The Measurement of Welfare

Theory and Practical Guidelines

Angus Deaton
LSMS Working Paper Series

No. 1. Living Standards Surveys in Developing Countries.
No. 2. Poverty and Living Standards in Asia: An Overview of the Main Results and Lessons of Selected Household Surveys.
No. 3. Measuring Levels of Living in Latin America: An Overview of Main Problems.
No. 5. Conducting Surveys in Developing Countries: Practical Problems and Experience in Brazil, Malaysia, and the Philippines.
No. 6. Household Survey Experience in Africa.
No. 9. Income and Expenditure Surveys in Developing Countries: Sample Design and Execution.
No. 10. Reflections on the LSMS Group Meeting.
No. 11. Three Essays on a Sri Lanka Household Survey.
No. 13. Nutrition and Health Status Indicators: Suggestions for Surveys of the Standard of Living in Developing Countries.
No. 15. Measuring Health as a Component of Living Standards.
No. 16. Procedures for Collecting and Analyzing Mortality Data in LSMS.
No. 18. Time Use Data and the Living Standards Measurement Study.
No. 20. Statistical Experimentation for Household Surveys: Two Case Studies of Hong Kong.
The Measurement of Welfare

Theory and Practical Guidelines
The Living Standards Measurement Study

The Living Standards Measurement Study (LSMS) was established by the World Bank in 1980 to explore ways of improving the type and quality of household data collected by Third World statistical offices. Its goal is to foster increased use of household data as a basis for policy decision making. Specifically, the LSMS is working to develop new methods to monitor progress in raising levels of living, to identify the consequences for households of past and proposed government policies, and to improve communications between survey statisticians, analysts, and policy makers.

The LSMS Working Paper series was started to disseminate intermediate products from the LSMS. Publications in the series include critical surveys covering different aspects of the LSMS data collection program and reports on improved methodologies for using Living Standards Survey (LSS) data. Future publications will recommend specific survey, questionnaire and data processing designs, and demonstrate the breadth of policy analysis that can be carried out using LSS data.
The Measurement of Welfare

Theory and Practical Guidelines

Angus Deaton

The World Bank
Washington, D.C., U.S.A.
This is a working document published informally by the World Bank. To present the results of research with the least possible delay, the typescript has not been prepared in accordance with the procedures appropriate to formal printed texts, and the World Bank accepts no responsibility for errors. The publication is supplied at a token charge to defray part of the cost of manufacture and distribution.

The World Bank does not accept responsibility for the views expressed herein, which are those of the authors and should not be attributed to the World Bank or to its affiliated organizations. The findings, interpretations, and conclusions are the results of research supported by the Bank; they do not necessarily represent official policy of the Bank. The designations employed, the presentation of material, and any maps used in this document are solely for the convenience of the reader and do not imply the expression of any opinion whatsoever on the part of the World Bank or its affiliates concerning the legal status of any country, territory, city, area, or of its authorities, or concerning the delimitation of its boundaries, or national affiliation.

The most recent World Bank publications are described in the annual spring and fall lists; the continuing research program is described in the annual Abstracts of Current Studies. The latest edition of each is available free of charge from the Publications Sales Unit, Department T, The World Bank, 1818 H Street, N.W., Washington, D.C. 20433, U.S.A., or from the European Office of the Bank, 66 avenue d’Iéna, 75116 Paris, France.

When this paper was first published Angus Deaton was professor of Econometrics at the University of Bristol, England.

Library of Congress Cataloging in Publication Data

Deaton, Angus.
The measurement of welfare.

(LSMS working paper, ISSN 0253-4517 ; no. 7)
"October 1980."
Bibliography: p.
1. Income. 2. Cost and standard of living. 3. Quality of life. 4. Welfare economics. I. Title. II. Series.
HC79.15D4 1985 339.2'2 85-12021
THE MEASUREMENT OF WELFARE: THEORY AND PRACTICAL GUIDELINES

TABLE OF CONTENTS

I. INTRODUCTION 1-2

II. THEORY AND MEASUREMENT OF INDIVIDUAL WELFARE 2-50

III. COMPARISONS OF WELFARE BETWEEN INDIVIDUALS 51-67

IV. FROM INDIVIDUAL TO AGGREGATE WELFARE 67-78

List of Works Cited 79-82
I. INTRODUCTION

1. In everyday discussion of practical welfare measurement the central concepts are income, real income and per capita real income, with occasional reference to a measure of inequality. These measures are seen as essentially objective statistical constructs and, as such, are used not only for individual welfare measurements, but also for comparisons between individuals as well as for combining or averaging welfare measures for groups of individuals. In apparent contrast, the economic approach to welfare works with such concepts as preferences, utility and social welfare functions and many economists take the view that, in these terms, even individual welfare measurement is arbitrary while welfare comparisons across individuals or combinations of welfare measures are completely unjustified. Not only does this view deprive economics of a major field of research, but it also deprives the statistical approach of both precision and methodological coherence.

2. The philosophy of this paper, in line with much of the most recent literature, is that, properly considered, the economic and statistical approaches are essentially identical. I shall argue that the measurement of income and its subsequent "correction" in various ways (real income, per capita income, etc.) is a procedure which is well-founded in economic theory, which is practical, and which seems to correspond well with what governments and laymen mean when they discuss welfare issues. In what follows, I shall almost always begin with the economics, from preferences and utility, and work towards the practical implications in terms of measurement of income, real income, or whatever. This is almost entirely a matter of style, and it would be just as feasible to start from the measurement and work back to its theoretical foundations. What is
important is that theory and measurement should match so that each can suggest developments in the other. Hence, the paper attempts not only to provide a systematic basis for measurement, but also to suggest areas in which further research is necessary, whether in collecting new data or in improving the theory.

II. THEORY AND MEASUREMENT OF INDIVIDUAL WELFARE

A. Preferences and Utility

3. The consumption of goods is the basis for economic welfare. This is true whether or not goods are provided through the market and since markets do not operate everywhere or for all goods anywhere, a non-market definition of welfare is the more basic. Since there are many goods and we want a single (or at least a small dimensional) measure of welfare, there is an immediate index number problem. One solution might be to take some standard or reference bundle of commodities (say a subsistence bundle determined by nutritional requirements) and measure welfare as a scalar multiple of this. This might work in a very poor society of consumers with very similar tastes, but if the composition of the actually chosen bundle varies strongly with income or tastes, the standard is not a very relevant one. Nevertheless, some such methodology is frequently relevant in measuring welfare for individuals close to poverty.

4. It is however, more satisfactory, to base a measure on the consumer's own tastes. Under very weak axioms of choice, we know that we can think of an individual as possessing a preference ordering which allows different bundles of goods to be ranked consistently relative to one another. With only very little loss of generality, this ordering can be represented by a utility function, e.g. $u(q)$,
for commodity vector \( q \), which has the property that if bundle \( q^1 \) is not dispreferred to bundle \( q^2 \), then \( u(q^1) \geq u(q^2) \). Clearly if \( u(q) \) does this job, so will \( f(u(q)) \) for any function \( f(\cdot) \) with a positive first derivative; hence, knowledge of preferences can tell us nothing about the choice of function \( f(\cdot) \). Since a complete knowledge of behavior in choosing among goods is equivalent to a complete knowledge of preferences, it follows that nothing in observable behavior can tell us anything about the appropriate choice of representing function. However, the preference ordering, by giving a family of scalar-valued functions, has partially solved the index number problem. To complete the solution, some particular representing function has to be chosen and it is important to realize that this choice must be based on some other criterion than observable behavior. I shall discuss a number of possible choices of units of measurement below, but whether or not these are acceptable as welfare measures is ultimately more a matter of philosophy than economics. However, the use of preferences is vital since it guarantees that whatever units are chosen the welfare measure will increase if and only if the consumer moves from a less to a more preferred position.

5. The existence of preferences allows us to draw indifference surfaces, a two dimensional example of which is shown in Figure 1. Note that these are
independent of whatever units are adopted and are essentially a graphical
representation of preferences, not utility. Any specific utility function
will attach numbers to each curve so that the problem of choosing a welfare
index can be seen as one of attaching meaningful numbers to indifference
surfaces. There are a number of different approaches. I start by discussing
those based on quantities (paragraphs 6-15), and then turn to approaches based
on prices and incomes (paragraphs 16-63).

B. Welfare Without Markets

6. Given the preference approach, we can revert to the idea of basing wel-
fare on some reference bundle of goods. Consider, for example, the point R
in Figure 1: This may or may not have some particular significance (e.g.
subsistence, "typical" choice, etc.). We can then join OR and label the four
illustrated indifference curves, or any others, by their distance from O
along OR. We can choose OR itself to be unity, so that R gets a welfare
level of 1.0, Q a welfare level of 0.5 and S one of about 1.2. This measure
of welfare I shall call quantity metric utility.

7. To formalize these ideas, we need a concept called the distance function;
this is just a convenient way of representing preferences which gives the
concepts we need. Let \( u(q) \) be any (quite arbitrarily chosen) representation
of preferences. This takes on the value \( u \) (for utility), i.e. \( u(q) = u \). Then
define the distance function \( d(u, q) \) by the implicit relationship

\[
d(u, q) = u \quad (1)
\]

Figure 2 illustrates. \( OZ \) is some quantity ray \( (q \) in \( (1) \) \) and \( u \) is the label
\( u(q) \) attaches to some quite arbitrary indifference curve (note that it is not
assumed that \( q \) generates \( u \), in general \( u(q) \neq u \)). Then \( d(u, q) \) is the amount
Figure 2
The Distance Function

by which \( q \) has to be divided to bring it on to \( u \). In the diagram \( d(u, q) \) is the ratio \( OZ/OX \). Note that, by the definition (1), \( d(u, q) = 1 \) if, and only if, \( u(q) = u \) so that, in fact, \( d(u, q) = 1 \) is just a different way of writing the utility function.

8. Quantity metric utility is then given by comparing each of the various bundles with some reference bundle, e.g. \( q^R \). For example, in Figure 1, if \( q^R \) is at \( R \) and \( q^S \) and \( q^Q \) are the other two bundles, then \( q^R \) has to be divided by 2 to bring it on to the indifference curve through \( q \) and by \( 1.2^{-1} \) to bring it on to that through \( S \), i.e. \( d[u(q^Q), q^R] = 2 \) and \( d[u(q^S), q^R] = 1.2^{-1} \).

