Estimating Support for Extremism and Its Correlates: The Case of Pakistan

Tarek Al Baghal
Institute for Social and Economic Research, University of Essex

The extent of support of extremist ideology is a major area of concern for both policy makers and academic researchers. Identifying the extent and correlates of a difficult to measure concept such as extremist ideology is often limited by the use of a single imperfect indicator. This paper outlines one approach, latent class analysis (LCA), to overcome this issue and uses the example of estimating support for such ideology in Pakistan. Using survey data from Pakistani men, the level of support is estimated using LCA employing several indicators related to extremism. The results suggest that although most Pakistanis are not supportive of extremist ideology, a substantively important portion of men are supportive. LCA also allows for class assignment, which is useful for understanding covariate relationships with the latent variable. Based on the results of the LCA, respondents are assigned to different classifications of extremist support, and a continuation-ratio logistic regression model is employed allowing for more covariates to be examined. The results suggest that there are a number of characteristics important in influencing support within this subset of the population. In particular, younger and less educated men are more likely to support extremism ideology. The results suggest a potentially useful methodology in understanding extremism, as well as a greater understanding of the problem of extremist support.

Keywords: extremist ideology, Pakistan, public opinion, latent class analysis

INTRODUCTION

Extremism and extremists are issues facing governments around the world, but identifying the extent of support for such ideology is difficult to ascertain, as are the correlates that drive this support. Support for extremist ideology is among many concepts and characteristics that researchers and policy makers alike would like to measure, but which are difficult if not impossible to capture. Attempts to measure support of extremism are often limited by the use of a single imperfect indicator. However, like many variables of interest, the construct of interest may be an underlying latent quality that is difficult to measure using single measures. As well as extremist ideology, these latent variables include concepts such as intelligence quotient (IQ), depression, or social class (Garret-Mayer and Leoutsakos 2010). One method that has been used to overcome the problem of single imperfect indicators is to use multiple indicators and employ a method to examine shared variance to estimate the latent variable. For indicators and latent variables that are categorical, latent class analysis (LCA) is the appropriate tool to use (McCutcheon 1987).

It is proposed that LCA is a possible tool to measure support of extremist ideology and identify its correlates. The current research explores the reasons LCA is a potentially useful tool, its extensions and potential limitations, using data from one of the most focused on areas for extremist beliefs, Pakistan. The goal is to suggest a potential methodological framework demonstrating the utility of LCA in analyzing this domain using exploratory methods. LCA has been used for a large number of purposes, including risk of drug use (James, McField, and Montgomery 2013), tolerance of nonconformity (McCutcheon 1985), groupings of political revolutionaries (Beissinger 2013), and the extent of measurement error (McCutcheon 2002), among others. LCA has also been used to examine the similar issue of fundamentalist belief in the Muslim world previously, although the focus of the indicator on was on religious belief rather than extremist actions or ideologies specifically (Blaydes and Linzer 2008). LCA can estimate the extent the latent variable exists in the population and the probabilities for each of the latent categorizations that a member of the population has given observed characteristics. These probabilities allow for a useful aspect of LCA, the assignment to categorizations which can be used to identify correlates of the underlying variable (Goodman 2007). The current research focuses on methods to estimate and examine support for extremist ideology, beginning with an overview of extremism in Pakistan and then LCA, then an analysis of the data employing four indicators in the LCA.

EXTREMISM IN PAKISTAN

Extremism in Pakistan and in the Islamic world has become an increasingly problematic issue both within and outside of the country. There is a desire to understand the causes of support for extremist ideology. Some researchers, frequently from the *clash of civilizations* perspective, seek to explain the motivations of individuals for supporting extremist ideology. This view argues that conflict and confrontation between the West and Islam is fueled by dislike of cultural values (Huntington 1996; Tessler and Robbins 2007). There also appears to be a motivation due to the perceptions toward the national government (Fair 2004; Tessler and Robbins 2007). However, other research has found that opinion towards the government or officials do not have much influence on support for extremist groups (Shapiro and Fair 2010). In addition, other potential psychological reasons underlie why some people support such ideology. These include: emotional vulnerability; dissatisfaction with their current ability to participate in and affect the political and societal spheres; identification and/or kinship with victims of perceived injustices; the belief that terrorism is not immoral; and a sense of reward from belonging to the group (Horgan 2008).

Additionally, some personal characteristics appear to be related to the support of extremism. In Pakistan, where there is a sizeable Shiite Muslim minority compared to the Sunni Muslim majority, extremists are largely sectarian (Haqqani 2005, 2006; Kaltenthaler, Miller, Ceccoli and Gelleny 2010). In particular, Sunni groups are more plentiful and more likely to make violent attacks than other sects or groups (Haqqani 2005). Shi'a extremism is on the decline (Kaltenthaler et al. 2010), with Sunni sectarian groups frequently attacking Shi'a communities, due to the view of the Shi'a as heretical (Haqqani 2006).

Those with lower education are also found to display higher average levels of support for extremism (Fair and Shepherd 2006; Haddad and Khashan 2002; Kaltenthaler et al. 2010; Tessler and Robbins 2006). Although some research finds that poorer respondents are less likely to support extremism (Blair, Fair, Malhotra, and Shapiro 2013; Fair and Shepherd 2006) income is not uniformly found as significantly influencing the support for extremism (see e.g. Haddad and Khashan 2002; Shapiro and Fair 2010). In other research, the related construct of religious fundamentalism is higher among employed men, but lower among employed women (Blaydes and Linzer 2008). Further, data on those who have carried out terrorist acts suggest that neither education nor economic status has a direct causal effect on participation in terrorism, but younger men are more likely to engage in terrorist acts (Krueger and Maleckova 2003). Blaydes and Linzer (2008) also find that men are far more likely to have fundamentalist religious beliefs than women. These findings are consistent with other research which suggests that supporters

of extremists largely consist of men, and appears to be essentially true for groups in Pakistan as well. In Pakistan, there is higher support among men for suicide bombing (Kaltenthaler et al. 2010), with only two women perpetrators in the current era of attacks (Khan 2010).

