THE MEAN VS. THE MEDIAN:
A CASE STUDY IN 4-R ACT LITIGATION

BY

D. A. FREEDMAN

TECHNICAL REPORT NO. 30
JUNE 1984

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY

Prepared for
J. of Business and
Economic Statistics

THE MEAN VS. THE MEDIAN:

A CASE STUDY IN 4-R ACT LITIGATION

By D. A. Freedman*
Statistics Department
University of California, Berkeley
California 94720

Abstract

In a recent case brought under the "4-R Act," the railroads argued that their effective property tax rate should be equalized to the median of the rates for other tax payers. However, the language of the statute compels the use of a weighted mean. The choice between the two measures of location is dictated not by their technical properties or by the nature of the data, but by the statute which defines the objectives of the statistical operations. Before choosing an estimator, we have to decide what parameter it is we are trying to estimate. This principle may seem obvious, but sad experience proves otherwise.

*This article is based on expert testimony given by Freedman on behalf of defendants in Southern Pacific et al vs. California State Board of Equalization et al. The computer work was done by D. Coster and W. Navidi.

Running head. The Mean vs. The Median.

TABLE OF CONTENTS

1. Introduction

Property taxes are levied on the basis of assessed value, for instance, at the rate of $1 per $1,000 of assessed value. However, the relationship of assessed value to market value is variable in nature, and in some jurisdictions different classes of property are assessed at different proportions of their market value. In effect, this creates different tax rates--per $1,000 of market value--for different classes of property. The railroads consider themselves to be the victims of such discrimination. Politically, it may be easier for state and local authorities to increase the effective tax rate on railroads than on voters; also, by comparison with other businesses, railroads find difficulty in relocating to avoid discriminatory property taxes. (This is a very brief summary of testimony given by James N. Ogden, Vice President and General Counsel, Gulf, Mobile and Ohio Railroad, before the Senate Subcommittee on Surface Transportation in 1967, during hearings on S.927, Ninetieth Congress, first session, serial no. 90-48.)

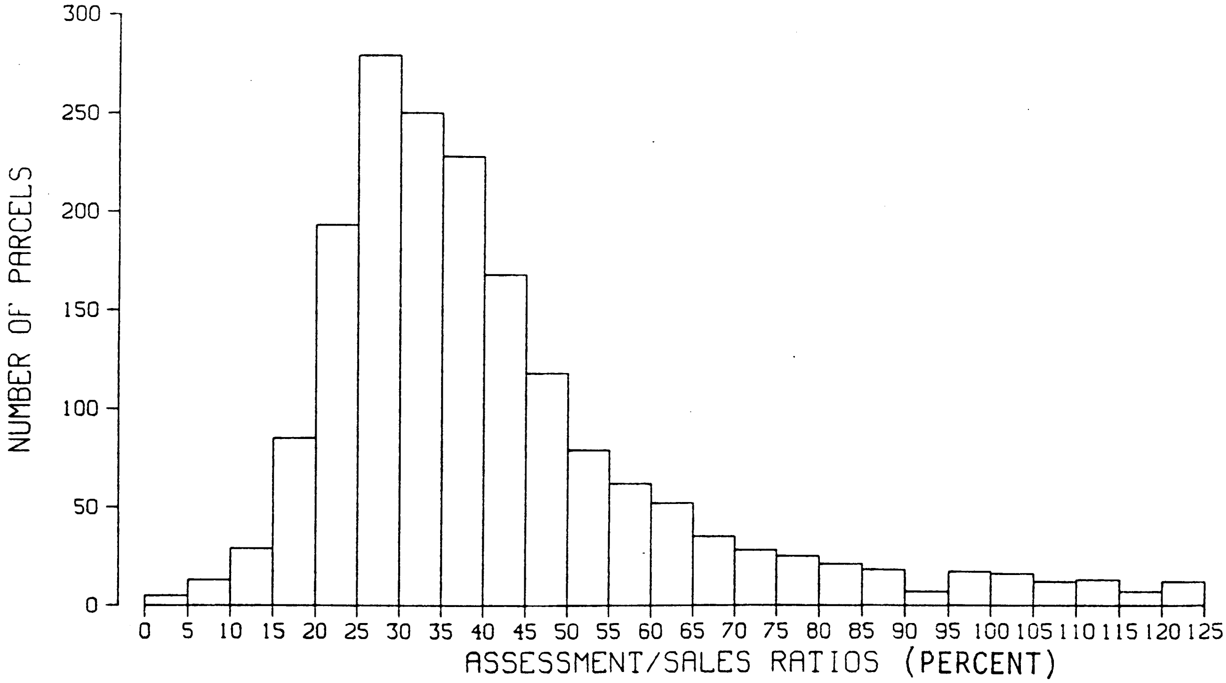
In 1976, recognizing the financial difficulties faced by most of the nation's railroads, Congress passed the Railroad Revitalization and Regulatory Reform Act, known colloquially as "the 4-R Act;" section 306 of that Act was intended to stop property tax discrimination against the railroads. In 1978, this section was recodified in Title 49 of United States Code, Section 11503, without substantive changes in meaning.

Section 11503 provides a standard to which railroad assessment rates must be equalized. However, there is some dispute as to what the standard means, and in such cases it is customary to look to the "legislative history," including the reports of congressional committees, and testimony offered at hearings. There does not seem to be much legislative history to this aspect of the 4-R Act itself; as a result, the courts have looked to the legislative history of similar predecessor bills which were not passed by Congress.

Statistical evidence is crucial in much 4-R Act litigation: for example, "assessment/sales ratio studies" are specifically referred to in the statute and in the legislative history. The idea is to study sales of parcels of property (lots, buildings, etc.) in a given tax jurisdiction in a given time period. For each bona fide "arms length" sale, it is possible to determine: (i) The assessed value of the parcel; (ii) the sales price; (iii) the "assessment/sales ratio," or ratio of assessed value to sales price. The data are usually edited to exclude non-arms-length sales and sometimes the sales prices are adjusted to make them better indicators of market value. For example, transfers between family members may take place at artificially low prices, and should be edited out. Similarly, adjustments to the sales price for inflation are common, as are adjustments to reflect the value of favorable financing terms.

The distribution of the assessment/sales ratios must be considered next. One such distribution is shown in Figure 1, and it has a long right hand tail, as is typical for this kind of data. The distribution is for sales during the tax year 1981-82 in Los Angeles County, for parcels in the "1975 base year." (Passed in 1978, Proposition XIII rolled California assessments back to their 1975 levels, and allowed them to increase only at 2% per year--except when the property is improved, or sold. Since inflation rates for some classes of property were appreciably higher than 2% in the late 1970's, this provision has had a material impact on assessment ratios. Property in California is stratified by county and base year, namely, the year of last prior sale; except that the 1975 base year includes all property last sold in 1975 or before. For example, take property sold during the tax year 1981-82 in Los Angeles County: if a parcel was previously sold in 1978-79, its base year is 1979; if the last previous sale was in 1973-74, the base year is 1975.)

Figure 1. Histogram for distribution of assessment/sales ratios, expressed as percents, Los Angeles County, 1981-82 sales, 1975 base year. The ratios are expressed in percent.



There is general agreement that the railroads should be equalized to some central value in the distribution of assessment/sales ratios. However, there is no unique way to pick out a central value for skewed distributions. Three measures of central tendency are considered in the sales assessment literature: the median, the arithmetic mean, and the weighted mean.

