ON THE SIZE DISTRIBUTION OF WEALTH IN CANADA*

BY J. B. DAVIES
University of Western Ontario

This paper is concerned with measurement of the size distribution of personal wealth in Canada. The only available estimates of this distribution are those provided on the occasions when Statistics Canada's Survey of Consumer Finance has surveyed assets and debts. Results of the latest "SCF" to do this, that of 1977, are not yet available. The paper shows that the previous study, conducted in 1970, indicated wealth-inequality as viewed by top quantile shares roughly of the same order as estimated by others for the U.S. and U.K. A comparison of asset and debt aggregates implied by the survey, however, with independent totals indicates that for almost all items the SCF likely under-estimated true holdings. The possible relative importance of sampling and non-sampling errors in explaining this distortion is considered, drawing on Monte Carlo evidence and American validation studies of survey response. It is concluded that sampling error is unlikely to provide the explanation for SCF discrepancies in aggregates, but that non-sampling error is capable of doing so. Finally the 1970 SCF distribution of wealth is re-estimated. First a correction is made for hypothetical differential response according to true net worth. Second an attempt is made to remove the effects of under-reporting by respondents. The "best-guess" re-estimated distribution exhibits mean net worth considerably greater than shown by the SCF but only a slightly greater degree of concentration. Under certain fundamental assumptions this result is surprisingly robust. The appropriate conclusion is not that survey estimates of the distribution of wealth are reliable, but that the strong non-sampling errors affecting the 1970 Canadian SCF wealth estimates may have been composed of almost completely offsetting sources of bias.

I. INTRODUCTION

Renewed interest has recently been shown, in both the United States and Great Britain, in refining estimates of the size distribution of personal wealth.[1] Surprisingly little effort has been made along these lines in Canada, however, despite Statistics Canada's series of five sample surveys of wealth-holding, conducted in the course of its frequent "Surveys of Consumer Finance" (SCF's).[2] The purpose of this paper is to examine critically the latest available SCF estimate of the distribution of wealth, obtained in 1970, to see whether it can form the basis for a reliable estimate of the size distribution of wealth in Canada. Independent evidence developed here on the balance sheet of the household sector indicates sizeable discrepancies between survey and independent aggregates for most assets and debts. Whether these discrepancies can plausibly be attributed to survey

*This paper is based on Chapters 4 and 5 of my thesis *Life-Cycle Saving, Inheritance and the Distribution of Income and Wealth in Canada* presented to the University of London. I have benefited from the comments of A. F. Shorrocks, as well as the help of Gail Oja and Roger Love of Statistics Canada, Michael Wolfson of the Canadian Department of Finance, and John Whalley who provided the initial encouragement for this project. Responsibility for all errors and omissions is my own.

[1]See, for example, Atkinson and Harrison (1978) for Great Britain, and Smith and Franklin (1974) and Natrella (1975) for the U.S.

[2]The valuable contribution of Wolfson (1978) is concerned with the explanation rather than the measurement of the distribution. It should be pointed out, however, that Wolfson made estimates of summary indexes of inequality for the SCF data available for the first time.

error, or can provide significant guidance in re-estimating the distribution of wealth in Canada, are the principal issues addressed.[3]

The paper is organized as follows. Section II examines the SCF estimate of the size distribution of wealth in Canada in 1970 and compares the indicated level of wealth-inequality with those estimated by others for the U.S. and U.K. Section III then presents strong evidence that the SCF suffered from important biases, in a comparison of the aggregate holdings of all assets and debts implied by the survey with carefully-estimated independent totals. In Section IV the plausibility of the indicated survey bias is considered. Evidence from a Monte Carlo study, and 15 validation studies of consumer response to sample surveys conducted mainly in the United States, is examined in order to assess the possible impact of sampling and non-sampling errors. Sampling error, it is concluded, is extremely unlikely to have caused the degree of survey error indicated, while non-sampling error *is* capable of causing such bias. In light of this evidence Section V "corrects" the SCF distribution assuming sampling error negligible. A correction for hypothetical *differential response* according to size of wealth, based on American evidence, is followed by an adjustment for *under-reporting* of assets and debts by respondents. The result is an indication that wealth-inequality in Canada was likely only slightly underestimated by the 1970 SCF. The net effect of non-sampling errors—in themselves severe—on apparent inequality may have been mild due to the largely offsetting nature of the various elements of error.

II. The 1970 Survey of Consumer Finance Estimates

Table 1 shows the size distribution of personal wealth in Canada in 1970 according to Statistics Canada's Survey of Consumer Finance. "Wealth" is here the sum of financial and tangible assets minus all debts. Equity in life insurance plans, pension rights, and consumer durables other than cars and houses are excluded. A "family unit" is a group of persons related by blood, marriage, or adoption sharing a common dwelling. The survey results are based on interviews with the members of about 10,000 such units.

By the standards of international evidence the SCF-estimated shares of top wealth-holders shown in Table 1—18, 39, and 53 percent for the top 1, 5, and 10 percent respectively—might at first seem low. This is largely a mistaken impres-

[3]The well-known discrepancy between estate-multiplier estimates of the upper tail of the wealth distribution and that provided by the Federal Reserve 1963 Survey of Financial Characteristics (SFC) in the United States gives a clear indication of the possibility that survey results may be seriously misleading. Estate multiplier estimates of the share of the top 1 percent of individuals in the early 1960's range from about 20 to 25 percent whereas the SFC figure for the top 1 percent of *families* (among whom, one would expect, wealth was more equally distributed) was 37 percent. (See Lindert and Williamson (1977, p. 90).) In a country like Canada where only survey results are available it is thus clearly necessary to examine their reliability closely.

The estate-multiplier technique estimates the upper tail of the distribution of wealth among the living on the basis of that among the dying. It is generally considered to be more reliable than the survey approach at high wealth levels. (See Harrison (forthcoming).) Decedents over a period of time are a sample of those living at the beginning of the period, and on the (strong) assumption that imminence of death does not affect behaviour the numbers with wealth of given size can be estimated by applying the inverse of age–sex–social-class–specific mortality rates to the number of decendents with estates of that size. (See Atkinson and Harrison (1978) for a sophisticated discussion of the method.)

TABLE 1

MEASURES OF CONCENTRATION AND INEQUALITY FOR SCF-ESTIMATE OF CANADIAN DISTRIBUTION OF WEALTH, 1970[a]

	Family Units	Per Adult Basis
Share of top 1%	18.0%	17.1%
5%	39.2	38.3
10%	53.1	51.9
20%	70.7	69.3
next 40%	28.6	29.3
bottom 40%	0.7	1.4
V^b	2.315	2.274
G^b	0.715	0.698
E^b	0.527	0.509
Mean	$18,164	$10,262
Median	7,581	4,717

[a]Approximately 2 percent of family units for whom the value of business equity could not be ascertained are excluded from this table. This explains the difference from Table 2 in mean wealth.

[b]V = coefficient of variation, G = Gini coefficient, E = "exponential" index defined in text.

Source: Calculations described in Appendix A based on tabulations provided by Dr M. C. Wolfson of the Canadian Department of Finance.

sion. It is true that estate multiplier estimates of the equivalent shares for *individuals* in the U.S. and U.K. are much higher. For example, Atkinson and Harrison (1978, p. 123) place these shares at 33, 56, and 69 percent in 1970 in the U.K., and the IRS figures in the U.S. indicate shares of 25, 44, and 53 percent in 1969.[4] The individual basis of these estimates, however, raises estimated shares relative to the Canadian SCF—principally because husbands and wives, in general, do not own equal portions of family wealth. Atkinson and Harrison have estimated that in moving from an individual to a family basis shares of the top 1 and 5 percent would fall *at the most* by 5 and 8 percentage points respectively in the U.K.[5] A correction of this order would indicate shares for the top 1, 5, and 10 percent of families of about 28, 48, and 59 percent in the U.K., and about 21, 41, and 51 percent in the U.S. These figures suggest that the 1970 SCF top shares might be roughly of the same order of magnitude as top shares in the U.S., while apparent concentration in the upper tail is lower in both the U.S. and Canada than in the U.K.[6]

[4]Harrison (forthcoming) computes these shares from the data provided by Natrella (1975).

