Online Appendix to "Mediating the Electoral Connection: The Information Effects of Voter Signals on Legislative Behavior"

John Henderson[*]

Assistant Professor

Dept. of Political Science

Yale University

John Brooks[†]

Post-Doctoral Associate

Social Science Research Institute

Duke University

April 15, 2016

[*]john.henderson@yale.edu, http://jahenderson.com/, Institution for Social and Policy Studies, Yale University, 77 Prospect Street, New Haven, CT 06511

[†]john.brooks@berkeley.edu,http://www.johnebrooks.com/, Social Science Research Institute, Duke University, Durham, NC 27708

I Introduction

This supplemental appendix to the paper "Mediating the Electoral Connection: The Information Effects of Voter Signals on Legislative Behavior" is intended for online publication. We first include tables that present our stratification results in greater detail than reported in the main draft of the paper. In the next section, we discuss the imputation approach taken in the paper, and some of the core instrumental variable (IV) diagnostics we employ. We then present descriptive statistics for the full list of covariates used throughout analysis in the paper, histograms of the two rain instruments, and additional descriptive and robustness results.

II Tables for Main Stratification Results

In the main draft, we discuss IV results stratifying by electoral security, partisanship and seniority to assess heterogeneity in incumbent adaptation. We present tables for these results here. Table I reports results for Safe and Competitive districts. Table II reports stratifications for Democratic and Republican incumbents. And seniority stratifications for Early Incumbency and Late Incumbency.

Table I: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning, by District Competitiveness

	Competitive		Safe	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day$_t$ (3)	Prior Weekend$_t$ (4)
Dem. Vote Margin$_t$	-1.751	-1.575	-0.714	-0.787
	(0.877)*	(0.745)*	(0.500)	(0.500)
Observations	1318	1318	4918	4918
Clusters	645	645	1485	1485
R^2	0.845	0.855	0.900	0.899

Specifications are 2SLS with additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table II: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning, by Incumbent Party

	Democratic		Republican	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day$_t$ (3)	Prior Weekend$_t$ (4)
Dem. Vote Margin$_t$	-1.101	-1.235	-1.559	-0.393
	(0.528)*	(0.538)*	(1.346)	(0.969)
Observations	3237	3237	3036	3036
Clusters	823	823	816	816
R^2	0.769	0.760	0.735	0.805

Specifications are 2SLS with *Incumbent* fixed-effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table III: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning, by Number of Terms in Office

	Early Incumbency		Late Incumbency	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day$_t$ (3)	Prior Weekend$_t$ (4)
Dem. Vote Margin$_t$	-1.018	-0.775	0.115	-1.220
	(0.424)*	(0.353)*	(1.113)	(1.574)
Observations	4523	4523	1714	1714
Clusters	1474	1474	567	567
R^2	0.896	0.902	0.892	0.870

Specifications are 2SLS with additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

III Multiple Imputation through Chained Equations (MICE)

To handle missing values in the control variables in our analysis we use multiple imputation.[1] Our specific approach is Multiple Imputation through Chained Equations (MICE) made available in the 'mice' package in **R** (van Buuren and Groothuis-Oudshoorn 2011). This approach estimates missing values on the precise scale of the data (i.e., continuous, binary), conditioning later imputations on previous ones under a Markov

[1] Note, we do not impute any missing values for the instruments or outcome variables in the main results in our study.

chain of constrained models (van Buuren and Groothuis-Oudshoorn 2011). To model the missingness process, we include the full list of controls and variables from Table VII. This imputation assumes that missingness is essentially random given information in $X_{it,d}$, a relatively weak assumption in this case. Except for competitiveness, the amount of imputed values is negligible (and essentially random). For competitiveness, nearly all missing values are from six whole years (1956 – 1962, 1970 and 1972), rather than across the time series. Cross-sectional district characteristics *cannot* correlate with this missingness, though year-to-year factors (e.g., president's party) might be relevant. We also find that missingness on the competitiveness variable does not correlate with any election or roll call outcomes at t and $t-1$. Imputed values for competitiveness are noisier than un-imputed values. However, these remain good predictors of non-missing data at $\rho^I = 0.662$ ($p < 0.001$) compared to $\rho^C = 0.897$ ($p < 0.001$) for complete cases. We also find that missingness on the (normalized) prior presidential preference measure is uncorrelated with roll call and election outcomes, and the amount of imputation here is small anyway at 3.8% of the data.

To produce our main data, we impute five values for each missing item, and average over these to produce the final imputed dataset. We find no difference between combining the imputations and then estimating the models, or estimating the models on the imputations and then combining the estimates. We also run additional models (a) imputing *all* missing values for the covariates, instruments and outcomes, and (b) imputing *no* missing values for any variables. These are presented below in Section VI. Our results remain robust regardless of whether none, some, or all of the missing data is imputed.

A final issue is that 484 incumbents lose reelection in our sample. We cannot observe the roll-call positions of these unelected members in the next Congress. Excluding these incumbents could attenuate our results by possibly conditioning on a post-treatment variable.[2] For the IV ratio estimator, attenuation in the denominator would be a con-

[2] Though importantly here, we show election rain does not influence incumbent defeat in our data.

cern. Our solution is to impute missing ideal points for losing incumbents by using their estimates from the previous Congress. Rainfall today cannot affect legislators' prior positions. This imputation assumes that losing incumbents are "standpat" on policy even in the face of defeat. The bias from this imputation can only be conservative, attenuating any effect rain has on repositioning. Yet, this will not bias the effect of rain on vote margins in the IV denominator. As such, our resulting IV estimates should be a *lower bound* on the true repositioning effect. More generally, we motivate this choice with a thought experiment. If incumbents could rerun the previous Congress knowing their eventual electoral outcome (i.e., are "oracles"), losers should be the *most* likely to change positions to increase their odds of survival. But, if incumbents were both oracles and standpats, they would choose to retire rather than run costly and ultimately unsuccessful campaigns. One might worry about retirement, but this would be uncorrelated with election rain, unless politicians had a forecasting ability eluding both satellites and meteorologists.

IV Assumptions of IV

Instrumental variables analyses typically make five assumptions: *exogeneity*, *inclusion*, *exclusion*, *'compliance' monotonicity*, and *SUTVA*. In this section, we provide additional details on the way we assess the first three core IV assumptions in the main analysis. The latter two are assumptions about counterfactuals, and do not have testable empirical implications. Hence we do not discuss these.

IV.1 Exogeneity

From our paper, we estimate this system of equations in a two-stage model:

$$D_{it,d} = \rho_0 + \rho_1 Z^{(j)}_{it,d} + \rho_2 Y_{it-1,d} + \rho_3 D_{it-1,d} + \kappa X_{it,d} + \eta_{i,d} + \nu \quad (1)$$

$$Y_{it,d} = \beta_0 + \beta_1 D_{it,d} + \beta_2 Y_{it-1,d} + \beta_3 D_{it-1,d} + \theta X_{it,d} + \delta_{i,d} + \epsilon, \quad (2)$$

for the ith incumbent in year t, in decade d. Ideal points at year t are $Y_{it,d}$ and prior ideal points at $t-1$ are $Y_{it-1,d}$. Election vote margins at year t and $t-1$ are $D_{it,d}$ and $D_{it-1,d}$, respectively. The models include relevant controls in $X_{it,d}$, and either *District* $\times$ *Decade* or *Incumbent* fixed effects, indicated by $\delta_{i,d}$ and $\eta_{i,d}$. $Z^{(j)}_{it,d}$ is each rain instrument, where j indicates prior weekend ($j = PW$) or election day ($j = ED$) rain.

Table IV: Placebo Test: Effect of Election Rainfall on Previous Democratic Vote Margin for House Incumbents, Using *District* $\times$ *Decade* Units

	Ordinary Least Squares			
	Dem. Vote Margin$_{t-1}$		Inc. Ideal Point$_{t-1}$	
	(1)	(2)	(3)	(4)
Election Day Rain$_t$	0.006		-0.008	
	(0.005)		(0.011)	
Prior Weekend Rain$_t$		0.006		-0.007
		(0.005)		(0.012)
Observations	7139	7139	7056	7056
Clusters	2920	2920	2883	2883
R^2	0.892	0.892	0.940	0.940

Specifications are OLS with *Decade* $\times$ *District* fixed effects, and additional state, year, and district controls. *Decade* $\times$ *District* cluster standard errors are in parentheses.