In addition to identifying the characteristics of those who support extremist ideology, it is important to identify the *extent* of support for such ideology. Prior research indicates that such extremists have some, although limited, support from the population in various nations. Moreover, most Pakistanis say that terrorism attacks by al-Qaeda are not justified; again, this is with the caveat that a sizable minority, around 15% report that it is either "often" or "somewhat" justified (Kaltenthaler et al. 2010). However, past research has been unable to provide estimates of the likely population size these groups can recruit from or rely on for popular support. Part of this problem is likely due to limitations and use of data available. Specifically, only a single indicator of support is used in each of the previous analyses, and direct measurement of support of extremism is fraught with difficulty (Bullock, Imai, and Shapiro 2011; Fair, Malhotra, and Shapiro 2012). Bullock et al. (2011) in particular discuss the problem of the single indicator to measure support for extremist ideology in Pakistan specifically, and use multiple indicators and a latent model (item response theory) to obtain a better measure.

The LCA in Blaydes and Linzer (2008) does suggest that 33% of those in Muslim-majority countries have fundamentalist beliefs, and particularly focus on these beliefs among women. While possibly related, these fundamental religious beliefs are not indicative of support of extremist ideology necessarily, and the while important, women take part in almost none of the violent action perpetrated by extremist groups. Further, the goal of most of the previous multivariate analyses on the topic of extremism was to estimate differences between characteristics of respondents or regional distribution, rather than estimate population size of support for extremism (Bullock et al. 2011; Fair et al. 2012; Haddad and Khashan 2002). However, estimating population size is not only of academic importance but also a practical once, as it can suggest the extent of the problem and potentially inform policy on how best to confront the issue.

LATENT CLASS ANALYSIS AND ASSIGNMENT

The current research suggests an alternative route in studying extremism in Pakistan, and more generally to the study of support for extremist ideology in many areas by employing LCA. By using a data set with multiple indicators, the limitations of any one indicator is potentially overcome, and increases the chance of better measurement of the construct. This technique allows not only for possible estimation of population sizes, but also the ability to assign

people to classes in order to identify covariates related to inclusion in those categorizations.

In addition to estimating population sizes and classifications, external observed variables may be used to understand the impact of these covariates. External grouping (i.e. non-indicator covariate) variables can be included in the in the latent class model in a straightforward manner. These additional categorical variables can influence only differences in the estimated latent class sizes, or can also influence the conditional response probabilities (Clogg and Goodman 1984). One potential limitation to using covariate groupings such as region, age and education directly in the LCA is the extent that these covariates could be used. Covariates have to be few in number and partitioned into broad categories to ensure that the data does not become overly sparse and the model is estimable. A potential way to overcome this problem is to assign respondents to one of the estimated latent classes based on their response distributions to the indicator variables, and then use other methods to examine covariates (Goodman 2007). This assignment can be done using the conditional response probabilities, converting these into posterior modal probabilities for each latent class given a specified response pattern, assigning each respondent to a class based on their observed responses.

However, this method potentially leads to biased and inconsistent estimates of the relationships between the latent classes and the external variables (Bolck, Croon and Hagenaars 2004). An alternative method would be to model the effect of external covariates on the estimated latent classes simultaneously in one step; however, this method also has problems (Asparouhov and Muthén 2014a; Vermunt 2010). Vermunt (2010) outlines an improved three-step method which corrects for classification errors and the bias. Using this improved three-step method to correct for the effect of classification error may still lead to somewhat underestimated standard errors for the third-step model (Bakk, Oberski, and Vermunt 2014).

While the use of such corrections is ideal, there are at the moment some potential limitations. First, when using the three-step correction, at least in some statistical packages, only multinomial logistic regression models are allowed to estimate covariate effects when there are multiple latent classes (Asparouhov and Muthén 2014a). The multinomial model is also the only one available in one-step programs as well (Linzer and Lewis 2011). The multinomial model does not account for any ordering or progression of the categories that may occur, and only allows for category comparisons to a baseline rather than any possible cumulative comparisons. Second, there is some debate about how to correct for the standard errors in the three-step model (Asparouhov and Muthén 2014b). Given the present issues in the implementations of these corrections, especially the lack of ordering possible in modeling classifications, the current research relies on the three-step method outlined in Goodman (2007), with the acknowledgement of the potential

biases this method introduces. It should be noted, however, that the biases in coefficients are in a downward manner; that is, relationships are underestimated rather than overstated. Also, possible biases in standard errors are relatively smaller and are minimized when the sample size is large or the entropy of the model is high (Bakk et al. 2014).

DATA AND METHODS

Data was collected as part of the Gallup World Poll, which has collected numerous surveys in many countries over the course of a number of years. All surveys are cross-sectional; different respondents are sampled at each data point. There were five separate surveys conducted in Pakistan: June 2007, June 2008, October 2008, December 2008, and May 2009. The Pakistani data was collected by personal interview surveys based on a national area probability sample, with those 15 years old or older eligible for selection. A complete description of the methodology and sampling can be found elsewhere (Gallup 2013).

Gallup data include a set of questions asked only to Muslim-majority countries, including Pakistan. Of particular interest are questions that may be indicative of support for extremist ideology. It is not possible, however, to directly ask survey respondents if they support extremists such as the Taliban or al-Qaeda. Few, if any, would openly admit to a stranger that they support a group that the government and external military forces are actively seeking out. Rather, support for extremism needs to be measured by proxy questions asking about support for beliefs and actions of extremists. The first three of these questions regards support for terrorist attacks, either specific events (e.g. September 11th attacks) or more general (e.g. attacks where civilians are the target). The first two questions also tap into anti-American sentiment, asking specifically about attacks on Americans. A final question, an indicator of intolerance toward Shiites is added to the model to capture the additional hallmark of extremism in Pakistan (Haqqani 2005, 2006). Pakistani extremists are intolerant of all non-Sunni Muslims, and consider Shiite Muslims to be heretics (Haqqani 2005). Although Pakistani extremists are also intolerant of other non-Muslims, questions about these groups may not discriminate well because of widespread unpopularity among the populace. The exact wording for each of these questions, which is the same for all waves, can be found in Appendix A.

