

Determining Who Voted in Historical Elections: An Aggregated Logit Approach

Jac C. Heckelman

Wake Forest University

The ecological fallacy literature suggests aggregate data cannot be used for microlevel inference. Building upon an aggregated logit model developed by Kelejian (1995), I am able to show that the estimated coefficients from aggregate data regressions are directly related to the true underlying microcoefficients, so meaningful interpretations can still be made. It is also suggested that in the cases where aggregation bias can be rejected, the microcoefficients can be directly estimated. An application of the model is shown using state-level data for historical elections in which survey data are unavailable. © 1997

Academic Press

Researchers are often interested in estimating how individuals behave in different scenarios. Unfortunately, there are many cases in which individual-level data, needed to observe how specific individuals react based upon their personal socioeconomic characteristics, are not available. Subsequently, these scholars are forced to rely upon aggregated data. In those situations where the original question was whether or not a specific action occurred, the aggregate dependent variable measures the percentages of those cases in which the action did occur. Logistic transformations of the aggregated dependent variable have been gaining popularity to control for biased estimates generated from standard ordinary least-squares regressions. Recent examples include studies of the percentage of people who registered (Knack, 1993) or voted (Filer et al., 1991).

There is, however, an inherent danger in running this sort of regression. Direct aggregation of the underlying logit model is not possible when individuals are heterogeneous. Furthermore, attempting to infer individual characteristics from

I acknowledge helpful clarifications from Harry Kelejian regarding some of the issues in this paper and two anonymous referees for suggestions that improved the original draft of this paper. Remaining errors are my own responsibility.

Address correspondence and reprint requests to Jac C. Heckelman, Wake Forest University, P.O. Box 7505, Winston-Salem, NC 27109. E-mail: heckeljc@wfu.edu.

aggregate data is problematic in its own accord. This is known as the ecological fallacy.

Robinson (1950) was the first to note this problem. He argued that aggregate data are descriptive of groups and not of individuals, and are therefore unreliable for predicting individual behavior. Statistical associations derived from aggregated data could differ in magnitude and sign from the underlying individual correlations. "Such reversals in the direction of the coefficients are statistically possible but contrary to the known effect found in the survey data" (Kim et al., 1975, p. 121).

The ecological fallacy literature suggests that macromodels cannot explain microlevel behavior (Klein et al., 1991; Huston, 1991; Matsusaka and Palda, 1993) but aggregate data are not completely useless for this task. If the original microcoefficients can be recovered, meaningful interpretations can still be made. The approach taken in this paper draws heavily upon an aggregation procedure developed by Kelejian (1995). His model identifies the proper macrolevel equivalent of a microlevel logit model.

The purpose of this study is to estimate the original microcoefficients from the macromodel to lend insights into microlevel behavior. Kelejian's macromodel results in coefficients which differ from the true microcoefficients we wish to estimate. Therefore, the macromodel cannot proxy for the true micromodel nor are the unit-level variables sufficient proxy for the individual-level variables. It will be shown, however, that the proper macromodel has parameters which directly relate to the original microcoefficients. I will also discuss the conditions under which the parameters of the micromodel can be estimated from the estimated macrocoefficients. Although the application here will be to voter turnout, it should become evident that the procedure is useful in many situations where only percentage data are available.

The empirical question posed in this study relates to historical election analysis. Election reform during the Gilded Age had an important impact upon the electoral arena. Poll taxes and literacy tests, which were in limited use prior to this time, gained widespread acceptance throughout the South to limit the participation of blacks and poor whites (Key, 1949; Kousser, 1974). Female suffrage was gaining in popularity, primarily due to the success of suffrage states in the West (Beeton, 1986). Finally, secret ballots were adopted by many states to curb rampant fraud and corruption (Rusk, 1974). Empirical evidence suggests that each of these laws reduced state turnout rates (Heckelman, 1995). In this paper I attempt to discover which individuals were most likely to vote in gubernatorial elections during this turbulent time period.

