ON TRENDS IN THE GAP BETWEEN RICH AND POOR IN LESS DEVELOPED COUNTRIES: WHY WE KNOW SO LITTLE

BY ALBERT BERRY
University of Toronto

This paper considers the problems involved in measuring trends over time in inequality in less developed countries. After considering some of the conceptual problems involved in choosing a measure of economic welfare, the period it should cover, and the statistical units to which it should be applied, the paper goes on to draw up a list of minimum data requirements for reaching reliable conclusions about such trends. It concludes that in many countries the available information falls well short of the minimum list, and it explores various sources of indirect evidence on trends in distribution. The central argument is that the available data permit no easy tests for trends in the level of economic inequality in less developed countries. At present, the best bet is to complement evidence on income distribution with available data on consumption distribution and on wage trends and production aggregates by occupational and sectoral groups. Inconsistencies will highlight problem areas, and their reconciliation should provide a firmer foundation on which to draw conclusions about distributional trends.

1. INTRODUCTION

Increasing concern with economic inequality and its trends in less developed countries has by now generated a reasonably large theoretical and empirical literature. Assessment of the current level of inequality with sufficient precision for many purposes is possible given the body of information now available on perhaps a third to a half of less developed countries, i.e. it is possible without many of the refinements which would in principle be desirable. But the equally important analysis of inequality trends over time is a different matter. Much evidence suggests that the level of inequality tends, in most countries and under most conditions, to change slowly. Hence, given the short periods of time for which data are typically available, a much higher level of precision is needed to identify trends with confidence. In the light of the problems reviewed below, it appears unlikely that we "know" (with 99 percent confidence, say) the direction of change in economic inequality in any less developed country. This review is directed towards drawing up a list of minimum data requirements for reaching reliable conclusions about such trends. We touch first on some of the basic conceptual issues, then discuss empirical aspects of the question. In many countries, and especially in the mainly agricultural countries of Sub-Saharan Africa, the available set of information falls well short of the minimum list, so the likelihood of accurately assessing trends in inequality in the near future is not high. Much of the discussion below is primarily relevant to countries somewhat farther along the development path, though such poor Asian countries as India, Bangladesh and Sri Lanka have relatively good data as well.

2. INEQUALITY IN THE DISTRIBUTION OF WHAT AMONG WHOM!

Different types of inequality are of interest in the analysis of different issues. Our concern here is the distribution of economic welfare among individuals and

the measurement of its changes over time. Since most distributional data on income and consumption have the family or the income earner as the unit of observation, they are not directly appropriate for the analysis of distribution among persons. Conceptual and empirical problems also arise when one addresses the question of the intra-family distribution of welfare.[1] There is, to begin with, a considerable amount of joint consumption; for consumption which is not of this kind, there is very little information on its distribution among family members. The distribution between spouses and between adults and children may vary considerably across cultures and even over time in a given culture.[2] This issue requires much more attention than has thus far been dedicated to it. A number of recent studies have opted for per capita family income (or consumption) as the measure of an individual's economic situation. Apart from the heroic assumption of equal distribution among the family members, another questionable aspect of this approach is the implicit assumption of equal needs as between adults and children. A preferable approach would be to apply coefficients reflecting relative expenditures on children of various ages in the calculation of per capita income.[3] Usually, though, the available data do not include the age structure of families by income (or consumption) level, so this refinement is not possible.

Apart from the definition of the person-unit and the appropriate allocation of intra-family income, another set of issues involve the variables used to measure economic welfare and the period covered by the measurement. The most conceptually plausible indicators of the relative material well-being of a person are measures of lifetime consumption or potential lifetime consumption (e.g. actual consumption plus wealth at end of life).[4] The distribution of current income (whether among families, income earners, or persons) has, however, been the focus of most discussions. This emphasis has partly resulted from the central position of that variable in national accounting, and perhaps also from the fact that potential consumption is closely related to income. The value in use of public consumption (i.e. expenditures undertaken on people's behalf by government) should be, in principle, added to private consumption to arrive at total consumption, or to private disposable income to arrive at total income. This is not frequently attempted due to data and measurement problems[5] and, in the case of true public goods, to more difficult conceptual problems. Practical factors greatly affect the relative usefulness of alternative indicators. Several considerations favour current

[1]Many of these questions also involve the problem of comparisons across generations and life-cycle issues in measurement of economic welfare. These matters have received some empirical discussion (see Gary Becker, *A Treatise on the Family*, Cambridge, Harvard University Press, 1981) but the distribution among members of the family at a point of time has not.

[2]Many cultural differences which may affect the relative well-being of various members of the household would not show up in economic variables so much as in matters of authority, acceptable behaviour patterns and the like.

[3]Such an approach would involve the implicit assumption that observed relative expenditures by age within the family reflect relative needs.

[4]Potential consumption may be greater than actual consumption if the individual places a high enough value on risk aversion, on passing wealth on to the next generation, or on simply being wealthy. Clearly it would be inappropriate to consider a person poor if his wealth is high, even if his consumption is low. In this sense, then, potential consumption is the better measure of material welfare.

[5]The allocation of public expenditures is hard to assess; further, there is less reason to presume a relation between cost and value in use than in the case of private expenditures.

(private) consumption over current income.[6] First, private consumption is usually better measured. In most income and expenditure surveys, the estimated rate of savings is below that suggested by other types of information such as changes in the stock of financial and other assets, implying that consumption is more fully recorded than income. Secondly, consumption varies less over time than does income, making it a better indicator of relative welfare than is income.[7]

A widely recognized source of ambiguity or error in point of time distribution estimates relates to life cycle phenomena and the fact that age may be an important correlate of income.[8] Because a distribution (whether of earners, persons or families) at a point of time reflects both the distribution among units at given age and differences in average income across age categories, the overall distribution is characterized by greater inequality than that of the typical age cohort. A distribution which reflects average income or consumption of individuals over lifetime would be more interesting.[9] Applying such a concept involves several decisions. Selection of the group within which one wishes to assess inequality is not obvious. One may wish to focus on the over life income (or consumption) of persons currently of a given age. Consideration also of inequalities across generations requires a decision as to which (if not all) existing cohorts should be included, with the corresponding data requirements stretching to the distant past and the distant future. Calculation of lifetime incomes requires also a formula for aggregating income over time (i.e. what discount rates to use) and a way to deal with different life spans. Since on average the rich live longer than the poor,

[6]Information on wealth is the hardest to obtain with a reasonable degree of accuracy, and fewer attempts have been made at the estimation of its distribution than for income and consumption.

