Volume Title: Income from Independent Professional Practice

Volume Author/Editor: Friedman, Milton and Simon Kuznets

Volume Publisher: NBER

Volume ISBN: 0-87014-044-2

Volume URL: http://www.nber.org/books/frie54-1

Publication Date: 1954

Chapter Title: Appendix A: The Reliability of the Department of Commerce Samples

Chapter Author: Milton Friedman, Simon Kuznets

Chapter URL: http://www.nber.org/chapters/c2333

Chapter pages in book: (p. 411 - 488)

APPENDIX A

The Reliability of the Department of Commerce Samples

THIS APPENDIX discusses in detail the reliability of the data on which the greater part of our analysis is based. These data were collected by the U. S. Department of Commerce as part of a study of the size of national income. Chapter 2 describes the samples we have used and summarizes the conclusions reached in this Appendix.

The relatively small size of the samples (see column 10 of Table 4) makes it specially important to investigate thoroughly their representativeness, to determine the biases to which they may be subject, and to find ways of eliminating these biases. We shall attempt to do this (1) by examining in detail the sampling methods used; and (2) by comparing (a) the distribution of the samples with the distribution of the universe, (b) different samples covering the same year with one another, (c) the Department of Commerce samples with other studies of the incomes of professional men.[1]

1 THE SAMPLING METHOD

The process of obtaining a sample involves first, the designation of a list of names to serve as the basis for sampling; second, the choice of the persons on the list to whom questionnaires are sent; third, the return of the questionnaires by the respondents; and fourth, the editing of the returned questionnaires before final use. Biases may enter at each stage: the list may be defective, the method of choosing names may not yield a truly 'random' sample, those who reply may differ from those who fail to reply, the answers of those

[1] Most of the computations cited in this Appendix do not incorporate minor revisions of the basic data made at a fairly late stage in the study. Correction of the computations would have altered none of the conclusions and hence did not seem justified. In consequence, however, there are minor discrepancies between some of the measures cited in this Appendix and supposedly identical measures in the text, text tables, or tables of Appendix B.

who reply may have a systematic bias, and the questionnaires rejected in the process of editing may differ systematically from those retained.

a The lists employed

i *The incompleteness of the lists for the earlier years.* All questionnaires requested information on income for several years from a sample of professional men selected from a list presumed to be comprehensive for the end of the period. For example, the questionnaires mailed to physicians in 1933 requested information on income for the four years 1929–32. These questionnaires were sent to a sample selected from a list of physicians in practice in 1932. Even if such a sample were entirely random for 1932 it would not be random for 1929, since the list would exclude the names of men who were in practice in 1929 but who died, retired, or left the profession for other reasons, between 1929 and 1932. The longer the period between the year for which information is requested and the year for which the list is comprehensive, the more incomplete the list will tend to be.

The resulting bias in average income depends not only on the number of men who leave the profession and hence are omitted from the list but also on their average income. If their average income were equal to the average income in the profession as a whole, there would be no bias; if it were higher than the average income in the profession, the sample average would tend to be too low; and conversely.

Since termination of independent professional practice is usually due to death or retirement, the persons omitted from the list are likely to be concentrated in the relatively high age groups. The average age of professional men at the time of death is about 69 years.[2] The average age of those omitted because of death was, therefore, less than 69 in the year in question—i.e., in one of the earlier years when they were in practice. The average age of those omitted because they retired or left the profession for other reasons was almost certainly still lower.

From the evidence in Chapter 6, it appears that average income rises for a time with increasing experience and then falls, eventu-

2 Clark, *Life Earnings in Selected Occupations*, p. 150. This is the expected average age at death of males who were 20 in 1930.

ally dropping below the average for the profession as a whole. In medicine and dentistry the average for the profession is reached after 33 or 34 years in practice, i.e., by men about 55 or 60 years old. In the other professions the rise continues longer and the fall is less rapid, so that the average for the profession is not reached until considerably later. These statements, based on data for all persons in practice, may not be valid for persons destined to leave practice in the next few years. Some men may be in a position to retire because their earnings have been better than average; others may retire or leave practice because they have been receiving such low incomes that there is no incentive for them to remain in practice.

While this evidence does not yield a clear-cut conclusion, it does suggest that those omitted from the list are concentrated in age groups whose average income is about the same as or somewhat higher than the average for the profession. If this is so, and if the average income at each age of men in practice can be assumed valid for men destined to leave practice in the next few years, the sample averages for the earlier years might be expected to be somewhat too low. Since the persons omitted from the list are likely to be more numerous and, on the average, younger the longer the period between the year for which income information is requested and the year for which the list is comprehensive, the bias in the sample averages will tend to be greatest for the earliest year covered. Hence the incompleteness of the list not only affects the absolute level of average income but also imparts an upward bias to the trend of average income over time.

In practice, this tendency is accentuated by the fact that a list purporting to be comprehensive for a given year rarely is. New entrants into a profession are difficult to trace and are almost always inadequately covered. Since the average income of new entrants is considerably below the average for the profession, their underrepresentation tends to make the sample average for the terminal year too high.

The existence of an upward bias in the trend of income over time is confirmed by our data. Such a bias would mean that the earlier of two samples for the same profession would tend to yield a higher average for an overlapping year. We show below (Sec. 2b) that this tendency is reflected in our data after correction is made for the specific biases discussed later. In the absence of detailed

data on the persons omitted from the list, it has not been possible to correct the data for this bias.

ii *The inclusion of persons not in independent practice.* All lists of professional men from which the samples were selected include not only men in independent practice but also salaried employees and persons not in professional practice.[3] Although the questionnaires emphasized that information was desired solely from men in independent practice, some of the returned questionnaires contained notations indicating that this instruction had been disregarded. It is unlikely, however, that a large number of persons not in independent practice are included in the final sample, since the many items on the questionnaires applicable only to persons in independent practice facilitated the identification of questionnaires inadvertently returned by others. Except for the elimination of questionable returns in the process of editing, no attempt has been made to correct the data for the inclusion of persons not in independent practice.

The inclusion of salaried employees on the list from which the sample was selected introduced an additional bias into the 1937 accountancy sample. The questionnaire for this sample differed from the questionnaires for the other accountancy samples and the other professions in that it requested a recipient who was a salaried employee of an accounting firm to hand the questionnaire to his employer. Consequently, an accountant with salaried professional employees would be more likely to be included in the final sample than an accountant without such employees. This would make for an overrepresentation of the more affluent accountants since accountants who have professional employees are likely to receive higher incomes than other accountants. Two bits of evidence suggest that this bias is unimportant. First, the 1937 accountancy sample yielded lower average incomes for the overlapping years than the earlier samples (Table 5). Second, it seems likely that a larger proportion of firms than of individual practitioners employ other accountants; yet the proportion of firm members in the 1937 sam-

[3] In addition, the lists include the names of some retired or deceased persons. Replies by the retired persons for the years prior to retirement would tend to counteract the bias discussed in the preceding section. In fact, however, it is clear from the returned questionnaires that most retired persons either do not reply or return the questionnaires with notations that they have retired but with no information for years prior to retirement.

ple is about the same as in the 1935 sample and decidedly lower than in the 1933 sample (Table 46).

iii *The medical, legal, and accountancy lists.* The lists used in selecting the medical, legal, and accountancy samples—the Directory of the American Medical Association, the *Martindale-Hubbell Law Directory,* and the mailing list of the American Society of Certified Public Accountants—seem excellent apart from the general defects already noted. Since in all three professions the right to practise is limited to individuals licensed by the state, the inclusion of new practitioners is relatively easy. The number of names on these lists checks very closely with the totals recorded in other sources.[4]

iv *The dental list.* The list used for dentists was restricted to members of the American Dental Association, which included only about 46 per cent of the dentists in practice when our samples were taken.[5] It is clear from the available evidence that members

[4] The 1931 medical directory lists 156,839 physicians; the 1930 *Census of Population,* 153,803. The 1930 issue of the legal directory (relating to lawyers in practice in 1929) lists 141,501 lawyers; the 1930 Census, 139,059 as engaged in professional service. (The Martindale-Hubbell directory makes no attempt to list lawyers employed by nonlegal enterprises. This explains why the comparison is made with the number of lawyers listed by the Census as engaged in professional service.) The mailing list of the American Society of Certified Public Accountants included between 13,000 and 15,000 names during the period in question. The American Institute of Accountants (with which the Society recently merged) estimated that there were approximately 16,500 certified public accountants in 1937.

[5] The estimate of 46 per cent is based on (1) data supplied by the American Dental Association on the number of members in 1932 and 1934, and (2) estimates of the total number of dentists in the United States in 1932 and 1934. The estimates of the total number of dentists are based on straight-line interpolation between 71,055, the number of dentists listed in the *Census of Population* for 1930, and 75,225, the estimate of the number of dentists in 1936 given in Table 1. The years 1932 and 1934 were used because our samples were drawn from the membership lists for those years.

According to these data, 45.5 per cent of all dentists were members in 1932 and 47.2 in 1934. The figure we use, 46.2, is an average of the two, with the 1932 and 1934 figures weighted respectively 4 and 3, the number of years covered by the corresponding samples.

These estimates are for all dentists, whereas our interest centers in independent practitioners. Some indication of the maximum error involved in using the same percentage for independent practitioners can be obtained by assuming that *all* members are in independent practice, and accepting unpub-

of the American Dental Association are not representative of all dentists. In a study of 1929 incomes made in 1930 and covering slightly over 5,000 dentists in twenty states, Leven found "that the net incomes of those who reported themselves as members were, on the average 30 per cent higher than those of the dentists who did not claim membership".[6] In a study made in California in 1934 and based on approximately 1,600 returns, the 1933 average net professional income of members of the American Dental Association was found to be 33.4 per cent higher than that of nonmembers.[7] Both percentages are based on fairly large samples. Their closeness, while not conclusive as to their reliability, gives some reason for confidence in them.

However, neither figure can be used, without further investigation, to adjust the average incomes from our samples. In both studies a dentist was classified as a member or nonmember on the basis of his answer to a question requesting him to indicate the societies to which he belonged. Leven found a wide discrepancy between the membership records of the American Dental Association and the information supplied by the dentists themselves: 49 per cent of the dentists in the 20 states covered by Leven's sample were carried on the membership rolls of the Association in 1929, whereas Leven estimates that 68 per cent would have classified themselves as members if all had returned questionnaires.[8] The California figures show a similar discrepancy. The discrepancy is presumably attributable to three groups: individuals who were formerly members of the Association but had been dropped for nonpayment of dues or for other reasons; individuals belonging to local or other dental societies but not to the national association;

---

lished estimates by the American Dental Association of the total number of dentists in independent practice (these estimates seem, if anything, slightly too low). These assumptions yield 57.8 as the percentage of independent practitioners who were members. Assuming that a smaller proportion of members than of nonmembers are in independent practice would of course yield a figure below 46.2 per cent; but there seems no particular numerical assumption that deserves special recognition as setting a lower limit.

[6] *Practice of Dentistry*, p. 200.

[7] *California Medical-Economic Survey*, Table 71. The average income of members was $3,022; of nonmembers, $2,265. These averages are based on 1,074 members and 541 nonmembers.

[8] *Practice of Dentistry*, pp. 12, 200.

individuals who had become members so recently that their names had not been entered on the membership rolls.

Since our samples were chosen from the membership rolls of the American Dental Association, only individuals listed as members by the Association could have been included. Hence, the relevant figure for our purposes is the percentage excess of the average income of this group of members over the average income of other dentists. The data available from the two studies cited do not yield a precise estimate of this figure. The best we can do is to set limits within which it may reasonably be supposed to lie. On the basis of Leven's figures, these limits are 17 and 42 per cent; on the basis of the California figures, they are 20 and 50 per cent. More or less arbitrarily, we have selected 30 per cent as the best estimate of the percentage excess of the average income of dentists on the membership rolls of the Association over the average income of other dentists.[9]

Since, as already indicated, approximately 46 per cent of all dentists were on the membership rolls of the American Dental Association when our samples were selected, a difference of 30 per cent between the incomes of members and nonmembers would imply that the average income of all dentists is 87.6 per cent of the average income of members alone. The final estimates of the average incomes of dentists given in Table 11 were computed by using this correction factor to adjust for the restriction of the samples

[9] Let k be the percentage excess of the average income of dentists on the membership rolls of the Association over the average income of the remaining dentists; x_m, the average income of dentists on the membership rolls of the Association; x_q, the average income of dentists who classify themselves as members but are not on the membership rolls of the Association; and x_n, the average income of all other dentists. The relation between k and the figure of 30 per cent cited by Leven or the figure of 33.4 per cent from the California study depends on the relation of x_q to x_m and x_n, and on the relative size of the three groups. According to Leven's figures, if x_q were equal to x_m, k would equal 17 per cent. On the other hand, if x_q were equal to x_n, k would equal 42 per cent. According to the California figures, these two extreme assumptions would give values of k of about 20 and 50 per cent respectively. Presumably, the correct value of k lies between these two extremes, since the self-designated members appear to be somewhat of a mixture of the other two groups and might be expected to have an average income between x_m and x_n. The figure of 30 per cent selected for k implies, on the basis of Leven's figures, that x_q is approximately 12 per cent greater than x_n and 17 per cent less than x_m; on the basis of the California figures, that x_q is 21 per cent greater than x_n and 13 per cent less than x_m.

to members.[10] This is the only table in which corrected data are presented. The data for dentists in all other tables should be interpreted as referring solely to members of the American Dental Association.

The correction factor used to adjust for the bias in average income is admittedly based on slender evidence. However, a recent study, the results of which became available only after our estimates had been made—and, indeed, published[11]—provides striking confirmation of their validity. This study was made by the Department of Commerce and is based on a sample of over 7,000 dentists. The sample was similar to those we analyze but covered both members and nonmembers. According to this study the average income of nonsalaried dentists was $4,267 in 1929, and $2,188 in 1933.[12] Our final estimates (Table 11) are $4,176 for 1929, and $2,178 for 1933; i.e., the difference is about 2 per cent for 1929, and less than 0.5 per cent for 1933.

The exclusion of nonmembers presumably affects not only average income but also other aspects of the frequency distribution of income by size. In particular it seems reasonable to expect that the

[10] Since the two figures on which the correction factor of .876 is based cannot be determined exactly but are selected from a range of possible values, it is of interest to investigate the effect on the correction factor of choosing different values. In the following table the values actually used are italicized; the other hypothetical values are approximately the largest and smallest values that, on the basis of the preceding analysis, could reasonably have been used.

PERCENTAGE DIFFERENCE BETWEEN INCOMES OF MEMBERS AND NONMEMBERS (k) TAKEN AS	VALUE OF THE CORRECTON FACTOR IF PERCENTAGE OF MEMBERS IS TAKEN AS		
	40	*46.2*	60
20	.900	.910	.933
30	.862	*.876*	.908
40	.829	.846	.886

[11] Incomes from Independent Professional Practice, 1929–1936, *Bulletin 72–73* (National Bureau of Economic Research, 1939), p. 10.

[12] Lasken, 'Incomes of Dentists and Osteopathic Physicians', Table 2. The returns included a disproportionately large number of members. Consequently the averages cited represent weighted averages of the incomes of members and nonmembers, the weights being the estimated total number in each group. Because of a marked increase in the membership of the American Dental Association between the dates when the samples we analyze were taken—1933 and 1935—and the date the later Department of Commerce sample was taken—1938—it is not possible to make a direct comparison between the percentage excess of the average income of members that we have used and that shown by the later sample.

frequency distribution will be more concentrated and less 'skew' than if a more comprehensive list had been employed. Unfortunately, no way could be found to correct for these deficiencies.

v *The engineering list.* The list used for consulting engineers appears seriously defective. It was compiled with the aid of the American Engineering Council, representatives of which examined the directories of four national engineering societies and checked the names of engineers thought to be consultants. The number of names obtained in this way totaled 3,286; yet, according to Table 1, there were in 1930 approximately 10,000 consultants. The list clearly excludes consultants who were not members of engineering societies. In addition, it excludes engineers whose status as consultants was not known by the persons who examined the directories. Both deficiencies operate in the same direction—to exclude the less prominent and well-known. Since these individuals might be expected to receive relatively low incomes, the final sample has a definite, and possibly fairly substantial, upward bias. Unfortunately, it has not been feasible to adjust the data for this bias.

b The selection of the persons to whom questionnaires were sent

A questionnaire was sent to every *n*-th dentist on the list of the American Dental Association. This procedure is entirely valid and should yield a 'random' sample of the list employed, though, for reasons noted in the preceding section, not of the universe of dentists. Similarly, no bias could have been introduced into the engineering sample at this stage, since a questionnaire was mailed to every person whose name was on the list. The method of selecting the persons to whom questionnaires were sent was less straightforward for the other professions and introduced a number of significant biases.

i *The nonrandomness of the 1937 medical and legal samples by states.* For both physicians and lawyers, questionnaires were sent to persons selected by taking a specified number of names from each page of the relevant professional directory. For the 1933 and 1935 medical samples and the 1935 legal sample, the same number of names was taken from each page. For the 1937 medical and legal samples, however, the number of names per page was deliberately

varied from state to state.[13] These two samples are therefore admittedly nonrandom as among states. To correct for this nonrandomness, all computations for these two samples have been made for each state separately, and the results weighted by the estimated total number of practitioners before being combined. (More exactly, each return has been weighted by the ratio of the estimated total number of practitioners in the state to the number in the sample for that state.)

For physicians, the weights used are the estimated number of physicians in active practice in each state in 1936.[14] The weights therefore include salaried physicians as well as those in independent practice whereas we use them in connection with data for independent practitioners alone. The proportion of all physicians in active practice in the United States in 1929 who were salaried employees has been estimated as about 15 per cent.[15] It is doubtful, therefore, that the inclusion of salaried physicians greatly affects the percentage allocation of the total among states; and, it is solely the latter, of course, that is relevant from the point of view of weighting. In any event, there are no data that could be used to estimate the proportion of salaried physicians by states.

For lawyers, the weights used are the number of lawyers in each state listed in the 1936 *Martindale-Hubbell Law Directory*.[16] This directory lists lawyers in practice in 1935 and includes salaried lawyers as well as independent practitioners. There seems little reason to suppose, however, that either deficiency seriously af-

[13] Since these samples were taken with the expectation that the data would be used in an analysis of income by states, it was desired to have a sample for each state sufficiently large to be used for this purpose. The same sampling ratio for all states would have necessitated a larger total sample than was feasible. Consequently, the sampling ratio was varied from state to state, a larger proportion of names being taken for smaller states.