Hence, the quantity metric utility levels are given by the reciprocals of \( d[u(q), q^R] \), i.e. 0.5 and 1.2 as before. Changing \( u(q) \) to \( f[u(q)] \) would not change the values of any of these expressions; given that the welfare level for \( q^R \) is unity, the values for the others are determined by the indifference map and not by the function chosen to represent utility. It is also true that by the definition (1), \( d[u(\lambda q^R), q^R]^{-1} = \lambda \) so that quantity metric utility is a generalization of the idea we started from, of measuring welfare in terms of multiples of some reference bundle. However, if \( q \) is not a scalar multiple
of \( q^R \), we select a \( \lambda q^R \) which is indifferent to \( q \) and use \( \lambda \) as the welfare measure.

9. In paragraph 8, (the reciprocal of) \( d(u,q) \) with \( q \) fixed was taken as a welfare measure. Alternatively, \( u \) can be held fixed and \( d(u,q) \) taken as a quantity index number (note that, by (1) \( d(u,q) \) is homogenous of degree one in \( q \)). But since \( d(u,q^1) > d(u,q^2) \) if and only if \( q^1 \) is preferred to \( q^2 \), this quantity index number is itself a perfectly good welfare measure.

Again take \( u^R \) (given by \( u^R = u(q^R) \) for example) as the reference indifference curve. Then \( d(u^R,q) \) tells us by how much \( q \) must be contracted (or expanded) in order to make it indifferent to \( q^R \), and in this context is to be interpreted as a welfare index of \( q \). In terms of Figure 1, if \( q^R \) is again the reference vector, the indices for \( Q \) and \( S \) would now be \( QQ^1/OQ \) and \( OS^1/OS \) respectively. Note that these measures are not, in general, the same as those given for \( Q \) and \( S \) by quantity metric utility in paragraph 8. Both procedures, however, measure the ratio differences between indifference curves; all that is different is the ray along which the comparison is taken. If, however, preferences are **homothetic** (i.e. so that each indifference curve is a blow-up or scaled-down version of every other one), then the indifference curves are the same distance apart along whatever ray from the origin is used. In this case, and in this case only, both measures give the same answer in all circumstances. Otherwise, we have what is sometimes called an "index-number problem". Of course, different answers are given only if different questions are asked and in specific contexts the appropriate choice of reference should always be clear.
C. Practical Evaluation

10. The exact practical evaluation of these welfare indices requires knowledge of the parameters of preferences. If, for example, preferences are those of the linear expenditure system, $u = \sum_k \log(q_k - v_k)$, then for reference vector $q^R$ and $\lambda = d\{u(q), q^R\}$, $\lambda$ can be solved (numerically) from

$$\sum_k \log \left\{ \frac{(q_k/\lambda - v_k)}{(q_k - v_k)} \right\} = 1. \quad (2)$$

However, knowledge of the parameters would usually (although not necessarily) come from estimating demand function with quantities as a function of prices and incomes, and, in this situation, there are other measures available which are probably preferable in practice. Even so, it may be possible in some non-market situations to find out enough about preferences, although solving (1) may not always be the best way of proceeding. (For further discussion, see paragraphs 59-63 below).

11. As with several of the index numbers I shall discuss, quantity-metric utility can be approximated by expressions which (a) require less information than a full knowledge of preferences and (b) bear known inequality relationships to the true index. To fix ideas, suppose that we wish to compare $q^0$ with $q^1$. Quantity metric utility of $q^1$ referenced on $q^0$ is $d\{u(q^1), q^0\}$ while that of $q^0$ referenced on $q^1$ is $d\{u(q^0), q^1\}$. (Note that the former is also the reciprocal of the quantity index for $q^0$ referenced on $u^1$ and the latter the reciprocal of the quantity index for $q^1$ referenced on $u^0$). Among the many useful properties of the distance function is that, if $q$ generates $u$, the first partial derivatives of $d(u, q)$ with respect to the $q$'s are the shadow prices of the $q$'s expressed as a fraction of total expenditure (i.e. $\sum_i \text{shadow price of } i \times q_i$). It can thus be used to generate a price/income configuration which
would cause \( u \) to be reached at \( q \) and also, by giving us the "marginal willingness to buy" for each good, tells us that aspect of preferences we most need to know. Hence, writing \( p_i^* \) (functions of \( u \) and \( q \)) for the shadow prices on \( u \) at \( q \), we have first-order approximations

\[
d(u^1, q^0) \approx d(u^1, q^1) + \frac{\Sigma p_i^*(q_i^1 - q_i^0)}{\Sigma p_i^* q_i^1}
\]

and

\[
d(u^0, q^1) \approx d(u^0, q^0) + \frac{\Sigma p_i^*(q_i^0 - q_i^1)}{\Sigma p_i^* q_i^0}
\]

Now \( d(u^1, q^1) = d(u^0, q^0) = 1 \) by (1), so that rearranging (3) gives quantity metric utility for \( q^1 \) reference \( q^0 \) as approximately

\[
\frac{\Sigma p_i^* q_i^1}{\Sigma p_i^* q_i^0}
\]

while quantity metric utility for \( q^0 \) reference \( q^1 \) is approximately

\[
\frac{\Sigma p_i^* q_i^0}{\Sigma p_i^* q_i^1}
\]

Expression (5) is the Paasche quantity index for \( q^1 \) over \( q^0 \) using the shadow prices as weights; (6) also uses shadow prices and its reciprocal is the Laspeyres quantity index, again for \( q^1 \) over \( q^0 \). Note that actual prices can be used if the consumer is in equilibrium in a market situation; otherwise shadow prices must be evaluated. This is never likely to be straightforward although information for other consumers whose preferences are revealed in the market situation can sometimes be used, see again paragraphs 59-63 below.

12. Given convexity of preferences, the distance function as defined by (1) is concave, so that the right-hand sides of (3) and (4) are always no less
than the left-hand sides so that both (5) and (6) are less than the theoretical magnitudes they approximate. Hence, the Paasche welfare/quantity index (5) is less than or equal to the quantity metric welfare index of $q^1$ over $q^0$ indexed on $q^1$, while the Laspeyres welfare/quantity index given by the reciprocal of (6) is no less than the quantity metric welfare index of $q^1$ over $q^0$ indexed on $q^0$. These inequalities mirror exactly the more familiar ones for true cost-of-living index numbers, see paragraph 27 below.

13. In practice, we can probably do better than either Paasche or Laspeyres quantity (or real consumption) indices by chaining weights in an appropriate way. However, this is not discussed here since the argument using the distance function is identical to that for the cost-function which is given in paragraph 31 below in the context of cost-of-living index numbers.

D. Assessment

14. The distance based measures are important because they give conceptually sound indexes of welfare without any assumptions on the institutional framework in which consumption takes place. All that is used are preferences and the quantities consumed; the types of constraints or limitations on behavior are irrelevant. However, since preferences are typically revealed by behavior subject to constraints (e.g. that the value of consumption not exceed some predetermined total), the measure may be difficult to evaluate without further knowledge or assumptions. Further, since markets are nearly always operative for some goods at least, it is often better to start from a market-based measure of welfare and to modify it as necessary for consumption which is not bought in
this way. As I shall argue below, this can be done in a way which is consistent with the approach just discussed.

E. Literature
15. Although labelling utility along some reference quantity ray is a standard textbook tool, the formalization of the approach through the distance function is not much appreciated. Nonetheless, the approach was fully worked out as early as 1953 in a classic paper by Malmquist; this has recently been picked up and expositions of the relationship between the Malmquist indexes and the more usual cost-of-living indexes are given in Deaton (1979) and Diewert (1980).

F. Welfare through Prices and Income
16. Consider now the case where the consumer buys everything in the market and does so at fixed parametric prices. There are no quantity constraints on purchases; as much can be bought as is desired, given the budget. The latter is a fixed money total, $x$, for total expenditure (or much more loosely, income). The question of what these quantities, prices and budget are will concern us in detail later. In particular, leisure, savings and durable goods require special treatment. For the moment, we assume that the quantity vector contains everything relevant to welfare. The consumer then behaves by maximizing $u = u(q)$ subject to the constraint $p \cdot q = x$, for market prices $p$. (The consumer is assumed to be non-satiated so that the equality will always hold).

The solution to this maximization problem is a system of $n$ equations linking demands to $x$ and $p$. Hence,

$$q_i = g_i(x, p) \quad i = 1, \ldots, n$$

(7)
constitute (Marshallian) demand functions $g_i$. Maximum attainable welfare is then given by substituting these functions into the utility function $u(q)$ giving
u = u[g(x, p)] = \gamma(x, p) \tag{8}

where \gamma(x, p) gives maximal attainable utility as a function of total outlay and prices. This is the indirect utility function.

17. To serve as an indicator of welfare, the indirect utility function, like the direct utility function above, has to be appropriately cardinalized. And once again, it helps to introduce another concept. This is the cost function, written c(u, p), which gives the minimum cost of reaching u at prices p, i.e.

\[ c(u, p) = \min_{q} \{ p \cdot q ; u(q) = u \} \tag{9} \]

Note that, for a utility maximizing consumer, c(u, p) must take the value x; if it were not so, the consumer could attain greater utility at the same outlay. Note also that by (9), c(u, p) is an increasing function of u for fixed p; it costs more to be better-off if you are already operating efficiently. Finally, note that c(u, p) contains exactly the same information as \gamma(x, p); c(u, p) = x if, and only if, \gamma(x, p) = u and one function can be obtained from the other by inversion.

18. The concepts so far can be illustrated with reference to the linear expenditure system. The direct utility function can be written in the form, \( \sum_{k} q_{k} = 1 \),

\[ u = \Pi(q_{k} - \nu_{k})^{\beta_{k}} \tag{10} \]

with demand functions (7), provided \( x \geq p \cdot \gamma \), given by,

\[ q_{i} = \nu_{i} + \frac{\beta_{i} (x - p \cdot \gamma)}{P_{i}} \tag{11} \]

so that the indirect utility function is, writing \( \beta_{o} = \Pi^{\beta_{k}} \),

\[ u = \beta_{o}(x - p \cdot \gamma) / \Pi_{p_{k}}^{\beta_{k}} \tag{12} \]
The cost-function can be solved directly from (9) and turns out to be
\[ c(u, p) = p \cdot \gamma + \beta^{-1} \cdot u^{\mathcal{R}_k} \] (13)
and we note that setting \( c(u, p) = x \) and solving for \( u \) gives (12), while
solving (12) for \( u \) yields (13).