All respondents to these surveys were self-identified Sunni Muslims, or identified simply as "Muslim" with no sect affiliation specified (assumed to be majority Sunni Muslims) as extremism is largely limited to the Sunni community (Haqqani 2005). Moreover, only data from men are analyzed, since many extremists, including those in Pakistan, are mainly men (Krueger and Maleckova 2003; Khan 2010)

and the desire is to identify those most likely to act on extremist ideology. Given the closeness in time of the different waves of the surveys, time is not considered to be a factor, and all data are combined over all collection points. Since each of the four questions of interest were not asked in all five waves of data collection, initial analysis considers differing sample sizes for the responses to each question. Each of the four questions were asked using five-point response options, ordered as scales going from opinions that are least supportive of extremist group views to most supportive. Table 1 presents the response distributions for the four questions.

Table 1 Responses to Four Indicators of Extremism

	1=Least Supportive	2	3	4	5= Most Supportive	Don't Know
Hope from Terrorist Attacks on US (n=1206)	33.7	15.6	16.9	10.5	4.1	19.3
Sept 11 Justified (n=1915)	55.0	18.9	19.6	4.3	2.1	0.1
Attacks on Civilians Justified (n=2326)	54.9	16.2	20.5	4.3	3.7	0.4
Attitudes Towards Shiite Muslims (n=1505)	24.5	25.5	25.1	7.8	9.2	8.0

Note: Numbers in table are percentages of respondents giving a certain response options

The data indicate that the majority of the respondents have views that are contrary to those of extremists. A significant number of men, however, do have beliefs that are more consistent with these extremists. Between 2% and 9% of male respondents gave the most extreme response on these scales. In addition, those giving responses directionally supportive of extremists is much higher. Extremist leaning responses, including those at the upper end of these scales beyond the mid-point, make up between 6% and 17% of all responses. While most responses clearly are not supportive of extremists, a sizeable proportion appears to be supportive.

A problem arises as to which measure, if any, of the above best captures support for these extremists. Additionally, survey questions are frequently measured with error (Lessler and Kalsbeek 1992). One way to overcome the problems of error and the limitations of any one measure is to use multiple indicators to estimate the construct of interest. Given the categorical nature of the data, latent class analysis (LCA) allows for an estimation of the level of error in each measure as well as estimating the true population size for the construct of interest, in this case, supporters of extremism. The LCA are estimated using Mplus 6.12 (Muthén and Muthén 2010).

EXTREMIST GROUP SUPPORT AS LATENT CLASSES

The only cross-sectional surveys that asked all four questions in one fielding are June and December 2008, and are thus the only data points available for further analysis in the LCA. Since each of the four indicators had five category responses, this leads to a (5^4=) 625 cell table to be analyzed. Such a large table frequently leads to problems of sparseness, i.e. problems with numerous zero or small number of observations in each of the cells of the full table (Agresti 2012). In order to counter these problems, as well as to be consistent with analysis of data using similar scales, the 5-point response scales for all four indicator items were collapsed to three point classifications. The mid-point of the original scale remained the mid-point in the new classification, with the valences of the original scales collapsed into consistent categories (i.e. 1 and 2 on the original scale collapsed together in the new classification, as were 4 and 5). Only respondents that outright refused to answer or had broken-off prior to all four questions being asked are not included in the analysis. All indicators are coded with the midpoint (=2) indicating neutral/ ambivalent, with 1 coded as positions more oppositional to extremism, and 3 coded as more supportive of views of extremist ideology. This rescaling of the four indicators leads to 743 observations in the (3^4=) 81 cells in the multi-dimensional data table. Although this collapsing is necessitated to reduce the sparseness in the table and make estimation feasible, doing so also makes the interpretation of the classes more difficult. In particular, by collapsing, the views of weaker and stronger strength opinions are combined.

Models employing two, three, four, and five latent classes were estimated as these were thought of as being the possible underlying categorical structure for support. It was possible that support was structured as a dichotomy, or with three categories with the middle category of those neither opposing nor supporting extremist ideology. The four-class model was proposed as having anchors most closely related to oppositional and supportive of extremist ideology, with two middle categories, representing those leaning to either opposition or support. The five-class model could be two points leaning oppositional or supportive each, with a middle category of those neither opposing nor supporting extremist ideology. As can be seen in Table 2, neither the two or three class model fit the data adequately. However, the four class model shows adequate model fit, as does the five class model. Of these two, the p-value is largest and the Aikake and Bayesian Information Criterions are smallest for the four class model and entropy highest indicating four classes provide the best model fit (Agresti 2012; Celeux and Soromenho 1996; McCutcheon 2002), and therefore is the final model used in the following analyses.

