PLSC 504: Replication Term Paper

Secular Party Rule and Religious Violence in Pakistan

Mario Belledonne

December 17, 2019

1 Introduction

Do terrorists cause violence in response to secular incumbency or does secular incumbency occur in response to terrorist violence?

RD in polysci

The period of 1998 $\rightarrow$ 2003 in Pakistan offers a "natural experiment" due to a plurality of first-past-the-post elections where both Islamist and secular leaders competed for local elections. These elections determine the Members of the National Assembly (MNA), who are responsible for implementing local policies at the behest of their constituency. The authors claim that secular victories within a MNA district attenuate violence at the district level, possibly moderated by police infrastructure [5].

The original authors develop a novel unit of observation (**??**) in an effort to mitigate geographical missmatch between the outcome variable, reports on religious violence, and treatment, secular victories. As a result, the authors employ a sophisticated answer strategy around these units. In this work, I both replicate the main findings and explore the implications for this strategy in simulation that support the original findings.

2 Design

The authors aim to measure the local average treatment effect (LATE) of secular victories on violence in democratic elections for close elections between secularists and Islamists. Due to certain complexities in the construction of observed units, the usual sharp regression discontinuity design (RD) where the margin of victory acts as the running variable is not feasable. Instead, the authors chose to employ an fuzzy RD estimation of the LATE using 2SLS under an instrumental variables (IV) approach. In the following section, I will visit the authors' defense for this design before replicating their main findings (**??**).

2.1 Data

2.1.1 Units

The main effect was obtained from reports from the BFRS Political Violence in Pakistan Dataset. This dataset tallied reports of political violence from a daily English-language newspaper, *Dawn*. The geo-political units of these reports are in terms of administrative districts. This immediately posses a challenge to identification as administrative units do not correspond in a one-to-one fashion to MNA constituencies and have re-organized over time.

Rather than interpolating violence at the district level to the constituency level, the authors define a novel geo-political unit of analysis, `joined-district`, that clusters administrative districts into groups that contain a complete set of constituencies 2.1.1.

> ...the smallest amalgamation of districts that encompasses complete MNA constituencies.

Importantly, the `joined-districts` have reorganized over time. Both administrative districts and MNA constituencies where partially reorganized in 2002 (TODO cite?). The authors note concern regarding autocorrelation of `joined-districts` that where previously members of the same unit. This leads to the introduction of a second unit, the `cluster districts`, that describes a second level of `joined-districts` defined as the smallest set of time-invariant administrative districts containing complete MNA constituencies between 1988 and 2013. In other words, `joined-districts` are subsets of `cluster districts` that the authors hope would capture any control for any error correlation among units at a given election year.

2.1.2 The Outcome Variable

The authors use a variety of outcomes for religious violence for a particular joined-district i at election t, $Y_{i,t}$.

1. Any Event: A binary variable that is *True* if any form of religious violence occured during the MNA's time in office for that district
2. Any Killed: Similar to *Any Event* but referring to any deaths
3. Event Count: The number of religious-violent events
4. Number Killed: The number of deaths caused by religious violence
5. Number of days: The number of days in which at least one instance of religious violence occured.

For each of these, a violent event was considered religiously motivated if there was no explicit evidence to suggest the contrary. The authors defend this criteria by noting that the majority of religious violence is labeled as politically motivated due to the

perpetrators belonging to religious political groups. Additionally, the authors perform robustness checks by removing units in the Karachi and Balochistan provinces as those regions have a high proportion of political violence arising from non-religiously motivated assassination (TODO cite, ref result).

2.1.3 The Treatment Variable

As described above in 2.1.1, the nature of MNA constituency aggregation leads treatment, $D_{i,t} \in [0, 1]$, to take on a rational value, the proportion of MNA seats won by secularist candidates in a `joined-district`.

2.2 Answer Strategy

The original authors motivate their 2SLS estimation procedure by first exploring the pitfalls of naive estimation via difference in means.