[7]Usually data on income refer to the previous month or even week; the period of payment frequently determines the period to which the income data refer. The level of measured inequality is higher the shorter the reference period, but there is at present no organized information on the sensitivity of the former to the latter. The sensitivity is probably enough to create problems for cross country or over time studies using surveys with different time periods. It is probably a more serious problem in analysis of distribution among earners than among families or persons. The question of how unemployed persons, unpaid family helpers and persons entering and leaving the labour force frequently should be treated in calculations of distribution among earners is related to this question. The unemployed constitute the extreme situation of current income below income averaged over a longer period. Several authors have explored the ways in which the permanent income of families can be estimated on the basis of a variety of observable variables (e.g., P. Musgrove and R. Ferber, "Identifying the Urban Poor", cited in footnote 20 below). This general procedure is perhaps the most promising way of dealing with the problems associated with fluctuating income over time, but it is analysis-intensive and therefore probably not feasible as a standard procedure to be applied to all the surveys one might wish to employ in a study of trends in a given country. Some simpler rules of thumb to guide the analyst would therefore be desirable, e.g., compilation of how distribution varies by period of observation.

[8]Attention was drawn by Paglin to the need for consideration of life cycle phenomena in the analysis of distribution trends in the U.S. (See Morton Paglin, "The Measurement and Trend of Inequality: A Basic Revision", *American Economic Review*, September 1975, pp. 598-609.

[9]One could also argue that the extent of fluctuations should be reflected and/or that some lifetime income trajectories are systematically preferred to others. For most people disposable income fluctuations are a disadvantage, as reflected in the widespread purchase of insurance to avoid them. There is presumably also some preference both for a rising income over most of one's life and for an income pattern which is reasonably close to the time profile of desired expenditures. Allowance for these considerations would probably widen the material welfare gap between the wage or salaried middle class with regular income, little unemployment after a stable job is once secured, and an upward trend of income during all or nearly all of the working career, and the lower income groups including independent farmers, daily wage labourers, etc. The highest income groups also undergo wide income fluctuations but here the welfare costs are less clear.

inclusion of this factor will lead to a higher estimate of inequality.[10] Apart from the length of life question, though, lifetime income (or consumption) is likely to be less unequally distributed than point of time income (or consumption) distribution because neither inequality due to normal age-income relationships nor inequality resulting from transitory income (or consumption) fluctuations is present.[11] Due to the practical problem of information, few studies have considered distribution of income over periods longer than a year, so there is less evidence on the basis of which to guess the extent of the needed adjustment. Most of the cross section information on age-income relationships refers to earners or to families (by age of household head),[12] but not to persons and their per capita family income. And virtually all of the evidence with which we are familiar on the over time age-income relationships refers to earners. The adjustment needed to move from a current to a lifetime distribution might or might not outweigh the effects of different length of lifetime, depending on how the latter is taken into account.[13] An adequate treatment would require some assumption about the trade-off in people's preference between high income and longer life; for the very poor a longer life might be no boon.

Use of evidence on income, private consumption, public consumption and wealth provides in principle much understanding of inequality and its level. But there may be better ways to get at certain aspects of it, and a number of suggestions for alternative measures of both welfare levels and the degree of inequality have

[10]See the discussion in Morrisson, pp. 249-254. (Christian Morrisson, "Income Distribution in Less Developed Countries: Methodological Problems", in *Personal Income Distribution*, International Economic Association, 1978). He notes that in France life expectancy at 30 for unskilled workers is 34 years while that of executives is 40 years. This would imply a difference of somewhat more than 6 years for total life expectancy. Comparisons across occupational groups or regions in L.D.C.'s frequently seem to indicate differentials of 10-15 years (e.g. rural areas of Algeria vs. Algiers or poorer districts of Argentina vs. Buenos Aires). The most extreme differentials occur between small white minorities and low income majorities, as in the former Rhodesia (70 vs. 37.5 years). Probably a gap of 15-20 years in life expectancy across the income scale is not infrequent in L.D.C.'s. In middle income L.D.C.'s the expectancy of the highest income groups is likely to approach developed country averages (e.g. 70 years), while that of the lower income groups lies below the national average (itself usually 55-65 years). See World Bank, *World Development Report, 1979*, pp. 166-169.

[11]It may be surmised that lifetime distribution would be less equal than an average of the observed age specific distributions, due to change over time in the relative ranking of different individuals. Such changes of rank are sure to happen, not only because of changing relative incomes for earners, but also because of shifts between the earner category and the non-earner category, and changes in the income position of the family to which one is attached (for non-earners), e.g. when a non-participating woman marries and moves from her parents' household to her husband's.

[12]A greater share of inequality at a point of time appears to be associated with the age of household head in developed countries than in L.D.C.'s. Morrisson (*op. cit.*, pp. 244-249) reports that household income differences across age cohorts (defined by age of household head) account for only 1-3 percent of the total value of the Theil coefficient, as compared to 11 per cent for the U.S. (partly because total inequality is less in the U.S., of course) But this difference may be offset by a closer association in L.D.C.'s than in developed countries between age of head and income on the one hand and family size on the other.

