EVALUATION REVIEW

A Journal of Applied Social Research

*Volume 15, Number 6* *December 1991*

THIS ISSUE IS DEVOTED TO:
**Statistical and Demographic Issues Underlying Voting Rights Cases**
*Edited by:* DANIEL L. RUBINFELD

CONTENTS



Sage Periodicals Press A Division of SAGE Publications, Inc. Newbury Park • London • New Delhi

*This article serves as an introduction to a symposium that was organized with the goal of explicating, discussing, and evaluating the validity of the most important statistical and demographic issues that have been raised in voting rights cases. The article outlines, discusses, and evaluates the issues surrounding a recent voting rights case in Los Angeles,* Garza v. County of Los Angeles.

# STATISTICAL AND DEMOGRAPHIC ISSUES UNDERLYING VOTING RIGHTS CASES

DANIEL L. RUBINFELD
*University of California, Berkeley*

The Voting Rights Act of 1965 explicitly forbids the imposition or application of any practice that would deny or abridge, on grounds of race or color, the right of any citizen to vote (42 U.S.C., 1973). As amended in 1982, the Act forbids not only intentional discrimination but any practice that is shown to have a disparate impact on minority voting strength. In other words, the amendment provides minority groups with a legal remedy if they can show that their voting strength has been diluted (42 U.S.C., 1973[b]).

The amended Voting Rights Act has been evoked frequently in the past decade as a means of overcoming disparate impacts on ethnic minorities such as African Americans and Hispanics.[1] Statistical and demographic evidence has been prominent in a number of these cases.[2] This symposium was organized with the goal of explicating, discussing, and evaluating the validity of the most important statistical and demographic issues that have been raised in voting rights cases. To make what might become an extremely abstract and theoretical debate more concrete, I have organized the symposium around a recent voting rights case. A brief description of the case is followed by an overview of the symposium issues.

On June 4, 1990, the U.S. District Court of the Central District of California filed an opinion in *Yolanda Garza et al. v. County of Los Angeles, Los Angeles Board of Supervisors et al.*, hereinafter *Garza* (CV 88-5143 KN, CV 88-5435 KN, 90 *Daily Journal* DAR 6141). The District Court's opinion

EVALUATION REVIEW, Vol. 15 No. 6, December 1991 659-672

concerning liability and remedy was upheld by the Ninth Circuit Court of Appeals on November 2, 1990 (90 CDOS 8138).[3] The U.S. Supreme Court recently denied the defendant's certiori petition.

The background of the case is as follows. Throughout this century the Los Angeles County Board of Supervisors has consisted of five members elected to serve 4-year terms in nonpartisan elections within the appropriate supervisory districts. Prior to 1991, a Hispanic had never been elected to a supervisorial position. In his District Court opinion in *Garza*, Judge David Kenyon ruled that plaintiffs had provided persuasive evidence showing that the Board of Supervisors had intentionally engaged in discriminatory redistricting that took place in 1959, 1965, and 1971. He also ruled that the most recent 1981 reapportionment, while not intentionally diluting voting strength, had the effect of preventing Hispanics from attaining a majority in any single district in the future. As a remedy, the District Court accepted a proposed reform plan that created a district in which the majority of the *voting citizen population* is Hispanic. As a consequence, in February 1990, the first Hispanic supervisor in over a century was elected in the newly created First Supervisorial District.

## LEGAL PRECEDENT

To understand what was involved in *Garza* specifically, and more generally to understand what it takes for a judge in a voting rights case to order that redistricting take place, some legal history is required. In an important case, *Thornburg v. Gingles* (1986; hereafter *Gingles*), the U.S. Supreme Court stated that discriminatory intent was not required for the plaintiff to be successful. Rather, the plaintiff must be able to show that it is possible for the minority group to elect a candidate or candidates of its choice. Specifically, the Court established three conditions that must be shown in order for the plaintiffs to prove liability based on discriminatory effect rather than discriminatory intent.

1. The minority group must be *geographically compact*. That is, there must be a geographical district that can be drawn in which the minority group will be a majority.
2. The minority group must be *politically cohesive*. That is, the minority political group must vote in a substantial majority for its desired candidate.
3. There must be a majority voting bloc. That is, the majority must be shown to vote sufficiently as a unified *racial voting bloc* so as to usually defeat the minority's preferred candidate.



