DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

MODELING THE PAST: A NOTE ON THE SEARCH FOR PROPER FORM

Ivy Broder
The American University

Allan J. Lichtman
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 358

November 1980

ABSTRACT

As historians move beyond the computation of predictive measures to fashion regression models of past behavior they need to consider not only the estimation of regression parameters, but also the specification of a functional form for the model itself. The choice of a particular form for expressing the relationships among variables often has important implications for interpreting the historical issues under investigation. This paper presents methodology developed by J. B. Ramsey and other theorists for comparing distinct functional forms of a regression model according to criteria other than differences in the capacity to account for variation in the dependent variable. Ramsey's procedure directly tests the hypothesis of correctly specified functional form through scrutiny of the pattern of residuals that would be expected for properly specified models. The methodology can thus provide useful information on the choice of computing models especially when measures of explained variation are too close in value to indicate a clear preference for a given equation.

MODELING THE PAST: A NOTE ON THE SEARCH FOR PROPER FORM

Ivy Broder
The American University

Allan J. Lichtman
California Institute of Technology

INTRODUCTION

Increasingly historians and other social scientists are turning to regression analysis as a means for assessing the influence of variables on a particular behavior of interest. Unlike correlation coefficients, which only indicate the strength and direction of the linear association between two variables, regression allows one to measure how the values of one variable (called the dependent variable) change in response to variations in the values of a set of possible explanatory variables (called independent or explanatory variables). Estimates of the coefficients for independent variables included in a linear regression model indicate the change in the behavior of the dependent variable for a unit change in any one of the independent variables (holding all other explanatory variables constant). Regression methods are thus appropriate whenever an investigator seeks a predictive equation for a form of behavior or attempts to separate and weigh the presumed causes of phenomena such as fertility, voting, social mobility, and economic growth.

The challenge of regression analysis, however, is not simply to estimate the parameters from a pre-selected regression equation,

but also to choose among alternative forms of the equation itself. Every multiple regression equation represents a distinct model of how to explain variation in the behavior of the dependent variable. Well specified regression models both include as many as possible of the actual determinants of behavior and portray the correct functional form of the relationship between the dependent and explanatory variables. Departures from correct specification may yield both unreliable predictors and misleading ascriptions of casual influence.

Historians are most familiar with misspecifications arising from the omission of independent variables related both to the behavior under investigation and to one or more of the included independent variables. Equations that omit relevant variables are not only incomplete, but yield misleading (biased) estimates of the regression coefficients for included variables that are correlated with any of the excluded variables. The magnitude and direction of this "specification bias" depends upon the relationship beween the excluded variable and the dependent variable and between the excluded variable and the included independent variable. If, for example, the behavior of a dependent variable, Y, is a linear function of two correlated independent variables, X_1 and X_2, the exclusion of X_2 would produce specification bias in the parameter estimate of X_1 equal to

$$b_{YX_2.X_1} \cdot b_{X_1X_2}$$

which is the regression coefficient of X_1 regressed on X_2 multiplied by the true regression coefficient of X_2 on Y (holding X_1 constant).

Less familiar to historians, but also important, are misspecifications that result not from the omission of variables but from using an incorrect functional form of the relationship between dependent and independent variables. The specification of functional form is particularly relevant for historians whose options to test new variables is often constrained by the availability of data, but can choose various functional forms for expressing relationships among measurable variables. The most common functional form of a multiple regression model is linear and additive, represented by the equation:

$$Y = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n + u$$

where Y is the dependent variable, each X_i an independent variable, a is a constant term, each b_i is a regression coefficient, and u is a random disturbance term. Competing forms of interest to historians include multiplicative power functions, in which independent variables are multiplied together and regression coefficients are powers (e.g., $Y = aX_1^{b_1}X_2^{b_2} \ldots X_n^{b_n}u$); nonlinear models in which polynomial functions of the original variables (e.g., X_1^2, X_1^3, etc.) are added to the regression equation; and interactive models in which the products of particular variables (e.g., X_1X_2, X_1X_3, $X_1X_2X_3$, etc.) are added to the regression equation.