Placebo Test

This model requires that $E[Z^{(j)}_{it,d}\nu] = 0$, or that no omitted factor correlating with rain influences vote margins. One way to assess this exogeneity assumption is through a placebo, estimating the effects of rain at election t on *prior* Democratic vote margins and legislative positions for incumbents at $t-1$. Since rain cannot influence prior election or roll call outcomes, any correlation in these tests would indicate that the instruments are likely confounded by some common prior factor. For this placebo we estimate the following model:

$$D_{it-1,d} = \delta_0 + \delta_1 Z^{(j)}_{it,d} + \kappa X_{it,d} + n_{i,d} + \upsilon \quad (3)$$

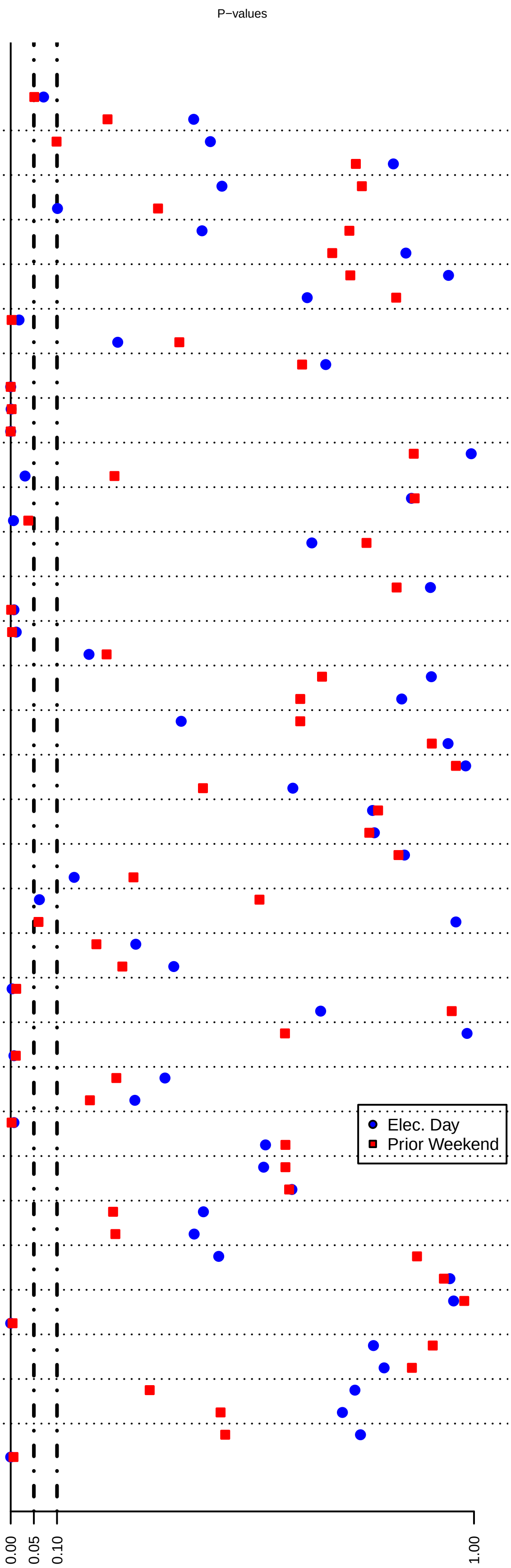
$$Y_{it-1,d} = \alpha_0 + \alpha_1 Z^{(j)}_{it,d} + \theta X_{it,d} + d_{i,d} + e, \quad (4)$$

We include the relevant controls in $X_{it,d}$ from the full list in Table VII below. We also include *Incumbent* or *District* $\times$ *Decade* fixed effects and cluster standard errors in these models to verify that our full models are unconfounded, regardless of how we isolate district variation. We present the results from this placebo tests for *Incumbent* units in Table 1 in the main draft. Importantly, we find here that election rain is uncorrelated with prior Democratic vote and incumbent ideology. We confirm these results in a *District* $\times$ *Decade* placebo test presented here in Table IV. Note, we include covariates and fixed effects throughout these tests to make weaker assumptions about the association between rainfall and districts, and cluster to appropriately measure error variance. Ultimately, we find the same results whether including or excluding controls.

In addition to these placebo tests, we also report the results from a series of balance tests. These regress election rain on a fuller model of 62 covariates that originate from the prior two election cycles ($t-2$ to t) in our data. Significant effects here could suggest that rain is correlated with some factors that may also influence electoral or legislative behavior. We also include previous rain measures to assess if past rain influences future weather. The results for these tests are displayed in balance plots in Figure I. As shown, the rain instruments are uncorrelated with the vast majority ($\approx 80\%$) of the covariates, including prior rain measures. Further, the pattern of association here appears idiosyncratic, mitigating the concern that election rain is confounded. In combination, both tests show that rainfall is not likely confounded with unobserved factors driving incumbent positions and election outcomes, conditional on included district-level factors.

Figure I: Balance Tests Assessing Election Rain Instrument Exogeneity

	Elec. Day T−statistic	Prior Weekend T−statistic
(Undeviated) Election Rain$_{t-1}$	1.806	1.956
(Deviated) Election Rain$_{t-1}$	−0.852	−1.257
Dem. Incumbent$_t$	0.787	1.65
Dem. Incumbent$_{t-1}$	0.22	−0.325
Dem. Incumbent$_{t-2}$	−0.746	−0.308
District Ideology$_{t-1}$	−1.639	−0.998
District Ideology$_{t-2}$	0.819	0.344
District Ideology$_{t-3}$	0.185	0.394
Competitiveness$_t$	−0.069	−0.341
Competitiveness$_{t-1}$	−0.468	−0.212
Competitiveness$_{t-2}$	−2.368	−3.047
Redistricted$_t$	1.198	0.908
Redistricted$_{t-1}$	0.413	0.484
Redistricted$_{t-2}$	−5.756	−6.087
Dem. Sec. of State$_t$	3.23	3.107
Dem. Sec. of State$_{t-1}$	−3.52	−3.846
Dem. Sec. of State$_{t-2}$	0.007	0.164
Dem. Governor$_t$	2.156	1.218
Dem. Governor$_{t-1}$	−0.17	0.161
Dem. Governor$_{t-2}$	2.763	2.08
Midterm Election$_t$	−0.454	−0.294
Midterm Election$_{t-1}$		
Midterm Election$_{t-2}$	0.118	−0.211
Presidential Co−Partisan$_t$	−2.72	−3.246
Presidential Co−Partisan$_{t-1}$	−2.502	−2.977
Presidential Co−Partisan$_{t-2}$	1.377	1.263
Median Income Per Capita$_t$	0.116	0.424
Median Income Per Capita$_{t-1}$	−0.197	−0.489
Median Income Per Capita$_{t-2}$	0.9	0.489
Log Population$_t$	0.07	0.114
Log Population$_{t-1}$	−0.022	−0.049
Log Population$_{t-2}$	−0.511	−0.816
% Black$_t$	−0.278	−0.262
% Black$_{t-1}$	0.273	0.288
% Black$_{t-2}$	0.189	−0.206
% Construction$_t$	1.488	1.114
% Construction$_{t-1}$	−1.87	−0.618
% Construction$_{t-2}$	0.048	−1.881
% Education$_t$	−1.102	−1.327
% Education$_{t-1}$	0.931	1.172
% Education$_{t-2}$	2.997	2.518
% Farmer$_t$	−0.428	0.06
% Farmer$_{t-1}$	0.019	−0.535
% Farmer$_{t-2}$	2.72	2.539
% Foreign Born$_t$	0.968	1.205
% Foreign Born$_{t-1}$	−1.108	−1.371
% Foreign Born$_{t-2}$	2.697	3.12
% Govt. Worker$_t$	−0.598	−0.534
% Govt. Worker$_{t-1}$	0.604	0.534
% Govt. Worker$_{t-2}$	0.514	0.523
% Manufacturing$_t$	0.813	1.223
% Manufacturing$_{t-1}$	−0.849	−1.211
% Manufacturing$_{t-2}$	0.758	0.155
% Unemployed$_t$	0.065	0.082
% Unemployed$_{t-1}$	−0.056	−0.026
% Unemployed$_{t-2}$	4.075	2.902
% Urban$_t$	−0.276	0.111
% Urban$_{t-1}$	0.245	−0.169
% Urban$_{t-2}$	0.328	1.036
% Retail Worker$_t$	−0.364	−0.75
% Retail Worker$_{t-1}$	0.313	0.735
% Retail Worker$_{t-2}$	3.664	2.724



Overidentification Test

We also present the results here of a Sargan overidentification test of our IV models. With k regressors and $r \geq k$ instruments, 2SLS estimators 'reduce' the orthogonal information in r to just that contained in k instruments. When $r > k$, the IV model is overidentified, which means there are $r - k$ additional restrictions left out in identifying $\hat{\beta}_{IV}$. A feature of such overidentification is that the residuals produced from 2SLS with $r > k$ instruments can be regressed on each instrument, something obviously not possible if the model is exactly identified $r = k$. In doing so, we are interested in observing whether the residuals from the 2SLS models are orthogonal to the instruments. If these are orthogonal, then we have additional reason to conclude the exogeneity assumption holds for the rain instruments. If these are not orthogonal, then we should be concerned that some omitted variation in $Y_{it,d}$ and $D_{it,d}$ is systematically correlated with rain $Z^{(j)}_{it,d}$.

The results of these tests indicate that both instruments are uncorrelated with any remaining systematic variation in the outcomes. The hypothesis test is against the null of no association between the IV residuals and the instruments (where overidentification stems from instruments at t and $t - 1$). The Sargan p-value for this test is 0.261 for the Election Day Rain instrument and 0.766 for Previous Weekend Rain. Both of these indicate that the instruments are valid under exogeneity.

IV.2 Inclusion

A frequent issue in IV analyses, including those using rain, is the issue of weak instruments. If $\hat{\rho}_1$ in the regression model above from Equation (1) is small, then the $\hat{\beta}_{IV}$ estimate for β_1 will be inconsistent and have a non-standard asymptotic distribution. While often a concern, this assumption is also testable. In the main paper we provide ample evidence that rain is an informative instrument in this analysis, and thus the IV estimates are consistent to identify the effects of shifts in vote margins on ideal-point positions. Here we provide more detail on these tests.

