the National Science Foundation, 1945–1992. New Brunswick, NJ: Transaction Publishers.

Morton, Rebecca B., and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.

Orsini, Michael. 2013. "'May I See your Color-Coded Badge?' Reflections on Research with 'Vulnerable' Communities." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, 2nd ed. Dvora Yanow and Peregrine Schwartz-Shea, eds. (Armonk, NY: M.E. Sharpe), in press.

Polanyi, Michael. 1966. *The Tacit Dimension*. New York: Doubleday.

Schön, Donald A. 1987. *Educating the Reflective Practitioner.* San Francisco: Jossey-Bass.

Schwartz-Shea, Peregrine. 1991. "Understanding Subgroup Optimization: Experimental Evidence on Individual Choice and Group Processes." *Journal of Public Administration Research and Theory* 1:1, 49–73.

Schwartz-Shea, Peregrine. 2002. "Theorizing Gender for Experimental, Game Theory: Experiments with 'Sex Status' and 'Merit Status' in an Asymmetric Game." *Sex Roles* 47:7/8, 301–319.

Schwartz-Shea, Peregrine. 2006. "Judging Quality: Evaluative Criteria and Epistemic Communities." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. Dvora Yanow and Peregrine Schwartz-Shea, eds. (Armonk, NY: M.E. Sharpe), 89–113.

Schwartz-Shea, Peregrine and Dvora Yanow. 2002. "'Reading' 'Methods' 'Texts': How Research Methods Texts Construct Political Science." *Political Research Quarterly* 55, 457–486.

Vaughan, Diane. 1996. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago: University of Chicago Press.

Voeten, Erik. 2012. "House Votes to Prohibit Political Science Funding." *The Monkey Cage* (10 May), http://themonkeycage.org/blog/2012/05/10/house-votes-to-prohibit-political-science-funding/.

Yanow, Dvora and Peregrine Schwartz-Shea, eds. 2013 [2006]. *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, 2nd ed. Armonk, NY: ME Sharpe.

Symposium: Reports from the Multi-Method Research Frontier

A Case for Case Studies: A Multi-Method Strategy for Ecological Inference

Anne Meng
University of California, Berkeley
ameng@berkeley.edu

Brian Palmer-Rubin
University of California, Berkeley
brianpr@berkeley.edu

Ecological inference (EI) is the process of using aggregate data to infer discrete individual-level relationships of interest when individual-level data are not available (Cho and Manski 2008). EI is one of the longest standing challenges to quantitative social science research, yet scholars continue to debate the best statistical methods to deal with this unit-of-analysis problem (Freedman 1999, King 1997). In political science research, ecological inference commonly comes into play when a scholar is interested in inferring the voting behavior of some subgroup but electoral data that distinguish among subgroups are unavailable. This problem is especially salient in studies of voting behavior analyzing developing or weakly democratic countries where reliable individual- or precinct-level polling is rare.

We propose a strategy to improve the validity of ecological inference through a combination of quantitative analysis of aggregate-level units and within-unit case studies. First we provide an overview of the ecological-inference problem, and the most prevalent statistical techniques that have been proposed to address it. Next, we outline a strategy of using qualitative case studies to gather evidence that allows the researcher to evaluate the assumptions that underlie ecological inference. Finally, we describe an application of these strategies to our own study of ethnic-minority voting in Mexico.[1]

The multi-method design that we propose builds on a recent wave of methodological scholarship that promotes case studies as a commonsense alternative or complement to complex statistical models for building and testing causal theories (Brady and Collier 2004, Freedman 1999, 2009, George and Bennett 2005, Gerring 2007, Mahoney 2010). We argue that qualitative case studies allow the researcher to collect fine-grained data on the micro-level mechanisms that underlie statistical relationships observed at the ecological level, and thus constitute a useful complement to quantitative EI approaches.

Aggregate Data and Disaggregated-Level Inference

A researcher faces the ecological-inference problem whenever she is interested in making inferences about the behavior of a particular sub-population, or "population of interest" (POI), such as an ethnic group, social class, or voters registered to a particular party, yet only has data that is aggregated at a higher level. The resulting "aggregation bias" is "the effect of the information loss that occurs when individual-level data are aggregated…The problem is that in some aggregate data collections, the type of information loss may be selective, so that inferences that do not take this into account will be biased" (King 1997: 17). Ecological fallacies occur when researchers assume that relationships observed at the aggregate level are the consequence of corresponding disaggregate-level relationships (e.g., *counties* with more Hispanic voters vote disproportionately for Democratic candidates, therefore Hispanic *voters* vote disproportionately for Democratic candidates).

Studies of minority voting behavior commonly deal with ecological-inference problems. In recent decades, institutional mechanisms such as majority-minority districts and legislative quotas have been created to promote political participation by ethnic minorities and descriptive or substantive representation of minority interests in policymaking. A number of studies use data that are aggregated at the district level (or higher) to attempt to gauge whether these institutions affect minority

Table 1: An Ecological Inference Hypothetical

	Democratic Candidate	Republican Candidate	Precinct Population (%)
Non-Hispanic Voters	? (*p*)	?	0.70
Hispanic Voters	? (*q*)	?	0.30
Total Votes (%)	0.80	0.20	

voting (Chandra 2004, Goodnow and Moser 2012). Due to the secret ballot, it is impossible to directly observe how given ethnic groups vote, however, and even when exit polls are available, their results are often not reliable in racially charged elections (Grofman, Handley, and Niemi 1992). Challenges abound in studies of developing or weakly democratic countries (e.g., Blaydes 2011, Chandra 2004, Goodnow and Moser 2012) or historical societies (e.g., Childers 1983, Hamilton 1982), as exit polls are rare and electoral data are usually only available at relatively coarse levels of aggregation.

Quantitative Strategies: The Status Quo

A number of quantitative techniques have been developed to generate estimates about disaggregate-level relationships when only aggregate data are available. In this section, we outline the Method-of-Bounds approach and regression-based approaches and discuss the limitations of these quantitative strategies.

Formalizing the Problem

Studies of voting behavior that face the EI problem often resemble the following hypothetical:

A congressional precinct in the United States is composed of Hispanic and non-Hispanic voters. There are two candidates running for office: a Democrat and a Republican. We have electoral data on the vote share each candidate received, as well as the percent of the precinct that is Hispanic. Due to secret ballots, we do not know which candidate each individual voter voted for. What percent of the votes for the Democratic candidate came from the Hispanic voters?

Table 1 summarizes the EI problem. Although we have data on the margins of the table, we would like to fill in the question marks—most relevantly, the question mark in the lower-left cell.

The Method-of-Bounds approach is the simplest method for estimating the missing values, providing scholars with a first-cut tool to evaluate the plausibility of their disaggregate-level inferences (Duncan and Davis 1953). It relies on the intuition that vote shares cannot be negative; therefore each missing cell in Table 1 must be bounded by a [0,1] interval. Let p denote the fraction of the non-Hispanic population that voted for the Democratic candidate and let q denote the fraction of the Hispanic population that voted for the Democratic. Since we know the fraction of the total precinct population that voted for the Democratic candidate, as well as the fractions of the population that are Hispanic and non-Hispanic, we know that p and q must satisfy the following equation:

$$0.70(p) + 0.30(q) = 0.80$$

Furthermore, since we know that p and q must be bounded by [0,1], we can derive upper and lower bounds for these terms. Let p=1. Then $0.70(1) + 0.30(q) = 0.8$. We then calculate q=0.33. Now let p=0. Then $0.70(0) + 0.30(q) = 0.80$. We then calculate q=2.6. However, since q is also bounded by [0,1], we can conclude that the percent of the precinct's Hispanic population that voted for the Democratic candidate is bounded by [0.33,1]. Bounds that are wider represent more acute EI problems, and they tend to occur as the POI represents a smaller proportion of the aggregate unit or as the outcome of interest occurs in close to 50 percent of the aggregate unit.

Regression-Based Approaches

Due to the limited inferential power of the Method-of-Bounds approach, many researchers have turned to regression-based strategies (e.g., Blaydes 2011, Chandra 2004, Goodnow and Moser 2012). Commonly used regression models for ecological inference include the Goodman regression (1953), neighborhood model (Freedman, et al. 1991), and King's model (1997). All of these techniques center on the same intuition, but differ in the complexity of the statistical model and underlying assumptions required to produce estimates. The Goodman model, the most basic ecological-regression approach, is set up as follows:

Returning to our voting example, let x denote the percent of the precinct population that is Hispanic, and y denote the percent of the total precinct vote share that went to the Democratic candidate. The subscript i in the equation indexes x, y, and ε by precincts

$$y_i = a + bx_i + \varepsilon_i$$

We can use Ordinary Least Squares (OLS) to estimate the parameters a and b. We interpret a as the estimate of the fraction of the non-Hispanic voters who voted for the Democratic candidate. Then $a + b$ represents the estimate of the fraction of the Hispanic voters who voted for the Democratic candidate.[2]

Notwithstanding its popularity, ecological regression is far from a silver bullet. The three most common regression-based techniques represent imperfect choices in dealing with the tradeoff between the plausibility of assumptions, the accuracy of estimates, and the simplicity of the model. Though the Goodman regression is easy to explain and understand, this technique is vulnerable to biases inherent to OLS regression and often produces nonsensical estimates of turnout or vote share that are not bounded by [0,1]. It also relies on an unrealistic "constancy assumption" that requires voting behavior across municipalities to be identical for minorities and non-minorities. The Freedman linear neighborhood model relaxes the constancy assumption by assuming that differences in voting behavior are independent of ethnicity, but is fundamentally based on the same OLS regression. Finally, while King's strategy offers certain advantages, such as the use of truncated distributions to compel respect for bounds and the incorporation of covariates to control for confounders, scholars debate whether this more complicated model produces more accurate estimates than its predecessors (Cho 1998; Freedman et al. 1991; King 1999).

A New Proposal: Using Case Studies to Evaluate Ecological Inference

Given the limitations in quantitative strategies, we propose a qualitative method for evaluating the plausibility of EI in causal research: carrying out case studies of units from a previously conducted large-N analysis wherein the scholar evaluates whether the observed ecological-level statistical relationship is driven by the hypothesized effect on the POI. These case studies have the goal of neither testing a hypothesis nor generating new hypotheses: Presumably, the researcher entered into the analysis with a plausible hypothesis to be tested and has moved onto these case studies because she has already found evidence on the aggregate level that supports this hypothesis.

Compared with statistical strategies, the case-study approach has the benefit of providing evidence about the specific mechanisms that underlie a causal relationship through the use of process tracing, which allows the scholar to "make strong inferences in just one or a few cases, based on one or a few pieces of the right kind of evidence" (Bennett 2008: 718). A drawback to the case-study approach, as with all small-n strategies, is the challenge to generalizability; it is impossible to prove that the POI-level relationships that one uncovers in a few studied units are the same that underlie the broader ecological-level relationship.

The goal of these case studies is to gather evidence that the ecological relationship that is observed in the original large-N analysis is driven by the hypothesized effect of the independent variable on the POI, rather than some other effect. For instance, in a study of minority voting, the scholar may seek to gauge whether a change in the *Democratic Party's campaign strategy* (independent variable) increased the *fraction of the Hispanic vote going to Democratic congressional candidates* (dependent variable). To this end, the scholar is looking for evidence to validate four criteria regarding the effect of the campaign strategy variable: causality, exclusivity, consistency, and non-interaction (Table 2).

These case studies do not necessarily require a large investment of time conducting interviews or archival research. These criteria can be evaluated in as few as two case studies—both cases that confirm the hypothesized relationship, one of which has a high prevalence of the POI and the other that has a low prevalence of the POI. (Of course, the greater the number of cases analyzed, the better case the scholar can make that the EI is valid across the dataset.) The scholar will make many "causal process observations" (Brady and Collier 2004: 277–78) within each case by gathering evidence on at least two within-unit subgroups (POI and non-POI). It is appropriate to forgo studies of hypothesis-refuting cases because the goal is not to measure the causal effect of the independent variable, but rather to observe how it operates on different subgroups within an aggregate unit.

Three types of evidence can be used to evaluate these criteria:

> *Direct evidence of mechanism*: First-hand information, gained through interviews or observations, demonstrating how the independent variable affects the POI. (Example: The researcher attends a Democratic Party rally and observes that issues important to the Hispanic community are touted more than at rallies in the past.)

Table 2: Ecological Inference Criteria to be Verified through Case Studies

Criterion	Description	Importance for Causal Inference	Seeking Evidence to Show that…
Causality	The independent variable affects the POI as hypothesized	Necessary	The new campaign strategy caused Hispanic voters to vote more for Democratic Party candidates than in previous elections
Exclusivity	The independent variable does not affect the non-POI	Preferable	The new campaign strategy has no effect on the vote choice of non-Hispanic voters
Consistency	The causal effect is not influenced by the proportion of POI in the aggregated unit	Preferable	The effect of the campaign strategy on Hispanic voting does not vary based on the proportion of Hispanic voters in a unit
Non-Interaction	The independent variable does not interact with some other unobserved variable	Preferable	The effect of the campaign strategy does not vary among subsets of Hispanic voters (e.g., across social classes)

Bystander account of effect: Media or other accounts that provide clues as to the voting behavior of the POI and non-POI. (Example: An interviewed election observer reports seeing more Hispanic attendees at Democratic Party campaign rallies after the strategy change.)

Disaggregated within-unit data: Medium- to large-N data within aggregated units demonstrating that the POI-level outcome corresponds to the initially observed aggregate-level relationship. (Example: A survey that breaks down voting by race within a district shows that Hispanic voters voted in greater proportion for the Democratic Party candidate since the adoption of the new strategy.)