[13]Estimates of the impact of allowing for life expectancy have been made by Morrisson for four L.D.C.'s, by estimating average life expectancy of different income cohorts and summing lifetime income; the Gini coefficients rose by 0.046 to 0.059 (or by 3.8 to 8.2 percent of their bases). These calculations implicitly assumed no change in the income ranking of persons over the life span, so in this sense they tend to provide an upper limit estimate of the increase in inequality as one shifts to this sort of over lifetime calculation.

been made. Life expectancy, level of education, adequacy of nutrition, availability of good water, electricity, and the like, and quality of shelter are among these. Were the assumptions which make data on income, wealth or private consumption distribution most informative—i.e. the same price vectors across persons and over time and rational decision making by buyers—to hold, then one could accept with little reservation the traditional argument that as far as items purchased privately are concerned (food, shelter, etc.) there is no need for direct measurements since the total income or total consumption figure is always a more meaningful indicator than any component of it. To the extent that these assumptions are not valid, a matter for empirical enquiry, measurement of individual items can help to draw a fuller picture of welfare distribution.[14] There are also practical reasons to consider such specific indicators as housing conditions, education and consumption of certain foods to be relevant and helpful. Some may be easier to measure accurately than income or total consumption; some may reflect public expenditures and thus not be reflected in private income or consumption figures (e.g., education). Also, since most individual measures of inequality are still quite deficient relative to any standard of perfection, it is important to have as much information as possible to enrich the detail of our picture of distribution and poverty and to provide internal consistency checks. With respect to the benefits to individuals of public consumption expenditures, the same argument would in principle be made as for private consumption, i.e., if the government allocated its expenditures in an efficient way,[15] knowing the total dollar benefits accruing to each individual would obviate any need to know those from education, health services or any other individual item. But here there is no easily attainable information on the distribution of benefits from public expenditures,[16] nor is the assumption of rational allocation of public expenditures so plausible as the parallel assumption for private expenditures. These facts add interest and merit to the use of other measures of welfare and its distribution than those items captured in the measures of income and private consumption. Perhaps the most developed attempt at a more direct measure of welfare than income or consumption expenditures is Morris' measure of the physical quality of life.[17] Many measures are more appropriate in the assessment of progress in poor countries than in middle or high income ones; life expectancy, nutrition, and levels of education discriminate better among L.D.C.'s than among developed countries, for example. When different indicators of different aspects of welfare are available the problem of how to summarize or integrate the diverse pieces

[14] They may help to avoid misleading interpretations of income or consumption figures which result from price differences across families (e.g. in the cost of food or shelter), misallocation of resources within the family (e.g. the drinking father) and the like.

[15] I.e., in such a way as to meet all the standard marginal conditions for utility maximization. For example, public expenditures on two different services would have to have the same marginal benefits per dollar spent for each individual.

[16] Recent studies by Selowsky in Colombia and by Meerman in Malaysia have broken new ground in the measurement and allocation of such benefits (Marcelo Selowsky, *Who Benefits from Government Expenditures? A Case Study of Colombia,* New York, Oxford University Press, 1979; Jacob Meerman, *Public Expenditure in Malaysia: Who Benefits and Why,* New York, Oxford University Press, 1979).

[17] Morris P. Morris, *Measuring the Conditions of the World's Poor: The Physical Quality of Life Index,* New York, Pergamon Press, 1979.

of information arises. It has been addressed recently by Atkinson and Bourguignon.[18]

Whatever the economic variable used to measure economic welfare and whatever the definition of the consumption or income receiving unit, the original information on inequality is the share of various quantiles. Much discussion surrounds the most useful or appropriate ways to summarize this information into a single number which reflects the level of inequality. The longest history is attached to the Gini coefficient, but other measures such as the Theil coefficient, the mean logarithmic deviation, the Kuznets index, and Atkinson's indices have increasingly come into use. The choice among these depends on the aspects of distribution in which one is most interested or the social welfare function one accepts. But it is not an issue of basic measurement of inequality but of how information may best be summarized into a convenient usable form. Since our concern here is with measurement *per se* we do not address that issue. We restrict the discussion to aggregate value measures like income and consumption as opposed to such indicators as life expectancy, level of education, etc. on the grounds that, at least for the time being, the former will continue to be used in the bulk of the effort to measure trends over time. We presume that, even if one opts not to adopt a full life cycle measure of economic welfare, one at least wishes to reasonably approximate permanent income or consumption, i.e. to remove short term fluctuations of income.

3. A Desirable Set of Information to Assess Trends in Inequality

We turn now to the major practical issues in estimating trends in economic inequality.

Without exception the data base in L.D.C.'s is so far from being adequate that each source must be carefully appraised and much attention must be given to the use of complementary pieces of information and cross-checks.

The above notwithstanding, most developing countries have had one or more household surveys which provide information on income and/or consumption of households as well as on the size and composition of the household. Where such surveys cover a decade or more it is reasonable to hope that they will provide useful evidence on trends in the level of inequality, but their use is far from a simple matter of comparing decile distributions at the two points of time, let alone comparing Gini or Theil coefficients. One must first assess the coverage, quality of reporting, and comparability of the sources. Seldom will one's mind be set at rest on all counts. Chances are best when coverage is nation-wide and the surveys corresponding to different points of time are at least conceptually comparable, a good example being the Philippines, for which a *Family Income and Expenditure Survey* has been undertaken at five-year intervals since 1956–57. The indicators of the distribution of reported income among households over the first four such surveys are shown in the first two rows of Table 1. They suggest a fairly definite widening of the gap between the poorest 20 percent and the rest, although the Gini coefficient remains nearly constant. There are several reasons

[18]Atkinson, A. B. and Bourguignon, F. The Comparison of Multi-Dimensioned Distributions of Economic Status, *Review of Economic Studies*, Vol. 49, 1982.

TABLE 1

INDICATORS OF ECONOMIC INEQUALITY AMONG HOUSEHOLDS, THE PHILIPPINES, 1956-71

	1956	1961	1965	1970-71
Distribution of Income Among Families Ranked by Family Income				
Gini Coefficient	0.48	0.50	0.51	0.49
% Share of Lowest 20%	4.5	4.2	3.5	3.7
Distribution of Consumption Among Families Ranked by Family Consumption				
Gini Coefficient[a]	0.45	0.40	0.41	0.40
% Share of Lowest 20%	5.0	6.0	5.7	5.9

[a]Based on seven quantiles.