Underlying all three of these requirements in voting rights cases are several important conceptual issues that the courts have not fully resolved: Is the relevant population to be considered in the evaluation of the three *Gingles* requirements (a) the total population or (b) the citizen voting-age population?[4] What is the appropriate criterion for deciding whether the Court's redistricting plan satisfies the U.S. Constitution's one person-one vote requirement?

In voting rights cases prior to *Garza*, the distinction between total population and citizen voting-age population was inconsequential. In Los Angeles, unlike most other major U.S. cities, there are a substantial number of noncitizen minorities. As a consequence, in *Garza*, the total population-citizen voting-age population distinction was crucial to the outcome of the case. (According to the 1980 Census, Hispanics make up 28% of the total population of Los Angeles County but only 15% of voting-age citizens.)

In *Garza*, the Court presumed, consistent with *Gingles* and with a number of earlier rulings, that citizen voting-age population rather than the total population is the appropriate ideal measure of geographical compactness. However, because the Court found that there had been intentional discrimination, it was not required to apply the *Gingles* disparate impact test. In particular, the Court chose not to apply a bright line test requiring just over 50% Hispanics in the voting-age population in order for the compactness test to be satisfied.[5]

There is even more controversy in the courts and in this symposium concerning the political cohesiveness and racial-bloc voting criteria.[6] The District Court's acceptable reapportionment plan, which must satisfy the second and third *Gingles* tests, was based on both total population and citizen voting-age population statistics. Thus in the plaintiff's proposed remedy, the *Hispanic opportunity district*, which was 71.2% Hispanic, was expected to contain 1,779,835 or 20% of the total population of 8,880,109 for Los Angeles County. (According to recent 1990 census figures, however, the population of this district is actually 12% lower than the population of the average district.[7]) However, the same district was expected to contain only 707,651 or 14.4% of the 4,897,971 voting-age citizens in the county. Los Angeles County argued unsuccessfully for citizen voting-age population and against total population as the basis for satisfying the one person-one vote criterion. It suggested that a redistricting plan based on population alone would weight the votes of residents of the district in which Hispanics are concentrated more heavily than those of citizens in other districts. The Court did agree, however, that citizen voting-age population was the appropriate basis for applying the compactness test.

With respect to the one person-one vote rule, Congress allocates congressional seats to states on the basis of number of persons, whether or not they are voting-age citizens. Within states, the issue is less clear, however, depending on the stated intent of the relevant state legislation (*Burns v. Richardson* 1966). In his dissent from the majority opinion of the Ninth Circuit Court of Appeals (1990) Judge Kozinski focused on this exact issue:

> Apportionment by raw population embodies the principle of equal representation; it assures that all persons living within a district – whether eligible to vote or not – have roughly equal representation in the governing body. A principal of equal representation serves important purposes: It assures that constituents have more or less equal access to their elected officials, by assuring that no official has a disproportionately large number of constituents to satisfy. . . .
>
> Apportionment by proportion of eligible voters serves the principle of electoral quality. This principle recognizes that electors – persons eligible to vote – are the ones who hold the ultimate political power in our democracy. This is an important power reserved only to certain members of society; states are not required to bestow it upon aliens, transients, short-term residents, persons convicted of crime, or those considered too young. (p. 8145, footnote omitted.)

The statistical and demographic issues that arise in *Garza* tie directly to the three prongs of the *Gingles* requirement. The question of whether there is geographic compactness has an important demographic component, as it involves issues of how best to use and interpret census data concerning total population and citizen voting-age population. The distinction between total population and citizen voting-age population is not at issue with respect to the second and third prongs involving political cohesiveness and racial bloc voting, respectively. What is of concern in all of *Gingles* are statistical issues that revolve around at the question of whether it is possible to make inferences about individual voter preferences from reported voting results that are aggregated by voter precincts? In short, an important statistical issue (absent a finding of intentional discrimination) is whether it is possible to redistrict in Los Angeles so as to satisfy the *Gingles* tests?