Applying the Case-Study Method: Minority-Concentrated Districts in Mexico

In this final section we describe the application of our case-study method to our study of minority voting in Mexico. We investigate whether minority-concentrated districts (MCDs)—an institution adopted in Mexico in 2005—influence the degree to which minority voters (the indigenous population) choose to vote for opposition parties.[3] Our study approximates a natural experiment by using a redistricting reform that took place between two congressional elections (2003 and 2009) to estimate the causal effect of living in an MCD on voter behavior at the municipal level.[4] We collected electoral data from the *Instituto Federal Electoral* (IFE), Mexico's national electoral institute, on congressional election outcomes. These data were recorded on the municipal level, the finest level of disaggregation available (districts are made up of several municipalities). Census data reporting the prevalence of indigenous and non-indigenous populations are also only available down to the municipal level. Therefore the aggregate unit for our large-N analysis is the municipality, and the POI is the block of indigenous voters in a given municipality.

Across the municipalities in our dataset—all of which are at least 50 percent indigenous—we found a significant negative effect of MCDs on the vote share for the dominant party (the PRI). These findings remain consistent across a variety of model specifications. Based on our quantitative evidence, and intuitions gleaned from newspaper accounts and previous research, we suspect that the creation of MCDs allowed opposition parties to penetrate populations that previously voted overwhelmingly for the dominant party. In this scenario, the purported mechanism is a change in opposition parties' electoral strategies following the adoption of MCDs: enterprising party leaders capitalized on these districts by nominating minority candidates and making targeted patronage appeals to indigenous communities.

We were satisfied to find a significant negative effect on the PRI's vote share on the aggregate level; however, we wanted to confirm that these findings reflect a change in indigenous voting behavior. *Are indigenous voters living in MCDs less likely to vote for the dominant party than indigenous voters not living in MCDs*? Answering this question, and providing evidence of the mechanism that underlies the effect of MCDs on indigenous voting, would make our causal inferences much stronger. Specifically, our quantitative analysis left four questions unanswered, corresponding to the four EI criteria described in the previous section:

1. *Causality*: Is the low vote share for the dominant parties in municipalities assigned to MCDs explained by the effect of MCDs on indigenous voting behavior?

2. *Exclusivity*: Did the assignment of municipalities to MCDs affect the voting behavior of non-indigenous voters in those municipalities? If so, did these populations respond in a way that strengthens our inference about indigenous voting behavior (voting less for opposition parties) or weakens this inference (voting more for opposition parties)?

3. *Consistency*: Did assignment to MCDs affect indigenous voting differently in municipalities with high indigenous populations (close to 100 percent) than it did in municipalities with relatively low indigenous populations (close to 50 percent)?

4. *Non-Interaction*: Did reassignment to MCDs have different effects among various subgroups within municipal indigenous populations (e.g., indigenous Catholics vs. indigenous Protestants or poor vs. non-poor indigenous)?

In order to respond to these questions, we undertook case studies that focused primarily on two municipalities in Chiapas, a highly indigenous state in southern Mexico. The two municipalities, Ocotepec and Simojovel, both received treatment (were assigned to MCDs) and both experienced significant decreases in the vote shares for the PRI from 2003 to 2009 (in Ocotepec, from 46 percent in 2003 to 29 percent in 2009 and in Simojovel from 52 percent in 2003 to 15 percent in 2009). They also belong to the higher and lower ends of the spectrum in the prevalence of indigenous populations in our dataset: Ocotepec's population is 97 percent indigenous and Simojovel's population is 71 percent indigenous. We conducted two methods of qualitative data collection: first, interviews with party leaders and indigenous authorities who reported on the mechanism underlying the relationship between MCDs and indigenous voting; and second, second-hand observations of the electoral behavior of the POI and non-POI, acquired through searches through newspaper archives and interviews with informed observers (academics and NGO workers).

The first three questions are oriented toward understanding the effect of the treatment on indigenous and non-indigenous voting. Is there evidence that reassignment to MCDs increased indigenous voting for opposition parties? Did this treatment have any effect on non-indigenous vote choice? Do these effects vary based on the proportion of indigenous voters in a municipality?

To address these questions, we interviewed indigenous authorities and leaders of the once-dominant party (PRI) and the main opposition party (PRD) in Chiapas, and gathered newspaper accounts of campaign activities by these parties. Interviews with party leaders uncovered affirmative evidence of the mechanism: PRD leaders reported that they targeted their campaigns to indigenous communities following the redistricting by nominating indigenous candidates and promoting social programs for indigenous communities. PRI leaders reported very little change in their campaign strategies following the redistricting. These reports were bolstered by newspaper reports of the intensification of indigenous-targeted appeals by the PRD in 2009. Furthermore, we found no evidence that these mechanisms operated differently in Ocotepec (our high-POI case) and Simojovel (our low-POI case).

The final question has to do with different treatment effects across subgroups of indigenous voters within a municipality. This issue is important because it is possible that the hypothesized effect on indigenous voters only occurs among a subset of the indigenous, which would suggest that we modify our initial hypothesis. For instance, scholars of indigenous activism have observed that Catholic and Protestant churches with indigenous congregations promote different forms of mobilization and respond to partisan appeals in different ways (Palmer-Rubin 2011, Trejo 2009). Reassignment to MCDs could also be more influential among poorer indigenous populations, which are likely to be more prone to clientelistic appeals than relatively better-off populations. To address these potential modifications to our findings, we interviewed subjects who were able to comment on the effect of MCDs on the electoral behavior of indigenous voters of different religions and different economic strata. Interviews with party leaders as well as both Catholic and Protestant indigenous authorities suggested that all indigenous populations were targeted by the PRD after 2005, regardless of religion. Both predominately Catholic and predominately Protestant indigenous organizations formed alliances with the PRD (although the switch to the PRD appeared to be more widespread by Catholic groups). Through interviews of leaders representing both relatively well-off (more urban) indigenous populations and relatively poor (more rural) indigenous populations, we found no compelling evidence that poverty mediated the effect of MCDs on indigenous vote choice. We also found newspaper reports of indigenous-targeted patronage in both rural and urban areas.

In sum, the case studies provided evidence to bolster our quantitative findings. The aggregate-level finding—that municipalities reassigned to MCDs demonstrated lower vote shares for the PRI than similar municipalities that were not reassigned—is grounded in the POI-level relationship that our hypothesis predicts, at least in the municipalities where we conducted case studies. We observed the mechanism that underlies the effect of MCDs on indigenous voting: namely, shifts in opposition-party strategies. We also found no compelling evidence that MCDs significantly affected the vote choice of non-indigenous voters, nor that the effect of MCDs on indigenous voting was mediated by the proportion of the municipality that is indigenous or by some other variable. Of course, we cannot be sure that these findings are generalizable to all other municipalities in our dataset. However, our case studies lend a great deal of plausibility to our causal argument by demonstrating that at least in a couple of cases, the hypothesized POI-level effect took place.

Concluding Thoughts

This paper addresses the challenges a researcher faces when working with aggregate data if she would like to make arguments about the behavior of some population at a lower level of aggregation. In response to the limitations of statistical techniques that have been proposed to alleviate the ecological-inference problem, we develop a case-study approach. Our proposed strategy does not differ markedly from other multi-method approaches that employ case studies to identify mechanisms that underlie causal relationships identified through large-N analysis. However, our approach is tailored to assist the scholar in observing the causal effect of an explanatory variable on the population-of-interest and to detect aggregation bias that may threaten the large-N findings. Compared with commonly used statistical strategies for addressing ecological inference problems, the primary advantage of case studies is that they provide the scholar with original evidence of mechanisms that underpin the hypothesized causal relationships.

While we believe that this strategy offers certain advantages over statistical approaches, we agree with Freedman's (1999: 4030) prediction that "the problems of confounding and aggregation bias…are unlikely to be resolved in the proximate future." Nonetheless, research using aggregate-level data will continue to be important and common, given the predominance of ecological-level data. Thus, in the interest of reaching the most defensible causal claims, scholars are advised to use all the evidence, both quantitative and qualitative, that it is practical to collect.

Notes

[1] The research project for which we developed this multi-method approach capitalizes on a redistricting reform undertaken between two congressional elections in Mexico to approximate a natural experiment, allowing us to measure the effect of indigenous-concentrated districts on voting outcomes. Due to space constraints, we do not go into detail about our large-N identification strategy here. The examples in this article, instead, resemble EI problems that would occur in the context of any large-N research that uses aggregate-level data.

[2] To see this, recall that a is the height of the regression line at x=0, and $a + b$ is the height of the regression line at x=1.

[3] We define the opposition parties in 2009 as the PAN and the PRD. Although the PAN won the presidency in Mexico in 2000, the PRI is the historically dominant party that ruled all of these municipalities before the insertion of the PAN and PRD in the 1990s.

[4] Treated units are municipalities that were redistricted to MCDs, and control cases are similar municipalities that were not redistricted to MCDs. To ensure balance between the two groups, we matched on redistricting criteria, socioeconomic variables, and pre-treatment electoral outcomes.

References

Bennett, Andrew. 2008. "Process Tracing: A Bayesian Perspective." In *The Oxford Handbook of Political Methodology*. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds. (New York: Oxford University Press), 702–721.

Blaydes, Lisa. 2011. *Elections and Distributive Politics in Mubarak's*

Egypt. New York: Cambridge University Press.
Brady, Henry and David Collier. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman and Littlefield.
Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed*. New York: Cambridge University Press.
Childers, Thomas. 1983. *The Nazi Voter: The Social Foundations of Fascism in Germany, 1919-1933*. Chapel Hill: University of North Carolina Press.
Cho, Wendy K. Tam. 1998. "If the Assumption Fits…: A Comment on the King Ecological Inference Solution." *Political Analysis* 7:1, 143–163.
Cho, Wendy and Charles Manski. 2008. "Cross-Level/Ecological Inference." In *The Oxford Handbook of Political Methodology*. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds. (New York: Oxford University Press), 547–569.
Duncan, Otis and Beverly Davis. 1953. "An Alternative to Ecological Correlation." *American Sociological Review* 18:6, 665–666.
Freedman, David. 1999. "Ecological Inference and the Ecological Fallacy." In *International Encyclopedia of the Social and Behavioral Sciences*. Neil J. Smelser and Paul B. Baltes, eds. (Oxford: Elsevier), 4027–4030.
Freedman, David. 2009. "On 'Solutions' to the Ecological Inference Problem." In *Statistical Models and Causal Inference*. David Collier, Jasjeet Sekhon, and Philip Stark, eds. (New York: Cambridge University Press), 83–96.
Freedman, David, Stephen P. Klein, Jerome Sacks, Charles A. Smyth, and Charles G. Everett. 1991. "Ecological Regression and Voting Rights." *Evaluation Review* 15:87, 673–711.
George, Alexander and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.
Gerring, John. 2007. *Case Study Research: Principles and Practices*. New York: Cambridge University Press.
Goodman, Leo 1953. "Ecological Regression and the Behavior of Individuals." *American Sociological Review* 18:6, 663–664.
Goodnow, Regina and Robert Moser. 2012. "Layers of Ethnicity: The Effects of Ethnic Federalism, Majority-Minority Districts, and Minority Concentration on the Electoral Success of Ethnic Minorities in Russia." *Comparative Political Studies* 45:2, 167–193.
Grofman, Bernard, Lisa Handley, and Richard G. Niemi. 1992. *Minority Representation and the Quest for Voting Equality*. New York: Cambridge University Press.
Hamilton, Richard F. 1982. *Who Voted for Hitler?* Princeton: Princeton University Press.
King, Gary. 1997. *A Solution to the Ecological Inference Problem*. Princeton: Princeton University Press.
King, Gary. 1999. "The Future of Ecological Inference Research: A Reply to Freedman, et al." *Journal of the American Statistical Association* 94:445, 352–355.
Mahoney, James. 2010. "After KKV: The New Methodology of Qualitative Research." *World Politics* 62:1, 120–147.
Meng, Anne and Brian Palmer-Rubin. 2012. "Ethnic Voting and Electoral Competition: The Effects of Minority-Concentrated Districts in Mexico." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago.
Palmer-Rubin, Brian. 2011. "From Frame to Campaign: Indigeneity in the Political Arena." Paper presented at the Annual Meeting of the American Political Science Association, Seattle.
Trejo, Guillermo. 2009. "Religious Competition and Ethnic Mobilization in Latin America: Why the Catholic Church Promotes Indigenous Movements in Mexico." *American Political Science Review* 103:3, 323–342.

More than the Sum of the Parts: Nested Analysis in Action

Craig M. Kauffman
University of Oregon
ckauffma@uoregon.edu

After reading Lieberman's (2005) article on nested analysis, I was eager to test the purported benefits by adopting this multi-method approach in my dissertation research (Kauffman 2012). After using it, I am even more convinced of Lieberman's assertion that both quantitative and qualitative methodologies "can inform each other to the extent that the analytic payoff is greater than the sum of the parts" (2005: 436). My purpose here is to provide an example of nested analysis in action to illustrate how quantitative and qualitative methods may be used in supplementary and complementary ways at various stages of a research design, from model specification, to case selection, to data analysis.

The puzzle that inspired my research was why some local governments in Ecuador pursued and successfully implemented a set of reforms known as Integrated Watershed Management (e.g., Gregersen, Ffolliott, and Brooks 2007), while others did not. Specifically, I was interested in the creation of two institutions designed to better manage local watershed resources: (1) a participatory decision-making structure in which multiple stakeholders jointly identify needs and develop watershed management plans and activities, and (2) a stable financing mechanism using local revenue sources to fund watershed management activities. Through previous fieldwork, I had seen how these reforms could change production practices, alter socio-political relations, and create new governance arrangements with characteristics commonly equated with good governance. I found the reforms curious in part because, while many attempts failed, they sometimes succeeded in unlikely places—cantons with poorly performing governments, high levels of poverty and inequality, corruption, clientelism, social and ethnic conflict, and little history of social organizing. I initially saw this as an opportunity to test competing hypotheses regarding improvements in local government performance.