Source: From or based on the same data cited in A. Berry, Income and Consumption Distribution Trends in the Philippines, 1950-70, *Review of Income and Wealth*, June, 1978, p. 316 and p. 318.

for which this might be an invalid conclusion to draw from the data, however; among the most obvious are (i) use of family income (or consumption) rather than personal income (or consumption) as the variable whose distribution is calculated; (ii) inappropriateness of income of the period used in the survey as an indicator of economic well-being; (iii) misreporting of income and other biases in the data as a reflection of the current level of inequality; (iv) the possibility that the price indices relevant to the different income groups have risen at different rates during the period under consideration. We consider each of these matters in turn.

a) *Which distributions?*

The first two problems both involve misranking of individuals due to mismeasurement of economic welfare. The problem with ranking by family income follows from the fact that while there is a marked positive association between family income and family size, per capita income bears a marked negative association with family size (Table 2). The distribution of income among families

TABLE 2

MEDIAN FAMILY INCOME AND INCOME PER CAPITA,[a] BY FAMILY SIZE, PHILIPPINES, 1970-71

(PESOS)

	Family Size		
	1	2-5	≥6
Median family income	1,095	1,995	2,934
Median family income per family member	1,095	692	374

[a]Whereas income per capita was the variable we sought to use here, it could only be estimated indirectly and approximately, so the variables "median family income" and "median family income per family member" are presented. For all families, mean income is 3,736 and median income 2,454.

Source: *Family Income and Expenditures, 1971*, BCS Survey of Households Bulletin No. 34, Manila, 1973.

thus tends to rank persons in small families low when in fact they are relatively high in the distribution of income among persons. While the positive association between family size and family income introduces spurious inequality, use of family income also hides some real inequality (large families with low per capita income do not appear in the low income categories); in the Philippines case the offset between the two biases was almost complete in 1970-71 (Table 3). More

TABLE 3

A COMPARISON OF DISTRIBUTION OF INCOME BY FAMILIES AND BY PERSONS, PHILIPPINES, 1970-71

(PERCENTAGES)

	Family Income	Per Capita Income (Estimate)
Lowest Quintile	3.7	3.9
2nd Quintile	8.2	8.5
3rd Quintile	13.2	13.2
4th Quintile	21.0	21.1
Top Quintile	53.9	53.4
Top 10%	36.9	36.7
Top 5%	24.3	24.5

Source: The distribution of family income is the same as that presented in Table 1. That of per capita income is calculated from *FIES* 1971, by using the classification of family income by family size to convert the figures on family income to per capita income by family income-family size cells and regrouping according to per capita income. Among families of a given size and income category, some positive correlation was assumed between size and family income and the final distribution appears little sensitive to plausible variations in this assumption.

generally, there is a tendency towards such offsetting, but it is of course not always complete.[19] Further, the relationship could change over time; we have not seen evidence on the extent to which it does.

The relationship between the distribution of income among families and the distribution of consumption among families appears disconcertingly variable over time, based on the few case studies at hand.[20] Were current consumption felt to

[19]Morrisson reports that for Taiwan and Trinidad-Tobago the Gini coefficients are, as in the Philippines, almost identical. In Hong Kong, that of personal distribution is greater (0.467 to 0.417) (Christian Morrisson, Income Distribution in Less Developed Countries, in *Personal Income Distribution*, International Economics Association, 1978, p.243). For Colombia the Gini of the personal distribution (1974) was 0.536 while that for households was 0.51 (author's calculations based on data appearing in M. Selowsky, *Who Benefits from Government Expenditures?*

[20]In Lima, Musgrove and Ferber report that only 41 percent of the persons found in the first (lowest) decile of the distribution of per capita permanent household income are also found in the first decile of the distribution of per capita household consumption. For the first four deciles together, this overlap is 71 percent (Philip Musgrove and Robert Ferber, Identifying the Urban Poor: Characteristics of Poverty Households in Bogota, Medellin, and Lima, *Latin American Research Review*, Vol. XIV, No. 2, p. 31). These overlaps were somewhat greater in Bogota and Medellin, but still the lack of coincidence was notable.

be a better indicator of welfare than income,[21] conclusions about trends could be seriously altered.[22]

Whereas the distribution of income among Philippine families suggests, as indicated above, a definite widening of the gap between the poorest 20 percent and the rest over 1956–1970/71 (even though the Gini coefficient did not increase, since the top 5 percent was also suffering a decline in its income share),[23] the consumption share of the bottom quintile of families ranked by consumption rose considerably while the Gini coefficient fell from 0.45 to 0.40.

Sri Lanka data present a similar but more dramatic contrast (Table 5). While the distribution of income among families improved markedly over 1963–73, judged both by the Gini coefficient and the share of the bottom quintile, the distribution of consumption among families (here the families are ranked by income, not consumption) worsened over the same interval. The point of these comparisons is not to argue that one indicator is more appropriate than the other but to highlight the apparent sensitivity of results to which indicator is chosen. This fact points to the need for careful selection of the conceptually most relevant indicator in a given analysis, where a choice is possible, and for better understanding of when and why the indicators behave differently. Such an understanding is a necessary condition to analyse distribution trends in the majority of L.D.C.'s, where, due to data limitations, the only indicator available over time is not the conceptually most desirable one. In the Philippines case, for example, while access to the original data at the household level would permit estimation of the distribution of income or consumption among persons ranked by per capita household income or consumption, these data are probably not accessible, either because they no longer exist in the original form or because the cost of obtaining them is prohibitive, so the researcher must use published tables on the family distributions and hope to draw on robust generalizations from studies in other

[21]In most countries many low income families consume well above their current income levels, whereas the consumption to income ratio is much lower for high income families. Thus the distribution of consumption among families ranked by income may be much less unequal than the distribution of income among families ranked by income. The former tendency is reflected in the 1970–71 data of Table 4 for the Philippines (the high absolute numbers, including the overall average, reflect the much greater underreporting of income than of consumption expenditures), the latter in Gini coefficients of 0.321 and 0.490 respectively for the two distributions. (This differential is towards the upper end of those observed in a set of nine countries. See A. Berry, "Empirical Relationships Among Income and Consumption Distributions: An Aid to Analysis of Inequality is Less Developed Countries", mimeo, 1982. The range of differences went from about 0.04 to over 0.2.) The distribution of consumption among families ranked by income is normally less unequal than the distribution of income both because some families who are low in the distribution of family income are not poor in any basic sense—they may be retired (or other) families with high levels of wealth but low income, or families who underreport their income—so their consumption is naturally high compared to their (reported) income, and because another subset of low income families does suffer from low income and low wealth but is able (and forced as much as possible) to consume more than they earn; they may use up savings, borrow from others or receive gifts of some sort. Partly because families who are not poor (as reflected by adequate consumption levels) appear low in the ranking, a distribution of consumption among families ranked by income may be considerably less unequal than a distribution of consumption among families ranked by consumption. This was the case in the Philippines as of 1970–71; the Gini coefficient of the latter distribution was 0.40 contrasted to 0.32 for the former.