[5]See Atkinson (1975), p. 130 and Atkinson and Harrison (1978), p. 247.

[6]This comparison is fairly crude since both U.S. and U.K. estimates, unlike the Canadian, include equity in life insurance and all consumer durables. These inclusions no doubt lower the U.S. and U.K. shares relative to the Canadian. In addition the U.K. estimates include an allowance for occupational pension rights which tends to increase the estimated shares (Atkinson and Harrison (1978, pp. 97–98)). On a fully-comparable basis the SCF top shares would likely be of the same order as the U.S. estate-multiplier estimates while the U.K. shares would remain significantly above the North American.

In addition to quantile shares, Table 1 presents a set of complementary summary indexes of 1970 SCF wealth inequality—the Gini coefficient (G), coefficient of variation (V), and an "exponential" measure (E).[7] The indexes G, V, and E share two important properties which make them attractive indicators of relative levels of inequality. They are all insensitive to uniform proportional changes in all holdings (that is, they are *mean independent*) and they all satisfy the important *principle of transfers*. The latter requires that an index should always decline as a result of a transfer from richer to poorer, and increase as a result of a donation in the opposite direction. It is perhaps the strongest principle of inequality ranking which can be accepted by all reasonable observers. Despite agreeing in these two important ways, however, the indexes considered each represent a special point of view. G is most highly sensitive to changes in the middle range of a distribution. On the other hand V is, in practice, most sensitive to the upper range, and E will always be most sensitive to the lower range.[8] The three measures therefore provide contrasting and complementary points of view, and a great deal of information when considered together.

The second column of Table 1 provides some insight into the causes of apparent wealth-inequality, by transforming the distribution from a family unit to a "per adult" basis. The family unit basis, although useful for international comparisons, is fundamentally misleading as it introduces a spurious source of wealth-inequality. Far too much weight is placed on unattached individuals, and they are in addition allowed to appear worse off than members of families whose wealth *per adult* may be the same. Ideally, it can be argued on welfare grounds, units should always be weighted according to their number of members and income or wealth should be calculated on a "per person" basis.[9] The second column of Table 1 implements this approach. It is assumed that all SCF families with more than two members have just two adults (a good approximation) and that they share their wealth equally.[10] The result is a decline of from 1.8 to 3.4

[7]The "exponential" measure is closely related to that suggested by Wolfson (1978) except that it is normalized to vary between 0 and 1

$$E = \ln\left\{\frac{1}{n}\sum_{i=1}^{n}\exp\left(1-\frac{w_i}{\mu}\right)\right\}$$

where μ is the mean and w_i wealth of the ith unit.

[8]V, defined as the ratio of the standard deviation to the mean, may easily be shown to exhibit constant sensitivity to a small equalizing transfer at all wealth levels. See Love and Wolfson (1976, p. 60). This means that its sensitivity to *proportional* transfers would increase with wealth. E, on the other hand, exhibits declining sensitivity both to equal absolute and proportional transfers as wealth rises. See Davies (1979, Ch. 3).

Most other attractive indexes (e.g., Atkinson's and Thiel's), it should be noted, cannot be applied in the present case as they are either not defined, or are badly behaved, for negative values. This is also true for the variance of logarithms, which although popular is unattractive as it violates the principle of transfers.

[9]There is a question of who should be considered a member of the unit. In income studies it would be natural to include children. Here, however, wealth is treated as if it were purely the property of parents and of no concern to children. Similarly, in income studies it would be most appropriate to make the "per person" correction by dividing by "adult equivalent" units. Economies of scale in wealth-holding have, however, never been measured and such a correction can not be made here.

[10]The latter assumption of course conflicts with the legal division of family wealth, implying that the distribution cannot be compared with those estimated on an *individual* basis in the U.S. and the U.K.

percent in the summary indexes and a decrease of about one percentage point in each of the top quantile shares. Very little of apparent wealth-inequality seems, therefore, to be due to the inappropriate treatment of units of varying size.

It might be thought that a further correction of the distribution to remove the influence of age should be made. Such a correction cannot, however, be pursued here rigorously to any significant extent. Of the three summary indexes presented, for example, none provides an additive decomposition of inequality into "within-" and "between-" group components.[11] It is interesting to note, however, that V^2 which does provide such a decomposition indicates that only 5 percent of total wealth-inequality on a per adult basis is due to differences between 10-year age groups.[12] Averages of V, G, and E obtained by weighting age groups by population are actually slightly *higher* than overall levels in each case. We can say loosely, then, that inequality between age groups appears insignificant compared to that within.[13]

III. Balance Sheet Evidence

As already mentioned it is highly revealing to compare the aggregates for household assets and debts estimated by the SCF with estimates from independent (largely institutional) sources. Such a comparison is provided in Table 2. Readers should be cautioned that the independent figures are themselves subject to significant sources of error. They do, however, almost all agree in pointing to SCF under-estimation, and, as will be shown, provide much the same contrast that has been observed in comparison of survey and independent aggregates in the U.S. and U.K.

For a large part the independent estimates of Table 2 are based on the "institutional" approach and are similar to those presented for 1970 by Emmerson (1974). Cash, deposits, bonds of all kinds, equity in life insurance and pension plans, and all forms of debt are derived from Statistics Canada's Flow-of-Funds estimate of the financial balance sheet of the personal sector by deducting the estimated holdings of non-profit organizations and unincorporated business.[14] A major drawback is that household totals are obtained as residuals after the deduction of the holdings of all other sectors. Errors in global totals and in deductions cumulate to cause possible large error in the household estimates. The likely size of such error will of course be related to the quality of the data providing

[11]The "decomposition" advocated by Paglin (1975) for the Gini coefficient (well known not to have the decomposition property in general) is valid only when no member of any age group has wealth exceeding that of any member of an age group with higher mean wealth than his own, that is, there is no "overlapping".

[12]V^2 may be decomposed as follows:

$$V^2 = \sum_j P_j \left(\frac{\mu_j}{\mu}\right)^2 V_j^2 + V_B^2$$

where P_j, V_j^2, and μ_j are the population, coefficient of variation squared, and mean for group j respectively, and V_B^2 represents "between-group" inequality. See Love and Wolfson (1976, p. 67).

[13]A similar conclusion has been emphasized by Atkinson (1971, p. 248).

[14]To an extent deductions were simply equal, proportionally, to the holdings for these sectors in the 1968 balance sheet of the United States. (See Goldsmith (1971).) Published Canadian data were available, however, for cash, deposits, and bank loans.

TABLE 2

SCF ESTIMATE OF THE BALANCE SHEET OF THE HOUSEHOLD SECTOR (SPRING 1970), AND AN INDEPENDENT ESTIMATE (END OF FIRST QUARTER, 1970)

	SCF Estimate			Independent Estimate Amount	Method[a]
Financial Assets					
1. Cash		$484m		$1,000m	I
2. Bank deposits		9,650		20,600	I
3. Savings deposits in other institutions		4,101		11,600	I
4. Bonds		6,581		17,000	I
—Canada Savings Bonds	4,259		6,600		
—Other	2,322		10,400		
5. Shares		5,416		26,700	$\frac{(I+M)}{2}$
6. Life insurance and funded pension plans		n.a.		26,500	I
7. Other financial assets[b]		5,709		12,100	O
8. Equity in business interests		24,069		29,400	M
Tangible Assets					
1. Owner-occupied houses and vacation homes		67,496		67,500	SCF
2. Other real estate		9,123		18,000	M
3. Automobiles		5,663		5,700	SCF
4. Other consumer durables		n.a.		8,800	P
Total assets		138,292		244,900	
Debts					
1. All mortgages		14,121		12,800	I
2. Personal debt		6,773		15,200	I
—Consumer	5,027		10,700		
—Other bank loans	425		1,700		
—Other loans	1,321		2,800		
Total debts		20,894		28,000	
Net worth		117,398		216,900	
Net worth per family unit		18,225[c]		33,672	

[a]I = institutional method, M = investment income multiplier method, O = altered in same proportion as all other financial assets, SCF = SCF estimate retained after examination of alternatives, P = perpetual inventory method. (See text for explanation of these methods.)