Adding Controls to Strengthen Instruments

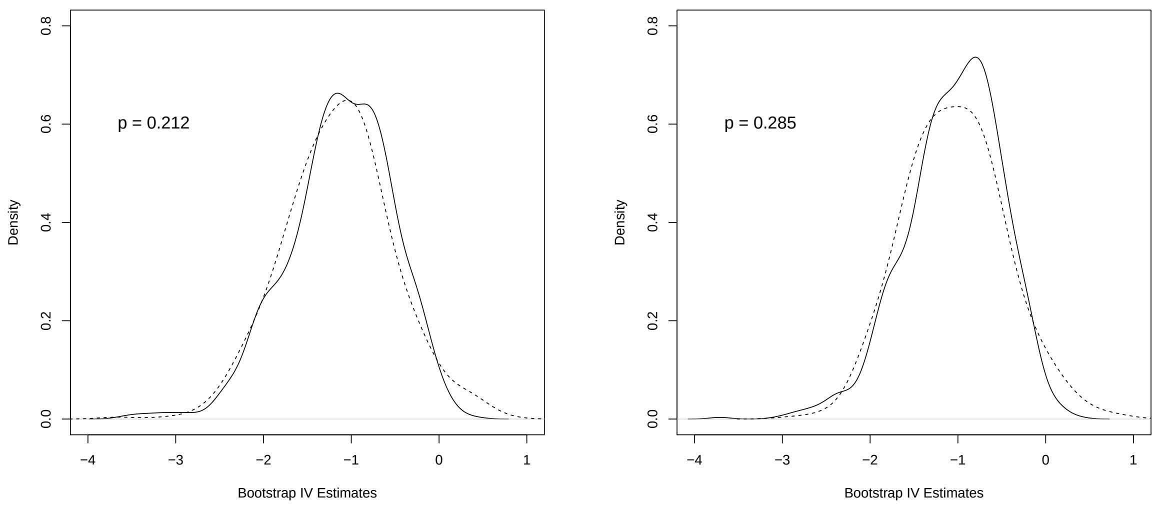
Conditioning on (pre-election-rain) covariates in modeling Equation (1) naturally reduces the imprecision of the estimates of ρ_1. Any covariates correlated with $Y_{it,d}$ and $D_{it,d}$, reduces unexplained variance in those outcomes (under the model linearity assumptions). Stated differently, variance in excluded covariates will propagate into additional variance in the outcomes and uncertainty into estimates since that information is not being included in the regression. In increasing precision (without introducing bias by including endogenous regressors), we can reduce bias in the estimates $\hat{\beta}_{IV}$ *even if* ρ_1 *is small.* For a similar reason (increasing sampling precision), increasing sample size in N (until $\rho_1 > 1/\sqrt{N}$) shrinks weak instruments' bias to zero for any given ρ_1.

F-Tests and Normality Tests of Instrument Strength

A common way to evaluate instrument strength is to inspect (un)conditional F-tests from a hypothesis test that the instruments are all zero. Stock and Yogo (2005) develop standard tests for weak instruments, identifying critical values for F-statistics that are associated with various levels of bias tolerated by researchers. For two instruments and one endogenous regressor (i.e., $D_{it,d}$), F-statistics greater than 19.93 indicate that first-stage effects are sufficiently influential (i.e., informative) for IV analysis. At this critical value, the rate of Type I Error is no more than 10%, for $\alpha = 0.05$ under a standard test of 2SLS coefficients under the no-effects null hypothesis. For these tests, we estimate cluster F-statistics, using the full model of covariates and fixed-effects. The cluster F-tests use Kleibergen-Paap rk Wald statistics (Kleibergen and Paap 2006). This statistic is equivalent to the standard Cragg-Donald Wald statistic when errors are i.i.d., but provide a correction for robust estimation of cluster errors otherwise. As discussed in the results in the main draft, each of our models surpasses (or nearly surpasses) this F-test threshold, indicating our instruments (and in particular using *Incumbent* units) are sufficiently informative for IV analysis.

This F-test approach is related to another test, following from the asymptotic properties of IV estimators. Given a sufficiently large F-statistic, the asymptotic distribution of $\hat{\beta}_{IV}^{(s)}$ from s repeated samples, sampled infinitely, is distributed normally, as $\beta_{IV} \sim N(\beta_1, \sigma^2_{\beta_1})$. An implication of this finding is that, if under s finite repeated samples of N districts, $\hat{\beta}_{IV}^{(s)}$ is not (approximately) normally distributed, then an associated F-statistic will be too small, since the precision of the first stage effect is too little. In such a case, either N or the first stage IV effect needs to be larger for its use as a valid instrument.

Figure II: Normality Tests Under Bootstrap Sampling IV Estimates



(a) Election Day Rain_t (b) Prior Weekend Rain_t

While we cannot repeatedly sample districts, we can bootstrap sample from the population of districts to produce a distribution of IV estimates. To do so, we fix N to be the number of districts in our main IV analysis (i.e., $N = 6237$), and we repeatedly sample $s = 1000$ each time sampling 6237 units with replacement. We estimate 1000 IV coefficients, $\vec{\beta}_{IV}$, for each rain instrument using the model in Equations (1) and (2), including covariates and *Incumbent* fixed effects. We then compare this surface of estimates to a hypothetical normal distribution centered at $E[\hat{\beta}_{IV}]$ with variance $V[\hat{\beta}_{IV}]$. If the quantiles of these two distributions are different, we conclude that the distributions $\vec{\beta}_{IV}$, are not normal, and the instruments could be weak.

Figure II presents the distributions of $\vec{\beta}_{IV}$ for Election Day Rain in II(a) and Previous Weekend Rain in II(b). The solid lines are densities of the distribution of bootstrap IV estimates, and the dashed lines are hypothetical normal distributions with the same moments as the bootstrap sample estimates. Clearly there is considerable overlap in the distributions, though both the bootstrap estimate densities have slightly longer left tails. This suggests some of the bootstrap sample first-stage estimates may have been small enough to inflate a small proportion of the IV estimates. To test similarity statistically, we use a *ks*-test. Such a test evaluates the null hypothesis that the quantiles of two distributions are the same, and in this case that the quantiles of $\vec{\beta}_{IV}$ suggest normality. The p-values for these tests are 0.212 (Election Day Rain) and 0.285 (Previous Weekend Rain) again suggesting that the instruments are sufficiently informative (and not weak).

IV.3 Exclusion

A final assumption in IV analysis is that the influence of rain on incumbent voting can only occur through its influence on election outcomes. This assumption is violated, for example, if election rain directly influences behaviors or attitudes outside of voting (e.g., if rain on election day changes beliefs about the plausibility of human caused climate or weather events). Generally, this assumption is untestable since it involves counterfactual quantities (i.e., the effect of rain on vote margins in the presence *and* the absence of its effect on some other factor). However, we think the exclusion restriction assumption is very plausible here. Rain on Election Day *immediately* impacts voting before it can impact virtually any alternative path to representation. Nearly all possible alternative mechanisms can only occur following rain's impact on elections. Second, the effects of rain (especially over just a couple days) are unlikely to be directly influential since rain is idiosyncratic, and modest in variation and impact. Whatever possible direct effects it

Table V: Effect of (Unnormalized) Election Rainfall on Democratic Vote Margin and Incumbent Reelection

	Democratic Vote Margin$_t$				Incumbent Reelection$_t$	
	Incumbent FEs		*District × Decade FEs*		*District × Decade FEs*	
	(1)	(2)	(3)	(4)	(5)	(6)
Election Day Rain$_t$	-0.041		-0.047		0.007	
	(0.017)*		(0.021)*		(0.010)	
Prior Weekend Rain$_t$		-0.048		-0.061		0.002
		(0.019)*		(0.024)*		(0.015)
Election Day Rain$_{t-1}$	-0.029		-0.034		-0.010	
	(0.017)$^+$		(0.023)		(0.010)	
Prior Weekend Rain$_{t-1}$		-0.023		-0.026		-0.010
		(0.019)		(0.020)		(0.012)
Dem. Vote Margin$_{t-1}$	-0.009	-0.008	-0.168	-0.168	-0.001	-0.003
	(0.026)	(0.026)	(0.029)***	(0.029)***	(0.014)	(0.015)
Inc. Ideal Point$_{t-1}$	-0.007	-0.007	0.002	0.002	0.003	0.004
	(0.010)	(0.010)	(0.011)	(0.011)	(0.005)	(0.005)
Dem. Incumbent$_t$	0.085	0.083	0.042	0.040	-0.001	-0.001
	(0.036)*	(0.035)*	(0.097)	(0.095)	(0.455)	(0.456)
District Ideology$_{t-1}$	-0.042	-0.036	-0.108	-0.104	0.016	0.015
	(0.019)*	(0.019)*	(0.023)***	(0.023)***	(0.017)	(0.017)
Competitiveness$_t$	-0.062	-0.062	-0.052	-0.052	-0.008	-0.008
	(0.005)***	(0.005)***	(0.005)***	(0.005)***	(0.004)*	(0.004)*
Observations	4177	4177	4177	4177	4567	4567
Clusters	1328	1328	1960	1960	2110	2110
R^2	0.928	0.928	0.942	0.942	0.255	0.255

Specifications are OLS with *Incumbent* [(1)-(2)] or *District* × *Decade* [(3)-(6)] fixed effects, and additional state, year, and district controls. *Incumbent* [(1)-(2)] or *District* × *Decade* [(3)-(6)] cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

might have are likely very small, and will fade away rather quickly, compared to rain's more sustaining impact on elections.

We also briefly discuss this exclusion assumption in Footnote 21, noting that normalized rain at $t - 1$ appears to have a small, but negative influence on elections at t. This effect is always smaller than the influence of rain at t, and we find that rain two elections out ($t - 2$) has no measurable influence on voting at t. Empirically, we find that this lagged effect stems, for the most part, from our rain normalization procedure. By using correlated average and variance moments to normalize rain at t and $t - 1$, we induce a weak association between our current and prior period rain instruments. This correlation amongst the instruments, in turn, may magnify a weak association between prior rain and current voting. To address this issue, we first rerun our results presented in Table 2 in the main draft using unnormalized, rather than normalized rain instruments. These regressions are presented here in Table V. From these results, we find that the influence of (unnormalized) prior rain on voting is typically half that recovered for current period rain. Further, our estimates of the influence of previous rain on Democratic voting at t are always statistically indistinguishable from zero at the $p = 0.05$ level. Given this finding, most (though not all) of the apparent lagged effect of rain is likely an artifact of how we control for trends in our rain measures.