[14] The estimate for each state was derived by multiplying the number of physicians in that state listed in the 1936 Directory of the American Medical Association (this count is given in the directory itself) by the 1931 ratio for the same state of the number of physicians in active practice to the total number of physicians. These ratios were based on Leland (*Distribution of Physicians*, p. 17) who gives, for each state, the total number of physicians listed in the 1931 directory, and the number listed as in active practice, retired, and not in practice. This tabulation is the most recent available.

[15] Leven, *Incomes of Physicians*, pp. 103–4.

[16] These figures were furnished by Martindale-Hubbell.

fects the relative weight assigned each state. And, as for physicians, there is no feasible alternative.

ii *The size of community bias in the medical and legal samples.* The selection of a specified number of names from each page of a directory—the procedure followed for physicians and lawyers—will yield a 'random' sample only if all pages contain the same number of names or if any variation in the number of names per page is independent of the characteristics to be studied. Examination of the medical and legal directories reveals that they satisfy neither requirement. The number of names per page varies considerably, and the variation is associated with size of community, which, in turn, is associated with income.

In both directories, the names of professional men are listed by communities. The communities are separated by a blank space and the name of each community and some information about it are given. More space per page is needed for this purpose, and hence less space remains for the listing of names, the smaller the communities listed on a page. By itself this would tend to make the number of names per page less for small communities than for large ones.

In the Directory of the American Medical Association, this tendency is more than counterbalanced by another: the number of lines devoted to each physician varies and tends to be greater for large communities. There are two reasons for this. First, a physician in a large community is likely to have a longer address, since it includes a street and number whereas the post office designation is ordinarily sufficient for a physician in a small community. Second, the medical directory lists the professional societies [17] of which the physician is a member, and, if he is a specialist, indicates by symbols his specialty. Specialists and members of the professional societies listed are concentrated in the larger communities.

The average number of names per column [18] of the medical directory is shown in Table A 1 for communities of various size. The number of names per column is approximately the same for all communities over 10,000 in population, but is considerably less

17 Other than the American Medical Association, membership in which is designated by printing the name in capital letters, but in the same size type.

18 There are three columns to a page.

in these communities than in smaller ones.[19] Such variation clearly tends to introduce a bias into a sample chosen by taking the same number of names from each page: communities for which the total number of names per page is relatively large tend to be underrepresented. Small communities therefore tend to be underrepresented in the medical samples.

TABLE A 1

Average Number of Names per Column of the Directory of the American Medical Association, by Size of Community

SIZE OF COMMUNITY (1930 Census)	AVG. NO. OF NAMES PER COLUMN
500,000 & over	44
100,000–500,000	45
50,000–100,000	47
25,000– 50,000	44
10,000– 25,000	46
2,500– 10,000	62
Under 2,500	54

Count of sample columns from 1936 Directory of the American Medical Association. In all, 70 columns containing 3,293 names were counted.

Though the variation in the number of names per page is small, the expected bias is revealed by a comparison of the distribution of the samples by size of community class with the corresponding distribution of all physicians. This comparison shows a slight underrepresentation of communities under 10,000 in population.[20]

[19] The differences between the averages in the two smallest community classes and the others cannot be attributed to chance. The significance of the differences among the mean values for the seven groups was tested by the analysis of variance. A value of 9.7 was obtained for F, the ratio of the mean square between groups to the mean square within groups. For the number of degrees of freedom available, a value of 3.1 would be exceeded by chance only once in a hundred times. On the other hand, the differences among the first five classes are not significant. The value of F for these five classes alone was 0.9.

[20] Leland, *Distribution of Physicians*, Table 42, gives the distribution of physicians in active practice in 1931 by size of community, based on a count of the 1931 Directory of the American Medical Association. The distributions by size of community of the physicians reporting income for 1932 in both the 1933 and 1935 samples were compared with this 1931 distribution for all physicians. (No attempt was made to compare the 1937 sample because of its nonrandomness among states.)

The effect of correcting for the size of community bias was tested by correct-

Comprehensive data are available, at least for 1931, on the distribution of physicians by size of community and geographic region.[21] These data could be used to correct for the size of community bias by weighting the sample data for each size of community class by the corresponding number of physicians in the universe. Before deciding to make this correction, however, a test was made, using 1932 arithmetic mean incomes from both the 1933 and 1935 samples, in order to see how much these would be altered by the correction. This test, summarized in Table A 2, suggests that the weighting does improve the results, but that the improvement is so slight as not to be worth the labor involved.[22]

---

ing the sample distributions on the basis of Table A 1 (see footnote 24 for the exact method used), adjusting the totals for the corrected distribution to make them equal to the totals for the uncorrected distributions, and computing χ^2 between the sample distributions and the distribution of all physicians. The smaller the value of χ^2, the less the discrepancy between the sample and the universe. For both samples, the correction for the size of community bias lessens the discrepancy between the sample and the universe.

	VALUE OF χ^2 FOR	
	1933 sample	1935 sample
Not corrected	65.77	16.96
Corrected	39.87	14.01

21 *Ibid.*

22 Table A 2 gives for the United States and the nine Census regions the unweighted and weighted arithmetic mean incomes and the difference between them. The unweighted means exceed the weighted means for the United States in both samples, for six out of nine regions in the 1933 sample, and for five out of nine, in the 1935 sample. Moreover, in both samples, the positive differences are, on the average, considerably larger in absolute value than the negative differences. These results conform to expectation. The two community classes that have the largest number of names per page and hence should be underrepresented in the sample include the smallest communities, in which the average income of physicians is relatively low (see Ch. 5). The use of correct weights should therefore raise the mean income.

Table A 2 therefore affords additional evidence of the existence of the size of community bias. At the same time, the difference between the weighted and unweighted means are all small. The correction raises the mean for the country from the 1933 sample by slightly over two per cent and the corresponding mean from the 1935 sample by about one-tenth of one per cent. The differences are not much larger for individual regions, for which the number of returns is much smaller and hence the possibility of random variation much greater. Comparable means from the two samples differ considerably more than weighted and unweighted means from the same sample.

TABLE A 2

Effect of Size of Community Bias on Arithmetic Mean Income

Physicians, 1932: 1933 and 1935 Samples

	1933 SAMPLE				1935 SAMPLE			
	No. of returns	Arithmetic mean income in 1932		Difference between unweighted and weighted means	No. of returns	Arithmetic mean income in 1932		Difference between unweighted and weighted means
		Unweighted	Weighted [1] (dollars)			Unweighted	Weighted [1] (dollars)	
New England	160	4,676	4,712	−36	104	4,161	4,121	+40
Middle Atlantic	480	3,977	3,907	+70	316	3,725	3,784	−59
E. N. Central	490	2,781	2,718	+63	301	3,089	3,107	−18
W. N. Central	242	3,599	3,656	−57	164	2,736	2,723	+13
S. Atlantic	233	3,406	3,214	+192	110	3,217	3,294	−77
E. S. Central	136	2,313	2,319	−6	76	1,434	1,410	+24
W. S. Central	173	2,646	2,494	+152	96	2,216	2,222	−6
Mountain	64	2,250	2,062	+188	57	3,198	2,957	+241
Pacific	282	4,202	4,025	+177	142	3,038	2,950	+88
U. S.	2,288 [2]	3,434	3,347	+87	1,392 [3]	3,107	3,104	+3

[1] Weights used are the number of physicians in 1931 in each region and size of community. R. G. Leland, *Distribution of Physicians in the United States* (American Medical Association, 1936), Table 42.

[2] Includes 28 returns for which either size of community or region was unknown.

[3] Includes 26 returns for which either size of community or region was unknown.

Hence the original sample has been used and no attempt has been made to correct for the size of community bias.

In the *Martindale-Hubbell Law Directory,* the tendency for the number of names per page to be greater for large communities than for small ones because of the space needed to separate communities and to describe them is reinforced by the variation in the amount of space allotted individual lawyers or firms: some names are printed in larger type than others, and some names are followed by a brief statement describing the activities of the lawyer

TABLE A 3

Average Number of Names per Page of the *Martindale-Hubbell Law Directory,* by Size of Community

SIZE OF COMMUNITY (1930 Census)	AVG. NO. OF NAMES PER PAGE
1,500,000 & over	148
500,000–1,500,000	141
250,000– 500,000	131
100,000– 250,000	109
Under 100,000 *	86

For communities over 500,000 with the exception of Milwaukee, Martindale-Hubbell supplied the number of lawyers listed in the 1936 directory. A count was made of the number of pages assigned each of these communities in the 1936 directory, and the number of names per page computed by division. The averages for the other size of community classes are based on counts of sample pages selected from the 1937 directory. In all, 47 pages containing 4,369 names were counted.

* In the substantive analysis by size of community, this class is broken down into four classes: 25,000–100,000, 10,000–25,000, 2,500–10,000, and under 2,500. There seemed, however, to be no significant differences among these classes in the average number of names per page. Hence, they were grouped together for the present purpose.

or firm and listing the members and associates of the firm. This larger amount of space is allotted only to lawyers who have been in practice a certain number of years and who have attained prominence in their communities. It so happens that the proportion of lawyers allotted the larger amount of space is higher in small communities than in large ones, presumably as a result of the concentration in large cities of both young lawyers and salaried lawyers.

These two factors make for wide differences in the average number of names per page. As Table A 3 shows, the average num-

ber of names per page decreases consistently with size of community and is 72 per cent larger for communities over 1,500,000 in population than for communities under 100,000. These differences are very much larger than in the medical directory and in the opposite direction. They might be expected to lead to a considerable underrepresentation of lawyers in large communities.

In order to eliminate this considerable bias, the data for both legal samples were grouped by size of community and all computations made for such groups.[23] In combining the size of community classes an adjustment was made on the basis of Table A 3.[24]

The conclusions that the two legal samples are subject to bias and that Table A 3 can be used to correct this bias, have so far been based solely on *a priori* reasoning. These conclusions should be tested empirically before being accepted, first, because other factors might conceivably have counteracted the presumed bias, second, because the analysis has been entirely in terms of the sample of questionnaires sent out, whereas the corrections must be applied to the sample of questionnaires returned.

Unfortunately, there are no comprehensive data on the number of lawyers by size of community with which to compare the distributions of the samples.[25] Indeed, if such data were available,

[23] In practice, the data for the 1935 sample were grouped by size of community and region, and for the 1937 sample, because of the necessity of weighting by states, by both size of community and state.

[24] This adjustment was made by multiplying the number of questionnaires in each size of community class by a factor proportional to the average number of names per page in Table A 3. For example, to get the same proportionate sample as from communities under 100,000, the sample from communities over 1,500,000 should have been 1.72 ($\frac{148}{86}$) times as large as it actually was. In order to adjust the sample, it is therefore necessary to treat each questionnaire from a community over 1,500,000 as if it represented 1.72 questionnaires; and similarly for other size of community classes. In combining averages, the average for each size of community was weighted by the adjusted number of questionnaires. In combining frequency distributions by size of income, the adjusted number of questionnaires, rather than the original number, was added for each income class.

[25] Offhand, it may seem that such data could be compiled from the *Census of Population,* at least for communities above 25,000. The main reason they cannot is that the Census classifies lawyers by residence, whereas the questionnaires were mailed to the business address. A comparison of the number of lawyers listed in the *Martindale-Hubbell Law Directory* for some of the larger cities with the number listed in the Census reveals large differences. For example, the

the bias could be corrected more adequately and directly: each size of community class could be weighted by the total number of lawyers in that class. In the absence of more comprehensive data, we have relied on a count made available by Martindale-Hubbell of the lawyers listed in the 1936 directory for each state and each city over 500,000 except Milwaukee. For the nine states containing one or more of these cities, the distribution of all lawyers by size of community classes was compared with the distribution of the sample both before and after correction for the size of community bias. This comparison confirms both the existence of the bias and the validity of the correction based on Table A 3.[26] A similar

Census gives a total of 1,898 lawyers for Boston, Martindale-Hubbell, 4,374. Yet the totals for the state of Massachusetts agree very well; the Census total is 6,940, the directory total, 7,150. This difference is, of course, consistent; the Census gives smaller totals for large communities and larger totals for small communities. Two additional minor difficulties are that the Census includes lawyers employed by both legal and nonlegal enterprises, and that it is for 1930, whereas the earlier of the two samples was mailed in 1933.

[26] For each of the nine states, the questionnaires from the 1935 sample reporting income in 1934 and the questionnaires from the 1937 sample reporting income in 1936 were grouped into the size of community classes in Table A 3. These distributions were then adjusted for the firm member bias discussed in Sec. 1b iv. Both the adjusted and the unadjusted distributions were then corrected for the size of community bias by the method outlined in footnote 24 above.

This procedure gave four distributions by size of community for each state and each sample: (1) not corrected for the size of community bias and not adjusted for the firm member bias, (2) not corrected but adjusted, (3) corrected but not adjusted, (4) corrected and adjusted. In order to make the four distributions comparable, their totals were made equal to the unadjusted and uncorrected total number of returns. The size of community classes were then combined to conform with the classification available for the universe; namely, communities over 1,500,000, 500,000–1,500,000, and under 500,000.

Even if the samples were entirely 'random', the ratio of the sample to the estimated universe should decrease as size of community increases, since the estimated number of lawyers includes not only independent but also salaried lawyers, and these tend to be concentrated in the larger communities. The resulting discrepancy between the sample and the universe is in the same direction as that arising from the size of community bias. Hence, while sample distributions corrected for the size of community bias should conform more closely to the estimated distribution of the universe than uncorrected distributions, they should still differ significantly from it.

Comparison of the four distributions for each state and sample with the distribution of the universe reveals that in practically all cases the discrepancy between the uncorrected distributions and the universe is in the expected

conclusion emerges if the distribution of all lawyers by states is

---

direction, and that correction for the size of community bias tends to reduce but not to eliminate the discrepancy. As a single objective measure of the extent of the discrepancy, χ^2 was computed for each of the four distributions available for each sample and each state (except Maryland, for which the sample was too small). The larger the value of χ^2, the greater the discrepancy between the sample and the universe. Hence the correction is to be considered successful if the corrected distributions yield smaller values of χ^2 than the uncorrected. (Since the totals of all distributions were arbitrarily made the same, the values of χ^2 measure solely the extent of the discrepancy between the distribution of the sample and of the universe.)

	NUMBER OF STATES IN WHICH	
	χ^2 for corrected distribution smaller than for uncorrected (i.e. correction successful)	χ^2 for corrected distribution larger than for uncorrected (i.e. correction unsuccessful)
1935 sample		
Not adjusted for firm member bias	7	1
Adjusted for firm member bias	8	0
1937 sample		
Not adjusted for firm member bias	4	4
Adjusted for firm member bias	7	1

On the whole, the correction was clearly successful. Moreover, since a correction introduced at random would be as likely to raise χ^2 as to lower it, the differences shown in the above table cannot reasonably be attributed to chance alone. Adding the values of χ^2 for the separate states yields a similar conclusion.

	SUM OF THE VALUES OF χ^2 FOR EIGHT STATES			
	1935 sample		*1937 sample*	
	Not adjusted for firm member bias	Adjusted for firm member bias	Not adjusted for firm member bias	Adjusted for firm member bias
Not corrected	169.67	169.22	38.47	40.98
Corrected	85.78	82.58	28.95	18.10

One final computation may be cited. For the 1935 sample, the distributions for nine states were combined, and χ^2 computed for each of the four sample distributions.

	χ^2 FOR NINE STATES, 1935 SAMPLE	
	Not adjusted for firm member bias	Adjusted for firm member bias
Not corrected	173.3	169.8
Corrected	79.3	77.0

This computation, like the others, indicates that the correction lessened the discrepancy between the sample and the universe. A similar computation for the 1937 sample could not validly be made, since the sample is not random among

compared with the distribution of the sample before and after correction for the size of community bias.[27]

iii *The overrepresentation of certain types of physicians and lawyers.* In selecting the medical and legal samples, the particular names taken from each page of the directory were determined by laying a straight edge marked off at equally spaced intervals along a column of names. The names that fell opposite the marks were included in the sample. The probability of a particular person being chosen is therefore proportional to the space his listing occupies. A person whose listing occupies two lines has twice as large a chance of being included in the sample as a person whose listing occupies only one line.[28] As noted in the preceding section, in both the medical and legal directories the number of lines devoted to each person varies. In the medical directory the variation is fairly limited: many names require one line and few names more than three. In the legal directory, on the other hand, the variation is much greater: from one line to ten or fifteen or even, occasionally, more.

The samples for both professions, but especially for lawyers, will therefore tend to overrepresent persons whose listings occupy

---

states. It should be noted that except for the unadjusted and uncorrected distributions, the probability of the computed χ^2 being exceeded by chance cannot be judged by the use of the ordinary sampling distribution of χ^2. The reason is that the totals of the remaining distributions were arbitrarily adjusted to equal the totals for the unweighted and uncorrected ones. Nor is there any sampling distribution available for testing the significance of the differences between the values of χ^2 computed in this way.

[27] As in the preceding footnote, χ^2 was used as a measure of the discrepancy and the totals of the sample distributions were arbitrarily made the same. Because the 1937 sample is not random among states, the test could be made only for the 1935 sample.

	χ^2 FOR 1935 SAMPLE BY STATES	
	Not adjusted for firm member bias	Adjusted for firm member bias
Not corrected	325.3	256.4
Corrected	239.3	176.9

[28] The statement that the probability of a particular person being chosen is proportional to the space his listing occupies is precise only for listings requiring a smaller number of lines than the space between the markings on the straight edge. Every person whose listing occupies more space than this must be chosen; the probability that he will be included in the sample is therefore unity no matter how much space his listing occupies.

more than the average amount of space.[29] As noted above, in the medical samples, these will tend to be specialists and physicians with long addresses; in the legal sample, the older and more prominent lawyers. Since both groups tend to have incomes higher than the average for all practitioners, the income data from the samples may be expected to have an upward bias.

Comparison of the percentage of specialists in the 1937 medical sample—the only sample for which this information is available—with the percentage indicated by other studies does not confirm the existence of this suspected bias (see Sec. 2c below). Similarly, comparison of our legal samples with studies of lawyers in Wisconsin and New York County fails to confirm its existence, since average incomes from our samples are lower than averages from the other studies. However, these tests are fragmentary and unsatisfactory and cannot be taken as establishing the absence of this bias. Unfortunately we have been unable to make more adequate tests. No correction has been made for this bias.

iv *The firm member bias in the legal and accountancy samples.* The legal and accountancy samples were selected from lists of persons, not of professional units (i.e., firms plus individual practitioners), but each person to whom a questionnaire was sent was requested, if a member of a firm, to reply for the firm as a whole. By the procedure followed, a firm had a greater chance of being included in the sample than an individual practitioner, since it was included if *any one* of its members was included; and the larger the firm, the greater its chance of being included. For example, suppose that a 2 per cent sample is taken from a universe of 5,000 individual practitioners and 5,000 members of two-member firms. The sample of persons to whom questionnaires are sent will tend to include 100 individual practitioners and 100 firm members. If all who receive questionnaires reply and if no firm has more than one of its members in the sample,[30] the questionnaires returned

29 Any variation in the amount of space devoted to each person that is associated with size of community has already been allowed for in the preceding section. What is relevant here is solely the variation among persons in communities of the same size.