At the core of the statistical issues is the technique of *ecological regression*, which was successfully used by the plaintiff's experts in this and other cases. In the context of voting rights cases, ecological regression allows one to make inferences about individuals from aggregate data from geographical areas such as voting precincts. (There are approximately 6,000 voting precincts in Los Angeles County.) If one had exit polls in a sample of precincts, then the inferences that ecological regression makes would not be necessary. Indeed, in the special case in which voting precincts are racially homogeneous, we could determine the preferences of racial groups simply by looking at precinct-level votes. It would then be relatively easy to aggregate these precinct results to obtain approximate estimates of voting patterns across supervisorial districts. Unfortunately, from a statistical point of view, most precincts (and districts) are not homogeneous; as a result, any inferences about individual preferences that are made from precinct vote tallies will involve potentially substantial statistical uncertainties.



## THE SYMPOSIUM

The five articles that follow are written (all or in part) by experts who testified in the *Garza* case. The first, "Ecological Regression and Voting Rights" by Freedman, Klein, Sacks, Smyth, and Everett (the first three of whom testified for the defendant), focuses almost entirely on the methodology involving ecological regression. Freedman et al. argue that it is inappropriate to apply the ecological regression model in evaluating the legal issues in *Garza*. To effectuate this criticism, they propose an alternative "neighborhood model" that they argue is at least as effective as the ecological regression model in explaining voting patterns.

The Freedman et al. article is followed by William Clark (a fourth defendant expert) and Peter Morrison's "Demographic Paradoxes in the Los Angeles Voting Rights Case," that explores some of the demographic issues involved in the case from the defendant's perspective. In addition, a brief postscript provides some recent new evidence from the 1990 census. "The Use of Demographic Data in Voting Rights Cases" by William O'Hare provides a different perspective from that of Clark and Morrison. "Statistics Without Substance: A Critique of Freedman et al. and Clark and Morrison" by Bernard Grofman, a plaintiff expert, responds to the criticisms of Freedman et al. and also comments critically on the demographic issues raised by Clark and Morrison. A more directed criticism of the ecological regression materials of Freedman et al. is given in the final article by Allan Lichtman, "Passing the Test: Ecological Regression Analysis in the *Garza* Case and Beyond." The symposium concludes with a final reply to the critiques of Grofman and Lichtman by Freedman et al.

These symposium pieces provide a unique opportunity to learn about the demographic and statistical issues in an important voting rights case directly from many of the players involved. Because many of these issues are complex, the sections that follow have been organized to highlight the most important areas of contention.

## GEOGRAPHICAL COMPACTNESS

The first *Gingles* requirement is that the minority group be geographically compact. If the 1980 Census data were used as the basis of analysis, the compactness criterion would not be met, since Hispanics would constitute only 41% of the eligible voters of the plaintiff's proposed Hispanic Opportunity District (HOD) district. Clark and Morrison point out that Hispanics are, on average, younger and less often citizens than non-Hispanics. Clark and Morrison find this prospect troubling because it suggests that because Hispanics are not the majority of voters in "their" HOD, that the compactness test is not met. Clark and Morrison do not, however, address the question of whether this particular "political" outcome is worse than the original supervisorial district it would replace, which the Court feels arose because of a pattern and practice of intentional discrimination.

Several important demographic issues follow from this starting point. First, should post-1980 Census data be used as the basis of analysis? The Court concluded that such data are likely to provide a more accurate point estimate of the current total number of Hispanics in Los Angeles County and are more likely to provide support for the view that the compactness test can be met (as the Court has done). Thus O'Hare cites county data that show an estimated increase in the number of Hispanics in Los Angeles from 2.032 million in 1980 to 3.251 million in 1990. However, the post-1980 data are generally not as reliable as the census data and therefore can create problems and inaccuracies when the unit of analysis is narrower than the county, say, at the census tract level. It is interesting to note that in *Garza*, Judge Kenyon was concerned with this issue, although arguably for legal rather than statistical reasons. In the end, Judge Kenyon ruled that the geographic compactness test was met using the post-1980 data. O'Hare provides an interesting discussion of the pros and cons of using postcensus data; Clark and Morrison do not directly confront the question of whether post-1980 Census data are appropriate. None of the articles raise the provocative question of whether the Court should have waited until the 1990 Census data were available before reaching a decision.