The rest of the paper is structured as follows. Kelejian's aggregation procedure is outlined in Section I. Data descriptions and macroresults are given in Section II. The conditions under which it is possible to estimate the original microcoefficients are developed in Section III along with an interpretation of the estimates. In the final section, I present a summary and comparison of the results.

I. AGGREGATION PROCEDURE

The vote choice can be represented by

$$\begin{aligned} Y_{jit} &= 1, & \text{if the } j\text{th person in state } i \text{ at election } t \text{ votes,} \\ Y_{jit} &= 0, & \text{otherwise.} \end{aligned}$$

Let P_{jit} represent shorthand notation for $\text{Prob}(Y_{jit} = 1 | X_{jit})$, where X_{jit} is a vector of personal and state characteristics for the j th person in state i at election t . Assuming the probability that an individual will vote follows a logistic distribution, then

$$P_{jit} = F(X_{jit}\beta) = \frac{e^{\alpha + X_{jit}\beta}}{1 + e^{\alpha + X_{jit}\beta}} \quad (1)$$

which is equivalent to

$$\ln\left(\frac{P_{jit}}{1 - P_{jit}}\right) = \alpha + X_{jit}\beta. \quad (2)$$

At this point, a logit model could be estimated if survey data were available. However, in those cases where only aggregated data are available, a model needs to be derived which will correspond to the aggregated data. It should be obvious that direct aggregation of (2) is not possible; thus another approach is necessitated.

Note $E(Y_{jit} | X_{jit}) = P(Y_{jit} = 1 | X_{jit})$ and define

$$\begin{aligned} \bar{Y}_{it} &\equiv \frac{1}{n_{it}} \sum_{j=1}^{n_{it}} Y_{jit} \\ \bar{X}_{it} &\equiv \frac{1}{n_{it}} \sum_{j=1}^{n_{it}} X_{jit}, \end{aligned} \quad (3)$$

where n_{it} is the number of eligible voters in state i at election t .

Under reasonable restrictions,¹ Kelejian (1995) demonstrates

$$\ln\left(\frac{\bar{Y}_{it}}{1 - \bar{Y}_{it}}\right) = \alpha + \bar{X}_{it}\beta + g(\bar{X}_{it}\beta) + \mu_{it} + O_p(n_{it}^{-1/2}), \quad (4)$$

where μ_{it} is a random error term such that

$$\begin{aligned} E[\mu_{it}] &= 0 \\ \text{Var}[\mu_{it}] &= \sigma_{it}^2. \end{aligned} \quad (5)$$

¹ For specific details on these restrictions, see Kelejian (1995, especially pp. 244–246).

Note the variance of the error term can vary both across states and over time. The last term in (4) is assumed to be of negligible importance and subsequently dropped from the model. If we denote $i = 1, \dots, S$ and $t = 1, \dots, T$ then this assumption is justified when S/n_{it} and T/n_{it} are “small” (i.e., the number of individuals vastly outnumbers the number of states and time periods under consideration).² Therefore Eq. (4) can typically be simplified to

$$\ln\left(\frac{\bar{Y}_{it}}{1 - \bar{Y}_{it}}\right) = \alpha + \bar{X}_{it}\beta + g(\bar{X}_{it}\beta) + \mu_{it}. \quad (6)$$

Kelejian’s derivation suggests the function $g(\bar{X}_{it}\beta)$ needs to be included in the macrolevel equivalent of (2). Thus $g(\bar{X}_{it}\beta)$ can be interpreted as the aggregation bias term which is caused by within-state variation of the microlevel variables. This reflects the distribution of X_{jit} over j (individuals).

The exact functional form of $g(\bar{X}_{it}\beta)$ is not known but Kelejian’s analysis suggests it should be nonlinear. His central result is that a model which is nonlinear in its variables is likely to have a nonlinear aggregation bias term. Therefore, $g(\bar{X}_{it}\beta)$ might reasonably be approximated by a general polynomial of degree D .

$$g(\bar{X}_{it}\beta) \approx \sum_{d=0}^D (\bar{X}_{it}\beta)^d b_d. \quad (7)$$