30 On the average, with the assumed figures, one firm would have both members selected: the probability of a particular person being selected is 1/50; of both members of a firm being selected, this number squared, or 1/2500; and there are assumed to be 2,500 firms.

will cover 100 individual practitioners and 200 firm members, since each firm member will reply not only for himself but also for his partner. The final sample will include twice as many firm members as individual practitioners, although there are an equal number of each in the assumed universe.

It follows that the method of selecting lawyers and accountants to whom questionnaires were to be sent, combined with the wording of the questionnaires, yield 'outgoing' samples that contain a known proportionate excess of firm members.[31] If, like the legal samples, the sample is small relatively to the universe, it will tend to contain approximately twice as many members of two-member firms, three times as many members of three-member firms, and so on, as it should for a representative sample. If, like the accountancy samples, the sample is fairly large relatively to the universe, the proportionate excess will be smaller, but still considerable.[32]

The overrepresentation of firm members in the 'outgoing' sample might be expected to lead to a similar overrepresentation in the sample of questionnaires returned—the 'incoming' sample. Whether the overrepresentation will be larger or smaller in the

[31] This bias does not affect the medical and dental samples because almost all physicians and dentists practise as individuals. It does not affect the engineering sample because questionnaires were sent to all persons on the list; consequently, every member of every firm on the list received a questionnaire.

[32] Let p be the proportion of all names on the list included in the sample and let $q = 1 - p$. Then q is the probability that a particular name will not be included in the sample, q^m, the probability that none of m specified names will be included, and $1 - q^m$, the probability that at least one of m specified names will be included, i.e., the probability that at least one member of a firm of m members, and hence the firm itself, will be included. If p is small, q^m is approximately equal to $1 - mp$, as can be seen by expanding $(1 - p)^m$. In this case, $1 - q^m = mp$. Since p is the proportion of individual practitioners included in the sample, it follows that if the sample is small relatively to the universe, the proportion of firms of size m in the sample will be m times the proportion of individual practitioners in the sample. Since $1 - mp$ is always less than q^m, $1 - q^m$ is always less than mp and $(1 - q^m)/p$ is always less than m. The difference is negligible when p is small, but increases as p increases. The difference reflects the increasing possibility that more than one member of the firm will be included in the sample. At the limit, when p is unity, i.e., when a 100 per cent sample is taken, q is zero, and $1 - q^m$ is unity, or equal to p. The formulae in this footnote are valid so long as the universe is fairly large. Otherwise, a slight correction is needed. See *Money Disbursements of Wage Earners and Clerical Workers in Five Cities in the West North Central-Mountain Region, 1934–36*, U. S. Bureau of Labor Statistics Bulletin 641, pp. 384–90.

incoming than in the outgoing sample depends on the relative proportion of firm members and of individual practitioners who refuse to reply.[33] Since we have no information on these proportions, we have tested the existence and magnitude of the presumed bias indirectly.

For lawyers, we have placed major reliance on a comparison between the estimated percentage of all lawyers who are firm members and the percentage of lawyers in our samples who are firm members, before and after correction for the firm member bias. The estimate of the percentage of all lawyers who are firm members is based on a sample count of the 1933 *Martindale-Hubbell Law Directory*.[34] Unfortunately the directory listings do not permit an entirely satisfactory classification. They distinguish between firm members and other lawyers, but not between lawyers in independent practice and salaried employees.[35] A survey made by the New

[33] If the probability that a firm will refuse to reply were independent of the number of members of the firm who received questionnaires and equal to the probability that an individual practitioner will refuse, it can easily be shown that the proportionate excess of firms of each size would be the same in the incoming as in the outgoing sample, i.e., that the probability of a firm of m members being included in the incoming sample is given by the formulae of the preceding footnote, with p, as there, equal to the ratio of the number of individual practitioners in the *outgoing* sample to the number in the universe. If the probability that a particular member of a firm will refuse were the same as the probability that an individual practitioner will refuse, and if, when more than one member of a firm received questionnaires, the probabilities that the different members would refuse were independent of one another (implying that a firm is more likely to reply the more members receive questionnaires), it can be shown that the probability that a firm of m members will be included in the incoming sample is different from the probability that it will be included in the outgoing sample and is given by the formulae of the preceding footnote, with p equal to the ratio of the number of individual practitioners in the *incoming* sample to the number in the universe. There is little basis on which to choose between these alternative sets of assumptions. Moreover, other plausible assumptions would give different results.

[34] The sample count included every 50th page of the directory. While involving much less work, such a sample is, of course, subject to larger sampling errors than one including every 50th name, because of the larger size of the sampling unit and consequently the smaller number of units. At the same time, it is not subject to any of the biases arising from taking the same number of names from each page.

[35] Only lawyers employed by other lawyers or by legal firms are listed in the directory, which makes no attempt to list lawyers employed by nonlegal enterprises.

York County Lawyers Association indicates that approximately 20 per cent of all lawyers in active practice in New York County are employed by legal firms.[36] Since the percentage of salaried employees is probably at least as large in New York County as elsewhere, we have used this figure to convert the results of the sample count into an upper estimate of the percentage of all lawyers who are firm members. The percentage computed directly from the sample count gives a lower limit.

The 1935 legal sample was adjusted for the firm member bias by weighting the number of members of two-member firms by one-half, the number of members of three-member firms by one-third, etc. The sum of these weighted numbers was used as the corrected number of firm members.[37] The 1937 legal sample was adjusted in the same way, except that more nearly exact weights were used and that the weights varied from state to state. This was necessary because in some states questionnaires were mailed to a sizable proportion of all lawyers.[38]

The comparison between the estimated percentage of firm members in the universe and the percentage computed from the samples (Table A 4) reveals that the 1935 sample conforms to expectations but the 1937 sample does not. The unadjusted percentage of firm members in the 1935 sample is much higher than the estimated percentage in the universe; the adjusted percentage of firm members, a trifle lower.[39] This indication of a firm member bias in the 1935 legal sample is confirmed by a comparison between the distribution of all lawyers by states and the distribution of the

[36] *Survey of the Legal Profession in New York County*, p. 12. The exact percentage is 21.14.

[37] The adjustment was made separately for each size of community class in order to correct for the size of community bias.

[38] The adjustment was made separately for each size of community class in each state in order to correct for the size of community bias and the nonrandomness among states.

[39] As a measure of the discrepancy between the estimated percentage of firm members in the universe and the percentage in the sample, we can use χ^2 computed from a 2x2 table giving the number of firm members and individuals counted in the directory and the corresponding numbers in the sample. The value of χ^2 before allowance for the inclusion of salaried employees in the directory is 174.6 for the unadjusted sample numbers, and .59 for the adjusted ones. The corresponding values of χ^2, after allowance for the inclusion of salaried employees, are 71.1 and 19.3, respectively.

TABLE A 4

Test of Existence of Firm Member Bias

Lawyers, 1935 and 1937 Samples

	NO. OF INDEPENDENT PRACTITIONERS	% OF INDEPENDENT PRACTITIONERS CLASSIFIED AS FIRM MEMBERS
Sample from 1933 *Martindale-Hubbell Law Directory*	8,109	24.0–30.0[1]
1935 sample[2]		
Not adjusted for firm member bias	1,332[6]	44.5
Adjusted for firm member bias[3]	933[7]	23.0
1937 sample[4]		
Not adjusted for firm member bias	1,168[6]	25.3
Adjusted for firm member bias[5]		12.9

[1] Lower limit assumes that no salaried employees are included in the directory. The upper limit assumes that 20 per cent of all practitioners in the directory were salaried employees and that these were all included in the sample count as individual practitioners.

[2] Compilation is of schedules reporting net income for 1933.

[3] Each size of firm was weighted by the reciprocal of the number of members.

[4] Compilation is of schedules reporting net income for 1936.

[5] Because the sampling ratio varied from state to state, different weights were used for each state. The weights used were $p/(1 - q^m)$, where p is the proportion of names in the particular state to whom questionnaires were sent, $q = 1 - p$, and m is the number of members in the firm.

[6] Total number of persons represented on questionnaires before correction for size of community or firm member bias.

[7] Total number of persons represented on questionnaires after adjustment for firm member bias but before correction for size of community bias.

1935 sample by states, before and after adjustment for the firm member bias. There is much closer agreement after adjustment than before adjustment.[40]

The unadjusted percentage of firm members in the 1937 legal sample is about the same as the estimated percentage in the universe; the adjusted percentages, very much lower. Adjustment for the presumed firm member bias, which yields closer agreement for the 1935 sample, has exactly the opposite effect on the 1937 sample. This difference between the two samples is puzzling; and we have

[40] See the values of χ^2 given in footnote 27 above.

been unable to explain it satisfactorily.[41] Hence the 1937 sample must be viewed with considerable scepticism.

In the light of these results, we have corrected the 1935 sample for the firm member bias by weighting each size firm by the reciprocal of the number of members in that size firm,[42] but have made no attempt to correct the 1937 sample. For the 1937 sample, firm members and individual practitioners have been combined without weighting.

For accountants, the absence of any evidence on the proportion of firm members rules out the kind of test for the existence of the presumed firm member bias in the incoming sample that we used for lawyers. We have relied instead on two even more indirect tests, both of which indicate that the bias in the outgoing samples is also present in the incoming samples.

The more satisfactory of these tests is based on data for New York State alone. (However, New York contains almost one-third of all certified public accountants.) These data were derived from sample counts of (1) the New York section of the 1933 yearbook

41 The only explanation that has occurred to us hinges on the content and form of the questionnaires. The questionnaire sent out in 1935 covered a single sheet and requested information solely about the legal enterprise—individual practitioner or firm, as the case might be. It therefore requested the same information from and involved the same amount of trouble for individual practitioners and firm members. The questionnaire sent out in 1937, on the other hand, was in two parts; the top part requested information about the enterprise, the bottom part about the practitioner. Both parts were to be filled out by individual practitioners. A firm member was supposed to fill out the top part for the firm as a whole, the bottom part for himself. Additional copies of the bottom part were enclosed and were to be filled out by firm members other than the one who received the questionnaire. (See Appendix C.) The trouble involved was therefore greater for a firm member than for an individual practitioner, since the former was asked to distribute additional copies of the bottom part. In addition, there may have been a reluctance to return the questionnaire unless all firm members filled out the bottom part; refusal by a single firm member might have resulted in a refusal by the firm as a whole.

Both factors would tend to make the refusal rate for firm members higher relatively to the refusal rate for individuals in the 1937 sample than in the 1935 sample. It seems dubious, however, that this factor alone could have produced a difference in results as great as that reflected in Table A 4.

42 These are not the exact weights, which, according to footnote 32, would be $p/(1 - q^m)$. For a sample as small as the legal sample, however, the difference between the two sets of weights is so slight that the gain in accuracy through the use of exact weights would not repay the extra labor involved.

of the American Institute of Accountants and (2) the 1938 register of the New York State Society of Certified Public Accountants.[43] Both lists contain enough information to permit the identification of persons in independent practice and of firm members. On the basis of this sample count, we estimated the number of certified public accountants in independent practice who were members of (1) the Institute, (2) the Society, (3) both the Institute and the Society, (4) the Institute but not the Society, (5) the Society but not the Institute. For each group, we estimated the percentage who were firm members. (See first five lines of Table A 5.)

Like most professional associations, these two include the more prominent and affluent practitioners. In accountancy, they would tend to include a larger proportion of firm members than of individual practitioners, since, on the average, firm members receive higher incomes and are older than individual practitioners (see Ch. 6, Sec. 3). At the time the count was made, the Institute was more biased in this direction than the New York State Society,[44] as is clear from the first two lines of Table A 5. The difference is even greater between accountants who were members of both associations and of only one: 71 per cent of the independent practitioners who belonged to both associations were firm members; but only about 40 per cent of those who belonged to one association. *A priori,* the percentage of firm members might be expected to be still smaller among certified public accountants who belonged to neither the Society nor the Institute.

This suggests a method of testing whether our samples overrepresent firm members: we can construct from our samples two estimates of the percentage of firm members among accountants who belonged to neither association—one, before adjusting the samples for the presumed firm member bias, a second, after adjusting the samples—and see which appears more consistent with this *a priori* presumption. These estimates, given in Table A 5, were computed as follows. The percentage of firm members in the original sample was applied to the estimated total number of certified public accountants in independent practice in New York State. This gave the total number of firm members and of individual

[43] Every third page of the former and every tenth page of the latter was included in the sample count.

[44] The Institute has since merged with the American Society of Certified Public Accountants.

practitioners in the state. Subtracting the number of firm members and of individual practitioners who belonged to one or both of the associations (as estimated from the sample counts) gave the number of firm members and of individual practitioners who belonged to neither association. The percentages in the column headed 'unadjusted' are based on these numbers. This process was then re-

TABLE A 5

Test of the Existence of Firm Member Bias

Certified Public Accountants

CERTIFIED PUBLIC ACCOUNTANTS IN NEW YORK STATE WHO BELONG TO	ESTIMATED % OF INDEPENDENT CERTIFIED PUBLIC ACCOUNTANTS WHO ARE MEMBERS OF FIRMS	
American Institute of Accountants *	63.1	
N. Y. State Society of Certified Public Accountants *	47.8	
Both Institute and Society	71.1	
Institute but not Society	37.8	
Society but not Institute	42.4	
	Samples	
Neither Institute nor Society: estimate based on	*Unadjusted*	*Adjusted*
1937 sample for 1936	49.5	24.0
1937 sample for 1934	43.3	18.8
1935 sample for 1934	38.9	16.7
All certified public accountants: estimate based on		
1937 sample for 1936	48.3	34.9
1937 sample for 1934	45.0	32.2
1935 sample for 1934	42.7	31.1

* Data derived from sample counts of the New York section of the 1933 yearbook of the American Institute of Accountants and the 1938 register of the New York State Society of Certified Public Accountants, respectively.

peated, except that the percentage of firm members applied to the estimated total number of accountants was based not on the original samples but on the samples after adjustment for the presumed firm member bias.[45] This yielded the percentages in the column headed 'adjusted'.

Table A 5 shows that accepting the original sample as valid

[45] The adjustment was made by weighting the number of members of firms of each size by $p/(1 - q^m)$, where p is the proportion of names included in the outgoing sample, q is $1 - p$, and m is the number of members in the firm. For the 1933 and 1935 samples p was taken as 0.5, for the 1937 sample, as 0.4. See footnote 32 above.

contradicts the *a priori* presumption stated above, since the percentage of firm members among certified public accountants who belonged to neither association is as large as or larger than the corresponding percentage among accountants who belonged to only one association. On the other hand, adjusting for the firm member bias yields percentages that are consistent with expectations: the percentage of firm members is uniformly smaller among accountants who belonged to neither association than among accountants who belonged to only one. This test therefore confirms the existence of the presumed bias in all samples.

The second test involved a comparison of the frequency distributions by size of firm computed from two different samples covering the same year. Two comparisons were made for each of two years: one, between distributions not adjusted for the firm member bias; a second, between distributions adjusted for this bias. For both years, the adjusted distributions differed less from one another than the unadjusted.[46]

These tests give no reason for doubting the existence of the firm member bias in any of the samples. Consequently, all three accountancy samples have been adjusted for the firm member bias by weighting each size firm inversely to the overrepresentation theoretically to be expected in the outgoing sample.[47]

c The return of questionnaires by the respondents

The problem of selecting a random 'outgoing' sample, to which Sections a and b have been devoted, is common to all investigations using samples, whether in the social, biological, or physical sciences. It is a problem treated extensively in the literature of

[46] χ^2 was used as a measure of the discrepancy between the two distributions. In order to make the χ^2's comparable, the totals of the adjusted distributions were made equal to the totals of the unadjusted distributions.

SAMPLES COMPARED	NO. OF DEGREES OF FREEDOM	VALUES OF χ^2 Unadjusted distributions	Adjusted distributions
1935 sample and 1933 sample for 1932 (i.e., for persons reporting income in 1932)	4	74.92	67.91
1937 sample and 1935 sample for 1934	5	26.96	16.48

[47] See footnote 45 for the weights used.

theoretical statistics, one that is susceptible to *a priori* analysis of the type suggested by the familiar model of balls in an urn.

In most fields other than the social sciences, the representativeness of the sample is determined entirely by its method of selection. Not so in the social sciences. If a ball is drawn from an urn it cannot refuse to become a member of a sample; but if an individual is sent a questionnaire, he can refuse to reply. The outgoing sample may be completely random, yet the incoming sample hopelessly biased if the persons who reply differ significantly and consistently from those who fail to reply.

Not only does the representativeness of a sample in the social sciences depend on the behavior of the objects sampled; so also may the validity of the information obtained. A ball chosen from an urn cannot tell the investigator that it is black when it is really white. The investigator may make a mistake; but his errors ordinarily are his own and subject to check. An individual asked to state his net income from independent professional practice may state his gross income, his income from both independent and salaried practice, or his income less personal expenses; if he thinks his answer will affect policy, he may deliberately overstate or understate his income; and so on.

This distinction between the outgoing and incoming samples that is so important for samples like ours has received little attention in the theoretical literature of statistics, first, because it does not arise in the kinds of samples with which theoretical statisticians have been mainly concerned, second, because it is not susceptible to *a priori* analysis. Whether individuals who reply differ from those who do not reply and whether the answers of those who reply are subject only to random errors or to consistent and biased errors are essentially empirical questions, the answers to which will depend greatly on the particular circumstances surrounding the inquiry—the agency sponsoring it, the purpose of the inquiry, the way the questions are worded, the methods used to encourage replies, and so on.

Unfortunately, there is little empirical evidence directly relevant to these questions. Our own samples give some; and prior studies of professional income add a bit more.

The only evidence from our samples on the characteristics of persons who fail to reply is provided by the 1937 medical and

legal samples, for which we know the number of questionnaires sent to persons in each state, as well as the number returned.[48]

These data reveal fairly conclusively that the refusal ratio (the ratio of questionnaires not returned to questionnaires sent) differs significantly among states.[49] Moreover, differences among the refusal ratios are fairly similar for the two professions: states that have a high proportion of refusals among physicians also tend to have a high proportion of refusals among lawyers.[50]

The question immediately arises whether the proportion of persons in a particular state who refuse to reply is correlated with the average income of the persons who do reply. Such a correlation between states would suggest a similar correlation within states. For example, if the average income of those who replied were relatively high in states that have a high proportion of refusals, it would seem reasonable to conclude that within each state separately persons with high incomes are less likely to reply than persons with low incomes.[51] For the 1937 medical and legal sam-

[48] Offhand, a comparison between the distribution by states, regions, etc., of the returned questionnaires and the corresponding estimated distribution of the universe might seem to provide evidence on the nature of the refusal bias. However, discrepancies revealed by such comparisons are the product of errors in the list employed, the method of selecting the sample, and the estimates of the universe as well as of differences in the willingness of various groups to reply. To disentangle the resultant composite is an almost impossible task. Such comparisons are valuable as evidence on the over-all reliability of the data, and are used for this purpose in Section 2a below.

