MEASURING MALAPPORTIONMENT*

GLENDON SCHUBERT AND CHARLES PRESS
Michigan State University

In addition to the legal and political implications of the case, the United States Supreme Court's decision in *Baker* v. *Carr*[1] brought sharply into public focus the technical problem of how to measure legislative malapportionment. The case itself exemplifies these various dimensions of the issue, for the majority opinion of Mr. Justice Clark and the dissenting opinion of Mr. Justice Harlan disagree not only regarding the law and the public policy of judicial intervention; they are also in manifest disagreement concerning "the facts" of malapportionment in Tennessee as of the time of the decision in March 1962. Our survey of the scholarly literature on the subject of apportionment, during the past decade, convinces us that the contributions of political scientists (and other commentators on the question) have made less than satisfactory progress thus far in the direction of devising an adequate metric to assist in the evaluation of what all concede is today a major problem in the theory and practice of democratic politics. In the absence of a reliable and valid method for measuring differences in apportionment along a common dimension, it is difficult to see how rational consideration of the normative aspects of the issue may be possible.

We are quite aware that the normative dispute hinges upon the question which factors *other than* population ought to be deemed relevant, and how much weight should be assigned to each of them. There appears, however, to be no substantial support for the proposition that population should not be considered a relevant factor in at least one house of all of the fifty state legislatures and of the Congress. Population is, therefore, the most common factor, and it cannot be ignored in either normative or empirical consideration of the apportionment of either the federal or the state legislatures. We propose, therefore, to confine our discussion to the measurement of apportionment in terms of this common factor; and the results should be of interest equally to those liberals who take their stand on the principle of "One Man—One Vote"[2] and those conservatives who favor the *status quo* and the representation of geography in addition to—or to some extent in lieu of—people.[3]

* For aid in the preparation of the empirical part of this paper, we wish to express our appreciation to our research assistant, Eric Carlson, doctoral candidate in political science at Michigan State University; to Richard M. Scammon, Director of the United States Bureau of the Census; to the All-University Research Fund, the Institute of Community Development, and the Computer Laboratory, all of Michigan State University; and to the many persons throughout the country who answered our correspondence and provided us with precinct maps, official descriptions from minutes of board meetings, and similar sources of raw data. We are particularly indebted to Francis M. Sim (who, during the formative period of our research, was Acting Director of the Bureau of Social and Political Research at Michigan State University) and to Professors Paul T. David and Ralph Eisenberg (both of the University of Virginia) for their comments upon earlier drafts of the theoretical and methodological parts of the paper.

In order to make our data available to other scholars, a set of our data, which include (as of March 26, 1962) population and representation totals for all legislative districts of both houses of the American national and fifty state legislatures, has been deposited with the Inter-University Consortium for Political Research, P. O. Box 1248, Ann Arbor, Michigan 41806.

[1] 369 U. S. 186 (March 26, 1962). For a summary of measures developed prior to this decision see extension of remarks of Senator Paul Douglas of Illinois, *Congressional Record*, Vol. 102, Part 4, 84th Cong., 2d sess., March 26, 1956, pp. 5536–48. For a similar discussion subsequent to the decision, see Arthur L. Goldberg, "The Statistics of Malapportionment," pp. 90–106 in "The Problem of Malapportionment: A Symposium on Baker *v.* Carr," *Yale Law Journal*, Vol. 72 (November, 1962). For an excellent article which focuses upon the normative question how simultaneously to maximize the three criteria (population equality, compactness, and contiguity) in order to produce optimally fair reapportionment, see James B. Weaver and Sidney W. Hess, "A Procedure for Nonpartisan Districting: Development of Computer Techniques," *Yale Law Journal*, Vol. 73 (December, 1963), pp. 288–308.

[2] Conference of Research Scholars and Political Scientists, *One Man—One Vote* (New York, Twentieth Century Fund, 1962).

[3] Alfred de Grazia, *Apportionment and Representative Government* (Washington, American Enterprise Institute for Public Policy Research, 1963).

I. CRITERIA

Before discussing any of the various measures of apportionment that have been proposed thus far, we shall first specify the criteria in terms of which we shall try to evaluate their adequacy. In the first place, we assume that what is to be measured is deviation from the ideal of equal representation for populations of equal size. Although we are aware that other definitions of "population" are not only possible but also the specified legal standard in some states, we mean here by population the gross number of human beings enumerated in the most recent decennial national census, irrespective of how many of them may be "legally qualified voters," "adults," "civilians," or otherwise members of restricted subsets of the general population. It is quite possible to use any of the measures that we shall discuss, and the one that we shall propose, to measure apportionment in relationship to whatever popular base may be posited. But it is desirable that we be explicit about the universe of raw data that we purport to observe.

In addition to observing general population differences among legislative representation districts, we must consider how many legislators are apportioned to each district; for as Klain demonstrated, in 1955 over 80 per cent of the states used multi-member districts in at least one house (as is also true, of course, of the United States Senate of the National Congress), and almost half of all seats in the lower houses of American state legislatures were apportioned to multi-member districts.[4] Our data show that multi-member districts are still widely used.

The precise boundaries of legislative districts must also be observed, in order to ascertain the subsets of the general population of any state that are represented by the legislator or legislators apportioned among the districts of either house of the state legislature. We have found that a common practice, in calculating urban district size, has been to divide county or city populations by the number of districts in the county or city. For a number of states we calculated some district populations from block census data. To secure maps, descriptions or published information, we wrote academic specialists, legislative service agencies, legislators, defeated candidates, city clerks, county clerks, county boards of supervisors, party headquarters, local newspapers and local librarians. Major problems occurred where apportionment was made by local agencies or where official descriptions used political boundaries such as wards, precincts, magisterial districts, etc., rather than county or township boundaries. However, we shall not be concerned in this article with gerrymandering, which is a question related to but quite independent of malapportionment. Gerrymandering consists of establishing district boundaries in relation to existing demographic patterns in such a way as to discriminate on the basis of some special parameter of the general population, such as political party, racial, or ethnic affiliation. Assuming equal proportions of Republican and Democratic voters among the general population of a state, a state legislature dominated by the Democratic Party—perhaps because of existing malapportionment—might redistrict the lower house in such a way as to concentrate large numbers of Republican voters in a minority of the districts, while at the same time establishing narrower but safe Democratic electoral majorities in a majority of the districts, or *vice versa* if Republicans were in control; and this might be done with strict adherence to the principle of equal *general* population subsets for all house districts in the state.

The raw data for the measurement of malapportionment consist, therefore, of the frequencies of general population and of representatives assigned to each legislative district, for both houses of a legislature. With such data in hand, one might entertain either of two somewhat divergent objectives in seeking to measure malapportionment. We shall call these Types I (intrastate) and II (interstate) objectives and indices.

The Type I objective might be to focus upon the relationships within a single jurisdiction (*viz.*, a state or the United States) in order to be able to specify the extent to which each resident of a subunit of government such as a city or a county (or, alternatively, all residents collectively of such a subunit) is over- or underrepresented in each house of the legislature or in both houses combined. The realization of this objective requires an index that measures for each subunit the deviation between the observed population/representation ratio, and the expected ratio under the assumption that each legislator represents an equal number of people. Such a measurement problem is a relatively

[4] Maurice Klain, "A New Look at the Constituencies: The Need for a Recount and a Reappraisal," this REVIEW, Vol. 49 (1955), pp. 1105–1119. Klain's findings were confirmed, and his data brought up to date, by Table 2 of David and Eisenberg, *State Legislative Redistricting: Major Issues in the Wake of Judicial Decision* (Chicago, Public Administration Service, 1962).

simple one, and an appropriate index requires little mathematical sophistication, either to derive or to apply. However—and we underscore this point because it seems not to have been noticed heretofore—such an index is of slight value for making *inter*state comparisons of malapportionment.

The reason why a Type I index has little utility for measuring Type II relationships is that for the latter objective, we require not a set of discrete observations of particularistic deviations from a state-wide norm, but rather a summary statistic which will characterize what is most important about the general pattern of malapportionment in each state. A Type I index which is a function of the unique general population total and of the disparate sizes of legislative chambers in each state cannot validly be employed to make interstate comparisons, because such an index may have the same value but different meaning for different states—or different values of the index may have the same meaning for different states. Alternatively, the simple ratio or percentage measures of deviation (which we shall describe below) sum to unity or zero for each state, and thus can hardly be used to make any meaningful statement about the degree of malapportionment in the state as a whole. There is also the question of bicameralism. This does not need to be resolved when the objective is intrastate measurement, because one can simply report the degree of malapportionment of subunits in each chamber. Some scholars, however, have attempted to devise a combined House and Senate measure of representation for each county. But the common assumption that, because equality of representation is posited as the ideal for fair apportionment, it is therefore appropriate to assume also that representation in the two houses of a legislature should be assigned equal weights, is even more dubious in the case of Type II indices. Our primary concern in the present article is to propose and to exemplify the use of a better Type II index, since this is the more interesting question from both a theoretical and a methodological point of view.

We would require two qualities in either type of index: objectivity, and comprehensiveness. By "objectivity," we mean that the method of employing the index be specified in operational form so that any researchers who possess some minimal level of professional competence ought to be able to observe the same raw data and report the same measurements on the basis of the index. By "comprehensiveness," we mean that the index can be used to measure the degree of malapportionment for all people living in districts represented in each house of any legislature.

In addition, we would make four other requirements of Type II indices: bicameral integration, interstate comparability, political realism, and statistical adequacy. By "bicameral integration," we mean that the index must lend itself to the computation of a single summary statistic which will describe the degree of malapportionment for any legislature. By "interstate comparability," we mean that such a single summary statistic will make it possible to scale all fifty state legislatures plus the Congress along a common metric which will measure the common dimension of representation of population that we have specified. An index that meets both of these requirements must be a pure number which is not affected by differences in the sizes of state populations or legislative chambers.

By "political realism," we require that the index should *not* be based upon arbitrary assumptions, which defy both our theoretical and empirical knowledge about the inequality of lower and upper chambers, or the way in which legislative majorities are formed, or the single and multi-member district systems. We reject as utterly unrealistic the legalistic assumption that lower chambers in American legislatures are either equal to or greater in power than upper chambers, because they are "closer to the people," or because "appropriation bills must originate in lower chambers," or because they are 'larger" than upper chambers. Typically, upper chambers are more important, in part because they share (as lower chambers generally do not) in the gubernatorial appointment power, thus exerting considerable influence over policy-making and implementation in the other-than-legislative governmental—including frequently, if indirectly, judicial—decision-making processes. A recognition of differential status is found in the thirty-three state and federal constitutional provisions that specify a longer term for members of one house: in all such cases, the longer term is for the upper house. Moreover, precisely because upper chambers are without exception smaller than lower chambers in American legislatures, senators typically represent (within any state) larger constituencies than members of the lower house, and their broader power base tends to assure that senators generally will be more influential politicians than members of the lower house. Therefore, we think that Type II indices ought to weight malapportionment in the upper chamber more heavily than malapportionment in the lower chamber, in most legislatures; and of course, in the absence of more precise empirical knowledge about the relative degrees of upper- and lower-chamber power in all states, it seems most reasonable at the present time to apply the same weighting formula to all states (and to the Congress).

As an example of an arbitrary assumption about the formation of legislative majorities, we can consider the Dauer-Kelsay[5] index, which is based upon a ranking of all districts (for a given house) from least to most populous, and then a summation of the population of the number of least populous districts equivalent to a minimal majority of representatives in the house. Such an index of the "minimum proportion of the state population necessary to control" the house in question is so patently unrelated to the relevant political characteristics of both district populations and their representatives that we must consider what significance can be attached to the numbers that it produces, for any particular set of empirical data. The assumption (which necessarily underlies the Dauer-Kelsay index) that the representatives from the districts with smallest populations—by no means necessarily rural districts—will frequently combine to vote in opposition to the remaining legislators who represent the districts with larger populations is contradicted by all of the recent empirical research in legislative voting behavior with which we are familiar.[6]

Political realism about the single-member district system requires that we adjust population differences among legislative districts according to the number of representatives apportioned to each district; we cannot, in other words, simply *assume* the prevalence of the single-member district system, and make comparisons between states on the exclusive basis of the (unweighted) populations of their legislative districts. Most students who recognize this problem have calculated persons represented by a representative in a multi-seat district by dividing the district population by the number of representatives in the district. We shall do the same, recognizing, however, that it is not politically realistic to treat the equal popular representation shares of the twenty-one members elected at large from one city within a single *county* (as was once the case in the Detroit area) as though they were the same as for representatives elected from single-member districts. Problems of representation raised by this type of gerrymander, devised to deprive sizeable popular minorities of their preferences in representation, are beyond the scope of this study.