2.2.1 Naive Estimation: Difference in Means

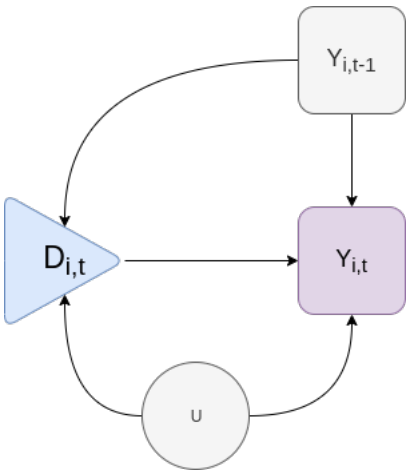

Figure 1: DIM Bias

The bivariate Difference in Means (DIM) estimator for the ATE takes the following form:

$$Y_{i,t} = \alpha + \beta * D_{i,t} + \epsilon_{i,t} \tag{1}$$

Where, $D_{i,t}$ is the treatment and $\epsilon_{i,t}$ describes the error term for that unit, (i, t).

Given the domain and design there are several probable sources of bias with this estimate.

1. Time lagged treatment and outcomes: `joined-districts` with histories of religious violence may seek out / distrust secularists.
2. Confounding variables: Covariates such as state capacity, education, economic stability may both impact popular opinion of secular vs religious candidates differentially
3. provincial FEs?
4. autocorrelation across temporally varying units: As mentioned in 2.1.1, since some group of units at time t to have been part of the same unit a $t-1$, it is possible for the error terms across those units to be correlated. (Note this would bias the estimate of SE)

2.2.2 Regression Discontinuity

Given that at some narrow margin of victory/defeat, the probability of treatment, a secular victory, would be as if random, regression discontinuity presents itself as an appealing answer strategy for estimating an `ATE` localized around the cutoff [1]. However as alluded to in 2.1.1 and 2.1.3, the aggregate nature of the data does not lend itself to a clearly identified running variable. Given $D_{i,t}$, the proportion of secularist victories for that unit, it is not immediately clear how the aggregate (average) election margins could be interpreted as a running variable.

2.2.3 2SLS IV estimation

The authors resort to 2SLS estimation strategy using a quasi-fuzzy RD argument. Rather than margin of victory serving as the running variable, the authors instead define and use the proportion of secularist victories where the margin of victory was $\leq 3\%$ as an exogenous regressor for treatment. This exogenous regressor, $Z_{i,t}$, cannot be used as a running variable as it does not have a clear cutoff without additional assumptions. Thus, the authors settle on instrumental variables estimation even though the instrument, in this case, covaries with treatment out of logical construction rather than causal relation.

In addition to the instrument, the authors attempt to control for unobserved covariates that would influence the proportion of close secularist victories by controlling for the proportion of close races either won or lost by secular canditates by $\leq 3\%$, $X_{i,t}$. In the Extensions section (4), I will present criticism of this answer strategy and offer alternatives revolving around binary treatment variables and single election districts.

The 2SLS answer strategy requires the following assumptions [4, 2, 1]. For outcome ($Y_{i,t}$), treatment ($D_{i,t}$), a running variable ($Z_{i,t}$), and a cutoff c:

1. Exogenous instrument/regressor
2. Non-zero first stage: $D_{i,t} \not\perp\!\!\!\perp Z_{i,t}$
3. Exclusion Restriction: $Y_{i,t} \perp Z_{i,t} \| D_{i,t}$

4. Monotonicity: $Pr(D_{i,t}(Z)|Z > c) > Pr(D_{i,t}(Z)|Z < c)$
5. Non-interference

The authors defend the exogeneity of the instrument by ensuring that the instrument did not predict a variety of pre-treatment covariates including state capacity, agricultural production, census data (education, utilities), and voter participation. These results are replicated in 3.2 along with a density test 5.

The `non-zero first stage` is replicated in table 4 by calculating the first-stage regression described in eq 1.