Having identified my general research question, I used both quantitative and qualitative methods to specify hypotheses and a model, assess the validity of indicators, evaluate rival explanations, check for omitted variable bias, and search for causal mechanisms behind correlations. Following the nested analysis approach, my research design involved a two-step process: (1) a statistical analysis of all 221 Ecuadorian cantons, and (2) a qualitative comparison of six case studies.

Model Specification

For the statistical analysis I used a logistic regression model to test competing explanations of local governance reform based on an original dataset I compiled. The dependent variable was dichotomous, indicating whether or not natural resource management reforms were attempted in each canton. While my interest was in watershed management, I looked at the broader category of natural resource management reforms in order to have a population of cases of sufficient size and variation to make statistically meaningful comparisons. Between 1997 and 2008 roughly half of Ecuador's municipalities (108 of 221) pursued reforms to better manage natural resources. An assumption of the study is that the relationships applying to natural resource management reforms generally also apply to the sub-category of watershed management reforms. My list of reforming municipalities came from Ecuadorian scholars and practitioners who catalogued innovative local government reforms in Ecuador since the 1990s (e.g., Ramón and Torres 2004; Asociación de Municipalidades Ecuatorianas 2004, 2008; Garzón 2009), as well as information from personal interviews. The dataset also included indicators of explanations commonly found in the literatures on local government performance and natural resource management. These include measures of political competition, political organization, citizen participation, social trust, technical assistance, central-local government relationships, ecosystem type, level of economic development, and demographic variables, among others. As expected, some explanations were ruled out while several variables had a significant effect (e.g., measures of participation, trust, and ecosystem type).

While statistical analysis could assess the strength of partial explanations, it could not illuminate the causal mechanisms or describe how different variables interacted to produce success. For this I turned to within-case analysis. Because my interest was initially in model testing, my first step in the qualitative portion was to conduct process tracing in a typical success case to check the validity of my indicators and uncover the causal mechanisms behind the correlations in the statistical study. To identify a typical success, I calculated the predicted probabilities of success for all 221 cantons and compared this with their actual scores. I then selected Tungurahua, the case with the highest predicted probability that had implemented the Integrated Watershed Management reforms described above.

By necessity, the indicator for my dependent variable in the statistical study had been relatively crude—a dummy variable that indicated reform or no reform. But government reform is rarely black or white. There are many types and degrees of success that must be empirically verified. One advantage of small-N analysis is the ability to develop more nuanced and precise indicators of success. To measure the extent to which Integrated Watershed Management reforms occurred, I developed an index of 15 indicators. These indicators measured the degree to which there was consensus among various stakeholders on the problem and proposed solutions, the degree to which institutions were created and action plans implemented, the degree to which stakeholder groups participated in the process, and the degree to which the reforms were institutionalized in a way that provided lock-in effects. Cases were scored on a scale ranging from 0 (low success) to 15 (high success). For each of the 15 indicators, a case was awarded 1 point if the condition was present and 0 points if it was absent. A half point was awarded when a condition was only weakly present or was present at one time but not sustained.

Four months of field research in Tungurahua produced surprising results. On one hand, it confirmed that Tungurahua was indeed an example of successful Integrated Watershed Management reform; it scored 13.5 out of 15 on the index. However, process tracing revealed that, instead of the structural explanations espoused in the literature and represented in my statistical model, the main story was the role played by transnational networks in setting the agenda for reform and mobilizing coalitions of local stakeholders with the motivation and capacity to pursue and implement reforms. In short, my qualitative analysis suggested that my statistical model was misspecified and potentially contained spurious correlations.

Navigating Between Model Testing and Model Building

At this point my endeavor turned from model testing to model building. The Tungurahua case presented a new hypothesis—that variation in the implementation of Ecuador's local natural resource management reforms (including watershed management reforms) was explained by the pattern of network ties and strategies employed by transnational coalitions advocating these reforms. I ultimately wanted to test this hypothesis through comparative case studies. But first I assessed the new hypothesis using a statistical model. I wanted more confidence that Tungurahua was not an anomaly and that the investment of time and resources needed to conduct case studies would be worthwhile.

I revised my quantitative study by adding a new independent variable, the presence of transnational environmental networks. The theory was that the presence of transnational environmental networks made reform attempts more likely since international actors carry new ideas and practices regarding natural resource management to cantons through these network ties. These networks also provide important resources to local advocates who embrace the reforms and seek to implement them. I measured the presence of transnational environmental networks using the amount of external environmental aid received by each canton from international actors. While environmental aid was not a perfect proxy for transnational networks, it was the best data available that would allow a large-N analysis. I assumed that if international organizations provided aid to particular cantons, particularly in a sustained manner, representatives of those organizations were interacting regularly with local actors to implement environmental programs. Theoretically, this regular interaction created ties between local actors and a transnational environmental network through which ideas and resources flowed. I also assumed that international environmental actors provided more aid to the cantons with which they had stronger ties, and a lack of aid reflected a lack of network ties.

The specific indicator for my transnational network variable was the average amount of environmental aid (in millions of dollars) donated by international actors in each canton between 2007 and 2009, as calculated by the Ecuadorian Agency for International Cooperation (AGECI). This indicator included the money spent by multilateral organizations (e.g., the World

Bank and the European Union), bilateral development organizations (e.g., U.S. Agency for International Development and German Technical Cooperation), international NGOs (e.g. The Nature Conservancy and Conservation International), and private businesses on local environmental programs. A key assumption was that levels of environmental aid during 2007–2009 (the only years for which data were available) were consistent with levels of similar aid during the previous seven years, during which most reform processes began. My six case studies later compensated for this weakness by providing evidence that the financing of Integrated Watershed Management reforms in the late 2000s typically followed a decade of similar funding on natural resource management programs.

It is beyond the scope of this article to describe the all details of my quantitative analysis. However, I want to share a few results that particularly influenced the qualitative portion of my project. The results supported the hypothesis regarding the importance of transnational networks. Controlling for all other explanations, the presence of transnational environmental networks (as measured through environmental aid) had both a large effect and was highly significant ($p = .002$). Various robustness tests confirmed these results. Many of the variables representing alternative explanations lost their significance, suggesting there may have been spurious correlations. In short, my new model allowed me to eliminate many alternative explanations and justified a narrower focus on transnational networks as the main explanatory variable of interest in my case comparisons.

The best fitting model—the one that combined parsimony with predictive power—contained three explanatory variables: the presence of transnational environmental networks (indicated by environmental aid), social trust (based on survey results conducted by the LAPOP project; Seligson et al. 2006), and a dummy variable, Sierra, indicting whether a canton is in Ecuador's mountainous Sierra region. I included this variable because the region is argued to have several conditions conducive to reform. Ecuador's Sierra region is commonly believed to have a more innovative political culture conducive to change, in part due to the presence of highly organized indigenous movements. Control over natural resources has been one of the banners around which indigenous movements have mobilized in recent decades. The region's ecosystem and pattern of water distribution are also important; water tends to be more scarce and unevenly distributed in the Sierra, particularly compared with the Amazon region.

Interestingly, the average marginal effect of environmental aid was very different between cantons inside and outside the Sierra region. Outside the region, a one standard deviation increase in environmental aid boosted the probability of reform attempts by 62 percent. By contrast, the average marginal effect of environmental aid inside the Sierra region was 34 percent. The lower marginal effect inside the Sierra region makes sense given the region's many propitious conditions. The fact that environmental aid's effect was much greater outside the Sierra than inside suggests that transnational environmental networks were able to compensate for the lack of propitious conditions in regions like the Amazon, which lacks a political culture conducive to social organizing and the water scarcity that might induce reforms. In this way, the quantitative analysis produced new hypotheses to be further tested through qualitative methods.

Case Selection for Comparative Case Studies

Once quantitative methods indicated that transnational networks mattered, I turned next to designing case comparisons that could reveal how and why they mattered. I selected six cases in four steps, using both quantitative and qualitative methods. First, I identified cases where similar Integrated Watershed Management reforms were attempted to ensure comparability. This narrowed the list to 26 potential candidates. Second, I selected six cases that ensured instances of both successful and unsuccessful reform, based on the 15-point index described above. Third, I used quantitative methods to select cases that were both "typical" and "deviant" (Seawright and Gerring 2008) from the perspective of existing explanations of local government reform. I did this using the predicted probabilities calculated from my original statistical model that excluded transnational networks.

Using these criteria, I selected six cases grouped into four categories: typical-success, typical-failure, deviant-success, and deviant-failure (see Figure 1). Two cases, Tungurahua and Celica, were predicted to successfully reform and did. Similarly, Zamora's efforts were predicted not to succeed, and they were less successful. I chose these "typical" cases to help identify the micro-causal processes leading to successful reform (George and Bennett 2005). By contrast, two cases, Pastaza and Ibarra, were strongly predicted to reform, but their attempts were unsuccessful. The sixth case, El Chaco, was predicted not to succeed, but it did. These "deviant" cases are useful for uncovering new explanations and causal mechanisms (Seawright and Gerring 2008). Comparing these four case types permitted several forms of analysis. The typical cases allowed me to evaluate the importance of transnational networks vis-à-vis alternative explanations and look for evidence that correlations in my original statistical model were spurious. The deviant cases allowed me to test whether the mechanisms relating to transnational networks facilitated reform even where propitious conditions did not exist (deviant-success), and whether their absence explained failure to reform even where propitious conditions did exist (deviant-failure). Together, these case comparisons constituted a harder test to provide more confidence in the generalizability of my theoretical model of how transnational networks explain variation in the success of local Integrated Watershed Management reforms.

Finally, I also selected these cases to control for alternative explanations of watershed management reform (see Table 1). These include ecosystem characteristics, quantity of available water, land use patterns (e.g., reflected by whether or not landowners live in the watershed, which affects their interest in and use of watershed resources), and demographic information such as the poverty rate, population size, and whether or not indigenous groups are among the watershed's stakeholders. This latter variable is important because indigenous groups tend to have higher levels of social organization, potentially

Figure 1: Four Case Types

	Typical	Deviant
Successful Reform	Tungurahua, Celica	El Chaco
Unsuccessful Reform	Zamora	Ibarra, Pastaza

Table 1: Case Study Characteristics

Case	Success	Ecosystem	Water Scarcity	Landowner Location	Indigenous Stakeholders	Population*	Poverty Rate*
Tungurahua	High	Andean	High	Watershed	Yes	329,856	50%
Celica	High	Andean	High	City	No	14,468	76%
El Chaco	High	Amazon	Low	Watershed	No	7,960	65%
Zamora	Mixed	Amazon	Low	City	No	25,510	61%
Pastaza	Low	Amazon	Low	City	No	62,016	67%
Ibarra	Low	Andean	High	Watershed	Yes	181,175	40%

*Source: 2010 Census, Integrated System of Social Indicators in Ecuador (SIISE), www.siise.gob.ec

making reform attempts easier. The cases also controlled for local political and economic conditions. Both successful and failed cases contained variation in political organization (e.g., relative party affiliations among mayors, municipal council persons, and the national government); the level and form of social organization (e.g., the existence of irrigation councils, indigenous movements, and environmental and development associations); as well as the particular stakeholders involved (e.g., whether hydroelectric companies, indigenous groups, or biodiversity conservationists, among others, were present). In sum, both quantitative and qualitative methods contributed to my case selection in complementary ways and allowed me to control for a wider array of alternative explanations.

Data Analysis

Having identified transnational advocacy networks as my main variable of interest, I needed both qualitative and quantitative methods to analyze if and why these networks explained variation in Integrated Watershed Management reform. The two methodological approaches were complementary in that they addressed different ways in which transnational networks matter. Transnational networks are a useful analytical concept because they combine elements of both structure and agency (Keck and Sikkink 1998). As a structure, they pattern the interactions and relationships among individuals and organizations. As an agent, they articulate and advocate specific policy changes. Quantitative methods can be useful for examining the structure of networks. For example, the positive correlation between environmental aid and reform attempts implied that the pattern of ties within transnational environmental networks—specifically, their geographic reach into different cantons—explains why reforms were attempted in some cantons but not others. Qualitative methods supplemented this quantitative analysis by verifying the validity of indicators and the correlations among them, as well as revealing the causal mechanism behind the correlations. Through network analysis I documented the variation in transnational network connections linking local watershed stakeholders to advocates of Integrated Watershed Management reform, as well as the information and resources flowing through these network ties. Process tracing revealed how differences in network ties produced variation in the reform processes and their outcomes.

Qualitative methods also complemented the quantitative analysis by analyzing the agency of transnational advocacy coalitions. Understanding the agential component of transnational networks was important because, while the diffusion of ideas and practices through network ties provided local actors the opportunity to reform, it did not guarantee success. Qualitative methods were needed to test hypotheses relating to the strategies used by transnational coalitions of advocates to explain why some reform attempts were more successful than others. Without going in to the details of my case comparisons, I used process tracing, network analysis and framing analysis to show that the relative success of reform attempts was explained by variation in the network construction and framing strategies employed in each reform campaign.

Another example of how combining quantitative and qualitative methodologies strengthens data analysis is the way both methodologies were used to overcome the endogeneity problem inherent in my statistical study. While there was undoubtedly a strong and robust correlation between external environmental aid and the tendency to attempt natural resource management reforms, the statistical model could not determine the direction of the causal arrow. It is possible that international actors chose to invest in cantons where they perceived reform attempts to be most likely, or where reforms were already initiated. The argument that external environmental aid led to reforms rather than the other way around rests on two crucial assumptions. First, that the idea to reform came from international rather than local actors. Second, that international environmental actors decided where to invest using criteria that were independent of local actors' desire to reform.