By "statistical adequacy," we mean that a Type II index should describe the entire frequency distribution of representational-unit population differences, giving consideration not only to the location of the mean, but also to the general shape of the curve, the direction and degree of its skewness, and the extent of its kurtosis or "peakedness." (We shall discuss these technical considerations in greater detail, below.) Again let us take the Dauer-Kelsay index as an example. It does tell us something about the difference between the mean of the population of representatives, and the mean of the general population. If 20 per cent of the general population of a state elects 51 per cent of the representatives in the lower house, this informs us that one or more representation districts are considerably overpopulated (*i.e.*, underrepresented) in terms of the norm of equality of representation that we have posited. The norm would require that 51 per cent of the general state population be the minimax proportion which can elect 51 per cent of the representatives. We can infer that some other districts must, therefore, be overrepresented. But only the relationship between the two means is observed. Yet an infinite number of frequency distributions can have the same mean. The information that the Dauer-Kelsay index can convey about the nature of malapportionment in any house for any state is, therefore, slight. A Type II index satisfying the statistical requirements that we have specified would do a much better job of identifying patterns of malapportionment in the various states, according to what they have in common and the ways in which they differ.

II. TYPE I INDICES

The simplest kind of Type I index consists of a listing, in parallel columns, of the proportion

[5] Manning J. Dauer and Robert G. Kelsay, "Unrepresentative States," *National Municipal Review*, Vol. 44 (1955), pp. 515–575, 587.

[6] David A. Derge, "Metropolitan and Outstate Alignments in Illinois and Missouri Legislative Delegations," this REVIEW, Vol. 52 (1958), pp. 1051–1065; John C. Wahlke, Heinz Eulau, William Buchanan and LeRoy C. Ferguson, *The Legislative System* (New York, 1962); John C. Wahlke and Heinz Eulau (ed.), *Legislative Behavior* (New York, 1959); Gilbert Steiner and Samuel Gove, *Legislative Politics in Illinois* (Urbana, University of Illinois Press, 1960); Murray C. Havens, *City vs. Farm* (University, Alabama, Bureau of Public Administration, University of Alabama, 1957), 57 pp.; and Thomas A. Flinn, "The Outline of Ohio Politics," *Western Political Quarterly*, Vol. 13 (September, 1960), pp. 702–21.

of the population of the state living in each district (or other governmental subunit), paired with the proportion of representatives apportioned to that district or other subunit in relation to the size of the house. This is calculated separately for each house. (In order to speak more precisely about the various indices that we shall discuss, we shall define each in mathematical terms.) Let X equal the population of any district or other subunit, and Y equal the number of representatives apportioned to the district from either chamber of the legislature. Let N equal the number of districts (or other subunits) whose sum is equal to the total territory of the state, and let sigma (Σ) denote the summing of all the items in a set, such as the population of all districts for a chamber. The arithmetic average or mean population for all districts of such a chamber can then be denoted by the symbol M, and defined as follows: $M=\Sigma X/N$. The simple proportion of the population of each district to the total population of the state may then be defined either as $P_a=X/\Sigma X$ or as $P_a=X/NM$. Similarly, the simple proportion of representation for the district may be defined as: $R_a=Y/\Sigma Y$. Frequently such ratios are multiplied by 100, and thus reported as percentages. This is the type of measure reported, on a selective basis, in Baker's early study of the rural-urban dimension of malapportionment.[7]

A slightly more complex index is the proportion of the population of any district to the mean population: $P_b=(X/M)-1$. The reason for subtracting unity is that an index value of 0 will then signify the ideal of population equality, while all positive values will signify overpopulation (underrepresentation), and all negative values will signify underpopulation (overrepresentation). It should be noted that this index approaches the negative limit of -1, a value which—if empirically attainable—would indicate a district with no population and therefore, logically, a non-resident representative. On the other hand, there is no fixed limit to the positive value that the index might attain, since the most populous district in a state might be several times as large as the mean. To convert to a percentage measure, one merely uses the alternative formula: $P_{b'}=(100X/M)-100$. Tyler has attempted to make interstate comparisons, by reporting the $P_{b'}$ values for the most and least populous *congressional* districts in each of the states.[8] Of course, there is no reason why this index could not be used to measure *all* of the districts, for each house, of the legislature for any state: what it permits one to observe is the proportionate population deviation of each district from the mean population for all districts.

A very similar kind of index, which has attracted somewhat more attention, is that of David and Eisenberg. Actually, from a mathematical point of view, the David-Eisenberg index of "the relative value of the right to vote" (or, as it has been called, of "vote-values") is no more complicated than the Tyler index discussed in the preceding paragraph; for it is basically the inverse ratio, without the subtraction of unity in order to produce negative values: $P_c=M/X$. In the percentage form in which David and Eisenberg have reported their empirical data, and for class intervals of counties ranging in population size from i to j,

$$P_{c'} = \frac{100M \sum_{i-j} Y}{\sum_{i-j} X}.$$

As the average size of county populations *increases* (in relation to the general mean), $P_{c'}$ will approach zero, and all index values of less than 100 signify varying degrees of underrepresentation, while all values of greater than 100 signify increasing degrees of overrepresentation. A value of 100 represents, of course, an ideal apportionment. David and Eisenberg used this index primarily for *intrastate* measurement of (initially) population class intervals of counties, for which they made historical trend comparisons. Subsequently they reported discrete rather than grouped county data, on the same basis.[9] But they also attempted to make interstate comparisons on the basis of this index, both for the largest standard metropolitan areas[10] and for the largest and smallest class intervals of counties that they had defined.[11] For these purposes, they averaged the house and the senate index values, in order to compute a single index value for any category of counties in relation to the legislature of a given state. They also computed "national averages" of "the relative value of the right to vote," for each of their four popula-

[7] Gordon E. Baker, *Rural versus Urban Political Power* (New York, 1955), pp. 16–17.

[8] Gus Tyler, "Court versus Legislature; The Socio-Politics of Malapportionment," *Law and Contemporary Problems*, Vol. 27 (1962), p. 402.

[9] Paul T. David and Ralph Eisenberg, *Devaluation of the Urban and Suburban Vote* (Charlottesville, Virginia, Bureau of Public Administration, University of Virginia, Vol. 1: 1961 and Vol. 2: 1962).

[10] *Ibid.*, Vol. 1, pp. 12–13.

[11] *Ibid.*, Vol. 1, p. 15. This table incidentally, does not consistently compare the *same* class intervals.

tion class intervals for counties, by calculating the mean value, weighted by population, for the set of fifty state index-value averages. The national index values for the class intervals are thus averages of state index values, which are in turn averages of the separate index values for each house of the state legislature, which are based upon values for the groups of counties included within the class intervals. It is apparent that such averages cannot convey valid information about apportionment differences *among* the ninety-nine state legislative chambers. In any event, the focus of the David and Eisenberg national index values is upon comparing population categories of counties, rather than states; and their comparison of states is limited to the index values for the largest and smallest population category of counties found within each state.

A variant approach, which focuses upon an attempt to measure the overall differences in the malapportionment of counties within individual states, has been proposed by Clem.[12] Clem's paper is very much concerned with the methodological aspects of measuring malapportionment, but his primary interest is in developing a Type I index. His discussion includes an evaluation of the criteria for valid and reliable indices, a brief critique of alternative indices to the one which he proposes, and an empirical exemplification (with data from five selected states) of the use of his own index, which he called the "County Ratio" index. Apart from his formula, the antecedent procedures for processing the raw data are identical to those employed for the use of the David-Eisenberg index; in fact, he proposes as an alternative to the County Ratio Index another formula which *is* the David-Eisenberg index, expressed in ratio rather than percentage form. (Clem wrote after the publication of the first of the David and Eisenberg volumes, with which he was familiar, but prior to the publication of the second volume which reported their data for individual counties; Clem's recommended "alternative procedure" was in fact the David-Eisenberg index divided by 100.) The mathematical equivalence of Clem's alternative and the David-Eisenberg index is readily demonstrated by comparing their formulae; and for the empirically rather than theoretically inclined, such a demonstration is no doubt reinforced by the observation that the index values reported by Clem for selected counties in five states[13] are identical to one per cent of the "vote values" reported for these same counties in David and Eisenberg's second volume, which was published less than two months after Clem's paper was given.

Clem's County Ratio index is the percentage of the algebraic difference (*i.e.*, retaining plus and minus signs) between the ratio of the county population to the total state population, and the average of the county's proportionate representation for both houses. In cases where a county is part of a multi-county district for purposes of representation in either house, Clem assigns to each county a fraction of representation equivalent to the ratio of county population to district population. Hence, if Y_h equals the number of lower house seats assigned to the house district of which a given county is a part, and Y_s equals the number of upper house seats assigned to the senatorial district of which a given county is a part, and $i-j$ designates the counties in multi-county district, then

$$P_d' = 100\left\{\frac{X}{\sum X} - .5\left[\left(\frac{Y_h}{\sum Y_h}\right)\left(\frac{X}{\sum_{i-j} X}\right) + \left(\frac{Y_s}{\sum Y_s}\right)\left(\frac{X}{\sum_{i-j} X}\right)\right]\right\}.$$

An index value of zero indicates the ideal of exact equivalence between county population and representation; negative values denote underrepresentation and positive values denote overrepresentation. There are no fixed limits to the values attainable by the index, in either direction. It should also be observed that basically, this is a very simple formula, for it is essentially:

$$P_d = P_a - R_a, \quad \text{or} \quad P_d = \frac{X}{\sum X} - \frac{Y}{\sum Y}.$$

Clem suggests the conversion of his Type I index into a Type II index by dividing the total number of counties in a given state into the number of counties whose over- or under-representation falls within proposed limits of maximum allowable deviation. The difficulty with this measure is that county units are treated as equal. For example, Michigan has 31 of its 83 counties outside the limits of maximum

[12] Alan L. Clem, "Legislative Malapportionment and the Mathematical Quagmire," a paper presented at the annual meeting of the Midwest Conference of Political Scientists, held at the University of Notre Dame, South Bend, Indiana, April 27, 1962; mimeo., pp. 1–33. See also by the same author, "Measuring Legislative Malapportionment: In Search of a Better Yardstick," *Midwest Journal of Political Science*, Vol. 7 (May, 1963), pp. 125–44.

[13] *Ibid.*, Appendix Table G.

allowable deviation. Four of these 31 contain 52.88 per cent of the state's population. Treating counties as if they were of equal importance seems to us to lead to gross distortion of the very factors we are attempting to measure—the degree to which population is the basis of representation.

As a final example of Type I indices, we might consider the formula suggested by Mr. Justice Clark, in his concurring opinion in *Baker* v. *Carr*, together with the modification for this formula advocated by Mr. Justice Harlan in dissent in the same decision.[14] Clark purported to evaluate differences in the apportionment of Tennessee legislators on the basis of what he described as an index of "total representation." Under the then existing Tennessee apportionment system, a county might be assigned both exclusive representation and a share of a representative as part of a multi-county district (called floterial seats) for either the lower house or the senate of the state legislature, or for both houses. There were 99 members of the lower house and 33 senators; and Clark proposed that a county's apportionment be evaluated as the *sum* of its shares of representation for both houses, the latter being defined as the integer or fraction of representatives assigned weighted by the size of the houses. The difference between Clark and Harlan related to the calculation of county representation shares for the multi-county districts; Clark assigned shares of representation proportionate to the number of counties in the same district, while Harlan (like Clem) assigned proportionate shares according to population in each county of a multi-county district.[15] To put this algebraically we shall need, in addition to the symbols previously used, some further terms defined as follows: let Y_a = the number of seats assigned exclusively to a county, and let Y_b = the number of seats assigned to a multi-county district, while N_h = the number of counties in a multi-county house district, and N_s = the number of counties in a multi-county senate district. Then Clarks' formula is:

$$R_b = \left(Y_{ah} + \frac{Y_{bh}}{N_h}\right) + 3\left(Y_{as} + \frac{Y_{bs}}{N_s}\right),$$

while the formula with Harlan's adjustment becomes:

$$R_{\underline{b}} = \left[Y_{ah} + (Y_{bh})\left(\frac{X}{\sum_{i-j} X}\right)\right] + 3\left[Y_{as} + (Y_{bs})\left(\frac{X}{\sum_{i-j} X}\right)\right].$$

An index value of 2 would signify—and according to either formula, since their difference lies in conceptualizing how fractional shares of representation *ought* to be weighed—that a county had joint representation in both houses equal to the average representation in each house. (Of course, such an index value could be produced by a combination of underrepresentation in one house and overrepresentation in the other house, as well as by an average representation in each house.) The index was used by Clark in the following way: all counties were listed according to the rank order of their populations, and the index values of "total representation" were reported in a parallel column. His assumption was that if the apportionment pattern were "rational," there would be a high positive rank correlation between the sizes of county populations and index values, at least when counties were partitioned into rural and urban subsets. Mr. Justice Clark did not compute rank correlation coefficients to make the comparison upon which, as he said, he based his judgment; instead, he relied upon "common sense" inferences based upon observations of selected portions of the data.