[49] The discrepancy between the geographic distributions of the returned questionnaires and the questionnaires sent was measured by computing χ^2 between the two distributions.

	NO. OF DEGREES OF FREEDOM *	χ^2	PROBABILITY THAT OBSERVED VALUE OF χ^2 WOULD HAVE BEEN EXCEEDED BY CHANCE
1937 medical sample	48	118.73	less than .00000005
1937 legal sample	47	95.86	less than .00005

* One less than the number of classes in the distribution. For the medical sample, the classes were the forty-eight states and the District of Columbia. For the legal sample, Delaware could not be included separately because there were too few questionnaires.

[50] The rank difference correlation between the refusal ratios for physicians and lawyers is .435 with a standard error of .144.

[51] This test is not vitiated by the use of sample rather than population average incomes unless (a) there is a significant relationship between income and willingness to reply and (b) this relationship differs among states (or other units),

ples, the correlation between the refusal ratio and average income is negligible, suggesting that willingness to reply is not related to income.[52]

For the particular samples analyzed—the 1937 medical and legal samples—the variation in the refusal ratio among states is of no importance, since these samples have been weighted by states. The findings are important for the inferences they suggest about the remaining samples and about the factors leading to refusals within states. For the other samples also, the refusal ratio presumably varies significantly among states and the incoming samples cannot be expected to be entirely representative of the universe even if the outgoing samples are. However, in the absence of any correlation with average income, the variation in the refusal ratio does not of itself introduce a bias into the average incomes computed from the samples. And the two samples analyzed give no direct indication of such a correlation between states; hence no indirect indication of such a correlation within states.

---

being direct in some and inverse in others. If the relationship is similar in all states, the sample averages will give the correct ranking of states by income. For example, suppose in all states nine-tenths of the individuals with incomes over $5,000 but only half of those with incomes below $5,000 refuse to reply; and suppose that the percentage of individuals with incomes over $5,000 is larger in state A than in state B. The sample averages will have a downward bias for both states, and this bias will be larger for state A than for state B. Nevertheless, the sample average for state A will be larger than the sample average for state B.

[52] The rank difference correlation between the refusal ratio and average income is +.023 for physicians, and —.064 for lawyers. The standard error of each is .144.

Another hypothesis tested was that differences among the refusal ratios were related to the political complexion of the states. Since the questionnaires were distributed by a government agency it might be supposed that persons in sympathy with the administration would be more likely to respond than persons not in sympathy with it. This hypothesis was tested by computing for both lawyers and physicians rank difference correlations first, between the refusal ratio and the percentage that the democratic presidential vote in 1936 was of the total vote, and second, between the refusal ratio and the change from 1932 to 1936 in the percentage that the democratic presidential vote was of the total vote. For physicians, the correlations obtained were —.04 and +.05 respectively; for lawyers, +.08 and +.07. None of these is significant, since the standard error of each is .146. While these correlations do not support the hypothesis, it should be noted that the refusal ratios were correlated with the political complexion of all voters in the state, whereas the relevant correlation would be with the political complexion of the specific professional group; and the former may be a poor index of the latter.

The evidence from our samples on the characteristics of those who fail to reply can be supplemented by evidence from three other studies of professional income—the studies of the 1929 incomes of physicians and dentists made by the Committee on the Costs of Medical Care, and the California Medical-Economic Survey sponsored by the California Medical Association.

For physicians, the Committee on the Costs of Medical Care made both a nationwide study and intensive surveys of particular communities. In connection with one of the latter—on incomes of Vermont physicians—"a supplementary study of a sample taken from the physicians who failed for one reason or another to participate in the state-wide survey disclosed the fact that the 57 per cent [who participated] were fairly representative of the total." [53] The nationwide sample of the Committee on the Costs of Medical Care yielded somewhat different results. A follow-up letter was sent to a sample of the physicians who failed to reply to the original questionnaire. "A comparison between the data received on the first letter and those received in response to the follow-up letter, for the same areas, show important differences in the two subsamples. The second contains a larger proportion of small incomes and its median income is about $1,000 lower than the median of the first returns." [54] Leven is disposed to explain this difference as being "due, to some extent at least, to the nature of the appeal in the follow-up letter. The physicians were urged to make returns even though they felt that their collections and charges in 1929 and 1930 were typical neither of their own practice nor of the incomes of physicians in general. This . . . undoubtedly resulted in returns from a disproportionately high number of those whose incomes were low because of special circumstances—sickness, old age, partial retirement, and the like." [55]

In the dental survey of the Committee on the Costs of Medical

[53] Leven, *Incomes of Physicians*, p. 8. In a footnote to this statement Leven indicates that "a test study was made in three Vermont communities. The gross incomes of the physicians who had not returned the mailed questionnaire were obtained by personal contact and, by adding these to the returns procured by mail, a 100 per cent sample was made available. The additional data changed the average by only $25. The reasons for the physicians' failure to reply to the mailed questionnaire were tabulated and it was established that failure to reply was not in any way associated with the size of income."

[54] *Ibid.*, p. 13.

[55] *Ibid.*, p. 14.

Care, questionnaires were sent to a 25 per cent sample of a presumably comprehensive list of all dentists in twenty states. Usable questionnaires were returned by 66 per cent of those circularized. The representativeness of this group was tested by comparing their age distribution with that of all dentists, by comparing the percentage of the respondents who were members of the American Dental Association or component societies with the corresponding percentage for all dentists, and, for three states, by obtaining independent estimates of the gross incomes of dentists who failed to reply.

The first test suggested that the only serious discrepancy was in the most recent graduating class, which seemed to be underrepresented in the sample. However, it is not clear whether this reflected a defect in the list used for sampling [56] or the unwillingness of young men to reply. In any event, correction of the sample for this deficiency would have lowered the estimated average income by only 1.4 per cent.[57]

The second test indicated that a larger proportion of members than of nonmembers replied. Since members have higher incomes than nonmembers, this bias resulted in an underrepresentation of dentists with low incomes.[58] The third test yielded similar results. In each of three states, Colorado, Georgia, and Wisconsin, the average gross income based on both the replies to questionnaires and the estimates of local dental committees for dentists who failed to reply was lower than the average based on the questionnaire data alone. The difference varied from $28 or less than one-half of one per cent to almost $500 or about eight per cent.[59]

It is clear that in this study of the Committee on the Costs of Medical Care, the refusal ratio was larger among dentists with low incomes than among dentists with high incomes, although the difference was not great. It is doubtful, however, that this result can be generalized. The study was conducted under the auspices of the American Dental Association; in addition, several of the devices used to stimulate replies must have been more effective with members of dental societies than with nonmembers. For ex-

[56] See Sec. 1a i above.

[57] Leven, *Practice of Dentistry*, p. 199.

[58] *Ibid.*, p. 200. The indicated correction in the average net income of all general practitioners was about 1.5 per cent.

[59] *Ibid.*, p. 203.

ample, special letters were sent by officers of various state societies; committees of component societies solicited individuals who had not replied; the purpose and plans of the study were presented and discussed at society meetings; and articles about the study appeared in a variety of dental journals.[60] That under these circumstances a larger percentage of members than of nonmembers should have replied is not surprising, and this in turn might explain the smaller percentage of returns from individuals with low incomes.[61]

The California Medical-Economic Survey sent two successive follow-up letters to physicians and dentists who did not return the original questionnaire. On the whole, the three successive samples differ little for either physicians or dentists. There are no significant differences among the arithmetic mean incomes, median incomes, standard deviations, or income distributions themselves.[62] The one consistent difference among the samples is that for both

[60] *Ibid.*, pp. 210–2.

[61] Unfortunately no data are available to test this hypothesis. The data published for the three states do not segregate members from nonmembers and it is therefore impossible to determine whether the lower average income for all dentists is due solely to a greater weighting of nonmembers.

[62] The measures for the different samples are summarized in the accompanying table.

	PHYSICIANS			DENTISTS		
INDIVIDUALS REPLYING TO	Median income	Arithmetic mean income	Standard deviation of income distribution	Median income	Arithmetic mean income	Standard deviation of income distribution
	(dollars)					
Original letter	2,700	3,651	3,523	2,300	2,745	1,918
First follow-up	2,700	3,483	3,311	2,400	2,710	1,836
Second follow-up	2,500	3,497	3,426	2,700	3,071	1,964
All letters	2,700	3,572		2,500	2,769	

California Medical-Economic Survey, pp. 71, 74. The standard deviations were computed from the frequency distributions.

The significance of the differences among the arithmetic mean incomes for each profession was tested by the analysis of variance and found not significant. In each set of standard deviations even the largest and smallest do not differ significantly and hence the whole set scarcely can.

The significance of the differences among the frequency distributions was tested by the χ^2 test. For physicians, χ^2 is 25.96; the number of degrees of freedom, 22; and the probability of the observed χ^2 being exceeded by chance, .25. For dentists the corresponding figures are 15.01, 14, and .38. For neither are the differences larger than might have been expected from chance.

physicians and dentists the percentage of general practitioners among those replying was larger for each successive letter; but even these differences are quantitatively so small that they do not appear to be statistically significant.

Though too meagre to justify a definite conclusion, the evidence presented on this important question of the refusal bias suggests first, that there are significant geographical differences among the percentage refusing to reply; second, that these differences are not associated with geographical differences in income; third, that while within each community or region there may be differences between the willingness of different income groups to respond, these differences cannot be large; and fourth, that if such differences do exist, the lower income groups are probably the least willing to reply.

The evidence from our samples on the second of the questions noted above—namely, the accuracy of the answers of those who reply—is qualitative rather than quantitative, and derives from a detailed study of individual questionnaires. This study suggests a number of possible sources of bias in the replies, none of which appear to be particularly serious.

Perhaps the most important source of bias arises from requesting information on income for a number of years. The difficulty of answering accurately for all years tends toward the insertion of approximate incomes, and, even more important, toward the insertion of the same or very nearly the same income for each year. While this tendency may not affect the average for the period as a whole, it does damp fluctuations in income over the period. This tendency is, of course, more important the longer the period for which information is requested. It therefore affects most the samples collected in 1937.

Some respondents seem to have been confused about the meaning of 'gross' and 'net' professional income and of 'professional expenses' to be deducted from the former in arriving at the latter. Although the instructions were fairly explicit even in the first set of questionnaires, and were made increasingly detailed and precise in each succeeding set, there are indications that some respondents included nonprofessional income and salaries from professional work with income from independent professional practice, or deducted living expenses as well as strictly professional expenses in deriving net income, or interpreted 'net' in the sense of 'net tax-

able income'. However, relatively few questionnaires seem to have been affected by such errors. So far as they affect the final results, they probably, on balance, make for a downward bias in net income.

The examination of the schedules suggests that the answers of accountants and dentists are probably most accurate, and those of lawyers and physicians least accurate. Because of the shorter period covered, the two earlier sets of questionnaires—those sent out in 1933 and 1935—are probably more accurate than the last set—those sent out in 1937.

d The editing of the questionnaires

The last possible source of bias in obtaining the sample of usable questionnaires is the processing of returned questionnaires before their actual use in the statistical analysis—the weeding out of the 'unusable' returns and the correction of 'obvious' errors. Many unusable questionnaires offer no serious problem—the respondent will have made a notation to the effect that he is 'not in practice', 'employed on salary', 'retired', or that he has reported income from 'incidental part-time work'; gross income will be less than net income; gross income less reported expenses will not equal net income; a firm member will indicate that he is reporting solely his own income; etc. Other questionnaires cannot be classified so simply. There is an inevitable tendency to regard as suspect a questionnaire that deviates widely in any respect from other questionnaires or from what is expected on the basis of other knowledge. The real problem of editing is how to decide whether each of these is really an error or simply an 'extreme' case. If, implicitly or explicitly, the rule 'exclude wherever doubtful' is followed, the extremes will inevitably tend to be eliminated, with resulting biases depending on their characteristics. On the other hand, if the rule 'include wherever doubtful' is followed, the extremes will be included, but so also will the erroneous returns.

The choice between these alternatives is necessarily arbitrary. Both involve errors of unknown magnitude and direction. In editing the present samples, an attempt was made to steer between the Scylla of excluding doubtful cases and the Charybdis of including them; but the course followed varied somewhat for the different samples. The 1933 samples were edited by different persons and at a different time than the later samples. In the editing

TABLE A 6

Number of Questionnaires Returned

Number Usable, and Number not Usable, by Reasons

SAMPLE & YEAR SENT OUT	TOTAL [1]	USABLE [2]	NOT USABLE	NOT USABLE BECAUSE: Deceased or retired	Not in independent practice [3]	No usable information	Other reasons [4]
Physicians							
1 (1933)	2,882	2,438	444	103	215	65	61
2 (1935)	1,686	1,588	98	17	66	14	1
3 (1937)	1,647	1,577	70	10	9	50	1
Dentists							
1 (1933)	1,609	1,499	110	8	68	20	14
2 (1935)	1,171	1,122	49	11	20	6	12
Lawyers							
1 (1935)	1,161	1,050	111	22	68	9	12
2 (1937)	1,260	1,063	197	18	79	91	9
Certified public accountants							
1 (1933)	977	679	298	21	171	19	87
2 (1935)	1,255	1,062	193	14	143	12	24
3 (1937)	853	752	101	15	57	19	10
Consulting engineers							
1 (1933)	804	415	389	59	227	43	60

[1] Excludes questionnaires returned by post office as undeliverable.

[2] Includes all questionnaires containing any usable information. The numbers in this column are therefore somewhat larger than the numbers on which most of the results in the text are based.

[3] Includes questionnaires on which the respondent indicated he was a salaried employee or 'not in public practice'. For accountants and engineers also includes questionnaires for incorporated firms.

[4] Includes questionnaires rejected for miscellaneous reasons such as: respondent indicated independent professional activity was purely incidental to other full-time work; questionnaire duplicated another returned by member of the same firm; respondent practised outside continental United States.

of the 1933 samples, Scylla exercised the stronger attraction while in the editing of the later samples the course was set somewhat closer to Charybdis. This difference is brought out clearly by Table A 6, which gives data on the number of questionnaires discarded for various reasons. For each of the professions for which more than one sample was taken, the percentage of schedules discarded is largest for the 1933 sample.

It is doubtful that the difference in procedure could have led to

serious discrepancies between the samples; or either procedure to a considerable bias in any one sample. Many of the questionnaires discarded were eliminated on the basis of explicit annotations on the returns. Most of the remainder, those in the discarding of which the judgment factor was decisive, are included in the category entitled 'no usable information', and even in the earlier samples are relatively few. The major differences among the professions are in the category 'not in independent practice' and these accord with the known characteristics of the professions: salaried employment is most frequent in accountancy and engineering, the two professions that show the largest percentage of questionnaires discarded for this reason.

2 TESTS OF THE RELIABILITY OF THE DATA

An examination of sampling methods can, at best, establish a presumption that certain biases are present and certain others, absent. But even though no biases are discovered, or whatever biases are discovered can be corrected, the final sample cannot be judged free from bias or the data accepted as accurate and reliable. If we can think of a dozen possible sources of bias, it would be naive to suppose that there were not a hundred overlooked. This need for additional tests of the reliability of our data is enhanced by our inability to evaluate properly certain recognized possible sources of bias—for example, differences between persons who return questionnaires and those who do not. We have applied three types of tests to our data: (a) comparison of the distribution of each sample by geographic units with the estimated distribution of all practitioners; (b) comparison of different samples for the same profession with one another; (c) comparison of our samples with other studies.

a Comparison with geographic distribution of all practitioners

The comparisons that we have made between the distributions of the samples and the estimated distributions of all practitioners are summarized in Table A 7. For physicians, the distributions compared are by size of community class and region;[63] for dentists, lawyers, and certified public accountants, by states. No compari-

[63] We have used the nine Census regions and seven size of community classes (see Ch. 5, footnote 3). The total number of classes is less than 63 because all size of community classes are not represented in all regions and because some classes containing few returns were combined for the χ^2 analysis.

TABLE A 7

Comparison between Distributions of Samples by Geographic Units and Estimated Distributions of All Practitioners

Physicians, Dentists, Lawyers, and Certified Public Accountants

PROFESSION & SAMPLE	GEOGRAPHIC UNITS	UNIVERSE WITH WHICH SAMPLE IS COMPARED	NO. OF DEGREES OF FREEDOM	χ^2	P[1]
(1)	(2)	(3)	(4)	(5)	(6)
Physicians					
1933	Region and size of community	No. in 1931 [2]	57	236.47	less than .000000001
	Regions only	"	8	55.55	less than .000001
	Size of community only	"	6	65.77	less than .0000001
	Size of community within each region [3]	"	49	164.50	less than .000000001
1935	Region and size of community	"	54	103.58	.00003
	Region only	"	8	31.29	.0001
	Size of community only	"	6	16.96	.009
	Size of community within each region [3]	"	46	71.37	.008
Dentists					
1933	States	Membership in A.D.A., Dec. 31, 1935 [4]	40	130.98	less than .000000001
	"	*No. from Census of Population*, 1930	40	179.07	"
1935	"	Membership in A.D.A., Dec. 31, 1935 [4]	33	86.99	.0000001
		No. on Apr. 1, 1936 [5]	33	135.11	less than .000000001
Lawyers					
1935	States	No. in 1935 [6]	40	325.26	"
Certified Public Accountants					
1935	States	No. in 1937 [7]	30	89.43	.000000006
1937	"	"	30	143.31	less than .000000001

[1] P is the probability that the observed χ^2 would be exceeded by chance.
[2] Leland, *Distribution of Physicians*, Table 42.
[3] χ^2 was computed separately for each region, and the resulting values summed.
[4] R. P. Thomas, 'Dental Survey', *Journal of the American Dental Association and the Dental Cosmos*, Jan. 1938, p. 158.
[5] *Ibid.*, p. 155. The figure for Illinois was reduced from 11,370 to 6,000 to correct for an obvious overestimate. See notes to Table 1, Ch. 1.
[6] Based on a count of the *Martindale-Hubbell Law Directory* for 1936 made available by Martindale-Hubbell.
[7] Estimates made by American Institute of Accountants.

sons have been made for consulting engineers, since there are no data on the number of consulting engineers either by regions and size of community classes, or by states.