Table 2 Model Fit for Different Latent Class Model Specification

Latent Class Model	G^2	DF	p-value	AIC	BIC	Entropy
2 Classes	120.184	63	0.000	4863.396	4941.778	0.735
3 Classes	92.313	54	0.001	4853.523	4973.401	0.711
4 Classes	50.620	45	0.266	4829.829	4991.204	0.774
5 Classes	41.008	36	0.260	4838.218	5041.089	0.653

While events may have occurred in between these months in 2008 that may have altered support, such as the resignation of President Pervez Musharraf in August or the bombing of the Marriot Hotel in Islamabad in September, there is no apparent impact on the LCA. Both may have had importance in influencing extremism: Musharraf strongly opposed extremist groups, often using military force and opposing any negotiations, while the bombing killed a number of Pakistanis and was blamed on Islamic extremist groups. However, the four-class model used fits equally as well when examining only July ($G^2_{45} = 53.68$, p = 0.176) or only December data ($G^2_{45} = 53.34$, p = 0.184) with similar class and conditional probability estimates (not shown). Given the lack of apparent differences, combining the data across surveys seems acceptable. Further, none of the bivariate residual covariances are significant in the four-class model, suggesting that the local independence criterion is likely met (see Appendix B). It is noted that the measures obtained may underestimate significant relationships (Oberski, van Kollenburg and Vermunt 2013). However, current limitations in the programming do not allow for further tests, and these measures have been used as accepted indicators elsewhere (Tay, Newman, and Vermunt 2011). The model estimates from this four class model are presented in Table 3.

As can be seen from this model, the largest estimated proportion of the population is the class labeled as those "oppositional" to extremist ideology, with the smallest proportion being those labeled "supporters" of extremist ideology. As noted, the necessitated collapsing of categories given the data also makes the interpretation of these classes more tentative, the response probabilities show categorizations of beliefs seemingly consistent with oppositional to supportive of extremist ideology. Those in the "oppositional" class consistently have high probability of responses contrary to the positions supported by what is known of extremist ideology. Conversely, those in the "supportive" class consistently have high probability of responses in-line with the above literature on extremist ideology. The middle two classes are in between those more "oppositional" or

"supportive" of extremist ideology. These are seen as those "leaning oppositional" (t=2) and those "leaning supportive" (t=3). The first of the two middle classes shows overall highly oppositional views toward extremism, except in showing some belief that attacks on the US provide hope for improving American policies. The second middle latent class consistently shows response probabilities more neutral, with a slightly higher probability generally of oppositional responses. These two middle classes are estimated to make up slightly more than 47% of the population. These two classes and the oppositional class share similar responses to the question about attitudes toward Shiite Muslims. As expected, only those supportive of extremism have high probabilities of negative views of Shiites, with the remaining three classes having the similar conditional response probabilities.

Table 3 Latent Class Estimates for Support of Extremist Groups, Unconstrained 4-Class model

Measure	Response	Oppositional	t = 2	t = 3	Supportive
Hope from Terrorist Attacks on US	Little Hope	0.92	0.00	0.32	0.15
	Neither	0.08	0.57	0.65	0.20
	Some Hope	0.00	0.43	0.04	0.65
Was Sept. 11 Justified?	Little Justification	0.89	0.88	0.38	0.00
	Neither	0.05	0.12	0.54	0.37
	Some Justification	0.06	0.00	0.08	0.63
Attacks on Civilians Justified?	Little Justification	0.92	0.92	0.06	0.00
	Neither	0.06	0.08	0.75	0.31
	Some Justification	0.02	0.00	0.20	0.69
Personal View on Shiite Muslims	Positive	0.55	0.58	0.61	0.12
	Neither	0.25	0.25	0.25	0.39
	Negative	0.18	0.18	0.18	0.50
Estimated Class Size		0.470	0.323	0.150	0.037

Note: G^2_{45} =50.60, p>0.25, n = 743

Although the "supportive" class makes up the smallest proportion of the population, the number this class possibly represents is still substantial. Three-plus percent of an entire population can equate to several million people, which can present serious problems to both national and international governments. This number also begins to give some perspective to the potential problem. As of 2012, there are 62,140,663 men 15 years old and older in Pakistan (Central Intelligence Agency [CIA] 2013). Taking 3.7% of this equals to 2,299,204.53 men, a sizeable figure. Further, although this is not an exact estimate, given potential limitations to the measures, LCA, and the survey method, there are reasons to expect that this number of more than 2 million may be underestimated. Most directly, those most supportive of extremist ideology are also likely to be the most reticent to respond to a survey request. More likely, this 3.7% is an estimate of those at risk of being most supportive of extremist ideology.

Additional analyses can further shed light on the scope of possible support for extremist ideology, with potential for guiding policy in Pakistan, such as for groups to focus attention on. Although covariates can be included directly into an LCA, as noted this is limited due to the possible sparseness adding these can cause. A potential way to overcome this problem is to assign respondents to one of the four latent classes based on their response distributions (Goodman 2007). This assignment can be done using the conditional response probabilities in Table 3, converting these into posterior modal probabilities for each latent class given a specified response pattern, then assigning each respondent to a class based on their observed responses. Although using this three-step procedure leads to potential bias in model estimates as it fails to account for uncertainty in classification, as noted, other methods as currently implemented limit the ability to select the most appropriate analytic model. Further, since the bias is generally to underestimate relationships, results from any model are more likely to be conservative rather than overestimates (Bolck et al. 2004).

Since the entropy is highest for the four class model, the classification from this model is likely superior to other model specifications. The classification table for the 743 cases used in the latent class analysis is presented in Table 4, along with the comparable latent class size estimate. The results show the possible error introduced by using deterministic assignment procedures. The diagonal shows that all of the average latent class probabilities for the most likely class membership for the expected class is above 0.8, with the lowest for the third class, "leaning supportive" (t=3).

Table 4 LCA Classification Table

	t=1	t=2	t=3	t=4	Number in assigned class
t=1	**0.925**	0.026	0.049	0.000	349
t=2	0.090	**0.882**	0.027	0.000	259
t=3	0.034	0.108	**0.812**	0.045	109
t=4	0.009	0.027	0.094	**0.870**	26

Using the assignment of individuals to the identified classes leads to a single outcome variable of support for extremist ideology, a four-category ordered measure ranging from "oppositional" to "supportive". The ordered nature is of potential importance, as there may be a progression in belief from more oppositional to supportive in general. Employing the multinomial logistic models implemented in procedures to adjust for classification error either simultaneously or in a corrected three-step method do not account for this ordering (Linzer and Lewis 2011; Asparouhov and Muthén 2014a). Rather, to account for the ordering a continuation-ratio logistic regression model is employed (Agresti 2012). This model compares each of the assigned class to all of lower assigned ordered classes (i.e. t=2 to t=1, t=3 to t=1, 2 and t=4 to t=1, 2, 3). The comparison of the "supportive" class (t=4) to the remainder of the sample importantly predicts the likelihood of being in the "supportive" class relative to any other class, which is possibly of most concern.