The seeming complexities of the Clark and Harlan formulae should not be permitted to

[14] 369 U. S. 186, 255n. 7, 262–264, 342–343.

[15] Clark's procedure for attributing representation in multi-county districts seems to be similar to that followed by David and Eisenberg, who apparently computed their index on the basis of an assignment of the *total* district population to each of the included counties; Clark, however, assigned equal shares of the total district population to each included county. Evidently, David and Eisenberg's procedure involves even more distortion than Clark's, as we shall exemplify with Michigan data. The 14th Michigan Senatorial District consists of two counties, Ingham and Livingston, whose populations are 211,296 and 38,233. David and Eisenberg assigned a vote value of 92 to each of these counties for the upper house of the Michigan legislature (*op. cit.* ftn. 9, *supra*, Vol. 2, pp. 76–77); using his formula for Tennessee, whose chambers of 33 and 99 are about the same as Michigan's 34 and 110, Clark would assign an index value of 1.5 to each county, while Harlan would assign 2.54 to Ingham and only 0.46 to Livingston. Moreover, and because of the not inconsiderable population differences, David and Eisenberg assign these two counties—with identical vote values for the senate—to different class intervals for purpose of their grouped county data analyses: Ingham is in the "100,000–499,999" category, while Livingston is in the "25,000–99,999" category.

obscure the fact that the underlying relationship that both these Supreme Court justices purported to observe is really the simplest of all of those we have considered thus far: $R_b = Y$. However, the other Type I indices measure underlying relationshps that are not much more complicated: Baker measures $X/\Sigma X$ and $Y/\Sigma Y$ while Clem measures their difference; and Tyler measures X/M while David and Eisenberg measure M/X. In other words, the Supreme Court justices simply count how many representatives are apportioned to each county; Baker and Clem compare the proportions of population and representation for the units of government that they observe; while Tyler, and David and Eisenberg, compute ratios of unit populations and population means. Questions of sampling procedure aside, all of these Type I indices are capable of being used objectively, and comprehensively for all of the districts or units of government of a state; so they meet the criteria that we specified for this type of index. None of them, however, are suitable for measuring interstate relationships. They all fail to meet two or more of the four criteria that we have posited for Type II indices, to a consideration of which we now turn.

III. TYPE II INDICES

We are familiar with two Type II indices that others have proposed, and we shall suggest a third. Tyler has reported extremity ratio measurements for each house in all of the states, and for congressional districts in each of most of the states;[16] Tyler, Dauer and Kelsay, and the National Municipal League have computed an index which measures the least population which can elect a majority of the members of a state legislative chamber;[17] and the index that we propose is a function of the coefficient of variation, which is a standard measure in parametric statistics. We shall first explain each of these, and then we shall evaluate them in terms of our criteria for Type II indices.

The formula for Tyler's extremity ratio is: $TE = X_{\max}/X_{\min}$; that is, the population of the most populous district divided by the population of the least populous district of the same house in any given state. The minimal value of the index is 1.0, signifying that all districts are equal in size; and there is no fixed limit to the maximal value that the index can attain—Tyler reported maximal values of 1081.3 for the lower house in New Hampshire, and of 4.5 for congressional districts in Michigan. The defects of the index, as the basis for making interstate comparisons of malapportionment, are readily apparent. The only one of our six criteria which it satisfies is the first, objectivity; it is clearly not a comprehensive measure for all districts in a chamber; it does not permit bicameral integration; it should not be used for interstate comparability, because the ratio is meaningful only in relationship to the particular chamber of the particular state from which the sample of two districts is drawn; it is not politically realistic; and as for statistical comprehensiveness, it provides no information about the mean, the variability of the distribution, skewness, or kurtosis. All that it does is to compare the most extreme cases in the two tails of the distribution.[18] It is a dramatic but essentially vacuous measure.

The Dauer-Kelsay index is computed as follows: all districts of a legislative chamber are listed in the rank order of their populations,

[16] *Op. cit.* ftn. 8, *supra*, pp. 393, 402.

[17] *Ibid.*, p. 391; *op. cit.* ftn. 5, *supra;* and William J. D. Boyd (ed.), *Compendium on Legislative Apportionment* (New York, National Municipal League, 1962 ed.), pp. iii–iv.

[18] Tyler's extremity ratio is only one step of complexity—since division is a slightly more complex mathematical operation than subtraction—removed from the *range*, the measure of dispersion deemed most appropriate (to characterize population differences among congressional representation districts in the states) by Mr. Justice Harlan, in the "Appendix" to his dissenting opinion in *Wesberry* v. *Sanders*, 32 L. W. 4142 at 4156–4157 (February 17, 1964). A typical sample of statisticians' comments upon the adequacy of the range as an index of variance follows; and these comments appear equally applicable to Tyler's ratio. Robert E. Chaddock, *Principles and Methods of Statistics* (Boston, 1925), p. 153: "*While this is the simplest measure of variability it is also the least informing.* It gives no idea of the nature of the distribution within these extreme limits. It is very unstable since by cutting off a single . . . item at either end of the scale or adding one the range may be entirely changed. It fails to characterize in a useful manner the series as a whole if stated alone, and ignores the degree of concentration almost entirely. It offers no basis for judging the typical character of the average itself." [Italics in the original.] Quinn McNemar, *Psychological Statistics* (New York, 2d ed. 1955), p. 19: "One may doubt whether the *range* (highest to lowest score) is of sufficient value in psychological research to justify its use as a measure of variation. It is, obviously, determined by the location of just 2 individual measures or scores and consequently tells us nothing about the general clustering of the scores about a central value."

from least to most populous, and the representation apportioned to each district is listed in a parallel column. If the size of the chamber is odd, the median item in the representation column is selected; otherwise, the $(N/2)+1$th item: this step identifies the district which elects the marginal member of the minimal majority of the chamber. The column of populations is then summed for the set of districts consisting of the first through the marginal member's district; that sum is then divided by the total state population, and the quotient is multiplied by a hundred in order to express it as a percentage. If j is defined as the marginal district with rank r in the second column such that

$$\sum_{i-j}^{\min} r_y > \frac{Y}{2}, \qquad \text{then } DK = \frac{100 \sum_{i-j} X}{\sum X},$$

which is the minimal population equivalent to a majority of representatives. The theoretical limits of the values attainable by this index are zero (signifying that a majority of the districts have no population) and >50 (indicating that all districts are of equal population). If legislative chambers were of infinite size, the maximum value would approach the limit of 50.00; but with the relatively small chambers that exist in American state legislatures, the theoretical maximal value is a function of the size of the chamber. For example, it is 50.25 for the lower house in New Hamsphire, which has 400 members; and it is 55.00 for the upper house in Alaska, which has 20 members—both Delaware and Nevada have smaller senates of 17 members, but the maximal value for these is only 52.94 because the N is odd. For the same reason, the maximal value for the lower house in Georgia in 1962 was 50.24, slightly less than that for New Hampshire, even though the Georgia lower house had only about half as many members (205). For the United States Congress, the maximal values are 50.11 for the House of Representatives and 50.50 for the Senate.

David and Eisenberg, on the basis of an articulate assumption about the substantial equality, in general, between the lower and upper chambers in bicameral legislatures,[19] have suggested that the index values for both houses of any state legislature simply be summed to produce a single value to measure the degree of malapportionment for the state. (It should be noted that Dauer and Kelsay also considered such a combined index but concluded that it probably would have validity only for states at the top and bottom of the scale.) The theoretical range of such a composite index would be from 0 to >100; for independent observations in the three years 1937, 1955, and 1962, the empirical minimal value was 24.0 for Florida in 1962, while the empirical maximal value was 100.8 for New Hampshire in 1937. In the case of Nebraska, the value for the unicameral legislative chamber was doubled. Apart from the question whether lower and upper chambers *should* be weighted equally—a point that we already have disputed—it is apparent that the summation scores may reflect very different kinds of malapportionment in individual states. Consider, for example, the 1962 chamber and composite Dauer-Kelsay values for the following states:

State	Lower House	Upper House	L+U	L−U
Louisiana	34.1	33.0	67.1	1.1
Washington	35.3	33.9	69.2	1.4
South Carolina	46.0	23.3	69.3	22.7
Missouri	20.3	47.7	68.0	−27.4

It is misleading to say that South Carolina and Missouri are equally malapportioned, or that the degree of malapportionment in either of these states is about the same as that in Louisiana or Washington. When one chamber is grossly malapportioned, it can function as a perpetual block to political changes favored by a majority of the population of the state, as the situation in Michigan throughout the decade of the 1950's exemplifies; the 1962 Dauer-Kelsay chamber values for Michigan are 44 and 29. Moreover, we are confident that gross malapportionment of the senate creates more blocking power than gross malapportionment of the house, and we expect that studies of political change, comparing states like Michigan and South Carolina with states like Missouri, would provide empirical support for this hypothesis.[20] Of course, the preceding discus-

[19] *Op. cit.* ftn. 9, *supra*, Vol. 1, p. 6.

[20] We have considered the possibility that one might rank the states according to the lowest Dauer-Kelsay index score observed for *either* chamber in each state. It might be assumed that this would rank states according to the maximal blocking power of minorities. Such an assumption is questionable on the grounds of political realism, because it omits one house from consideration. A state in which D-K indices for both houses are 17.0 and 17.0 certainly is different politically from one in which the two houses are 17.0 and 47.0, respectively. In the second state one would expect the activities of the malapportioned house to be spotlighted. We are not aware that any comparative empirical studies have been made of such

sion has assumed the validity of the Dauer-Kelsay index values for individual chambers, which is a question that we must now examine.

The Dauer-Kelsay index is objective, and it meets the criteria of comprehensiveness, bicameral comparability, and interstate comparability. But it fails to meet our criterion of political realism for two reasons. We have no empirical knowledge which suggests that legislative majorities typically—or, indeed, ever—combine in the manner specified by the D-K index. To the contrary, the available studies of legislative voting behavior indicate that the *ad hoc* majorities that typically form are complex functions of such variables as political parties, general interest group activity, specific constituency interests, and regional (sub-cultural) differences.[21] Urban as well as rural districts can be considerably less populous than the state mean for district populations. In Michigan, for example, and for both the House and Senate, most of the largest and several of the smallest districts (including the smallest) are Democratic, while Republican districts (which include the most populous Senate district) also range from gross under- to gross over-population (in relationship to the mean). So what the D-K index measures has little relation to the empirical realities of legislative voting behavior. Moreover, the imputed equality of power for lower and upper chambers is not, in our opinion, empirically justifiable, for reasons that we already have explained. From a statistical point of view, the D-K index identifies a cutting point somewhere on the curve for the frequency distribution of district population sizes for a chamber. We have already questioned the empirical significance of the identification of such a point; we now question its theoretical significance. The D-K index provides no information about the population mean, or the variability of the curve, or its skewness or kurtosis. If the D-K index value *were* compared with the mean value for such a curve, then it might provide a rough measure of skewness; but there are better ways than this to measure skewness.