[b]Includes annuities, mortgage holdings, interest in trust funds or estates, loans to persons, and royalties, copyrights, etc.

[c]This figure differs from mean wealth in Table 1 since Table 1 excludes families for which the value of business equity was not ascertained.

Source: Most of the SCF estimates are from Podoluk (1974, p. 207). The remainder are calculated from mean values given in Statistics Canada (1974), # 13-547. The independent estimates are derived by procedures described in the text.

global totals and deductions, and to the number and size of deductions. In view of these considerations, of the institutional estimates mentioned, that for non-savings bonds is perhaps the least reliable.

In order to supplement the institutional estimates national accounts investment income figures and data on average yields by type of asset were used to form investment income multiplier estimates wherever possible. For equity in business

interests and "other real estate" the estimates provided are the only ones available. The business equity estimate assumes the relationship between net income and net worth of unincorporated business in Canada the same as that in the U.S.[15] The real estate figure assumes equivalence of the after-tax rate of return on owner-occupied and other real estate.[16] Finally, the value of shares held by Canadian residents was estimated by applying the Toronto Stock Exchange dividend yield series to the national accounts estimate of Canadian dividends received, and Standard and Poor's yield for U.S. industrials to an estimate of foreign dividends.[17]

Estimation of share-holding is difficult, and crucial in an evaluation of the SCF distribution of wealth. Fortunately, totals obtained by the institutional and investment income methods are quite similar. The institutional figure for Canadian shares was $27.6 billion, the investment income figure $25.1 billion. Taking a simple average of these estimates and adding the investment income estimate for foreign stock of $0.4 billion the figure of $26.7 billion shown was obtained. Although this may appear outlandish in comparison with the SCF figure of $5.4 billion, previous estimates have placed Canadian share-holding in the $40–$60 billion range over the period 1966–68.[18] The present estimate, therefore, may be regarded as conservative.

Finally, perpetual inventory estimates were constructed for all tangible assets using national accounts expenditures and implicit price indexes, U.S. flow of funds service lives, and straight-line depreciation. Estimates of $51.0, $14.6, and $22.6 billion were obtained for houses, autos, and other consumer durables respectively. Both the housing and auto estimates were rejected in favour of the SCF, the former in view of strong evidence that house values are well estimated by the survey method (see Table 8), the latter in view of the fairly sophisticated method of survey estimation, which used reports of type of auto with subsequent valuation at standard used-car prices, rather than self-assessment. While the perpetual inventory estimate for housing is likely not reliable as a result of the use of a price index reflecting changes in construction costs rather than actual house prices, the figure for autos may be quite good if one wishes an estimate on a "going concern" basis. The preference for the SCF estimate reflects the choice, here, of a "realization" basis of valuation. The perpetual inventory estimate is only available for durables other than autos. It has been reduced by 61 percent in the independent balance sheet, on the hypothesis that the proportional difference

[15]A weighted average of net income of unincorporated business in Canada in 1969 and 1970, giving an estimate for Spring 1970 (the time of the SCF survey), is $4,947 million. Goldsmith's balance sheets for 1960 and 1968 in the U.S. indicate ratios of net income to net worth of this sector of 0.172 and 0.164 respectively. Taking the arithmetic average as an estimate of the Canadian ratio in 1970, estimated business equity of $29,400 is obtained.

[16]The rate of return on owner-occupied dwellings was estimated as the ratio of the national accounts imputed rent to the SCF-estimated equity in owner-occupied houses (7.6 percent). This was increased to 9 percent to give an estimate of the before-tax rate on other real estate.

[17]The national accounts total for foreign investment income is not available in disaggregated form. It was assumed, arbitrarily, that one-half of this income was from shares.

[18]Bhatia (1972, p. 94) estimated, on the basis of U.S. relationships between dividends and household share-ownership, that Canadian residents' shareholding totalled $54.9 billion at the end of 1968. Bhatia also quotes a study by Conway (unpublished) based on the value of outstanding shares of 101 large Canadian corporations which appears to indicate residents' holding of Canadian stock between $35 billion and $44 billion in 1966.

between "going concern" and "realization" bases might be similar to that for autos.

Table 3 allows us to examine differences in the composition of SCF-indicated net worth for "low", "middle", and "high" wealth-holders, and the proportional

TABLE 3

RELATIVE IMPORTANCE OF COMPONENTS OF NET WORTH FOR LOW, MIDDLE AND HIGH WEALTH HOLDERS, AND DEGREE OF SCF UNDERESTIMATION

Component	Wealth Groups[a] (1) Low	(2) Middle	(3) High	(4) Total	SCF Under-estimation[b]
1. Cash and chequing accounts	13.4%	2.7%	1.6%	2.7%	53%
2. Savings accounts and Canada Savings Bonds	34.0	13.5	11.3	13.4	54
3. Other bonds	1.9	1.3	2.5	1.9	78
4. Shares	4.8	2.5	4.0	3.3	80
5. Other financial assets	4.3	2.6	5.8	4.1	n.a.
6. Business equity	5.1	8.9	41.5	23.5	18
7. Owner-occupied houses and vacation homes	115.1	80.5	25.5	57.1	0
8. Other real estate	8.6	6.2	14.1	9.9	49
9. Automobiles	38.7	5.0	1.6	5.0	0
Total assets[c]	226.1	123.1	107.7	120.7	34
10. Mortgage debt	−60.0	−19.2	−5.7	−14.9	−10
11. Personal debt	−66.1	−3.9	−2.0	−5.8	55
Total debt	−126.1	−23.1	−7.7	−20.7	25
Mean wealth	$1,510	$20,570	$109,847	$18,225	35

[a] The low wealth group includes families with less than $8,000 net worth (51.1 percent of units), the middle group families with between $8,000 and $50,000 (41.8 percent of units), and the high wealth group families with $50,000 or more (7.1 percent of units).

[b] Calculated from same data as Table 2. No figure is shown for "other financial assets" as this paper does not, in fact, provide independent evidence on the underestimation of this category.

[c] Excluding equity in life insurance and pension plans, and consumer durables other than autos.

Sources: First four columns, same as Table 1. Final column, calculated from same data as Table 2.

rate of under-estimation indicated by the independent balance sheet for the various assets and debts. Four of the eight assets for which under-estimation is measured are both badly under-estimated and of disproportionate importance for high wealth-holders. These include non-savings bonds, shares, business equity, and other real estate. Paralleling this observation, personal debt is both very badly under-estimated and of far greater importance for low wealth-holders than high. If all under-estimation was due to non-sampling error such observations would almost certainly suggest that the true distribution would show a higher level of inequality than the SCF, unless under-reporting was strongly inversely correlated with the size of an asset or debt.

In contrast to the group of badly under-estimated assets which are "luxuries", two important assets are both severely under-estimated and clearly "necessities". Cash and chequing accounts, and savings accounts and Canada Savings Bonds, are both of uniformly declining importance as wealth rises, and are under-estimated by 53 and 54 percent respectively. In addition, mortgage debt, which appears to have been slightly *over-estimated* is of declining importance as wealth rises. These observations suggest the presence of sources of bias in estimated wealth-inequality that compete with those causing the under-estimation of stocks, bonds, etc. Correcting the survey estimation of these items would almost certainly decrease apparent wealth-inequality.