19. The quantity \( c(u, p^R) \) for fixed reference price vector \( p^R \) is known
as money metric utility, after Samuelson (1974), and I shall adopt it here
as the basic measure of welfare in market situations. Much of the rest of
the paper is concerned with elaborations and modifications of this concept.
Essentially \( c(u, p^R) \) labels indifference curves by the minimum cost necessary
to reach them at some reference prices. For example, an individual might be
asked what is the minimum amount of money necessary to maintain him at his
actual living standards for each of the years 1970 to 1979 if (counterfactually)
prices had remained constant at their 1970 level. This measure would coincide
with actual expenditure only in 1970; in later years it is the natural measure
of real expenditure. However, more is being done than the repricing of the
actual consumption bundle; rather the actual welfare levels are being repriced
and it is this which justifies money metric utility as a measure of welfare
itself.

20. On the basis of money metric utility, ratio index numbers of welfare can
be established. Hence comparing \( u^1 \) and \( u^0 \), the relative welfare index is
\[ \frac{c(u^1, p^R)}{c(u^0, p^R)} \] (14)
which, as illustrated in Figure 3 is given by any of the ratios \( \text{OB}/\text{OA} \),
\( \text{OF}/\text{OE} \) or \( \text{OD}/\text{OC} \). In many circumstances the natural choice of \( p^R \) would be
either \( p^1 \) or \( p^0 \); for the former we have a current weighted real income index
\[
c(u^1, p^1) / c(u^0, p^1)
\] (15)
while for the latter we have the base-weighted real income index
\[
c(u^1, p^0) / c(u^0, p^0).
\] (16)

21. It is convenient to analyze money metric utility by separating it into a money measure and a price index. For any \( c(u, p^R) \) we can write
\[
c(u, p^R) = c(u, p) + c(u, p) \frac{c(u, p)}{c(u, p^R)}.
\] (17)
The first term on the right hand side of (17) is simply total expenditure for a utility maximizing consumer, while the second is a true cost-of-living index comparing \( p \) with \( p^R \) referenced on the utility level \( u \). In particular, for \( u^1 \) referenced on period 0's prices \( p^0 \)
\[
c(u^1, p^0) = x^1 + c(u^1, p^1) / c(u^1, p^0)
\] (18)
and for (the backward-looking) \( u^0 \) referenced on \( p^1 \)

\[
c(u^0, p^1) = x^0 \cdot c(u^0, p^1) / c(u^0, p^0). \tag{19}
\]

The price indices on the right-hand side of (18) and (19) are the current-weighted (or Paasche) true cost-of-living index and the base-weighted (or Laspeyres) true cost-of-living index respectively. Hence, if we can calculate the true cost-of-living indices or approximations for them, then they can be used to deflate (or reflate) money expenditures to arrive at money metric utility. We can thus approach welfare measurement through the more familiar operation of measuring prices relying on (17) to give the link with utility. Hence, in section G below, I turn to the theory and practice of price indices.

22. Although, for most of the analysis, I shall work with ratio index numbers such as price indices or real-income indices like (15) and (16), there is no reason why two welfare levels cannot be compared by the differences in costs of reaching them. For example, \( u^1 \) and \( u^0 \) might be compared by

\[
c(u^1, p^R) - c(u^0, p^R) \tag{20}
\]
rather than by the ratio (14). Note that, since \( c(u, p^R) \) is increasing in \( u \), (14) will be \( \geq 1 \) as (20) is \( \geq 0 \) so that both "indicate" welfare changes equivalently (and correctly.) The expression (20) tells us how much money the individual would need at reference prices to make the transition from \( u^0 \) to \( u^1 \). Once again, the cases of \( p^R = p^0 \) and \( p^R = p^1 \) are instructive and there exist analogues to (18) and (19). Specifically,

\[
c(u^1, p^0) = c(u^1, p^1) - \{c(u^1, p^1) - c(u^1, p^0)\}
= x^1 - EV^1 \tag{21}
\]
where EV, Hicks' (1956) equivalent variation, is the maximum amount the consumer with \( x^1 \) would pay to avoid the change from \( p^0 \) to \( p^1 \). Similarly,

\[
c(u^0, p^1) = c(u^0, p^0) + \frac{1}{2} c(u^0, p^1) - c(u^0, p^1)
\]

\[
= x^0 + CV
\]

(22)

where CV, Hicks' compensating variation, is the minimum amount the consumer at \( x^0 \) would have to be paid after a change from \( p^0 \) to \( p^1 \). Hence, ratio comparisons lead to conventional index numbers and difference comparisons to consumer surplus concepts. It should be obvious from (18), (19), (21) and (22) that, properly stated, the two sets of concepts are different ways of presenting the same information; there is no substantive difference. If we can calculate CV and EV exactly we can calculate true cost-of-living indices exactly and vice versa; the same is also true for approximations. In practice, one can choose the concept which most naturally suits the problem at hand. For example, for the evaluation of competing and mutually exclusive projects, welfare gains should be measured by differences and not by indices, since the object is to increase welfare by as much as possible rather than to select the project with the highest rate-of-return (which may be small and may exclude a larger, more beneficial project). However, for general welfare measurement, as in national accounts or other summary statistics, ratio index numbers are well-established and seem the appropriate tools.

23. I note finally a rather technical point. The concept of money metric utility is the precise "dual" of the quantity metric utility with which I began. If a diagram is drawn with \( p_1 \) and \( p_2 \) on the axes and indifference curves from the indirect utility function \( \phi(x, p) \) with \( x \) set at some arbitrary
number, say unity, then money metric utility is equivalent to marking off indifference curves by their distances from the origin along some reference price ray \( p^R \). This, of course, does not mean that quantity-metric and money-metric utility are the same. Comparing Figures 1 and 3 should show that this will only be the case (in all circumstances) if preferences are homothetic.

G. Price Indices

24. The "true" cost-of-living index is derived by comparing the cost of reaching a given indifference curve at different price vectors. Hence, using the notation \( P(p^1, p^0; u^R) \) for a price index (\( P \)) comparing \( p^1 \) with \( p^0 \) with the referencing variable after the semicolon, we have

\[
P(p^1, p^0; u^R) = \frac{c(u^R, p^1)}{c(u^R, p^0)}. \tag{23}
\]

These indices were first properly defined and analyzed in papers by Konüs (1924), (1939).

25. The dependence of \( P(p^1, p^0; u^R) \) on \( u^R \) is important. If preferences are homothetic, then utility can always be labelled so that the minimum cost of reaching \( u \) is proportional to \( u \). In this case, and in this case only, the cost function factors into two functions, one in \( u \) and one in \( p \), i.e. under homotheticity

\[
c(u, p) = g(u) a(p) \tag{24}
\]

for some increasing function \( g \) and some linearly homogeneous function \( a(p) \).

Hence, for this case, (23) reduces to

\[
P(p^1, p^0; u^R) = \frac{a(p^1)}{a(p^0)} \tag{25}
\]

which is independent of \( u^R \). In this case we have a unique cost-of-living or "the" price index.
26. As can easily be seen by considering an expansion path for a consumer with homothetic preferences, the pattern of demand is independent of the level of welfare (or level of income). In other words, all total expenditure elasticities are unity. In general, empirical evidence refutes such a possibility so that it is not true that, even for a single individual, a single price index can represent the true cost-of-living at different welfare levels. Similarly, if two consumers have identical non-homothetic tastes but different welfare levels, they will have different consumption patterns and will require different price indices to represent their respective costs-of-living. I shall return to this issue in paragraphs 71-2 below.

27. Not only does the cost-of-living index depend on the reference utility level chosen, but it should be realized that the index prices utility at different price vectors, not fixed quantity bundles. As relative prices change, the consumer can attain the same utility by re-arranging purchases, and in calculating price indices, this substitution must be taken into account. When this is done with reference utility levels $u^0$ and $u^1$ used for $u^R$ in (23), the famous Konüs inequalities immediately appear. For $c(u^0, p^1)$, we can write

$$c(u^0, p^1) \leq c(u^0, p^0) + (p^1 - p^0) \cdot \nabla c(u^0, p^0)$$

(26)

where $\nabla c(u^0, p^0)$ is the vector of first partial differentials of $c(u, p)$ at $u^0, p^0$ and $x \cdot y$ denotes a dot product. The right-hand side of (26) is a first-order Taylor approximation to $c(u^0, p^1)$; it is also no less than it since $c(u, p)$, as a minimum value function, is concave in $p$. Note that (26) immediately gives a first-order approximation and an upper bound -

$$(p^1 - p^0) \cdot \nabla c(u^0, p^0)$$

—for $C^V$ in (22), but if we divide through by $c(u^0, p^0)$,
using \( c(u^0, p^0) = x^0 = p^0 q^0 \) and the derivative property of the cost function
\( \gamma c(u^0, p^0) = q^0 \) (see Appendix), then
\[
P(p^1, p^0; u^0) \leq (\approx) p^1 q^0 / p^0 q^0 = P(p^1, p^0; q^0)
\]
where \( P(p^1, p^0; q^0) \) is my notation for the Laspeyres price index. By (27), this Laspeyres index is both a first-order approximation and an upper-bound for the base-referenced true cost-of-living index. Similarly \( (p^1 - p^0) \cdot q^0 \) is a first-order approximation and upper-bound for the Hicksian compensating variation. Repeating the analysis for \( u^1 \) rather than \( u^0 \), the analogue of (26) is
\[
c(u^1, p^0) \leq (\approx) c(u^1, p^1) + (p^1 - p^0) \cdot \gamma c(u^1, p^1)
\]
so that, corresponding to (27)
\[
P(p^1, p^0; q^1) = p^1 q^1 / p^0 q^0 \leq (\approx) P(p^1, p^0; u^1).
\]
Hence the Paasche index \( P(p^1, p^0; q^1) \) is no greater than the current-referenced true cost-of-living. Similarly, the equivalent variation \( EV \) in (21) is bounded below by \( (p^1 - p^0) \cdot q^1 \). Note that it is not true that the true index lies between the Paasche and Laspeyres; there is no (single) true index except when preferences and homothetic and the Paasche and Laspeyres indices are bounds to different concepts. Neither is it true that "consumer surplus" lies between \( (p^1 - p^0) \cdot q^1 \) and \( (p^1 - p^0) \cdot q^0 \); \( EV \) and \( CV \), like the two price indices, are distinct concepts. Note finally that the approximations in (27) and (29) will be exact if the cost functions is linear in prices, i.e. if
\[
c(u, p) = \sum_k k x_k(u) p_k.
\]
This is the case of non-homothetic Leontief preferences where indifference curves are right-angles so that there is no substitution in response to relative price changes. (Note, however, that this is not the only case where Paasche and/or Laspeyres indices are exact, see Diewert (1980) for examples.)