Our own approach is premised upon the assumption that since the basic problem of measuring malapportionment in state legislatures is a statistical problem, political scientists and lawyers and other social scientists who are concerned with this problem would be well advised to consider the use of standard statistical measures before trying to devise *ad hoc* indices. There has been an immense accumulation of experience in the use of most of the standard parametric measures, and their properties have been studied by mathematicians and statisticians who, generally speaking at least, are expert in such matters to a degree far beyond what can be the reasonable aspirations of most political scientists and lawyers. Moreover, other social sciences—such as psychology and sociology—long since have based much of their empirical research upon the assumption that the standard measures of both parametric and non-parametric statistics will be invoked, where appropriate, as a matter of course. Perhaps one reason why wider use of such methods has not been made by political scientists and lawyers in their attempts to measure apportionment relates to sub-cultural factors such as the traditions and methods of graduate training for these professions.

David and Eisenberg commented in a footnote[22] that:

> Statistical purists will recognize that other methods of categorizing the counties might have been possible, based on average county size and a computed standard deviation for each year studied. Unfortunately, the results of an analysis based on this kind of an approach would have been relatively inaccessible to all except professional statisticians—not to mention the increases in computation time and labor that would have been involved.

We think, to the contrary, that it is worth the extra time and labor and costs necessary to undertake an adequate measurement of state malapportionment. As we shall discuss in greater detail below, the really costly aspect of empirical research in this subject is the collection of the raw data; the difference in the costs for the computation of precise and crude indices is, by comparison, minimal. (Our computer time, for the basic calculations of the indices reported in Table 2, *infra*, was three minutes.) We also reject the conclusion that only professional statisticians can understand

situations. It is reasonable to hypothesize, however, that the combined pressures of a governor and a well-apportioned house might force political concessions from the malapportioned chamber. Where both houses are badly malapportioned, we hypothesize that still a different political situation would exist.

Cf. the discussion of the blocking power of legislative chambers, in L. S. Shapley and Martin Shubik, "A Method for Evaluating the Distribution of Power in a Committee System," this REVIEW, Vol. 48 (1954), pp. 787–92.

[21] See note 6 *supra*.

[22] *Op cit.* ftn. 9, *supra*, vol. 1, p. 10n. 9.

such matters as chi squares and standard deviations. To the extent that readers of this article do not employ elementary statistical concepts as part of their workaday professional tools, only a slight effort will be required to follow the discussion below in which we shall explain our Inverse Coefficient of Variation (ICV), which combines such common statistical measures as the mean and standard deviation.

IV. THE INVERSE COEFFICIENT OF VARIATION; SKEWNESS; AND KURTOSIS

ICV also is a Type II index. It is computed as follows:

1. List the population for each district of a state legislative chamber, in the sequence of the assigned district numbers.
2. Partition the population of multi-representative districts into representational units of equal size, by dividing the district population by the number of district representatives. (Each single-member district has its population divided by one, and therefore becomes a representational unit.)
3. List all representational units in rank order from least to most populous. (This step can be omitted for computer processing, for which special data coding and format will be required in any event; but it is a convenience when operations are to be performed with a desk calculator.)
4. Let X, M, and N have the same meanings as in the previous examples; *i.e.*, X = the population of any representational unit; N = the total number of representational units = the size of the chamber; and M equals the mean, or $M = \Sigma X/N$.
5. Let $x = X - M$, the algebraic difference between the population of any district and the mean population; list these x values in a column parallel to the X column; and calculate x^2, x^3, and x^4 for each representational unit and list these values in subsequent parallel columns. (To calculate x^3 for any representational unit, multiply x and x^2, and to calculate x^4, square the x^2 value.)
6. Sum the x, x^2, x^3, and x^4 columns. (The x column should sum to zero, within rounding error. The x^2 and x^4 sums must be positive—negative sums are impossible; a sum of zero for these columns would indicate absolute population equality among all representational units of the chamber. The x^3 sum may be either positive or negative, and will be zero when the distribution is symmetric.)
7. Calculate the standard deviation (sigma):

$$\sigma = \sqrt{\frac{\sum x^2}{N}}$$

8. Calculate the inverse coefficient of variation:

$$\text{ICV} = \frac{1}{1 + \frac{\sigma}{M}}$$

9. Calculate the index of skewness (g_1), so that

$$g_1 = \frac{\sum x^3}{\delta \sum x^2}$$

10. Calculate g_2, the index of kurtosis (*i.e.*, the degree of peakedness of the distribution):

$$g_2 = \frac{N \sum x^4}{(\sum x^2)^2} - 3$$

The mean (M) defines the average population for all representational units, given the total population of the state and the size of the legislative chamber. Under our assumption that equality of population among all representational units is the norm that provides our criterion, then the mean represents the ideal apportionment, deviations from which signify varying degrees of malapportionment. The greater the deviation from the mean, the greater is the malapportionment in regard to any particular representational unit. Obviously, units that are underpopulated will have negative deviations from the mean, indicating that they are overrepresented; while units that are overpopulated will have positive deviations from the mean, indicating that they are underrepresented.

The x scores provide an absolute measure of variation, but their direct use as a measure of malapportionment has certain obvious disadvantages—the fact that their sum is zero, for example. As Guilford notes,[23]

> The standard deviation, or σ, is the most commonly used indicator of degree of variability, and . . . it is usually the most reliable. That is, it varies least from sample to sample drawn at random from the same population. It is therefore more dependable and, as an estimate of the dispersion of the population, it is more accurate.

From one point of view, of course, we are measuring the entire population, in a statistical sense of that word, but it seems more realistic to conceive of our operation as a kind of sampling, since the census data are themselves really an estimate of the true population of any state or district thereof, and even these estimates (in the form of the decennial "enumeration" data commonly employed in malapportionment studies) are static while population

[23] Joy Paul Guilford, *Fundamental Statistics in Psychology and Education* (New York, 1956, 3d ed.), p. 85.

growth is dynamic. In any event, we require the sums of the second, third and fourth powers of the deviations in order to relate the variance to the mean in a summary statistic which will measure the entire series of representational unit deviations, and also in order to measure skewness and kurtosis. We divide sigma (the standard deviation) by the mean in order to compute the coefficient of variation because "Measures of variability are not directly comparable unless they are based upon the same scale of measurement with the same unit."[24]

We are justified in using CV with malapportionment data, because our measuring scale is one of equal units (the individual human beings who comprise a state, district, or representational-unit population are given equal weight in the census count) and it has an absolute zero point: no population (one New Hampshire town, which constituted a representational unit—indeed, a district—for the lower house in 1962, had a population of 3). The formula for the direct coefficient of variation is: $CV=\sigma/M$. Since sigma is a general measure of the variance, while M defines the apportionment ideal, the smaller the sigma, the better the apportionment; and the larger the sigma, in relation (of course) to its own mean, the worse the malapportionment. The effect of dividing sigma by the mean is to produce a pure number which is directly comparable with other CV's, irrespective of the differences in the sizes of the state populations or legislative chambers upon which they are based.[25] Therefore, CV provides a measure which will permit us to compare the two houses of the same legislature, or houses of legislatures in different states. For a perfectly apportioned state chamber, the value of CV would be zero, signifying no deviance. But for a state whose chamber was badly malapportioned, the value of CV could exceed 1—in fact, the value of CV for the United States Senate in 1962, using 1960 census data, is 1.04—because sigma can be greater than the mean. CV has the slight disadvantage, from the point of view of interpretation in relation to common-sense intuitions, that low values of CV indicate good apportionment, while high values indicate malapportionment. Therefore, it seems desirable to invert the CV index, so that it will range only between 0 and +1.00 and so that low values will indicate malapportionment (when our criterion is the single dimension of variance) while a value of +1.00 will indicate perfect apportionment. This is precisely what the ICV does, and this is the interpretation that we associate with ICV scores.

Even with a normal curve,[26] there could be discrimination in favor of a few and against a large number of persons, due to variance. Let us assume a hypothetical distribution for which the ordinate is a measure of the frequency, ranging from zero upwards, of representational units of any given size; and the abscissal axis is a metric, ranging from left to right, of population size of units. Since the curve for this distribution is normal, it is mesokurtic (*i.e.*, it shows neither positive nor negative kurtosis) and it is not skewed; and about a third of the units lie within a range of one standard deviation from either side of the mean.[27] Since

[24] *Ibid.*, p. 101.

[25] *Cf.* McNemar, *op. cit.* ftn. 18, *supra*, p. 29: "the ratios [for g_1 and g_2] are pure numbers, *i.e.*, are not inches or pounds or IQ's or minutes. If we have the distribution of the weights and of the heights for 1000 individuals, the measure of skewness for the height distribution may be compared directly with that for the weight distribution. This is true by virtue of the fact that for each we are expressing *the third moment relative to the* amount of variability, both in inches for one distribution, both in pounds for the other. Likewise, it can be reasoned that the measures of kurtosis for different distributions are comparable, although the distributions involve different measurement units."

[26] For good introductory discussions of the normal curve, see Margaret J. Hagood and Daniel O. Price, *Statistics for Sociologists* (New York, rev. ed., 1952), ch. 14; or Helen M. Walker and Joseph Lev, *Elementary Statistical Methods* (New York, rev. ed., 1958), ch. 12.

[27] In theory, the tails of a normal curve extend to infinity, but this assumes (in terms of our problem) a legislative chamber that consists of an infinitely large number of members. In practice, the range of empirical approximations of the normal curve is limited by the size of the chamber, and with an N of 400 to 500 usually will extend over ± 3 standard deviations; with an N of 100, the mean range drops to about ± 2.5 standard deviations; for 67 it is about $\pm 2.4\ \sigma$; and with an N of 35 the mean range is only about ± 2.1 standard deviations, while with an N of 17 the range averages about $\pm 1.8\ \sigma$. See George W. Snedecor, *Statistical Methods: Applied to Experiments in Agriculture and Biology* (Ames, Iowa, The Collegiate Press, 4th ed., 1946), p. 98. For our own empirical analysis, based upon the sizes in March 1962, none of the "lower" chambers of state legislatures was smaller than 35, although 17 of the state senates included between 17 and 33 members. Conversely, none of the "upper" chambers was larger than 67, although three-fourths (37) of

our ideal of absolute equality of population among representational units has no possibility of realization in the empirical world, there will always be some variance around the mean of any empirical distribution of units; and there might even be substantial agreement that a normal curve with a CV of .20 (and therefore an ICV of .83) is the model[28] for the closest approximation of the ideal that offers any realistic possibility of being attained for any American legislature, at least for some time to come. Even so, up to a sixth of the representational units, under a normal curve, will fall in the left (or negative) tail of the distribution: these will be thinly populated units which are *over*represented in the hypothetical legislative chamber; while another sixth of the units will fall in the right (or positive) tail, and these will be densely populated units which are *under*-represented. Since many more persons reside in the units that lie in the right tail, as compared to the units that fall in the left tail, the discrimination against the overpopulated units that are underrepresented is—in terms of our postulated ideal of equality of popular representational shares—a much more serious form of malapportionment than the discrimination in favor of the underpopulated units of the left tail which are overrepresented. The many persons who receive less than equal representation constitute a much more important deviation from the ideal than do the relatively few persons who get much more than equality of representation. These are characteristics, we repeat, of a *normal curve,* which has relatively little to moderate variance, no skewness, and moderate kurtosis. Our point is that there will be a not insubstantial amount of both kinds of discrimination under the criterion curve which will function, in the following evaluation of empirical data for American legislatures, as our model of a *relatively* well-apportioned chamber. Evidently, to the extent that the coefficient of variation (the proportion of the standard deviation to the mean) exceeds about .67, one or the other or both kinds of discrimination must also be greater than that which obtains with a normal curve.[29] There is, however, a better way than the measurement of relative variance to test for normality in the form of distributions: the use of the indices of skewness and of kurtosis.[30] The skewness index will inform us which kind of discrimination is increased, and the kurtosis index will inform us by how much such discrimination is increased, as well as the extent to which an empirical distribution deviates from its best-fitting normal curve.

The index of skewness[31] (g_1) indicates whether the variance for units on one side of the mean is less than that for units on the other side; and if imbalance exists, the index identifies which side has the greater variance. For a normal curve, g_1 is zero. If g_1 is within the limits ± 1, this shows that the distribution is

the lower chambers were larger than 67.

[28] McNemar, *op. cit.* ftn. 18, *supra,* p. 33: "In order to write the equation of a particular normal curve, i.e., one which corresponds to a particular distribution, we need to know N, M, and σ. This is the basis for the fact that, when we have the usual bell-shaped distribution, we need only the mean and standard deviation to describe it adequately. But in order to say that a given distribution is really normal, it is necessary to show that the g's . . . are zero or approximately zero."