Usually, in assessment/sales ratio studies the distribution of ratios has a long right hand tail. Then, the median will be smaller than the mean. Not surprisingly, railroads argue that they should be equalized to the median; taxing authorities, on the other hand, argue for the weighted mean. This is one of the most common issues to have arisen in 4-R Act litigation; and the courts have divided on this question.

Several such cases came to trial in Federal Court, Northern District of California, in September, 1983. The plaintiffs included all the major railroads and carlines in California, while the defendants were the State of California, the State Board of Equalization, and the 52 California counties with railroad operating property. Fifty counties were jointly represented by an outside law firm; one county was represented by its own county counsel; and the 52nd county, Los Angeles, largely entrusted its defense to the first-mentioned 51. (I gave expert testimony on behalf of the 50 counties.) In this case, the difference between the median and the mean was estimated by plaintiffs as being worth several million dollars a year in tax revenue.

2. The Statistical Issues

In the first phase of the trial, the key issue was framed by the railroads as a technical one: the choice of the appropriate measure of central tendency for the distribution of ratios in data derived from an assessment/sales ratio study. Plaintiffs favored the median, their arguments being largely couched in terms of technical properties of this measure.

From the perspective of a statistician, however, this puts the cart before the horse. In order to decide what statistical operation should be done on the data, we need to have the objectives defined: To what standard should the railroads be equalized? It is the statute and the legislative history which define the standard; the data are then used to get a numerical approximation to that standard. The standard is not to be elicited from technical properties of various statistics, or from the data, but from the law. This argument was made explicitly by the defendants, and never really challenged by the plaintiffs.

In this regard, the key passage in 49 USC 11503 is that a state may not

> assess rail transportation property at a value that has a higher ratio to the true market value of the rail transportation property than the ratio that the assessed value of other commercial and industrial property in the same assessment jurisdiction has to the true market value of the other commercial and industrial property.

Thus, two ratios are contemplated. The first is

> [the ratio of the assessed value] to the true market value of the rail transportation property

The second is

> the ratio that the assessed value of other commercial and industrial property...has to the true market value of the other commercial and industrial property

Railroads are entitled to sue for relief if the first ratio exceeds the second by 5 percentage points or more.

The main controversy was about the interpretation of the ratio for "other commercial and industrial property." In my opinion, however, the statutory language defining that ratio is clear and unambiguous. For all such property in the assessment jurisdiction, the idea is to find: (i) the total assessed value of this property, and (ii) the total true market value of this property, and then (iii) form the ratio

$$\frac{\text{total assessed value}}{\text{total true market value}}$$

In the sales assessment literature, this is called a "ratio of aggregates" or an "aggregate ratio." (The meaning of "commercial and industrial property" is defined in the statute; to avoid tedious repetition, from now on by "property" we will mean commercial and industrial property, other than railroad transportation property.)

At this point, an example may be useful. Consider a hypothetical jurisdiction in which there are only 6 parcels of property, with assessed values and market values as shown in Table 1.

Table 1. Six parcels of property, in a hypothetical jurisdiction.

Parcel	Business	Assessed Value	True Market Value	Ratio
A	Grocer	10,000	40,000	.25
B	Wholesaler	17,500	50,000	.35
C	Beautician	6,000	15,000	.40
D	Factory	600,000	1,200,000	.50
E	Shopping Mall	3,250,000	5,000,000	.65
F	Utility	9,000,000	10,000,000	.90

SOURCE: Affidavit of plaintiffs' statistical expert.

The true market value of these 6 parcels taken together is $16,305,000, namely, the sum of the 6 individual true market values. The assessed value of the 6 parcels is, in the aggregate, $12,883,500. The ratio of aggregates is then

$$\frac{\$12{,}883{,}500}{\$16{,}305{,}000} \approx 79\%$$

In short, I believe "the true market value of the other commercial and industrial property" is just the sum of the true market values of the individual parcels.

For the parcels in Table 1 taken together, the aggregate assessment level is 79% of the aggregate market value. The interpretation in terms of tax rates: if property taxes are levied at the rate (say) of $1 for every $1000 of assessed value, then the owners of the 6 properties in Table 1 are as a class paying taxes at the effective rate of $0.79 for every $1000 of the true market value of their property.

The last column of Table 1 shows the ratio of assessed value to true market value for each parcel; for instance with parcel A,

$$\frac{10{,}000}{40{,}000} = .25$$

Thus, parcel A is assessed at 25% of its true market value. Of course, the ratio of aggregates (as defined above) is the weighted average of the individual assessment ratios, the weights being the true market values.

My reading of the statute as calling for the ratio of aggregates is confirmed by the legislative history, especially the testimony of the spokesman for the railroads, Mr. Ogden. Here is a portion of his testimony:

> Therefore, in order to make possible the fairest comparison--that of the carrier with the hypothetical "average" taxpayer--the unit to be used is that of <u>all</u> parcels of property in the district, considered in the aggregate.
>
> Proof of the composite level of assessment of this property would customarily be submitted by means of the results of a so-called "sales/assessment ratio test." In such a test, various data--the assessed value, sales price, type of property and other relevant factors--are recorded from a random selection of parcels of real estate which have been sold in arm's-length sales between willing

> buyers and willing sellers. From eligible transactions comprising the sample various ratios are computed, with consideration being given to both urban and rural property, and a weighted average is determined and expressed in terms of an assessment ratio, or, in other words, the percentage of true market value at which the property in the sample is assessed. This result can be statistically demonstrated to be accurately representative of the level of assessment of "all other property" in the geographical area (i.e., taxing district) represented by the sample.
>
> It follows, therefore, that a carrier may obtain the relief provided for in the bill by demonstrating that its property is assessed at a higher proportion of its true market value than the proportion at which all other property in the taxing district, <u>in the aggregate</u>, is assessed.

Given data on all parcels in the jurisdiction, the quote seems clear enough: compute the weighted average, construed as

> the proportion at which all other property in the taxing district, in the aggregate, is assessed.

Consider next the statistical problem of estimating the ratio of aggregates

$$\frac{\text{total assessed value}}{\text{total true market value}}$$

The numerator of this ratio, the total assessed value, can be determined from the tax rolls. On the other hand, the denominator is not known and must be estimated from data: that is where the sales data comes in. (The assessed value of a parcel is an administratively determined figure recorded on the tax rolls. The market value is a question of fact, and not known a priori to the tax authorities.) This makes clear the idea we began with: the statute defines the standard conceptually, the data are then used to get a numerical estimate for that standard. The crucial statistical issue is to estimate the total true market value of "the other commercial and industrial property."

Suppose that adjusted sales price is an unbiased estimate of the true market value. And to begin with the easiest case, suppose too that the parcels sold during the study period can be regarded as a random sample of all parcels in the class of commercial and industrial property in the jurisdiction, other than railroad transportation property. (These assumptions are discussed in the Technical Appendix.)

In such circumstances, the ratio of aggregates in the population can be estimated by the ratio of aggregates in the sample:

$$\frac{\text{total assessed value of sample property}}{\text{total sales value of sample property}}$$

If desired, the total market value for all property in the population can be estimated as

$$\begin{matrix}\text{total assessed value of all} \\ \text{property in the population}\end{matrix} \div \begin{matrix}\text{ratio of aggregates} \\ \text{in the sample}\end{matrix}$$

These ratio estimators are very widely used; see (Cochran, 1977, Chapter 6).

In general, however, the parcels of property sold during the study period do not constitute a random sample of the population of all parcels. For example, urban property may sell more easily than rural, so urban property may be more heavily represented in the sample than in the population. Likewise, there may be differences across counties within a state. It is therefore customary to stratify, that is, to do the statistical analysis separately for different subclasses or strata of property.