28. The inequality-approximations (27) and (29) can be translated into money-metric utility relationships using the identities (18) and (19). Specifically,

\[ c(u^1, p^0) = x^1/P(p^1, p^0; u^1) \leq (\simeq) x^1/P(p^1, p^0; q^1) \]

so that, cf. (6)

\[ c(u^1, p^0) \leq (\simeq) p^0.q^1 \]  

(31)

while, analogously

\[ c(u^0, p^1) \leq (\simeq) p^1.q^0. \]  

(32)

Of course, both (31) and (32) are immediately obvious from the definition (and derivative property) of the cost function.

29. How good are the Paasche and Laspeyres approximations in practice and how can they be improved? Consider one further term in the Taylor expansion in (26). Since the Hessian of the cost function is the Slutsky substitution matrix which I shall write as \( S \) with typical term \( s_{ij} \), the second order expansion is

\[ c(u^0, p^1) = p^1.q^0 + \sum_{i,j} s_{ij}^0 (p_i^1 - p_i^0) (p_j^1 - p_j^0), \]

(33)

where the superscript on \( s_{ij} \) indicates that the matrix is evaluated at \( u^0, p^0 \). Hence, knowledge of \( S \) can be used via (33) to improve the approximations.
In particular, we may well be interested in situations when only one or a small number of prices have changed (say in project appraisal) and the relevant elements of $S$ may well be known. In this case (33) can be used to measure CV or the welfare change given only knowledge of the crucial price elasticities.

30. Since the last term on the right-hand side of (33) is a first approximation to the error in the Laspeyres index, the expression also indicates how far the latter may be misleading. First, note that if $s_{ij}\approx 0$ or small, the quadratic form will be small and the error low independently of the price changes which occur. For broad groups of goods, substitution is likely to be small so that the Laspeyres will be accurate. Similarly, if preferences are those of the linear expenditure system or similar, little substitution is allowed and, once again, the Laspeyres will be accurate. (N.B. The fact that the Laspeyres is close to the true index based on the linear expenditure system is something which essentially happens by construction; it cannot be used as evidence in favor of the accuracy of the Laspeyres to the unknown true index). Second, if $p^1$ is approximately proportional to $p^0$, the quadratic form will be small. This occurs since $\Sigma s_{ij}p_j^0$ is zero by homogeneity, so that writing $p^1_i = \nu p^0_i + \epsilon_i$, and quadratic form becomes

$$\frac{1}{2} \sum_{ij} s_{ij}^0 \epsilon_i \epsilon_j$$

which, although always non-positive ($S$ is negative semi-definite), will be small if the $\epsilon_i$'s are. Since, in much time-series data, $p^1 = \nu p^0 + \epsilon$ is a plausible description of prices (and wages), the smallness of (34) is a good reason for believing Laspeyres indices to be accurate in that context. The same, of course, applies to the Paasche approximation. In contexts where $p^1$ is not approximately proportional to $p^0$, or where substitution possibilities
are large, the Laspeyres/Paasche approximations may be very poor. Perhaps the most important case here is in computing price indices between countries, e.g. using the sort of price information collected by Kravis et al (1975). In these comparisons, relative prices are very different and there is some evidence, see also Gilbert et al (1958), that there are large long-term substitution effects in response to these long-standing differentials. In consequence, the true indices, if they could be calculated, are likely to be very different from their Paasche or Laspeyres approximations.

31. The Paasche and Laspeyres indices are also attractive in that they require knowledge only of \( p^1, p^0, q^1 \text{ and } q^0 \), in contrast to the true indices which require knowledge of preferences, or even the approximation (33) which requires the substitution matrix. These latter will always be subject to greater uncertainty in practice than the prices and quantities. It is thus worth exploring other indices based only on these. In an important paper, Diewert (1976) calls price index numbers superlative (after Irving Fisher) if they are exact for a utility function general enough to provide a second-order approximation to whatever the unknown true utility function may be. Such utility functions (direct, indirect, or equivalently cost-functions) are known as "flexible" functional forms. For example, Fisher's own "ideal" index, the geometric mean of the Paasche and the Laspeyres, is superlative conditional on homotheticity since if the direct utility function is quadratic homothetic, the ideal index is the true cost-of-living index, an elegant theorem originally due to the Russian economist Buscheagueuce (1925). However, homotheticity is not an attractive assumption given its contrafactual implication.
of unitary income elasticities. Diewert, however, proves the following result.

Consider the cost function given by

\[ \log c(u, p) = c_0 + \sum_k x_k \log p_k + \frac{1}{2} \sum_{k,j} \alpha_{kj} \log p_k \log p_j + \beta \log u \]

\[ + \alpha (\log u)^2 + \frac{1}{2} \sum_k \gamma_k \log p_k \log u \]  

(35)

which is clearly a second-order approximation to whatever \( c(u, p) \) may be. Then, given (35), the Tornquist (1936) index given by

\[ \log P_T = \frac{1}{2} \sum_k (w_k^0 + w_k^1) \log \left( \frac{p_k^1}{p_k^0} \right), \]

where \( w_k = p_k q_k / x \) is the budget share for good \( k \), is exact for a reference utility level \( u^* = \sqrt{u^0 u^1} \). Of course, \( u^* \) may not be the reference level we are interested in, but if \( u^0 \) is close to \( u^1 \), this may not matter. In any case, index numbers similar to \( P_T \) are likely to work very well for measuring price indices over time at least as far as approximating the substitution effects of price changes. In particular, the chaining principle embodied in (36) is likely to be useful either as in the Tornquist index or in the corresponding Laspeyres type index

\[ \frac{1}{2} \sum_k (w_k^0 + w_k^1) \left( \frac{p_k^1}{p_k^0} \right). \]

In many practical contexts these will be quite adequate.

32. The Laspeyres and Paasche indices, like the true price indices, vary with income or welfare level. To illustrate, write the Laspeyres index in the form

\[ P(p^1, p^0; q^0) = \frac{p^0}{p^1} \sum_k w_k^0 \left( \frac{p_k^1}{p_k^0} \right). \]  

(37)

For the same household with different welfare levels, or for different households, \( w_k^0 \) will change (unless preferences are homothetic). Hence, approximating the true index by the Laspeyres carries no commitment to a single index; just as with the true index, the index will vary with the pattern of consumption. As we shall see later, it is often very important to allow for this, especially in
dealing with welfare comparisons across households. But even for a single consumer, welfare may change over time so that with non-homothetic tastes, systematic variations in the weights must be built into the indices. I shall give specific examples of how this might be done in paragraph 77 below.

For the moment, I wish only to emphasize that there are two quite separate issues in calculating price indices. The first is allowing for substitution as relative prices change; as argued above this may not be very important in a time-series context and can often be adequately captured using, for example, Tornquist indices provided that there is little welfare change over the period of change. The second issue is the variation in the index with welfare, which is probably the more important of the two issues, especially in making comparisons between households or individuals.

H. Separability and Subindices

33. So far I have simply assumed that the vector $q$ contains everything of interest to the consumer and it is now time to put some structure on this. In principle, there would be an element of $q$ for each of however many millions of goods are potentially available and for each of every conceivable type of labor supply. In turn, each of these would be represented for all periods in the consumer's life-cycle or even all the way till doomsday if the consumer is concerned with the welfare of his descendants. Clearly such a massive array of possibilities has to be subdivided into manageable sections. I shall develop this theory first, turning later to specific applications.

34. Preferences are said to be (weakly) separable if the direct utility function can be written in the form

$$u = \xi\{u_1(q^1), u_2(q^2), \ldots, u_N(q^N)\}$$

(38)
where the subvectors $q^1, q^2, \ldots, q^N$ form a partition of $q$, i.e. each good belongs to one and only one subvector. There is no restriction on the number of goods in each group, from one upwards. To fix ideas, one group might be leisure, another goods; one might be public goods, another private goods; or each subvector might correspond to a particular time period in the consumer's life. Each of the individual subutility (or felicity) functions $u_g(q^G)$ represents a preference ordering on the subvector $q^G$, the crucial point about weak separability being that these sub-orderings exist, so that preference rankings between $q^{G0}$ and $q^{G1}$, i.e. two distinct bundles within group $G$, are not affected by the quantities consumed outside the group. Of course this may be false; my ranking of what to eat today is unlikely to be independent of what I ate yesterday. Indeed, weak separability is in principle empirically testable through its implications for the structure of the substitution matrix (see e.g. Deston and Muehlbauer (1980, pp. 128-9)).

Even so, substitution matrices are hard to measure accurately, and weak separability is sufficiently weak for there to be little accumulated empirical evidence on the matter. In consequence, it is largely inevitable that we rely on intuition as to what does or does not constitute an appropriate separable group. It is worth noting also that weak separability is very far from being the only type of separability, indeed there are a very large number of possibilities. Each of the types has its own distinct implications for behavior and yet all correspond at least loosely to intuitive notions of separability. Even so, I shall confine myself to weak separability here.

---

1/ See, for example, Blackorby et al (1978).
35. Corresponding to the subutility functions $u_G(q^G)$ in (38) we can define subcost functions $c_G(u_G, p^G)$ for price subvectors $p^G$ and subutility levels $u_G$ which for utility maximizing consumers must equal the total expenditure on the group $x_G$, i.e.

$$c_G(u_G, p^G) = x_G. \tag{39}$$

If (39) were not true, greater utility could be attained for at least one of the groups at no increase in group expenditure so that, since $f$ in (38) is increasing in each of its arguments, overall utility cannot be maximal. The subcost functions can then be used to define money-metric utility and group price indices in exactly the manner discussed in Section G above. For reference group price vector $p^{GR}$, $c_G(u_G, p^{GR})$ measures the minimum expenditure on, say, food, to attain a given food standard of living ($u_G$) at reference food prices $p^{GR}$. This is the "food" welfare index or index of real food expenditure. Similarly $c_G(u^R_G, p^{GL})/c_G(u^R_G, p^{GO})$ is the true cost-of-living for the group index referenced on some fixed group welfare level $u^R_G$. In fact, these are the price and quantity indices which we can think of as being used by the consumer to allocate income or total expenditure to each of the $N$ groups.