[29] With a normal curve for which the abscissal value of sigma is 5% of the mean, ICV would be .95; but with a sigma equal to two-thirds of the mean, ICV would drop to .60. Obviously, an ICV of much less than .60 (which is equivalent to a CV of .67) would not be associated with a normal curve, because an ICV of .50 implies a CV of 1 (with sigma equal to the mean); and with a symmetrical distribution and an absolute zero point for abscissal values, this would signify a range of ± 1 standard deviation; but the expected minimal range for any of our empirical distributions—if they are normal—is ± 1.8 sigma (when $N=17$).

[30] We are not unaware of McNemar's admonition that "The nature of the research, the type of variable being studied, and also the size of the sample are factors which need to be considered in making a decision as to the necessity for computing measures of skewness and kurtosis. It is seldom advisable to compute these measures when N is less than 100." *Op. cit.* ftn. 18, *supra,* p. 30. *Cf.* Hagood and Price, *op. cit.* ftn. 26, *supra,* p. 214: "The reason that higher moments [than the second] are not frequently computed for distributions observed among only a small number of cases is that they are summarizing measures less stable than the mean and standard deviation." As our footnote 27, *supra,* suggests, over a third of the lower and all of the upper chambers of the fifty state legislatures each had fewer than 100 members at the time for which our empirical data were collected. But all except 3 of the lower chambers each had more than 50 members, and 90% of the senates had at least 25. Consequently, we think that more confidence ought to be reposed in the g scores for the lower chambers than in those for the senates.

[31] For an excellent elementary discussion of the g indices and their formulas, see Hagood and Price, *op. cit.* ftn. 26, *supra,* pp. 210–17; and *cf.* McNemar, *op. cit.* ftn. 18, *supra,* pp. 27–31.

generally symmetrical, although such distributions might well include a number of cases of extreme variance in one tail which are offset by an equivalent number of cases of equally extreme variance in the other tail, as we shall exemplify in a later discussion of cutting points for deviations in variance, using hypothetical distributions in which all items lie in the tails. The index of kurtosis will inform us, however, whether or not such an extremely platykurtic (flat or rectangular) distribution obtains. Any negative g_1 index value shows a distribution which is humped to the right, and a value of -1 or greater would signify that a few (probably rural) units are very much overrepresented. Although theoretically possible, negative skewness is not to be anticipated for apportionment data on American legislatures, for in order for this to occur empirically, a majority of the representational units must be relatively (for the particular state) densely populated, and the minority of rural units are then given equal representation with the more populous urban and suburban units. This implies a state in which there are no extensive metropolitan areas, and in which there is no sharp contrast between densely populated urban units and thinly populated rural units. Most of the American states are quite to the contrary, so we might anticipate that, when representational unit distributions are skewed, they are going to be positively skewed.[32] Positive skewness constitutes the more serious kind of discrimination referred to above, which results from a distribution under which the bulk of the units are thinly populated and lie therefore in a hump to the left, and with a small variance to the left side of the mean; while there is considerable variance to the right or positive side of the mean, where there are a few underrepresented units in which reside the bulk of the population of the state. It is only through distributions that are quite positively skewed that a rural-based minority can really dominate a state legislative chamber.

[32] As Hagood and Price point out, "For a distribution to approach normality closely the range of its possible values must extend several standard deviation units on either side of the mean. For many characteristics, the measures of which can take only positive values, the range is cut off on the left side within two or three standard deviation units of the mean, causing a skew to the right." *Op. cit.* ftn. 26, *supra*, p. 270. For some of our empirical distributions for representation in state legislative chambers, the range is cut off at less than two standard deviations on the left or negative side of the mean.

For the purpose of evaluating apportionment distributions, we suggest that skewness index values be given the following interpretation:

(1) if g_1 is greater than $+1.0$, this signifies relatively serious malapportionment, in the form of underrepresentation of overpopulated units;
(2) if g_1 is more negative than -1.0, this signifies a less serious form of malapportionment, under which there is over representation of underpopulated units;
(3) if g_1 is within the range from -1.0 to $+1.0$, this signifies a relatively normal or symmetrical distribution, and not much greater discrimination of either type than has been described above as inherent in the postulated criterion curve.

The index of kurtosis (g_2) tells us whether the distribution is relatively flat or peaked, in comparison to the normal curve, for which g_2 has a value of zero. A negative kurtosis value indicates a flat or twin-peaked or rectangular curve; there is no single peak signifying a concentration of many units of similar size. Positive kurtosis indicates that there is such a peak, and the greater the concentration of a higher proportion of the units comprising the sample, at or near the same population size, the higher will be the positive kurtosis index. Evidently, the peak of a curve with a high positive g_2 index is likely to lie close to the mean of the distribution of representational units; and it is also apparent that the evaluation of g_2 indices, in relation to apportionment curves, must depend upon where the peak is in relation to the distribution. If a distribution is not skewed, then a high positive g_2 index shows that most of the units are of average size and they are given equal representation; and as we remarked earlier, although we must reject such an ideal as unattainable on empirical grounds, a perfect apportionment would have neither variance (*i.e.*, it would have an ICV of $+1.0$) nor skewness, ($g_1=0$), so its kurtosis would approach positive infinity. In relation to the interpretation that we have suggested for the skewness index, if skewness is normal (*i.e.*, g_1 is within the range ± 1.0) then the higher the positive kurtosis value, the better the apportionment; and with either negative or positive skewness, the higher the positive kurtosis value, the greater the malapportionment. With a skewed distribution, a high positive kurtosis value indicates that the concentration of units in one or the other tail—depending upon the direction in which the distribution is skewed—is relatively extreme. Since we anticipate that, for apportionment distributions, skewness will be positive, we can infer that high positive g_1 values are likely to be associated with high positive g_2 values; and that their joint occurrence signifies relatively extreme malapportionment. We think that when an apportionment curve has high positive index values for both skewness and kurtosis, this means that a rural popular minority is dominating the legislative chamber, irrespective of the degree of variance, because the indicated relationships could obtain even when the range of differences, among the size of population of representational units, was relatively small—corresponding, let us say, to a deviation of 25% from the mean. But we expect that empirically, high positive skewness and kurtosis are likely to be positively correlated with variance: that is, that g_1 and g_2 will both be positive and higher in value, as ICV becomes smaller.

We are now in a position to suggest an operationalized definition of an ideal apportionment which maximizes the principle of equality of representation, and also one for the antithetical ideal of malapportionment. The affirmative ideal is a legislative chamber for which the distribution of representational units is characterized by zero variance, zero skewness, and maximal positive kurtosis: ICV$\rightarrow +1$, $g_1 \rightarrow 0$, $g_2 \rightarrow +\infty$. The negative ideal is characterized by maximal variance, extreme positive skewness, and extreme positive kurtosis: ICV$\rightarrow 0$, $g_1 \rightarrow +\infty$, $g_2 \rightarrow +\infty$.

We have not yet considered the question of bicameralism, and the use of ICV values in order to make interstate comparisons. We shall suggest two alternative ways of proceeding, both of which will be illustrated in the empirical exemplification which constitutes the latter section of this article. The first approach is to prepare a two-dimensional scattergram, using the lower house ICV values for each state (and for the national House of Representatives) as a set of abscissal coordinates, and the upper house ICV values as coordinates for the ordinate dimension. Each state (and the Congress) will then be represented by a single point in the two-dimensional space. (The scattergram procedure could also be used, of course, substituting either g_1 or g_2 values in place of ICV values; and the resulting graphs of the distribution of legislatures, according to skewness or kurtosis, could be compared with the graph of the distribution according to variance such as we shall exemplify in Figure 2.) We can then superimpose a set of orthogonal axes in the space, using the median value of the distribution of upper house ICV values to locate the horizontal axis, and the median value of the distribution of lower house ICV values to locate the vertical axis. These two axes will then divide the space into quadrants, and we can then characterize in a general way the apportionment of all fifty-one legislatures, according to the quadrant in which their points fall. The design for such a scattergram is shown in Figure 1.

We would justify this characterization of quadrants on the following basis. In order to fall in the A quadrant, both the upper and lower house of the legislature must be in the upper halves of their respective distributions, so that on a relative basis, this set includes the most consistently (for both houses) well apportioned legislatures. Conversely, those legislatures, both of whose houses are in the lower halves of their respective distributions, are relatively those which are most malapportioned, and are found in Quadrant D. We distinguish between the second and fourth quadrants, and the first quadrant, on the ground that no matter how high the ICV value for one house, if the corresponding value for the other house is below the median, the relatively malapportioned house will have the blocking power that we previously have discussed; and hence, legislatures with one poorly apportioned house must be considered to be less well apportioned than those which fall in Quadrant A. We distinguish between the second and fourth quadrants on the ground, previously discussed, that malapportionment of the upper house of a legislature (Quadrant C) should be evaluated as a greater handicap to fair representation than malapportionment of the lower house (Quadrant B); therefore, the states that fall in the second quadrant, with relatively well apportioned upper houses but poorly apportioned lower houses, are considered to be better apportioned than the states in the fourth quadrant, where the converse relationship holds. There is an additional historical consideration: Quadrant B legislatures ought to result from the "silent gerrymander," *i.e.*, the failure to reapportion over a long period of time; while Quadrant C legislatures are apt to be the result of deliberate affirmative acts (such as that of the Philadelphia Convention of 1787) to create malapportionment in the Senate. We can also observe that in the B Quadrant, the house blocks the better-apportioned Senate; while in the C Quadrant, the Senate blocks the better-apportioned House.

The scattergram also can readily be used to classify legislatures according to which of their chambers is the more malapportioned, and the extent of difference between the degree of malapportionment in each chamber of the legislature. If the diagonal is drawn from the origin

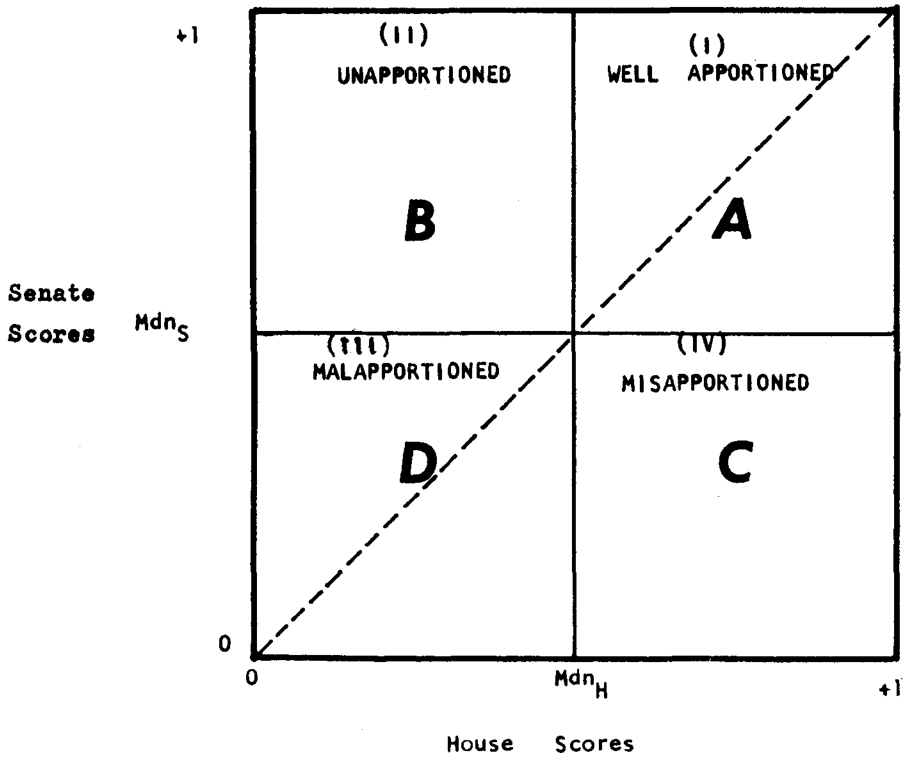

FIG 1. Design for measuring joint variance in apportionment of Houses and Senates of American Legislatures.

to the opposite corner of the space, then the points for all states in which the lower house is more malapportioned than the senate lie above the diagonal, and the points for all states (and, of course, for the United States), in which the senate is more malapportioned than the lower house, lie below the diagonal. The farther away any point is from the diagonal, the greater the difference in the relative degree of malapportionment of the two chambers.