I consider here the simple case of a quadratic which corresponds to $D = 2$ in (7). Thus,

$$g(\bar{X}_{it}\beta) \approx b_0 + (\bar{X}_{it}\beta)b_1 + (\bar{X}_{it}\beta)^2 b_2. \quad (8)$$

Substituting (8) into (6) and ignoring any approximation error yields

$$\ln\left(\frac{\bar{Y}_{it}}{1 - \bar{Y}_{it}}\right) = \alpha + \bar{X}_{it}\beta + b_0 + \bar{X}_{it}\beta b_1 + (\bar{X}_{it}\beta)^2 b_2 + \mu_{it} \quad (9)$$

which can be condensed into

$$\ln\left(\frac{\bar{Y}_{it}}{1 - \bar{Y}_{it}}\right) = a + \bar{X}_{it}\lambda + (\bar{X}_{it}\beta)^2 b_2 + \mu_{it}, \quad (10)$$

where $a = \alpha + b_0$, and $\lambda = \beta + \beta b_1$, which implies $\beta = \lambda(1 + b_1)^{-1}$. Substituting for β in (10) yields the final form representation

$$\ln\left(\frac{\bar{Y}_{it}}{1 - \bar{Y}_{it}}\right) = a + \bar{X}_{it}\lambda + (\bar{X}_{it}\lambda)^2 \gamma + \mu_{it}, \quad (11)$$

² Specifically, the condition requires $S/n_{it} \rightarrow 0$ and $T/n_{it} \rightarrow 0$.

where $\gamma = b_2(1 + b_1)^{-2}$. Under the given assumptions, Eq. (11) represents the proper macromodel aggregation of (2) and Kelejian's suggested test for aggregation bias relates to $\gamma = 0$.

II. DATA AND REGRESSION RESULTS

Aggregated data can now be used to estimate (11). The dependent variable is the logit of the state turnout rate. Total votes cast are available on tape from the Inter-University Consortium for Political and Social Research. The sample used here covers all gubernatorial elections held between 1870 and 1910. South Carolina elections are not included because there is an election (1876) in which the recorded turnout rate exceeds 100%. This estimate is clearly incorrect or fraudulent, and the logistic transformation cannot be computed.

Political and economic theory have been of little help in developing a standard turnout model and the empirical evidence has been mixed. Income, education, age, race, and residence seem to be the standard variables employed in regression analysis.³ I would therefore like to include as explanatory variables, the individual's wage, and whether or not the person was illiterate, elderly, black, or lived in an urban area. These variables represent the vector X_{jit} in (2). The corresponding aggregate-level variables in (11) are the average wage per worker and the percentage of illiterates, elderly, black, and urban in the state.

Population data are taken from Historical Statistics and interpolated for noncensus years. Data are not available for the Montana, North Dakota, and South Dakota elections of 1879. Real wage per worker is indexed to real GNP per capita.

McGovney (1949) has compiled a list of all states using poll taxes and literacy tests during this time period. Beeton (1986) describes states which granted full suffrage to women prior to the Nineteenth Amendment. Ludington (1911) lists states which had adopted a secret ballot by 1910.

A Southern dummy is included to control for turnout differences which might relate to Reconstruction. The drop-off effect, whereby fewer people tend to vote for governor in off-year elections, will be captured by a dummy variable for presidential election years. The data set contains 688 observations. Sample means and standard deviations are listed in Table 1.

Equation (11) is estimated by a nonlinear least-squares (NLLS) iterative procedure. Since the error term variance given in (5) is not assumed to be constant but the exact functional form is not specifically known, robust standard errors are used to account for possible unknown heteroskedasticity (White, 1980a).

There is also the possibility of autocorrelated errors from this aggregation model, but testing for the presence of autocorrelation is hampered by the nature of the data set. The timing of gubernatorial elections differs by state, and hence there are an unequal number of observations for each state. Standard tests for serially correlated errors require a balanced panel and for some of the states there are not

³ See for example, Wolfinger and Rosenstone (1980), Cassel and Hill (1981), Patterson and Caldeira (1983), Filer et al., (1991), Matsusaka and Palda (1993), and Knack (1993).