Based on current understandings, several groups were examined more in-depth by including external variables into the base model. There are possible regional differences in Pakistan, with the Northwest Frontier Province (NWFP) being identified potentially more problematic and home of more extremists than other regions (Rashid 1999; Kaltenthaler et al. 2010; Kemp 2010), although Bullock et al. (2011) suggest Punjab as having higher levels of support for extremist ideology. The NWFP is compared to the other regions of Baluchistan and Sindh, while using Punjab as the baseline comparison. Additionally, age and education are possible important correlates of support for extremism (Fair and Shepherd 2006; Kaltenthaler et al. 2010; Krueger and Maleckova 2003; Tessler and Robbins 2007). Age is measured as a continuous variable, while education is measured categorically, as that is how it was captured in the survey, comparing those with eight years of education or less to those with more education.

Other variables were also sought out that might explain the position on the oppositional-supportive scale (complete question wordings can be found in

Appendix A). First, the belief that Shari'a should be the only source of law is expected to increase the probability of support of extremist ideology, as increased religiosity and desire for religion to be reflected in government is a hallmark of such extremism (Fair and Shepherd 2006; Haqqani 2006). Belief that Pakistan's leadership is taking the country in the right direction is also hypothesized to decrease support for extremist ideology, as dissatisfaction with the political process may be an important psychological factor (Horgan 2008). The same is possibly true about identification with victims, real or imagined, by powers such as the United States (Horgan 2008). As such, it is expected that belief that the West respects the Muslim world will also decrease the likelihood of support for extremist ideology. Finally, income (measured in quintiles of the population) was included in the model, but similar to some previous findings, income does not have a significant effect (Haddad and Khashan 2002; Krueger and Maleckova 2003). Income is also a large source of data missingness (212 of the 743 cases missing) and is not included in the final model.

Other variables have some missingness, causing some cases to be dropped from the final model. The question about the West's respect for the Muslim world has 178 missing cases, leading to the most dropped cases in the final model. There are 99 missing responses to the question about the direction Pakistan's leadership is taking the county. The question on Shari'a as the only source of law has 53 missing cases. Dropping any of these alone or in combination from the final model does not change the substantive results of the other variables in any way (not shown), and given the possible importance, are retained in the final analysis. The only other variable with missingness is education, with only one missing case. In combination, including these variables led to 500 complete cases for analysis in the final model. The cases dropped were near proportional to the total across the four classes. Of the 243 cases, 106 (43.6%) are from the "oppositional" (t=1) class, 95 (39.1%) from the "leaning oppositional" (t=2) class, 39 (16.1%) from the "leaning supportive" (t=3) class, with the proportionally least dropped from the "supportive" (t=4) class (3 cases, 1.2%). The final model using complete cases is presented in Table 5.

The results in Table 5 suggest that few of the variables explain change between categories among the lower assigned classes. Believing that the West respects the Muslim world leads to the prediction that a man is more likely to be in the "oppositional" class compared to the "leaning oppositional" class (t=2). Similarly, those in the Sindh and Baluchistan provinces are less likely to be in the "leaning oppositional" class compared to the "oppositional" class. Education is not significant in predicting membership in either of the two middle classes, "leaning oppositional" and "leaning supportive". Belief that the West respects the Muslim world leads to a decreased likelihood of being classified in the "leaning

supportive" class relative to those "leaning oppositional" and "oppositional" classes. Older respondents are less likely to being in the "leaning supportive" category compared to the two more "oppositional" to extremist ideology classes, and again people in Sindh province shows a lower probability of classification in the "leaning supportive" class. This portion of the model is the only level that finds a significant effect for being in the NWFP as well, with people in NWFP being more likely to be in the "leaning supportive" category relative to the more "oppositional" classes.

Table 5 Continuation-Ratio Logistic Regression on Assigned Extremist Support

Variable	Class 2 (s.e.)	Class 3 (s.e.)	Class 4 (s.e.)
Shari'a Only Source	-0.247 (0.240)	-0.313 (0.310)	-0.096 (0.565)
Leadership	-0.048 (0.238)	-0.273 (0.291)	-1.212** (0.598)
Respect from West	-1.512** (0.285)	-0.619** (0.309)	-2.602** (0.553)
Low Education	-0.328 (0.236)	0.382 (0.285)	1.593** (0.609)
Age	-0.006 (0.007)	-0.017* (0.009)	-0.043** (0.018)
NWFP	0.086 (0.307)	0.548* (0.329)	-0.072 (0.861)
Sindh	-1.690** (0.296)	-1.229** (0.439)	-0.172 (0.665)
Baluchistan	-1.162** (0.426)	-0.379 (0.523)	-1.126 (0.700)
Intercept	1.924** (0.462)	-0.878* (0.527)	-1.118 (0.936)

Note: n=500, G^2_{24} = 153.71, *p<.10,**p<.05

In examining the final column of Table 5, it is possible to identify the factors influencing those most likely to be in the "supportive" of extremist ideology class. First, unlike for other categories, provincial differences are not found. That the NWFP is not significantly different from the other provinces is contrary to expectation. In addition, the effect of belief in the necessity of Shari'a law is not

significant, consistent with Shapiro and Fair (2010) finding that attitudes towards Shari'a does not discriminate support for extremist groups well.[1] Unlike Shapiro and Fair (2010), significant effects are found, however, for the effect of beliefs that the leadership of Pakistan is taking the country in the right direction and that the West respects the Muslim world. Holding a negative view of the leadership of Pakistan leads to an increased probability of being classified as "supportive" of extremist ideology, as does having a negative view of the West's respectfulness toward the Muslim world. Further, those with low education and younger men have an increased estimated probability of being in the "supportive" class compared to the other classes. These findings may be of particular interest given the relative low education and youth of the population: Pakistani males average eight years of schooling and the median age for men is 21.9, with the large majority of the male population being under 40 (CIA 2013).