The choice of a given functional form may entail significant differences in the interpretation of past behavior. Soares and Hamblin, for instance, found that a multiplicative, power function of the form portrayed above better explained social and economic influences on voting for radical left candidate Salvador Allende in

Chilean elections of the 1950s than did a linear, additive model. Noting that this multiplicative model of voting behavior was identical in form to the equation used by psychologists for gauging responses to involuntary stimuli, Soares and Hamblin argued that the Allende vote arose from feelings of frustration with the status quo and that the intensity of such frustration increased as a power function of the social and economic stimuli acting upon the voter.[1]

Applying the work of Soares and Hamblin to American politics, Burnham and Sprague found that for Pennsylvania counties in the presidential election of 1968 an additive, linear model best represented the influence of social and economic variables on voting for major party candidates, but a multiplicative model best represented voting for third-party candidate George Wallace. The authors interpreted this finding to mean that the vote for Wallace reflected alienation from the broad consensus of values forged by the two major parties and was an "'act of aggressive hostility' against the leaders, parties, and policies associated with the existing political order." Burnham and Sprague further surmised that linear, additive models should best portray American voter behavior during eras of stable politics and that multiplicative models should characterize voter behavior during periods of realigning change.[2]

The most familiar method used by social scientists for evaluating different functional forms of a regression equation is to compare the coefficients of multiple determination (R^2) that reveal the proportion of the variation in the behavior of the dependent variables that can be explained by a particular regression model.

When competing models include different numbers of independent variables, the proper measure is $\bar{R}^2$, the coefficient of multiple determination adjusted by the number of variables and the number of observations in the regression model.[3] Both Soares and Hamblin and Burnham and Sprague used the examination of coefficients of determination for choosing between additive, linear and multiplicative models.

This paper further explores the question of how to decide which of several distinct functional forms of a regression model best fits the historian's data. The choice of functional form, of course, depends not only on the results of a posteriori analysis, but also on the prior knowledge that an investigator brings to a study and upon the simplicity and elegance of alternative models. Nonetheless, a researcher typically may be uncertain of precisely how to specify a regression model and may use the data not simply to estimate the parameters of a preselected model, but also to help choose an appropriate specification. The goal is not generally to find the single "correctly specified" model, but rather to choose the best model from a set of competing alternatives.

We will present methodology developed by J. B. Ramsey and other statisticians for comparing distinct functional forms of a regression model according to criteria other than differences in the capacity to account for variation in the dependent variables (as measured by coefficients of multiple determination, R^2 and $\bar{R}^2$). Ramsey's methodology can be used in conjunction with a scrutiny of R^2 or $\bar{R}^2$ to help historians choose among competing regression models.

The examination of coefficients of determination is of limited value in choosing among various specifications, since it gives only indirect indications of proper form through the predictive power of the equation. Ramsey's procedure, in contrast, directly tests the hypothesis of correct specification through scrutiny of the patterns of residuals that would be expected for properly specified equations. The test can thus provide useful information on the choice of models especially when coefficients of determination are too close in value to indicate a clear preference among competing equations. Ramsey found, for example, that the procedure was able to discriminate among various specifications of production functions all of which had R^2s of .9 or greater, indicating almost perfect fit.[4]

METHODOLOGY

In several articles, Ramsey has presented specification tests designed to distinguish between competing regression models. The test used here is a modified version of Ramsey's original test labelled RESET, which probes for specification bias arising from either omitted variables, incorrect functional form, or simultaneous equation problems.[5] The test is thus able to discriminate between different models that set forth the same set of original variables in different functional form.

The variant of RESET used here is a computationally simple test procedure based on the insight that under conditions of perfect specification a vector of residuals (the difference between the actual values of the dependent variable and the values estimated from the regression equation) for all observations ($\tilde{u}$) has an expected value (the mean value under repeated sampling) equal to a vector of zeros termed the null vector: $E(\tilde{u}) = \emptyset$. Each element of the residual vector is expected to converge to a mean of zero under repeated sampling because, given perfect specification, positive and negative deviations between actual and estimated values of the dependent variable are equally likely and will cancel one another. If, however, specification bias arises from omitted variables, incorrect functional form or simultaneous equation problems, errors will no longer be expected to cancel one another and elements of the residual vector should be unequal to zero. The entire vector will then have an expected value unequal to the null vector: $E(\tilde{u}) \neq \emptyset$.

A test of the null hypothesis that $E(\tilde{u}) = \emptyset$ against the alternative hypothesis that $E(\tilde{u}) \neq \emptyset$ therefore becomes a test of the null hypothesis of correct specification against the alternative hypothesis of specification error. Since $E(\tilde{u})$ is based on repeated sampling and thus is unobservable to the investigator, its elements must be estimated from information that is available from the regression model being tested.[6] RESET derives these estimates from a simple two-stage regression procedure and relies on the F-statistic to test the null hypothesis that their value is zero and the regression equation is therefore properly specified. This F-test for the expected value of the residual vector should not be confused with other F-tests that typically are used to test the statistical significance of a regression equation as a whole (which is equivalent to testing the hypothesis that $R^2 = 0$) or the significance of

individual regression coefficients (which is equivalent to testing the hypothesis that $b_i = 0$).

