Comment and Debate

On Flanigan's and Zingale's "Alchemist's Gold"

Speculation or Specification? A Note on Flanigan and Zingale

J. MORGAN KOUSSER

IN A RECENT article in *Social Science History*, Flanigan and Zingale reviewed the old problem of inferring individual relationships from aggregate data, cast doubt on Goodman's ecological regression technique as a method of estimating such relationships, and contended that the specification analysis approach to the issue was misleading. Instead, they argued that an *ad hoc* procedure suggested by Shively in a 1974 article was superior to Goodman's point estimates because it avoids dubious assumptions and forces investigators to make their premises clear. Comparing results from analyses of state-level data on major party voting in the 1968 and 1972 presidential elections with figures from the Michigan surveys, they concluded, to their satisfaction at least, that the assumptions required for ecological regression were badly violated and that it was preferable in this and other cases to use the system that Shively, without making any grandiose claims for it, had originated (Flanigan and Zingale, 1985; Goodman, 1959; Hanushek *et al.*, 1974; Shively, 1974; Langbein and Lichtman, 1978).

While I am always in favor of increasing the number of useful methods available to historians and political scientists, I will argue in this note: first, that Flanigan's and Zingale's comments on the

J. Morgan Kousser is Professor of History and Social Science at the California Institute of Technology. This academic year he is at the Wilson Center, Old Smithsonian Building, Washington, DC 20560.
Social Science History 10:1 (Spring 1986). CCC 0145-5532/86/$1.50.

specification error strategy are misconceived; and second, that as a reanalysis of the example that they used will demonstrate, ecological regression, sensitively employed, can make better use of the available knowledge than does the tactic that they recommend. Consequently, it would be a mistake for political investigators who have access only to aggregate data to substitute Shively for Goodman.

The inescapable problem in inferring individual from aggregate relationships is that we simply do not have data below the state, county, or, in some cases, township level. The lowest units for which information has been collected are black boxes that we cannot see into, and nothing in either Shively's or Goodman's methods guarantees us certain knowledge of their contents. Even if we luckily found a few unclosed cartons—for instance, by discovering records of open ballots or party lists for scattered townships, counties, or constituencies—we would still be able to guess the allegiances of a larger voting universe only by assuming that the relationships were the same in the concealed as in the uncovered containers (Bourke and DeBats, 1978 and 1980; Elklit, 1985; Austin *et al.*, 1981: 94–95). But we do have hunches, which may be more or less well founded, about what has gone on inside the bins in which we receive the data, and we can observe various facts about the relations between one box and another. The question is, how should we seek to use our knowledge—in ways that are partially testable, or through means that are fundamentally indeterminate?

In a brilliant paper in 1974, Hanushek *et al.* questioned Robinson's famous 1950 derogation of aggregate data by showing that a multivariate regression model produced estimates of the individual relations between literacy and race or foreign birth that were much closer to the true individual relationships than were those based on the bivariate equations that Robinson had earlier employed. A more properly specified model—that is, one that took into account all the available information on the correlations not only between the dependent and independent variables but also between each of these and other possible intervening variables as well could not pry open the black boxes, but it could lead to better-warranted conjectures about their internal composition. If the coefficients for the independent variables of interest changed markedly and their standard errors shrank when more independent variables were added and if the proportions of variance explained by the models rose appreciably, then the investigator had at least some evidence that the expanded regressions had produced more valid results.

Flanigan and Zingale question this way of proceeding on the grounds that Hanushek *et al.* merely estimated a different model and that setting their results against the individual-level bivariate relationships was "comparing apples and oranges—or, better yet, cooked cabbage and raw carrots" (Flanigan and Zingale, 1985: 79). To the extent that investigators are interested solely in causal inferences, such as the question of what accounts for the observed relations between race and literacy, their point is well taken. But if observers are concerned as well with purely descriptive matters —for instance, as Robinson was, with discovering the percentages of illiterate whites and blacks, or native- and foreign-born persons who were able to read and write in 1940—then any method that improves the estimates of those numbers will be useful, whatever the number of variables on the right hand side of the equations.