DISCUSSION AND CONCLUSIONS

The current research presents several important findings based on data from the Gallup World Poll. First is that given the right set of questions and appropriate statistical techniques, estimation of population size of at those possibly supportive of extremist ideology is possible. The use of Pakistani public opinion data and the use of latent class analysis suggest a possible methodology that can be employed in a number of other countries of interest where violent extremism is a problem or potential problem. Based on information from these models, respondents can be assigned to classes and additional covariates used to identify potential causes of extremist support. Second, it appears that it is not only the importance of questions related to terrorist acts, such as attacks on civilians generally, Americans, or support for the attacks on September 11, 2001, but also the importance of attitudes towards different Islamic sects that determine support for extremist ideology. In particular, Sunni Muslims' attitudes towards Shiite are important in differentiating extremist-leaning men from others. That this indicator has similar measurement properties and conditional response probabilities for all groups except those "supportive" of extremist ideology suggest its importance as a marker of potential extremism.

Third, the findings suggest that overall, most Pakistani men are not overly supportive of extremist ideology. Still, although only slightly more than three percent are estimated as "supportive" of extremist ideology, an additional fifteen percent are classified as "leaning supportive" of extremist ideology. Care may be needed to ensure actions, events, or communications do not sway these men into the supportive category. Another important finding is that surprisingly, those in the Northwest Frontier Province do not appear different in belief than those in other parts of the country. However, this may be due to differential survey response,

with those more supportive of extremism both being more likely to reside in the NWFP and more likely to refuse to participate in the survey, leading to decreased estimates of support in this region.

Conversely, age and education are both important covariates, with younger and less educated men more likely to be classified as "supportive" of extremist ideology than those who are older and better educated. Attitudes were also examined, and results show the effect of beliefs about Pakistani leadership and whether the West respects the Muslim in regards to the level of support. Importantly, those with a poor view of the Pakistani leadership and the West's attitude toward the Muslim world are more likely to be assigned to the "supportive" of extremist ideology relative to other classes.

These findings suggest several implications and continuing problems with extremism in Pakistan, which can help not only future research but also help guide pragmatic policy decisions in Pakistan. For example, knowing the approximate extent of the problem potentially allows for better allocation of resources. Additionally, efforts of outreach may be better focused on young and low educated people across the country, rather than on other groups, or focusing on only one province. Examination is needed of whether there is greater nonresponse bias or measurement error in some areas, such as the NWFP, where more extremist-minded men may be more likely to not respond to a survey or misrepresent their true beliefs.

Although this research suggests the usefulness of surveys and LCA as methods to identify support for extremist ideology, there are some limitations with the current research that should be noted. First, this research is largely exploratory; a more confirmatory approach is useful and is possible using LCA (Finch and Bronk 2011). Second, the latent classes are not necessarily perfect indicators of support for extremist ideology. Further, given the nature of the survey design, the overall number of cases available for use was someone limited. This contributed, along with generally small portions of the population endorsing extreme beliefs, to sparseness in the data table. In part to overcome this sparseness, response categories had to be collapsed. Given the collapsing of categories of those who hold beliefs of varying strengths, it is possible that the size of the "supportive" (or any) class is overstated. Hopefully future research will be able to overcome issues of table sparseness when estimating population sizes. Still, the facts that the indicators used to estimate the latent classes "load" in a consistent manner, and that findings about the effects of the correlates identified are consistent with expectation lends support to the efficacy of the model.

Third, the assignment procedure used here in conjunction with the LCA has been shown to lead to underestimated relationships between the classifications and covariates (Bolck et al. 2004). Procedures allow for simultaneous estimation of the

LCA and covariate model (e.g. Linzer and Lewis 2011) or corrections to the three-step procedure (Asparouhov and Muthén 2014a; Vermunt 2010). Further, standard errors for models using assigned classes may be underestimated in certain cases (Bakk et al. 2014). However, implementations of correction methods in available statistical packages largely limit the models available to identify relationships between classes and covariates. In particular, taking into account the potential importance of ordering is not widely available; however, the expectation is the coefficients are biased downward rather than overstated. While the standard errors may also be biased somewhat downward, this is more of a problem when the sample is small and/or entropy is low (Bakk et al. 2014). Although the sample size here is not large, the LCA used was selected through a comparison of models selected the one with the highest entropy. Future research should explore the use of LCA in estimating support for extremist ideology, but should do so with the goal of incorporating these corrections. Even with these potential limitations, however, the findings presented here appear to be of practical significance.

NOTES

1 Rather, Shapiro and Fair (2010) find it is desire for a change in the role of Shari'a in Pakistan.

REFERENCES

Agresti, Alan. 2012. *Categorical Data Analysis, Third Edition*. New York, NY: Wiley

Asparouhov, Tihomir and Bengt Muthén. 2014a. Auxiliary Variables in Mixture Modeling: 3-Step Approaches Using Mplus. Mplus Web Notes: No. 15 https://www.statmodel.com/download/webnotes/webnote15.pdf, accessed 29 October 2014

Asparouhov, Tihomir and Bengt Muthén. 2014b. Response to "Relating Latent Class Assignment to External Variables: Standard Errors for Correct Inference" Z. Bakk, D.L. Oberski, and J.K. Vermunt. Unpublished manuscript, September 24, 2014.