A second demographic issue relates to the reliability of the Census reports of who is and is not a voting-age citizen. Thus, according to Clark and Morrison, many Mexican-born noncitizens reported themselves as naturalized citizens of Mexican birth on the 1980 Census and were counted as citizens when in fact they were not. Clark and Morrison would correct the published census data for overreporting of Hispanics, but others dispute the wisdom of doing so.



Third, one can question the best way in which county data concerning citizenship can be disaggegated to obtain the relevant supervisorial district data. Thus with respect to citizenship, Clark and Morrison point to the fact that published census data do not distinguish voting-age citizens and noncitizens, even at the precinct level. To design a redistricting plan that satisfies the compactness criterion, however, one must make a reasonable assumption about the extent, if any, that the proportion of citizens in the population varies across supervisorial districts. One plausible approach is to work with census tract data that can be aggregated into the appropriate districts. For the demographers, then, how citizenship proportions vary by census tract during and after the 1980 Census is a critical issue.

## POLITICAL COHESIVENESS AND RACIAL VOTING BLOCS

Recall that to satisfy the *Gingles* test, the plaintiff must show that the minority group votes in a substantial majority for its desired candidate and that, absent redistricting, the minority group would be outvoted by a majority voting bloc. Because individual votes (and ethnic status) are not publicly available,[8] statistical inferences must be made. One plausible approach that should probably be seriously considered by the courts is the use of extensive exit polling, as suggested by Freedman et al. While exit polling can be costly, these costs are certainly small in relation to the costs of litigating cases such as *Garza*. However, because exit polling data were not available for the relevant elections, the *Garza* court did not treat this option, and it will not be pursued further here.

The approach used by plaintiffs in *Garza* builds on the fact that one can obtain data by voter precinct that describe voting outcomes and ethnic makeup but not the correlation between the two. They used the statistical technique of *ecological regression* to make inferences about individual voting behavior by ethnicity from aggregated voter precinct data. Formally, ecological regression assumes that the preferences of racial groups are identical across precinct boundaries. Practically, Grofman, Lichtman, and Freedman et al. would most likely agree that ecological regression is most effective when the preferences of racial groups are reasonably similar across political boundaries, and least effective when preferences vary substantially across these boundaries. However, there is very substantial disagreement as to the practical validity of the procedure in the context of the *Garza* case.

To see how ecological regression works in its simplest form, suppose that there are two candidates, one of which is Hispanic and the other non-

Hispanic, and that turnout rates across candidates are identical for Hispanics and non-Hispanics.[9] Let

$y_i$ = percentage of votes received by the Hispanic candidate in precinct i
$x_i$ = percentage of the voters who are Hispanic
$\varepsilon_i$ = a random error term that reflects differences in turnout rates for political candidates among precincts.

The ecological regression model assumes that voter preferences are determined by ethnicity and not other district-specific factors. It is given by $y_i = \alpha + \beta x_i + \varepsilon_i$ in which $\alpha$ measures the percentage of non-Hispanics who vote for the Hispanic candidate and $\alpha + \beta$ represents the percentage of Hispanics who voted for that candidate.[10] Note that a low value of $\alpha$, which tells us that relatively few non-Hispanics supported the Hispanic candidate, is consistent with the ethnic voting bloc requirement. In addition, a high value of $\alpha + \beta$, which tells us that most Hispanics voted for the Hispanic candidate, is consistent with the political cohesiveness test.

Using a more sophisticated version of this ecological regression model, plaintiffs' experts found, in effect, that $\alpha$ was relatively low and $\alpha + \beta$ relatively high. The court used this evidence (as they had in previous voting rights cases) as the basis of their conclusion that the political cohesiveness and racial voting bloc tests were satisfied.

Freedman et al. criticize the use of this particular ecological regression model. In statistical terms, their criticism is that it is invalid to assume that $\alpha$ and $\beta$ are constant across voter precincts and that $\varepsilon$ is an identically distributed random variable with zero mean. In terms of voting behavior, the criticism is that it is invalid to assume that voter preferences are independent of the fraction of Hispanics in each precinct, or to put it differently, that individual Hispanics in each precinct have the same voting preferences. Freedman et al. raise the possibility that a positive $\beta$ occurs because non-Hispanics in largely Hispanic precincts are more likely to support the Hispanic candidate than are non-Hispanics in largely non-Hispanic precincts.