Table VI: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning – Robustness to Including Prior Rain Variable

	No Prior Rain		Prior Rain as Control	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day$_t$ (3)	Prior Weekend$_t$ (4)
Dem. Vote Margin$_t$	-2.707	-1.817	-2.263	-1.595
	(1.318)*	(0.941)$^+$	(0.965)*	(0.730)*
Observations	6257	6257	6237	6237
Clusters	1615	1615	1609	1609
R^2	0.860	0.899	0.882	0.907

Specifications are 2SLS with *Incumbent* fixed-effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

To further assess this concern about lagged effects, we rerun our main analysis entirely excluding prior rain from the models and including prior rain as a control rather than as an instrument. These results are presented here in Table VI.[3] As can be seen, our results are generally robust to both alternative specifications. Most notably, in models (3) and (4), we find very similar estimates of adaptation to those in our main regressions, even after controlling for the potential lag effect of past rain on current election outcomes. This suggests additional evidence that any lingering influence of previous election rain

[3]Note, these models use fully normalized rain measures.

on voting does not account for the IV effects we uncover.

Generally speaking, this potential prior rain effect might be a concern in our IV analysis, under the strict exclusion assumption that *all* of rain's influence on elections is captured immediately in Democratic voting on Election Day. Though we show our results are robust to controlling for prior variation in election rain, it is possible that some of the elite adaptation behavior we observe is due to trends in Democratic voting stemming from multiple elections. This would slightly alter the interpretation of our results, but would not fundamentally change what the findings say about our basic theoretical account. Specifically, this could indicate that elites adapt to trends in voting that occur over the last two cycles, rather than just the previous election, with our main IV estimates capturing the total adaptation occurring over both periods.[4] Under this expanded interpretation, the results still support our theoretical claim that incumbents adapt to new information about trends in opinion in their districts signaled by previous election-margins.

V Descriptive Tables and Figures, and Additional Simulations

This section includes a number of additional tables and figures containing descriptives and data, density figures of the rain instruments, and further simulation results.

V.1 Descriptives Statistics and Density Figures

Table VII contains descriptive data on the full list of controls used throughout the analyses. These are listed under CONTROLS. The table also includes descriptive data on the rain instruments and the two outcome measures. Column 5, 'Impute,' contains the

[4]If prior rain has as lagged influence on voting, our IV estimates would capture both the present (t) and the lagged effect ($t-1$) of rain on incumbent adaptation. These effects would be additive under the assumption that incumbents do not condition the magnitude or direction of their adaptation depending on whether they are responding to signals at $t-1$ rather than at t. If incumbents do condition their adaptation and thus effects attenuate, then it would be the case that the rain effect at t is even greater than what we report in our main results. On the other hand, this could evidence adaptation behavior that is more complex than that depicted in our theoretical account.

Table VII: Descriptive Statistics for House Elections Data, 1956 to 2008

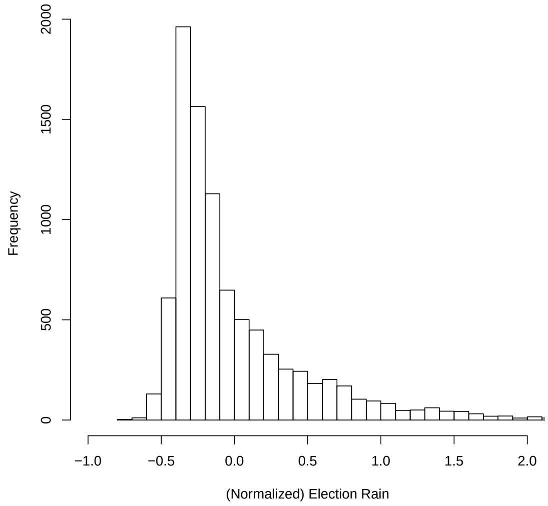
	N	Mean	SD	Impute
	CONTROLS			
Dem. Vote Margin$_{t-1}$	7801	0.057	0.342	0.000
Inc. Ideal Point$_{t-1}$	7717	-0.121	1.002	0.000
Dem. Incumbent$_t$	9812	0.568	0.495	0.000
District Ideology$_{t-1}$	9319	0.082	0.387	0.038
Competitiveness$_t$	7760	3.618	2.697	0.210
Redistricted$_t$	9716	0.231	0.421	0.001
Dem. Sec. of State$_t$	9755	0.594	0.491	0.005
Dem. Governor$_t$	9812	0.530	0.499	0.000
Midterm Election$_t$	9812	0.481	0.500	0.000
Presidential Co-Partisan$_t$	9812	0.363	0.481	0.000
Median Income Per Capita$_t$	9805	0.037	0.025	0.001
Log Population$_t$	9812	13.111	0.231	0.000
% Black$_t$	9812	0.113	0.145	0.000
% Construction Worker$_t$	9797	0.027	0.010	0.002
% High School$_t$	9812	0.191	0.039	0.000
% Farmer$_t$	9812	0.016	0.022	0.000
% Foreign Born$_t$	9812	0.069	0.082	0.000
% Govt. Worker$_t$	9812	0.060	0.022	0.000
% Manufacturing$_t$	9683	0.090	0.041	0.009
% Unemployed$_t$	9812	0.024	0.010	0.000
% Urban$_t$	9806	0.704	0.257	0.001
% Retail Worker$_t$	9804	0.080	0.018	0.001
	INSTRUMENTS			
Election Day Rain$_t$	9043	-0.005	0.486	0.000
Prior Weekend Rain$_t$	9043	-0.002	0.439	0.000
Election Day Rain$_{t-1}$	8568	0.020	0.517	0.000
Prior Weekend Rain$_{t-1}$	8568	0.023	0.465	0.000
	OUTCOMES			
Dem. Vote Margin$_t$	8234	0.057	0.349	0.000
Inc. Ideal Point$_t$	7812	-0.133	1.010	0.000

Impute indicates the proportion of values imputed for all 9812 races contested by sitting incumbents in the IV main analysis.

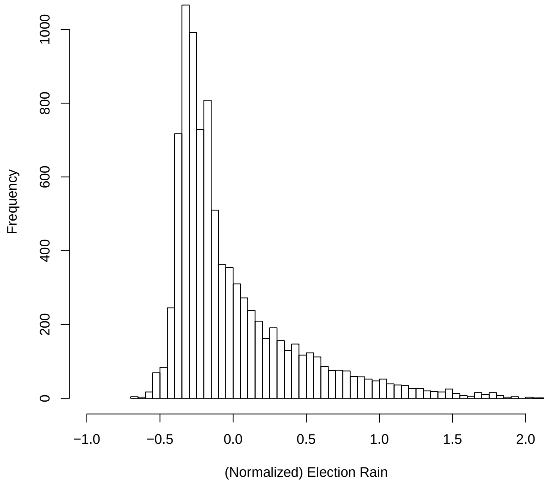
proportion of the data that is imputed in the final analyses. Recall in the main analysis we do not impute rain or vote or ideal-point outcomes. In the robustness results below, we impute these measures in order to test whether doing so modifies the results.

Figure III presents histograms for the Election Day III(a) and Prior-Weekend Rain III(b) instruments. Given the deviation and normalization of our rain measures, we see

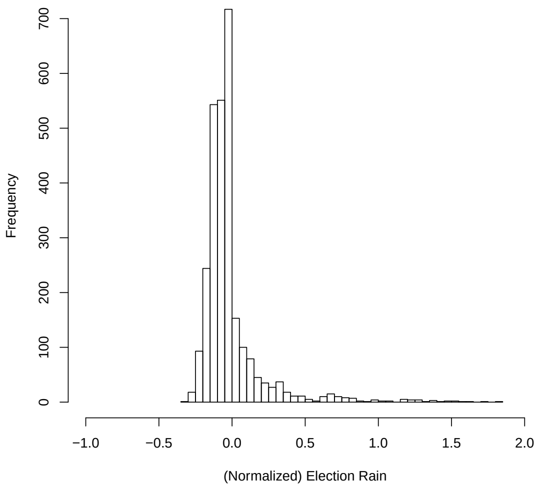
Figure III: Histograms of Daily Average Rain Instruments



(a) Election Day Rain_t



(b) Prior Weekend Rain_t

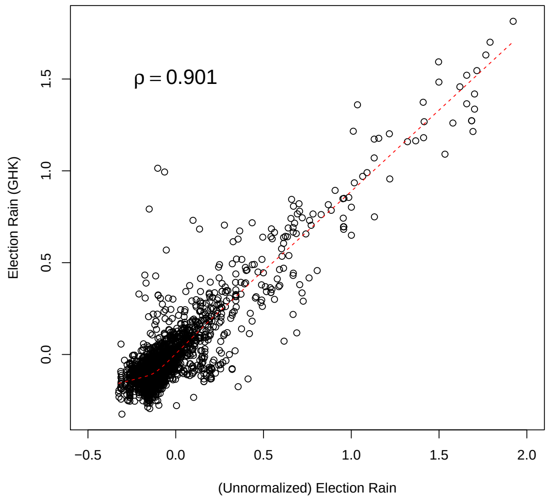


(c) Gomez, Hansford and Krause (2007) Measure

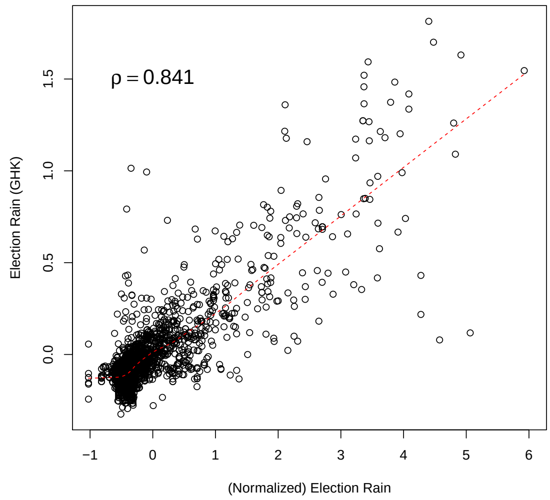
some skew in rain distributions with longer right tails. We also see somewhat lower variance in the Prior-Weekend Rain instrument, as it collects average rain over four (Sat. – Tues.) instead of two days (Mon. – Tues.).