The six case studies provided strong evidence that these assumptions were valid. Process tracing showed that while local actors identified problems related to natural resource management, in each case it was external actors who identified IWM reforms as the best solution to these problems. Furthermore, these international actors often worked for years persuading local politicians and social groups before the reforms were implemented. The case comparisons demonstrated how Integrated Watershed Management reforms grew out of environmental projects financed by international actors a decade earlier. Furthermore, they showed that the financing of these reforms in the late 2000s generally followed many years of similar funding, increasing confidence that my environmental aid indicator was valid.

The counter argument—that the idea to reform came primarily from local actors—was tested through quantitative analysis. A dummy variable in my dataset, *Decentralization Requested*, indicated those cantons where local governments requested environmental decentralization, showing a desire to take on additional environmental management responsibilities. It is reasonable to expect that these cantons would be more likely to pursue natural resource management reforms if such reforms primarily emerge from local actors. If this argument were true, we would expect the variable *Decentralization Requested* to be positively correlated with reforms being attempted. However, the data did not support this.

Regression analysis similarly undermined the argument that transnational actors allocated environmental aid based on which cantons were most likely to reform. If this were true, we would expect *Decentralization Requested* to be a significant predictor of environmental aid. That is, we would expect transnational actors to steer their investments toward local governments that expressed an interest in taking on environmental management responsibilities. Again, the data do not bear this out.

The case comparisons provided further evidence that international actors selected their sites based primarily on criteria other than local desires to reform. While local political will was important, the evidence showed that international actors financing natural resource management activities selected their sites based primarily on the importance and vulnerability of the ecosystem. Other criteria included social needs, such as poverty and water conflicts.

In sum, both quantitative and qualitative data indicated that the assumptions behind interpreting environmental aid as an explanation for reform attempts were valid. All the evidence showed international actors chose where to promote environmental programs based on factors that were independent of a canton's inherent propensity to pursue reforms. It also showed that the impetus to pursue reforms came primarily from international rather than local actors. Together, the evidence indicated that the likelihood of an endogeneity problem was small.

Conclusion

I initially viewed nested analysis as a linear process in which quantitative and qualitative methods were used sequentially, each adding different pieces to a single puzzle. In some sense this is true—seen, for example, in the way each methodology provided evidence related to a different aspect of my explanatory variable and helped resolve issues of endogeneity. However, I have come to realize that the real power of combining quantitative and qualitative methodologies is the way they can inform each other through an interactive process that produces analytic insights greater than the sum of its parts. For me, the power of this interaction was most evident in the way it informed my model specification and case selection, producing a research design that would have been unlikely had I used only one methodology or the other.

References

Asociación de Municipalidades Ecuatorianas. 2004. "Premio a las Mejores Prácticas de la Gestión Pública Seccional-Primera Edición (2003–2004)." Quito, Ecuador: AME.

Asociación de Municipalidades Ecuatorianas. 2008. "Reportes de Gestión-Premio MPS." Quito, Ecuador: AME.

Chen, Xiao, Phil Ender, Michael Mitchell, and Christine Wells. 2007. "Logistic Regression with Stata." Los Angeles: UCLA: Academic Technology Services, Statistical Consulting Group.

Garzón, Andrea. 2009. "Estado de la Acción sobre los Mecanismos de Financiamiento de la Protección y Recuperación de los Servicios Ambientales Hidrológicos." In *Informe Final de la Síntesis Regional Sobre Servicios Ambiental Hídricos en Los Andes*. Lima, Peru: Consorcio para el Desarrollo Sostenible de la Ecorregión Andina.

George, Alexander L. and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Gregersen, Hans M., Peter F. Ffolliott, and Kenneth N. Brooks. 2007. *Integrated Watershed Management: Connecting People to their Land and Water*. Cambridge: CABI Publishing.

Kauffman, Craig. 2012. "Global Governors and Local Governance: Transnational Networks and the Decentralization of Watershed Management in Ecuador." Ph.D. diss. The George Washington University, Washington, DC.

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists Beyond*

Borders: Advocacy Networks in International Politics. Ithaca, NY: Cornell University Press.

Lieberman, Evan. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99:3, 435–452.

Ramón, Galo and Victor Hugo Torres. 2004. *El Desarrollo Local en el Ecuador: Historia, Actores y Métodos*. Quito, Ecuador: Abya-Yala.

Seawright, Jason and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61:2, 294–308.

Seligson, Mitchell A., Juan Carlos Donoso, Daniel Moreno, Diana Orcés, and Vivian Schwarz-Blum. 2006. "Democracy Audit: Ecuador, 2006." Nashville: Vanderbuilt University.

A Horse of a Different Color: New Ways to Study the Making of Citizens

Calvert W. Jones
Yale University
calvert.jones@yale.edu

To what extent can contemporary state leaders use social engineering[1] to produce the citizens they want? Under what conditions do they succeed or fail? The question of how states shape citizens is a classic one, yet it is also one that has taken on additional complexity and alternate "colors" in the contemporary era. Most scholars understandably link the challenge of citizen-building with the effort to foster a common national identity, among the first and most fundamental tasks in state-building. Today, however, and in contrast to earlier eras, billions of people already recognize themselves as citizens of a state, and take its authority for granted. Shifts in the international system, such as the decline of major war and intensification of global economic competition, may also be influencing state leaders' priorities. In these conditions, the willingness and ability of citizens to fight in battle for the state may be less important than their willingness and ability to "fight" in markets by contributing to their nations' economies.

My dissertation ("Bedouins into Bourgeois") investigates how state leaders are re-interpreting the challenge of citizen-building, what outcomes they are achieving, and the conditions for their success and failure. To answer these questions, I use the United Arab Emirates (UAE) as a data-rich empirical laboratory for the contemporary "making of citizens." As in many countries, UAE state leaders are struggling to build more *entrepreneurial citizens*: individuals who will demand less from the state in terms of social and economic welfare, while showing a greater willingness to contribute to market-driven economies and take risks to build vibrant private sectors. With a strategy of "soft" social engineering through education reform, state-sponsored spectacles, and other instruments, they are hoping to create what they see as a more enlightened citizenry, motivated to achieve but still loyal, without provoking unrest of the sort we have seen by way of the Arab Spring as well as the Greek crisis. In the international community, these efforts to build a "new Arab citizen," one who is better educated and equipped to compete in the global economy and less susceptible to radicalism, have been heralded as a way forward for the troubled region (UNDP 2003; World Bank 2007; Faour and Muasher 2011).

Based on my findings, however, "pro-globalization" social engineering is failing in some intriguing and unexpected ways: instead of building entrepreneurial citizens, the data suggest it is giving rise to "*super*-entitled citizens," cultivating youth with expectations of elite status and little interest in private sector work. To explain this outcome, I highlight the role of the middle-men carrying out the social engineering campaign on behalf of autocrats. Largely Western-educated professionals embedded in transnational networks of expertise, they operate in a political context, I argue, that produces behaviors on their part that undermine macro-level strategy. The counter-intuitive result has been an intensified culture of rentierism among the Arab youth who are the target of the campaign, quite the opposite of what was intended.

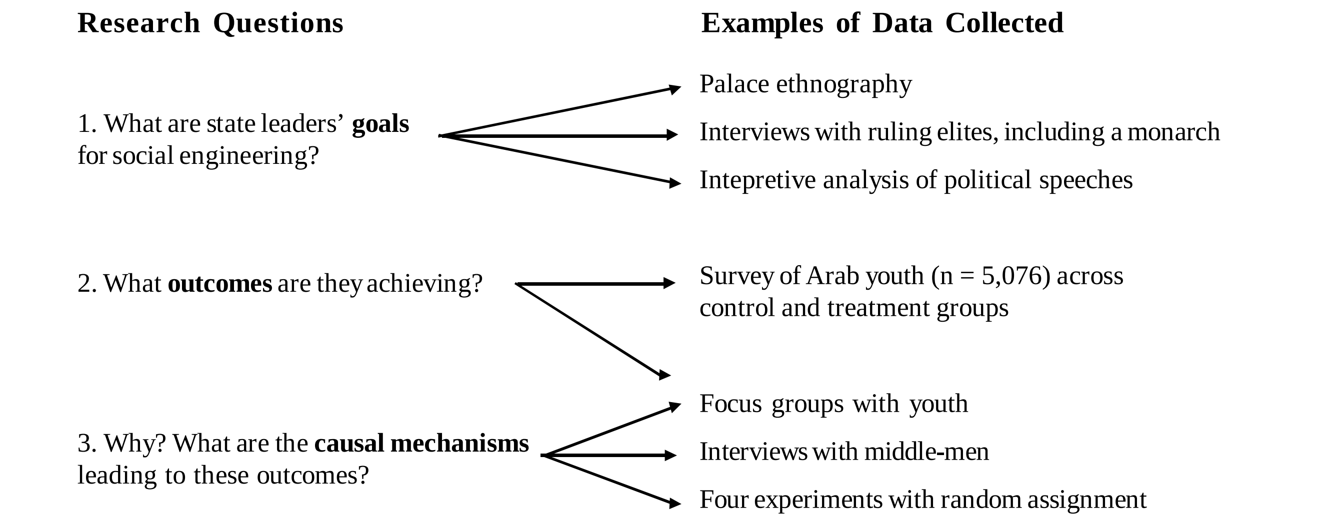
To substantiate these arguments, I use several methods uncommon in the existing literature, aiming to bring new approaches and data to bear the classic question of how states shape citizens. In addition to conducting over a hundred interviews, for example, I surveyed 5,076 Arab youth across the UAE across treatment and control groups. I also gathered data through rare in-palace ethnography, interviews with ruling elites (including a monarch), interpretive analysis of state-sponsored spectacles, focus groups, and four experiments with random assignment. In my contribution to this symposium, I discuss this mix of methods, focusing on their integration in the service of an overarching research design, the substantive payoffs I see arising from that design, and the surprises and lessons learned along the way. A key goal is to illustrate the added value of multi-method research designs for addressing classic questions in new and multi-faceted ways. I suggest that such approaches, like lens focal lengths, can help scholars to see phenomena from different angles, promoting creativity and innovation within broader research streams as well as methodological thoroughness.

Multi-Method Research Design

For the state, the making of citizens is a classic challenge. In the modern era, state leaders have placed particular emphasis on molding the citizen, defined at a minimum by a sense of national identity, recognition of the state as a legitimate political authority, and willingness to obey its rules (Merriam 1931; Bendix 1964; Gellner 1983; Hobsbawm 1992). Beyond these fundamentals, state leaders have also sought to mold the citizen in rather more specific and sophisticated ways, aiming to enhance state power, speed up economic growth, or conform to certain ideas of progress.[2] One of the key comparative questions in these areas concerns the extent to which state leaders succeed or fail in their efforts to shape citizens, engaging in what Rogers Smith has called the politics of "people-building" (Smith 2001). Any attempt to answer this question, however, raises several key challenges.

One challenge involves how to conceptualize and measure success and failure. Although the success or failure of social engineering is often presented in historical hindsight as

Figure 1: Integrating Methods



self-evident, the question is a deceptively simple one, and the answer is rarely obvious. For example, Soviet Russia is frequently given as an example of the failures of social engineering (Alexander and Schmidt 1996; Scott 1998). Yet scholars also argue that many of the attitudes, values, and behaviors forged by the Soviet system have persisted, suggesting Soviet social engineering to create the "new Soviet man" may have succeeded in important ways (Kenez 1985), even if the Soviet regime did not. Another challenge is identifying the effects of social engineering. Makiya (1998), for instance, suggests that Baathist social engineering played an important role in stamping out the ability of Iraqi citizens to think independently, but he also highlights the "culture of fear," driven not by social engineering but by coercion and violence by the state, which contributed to that same outcome. Eugen Weber's famous account of social engineering and cultural change in nineteenth century France (1976) provides another excellent example. In what ways did "peasants" really become "Frenchmen"? What role did the deliberate activities of elites involved in the Paris-based social engineering campaign play in this transformation, as opposed to other causal forces, such as modernization?

In my dissertation, I use a multi-method research design to respond to these challenges, aiming to investigate classic questions of citizen-building in novel ways. A central goal of the study is to help build a more up-to-date and nuanced theory of social engineering for the purposes of citizen-building. To do this, I argue that social engineering goals, outcomes, and causal mechanisms connecting goals to outcomes must be identified with greater conceptual and methodological precision than is typical. A more systematic investigation is needed, in other words, of what state leaders in different contexts *intend* as well as what they *achieve* in their efforts to mold the citizen. For this purpose, I use the United Arab Emirates as a multi-method empirical laboratory (Figure 1). The UAE is a valuable, data-rich context for the study of top-down social engineering, allowing opportunities for ethnography, experiments, process-tracing, congruence-testing, and other modes of analysis. Not only are state leaders unusually open to policy experimentation in the making of citizens, but they also have the political and budgetary flexibility to move beyond rhetoric about change, providing a rich testing ground for theory development. I investigate the following three research questions:

What are state elites' goals for social engineering, and how are those goals being pursued?

What outcomes have been achieved, thus far?

What causal mechanisms best explain why and how these outcomes have occurred?

First, to identify goals, I use a conceptual framework adapted from the literature on nationalism and citizenship (Bendix 1964; Marshall 1964; Kymlicka and Norman 1994). As Figure 2 illustrates, the framework "unbundles" the concept of the citizen and his or her relationship to the state into four component parts: economic, national, political, and civil. Disaggregation allows more precise goals in citizen-building to be identified, taking into account the possibility of multiple, complex, and conflicting goals at the macro-level.