TABLE 1
Means and Standard Deviations

| Variable | Mean | Standard deviation |
|------------------------------|-------|--------------------|
| Turnout rate | .596 | .165 |
| Logistic turnout | .435 | .780 |
| Secret ballot | .465 | .499 |
| Poll tax | .185 | .388 |
| Literacy test | .151 | .358 |
| Female suffrage | .0392 | .194 |
| Presidential election | .395 | .489 |
| Log (wage/worker) | 8.042 | .282 |
| Urban percentage | .342 | .238 |
| Black percentage | .106 | .160 |
| Illiteracy rate | .122 | .117 |
| Over 65 years old percentage | .0391 | .0681 |
| South | .180 | .385 |

enough observations to implement these tests.⁴ But the nature of elections would not lead one to expect a systematic correlation of the errors over time in each state. Consider the cases of Virginia where elections were held only once every four years, and Massachusetts which held annual elections. If turnout follows an AR(1) process, this implies information from an election four years ago is an important indicator of turnout in Virginia, but not in Massachusetts where only last year's election would be considered relevant. A single AR process forces different error structures across the states, which does not seem reasonable. Furthermore, to the best of my knowledge, voting studies are not concerned with this possible problem, although heteroskedastic errors are often considered in different forms (Huston (1991), Knack (1993), Heckelman (1995)).

Despite these reservations, as a first approximation I estimated the Durbin-Watson statistic in each state based upon the residuals from the pooled sample (as mentioned above, individual state regressions were not possible in every case due to lack of observations). The estimated value was below the lower bound critical value in only three states. Although this is not a proper test for autocorrelation, in the absence of overwhelming evidence to support possible autocorrelation, I will assume the absence of time-dependent errors in this sample.⁵ Since the purpose of this study is primarily concerned with the ability to differentiate between the macro- and microestimates, further development of this issue here is beyond the scope of investigation. Thus, for the rest of this paper I focus only on possible heteroskedasticity, but it should be recognized that not correcting for autocorrelation, if it does in fact exist, will lead to inconsistent parameter estimates.

The coefficients (λ) for the macrolevel model are listed in Table 2.⁶ First note

⁴ See Judge et al. (1985) for a survey of serial correlation tests.

⁵ Removing these states from the sample does not substantially alter the regression estimates.

⁶ The Wald Statistic, which tests the hypothesis that the coefficient is not significantly different

TABLE 2
Macrolevel Regression Coefficients

| Variable | Estimates | Wald statistic ^a |
|-------------------------------|------------------|-----------------------------|
| Intercept | 1.937 (1.748) | |
| Secret ballot | .0598 (.103) | .337 |
| Poll tax | .282 (.473) | .355 |
| Literacy test | .0625 (.103) | .368 |
| Female suffrage | .0983 (.171) | .575 |
| Presidential election | -.208 (.355) | .586 |
| Log wages | .244 (.435) | .561 |
| Urban | .365 (.629) | .337 |
| Black | .904 (1.530) | .349 |
| Illiteracy | -.706 (1.166) | .367 |
| Over 65 | .349 (.630) | .307 |
| South | .169 (.291) | .337 |
| Aggregation bias (γ) | -.780 (2.318) | .113 |
| Log likelihood function | -547.916 | |
| Chi-square statistic | 11000.168* | |

Note. Model: $\ln(Y/1 - Y) = a + X\lambda + (X\lambda)^2\gamma + \mu$. Estimation by NLLS with robust standard errors using White (1980a) procedure. Standard errors are in parentheses.

^a Critical value of Wald statistic = 3.84 at 5% level.

* Significant at 5%.

that the aggregation bias term coefficient (γ) is not significantly different from zero. Aggregation bias does not appear to significantly affect the variable relationships. The remaining estimated coefficients properly describe the macrorelationship between the variables. For example, state turnout rates were higher for states with a high relative percentage of blacks but the difference was not statistically significant. We can also use the information from the macromodel to

from zero, is distributed as a chi-square with 1 degree of freedom. The critical value for a 95% confidence level is 3.84. None of the coefficients are significant by this criterion.

discuss the expected probability of voting for a generic individual, conditional upon the observed macrodata.