Figure IV(a) presents a histogram of the rain measure collected by Gomez, Hansford, and Krause (2007) at the county level in presidential election years. Thus, to form

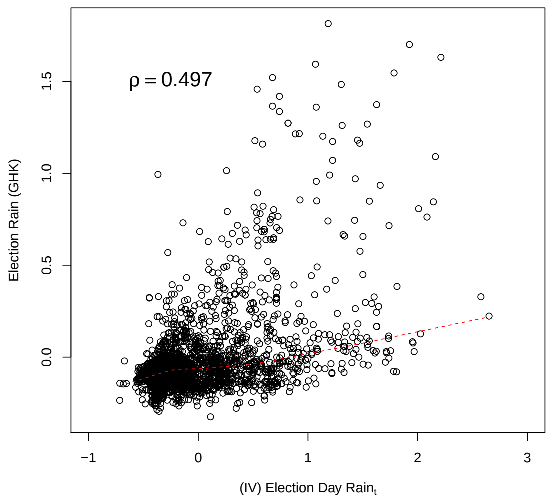
Figure IV: Histogram and Scatterplot of Gomez, Hansford and Krause (2007) District-Level Rain Measure



(a) (Unnormalized) Rain Just On Election Day



(b) (Normalized) Rain Just On Election Day



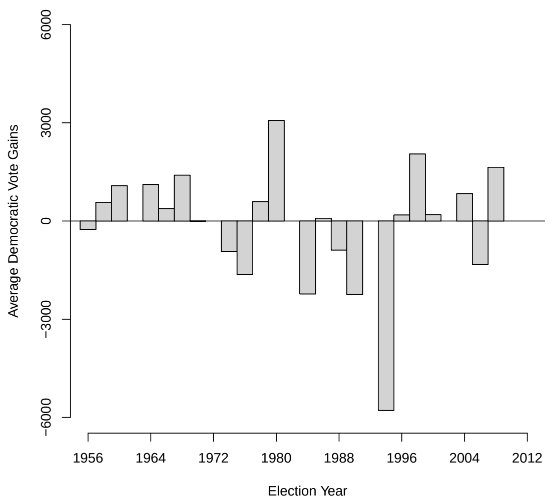
(c) Election Day Rain Instrument

this measure, we use county-to-district linking as outlined in the data section in the manuscript. Figure IV(b) presents the correlation plot between the transformed Election-Day rain instrument used by Gomez, Hansford, and Krause (2007), and our Election-Day instrument. These are correlated at a 0.84, likely reflecting the different ways we measure average rain in the districts.

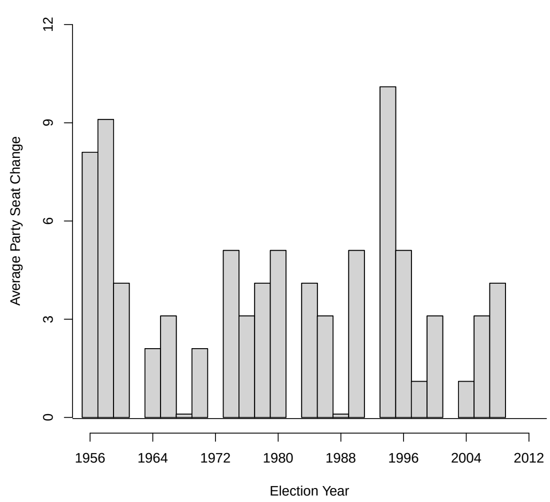
V.2 Additional Simulation Results

We present here a few additional results from our simulations described in the paper. The first of these imputations estimates the net number of Election-Day votes that swing from the Democrats to the Republicans due to rain. We estimate these from our calculation of rain's impact on Democratic voting, and the distribution of rain across districts. Figure V(a) shows the net partisan effect of rain on Election Day due to the average amount of rain in a given election year. Notably these are much smaller than the total changes in party voting occurring given rain as it deviates from its average across all districts. Hence we present this total party-vote impact per year in V(b). Clearly, we see much larger vote effects across districts given rain. However, Figure V(a) shows these have relatively modest net impacts on Congress as whole.

Figure V: Estimated Vote and Seat Swings Per Year Given Average Rain Across Districts



(a) Vote Swings



(b) Seat Changes

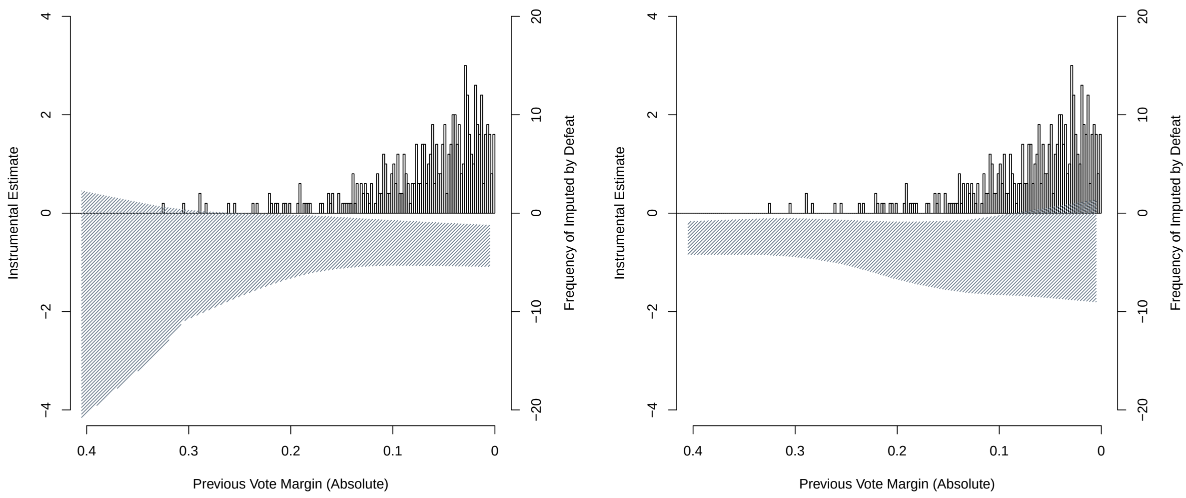
Another way to illustrate this is to measure the number of seats in a given election that could have changed parties had districts all experienced only average amounts of rain. We calculate these by taking the difference between the expected number of votes given average rain, and the votes received during actual rain in the district. Districts are included if their elections were decided by fewer votes than this difference calculated from counterfactual amounts of rain. As shown in Figure V(c), while quite modest, rainfall could have altered the election outcomes of as many as 11 (1994) and as few as 0 (1968, 1988) seats in any given election. On average, election rain could have altered 4.3 seats per election year, and perhaps as many as 116 over the total period, or about 1% of the total House seats (contested and uncontested).

VI Robustness Results to Imputing Data

In this section, we present robustness results for imputing missing data. In Section VI.1, we briefly explore whether our choice to impute losing incumbents' ideal points with their previous values is likely to attenuate our overall estimates. As explained below, we do so by varying the threshold that defines competitive and safe districts, and observe how our IV estimates change as the threshold converges to zero. We also consider whether our results are robust to imputing no missing data in the controls, outcomes and instruments, and also imputing all missing data. We present these results in Table VIII in Section VI.2. As can be seen, our results remain robust irrespective of whether we impute all missing data, or none at all.

In summary, we recover virtually identical results across all alternative specifications, including whether we impute some, all, or none of the outcome and instrument measures. This suggests our findings of adaptation to rain-induced shifts are quite robust overall.

Figure VI: Effect of Imputing Losing Incumbents' Ideal Points on Instrumental Estimates, by Previous Vote Margin Threshold for Defining Competitiveness



(a) Estimate for Safe Races (b) Estimate for Competitive Races

VI.1 Attenuation from Imputing Losing Incumbents Ideal Points

Our findings on marginality suggest that defeated politicians in competitive districts would have been *more* likely to moderate had they managed to hold onto their seats. Unless losing politicians actively reposition themselves against the wishes of their electorates, our choice to impute losing incumbents as standpats is likely to attenuate estimates of adaptation. To examine the degree of attenuation, we iteratively rerun our analysis for marginal and safe districts from the main study, each time widening the competitiveness threshold from a 0.5% to 40% margin of the Democratic vote. As the threshold increases, we include more non-competitive races along with competitive ones in estimating effects in safe and competitive districts. Thus, we should observe less responsiveness in the set of 'competitive' races as the threshold increases, assuming only truly marginal incumbents adapt. In the limit, safe and competitive districts should exhibit indistinguishable levels of adaption. In comparison, as the threshold shrinks, 'safe' districts should increasingly resemble all districts. Thus, the IV estimates in this subset should converge on the main full-sample results.

If incumbents are increasingly responsive due to being more marginal, we would expect IV estimates to decline for competitive races as the competition threshold shrinks. As competitive races are defined to be more marginal, fewer incumbents in these districts may safely avoid adapting to voters. However, shrinking the threshold also includes more defeated incumbents in the set of competitive races. Including more of these candidates should attenuate the resulting IV estimates toward zero, since we assume they standpat on policy. The simulations can identify which competing effect dominates.

