

Gambling with Elections: The Problems of Geodemographics

Mark Atlas

The introductory essay in this volume asked whether the new campaign technology is "magic . . . or blue smoke and mirrors?" The previous article by Jonathan Robbin proclaimed the Claritas Clusters targeting system to be "the new magic." This article by Mark Atlas presents a very different perspective on Robbin's creation. Atlas critically evaluates the application of the Cluster system of geodemographics to political campaigns. (Using the criteria and data supplied by Claritas, the author concludes that the claims made about the system's performance in a 1978 election in Missouri are unsubstantiated.) The system was touted for its ability to turn a seemingly lost election into a victory and accurately to project what the final election results would be in each Missouri county. A reanalysis of the data suggests, however, that neither assertion is correct. The article also points out serious problems with the system's data collection procedures and potential difficulties with the quantitative technique used to process the data.

There is an old saying among the managers of Las Vegas casinos: "We welcome all gamblers coming to our casinos, but for gamblers with a 'system' we'll send them a free plane." Just as prospective gamblers may lunge too eagerly at "systems" that appear to offer the magic solution to their problems but that are in reality only sophisticated methods for losing their funds, so too political campaigners may grasp purported "revolutionary" improvements in campaign technologies too quickly for their own good. While no research methodology attempting to analyze political attitudes and behaviors can be perfect, it is crucial for the research supplier and the

client-candidate to be aware of the limitations of any technology used. Such an awareness can ensure the safe interpretation of any resulting data and enhance its usefulness in decision-making.

The purpose of this article is to examine the application of the Claritas Cluster system of geodemographics to political campaigns. After providing a brief description of the operation of the system, this analysis will follow the Claritas presentation through its major intellectual components, starting with an evaluation of the empirical evidence offered to verify the system's "magical" effect on a 1978 election, next considering the gathering and manipulation of the system's data inputs, and concluding with a discussion of the methodological underpinnings of the system.

Mark Atlas is a lawyer and researcher who has served on the staffs of several political campaigns.

As will be shown, the methodology underlying the system seems plausible enough to warrant further research to assess the procedure's validity. However, the methods for collecting and processing the data used, at least with respect to the 1978 "test" election, resulted in unreliable analyses. Furthermore, the figures provided allegedly proving the system's effectiveness actually show something quite different.

Overview of the Claritas Cluster system

(A brief recap of the relevant parts of the system's structure and operation is in order.) The system uses U.S. Census data as its initial information base. These data come aggregated into about 250,000 geographic units (called Block Groups/Enumeration Districts, or BG/EDs), each containing data on about 900 persons. Using factor analysis, a quantitative technique that combines a large number of individual variables into a smaller number of groups of those variables for easier analysis, 535 Census variables measuring the demographic characteristics of the people within each of the BG/EDs are divided into thirty-four groups (or factors) of composite variables. Then, using cluster analysis, a quantitative procedure that divides a large number of individual items (in this case, BG/EDs) into a smaller number of fairly homogeneous groups of those items, the 250,000 BG/EDs are divided into forty groups (or Clusters). Each BG/ED is mathematically assigned to a particular Cluster on the basis of that BG/ED's numerical values for the variables in the thirty-four factors previously generated. (The end result, then, is that within each Cluster are Census geographic units from across the country in which people with similar demographic characteristics reside.) ✓

Therefore, on the basis of distinct demographic profiles, each of these Clusters could be thought of as containing neighborhoods in different parts of the United States wherein residents have unique demographic characteristics and—according to Claritas' theory—

a unique lifestyle, set of attitudes, and so on. Furthermore, every person in the country is assigned to a Cluster, by determining which BG/ED he/she resides in (by simply finding the person's address on a map of all BG/EDs) and then to which specific Cluster that that particular BG/ED has been assigned as a result of the original cluster analysis. Thus, a user could select a particular state, county, or electoral precinct and determine what general types of people inhabit it by analyzing the Cluster composition of the area (i.e., what proportion of the population in the area is in BG/EDs included in Cluster 1, in Cluster 2, etc.). ✓