Suppose then we have a ratio of aggregates for each stratum:

$$\frac{\text{total assessed value of sample property in the stratum}}{\text{total market value of sample property in the stratum}}$$

How are these several ratios to be combined? Given that the objective is to estimate the ratio of aggregates in the population, the right procedure is as follows: For each stratum, form the quotient

$$\begin{matrix}\text{total assessed value of all} \\ \text{property in the stratum}\end{matrix} \div \begin{matrix}\text{the ratio of aggregates for} \\ \text{sample property in the stratum}\end{matrix}$$

This gives an estimate of the total market value of all parcels in the population in that stratum. Now add these estimates over all the strata, to get an estimate of the total market value of all parcels in the population. Finally, take the ratio of total assessed value for all parcels in the population to the estimated total market value. In the sales assessment literature, this is described as a value-weighted average of the separate ratios for the various strata.

Table 2. Stratification

	Assessed Value	Market Value	Ratio of Aggregates
Urban (Total population)	100,000,000	?	?
Rural (Total population)	50,000,000	?	?
Urban (Total sample)	12,000,000	24,000,000	50%
Rural (Total sample)	4,000,000	10,000,000	40%

At this point, an example may be useful. Suppose we have two strata, urban and rural, as shown in Table 2. From the tax rolls, the total assessed value of all urban property is \$100,000,000; of all rural property, \$50,000,000. The corresponding market values are unknown and to be estimated. For the sample, we now take all properties sold during a given period, and classify each parcel as urban or rural. In the sample, the total assessed value of urban property is \$12,000,000, with a total market value of \$24,000,000. (Assessed values are known from the tax rolls, market values are estimated from sales prices.) The sample ratio of aggregates, for urban property, is 50%. We can now estimate the total market value of all urban property in the jurisdiction as

$$\$100{,}000{,}000 \div .50 = \$200{,}000{,}000$$

Likewise, the total market value of all rural property in the jurisdiction is estimated as

$$\$50{,}000{,}000 \div .40 = \$125{,}000{,}000$$

The ratio of aggregates for all property (urban and rural) is then estimated as

$$\frac{\$100{,}000{,}000 + \$50{,}000{,}000}{\$200{,}000{,}000 + \$125{,}000{,}000}$$

or 46%. To set out the reason explicitly: Let T be the unknown total market value of all urban property in the jurisdiction. We estimate the ratio of aggregates for such property to be .50, from the sample. So \$100,000,000/T = .50, and T = \$100,000,000/0.50; likewise for rural property. The statistical procedure described and illustrated above is quite standard; for instance, it is used by the U.S. Bureau of the Census: and its logic was not seriously disputed by plaintiffs.

The real dispute was about the interpretation of the statutory standard. On plaintiffs' interpretation, the railroads should be equalized not to the weighted mean but to the median ratio in the population of all sales assessment ratios. Plaintiffs claimed that the statutory language is ambiguous, so the legislative history must be consulted. The passage they focused on (again from Mr. Ogden's testimony) is as follows:

> ...the bill contemplates the relationship between a common carrier's property and that of the "average" tax payer in the taxing district. However, the word "average" has a precise arithmetical connotation which makes it unsuitable in this context.
>
> For simplicity, therefore, the phrase "all other property in the taxing district" has been used as the equivalent of the property of the "average" taxpayer there. Thus the words "all other property" are to be construed as meaning property in the aggregate...

Plaintiffs argued that the "average" has to be construed as the median. For this choice, they gave technical reasons, and reasons based on linguistic usage.

Given plaintiffs' construction of the statute, their statistical procedures follow as a matter of logic. If the parcels sold during a given year can be viewed as a random sample from the population of all parcels, the sample median is a natural estimator of the population median. Similarly, with a stratified sample, the appropriate procedure is to "parcel-weight" the median as follows: (i) for each parcel in the sample, compute the ratio of its assessed value to its sales price; (ii) array these parcels, so the ratios increase from left to right; (iii) attach a weight to each parcel, being the number of parcels in the corresponding stratum of the population divided by the number of sample parcels in that stratum; (iv) find the ratio in the array which has half the total weight to the left, and half to the right. Again, there was no serious disagreement among the parties as to the logic of this statistical procedure, given the objective of estimating the median assessment ratio in the population of all such ratios. Thus, the controversy was not about technical statistical matters, but about the meaning of the statute.

This completes an outline of the statistical argument, but does not reveal one of the major financial issues behind the dispute: the treatment of other "centrally assessed" property, namely, utilities. On my reading, the statute and the legislative history dictate that such property be included in "the other commercial and industrial property," for purposes of determining the overall assessment level.

Utility and railroad property is assessed "centrally," by the California State Board of Equalization. Other properties are assessed "locally," by the counties. Proposition XIII applies only to locally-assessed property. In an assessment/sales ratio study, centrally-assessed property is treated as a separate stratum. Absent proof to the contrary, there is a presumption that the utilities are assessed at 100% of market value.

The plaintiffs conceded the foregoing, but argued that the utilities should be factored in using the parcel-weighted median approach: one utility, one parcel. Since there were only a few hundred utilities in California, weighting them in on such a basis would only change the median by a few tenths of a percentage point. On the other hand, the utilities accounted for 14% of the assessed value of "other commercial and industrial property" in California, so weighting them in by value would change the numerical standard by a substantial amount. This was an issue with a lot of horsepower.

3. Plaintiffs' Technical Arguments

Plaintiffs' technical arguments will be reviewed here. The main ones were as follows: (i) the median has a smaller sampling error than the weighted mean; (ii) the weighted mean can be skewed by a few large properties; (iii) the median minimizes tax discrimination; (iv) the median is more equitable than the weighted mean; (v) the median is the generally preferred measure of central tendency in the sales assessment literature.

i) Sampling error. Plaintiffs' assertion about sampling error is simply false, as a general proposition. Whether the median or the weighted mean will have a smaller sampling error depends on the population, and either measure can have the advantage. However, these considerations are virtually irrelevant. If the weighted mean of the population is the legally-mandated standard, the sample median should not be used, because it is badly biased as an estimator of the weighted mean. Likewise, if the median of the population is the standard, the weighted mean of the sample is the wrong estimator to use.

ii) Skewing. In Table 1, the weighted average is 79%; so 5 out of the 6 parcels are assessed below the weighted average assessment level. This "skewing" is due to the influence of parcel F on the weighted average. Parcel F is large in dollar terms and has a high ratio. The argument about skewing may have some force if it refers to the population, in an argument about the equities; although to me it seems fair enough to compare railways to utilities and shopping malls rather than grocers and beauticians. If the implication is drawn that the sample results will be distorted by an occasional large sale, then this simply restates the first point on sampling error, and is wrong. The sample results could be distorted, but this is unlikely to be the case.

iii) Discrimination. Consider a set of assessment ratios for various parcels, as in Table 1. We wish to equalize the railroads' ratio to some central value. For each such possible central value, consider the sum of the absolute values of the deviations between that central value and the six values in the table. Plaintiffs' expert considered this sum to be a measure of tax discrimination; and the sum is minimized when the central value is the median.