To view the matter from another angle, what do we assume about the other traits of the persons involved when we cross-classify two characteristics of individuals? Suppose we have a table which shows how those who cast ballots for Humphrey or Nixon in 1968 voted in 1972. In principle, and perhaps in fact, we know a great deal more about the subjects, which may allow us to construct sub-tables showing the political behavior of, say, whites and blacks, men and women, managers, professionals, union and non-union manual workers, and the unemployed, etc. Some if not all of the choices of the groups captured in the sub-tables would differ from each other. But all these various relationships are, in effect, taken into account before being added together to form the basic two by two table for the whole population. Likewise, when we reduce specification errors by adding variables explicitly to the equation for aggregate data, we are just allowing for those sub-categorical relationships. After assessing the influence of the other traits, we may, if we wish, derive estimates of the entries of cross-classification tables that are of particular interest by operations very similar to collapsing sub-tables of individual-level data.

To anyone working with pre-survey era data, the question of how this "summation" should be carried out is of some moment. Since it has not, to my knowledge, ever been entirely explained in the literature of historical statistics, it may be useful to spell it out. Consider equation (1), which involves only three variables. The dependent variable, N72, is the percentage who voted for Nixon in 1972; the independent variables are the proportions for Nixon in 1968 (N68) and a "dummy" variable for region, which takes on the value 1 if the voter is southern, and 0 otherwise; the B's are the usual OLS parameters; and e is an error term.

$$(1)\quad N72 = B_0 + B_1 N68 + B_2 S + e.$$

We proceed by first constructing separate two by two tables for each region. Northern voters who favored Humphrey in 1968 and McGovern in 1972 are given by B_0, since N68 and S are both zero by definition for such persons. The estimate of southern Democratic loyalists is $B_0 + B_2$, because N68 is zero for them and B_2 taps the deviation of southern from northern electoral patterns. Northerners who stayed with Nixon both times are equivalent to $B_1 - B_0$, the Nixon effect less the yellow-dog Democrats, and southern Nixon voters are measured by $B_1 - B_0 + B_2$. To get the national two by two table, we then multiply each regional estimate by the proportion of total voters in the south and north, as appropriate, and add them. For instance, the total percentage of persevering Democrats is their southern percentage times the proportion of national votes that were cast in the south, plus the percentage of Yankees who remained Democratic times the proportion of total votes recorded in the north. After determining the "stayers," we can calculate the "movers" by subtracting the percentages of the party faithful from the overall percentages for McGovern and Nixon in 1972. The example gives concreteness to the analogy with collapsing categories in tables of individual or survey voting records to form more general tables.

When we have some continuous independent variables, for example, income, instead of only dichotomous "control" variables, such as in equation (2), where I stands for income, the proper procedure is to estimate the relationship and then determine the predicted percentage of the devoted and the switchers for persons of mean or perhaps median income.

$$(2)\quad N72 = B_0 + B_1 N68 + B_2 I + e.$$

By estimating the equation, we take into account the effect of I on the relationship between N68 and N72. To describe the most common or overall defection and stand-pat rates, those for the average person, therefore, we should estimate the rates for those with average incomes (naturally, for other purposes, it may be important to consider the rates for those of low or high incomes). For instance, B_0 added to B_2 times the average income yields the estimate of those who pressed the Democratic lever in both years.

The same basic principles apply to cases in which the models are more complex, as in equation (3), in which variables for the percentage of each unit which is Catholic (C) and urban (U) have been added to the Nixon in 1968, income, and regional variables.

$$(3) \quad N72 = B_0 + B_1 N68 + B_2 I + B_3 C + B_4 U + B_5 S + e.$$

Now, if we had information on each individual, C and U would be dummy variables, as S is. But if the only accessible data describe collections of voters, C and U become interval-level variables. If we focus on the stable and crossover rates for Humphrey and 1968 Nixon supporters, we should perform our regressions and then calculate those rates by setting the Catholic and urban proportions at their means and then multiplying by the values of the relevant coefficients. For example, northern Humphrey-McGovern support would be figured by adding B_0 to the products of three coefficients with means of the associated variables: B_2 times the average income, B_3 times the average percentage Catholic, and B_4 times the average proportion urban. Other cell entries and even more complicated models can be dealt with using similar principles.