(A major purported benefit of this process is that one could do a sample survey of people in a state (or nation) and then project the results of this statewide survey down to smaller geographic areas within the state.) To do this one must assume that the universe of people actually residing in BG/EDs assigned to a particular Cluster would have responded identically to the survey questions as the sample of people who were actually interviewed during the survey. Thus, if a statewide survey showed that of those people actually interviewed who resided in BG/EDs assigned to Cluster 1 (assigned by the original cluster analysis done on all of the BG/EDs in the U.S.), 60 percent favored a particular ballot proposition, a BG/ED anywhere in the state assigned to Cluster 1 would be assumed to contain residents of whom 60 percent favored the proposition. Therefore, to project a statewide survey's results to a county, for example, one would first divide all of the statewide survey respondents into their respective Clusters, by using their addresses to locate the BG/EDs in which they reside and then determining the Clusters to which those BG/EDs are assigned. Next, one would calculate the overall response to each survey question within each Cluster, by simply cumulating the answers of all of the respondents who are members of a particular Cluster. Then one would merely assign each separate Cluster's sample survey responses to any of

the Cluster's BG/EDs within that county. Finally, one would add up the resulting figures for each of the county's BG/EDs (weighting each BG/ED's responses by its proportion of the county's population) to derive the overall county total. This procedure could then be followed anytime when one wished to project survey results from larger geographic areas to smaller ones.

(Obviously, any procedure that can accurately generate numbers detailing the opinions of small groups of people in relatively minute geographic areas by simply using a statewide survey, rather than accumulating a statistically acceptable sample from each area, has enormous potential. This is the potential that the Claritas Cluster system claims to tap, enabling campaigns to pinpoint very small population areas that, on the basis of *projected* survey results, seem receptive to campaign appeals.) By clustering and projecting, one could allegedly pinpoint small areas where, for example, voters are less aware of the candidate's existence, where the candidate has only lukewarm support, or where a particular issue position could be exploited for votes—all of which after interviewing people from different parts of the state but not necessarily anyone in the particular geographic areas selected for targeting.

(The ultimate question, of course, is whether the system as implemented actually delivers its promised benefits.) Most of the empirical evidence Claritas offers to prove the system's effectiveness is from the 1978 election battle in Missouri over a right-to-work ballot proposition. To briefly summarize Claritas' stated role in that campaign, the company was brought in by Matt Reese & Associates, a political campaign management firm, to attempt to target on a very selective basis the geographic areas where there were voters who were susceptible to appeals to vote against the right-to-work proposition. A statewide survey was thereafter completed in September 1978 and questions regarding voters' stand on the ballot issue were then projected from the statewide survey down to Missouri's BG/EDs,

in the manner described earlier. These projected numbers were then used to pinpoint areas containing voters worthy of intensive contact, and subsequently direct-mail and telephone and door-to-door canvassing were used to try to persuade those people to vote against the proposition.

Evaluating Claritas' effectiveness: impact on the election

One of the two general criteria that Claritas used to assess the effectiveness of its system in Missouri in 1978 was the system's impact on the election results. This impact was supposedly shown by a comparison between the statewide percentage of voters opposing the ballot issue before the Cluster system's use in the campaign and the final election results after its use. Claritas claims that this comparison proves the system's powerful effect on the election since surveys in February and July of 1978 showed the ballot issue winning by about 60 percent to 40 percent, while the final election results had the proposition losing 59 percent to 41 percent. Claritas claims this "stunning turnaround" was at least partly its doing.

(By examining Claritas' own data it is clear that, far from turning around the election, the system at best merely helped stabilize the status quo.) Regardless of what earlier surveys indicated, the September 1978 survey shows that the right-to-work proposition was *already behind* by about 56 percent to 44 percent—nearly the margin by which it eventually lost in November. The clearest proof of the advantageous position of the anti-right-to-work forces is provided by Exhibit 1, a table reproduced from a sales document. That table displays what the attitude toward right-to-work was within some of the Clusters and, most importantly, the overall statewide situation according to the September 1978 survey. As the table's first column shows, 40 percent of the sample *opposed* the ballot issue (27 percent strongly plus 13 percent moderately), 29 percent were mixed, and 32 percent favored

it (22 percent strongly plus 10 percent moderately)—far from the three to two margin in favor Claritas claimed existed before its system's use. In fact, if we then divide the 29 percent mixed in the same proportion as the opposed (40 percent) is to the favored (32 percent), the method used by Claritas in all of its predictions, the proposition loses 56 percent to 44 percent—strikingly similar to the final election result of 59 percent to 41 percent.