Here again there was no dispute about the mathematics, but the interpretation is open to question. Why is this sum the right measure of discrimination? If we measure discrimination in dollar terms rather than percentage points, it is the "value-weighted median" which gives the minimum. If we combine the deviations by taking their root-mean-square instead of the sum, the value-weighted mean gives the minimum. Plaintiffs' choice of total absolute deviation as a measure of discrimination was arbitrary.

iv) Equity. If the railroads are equalized to the median, the tax rate on half the parcels will be lower, and on half the parcels the rate will be higher. On the other hand, if the railroads are equalized to the mean, the railroads wind up paying tax at a higher rate than over half the taxpayers tax rate being in relation to market value not assessed value.

This is true, but there is also an argument from equity in favor of the weighted mean: Tax inequity arises from the fact that different taxpayers are assessed at different fractions of market value. From a social viewpoint, the fairest equalization standard is that which, if applied to all taxpayers, would not affect the total revenues to the taxing jurisdiction, but would merely re-allocate tax burdens by raising some assessments and lowering others If all taxpayers in a jurisdiction were equalized to the weighted mean, total tax revenues would be unaffected. If, however, all taxpayers were equalized to the median, total tax revenues would be affected, and in general would be reduced. Again, these conflicting arguments seem to be resolved by the statute.

v) The weight of opinion. Plaintiffs argued that the weight of opinion favors using the median as a measure of central tendency in sales assessment ratio studies. This may be so within a stratum, although the preference is a guarded one. See, for example, the National Association of Tax Administrators (1954, p. 24). For combining strata, however, a value-weighted mean of the measures for the separate strata is recommended by all the authorities. The National Association of Tax Administrators is very clear about this. Also see the International Association of Assessing Officers (1980, pp. 5-6). Conventional opinion, then, favors a value-weighted mean of medians. This is quite different from the "parcel-weighted median" urged by the plaintiffs. In particular, with a value-weighted mean of medians, centrally-assessed property will come in as recommended by defendants.

Again, however, given the objective of estimating the total market value of "the other commercial and industrial property," this is all irrelevant. Plaintiffs conceded, in response to interrogatories and in deposition testimony of their experts, that for such purposes the weighted mean must be used. Indeed, this seems to be the consensus opinion, as expressed for example by the International Association of Assessing Officers (1978, p. 127);

> This weighting feature makes [the weighted mean] without question the single most appropriate measure of the assessment level for estimating the full cash value of all real property in a particular use class or stratum of properties, as might be done for equalization purposes. The appropriateness of the weighting feature, however, is a matter of opinion with respect to the measurement of assessment performance.

Why then is the median so popular? The answer may be in the last sentence of the passage just cited. Historically, one main purpose of sales assessment ratio studies seems to have been quality control: gauging the "typical" level of assessments in a jurisdiction, and the dispersion around that level. For such purposes, the median has intuitive appeal as a measure of central tendency; so does the average absolute deviation around the median, as a measure of dispersion. In California, after Proposition XIII, this quality-control function is irrelevant. In any jurisdiction, this function seems largely beside the point in 4-R Act litigation.

4. Issues of Terminology

There was considerably dispute over the meaning of various terms of art in statistics, starting with the word "average," however odd this may seem. A 29-page reply brief filed on behalf of one plaintiff railroad, for example, devoted 7 pages to this issue, including 3 pages of citations from Webster's Third New International Dictionary and from The Mathematics Dictionary. (This choice of mathematical authority may have been unfortunate. The "generalized formulation" there, when applied to a data set consisting of the two elements -1 and 1, typically produces an infinite string of imaginary numbers, rather than the desired 0: perhaps the first time that a Federal Court has been faced with imaginary numbers.)

To me, however, the facts seem clear enough. The word "average" is used two ways, generically and specifically. In its generic sense, it means "not unusual:"

Q. How are you feeling today?

A. Oh, just average.

In its generic sense, the average could refer to the median, or the mean, or the weighted mean, or any other number somewhere in the central part of a distribution: all such distinctions are blurred when the term "average" used generically. In its specific sense, the average means one and only one thing: you add up all the numbers, and divide by how many there are. That is even what plaintiffs' cited authorities do, when they get down to doing arithmetic.

There was also some controversy about the term "weighted average" in the legislative history. Since "average" used generically could mean the median, one plaintiff argued that "weighted average" meant the median too:

> The median, and the mode, as well as the mean, can in general parlance be weighted averages.

However, in contrast to the term "average," the phrase "weighted average" is not used generically by statisticians or anyone else, as far as I can tell. The weighted average is strictly synonymous with the weighted mean, and it is never used to refer to the median, weighted or otherwise.

Another of the plaintiffs took a different tack, starting from their favorite sentence in the legislative history:

> However, the word "average" has a precise arithmetical connotation which makes it unsuitable in this context.

Congress must therefore have intended to banish the arithmetic mean, and by extension the weighted mean, for that too involves doing arithmetic. As that plaintiff put the matter, "the value weighted mean is an arithmetic average," and therefore taboo.

We close with yet another of plaintiffs' linguistic arguments for the median over the mean. Both, after all, are "measures of central tendency." And, continuing from the deposition transcript of plaintiffs' statistical expert:

> I suggest to you that when you deal with measures of central tendency, you must be talking about numbers near the center. Otherwise, why use central tendency? And I don't know of any number closer to the center than the center, which is the median.

The sad fact is that for the skewed distributions characteristic of sales assessment ratio data, there is no unique center. The mode defines the center in one sense; the median, in another; the mean, in still another. To decide which definition of center is appropriate, we have to look not to the data, not to the theory of statistics nor its terms of art, but to the legislation.

5. What the Court Held

The Court ruled for defendants on all issues in dispute in phase I of the litigation. (Phases II and III dealt with other issues, and had not come to trial when this article was prepared.) In particular the Court found that the weighted mean was the appropriate measure of central tendency within each stratum; that strata should be combined by value; and centrally-assessed property should be factored in by value. The Court rejected the median and parcel weights. For a brief review of other pertinent cases, see the Appendix on the Case Law.

TECHNICAL APPENDIX

This appendix covers the following topics: (i) the randomness assumption; (ii) standard errors for weighted means; (iii) standard errors for weighted medians; (iv) analysis of some sample data; (v) exceptional cases. It is intended as a brief review of technical properties of the main estimators considered by the parties and does not bear on the construction of the statute.

The Randomness Assumption

A bedrock assumption in an assessment/sales ratio study is that, after stratification, the sample is random. On its face, the randomness assumption is doubtful: for example, there is opinion that relative to smaller properties, large commercial properties sell rarely, appreciate slowly, and are assessed at lower proportions of their market value. Stratification by assessed value might help, but this is not standard. George Mitchell, the intellectual father of assessment/sales ratio studies, had this to say on the randomness assumption (National Association of Tax Administrators, 1954, p. 36):

> That the conditions of random sampling are met by sales samples is usually implicitly assumed without study or test.

In the balance of this appendix I will assume, faute de mieux, that the sample is random within strata. I will also assume that market values are known accurately for parcels in the sample; however, the component of variance due to measurement error in determining market values is largely picked up in the formulas.

Standard Errors for Weighted Means

We consider stratified samples. Index the strata by s. Let a_s be the total assessed value of all parcels in the population of type s, which is known; and v_s the corresponding market value, which is unknown. Let $a = \Sigma_s a_s$. Let A_{sj} be the assessed value of parcel j of type s

in the sample, and V_{sj} the market value. The sample ratio of aggregates for stratum s is $R_s = \Sigma_j A_{sj}/\Sigma_j V_{sj}$, and its SE may be approximated as follows:

$$(1)\qquad \begin{cases} \text{SE of } R_s \doteq R_s \cdot F_s \\ F_s = \left[\dfrac{\Sigma_j A_{sj}^2}{(\Sigma_j A_{sj})^2} + \dfrac{\Sigma_j V_{sj}^2}{(\Sigma_j V_{sj})^2} - 2\,\dfrac{\Sigma_j A_{sj}V_{sj}}{(\Sigma_j A_{sj})(\Sigma_j V_{sj})}\right]^{\frac{1}{2}} \end{cases}$$

This formula, like subsequent ones, is based on the delta-method: see Cochran (1977, Chapter 6). It assumes randomness within strata, with the sample being only a small fraction of the population.