36. The most obvious application of these ideas is to indices for subgroups of household expenditure. In the development context, we often want a measure of food consumption or of housing services; and these, like the overall measures, can be thought of as welfare indicators. In this context, price indices such as the Tornquist or Laspeyres have a useful property of aggregating from subindices to the overall index. For example, the Laspeyres can be written as

---

so that using within-group budget shares to derive group Laspeyres indices, the latter can be weighted using group budget shares to give an overall Laspeyres index. The same device works for the Tornquist index but does not do so in general; e.g. the true subindices cannot be aggregated to give the true overall index /1/.

I. Intertemporal Welfare, Income, Consumption and Assets

37. To consider the definition of welfare in the context of intertemporal choice, it is necessary to look at the consumer's choice problem. We do this at the beginning of the life-cycle and write life-time utility as

\[ u = u(q_1, q_2, \ldots, q_t, \ldots, q_L, A_{L+1}/P) \]  

(41)

where \( q_t \) is the consumption vector at age \( t \), \( L \) is the date of death, \( A_{L+1} \) are bequests and \( P \) is an index of prices comprising all goods in all periods from death to doomsday. This last term can be thought of as representing the indirect (vicarious) utility of the consumer's heirs generated from bequests. As the consumer ages, (41) gets "filled-in" term by term so that no matter at what point in the life-cycle we observe, (41) is the fundamental function for calibrating life-time welfare. To move to the cost function for (41), we need the budget constraint relative to which utility is maximized. Let \( A_0 \) be financial assets at the end of period 0 (starting assets) and I assume that assets earn interest at rate \( r_t \) at the beginning of period \( t \). Hence, the periods are linked by

\[ A_t = (1 + r_t) A_{t-1} + r_t - P_t q_t \]  

(42)

/1/ See Pollak (1975) for further discussion.
for exogenous labor income $y_t$ (I deal with the case of endogenous labor supply in Section K below). This period-by-period constraint can be used to link $A_{L+1}$ to $A_L$ and thence to $A_{L-1}$ and so on back to $A_0$. Doing so gives the full intertemporal constraint

$$
\sum_{t=1}^{L} \rho_t p_t q_t + \rho_L p_f \frac{A_{L+1}}{p_f} = A_0 + \sum_{t=1}^{L} \rho_t y_t
$$

(43)

where $\rho_1, \rho_2, \ldots, \rho_L$ are discount factors defined by

$$
\rho_t = \frac{1}{(1 + r_1)(1 + r_2) \ldots (1 + r_t)}.
$$

(44)

The constraint (43) is a linear budget constraint in which the choice variables $q_1, \ldots, q_L$ and $(A_{L+1}/p_f)$ have prices $\rho_1 p_1, \ldots, \rho_L p_L$ and $\rho_L p_f$ and where there is an overall expenditure constraint $A_0 + \sum \rho_t y_t$ given by assets and discounted future incomes.

38. The intertemporal cost function takes the value $A_0 + \sum \rho_t y_t$ and is written

$$
c(u_0, \rho_1 p_1, \ldots, \rho_L p_L, \rho_L p_f) = A_0 + \sum \rho_t y_t
$$

(45)

and this can be used, as before, to give money-metric utility, price and welfare indices. The price indices will be defined on current and future discounted prices and the income concept that is to be deflated comprises assets plus the sum of present and future discounted incomes.

39. It is undoubtedly possible to make a case for measuring welfare on a full intertemporal basis. Indeed, when doomsday comes, the recording angel will have all the information to fill in all the slots in (41) and (45) and will be able to complete a full assessment of earthly welfare for each individual before him. Until then, there are severe practical problems in implementing welfare or price indices based on the intertemporal cost function. The actual, objective, or ex post, welfare level is unknown and, in practice, future prices, discount rates and
incomes can only be estimated on a subjective basis. At any time, we can think of each individual as having expectations of all the required magnitudes, but if we use these, individuals in the same apparent objective circumstances are likely to have widely different welfare levels. This is very much against the spirit of the approach taken in this paper; moreover, I do not think that such concepts are what we usually mean when we talk about welfare.

An alternative approach is to use weak separability and assume that the utility function (41) is separable with each period’s consumption vector forming a group. Welfare would then be measured with reference to the subcost function for the current period which is a function of current prices and current utility alone. The welfare and price indices which result would then measure current welfare and prices alone and these more limited measures seem to me to be closer to the practical concepts we want.

40. This approach, based on intertemporal choice with intertemporal weak separability, is essentially one of deflating total consumption expenditure in the current period by the appropriate current cost-of-living index. This is, of course, not the same as deflating current income by a price index; labor income is taken to be exogenous and can follow any sort of pattern over the life-cycle so that there is no reason to suppose that income in any one period is indicative of welfare in that or any other period (but see paragraph 41 below). Total consumption expenditure in each period, by contrast, is determined according to the model by all the prices, incomes and discount rates, both present and future. Even so, there is no reason in general to suppose that the consumer will choose to spread welfare (or real consumption) evenly over time, especially since needs are not evenly distributed over the life cycle. This, of course, is the main argument in favor of using full inter-temporal welfare rather than the subutility for a single period. On the
other hand, models such as the permanent income or life-cycle hypotheses, at
least in their stricter forms, assume that welfare is evenly spread and if
this is so, not only is the one-period welfare measure attractive in its
own right, but it also has the advantage of indicating full intertemporal
welfare. There is certainly a good deal of empirical evidence, accumulated
over the years, which suggests that current expenditure is indeed responsive
to the variables predicted by the model of intertemporal choice. However,
for both theoretical (cf. paragraph 41 below) and empirical reasons, it would
be rash to assume a strict proportionality between real current consumption
and the real value of assets and discounted future incomes. Nevertheless,
the current welfare measure is theoretically valid as a measure of welfare
now and is almost certainly at least positively related to the wider life-
cycle concept.

41. One of the most important reasons why consumers may not plan intertemporal
choice as indicated above is the presence of constraints on borrowing, or
liquidity constraints for short. Poor consumers or consumers who live in
countries with underdeveloped financial systems often cannot borrow so that
if, for example, higher incomes are anticipated later, the optimal consumption
plan may require borrowing and thus cannot be realized. Like all ration or
quantity constraints (cf. paragraphs 55-63 below), borrowing constraints reduce
welfare. They also enforce a connection between current income and current
consumption (in the extreme case, consumption equals current income) which is
absent from the full intertemporal model. For such consumers (and the young
and poor are likely to be particularly affected), current income, appropriately
deflated, is a good measure of current welfare, as, of course, is current consumption. In this case, however, current consumption does not reflect the wider life-cycle welfare as it does when there are no liquidity constraints. Yet, it may well be the case that many poor consumers are not constrained in this way since their future expectations are no different from their current experience. For such consumers, current income, current consumption, and life-time consumption are all too closely related.

42. From all this, the best measure in my view is deflated current consumption expenditure. Even so, some time unit is required and here the obvious choice, that of the year, seems best. Anything shorter involves both the consumer and the measurer in correcting for seasonal fluctuations. Anything longer has no obvious basis in the planning of the consumer, and the detection of life-cycle effects would require much longer periods than are likely to be practical. It may also be possible to allow specifically for the effects of age on welfare by the methods of Part III so that correcting real annual consumption could help remove bias due to omitted life-cycle effects.

J. Durable Goods

43. According to the neoclassical theory of durable goods¹, durable goods can be treated as ordinary goods within an intertemporal budget constraint such as (43) provided the appropriate prices are used. The basic theory assumes that it is stocks of durable goods which generate welfare and that these suffer physical deterioration through time. This takes the special

¹ See, for example, Diewert (1974) or Deaton and Muellbauer (1980), pp. 105-9 and Chapter 13.
form that some proportion of the stock simply evaporates leaving a stock which is smaller than but physically indistinguishable from the original. This can be sold on second-hand markets at the same price per unit as new stock. If so, the cost of holding one unit of stock for one period is given by the user cost or rental equivalent price of

\[ p_t^* = p_t - p_{t+1}(1 - \delta_t)/(1 + r_{t+1}) \]  

(46)

where \( p_t \) is the market price in period \( t \), \( \delta_t \) is the rate of physical depreciation in period \( t \) and \( r_{t+1} \) is the discount rate linking \((t+1) \) and \( t \). Hence the user cost is the (present value) cost of holding one unit of stock for one period, reselling it in the next.

44. It is not necessary to take this whole neoclassical story completely seriously to recognize the essential validity of the user cost concept. In practice, one might wish to substitute some discounted second-hand prices for \( p_{t+1} \), but nevertheless for goods where second-hand markets are relatively good (houses, automobiles), it undoubtedly makes sense to price durables according to user cost in evaluating total expenditure or price indices. For goods that have little second hand value (household hardware etc.), the standard national income accounting practice of charging such goods as non-durables is not seriously misleading. However, durable goods will often have an important role as producer goods in households in poor countries, an aspect which I return to in paragraph 55 below.

K. Labor Supply and Leisure

45. So far, I have said nothing explicit about leisure or labor supply; implicitly, therefore, leisure is treated simply as another good. However, even this approach introduces a number of minor complexities. I shall examine first the case where the consumer is free to allocate his time between work and leisure
facing a fixed wage per unit of time worked. Not surprisingly, this can be handled in the standard way. I shall then deal with the (perhaps) more realistic case where hours are not flexible so that the consumer must work a fixed number of hours (including perhaps zero hours, i.e. unemployment).

In the neoclassical model of labor supply, the consumer is supposed to have a fixed number of hours available for work or leisure. This is denoted by $T$ and can be thought of as 8760 (hours per year) less those hours required for sleeping and minimal personal maintenance. Let leisure be $q$, so that hours worked are $h = T - q$, which can range from 0 to $T$. If the (fixed exogenous) wage rate is $w$, the utility maximization problem is to maximize

$$u = u(q, p)$$

subject to the budget constraint

$$p \cdot q = w(T - q)$$

where $w$ is (exogenous) nonlabor income. The budget constraint can be put in the standard form by defining $X$, full income, by $X = wT + u$, the value of the time endowment plus nonlabor income, so that (48) becomes

$$wq + p \cdot q = wT + u = X$$

Given this, we can define a "full" cost function $c(u, w, p)$ as the minimum expenditure on goods and leisure necessary to attain the indifference curve $u$, and for a utility maximizing consumer, this takes the value $X$. Then, just as before, money metric utility is defined by $c(u, w, p)$ for reference wage rate $w$ and price vector $p$, and, as before, we might approximate this by revaluing actual leisure and goods in a number of periods at base periods wages and prices. Similarly, the cost-of-living price indices, including the cost-of-leisure, would be given by $c(u, w, p) / c(u, w, 0)$ for suitable choice (s) of $u$.