To perform the RESET test, the investigator specifies the regression model,[7] estimates its parameters and then calculates estimated values of the dependent variable for all observations ($\hat{Y}_i$).[8] The second, third, and fourth powers of these estimated values ($\hat{Y}_i^2$, $\hat{Y}_i^3$, $\hat{Y}_i^4$) are computed and added to the original regression equation as three additional independent variables. The parameters of this augmented model are estimated from the original data. The parameter estimates of the regression coefficients (b_1^*, b_2^*, b_3^*) for the three additional variables $\hat{Y}_i^2$, $\hat{Y}_i^3$, and $\hat{Y}_i^4$ are estimates of the elements of the vector $E(\tilde{u})$. Thus a test of the null hypothesis that these coefficients all have values of zero ($b_1^* = b_2^* = b_3^* = 0$) against the alternative hypothesis that at least one of the coefficients is unequal to zero is a test of the null hypothesis that $E(\tilde{u}) = \emptyset$ against the alternative hypothesis that $E(\tilde{u}) \neq 0$. By simple inference, the test on the three regression coefficients for the powers of the fitted values becomes a test of the null hypothesis of correct specification against the alternative of misspecification.

After the second-stage regression is carried out, an F-test is used to test the hypothesis that the coefficients of the powers of $\hat{Y}_i$ are zero. The degrees of freedom for the F-statistics are 3 in the numerator and n-k-4 in the denominator. This is equivalent to an analysis of variance of the incremental predictive power of the three new "variables." The F value is obtained by the ratio:

$$\frac{\Delta R^2/3}{1-R^2/n-k-4}$$

where ΔR^2 is the additional to R^2 of the three powers of $\hat{Y}_i$ and R^2 is the multiple coefficient of determination for the entire (new) equation.

The resulting F-statistics, of course, can be used with a single equation to test the null hypothesis of correct specification at a given level of statistical significance. Values of an F-statistic greater than the critical value established for 3 and n-k-4 degrees of freedom would lead to rejection of the null hypothesis. Given, however, that historical models will virtually always be misspecified, both for lack of adequate data and sufficiently precise theory, the result of testing single equations for specification bias would largely become a function of sample size. As sample size increases—a usually desireable property since it shrinks the confidence bands of parameter estimates and permits more extensive analysis—historians may more frequently reject the null hypothesis of correct specification in favor of the alternative hypothesis of misspecification. Usually null hypotheses are framed so that they represent random associations among variables whereas investigators are seeking to find systematic relationships. For Ramsey's test, in contrast, the null hypothesis represents precisely the goal that the historian is seeking to achieve.

Ramsey's procedure, we believe, is most useful for comparing several different forms of regression equations, calculated for the same number of observations. In this case, the F-statistic becomes a criterion for choosing the best model from a set of competing

possibilities. The higher the value of F (at a fixed sample size) for a given equation the less likely it is to approach the correctly specified form. Thus the test would favor selection of whichever model has the lowest score on the F-statistic.

ILLUSTRATION

To illustrate the methodology for investigating the specification of regression models, we examined several models for each of two measures of voting for president in the presidential election of 1928: the Democratic candidate Al Smith's percentage of the vote cast for president (% DEM 1928) and the difference between this percentage and the percentage of the vote gained by John W. Davis, the Democratic nominee for president in 1924 (% DEM 1928 minus % DEM 1924). Drawing on a data base of all 2058 counties outside the formerly Confederate South, we selected five independent variables: the percentage of Catholics residing in a county, the percentage of urbanites, the percentage of those foreign born or with a foreign born parent, the percentage of home owners, and the percentage of families owning radios.[9]

Our objective was to examine the patterns of R^2 and F statistics to see which of the models most closely approached the properly specified form of the regression equation. We did not attempt to test the null hypothesis of perfect specification at a given level of significance, since, as Ramsey and Zarembka note, even researchers using contemporary rather than historical data "are always aware of the fact that their models are likely to be misspecified in one way or another."[10]