The total market value of all property in the population is estimated as $\hat{v}_s = a_s/R_s$, with the standard error given as follows:

$$(2)\qquad \text{SE of } \hat{v}_s \doteq \hat{v}_s \cdot F_s$$

Then $\hat{v} = \Sigma_s \hat{v}_s$, with the standard error given as follows:

$$(3)\qquad \text{SE of } \hat{v} = [\Sigma_s (\text{SE of } \hat{v}_s)^2]^{\frac{1}{2}}$$

The population ratio of aggregates is estimated as $R = a/\hat{v}$. The SE of R may be estimated as follows:

$$(4)\qquad \frac{\text{SE of } R}{R} \doteq \frac{\text{SE of } \hat{v}}{\hat{v}}$$

Standard Errors for Weighted Medians

To begin with, we view the sample ratios A_i/V_i as independent, identically distributed observations on the population density f. This is a reasonable approximation, if the sample is random and only a small part of the population. Let $\hat{\theta}_n$ be the sample median of $A_1/V_1,\ldots,A_n/V_n$. Let θ be the median of f. Then $\hat{\theta}_n$ may be used to estimate θ, with standard error

$$\text{(5)} \qquad \text{SE of } \hat{\theta}_n \doteq \frac{1}{\sqrt{n}} \cdot \frac{1}{2f(\theta)}$$

To use this formula, we must estimate $f(\theta)$. This can be done (crudely) as follows: take the frequency distribution of sample ratios; find the class interval I containing the sample median; then

$$\text{(6)} \qquad f(\theta) \doteq \frac{\text{number of ratios in I}}{n \cdot \text{length of I}}$$

A reference is Lehmann (1983, p. 354).

We turn now to stratified samples. Suppose the ratios in stratum s follow the density f_s . Suppose in the population there are N_s parcels in stratum s , and $N = \Sigma_s N_s$. The overall density is $f = \Sigma_s N_s f_s/N$, and we seek the median θ of f . We view $X_{sj} = A_{sj}/V_{sj}$, $j = 1,\ldots,n_s$ as n_s independent observations on f_s , where n_s is the number of parcels of type s in the sample, and $n = \Sigma_s n_s$. We attach weight N_s/n_s to X_{sj} . Let $\hat{\theta}$ be the value such that half the weight lies to the left, and half to the right; this the "parcel-weighted median." Then $\hat{\theta}$ estimates θ .

The sampling theory of $\hat{\theta}$ is not so well developed, but the following approximations may be useful. Let

$$p_s = \int_{-\infty}^{\theta} f_s(x)dx$$

Then the SE of $\hat{\theta}$ may be approximated as follows:

$$\text{(7)} \qquad \text{SE of } \hat{\theta} \doteq [\Sigma_s \frac{N_s^2}{n_s} p_s(1-p_s)]^{\frac{1}{2}}/[\Sigma_s N_s f_s(\theta)]$$

Here, p_s may be estimated as the fraction of the X_{sj} , $j = 1,\ldots,n_s$ to the left of $\hat{\theta}$, while $f_s(\theta)$ can be estimated as the fraction of the X_{sj} , $j = 1,\ldots,n_s$ falling into the class interval containing $\hat{\theta}$, divided by the length of that interval.

Data Analysis

The object is to illustrate the discussion using data for sales in Los Angeles in 1981-82. Table 3 shows the total population parcel counts and values; data for the sample (property sold during 1981-82) are shown in Table 4.

Table 3. Commercial and industrial property in Los Angeles County 1981-82. Population figures.

Base Year s	Number of Parcels N_s	Total Assessed Value a_s (millions of dollars)
1975	105,382	22,062
1976	11,282	2,209
1977	15,793	3,290
1978	18,410	4,622
1979	19,394	5,433
1980	21,901	6,671
1981	21,088	7,283
Total	213,250	51,570

DATA SOURCE: Plaintiffs

Table 4. Commercial and industrial property sold in Los Angeles County 1981-82; values are in millions of dollars.

Base Year s	Number of Parcels n_s	Total Assessed Value $\Sigma_j A_{sj}$	Total Sales Value $\Sigma_j V_{sj}$
1975	1818	216	535
1976	116	19.1	38.2
1977	172	33.5	66.4
1978	268	55.1	101
1979	363	79.6	131
1980	419	105	150
1981	310	100	142

DATA SOURCE: State Board of Equalization

We begin with a rough test for randomness. For the 1975 base year, the mean assessed value for all parcels of commercial and industrial property in Los Angeles is (Table 3)

$$\frac{\$22,062,000,000}{105,382} \approx \$209,000$$

In the sample, the corresponding figure is (Table 4)

$$\frac{\$216,000,000}{1818} \approx \$119,000$$

Thus, large parcels are underrepresented in the sample. This is hardly due to sampling error: the standard error of the sample mean is only about \$6,000. In other strata, the problem does not appear to be serious. The 1975 stratum is the most heterogeneous; further stratification might be helpful in reducing bias, for instance by year of last sale, type of use, or size class.

Assuming randomness, results for the various base years are shown in Table 5. The ratio of aggregates increases markedly with the base year:

this appears to be due to the impact of Proposition XIII on assessments.

Table 5. Estimated ratio of aggregates and standard error, for each base year.

Base Year	Estimated Ratio of Aggregates	Standard Error
1975	.404	.011
1976	.500	.033
1977	.505	.017
1978	.546	.027
1979	.608	.038
1980	.700	.021
1981	.704	.084

Table 6 shows the estimated population market values by base years. The overall ratio of aggregates (all base years combined) is estimated as 50.1% with a standard error of 1.0 percentage points. The results for the medians are summarized in Table 7. Overall, the parcel-weighted median is estimated as 47.8% ± 0.5%.

Table 6. Population assessed value and estimated market value, by base year, in millions of dollars.

Base Year	Assessed Value	Estimated Market Value	Standard Error
1975	22,062	54,609	1,487
1976	2,209	4,418	292
1977	3,290	6,515	224
1978	4,622	8,465	420
1979	5,433	8,936	553
1980	6,671	9,530	286
1981	7,283	10,345	1,234
Total	51,570	102,818	2,106

A bootstrap experiment was performed to test the validity of the approximation (1-4). The experiment used the actual sample (Table 4) as an artificial population, stratified by base year: for instance, there were 1,818 parcels in base year 1975. Then, 100 artificial data sets were created by Monte Carlo. Each data set involved drawing an independent sample (at random, with replacement) from each stratum, the sample size for stratum s being n_s . For each artificial data set and each stratum, two quantities were computed in the first instance: the ratio of aggregates, and its standard error from (1). The average and standard derivation of the 100 ratios were computed, and the root-mean-square of the 100 standard errors. The average was compared to the assumed value (column 2 of Table 5); the rms SE to the SD. Likewise, for each artificial data set, a pooled ratio of aggregates was computed, along with its standard error from (4), and analyzed as before.