What kind of a citizen do state leaders want to create? To answer this question, I combined evidence from rare in-palace ethnography, interviews with ruling elites, and content analysis of government strategy documents. Official documents can be an excellent source of information about the type of citizen that state leaders want to produce, especially in times of reform when the "citizen of the future" is often described in rich detail. Yet official documents rarely tell the whole story, especially within secretive authoritarian regimes. Thus, I gathered additional ethnographic evidence as a frequent guest at pal-

Figure 2: "Unbundling" The Citizen

Dimension	Empirical Referents
1. Economic	Demands for state-provided welfare
2. National	Attitudes toward national identity/allegiance
3. Political	Demands for political participation
4. Civil	Demands for civil liberties, individual freedom

ace dinners, meetings, and other events attended by ruling elites, where the problems of the youth and strategies for developing them in new ways were a common topic of conversation. I also conducted semi-structured interviews with ruling elites, including one of the country's seven ruling monarchs. In addition, I used an interpretive approach to examine the regime's use of symbolism and spectacle to motivate and model the new entrepreneurial citizen. For example, I visited new schools built as part of the student-centered education reform movement to interview reformers on the ground and attend classes. Following in the footsteps of George Mosse (1975) and Lisa Wedeen (1999), I analyzed major events targeted at youth as political spectacles, including the Young Entrepreneurs Competition, Festival of Thinkers, the Abu Dhabi Science Festival, the Summer of Semiconductors, and the Celebration of Entrepreneurship. At such events, concepts such as "work" are being invested with new meanings that serve a political purpose. For instance, the idea of private sector work is being tied to personal fulfillment and self-discovery in an effort to create a more market-friendly culture with hints of Weber's spirit of capitalism. I also collected over a hundred photographs of installations, artistic exhibits, posters, slogans, and other forms of visual propaganda that reflect the goals of the social engineering campaign.

To answer my second research question, I surveyed over 5,000 Emirati Arab youth across the country, and used a quasi-experimental design to build knowledge about the micro-level outcomes of the campaign. I selected as the "treatment" a new and celebrated public high school, which has served as an important policy experiment in the fostering of the new citizen. My purpose in designing the survey was to uncover the effects of social engineering in a more nuanced and precise way than is typical of the existing literature on the making of citizens. To estimate these effects, I surveyed Arab youth in control and treatment school types as well as earlier ("pre-treatment") and older ("post-treatment") grade cohorts. Using a difference-in-differences (DD) causal identification strategy, I compared differences in students' attitudes across younger and older cohorts within control schools against the differences in those same attitudes across the same cohorts in treatment schools.[3] By identifying micro-level outcomes in the four areas of citizenship reflected in my framework, I was able to test hypotheses about the intended and unintended effects of social engineering.

Finally, I used several methods to explore causal mechanisms in response to my third research question, aiming to unearth the reasons for success and failure in top-down social engineering. I used congruence-testing and process-tracing to examine the "fit" of existing theory in these areas, exploring the extent to which prominent explanations in the literature actually explain the outcomes that I identified. I also triangulated evidence from focus groups with students and interviews with the Western-educated middle-men actors who are carrying out the campaign on behalf of rulers. To supplement this qualitative evidence, I conducted four experiments with random assignment, which I designed to help disentangle the role played by potential causal mechanisms.

Substantive Payoffs

Several substantive payoffs emerge from this approach. First, combining methods allowed me to paint a richer and more empirically accurate portrait of leaders' goals than would have been possible, I think, through the use of one method. When I present my work, an excellent question that often arises is, "Do autocrats really want to change citizens like this? Our models don't predict that, since we assume autocrats want to maintain social stability." Combining palace ethnography, interviews with ruling elites, and document analysis allowed me to "see" that, in fact, UAE autocrats *do* desire changes in citizens that may be surprising from the perspective of existing theory. For example, it was only through ethnography at the palace and other venues in which ruling elites could speak their minds less publicly that I realized the role played by factors like embarrassment. Poor performance by young citizens in school and on international tests—and the idea that the world may not sufficiently respect the UAE as a result—have caused embarrassment and concern about status. Along with other factors, such as over-reliance on hydrocarbons, the questions of how to obtain respect in the world and project modernity, despite authoritarianism, have helped motivate social engineering. I suspect that, on their own, document analysis and semi-structured interviews may not have revealed such rich information about macro-level goals and motivations.

A second substantive payoff has been a set of insights about what leads to success and failure in the use of social engineering to mold the citizen in these ways. In pursuit of authoritarian neoliberal enlightenment, UAE ruling elites have sought to create a citizen who is economically, socially, and culturally more conscious, self-reliant, and hard-working, but nevertheless remains politically passive. Having hired armies

Figure 3: Actual Outcomes of Social Engineering—The *Super*-Entitled Citizen

Dimension	Intended Change	Theorized Change	Actual Change
1a. Economic: *Entrepreneurialism*	↑		
1b. Economic: *Rentierism*	↓		↑
2. National *Nationalism & cultural pride*		↑	↑
3. Political *Demand for political participation (self)* *Demand for political participation (others)*		↑ ↑	↑ ↓
4. Civil *Demands for civil liberties*		↑	

of Western-educated professionals to design and implement "soft" social engineering, they have undertaken a variety of ambitious reforms, ranging from major investments in research and science-based innovation to the building of a new cultural and educational district featuring branches of the Louvre and the Guggenheim, to a radical overhaul of public schools. These types of changes are, in many ways, exactly what critics both inside and outside of the Middle East have long said are necessary for the region's renewal and revitalization.[4]

Yet I find that social engineering is failing and even backfiring in unexpected ways. Instead of building entrepreneurial citizens as desired, the data suggest it may be giving rise to "*super*-entitled citizens," thus reinforcing the very rentier citizenship mentality and sense of entitlement that rulers wish to change. Rather than displaying greater entrepreneurialism and self-reliance, youth subject to school-based social engineering reported *stronger* economic claims on the state, especially through the perceived right to a government job. In the sample, such "treated" youth were also less willing to pay an income tax to support the country's development than were their same-age counterparts across the same grade levels in regular government schools that are not part of the reform movement. In addition, the "treated" youth reported heightened levels of national and cultural pride, and higher levels of interest in political participation for themselves. At the same time, however, they reported lower levels of support for the right of *all* UAE citizens to have a say in government policymaking. I summarize these changes in Figure 3, illustrating the growth of the *super*-entitled citizen with expectations of elite status and little interest in private sector work.

The use of a multi-method research design has allowed not only a more precise and nuanced investigation of social engineering goals and outcomes, but also a means of identifying causal mechanisms. Why has social engineering failed in these ways, and why might it, unintentionally, be producing *super*-entitled citizens? To explain this outcome, I offer an alternative to the conventional wisdom that top-down planners of all stripes fail because they lack local knowledge (Scott 1998; Mitchell 2002). Rather, I argue, the reasons for failure and perverse outcomes in citizen-formation can be found in the political context of implementation by a professional class of middle-men, who are carrying out the social engineering campaign on behalf of rulers. Largely Western-educated experts specializing in areas such as education, business training, and youth development, these middle-men are key players, both whispering into the ear of monarchs and operating at the local level as teachers, principals, and trainers. They operate in a political context producing behaviors on their part that undermine macro-level strategy. At the local level in schools, for instance, they deliver undue praise, inflate grades, and flatter the culture, both pursuing self-interest and job security and conforming to cosmopolitan norms of political correctness. Rarely, moreover, do they "speak truth to power" by telling the autocrats who hired them what they really think. As a result, although wealth, political will, and expertise are not lacking, macro-level strategy is failing.

Evidence from different sources has been crucial in building and substantiating these arguments about causal mechanisms. First, focus groups and interviews with youth helped to illustrate the substantive significance of survey findings. When treated students report stronger support for the right of citizens to receive a government job, why do they feel this way? When they report heightened interest in political participation for themselves, but *less* support for the right of all citizens to participate in politics, what do they mean, exactly, and why? In response to such questions, subjects gave very revealing answers: “Because we are leaders,” “I want to have a good position [in government], a high one, and to have a good salary that fits me,” and “We work hard, we put extra effort in studying. So why would we be equal to them?” Interviews with the Western-educated middle-men also served to clarify the behaviors on their part, influenced by the political context, which are leading to such micro-level outcomes. For instance, these interviews made it clear that the risks of being downgraded, transferred, or even deported were leading new teachers, trainers, and principals to offer excessive praise and flattery, without requiring students to earn it. In a tribal authoritarian context such as this, one text message from an angry high school student to a relative in government can get a teacher fired.

Finally, four experiments with random assignment also helped to elucidate causal mechanisms. The results suggest that even small doses of praise and the related increase in one’s perceived status can affect attitudes in these areas, in this case helping to foster what I call *super*-entitled citizens. Since the large-N survey did not differentiate between students’ attitudes toward their *own* right to a government job and the right of *all* citizens to a government job, these experiments and qualitative data enriched my overall findings about the unintended effects of social engineering, reasons for failure, and the causal mechanisms leading to perverse and unexpected outcomes at the micro-level.

Surprises and Lessons Learned Along the Way

Using multiple methods like this can be a challenge. Yet, while gathering data of different types can increase complexity and attract charges of dilettantism, it can also add value in very significant ways. As I suggest above, a multi-method research design can help to develop, qualify, and enrich findings, keeping the researcher anchored to the overarching question of interest rather than any particular methodology. Beyond this, however, I suggest that multi-method research designs can offer more than, so to speak, the sum of their parts. Such an approach, like experimenting with lens focal lengths, can help scholars to “see” phenomena in new ways, promoting creativity and innovation within broader research streams.

For example, once I had committed myself to a multi-method research design for my dissertation, I was surprised by the relative ease with which I was able to formulate and test interpretive hypotheses about symbolism and spectacle, both important tools within the UAE social engineering campaign. As many a graduate student knows, despite talk of a new era in which all methodologies are treated as equal, significant divides remain. Arguments about potentially differing logics of causal inference are very much alive. Before I embarked on my fieldwork, I recall experimentalists arguing that “You can’t test interpretive claims. They’re not falsifiable.” I also recall consulting interpretive social scientists, who frowned at the use of a large-N survey and experiments to investigate symbolism and spectacle.

As a result, when I began my fieldwork, one surprise was that linking these methodologies was nowhere near as challenging as I had been led to imagine. For me, interpretive methods were essential for understanding what state leaders wanted and how they defined the ideal citizen. They were also critical for understanding how evolving symbolism, regime rhetoric, and spectacles aimed at the rising generation were being put to use for the purposes of social engineering. To employ the typical three-part model of communication, these approaches were most valuable in uncovering (1) the message leaders wish to communicate and (2) the ways in which that message has been communicated to the Arab youth who are the “objects” of the social engineering campaign. Yet, how has this message been (3) received and interpreted? For interpretive political scientists, I believe that a key challenge is how to determine, with rigor and precision, how the symbols, spectacles, and other ideational phenomena involved in the manipulation of meaning actually affect targeted audiences. In this area of inquiry, I found quasi-experimental and experimental approaches especially helpful, supplemented by focus groups and interviews.

Another surprise was how a multi-method approach helped me to “see,” not just the payoffs of combining methods, but the challenge of citizen-building itself in new ways. Before I embarked on my fieldwork, a common question was, “But is this really citizen-building? It’s not a democracy.” That query made me realize how very theory-laden and potentially narrow our conceptualization of this phenomenon may be. For instance, we understandably link the term “citizen” with the Western historical experience of liberal democracy. However, as Aristotle pointed out, the relationship between the individual and the state can be conceived in far more multi-dimensional ways. Second, we associate the challenge of citizen-building with the effort to foster a common national identity, among the first and most fundamental tasks in state-building. While this task remains important, circumstances today are different from earlier eras, since billions of people already have a national identity. A key question for citizen-builders, then, is what comes next? The challenge of citizen-building today may thus be a “horse of a different color,” similar in fundamentals to what it was before but imbued with new meanings. Multi-method research can help us explore such classic questions in new ways, facilitating not just methodological rigor but conceptual creativity and theoretical risk-taking.

Notes

[1] I define “social engineering” as activities consciously undertaken by state leaders to shape the hearts and minds, and ultimately the culture, of their own citizens. My definition restricts these activities to socialization and the management of meaning through educational initiatives, state-sponsored spectacles, media campaigns, and other non-violent methods. (Of course, some regimes have also used force to shore up their efforts at “making” citizens, but I treat coercive power as a separate causal factor that can be distinguished conceptually from social engineering.)

[2] See, for example, Kenez (1985) on Soviet Russia, Hanioglu (2011) on Turkey under Ataturk, Garon (1997) and Dower (1999) on post-1945 Japan, and Alston (1969) on Tsarist Russia.

[3] This approach has several advantages for causal inference. First, it controls for selection bias in treatment assignment. Selection bias is a well-known challenge to causal inference in these areas; students, of course, are not randomly assigned to schools. The DD approach removes this type of selection bias by subtracting out initial differences in outcomes between control and treatment populations, preventing any unobserved factors that remain constant over time, which correlate with treatment assignment and affect the outcome variables, from biasing treatment effect estimates. Such factors may include income levels, levels of parental education, and other demographic differences. Another advantage is the removal of bias stemming from aggregate factors that would cause change in the outcome variables over time or across grade cohorts even in the absence of the treatment. Such factors include age or maturation, broad socio-economic changes, and national or regional political context.

[4] See, for example, the UN's Arab Human Development Report, "Building a Knowledge Society" (2003) and Nasr (2009).

References

Alexander, Jon, and Joachim K.H.W. Schmidt. 1996. "Social Engineering: Genealogy of a Concept." In *Social Engineering*. Adam Podgórecki, Jon Alexander, and Rob Shields, eds. (Ottawa: Carleton University Press), 2–19.

Alston, Patrick L. 1969. *Education and the State in Tsarist Russia*. Stanford: Stanford University Press.

Bendix, Reinhard. 1964. *Nation-Building and Citizenship*. New York: Wiley.

Dower, John. 1999. *Embracing Defeat: Japan in the Aftermath of World War II*. London: Penguin Books.