[22]In some cases of course one might find that all estimated distributions move in quite similar ways over time, providing some reassurance that even though there is some uncertainty as to which is superior from conceptual or empirical points of view the result is not sensitive to which one is chosen.

[23]Whether the trends of income distribution among persons would be the same is not clear.

TABLE 4

EXPENDITURE/INCOME RATIOS OF PHILIPPINE FAMILIES, BY FAMILY INCOME, 1970-71

Family Income (Pesos)	Percent of Families	Expenditure/Income
<500	5.2	4.87
500-999	12.1	2.66
1,000-1,499	12.2	1.97
1,500-1,999	11.8	1.69
2,000-2,499	9.6	1.50
2,500-2,999	8.1	1.37
3,000-3,999	12.5	1.25
4,000-4,999	7.5	1.16
5,000-5,999	5.0	1.12
6,000-7,999	6.3	1.08
8,000-9,999	3.6	1.02
10,000-14,999	3.7	0.97
15,000-19,999	1.1	0.93
≥20,000	1.3	0.60
Total	100.0	1.20

Source: *FIES*, 1971, p. 1.

TABLE 5

INDICATORS OF ECONOMIC INEQUALITY AMONG HOUSEHOLDS, SRI LANKA, 1953-73

Indicator	1953	1963	1969-70	1973
Distribution of Income Among Families Ranked by Family Income				
Gini Coefficient	0.49	0.47	0.36-0.38	0.35
% Share of Lowest 20%	5.2	4.5	7.4	7.3
Distribution of Family Consumption Among Families Ranked by Family Income				
Gini Coefficient	0.40	0.35	0.39	0.39
% Share of Lowest 20%	8.7	10.0	n.a.	8.0

Source: A. Berry, Canadian Foreign Aid and Income Distribution, Economic Council of Canada, Working Paper No. 133, July, 1968, Appendix F, p. 20 and p. 24.

countries on the relationship of trends in these distributions to trends in the personal distributions. No attempt appears yet to have been made to systemize evidence on the relationships of the trends of different distributions. Not infrequently the researcher has access to the last in a series of data sets but not the earlier ones. It may often be productive in such a situation to analyse the relationship among distributions in much detail for the last data set, since together with other information (e.g. on demographic trends such as the size distribution of families) this may give guidance on whether the more relevant distributions are likely to have evolved in the same way as the one for which data are available. At present, even the most careful use of such information would probably leave

considerable ambiguity as to how inequality has moved over time in the Philippines, but it would undoubtedly reduce the level of such uncertainty.[24]

b) *Quality of Data*

Quality of the data in the available surveys is as important as the conceptual appropriateness of the distributions available; changes over time in coverage and accuracy can easily lead to wrong interpretations of the course of events. The income levels reported in household surveys are usually well below those implicit in national accounts estimates of personal income.[25] The ratio of the former to the latter falls as low as a third in some cases and reaches 75 or 80 percent in others, with many cases falling in the range 60–75 percent.[26] Since national accounts figures are usually more likely to understate[27] true income than to overstate it, the underreporting of the surveys is usually even greater than indicated by the comparison with the national accounts.[28]

Incomplete income reporting in household surveys involves both deliberate underreporting and/or problems of recall and failure to include certain types of

[24] As of 1971 the distribution of income among persons in the Philippines was quite similar to that among households, even though the ranking of individuals was rather different. With respect to consumption, a similar result might be surmised though very little evidence is available. But such evidence is not sufficient for the purpose of assessing trends over time. The question in that context is whether one distribution is likely to move differently from the other; a rather modest change in the relationship between the distribution of income among persons and that among families, for example, could mean that inequality was increasing according to one indicator and decreasing according to the other.

[25] The comparison involves blowing up the sample figures to a national aggregate, i.e. multiplying total income reported in the sample by the inverse of the share of all households which were included in the sample.

[26] In the four Philippine surveys reported in Table 1, the ratio ranged between 0.65 and 0.71 (A. Berry, Income and Consuption Distribution..., p. 315). For various Colombian sources the range is from about 0.45 in the population census of 1973 to 0.73 in DANE's 1970 household survey (see A. Berry, "Recent Trends in the Distribution on Income in Colombia: Possible Factors", mimeo, 1978, p. 22). Webb and Pfefferman cite ratios of 73 percent for Mexico (1963), 66 percent and 73 percent for Korea (1965 and 1976 respectively), and 60 percent for Turkey (non-agricultural income), noting that in all these cases the data are from expenditure surveys and that in employment and other non-expenditure surveys the ratios are usually lower still (G. Pfefferman and R. Webb, *The Distribution of Income in Brazil*, World Bank Staff Working Paper No. 356, 1979).

[27] National accounts methodology differs across countries, but in most it is based on production data of one sort or another (rather than income data). Data are naturally more precise for some sectors than for others. In some sectors the estimation methodology may be unbiased, but in others it is likely to be downward biased due to a tendency to assume the reporting system covers all value added in the sector. The production generated in second jobs and the like may often escape the reporting system *and* the estimation process.