Simulation results generally confirm the above expectations, and especially that the stand-pat imputations attenuate estimates overall and for competitive districts. Figure VI plots a histogram of losing incumbents imputed with their previous ideal-points, and overlays the instrumented effect of vote margin on responsiveness in (a) safe or (b) competitive districts, as the absolute margin defining competitiveness shrinks to zero.[5] As shown in Figure VI(a), we consistently fail to reject the null of no effect in safe districts, except at very small threshold values (<5%), when many truly competitive elections are misclassified as safe. In competitive races (Figure VI(b)), however, we obtain statistically significant negative effects until the threshold declines to a margin of roughly 10%. From there, the slope of the estimates converges to zero, mirroring the increased proportion of losing incumbents in that subset. This last finding strongly suggests attenuation, which is apparent here, may also be present in our other results. Given the findings in the main paper about marginal winners, these losers may very well have been *more* likely to respond to changes in their election margins.

VI.2 Imputing *No* and *All* Missing Data

Here we rerun the main analysis, without imputing any data prior to estimation. As seen in Table VIII, in columns (1) and (2), we recover identical (though less precise)

[5]Surfaces are 95% confidence intervals for smoothed lowess estimates of IV coefficients.

Table VIII: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning – Imputing *No* or *All* Missing Data

	No Missing Data		All Missing Data	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day$_t$ (3)	Prior Weekend$_t$ (4)
Dem. Vote Margin$_t$	-1.409	-1.484	-2.223	-1.939
	(0.575)*	(0.623)*	(0.658)***	(0.578)***
Observations	4610	4610	9812	9812
Clusters	1321	1321	1989	1989
R^2	0.918	0.916	0.873	0.885

Specifications are 2SLS with *Incumbent* fixed-effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

results if we do not impute any data. We rerun the main analysis again, imputing all missing data prior to estimation. Again, as seen in columns (3) and (4) in Table VIII, we recover identical (and much more precise) results if we impute all data.

VII Additional Robustness Tests

We present additional robustness checks in this section. We assess whether our results replicate using alternative ideal-point outcomes (Section VII.1) and rain measures (Section VII.2). We measure adaptation behavior in states allowing voting before Election Day (Section VII.4), through vote-by-mail (Section VII.4.1 and VII.4.3) or early voting (Section VII.4.2). We also check if incumbents adapt to rain-induced changes in vote margins in House elections in Presidential battleground states (Section VII.5). Finally, we assess if our results are robust to sequentially removing each election year from our sample, to test if particularly rainy election years drive our findings (Section VII.3). Generally, our results replicate across all these tests, and thus accord with our theoretical expectations about incumbent adaptation to new voter signals.

VII.1 Nominate Ideal Point Positions

We first replicate our results using Nominate scores rather than our Bayesian-IRT ideal-point estimates (Clinton, Jackman, and Rivers 2004; Poole and Rosenthal 1997). Nominate has come to be a popular way for political scientists to scale legislators in a common space, and is the main alternative to the item-response theory (IRT) ideal-point estimation used in our study. Both Nominate and IRT approaches are based on a similar underlying model of latent preferences. Yet, each makes different assumptions about how to model the error structure in choice, as well as how to identify and estimate the models. One major difference is that Nominate achieves identification by restricting parameters to lie within the unit sphere in multidimensional space, which constrains any additive bias in parameters by setting the maximum ideal-point to be at most ± 1. In contrast, Bayesian IRT models achieve identification through the use of prior parameters that decrease the posterior probability of drawing extreme estimates given the data.

Table IX: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning – Unbridged and Bridged Nominate Ideal Points

	Unbridged Nominate		Nokken-Poole Nominate	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day (3)	Prior Weekend (4)
Dem. Vote Margin$_t$	-0.998	-1.070	-1.298	-1.144
	(0.286)***	(0.276)***	(0.531)*	(0.499)*
Observations	6280	6280	6237	6237
Clusters	1620	1620	1609	1609
R^2	0.874	0.869	0.937	0.940

Specifications are 2SLS with *Incumbent* fixed effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Using different estimation approaches could affect our results. For example, if there is measurement error in the ideal-point estimates due to the competing model assumptions, our IV confidence intervals might be overconfident. To assess this concern, we estimate

Nominate scores using the 'wnominate' package and function in **R**. Here we estimate Nominate scores *separately for each Congress*, using the same roll-call data as that used in the IRT estimates for our main responsiveness outcomes. We constrain the polarity of these by restricting the signs of four legislators (Reps. Gwynne, Gross, Hyde and Shadegg). Notably, unlike DW-Nominate, we first model each Congress independently. We do not attempt to bridge different Congresses, either by constraining legislator trends over time to a fixed point, or parameterizing these trends linearly as in (D)W-Nominate. This is because imposing *any* fixed constraint on legislators' ideal-points between Congresses, by construction, reduces the range of possible adaptation. Indeed, Poole and Rosenthal (1997) utilize such constraints under the theoretical view that legislators do not adapt once in office. When estimating adaptation, we want to make as few additional assumptions about legislative behavior as possible, especially as it changes, and thus require a more flexible model. Further, there is no loss in leaving Congresses unbridged since election rain across districts cannot depend on the (floating) average ideological preferences of legislators in the prior session.

Nevertheless, some sort of ideal point bridging may be motivated given our theoretical account of adaption. We argue that incumbents are changing their voting behavior given new electoral information and in light of their past voting record. Thus, it is important to ensure our findings are robust to imposing some common scale on the ideal points over time, and especially to check that the floating dimensions do not correlate with overall election or rain variance in troublesome ways. To assess this concern we rerun our main analyzing using Nokken-Poole (NP-Nominate) estimates (Nokken and Poole 2004). The approach in NP-Nominate is to first fix the overall two-dimension ideological space over all Congresses, by estimating a model with a single, constant ideal point for each legislator. Next, a new set of ideal points is estimated for each legislator in each Congress, using the fixed item parameters (technically the bill cutting lines) as the cutpoints. This approach ensures that any trends in legislator ideal points all are moving in the same dimension (i.e., the fixed bill space) at any point in time (Nokken and Poole 2004).

We present the results of our full analysis using both Nominate and NP-Nominate outcomes in Table IX. As can be seen in models 1 and 2, we find similar (and more precise), estimates of rightward adaptation following conservative shifts in prior vote margins, using standard Nominate. And we find very similar results using the bridged ideal point NP-Nominate scores as shown in models 3 and 4. Thus, our results do not depend on whether we use alternative or 'off-the-shelf' approaches to modeling legislators' preferences.

VII.2 Gomez, Hansford and Krause (2007) Rain Measure

We next assess whether our results depend on the specific rain measures we use, and in particular if we recover different findings using the data collected by Gomez, Hansford, and Krause (2007) (GHK). Both our rain data, and the measures used by GHK come from weather station daily readings produced by the National Climatic Data Center (NCDC). GHK collected data from around 20,000 stations located in counties in the U.S. We managed to expand this effort to include daily measures taken from 36,568 stations (though not all of these were in continual operation over our sample). The central reason we collected our own rain data, rather than used those produced by GHK, is that they did not collect any data for midterm election years, since they focused exclusively on studying rain's effect on presidential races. For our analysis, we required rain data covering both midterm and presidential years, which necessitated us to extend the data collection for these cycles. In doing so, we were able to collect anywhere between 50% and 80% more rain data than Gomez, Hansford, and Krause (2007). As added benefit, these additional stations may help improve the quality of our rain data (i.e., lower measurement error) by increasing the sample of daily readings we used to form our aggregate rain instruments.

Gomez, Hansford, and Krause (2007) were concerned with measuring precipitation at the county-level to be used in models predicting county presidential vote. In the data, some counties have multiple stations, and others few or no stations. Thus, to produce (i.e., impute) county-level rain measures, the authors used a spatial modeling procedure, kriging, to estimate weather over geographies. Kriging is a local linear regression technique that models rainfall for units given rain in neighboring geographies, with influence decreasing in distance following a Gaussian density. In contrast, we are interested in measuring rain at the congressional district level. However, moving from counties to districts adds the complexity of how to link particular stations to their districts when these do not have district identifiers. This is an issue since counties do not wholly reside within districts, and often overlap legislative boundaries. Our approach is to average station readings over the counties in which these reside, to produce county-level rain measures. Then we average over the county readings for all those with some part of the county residing within a district. This approach is a much coarser sort of spatial imputation. The difference is that 'far away' rain has zero influence on our local estimates once we go beyond all county borders for counties that touch some part of the district. Rain 'nearby,' that is taken from stations within counties linked to districts, has equal influence within counties, and then county rain has equal influence within districts (proportional to county size).

Compared to our approach, kriging likely produces rain data with lower measurement error *at the county-level*. Yet, measurement error in the rain variables will be introduced in aggregating from counties to districts, regardless of how county-level rain is measured. We mainly chose our approach based on its simplicity to implement, and under the assumption that any measurement error would produce conservative findings through attenuation. Interestingly, despite the different approaches, the GHK kriging measures are actually quite similar to those we collected. Scatterplots in Figure III above present the GHK measure (aggregated to districts through county averaging) on the y-axis, and analogous measures of rain using our averaging approach. These plots compare the GHK measure, and (a) Unnormalized Rain Just on Election Day, (b) Normalized Rain Just on Election Day, and our main (c) Election Day Rain instrument (with Monday and

Tuesday rain). The main finding here is that the GHK measure correlates with the analogous measure we produced using our coarser approach at $\rho = 0.901$. Perhaps this difference is due to the use of kriging versus station averaging to produce county measures. Yet, it is worth noting that some of the difference could be due to the inclusion of 16,568 additional stations in our measures.