As a measure of the discrepancy between the sample and universe distributions, we use the statistic χ^2, entered in column 5 of Table A 7. The values of χ^2 are not directly comparable, since they are affected by the number of classes in the distributions. They can be compared by reference to column 6, which gives *P*, the probability that the observed values of χ^2 would be exceeded by chance. The smaller the probability, the larger the discrepancy between the sample and universe distributions.

For all professions, the estimated distributions of the universe have been compared with the original sample distributions, before correction for biases. The original distributions have been used because values of χ^2 based on the corrected distributions cannot be interpreted in probability terms. In judging Table A 7, however, it should be borne in mind that the discrepancies it reveals for lawyers and accountants are partly explained by the size of community and firm member biases noted above, and that the data on which the substantive analysis in the text is based have been corrected for these biases.[64] No comparisons have been made for the 1937 legal and medical samples because they are not random among states.

The striking feature of Table A 7 is the extreme discrepancies between the sample and universe distributions. The two differ least, as measured by the probabilities associated with χ^2, for the 1935 medical sample. Yet even for this sample, the largest value of *P* is .009, i.e., the discrepancy between the size of community distributions of the sample and universe is so large that it would be exceeded by chance alone only 9 times in a thousand. For the other medical sample and the other professions, the smallest discrepancy is so large that it would be exceeded by chance alone less than one time in a million.

[64] Comparisons for lawyers and accountants based on both the original and adjusted distributions have already been presented above (see footnotes 26, 27, and 46). For accountants the differences between the comparisons based on the original distributions and those based on the adjusted distributions are relatively small, but for lawyers adjustment for the firm member and size of community biases accounted for almost half of the discrepancy between the sample and the universe.

These large discrepancies are not subject to a clear and unambiguous interpretation; they may reflect errors not in the samples but in the estimated distributions of the universe. The samples supposedly include only persons in independent practice; the available data on the distribution of all practitioners uniformly include not only persons in independent practice but also salaried employees and some persons who are not in practice either because they have retired or for other reasons. Moreover, the distributions of all practitioners are estimates subject to a considerable margin of error and are available for the year in which the sample was chosen for only three comparisons. The inclusion of salaried employees is the most serious deficiency: they not only number almost one-sixth of all physicians and dentists and almost one-third of all lawyers and certified public accountants, but also are more concentrated in large and prosperous communities.

The comparisons in Table A 7 for dentists are instructive in indicating the importance of using the correct universe. Though our samples were restricted to members of the American Dental Association, they were compared with both members and all dentists. For each sample, χ^2 is lower for the former comparison—27 per cent lower for the 1933 sample, 36 per cent lower for the 1935 sample.[65]

Errors in the estimated distributions of all practitioners may therefore explain part, possibly a large part, of the observed discrepancies between the sample and universe distributions. It seems unlikely, however, that they are entirely responsible for these discrepancies. If they were, the discrepancies might be expected to be least for physicians and dentists, the professions which contain fewest salaried employees. Yet, according to Table A 7, the discrepancies are of about the same order of magnitude for these professions as for lawyers and accountants.

For each profession included in Table A 7, we can compute the ratio of the number of practitioners in the sample from a particular state to the estimated total number of practitioners in that state. The existence of discrepancies between the sample and universe distributions means that these ratios vary from state to state. To the extent that the discrepancies reflect factors common to all professions, the ratios for different professions will be corre-

[65] The χ^2's are directly comparable because the number of degrees of freedom is the same.

lated. The inclusion of salaried employees in the estimated total number of practitioners is one such factor, since salaried employees in the different professions are presumably concentrated in the same states. Similarly there may be some factors affecting the willingness of individuals to reply that differ in strength from state to state but affect all professions alike.[66]

To test the correlation, the ratios for each of the four professions were ranked in order of magnitude.[67] Though by no means identical, the four sets of ranks show a degree of similarity that would be exceeded by chance only 16 times in 100,000. The measure of similarity, χ_r^2, is almost half as large as its maximum possible value, i.e., as the value that it would attain if the ranks were identical.[68] This finding suggests first, that much of the apparent nonrepresentativeness of the samples is attributable to factors affecting all samples alike, such as errors in the estimated distributions of all practitioners or general forces affecting willingness to reply; second, that such factors are not entirely responsible for the apparent nonrepresentativeness of the samples; and third, that much of the apparent nonrepresentativeness is attributable to factors affecting particular professions or samples rather than to factors affecting all professions and samples alike.

The nonrepresentativeness of our samples means that no great confidence can be placed in our data on the number and proportion of professional men in various states, regions, or size of community classes. This presumptive unreliability of the geographic distributions of our samples is in itself not serious, since we have little interest in using our samples to study the geographic distribution of professional men. But it inevitably arouses suspicion

[66] Since the data were collected by a government agency, the attitude toward the party in power might be such a factor.

[67] Ratios based on the 1935 samples were used for all professions. For dentists, the number of members of the American Dental Association rather than the total number of dentists was the denominator of the ratio. For accountants, the numerator of the ratio was the number of accountants in the sample after correction for the firm member bias, and for lawyers, the number of lawyers in the sample after correction for both the firm member and size of community biases.

[68] For the test of similarity, see Friedman, 'The Use of Ranks', pp. 675–701, especially pp. 675–80, 694–5. The value of χ_r^2 is 77.37; the mean value on the chance hypothesis, 46; and its standard deviation, 8.31. The probability stated above assumes χ_r^2 normally distributed. The value of χ_r^2 that would be obtained if the ranks were identical is 184.

about the reliability of the data on average income or the distribution of income by size. If the states (or other geographic units) underrepresented in the samples tended to have relatively many professional men with low incomes, income data for groups of states would have an upward bias, and conversely. Moreover, if there were such a bias for groups of states, it would be reasonable to infer a similar bias within each state separately.

Fortunately, there appears to be no correlation between average income and biases in the geographic distributions of the samples. Rank difference correlation coefficients between the sample average income and the ratio of the sample to the estimated universe are uniformly small (Table A 8).[69] Four of the eight coefficients computed are less than their standard errors, and all are less than twice their standard errors; i.e., all might easily reflect merely chance variation.[70] This result was foreshadowed by similar correlations described in Section 1c, and it is confirmed by others described below. The apparent absence of any relation between average income and the biases in the geographic distribution of our samples suggests that income figures for groups of states will not be contaminated by the non-randomness of the geographic distributions; such figures will of course be subject to random errors but not, on this score at least, to bias.

[69] For physicians, the correlations were computed both by states and by region and size of community; for the other professions, only by states. See footnote 51 for a discussion of the validity of the test.

[70] Despite the risk of attempting to explain too much, it may be worth noting that the observed coefficients are entirely consistent with known biases in the samples and known deficiencies in the estimates of the universe. The medical sample underrepresents small communities, where incomes are relatively low. Hence, this factor alone would tend to produce a positive correlation between average income and the ratio of the sample to the universe, when the correlation is by region and size of community. Both observed correlations of this type for physicians are positive. In all professions, salaried employees tend to be concentrated in large and prosperous communities, where incomes are relatively high. Our estimates of the universe are too large, and our ratios of sample to universe too low, for states that contain many salaried employees. This factor, then, would tend to produce a negative correlation. Four of the five coefficients for the professions other than medicine are negative. For physicians, the known bias in the sample would be less important relatively to the deficiencies in the universe in a distribution by states than in a distribution by region and size of community. The correlation coefficient for physicians by states not only is less than by region and size of community, but also is negative.

TABLE A 8

Rank Difference Correlation Coefficients between Sample Average Income and Ratio of Sample to Universe

Physicians, Dentists, Lawyers, and Certified Public Accountants

PROFESSION & SAMPLE	GEOGRAPHIC UNITS	UNIVERSE	YEAR OF AVERAGE INCOME	NO. OF ITEMS CORRELATED	RANK DIFFERENCE CORRELATION COEFFICIENT	STANDARD ERROR
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Physicians						
1933	Region and size of community	No. in 1931 [1]	1932	59	.239	.131
1935	"	"	1932	59	.158	.131
1935	States	No. in active practice, 1934 [2]	1934	49	—.015	.144
Dentists						
1935	"	Membership in A.D.A., Dec. 31, 1935 [3]	1932	49	—.216	.144
1935	"	"	1934	49	—.245	.144
Lawyers [4]						
1935	"	No. in 1935 [5]	1934	47	—.057	.147
Certified public accountants [6]						
1935	States	No. in 1937 [7]	1934	47	.075	.147
1937	"	"	1936	47	—.112	.147

[1] Leland, *Distribution of Physicians*, Table 42.

[2] Total number of physicians in each state as given in the Directory of the American Medical Association, multiplied by 1931 ratio for that state of total number of physicians in active practice to total number of physicians (Leland, Table 14).

[3] Thomas, 'Dental Survey', p. 158.

[4] The ratios of sample to universe are based on the number of individuals in the sample after correction for firm member and size of community biases. Average income is solely for individual practitioners.

[5] Based on a count of *Martindale-Hubbell Law Directory* for 1936 made available by Martindale-Hubbell.

[6] The ratios of sample to universe are based on the number of individuals in the sample after correction for firm member bias.

[7] Estimates made by American Institute of Accountants.

For physicians, the availability of data from the 1937 sample on type of practice permits one additional comparison with the universe. According to this sample, 22 per cent of all physicians considered themselves complete specialists, 38 per cent, partial specialists, and 40 per cent, general practitioners. According to a comprehensive count of the 1931 Directory of the American Medical Association, 16.5 per cent of all physicians in active prac-

tice were complete specialists, 15.6 gave 'special attention' to a subject, and 67.9 per cent were in general practice.[71] The large difference between our sample and the directory count in the percentage classified as partial specialists is of little significance, since the distinction between partial specialists and general practitioners is extremely vague.[72] Of more interest is the considerably higher percentage classified as complete specialists in the sample. One possible explanation is the bias, noted in Section 1b iii, which is due to selecting the sample by laying a straight edge marked off at equally spaced intervals along a column of names. However, this explanation is contradicted by the comparisons made below between our samples and other studies that presumably do not have this bias. The percentage of complete specialists in our sample is fairly close to the percentage in the other studies and for three of four comparisons is below rather than above the latter. The observed difference may therefore more reasonably be attributed to (1) a general tendency for specialists to be more willing to reply, (2) a difference in the wording and interpretation of the question on type of practice, (3) the inclusion among those listed in the directory as 'in active practice' of some persons who have retired or are not in practice and who tend to be classified as general practitioners, (4) the inclusion in the directory count of unclassified nonmembers as general practitioners,[73] or (5) an increase in the percentage of specialists from 1931 to 1936. Though there is no evidence on the relative importance of these factors, the slight tendency for the percentage of specialists to be smaller in our sample than in other samples suggests that point (1), if it applies at all, probably applies to our samples less than to the others. This would be consistent with the fact that the other studies were all conducted by medical associations and hence probably evoked a better response from members, who include a larger percentage of specialists than of general practitioners.

b Comparison of different samples for the same profession

Two or three independent samples are available for each profession other than engineering. These samples cover spans of years that overlap. Comparisons between the distributions of the

[71] See Leland, *Distribution of Physicians*, p. 32.

[72] See footnote 12 of Ch. 6.

[73] See Leland, *Distribution of Physicians*, p. 17.

samples for the overlapping years reveal larger discrepancies than can reasonably be attributed to chance alone (Table A 9).[74] At the same time, the discrepancies are considerably smaller than those between the sample and universe distributions (compare Tables A 7 and A 9). In part, the smaller discrepancies merely confirm our earlier finding that a large part of the observed discrepancies between the sample and universe distributions is attributable to errors in the estimated distributions of the universe or

TABLE A 9

Comparisons between Distributions of Successive Samples for Overlapping Years

Physicians, Dentists, and Certified Public Accountants

PROFESSION & SAMPLES COMPARED	GEOGRAPHIC UNITS	YEAR OF COMPARISON	NO. OF DEGREES OF FREEDOM	χ^2	P [1]
(1)	(2)	(3)	(4)	(5)	(6)
Physicians					
1933 & 1935	States	1932	36	54.59	.022
	Size of community	1932	6	14.12	.029
Dentists					
1933 & 1935	States	1932	31	66.67	.00009
Certified public accountants [2]					
1933 & 1935	States	1932	34	48.52	.047
1937 & 1935	"	1934	29	39.03	.10

[1] P is the probability that the observed χ^2 would be exceeded by chance.
[2] Comparisons are between distributions not corrected for firm member bias.

to general forces affecting all samples alike. In part, also, they are explained by the fact that similar methods were used to choose the different samples for the same profession. Deficiencies in the sampling methods would affect all samples alike and hence would not produce discrepancies among them.

The observed discrepancies must reflect either changes in the willingness of individuals to reply that were not uniform for all states (or other units), or differences in the universes from which the samples were drawn. Differences in the universe can hardly be important for physicians or accountants since the samples for

[74] No comparison was made for law because the 1937 legal sample is not random among states. For the same reason, the 1937 medical sample was not compared with the other medical samples.

these professions were chosen from comprehensive lists at intervals of only two years, and the distribution of all persons in the professions is very stable. The dental samples, on the other hand, were chosen from the membership lists of the American Dental Association, which experienced a decline in membership of about 8 per cent from 1932 to 1933 and then an increase of about 15 per cent from 1933 to 1934.[75] The decline and subsequent growth were doubtless not the same in all states. This greater change in the universe may be the reason why the dental samples differ more from one another than the samples for the other professions.

The relatively small discrepancies between the geographic distributions of the samples appear uncorrelated with differences between average incomes. For each pair of samples in Table A 9, we have computed rank difference correlation coefficients by states between (a) the ratio of the number in one sample to the number in the other and (b) the difference between the average incomes from the two samples. The correlation coefficients are uniformly small (Table A 10). Three are smaller than their standard errors, and the fourth is only slightly larger than its standard error.

The differences between the nationwide average incomes for years covered by more than one sample for the same profession confirm the bias discussed in Section 1a i. It was suggested there that requesting information for a period of years from a sample chosen from a list presumed to be comprehensive for the end of the period would tend to impart an upward bias to the trend of income over time. In consequence, the earlier of two samples would tend to yield a higher average for an overlapping year. Seven of the ten differences summarized in Table 5 (Ch. 2) are in the expected direction, and two of the three exceptions are for lawyers, for whom the 1937 sample is suspect on other grounds (see Sec. 1b iv above). Except for one of the accountancy comparisons, the positive differences are moderate, the largest being 13 per cent. Even the 35 per cent difference between the 1929 averages from the 1933 and 1937 accountancy samples is not disturbing since the bias in question should be larger the longer the time elapsing between the selection of the samples, and between these dates and the date for which the comparison is made. It is fair to conclude, there-

[75] Based on data from the American Dental Association. These changes are *net*. Others may well have taken place and may have affected the distribution of members by states.

fore, that these comparisons give no evidence of the existence of any biases other than those already noted.

A more detailed comparison between the average incomes from successive samples is summarized in Chart A 1, which contains three panels. The panels for physicians and dentists compare the 1932 average incomes from the 1933 and 1935 samples; the panel for accountants compares the 1934 average incomes from the

TABLE A 10

Rank Difference Correlation Coefficients between the Ratio of the Number of Persons in One Sample to the Number in Another Sample for the Same Profession and the Difference in Average Income, by States

Physicians, Dentists, and Certified Public Accountants

PROFESSION & SAMPLES COMPARED	YEAR OF COMPARISON	NO. OF STATES	RANK DIFFERENCE CORRELATION COEFFICIENT	STANDARD ERROR
(1)	(2)	(3)	(4)	(5)
Physicians				
1935 & 1933	1932	48	.1383	.1459
Dentists				
1935 & 1933	1932	48	—.23	.1459
Certified public accountants *				
1935 & 1933	1932	46	.0944	.149
1937 & 1935	1934	46	.1148	.149

* The first correlation is based on the original number of persons, uncorrected for the firm member bias, and on average incomes, also uncorrected. The second correlation is based on the number of persons after correction for the firm member bias, and on the average incomes of individuals practising alone (i.e., on averages that exclude firm members).

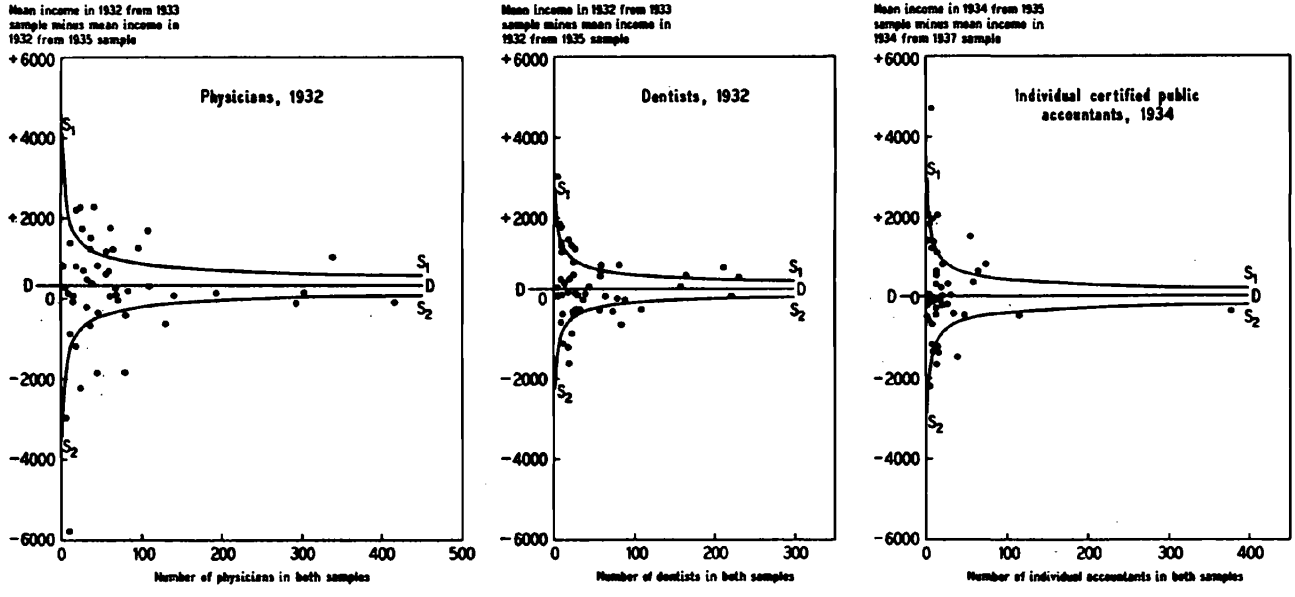
1935 and 1937 samples.[76] Each point in the chart is for an individual state. The vertical distance of the point from the zero axis measures the difference between the two averages for a state (the average from the earlier sample minus the average from the later sample); the horizontal distance, the total number of persons from that state who reported their incomes in the two samples.