[28] This is more true given that capital gains which should be included as part of income in the analysis of distribution are not included in the national accounts concept of income. While capital gains are not usually high relative to conventional national income (e.g. less than 10 percent), the potential effects on income distribution of capital gains and losses together may be significant. If net capital gains were 8 percent as large as conventional national income and gross capital losses were 4 percent (so that gross capital gains were 12 percent of it) then gross gains and losses would total a healthy 16 percent of conventional national income. For the U.S. over the period 1948–64 Bhatia estimated that (net) accrued capital gains averaged about 12 percent as much as reported personal income (K. B. Bhatia, Capital Gains and the Aggregate Consumption Function, *American Economic Review*, Vol. XII, No. 5, Dec. 1972, p. 869). In developing countries the physical capital share is usually higher than in developed countries, though the assets may be less prone to appreciate.

income due to misunderstanding by the respondent or oversight in the survey.[29] In most cases underreporting of labour incomes is moderate, at least relative to national accounts figures; capital incomes, however, are usually seriously understated even in developed countries and, one would guess, more so in developing ones.[30] Components like imputed rent on owner occupied dwellings are not normally included in principle and thus go unrecorded. Self employment income (a combination of income from labour and from capital) is likely to suffer a degree of underreporting between that of earnings from labour and capital income. Often an accurate estimate of such income is difficult to make even with the best of will. In a detailed discussion of Latin American surveys and population censuses which provide income data, Altimir reports that they "give significantly lower estimates of entrepreneurial income than the national accounts. In only one-third of the cases is the minimum discrepancy between 15 and 20 percent: in most it is as high as 30-50 percent."[31]

Where comparable methodology has been used over a series of surveys one might expect similar levels of underreporting, so that at least the trends in income distribution would not be in error. Unfortunately comparability in practice is hard to test for; it involves not only the set of questions applied but the way in which they are asked, follow-up discussions, etc. In the Philippines case, although each family income and expenditure survey has had the same format, the survey based consumption estimate rose from two-thirds or less of the national accounts figure in 1956 to 86-87 percent in 1970/81.[32] Usually the surveys are less comparable over time than those of the Philippines and the difficulties accordingly greater. The available Colombian sources (1965-78) show a wide range of coverage relative to the national accounts (Table 6). Comparing distributions between the 1970 survey with fairly good coverage (71 percent of the national accounts figure) and that of 1978 (coverage of 45 percent or so) obviously involves serious risks. One might expect a downward misreporting bias in the Gini coefficient to be related to the degree of underreporting, on the assumption that relative underreporting of the high income categories is positively associated with total underreporting.[33] But the relationship between overall underreporting and that of various income groups and various types of income is not at all clear. As overall reporting completeness falls from the upper end of the normally observed

[29]Some respondents, by some combination of their own inclination and lack of clarity from the questioner, think of income as including only money income; others think of it as including only earnings from labour. Where attention is drawn to the other forms of income, reporting is likely to be higher but still far from complete. A further source of underreporting in most surveys lies in the fact that the chance that they will include any of the very top income families is very small.

[30]Sawyer (M. Sawyer, Income Distribution in OECD Countries, *OECD Economic Outlook, Occasional Studies*, July 1976) has estimated the completeness of reporting of income by source in household surveys in OECD countries. The ratio of income reported in surveys to that estimated in the national accounts ranges from a little over 80 percent to almost complete for wages and salaries, from a third to almost complete for entrepreneurial income though usually over half, and from about a third to 87 percent (U.K.) for property income with most figures about or below 50 percent.

[31]Oscar Altimir, Income Distribution Estimates from Household Surveys and Population Censuses in Latin America: An Assessment of Reliability, World Bank, Development Research Center, mimeo, 1977, p. 69.

[32]A. Berry, Income and Consumption Distribution Trends, p. 315.

[33]In that case, with severe underreporting the Gini would be significantly downward biased.

TABLE 6

INCOME COVERAGE OF HOUSEHOLD SURVEYS AND CENSUSES RELATIVE TO NATIONAL ACCOUNTS, COLOMBIA, VARIOUS YEARS

Source	Year	Ratio of Survey or Census Income Estimate to National Accounts Income[a]		
		All Income	Labour Income	Income from Capital and Self Employment
Ministry of Health —ASCOFAME	1965-66	0.554		
DANE, Survey				
EH1	1970	0.676	0.911	0.455
EH4	1971	0.560	0.696	0.418
EH5	1971	0.526	0.689	0.428
Population Census[b]	1973	≈0.45	0.608-0.667	0.33-0.39
Selowsky Survey	1974	0.596		
DANE, EH19	1978	0.450	0.542	0.364

[a]National accounts income is defined here as the sum of remuneration to labour and income of unincorporated enterprises, to facilitate its division into income of paid labour and other income.

[b]The 1973 figures are particularly open to question because for many members of the employed labour force (8.9 percent), job position was not reported, and among those for whom it was, an unusually high share (70 percent) report that they were paid employees. Hence the wide range in the estimates of coverage by earnings category.

Source: A. Berry, "Recent Trends", p. 22 and p. 24. For some sources the cited study gives a range rather than a single estimate of the relative coverage of the two sources; here the mid-point of that range is used, except for the 1973 census where a best estimate was provided along with the range. That best estimate is used here.

share of national accounts income, say 80 percent, to 40-50 percent, it may be that coverage falls no more for capital than for labour income.[34] The Colombian results presented in Table 6 suggest, in fact, greater variance in the reporting of labour than of capital income, but no evident relationship between overall coverage and the Gini coefficient. Across countries it does appear that the completeness of reporting of capital income varies more than that of labour income, and the Colombian figures of Table 6 notwithstanding, this seems the most plausible expectation across surveys in a given country. But the data of Table 6 make it clear that one cannot take such a relationship for granted; much research will have to be dedicated to this issue before any generalization will be possible. In practice most analysts have opted to assume an income elasticity of underreporting which is above zero but the same across different surveys in a given country. Any specific assumption is open to question, though sensitivity

[34]It seems likely that when overall reporting is high much of the discrepancy between the survey and national accounts figures must be associated with underreporting of capital income. This proposition is supported in the Colombian data by the fact that as between CEDE's 1967-68 budget surveys, where income reporting was unusually accurate, and the 1970 DANE survey with good coverage ratio of nearly 70 percent, the relative income of the self employed, including employers, was much higher for the former, at least for Bogota. The ratio was never as high in DANE surveys of the 1970's as in the CEDE survey.