Table X: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning – Gomez, Hansford and Krause (2007) Rain Measure

	Not Imputing Rain		Imputing Rain	
	GHK Rain (1)	Analogous Measure (2)	GHK Rain (3)	Analogous Measure (4)
Dem. Vote Margin$_t$	–	-1.422	-2.104	-2.390
	–	(1.023)	(1.066)*	(1.362)$^+$
Observations	0	6183	9812	9812
Clusters	0	1607	1989	1989
R^2	–	0.911	0.878	0.865

Specifications are 2SLS with *Incumbent* fixed effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

The main test, however, is whether we recover similar findings in using the GHK measure, rather than our rain instruments. We present results from this analysis in Table X. We first present results for the GHK Rain instrument in column (a) without imputing missing data. Unfortunately, there are no districts with complete measures for all our controls *and* the GHK instruments, when we include the prior GHK rain measure in the model. This is because prior rain measures (from midterm elections) are always missing. There are 2,345 complete units just for presidential races. When we focus on these races, we recover a similar result from our main analysis (-2.044), though one that just misses statistical significance. In light of this, we examine the rain instrument we produced for presidential and midterm elections that is as similar as possible to the GHK measure. This 'Analogous Measure' is our Normalized Rain Just on Election Day variable.[6] This

[6]Note this measure is different from our main Election Day Rain instrument in using rain data only

result in presented in column (2) in Table X. Again, this finding (-1.422) is consistent with our main results, but just misses statistical significance under a two-tailed test. If we changed this to a one-tailed test, and assessed the null against the alternative that rightward vote margins led to conservative roll-call shifts, both of these coefficients would be statistically significant at the $p < 0.1$ level (p-values are 0.056 and 0.089).

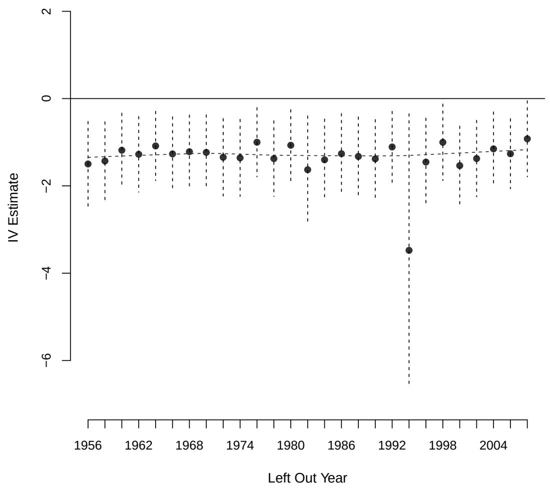
Missing data is a central limitation in both of these analyses, and especially in column (1). Thus, we now turn to tests that impute all missing data, including missing values for the (c) GHK and (d) Analogous rain instruments. For both of these tests, the coefficients are negative (-2.104, -2.390) and statistically distinct from zero at standard levels (p-values are 0.048 and 0.079), for a two-tailed test with cluster errors. These results (and especially the latter imputation findings) add additional confidence in the main findings in our study. These robustness findings show that using alternative measures for the rain instruments does not alter our conclusions about elite adaptation in light of conservative shifts in prior vote-margins.

VII.3 Leave-One-Year-Out Analysis

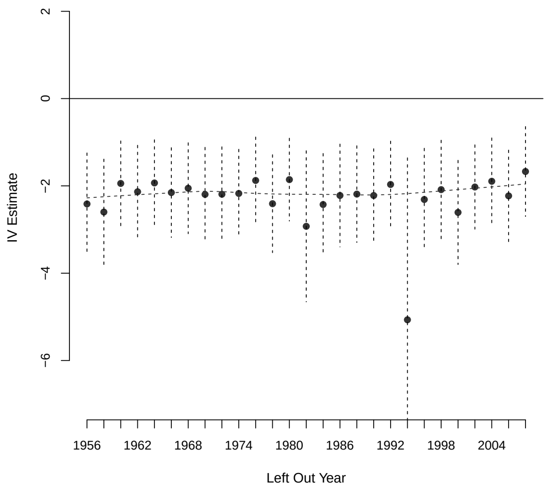
We next assess whether our results are driven by the inclusion of particular election years. For example, 1994 was an especially rainy year, and resulted in the defeat of many Democrats. By leaving each year out of the analysis, we can ensure that our adaptation findings generalize beyond one or a few exceptional years. To do this, we rerun our main IV analysis sequentially, removing each election year during estimation. These models include the full controls and *Incumbent* fixed effects, with *Incumbent* cluster errors. The results for these tests are presented in Figure VII for the Election-Day instrument, and in Figure VIII for Prior-Weekend rain. For these analyses, we first estimate the *leave-one-out* models without imputing any data. These models are presented in Figures VII(a) and VIII(a). We then rerun the analysis imputing all missing items, reported in Figures VII(b)

from Election Day Tuesday.

Figure VII: Robustness of IV Estimates to Removing Each Year From the Analysis, Election Day Rain Instrument

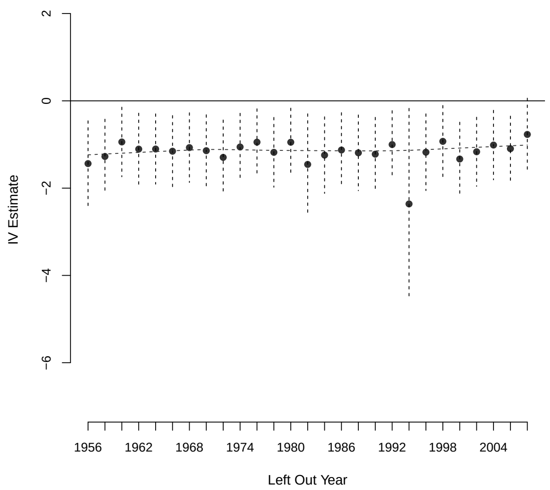


(a) No Missing Data Imputed

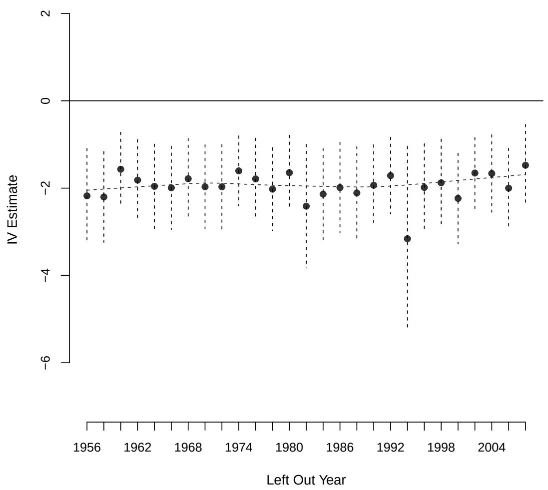


(b) Missing Data Imputed

Figure VIII: Robustness of IV Estimate to Removing Each Year From the Analysis, Prior Weekend Rain Instrument



(a) No Missing Data Imputed



(b) Missing Data Imputed

and VIII(b). In these Figures, dashed vertical-lines indicate 95% confidence intervals around 2SLS IV estimates.

Turning to the unimputed results for the Election-Day instrument in Figure VII(a), we find that our IV estimates remain consistently negative and statistically distinct from zero when we leave each year out. In other words, our IV results do not depend on whether or not we include any particular year in the analysis, including 1994. We find very similar results when looking at the leave-one-year-out results for the Prior-Weekend rain in Figure VIII(a). Here again, all of our estimates are statistically different from zero, with the exception of excluding 2008. When we leave the 2008 elections out, our cluster results just miss statistical significance under a two-tailed test, but are significant at $p = 0.064$ under a one-tailed test. Similar to the above, these results exclude around 3,000 elections due to missing data. We rerun this leave-one-out analysis, imputing missing data to maximize the power of our tests. Imputation is helpful here since, in excluding particular years, we lose additional statistical information through a smaller sample. After imputing missing data, we recover similar and stronger results, in comparison to our unimputed tests. As seen in Figure VII(b) for Election-Day rain, and in Figure VIII(b) for Prior-Weekend rain, all of our results are negative and statistically distinct from zero at the $p < 0.05$ level. In summary, these tests generally show that our findings do not depend on whether or not we include any particular years in our sample.

Though we still recover negative and significant estimates leaving out 1994, the year is particularly influential in our results, as seen by the much wider confidence intervals for the associated test. Interestingly, this result suggests that 1994 was particularly informative in our analysis in measuring elite adaptation behavior. Such a result seems both theoretically reasonable and intriguing. While Democrats lost a large number of seats, many still returned to office. These incumbents may have viewed large vote-swings as strong signals that their districts demanded more conservative policy action. These swings would appear especially credible, given the Republican takeover in Congress. Rainy elections could have augmented these swings locally, inducing stronger conservative shifts in voter signals. In light of this, perhaps winning Democrats were more sensitive to shifts in vote margins, since these mirrored the national tide. Though somewhat speculative, in line with our information account, this finding may suggest that elites attune their adaptation more significantly when voter signals attain additional credibility from nationalization. We intend to explore this conjecture in the future.

VII.4 Voting Before Election Day

In a penultimate series of robustness checks, we examine whether our results replicate in states with various in-person or early voting laws. If states allow significant voting prior to elections, then rain on Election Day should have much less influence on Democratic turnout and voting in those districts. In turn, this means that rain should not correlate with incumbent roll-call voting since it conveys no information about changes in voter attitudes through prior vote margins. We assess this through three tests. We first estimate adaptation just for Oregon before and after the state's adoption of all vote-by-mail elections in 2000 in Section VII.4.1. We then stratify our estimates on states with and without no-excuses required, early in-person voting in Section VII.4.2. Finally, we stratify by states with and without no-excuses needed, absentee voting in Section VII.4.3.