Faour, Muhammad and Marwan Muasher. 2011. "Education for Citizenship in the Arab World: Key to the Future." In *The Carnegie Papers*. Carnegie Endowment for International Peace.

Garon, Sheldon M. 1997. *Molding Japanese Minds: The State in Everyday Life*. Princeton: Princeton University Press.

Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca: Cornell University Press.

Hanioglu, M. Sükrü. 2011. *Atatürk: An Intellectual Bibliography*. Princeton: Princeton University Press.

Hobsbawm, Eric J. 1992. *Nations and Nationalism since 1780: Programme, Myth, Reality*, 2nd ed. Cambridge: Cambridge University Press.

Kenez, Peter. 1985. *The Birth of the Propaganda State*. Cambridge: Cambridge University Press.

Kymlicka, Will and Wayne Norman. 1994. "Return of the Citizen: A Survey of Recent Work on Citizenship Theory." *Ethics* 104: 352–381.

Makiya, Kanan. 1998. *Republic of Fear: The Politics of Modern Iraq*. Berkeley: University of California Press.

Marshall, Trevor H. 1964. *Class, Citizenship, and Social Development*. New York: Doubleday.

Merriam, Charles Edward. 1931. *The Making of Citizens: A Comparative Study of Methods of Civic Training*. Chicago: The University of Chicago Press.

Mitchell, Timothy. 2002. *Rule of Experts: Egypt, Techno-Politics, and Modernity*. Berkeley: University of California Press.

Mosse, George L. 1975. *The Nationalization of the Masses: Political Symbolism and Mass Movements in Germany from the NapoleonicWars through the Third Reich*. New York: H. Fertig.

Nasr, Vali. 2009. *Forces of Fortune: The Rise of the New Muslim Middle Class and What It Will Mean For Our World*. New York: Free Press.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press.

Smith, Rogers M. 2001. "Citizenship and the Politics of People-Building." *Citizenship Studies* 5:1, 73–96.

UNDP. 2003. *Building a Knowledge Society: The Arab Human Development Report*. New York: UNDP.

Weber, Eugen Joseph. 1976. *Peasants into Frenchmen: The Modernization of Rural France, 1870–1914*. Palo Alto: Stanford University Press.

Wedeen, Lisa. 1999. *Ambiguities of Domination: Politics, Rhetoric, and Symbols in Contemporary Syria*. Chicago: University of Chicago Press.

World Bank. 2007. "The Road Not Traveled: Education Reform in the Middle East and North Africa." In *MENA Development Report*.

Using Formal Theory in Multi-Method Research: National Democratic Support for Subnational Authoritarianism

Juan Rebolledo
Yale University
juan.rebolledo@yale.edu

Beginning my third year of graduate school, I faced the daunting task of going from a scrutinizer of academic work to a producer of such work. Coming from Mexico shaped my general areas of interest, but narrowing my project down to a manageable yet still relevant question was a tough task. I was concerned with the political development of young democracies and intrigued by the existence, in Mexico, of sub-national governments with authoritarian characteristics that had successfully maintained power despite national democratization. After reading V.O. Key's canonical text *Southern Politics in State and Nation* (1949), I realized this was not a Mexico-exclusive phenomenon. Upon further research, I recognized that similar dynamics were occurring in countries as diverse as India, Brazil, the Philippines, and Argentina. The very existence of these authoritarian pockets in democracies was puzzling, yet it is even more puzzling that a national democratic government seems to be unable or unwilling to challenge subnational authoritarian practices, even when those regions are controlled by the opposition. I successfully narrowed the topic to an intriguing question: In young democracies, why do democratic national governments, which recently fought authoritarian abuses, seem to be unwilling or unable to act against regional autocrats?

Given the relatively little work (Gibson 2005, Gervasoni 2010, Giraudy 2010) done in this area of research, and the complications that empirical scholars have faced for years when studying democracy, it became clear to me that demonstrating an unequivocal causal relationship would be nearly impossible. There would be no perfect instrument for level of democracy. Nor would I find a natural experiment in which some states were "as if" randomly assigned to retain authoritarian characteristics while others became more democratic. In addition, the complexity and many moving parts of the question made it likely I would miss interesting and important implications of the argument, or even worse run the risk of falling into logical inconsistency, if I forced a single answer onto the question. At times the issue felt intractable, but it became clear that I could only make the project both theoretically manageable and causally explanatory by bringing multiple methods to the task. I thus combined a formal model, quantitative analysis, and qualitative case studies in the project.

The Formal Model

Timing presented the biggest challenge of using a multimethod approach. On one hand, I wanted every aspect of the research project to be informed by the different methods; on the other hand, I wanted the project to be deductive. My knowledge of the cases had to illuminate the assumptions I was prepared to make, but not the hypothesis itself. I had spent time doing fieldwork in Latin America on other projects, which helped me get a sense of the relevant political actors and the strategies available to them. Extensive background reading into other country cases suggested these phenomena were generalizable. However my theory at this point had an excessive number of moving parts. I began working on the formal model, which brought clarity to my theory. I knew who the key actors were: the central government, the governors, and the electorate. The model helped me make my assumptions explicit, and forced me to sort those assumptions I was willing to make from those I was not. It allowed me to incorporate aspects I felt had to be part of the story (governmental resources, resource exchange among different levels of government, expenditure on electoral cycles, the level of democracy and policy making), while forcing me to leave aside inessential aspects of the theory. While many of the inessential aspects are still substantively important, they had to be removed to make manageable an otherwise unmanageable and irreducibly complex reality. In the end, the model generated interesting, counterintuitive, and testable propositions that I would not have uncovered without it.

Let me briefly describe the model and the main propositions derived from it. To understand the endurance of regions with high levels of authoritarian practices, we must pay close attention to the strategic interaction between the central government and the different regions. In particular, in the dissertation I propose that because of support the central government needs from the opposition to get its preferred policy approved, it tolerated states controlled by the opposition with high degrees of authoritarian practice persistence. The formal model consists of subnational regions belonging to different parties, the central government from one of these parties, and the citizens. In my model, regions can vary in their levels of authoritarian practices.[1]

In this probabilistic voting model, citizens can vote for either party. Because a citizen can cast only one vote in the model, citizens from the region in which the central government and the regional government are controlled by different parties must support one at the expense of the other. The central government has two distinct concerns: its party's electoral fortunes (obtaining a large vote share in both regions) and its party's policy agenda. The central government has resources it can spend in any region to sway the electorate to support its party and must choose how to divide these resources among the regions. When the central government spends resources in a region controlled by an opposing party, the governor of that region and his party will have to spend more resources trying to stay in power.

A central government wishing to advance its policy agenda must go through the political process (e.g., a legislative bill in the national congress) and will, therefore, need the support of some members of the opposing party. A central government that does not need support from an opposing faction to implement its desired policy is beyond the scope of this theory. However, I contend that, in most cases of national democracies, no single faction can determine policy unilaterally.

Regional governors are able to offer such support. Governors often have undue influence over legislators for a variety of reasons: because legislators owe their nomination to the governor, as is the case in modern Argentina (Jones and Hwang 2005, Behrend 2011), or because support for systematic disenfranchisement in a region assisted the power holders in being elected, and unites them across an array of policy issues, as was the case in the U.S. South during the Jim Crow era.[2] Crucially, the central government can achieve its preferred policy by negotiating with regional opposition leaders, rather than opposition leaders at the national level who have already expressed opposition to the policy in question. If the government cares about policy, it may offer to reduce electoral intervention in regions controlled by an opposing party in exchange for the political support of that region in advancing its policy agenda.[3]

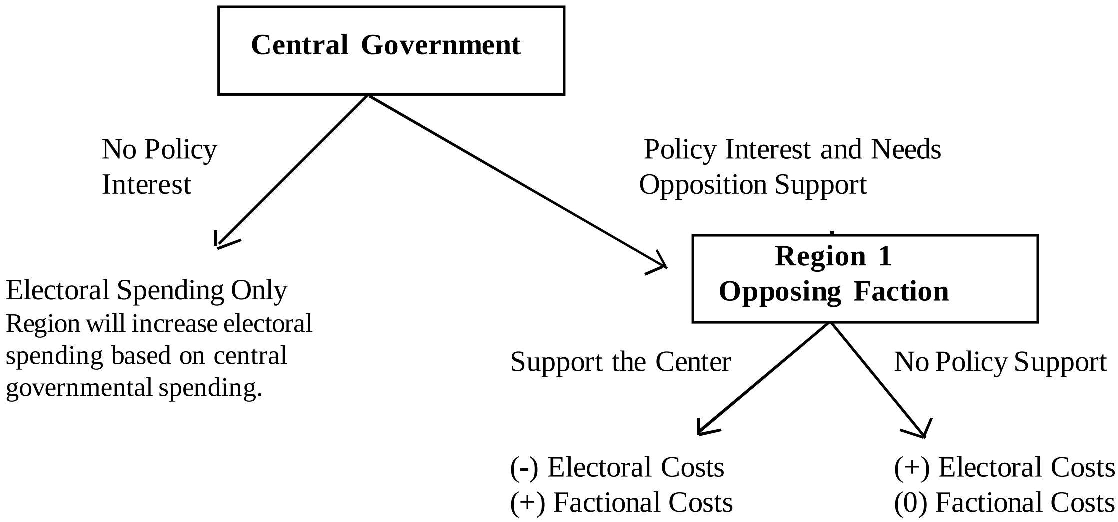
Regional governors face a lose/lose situation. On the one hand, they face an electoral threat from the central government, which, unless they offer their support to the opposing party's agenda, will spend high levels of resources in their region attempting to oust them. Even if the central government does not succeed, its spending resources in this manner entails higher electoral costs for the regional governors, who must react to increased resources spent against them by increasing the resources used to defend themselves. On the other hand, regional governors also face a cost imposed by their party if they decide to support a central government led by an opposing party. The cost imposed by their own party for defecting can manifest itself in a variety of ways, including resource flow, public perception, and electoral support.[4]

A region controlled by an opposition party will make a cost-benefit comparison between the party-imposed costs of supporting the central government and the electoral costs of party loyalty.

Figure 1 presents a stylized version of the formal model based on the current presentation of the argument.

The model shows that an equilibrium exists (out of two possible equilibriums) in which the central government offers to reduce electoral spending in a region controlled by an opposing party in exchange for policy support at the national

Figure 1



level. For this equilibrium to exist, two conditions must hold: First, as the chart shows, the central government must have policy interests; second, trade-offs will occur only if the region can overcome the cost imposed by its national party. What makes these agreements more likely, especially if the electoral and party costs are held constant? Leaders of regions with low levels of democracy have a higher probability of electoral victory, inducing them to spend less on elections. As a result, they can compensate for costs imposed by their own national party more easily than their counterparts in high-democracy regions. In addition, low-democracy regions are willing to accept a lower concession from the central government precisely because their entrenched position and control allows them to more easily overcome party costs. Ceteris paribus, regions with lower levels of democracy have a higher probability of absorbing their national party's costs, are less affected by central government intervention, and are thus more likely to meet the second condition for defection in a much broader set of cases.

The security and control of the governing group in a low-democracy region allows it to be more independent of the national party, and therefore more likely to reach an agreement with the central government. Nevertheless, party defection by a regional governor to support the central government is still a rare event. If party costs are sufficiently high, no degree of authoritarian continuity will be enough for an agreement to be reached.

The theory indicates not only that low-democracy regions have a higher probability of supporting the central government, but also, precisely because they supported the central government, electoral efforts by the central government's party against the regional governor from an opposing party will decrease. Regions governed by an opposing party that supported the central government parties' agenda will observe lower levels of expenditures against them by the central government in their region. We are left with two very unlikely allies, with the national democratic party allying itself with the regional low-democracy governors from an opposing party.

Quantitative Analysis

Having used the formal model, I developed clear propositions which were not obvious at first. Testing the theory clearly required evidence not only of the causal claim I was making, but also of the assumptions that led me to the prediction in the first place. This meant generating a new dataset. The reinforcing nature of the multi-method project became very handy. Just as my knowledge of the key strategic actors and relevant variables, based on previous fieldwork, were fundamental to designing the model, now I began with a concrete list of variables, with specific meanings in the context of the model. I began searching for proxies of these variables. Using multiple methods helped set a clear direction to the data collection process.

For the empirical analysis, I constructed a panel dataset of all 32 states in Mexico for the period 1997–2008. To the best of my knowledge this is the only study that includes all 32 states and covers not just one presidential administration but three distinct administrations, including both PRI and PAN national governance. Mexico lends itself to the study of cross-regional variation in the persistence of subnational authoritarian institutions. Mexico is a federation composed of 32 subunits, each with its own executive, legislative, and judiciary branch. Nationally the country had an autocratic political system, characterized by a single-party hegemonic regime, which began crumbling in the late 1980s. There is much literature on how the Institutional Revolutionary Party (PRI) stayed in power for so long. However, a rough consensus holds that a combi-nation of good economic performance in the 1960s and 1970s, in conjunction with a rupture-preventing strategy, helped the PRI maintain hold (Magaloni 2006, Ames 1970). After the 1993 and 1996 round of electoral reforms that "leveled the playing field" for the opposition, there was finally a guarantee of free and fair elections. The advent of national democracy in Mexico is usually attributed to the growing support for the opposition after recurrent economic crises during the 1980s and 1990s. These crises, in conjunction with economic development that had occurred in the previous two decades, freed citizens from reliance on the state and allowed them to make ideological investments. Regardless of the theory of the national democratization of Mexico, the important point is that it was an exogenous shock to subnational units' political power.[5]

The democratizing reforms crystallized in 1997 when the PRI lost majority control of the congress and needed the support of the opposition to pass legislative initiatives. In 1997, under President Ernesto Zedillo of the PRI, Mexico was living in a new democratic reality, in which the President needed opposition support to pass his legislative agenda. However subnational alternation in office had begun some years before. In 1989, an opposition governor first took office in Baja California after a series of post-electoral conflicts. Research has already been conducted on subnational alternation in an authoritarian setting, and it is generally understood that alternations in gubernatorial offices before 1994 were not the product of elections (e.g., Eisenstadt 2004). Since national democratization, some states followed the national trend and became democratic. In other states, governors successfully concentrated their power and were able to sustain authoritarian practices and institutions in the newly democratic setting. Thus, as is widely acknowledged, Mexican states vary significantly in their level of democracy (Fox 1994, Giraudy 2010).