The assumption of the `exclusion restriction` is defended by the construction of their running variable. By conditioning on the incidence of close secularist races (X), the only logical path that the proportion of close secular victories could effect violence would be through the proportion of victories. While the original work does not explicitly mentioned the assumption of `monotonicity`, this assumption is defended on similar grounds. As the proportion of close secular victories increase, so must the proportion of secular victories regardless of the margin.

Fixed Effects The original compilers of the BFRS dataset recommend using province level fixed effects, θ_p in order to control for regional differences in reporting behavior. However, the authors note that excluding FE did not change the main estimates. (TODO add figure in results)

In sum, figure 2 illustrates the model identifying the LATE. They formally define a two stage estimation strategy as follows:

$$\widehat{D_{i,t}} = \mu + \lambda * Z_{i,t} + \kappa * X_{i,t} + \theta_p + v_{i,t} \tag{2}$$

$$Y_{i,t} = \alpha + \beta * \widehat{D_{i,t}} + \gamma * X_{i,t} + \theta_p + \epsilon_{i,t} \tag{3}$$

Where

- $Z_{i,t}$ describes the proportion of MNA seats won by secular candidates when in close competition with Islamist candidates (within $\pm 3\%$).
- $\widehat{D_{i,t}}$ is the predicted proportion of MNA seats won by secular candidates across all races in that district.
- θ_p describes a geo-spatial fixed effect of violence reporting at the province level.

Here 2 refers to the first stage and 3 refers to the estimator of the `LATE` localized to close races between secular and Islamist MNA elections.

To give insight to how the outcome variables, treatment, and covariate (X) vary with the instrument, I visualized figure 3 4.

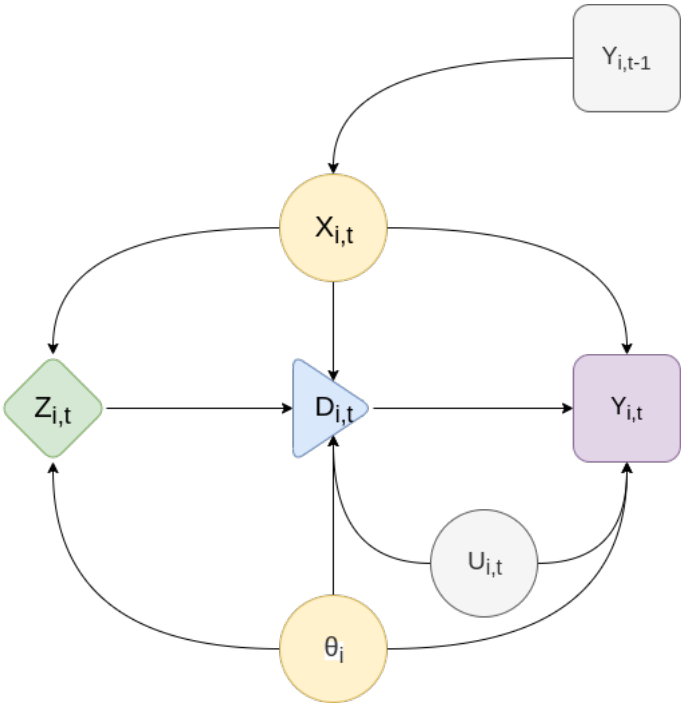

Figure 2: IV Model

3 Results

3.1 Estimates for the LATE

	Replication Any Event	Original
Prop. Secular Win	−0.660***	−0.660***
	(0.215)	(0.212)
Prop. Secular Close Race	0.031	0.031
	(0.085)	(0.084)
Province FEs	Y	Y
Num. obs.	437	437

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.
Robust SEs clustered by cluster-district area, in brackets

Table 1: 2SLS LATE with STATA SE estimation

In table 1, we replicate the main estimate, the LATE of the proportional of seats won by secular leaders on religious violence for compliers. The full replication table is reported in 8

However, it is important to note that the significance of these estimates are sensitive to the standard error estimation strategy (`stata`). In table 2, I recalculated the effects using