This is not to say that more right hand side variables are always better than fewer, or that adding information will lead to very different estimates in every instance. Indeed, it will often be the case that many potential characteristics will have no substantively or statistically significant relationships with the dependent variable, after the traits that we are mainly curious about have been entered into the equations, or that the intercept and the coefficients for the most crucial indexes will be only marginally affected by introducing other independent variables. But it is surely proper to reduce the error variance and the confidence intervals of the parameters of interest if we have other information at hand rather than to neglect to do so.[1] And if the analyst is primarily concerned with determining the relations among only a subset of all the measured variables, then the procedures outlined for equations

(1) to (3) provide the proper parallel with the treatment of data on individuals. For those chiefly intent on description, as many historians are, Flanigan and Zingale's dismissal of the specification approach is simply mistaken.

Flanigan and Zingale endorse what they describe as a "conservative, controlled method for reducing the range of the individual relationship." Faced with Robinson's seemingly devastating dismissal of aggregate data, Duncan and Davis pointed out in 1953 that even the grouped information placed some limits on the possible individual level correlations. Every state had some Humphrey and some McGovern voters, for example, and it is easy to show that in such a situation, there must have been at least a few people in each state who opted for the Democrats in both elections (Duncan and Davis, 1953; Flanigan and Zingale, 1985: 82 for the quotation and 82–87 for more detail on the Duncan-Davis and Shively rationales). The trouble is that these constraints are usually so loose that we cannot even tell whether the individual correlations are negative or positive. To constrict these bounds, Shively advanced the proposition that observers might be willing to agree on some rules of thumb in analyzing the relationships between particular variables.

Taking data from state-level returns from the 1968 and 1972 presidential elections, Flanigan and Zingale showed that the Duncan-Davis method yields a range in the individual (phi or tau-beta) correlation of from +0.80 to −0.47, but that with "reasonable" assumptions, Shively's strategy constricts that to the interval between +0.76 and +0.48. The more restrictive the assumptions, the narrower the limits. Using survey data from 36 states, they demonstrated that the "true" phi is +0.60, while the estimate derived from bivariate ecological regression on state-level election returns is +0.74. Moreover, when they divided the survey data by states, they found considerable variation in phi coefficients from state to state, throwing doubt on an assumption crucial to ecological regression—that the within-area relationships differ only randomly from each other. Like Robinson, they seemingly invalidated a fairly widely used technique with a striking example, and they improved on his performance by offering a less drastic solution than the total abandonment of what is usually for historians the only available information.

Fortunately for those who have used ecological regression or who intend to in the future, Flanigan's and Zingale's arguments are unsound. There are four difficulties particular to their discussion of the example. First, answers to surveys do not, as Flanigan and Zingale recognize, precisely reflect what happens in the election booth.[2] The observed discrepancy between the results from home-administered questionnaires and voting returns probably partly reflects the fact that each source describes a related but different action. Second, since a national sample is not drawn to be representative of state electorates but only of the nation, survey data broken down by states do not provide a reliable standard against which to test the assumptions of ecological regression. Third, state-level data are so highly aggregated that it is hardly surprising that they yield imprecise estimates of personal voting patterns. Fourth, the proportions of loyalists and crossover voters estimated by bivariate ecological regression in this case are not really all that different from the survey results, even with grossly consolidated data. Relative to the "true" proportions found in the survey, the simplest Goodman-type procedure overestimates Nixon loyalism by five percent and underestimates Democratic defections from McGovern by 11%.[3]

There are also two larger problems with the Shively technique in general that greatly diminish its usefulness. First, its weak tests apply only to the units that are the most behaviorally homogeneous. Second, it does not make use of all the available information.

Consider Table 1, which shows hypothetical individual voting patterns (the "partials" or internal cell entries) grouped into units by area (the totals or "marginals"). If both elections are very close, as in Panel A, and if nearly all units have roughly equal marginal percentages, then a wide variety of assumptions about the internal cell entries will be consistent with the data, and the Shively bounds will be extremely broad (similarly, the bivariate ecological regression coefficients will have large standard errors). In this instance, from 4% to 100% of the Democrats and 0% to 100% of the Republicans might have voted the same way in both elections. Since the only test on the assumptions (e.g., that at least half of Humphrey's voters supported McGovern) that his tactic allows is whether the posited partials could mathematically add up to the marginals, there will be few clear tests in such a situation. Any attempt to impose arbitrary and untestable constraints

Table 1 Hypothetical examples of individual votes aggregated into areal units

	D1*	R1	Total	D1	R1	Total
Panel A: Cliffhanger elections mean wide bounds						
D2	51	0	51	2	49	51
R2	0	49	49	49	0	49
Tot.	51	49	100	51	49	100
Panel B: One landslide election narrows the estimates						
D2	51	39	90	41	49	90
R2	0	10	10	10	0	10
Tot.	51	49	100	51	49	100
Panel C: Extreme and divergent patterns also lead to wide bounds						
D2	85–90	0–5	90	5–10	0–5	10
R2	0–5	5–10	10	0–5	85–90	90
Tot.	90	10	100	10	90	100

*Entries are percentages. Categories are Democratic in first election (D1), Republican in first election (R1), Democratic in second election (D2), Republican in second election (R2), and totals for each election (tot.). In Panel C, entries are possible ranges of percentages.

will be controversial, for what appears reasonable to one person may be ludicrous to another.