The 31 Mexican states, and the federal district, enjoy a great deal of autonomy. They all drafted their own constitutions. Governors are all elected for six-year terms and local legislators for three-year terms, with no reelection allowed in either branch. Since 1997, all gubernatorial candidates have been products of some sort of real electoral process taking place within the state. Fiscally, states also enjoy a fair degree of autonomy. They receive both an automatic transfer from the central government and a potential discretionary transfer; since the 1990s, the *automatic* transfers have been large enough that states are not financially dependent on the central government. This arrangement means that local taxation is quite low and deficient (around 80% of a state resources come from federal transfers).

Taking advantage of state fixed-effects, I find evidence that, as the theory suggests, lower levels of democracy in an opposition state mean a higher likelihood that state will defect from party lines and support the central government. The evidence is not only statistically significant, but substantively significant as well: States that retain the most authoritarianism are 37.9% more likely to support the central government than states with a mean level of democracy.

I also find evidence that opposition governors who defected from party lines to support the central government experience lower levels of hostile spending from the central government in their regions (measured as media expenditures by the federal government in the region). The measure is also both statistically and substantively significant. The effect is large: having supported the central government in the year prior to an election decreases the media expenditures by the central government in the state by about 64%. Both of these results are robust in terms of distinct specifications and conceptualizations of the variables, adding certainty to the results.

Though unequivocal causation cannot be established by quantitative evidence alone, these findings invite further investigation into both the conditions under which these unlikely alliances are made, and the benefits to the parties involved. This study also makes explicit that researchers should pay attention to the legislative incentives behind such agreements.

Qualitative Case Studies

The quantitative evidence I present in the dissertation clearly shows a link between opposition regions with low levels of democracy and the likelihood they would defect from party lines to support the democratic national government. Nonetheless, more evidence needed to be presented to show that the mechanism driving this association was the one I proposed through the aforementioned formal model. Just as I attempted to bring the advantages of qualitative work to my quantitative analysis—by emphasizing concept development and choosing proxies based on a deep understanding of recent political history—I tried to bring the rigor of quantitative analysis to my qualitative research. To that end, I pursued both within-case analysis and across-case variation.

To better understand the mechanism and the different equilibriums, it was important to look at cases that present variations in the key independent and dependent variables, in other words variations in level of democracy, support for the central government, and level of central intervention in the region. Additionally, because elections in Mexico are staggered by state, I wanted to examine states facing the same relevant actors and actions during their electoral years. For this reason I chose four states that had elections in 1998, 2004, and 2010. Puebla, Veracruz, and Oaxaca have low levels of democracy, whereas Chihuahua is widely held to have a high level of democracy.

First, I undertook within-case analysis of each state, presenting evidence both that the assumptions of the model hold, and that the suggested causal mechanism is at work. For this I conducted over 50 interviews with elites and non-elites in the states of Puebla, Veracruz, Chihuahua, Queretaro, Michoacan, and Mexico City. Among them were: the heads of all important parties in the region, the campaign managers, heads of newspapers, journalists, members of the electoral commission, presidents of local universities, former local legislators, current local legislators, and former national legislators, among others. Archival research was conducted to analyze newspaper clippings in the different states.

Using a combination of the primary evidence from fieldwork and research on the recent political history of each state, I presented a detailed account of the process that takes temporality, as well as the sequence of events, as a crucial part of the evidence. After presenting within-case studies for all four states, I pursued a case study of a particular bill, the 2003 attempted fiscal reform, to analyze cross-state variation in gubernatorial defection and central government intervention during the next election. Careful case selection of all four states allowed me to isolate the variables of interest by holding other factors constant: In addition to having elections in 2004, all four states were governed by the PRI while Vicente Fox of the PAN was president. This bill case study allowed me to identify cross-state variation, which had so far been absent from the quantitative analysis.

Conclusion

In the end, while perhaps no individual piece of evidence was a silver bullet, the combination of methods presents a strong case for the theory proposed. Different types of evidence all pointing in the same direction reduce skepticism. However, a word of caution is warranted. Issues of timing in the use of the different methods were a constant concern. I was especially concerned about producing deductive work. I did not want to make the process of theory generation and testing identical by generating a theory based on Mexico, and then testing that theory using Mexico as a case.

Researchers using multiple methods that include quantitative and qualitative work confront a tension between needing to understand cases in order to identify the relevant actors and their possible strategies, and generating a theory independent of evidence gleaned from fieldwork—*unless* the researcher plans to test the resulting theory in a different context. The testable predictions I presented were not generated by my original fieldwork, which informed the model but could not have anticipated the comparative statics that it generated. In addition, I incorporated Argentina as a shadow case in the concluding chapter of the dissertation. This both tests the generalizability of the theory and serves as an out-of-sample test.

Moving from the dissertation to a book project, I intend to continue using a combination of quantitative and qualitative methods to fully analyze the theory in three different countries: Brazil, Argentina, and the United States. I have begun collecting data for these countries, which I plan on complementing with in-depth case studies in each. In addition, now that I will be making explicit comparisons across countries, some of the variables of the model that were fixed for the case of Mexico will vary, allowing me to test further implications of the theory.

Notes

[1] Examples of regions that maintain high levels of authoritarian practices are, in Argentina, La Rioja, San Luis, Santiago del Estero, Santa Cruz, Rio Negro, and Formosa; in Mexico, Oaxaca, Hidalgo, Veracruz, and Puebla; and in Brazil, Maranhao, Para Piauy, and Bahia until 2006.

[2] Brazil is another example of a country in which the central government goes through the regional bosses to obtain congressional support. Samuels (2003) defends at length this position, on the other hand, Cheibub et al. (2009) find that the regional effect is not as large as expected. The theory in this present work is consistent with Cheibub because we would expect parties to have the greatest influence and that state effect would be rare, as predicted by the model.

[3] Central government intervention in a region can take a variety of forms, including declarations of federal intervention that remove the governor, resources spent garnering votes, deploying the central government's intelligence, and influencing the Supreme Court in sanctioning (or not) governors that continue authoritarian practices.

[4] This was the case with the UCR in Argentina from 2006–2010 in threatening to expel from the party legislators and governors who defected from party lines to support the central government. They also were threatened with intervention in the provincial party committees.

[5] Perhaps the only theory of Mexican democratization that would be problematic is presented by Lujambio and Segl (2000) who claim that national democratization in Mexico was only possible via local democratization as a first step. However, this view is not widely held. Local democratization in a few municipalities is usually seen as temporary exceptions.

References

Ames, Barry. 1970. "Bases of Support for Mexico's Dominant Party." *American Political Science Review* 64:1, 153–167.

Behrend, Jacqueline. 2011. "The Unevenness of Democracy at the Subnational Level: Provincial Closed Games in Argentina." *Latin American Research Review* 46, 150–176.

Cheibub, José Antonio, Argelina Figueiredo and Fernando Limongi. 2009. "Political Parties and Governors as Determinants of Legislative Behavior in Brazil's Chamber of Deputies 1988–2006." *Latin American Politics and Society* 51:1, 1–30.

Eisenstadt, Todd A. 2004. *Courting Democracy in Mexico: Party Strategies and Electoral Institutions*. London: Cambridge Universtiy Press.

Fox, Jonathan. 1994. "Latin Americas's Emerging Local Politics." *Journal of Democracy* 5:2, 105–116.

Gervasoni, Carlos. 2010. "A Rentier Theory of Subnational Regimes: Fiscal Federalism, Democracy and Authoritariansim in the Argentine Provinces." *World Politics* 62:2, 302–340.

Gibson, Edward L. 2005. "Boundary Control: Subnational Authoritarianism in Democratic Countries." *World Politics* 58:1, 101–132.

Giraudy, Agustina. 2010. "The Politics of Subnational Undemocratic Regime Reproduction in Argentina and Mexico." *Journal of Politics in Latin America* 2:2, 53–84.

Jones, Mark P. and Wonjae Hwang. 2005. "Provincial Party Bosses: Keystone of the Argentine Congress." In *Argentine Democracy: The Politics of Insitutional Weakness*. Steven Levitsky and María Victoria Murillo, eds. (University Park, PA: Pennsylvania State University Press), 115–138.

Key, V. O. 1949. *Southern Politics in State and Nation*. Knoxville: University of Tennesee Press

Lujambio, Alonso with Horacio Vives Segl. 2000. *El Poder Compartido: Un Ensayo sobre la Democratización Mexicana*. Mexico: Oceano.

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico*. London: Cambridge University Press.

Samuels, David. 2003. *Ambition, Federalism and Legislative Politics in Brazil*. London: Cambridge University Press.

Multi-Method Fieldwork in Practice: Colonial Legacies and Ethnic Conflict in India

Ajay Verghese
Stanford University
ajayv@stanford.edu

It seems odd these days for a young graduate student to suggest a dissertation project which does not include a multi-method research design. The multi-method movement in political science has become central to the way we structure our research. A carefully executed multi-methods project is the holy grail to which many a scholar aspires. And yet, graduate students have only limited advice in carrying out this kind of research. My goal in this article is to lay out the inner workings of a multi-method project. By presenting the components of my dissertation, which combined a large-N analysis of 589 districts in India with 15 months of archival research and elite interviews carried out in six case studies, I explore: (a) how a multi-method project is created and implemented, and (b) its potential pitfalls and payoffs.

My project (Verghese 2012) examines the puzzle of why ethnic conflicts in multi-ethnic states revolve around one identity rather than another. So, for example, why do conflicts sometimes revolve around religion but at other times around language? What explains *patterns* of ethnic conflict in a multi-ethnic state? This is an important question for plural states around the world struggling to limit ethnic violence. I had chosen this research question in part because of important recent work published on ethnic violence in India (Brass 1997, 2003; Varshney 2002; Wilkinson 2004). While these books had advanced our knowledge of ethnic conflict in the country considerably, I was rather surprised that they all omitted a large potential explanatory variable: the legacy of British colonial rule. After all, India was the crown jewel of the British Empire, and the political science literature linking colonialism and contemporary ethnic violence is immense, much of it focused on African cases (Horowitz 1985; Laitin 1986; Young 1994; Mamdani 1996; Posner 2005).

Furthermore, India presented a unique opportunity to study the impact of colonialism on ethnic conflict because only three-fourths of the population of the country ever came under direct colonial rule. These areas, known as *provinces*, were governed by British administrators. But the rest of the country remained under the control of independent native kings in territories called *princely states*. By comparing conflict outcomes across provinces and princely states, I hoped to isolate the effects of colonialism on ethnic conflict.

But first: why use multi-methods in this project at all? I admit that my initial intentions were hardly honorable; I wanted to utilize multi-methods because that's what you were supposed to do, especially if you wanted to write a well-received dissertation. Luckily, I later came to realize that a project as broad as this would have suffered had it utilized merely one methodology. Imagine only a statistical analysis—you would instantly say, "But you never spent a day in the field or in the archives!" Likewise, imagine fieldwork in two case studies; you might here rightly ask: "But does the argument travel?" Multi-methods have achieved a place of prominence in political science research not because it is simply the newest fad, but because it allows researchers to examine questions in a more complete and exhaustive fashion.

But the larger problem with combining methodologies was that the kind of work I really admired and wanted to do—comparative-historical analysis—was rarely married together with statistics.[1] What interested me was colonialism in India. But how exactly did one go about combining, for instance, archival research on the colonial period with a regression analysis? The answer was not obvious to me. An especially vexing problem was collecting data. I quickly realized that finding reliable figures on ethnic conflict *during* the colonial period in India would be almost impossible, so I would have to restrict my statistical analysis to the contemporary (post-independence) period.

However, there is no inherent contradiction between doing comparative-historical work *and* statistically-oriented research. I planned to run a statistical analysis of the broad pattern of ethnic conflict in contemporary India, but the comparative-historical analysis, on the other hand, would be situated within a number of targeted case studies, aimed at uncovering the mechanisms at work in producing specific violence outcomes. The two methodologies seemed complementary rather than conflicting, and together could help explain not only contemporary outcomes but also their historical causes.

Dataset Construction and Statistical Analysis

Once I had decided that I wanted to pursue a multi-method dissertation, I then set about figuring out how to actually do it. I started with the large-N analysis. I wanted to examine the broad pattern of ethnic conflict throughout modern India and its potential causes, and statistical analysis offered the best opportunity to see the big picture. I began by spending an inordinately long time collecting data on a variety of variables for my study.

I first considered the unit of analysis which I wanted to examine for this research project. Because I was interested in the effects of colonial rule, I decided to pursue a district-level analysis. The entire system of district administration in present-day India was a legacy of the British period; districts, for the most part, were either completely part of a former British province or a former princely state. Looking at states, on the other hand, was much more problematic: a state like Kerala in southern India, for example, was half-British and half-princely, which posed a major coding problem. But looking at the districts *within* Kerala made identifying British and princely areas much easier. I used the list of 2001 districts from the Indian census and ended up with a total of 589 for the analysis.

Then I coded the primary independent variable: the type of colonial rule. This was the most painstaking process of all: I had to determine whether every district in India was either part of a former province or princely state.[2] This necessitated re-searching district websites, reading British colonial reports on individual districts, and comparing geographical coordinates between the two. I coded the type of colonial rule in two ways: a dummy variable (1 if a province), and a variable recording the number of years a district was under British rule (0 for all princely states). I hoped that using two different measures would increase the confidence in my coding.