1/ Becker (1965)
47. This model can be (and has been) complicated in a number of ways. For example, family labor supply can be modelled by recognizing different leisures for different family members each with their own wage rates. Alternatively, each individual can have different types of leisure or we can recognize various household "production functions" which combine time with goods to yield outputs which in turn generate welfare. In all these cases, suitable new linear budget constraints can be defined, with suitable modifications to the prices, and these will yield cost functions for welfare measures or price indices. Most importantly, perhaps, an intertemporal version of (47) can be set up to match (41) so that the consumer would simultaneously plan life-time consumption, hours worked and labor force participation, particularly retirement. Such a model is an essential ingredient in studying life-cycle labor supply variations or retirement decisions, but for the same reasons as in the intertemporal consumption analysis, I doubt whether the full intertemporal analysis with labor supply is either feasible or relevant.

48. Even so, the intertemporal analysis also emphasizes another aspect of behavior, that of educational choice and the endogeneity of the wage rate. Indeed, in the human capital literature, the wage rate is fully determined by the amount of human capital acquired through education and training. If this is so, then, except in the very short run, the wage rate is not a measure of the economic opportunities available to the consumer and cannot be used to measure welfare as in paragraphs 45-46 above. Faced with this, either we must fall back back on a short-run fairly narrow definition of welfare (and this is not

1/ See particularly the papers by Hanock (1976) and by Becker (1965).
2/ See, for example, Ghez and Becker (1975).
3/ See in particular the work of Mincer, e.g. (1974), (1976).
unattractive, see also paragraphs 40 and 42 above) or we must seek some wider measure of individual opportunities in the long-run, e.g. availability of educational opportunities. My own view is that this is unlikely to lead anywhere very useful. If the human capital-household production story is followed to its limits, we come to the position advocated by Stigler and Becker (1977): all consumers have identical tastes and if they behave differently it is because, with different and differently differentiated stocks of human capital, the implicit prices of different activities vary from consumer to consumer. Market prices are, of course, the same for everyone, so that, since human capital stocks are themselves chosen in response to price incentives, differences in behavior can ultimately only be due to random factors or to differences in birth (although this only shifts the problem back a generation). In such a world, welfare measurement can perhaps only be achieved through IQ tests. Although there is indeed a voluminous literature on the genetic determination of earnings, I do not believe that such an approach has much practical relevance in the present context.

49. Even if we return to the single period labor supply model of paragraph 46, there are a number of problems of implementation. First, even if labor supply is freely variable, it is far from obvious how to measure leisure, i.e. hours which could potentially be offered for work at w but are voluntarily retained. In principle, T can be treated as a parameter and estimated in a labor supply equation, 1/ but such attempts have not been notably successful, presumably because such models are not a reasonable representation of reality. Second, there is an inequality restriction, that hours worked be no less than

1/ See, for example, Abbott and Ashenfelter (1976).
zero, which induces a 'kink' in the budget constraint, see point A in Figure 4. For consumers at this point, those who choose not to participate in the labor force (e.g. housewives, retirees or other secondary workers), the linear portion of the budget constraint does not fully determine welfare and the cost function approach no longer works without modification. Consider the consumer on u at point A. The slope of the indifference curve here is steeper than the wage rate; alternatively, the shadow price of leisure (the slope of BB) is greater than the wage. For such an individual, an "artificial" budget set can be defined (i.e. BB in the figure) and the implicit shadow prices used to measure welfare. Once again, there is a practical problem in measurement since, by definition, the shadow wages are not observed in the market. Nevertheless, data on labor force participation and hours worked can be used simultaneously to estimate shadow wage rates and a great deal of extraordinarily high quality econometric work has gone into doing so (see in particular Heckman (1974) (1978)). Consequently, if we wished to use these shadow wages, the methodology exists to calculate them, albeit with some complexity.
50. As for the consumption case, an alternative approach is based on separability. If leisure is separable from goods (and for some goods at least this is certainly not true, see e.g. Barnett (1979)), then a sub-measure of welfare can be defined on goods alone and, for some purposes, this narrow definition of welfare, excluding leisure, may be thought to be desirable. Note, however, that if leisure and goods are indeed not separable (for example, if leisure time is required to consume goods), then the consumer's welfare ranking of different commodity bundles is not independent of the number of hours worked, so that an attempt to base a welfare measure on goods alone is ill-founded.

51. Perhaps the major difficulty with the neoclassical labor supply model is its descriptive inaccuracy. Although there are many economists who would not accept this, common observation suggests that many consumers are not free to vary their hours worked (or at least not with corresponding changes in remuneration) while others are 'involuntarily' unemployed. Such non-market constraints invalidate the cost function approach (shadow prices are not equal to market prices), although in paragraphs 55-63 below I shall discuss how the appropriate modifications can be made. But it is worth making a few points here. First, under the separability assumption of paragraph 50, commodity welfare can be measured independently of hours worked, so that the cost function approach can still be applied to commodity welfare, independently of whether or not labor market constraints are operating. Second, in evaluating leisure, 'involuntary' leisure has a shadow price less than, not greater than or equal to, the market wage. Indeed the shadow price of leisure for the involuntarily unemployed may well be zero or close to it. Similarly, to impute a shadow
wage of the market wage or above to a housewife on the assumption that she chooses to stay at home when, in fact, she is at home because there are no market opportunities for her, is to grossly inflate the measure of welfare. This is a good reason for caution in advocating the addition of the value of time in the home to standard welfare measurement. A similar problem arises in imputing values to the leisure activities of persons who do work. For example, is the gardening of the economics professor to be valued at the wage rate for gardeners (or less) or should it be valued at the wage rate for economists? Once again, the answer depends on the opportunities available to the individual. If the economist can earn the economists wage for any and all hours supplied, but chooses to garden, he is purchasing that time at the economist's wage and it should be charged as such with offsetting allowance for the value (if any) of outputs. Alternatively, it may be that the market wage falls with hours supplied, or that there are only a limited number of hours available which pay the higher rate. If so, a lower time valuation is appropriate.
L. The Consumer as Producer

52. In the standard model of consumer choice, an increase in a single price decreases welfare in proportion to the amount of the good consumed. In the labor supply analysis, this is no longer the case for one "price"; the wage rate, since although an increase in \( w \) makes leisure more expensive, the consumer has an endowment of leisure, \( T \), at least as large as the amount consumed so that the revaluation in the endowment more than offsets the increased cost of leisure. Many consumers have endowments beyond leisure time; there are stocks and property rights of various sorts. A similar phenomenon occurs when consumers purchase negative quantities of goods, i.e. when they are suppliers. In developed economies, the national accounts attempt to draw a line between the consumer as consumer and the consumer as producer or small businessman but this is to some extent artificial and in any case is unlikely to be practical for developing countries.

53. As a starting point consider the simplest case of a consumer/producer whose consumption and production activities can be entirely separated. The utility function is \( u(q) \) as before and, although the production part of the agent may produce some (or all) elements of \( q \), we assume that \( q \) may be bought and sold at a fixed price vector \( p_q \). Assume that at a vector of inputs \( z \), prices \( P_z \) are used to produce the outputs. Define the profit function over \( p_q \) and \( p_z \) by
\[ \pi(p_q, p_z) = \max_{q, z} \left\{ p_q q - p_z z; d(q, z) = 1 \right\} \]  

(50)

where \( d(q, z) \) is an implicit representation of the technology linking inputs and outputs. (N.B. many elements of \( q \) may be zero). Now since the elements of \( z \) do not enter the utility function, the consumer can take \( \pi(p_q, p_z) \) as the constraint on purchases, so that \( u(q) \) is maximized subject to

\[ p_q q = \pi(p_q, p_z). \]  

(51)

The cost function \( c(u, p_q) \) then takes the value \( \pi(p_q, p_z) \) for a utility-maximizing profit-maximizing consumer/producer so that money-metric utility is given by deflating maximal profits \( \pi(p_q, p_z) \) by the cost-of-living index \( c(u, p_q)/c(u, p^*_q) \). Note that increases in \( p_q \) will now increase or decrease welfare as profits \( \pi(p_q, p_z) \) increase more or less than the cost-of-living \( c(u, p_q) \). Since \( \partial \pi/\partial p_q \) is the vector of outputs of \( q \) and \( \partial c/\partial p_q \) is the vector of consumption \( q \), price increases increase or decrease welfare as the agent is a net producer or consumer of the good in question.

54. The foregoing example is artificial in that none of the inputs to the firm side of the consumer enter the utility function. Hence the consumer can be thought of as an owner or stockholder who simply takes the firms profits as his income constraint. This hardly represents the situation of a small farmer or retailer whose own labor in particular is an essential input to production. To model this more important case let \( q_0 \) be a vector (or scalar) of leisure times or other inputs which enter utility, which are inputs in production, but which are not sold on the open market. Define then the restricted profit function over \( p_q, p_z \) and \( q_0 \) as
\[ \pi^*(p_q, p_z; -q_0) = \max_{q, z} \left\{ p_q \cdot q - p_z \cdot z; d(q, z, -q_0) = 1 \right\} \] (52)

so that maximization takes place with respect only to \( q \) and \( z \) with \( q_0 \) fixed.