All went well, with two minor exceptions:

- In base year 1980, the sample ratio of aggregates was biased by an amount which was small but statistically significant.
- In base year 1981, the formula (1) is biased downward, by about 15%. This seems to be due to one very bad outlier in the data set for 1981.

Furthermore, the distribution of the ratios is close to normal: a p-p plot for base year 1979 is presented in Figure 2. Thus, the approximations are very good, despite the skewed and long-tailed data. For a discussion of the bootstrap, see Freedman and Peters (1984); on the impact of extreme values, see Bickel and Freedman (1981, p. 1210).

Figure 2. A p-p plot for Monte Carlo ratios of aggregates: Los Angeles County, 1981-82 sales, 1979 base year. Let Φ be the standard normal distribution function, and Ψ the standardized empirical distribution of 100 ratios from the bootstrap experiment. The horizontal axis shows $\Phi(x)$; the vertical, $\Psi(x)$.

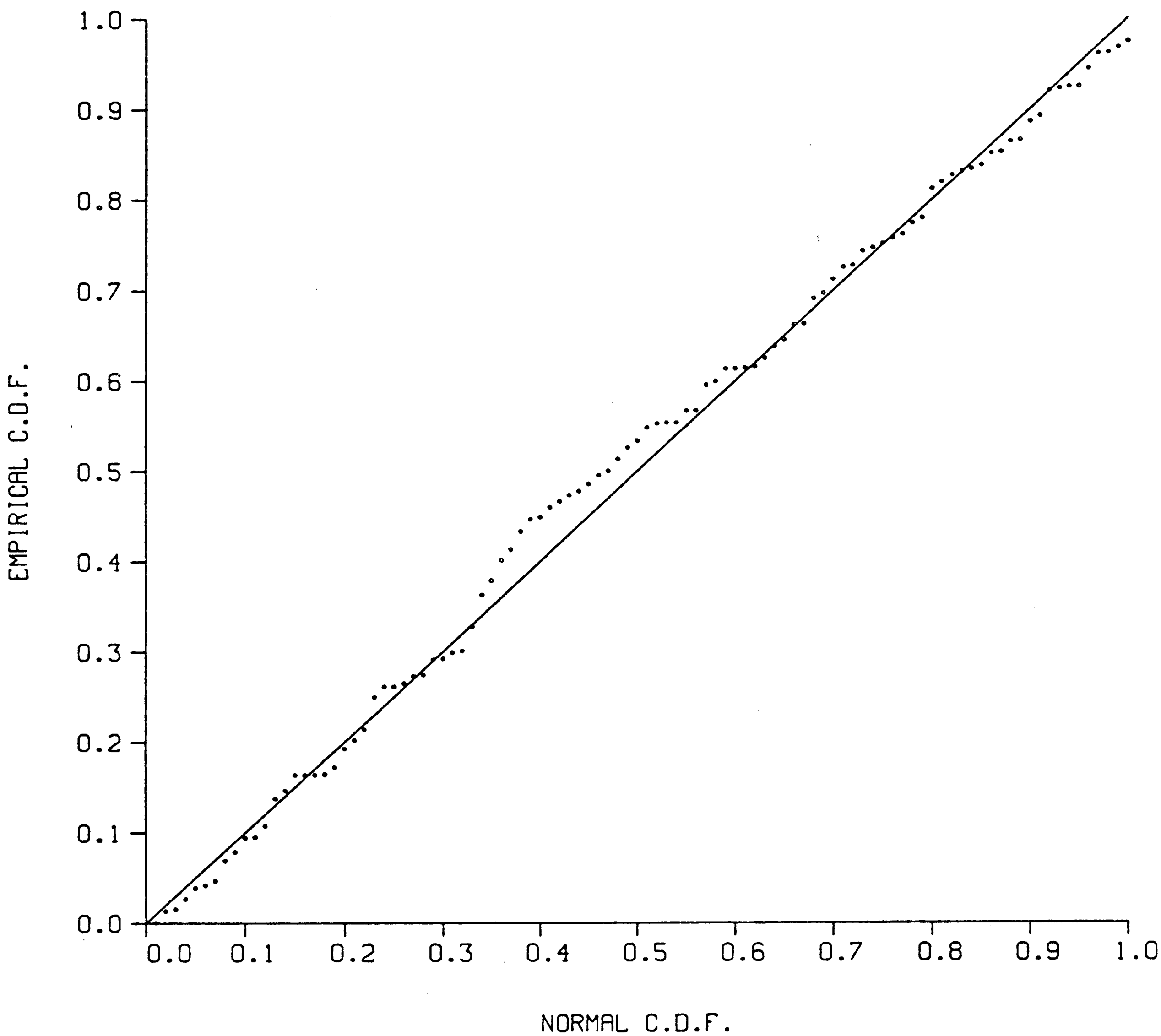
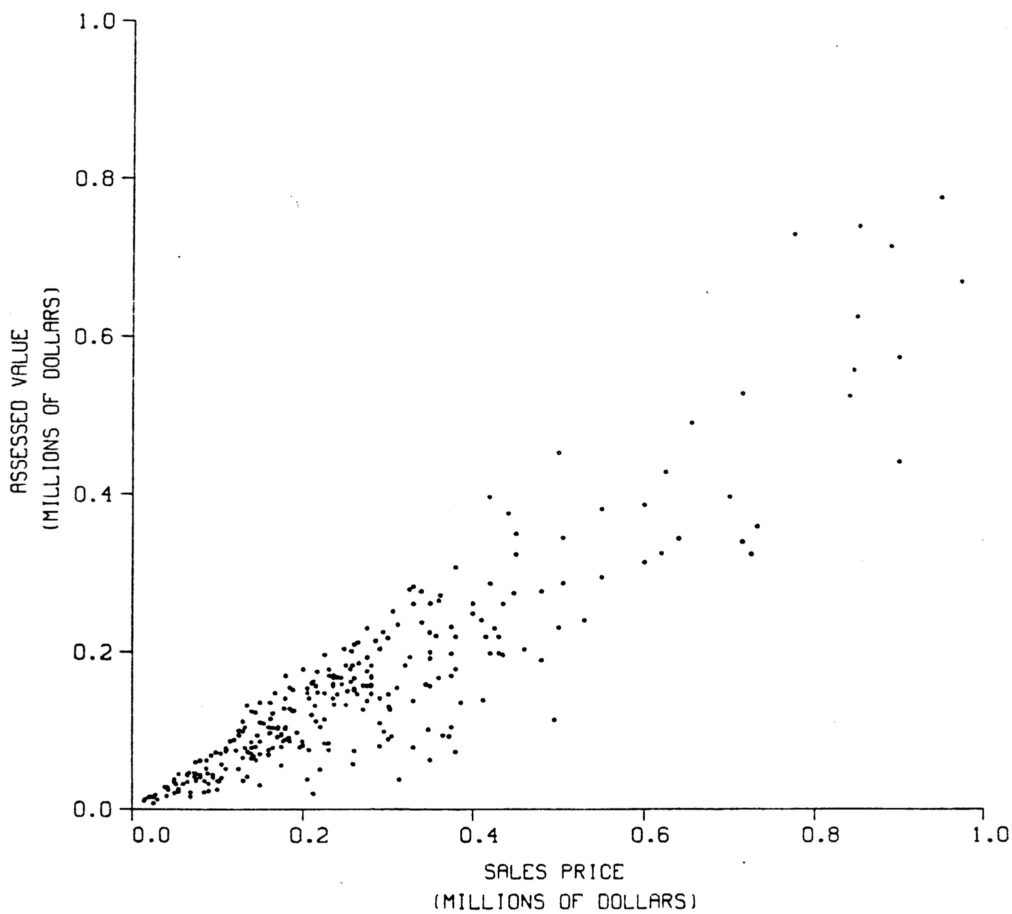


Table 7. Results for the parcel-weighted median.