For each of our two dependent variables we constructed first the following two models with the same numbers of independent variables: an additive linear model of the form: $Y = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n + u$ and a multiplicative, power function of the form used by Soares and Hamblin, and Burnham and Sprague: $Y = aX_1^{b_1}X_2^{b_2} \ldots X_n^{b_n}u$.. As did these other researchers, we transformed all variables of the multiplicative model into logarithms so that it could be set forth in additive, linear form and its parameters estimated through ordinary least squares: $\log Y = \log a + b_1\log X_1 + b_2\log X_2 + \ldots b_n\log X_n + \log u$. The multiplication of a logarithm by a constant is equivalent to raising the numerical counterpart of the logarithm to the power of a constant, and the addition of logarithms is equivalent to multiplying their numerical counterparts.

In addition to computing R^2 for each regression equation, we compared the linear and logarithmic models using the Ramsey specification test RESET. As described previously, the powers (second, third, and fourth) of the fitted values of the dependent variables were calculated for each model and then added to the original regression equations, increasing the number of independent variables in the second round of regressions to K + 3. Coefficients of determination for the initial equations as well as joint F-statistics for the contribution of the three new variables are reported in Table 1. The higher the F-statistic, the more likely is the departure from proper specification.

TABLE 1

COMPARISON OF LINEAR AND MULTIPLICATIVE REGRESSION MODELS

Dependent Variable	Model Form: Linear R^2	Linear F-Statistic	Multiplicative R^2	Multiplicative F-Statistic
% DEM 1928	.21	7.5	.09	24.2
% DEM 1928 minus % DEM 1924	.70	12.5	.51	143.6

For both dependent variables, R^2 and the F-statistic give the identical message that the linear model is preferable to its logarithmic competitor. For Al Smith's percentage of the 1928 vote the linear model accounts for 21 percent of the county-to-county variance, compared to 9 percent for the logarithmic model. The linear model has an F-statistic of 7.5 compared to 24.2 for the logarithmic alternative. For the difference between the vote for Smith in 1928 and the vote for Democratic candidate John W. Davis in 1924, the linear model accounts for a very substantial 70 percent of the variance, compared to 51 percent for the logarithmic form. The linear model has an F of 12.5 compared to 143.6 for the logarithmic model.[11]

The choice of the additive, linear model over the multiplicative alternative has important implications for the interpretation of voting behavior in 1928. If Burnham and Sprague are correct in suggesting that periods of realignment are characterized by the multiplicative influence of social and economic variables on voter choice, then this finding would further bolster the contention that 1928 was not part of the realigning process that ended an era of Republican hegemony and led to the formation of the Roosevelt coalition. The superiority of an additive, linear model over the multiplicative competitor also indicates that the presidential election of 1928 did not polarize the electorate into two distinct groups of voters characterized by mutually reinforcing scores on the variables included in the regression analysis. This finding challenges the supposition of many historians that the encounter between Hoover and Smith tapped a fault line in American society that

neatly divided the polity into two sets of antogonists: rural, Protestant, native-stock, middle-class, traditionalists vs. urban, Catholic, foreign-stock, lower-class, cosmopolitans.[12]

Although the coalitions forged by Al Smith and Herber Hoover did not reflect a neat duality in the American electorate—the minds sets of the old and the new America, a linear, additive model of voter behavior may not fully capture the distinct and intersecting paths of the different themes that run through the politics of 1928. To further explore specifications of the vote for president in 1928, we constructed two additional models: an interactive model that added to the independent variables all first order interaction terms (the product of every pair of explanatory variables) and a nonlinear model that added the squares of each explanatory variable. The interactive model suggests a focus on dependencies among particular explanatory variables: rural Protestants, for example, may have been more hostile to the Catholic candidate Al Smith than their counterparts in the city. The nonlinear model discloses "contextual" effects arising from the concentration of various kinds of voters in the counties being studied: urbanism, for example, may begin to have an important influence on voter behavior only after counties reach a certain proportion of urban residents. Table 2 reports $\bar{R}^2$ (coefficients of determination adjusted for degrees of freedom since each model has a different number of independent variables)[13] and F-statistics for the linear, interactive and nonlinear models, respectively.