To illustrate this criticism, Freedman et al. put forward for consideration an alternative *neighborhood* model. In its simplest form, the neighborhood model makes the assumption that within each individual precinct there is no difference between the voting behavior of Hispanics and non-Hispanics—neighbors vote alike. Differences in votes across precincts arise because preferences differ, not because of differences in the ethnic makeup of the precincts. Freedman et al. criticize the ecological regression model by focusing on the "constancy assumption," which holds that voting behavior within ethnic groups does not vary across precincts.



Both the ecological regression model and the neighborhood model have the same goal: to make inferences about individuals on the basis of aggregated precinct-level data. To get a clearer sense of how the models compare to each other, it is useful to go back to basics to see what assumptions about individual, or micro, behavior underlie each. To begin, consider the micro version of the ecological regression model:

$$y_{jn} = \alpha + \varepsilon \tag{1}$$

$$y_{jh} = (\alpha + \beta) + \varepsilon_j, \tag{2}$$

where $y_{jn}$ is the probability that the $j^{th}$ non-Hispanic will vote for the Hispanic candidate, and $y_{jh}$ is the probability that the $j^{th}$ Hispanic will vote for the Hispanic candidate.

Equations 1 and 2 describe linear probability models; in each case, the probability that a particular individual will vote for a Hispanic candidate is a function of a fixed component and a random error term. When aggregated over a precinct with $N$ voters (h = Hispanics and $N$-h = non-Hispanics), the linear probability model becomes

$$y = (hy_{jh} + ny_{jn})/N = \alpha + \beta(h/N) + \varepsilon = \alpha + \beta x + \varepsilon \tag{3}$$

Note that in the ecological regression model, the parameter $\beta$ is a *micro*-parameter; it reflects differences in the behavior of Hispanic and non-Hispanic individuals. In this version of the ecological regression model, there are no *macro*parameters that reflect differences across precincts because all individuals are assumed to behave the same wherever they are located.

It is clear that the aggregate version of the ecological regression model given in Equation 3 can easily be estimated by using a linear regression of y on x. What is not so clear is how one should interpret the coefficients of the model. To pursue this interpretation issue further, consider the assumptions underlying the micro version of the neighborhood model:

$$Y_{jnp} = y_{jhp} = (\alpha + \beta)_p + \varepsilon_j, \tag{4}$$

where $y_{jnp}$ is the probability that the $j^{th}$ Hispanic in the $p^{th}$ precinct will vote for the Hispanic candidate, while $y_{jhp}$ is the probability that the $j^{th}$ non-Hispanic in the $p^{th}$ precinct will vote for the Hispanic candidate, and $(\alpha + \beta)_p$ is a set of parameters that varies across precincts.

This linear probability model assumes that voter preferences are independent of race; more important, it assumes that preferences vary systematically from precinct to precinct. When aggregated over a precinct with $N$ voters, the neighborhood model becomes

$$y_p = (hy_{jhp} + ny_{jnp})/N = (\alpha + \beta)_p + \varepsilon_p. \quad (5)$$

To complete the specification of the model, Freedman et al. must describe how the parameters of the neighborhood model vary across precincts. In their preferred *nonlinear* model, no functional form is forced on these parameters; essentially, one parameter is calculated for each precinct. However, in the special case of a *linear* model, the authors add structure by assuming that the fraction of voters supporting the Hispanic candidate will vary from precinct to precinct *as a function of the fraction of Hispanics in that precinct.* In the case of the linear neighborhood model, the fraction supporting the Hispanic candidate is presumed to be a linear function of the number of Hispanics in the district; that is,

$$(\alpha + \beta)_p = \alpha + \beta(h/N) \quad (6)$$

This relationship might come about, for example, if the probability of a non-Hispanic voting for a Hispanic candidate increases as the proportion of Hispanics in the precinct increases. (Perhaps voters in "integrated precincts" are more "liberal" than voters in homogeneous non-Hispanic precincts.[11])

With the additional assumption given by Equation 6, the modified "linear" version of the neighborhood model becomes

$$y = \alpha + \beta x + \varepsilon \quad (7)$$

In Equation 7, the parameter $\beta$ is a *macro*parameter because it reflects the differences across precincts in voting behavior, not the differences among individuals. Even though the specification in Equation 7 is identical to that of Equation 3, the underlying assumptions are different; for example, in this version of the neighborhood model, there are no microparameters.