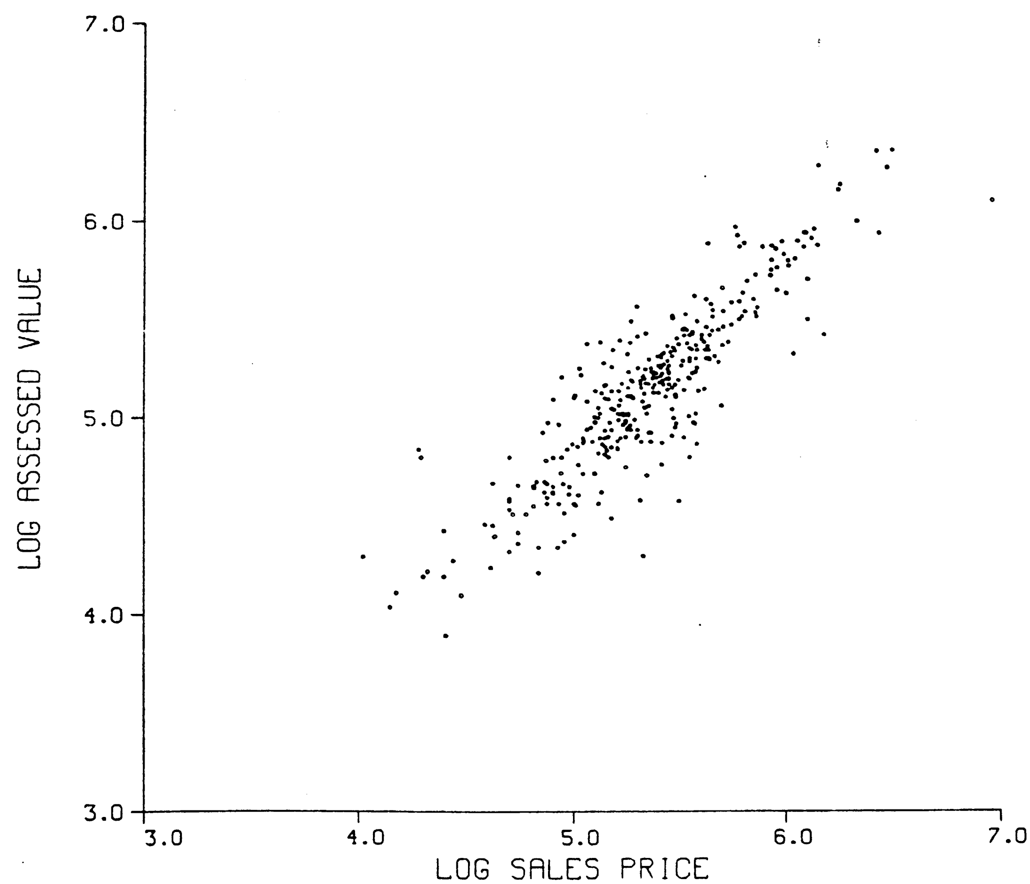
Base Year s	n_s	N_s	p_s	$f_s(\hat{\theta})$
1975	1,818	105,382	.72	.013
1976	116	11,282	.58	.017
1977	172	15,793	.44	.033
1978	268	18,410	.36	.017
1979	363	19,394	.25	.018
1980	419	21,901	.15	.0091
1981	310	21,088	.10	.0032

Figure 3. A scatter plot of assessed value against sales price: Los Angeles County, 1981-82 sales, 1979 base year. Parcels with sales price exceeding $1,000,000 or assessment/sales ratio exceeding .99 are excluded.



The next topic is a description of the data. A plot of assessed value against sales price in Figure 3 shows marked heteroscedasticity. However, a logarithmic transformation more or less induces normality, as in Figure 4; the logarithms are to base 10, so 3 on the axis corresponds to $1,000, and 7 to $10,000,000.

Figure 4. A scatter plot of $\log_{10} A$ against $\log_{10} V$: Los Angeles County, 1981-82 sales, 1979 base year.



If X is normal with mean μ and variance σ^2, then $\exp(X) = e^X$ has expectation $\exp(\mu+\frac{1}{2}\sigma^2)$. With stratum s, the log normal idea therefore suggests an alternative estimator for the ratio of aggregates, namely,

$$\tilde{R}_s = \exp\{\hat{\mu}_{sA} + \tfrac{1}{2}\hat{\sigma}^2_{sA} - \hat{\mu}_{sV} - \tfrac{1}{2}\hat{\sigma}^2_{sV}\}$$

where

$$\hat{\mu}_{sA} = \frac{1}{n_s} \sum_{j=1}^{n_s} \log A_{sj}$$

$$\hat{\sigma}^2_{sA} = \frac{1}{n_s} \sum_{j=1}^{n} (\log A_{sj})^2 - \hat{\mu}^2_{sA}$$

and likewise for V. We tried this out in a bootstrap experiment for 1979 base year (resampling the empirical distribution), and found it worked reasonably well: $\tilde{R}_s$ is somewhat biased, but this problem is more than offset by the reduction in sampling error, relative to R_s. We do not recommend $\tilde{R}_s$ for general use, without further testing.

How does the assessment/sales ratio A/V depend on the assessed value V? In general, A/V seems to decrease slowly with V. Table 8 below shows fitted regression equation for each base year, of the form

$$\log A/V = \alpha + \beta \log V + \text{error}$$

In 1975, for example,

$$\log A/V = .29 - .14 \log V + \text{error}$$

In round numbers, the equation suggests that in the 1975 base year, the median value of A/V will be around $2/V^{1/7}$: parcels with a sales price around $100,000 will have a median assessment/sales ratio of around 40%, while parcels with a sales price near $1,000,000 will have a median assessment/sales ratio of around 30%. The effect for other base years is similar.

Table 9. Regression of log A/V on log V.

Base Year	Intercept	SE	Coefficient	SE
1975	.29	.06	-.14	.01
1976	.51	.28	-.16	.05
1977	.21	.18	-.10	.03
1978	.15	.16	-.08	.03
1979	.27	.14	-.09	.03
1980	.91	.14	-.20	.03
1981	.45	.14	-.10	.03

Exceptional Cases

In some cases, it may be advisable to estimate the median by the mean, or vice versa. The following model is intended to illustrate the point; it is not of practical interest. The model is as follows:

(8) $$A_i/V_i = r + \varepsilon_i$$

(9) The ε_i are independent with a common distribution symmetric about 0.

Thus, A_i is rV_i give or take a symmetric random error, whose size is proportional to V_i. The aggregate assessment level r is to be estimated. If (say) the ε_i are normally distributed, then r should be estimated as the arithmetic mean of the ratios A_i/V_i : the mean is used to estimate the median. On the other hand, if (say) the ε_i have a double exponential distribution with density proportional to $e^{-|x|}$, then r should be estimated as the median of the ratios: the median is used to estimate the mean. Squared-error loss is assumed. For a discussion of loss functions, see Varian (1974).

APPENDIX ON CASE LAW

This appendix gives a brief summary of some pertinent cases, prepared by counsel for defendants in Southern Pacific v. State Board of Equalization.