Another indicator of the strength of the early opposition to the ballot issue is provided in Exhibit 2, a graph reproduced from a Claritas sales document. The graph, again based on September 1978 survey data, plots each of the forty Claritas Clusters according to the percentage of the respondents in the Cluster that was opposed to the proposition (vertical axis) and the percentage that had mixed feelings about it (horizontal axis). Thus, for example, Cluster 29 is located at the intersection of 56 percent and 30 percent—56 percent of the respondents in this Cluster opposed the ballot issue, while 30 percent were mixed. Obviously, by subtracting each Clust-

er's percentages of opposed and mixed from 100 percent, the remaining portion is the percentage of each Cluster's respondents that favored the proposition. By carrying out this calculation for each of the Clusters, one can determine in how many Clusters and by what margins the forces opposed to right-to-work were ahead. The heavy line superimposed on the graph by the author (originating at the 50 percent point on the vertical axis and extending downward) serves as a dividing line showing each Cluster's overall position on the proposition—Clusters located above the line contain at least a plurality opposing the ballot issue, while Clusters below the line favor the ballot issue. Thus, as can be seen, right-to-work was opposed by voters in at least twenty-eight Clusters and favored by voters in only at most twelve Clusters, even before the system's "new magic" went to work.

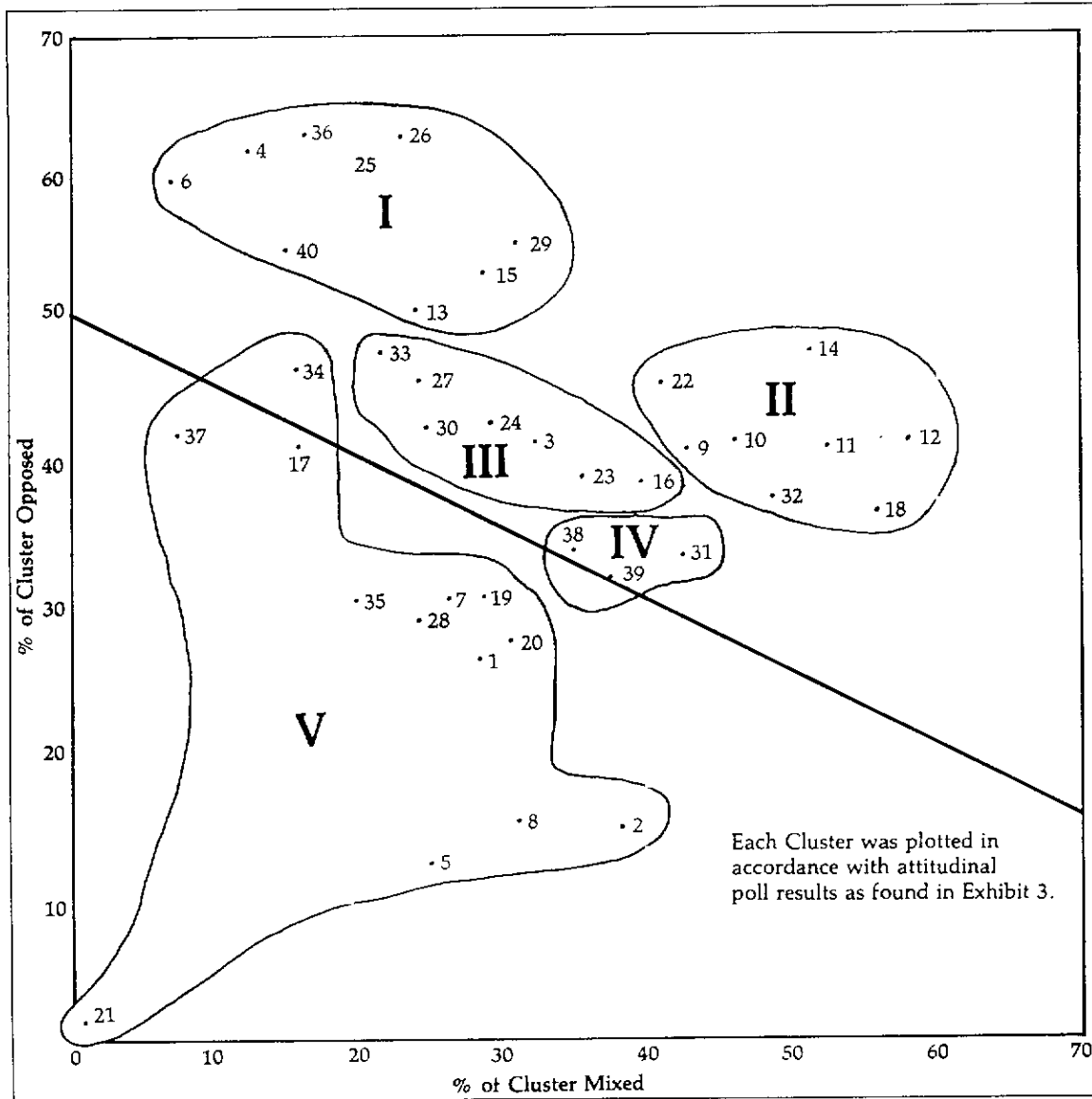
Thus, the empirical evidence offered by Claritas does not support its claim of providing a "stunning advantage" to its users in Missouri; at best, it merely avoided squandering a substantial lead. Comparing the final election results to surveys done early in 1978 rather