That the patterns of under-estimation identified in the SCF are not implausible is shown by Table 4 which presents comparisons between survey and

TABLE 4

COMPARISON OF SURVEY AND NATIONAL BALANCE SHEET AGGREGATES, VARIOUS STUDIES

Year	Asset or-Debt	Apparent Error in Survey Estimate[a]
A. Oxford Savings Surveys (U.K.)		
1952	liquid assets[b]	39%
1953	liquid assets[b]	48
1954	liquid assets[b]	34
1955	liquid assets[b]	35
B. Federal Reserve Board Survey of Financial Characteristics (U.S.)		
1958	liquid assets[b]	52
	corporate stock	74
	debt[c]	19
1963	savings accounts and bonds	46
	corporate stock	21
	debt[c]	15

[a]1 – (Survey Estimate/National Balance Sheet Estimate).
[b]demand and savings deposits in all institutions, and bonds.
[c]excludes farm debt.
Sources: Oxford Savings Surveys and 1958 Federal Reserve Board SFC, see Ferber (1966, pp. 26 and 35). 1963 Federal Reserve Board SFC, see Projector and Weiss (1966, p. 61).

national balance sheet aggregates in both the U.S. and U.K. Under-estimation of liquid assets—demand and savings deposits and all bonds—in the four U.K. surveys averages 39 percent, and in the two U.S. studies 49 percent, figures not too much less than the 59 percent recorded here. Under-estimation of corporate stock in the 1958 Federal Reserve SFC, at 74 percent, was close to the figure of 80 percent estimated here for the SCF.[19] Finally, the under-estimation of total debt

[19]The great improvement for this item in 1963 reflects the adoption of a very detailed questionnaire and the use of actual stock exchange prices to value shares named by respondents. Such procedures were not used in the 1970 Canadian SCF.

in the two Federal Reserve studies, averaging 17 percent, is not too much less than the 25 percent attributed to the SCF.

IV. Evidence on the Possible Importance of Sampling and Non-Sampling Error

If the independent balance sheet developed in the previous section is to be believed, the 1970 Canadian SCF badly under-estimated holdings of most assets and debts, and under-estimated overall net worth by 35 percent (excluding assets not covered by the SCF). This section asks how such severe error could possibly have occurred. It is necessary to take this step in order to justify, and prepare for, possible re-estimation of the distribution.

It has frequently been pointed out that under-estimation of both means and variances is likely in sampling from a highly-skewed distribution merely as a result of *sampling error.* In spite of the large sample employed by the SCF (approximately 10,000) it has been suggested that such error might explain an important part of the discrepancy between SCF and independent estimates of household net worth. It is not difficult to check such an assertion. Several theoretical frequency distributions are available which display shapes similar to observed distributions of wealth under suitable parameterizations. These can be given whatever degree of skewness one desires. A large number of repeated random samples from such a distribution may be taken with the help of the computer, using a sample of SCF size, in order to gain some idea of the approximate sampling distribution of mean SCF net worth.[20]

The three parameter Champernowne distribution (Champernowne (1952, p. 592))

$$F(w) = 1 - \frac{1}{\theta}\tan^{-1}\left[\frac{\sin\theta}{\cos\theta + (w/m)\alpha}\right] \tag{1}$$

was chosen for the Monte Carlo exercise in view of its easy parameterization. Here "w" stands for wealth, $F(w)$ is the cumulative distribution function, and "m" is the median, which may be set equal to unity in the following exercise without loss of generality. The distribution is asymptotic to the Pareto in its upper tail, the shape of the tail being determined by the parameter α. Pareto distributions fitted to the upper tail of income or wealth distributions typically give values of α in the neighbourhood of 1.5.[21] Using this value the parameterization may be completed by finding a suitable value of θ from the relationship

$$\bar{w} = \frac{\pi}{\theta}\left[\frac{\sin(\theta/\alpha)}{\sin(\pi/\alpha)}\right] \tag{2}$$

[20] Statistics Canada, of course, does not use a random sample, rather a stratified clustered scheme. Hence we may tend to underestimate, e.g., the sampling variance of the mean here. In view of the nature of the results this is a fairly minor consideration.

[21] Pareto himself believed α was typically 1.5 for the upper tail of the distribution of income. Atkinson and Harrison (1978, p. 25) suggest that a value of 1.5 is not unreasonable for Britain in the later 1960s. See Shorrocks (1975, Figure 1, p. 156) for an indication that upper tails within age groups are not as closely Pareto for wealth as might be supposed.

where θ is measured in radians (Champernowne (1952, p. 596)), given a hypothetical ratio of mean to median wealth. A ratio of 2.25 was taken giving $\theta = 0.982$.[22]

The result of the Monte Carlo exercise was a range of sample means from 1.98 to 2.92 times median, with under-estimation of the true mean (2.25 times median, of course) occurring in 71 of 100 samples. Under-estimation of more than 10 percent, however, occurred in only 4 cases. Maximum error was 12 percent. While under-estimation due to sampling error is therefore likely, it does not appear likely to be severe, given realistic skewness and a sample of SCF size.

The sampling distribution of the coefficient of variation, V, was also examined. This displayed far greater variability than that of the mean. The standard deviation of V, as a ratio of its mean, was more than ten times that of sample means (taken as a ratio to their mean). V's ranged all the way from 2.4 to 25.7! Sampling variation in other inequality indexes has not been investigated due to computing cost but it seems likely that although it would be less than in V (recall the sensitivity of V to extreme values) it would likely remain significant.

If SCF sampling error was indeed negligible, under-estimation would have to be explained by non-sampling error. This may take two forms. Error may occur, even if respondents report correctly, if response rates differ with wealth and there is no correction in the weighting procedure. This is referred to below as error due to *differential response.* On the other hand, even if response rates are uniform, error may occur as a result of *mis-* or *under-reporting.* The question which must now be asked is whether it is reasonable to suppose that errors of these types could explain the severe under-estimation of net worth attributed here to the SCF.

The only available evidence on variation of response rates with "economic status" in the SCF is a published breakdown by region and urbanization category. This is shown in Table 5. A clear but imperfect negative correlation of response rate with the prosperity of the different areas is apparent. Across regions the correlation coefficient for mean SCF family income and response rate, for

TABLE 5

RESPONSE RATES TO THE 1970 CANADIAN SCF BY REGION AND URBANIZATION CATEGORY

	Urban Areas 15,000+	Small Urban Areas	Rural Areas	All Areas	Mean Family Income (1969)
1. Atlantic Provinces	75.0%	77.7%	81.6%	78.0%	$5,864
2. Quebec	76.4	78.3	85.1	78.1	7,494
3. Ontario	68.9	80.2	73.7	70.8	8,559
4. Prairie Provinces	76.0	79.0	79.2	77.3	6,937
5. British Columbia	68.6	72.5	74.5	69.9	7,770
Canada	72.7	78.3	79.7	74.9	7,686

Source: Statistics Canada (1973, No. 13-547, pp. 80 and 175).

[22]The SCF itself indicates a ratio of 2.40 on a family basis, and one of 2.17 on a per adult basis.

example, is a sizable −0.74. Published estimates of family income within the more detailed categories are not available. However, in four of the five regions the likelihood of response steadily increases as we move from "metropolitan" to rural areas, providing further strong evidence of a negative association of response rates and "economic status" given the well-known income differentials between large and small towns, and rural areas.[23]

Better information on the crude relationship between response rates and economic status, again as measured by income, was provided by Projector and Weiss (1966) for the 1963 Federal Reserve SFC. Prior information from both census and tax sources was used to study the variation of response rates across income groups. The results are shown in Table 6. Aside from an initial rise,

TABLE 6

RESPONSE RATES TO THE 1963 FEDERAL RESERVE BOARD SURVEY OF FINANCIAL CHARACTERISTICS, BY INCOME CLASS

Income Class ($000)	Response Rate	Income Class	Response Rate
0–3	86.9%	15–25	72.9
3–5	89.1	25–50	65.5
5–7.5	85.0	50–100	50.3
7.5–10	82.9	100+	37.1
10–15	74.1		
		All	82.9

Source: Calculated from Projector and Weiss (1966, Table 15, p. 52).

response rates fell continuously with income, declining to very low levels for the highest income brackets. If a similar pattern held for the 1970 Canadian SCF one might expect distortion since the SCF weighting does not use prior information on income or wealth. Observations are weighted only according to family size, the labour force status of family head, and region.[24]

Considerable evidence on the impact of both differential response and under-reporting was compiled in a series of 15 validation studies on consumer response to financial surveys conducted mainly in the U.S. in the 1950s and 60s. These are surveyed in Tables 7 and 8.