Equation 7 can be estimated by least squares to obtain the same coefficients as in the ecological regression model. The essential difference between the ecological regression and neighborhood models lies in the interpretation of the coefficients of Equations 3 and 7. In the ecological regression model, $\alpha$ is the percentage of non-Hispanic votes for the Hispanic candidate; in the neighborhood model, $\alpha$ is the percentage of votes for the Hispanic candidate in a precinct with no Hispanics. In the ecological regression model, $\alpha + \beta$ is the percentage of Hispanic votes for the Hispanic candidate; in the neighborhood model, $\alpha + \beta$ is the percentage of votes for the Hispanic candidate in a precinct with all Hispanics.

From the point of view of the statistician, the two models cannot be distinguished using aggregate data because they make the same predictions (albeit for different reasons). Thus if there are h Hispanics is a supervisorial district of $N$ voters, then the ecological regression model predicts $(\alpha + \beta)h = \alpha h + \beta h$ Hispanic votes from Equation 1, as well as $\alpha(N - h)$ non-Hispanic votes from Equation 2, and therefore, $\alpha N + \beta h$ total votes for the Hispanic candidate. The neighborhood model would predict $(\alpha + \beta[h/N])N = \alpha N + \beta h$ votes for the Hispanic candidate, in this case from Equation 7. If estimated using ordinary least squares, both models generate the same coefficients; as there are at least two possible interpretations, the two models are unidentified. Without looking at microlevel data, one cannot tell which model more appropriately describes individual voter preferences.



To see how the models differ in their implications, suppose that the least squares regression yields the following parameter estimates:

$$y = 30\% + 40\%x. \quad (8)$$

Consider how each model would predict the votes in a district that is 60% Hispanic (and 40% non-Hispanic). The ecological regression model would predict that 30% of the non-Hispanics and 70% of the Hispanics (30% + 40%) vote for the Hispanic candidate; overall, 54% [30%(.4) + 70%(.6)] of voters would be predicted to vote for the Hispanic; the neighborhood model makes the identical prediction – that 54% [30% + 40%(.6)] of all voters (Hispanic or otherwise) would vote for the Hispanic candidate.

To summarize, the ecological regression and neighborhood models differ in their underlying assumptions, but they generate the same predictions with respect to the actual aggregated voting data. As a result of the important basic differences in the models, the model choice that one makes can affect the conclusion that one reaches concerning polarized voting. Freedman et al. point out that if the ecological regression model is correct, the neighborhood model may not find polarized voting when it exists. On the other hand, if the neighborhood model is correct, the ecological regression model may overestimate the differences between Hispanics and non-Hispanics, and therefore may find polarized voting when it is not present. Freedman et al. go on to suggest that the neighborhood model might better characterize individual vote preferences than the ecological regression model. In doing so, they rely primarily on evidence from other elections and other studies.

The presentation of an alternative to the ecological model is a useful pedagogic device, as it serves to highlight the limitations of the ecological model. However, the neighborhood model can be criticized on a number of grounds – in the end it may or may not be more effective than the ecological regression model for explaining voting. Finally, both the neighborhood

model and the ecological regression model as used in *Garza* lack a fully satisfactory conceptual underpinning. Neither Lichtman nor Grofman explain why they believe that Hispanic preferences for Hispanic candidates are identical across districts, nor do they explain why other important socioeconomic factors do not explain voting preferences. Freedman et al. do not fully explain why the underlying preferences of voters are such that the fraction of voters who support Hispanic candidates is related linearly to the fraction of Hispanics in each district.