analysis can at least test the robustness of any conclusion with respect to the pattern of relative underreporting.[35]

Surveys which present data on consumption expenditures as well as on income provide both the raw material for a probably more interesting distribution and a source against which to compare the income figures. The relationship between total reported consumption and total reported income gives clues on the extent of income underreporting. Reporting of consumption is usually (perhaps always) more complete than that of income, partly because of the necessarily detailed nature of the questioning, and probably also because of the lesser felt need to hide information. The importance of cross-checking survey income and consumption data with those of the national accounts cannot be over-emphasized, though the weakness of the accounts must be borne in mind as well. Where accurate surveys are difficult to carry out the national accounts are likely to face somewhat parallel, albeit generally less serious problems, so it is important to look into national accounts methodology to form some idea of their possible biases. In some cases they are not methodologically independent of the household surveys.[36] Their main weakness usually involves the less modern sectors of the economy. Although national accounts may misestimate true values considerably, the methodology is normally consistent over time so the degree of bias should not be subject to sharp changes. Hence the ratio of survey income to national accounts income should reflect changes, especially abrupt ones, in the adequacy of reporting of the survey.

c) *Income Specific Cost of Living Indices*

Another requirement for satisfactory analysis of inequality trends is a set of income-level-specific price series. Although few countries publish more than two cost of living series (e.g. for middle income and low income urban groups) the raw materials for construction of additional series do exist in countries where consumer surveys reveal the consumption baskets of various income groups and price series are available for specific items. In some countries more or less systematic differences in price increases across income levels are large enough in the course of a decade, say, to significantly affect the measured trends in income or consumption distribution. In other cases, abrupt changes in relative food prices occur; such changes may be most likely when sharp changes in the real exchange rate take place, or when inflation accelerates, as was the case for example in Colombia in the early 1970's. During the 1960's when the rate of inflation was moderate the cost of living rose by almost the same percent for low, middle and high income groups.[37] But in the early 1970's (1971-75) when

[35]Such a test has been used by various authors, e.g. Joel Bergsman, *Income Distribution and Poverty in Mexico*, World Bank Staff Working Paper No. 395, June, 1980.

[36]Estimates for private consumption, for example, usually make use of household survey data on consumption to income ratios. Estimates of output in some informal sectors may be based in part on household survey income information. In many countries the national accounts estimate private consumption as the residual between G.N.P., investment and government expenditures, so that errors in the estimation of those aggregates are fed into that of consumption. This also must be taken into account.

[37]A. Berry and M. Urrutia, *Income Distribution in Colombia*, New Haven, Yale Press, 1976, p. 122.

food prices led the increases, the cost of living index rose by up to 25–30 percent faster for the poorest groups than the richest.[38] Changes in relative prices between rural and urban areas also warrant monitoring as they may produce spurious trends in distribution.

4. Indirect Evidence on Trends in Distribution

a) *Composition of Consumption*

While there is no substitute in the study of distribution trends for a set of comparable household surveys, such a resource is so rare that all possibly relevant evidence of other types must normally be drawn on. One useful consistency check on estimated trends in the per capita income of quantiles of the distribution is an analysis of how the composition of consumption expenditures has changed by quantile. Rising income should produce a shift away from necessities towards luxuries and vice versa.[39,40]

In most developing countries, the data available to analyse distribution trends are less complete and satisfactory than those for the Philippines (and in many cases even than those for Colombia) and the difficulties in drawing robust conclusions correspondingly greater. Often survey data are for urban areas only or for the capital city only and/or for the modern sector only; usually they do not include consumption expenditures and usually there are serious problems of comparability over time. With much attention to these deficiencies, and drawing on evidence from other countries (e.g. on the relationship between urban distribution and overall distribution), it may still be possible to draw tentative conclusions.

b) *Wage Trends and Other Labour Market Data*

The major data apart from household surveys which can contribute to the analysis of income trends of economic groups are those on wage rates and participation rates. These have the advantage, usually, of methodological consistency over time. When apparent inconsistencies arise between a wage series and the income trend of those quantiles earning wages in the industries in question, search for the source of the inconsistency often yields evidence of value in interpreting the distributional data. This process is greatly aided if the household surveys provide income data by occupational groups, as in the case of the 1965 and 1970–71 Philippines surveys, whose data are compared with official wage series in Table 7. While the two sources are not directly comparable in that the

[38]A. Berry, The Effects of Inflation on Income Distribution in Colombia: Some Hypotheses and a Framework for Analysis, in A. Berry and R. Soligo (eds.), *Economic Policy and Income Distribution in Colombia*, Boulder, Colorado, Westview Press, 1980.

[39]Other factors, such as changes in relative prices, could of course affect the trends in consumption composition and such possibilities would have to be looked into in order to assess exactly the implications of observed trends.

[40]Consumption composition data are revealing in other ways. One evidence, for example, of the misranking of families in the family income distribution is the tendency for many of the low income families to have consumption bundles suggestive of higher welfare levels than those of higher ranked families. This is consistent with the considerably different ranking of families as between the distributions of consumption and of income per person, noted above.