EXHIBIT 1 Poll Results on Right-to-Work Attitude by Cluster

	State-wide	Cluster 24	Cluster 25	Cluster 26	Cluster 27	Cluster 28	Cluster 29	Cluster 30
Total Sample Size	1,367	66	24	24	56	24	28	104
<i>Attitude on Right-to-Work Issue (Percent)</i>								
Strongly Oppose	27	33	53	52	33	9	52	40
Moderately Oppose	13	10	9	11	11	21	4	7
Mixed	29	29	19	23	24	25	30	24
Moderately Favor	10	7	9	5	3	21	2	9
Strongly Favor	22	21	10	8	30	25	12	19
Pro-Union	23	24	46	46	23	6	58	31
Anti-Union	16	15	10	4	14	26	8	19

The anti-right-to-work people selected Clusters 25, 26, and 29 targets. The Cluster system then provided the campaign with the names, addresses, and phone numbers of people living in those Clusters. As a result, when the campaign contacted these people, they reached households that had respectively an 81 percent, 87 percent, and 86 percent likelihood of not being in favor of right-to-work.

EXHIBIT 2 Targeting Chart of Attitude toward
Right-to-Work, Missouri, 1978, by Cluster*



- | | | | |
|-----|--------------------------------|----|----------------------------|
| V | Likely losers | II | Wide margin for persuasion |
| IV | Narrow margin for persuasion | I | Likely winners |
| III | Moderate margin for persuasion | | |

The Anti-right-to-work campaign targeted those people living in neighborhood/block groups in Group II as highly persuadable since those Clusters polled high on both opposition to right-to-work and mixed attitude toward (uncertain about) right-to-work.

*CLARITAS CORP. #5172, 9/28/78

than to surveys done immediately before the system was used in September serves only to mask the system's true impact. Why these early poll results differed so significantly from the September survey is beyond the scope of this article. Perhaps partial answers are data handling errors, changes in question-wording, or actual changes in people's opinions as time passed. The key finding, however, is that no sharp turnaround in voting behavior occurred after September that is disclosed by the data.

Evaluating Claritas' effectiveness: predictive ability

The second general criterion that Claritas uses to assess the effectiveness of its system is the accuracy of its projections of county election results. This was supposedly indicated by a comparison between the percentage of voters opposing the ballot issue in each county as projected from the September 1978 statewide survey and the actual final election results in each county. However, again by examining Claritas' own data, it is clear that the system did not produce exceptionally accurate estimates of county voting results. Claritas offers two items as its primary proof for the predictive ability of its system: (1) the system's ability to correctly predict which counties would have a majority oppose the ballot issue and which counties would have a majority favor it, and (2) the 0.628 correlation between the final election results and Claritas' projected vote totals for each of Missouri's counties (a correlation measures the degree of predictability that exists between two sets of data).

For the first point, Claritas claims that it correctly predicted which side would win the ballot issue in 83 of the 115 counties; however, this is an unnecessarily imprecise way of evaluating the county projections' accuracy. For example, if the actual election result for a county was 52 percent opposed to right-to-work, a Claritas projection of 80 percent opposed would thus be deemed more accurate than an estimate of 49 percent opposed, simply because the former projection was over 50

percent while the latter was under 50 percent, even though the 49 percent figure was much closer to the actual result. On the second point, the 0.628 correlation claimed by Claritas between the predicted and the actual county vote is not especially strong. By squaring this correlation, one obtains the "coefficient of determination," which is a measure of the percentage of the variance in the actual county election results that is explained by the Claritas projections and is, therefore, an indicator of the latter's predictive ability. By such a calculation, only 39 percent of the variance in the actual election returns is explained by the estimates—not a very substantial proportion.