Since $E(Y_{jit} | \bar{X}_{it}) = \text{Prob}(Y_{jit} = 1 | \bar{X}_{it})$, by reversing the logistic transformation on (11) we can estimate,⁷

$$\text{Prob}(Y_{jit} = 1 | \bar{X}_{it}) \approx \frac{e^{a + \bar{X}_{it}\lambda + (\bar{X}_{it}\lambda)^2\gamma}}{1 + e^{a + \bar{X}_{it}\lambda + (\bar{X}_{it}\lambda)^2\gamma}}. \quad (12)$$

In essence, (12) represents the weighted average of the individual-level probabilities, where the weights are determined by the state-level information. Note that this is not the same as directly estimating P_{jit} , which is the expectation conditional upon the unknown vector X_{jit} . Since P_{jit} is a function of unknown personal socioeconomic characteristics given in (1), the best estimate for any randomly chosen individual residing in a particular state during a given election is found by a weighted average of all the possible heterogeneous voting probabilities in that state at that time.

For example, we would predict a person from a nonsouthern state without any of the considered electoral laws, that has mean values for wage/worker, and percentage of urban, black, illiterate, and elderly, would have a probability of voting on average equal to .639. In a presidential election year, this probability rises to .734. These predictions are found by substituting the mean values listed in Table 1 and the estimated coefficients in Table 2 into (12). The dummy variables corresponding to the various election laws and the south are not included since they would have a value of zero in this example. Although none of the estimated coefficients are individually statistically significant suggesting marginal effects are not different from zero, there may still be qualitatively important prediction differences for different states or time periods as the above example will attest. A presidential election generates almost a 15% difference in predicted voting probability. The reader is invited to consider alternative state values that may be of personal interest.

The coefficients from this model cannot, however, be used to explain how, for example, an *individual's* race affects his own probability of voting. The ecological fallacy is still present. The macrocoefficients do not exactly correspond to the original microcoefficients. This does not mean, however, that the individual relationships are uninterpretable.

Note the construction of the λ coefficient vector. The only difference between λ , the macrocoefficients, and β , the microcoefficients, is the multiplicative factor $(1 + b_1)$. If $\lambda = (\lambda_1, \dots, \lambda_k)'$, and $\beta = (\beta_1, \dots, \beta_k)'$ then for all $1 \leq v, w \leq k$,

$$\frac{\lambda_v}{\lambda_w} = \frac{\beta_v(1 + b_1)}{\beta_w(1 + b_1)} = \frac{\beta_v}{\beta_w}. \quad (13)$$

Equation (13) shows the ratios of coefficients from (11) can also be used to

⁷ See also Kelejian (1995, p. 246) for details on this approximation technique.

describe the ratios of coefficients from (2). While conclusions about an individual's race or residence cannot be independently interpreted, the information in Eq. (13) and Table 2 suggests

$$\frac{\lambda_{\% \text{ black}}}{\lambda_{\% \text{ urban}}} = \frac{.904}{.365} = 2.477 = \frac{\beta_{\text{black}}}{\beta_{\text{urban}}}.$$

We can therefore conclude that a person's race outweighs his residential characteristics in impacting upon the probability of voting.

Using (12) and (13) to interpret the regression in (11) can lend important new insights into voting probabilities without resorting to treating the state as the actor (see for example Patterson and Caldeira (1983)).⁸ Although these procedures may be helpful in answering certain types of questions about individual behavior when only aggregate data are available, they do not directly address the questions raised in the introductory section. To speak specifically about which individuals were most likely to have voted, it is necessary to invoke another assumption about the macromodel.

III. RECOVERING THE MICROCOEFFICIENTS

The aggregation bias in (6) took the form of a nonlinear component added to the original microequation. It has been shown that the quadratic term in (11) is not significant (γ in Table 2). According to (8) and (11), the only remaining bias must be of a linear form.⁹ Since $\lambda = \beta(1 + b_1)$ a test for the remaining aggregation bias hinges upon $b_1 = 0$.

$$\begin{aligned} H_0: & \quad b_1 = 0 \\ H_A: & \quad b_1 \neq 0 \end{aligned} \tag{14}$$

Unfortunately, λ cannot be disaggregated to isolate b_1 .¹⁰ There simply is no specification test for $b_1 = 0$. But is it reasonable to assume $b_1 \neq 0$ when $\gamma = 0$?