TABLE 7

INCOME TRENDS BY OCCUPATIONAL CATEGORIES, PHILIPPINES, 1965 TO 1971: FIES FAMILY INCOME DATA AND SELECTED WAGE SERIES FROM OTHER SOURCES

Occupational Category	Percent Increase, 1965 to 1970–71: Income	Consumption
1. Agricultural labour		
a) Daily wage	−12.2	
b) Annual income, families whose household head is an agricultural labourer	24.2	11.3
2. Manufacturing labour		
a) Manufacturing wage earners, monthly earnings	1.7	
b) Annual income, families whose household head is craftsman, factory operative, etc.	1.3	11.7
3. Transport workers		
a) Wage earners, monthly earnings	−12.0	
b) Annual income, families whose household head is transport worker	7.6	2.7
4. Commerce		
a) Wage earners, monthly earnings	−10.8	
b) Annual income, families whose household head is sales worker	0	10.1
5. White collar		
a) Salaried employees, monthly earnings	−10.9	
b) Annual income, families whose household head is:		
i) professional, technical, etc.	−28.5	−18.7
ii) administrative, executive and management	−39.2	−26.0
iii) clerical	7.3	20.2
All three—	−18.2	
Families weighted by relative numbers of wage workers in the 1971 BCS Labour Force Survey	−14.4	

Source: A. Berry, Income and Consumption Distribution in the Philippines, p. 330.

wage data refer to individual earners and the family survey data to family income (with families classified by occupation of household head), similar trends would nevertheless seem probable in most cases. The apparent inconsistency in the case of agricultural labourers is particularly fascinating, since this is a large low-income group. Several hypotheses besides the obvious one—errors of observation in one or both sources of information—come to mind. The wage figure is daily; if days worked per year increased, annual income could rise though the daily wage was falling. Or an increase in income from other sources than agricultural wage income could explain the discrepancy;[41] the increasing importance of non-agricultural income for farm families as the process of development proceeds has been noted in many countries, e.g. the U.S., Japan, Korea. In a case like the Philippines, access to detailed analysis of the household survey data would permit

[41]According to *FIES*, 1971, while 14.0 percent of rural families had agricultural wages as their main source of income (p. 12), 28.9 percent of families received some agricultural wages (p. 13). For 6.5 percent of families non-agricultural wages were a secondary source of income, for 9.3 percent trading was, for 8.4 percent manufacturing was, for 26.8 percent farming, for 63.7 percent fishing, forestry, and hunting, for 37.2 percent production of articles for own use. While the data do not permit one to sort out which of these secondary incomes were important for families whose main source was agricultural wages, they suggest the range of possibilities.

the testing of such hypotheses and would very probably resolve the question of their consistency or not with the wage series data. Until such analysis is carried out, any conclusion about trends in Philippines income distribution would be premature. In Colombia's rural areas a similar issue arises. Whereas the household surveys of 1970 and 1978 would, if taken at face value, imply a near constancy of rural distribution (Ginis of 0.421 in 1970 and 0.43 to 0.45 in 1978),[42] this is hard to square with the failure of real wages to rise nearly as fast as average agricultural output per capita over this period. The household survey data on incomes of rural wage earners support the wage series, and imply that increasing underreporting of non-wage incomes is the reason why an apparent worsening of rural distribution did not show up in the observed Gini coefficients.

The importance of studying trends in earnings of various groups over time is strengthened by the increasing importance of labour earnings as a share of national income over the process of development[43] and especially of the increasing share of national income going to what we may call human capital. In developed countries, where the labour share reaches 70–80 percent of G.N.P., it is evident that the distribution of labour earnings is the major determinant of the shape of much of the distribution of income, and that much can be learned about changes if one knows how the occupational composition of the labour force changes and how wage differentials change. During some stages of development the participation of women changes significantly, so participation rates must be analysed, as must changes in family structure.

As less developed countries become more developed and more urbanized, the family unit tends to become smaller; parents are less likely to live with children, divorces and separations become more common, and so on. These confuse the analysis of distributional trends. The assumption of equal distribution of family income or consumption among family members (or among adult-equivalents), while not expected to be really accurate, would probably not lead to serious errors in the assessment of trends in distribution as long as the composition of family units does not vary over time and the intra family distribution does not change. When the composition of family units does change, one's estimate of distribution may change even though no real change occurred (or vice versa). Independent of this problem, family composition changes may lead to real changes in welfare distribution which our available measures of welfare distribution cannot pick up, because of lack of knowledge of and recognition in our measures of how family size and composition affect welfare of members, via economies and diseconomies of scale in consumption for example.[44]

[42]The former figure comes from DANE, *Boletin Mensual de Estadistica* #237, April, 1971, p. 71. The latter is calculated by the author on the basis of the rural distribution of income presented in DANE, *Boletin Mensual de Estadistica*, #332, Marzo, 1979, p. 33.

[43]This relationship is discussed by J. Lecaillon and D. Germidis, Economic Development and the Wage Share in National Income, *International Labour Review*, Vol. 3, No. 5, May, 1975.

[44]A different problem related to family composition is that of the treatment of live-in domestic servants. In some household surveys they are treated as members of the families with which they reside and would therefore be assigned the same per capita income as other members when family income is converted to personal income. This group is large (over 10 percent of the labour force in some cities) and low income but living with relatively high income families, so such a procedure would constitute a serious downward bias in the estimate of income inequality. Some adjustment would be needed to take account of it.

c. *Production Aggregates*

In most countries one of the major problems (often the major one) in tracing distributional trends involves the agricultural sector. In the extreme case there is virtually no distribution information on it (e.g. no household survey undertaken there). In other cases the agricultural (or rural) data are less precise, and it is difficult but important to assess the meaning of a large gap between the reported average income in rural and in urban areas.

A major component of the inequality in many countries, especially African ones, is believed to be that between urban and rural areas, so much attention has focussed on the trends in aggregate income or consumption in rural and urban areas, estimated on the basis of the corresponding production aggregates. Sometimes the income of rich elites is also assessed primarily from the production side.

5. Conclusions

The central argument of the above discussion is that available data permit no easy tests for trends in the level of economic inequality in less developed countries. The obvious source—a series of household surveys with data on income and consumption—is fraught with many problems and pitfalls in the typical L.D.C. and is unlikely by itself to provide reliable conclusions. Uncertainty can, indeed, be greatly diminished by careful study of how the surveys were undertaken, their completeness of reporting, and so on. Further assistance could in principle come from systemized evidence for other countries on the likely effects of certain deficiencies or more generally, characteristics, of a sample on the results which emerge from it; but no such compilation is available as yet in any accessible form. At present, the best bet is to complement evidence on income distribution with available data on consumption distribution and on wage trends and production aggregates by occupational and sectoral groups. Inconsistencies will frequently raise red flags with respect to the household survey data; reconciliation of such inconsistencies should provide a firmer foundation on which to draw conclusions about distributional trends.