If one or more elections is a landslide, or if most units are all close to unanimous *in the same direction* in at least one of the two elections, as in Panel B, then the limits will be much more binding and the tests much more useful. In this case, Democratic crossovers could have amounted to no more than 20%, and Republican defections to no more than 22%. Both the Shively tactic and bivariate regression will perform well in this circumstance.

A third idealized case, given in Panel C, in which the internal cell entries give the range of possible partials that could satisfy the marginals, seems to me the one most likely to occur in county or township level data. In this instance, the marginals for both units are entirely stable, and while one is overwhelmingly Democratic, the other is just as solidly Republican. Defection rates for each party could range from 0% to 50%. When there are only a few such counties, the analyst must choose between admitting that his assumptions are violated a certain number of times without being able to determine how many is too many, or publishing only a very broad estimate, which will not go very far towards answering the questions that he or she is presumably most interested in. In other words, in what seems likely to be a fairly typical situation confronting historical investigators, Shively's weak tests will either have to be jettisoned or the results will be vague and unsatisfying.

In contrast, ecological regression may be able to produce accurate estimates in this third case. Shively's method focuses entirely on the relationships between the partials and marginals for only two variables, ignoring whatever knowledge we may have of other variables. If the historian instead uses multiple regression, he or she can employ such data and can make use of much more powerful means of evaluating the validity of his or her models. The key to seeing this is to ask why the voting patterns in the units in Panel C of Table 1 diverge so much from each other. The answer is that the areas differ in other respects as well, which in many cases will be measurable. One is, say, nearly all white, one predominantly black, or one Protestant, the other Catholic, or one urban, the other rural, and so on. If we have sufficient data, we can take these extraneous variables into account in a multiple regression equation and obtain better estimates of the overall cohesion of the two parties over two elections than with bivariate regression and much more precise estimates than with the strategy of intuitive and largely untestable hunches that Flanigan and Zingale propose.

Let us focus on their example. The single most obvious fact about the comparison between the 1968 and 1972 elections is that in the former there were three significant candidates but in the latter only two. While George Wallace's disproportionate appeal in the south muddles estimates of the major party voting patterns from aggregate data, it suggests at the same time better models, that is, models including regional dummy variables.[4] Equation (1), discussed earlier, and equation (4), given below, represent two ways to operationalize regional effects. In equation (1), it was assumed that equations for the south and the non-south merely had different intercepts. Equation (4) may be used to test the hypothesis that within each region the slope of the relationships

Table 2 Ecological regression equations for 50 states and the District of Columbia, 1968 and 1972 U.S. presidential elections*

Model #	1	2	3
Constant	.0498 (0.86)	.0891 (1.98)	.0149 (0.37)
Nixon '68	.6541 (5.57)	.6278 (6.97)	.7796 (9.61)
South		−.1041 (−5.95)	.3723 (3.70)
S * N68			−.9905 (−4.78)
R^2	.388	.648	.763
adj. R^2	.375	.633	.748

*Entries are regression coefficients with t-statistics in parentheses. Variables are:
Nixon '68 = % of two-party vote for Nixon in 1968,
South = 1 for 11 ex-Confederate states, Kentucky, and Oklahoma, 0 otherwise,
S * N68 = Nixon 1968 percentage if state was in South, 0 otherwise,
R^2 = percentage of variance explained,
adj. R^2 = R^2 adjusted for degrees of freedom, dependent variable in all equations is % of two-party vote for Nixon in 1972.

between voting for Nixon in 1968 and in 1972 differed as well. Thus, the variable S times N68 in equation (4) reflects the interaction between Nixon support and the regional effect.

$$(4) \quad N72 = B_0 + B_1 N68 + B_2 S + B_3 SN68 + e.$$

The important advantage of such a model, which is equivalent to dividing the data into sets of the northern and southern states and running separate regressions on each group, is that it allows tests of the hypotheses.[5] If B_2 and B_3 are statistically significant and the proportion of variance explained, adjusted for the diminished degrees of freedom, rises, then we know that we have better estimates of the relationships between the marginal proportions. There is no counterpart in the Shively method.