Table 7 reports the results of validation studies where the true holding of a specific asset or debt was known in advance for those in the initial sample through contact with an institution such as a bank or loan company. In these cases the effect of differential response on the survey-estimated mean could be assessed by comparing the true mean of the initial sample with the true mean of those responding to the survey. The remaining error, the difference between the true and reported mean of respondents, is due to under-reporting. As the table shows,

[23]Statistics Canada (1973, No. 13-547, p. 182).

[24]I have sometimes received the comment that drawing attention to the strong correlation of regional response rates and income is misleading since the Statistics Canada weighting procedure *corrects* for differential response by region. This misses the point. The reason the regional correlation is interesting is that it suggests a *general* relationship, including a negative association of response rates and economic status *within* regions.

TABLE 7

VALIDATION STUDIES OF NON-RESPONSE AND REPORTING ERROR IN CONSUMER FINANCIAL SURVEYS

Organization[a]	Item	Year	Sample Size	Total Error[b]	% of Error Due to Non-Response	% of Error Due to Mis-Reporting
1. FRB	New Car Debt	1956	1,900	2.0%	28%	72%
2. SRC	New Car Debt	1957	105	6.2	−84	184
3. CSP	Farm Debt	1960	350	2.1	−88	188
4. SRC	Time Deposits	1958	170	41.1	23	67
5. CSP	Time Deposits	1959	300	51.9	48	52
6. CSP	Time Deposits	1960	350	66.8	44	56
7. NPO	Time Deposits	1958	3,300	13.7	20	80
8. FRB	Time Deposits	1963	1,100	46.0	40	60
9. CSP	Demand Deposits	1960	350	24.2	44	56
10. FRB	Corporate Shares[c]	1963	639	35.0	38	62

[a]FRB = Federal Reserve Board, SRC = Survey Research Center, CSP = Consumer Savings Project, NPO = Netherlands Post Office.

[b]1 − (Respondents' mean/true sample mean).

[c]Number of shares in selected corporations.

Sources: Studies 1–7: Ferber (1966, pp. 46, 50, 59, 66, 99, 111, 120). Study 8: Ferber *et al.* (1969, Table 4). Study 9: Ferber (1966, p. 143). Study 10: Ferber *et al.* (1969a, p. 417).

TABLE 8

SUMMARY OF STUDIES OF REPORTING ERROR IN CONSUMER FINANCIAL SURVEYS

Organization[a]	Item	Year	Sample Size	Reporting Error[b]
1. SRC	New Car Debt	1957	25	3.4%
2. SRC	New Car Debt	1957	33	30.1
3. FRB	House Value	1950	568	−3.8
4. Kain and Quigley	House Value	1967	251	−2.2
5. CSP	Farm Assets	1960	350	−6.6

[a]SRC = Survey Research Center, FRB = Federal Reserve Board, CSP = Consumer Savings Project.

[b]1 − (Respondents' mean/respondents' true mean).

Sources: Studies 1, 2 and 5: Ferber (1966, p. 59). Study 3: Kish and Lansing (1954, p. 520). Study 4: Kain and Quigley (1972, p. 804).

the two sources of error together can be very serious, causing error ranging up to 67 percent and averaging about 25–30 percent. Error in debt is small compared to that for deposits or shares, in agreement with the pattern implied by Table 2. Interestingly, the error in repeated studies of deposits is very similar to the error implied by the aggregate comparisons in the U.S. and U.K. reported in Table 4. Error due to differential response accounts, on average, for only about 10 percent of the total error, although in the majority of cases where it is positive it causes about 35 percent of the error on average.

Table 8 shows results of five further studies which obtained independent information only for respondent units. In two widely separated studies of home-ownership respondents' self-assessments were on average 2 and 4 percent above the values given by independent appraisers. This provides striking support for the suggestion made above that the 1970 SCF aggregate should be preferred to the perpetual inventory estimate. It is hard to believe that Canadian home owners make less shrewd assessments of their home values than their American neighbours. The net *over-reporting* of farm assets is also of interest in view of the rather small upward revision of business equity called for by the independent balance sheet of the previous section.

Finally, the validation studies provide important information on the nature of the under-reporting problem. It has been shown repeatedly that most of the reporting error results from complete non-reporting of holdings by some respondents. Broadly speaking, values, when reported, tend on average to be correct.[25] (The explanation of the good reporting of house values may well then be the difficulty of denying, or forgetting, home ownership.) This provides a clue as to the best method for obtaining accurate survey results. Considerable effort must be made to secure a report by the respondent. For this purpose a series of detailed questions, rather than a simple question like "How much stock do you own?", would seem to be appropriate. Secondly, a number of the studies provide evidence on the important issue of whether the degree of under-reporting varies by the size of holding. The two studies of house values indicate a lack of relationship. In view of the importance of housing equity this is significant, but implies little about the likely relationship for other assets, given the typical accurate reporting of average home value. Two studies of time deposits by Ferber (1966, pp. 101 and 112) indicate a moderate effect of true value of deposit on the frequency of complete non-reporting (which here, as elsewhere, accounts for most of the reporting error), while a slightly stronger relationship has been observed in the Federal Reserve 1963 national validation study on corporate stock (Ferber, *et al.* (1969a, Tables 3 and 5)). In contrast the 1963 national study on time deposits found a *negative* relation between non-reporting and true size of account (Ferber, *et al.* (1969, p. 439)). Since among those actually reporting holdings a fairly strong relationship between the under-reporting rate and true holding has been found for both time deposits and shares in the 1963 national studies (Ferber, *et al.* (1969, p. 441) and (1969a, p. 426)), it might be concluded on balance that under-reporting rates, although exhibiting no relationship to true values for housing, probably in general increase at least moderately with true holdings of financial assets.

In summary, there is ample evidence that survey errors can occur on the scale required to explain the balance sheet discrepancies indicated in the previous section. These are unlikely to be due to sampling error. Differential response typically explains a significant amount of error, but one which is considerably smaller than that due to under-reporting. The principal component of under-reporting is complete non-reporting of holdings, suggesting that one way to secure good survey reporting may be to make a strong effort to obtain some report where

[25]See, for example, Ferber (1966, p. 262) and Ferber *et al.* (1969, p. 444).

an actual holding is present. Finally the proportional increase in under-reporting with rising true holding may, in general, be moderate for financial assets, although further evidence on this score is needed.

V. Re-estimating the Distribution of Wealth in Canada in 1970

The previous two sections have provided sufficient information to guide and motivate a tentative "correction" of the SCF distribution of wealth. This section develops experimental corrections for both differential response and reporting error, and presents alternative re-estimated distributions of wealth for Canada in 1970. The re-estimates are based on alternative assumptions (from lower to upper bounds) on the severity of the two types of non-sampling error.

In order to determine the plausible range of required correction for differential response it is useful to examine the relationship between income and response rates in the 1963 Federal Reserve SFC, discussed above, in greater detail. Figure 1 shows that in this survey response rate was approximately a linear function of the log of income, above the lowest income group. This regularity suggests the possibility of making a tentative correction to the 1970 SCF for differential response according to size of *wealth*. It does not seem unattractive to assume that the Canadian relationship might have been similar, that is, that response rates may have declined linearly with the log of wealth.