The results of analysis suggest the value of exploring nonlinear and interactive relationships, but indicate no clear choice

TABLE 2

COMPARISON OF LINEAR, INTERACTIVE, AND NONLINEAR REGRESSION MODELS

Dependent Variable	Linear R^2	Linear F-Statistic	Interactive R^2	Interactive F-Statistic	Nonlinear R^2	Nonlinear F-Statistic
% DEM 1928	.21	7.5	.25	4.3	.24	4.7
% DEM 1928 minus % DEM 1924	.70	12.5	.71	11.8	.70	5.3

of a preferred model. As Table 2 reveals, the values of adjusted coefficients of determination cluster fairly closely together for the linear, nonlinear, and interactive models. For Al Smith's percentage of the 1928 vote, the interactive model has an $\bar{R}^2$ of .25 compared to .24 for the nonlinear and .21 for the linear equation. For the difference beween the vote for Smith and the vote for Davis in 1924, the interactive model is again in first place with an $\bar{R}^2$ of .71, closely followed by the linear and nonlinear models with identical $\bar{R}^2$s of .70. The F-statistics indicate that for Smith's percentage of the 1928 vote either the interactive or nonlinear equation with values of 4.3 and 4.7 respectively is preferable to the linear model with a value of 7.5. For the difference between voting for Smith and voting for Davis four years earlier, the F-test suggests that the nonlinear model with an F-statistic of 5.3 is preferable to either the interactive model with a value of 11.8 or the linear model with a value of 12.5.

These ambiguous results suggest that, given our data, a model that incorporates linear, nonlinear, and interactive effects might be most appropriate. To test this supposition we constructed a regression equation that included the five initial independent variables, their squares, and all of their first order interactions. Although Table 3 reveals that the combination of all three types of effects did not add appreciably to the ability to predict variation in the dependent variables, the F-statistics for the combined model were lower than for any of the previously constructed alternatives. Thus Ramsey's test for appropriate specification points to the combined

TABLE 3

COMPARISON OF LINEAR, INTERACTIVE, NONLINEAR, AND FIXED REGRESSION MODELS

Dependent Variable	Model Form: Linear R^2	Linear F-Statistic	Interactive R^2	Interactive F-Statistic	Nonlinear R^2	Nonlinear F-Statistic	Mixed R^2	Mixed F-Statistic
% DEM 1928	.21	7.5	.25	4.3	.24	4.7	.26	2.6
% DEM 1928 minus % DEM 1924	.70	12.5	.71	11.8	.70	5.3	.71	3.7

model as the best among our several regression models, even though examination of $\overline{R}^2$ failed to distinguish among them.

The overall superiority of the mixed model suggests that even when investigators of voting behavior and other historical phenomena find a linear, additive equation preferable to a multiplicative, power function they cannot afford to overlook the possibility of nonlinear relationships within their data or interaction among particular independent variables. Lichtman, for example, found that analysis of a combined linear, interactive, and nonlinear model did not reveal interactions between religion, foreign-stock heritage and urban-rural residence that would be expected from the thesis that the election divided the nation into two distinct cultures. But he did discover interaction between religion and foreign-stock heritage that reflected both the assimilation of Catholics with an extended lineage in the United States and the anti-Catholicism of the Protestant immigrants who continued to enter the country in large numbers during the late nineteenth and early twentieth centuries.[14] A full discussion of the implications of the mixed model for interpreting the presidential election of 1928, however, is beyond the scope of this article.[15]

CONCLUSION

As historians move beyond the use of strictly predictive measures to consider how variables influence one another, a sensitivity to alternative specifications becomes increasingly important. Not only may variables be added to or subtracted from regression equations, but choices can be made from among various functional forms for expressing the relationships between dependent and independent variables. The choice of an appropriate regression model, we have shown, is not a heuristic exercise, but can have significant implications for the interpretation of historical issues.

We have set forth methodology borrowed from Ramsey and other theorists for directly testing the specification of regression equations. The simple test procedure is to specify an equation; estimate the values of the dependent variable; take the second, third, and fourth powers of these estimates; add the three power functions to the original equation; estimate the parameters of the augmented equation; perform an F-test of the hypothesis that the regression coefficients for the three additional variables are zero (equivalent to the analysis of variance test that the three new variables add nothing to the predictive power of the original equation). The higher the values of the F-statistic the more suspect is the specifiction being tested.

The quest for an appropriate specification, however, should not devolve into an indiscriminate testing of a procession of mechanically generated models. Before using his data to compare specifications, the historian should have good reasons for the choice of competing models, grounded in relevant theory as well as an expert´s knowledge of the historical problem being addressed. The more indiscriminate the search, the more likely are specious results generated by random processes.

The results of a posteriori statistical testing do not automatically determine of the choice among possible specifications.