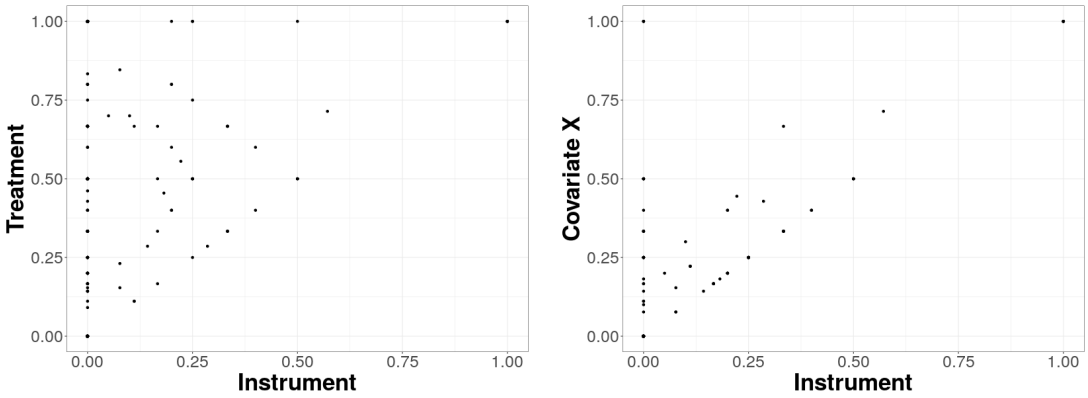

Figure 3: Visualized the data

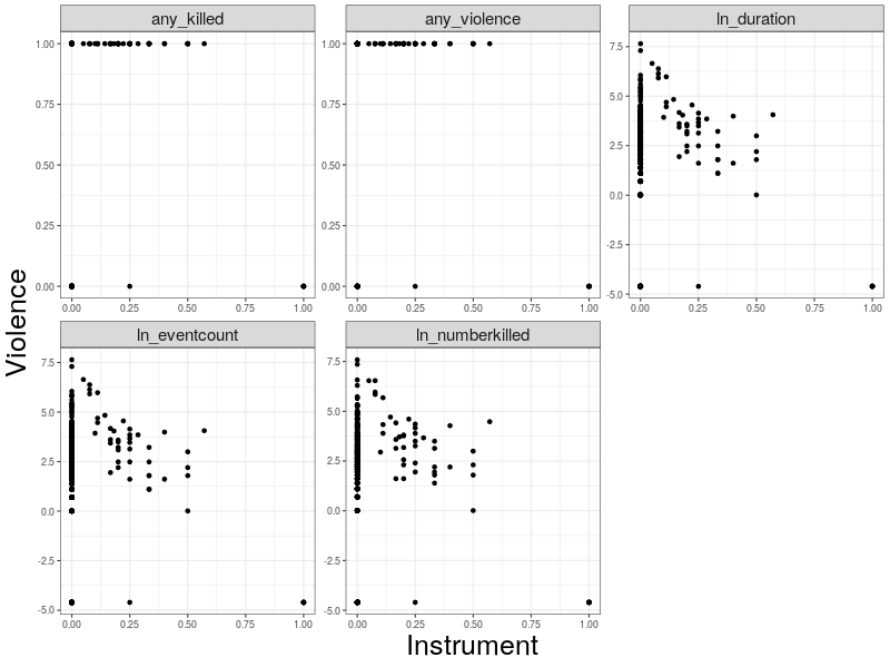

Figure 4: Religious Violence Across instrument

	Replication Any Killed	Original
Prop. Secular Win	−0.477	−0.477*
	(0.296)	(0.271)
Prop. Secular Clost Race	0.004	0.004
	(0.173)	(0.158)
Province FEs	Y	Y
Num. obs.	437	437

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Robust SEs clustered by cluster-district area, in brackets

Table 2: 2SLS LATE with CR2 SE estimation

`CR2` standard error estimation and the effect for *Any Killed* was no longer significant. The full table is reported in 9.