The negative signs on \( q_0 \) indicate that although \( q_0 \) enters utility positively, u = \( u(q_0, q) \), increases in \( q_0 \) decrease profits and production (again, think of \( q_0 \) as leisure). The consumer/producer's problem is then

maximize \[ u = u(q, q_0) \]
subject to \[ p_q \cdot q = \pi^*(p_q, p_z; -q_0) \]. (53)

At an interior solution, this will involve, for positive Lagrange multiplier \( \lambda \)
\[ \frac{\partial u}{\partial q_0} = \lambda \frac{\partial \pi^*}{\partial h_i} ; h_i = -q_{0i} \] (54)

Since the derivatives of a restricted profit function with respect to conditioning variables are the shadow prices (in production) of those variables (marginal profit products), (54) tells us that the shadow prices of the \( h_i \)'s (inputs) act directly as the prices of the utility yielding \( q_{0i} \)'s. In other words, by linearizing around the optimum \( q_0^* \) we can rewrite the budget constraint (53) as

\[ p_q \cdot q + \lambda \frac{\partial \pi^*}{\partial h} \cdot q_0 = \pi^{**} \] (55)

where \[ \pi^{**} = \pi^*(p_q, p_z; -q_0^*) + \lambda \frac{\partial \pi^*}{\partial h} \cdot q_0^* \] (56)

These equations can then be used in exactly the same way as was (51) above provided the concepts are appropriately modified. \( \pi^{**} \) is total business profit comprising money profits from sales plus the (shadow) value of output of \( q_0 \). This has to be deflated by a price index whose constituents are the commodity prices and the prices of non-market inputs, the latter evaluated by their
marginal profit values on the production side. The shadow prices involved here are essentially producer shadow prices unlike the consumer shadow prices in paragraph 11 or in Section (M) below and should therefore be considerably easier to attach values to in practice.

55. The methods of the preceding two paragraphs can also be applied to the measurement of welfare resulting from the presence in the household of durable goods the primary purpose of which is for producing other goods, either for domestic consumption or for external sale. In the simplest case, where the outputs of production have a clear market value, where the durable good does not contribute directly to welfare, and where equivalent rental markets exist for the durable, the analysis is that of paragraph 53 with the cost of the durable good’s services priced at user cost or rental equivalent price, see paragraph 43 above. Household income thus includes the value of output from the durable net of the cost of inputs including the user cost of the durable. If, on the other hand, the second-hand markets do not exist, the durable goods are appropriately treated as fixed factors in the restricted profit functions of paragraph 54. They then have an imputed value equal to their shadow price in production (marginal contribution to profits) allowing both for 'business' and 'personal' gain. For specific types of household capital (a plow, a last, and so on) these shadow prices may be readily available. In other cases, and especially for poor households, the presence of the durable good may be of greater importance than its precise valuation. This is important to realize since presence or absence can be straightforwardly and accurately assessed by survey personnel. Various alternative imputations can then be investigated at the analytical stage of welfare measurement studies.
56. The household production approach can, of course, be extended to all of consumer behavior by regarding goods as inputs into activities with the latter generating utility. Such a procedure has been used to great effect in the theoretical analysis of behavior in a wide range of circumstances. For practical work, however, it is necessary that the activities and their production functions are clearly defined (as, for example, when actual production takes place) otherwise the approach is not effectively different from the standard analysis discussed above. One possible exception, however, is the recognition that many purchases by consumers are of little or no direct benefit to themselves, but rather are necessary if other activities are to be performed. Obvious examples of such "regrettable necessities" are transport to work, or the fixed costs which a housewife may have to overcome to be able to participate in the labor force (child nursery care, etc.) If these are well defined, the money-metric utility estimates can straightforwardly be modified to account for them, simply by a shift of origin. To take a specific example, the $\gamma_i$ parameters in the linear expenditure system can be interpreted as "regrettable necessities" since they yield no utility in themselves. Money metric utility for the linear expenditure system without modification is the constant price value of the $\gamma$'s plus the real value of total expenditure less subsistence. After correction, only the latter remains. More generally, an appropriate correction would be to use the cost-of-living index to deflate, not total expenditure, but total expenditure exclusive of the non-productive expenditures.
M. Welfare with Quantity Restrictions

57. I deal finally with the important mixed case where some goods are bought in the market at parametric prices while others are supplied in fixed quantities outside the consumer's control. A number of examples have appeared already. Among the most important are public goods (schools, hospitals, parks, law enforcement, defence, social insurance) and involuntary unemployment or restrictions on hours. Rationing as in wartime is another example and, although the whole subject is frequently dealt with under this heading, it is important to realize that, in general with quantity restrictions, the consumer may have to consume more as well as less than he would optimally choose. The presence of transactions costs or imperfect markets may also effectively constrain consumers to retain their current consumption levels, particularly of durable goods, e.g. housing, heating plant, etc. All such phenomena, by placing an additional constraint on behavior, reduce welfare below what it would have been with free choice, parametric prices, perfect markets and absence of transactions costs.

58. I denote the vector of fixed quantities by $\mathbf{z}$, provided at prices $p_z$ (which may be zero). The consumers choice problem is then to maximize utility $u(q, z)$ with respect to $q$ only and subject to the budget constraint $pq = x - p_z \cdot z$. Corresponding to this there is a restricted cost function $c^*(u, p_z, p; z)$ which gives the minimum cost of reaching $u$ at $p_z$ and $p$ given that the vector $z$ must be bought. Note that we must have
for unrestricted cost function \( c(u, p_z, p) \); since \( z \) is one of the choices available to the consumer, the minimum cost of \( u \) must be no greater if \( z \) and all other vectors can be chosen. For many purposes, e.g. in measuring welfare in the presence of public goods, the restricted cost function is all that we need since the composition of the \( z \) vector remains constant (i.e. public parks are never available in the market). The only new factor in using \( c^* \) for welfare measurement compared with \( c \) is the presence of the \( z \) vector. For example, the deflator for expenditure in the face of a change in prices and public goods would be \( c^*(u, p_z^1, p^1; z^1)/c^*(u, p_z^0, p^0; z^0) \) or, in consumer surplus terms, the amount needed to compensate a consumer, originally at \( u^0 \), for the change in \( p \) and \( z \) would be

\[
CV = c^*(u^0, p_z^1, p^1; z^1) - c^*(u^0, p_z^0, p^0; z^0). \tag{58}
\]

Once again, the derivatives of the cost function can be used to give approximations to these expressions; as before \( \partial c^*/\partial p \approx q, \partial c^*/\partial p_z = z \) (since \( p_z \) only appears in \( c^* \) in the additive term \( p_z^1z \)) and, as appears below, \( \partial c^*/\partial z = -p_z^* \), the shadow prices of \( z \). Hence, for example, (58) can be approximated by

\[
CV \approx z^0(p_z^1 - p_z^0) + q^0(p^1 - p^0) - p_z^*(z^1 - z^0) \tag{59}
\]

so that price increases decrease welfare (increase required compensation) while quantity increases increase welfare in proportion to their shadow prices. As before \( CV \) is overstated by the right-hand side of (59), since \( c^* \) is concave in both \((-z)\) and \( p \). Corresponding approximations exists for the true cost-of-living indices, e.g. the base-weighted index is related to the Laspeyres by
so that the ordinary Laspeyres index of price change has to be modified by the shadow evaluation of the change in quantities as a percentage of original expenditure. Such indices would be appropriate measures of the change in the cost-of-living if, for example, a government were to cut indirect taxes and reduce social services simultaneously.

59. In other situations, welfare measurement is rather different in that we wish to compare a quantity-constrained with an unconstrained position, for example, if we wish to discuss the welfare costs of unemployment. The natural measure of this is the difference between the two sides of the inequality (57) taking \( u, p_z \) and \( p \) as fixed; indeed this is the compensating variation required for an exogenous change in \( z \). The relationship between the restricted and unrestricted cost functions turns out to be

\[
c^*(u, p, p; z) = c(u, p^*, p) + (p_z - p^*_z) \cdot z
\]

where

\[
\frac{\partial c(u, p^*_z, p)}{\partial p_z} = z.
\]

Note that, contrary to its immediate appearance, (61) is exact, not an approximation; indeed (62) gives \( p^*_z \) as a function of \( u, p \) and \( z \) so that (61) is an identity in \( u, p, p_z \) and \( z \) which defines \( c^*(u, p_z, p; z) \) for any given preferences as represented by \( c(u, p_z, p) \). Note too that (62) is just a

\[\text{See Neary and Roberts (1980) or Deaton (1980).}\]
formal definition of the shadow prices $p_z^*$. Since $\frac{\partial c}{\partial p_z}$ gives the unrestricted demands for $z$ at $u$, $p_z$ and $p$, $p_z^*$ is that price which would cause a consumer at $u$ facing prices $p$ to choose $z$. The two equations (61) and (62) can also be used straightforwardly to show that $\frac{\partial c^*}{\partial z} = -p_z^*$.

60. The cost of the quantity constraint is then given by deducting $c(u, p_z^*, p)$ from the right hand side of (61). Hence, for a consumer, originally constrained, who is subsequently forced to buy $z$, the compensating variation is

$$CV = c(u, p_z^*, p) - c(u, p_z, p) + (p_z - p_z^*) z$$

(63)

which can be approximated by (and is bounded above by),

$$CV \approx (z) (p_z - p_z^*) (z - z^*) \geq 0$$

(64)

where $z^*$ is the free demand for $z$ at $u$, $p_z$ and $p$. Since $p_z > p_z^*$ as $z > z^*$ (if the consumer is forced to buy more than is desired, the market price is above the shadow price, and vice versa), the approximation is always positive as required. As an example, for an individual who becomes wholly unemployed, for whom leisure has a zero shadow price and who desires to work full-time, the compensation required by (64) would be the full income foregone.

61. The major practical problem in all this is the estimation of the shadow prices $p_z^*$. There are a number of possibilities, none of which is likely to be entirely satisfactory. Nevertheless, as with the evaluation of producer prices, a combination of techniques applied to specific problems may well yield satisfactory answers. It should also be emphasized that it is important not to ignore the problem. Substantial contributions to welfare are provided by non-market goods and services almost everywhere and major bias will result if these contributions are ignored.
62. The first possibility is that unconstrained demand functions can be observed for periods or consumers who are not subject to the quantity constraint. For example, labor supply functions can be used to estimate the shadow price of leisure (cf. paragraph 49 above) and this can be used in a subsequent period of unemployment to evaluate its welfare cost. Alternatively, in for example the housing market, it has been widely observed that poor households spend around one third of their incomes on housing and this figure can be used as a shadow price for evaluating subsidized housing programs. More generally, any demand function or system of demands estimated for the relevant consumers or some group of similar consumers can be "inverted" to yield prices as a function of quantities for the relevant commodities. In some cases, where markets clear by price adjustment in the short-run, demand functions may even have been estimated in this form with price as a function of exogenously supplied quantities.