If we accept the hypothesis suggested for the sake of experiment, there is still a *range* of plausible required corrections as a result of Statistics Canada's weighting by family size, labour force status of head, and region. To the extent that these variables correlate with the log of wealth the error introduced in the estimated distribution by differential response will already have been reduced. Suppose, for example, that the true relationship determining the response rate, r_j, in wealth range j is

$$r_j = \alpha \log (W_j) \tag{3}$$

where W_j is mean wealth in range j, and both r_j and $\log (W_j)$ are measured as deviations from overall means. Also suppose that the SCF estimate of response rate in range j, $\hat{r}_j$, is a linear function of a vector X reflecting the weighting characteristics listed above

$$\hat{r}_j = \beta X. \tag{4}$$

Then the accuracy of the $\hat{r}_j$'s clearly depends on the correlation of X and $\log (W_j)$. Empirical studies of the determinants of wealth-holding suggest that this correlation would not show R^2 above about 0.25.[26] In the suggested, oversimplified, model of underlying relationships this would imply R^2 of less than 0.25

[26]Projector and Weiss (1966, p. 7) report a regression of log wealth on log income, age, employment status of head, and inheritance status, which shows R^2 of 0.38. Statistics Canada (1974, No. 13-551, pp. 67–71) reports a regression of 1970 SCF net worth (not in log form) on income per person in a family, age of head, and home-ownership status, for Canadian-born and immigrant families. R^2 in the first case was 0.28, in the second, 0.17. The high correlation of income and net worth suggests that if income were omitted from these regressions R^2 would fall well below the figure of 0.25.

for $\hat{r}_j$ and log (W_j) as well, and therefore that at the most

$$\sum_j r_j^2 \cong 4 \sum_j \hat{r}_j^2. \tag{5}$$

Assuming arbitrarily that the proportional error was constant across all j this would give

$$r_j = 2\hat{r}_j, \qquad j = 1, \ldots, n \tag{6}$$

or sufficient information to correct the SCF weight given to each wealth class in a "best-guess" re-estimate. The lower-bound assumption, in contrast, is that $r_j = \hat{r}_j$ (all differential response corrected by weighting), while the upper-bound assumption is that $\hat{r}_j = 0$ (no differential response corrected by the weighting procedure).

In order to "correct" the SCF weighting scheme for differential response it is necessary to calculate the hypothetical r_j's. To do this an OLS regression was first run using the data of Figure 1, explaining response rates in the 1963 SFC as a function of the log of income, omitting the lowest income group. ($R^2 = 0.98$). Percent deviations of the predicted response rates from their mean at both median

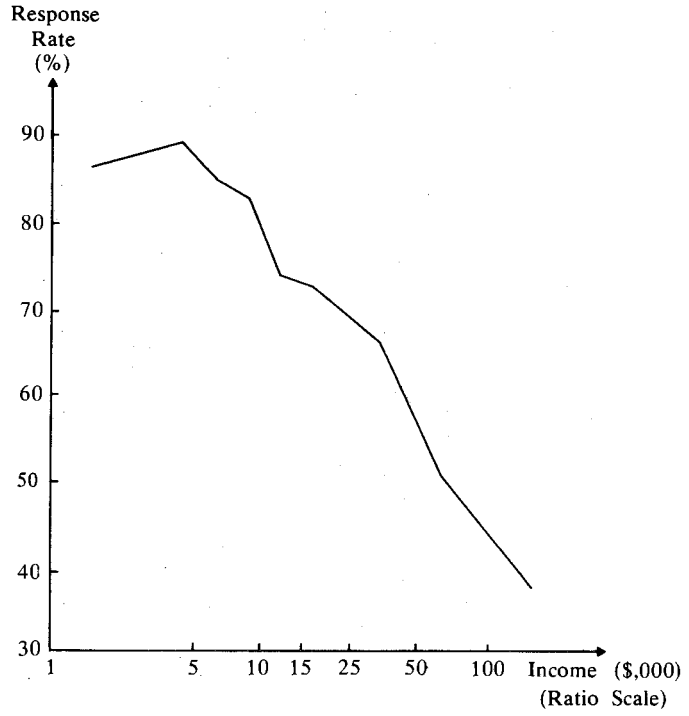

Figure 1. Relationship Between Response Rate and Income, Federal Reserve SFC, 1963. Source: Data required are in Projector and Weiss (1966, pp. 52 and 60).

and upper quartile income were noted. The percent deviations of Canadian response rates at both median and upper quartile *wealth* were then assumed the same, thus defining a linear relationship between log of wealth and the probability of response in the Canadian case. Finally, the response rate for those with negative wealth (13 percent of adults) was set at 82.1 percent, deviating from the mean Canadian rate to the same extent as the American observation for the lowest income group does from the mean American rate.

If under-reporting of assets and debts were largely asset-specific the greater part of non-sampling error remaining in the SCF estimate after the differential response correction could be removed by adjusting all holdings of a particular asset or debt by a uniform multiplier. Letting a_{ik} stand for the holdings of asset k by unit i the correction suggested would give re-estimated wealth

$$w_i' = \sum_k \frac{A_k'}{A_k} a_{ik} \tag{7}$$

where A_k' and A_k are the aggregates indicated by the independent balance sheet and SCF respectively for asset k.[27] This should provide a lower-bound on the possible influence of under-reporting in view of the evidence reviewed in the previous section that under-reporting increases in severity with the size of true holding for at least some financial assets. An upper bound, on the other hand, may perhaps be obtained by assuming that the proportion of an asset or debt reported declined steadily as the true holding increased, for *all* assets and debts. For the sake of experiment it was therefore hypothesized that

$$\frac{\hat{a}_{ik}}{a_{ik}} = \gamma_0 a_{ik}^{(\gamma_1 - 1)}; \qquad 0 < \gamma_1 < 1 \tag{8}$$

where $\hat{a}_{ik}$ and a_{ik} are the reported and true holdings of asset k by unit i respectively. Equation (8) may be written, equivalently, as

$$\hat{a}_{ik} = \gamma_0 a_{ik}^{\gamma_1}; \qquad 0 < \gamma_1 < 1 \tag{9}$$

calling attention to the fact that not only the elasticity of the reporting rate, but also that of reported holding, is assumed constant with respect to the true holding. With a value of $\gamma_1 = 0.9$, one might suppose, equation (9) would suggest something like an upper bound correction for the effect of under-reporting in the 1970 SCF.[28] For the sake of producing an intermediate re-estimate a "best-guess" correction was also made using $\gamma_1 = 0.95$. By imposing a uniform pattern of

[27]Data are in fact not available for individual family units. The correction is made at the level of the age–family size–wealth class and all families' wealth within that class is revised in equal proportion. See Appendix A for description of the method used to create a sample of families from the tabulations of 1970 SCF data made available.

[28]The implications of this choice of parameter value may be shown as follows. Assume that at mean true holding of some asset the under-reporting rate was equal to the overall rate, say, 25 percent for the sake of example. Then at twice mean true holdings under-reporting would be 30 percent, at five times the mean it would be 36 percent, and at ten times the mean the rate would be 40 percent. This may seem a moderate rate of increase. However the evidence that under-reporting becomes proportionally more severe as true holdings rise does not appear strong enough to warrant a more rapid increase, especially since imputed true holdings of stocks and bonds in the upper tail may be up to 20–50 times average, leading to quite a high rate of hypothetical under-reporting with $\gamma_1 = 0.9$.

worsening reporting as true holdings increase this correction may well itself be somewhat extreme, but, as the discussion below indicates, choosing a different "best-guess" correction would have little impact on the nature of the results.

The results of five alternative corrections to the SCF estimates on a family basis, using the upper and lower bound, and "best-guess" assumptions discussed, are shown in Table 9.[29] Perhaps most immediately striking is the indication that differential response error, if not completely corrected by the survey weighting, may have caused the SCF-estimated distribution to be *more*, rather than *less*, unequal than it would otherwise have been. The explanation is simply that, following the 1963 Federal Reserve SFC pattern of differential response according to income (see Figure 1), we have assumed low wealth-holders had exceptionally high response rates as well as that the wealthy had very low response. While correcting the latter distortion tends to make the re-estimated distribution more unequal, removing the former hypothetical bias is strongly equalizing. Although with a different assumed pattern of differential response the direction of the net effect on re-estimated inequality could easily be reversed, it is clear that the differential response problem does not necessarily reduce SCF-estimated wealth-inequality. This is an important conclusion since it is sometimes supposed that the effect of differential response in surveys of wealth-holding must be to produce downward bias in estimated inequality.