1. ACF Industries, Inc. v. Arizona, 561 F. Supp. 595 (D. Arizona 1982), aff'd, 714 F.2d 93 (9th Cir. 1983). Arizona's scheme for compliance with The 4-R Act sets the assessment ratio of railroad property using a weighted mean of the ratios of other commercial and industrial property. The federal district court approved the use of the weighted mean in these circumstances, stating that in *de facto* discrimination cases, the median might be preferred but that in a *de jure* case such as this, "the 'hypothetical average taxpayer' is more appropriately selected by reference to a weighted mean...."

The court also found that it was proper to exclude leased or rented residential property since it is the *use* of the property rather than whether it produces income for its owner which determines whether it is commercial or industrial property. The court further held that the term "commercial and industrial property" only includes property that is subject to a property tax under State law, so that business inventory, which is not taxed in Arizona, should not be included. Finally, the court found that the plaintiffs had failed to establish any legal basis for distinguishing between centrally-assessed and locally-assessed property for purposes of calculating the assessment ratio of "commercial and industrial property" in a *de jure* case.

2. Clinchfield Railroad Company v. Lynch, 527 F. Supp. 784 (E.D.N.C., 1981), aff'd, No. 82-1049 (4th Cir. Feb. 3, 1983). The federal district court for the Eastern District of North Carolina determined that the assessment

ratios of centrally-assessed property, but not locally-assessed personal property, should be included in the determination of the assessment ratio of "other commercial and industrial property" for 4-R Act purposes. The court further found that the proper method for incorporating centrally-assessed property ratios was to calculate the median. On appeal, the Fourth Circuit Court of Appeal affirmed that centrally-assessed public service companies should be included in the calculation for determining the existence of discrimination against the railroads. It also agreed that the median was the appropriate measure of central tendency. However, it disagreed that personal property ratios could be ignored, except in extraordinary circumstances.

3. Ogilvie v. State Board of Equalization, 492 F. Supp. 446 (D.N.D. 1980), aff'd, 657 F.2d 204 (8th cir.), cert. denied, 454 U.S. 1086 (1981). The federal district court for North Dakota ruled that the railroads should be assessed at the aggregate level of assessment for all commercial and industrial property regardless of whether centrally or locally-assessed. The court further found that including personal property in the assessed value of railroads, but not other industrial and commercial property, violated The 4-R Act since, in effect, it imposed a personal property tax on railroads that was not imposed on locally-assessed business. On appeal, the Eighth Circuit Court of Appeals affirmed the district court in both respects.

4. Tennessee v. Louisville & Nashville Railroad Co., 478 F. Supp. 199 (M.D. Tenn. 1979), aff'd, 652 F.2d 59 (6th Cir.), cert. denied, 454 U.S. 834 (1981). Tennessee argued that its property tax classification system was compatible with Section 11503 because it treated railroad property similarly to some (but not all) other types of commercial and industrial property. The court rejected this argument.

5. The Atchison, Topeka and Santa Fe Railroad Co. v. Lennen, 640 F.2d 255 (10th Cir. 1981), on remand Atchison, Topeka and Santa Fe Railway Company v. Lennen, 315 F. Supp. 220 (D. Kan. 1981). The Tenth Circuit Court of Appeals decided that the State, rather than each county, was the appropriate assessment jurisdiction for purposes of determining the comparative assessment ratios, and declined to divide railroad property into real property and personal property, respectively, for purposes of comparing its assessment to other real and personal commercial and industrial property. The court further decided that it could not determine an appropriate ratio for commercial and industrial property in Kansas, and thus adopted an assessment ratio for "all other property" in the State. In determining this latter ratio, the court expressly included personal property and centrally-assessed utility property, excluding railroads.

6. The Atchison, Topeka and Santa Fe Railway Company v. Arizona, Nos. Civ. 81-1279 PHX CLH and Civ. 82-1792 PHX CLH (Consolidated) (D. Ariz., May 2, 1983); Southern Pacific Transportation Company v. Arizona, Nos. Civ. 81-1298 PHX CLH and Civ. 83-185 PHX CLH (Consolidated) (D. Ariz., May 2, 1983). The federal district court affirmed that a sales assessment ratio study should include centrally-assessed property; rejected the use of the weighted mean in favor of the median; held that the state, rather than each county, was the appropriate assessment jurisdiction; and required the inclusion of leased residential property and personal agricultural property in calculating the assessment ratio of "other commercial and industrial property."

7. The Atchison, Topeka & Santa Fe Railway Company v. Lennen, 552 F. Supp. 1031 (D. Kan. 1982), The federal district court for Kansas found that the use of the median ratio was proper and further held that, while centrally-assessed

property should be included for purposes of calculating the assessment ratio of other commercial and industrial property, personal property need not be.

REFERENCES

BICKEL, P., and FREEDMAN, D. (1981), "Some Asymptotic Theory for the Bootstrap," *Annals of Statistics*, 9, 1196-1217.

BUREAU OF THE CENSUS (1979), *1977 Census of Governments, Taxable Property Values and Assessment/Sales Price Ratios*, 2, (especially pp. 14-15, 60, 94), U.S. Department of Commerce.

COCHRAN, W. G. (1977), *Sampling Techniques*, New York: John Wiley.

FREEDMAN, D., and PETERS, S. (1984), "Bootstrapping an Econometric Model: Some Empirical Results," *Journal of Business and Economic Statistics*, 2, 150-158.

INTERNATIONAL ASSOCIATION OF ASSESSING OFFICERS (1978), *Improving Real Property Assessment: A Reference Manual*, (see especially Chapter 5), Chicago: 1313 E 60th Street.
-----(1980), *Standard on Assessment Ratio Studies*, Chicago: 1313 E 60th Street.

JAMES, G., and JAMES, R.C. (1959), *Mathematics Dictionary*, Van Nostrand.

LEHMANN, E. (1983), *Theory of Point Estimation*, New York: John Wiley.

NATIONAL ASSOCIATION OF TAX ADMINISTRATORS (1954), *Guide for Assessment Sales Ratio Studies*, Chicago: Federation of Tax Administrators, 1313 E 60th Street.

VARIAN, H. (1974), "A Bayesian Approach to Real Estate Assessment," in *Studies in Bayesian Econometrics and Statistics*, eds. S. E. Fienberg and A. Zellner, North Holland.

LITERATURE CITATIONS

The following articles are cited only as examples of other treatments of related questions.

BEHRENS, J. (1983), "The General Nature of 'the' Property Tax Today," in *The Property Tax and Local Finance*, by Harris, C., New York.

CARR, F. (1959), "Measures of Central Tendency in Assessment Ratio Studies," in *Revenue Adminstration, 1959*, Proceedings of the Twenty-Seventh Annual Conference, National Association of Tax Administrators.

CHENG, P. L. (1970), "The Common Level of Assessment in Property Taxation," *National Tax Journal*, 23, 50-65.

GASTWIRTH, J. (1982), "Statistical Properties of a Measure of Tax Uniformity," *Journal of Statistical Planning and Inference*, 6, 1-12.

GREEN, H. (1981), "Assessment Equalization and Standard Statistical Methods: An Appraisal of the Sales-Ratio Approach," *Assessors' Journal*, 16, 175-185.

REINMUTH, J. (1976), "Recent Advances in Sales-Ratio Analysis," *Assessors' Journal*, 11, 101-117.

SHENKEL, W. (1972), "Property Tax Assessment Ratios: A Critical Review," *Assessors' Journal*, 7, 3-19.

WEIL, R. (1953), "Property Tax Equalization in Illinois," *National Tax Journal*, 6, 157-167.