[76] The panel for accountants refers solely to individuals practising alone, i.e., excludes firm members. The latter were excluded because of the firm member bias.

CHART A 1

Difference between Arithmetic Mean Incomes from Successive Samples

Each point is for an individual state
DD = difference between countrywide means
S_1S_1 = DD plus one standard deviation of the difference between the means
S_2S_2 = DD minus one standard deviation of the difference between the means



The general bias arising from the time elapsing between the selection of the samples is reflected in the tendency for the points to be above the zero axis. The chart reveals nothing further about this bias. Its purpose is rather to test whether, aside from this general bias, the differences between the samples can be attributed to random sampling fluctuations. If the differences were random, they would tend to vary most from the average nationwide difference for states contributing the fewest returns to the samples, and to vary least for states contributing the largest number of returns. The rapidity with which the state differences might be expected to converge toward the nationwide difference is indicated by the heavy curved lines in each panel, which represent approximately the nationwide difference plus and minus one standard deviation.[77] On the other hand, if the differences were attributable not to random factors but to differential biases varying from state to state, there would be no tendency for the state differences to approach the nationwide difference as the number of returns increased.

The chart is entirely consistent with the hypothesis that the differences are random. For each profession, the points converge toward the nationwide difference as the number of returns increases. Moreover, the points converge in the manner indicated

[77] The plotted standard deviation lines are rough approximations computed from the formula $\sigma^2_{\bar{x}_1-\bar{x}_2} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} - 2r_{\bar{x}_1\bar{x}_2}\sigma_{\bar{x}_1}\sigma_{\bar{x}_2}$, where $\bar{x}_1$ and $\bar{x}_2$ are the arithmetic mean incomes for a state from the earlier and later samples respectively, and $r_{\bar{x}_1\bar{x}_2}$ is the product moment correlation coefficient between the arithmetic means. $\sigma^2_{\bar{x}_1}$ was assumed equal to $\sigma^2_{\bar{x}_2}$ and each equal to $\frac{2\sigma^2}{n}$, where σ^2 is the variance of an individual observation, i.e., the variance of the sample as a whole, and n, the total number of returns from a state in both samples. These assumptions imply that the two samples contain the same number of returns from each state and have equal variances. Neither implicit assumption is correct but both are fair approximations. The correlation coefficients were computed by converting the rank difference correlation coefficient between the means (r') into the product moment correlation coefficient (r) by the formula $r = 2 \sin \frac{\pi r'}{6}$. The numerical values assigned the standard deviations for the sample as a whole and the correlation coefficients are:

	σ	r
Physicians	4,100	.6
Dentists	2,500	.5
Accountants	3,300	.6

by the standard deviation lines, and their distribution about these lines does not deviate significantly from the distribution that random factors would produce.[78] The chart therefore gives no evidence that, aside from the general bias noted, there are any other biases that affect the successive samples for the same profession differently.

c Comparison of our samples with other studies

Data on professional income for the period covered by our samples are rare. There are a fair number of studies for physicians and dentists, two fragmentary studies for lawyers, one for consulting engineers, and none for certified public accountants. Three of these studies are sufficiently extensive to warrant and permit detailed comparison with our data. These are the studies of the 1929 incomes of physicians and of dentists made under the auspices of the Committee on the Costs of Medical Care; and the California Medical-Economic Survey, which covered both physicians and dentists for the years 1929–34. Only very general comparison with the other studies is feasible.

i *Incomes of physicians in 1929.* Data on 1929 income was obtained by the Committee on the Costs of Medical Care from over 5,000 physicians in independent practice who replied to a questionnaire mailed to 20,000 physicians throughout the nation. The results of this study were presented and analyzed by Leven.[79] His book includes, in addition, a brief analysis of more than 6,000 questionnaires obtained by the American Medical Association in

[78] The observed distributions of the differences compared with the expected distributions, computed on the assumption that the differences are normally distributed, are:

	NUMBER OF DIFFERENCES FOR					
	PHYSICIANS		DENTISTS		ACCOUNTANTS	
DIFFERENCE	Observed	Expected	Observed	Expected	Observed	Expected
Larger than σ	12	7.6	11	7.8	7	7.3
σ to mean	9	16.4	13	16.7	15	15.7
Mean to $-\sigma$	16	16.4	17	16.7	16	15.7
Less than $-\sigma$	11	7.6	8	7.8	8	7.3
Total	48	48	49	49	46	46

The discrepancy between observed and expected distributions is greatest for physicians. Yet even this discrepancy would be exceeded by chance alone more than one time in 20.

[79] *Incomes of Physicians.*

a study, restricted mainly to its members, covering 1928. The questionnaires used by the Committee on the Costs of Medical Care asked only for gross income; net income was estimated on the basis of the relation between net and gross income revealed by the American Medical Association sample.

Arithmetic mean and median gross and net incomes from the other samples are reasonably consistent with those from ours (Table A 11). The averages from the American Medical Association sample are uniformly highest, as might be expected from its overrepresentation of members. The averages from our 1933 sample (referred to in the table as the Department of Commerce sample) are somewhat higher than those from the Committee's sample, although, on *a priori* grounds, the differences might be expected to be in the opposite direction. Our sample was taken at a later date and hence should have a downward bias relatively to the Committee's. However, the differences are small—the largest is less than four per cent—and, for the arithmetic means, are only slightly larger than their standard errors.[80] Moreover, the averages presented by Leven are weighted averages of the averages for size of community classes, the weights being the estimated total number of physicians in each size of community class. This adjustment was made because the Committee's sample seemed to underrepresent both very small and very large communities. The average gross income yielded directly by the sample can be computed from the data Leven presents and is given in Table A 11. It is almost identical with the average from our 1933 sample.[81]

The distribution of the Committee's sample by size of gross income differs somewhat from the corresponding distribution for 1929 of our 1933 sample. Although the standard deviations of the two distributions are very close, $11,790 for the Committee's sample and $11,835 for our sample,[82] the observed difference between the distributions as a whole would be exceeded by chance

[80] The standard error of the difference between the arithmetic means is approximately $300 for gross income and $173 for net income.

[81] It was noted in Sec. 1b ii above that our samples underrepresent the smaller communities. The resulting upward bias in average income seemed too slight to justify correction. This upward bias may, however, have been larger in 1929 than in 1932, the year for which it was tested, and may partly explain the failure of our averages to be below those from the Committe's sample.

[82] The difference between the two standard deviations is not statistically significant.

TABLE A 11

Arithmetic Mean and Median Gross and Net Incomes

Physicians: Department of Commerce, Committee on Costs of Medical Care, and American Medical Association Samples

SAMPLE	YEAR	NO. OF RETURNS	ARITHMETIC MEAN			MEDIAN		
			ACTUAL	DIFFERENCE FROM 1933 D. OF C. SAMPLE		ACTUAL	DIFFERENCE FROM 1933 D. OF C. SAMPLE	
				Absolute	Percentage		Absolute	Percentage
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			(dollars)			(dollars)		
					Gross Income			
Dept. of Commerce, 1933	1929	2,226	9,778			7,275		
American Medical Association	1928	5,475	10,440	+662	+6.8	8,090	+815	+11.2
Comm. on Costs of Med. Care								
Original	1929	5,380[1]	9,764[2]	−14	−.1			
Weighted by size of community	1929	5,380[1]	9,461	−317	−3.2	7,026	−249	−3.4
					Net Income			
Dept. of Commerce, 1933	1929	2,139	5,916			4,223		
Dept. of Commerce, 1937	1929	912[3]	5,493[4]	−423	−7.2			
American Medical Association	1928	5,475	6,499	+583	+9.9	4,938	+715	+16.9
Comm. on Costs of Med. Care, weighted by size of community	1929	5,380	5,700	−216	−3.7	4,100	−123	−2.9

Maurice Leven, *Incomes of Physicians* (University of Chicago Press, 1932), pp. 20, 24, 105, 109, 111.

[1] Leven cites a figure of 6,989. But this is adjusted to allow for individuals who failed to reply and to whom no follow-up letter was sent. The figure given above, 5,380, is an estimate of the actual number of returns on which the averages are based. It is derived from *Incomes of Physicians*, pp. 13, 15, and 109.

[2] Computed from *ibid.*, p. 111.

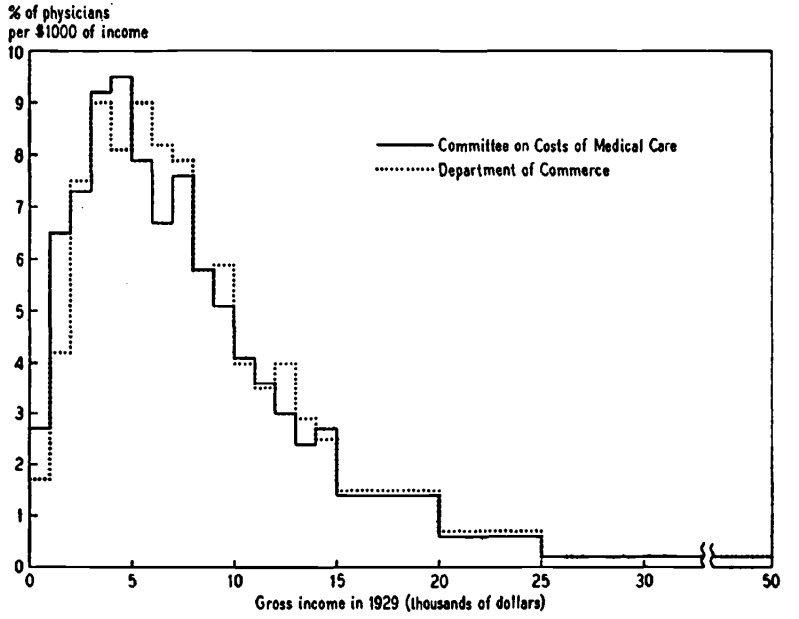
[3] Actual number of returns before weighting.

[4] Weighted average of the state averages.

CHART A 2

Distribution of Gross Income by Size

Physicians, 1929: Committee on Costs of Medical Care and Department of Commerce Samples



alone less than one time in a thousand.[83] The difference between the two distributions is therefore *statistically* significant. At the same time, as Chart A 2 shows, the difference is so small that it is hardly significant in any other sense.[84] The Committee's sample contains relatively more very low incomes and relatively fewer intermediate incomes. These differences may reflect the underrepresentation of small communities in our sample. The distributions of net income differ somewhat more than the distributions of gross income, but since the Committee's distribution of net income was derived from its distribution of gross income by using ratios of net income to gross income, this comparison is of slight importance.

[83] χ^2 is 46.8, with 18 degrees of freedom.

[84] The reason so minor a difference is statistically significant is, of course, that the samples are so large.

Comparison is possible not only between the nationwide data from the two samples but also between the arithmetic mean gross incomes in each size of community class within each region. While our size of community classes are not identical with those used by Leven, the differences are slight.[85] Chart A 3, constructed in the same way as Chart A 1, summarizes the comparison. Each point in the chart is for a size of community and regional cell, or for a size of community class or region as a whole. The vertical distance of the point from the zero axis measures the difference between the average from the Committee's sample and the average from our sample; the horizontal distance, the square root of the number of physicians in our sample. The heavy curved lines include a range of one standard deviation on either side of the zero axis, since zero is the expected difference on the hypothesis that the two samples represent the same universe.[86] If the differences between the samples reflected sampling fluctuation alone, the points should converge toward the zero axis as the number of returns increased. On the other hand, if the differences reflected bias, there

[85] The two classifications differ at the extremes. We use the classes under 2,500, and 2,500–10,000, whereas Leven uses under 5,000, and 5,000–10,000; at the other extreme, Leven uses the classes 500,000–1,000,000, and 1,000,000 and over, whereas we use the one class, 500,000 and over. Since Leven does not give the number of returns in each cell, we were unable to combine his class intervals. Consequently we compared our class 500,000 and over with his class 500,000–1,000,000 and at the lower extreme compared the overlapping class intervals.

[86] The standard deviation is approximate. Since Leven does not give the number of returns in each cell, the standard deviation is computed on the assumption that the ratio of the number in the Committee's sample to the number in our sample is the same for each cell as for the entire country. For simplicity, the standard deviation of an individual observation is assumed equal for the two samples, an assumption that is approximately correct. On these assumptions, the standard deviation of the difference between the averages in any cell is

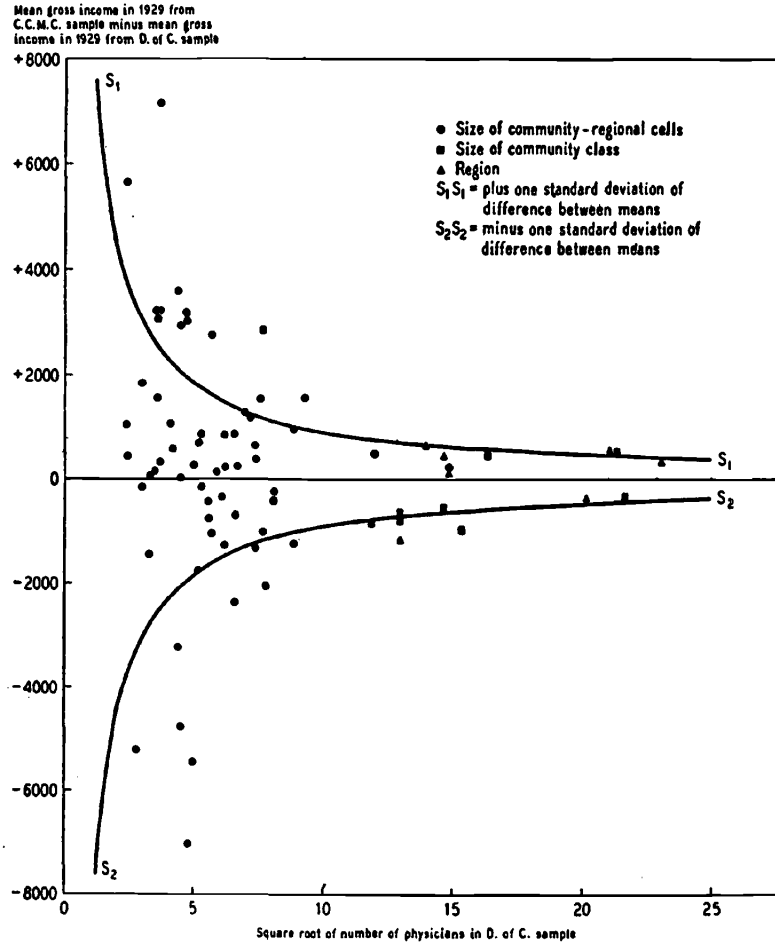
$$\frac{\sigma}{\sqrt{n_1}}\sqrt{\frac{p+1}{p}-\frac{2r}{\sqrt{p}}}$$

where σ is the standard deviation of an individual observation, i.e., the standard deviation of the sample, p is the ratio of the number in the Committee's sample to the number in our sample, and r is the correlation between the cell averages from the two samples. σ was taken as $11,790, p as 2.46, and r as .64. The product moment correlation coefficient was estimated from the rank difference correlation coefficient.

CHART A 3

Difference between Arithmetic Mean Gross Incomes from Committee on Costs of Medical Care and Department of Commerce Samples

Physicians, 1929



would be no tendency for the points to converge. It is clear from the chart that the points do converge in the manner indicated by the standard deviation lines. The distribution of the points about the standard deviation lines does not differ significantly from

the distribution that chance alone would produce.[87] Interestingly enough, the average from the Committee's sample is larger than the average from our sample for more than half the cells, 37 out of 59, although the difference between the nationwide averages is in the opposite direction.

Both the Committee's sample and our 1937 sample furnish data on the percentage of physicians who consider themselves general practitioners, partial specialists, and complete specialists. While

	PERCENTAGE OF ALL PHYSICIANS	
	Comm. on Costs of Med. Care	1937 D. of C. sample
General practitioners	56.7	38.6
Partial specialists	20.7	39.9
General practitioners and partial specialists	77.4	78.5
Complete specialists	22.6	21.5
Total	100.0	100.0

the percentage of physicians who classified themselves as complete specialists is very similar for the two samples, the division of the other physicians between general practitioners and partial specialists is very different. However, as noted above, the vagueness of the distinction between the general practitioner and the partial specialist makes this difference of little significance.

These comparisons between our medical samples and that of the Committee on the Costs of Medical Care are comforting. The samples are remarkably similar in practically all respects despite marked differences between the agencies conducting the surveys, the dates at which they were made, the sampling methods, and the questionnaires used.

ii *Incomes of dentists in 1929.* The study of dental incomes in 1929 made by the American Dental Association in cooperation

[87] The expected and observed distributions of the points for the individual cells are:

	OBSERVED	EXPECTED
Larger than σ	12	9.4
σ to 0	25	20.1
0 to $-\sigma$	13	20.1
Less than $-\sigma$	9	9.4
Total	59	59.0

χ^2 between the observed and expected is 4.46 and would be exceeded by chance in more than one-fifth of random samples. The regional and size of community averages are, of course, not independent of the averages for the separate cells and hence have been excluded from the above comparison.

with the Committee on the Costs of Medical Care is less directly comparable with our samples than the Committee's study of the incomes of physicians. A minor difficulty is that the Committee's dental study was restricted to 20 states. About 9,000 dentists were circularized in these states and over 5,000 returns containing usable information on income obtained. A more important difficulty is that our sample is restricted to members of the American Dental Association, whereas the Committee's sample is not.

To facilitate comparability, arithmetic mean net and gross incomes in 1929 have been computed from our 1933 sample for the twenty states covered by the Committee's sample and have been corrected for the restriction of our sample to members (Table A 12). The corrected averages were computed from the original averages by multiplying the latter by the correction factor derived in Section 1a iv for the country as a whole. The same correction factor was used for both net and gross income, though it was derived from net income data alone. Consequently, the estimate of the average gross income of all dentists is even rougher than the estimate of the average net income.

Table A 12 also includes two sets of averages from the Committee's sample: the original averages derived directly from the sample; and the final estimates for the 20 states presented by Leven, who was responsible for the analysis of the dental sample as well as the medical sample. These final estimates embody two adjustments. First, the net incomes of dentists who reported gross income but not net income were estimated on the basis of the relationship between net income and gross income revealed by the remaining returns, and these estimated net incomes were added to the original sample.[88] Second, a correction was made for the inclusion in the sample of a larger percentage of members than of nonmembers.