Table 9 shows that the plausible range of impacts of non-sampling error, given our basic assumptions, is surprisingly narrow. The difference between the two most extreme corrections in terms of the top shares and summary indexes does not exceed 15 percent. Under the fundamental assumptions made, the re-estimated distribution is strikingly robust.[30] It is therefore very interesting that the "best guess" figures indicate the need for only a small upward revision in the 1970 SCF view of wealth-inequality in Canada. The explanation is that although non-sampling error is no doubt large, offsetting effects on estimated inequality are here attributed to different aspects of this error. The impact of the low response of the rich, for example, is more than offset by the high response of the poor, as pointed out above. Similarly the severe under-reporting of some assets, such as bonds and corporate shares, competes with the also significant under-reporting of assets like cash, and bank and other accounts, which are of disproportionate importance for those with low net worth.

VI. Conclusion

This paper has shown, first, that the aggregate holdings of assets and debts by the household sector in Canada indicated by the 1970 Canadian Survey of Consumer Finance are far below those which can be derived by an independent,

[29]The corresponding results on the "per adult" basis are shown in Table 11 of Appendix B. Note that a comparison of the 1970 SCF and "best guess" re-estimated frequency distributions is also given in this appendix, in Table 10.

[30]These fundamental assumptions are (a) that the independent balance sheet of Table 2 is correct, (b) that differential response by wealth size before Statistics Canada's weighting had the same pattern as that according to income found in the 1963 Federal Reserve SFC, and (c) that proportional under-reporting does not decline with rising size of true holding of any asset or debt.

TABLE 9

ALTERNATIVE RE-ESTIMATES OF THE DISTRIBUTION OF WEALTH[a] IN CANADA BASED ON THE 1970 SCF AND INDEPENDENT EVIDENCE, FAMILY UNITS

Differential Response Assumption[b] / Under-reporting Assumption[c]	Lower Bound	Upper Bound	Lower Bound	Upper Bound	"Best Guesses"
	Lower Bound		Upper Bound		
Share of top 1%	19.2%	18.4	21.2	20.2	19.6
5%	42.8%	41.6	45.7	44.6	43.4
10%	57.5%	56.6	59.8	59.0	58.0
20%	74.2%	73.2	75.3	74.4	74.0
next 40%	25.9%	26.5	24.7	25.2	25.8
bottom 40%	−0.1%	0.3	0.0	0.4	0.2
Coefficient of Variation	2.463	2.403	2.686	2.620	2.519
Gini Coefficient	0.748	0.737	0.758	0.747	0.746
"Exponential" Index	0.562	0.551	0.573	0.562	0.560
Mean[d]	$27,600	27,600	27,600	27,600	27,600
Median	$10,900	11,200	10,200	10,800	11,000

[a]"Wealth" is defined as in the 1970 SCF except that consumer durables omitted from the survey have been imputed.

[b]Lower bound assumes all differential response has been corrected by SCF weighting procedures, upper bound that none has been corrected, and "best guess" that one-half has been corrected.

[c]Lower bound assumes the percentage reporting rate for a given asset or de!• is uniform across families; upper bound that it declines with respect to the true holding with an elasticity of −0.1 for financial assets, real estate, business equity and debts (no decline for homes and durables); and "best gue•;" that it declines with elasticity of −0.05.

[d]Re-estimated mean wealth differs from mean net worth of the independent balance sheet of Table 2 since (a) the re-estimated distributions omit families for which the value of business equity was not ascertained, and (b) insurance equity and pension rights are excluded here.

Source: Computations described in text.

"institutional" approach. A carefully estimated complete independent balance sheet for the household sector at the time of the SCF survey has been provided. This indicates overall survey under-estimation of net worth by 35 percent, with rates of under-estimation ranging up to 80 percent for certain assets. Similar discrepancies between survey and national balance sheet aggregates in the U.S. and U.K., it is pointed out, have been observed in at least six cases.

The question of how SCF under-estimation on the indicated scale could have occurred has been examined closely. A Monte Carlo exercise has led to the dismissal of sampling error as a likely major component of the explanation, although the sizable impact of this type of error on estimates of overall inequality has been pointed out. Evidence from 15 validation studies of consumer response to sample surveys of asset-holding has also been examined. This shows that non-sampling errors can be as large as those imputed here to the SCF. Errors vary considerably according to the characteristics of assets and debts, and the type of variation is similar to that indicated by the balance sheet exercise of the present paper. The greater part of non-sampling error apparently may be attributed to under-reporting, specifically to complete non-reporting by some respondents. Finally it was found that there is evidence of a moderate positive relationship between under-reporting rates and size of holding for financial assets.

The paper has concluded by presenting a range of alternative re-estimates of the distribution of wealth in Canada. These draw on the pattern of differential under-estimation indicated by the balance sheet comparison as well as information gained from our review of foreign studies of non-sampling error. An attempt has been made to place upper and lower bounds on the required revision of estimated wealth-inequality. Under the fundamental assumptions that the independent balance sheet presented in this paper is correct, that the pattern of differential response according to *wealth* for the 1970 Canadian SCF was similar to that according to *income* in the 1963 Federal Reserve SFC (i.e., before Statistics Canada's weighting procedure), and that proportional under-reporting does not decline with the true size of any asset or debt, we have found that the range of plausible impacts of non-sampling error is quite narrow. The re-estimated distribution is therefore surprisingly robust, and it is extremely interesting that it indicates a fairly small required upward revision of estimated wealth-inequality in Canada according to conventional indicators. As pointed out in the final section, the reason is *not* that the 1970 SCF did not suffer badly from non-sampling error. Rather, the errors were undoubtedly severe, as the balance sheet comparisons of this paper show, but the separate impacts of different aspects of non-sampling error were largely offsetting.

The exercise carried out in this paper has important implications for current and future research. First, the balance sheet comparisons and evidence on differential patterns of response and under-reporting for many assets make clear that the reliability of studies using estimated distributions of individual assets and debts from even sophisticated large-scale surveys of consumer finance is problematic. This is an important point since an increasing number of applied studies of personal distribution with significant asset and debt data requirements are being conducted. The only possible exceptions to this conclusion are studies using survey evidence on owner-occupied houses, and perhaps mortgages and auto-ownership. Second, while the re-estimation procedures of the final section have shown that in the case of the 1970 Canadian SCF the effects of non-sampling errors on estimated wealth-inequality might well have been largely offsetting, there can clearly be no reliance on this kind of "self-correction" in the long run. In a country like Canada where household surveys provide virtually the only method of ascertaining the distributions of assets, debts, and net worth, there must evidently be a constant search for new means of reducing survey error. It is quite possible that an acceptable degree of accuracy will not be achieved without pecuniary incentives to respondents or, indeed, some form of compulsory compliance. Nonetheless it is encouraging that Statistics Canada's 1977 survey of assets and indebtedness—preliminary results of which are now available—has taken steps to achieve significantly greater reliability than the 1970 SCF. It is to be hoped that similar renewed efforts will be made at frequent intervals in the future.

BIBLIOGRAPHY

Aigner, D. J. and Goldberger, A. S., "Estimation of Pareto's Law from Grouped Observations," *Journal of the American Statistical Association* June, 1970.

Atkinson, A. B., "On the Measurement of Inequality," *Journal of Economic Theory*, September, 1970, also in Atkinson (1973).

———, "The Distribution of Wealth and the Individual Life Cycle," *Oxford Economic Papers*, July, 1971.

———, (ed.) (1973), *Wealth, Income and Inequality* (Penguin; London).

———, (1975), *The Economics of Inequality* (Clarendon Press, Oxford).

Atkinson, A. B. and Harrison, A. J., *Distribution of Personal Wealth in Britain*, Cambridge University Press: Cambridge, 1978.

Bhatia, K. B., "Estimation and Taxation of Capital Gains in Canada," *Canadian Tax Foundation 1971 Conference Report*, 1972.

Champernowne, D. G., "The Graduation of Income Distributions," *Econometrica*, July 1952.