Also, the whole process of producing the Claritas Clusters and projecting the results to individual counties seems hardly to be worthwhile from the viewpoint of enhancing predictive ability. If one calculates the absolute difference between the actual election results and the projected figures for each county, the mean prediction error over all counties was seven to eight percentage points. If one instead had foregone all of the clustering and projecting, but had simply predicted for ease and convenience that each county would have the same final election result as the September 1978 survey figure for the entire state (about 56 percent opposed), the mean prediction error for all counties would have been about nine to ten percentage points. Thus, all of the extensive processing techniques only raised the prediction accuracy by two percentage points, probably not enough to affect decision-making or to justify the additional expense.

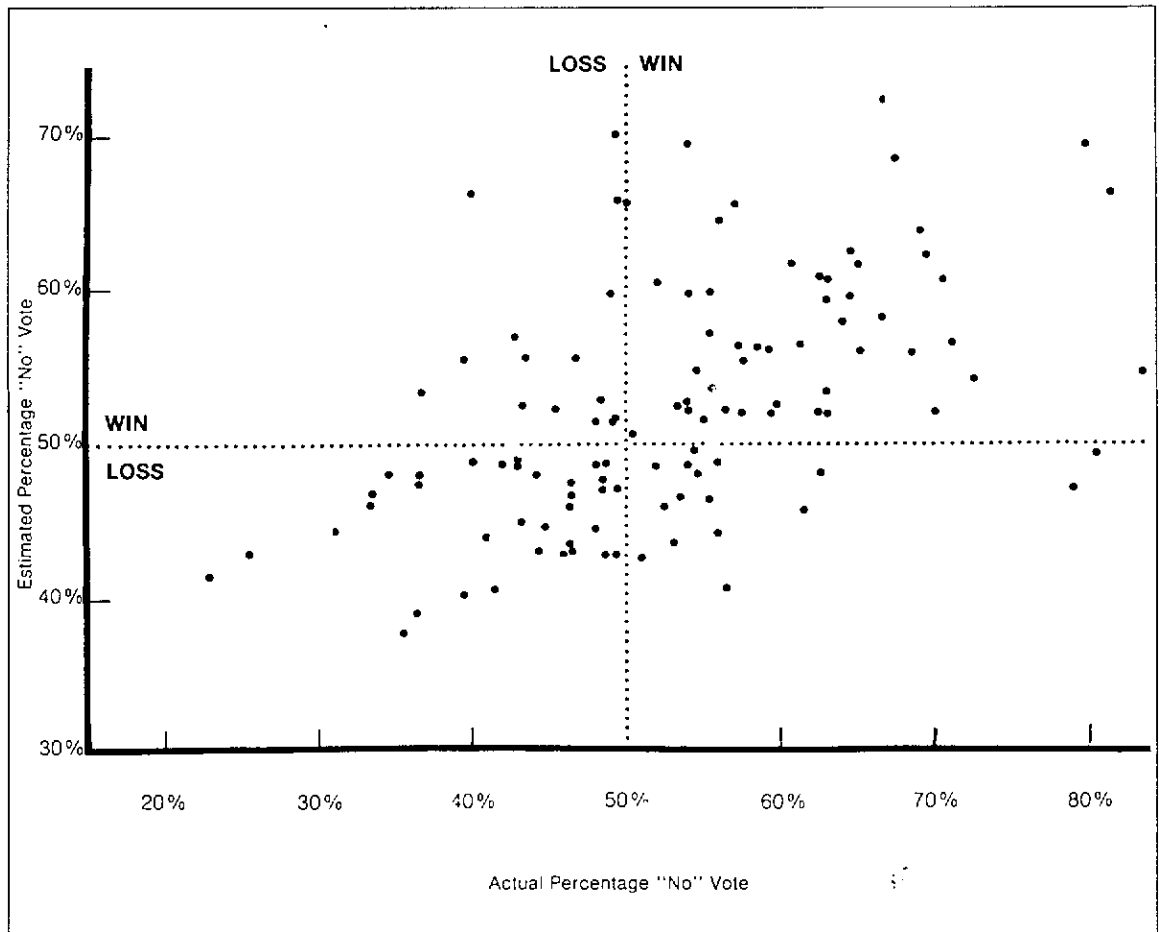
Another criterion against which to measure the validity of a predictive technique is whether its estimates contain systematic biases—or, in other words, whether the accuracy of the predictions vary in some sort of pattern. Generally, it is desirable that the size and direction of any errors in predictions be randomly distributed among the predictions rather than displaying any consistent trends. By examining Claritas' own data presented in Exhibit 3, a graph reproduced from the