The final estimates from the Committee's sample are somewhat higher than those from our sample, as might be expected from the general tendency for the earlier of two samples to yield a higher average for the overlapping year. The average net income is $373, or 8.4 per cent, higher; the average gross income, $508, or 6.7 per

[88] No such adjustment was made for the Department of Commerce sample. However, the percentage of respondents who reported gross income but not net income was five times as large in the Committee's sample as in the Department of Commerce sample.

TABLE A 12

Arithmetic Mean Gross and Net Incomes

Dentists in 20 States: Department of Commerce and Committee on Costs of Medical Care Samples, 1929

	GROSS INCOME		NET INCOME	
	No. of returns	Arith. mean income (dollars)	No. of returns	Arith. mean income (dollars)
Dept. of Commerce 1933 sample, for 20 states covered by Comm. on Costs of Med. Care sample				
Original	879	8,702	855	5,076
Adjusted for restriction to Amer. Dental Assn. members [1]		7,623		4,447
Comm. on Costs of Med. Care sample for 20 states				
Original [2]	5,493	8,279	4,705	5,011
Final estimate (corrected for returns giving gross income but not net income and for over-representation of members) [3]		8,131		4,820

[1] Averages in first line multiplied by .876.
[2] Maurice Leven, *Practice of Dentistry and the Incomes of Dentists in Twenty States: 1929* (University of Chicago Press, 1932), pp. 75–7.
[3] *Ibid.*, p. 207.

cent, higher. Part of this difference is doubtless due to the underrepresentation of new entrants in the Committee's sample noted by Leven.[89] His final estimates do not allow for this underrepresentation, although he indicates that correction for it would lower the average net income by about $65.[90] The rest of the difference in net income, $308, is slightly less than twice its standard error, while, even without correction for this factor, the difference in gross income is also slightly less than twice its standard error.[91] On the whole, then, the agreement is fairly good.

[89] *Practice of Dentistry*, p. 199.
[90] *Ibid.* Leven states that this would be the effect on the income of *general practitioners*. However, since 89 per cent of the dentists were general practitioners the same figure may without serious error be used for all dentists.
[91] The standard error of the difference between the means is $155 for net income, and $271 for gross income.

The difficulty of correcting characteristics other than the arithmetic mean for the restriction of the Department of Commerce sample to members makes precise comparison of the samples with respect to these characteristics impossible. The uncorrected frequency distributions of net and gross income by size naturally differ more than would be expected from chance alone; so also do the distributions of the samples by states.[92] As Chart A 4 indicates, the differences between the frequency distributions of net and gross income are consistent with the tendency for nonmembers to have lower incomes than members. The distribution from the Committee's sample includes relatively more low incomes and relatively fewer intermediate incomes. This divergence appears more marked in the gross income distribution than in the net income distribution because of the wider dispersion of the former; it is definitely present in both. Both the gross and net income distributions from the Committee's samples have smaller standard deviations than the corresponding distributions from our sample.[93] Though not particularly large in magnitude, this difference is in the opposite direction from what might be expected. As noted in

92 The values of χ^2 and the number of degrees of freedom for these comparisons are as follows:

	χ^2	NO. OF DEGREES OF FREEDOM
Distributions of income by size		
Gross income	50.1	19
Net income (C.C.M.C. original distribution)	30.9	18
Net income (C.C.M.C. distribution adjusted to include dentists reporting gross but not net income)	35.8	18
Distributions of returns by states	74.9	19

For 19 degrees of freedom χ^2 will exceed 30.14 once in twenty times from chance; 36.19, once in a hundred times; and 43.82, once in a thousand times. For 18 degrees of freedom the corresponding values are 28.87, 34.80, and 42.31. The C.C.M.C. distributions were taken from *Practice of Dentistry*, pp. 84, 88, 189. The Department of Commerce distributions compared with them include only dentists in the 20 states covered by the C.C.M.C. study.

93 The values of the standard deviations are:

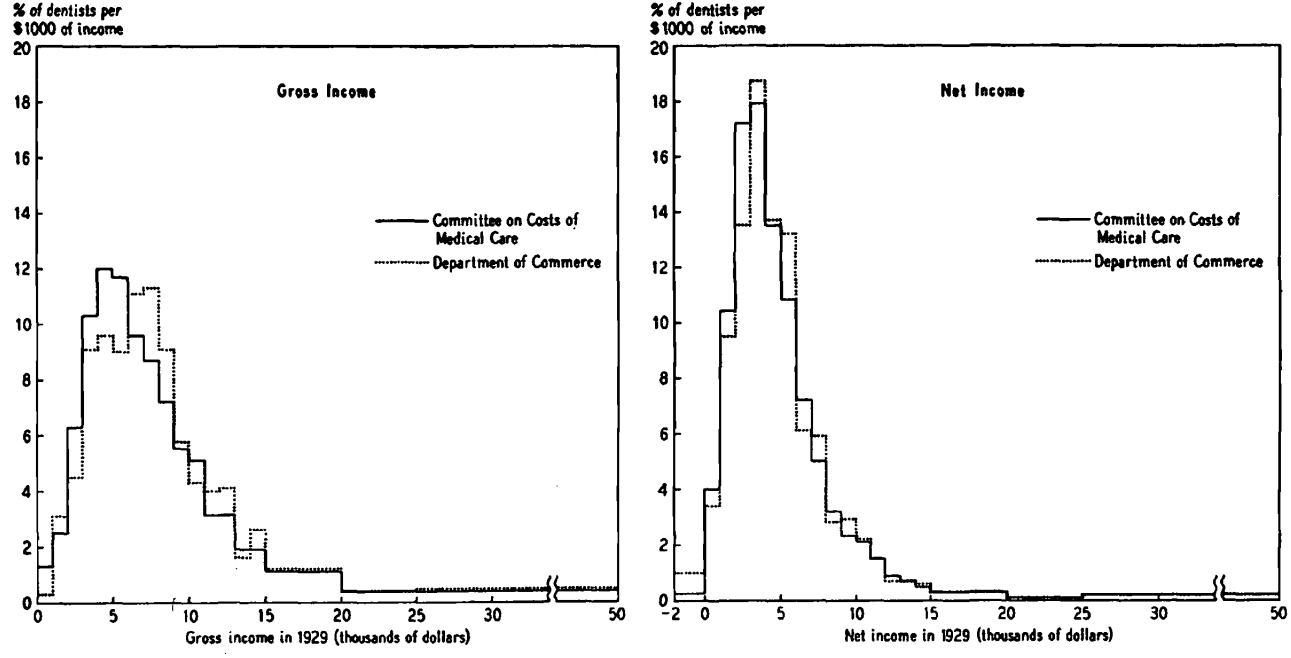
	COMM. ON THE COSTS OF MED. CARE	DEPT. OF COMMERCE
	(dollars)	
Gross income	6,817	7,456
Net income	4,107	4,217

The difference between the standard deviations is statistically significant for gross income but not for net.

CHART A 4

Distributions of Gross and Net Income by Size

Dentists in 20 States, 1929: Committee on Costs of Medical Care and Department of Commerce Samples



Section 1a iv, a sample restricted to members of the Dental Association might be expected to be more concentrated than a sample including all dentists.

iii *Incomes of physicians and dentists in California.* The study sponsored by the California Medical Association is probably the most intensive study of its kind ever made. In January 1935, questionnaires were mailed to every physician and dentist in California.[94] The initial mailing list included over 9,000 physicians and almost 6,000 dentists. Completed questionnaires were finally received from over 3,000 physicians and almost 2,000 dentists,[95] or from about a third of all physicians and dentists in the state.

The questionnaires requested information on income for each year from 1929 to 1933, and for the first nine months of 1934. The incomes reported for 1934 were raised to full-year equivalents. Physicians and dentists on salary as well as those in independent practice were included in the sample. Most of the data presented seem to be for both groups together, although this is not entirely clear. In this respect, as well as in the inclusion of income from salaried positions along with income from independent practice, the California sample is not comparable with the Department of Commerce sample. It is difficult to evaluate the importance of this noncomparability. The inclusion of full-time salaried employees would tend to lower arithmetic mean incomes; but the inclusion of supplementary salaries of persons in independent practice would tend to raise them. An additional source of noncomparability is that the Survey's dental sample covers both members of the American Dental Association and nonmembers, whereas our

[94] Osteopaths, hospitals, and public health agencies were also included in the study. In addition a large sample of families was interviewed and information obtained on their incomes, expenditures for medical and dental care and other purposes, and their need for medical and dental care.

Two reports dealing with the study have been issued: *California Medical-Economic Survey,* and Dodd and Penrose, *Economic Aspects of Medical Service.* The first contains the statistical tables and graphs included in the second, but none of the text of the report by Dodd and Penrose.

"Every person whose name appeared in the March, 1934 Roster of Physicians and Surgeons, and in the September, 1933 Dentists' Directory, . . . was solicited by means of questionnaires", Dodd and Penrose, p. 8.

[95] In addition to the original mailing, individuals failing to reply were circularized a second time; and those failing to reply to the follow-up were circularized once again.

sample covers members only. This is less serious, however, than may appear, since members were overrepresented in the Survey's dental sample.[96]

Our 1933 sample of California physicians yields average gross incomes between $847 and $1,321 *higher* than the corresponding averages from the Survey (Table A 13). Our later medical samples, on the other hand, yield average gross incomes between $203 and $1,278 *lower* than the corresponding averages from the Survey. While these differences seem fairly large, the number of physicians in our samples for California is so small that all of the differences might easily have arisen from chance.[97] Moreover, that the earlier sample is higher, and the later samples lower accords with what might be expected from the oft-repeated characteristics of samples taken at different dates.

The differences between the average net incomes from the two studies are naturally similar to the differences between the average gross incomes. The averages from our 1933 sample are higher, and from our 1935 and 1937 samples lower, than the corresponding averages from the Survey. The differences between the 1933 sample and the Survey are again sufficiently small to be attributable to sampling fluctuations.[98] The differences between the later samples and the Survey, however, cannot be interpreted in this way. The smallest of the three differences for the 1935 sample is 2.2 times its standard error; the largest, 3.4 times its standard error.[99] Three

[96] Dodd and Penrose infer that members are seriously overrepresented from the fact that 66.5 per cent of the respondents reported themselves as members, whereas, according to the Association's records, only 41.3 per cent belonged to the American Dental Association. This discrepancy may, however, merely reflect the tendency noted by Leven for individuals not listed on the Association's membership rolls to report themselves as members. (See Sec. 1a iv above; Leven, *Practice of Dentistry*, pp. 200–1; Dodd and Penrose, p. 170.) It seems likely that both factors were at work, and hence that the overrepresentation of members, while present, is less serious than Dodd and Penrose believe.

[97] The standard error of the difference between the means is larger than $540 for the 1933 sample, $720 for the 1935 sample, and $970 for the 1937 sample.

[98] The relevant standard error is larger than $210.

[99] The relevant standard error is about $270. The smaller standard error for net income than for gross income explains why the differences between the average net incomes are, and the differences between the average gross incomes are not, statistically significant even though the former are numerically smaller than the latter.

TABLE A 13

Arithmetic Mean Gross and Net Incomes, and Number of Persons Covered

California Physicians: Department of Commerce Samples and California Medical-Economic Survey, 1929–1934

	1929	*1930*	*1931*	*1932*	*1933*	*1934*
				Gross Income		
Arithmetic mean (dollars)						
Cal. Med.-Ec. Survey	11,049	10,092	8,868	7,195	6,456	6,911
Dept. of Commerce						
1933 sample	11,896	11,021	10,189	8,177		
1935 sample				6,067	5,537	5,940
1937 sample	9,771	9,610	8,649	6,686	6,253	6,607
Difference between means (dollars)						
Cal. Med.-Ec. Survey minus						
Dept. of Commerce						
1933 sample	−847	−929	−1,321	−982		
1935 sample				+1,128	+919	+971
1937 sample	+1,278	+482	+219	+509	+203	+304
Number of persons covered						
Cal. Med.-Ec. Survey	1,154	1,212	1,285	1,361	1,470	1,581
Dept. of Commerce						
1933 sample	202	214	219	223		
1935 sample				111	115	119
1937 sample	52	52	54	62	68	76
				Net Income		
Arithmetic mean (dollars)						
Cal. Med.-Ec. Survey	6,657	5,984	5,069	4,146	3,572	4,068
Dept. of Commerce						
1933 sample	6,680	6,166	5,500	4,308		
1935 sample				3,157	2,969	3,254
1937 sample	5,336	5,221	4,424	3,104	3,061	3,489
Difference between means (dollars)						
Cal. Med.-Ec. Survey minus						
Dept. of Commerce						
1933 sample	−23	−182	−431	−162		
1935 sample				+989	+603	+814
1937 sample	+1,321	+763	+645	+1,042	+511	+579
Number of persons covered						
Cal. Med.-Ec. Survey	1,445	1,524	1,645	1,756	2,737	1,871
Dept. of Commerce						
1933 sample	200	210	215	216		
1935 sample				110	112	115
1937 sample	49	47	50	56	60	76

California Medical-Economic Survey, Formal Report on Factual Data (California Medical Association, 1937), pp. 94–5.

of the six differences for the 1937 sample are less than twice their standard errors, but the other three are all more than twice their standard errors.[100] The later date at which the 1937 sample was selected makes the differences between it and the Survey understandable. But this explanation will not serve for the 1935 sample.

Comparison of the distributions of net income by size supports the view that the 1933 sample does not differ significantly from the Survey but that the 1935 sample does. While the discrepancy between the 1929 distributions of income from the two studies would be exceeded by chance more than half the time, the discrepancy between the 1933 distributions from our 1935 sample and the Survey would be exceeded by chance less than one time in a hundred.[101] We have been unable to find any explanation of the difference between our 1935 sample and the Survey.[102]

The percentage of complete specialists in our 1937 sample, 25 per cent, is considerably lower than the percentage in the Survey, 34 per cent. So far as this discrepancy represents more than a difference in the interpretation of the term 'complete specialists', it reflects upon the adequacy of the Survey data rather than ours, since only 20 per cent of California physicians were listed as complete specialists in the 1931 Directory of the American Medical Association.[103]

The average gross and net incomes of California dentists (Table A 14) give no reason to suspect the reliability of either of our dental samples. The original averages from our samples are consistently above the corresponding averages from the Survey, as is to be expected from the restriction of our samples to American Dental Association members. With two exceptions, the averages

100 The relevant standard error is larger than $360.

101 χ^2 for the 1929 distributions is 10.6 with 12 degrees of freedom; a χ^2 of 11.3 would be exceeded by chance half the time. χ^2 for the 1933 distributions is 15.4 with 5 degrees of freedom; a χ^2 of 15.1 would be exceeded by chance only once in a hundred times.

102 One hypothesis tested was that the difference was attributable to a discrepancy between the distributions of the samples by size of community. However, there appears to be no such discrepancy. χ^2 between the size of community distributions is 5.5 with 5 degrees of freedom; the probability of this value being exceeded by chance is 0.4.

103 See Leland, *Incomes of Physicians*, p. 17.

TABLE A 14

Arithmetic Mean Gross and Net Incomes, and Number of Persons Covered

California Dentists: Department of Commerce Samples and California Medical-Economic Survey, 1929–1934

	1929	1930	1931	1932	1933	1934
				Gross Income		
Arithmetic mean (dollars)						
Cal. Med.-Ec. Survey	8,680	8,308	7,163	5,633	4,936	5,444
Dept. of Commerce, original						
1933 sample	9,591	9,439	8,416	6,436		
1935 sample				6,027	5,406	5,719
Dept. of Commerce, adjusted for restriction to members of Amer. Dental Assn.*						
1933 sample	8,296	8,165	7,280	5,567		
1935 sample				5,215	4,676	4,947
Difference between means (dollars)						
Cal. Med.-Ec. Survey minus Dept. of Commerce, original						
1933 sample	−911	−1,131	−1,253	−803		
1935 sample				−394	−470	−275
Dept. of Commerce, adjusted						
1933 sample	+384	+143	−117	+66		
1935 sample				+420	+260	+497
Number of persons covered						
Cal. Med.-Ec. Survey	720	781	839	894	965	866
Dept. of Commerce						
1933 sample	110	122	125	129		
1935 sample				93	99	99
				Net Income		
Arithmetic mean (dollars)						
Cal. Med.-Ec. Survey	5,095	4,772	3,968	3,012	2,769	2,956
Dept. of Commerce, original						
1933 sample	5,408	5,478	4,721	3,299		
1935 sample				3,247	2,792	3,023
Dept. of Commerce, adjusted for restriction to members of Amer. Dental Assn.*						
1933 sample	4,678	4,738	4,084	2,854		
1935 sample				2,809	2,415	2,615
Difference between means (dollars)						
Cal. Med.-Ec. Survey minus Dept. of Commerce, original						
1933 sample	−313	−706	−753	−287		
1935 sample				−235	−23	−67
Dept. of Commerce, adjusted						
1933 sample	+417	+34	−116	+158		
1935 sample				+203	+354	+341

* Original arithmetic means multiplied by .865. This yields arithmetic mean of a hypothetical sample containing the same percentage of American Dental Association members as listed on the 1933 membership rolls (41.3 per cent) on the assumption that the average income of members exceeds that of nonmembers by 30 per cent.

TABLE A 14 (cont.)

	1929	1930	1931	1932	1933	1934
			Net income (cont.)			
Number of persons covered						
Cal. Med.-Ec. Survey	852	920	979	1,048	1,615	976
Dept. of Commerce						
1933 sample	108	118	121	128		
1935 sample				93	100	100

California Medical-Economic Survey, pp. 99–100.

corrected for this bias are below those from the Survey. The differences are all sufficiently small to be attributable to sampling fluctuations.[104] Moreover, the Survey overrepresents members and hence has a slight upward bias. Since the correction applied to our samples purports to make full allowance for their restriction to members, the upward bias in the Survey data would explain the lower corrected averages from our samples.[105]

Comparisons of the distributions of the samples by size of community and by size of net income on the whole confirm the absence of any significant differences between our samples and the Survey.[106]

In sum, the only significant discrepancy revealed by these comparisons is between our 1935 medical sample and the Survey. This sample appears, at least for California, to have a downward bias of some importance.

iv *Other studies of the incomes of physicians.* Three additional studies of medical incomes with which our samples can be compared are for Wisconsin, Michigan, and Utah.

The Wisconsin study, made by a committee of the State Medical Society, covered 1930 incomes as reported on state income tax re-

104 Minimum estimates of the relevant standard errors of the difference between the arithmetic mean gross incomes are $400 for the 1933 sample and $430 for the 1935 sample; of the difference between the arithmetic mean net incomes, $220 for the 1933 sample and $230 for the 1935 sample.