Bakk, Zsuzsa, Daniel L. Oberski, and Jeroen K. Vermunt. 2014. Relating latent class assignments to external variables: standard errors for correct inference. *Political Analysis*, *22*, 520-540.

Beissinger, Mark R. 2013. The Semblance of Democratic Revolution: Coalitions in Ukraine's Orange Revolution. *American Political Science Review, 107,* 574–92.

Blair, Graeme, C. Christine Fair, Neil Malhotra, and Jacob N. Shapiro. 2013. Poverty and Support for Militant Politics: Evidence from Pakistan. *American Journal of Political Science, 57,* 30-48.

Blaydes, Lisa and Drew A. Linzer. 2008. The Political Economy of Women's Support for Fundamentalist Islam. *World Politics, 60,* 576-609.

Bolck, Annabelle, Marcel Croon and Jacques Hagenaars. 2004. Estimating Latent Structure Models with Categorical Variables: One-Step versus Three-Step Estimators. *Political Analysis, 12,* 3–27.

Bullock, Will, Kosuke Imai, and Jacob N. Shapiro. 2011. Statistical Analysis of Endorsement Experiments: Measuring Support for Militant Groups in Pakistan. *Political Analysis, 19,* 363–84.

Celeux Gilles and Gilda Soromenho. 1996. An Entropy Criterion for Assessing the Number of Clusters in a Mixture Model. *Journal of Classification, 13,* 195–212.

Central Intelligence Agency. 2013. World Fact Book: Pakistan https://www.cia.gov/library/publications/the-world-factbook/geos/pk.html, accessed 13 October 2013

Clogg, Clifford C., and Leo A. Goodman. 1984. Latent Structure Analysis of a Set of Multidimensional Contingency Tables. *Journal of the American Statistical Association 79,* 762-771

Fair, C. Christine. 2004. Militant Recruitment in Pakistan: Implications for Al Qaeda and Other Organizations *Studies in Conflict and Terrorism, 27,* 489-504.

Fair, C. Christine and Bryan Shepherd. 2006. Who Supports Terrorism? Evidence from Fourteen Muslim Countries *Studies in Conflict and Terrorism, 29,* 51–74

Fair, C. Christine, Neil Malhotra, and Jacob N. Shapiro. 2012. Faith or Doctrine? Religion and Support for Political Violence in Pakistan *Public Opinion Quarterly* 76, 688-720

Finch, W. Holmes, and Kendall C. Bronk. 2011. Conducting Confirmatory Latent Class Analysis Using Mplus. *Structural Equation Modeling 18*, 132-151.

Gallup. 2013. Gallup World Poll Methodology. PDF monograph https://worldview.gallup.com/content/methodology.aspx

Garret-Mayer, Elisabeth and Jeannie-Marie Leoutsakos. 2010. Latent Class Analysis. In Chow, S-C (ed.) *Encyclopedia of Biopharmaceutical Statistics, 3rd Edition,* 700-707 Informa Healthcare, New York.

Goodman, Leo A. 2007. On the Assignment of Individuals to Latent Classes *Sociological Methodology 37,* 1-22.

Haddad, Simon and Hilal Khashan. 2002. Islam and Terrorism: Lebanese Views on September 11. *Journal of Conflict Resolution, 46,* 812-828.

Haqqani, Husain. 2005. The ideologies of South Asian Jihadi groups. *Current Trends in Islamist Ideology, 1,* 12–26.

Haqqani, Husain. 2006. "Weeding out the heretics:" Sectarianism in Pakistan. *Current Trends in Islamist Ideology 4,* 1–12.

Horgan, John. 2008. Psychology on Radicalization into Terrorism From Profiles to Pathways and Roots to Routes: Perspectives *The ANNALS of the American Academy of Political and Social Science, 618,* 80-94.

Huntington, Samuel P. 1996. *The Clash of Civilizations and the Remaking of World Order*, New York: Simon and Schuster.

James, Sigrid, Edward S. McField, and Susanne B. Montgomery. 2013. Risk Factor Profiles Among Intravenous Drug Using Young Adults: A Latent Class Analysis (LCA) Approach. *Addictive Behaviors, 38,* 1804-1811.

Kaltenthaler, Karl, William J. Miller, Stephen Ceccoli, and Ron Gelleny. 2010. The Sources of Pakistani Attitudes toward Religiously Motivated Terrorism *Studies in Conflict and Terrorism*, *33*, 815–835.

Kemp, Robert. 2008. Religious Extremism and Militancy in the Pashtun Areas of Afghanistan and Pakistan. *BC Journal of International Affairs*, *11*.

Khan, Ismail. 2010. "Female Bomber Kills Dozens in Pakistan, Official Says" *New York Times*, Dec. 26, 2010, http://www.nytimes.com/2010/12/26/world/asia/26pstan.html, accessed 29 October 2014.

Krueger, Alan B. and Jitka Maleckova. 2003. Education, Poverty and Terrorism: Is There a Causal Connection? *Journal of Economic Perspectives, 17,* 119-144.

Lessler, Judith T. and William D. Kalsbeek. 1992. *Nonsampling Errors in Surveys*. New York: Wiley.

Linzer, Drew A. and Jeffrey B. Lewis. 2011. poLCA: An R Package for Polytomous Variable Latent Class Analysis *Journal of Statistical Software, 42*.

McCutcheon, Allan L. 1985. A Latent Class Analysis of Tolerance for Nonconformity in the American Public *Public Opinion Quarterly 49,* 474-488.

McCutcheon, Allan L. 1987. *Latent Class Analysis*. Newbury Park, CA: Sage

McCutcheon, Allan L. 2002. Basic Concepts and Procedures in Single- and Multiple-Group

Latent Class Analysis. In J.A. Hagenaars and A.L. McCutcheon (eds.) *Applied Latent Class Analysis* 56-88 Cambridge, MA: Cambridge University Press.

