., J. May, and D. A. Armstrong (2010). Coalition-directed voting in multiparty democracies. *American Political Science Review 104*(4), 678–719.

Eifert, B., E. Miguel, and D. N. Posner (2010). Political competition and ethnic identification in africa. *American Journal of Political Science 54*(2), 494–510.

Feddersen, T., S. Gailmard, and A. Sandroni (2009). Moral bias in large elections: Theory and experimental evidence. *American Political Science Review 103*(2), 175–192.

Fukumoto, K. (2009). Systematically dependent competing risks and strategic retirement. *American Journal of Political Science 53*(3), 740–754.

Gehlbach, S. and E. J. Malesky (2010). The contribution of veto players to economic reform. *Journal of Politics 72*(4), 957–975.

Gent, S. E. and M. Shannon (2010). The effectiveness of international arbitration and adjudication: Getting into a bind. *Journal of Politics 72*(2), 366–280.

Gerber, A. S., G. A. Huber, D. Doherty, C. M. Dowling, and S. J. Hill (2013). Do perceptions of ballot secrecy influence turnout? results from a field experiment. *American Journal of Political Science 57*(3), 537–551.

Greene, K. F. (2011). Campaign persuasion and nascent partisanship in mexico's new democracy. *American Journal of Political Science 55*(2), 398–416.

Harris, M. N. and X. Zhao (2007). A zero-inflated ordered probit model, with an application to modelling tobacco consumption. *Journal of Econometrics 141*(2), 1073–1099.

Hart, A. (2013). Can candidates activate or deactivate the economic vote? evidence from two mexican elections. *Journal of Politics 75*(4), 1051–1063.

Huth, P. K., S. E. Croco, and B. J. Appel (2011). Does international law promote the peaceful settlement of international disputes? evidence from the study of territorial conflicts since 1945. *American Political Science Review 105*(2), 415–436.

Kalyvas, S. N. and L. Balcells (2010). International system and technologies of rebellion: How the end of the cold war shaped international conflict. *American Political Science Review 104*(4), 414–429.

Kam, C. D. and D. R. Kinder (2012). Ethnocentrism as a short-term force in the 2008 american presidential election. *American Journal of Political Science 56*(2), 326–340.

Kam, C. D. and E. N. Simas (2010). Risk orientations and policy frames. *Journal of Politics 72*(2), 381–396.

Kardesheva, R. (2013). Package deals in eu legislative politics. *American Journal of Political Science 57*(4), 858–874.

Kayser, M. A. and M. Peress (2012). Benchmarking across borders: Electoral accountability and the necessity of comparison. *American Political Science Review 106*(3), 661–684.

Liu, A. H. (2011). Linguistic effects of political institutions. *Journal of Politics 73*(1), 125–139.

Lupu, N. (2013). Party brands and partisanship: Theory with evidence from a survey experiment in argentina. *American Journal of Political Science 57*(1), 49–64.

Noel, H. (2012). The coalition merchants: The ideological roots of the civil rights realignment. *Journal of Politics 74*(1), 156–173.

Panagopoulos, C. (2011). Thank you for voting: Gratitude expression and voter mobilization. *Journal of Politics 73*(3), 707–717.

Panagopoulos, C. (2013). Extrinsic rewards, intrisic moviation and voting. *Journal of Politics 75*(1), 266–280.

Potter, P. B. and M. A. Baum (2013). Looking for audience costs in all the wrong places: Electoral institutions, media access, and democratic constraint. *Journal of Politics 76*(1), 167–181.

Prior, M. (2009). Improving media effects research throguh better measurement of news exposure. *Journal of Politics 71*(3), 893–908.

Scott, J. T. and B. H. Bornstein (2009). What's fair in foul weather and fair? distributive justice across different allocation contexts and goods. *Journal of Politics 71*(3), 831–846.

Sinclair, B., M. McConnell, and D. P. Green (2012). Spillover effects: Design and analysis of multilevel experiments. *American Political Science Review 56*(4), 1055–1069.

Stanton, J. A. (2013). Terrorism in the context of civil war. *Journal of Politics 75*(4), 1009–1022.

Stasavage, D. (2010). When distance mattered: Geographic scale and the development of euroepean representative assemblies. *American Political Science Review 104*(4), 625–643.

Thomson, R. (2011). Citizens' evaluations of the fulfillment of election pledges: Evidence from ireland. *Journal of Politics 73*(1), 187–201.

Tir, J. (2010). Territorial diversion: Diversionary theory of war and territorial conflict. *Journal of Politics 72*(2), 413–425.

Utych, S. M. and C. D. Kam (2013). Viability, information seeking, and vote choice. *Journal of Politics 76*(1), 152–166.

Wright, J. (2009). How foreign aid can foster democratization in authoritarian regimes. *American Journal of Political Science 53*(3), 552–571.