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 unreal-

istic "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