Since Kelejian's analysis suggests the aggregation bias of $g(\bar{X}_{it}\beta)$ should be nonlinear, it is unlikely that the aggregation bias is actually linear even if the nonlinear terms in $g(\bar{X}_{it}\beta)$ are not significant. If we are willing to accept the notion that the aggregation bias should be of a nonlinear form, we should be willing to rule out the importance of the linear terms when the nonlinear terms are found to be insignificant. Acceptance of the null hypothesis in (14) would then immediately follow.

Since Kelejian's aggregation bias test rejects the statistical significance of γ in our sample, then conditional upon $b_1 = 0$,

$$\ln\left(\frac{\bar{Y}_{it}}{1 - \bar{Y}_{it}}\right) = \alpha + \bar{X}_{it}\beta + \mu_{it}. \tag{15}$$

⁸ For a detailed criticism of the state-actor approach, see Matsusaka and Palda (1993).

⁹ Since $\gamma = b_2(1 + b_1)^{-2}$, nonsignificance of γ could imply either $b_2 = 0$ or $b_1 = \infty$. The former seems the more plausible assumption to make.

¹⁰ If b_1 could be isolated, it would be a simple task to solve for β regardless of the value of b_1 .

In this case, the microcoefficients can be directly estimated using the aggregate data variables. Thus a serious limitation to further analyzing aggregated data in the logit framework is the implicit assumption that $b_1 = 0$ when aggregation bias has been rejected, even if there is not direct empirical justification. The relevance of the rest of this section hinges upon the reader's acceptance of this notion.

Model (15) is estimated using ordinary least squares but the possible heteroskedasticity suggested in (5) remains.¹¹ In a linear model, consistent estimates of the variance-covariance matrix can be generated using the procedure described in White (1980b).

According to the estimated microcoefficients listed in the second column of Table 3, each of the state electoral laws considered does reduce the odds of an individual voting. Controlling for these laws, blacks, urbanites, the rich, and the elderly, were all less likely to vote. Illiterates were more likely to vote and presidential elections spurred voters to the polls to vote for governor.

The estimated coefficients explain the marginal impact each variable has on the log of the odds of voting. The importance of these results is easier to decipher after converting back to the probability of voting formula using Eq. (1). But in this expression, the model is no longer linear so the marginal impacts will not be constant. However, since the transformation in (1) is monotonic, variables with larger estimated coefficients will always generate larger probability changes, and the direction of these changes are known from the sign of the coefficient.

To estimate these probability changes, we can consider changes to an individual with specific socioeconomic characteristics. For simplicity, the default person is assumed to be a white literate under the age of 65 with mean income, who lives in the rural portion of a nonsouthern state where there are no secret ballot laws, poll taxes, literacy tests, or female suffrage laws. The probability of this person voting for governor in a nonpresidential election year is found by coding each variable as a zero value, except for log wages which is assigned the mean value (listed in Table 1). The predicted value using only the intercept term and mean income is then transformed as in Eq. (1). A person with these characteristics is estimated to have a probability of voting of .702.

Alternative probabilities of voting can be then found by altering this person's assumed characteristics by recoding each variable in turn. The changes induced by analyzing different types of characteristics, one at a time, are listed in the last two columns of Table 3. For example, poll taxes reduce the probability of this person voting by almost 25%, but he would be 13% more likely to vote during a presidential election year (assuming there are no poll taxes). If this person were instead from an urban area (assuming the absence of poll taxes and presidential elections), his probability of voting would be $-.180$ percentage points lower, or a 25% reduction. Similarly, a black with the above listed characteristics would be

¹¹ Again, we ignore the possible problems associated with serially correlated errors for the reasons discussed in Section II.