An alternative approach consists of the assignment of weights for lower and upper house ICV scores. Although any set of weights, based upon subjective inferences from the inadequate empirical knowledge presently available, necessarily reflects an arbitrary judgment, we think that an assignment of 1.5 to the senate and 1 to the house is probably more realistic than the assignment of equal weight to each. It may be that, at some future time, empirical investigations of the relative power of the two chambers of the legislature, in each state, corrected for developmental change by repeated observations, will make it possible to apportion weights, for purposes such as ours, on a more discriminating and better-informed basis; for there is no reason to assume that the same weights should apply to each legislature, and there is certainly no technical reason why different sets of weights could not be substituted for the constants in our formula below. But for the present, we prefer the arbitrary assignment of greater weight to the senates, to the equally arbitrary assignment of equal weight to both chambers.[33] In the case of unicameral Nebraska, we shall assign a value of +1.0 for the ICV_h, both for the computation of the summary apportionment score (to be discussed presently) and for the plotting of an abscissal coordinate value in Figure 2, and we shall consider the unicameral ICV to be a senate value; a non-existent house cannot be malapportioned, nor can it block the decisions

[33] We do not insist upon the validity, but only upon the plausibility of our political assumptions; any reader who disputes them is free, of course, to take our ICV and *G* score data and assign any weights he chooses. It was quite laborious to produce the data, but it is only a matter of simple arithmetic and an hour or less at a calculating machine to compute another set of apportionment scores, based on different weights.

of the unicameral chamber which performs the usual functions of an upper chamber—and is called "the Senate"—in Nebraska.

In constructing a scale of apportionment scores, we could use only the weighted ICV indices for each legislature, or we could also take skewness and kurtosis into consideration. Even though we expect these three attributes of representational distributions to be positively intercorrelated, there is in theory a considerable difference—and there is likely in practice to be a considerable difference as well—between two legislatures, both with senate ICV indices of .5, if the one senate has low negative skewness and kurtosis and the other has high positive skewness and kurtosis. Certainly, we require all three indices for an adequate description of the curve for the apportionment distribution of a chamber.[34] One way to measure the legislatures along all three dimensions simultaneously would be to construct a three-dimensional space, with one of the reference axes representing variance, another skewness, and the third kurtosis. We could either plot the two chambers of the legislatures as separate points in such a space; or we could compute a composite score, for the two chambers of each legislature, which could then be plotted as a single point in the space. A somewhat simpler approach, which better serves our present objective, is to combine the ICV, g_1, and g_2 scores for each chamber, and to weight the composite scores for the two chambers of each legislature, in order to summarize in a single index value the essential characteristics of the curves for both chambers of each legislature. A set of such weighted composite indices would permit us to rank all American legislatures along the single dimension of apportionment.

Although we knew that for apportionment distributions, negative values could not be large for either skewness or kurtosis, there was no way of anticipating how large positive values might be for either index. We could and did postulate ± 1.0 as the limits of the normal range for both g_1 and g_2, but we were able to establish an upper cutting point, between large and extreme values, only after observing the empirical data reported in Table II. On the basis of these observations, we defined a G score matrix, which would permit us to weight skewness and kurtosis jointly, for each chamber of each legislature. As Table I shows, we distinguished three segments of the continuum for each distribution of index values: normal, considerable, and extreme; the normal range is the same for both indices, and was defined in advance of the analysis of the empirical data, while the ranges for the other two categories reflect the differences between the two G score distributions. Our assignment of weights to the cells of the G matrix is arbitrary in detail, but not in principle: we already have stated our theoretical reasons for assigning positive weights to positive kurtosis when it is associated with normal (*i.e.*, minimal) skewness, and also for assigning increasingly negative weights to positive kurtosis which is associated with increasingly positive skewness. Actually, it makes little difference what weights we assign, so long as they differentiate consistently between the various combinations of skewness and kurtosis on the basis of the principle stated above, and so long as the weights chosen will establish, in relation to ICV scores in the apportionment index formula, a weight for skewness and kurtosis and a weight for variance that reflects the relative importance of these three characteristics of representational distribution curves. We have assumed that variance is more important than either skewness or kurtosis; in fact, we have assumed that variance is more important than both skewness and kurtosis together. Since the range of our ICV indices is from 0.0 to +1.0, we have assumed that the G (joint) scores for g_1 and g_2

[34] McNemar states that "a typical frequency curve (or polygon) or a frequency distribution can be roughly characterized as one which shows 4 chief features: a clustering of individuals toward some central value, dispersion about this value, symmetry or lack of symmetry, and flatness or steepness. Many variables or traits yield distributions which are said to be approximately bell-shaped, but such a description is not adequate for scientific purposes. One wishes to know about what particular value and with how much scatter the individual scores are distributed, to what extent the distribution is symmetrical, and to what extent it is peaked or flat. That is, we need measures of central value or tendency, measures of scatter or dispersion or variability, and measures of skewness (lack of symmetry) and of kurtosis (peakedness or flatness.) With such measures, one can describe the distribution mathematically, and in such a way that a statistically trained contemporary, say in Melbourne, can picture to himself the frequency distribution." *Op. cit.* ftn. 18, *supra*, p. 13. Hagood and Price are in accord: "In a thorough analysis and description of a quantitative distribution, the aspect of form must be treated just as the aspects of central tendency and dispersion are treated, and if there is to be a generalization of the results, tests of hypothesis about form must be made." *Op. cit.* ftn. 26, *supra*, p. 270.

TABLE I. *G* SCORE MATRIX FOR SKEWNESS AND KURTOSIS

KURTOSIS: \ SKEWNESS:	Extreme	Considerable	Normal	(g_2 scores)
Normal	-.3	-.1	+.1	-1.00* to 1.00
Considerable	-.4	-.2	+.2	1.01 to 8.99
Extreme	-.5	-.3	+.3	≥ 9.00
(g_1 scores)	≥ 3.00	1.01 to 2.99	-1.00 to 1.00	

* As explained in the text, the range for the normal category was defined in advance of the empirical analysis. Included in the Normal/Normal (or the +.1) cell of the empirical *G* score matrix (Table III) are the lower houses of three states for which the skewness scores were positive and less than .65, and the kurtosis scores were as follows: North Carolina, −1.02; Wyoming, −1.03; and Akansas, −1.05. We decided to classify these three with the normal kurtosis scores, rather than to complicate the analysis by almost doubling the number of cells in order to define a four-by-four matrix with a new category, "considerably negative."

ought to have somewhat lesser weight in the apportionment score formula; and as Table I shows, *G* scores are confined to a range of .9 that of the ICV scores.

Our formula for combining ICV and *G* scores. for both chambers of each legislature is:

$$\text{Apportionment Score} = 60(\text{ICV}_s + G_s) + 40(\text{ICV}_h + G_h).$$

By using the constants 60 and 40, to bring about the 1.5 to 1 weighting of senates and houses, we convert the decimal indices for ICV and for *G* into a set of apportionment scores with an approximate range from about zero to about a hundred.

V. AN EMPIRICAL STUDY OF VARIATION IN MALAPPORTIONMENT

Table II reports the means and standard deviations, and the ICV, skewness, and kurtosis indices for both chambers of all fifty state legislatures and for the United States Congress. We used 1960 census data, and the apportionments as of March 26, 1962, the date of *Baker v. Carr*. So many changes have occurred since that date, and the prospects for further changes seem so likely with another major Supreme Court decision pending in the late spring of 1964 or during the 1964 term, that we think the most useful function our empirical study can perform—in addition to that of exemplifying the use of the suggested indices—will be to report the base-line of malapportionment on a nationwide basis; this can then serve as the criterion against which subsequent changes, in any individual state or group of states, can be measured.

Table II reports seven-place decimals, not because we think that such detail would change our own present interpretation in the least from that which we would make with two-place decimals, but rather as a possible convenience to others who might wish to use these data in subsequent analysis. For similar reasons, we report at the end of this table the indices for certain hypothetical distributions with criterion coefficients of variation, because quite a few scholars have shown an interest in postulating norms of "maximum permissible variance" from the ideal of equality of representation. We illustrate the method of procedure with eight such norms; anyone who wishes to extend the analysis to other norms can readily do so, as we shall explain presently. We used a hypothetical sample of ten million voters divided among one hundred representational units.

There were two subsets of fifty representational units each. All units in the same subset were equal in population; but the total populations for the two subsets varied according to the percentage of maximum variation. The stipulated percentage of variation is precisely equivalent to the ratio of the standard deviation to the mean: *i.e.*, the CV is .05, for a maximum variation of five percent; it is .10 for ten percent; and so forth. Thus with a five percent maximum variation

$$\left(\text{and with } CV = \frac{5{,}000}{100{,}000} = .05\right),$$

the population of half of the representational units would be 5,250,000; and the population of the other subset of 50 units would be 4,750,000. The calculation of the ICV for thirty-five percent (for example) reduces, therefore, to

$$\frac{1}{1 + .35} = .741,$$

approximately. We know that these distributions cannot be skewed, so the g_1 index value has to be within rounding error of zero. Similarly, we know that our hypothetical distributions are maximally platykurtic, so the g_2 index value has to be within rounding error of minus two.

In Figure 2 we present the graph, based upon the design of Figure 1, which results from the joint plotting of senate and house ICV indices. In terms of the criterion values for maximum variation (as specified in Table II), none of the chambers was within the 15 per cent range, but ten senates and six lower houses were within 35 per cent. Among the senates, Massachusetts and Arkansas were within the 20 per cent range; Wisconsin, 25 per cent; Ohio, 30 per cent; and New York, New Hampshire, Vermont, Oregon, Maine, and West Virginia were within 35 per cent. None of the lower houses was within the 20 per cent range, but South Carolina was within 25 per cent; Oregon, 30 per cent; and Massachusetts, Illinois, Wisconsin, and the United States House of Representatives all were within 35 per cent. In only three states were both houses of the state legislature within the 35 per cent range: Massachusetts, whose senate ranked first and lower house third; Wisconsin, whose two houses ranked third and fifth; and Oregon, eighth and second. In terms of the criterion of variance, it is clear that Massachusetts ranked as the best apportioned legislature in the country. The remaining 85 chambers fall short—and mostly by a considerable margin—of a "maximum permissible variation" of thirty-five percent. Either that norm is unrealistically high, or else American legislatures generally had a long way to go, in order to realize it, at the time the Supreme Court decided *Baker* v. *Carr*.