	Any Event	Event Count	Any Killed	Number Killed	Number Days
Secularist Close Win	−0.176*	−1.265*	−0.141	−0.769	−1.265*
	(0.092)	(0.654)	(0.089)	(0.647)	(0.654)
R^2	0.330	0.393	0.398	0.390	0.393
Adj. R^2	0.267	0.336	0.341	0.332	0.336
Num. obs.	59	59	59	59	59
RMSE	0.279	2.303	0.280	2.435	2.304

Robust SEs clustered by cluster-district area, in brackets

Table 3: Difference in Means Estimate

DIM estimates of ATE The authors also compute a difference-in-means (DIM) estimate of the ATE by selecting units with a single election. These results are replicated in table 3. To this end, the authors restrict the dataset to 59 years and joined-district points where there was a single close election between secular and Islamist candidates. I was able to replicate the original findings, with a significant negative effect for *Any Event, Event Count* and *Number of Days*. Table 10 shows the results calculated with `CR2` standard errors.

3.2 Design Checks

Non-zero First Stage A key component to an IV design is a non-zero first stage 2.2.3. In other words, the exogenous instrument, $Z_{i,t}$ must have a statistically significant effect on treatment, $D_{i,t}$. Table 4 replicates the original work where the proportion of close wins by secular candidates has a strong effect on the proportion of secular wins by any margin (eq 2).

	Prop. Secular Win
Prop. Secular Close Win	0.903***
	(0.123)
Prop. Secular Close Race	−0.098
	(0.103)
Num. obs.	437
F statistic	54.446
RMSE	0.346

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Robust SEs clustered by cluster-district area, in brackets
F-statistic reported for Prop.Secular Close Win

Table 4: First Stage Regression

	Any Event	Event Count	Any Killed	Number Killed	Number Days
Prop. Secular Win	−0.066	−1.165	−0.127	−1.404	−1.162
	(0.251)	(1.947)	(0.245)	(1.969)	(1.947)
Prop. Secular Close Race	−0.364**	−1.802	−0.313*	−1.696	−1.811
	(0.163)	(1.335)	(0.168)	(1.398)	(1.335)
Num. obs.	437	437	437	437	437
RMSE	0.301	2.535	0.355	2.904	2.540

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Robust SEs clustered by cluster-district area, in brackets

Table 5: Placebo Check — Can Secular Victory in Close Elections at Time t Predict Prior Violence

Exogenous Instrument In order to address the potential influence of reverse causality, the authors attempted to predict the previous outcome for a given joined-district $Y_{i,t-1}$. I was able to reproduce the authors' null result 5. However, it is important to note that the instrument can predict $Y_{i,t-1}$. The authors explore this correlation in 6, showing that close secular races tend to occur in low violence joined-districts. Further analysis, shows that the instrument does not predict a variety of pre-treatment covariates such as agricultural production, education, and civil infastructure 11.

Monotonicity In an effort to obviate concerns of sorting described in 2.2.3, the authors perform a density test that shows a null result (fig 5) [3]. While this cannot conclusively rule out sorting or attrition, a null result suggests that treatment is as good as random under the instrument (TODO: cite McCrary).

3.3 Mechanisms

The later section of the source, the authors perform explanatory analysis to illustrate potential mechanisms of secular MNA seats on religious violence. One explored avenue

	No Fixed Effects	Disctrict Cluster FE	Disctrict Cluster + Province-Year FEs
Secularist Close Race	-0.355^{***}	-0.386^{***}	-0.257^{***}
	(0.119)	(0.111)	(0.091)
Num. obs.	437	437	437
RMSE	0.371	0.337	0.294

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Robust SEs clustered by cluster-district area,
in brackets

Table 6: Correlation Between Close Secular/Nonsecular Elections and Violence at Time t-1

was the electoral accountability. The authors claim that secular leaders often include diminished religious violence as a campaign promise. The authors then predict that secular MNA candidates expect to suffer in future elections if religious violence does occur during their tenure.