Both the Lichtman and the Grofman articles defend the use of ecological regression and the conclusion that there is political cohesiveness and racial bloc voting in Los Angeles County. One important response of the authors goes broadly to the question of policy relevance: The issue is not whether ecological regression can perfectly explain individual voting behavior—it clearly cannot; rather, the issue is whether the conclusion of political cohesiveness is appropriate in the Los Angeles case. Lichtman and Grofman believe that it is. Lichtman supports this view by showing voting predictions that result when the ecological regression model is applied to 20 elections in Los Angeles in the 1980s (but see Freedman et al.'s answer to the Lichtman evidence).

Lichtman and Grofman also point to the Court's opinion as supporting the use of ecological regression; the support of the Court is based on the belief that ecological regression provides useful estimates of average voting behavior in most jurisdictions. They also praise ecological regression because it has been effective in prior legal procedures. This line of argument is not persuasive, however, because most prior situations involved African-American rather than Hispanic cases. Neither Lichtman nor Grofman presents any evidence to support the view that the correlation between ethnic voting and other important determinants of voting behavior (e.g., party identification and socioeconomic status) is similar for Blacks and Hispanics.

Lichtman points to another problem associated with the neighborhood model, at least in its simplest form. The neighborhood model assumes that the total number of votes for a Hispanic candidate within each precinct is a function of the percentage of the precinct that is Hispanic. However, the model does not specify which variables, such as political affiliation and socioeconomic status, explain these variations in voting behavior. This criticism seems inappropriate; one can make a strong case for the use of a multiple regression model that controls for certain socioeconomic and political factors, rather than either the ecological regression model or the neighborhood model. If one wants to make the most accurate inferences about individual preferences as possible, then controlling for other variables can be important.



Lichtman and Grofman would object to the use of a multiple regression version of the ecological model because the correlation among independent variables (including ethnicity) can complicate the process of making predictions about district-level voting by ethnic group. Freedman et al. would be more sympathetic to a multiple regression approach because it allows for a more complex specification of the manner in which voting for Hispanic candidates varies across voter precincts. Defendants apparently chose not to pursue this approach, in part to simplify the voting analysis.

## NOTES

1. Thernstrom (1987) provides a review of the case law.
2. See Jacobs and O'Rourke (1986), Part 2, for a review of the use of various statistical techniques in voting rights cases.
3. The panel consisted of Judges Schroeder, Nelson, and Kozinski.
4. The article by O'Hare in this symposium makes an effective case for a third alternative: citizens who are eligible to vote. This definition would rule out recent movers and those in jails and prisons as well as convicted felons out of prison.
5. This avoidance of a bright-line test seems appropriate from a policy perspective. As O'Hare points out, if one's goal is to have a substantial probability of winning an election, variations in voter registration and voter turnout can allow a "minority" to elect a candidate of their choice.
6. Clark and Morrison presume that citizen voting-age population is the appropriate norm, Grofman, however, points out that the courts have not generally supported this view.
7. The details are presented in the final rejoinder of Freedman et al.
8. Data on ethnicity can be inferred by searching for Spanish surnames.
9. Allowing for turnout complicates the ecological regression formula without adding substantively to the issues discussed here.
10. A second equation is required to account for differences in turnout rates among precincts.
11. Perhaps the percentage of Democrats among non-Hispanics increases proportionately with the percentage of non-Hispanics.

## REFERENCES

Burns v. Richardson, 384 U.S. 73, 91-91, 86 St. Clair 1286, 1296-97 (1966).

Jacobs, Paul W. II, and Timothy G. O'Rourke. 1986. Racial polarization in vote dilution cases under section 2 of the Voting Rights Act: The impact of *Thornburg v. Gingles*. *Journal of Law & Politics* 3:295-354.

Ninth Circuit Court of Appeals. 1990. Yolanda Garza et al. v. County of Los Angeles et al., 90 CDOS 8138ff.

Thernstrom, Abigail M. 1989. *Whose votes count? Affirmative action and minority voting rights* Cambridge, MA: Harvard University Press.
Thornburg v. Gingles, 478 U.S. 30, 106 S. Ct. 2752 (1986).

*Daniel L. Rubinfeld is Professor of Law and Professor of Economics at the University of California, Berkeley. He is coeditor of* International Review of Law and Economics *and former Chair of the Program in Jurisprudence and Social Policy. His research interests include quantitative methods in law, the economics of litigation, and federalism.*