105 See footnote 96 above. The correction we make accepts the estimate of 41.3 per cent cited in the Survey as the percentage of dentists who were members of the dental association in 1933 (see *California Medical-Economic Survey*, p. 73). The basis of our estimate of 30 per cent as the excess of the income of members over that of nonmembers is given in Sec. 1a iv above.

106 A comparison of the distribution of the samples in 1933 by size of community yields a χ^2 of 2.9 with 5 degrees of freedom. The probability of this value

turns filed by physicians.[107] Ordinarily, studies based on income tax returns are subject to a serious upward bias because they omit persons whose incomes were below the exemption limit. This difficulty is, however, less serious for Wisconsin than for most other states because a much larger percentage of persons are required to file returns.[108] Income tax returns were found for 2,129 of an estimated 2,836 physicians in active practice. The committee in charge of the study estimated that considerably fewer than 500 of the 707 physicians omitted failed to file because of low incomes. The downward bias arising from the tendency toward understatement of income on tax returns probably offsets, in this study at least, most if not all of the upward bias arising from the omission of persons who did not file tax returns.

Our samples agree remarkably well with the Wisconsin study (Table A 15). The largest difference between the arithmetic means for all physicians is between mean net incomes from the Wisconsin

---

being exceeded by chance is 0.7. The results of the comparisons of the net income distributions are as follows:

	NO. OF DEGREES OF FREEDOM	χ^2	PROBABILITY THAT χ^2 WOULD BE EXCEEDED BY CHANCE
Cal. Med.-Ec. Survey compared with 1933 sample			
1929	8	9.2	.3
1930	9	10.7	.3
1931	8	6.6	.6
1932	7	4.5	.7
Cal. Med.-Ec. Survey compared with 1935 sample			
1932	6	9.4	.15
1933	6	4.6	.6

This last set of results, though showing no significant differences, is somewhat disturbing since the distributions from the Department of Commerce samples were not corrected for the restriction of the samples to members.

107 The results of the study appear in *Wisconsin Medical Journal*, Supplement, Dec. 1932, pp. 939–67.

108 The exemption limits in Wisconsin in 1930 were $800 for single individuals and $1,600 for married persons. In addition, the income tax authorities can request returns from any individual, regardless of his income. Typically, the assessors retain on their mailing lists individuals who have paid a tax in any of the preceding few years.

study and our 1933 sample; and even this difference, $408, might easily have arisen from chance.[109] Despite the slenderness of our sample, the averages for the separate groups of physicians—general practitioners and partial specialists, and complete specialists—are

TABLE A 15

Arithmetic Mean Gross and Net Incomes, and Number of Persons Covered

Wisconsin Physicians: Department of Commerce Samples and Study of Wisconsin State Medical Society, 1930

	GROSS INCOME			NET INCOME		
	All physicians	General practitioners and partial specialists	Complete specialists	All physicians	General practitioners and partial specialists	Complete specialists
Arithmetic mean income (dollars)						
Wisconsin study	7,553	5,323	12,468	4,704	2,945	8,542
Dept. of Commerce samples for Wisconsin						
1933 sample	7,767			4,296		
1937 sample	7,575	5,515	12,932	4,555	2,986	8,318
Number of persons covered						
Wisconsin study	2,129	1,460	669	2,129	1,460	669
Dept. of Commerce samples for Wisconsin						
1933 sample	43			42		
1937 sample	18	13	5	17	12	5

Wisconsin Medical Journal, Supplement, Dec. 1932, pp. 959–67.

very close. Finally, the percentage of physicians in our 1937 sample who designated themselves complete specialists—28 to 30 per cent—is almost exactly the same as the percentage of physicians in the Wisconsin sample who were listed as complete specialists—31 per cent.

The Michigan data are from a questionnaire survey conducted by the State Medical Society.[110] Information on 1931 income was

[109] The standard error of the difference between the arithmetic mean net incomes is approximately $600 for the 1933 sample and $1,000 for the 1937 sample. The corresponding standard errors for gross income are even larger.

[110] See *Report of the Committee on Survey of Medical Services and Health Agencies* (Michigan State Medical Society, 1933), Ch. V, VI.

TABLE A 16

Arithmetic Mean Gross and Net Incomes, and Number of Persons Covered

Michigan Physicians: Department of Commerce Samples and Study of Michigan State Medical Society, 1929 and 1931

	GROSS INCOME		NET INCOME	
	1929	*1931*	*1929*	*1931*
Arithmetic mean income (dollars)				
Michigan study				
All returns	9,976	6,590	6,306	3,876
Returns reporting for both 1929 and 1931	9,976	7,546	6,306	4,497
Dept. of Commerce samples for Michigan				
1933 sample	9,997	7,188	6,801	4,282
1937 sample	10,385	7,000	5,722	4,293
Number of persons covered				
Michigan study				
All returns	592	1,289	592	1,289
Returns reporting for both 1929 and 1931	592	592	592	592
Dept. of Commerce samples for Michigan				
1933 sample	73	78	74	78
1937 sample	13	15	13	13

Report of the Committee on Medical Services and Health Agencies (Michigan State Medical Society, 1933), Ch. V, VI.

reported by 1,289 of the almost 5,000 physicians estimated to be in private practice. Only 592 of these, however, reported income in 1929. The average incomes from this study are reasonably similar to those from our samples for Michigan (Table A 16). The differences vary from $21 to about $600 and all are well within the range of sampling variation.[111]

The Utah data are from a questionnaire survey made by the Utah State Medical Association and cover 1929–31, and 1933.[112] Although returns were received from only 94 physicians, these constitute almost a fifth of all physicians practising in the state. Of the three studies under discussion, this is the only one with which our data fail to check satisfactorily. While the 1937 and

111 The standard error of the difference between the arithmetic mean net incomes is approximately $500 for the 1933 sample and $1,000 for the 1937 sample. The corresponding standard errors for gross income are even larger.

112 See *A Survey of Medical Services and Facilities of the State of Utah—1934, December* (Utah State Medical Association).

1935 samples do not differ from the Utah sample more than might be expected from chance, the 1933 sample does (Table A 17).[113] The arithmetic mean net incomes from our 1933 sample are substantially below those from the Utah study in all three years for which the comparison can be made. The difference is $2,826 for 1929, $1,964 for 1930, and $1,836 for 1931. Each difference is considerably more than twice its standard error. While the arithmetic mean gross incomes differ somewhat less, those from our 1933 sample are consistently lower. It seems clear that the 1933 sample for Utah has a sizable downward bias. The other samples seem satisfactory. A further test of the 1937 sample is that the percentage of physicians reporting themselves as complete specialists, 17 per cent, is almost identical with the corresponding percentage from the Utah study, 16 per cent.

v *Incomes of dentists in Minnesota.* As part of a general study of the economics of dentistry in Minnesota, the University Relations Committee, comprising representatives from the State Dental Association and the University of Minnesota School of Dentistry, obtained completed questionnaires from about 600 of the 2,000 dentists in Minnesota.[114] The questionnaires asked for gross and net income in 1933 and 1934. Our 1935 dental sample checks closely with this more extensive sample (Table A 18). The original arithmetic means from our sample are slightly higher than those from the Minnesota study. However, these are not entirely comparable since our sample is restricted to American Dental Association members. Arithmetic means corrected for this bias are between $100 and $160 lower than the corresponding averages from the Minnesota study. These differences are all well within the range of sampling fluctuation;[115] moreover, since the Minnesota

[113] The approximate standard errors of the differences between arithmetic mean net incomes are:

1933 sample: 1929, $1,000; 1930, $800; 1931, $450;
1935 sample: 1933, $500;
1937 sample: 1929, $630; 1930, $530; 1931, $300; 1933, $300.

The corresponding standard errors for gross income are larger.

[114] See 'Report of the University Relations Committee', *North-West Dentistry*, April 1936, pp. 79–88.

[115] Approximate standard errors of the differences are: net income, $200; gross income, $300.

TABLE A 17

Arithmetic Mean Gross and Net Incomes, and Number of Persons Covered

Utah Physicians: Department of Commerce Samples and Utah State Medical Association Study, 1929–1931 and 1933

	1929	*1930*	*1931*	*1933*
		Gross Income		
Arithmetic mean income (dollars)				
Utah study	9,644	8,683	7,151	5,661
Dept. of Commerce sample for Utah				
1933 sample	8,073	6,984	5,688	
1935 sample				5,428
1937 sample	9,994	9,248	7,940	5,853
Number of persons covered				
Utah study	94	94	94	94
Dept. of Commerce sample for Utah				
1933 sample	10	11	11	
1935 sample				8
1937 sample	31	27	28	31
		Net Income		
Arithmetic mean income (dollars)				
Utah study	6,532	5,202	4,139	3,038
Dept. of Commerce sample for Utah				
1933 sample	3,706	3,238	2,303	
1935 sample				3,032
1937 sample	6,883	5,702	4,685	3,369
Number of persons covered				
Utah study	94	94	94	94
Dept. of Commerce sample for Utah				
1933 sample	11	11	11	
1935 sample				8
1937 sample	29	27	28	30

A Survey of Medical Services and Facilities of the State of Utah—1934, December (Utah State Medical Association).

study was sponsored by the State Dental Association, it is likely that it overrepresents Association members. The quartiles and medians, as well as the arithmetic means, are similar. These measures are not corrected for the restriction of our sample to members. Presumably for this reason, the measures from our sample are higher in eight out of twelve comparisons.

TABLE A 18

Arithmetic Mean, Median, and Quartile Gross and Net Incomes, and Number of Persons Covered

Minnesota Dentists: Department of Commerce 1935 Sample and Study of Minnesota University Relations Committee, 1933–1934

	GROSS INCOME		NET INCOME	
	1933	*1934*	*1933*	*1934*
Arithmetic mean income (dollars)				
Minnesota study	3,739	3,981	2,149	2,378
Dept. of Commerce 1935 sample for Minnesota				
Original	3,886	4,205	2,168	2,394
Adjusted for restriction to members •	3,593	3,889	2,005	2,214
First quartile (dollars)				
Minnesota study	2,350	2,496	1,304	1,507
Dept. of Commerce 1935 sample for Minnesota	2,500	2,375	1,357	1,250
Median (dollars)				
Minnesota study	3,426	3,695	1,992	2,260
Dept. of Commerce 1935 sample for Minnesota	3,666	3,500	2,115	2,400
Third quartile (dollars)				
Minnesota study	4,759	5,073	2,822	3,112
Dept. of Commerce 1935 sample for Minnesota	4,900	5,875	2,500	3,416
Number of persons covered				
Minnesota study	600	600	600	600
Dept. of Commerce 1935 sample for Minnesota	44	49	42	47

'Report of the University Relations Committee', *North-West Dentistry*, April 1936, pp. 79–88.

• Original averages multiplied by .925. This correction factor assumes that 67.4 per cent of all dentists in Minnesota are members and that the income of members is 30 per cent greater than that of nonmembers. The percentage of members is based on Thomas, 'Dental Survey', pp. 155, 158.

vi *Incomes of lawyers.* The only studies of legal incomes with which we have compared our data are for Wisconsin and New York County.

The Wisconsin study was conducted by Dean Lloyd K. Garrison and, like the Wisconsin study of physicians' incomes, is based on income tax returns.[116] Returns were found for 2,161 of an esti-

[116] See 'Survey of the Wisconsin Bar', pp. 150–61.

mated total of 3,027 lawyers. The percentage of lawyers for whom returns were not found is somewhat larger than the corresponding percentage of physicians, and hence the resulting upward bias may be somewhat larger. Although primarily restricted to lawyers in independent practice, some lawyers in salaried posts were included. The study covers the period 1927–32, but since we have no data for lawyers prior to 1932 the comparison is necessarily limited to that year. The arithmetic mean net income from our 1935 sample is very close to that from Garrison's sample—it is less

TABLE A 19

Arithmetic Mean Net Income, and Number of Persons Covered

Wisconsin Lawyers: Department of Commerce Samples and Garrison's Sample, 1932

	NO. OF PERSONS COVERED	ARITH. MEAN NET INCOME, 1932 (dollars)
Garrison's sample	2,161	3,517
Dept. of Commerce sample for Wisconsin		
1935 sample	45[1]	3,327[2]
1937 sample	13[1]	1,865[3]

L. K. Garrison, 'A Survey of the Wisconsin Bar', *Wisconsin Law Review*, Feb. 1935.

[1] Number of individuals before weighting.
[2] Weighted for firm member and size of community bias.
[3] Weighted for size of community bias. No replies were received from Wisconsin firms.

than $200 below the latter (Table A 19). The average from the 1937 sample, on the other hand, is almost $1,700 lower. Although the variability of the incomes of lawyers is large and consequently very sizable differences might arise from chance in a sample containing so few as 13 lawyers,[117] these results seem to support our earlier conclusion that the 1937 legal sample is considerably less reliable than the 1935 sample.

The data for New York County, on the other hand, indicate exactly the opposite, since the 1937 sample agrees with the New York County sample better than the 1935 sample (Table A 20).

[117] The standard error of the difference between the means is about $1,000 for the 1937 sample and $700 for the 1935 sample.

The New York data were compiled from questionnaires sent to all lawyers in New York County (i.e., Manhattan).[118] Data on 1933 income were reported by 3,210 of an estimated total of 15,000 lawyers.[119] Of these, 2,667 specified that they were in 'private practice'. The New York survey was restricted to Manhattan and included income from salaried employment as well as from independent practice, whereas our samples are for all five boroughs of New York City and cover only income from independent prac-

TABLE A 20

Arithmetic Mean and Median Net Incomes, and Number of Persons Covered

New York Lawyers: Department of Commerce Samples and New York County Sample, 1933

	NO. OF PERSONS COVERED	NET INCOME IN 1933	
		Arith. mean	Median
		(dollars)	
New York County sample for lawyers in private practice[1]	2,667	6,664	3,210
Dept. of Commerce samples for New York City			
1935 sample[2]	51	3,897	2,000
1937 sample	52	5,961	2,750

[1] *Survey of the Legal Profession in New York County* (New York County Lawyers Association, 1936), p. 18. Arithmetic mean is not presented but was computed by us from frequency distribution.

[2] Number of persons covered represents actual number for whom information was reported. The mean and median incomes are corrected for firm member bias.

tice. These discrepancies between the samples may possibly account for the $703 difference between the arithmetic means from our 1937 sample and the New York County study; they are hardly sufficient to account for the $2,767 excess of the latter over the mean from our 1935 sample. However, an examination of the variability of income makes this difference appear less serious. The variability of income is so large that, for a sample the size of our 1935 sample, a difference of $2,767 would be exceeded by chance alone

[118] See *Survey of the Legal Profession in New York County.*

[119] 19,000 questionnaires were mailed.

between 9 and 16 times in a hundred.[120] The observed difference may therefore be attributable to sampling fluctuations. An additional check on our samples is that they show about the same percentage division of lawyers between individual practitioners and firm members as the New York County study. According to the latter, 70 per cent of all lawyers in New York County are individual practitioners and 30 per cent are firm members.[121] According to our 1935 sample, 66 per cent of all lawyers are individual practitioners; according to our 1937 sample, 69 per cent.[122]

vii *Incomes of consulting engineers.* The Bureau of Labor Statistics has recently published the results of an intensive study of the incomes of engineers.[123] This study covered engineers of all types, among them independent consultants. Questionnaires were sent to 173,151 professional engineers, 52,589 of whom replied. Data on income were requested for 1929, 1932, and 1934. Unfortunately, few data were prepared for consulting engineers separately. The only data on the income of consulting engineers in the final report are the median, quartiles, and first and ninth deciles of 'monthly engineering earnings'. In Table A 21 we have converted the medians and quartiles into annual earnings by simply multiplying by 12. This would clearly be an invalid conversion if the initial data were really full-time monthly earnings, the information requested, since annual earnings depend not only on the monthly rate but also on the fullness of employment. However, as the Bureau of Labor Statistics report states, "it is questionable if such a thing as a rate of compensation can be applied to this field of engineering service [consulting engineering], for, unlike the other

120 The reason for the range is that the mean from our sample is a weighted mean and hence its standard error is difficult to compute exactly. The figures cited are based on a standard error between $1,650 and $1,980.

121 *Survey of the Legal Profession in New York County*, p. 12. The data on which these percentages are based include all lawyers who reported organization of practice, whether or not they reported net income.

122 This comparison is an additional check on our conclusion (Sec. 1b iv) that the 1935 sample requires correction for a firm member bias but the 1937 sample does not. The percentage cited for the 1935 sample is corrected; the percentage for the 1937 sample is not. The uncorrected percentage from the 1935 sample is 45; the corrected percentage from the 1937 sample, 86.

123 Andrew Fraser, Jr. under the direction of A. F. Hinrichs, *Employment and Earnings in the Engineering Profession, 1929 to 1934*, Bul. 682 (Bureau of Labor Statistics, 1941).

TABLE A 21

Quartile and Median Net Incomes,
and Number of Persons Covered

Consulting Engineers: Department of Commerce and
Bureau of Labor Statistics Samples, 1929 and 1932

	1929		1932	
	Dept. of Commerce sample	BLS sample*	Dept. of Commerce sample	BLS sample*
Third quartile (dollars)	14,805	8,772	4,785	5,160
Median (dollars)	7,943	5,268	2,178	2,940
First quartile (dollars)	3,570	3,492	33	1,728
No. of persons covered	471	997	474	1,059

Andrew Fraser, Jr., *Employment and Earnings in the Engineering Profession, 1929 to 1934*, Bul. 682 (Bureau of Labor Statistics, 1941), pp. 184, 189.

* Medians and quartiles are corresponding measures for 'monthly engineering earnings' multiplied by 12.

kinds of engineering employment, the rates reported were almost necessarily derived directly from the earned annual incomes reported".[124] Despite this comment, the conversion may still not be entirely valid since some consulting engineers may have entered the amount normally earned during a month of reasonably full employment. Any error in the conversion would presumably be much more serious for 1932 than for 1929. Since the questionnaire was designed primarily for salaried employees, no explicit instructions were given that net rather than gross income be entered.[125] In consequence, some engineers may have reported gross income.

There is little similarity between the measures from the two samples. The measures from our sample are higher for 1929 but lower for 1932. The differences for 1929 are consistent with the

124 *Ibid.*, p. 187.

125 The actual question was:

"Earned income (please give data for each year):

From salaries or personal services in both engineering and nonengineering work

Average monthly rate from engineering work for time actually employed". *Ibid.*, p. 215.

known upward bias in our consulting engineering sample arising from the inadequacy of the sampling list (see Sec. 1a v). The differences for 1932, however, are in the opposite direction. Conceivably, these may reflect the incorrectness of computing the annual measures by multiplying the measures for 'monthly engineering earnings' by 12; measures computed directly from the Bureau of Labor Statistics data on annual earnings might be lower than those from our sample. We have been unable to test this possible explanation.