VII.4.1 Oregon Before and After Adoption of Vote-by-Mail Elections in 2000

Our results just for Oregon before and after its adoption of all vote-by-mail elections in 2000 are presented in Table XI. We again rerun our models using the full list of controls and *Incumbent* fixed-effects, with cluster errors. Unfortunately, there are very few elections to help us examine whether or not adaptation behavior is evident before or after the change in election law. Due to this, our estimates are very unstable and uninformative. We find conflicting estimates for our (1) Election-Day and (2) Prior-Weekend instruments prior to 2000. But note these estimates are drawn from only 15 incumbent legislators. Post-2000 there are only four incumbents, which produce estimates with very large confidence intervals in (3) and (4). Thus, it is difficult to draw deeper conclusions just from looking at one state.

Table XI: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning – Before and After Introduction of Vote-by-Mail Elections in Oregon

	Before Vote-by-Mail		After Vote-by-Mail	
	Election Day_t (1)	Prior Weekend_t (2)	Election Day_t (3)	Prior Weekend_t (4)
Dem. Vote Margin_t	0.133	-0.692	-114.879	-57.960
	(2.045)	(2.004)	(626.046)	(255.968)
Observations	66	66	17	17
Clusters	15	15	4	4
R^2	0.907	0.907	–	–

Specifications are 2SLS with *Incumbent* fixed effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

VII.4.2 States With Early In-Person Voting

We present our results for states with early in-person in Table XII.[7] In columns (1) and (2), we present IV estimates for districts that reside in states with early voting laws in place. As expected, the findings for both of these models are not distinguishable from zero. Further, results are null both for the first-stage effect of rain on Democratic voting *and* the effect of rain on incumbent adaptation. In contrast, we find very similar results to our main analysis in states without early voting laws. As shown in models (3) and (4), results are negative and statistically significant using both rain instruments.

[7]States early voting, with year of adoption, are: AK (1960), AR (1972), AZ (2000), CA (1976), CO (1992), FL (2006), GA (2004), HI (1996), IA (1992), ID (1972), IL (2006), IN (2002), KS (1998), LA (2006), ME (2006), MT (1988), NC (2000), ND (2004), NE (2000), NM (1996), NV (1992), OH (2006), OK (1992), OR (1980), SD (2004), TN (1994), TX (1988), UT (2004), VT (1980), WA (1980), WI (2000), WV (2004), and WY (2006).

Table XII: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning – States With and Without Early Voting

	Early Voting		No Early Voting	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day$_t$ (3)	Prior Weekend$_t$ (4)
Dem. Vote Margin$_t$	-2.714	-5.340	-1.196	-1.073
	(9.446)	(9.927)	(0.461)**	(0.434)*
Observations	1064	1064	5173	5173
Clusters	379	379	1351	1351
R^2	0.910	0.806	0.913	0.916

Specifications are 2SLS with *Incumbent* fixed effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

VII.4.3 States With Absentee Voting

We next present results for states with and without absentee voting in Table XIII.[8] Similar to the above, we find null IV results using (1) Election-Day and (2) Prior-Weekend instruments to assess adaptation in states that allow absentee voting. As expected, we recover conservative and significant adaptation behavior following shifts in Democratic vote-margins in states without (or with restricted) absentee voting.

VII.5 Presidential Battleground States

We finally analyze whether incumbents adapt to election rain in presidential battleground states in presidential election-years. Electoral turnout in these states is likely to be heightened by the voter mobilization underway by presidential campaigns. Thus, any influence rain may have on Democratic voting should be dampened, also diminishing any possible incumbent adaptation. To assess this, we stratify states based on whether the

[8] States, with year of adoption, are: AK (1996), AR (2002), AZ (1992), CA (1980), CO (1992), FL (1996), GA (2004), HI (1996), ID (1972), IN (2002), KS (1996), MD (2008), ME (2000), MN (1992), MT (1988), NC (2000), ND (2000), NE (2000), NJ (2008), NM (1996), NV (1992), OH (2008), OK (1992), OR (1980), SD (2004), TN (1972), UT (2000), VT (1980), WA (1976), WI (2000), and WY (1992).

Table XIII: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning – States With and Witout Absentee Voting

	Absentee Voting		No Absentee Voting	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day$_t$ (3)	Prior Weekend$_t$ (4)
Dem. Vote Margin$_t$	-4.200	-3.775	-1.223	-1.099
	(7.979)	(5.995)	(0.474)**	(0.449)*
Observations	868	868	5369	5369
Clusters	310	310	1409	1409
R^2	0.869	0.885	0.912	0.916

Specifications are 2SLS with *Incumbent* fixed effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

overall final presidential vote-margin was less than 8%.[9][10] IV estimates for incumbents in battleground states are presented in Table XIV, in columns (1) and (2). As expected, the estimates are not statistically different from zero. Turning to the results for non-battleground states, we find evidence of adaptation. Using (3) Election-Day and (4) Prior-Weekend rain, we find negative and significant IV estimates for adaptation, again replicating our main analysis and results.

[9]See Fraga and Hersh (2010) for a justification for using this margin for battleground competition.

[10]The full list of battleground states is presented below in Table XV.

Table XIV: Instrumented Effect of Democratic Vote Margin on Subsequent Incumbent Roll Call Positioning – By Battleground Presidential States

	Battleground		Non-Battleground	
	Election Day$_t$ (1)	Prior Weekend$_t$ (2)	Election Day$_t$ (3)	Prior Weekend$_t$ (4)
Dem. Vote Margin$_t$	-2.408	-13.008	-0.925	-0.838
	(16.352)	(44.622)	(0.455)*	(0.455)$^+$
Observations	1333	1333	4904	4904
Clusters	896	896	1517	1517
R^2	0.882	0.592	0.924	0.925

Specifications are 2SLS with *Incumbent* fixed effects and additional state, year, and district controls. *Incumbent* cluster standard errors are in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table XV: Presidential Battleground States

	Year	States
1	1956	Arkansas, Minnesota, Missouri, North Carolina, Tennessee
2	1960	Alaska, Arkansas, California, Connecticut, Delaware, Florida, Hawaii, Idaho, Illinois, Kentucky, Maryland, Michigan, Minnesota, Mississippi, Missouri, Montana, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oregon, Pennsylvania, South Carolina, Tennessee, Texas, Virginia, Washington, West Virginia, Wisconsin
3	1964	Arizona, Florida, Idaho, Nebraska, Virginia
4	1968	Alaska, Arkansas, California, Connecticut, Delaware, Illinois, Kentucky, Maryland, Michigan, Missouri, New Jersey, New York, Ohio, Oregon, Pennsylvania, South Carolina, Tennessee, Texas, Washington, Wisconsin
5	1972	Minnesota, Rhode Island
6	1976	California, Connecticut, Delaware, Florida, Hawaii, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Michigan, Mississippi, Missouri, Montana, Nevada, New Jersey, New Mexico, New York, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, South Dakota, Texas, Virginia, Washington, Wisconsin
7	1980	Alabama, Arkansas, Delaware, Hawaii, Illinois, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, New York, North Carolina, Pennsylvania, South Carolina, Tennessee, Vermont, West Virginia, Wisconsin
8	1984	Iowa, Maryland, Massachusetts, Minnesota, Pennsylvania, Rhode Island
9	1988	California, Colorado, Connecticut, Illinois, Maryland, Massachusetts, Michigan, Minnesota, Missouri, Montana, New Mexico, New York, Oregon, Pennsylvania, South Dakota, Vermont, Washington, West Virginia, Wisconsin
10	1992	Alabama, Arizona, Colorado, Connecticut, Florida, Georgia, Indiana, Iowa, Kansas, Kentucky, Louisiana, Michigan, Montana, Nevada, New Hampshire, New Jersey, North Carolina, Ohio, South Dakota, Tennessee, Texas, Virginia, Wisconsin, Wyoming
11	1996	Alabama, Arizona, Colorado, Florida, Georgia, Indiana, Kentucky, Mississippi, Missouri, Montana, Nevada, New Mexico, North Carolina, North Dakota, Ohio, Oklahoma, South Carolina, South Dakota, Tennessee, Texas, Virginia
12	2000	Arizona, Arkansas, Florida, Iowa, Louisiana, Maine, Michigan, Minnesota, Missouri, Nevada, New Hampshire, New Mexico, Ohio, Oregon, Pennsylvania, Tennessee, Washington, West Virginia, Wisconsin
13	2004	Colorado, Delaware, Florida, Iowa, Michigan, Minnesota, Missouri, Nevada, New Hampshire, New Jersey, New Mexico, Ohio, Oregon, Pennsylvania, Washington, Wisconsin
14	2008	Florida, Georgia, Indiana, Missouri, Montana, North Carolina, Ohio, Virginia

States are included if presidential election margin is less than 8%.

References

Clinton, Joshua D., Simon Jackman, and Doug Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98 (2): 355–370.

Fraga, Bernard, and Eitan Hersh. 2010. "Voting Costs and Voter Turnout in Competitive Elections." *Quarterly Journal of Political Science* 5 (4): 339–356.

Gomez, Brad T., Thomas G. Hansford, and George A. Krause. 2007. "The Republicans Should Pray for Rain: Weather, Turnout, and Voting in U.S. Presidential Elections." *Journal of Politics* 69 (3): 649–663.

Kleibergen, Frank, and Richard Paap. 2006. "Generalized Reduced Rank Tests Using the Singularvalue Decomposition." *Journal of Econometrics* 127 (1): 97–126.

Nokken, Timothy P., and Keith T. Poole. 2004. "Congressional Party Defection in American History." *Legislative Studies Quarterly* 29 (1): 545–568.

Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A Political-economic History of Roll Call Voting*. New York, NY: Oxford University Press.

Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models*, ed. James H. Stock and Donald W.K Andrews. Cambridge, UK: Cambridge University Press.

van Buuren, Stef, and Karin Groothuis-Oudshoorn. 2011. "MICE: Multivariate Imputation by Chained Equations in **R**." *Journal of Statistical Software* 45 (3).