What results do we get if we apply this method to the 1968 and 1972 state-level election data? Table 2 contains estimates of three ecological regression equations: a bivariate model and the more complex models given in equations (1) and (4). The increases from Model 1 to Model 3 in the variance explained and in the t-statistics for N68, and the fact that the t-statistics for both the intercept and the slope for the regional dummies are greater than 2.0, which corresponds to a level of significance greater than 0.05, show that the more complex equations fit the aggregate marginals better than Model 1 does.

They also yield estimates of the individual transition matrix

Table 3 The estimates implied by the Models in Table 2 compared to the "true" survey picture of the 1968–1972 elections*

Panel A: The survey estimates

	1972 Election		
	Nixon	McGovern	Total
1968 Election			
Nixon	90.2	9.8	100.0
Humphrey	32.4	67.6	100.0

Panel B: Estimates from Model 1 (bivariate)

	1972 Election		
	Nixon	McGovern	Total
1968 Election			
Nixon	95.0	5.0	100.0
Humphrey	29.6	70.4	100.0

Panel C: Estimates from Model 2 (regional intercept differences)

	1972 Election		
	Nixon	McGovern	Total
1968 Election			
Nixon	93.4	6.6	100.0
Humphrey	30.9	69.1	100.0

Panel D: Estimates from Model 3 (regional intercept and slope differences)

	1972 Election		
	Nixon	McGovern	Total
1968 Election			
Nixon	90.1	9.9	100.0
Humphrey	34.6	65.4	100.0

*Entries are percentages, which add to 100% across rows.

that are closer to the survey results. Table 3 compares the percentage estimates implied by the equations in Table 2 with the "true" percentages from the Michigan survey given in Flanigan and Zingale's article. While Model 1 overestimates the Nixon admirers by 4.8% and underestimates the McGovern haters by 2.8%, Model 2 and especially Model 3 are nearer the mark. Model 3 reproduces the 1972 behavior of 1968 Republican voters almost exactly and overestimates Democratic loyalty by only 2.2%. Flanigan's and Zingale's bounds, on the other hand, imply percentage estimates for the Nixonians of from 85.5% to 95.0%, and of Democrats disgruntled with McGovern of from 21.1% to 39.5%. While the near misses of Model 3 would, I believe, satisfy most historians, a range of Humphrey defectors of from one in five to two in five would cause considerable consternation.[6]

When Shively first introduced the modified Duncan-Davis bounds procedure, he claimed only that it "may be a useful supplement to ecological regression" (Shively, 1974: 71). Flanigan and Zingale were less reserved and sought to replace ecological regression entirely. But their critique of the specification approach to the problem was misleading, and a more subtle analysis of their example demonstrates why and how ecological regression can be made much more useful to historians than the technique that Flanigan and Zingale favor. Their article will have served a useful purpose, however, if it contributes to a more complete understanding of the problems of using aggregate data to estimate the behavior of individuals and if it leads historians to employ more sensitive and complex models.

NOTES

1 I freely admit that before the Hanushek *et al.* article, I did not see this solution as clearly as I have since, and that the estimates in Kousser (1974) might in some cases have been improved if I had understood it better.

2 Flanigan and Zingale (1985: 72) starkly contrast the "imaginative" use of survey data on the recall during the 1950s of their votes during the 1920s by Andersen (1979) and Butler and Stokes (1969), with the "uncritical" use of ecological regression by unnamed historians. They fail to note the devastating critiques of the recall data in those works by Niemi *et al.* (1980) and Katz *et al.* (1980) and perhaps underestimate the care with which at least some political historians have tried to test the validity of the methods that they used for their particular data sets. In fact, after a long honeymoon with surveys, political scientists have recently begun to see many virtues in more aggregated data. See, e.g., Kramer (1983); Ingelhart (1985).

3 The estimates below in Table 2, Model 1 are even closer to the survey results because I used all 50 states plus the District of Columbia, whereas they used only the aggregate data from 36 states. If the survey sample was

chosen to be nationally representative, then it broke the analogy for Flanigan and Zingale to restrict the aggregate estimates to 36 states.