TABLE 3
Microlevel Regression Results

| Variable | Estimates | Marginal impact on probability of voting ^a | |
|-----------------------|-------------------|---|--------------------|
| | | Absolute change | Percent change |
| Intercept | 5.529 (.838) | | |
| Secret ballot | -.145* (.0461) | -.031 | -4.44 |
| Poll tax | -.744* (.0692) | -.174 | -24.76 |
| Literacy test | -.189* (.0613) | -.041 | -5.84 |
| Female suffrage | -.232* (.0887) | -.051 | -7.22 |
| Presidential election | .505* (.0441) | +.094 | +13.40 |
| Log wages | -.581* (.101) | -.035 ^b | -5.04 ^b |
| Urban | -.769* (.111) | -.180 | -25.65 |
| Black | -2.021* (.545) | -.464 | -66.11 |
| Illiteracy | 1.623* (.648) | +.221 | +31.44 |
| Over 65 | -.796* (.209) | -.187 | -26.61 |
| South | -.384* (.116) | -.086 | -12.24 |
| <i>R</i> ² | .523 | | |
| <i>F</i> statistic | 67.308* | | |

Note. Model: $\ln(Y/1 - Y) = \alpha + X\beta + \mu$. Estimation by OLS with robust standard errors using White (1980b) procedure. Standard errors are in parentheses.

^a Marginal effects are calculated for default individual whose characteristics are denoted by having all values set equal to zero except for log wages which has mean value. This individual has an expected probability of voting of .702. See text (Section III) for details.

^b Calculated for a one standard deviation increase in log wages.

* Significant at 5%.

expected to be 66% less likely to vote than a corresponding white (-.464 lower probability of voting).

Caution must be taken in interpreting the results. Recall the probability function is not linear, so one cannot simply add together two different marginal changes to estimate, for example, the change in probability associated with an urban black with otherwise identical characteristics. This would lead to an

overestimate of the induced probability change. To find the change in voting probability from a person with multiple characteristic differences, it is necessary to add the estimated regression coefficients prior to transforming the predicted value. For example, an urban black with otherwise identical characteristics as the default individual has an expected probability of voting of .126, which represents a decrease in probability of $-.576$, not $-.644$ which would be found by adding the *ceteris paribus* changes listed in Table 3 (and given in the above examples).

IV. CONCLUSIONS

Most aggregate-level voting studies ignore the ecological fallacy (e.g., Filer et al., 1991). Others have tried to skirt the issue by examining only across-unit variation (e.g. Patterson and Caldeira, 1983). The results generated in this paper suggest doing so may lead to the wrong conclusions. The macroresults in Table 2 imply none of the state laws examined had a significant impact on state turnout. The microresults in Table 3 suggest quite the opposite. Each of the laws did have an important influence on the probability of an individual in that state voting. How can these two results be reconciled?

As was noted in Section II, none of the individual-level variables have any importance at the state level. None of these variables are significant at the standard 5% level. This would seem to suggest that direct aggregation may not reveal the proper aggregate-level relationships. While income may be an important contributing factor to an individual's voting decision, and for all individuals in the state, it may not translate to per-capita income being important in determining turnout rates. The macromodel has an added component due to the distributional asymmetries which has no relevance in the micromodel. States with low per capita income need not be states with the greatest *number* of poor, which would be needed for the macrocoefficient to reflect the microcoefficient (Roth (1986)). Thus, if one is solely interested in estimating the macrorelationships, this model may need to be supplemented with pure macrovariables, such as gini coefficients. The macromodel developed in Section I, while superior to estimation techniques which simply employ state-level proxies, may still be misspecified so the results may be misleading.

It should also be remembered that the macroresults do not in general present a sufficient proxy for the true microcoefficients. Relying upon the macroresults yields biased estimates of the microresults. Poll taxes and literacy tests did in fact reduce the likelihood for individuals to vote. Table 3 makes this clear. Generating an aggregate model is not sufficient to discuss microimplications, even when only macrodata are available.