To test their prediction, the authors estimate the causal effect of religious violence in the previous term on the proportion of secular MNA seats in the following election. I was able to reproduce these estimates in full (table 7)

OLS				
Prop. Secular (t-1) x Any violence	-0.115^{***}		-0.102^{**}	
	(0.040)		(0.043)	
Prop. Secular (t-1) x Event count (ln)		-0.018^{***}		-0.014^{*}
		(0.006)		(0.007)
Any violence	0.105^{***}		0.086^{***}	
	(0.028)		(0.024)	
Prop. secularist wins (t - 1)	0.038	-0.053^{**}	0.050	-0.030
	(0.043)	(0.026)	(0.050)	(0.031)
Event count		0.018^{***}		0.015^{***}
		(0.004)		(0.004)
Num. obs.	344	344	344	344
RMSE	0.139	0.138	0.126	0.127

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Robust SEs clustered by cluster-district area,
in brackets
Province Fixed effects omitted

Table 7: Mechanisms - Electoral Incentives

4 Extensions

4.1 Fuzzy RD 2 - Electric Bugaloo

ITT

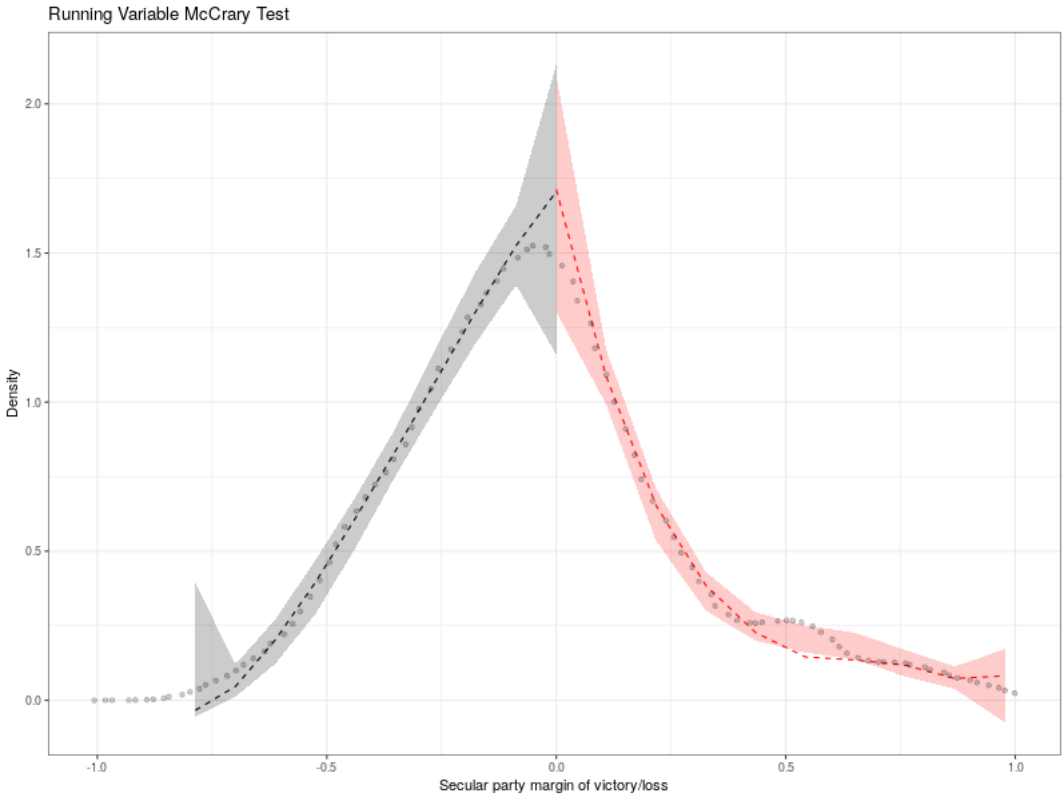

Figure 5: McCrary Density Test

4.2 A case for Binary Treatment

5 Discussion

6 Supplementary

	Any Event	Event Count	Any Killed	Number Killed	Number Days
Prop. Secular Win	-0.660^{***}	-4.654^{**}	-0.477^{*}	-3.266	-4.700^{**}
	(0.215)	(1.744)	(0.275)	(2.144)	(1.767)
Prop. Secular Close Race	0.031	0.837	0.004	0.281	0.947
	(0.085)	(0.875)	(0.161)	(1.276)	(0.884)
Num. obs.	437	437	437	437	437
F statistic	9.831	11.564	4.500	2.597	10.839
RMSE	0.357	3.025	0.389	3.194	3.114