Tennessee was the subject—or object—state in that decision; and Figure 2 shows that Tennessee then had an exceptionally well balanced legislature as between its house and senate, though at a very high level of variance in the population of representational units, with an ICV of about .5 in both chambers. So far as we are aware, all of the notorious examples *au courant* in the literature on reapportionment assume in Figure 2 the positions that have been attributed to them: the California senate is indeed the worst in the country, as is the Vermont house (in terms of the stipulated criterion), with Connecticut a close second. If the Senate of the United States showed no more variance than the House of Representatives, the point for the Congress would be located up near Maine, instead of near the center of the C quadrant. Incidentally, the effect of the "federal analogy" argument can be seen in Figure 2 in the proximity, to the United States, of three states which have recently written or revised their constitutions: Alaska, Hawaii, and New Jersey. In the constitutional conventions of the forty-ninth and fiftieth of these United States, the United States Senate was pointed to as a model for emulation, by conservatives who sought to checkmate the "popular" lower house, in strict accord with the precedent established by the

TABLE II. SCORES FOR AMERICAN LEGISLATIVE CHAMBERS ON APPORTIONMENT INDICES OF CENTRAL TENDENCY, DISPERSION, VARIATION, SKEWNESS, AND KURTOSIS

State	Chamber	Mean	Standard Deviation	ICV	Skewness	Kurtosis
Alabama	S	93,335	106,513	.4670299	3.9555078	16.6654127
	H	26,497	18,088	.5942976	2.0111584	5.2508747
Alaska	S	10,396	10,299	.5023477	2.3280182	4.1848776
	H	3,679	1,690	.6851541	1.0663697	0.8958056
Arizona	S	46,505	85,034	.3535464	2.7200651	6.1425738
	H	16,272	8,781	.6495123	1.0338302	1.4462045

TABLE II. (*Continued*)

State	Chamber	Mean	Standard Deviation	ICV	Skewness	Kurtosis
Arkansas	S	49,024	9,771	.8338126	1.5577501	2.4211067
	H	15,926	7,082	.6921997	0.2136530	−1.0504671
California	S	392,780	938,308	.2950822	5.4394092	29.7892981
	H	195,478	79,322	.7113456	0.8873093	0.1960401
Colorado	S	46,176	32,288	.5884998	1.4120893	0.8372165
	H	22,070	14,390	.6053289	1.6462569	1.9114613
Connecticut	S	70,423	41,320	.6302255	1.3263635	0.6923563
	H	7,830	12,722	.3809758	3.7416749	16.0875521
Delaware	S	26,252	21,308	.5519821	0.8990674	−0.8467880
	H	12,751	13,497	.4857966	1.8861561	2.7897886
Florida	S	131,881	177,165	.4267358	2.7307594	8.7673628
	H	36,951	51,968	.4155573	2.9664300	10.4490350
Georgia	S	73,022	85,985	.4592377	3.7847716	16.8197286
	H	14,666	17,134	.4611929	6.9471428	61.6739954
Hawaii	S	22,567	20,943	.5186621	1.1141456	−0.3873740
	H	11,687	5,122	.6952620	1.1366190	0.0982799
Idaho	S	15,163	17,925	.4582692	2.4660653	6.7240682
	H	9,578	5,966	.6161703	0.5557035	−0.5172573
Illinois	S	173,810	108,724	.6151829	1.4598226	2.4230950
	H	56,872	18,127	.7582984	1.7462971	3.0916962
Indiana	S	111,758	109,131	.5059454	4.4688159	22.3989337
	H	65,963	105,932	.3837412	5.0122232	26.2384827
Iowa	S	55,149	39,820	.5807067	3.2766471	3.6299987
	H	22,676	16,816	.5742019	3.4766969	17.4060922
Kansas	S	54,446	57,735	.4853394	3.5405495	13.6147078
	H	17,954	18,336	.4947344	2.2296437	5.5753175
Kentucky	S	80,077	38,727	.6740269	4.1624178	19.3547055
	H	30,373	21,456	.5860231	4.7644850	29.0813696
Louisiana	S	83,088	49,700	.6257212	1.7573380	3.1323713
	H	29,446	20,995	.5837671	2.5250115	7.3668008
Maine	S	26,211	8,807	.7485068	0.8096811	−0.4027144
	H	5,955	2,161	.7337001	0.9859042	3.3689857
Maryland	S	107,030	119,919	.4716037	1.7185467	2.1288223
	H	21,232	18,222	.5381452	1.8467339	2.7182323
Massachusetts	S	127,405	19,780	.8656090	0.7189994	0.5169948
	H	21,825	6,913	.7594519	0.9061084	1.3235289
Michigan	S	230,118	148,635	.6075671	1.6136593	2.2277965
	H	73,968	34,735	.6804599	1.1923953	1.3071320
Minnesota	S	51,370	45,364	.5310429	4.9967158	29.7465438
	H	26,763	28,210	.4868361	4.9053185	31.0552267
Mississippi	S	43,511	22,361	.6605446	1.4414026	2.5126735
	H	16,085	12,535	.5620148	2.4338574	6.0446013
Missouri	S	126,559	45,005	.7376789	1.8501350	3.3171122
	H	27,400	22,615	.5478380	2.0165619	6.2159266
Montana	S	21,049	15,652	.4349539	2.8133181	8.0989704
	H	6,183	3,112	.6651848	0.1302020	−0.9580776
Nebraska	S	32,822	15,364	.6811493	2.2854695	6.9409366
Nevada	S	16,104	31,297	.3397351	2.4534278	4.3676365
	H	4,208	4,098	.5065894	1.9814647	3.3576332
New Hampshire	S	25,285	7,947	.7608633	0.5390549	0.1762299
	H	1,343	618	.6847506	1.1859368	2.9099586
New Jersey	S	288,893	243,806	.5423191	1.1110387	0.4053261
	H	106,539	43,806	.7086302	0.6902714	0.2968090
New Mexico	S	29,719	45,502	.3950909	4.1341645	18.3434493
	H	11,083	6,276	.6384566	0.7682322	0.0918301

TABLE II. (*Continued*)

State	Chamber	Mean	Standard Deviation	ICV	Skewness	Kurtosis
New York	S	280,014	87,676	.7615487	1.8785133	4.8034630
	H	108,272	49,110	.6879584	0.7819559	1.7476650
North Carolina	S	99,009	51,116	.6595091	1.9167141	3.5592471
	H	34,610	20,663	.6261604	0.3760864	−1.0172762
North Dakota	S	12,877	8,016	.6163179	2.0505162	4.0352867
	H	5,898	3,620	.6196405	2.6064888	6.9992922
Ohio	S	266,788	79,625	.7701434	0.7763989	0.0224785
	H	54,077	32,566	.6241370	0.7459419	−0.4307439
Oklahoma	S	50,946	61,722	.4521801	3.6939135	3.6047136
	H	14,780	9,528	.6080271	2.6898505	9.3448576
Oregon	S	53,425	17,392	.7544100	0.4822606	−0.6204543
	H	27,716	7,774	.7809573	0.0653482	−0.7192202
Pennsylvania	S	226,380	119,542	.6544244	1.2429151	0.8493574
	H	51,210	18,251	.7372438	0.5084171	3.0602500
Rhode Island	S	19,299	18,013	.5172321	1.0935368	0.2608778
	H	8,331	5,011	.6244422	1.1642417	1.9587987
South Carolina	S	51,790	49,648	.5105543	2.1870522	4.1363029
	H	18,164	3,882	.8239129	0.0333626	−0.2751049
South Dakota	S	18,793	8,988	.6764667	3.0599798	9.9592973
	H	8,833	3,113	.7394452	0.5438431	−0.1616885
Tennessee	S	120,530	122,175	.4966112	2.9621367	9.5570304
	H	35,449	35,393	.5003961	4.2531467	24.6328687
Texas	S	308,531	241,503	.5609312	2.5115349	5.8760006
	H	55,262	33,363	.6235514	3.1492151	12.6480038
Utah	S	25,704	19,326	.5708154	1.0373689	−0.5113443
	H	9,122	6,792	.5731987	1.4001175	2.4790063
Vermont	S	12,033	3,892	.7555925	−0.6916369	0.2508094
	H	1,585	3,130	.3361228	6.5547544	58.7330107
Virginia	S	99,383	45,576	.6855906	2.2108364	5.7752989
	H	39,978	20,467	.6613963	2.3672073	7.4592624
Washington	S	57,636	26,563	.6845204	1.3975892	2.3238180
	H	28,362	10,002	.7392864	0.8924947	0.2523954
West Virginia	S	58,138	19,536	.7484917	2.5834058	6.4919247
	H	17,482	7,243	.7070628	0.4962934	0.2532985
Wisconsin	S	119,690	28,859	.8057295	1.1927362	1.4330474
	H	39,529	12,913	.7537712	1.7873667	3.1621984
Wyoming	S	11,142	7,520	.5970433	1.1960926	0.2863041
	H	5,439	2,204	.7115949	0.6471677	−1.0312947
United States	S	1,780,193	1,845,698	.4909672	1.8815122	3.3116625
	H	408,644	135,327	.7512243	1.4026598	2.6053311
Criterion values for % of maximum variation						
(%)						
5		100,000	5,000	.9523810	0.0000017	−2.0001151
10		100,000	10,000	.9090909	−0.0000001	−1.9999985
15		100,000	15,000	.8695652	0.0000000	−2.0000002
20		100,000	20,000	.8333333	0.0000000	−2.0000003
25		100,000	25,000	.8000000	−0.0000000	−2.0000000
30		100,000	30,000	.7692308	−0.0000000	−2.0000000
35		100,000	35,000	.7407407	−0.0000000	−2.0000000
40		100,000	40,000	.7142857	0.0000000	−2.0000000

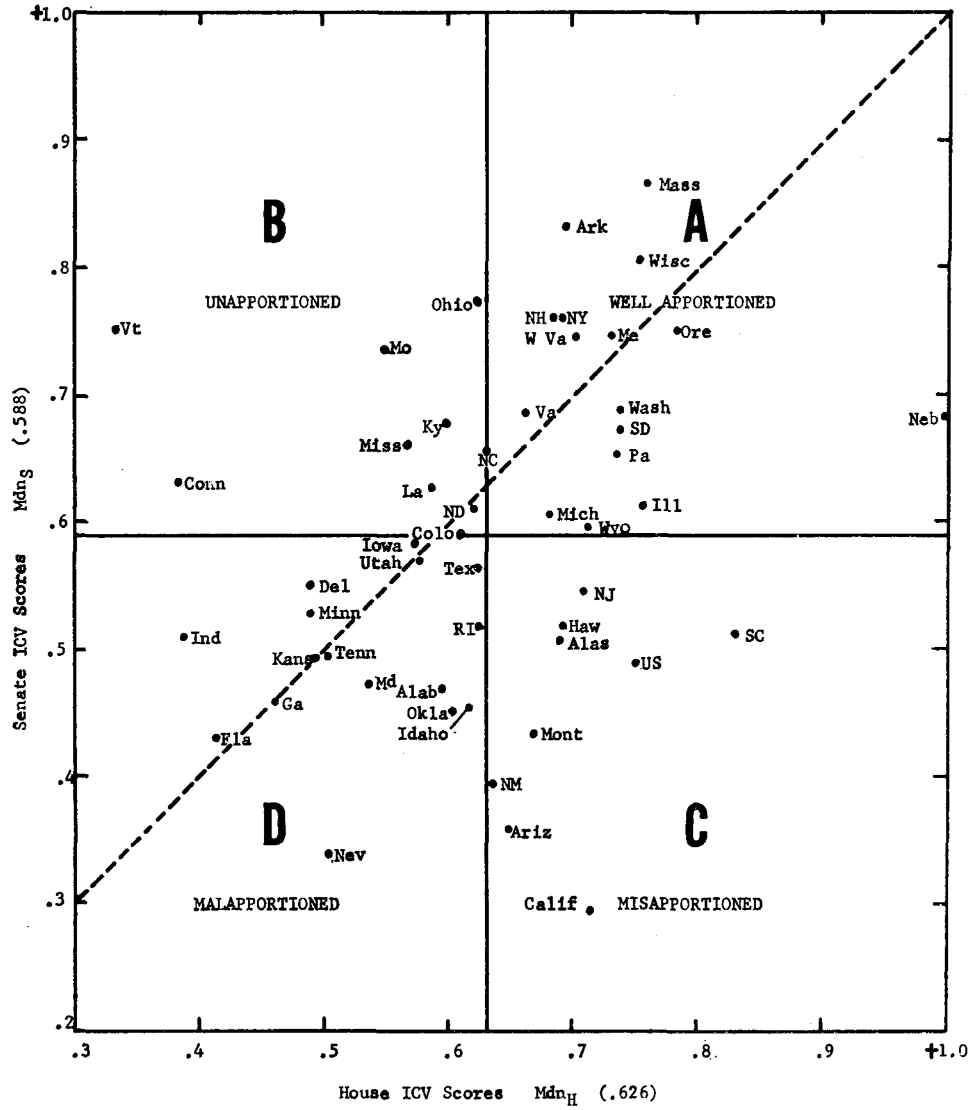

FIG. 2. The apportionment of American Legislatures, as measured by variance in the population of representational units (as of March 26, 1962).

Federalists of 1787. We think that the empirical data support our attribution—on the basis of theory, a year before we saw any computer results—of the label "misapportioned" to the C quadrant. Two of these are states which have taken affirmative steps to assure that the senate will *not* be apportioned according to equality of population, but will instead be able to exercise a veto power over a lower house that is, relatively speaking, apportioned on such a basis; the New Jersey convention was specifically precluded from changing such an existent system—a concession which liberals made in order to change other parts of an archaic state constitution. In the B quadrant, on the other hand, the houses of Vermont and Connecticut most conspicuously reflect the silent gerrymander, retaining in the mid-Twentieth Century the patterns of apportionment that were established in the Eighteenth Century and the early Nineteenth Century, respectively. The lower houses of the states in this quadrant remain unapportioned. Our characterization of the quadrants, of course, is most clearly true of the four corners of the graph; they do not differentiate reliably among the dozen or so states that cluster in the middle around the intersection of the three medians.