I then began to compile figures on ethnic conflict using two different existent datasets. The first source of data was the Worldwide Incidents Tracking System (WITS), which used national and international press reports to provide figures on caste and tribal conflict throughout India during the period 2005–2009. The second source was the Varshney-Wilkinson dataset on Hindu-Muslim riots in India covering the period 1950–1995. Together, these two sources of data gave me a broad view of ethnic conflict throughout the post-independence Indian republic. I couldn't document what had happened during the colonial period, but I could at least know about ethnic conflict today, and with smartly selected case studies, I hoped to be able to uncover the deeper causes behind it.

I finally compiled dozens of control variables from various sources: the Indian census, a private statistical firm (IndiaStat), and the Indian Human Development Survey, carried out by researchers at the University of Maryland. Examples of some of these variables were contemporary data on population, geography, the economy, and infrastructure. These variables allowed me to account for a number of alternative arguments about the causes of ethnic violence. Because the conflict data for my dependent variables were count variables (i.e., number of deaths and injuries in ethnic conflicts), I utilized a negative binomial regression model.

I hoped that by constructing this highly detailed database, both the specific independent variable in which I was interested (the type of colonial rule) and the specific dependent variable of interest (ethnic conflict) could then be used by other scholars in future studies. My hope was that a scholar interested in, for example, the effect of colonial legacies on Indian political parties could use my coding scheme of provinces and princely states; or, a scholar interested in the effect of poverty on caste riots might find my compilation of WITS conflict data to be helpful in that regard.

The results of my analysis confirmed that colonial rule had a major impact on ethnic violence outcomes in modern India, but not in the way I had initially expected. I found that in former provinces, *caste and tribal conflict* was the major problem; however, in former princely states, *religious conflict* was endemic. I had simply expected that *all* British provinces would be worse in terms of ethnic violence, but there was an important dichotomy in violence outcomes which I had not anticipated. My statistical analysis therefore confirmed that my chief independent variable of interest—colonial rule—was important, and it had likewise allowed me to rule out a number of potential alternative explanations. It gave me support for my working hypothesis, which I could then further investigate using qualitative fieldwork in India.

Comparative Case Study Fieldwork

The next step in my project was qualitative fieldwork. The statistical analysis alone was not enough. How could I explain the result which I had found? What *specifically* about colonial rule created this apparent dichotomy in ethnic violence outcomes? What were the mechanisms at work? Although you could certainly test mechanisms using certain kinds of advanced quantitative techniques, I felt the need to get my hands dirty and spend some time in the field. I wanted to unpack the logic at work that drove ethnic violence. Furthermore, I believed that qualitative fieldwork would lend a certain credence and believability to the project which a statistical analysis alone could never do.

First, I needed to carefully pick cases to study. This proved a rather daunting task. My working hypothesis, supported by my statistical analysis, was that variations in colonial rule effected contemporary patterns of ethnic conflict. So ideally what I wanted were paired comparisons—that is, two cases which were similar in almost every regard except for variation on a key independent variable of interest: colonial rule. I had the image in my head of exactly what I hoped to find: one princely state situated right next to a British province.

As I had spent some time in a famous princely state of north India while learning Hindi, this naturally became my first case (Jaipur). And lo and behold, right next door was a former British province (Ajmer). I called these two cases a "paired historical comparison," adapting terminology used by George and Bennett (2005: 151). I began my research in this area in the fall of 2010. The first thing I needed to do was figure out whether the pattern from my statistical analysis was also evident at the small-N level of analysis. That is, did Jaipur, as a former princely state, experience more religious conflict than Ajmer? And did Ajmer, as a former British province, experience more caste and tribal conflict than Jaipur? This is what my theory would predict, so I viewed qualitative fieldwork as both an opportunity to investigate mechanisms *and* an opportunity to re-confirm my broader hypothesis.

Again, it is worth pausing for a moment to acknowledge the anxiety you feel in the field when you realize that your hypothesis may not be supported. And this was how I felt while collecting data in Jaipur and Ajmer about contemporary patterns of ethnic conflict. I had a sneaking suspicion every case I studied might turn out to be a deviant case. My qualitative research in the area consisted of two components: elite interviews (to figure out the state of violence in the contemporary period), and archival research (to figure out the underlying cause of this violence).

So I began with interviews, and I had been told that you always begin with journalists. No one knows more about the broad politics of an area than a journalist. Then I expanded my interviews to include police officers, government officials, NGO workers, ethnic group leaders, and a wide variety of other respondents. What I found was that the same pattern of ethnic conflict which I had uncovered at the large-N level of analysis was also evident when looking at Jaipur and Ajmer. Jaipur was indeed a major area of religious conflict, but in Ajmer the viol-

Table 1: Verghese Dissertation Case Studies

Case Study	Colonial History	Selection Criteria	Predominant Violence
Jaipur	Princely	Northern Case	Religious
Ajmer	Province	Northern Case	Caste and Tribal
Malabar	Province	Southern Case	Caste and Tribal
Travacore	Princely	Southern Case	Religious
Bastar	Princely	Deviant Case	Caste and Tribal
Hyderabad	Princely	Deviant Case	Caste and Tribal

ence revolved around caste. A wealth of interviews provided strong evidence that this was the case.

I then shifted to archival research. This really got to the crux of the issue—what made Jaipur experience more religious conflict? After spending weeks at the Jaipur City Palace archives, I uncovered a long colonial history of religious riots, most of them due to the discriminatory policies of the Hindu kings who ruled over the princely state. Muslims had been brutally repressed in the area, leading to long-term antagonisms between the two communities. In Ajmer, however, British administrators enforced discriminatory policies not toward Muslims, but lower castes and tribal groups. New land policies increased rural taxes and strengthened the power of local landlords. Therefore, there was little religious conflict in Ajmer, but a lot of violence which revolved around caste and tribal identities.

So in short, my quantitative analysis had confirmed that colonial legacies did matter, but my qualitative analysis finally gave me a plausible *mechanism*: contemporary patterns of ethnic conflict were caused by legacies of discriminatory policies dating from the colonial period. Combining the two methodologies together, it finally started to make sense.

My next step was to carefully select another paired historical comparison, this time from south India. As I had already worked in the north, I traveled southward to try to account for the enormous regional diversity of India. I found that in the small southern state of Kerala, the entire northern region (Malabar) had been under the control of the British, but the entire southern region (Travancore) had remained under the control of a Hindu dynasty. Better yet, the British themselves had called this political system an "accident" of history. I then embarked on the same fieldwork which I had carried out in Jaipur and Ajmer: a number of interviews to determine the contemporary pattern of conflict, and then extensive archival research to address underlying historical causes. Just as in the north, I found that British Malabar experienced more caste and tribal conflict whereas princely Travancore experienced more religious conflict. And again, archival research revealed long legacies of discrimination which continued to reverberate into the modern period. Because I had found similar patterns in north and south India, I felt reassured that I was onto something.

Finally, I selected two deviant cases: princely states with enormous amounts of caste and tribal conflict. These two cases, Hyderabad and Bastar (both located in eastern India), posed a major problem for the theory underlying my dissertation. And when I set foot in Hyderabad, I realized that I really had no explanation whatsoever to account for the deviant nature of these cases. Why should Hyderabad and Bastar experience such immense caste and tribal bloodshed, especially when no other princely states were similar? Interviews in the region were helpful in explaining contemporary violence, but I still couldn't understand why the two regions were so violent, which is exactly the opposite from the outcome that my theory would predict.

The major breakthrough came after spending a lot of time in regional archives. I discovered that Hyderabad had initiated the same land reforms which had occurred throughout British India, and was one of the few princely states to do so. Similarly, Bastar had come under heavy British intervention during the colonial period, much more than most other princely states of similar size. That's why these cases were *idiosyncratic*. In both cases, it also looked like the British were the culprits behind the scene. Therefore, I at the very least had an explanation for why these were deviant cases for my theory.

By the fall of 2011 I was ready to return home to America. I had spent a year in the field, had visited five archives, and conducted around 75 interviews. Only after a brief period of not thinking about political science at all was I then able to return to my project and begin to unpack what I had discovered, and how it all fit together.

Pitfalls and Payoffs of Multi-Method Research

Critics of multi-method projects often note that using multiple methodologies is quite different from using them *well*. This is a good point. Most people who use multiple methodologies do not become experts on two kinds of methodologies; rather, they learn basic competency in two areas. And it's an open question as to whether or not that is preferable to proficiency in one.

While I felt quite proficient at carrying out interviews in Hindi and poring over centuries-old archival documents, I felt somewhat less confident in my statistical analysis. What if I had omitted a critical variable and re-running the regression with said variable changed everything? What about endogeneity? Or robustness checks? I recall sitting down with a professor of history in India to explain my project, and after I detailed my statistical results she interrupted and said: "In history we don't really use statistics, but I'm guessing you are confident your results are correct?" Not without reservations.

I suppose that every scholar to some extent must grapple with this question. But those undertaking multi-method projects open themselves up to criticism from all fronts—the ethnographically-inclined are not pleased with only six months of fieldwork, the historically-inclined might like further archival work, while the statistically-inclined are similarly unimpressed with your rather basic model. So, what to do?

Considering the diversity of the political science discipline, I hardly think the answer is to completely ignore either qualitative or quantitative work in our projects. That is simply no longer a tenable position. So rather than accept competency in two methodologies as the basic criteria for doing a multi-method project, strive to do two well. Certainly this is easier said than done. I never intended to do any statistical analysis when I got to graduate school, so having a large-N chapter in my dissertation was a challenging but good step in the right direction. It is far from perfect, but getting further quantitative training under my belt is entirely within my control. There's no reason I cannot become as skilled in quantitative methods as I am with qualitative methods, and there's no reason a multi-method project can't make more than one of the methodologically diverse political science audiences (relatively) happy.

Part of the problem may also be that multi-method work and the multi-method movement are relatively new within political science. Therefore, departments still are in the process of adapting and ensuring that graduate students receive adequate training in how to carry out both qualitative and quantitative research. As multi-methods continue to gain popularity, more graduate students will be equipped with the tools to carry out these kinds of projects successfully.

The payoff of a smart multi-method dissertation is obvious: You have a variety of evidence that bolsters the strength of your central argument. My belief that colonialism matters in promoting patterns of ethnic conflict in India is borne out not merely by a large-N statistical analysis, but also by interviews, archival research, and lots of time spent in the field. By triangulating various techniques, I feel more confident in my argument than I would had I used merely one kind of methodology. This is not to disparage the work of those who do—but scholars are always left answering one of the questions which I stated earlier: Why didn't you go into the field or archives? Or, does your argument travel? It seems like the only way to offer sufficient answers to these questions (whether the questioner is a colleague, committee member, or potential employer) is to employ multi-methods.

I have an idea about my next research project, although I have not yet thought in detail about its methodology. But I do know that if I have enough data available, I'll employ a multi-method research design. And I'll continue to work at getting better at any kinds of methods I utilize. This is the best way to persuasively tackle the research problems that face us as political scientists.

Notes

[1] Most of the work in the comparative-historical tradition (see Mahoney and Rueschemeyer 2003) does not use statistical analysis. In fact, scholars like James Mahoney (2004) have argued that statistical analysis is poorly suited to comparative-historical research for a variety of reasons. Recent work by economists has sought to combine econometric analysis with historical research; see, for example, Acemoglu et al. 2001. However, the historical research in question is almost always limited to the brief use of secondary sources, and rarely entails in-depth archival or case study fieldwork. As Marcus Kreuzer notes about this kind of work, "the quality of quantitative research directly depends on the closeness of its *dialogue* with historical knowledge" (2010: 383, emphasis added). Too often, quantitative scholars use history only to grasp for and sketch out plausible causal mechanisms.

[2] Iyer (2010) also constructed a dataset of colonial India. However, she largely compared colonial and post-colonial maps, whereas I used actual geographical data and district reports from both the colonial and contemporary period to match districts.

References

Acemoglu, Daron, Simon Johnson, and James Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation" *American Economic Review* 91:5, 1369–1401.

Brass, Paul R. 1997. *Theft of an Idol: Text and Context in the Representation of Collective Violence*. Princeton: Princeton University Press.

Brass, Paul R. 2003. *The Production of Hindu-Muslim Violence in Contemporary India*. Seattle: Washington University Press.

George, Alexander and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Iyer, Lakshmi. 2010. "Direct versus Indirect Colonial Rule in India." *Review of Economics and Statistics* 92:4, 693–713.

Kreuzer, Marcus. 2010. "Historical Knowledge and Quantitative Analysis: The Case of the Origins of Proportional Representation" *American Political Science Review* 104:2, 369–392.

Laitin, David D. 1986. *Hegemony and Culture: Politics and Religious Change Among the Yoruba*. Chicago: University of Chicago Press.

Mahoney, James. 2004. "Comparative-Historical Methodology" *Annual Review of Sociology* 30, 81–101.

Mahoney, James and Dietrich Rueschemeyer. 2003. *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.

Mamdani, Mahmood. 1996. *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton: Princeton University Press.

Posner, Daniel. 2005. *Institutions and Ethnic Politics in Africa*. Cambridge: Cambridge University Press.

Varshney, Ashutosh. 2002. *Ethnic Conflict and Civic Life: Hindus and Muslims in India*. New Haven: Yale University Press.

Verghese, Ajay. 2012. *Colonialism and Patterns of Ethnic Conflict in Contemporary India*. Ph.D. Dissertation, The George Washington University.

Wilkinson, Steven I. 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. Cambridge: Cambridge University Press.

Young, Crawford. 1994. *The African State in Colonial Perspective*. New Haven: Yale University Press.