First, the results of testing only apply to the particular functional forms being examined by the investigator and not to any other variants. Second, the final choice of a regression model also depends on the historians a priori knowledge as well as the simplicity and elegance of the models themselves.[16] The results of statistical testing, for instance, may narrowly favor a multiplicative model over an additive, linear model of the social and economic influences on voting in a particular election. But the historian may still use the linear model to analyze sources of the vote because it better fits his understanding of the dynamics of the election and is easier to present and explain.

FOOTNOTES

1. Glancio Soares and Robert T. Hamblin, "Socio-Economic Variables and Voting for the Radical Left: Chile, 1952." American Political Science Review 61 (1967), pp. 1062-1063.

2. Walter Dean Burnham and John Sprague, "Additive and Multiplicative Models of the Voting Universe: The Case of Pennsylvania, 1960-68." American Political Science Review 64 (1970), p. 486.

3. $\bar{R}^2$ is the adjusted coefficient of multiple determination. It is used to compare models with the same dependent variable but with a different number of independent variables. Looking at unadjusted R^2 to determine which model is better will not necessarily be helpful because R^2 almost always increases when more variables appear in the equation (and will never decrease). Using the formula

$$\bar{R}^2 = 1 - \left(\frac{N-1}{N-K} \times (1-R^2)\right)$$

where N is the number of observations, K is the number of independent variables, adjusts R^2 for the number of independent variables and is a better statistic for comparative statements.

4. J. B. Ramsey and P. Zarembka, "Specification Error Tests and Alternative Forms of the Aggregate Production Function." Journal of the American Statistical Association 66 (1971), pp. 471-477.

5. J. B. Ramsey, "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis." Journal of the Royal Statistical Society, Series B 31 (1969), part 2, pp. 350-371; J. B. Ramsey and P. Schmidt, "Some Further Results on the use of OLS and BLUS Residuals in Specification Error Tests." Journal of the American Statistical Association 71 (1976), pp. 389-390. Simultaneity problems occur when there are mutually dependent causal relationships between dependent and independent variables.

6. Technically, the vector of expected values $E(\tilde{u})$ that is estimated in this test is not derived from the ordinary least squares estimate of residuals, but from an alternative vector of residuals developed by Theil and termed the BLUS vector (for Best Linear Unbiased Scalar covariance matrix). The BLUS estimates of the residual terms are preferred to the OLS estimates because the former are independently distributed whereas the latter are not. See H. Theil "The Analysis of Disturbances in Regression Analysis," Journal of the American Statistical Association 60 (1965), 1067-1079.

7. For purposes of computation various functional forms of regression equations can be converted into linear, additive form through transformations of variables.

8. Aside from problems of specification it is assumed that the standard assumption of regression analysis hold, including the normal distribution of error terms with no serial correlation; homoskedasticity, and no exact linear relationship between any

two or more independent variables.

9. Presidential election returns were obtained from the Inter-University Consortium for Political and Social Research, Ann Arbor, Michigan; the percentage of urbanites and foreign-stock Americans from U.S. Department of Commerce, Bureau of the Census, Fifteenth Census of the United States, Population, 1930, Vol. III, Table 13; the percentage of home owners and radio owners from Fifteenth Census of the United States, Families, 1930, Vol. VI, Tables 19-20; and the percentage of Catholics from U.S. Department of Commerce, Bureau of the Census, Religious Bodies, 1926, Vol. I, Table 32.

10. Ramsey and Zarembka, "Specification Error Tests," p. 472.

11. R^2s for the multiplicative model continue to be substantially smaller than those for the additive, linear model when computed by the alternative method of estimating from the logarithmic model, inserting these parameter estimates in the original multiplicative model with untransformed variables, and using the resulting equation to estimate the untransformed dependent variable.

12. See Allan J. Lictbman, Prejudice and the Old Politics: The Presidential Election of 1928. (University of North Carolina Press: Chapel Hill, 1979), pp. 16-18, 50, 231.

13. Given 2058 counties, the value of R^2 and $\bar{R}^2$ for these equations actually differ only in the third decimal place.

14. Lichtman, Prejudice and the Old Politics, pp. 50-51.

15. For a detailed analysis of a mixed linear, interactive, and nonlinear model of voter behavior in 1928 see Licthman, Prejudice and the Old Politics, pp. 257-264.

16. For a sophisticated discussion of a Bayesian model for specification searches see Edward E. Leamer, Specification Searches: Ad Hoc Inference with Nonexperimental Data, (New York: John Wiley and Sons, 1978).