Davies, J. B., *Life Cycle Saving, Inheritance and the Distribution of Income and Wealth in Canada*, Thesis presented to the University of London, 1979.

Davies, J. B. and Shorrocks, A. F., "The Quantitative Importance of Inheritance in the Distribution of Wealth," *Oxford Economic Papers*, March, 1978.

Emmerson, F. W. (1974), Appendix to Podoluk (1974).

Ferber, R., *The Reliability of Consumer Report of Financial Assets and Debts*, Bureau of Economic and Business Research, University of Illinois, Urbana, 1966.

Ferber, R., Forsythe, J., Guthrie, H. W., and Maynes, E. S., "Validation of a National Survey of Consumer Financial Characteristics: Savings Accounts," *Review of Economics and Statistics*, November 1969.

———, "Validation of Consumer Financial Characteristics: Common Stock," *Journal of the American Statistical Association*, 1969a.

Goldsmith, R. W. (ed.), *Institutional Investors and Corporate Stock*, Columbia University Press: New York, 1971.

Harrison, A. J., *Frequency Functions to Describe the Distributions of Income and Wealth*, Thesis presented to the University of Essex, 1978.

———, *The Distribution of Wealth in Ten Countries*, Royal Commission on the Distribution of Income and Wealth: London, forthcoming.

Kain, J. F. and Quigley, J. M., "Note on Owner's Estimate of Housing Value," *Journal of the American Statistical Association*, December 1972.

Kish, L. and Lansing, J. B., "Response Errors in Estimating the Value of Houses," *Journal of the American Statistical Association*, September 1954.

Lindert, P. H. and Williamson, J. G., "Long-Term Trends in American Wealth Inequality," University of Wisconsin, Institute for Research on Poverty: Discussion Paper No. 472-77, 1977.

Love, R. and Wolfson, M., *Income Inequality: Statistical Methodology and Canadian Illustrations*, Statistics Canada, No. 13-559, 1976.

Natrella, V., "Wealth of Top Wealth-Holders," paper presented to the 135th Annual Meeting of the American Statistical Association, 1975.

Nordhaus, W. D., "The Effects of Inflation on the Distribution of Economic Welfare," *Journal of Money, Credit and Banking*, February, 1973.

Paglin, M., "The Measurement and Trend of Inequality: A Basic Revision," *American Economic Review*, September 1975.

Podoluk, J. R., "Measurement of the Distribution of Wealth in Canada," *Review of Income and Wealth*, June 1974.

Projector, D. S. and Weiss, G. S., *Survey of Financial Characteristics of Consumers*, Federal Reserve Board: Washington, D.C., 1966.

Sen, A. K., *On Economic Inequality*, Clarendon Press: Oxford, 1973.

Shorrocks, A. F., "The Age-Wealth Relationship: A Cross-Section and Cohort Analysis," *Review of Economics and Statistics*, May, 1975.

Singh, S. K. and Maddala, G. S., "A Function for Size Distribution of Incomes," *Econometrica*, September, 1976.

Smith, J. D. and Franklin, S. D., "The Concentration of Personal Wealth, 1922–1969," *American Economic Review*, Papers and Proceedings, Vol. 64, 1974.

Statistics Canada (1973, No. 13-547), *Income, Assets, and Indebtedness of Families in Canada, 1969.*

———, (1974, No. 13-551), *Survey of Consumer Finances, Volume II, Selected Reports, 1970.*

———, (1976, No. 13-002), *Financial Flow Accounts, System of National Accounts, 4th Quarter 1975.*

Wolfson, M. C., *The Causes of Inequality in the Distribution of Wealth, A Simulation Analysis*, Ph.D. thesis, Cambridge University, 1977.

Appendix A

Method of Estimating Inequality Measures

This appendix describes the method used to calculate the inequality measures presented in this paper. A detailed tabulation giving the composition of wealth for 18 wealth classes within seven age groups and two family size groups was made available by Dr M. C. Wolfson of the Canadian Department of Finance. Estimation of the inequality indexes on the assumption of degenerate distributions within wealth classes was rejected as liable to introduce significant bias. Instead a theoretical distribution suggested by Singh and Maddala (1976),

$$F(w) = 1 - \frac{1}{(1 + a_1 w^{a_2})^{a_3}}$$

was fitted for each age–family-size distribution by the method of χ^2 minimization (see Aigner and Goldberger (1970) and Harrison (1978, Ch. 4)). (Fitted three parameter Champernowne distributions were in each case inferior.) Samples of size 500 were then taken for each age–family-size group in such a way that the SCF-indicated frequencies and means were retained for wealth classes. The purpose of the fitted distribution is merely to interpolate within wealth classes and allow extrapolation into the open-ended upper and lower extreme classes.

Samples of size 500 were chosen as this is the size of sample used in the simulation models of Davies (1979) and it was desired to keep grouping error as constant as possible in that study. It should be noted that as sample size increases so do estimated inequality indexes. It is to be hoped, however, that with overall sample size of 7,000 this problem has largely been eliminated here.

Appendix B

Selected Estimated Frequency Distributions and Re-estimated Per Adult Distributions

This appendix provides further detail on both the SCF and re-estimated distributions of wealth in Canada in 1970. Table 10 shows frequency distributions on both family and "per adult" bases indicated by the 1970 SCF and the "best guess" re-estimated distribution described in the text. Note, as pointed out in the text, that the "per adult" basis is not equivalent to an *individual* basis, since it divides the wealth of couples equally between husband and wife whereas the individual basis would follow the split dictated by legal ownership. It should also be noted that the per adult version of the SCF distribution is approximate as its derivation required the assumption that all SCF families of two or more members had precisely two adult members.

Table 11 shows inequality measures for the experimental re-estimates of Table 9 on a *per adult* basis. Inspection of the table reveals a pattern very similar to that found using family units, except that the inequality measures are in all cases somewhat lower than for families.

TABLE 10

FREQUENCY DISTRIBUTIONS—1970 SCF AND "BEST GUESS" RE-ESTIMATED DISTRIBUTIONS OF WEALTH[a]

Wealth Class	1970 SCF		"Best Guess" Re-estimates	
	Family Units	Per Adult	Family Units	Per Adult
Negative	12.6%	12.5%	15.4%	15.6%
Under $4,999	31.2	38.9	23.3	29.1
$5,000–9,999	11.6	18.6	9.7	17.5
$10,000–19,999	17.8	17.8	16.8	19.7
$20,000–49,999	19.6	9.8	22.4	12.3
$50,000–99,999	5.3	1.6	7.1	3.4
$100,000 and over	2.0	1.0	5.2	2.4
Mean	$18,200	$10,300	$27,600	$15,600
Median	7,600	4,720	11,000	6,300

[a]In the re-estimate "wealth" is defined in the same way as in the SCF except that an imputation is made for omitted durables.

Sources: The same calculations on which Tables 1 and 9 are based.

TABLE 11

ALTERNATIVE RE-ESTIMATES OF THE DISTRIBUTION OF WEALTH[a] IN CANADA
Based on the 1970 SCF and Independent Evidence, Per Adult Basis

Differential Response Assumption[b]	Lower Bound	Upper Bound	Lower Bound	Upper Bound	
Under-reporting Assumption[c]	Lower Bound		Upper Bound		"Best Guesses"
Share of top 1%	18.3	17.6	20.4	19.5	18.8
5%	42.1	41.1	45.1	44.2	42.9
10%	56.6	55.6	59.0	58.1	57.1
20%	73.4	72.6	74.5	73.9	73.3
next 40%	26.1	26.5	24.9	25.1	25.8
bottom 40%	0.6	1.0	0.7	1.1	0.8
Coefficient of Variation	2.438	2.422	2.688	2.677	2.529
Gini Coefficient	0.736	0.726	0.746	0.737	0.734
"Exponential" Index	0.550	0.540	0.561	0.551	0.548
Mean	15,600	15,600	15,600	15,600	15,600
Median	6,300	6,600	6,200	6,500	6,300

[a,b,c]See Table 9.

Source: Computations described in text.