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Robust SEs clustered by cluster-district area, in brackets
F-statistic reported for Prop. Secular Win

Table 8: Instrumental Variable Results

	Any Event	Event Count	Any Killed	Number Killed	Number Days
Prop. Secular Win	-0.660^{**}	-4.654^{**}	-0.477	-3.266	-4.700^{**}
	(0.229)	(1.866)	(0.296)	(2.309)	(1.892)
Prop. Secular Clost Race	0.031	0.837	0.004	0.281	0.947
	(0.086)	(0.907)	(0.173)	(1.362)	(0.916)
R^2	-0.580	-0.486	-0.164	-0.169	-0.482
Adj. R^2	-0.602	-0.507	-0.180	-0.185	-0.503
Num. obs.	437	437	437	437	437
RMSE	0.357	3.025	0.389	3.194	3.114

Robust SEs clustered by cluster-district area, in brackets

Table 9: IV with CR2 SE estimation

	Any Event	Event Count	Any Killed	Number Killed	Number Days
Secularist Close Win	-0.176^{*}	-1.265^{*}	-0.141	-0.769	-1.265^{*}
	(0.095)	(0.670)	(0.092)	(0.663)	(0.670)
R^2	0.330	0.393	0.398	0.390	0.393
Adj. R^2	0.267	0.336	0.341	0.332	0.336
Num. obs.	59	59	59	59	59
RMSE	0.279	2.303	0.280	2.435	2.304

Robust SEs clustered by cluster-district area, in brackets

Table 10: DIM ATE with CR2 SE Estimation

	Area	Pacca Prop. HHs	Elecriticy	Gas	Total Literacy	Female Literacy	Primary Schools
Prop. Secular Win	0.104	0.070	−0.073	−0.002	−0.044	−0.045	0.000
	(0.528)	(0.124)	(0.170)	(0.028)	(0.054)	(0.045)	(0.000)
Prop. Secular Close Race	−0.319	−0.125	−0.080	−0.021	−0.047	−0.024	−0.000*
	(0.373)	(0.084)	(0.127)	(0.027)	(0.039)	(0.036)	(0.000)
Num. obs.	435	421	421	417	425	425	419
F statistic	18.812	1192.829	180.434	1414.017	348.226	504.437	4.095
RMSE	0.844	0.253	0.263	0.103	0.125	0.113	0.000

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Electoral outcomes for 1988, 1990, 1993, and 1997 are used to predict (as a falsification test) census outcomes measured in 1981; electoral outcomes for 2002 and 2008 are used to predict census outcomes measured in 1998. Sample sizes vary somewhat across models due to missingness in some census data. Missingness is minimal and appears to be unsystematic.
Robust SEs clustered by cluster-district area, in brackets
F-statistic reported for Prop. Secular Win

Table 11: Lagged Joined-District Characteristics Measured in Census of Pakistan 1981 and 1998

References

[1] Matias D. Cattaneo, Nicolás Idrobo, and Rocío Titiunik. A practical introduction to regression discontinuity designs. Nov 2019.

[2] Guido W. Imbens and Thomas Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615 – 635, 2008. The regression discontinuity design: Theory and applications.

[3] Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698 – 714, 2008. The regression discontinuity design: Theory and applications.

[4] Stephen L Morgan and Christopher Winship. *Counterfactuals and causal inference*. Cambridge University Press, 2015.

[5] GARETH NELLIS and NILOUFER SIDDIQUI. Secular party rule and religious violence in pakistan. *American Political Science Review*, 112(1):49–67, 2018.