The clustering of states along the median diagonal simply indicates a general tendency of many states to apportion both chambers of their legislature—at whatever level of variance—about the same. A general tendency also emerges, however, for the houses to be better apportioned than the senates: 21 states are to the left of the median diagonal, and 28 (not including Nebraska) are to its right. Thus there are a third again as many states in which the senate is less well apportioned than the house as there are states in which the opposite is true. The medians suggest a similar finding: the senate median ICV score is .59, while that for the house scores is .63; and since these axes are medians, they necessarily partition the legislatures symmetrically among the quadrants: sixteen and a half each in A and D, and nine each in B and C (see Table V). None of the states that appear in the D quadrant, with the exception of Rhode Island, is primarily urbanized and industrial; while several such states—New York, Massachusetts, Illinois, Pennsylvania, and Michigan—do appear in Quadrant A.

Table III, which utilizes the categories of skewness and kurtosis on the basis of the g_1 and g_2 score ranges defined in Table I, provides summary information about the distributions of the representational curves. Considerable and extreme skewness are always positive; the solitary negative value lies within the normal range. Negative kurtosis is much more common than negative skewness, but also only within the normal range. Considerable and extreme kurtosis are always leptokurtic (positive); and with the exception of four lower chambers, considerable or extreme skewness always accompanies considerable or extreme kurtosis. The ten chambers with the highest skewness scores, in each of the two distributions (house and senate), are also the chambers with the ten highest kurtosis scores. The combination of extreme positive skewness with extreme leptokurtosis shows a piling up of many representational units with low population, with consequent severe underrepresentation of heavily populated districts. The marginal column and row totals of the table show that the senates tend to be more skewed than the lower houses, but that they show generally no more kurtosis than the houses do. We have indicated already that, in theory, skewness and kurtosis ought to be positively correlated; and Table III shows that, for our data, they are so associated in practice.

In Table IV we present a scale which ranks all

TABLE III. SKEWNESS AND KURTOSIS IN POPULAR REPRESENTATION IN AMERICAN LEGISLATIVE CHAMBERS

KURTOSIS:	SKEWNESS: Extreme	Considerable	Normal	Totals
Normal		8 / 2	7 / 14	15 / 16
Considerable		24 / 19	/ 4	24 / 23
Extreme	11 / 10	1 / 1		12 / 11
Totals	11 / 10	33 / 22	7 / 18	51 / 50

Senates

Houses

TABLE IV. A SCALE OF THE APPORTIONMENT OF AMERICAN LEGISLATURES, WITH SUMMARY EVALUATION OF THE SKEWNESS, KURTOSIS, AND VARIABILITY IN THE CURVES FOR THEIR DISTRIBUTION ACCORDING TO THE POPULATION OF THEIR REPRESENTATIONAL UNITS

Rank	State	Senate:		House:		Joint Vari-ability*	Appor-tionment Score
		Skewness	Kurtosis	Skewness	Kurtosis		
1	Mass.	Normal	Normal	Normal	Considerable	A	96.3
2	Ore.	Normal	Normal	Normal	Normal	A	86.5
3	Ohio	Normal	Normal	Normal	Normal	B	81.2
4	Neb.	Considerable	Considerable	——NONE——		A†	80.9
5	Maine	Normal	Normal	Normal	Considerable	A	80.3
6	N.H.	Normal	Normal	Considerable	Considerable	A	71.0
7	Pa.	Considerable	Normal	Normal	Considerable	A	70.8
8	N.Y.	Considerable	Considerable	Normal	Considerable	A	69.2
9	Ark.	Considerable	Considerable	Normal	Normal	A	66.1
10	W.Va.	Considerable	Considerable	Normal	Normal	A	65.2
11	Wash.	Considerable	Considerable	Normal	Normal	A	62.6
12	Wyo.	Considerable	Normal	Normal	Normal	A	62.3
13	N.J.	Considerable	Normal	Normal	Normal	C	58.9
14	Wisc.	Considerable	Considerable	Considerable	Considerable	A	58.5
15	N.C.	Considerable	Considerable	Normal	Normal	B/A	56.6
16	S.C.	Considerable	Considerable	Normal	Normal	C	55.6
17	Del.	Normal	Normal	Considerable	Considerable	D	50.6
18	Idaho	Considerable	Considerable	Normal	Normal	D	50.1
19	Haw.	Considerable	Normal	Considerable	Normal	C	48.9
20	Va.	Considerable	Considerable	Considerable	Considerable	A	47.6
21	Ill.	Considerable	Considerable	Considerable	Considerable	A	47.2
22	Mo.	Considerable	Considerable	Considerable	Considerable	B	46.2
23	Colo.	Considerable	Normal	Considerable	Considerable	D/B	45.5
24	Vt.	Normal	Normal	Extreme	Extreme	B	44.8
25	Mont.	Considerable	Considerable	Normal	Normal	C	44.7
26	S.D.	Extreme	Extreme	Normal	Normal	A	44.2
27	Mich.	Considerable	Considerable	Considerable	Considerable	A	43.7
28	Utah	Considerable	Normal	Considerable	Considerable	D	43.2
29	Miss.	Considerable	Considerable	Considerable	Considerable	B	42.1
30	R.I.	Considerable	Normal	Considerable	Considerable	D	42.0
31	N.D.	Considerable	Considerable	Considerable	Considerable	B	41.8
32	Alas.	Considerable	Considerable	Considerable	Normal	C	41.5
33	La.	Considerable	Considerable	Considerable	Considerable	B	40.9
34	U.S.	Considerable	Considerable	Considerable	Considerable	C	39.5
35	Md.	Considerable	Considerable	Considerable	Considerable	D	29.8
36	Ariz.	Considerable	Considerable	Considerable	Considerable	C	27.2
37	Conn.	Considerable	Normal	Extreme	Extreme	B	27.1
38	Tex.	Considerable	Considerable	Extreme	Extreme	D	26.6
39	N.M.	Extreme	Extreme	Normal	Normal	C	23.2
40	Cal.	Extreme	Extreme	Normal	Normal	C	20.2
41	Nev.	Considerable	Considerable	Considerable	Considerable	D	16.6
42	Ky.	Extreme	Extreme	Extreme	Extreme	B	13.9
43	Alab.	Extreme	Extreme	Considerable	Considerable	D	13.8
44	Tenn.	Considerable	Extreme	Extreme	Extreme	D	11.8
45	Kans.	Extreme	Extreme	Considerable	Considerable	D	10.9
46	Fla.	Considerable	Considerable	Extreme	Extreme	D	10.2
47	Okla.	Extreme	Extreme	Considerable	Extreme	D	9.5

* The letters correspond to the quadrant designations in Figure 2.

† A maximal ICV score of +1.0 and a maximally positive G score of +0.3 were assigned to Nebraska's non-existent lower chamber.

TABLE IV. (*Continued*)

Rank	State	Senate: Skewness	Senate: Kurtosis	House: Skewness	House: Kurtosis	Joint Vari-ability*	Appor-tionment Score
48	Iowa	Extreme	Extreme	Extreme	Extreme	D	7.8
49	Minn.	Extreme	Extreme	Extreme	Extreme	D	1.3
50	Ga.	Extreme	Extreme	Extreme	Extreme	D	−4.0
51	Ind.	Extreme	Extreme	Extreme	Extreme	D	−4.3

Summary Indices for the Apportionment Score Distribution:

Median = 44.2 ICV = .64
Mean = 42.5 $g_1 = -.02$
$\sigma = 24.2$ $g_2 = -.62$

of the legislatures in our sample according to "apportionment scores" which combine the ICV scores (that measure variance) with the *G* scores (that measure skewness and kurtosis). This scale is more comprehensive, therefore, than Figure 2, which measures variance alone, although as Table V shows, the quadrant classifications of Figure 2 are closely and positively associated with the tails of the distribution for the apportionment score ranking. In the middle ranks, however, the scale is quite independent (in a statistical sense) of the ICV categories. What this shows is that when both chambers of a legislature are similar in variance, then variance scores are a fairly good predictor for skewness and kurtosis, and hence for the composite apportionment scores; but when the two chambers of a legislature are quite dissimilar in variance, then many different combinations of variance-skewness-kurtosis can result in similar apportionment scores. We might say that the road to relatively good, or to relatively bad apportionment, is straight and narrow; while there are many ways in which a mediocre apportionment can be produced.

According to our scale, Massachusetts was the best-apportioned legislature in the country, and Indiana was the worst, at the time the Supreme Court decided *Baker* v. *Carr*. Georgia, however, ranks fiftieth; and Georgia has been conspicuous in recent reapportionment litigation before the Supreme Court.[35] There are only four states in which normal skewness is accompanied by considerable kurtosis, a condition which indicates a piling up of many average sized representational units near the mean and midpoint of the population range. We postulated that this was an essential condition for good apportionment; and it occurs empirically only in the lower chambers of Massachusetts, Maine, Pennsylvania, and New York—states which rank 1, 5, 7, and 8 in our scale. Not a single legislature apportions both of its chambers so well. Only seven states have normal *G* scores for their senates, and five of these rank among the first six in the scale. The other two are Delaware and Vermont, and Delaware ranks lower because of the high variance for both of its chambers, while Vermont ranks lower because of the extreme skewness and kurtosis, as well as the exceptionally high variance in the apportionment of its house. Nebraska ranks very high (fourth) because of our statistical—and political—assumption that

TABLE V. ASSOCIATION BETWEEN ICV QUADRANTS (FROM FIGURE 2) AND APPORTIONMENT RANKS (FROM TABLE IV)

ICV Quadrants:	Ranks: Top Third	Ranks: Middle Third	Ranks: Bottom Third	Total
A	12½	4	0	16½
B	1½	5½	2	9
C	2	4	3	9
D	1	3½	12	16½
Total	17	17	17	51

[35] *Gray* v. *Sanders*, 372 U. S. 368 (March 18, 1963), declaring illegal Georgia's "county unit" system for statewide primary elections; and *Wesberry* v. *Sanders*, 32 L.W. 4142–4157 (February 17, 1964), invalidating Georgia's districting for the national House of Representatives.

the lower house of a unicameral legislature must be perfectly apportioned; there are a few states (actually less than a dozen) with senates better apportioned than Nebraska's, but none with a more equally apportioned house. Only five states have extreme G scores for both chambers of their legislatures; four of these states also are in the D variance category, and they occupy the bottom four ranks of the scale.

As Table IV reports, the set of apportionment scores shows considerable variance, but it evidently is quite normally distributed since it is not skewed, and it is only slightly platykurtic with a negative kurtosis value within the normal range. The mean and the median for the distribution of apportionment scores both fall between 42 and 45. Of the three states that are closest to the median of the distribution—Montana, South Dakota, and Michigan—two in one way or another are not a good choice as a "typical" pattern of apportionment; Montana's senate shows too much variance, and South Dakota's senate has too much skewness and kurtosis to be considered typical. However, Colorado (which ranks twenty-third and had the median senate) and Michigan both fall near the intersection of the median axes in Figure 2; and in spite of the fact that they are classified in different quadrants (D/B and A) of the figure, both have about average variance, which is accompanied by considerable skewness and considerable or normal kurtosis, in both houses of their legislatures. We think that both of those states (the one primarily rural and non-industrialized and the other primarily urban and industrialized) better typify the apportionment of American legislatures as of March 1962.

As Table II shows, and somewhat contrary to the usual understanding, the national House of Representatives, although marginally within the 35 percent range of variation, is not really very much better apportioned than is the United States Senate in terms of the criteria of skewness and kurtosis; and this is reflected in the rank of the national legislature on our scale.[36] The Congress is in rank 34, below two-thirds of the states.

[36] The Supreme Court ruled, in *Wesberry* v. *Sanders*, *ibid.*, that Georgia's districting violated the principle of "one person, one vote" and therefore the popular election clause (Art. I, sec. 2) of the Constitution. Mr. Justice Harlan explicitly argued, in dissent, that if the conclusion of the Court's seven-man majority were correct, then the congressional districting in most of the other states also would be unconstitutional if appraised under the same standard that was applied to Georgia. The considerable skewness and kurtosis, for the distribution of representational units of the national House of Representatives, tend to support Justice Harlan's inference.