Muthén, Linda K. and Bengt O. Muthén. 1998-2010. Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Oberski, Daniel L., Geert H. van Kollenburg, and Jeroen K. Vermunt. 2013. A Monte Carlo Evaluation of Three Methods to Detect Local Dependence in Binary Data Latent Class Models. *Advances in Data Analysis and Classification*, 7, 267-279.

Rashid, Ahmed. 1999. The Taliban**:** Exporting extremism**.** *Foreign Affairs, 78,* 22-35.

Reidel, Bruce. 2008. Pakistan and Terror: The Eye of the Storm *The ANNALS of the American Academy of Political and Social Science, 618*, 31-45

Shapiro, Jacob N., and C. Christine Fair. 2010. Understanding Support for Islamist Militancy in Pakistan. *International Security, 34,* 79–118.

Tay, Louis, Daniel A. Newman, and Jeroen K. Vermunt, 2011. Using Mixed-Measurement Item Response Theory With Covariates (MM-IRT-C) To Ascertain Observed and Unobserved Measurement Equivalence. *Organizational Research Methods, 14,* 147-176.

Tessler, Mark. and Michael D.H. Robbins. 2007. What Leads Some Ordinary Arab Men and Women to Approve of Terrorist Acts Against the United States? *Journal of Conflict Resolution, 51*, 305-328.

Vermunt, Jeroen K. 2010. Latent Class Modeling with Covariates: Two Improved Three-Step Approaches. *Political Analysis, 18,* 450–469.

Tarek Al Baghal is Senior Researcher, Institute for Social and Economic Research, University of Essex. His research focuses on psychology and measurement in surveys as well as applications of measurement models to survey data. His most recent publication is Tarek Al Baghal (2014) "Is Vague Valid?: The Comparative Predictive Validity of Vague Quantifiers and Numeric Response Options." *Survey Research Methods*, 8: 169-179.

Email: talbag@essex.ac.uk

APPENDIX A: QUESTIONS USED

"Many people have discussed different approaches to improving American policies. How much hope do you place in the role of each of the following in improving American policies toward your country: Please indicate your response using a 5 point scale where a 1 signifies no hope at all and a 5 signifies a great deal of hope. Terrorist attacks on American citizens"

"There are many acts some people may do in life. I will read out to you a number of these acts. I would like you to indicate to which extent it can be morally justified. Events of Sept 11th in USA, that is, the attack on the World Trade Center and the Pentagon." 1= Not at all justified, 5= Totally justified.

"There are many acts some people may do in life. I will read out to you a number of these acts. I would like you to indicate to which extent it can be morally justified. Other attacks in which civilians are the target." 1= Not at all justified, 5= Totally justified.

"Finally, we'd like to know how you feel about people of different religious or spiritual groups. Please say whether your overall view of each group is very positive, somewhat positive, neutral, somewhat negative, or very negative. How about Shiite Muslims"

"Sharia is an Arabic word which means Islamic religious principles. In general, which of these statements comes closest to your own point of view?

-Sharia must be the only source of legislation
-Sharia must be a source of legislation, but not the only source
-Sharia should not be a source of legislation"

"Do you believe that Western world respects Muslim world?"

"Is the leadership of Pakistan taking it in the right or wrong direction?"

APPENDIX B: BIVARIATE RESIDUALS FROM FOUR-CLASS LCA

Hope from Attacks on US	Sept 11 Justified	Residual Z-score
Category 1	Category 1	0.046
Category 1	Category 2	0.17
Category 1	Category 3	-0.333
Category 2	Category 1	0.016
Category 2	Category 2	-0.334
Category 2	Category 3	0.702
Category 3	Category 1	-0.09
Category 3	Category 2	0.328
Category 3	Category 3	-0.194

Hope from Attacks on US	Attacks on Civilians Justified	Residual Z-score
Category 1	Category 1	-0.003
Category 1	Category 2	0.069
Category 1	Category 3	-0.107
Category 2	Category 1	0.018
Category 2	Category 2	-0.218
Category 2	Category 3	0.36
Category 3	Category 1	-0.017
Category 3	Category 2	0.312
Category 3	Category 3	-0.309

Hope from Attacks on US	Attitudes Towards Shiite Muslims	Residual Z-score
Category 1	Category 1	-0.183
Category 1	Category 2	0.266
Category 1	Category 3	-0.037
Category 2	Category 1	-0.061
Category 2	Category 2	-0.242
Category 2	Category 3	0.346
Category 3	Category 1	0.381
Category 3	Category 2	-0.153
Category 3	Category 3	-0.374

Sept 11 Justified	Attacks on Civilians Justified	Residual Z-score
Category 1	Category 1	-0.039
Category 1	Category 2	-0.068
Category 1	Category 3	0.28
Category 2	Category 1	-0.072
Category 2	Category 2	0.509
Category 2	Category 3	-0.708
Category 3	Category 1	0.221
Category 3	Category 2	-0.837
Category 3	Category 3	0.544

Sept 11 Justified	Attitudes Towards Shiite Muslims	Residual Z-score
Category 1	Category 1	0.088
Category 1	Category 2	0.068
Category 1	Category 3	-0.2
Category 2	Category 1	-0.691
Category 2	Category 2	0.584
Category 2	Category 3	0.461
Category 3	Category 1	0.955
Category 3	Category 2	-0.999
Category 3	Category 3	-0.081

Attacks on Civilians Justified	Attitudes Towards Shiite Muslims	Residual Z-score
Category 1	Category 1	0.022
Category 1	Category 2	0.163
Category 1	Category 3	-0.213
Category 2	Category 1	-0.409
Category 2	Category 2	-0.07
Category 2	Category 3	0.761
Category 3	Category 1	0.714
Category 3	Category 2	-0.356
Category 3	Category 3	-0.472