The ecological fallacy, however, has been overstated. Proper aggregation of the micromodel can yield meaningful interpretations of the microrelationships in the form of ratios.¹² Furthermore, as shown in Section III, when aggregation bias is not

¹² Although the macro- and microcoefficient estimates are quite different, their ratios are very similar in both sign and magnitude.

present it may also be possible under certain additional assumptions to recover the original microcoefficients, in which case the ecological fallacy is eliminated entirely.

One ground for caution remains. The functional form of the aggregation bias tested was not formally derived; the true distribution remains unknown. A simple functional form was suggested and modeled. Nonsignificance of the aggregation bias coefficient is thus consistent with two alternative hypotheses: either aggregation bias is not an important consideration to the regression, or the form of the bias is not correlated with a quadratic. It is possible that other distributions might better capture the bias, but which functions these are remain a mystery. A promising avenue for future work would be to develop alternative specifications for possible aggregation bias. The results derived here should be viewed only as preliminary, but it is believed that other aggregate data studies would benefit from taking even these preliminary steps.

REFERENCES

- Beeton, B. (1986). *Women Vote in the West: The Woman Suffrage Movement 1869–1896*, Garland Publishing, Inc., New York.
- Cassel, C. A., and Hill, D. B. (1981). "Explanation of turnout decline: A multivariate test," *American Politics Quarterly* **9**, 181–195.
- Filer, J. E., Kenny, L. W., and Morton, R. W. (1991). "Voting laws, educational policies, and minority turnout," *Journal of Law and Economics* **34**, 371–393.
- Heckelman, J. C. (1995). "The effect of the secret ballot on voter turnout rates," *Public Choice* **82**, 107–124.
- Historical Statistics of The United States (1975). *Colonial Times to 1970, Parts I and II*, Washington, DC, Government Printing Office.
- Huston, J. L. (1991). "Weighting, confidence intervals, and ecological regression," *Journal of Interdisciplinary History* **21**, 631–654.
- Judge, G. G., Griffiths, W. E., Hill, R. C., Lütkephol, H., and Lee, T.-C. (1985). *The Theory and Practice of Econometrics*, Wiley, New York.
- Kelejian, H. (1995). "Aggregated heterogeneous dependent data and the logit model: A suggested approach," *Economics Letters* **47**, 243–248.
- Key, V. O. (1949). *Southern Politics in State and Nation*, Vintage Books, New York.
- Kim, J., Petrocik, J. R., and Enokson, S. N. (1975). "Voter turnout among the American states: Systemic and individual components," *American Political Science Review* **69**, 107–131.
- Klein, S. P., Sacks, J., and Feldman, D. A. (1991). "Ecological regression *versus* the secret ballot," *Jurimetrics* **31**, 393–413.
- Knack, S. (1993). "The voter participation effects of selecting jurors from registration lists," *Journal of Law and Economics* **36**, 99–114.
- Kousser, J. M. (1974). *The Shaping of Southern Politics: Suffrage Restrictions and the Establishment of the One-Party South, 1880–1910*, Yale Univ. Press, New Haven, CT.
- Ludington, A. (1911). "American ballot laws, 1888–1910," *Education Department Bulletin (New York)* **448**, 10–207.
- Matususaka, J. G., and Palda, F. (1993). "The Downsian voter meets the ecological fallacy," *Public Choice* **77**, 855–878.
- McGowney, D. (1949). *The American Suffrage Medley*, Univ. of Chicago Press, Chicago.
- Patterson, S. C., and Caldeira, G. A. (1983). "Getting out the vote: Participation in gubernatorial elections," *American Political Science Review* **77**, 675–689.

- Robinson, W. S. (1950). "Ecological correlations and the behavior of individuals," *American Sociological Review* **15**, 351–357.
- Roth, R. A. (1986). "Ecological regression and the analysis of voter behavior," *Historical Methods* **19**, 103–117.
- Rusk, J. (1974). "The American electoral universe: Speculation and evidence," *American Political Science Review* **68**, 1028–1049.
- White, H. (1980a). "Nonlinear regression on cross-section data," *Econometrica* **48**, 721–746.
- White, H. (1980b). "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica* **48**, 817–838.
- Wolfinger, R., and Rosenstone, S. (1980). *Who Votes?* Yale Univ. Press, New Haven, CT.