4 It seems somewhat strange that Flanigan and Zingale themselves did not attempt to estimate a more complex ecological regression model with southern dummy variables, for they noted that they might have applied different assumptions to each region in a Shively-type analysis, and they have elsewhere remarked that "Perhaps dummy variable analysis offers the simplest and most promising expansion in the use of multiple regression introduced in recent years." Flanigan and Zingale (1985: 86; and 1981: 260).

5 For a fuller description of the dummy variable procedure and proof that equation (4) is equivalent to running separate regressions on subsets of variables, see Johnston (1984: 225–27).

6 Flanigan and Zingale state their bounds as individual tau-beta coefficients, which for two by two tables are equivalent to phi coefficients, but the range of individual percentages corresponding to those bounds can be recaptured for a table with given marginals by guessing at the percentages, calculating the phi's, and then repeating the process until one gets the correct phi's. It is unclear to me why anyone would be interested in the range of tau betas, which signify only proportionate error reduction, rather than in the percentages, which have a meaningful substantive interpretation. Since few historians are familiar with tau betas—those with memories as bad as mine have to rush off to old statistics texts to review the meaning and formulas for such statistics each time that they run into them—using them only adds mystification to an already sufficiently arcane practice.

REFERENCES

Andersen, K. (1979) The Creation of a Democratic Majority, 1928–1936. Chicago: University of Chicago Press.

Austin, E. A., J. M. Clubb, and M. W. Traugott (1981) "Aggregate Units of Analysis" in J. M. Clubb, W. H. Flanigan, and N. H. Zingale (eds.) Analyzing Electoral History: A Guide to the Study of American Voter Behavior. Beverly Hills, Ca.: Sage Publications.

Bourke, P. F. and D. A. DeBats (1979) "Identifiable Voting in Nineteenth Century America: Toward a Comparison of Britain and The United States Before the Secret Ballot." Perspectives in American History 11: 259–88.

——— (1980) "Individuals and Aggregates: A Note on Historical Data and Assumptions." Social Science History 4: 229–49.

Butler, D. and D. Stokes (1969) Political Change in Britain: Forces Shaping Electoral Choice. New York: St. Martin's Press.

Duncan, O. D. and B. Davis (1953) "An Alternative to Ecological Correlation." American Sociological Review 18: 665–66.

Elklit, J. (1985) "Nominal Record Linkage and the Study of Non-Secret Voting: A Danish Case." Journal of Interdisciplinary History 15: 419–43.

Flanigan, W. H. and N. H. Zingale (1985) "Alchemist's Gold: Inferring Individual Relationships from Aggregate Data." Social Science History 9: 71-92.

______ (1981) "Relationships Among Variables." in J. M. Clubb, W. H. Flanigan, and N. H. Zingale (eds.) Analyzing Electoral History: A Guide to the Study of American Voter Behavior. Beverly Hills, Ca.: Sage Publications.

Goodman, L. (1959) "Some Alternatives to Ecological Correlation." American Journal of Sociology 64: 610-25.

Hanushek, E. A., J. E. Jackson, and J. F. Kain (1974) "Model Specification, Use of Aggregate Data, and the Ecological Correlation Fallacy." Political Methodology 1: 87-106.

Ingelhart, R. (1985) "Aggregate Stability and Individual-Level Flux in Mass Belief Systems: The Level of Analysis Paradox." American Political Science Review 79: 62-78.

Johnston, J. (1984) Econometric Methods. 3rd ed., New York: McGraw-Hill.

Kousser, J. M. (1974) The Shaping of Southern Politics: Suffrage Restriction and the Establishment of the One-Party South, 1880-1910. New Haven, Ct.: Yale University Press.

Kramer, G. H. (1983) "The Ecological Fallacy Revisited: Aggregate- Versus Individual-Level Findings on Economics and Elections, and Sociotropic Voting." American Political Science Review 77: 92-111.

Langbein, L. I. and A. J. Lichtman (1978) Ecological Inference. Beverly Hills, Ca.: Sage Publications.

Robinson, W. S. (1950) "Ecological Correlation and the Behavior of Individuals." American Sociological Review 15: 351-57.

Shively, W. P. (1974) "Utilizing External Evidence in Cross-Level Inference." Political Methodology 1: 61-74.