previous article, it is obvious that very strong systematic biases occurred in its predictions. In the twenty-six counties in Missouri where opponents of right-to-work ultimately gathered under 45 percent of the vote, the Claritas predictions *overestimated* the vote percentages in twenty-four of those twenty-six counties (by a mean of about eleven percentage points). In the forty-seven counties where the opponents of the ballot issue finished with *over* 55 percent of the vote, the Claritas predictions *underestimated* the vote percentages in forty of those forty-seven counties (by a mean of about ten percentage points). Obviously, these are

extremely strong systematic biases indicating a flawed prediction technique. Only in those closely fought counties where the right-to-work opponents ultimately garnered 45 percent to 55 percent of the vote did Claritas' predictions split fairly equally between over- and underestimates.

Thus, on the whole, the empirical evidence offered by Claritas to prove its system's predictive capability shows instead that there exist within it strong biases, that the level of the statistical relationship between the predictions and the actual results is at best moderate, and that predictions of only slightly

EXHIBIT 3 Estimated Versus Actual "No" Vote for 115 Counties of Missouri on Amendment 23 ("Right to Work" Initiative), November 7, 1978



less accuracy could have been obtained without clustering and projecting the statewide survey results.

Data collection and decision-making

A major problem inherent in the manner of survey data collection and subsequent decision-making for the Claritas Cluster system is apparent in Exhibit 1. The first row of figures underneath the Cluster number headings shows the respondent sample sizes obtained for the Clusters and upon which the subsequent analyses and targeting decisions were based. The sample sizes shown are generally far too small to base reliable, quantitatively supported decisions upon. In fact, having forty Clusters to sample for and having interviewed only 1,367 people in Missouri, it is obvious that the mean number of respondents per Cluster was only thirty-four. Dealing with samples of this small size clearly crosses the line from quantitative to qualitative research. Statistically reliable results usable for decision-making are rarely attainable from just a few dozen respondents. For example, referring to Exhibit 1, Claritas selected Clusters 25, 26, and 29 as desirable because they "had respectively an 81 percent, 87 percent, and 86 percent likelihood of not being in favor of right-to-work." However, just due to normal sampling error, these numbers could be over a dozen percentage points too high (based on a 95 percent confidence interval) simply because there is a much greater likelihood of selecting an unrepresentative sample when only twenty-four or twenty-eight people are interviewed.

This finding has unsettling possibilities for the decision-making that was based on the Clusters' configurations shown in Exhibit 2. Since the Clusters that were targeted for intense contact were apparently selected on the basis of the numbers represented in the graph, it would have been important to keep in mind that due to normal sampling fluctuations a Cluster's position could actually be ten to twenty percentage points off (based on a 95

percent confidence interval) on the vertical and/or horizontal axis from where it would have been had an ideally representative survey been done. Thus, one should be extremely cautious when using such small sample studies for making decisions. What is doubly troubling is the notion that these possibly volatile small sample Cluster results were directly projected down to BG/EDs without identifying the potential error and were then used to predict county vote results. Even more troubling is the use of such small samples for designing advertising approaches adapted to each of the targeted Clusters—for something as complex as writing advertising appeals, samples of this size range are clearly inadequate and could yield very biased results.

(One of the critical problems with using the Cluster system in politics is the need for enormous survey samples and thereby substantial survey costs. A general rule of thumb is that to analyze any subsample one generally needs at least fifty respondents to be contained therein, though the potential statistical error involved would be quite substantial even with fifty. If one tried, though, to ensure that each of the forty Clusters would have even fifty people in it, a total sample of 2,000 would be required—as would a large budget to pay the bill. While using forty Clusters may have generated adequate Cluster samples when Claritas was dealing with national marketing surveys of at least 15,000 people, changes need to be made for the procedure to be used effectively in politics.) One obvious solution is reducing the number of Clusters by combining two or more very similar Clusters into new Clusters. While conceptually some degree of specificity would be lost, it seems better to have valid but more aggregated data than very finely differentiated data that cannot reliably be used for decision-making

Claritas' methodological foundations

The Cluster system's methodology primarily rests upon its use of cluster analysis, both as a means of processing the Census data to

generate distinct geodemographic segments and as a theoretical principle underlying the desirability of targeting on a multivariate basis. Unfortunately, very little information regarding the precise manner of clustering is provided in any of the system's literature; thus, evaluating the quality of the approach is difficult. Also, except for the 1978 Missouri election, there are few indications that any of the theoretical assumptions of the system have been tested, much less that any test results have been disclosed. Therefore, the following points address some of the issues surrounding the system's methodological foundations but generally will not—due to a lack of information—be able to decisively evaluate the quality of the analytic techniques and assumptions utilized.

One point brought out in Claritas documents that may give potential clients a false sense of security is the focus on the Cluster system explaining 87 percent of the demographic variance in the United States. This percentage refers not to the results of the cluster analysis that led to the forty Clusters, but rather to the factor analysis preceding it that combined the 535 demographic variables dividing a BG/ED into thirty-four composite groups. This is an important distinction because a "good" factor analysis result is a necessary, *but not sufficient*, indicator of a "good" cluster analysis since the latter segments the items through a different procedure and on different criteria. Also, the fact that 87 percent of the variance among demographic variables is explained is not surprising, since one would expect them to be strongly linked and since thirty-four factors—a fairly large number—are required to explain this percentage of the variance.

There are some additional aspects of cluster analysis that should be approached with caution, though it is unclear whether their associated problems affect the Cluster system. First, the number of clusters arrived at is fairly arbitrary—the user instructs the computer to divide the data into a designated number of

clusters and then the user interprets the output. By trying different numbers of clusters (Claritas tried variations from twenty to one hundred) the user can then select the results that seem most interpretable, sensible, and usable; there is, however, no magic formula for determining the appropriate number of clusters. Second, most, if not all, cluster analysis computer programs can generate substantially different results even if only the ordering of the data used in the analysis is altered. For example, if data are processed in a clustering program in one order and results are obtained, the mere changing of the order in which the data are entered into the computer can cause a substantially different set of clusters to emerge when the data are processed again. Thus, the stability of the Clusters is always a matter of concern. Finally, one would expect that at least somewhat different cluster compositions would result from analyzing national data as opposed to data from a particular state or locality. Thus, the forty Clusters derived from the entire nation's Census data may be very different from the Clusters that would be generated if each state's data had been clustered individually. The degree of similarity of such sets of Clusters would indicate the generalizability of the Claritas Clusters. Given these potential problems that might undermine the efficacy of the Cluster system, it would be very useful for the results of any tests of the system's stability, validity, and generalization potential to be disclosed so that potential clients can evaluate the relative quality of the service offered.

Conclusion

As this article has demonstrated, using the same criteria and data supplied by Claritas, it is clear that the claims made about the Cluster system are not supported, and possibly are contradicted, by the facts. If added to this lack of proven effectiveness are the inherent potential methodological difficulties associated with cluster analysis, it is obvious that

much more needs to be known and shown about the system before it can be used indiscriminately. The allocation of a campaign's resources is too crucial a task to be undertaken when there is substantial uncertainty about the targeting procedure's validity both in theory and in practice. Greater disclosure of the methodology used in clustering and of data from the system's use in other elections would help to either ease some doubts or clarify problems sufficiently to enhance their resolution. Also, it would be highly desirable to test whether projections from surveys done in larger geographic areas down to smaller areas accurately reflect the

latter's opinions. (This could be done by comparing the results of such projections to the results of adequately sized surveys actually carried out in the smaller area, at the same time asking some of the same questions.) Finally, assuming that the doctrine of caveat emptor holds true in elections, it is the responsibility of political campaigners to closely scrutinize systems offered for their use, to demand proof for any claims, and to expect to spend considerable time verifying such proof.

Otherwise, such campaigners should always expect free planes to be sent for them.