63. However, for many goods (including most public goods), the market may never have functioned so that the information needed to estimate demand functions does not exist. One possibility in this case is to estimate demand functions for market goods conditional on the public goods and use the effect of the latter on the former to evaluate shadow prices. This works in principle because the free demands \( q \) are the derivatives of the rationed cost function \( c^*(u, p_z, p; z) \) with respect to \( p \) so that knowledge of the demands can be turned into knowledge of \( c^* \) which, in turn, gives \( p_z^* \) via \( \frac{\partial c^*}{\partial z} \). (I shall discuss essentially similar techniques for studying family composition effects in Part III.)
Put another way, we can observe \( \frac{\partial q_i}{\partial z_j} \) directly (e.g. by observing the effects of state hospitals on private medical expenditures). But \( q_i = \frac{\partial c^*}{\partial p_i} \), so that \( \frac{\partial q_i}{\partial z_j} = \frac{\partial^2 c^*}{\partial z_j \partial p_i} \) which, by Young's Theorem, is equal to \( \frac{\partial^2 c^*}{\partial p_i \partial z_j} \) which is \( \frac{\partial}{\partial p_i} (p^*_j) \). Integrating this gives \( p^*_j \). The problem with the technique is that we only pick up those effects of changes in \( z \) which affect behavior on the other goods. If, for example, \( z \) is weakly separable from \( q \), changes in \( z \) will only have income effects on \( q \) so that the compensated demands for \( q \) are independent of \( z \); technically, this implies that \( p^*_2 \) is independent of the prices of the \( q \)'s, so that the whole of \( p^*_2 \) is lost in differentiating with respect to \( p_i \) and 

\[
\frac{\partial q_i}{\partial z_j} = \frac{\partial^2 p^*_j}{\partial p_i} = 0 \text{ identically.}
\]

More bluntly, imagine that to an original direct utility function \( u(q, z) \) we add some function of \( z \) alone, e.g. \( f(z) \). Then since \( z \) is not under the consumer's control, \( u = u(q, z) \) and \( u = u(q, z) + f(z) \) have identical behavioral implications for the \( q \)'s and we have no means of identifying the impact of \( z \) on welfare.

64. In the next part, I shall argue the case for not recognizing welfare effects which have no counterpart in observable behavior. For some public goods, the same argument might work if it is insufficiently plausible that they form a separable group. For example, if the government introduces state medical care and subsequent consumer behavior shows no change whatsoever, I think that there is a very good case for assigning a zero shadow price to the program. However, if we take defence expenditure as a counterexample, a decrease in the armed forces is unlikely to induce individual consumers to place cannons on their rooftops, and we would not expect to be able to assign a shadow price to defence on the basis of observable market behavior. Once again, this technique, like
the last, is likely to work for some goods but not for others.

65. The first possibility is to attempt to design survey questions which elicit the relevant prices. A certain amount of experimentation has been done in this area, 1/ and this is likely to be a fruitful area for further research and discussion, especially between survey experts and public economics theorists.

N. The Generation of Welfare

66. The focus so far has been almost entirely on the measurement of welfare and I have abstracted from the factors determining it. However, measurement is a prelude to the improvement of economic conditions and in survey work it is crucial not only to assess the current state of affairs, but also to measure the factors which will have to be affected when improvement is sought. The next two paragraphs list the most important of these; they are also the variables which will be required in analytical work on survey results.

67. In paragraphs 37–42 above, I argued in favor of consumption rather than income as a basis for welfare measurement. But it is income which provides the means for consumption, and survey results on sources of income are a vital ingredient in the analysis of the level and distribution of consumption. To list the most obvious contributing factors requiring both measurement and analysis: employment opportunities, especially between formal and informal, rural and urban sectors; unemployment and its geographical incidence; the

1/ See, for example, Maital (1979), Mueller (1963), Watts and Free (1973) and Weisbrod (1978).
labor force participation behavior within the household, and the point at which children and mothers seek formal or informal employment; the determination of wages and the influence of education and experience on market wages and on productivity; the pattern of wages, productivity and earnings over the life-cycle. All these are major topics of research in their own right, each in an important part of the welfare story, and each imposes its own requirement in terms of data collection and survey design.

68. A second set of factors interact with income in the generation and distribution of welfare. Prices constitute one such factor, in particular their variation across regions and between rural and urban areas. Transport is a crucial factor here, as well as the availability of income in kind to agricultural workers. The distribution of needs is a second factor, especially the role of demographic variables such as family size, life expectancy and the variation of needs with age. The relationship between family composition and welfare is discussed in the next section, but the independent causal part played by the demographics is a subject in its own right. The third important group of variables are those facilities which are publicly provided, such as water, drainage, minimal housing, education as well as the availability of social security and health systems. Many of these would be directly measured in measuring welfare, but those which exert an indirect influence should also be covered in survey work.
III. COMPARISONS OF WELFARE BETWEEN INDIVIDUALS

A. A Unit for Comparisons

69. In Part II above, I argued that money metric utility, to be measured by deflating expenditure by a cost-of-living index, was a suitable measure of individual welfare. In this subsection, I shall attempt to justify the same measure for comparisons between individuals. Later in the section I shall turn to the specific issues which arise in its calculation in this context.

70. It is possible to take the view that no unit exists which gives a satisfactory measure of welfare between different individuals. Consumers have different tastes, face different circumstances, and are of different ages and abilities. And even if this were not so, if we could find two identical individuals who respond identically to every stimulus to which we can subject them, can we be sure that given the same income and the same prices they have the same welfare? For all practical purposes, I believe the answer to this question to be yes. Indeed I shall take it as axiomatic that for two individuals with identical preferences (i.e. identical behavior in all circumstances), the same economic constraints give rise to the same welfare. Even if, in some fundamental sense this is not true, if one individual can extract more welfare from the same consumption than can another identical individual, I believe that practical welfare measurement should be fundamentally based on opportunities rather than on their untestable consequences. No government is going to give special treatment to an individual who claims his extra sensibilities require special facilities, at least not without some objective evidence of why money means something different to him than to anyone else.

71. Welfare measurement however, cannot be confined to individuals with identical tastes; instead some way must be found to compare consumers who have different preferences. The most fruitful way of doing this is to "explain"
differences in tastes in terms of observable differences in household characteristics (and perhaps random terms to pick up unobservables, depending on the purpose at hand). Hence, fundamentally, all households are assumed to have the same utility function but this function has as many arguments as are necessary to explain variations in behavior. Obvious examples of conditioning variables are age, numbers and ages of children, education, social status, residential location, religious affiliation, and so forth. Provided this list is long enough, modelling preferences in this way is not very restrictive, although Arrow (1977) had described the procedure as "denying (his) individuality in a deep sense." In practice, however, only a few conditioning variables can be allowed for, and this embodies much stronger assumptions. But for those conditioning variables which can be included in demand functions we can, in principle, calculate their welfare consequences. For example, in paragraphs 75-83 below, I shall discuss methods for estimating how much extra outlay a household with, say, two children, would have to have to be as well-off as a family with no children. In this sort of way, we can calculate the compensations required for each of the various characteristics.

72. Two separate problems must now be faced. Firstly, there may be some observable characteristics which, although conceivably contributing to welfare, have no apparent effect on behavior. For example, two religious sects might each regard their own welfare as infinitely higher than the eternal damnation which is the lot of the other. However, religion may have no detectable effect on behavior once other conditioning variables are allowed for. In such cases I would propose to make the strong assumption that if an observable characteristic has no effect on behavior then it has no effect on welfare. No doubt counterexamples to this can be constructed (see the parallel discussion in paragraph 62), but I see no
other acceptable way of proceeding. The second problem is the opposite one; some characteristics may affect behavior, and thus presumably welfare, yet we do not wish to officially recognize the welfare loss in measurement. For example, we might well be able to recognize "greediness" as a characteristic and certain behavioral patterns are likely to be well explained by the presence or absence of it. Presumably, too, greedy people obtain less welfare from the same quantities than do more moderate individuals. Even recognizing all this, I think there is a strong case for not allowing the presence of greediness as a characteristic to depress observable welfare. Particularly if we believe that measuring welfare is only a first step in its promotion, most people would agree that greedy people ought not to be compensated for their greed. I use this artificial example as a stalking horse for disallowing the much more serious phenomena of altruism and envy. Envy ought not to be compensated nor should altruists be taxed. Neither ought to be a component in measurable welfare.

73. On these arguments, if taste variation is correctly modelled, correcting expenditure not only for prices but also for welfare-affecting characteristics will give a correct indicator of welfare across different households. But what if tastes vary in some other way or if the modelled variation only accounts very partially for the great diversity in reality? Then the measure of the welfare of household $h$, which I shall propose, can be interpreted as the amount of money a typical household would require at reference prices and its own (reference) characteristics to reach the welfare level it itself would reach were it to have the income and characteristics of household $h$. If the typical household and household $h$ do indeed have the same tastes, allowing for conditioning variables, this measure is simply the money metric utility for household $h$. If not, it is the welfare a typical household would obtain in household $h$'s circumstances.
Once again, I believe that this is a correct representation of the implicit considerations usually present in practical welfare measurement. Welfare is once again assessed by opportunities, standard or typical tastes being used to translate these opportunities into a scalar measure of welfare.

74. Finally, something should be said about the unit whose welfare is being measured. In principle, this could be an individual, a household or a family. There are strong arguments for using individuals as the basic atoms of the economic structure. Even so, such models as we have of the relative consumption patterns of adults and children work almost entirely at a household or family level. In consequence, theory and practical data collection probably only coincide for a household definition. The problem with this is that a household may regard itself as reasonably well-off even though, by social standards, the children may be badly treated. Social welfare is essentially defined over individuals and this has to be allowed for when using measures based on household measurement, e.g. by separate studies of children's welfare. Household and individual welfare can only be fully reconciled by means of a model of allocation within the household and none such has been fully worked-out and tested as far as I am aware, but see Samuelson (1956) and Muellbauer (1976) for first attempts.

B. Money Metric Utility Across Households

75. The use of money metric utility across households involves labelling their standards of living by the money needed to reach them at some reference prices. This is a device employed in everyday parlance: a family is described as a $10,000 a year family and so on. If all households face the same prices, then this is equivalent to measuring instantaneous welfare by expenditure. If more than one period is involved, or different households face different prices, then real measures must be used and, once again, we can deflate money expenditures by the appropriate cost-of-living index.
76. It is important to realize that it is not appropriate to use the same cost-of-living index for all households. First, different households may face different prices perhaps because they live in different areas and some factors are immobile or because there are transport costs. Second, different households may have different tastes; this issue is discussed in paragraph 78. Third, even if all tastes are identical, as was pointed out in paragraph 25, the true cost-of-living index \( c(u, p^1)/c(u, p^0) \) depends on \( u \), so that better-off consumers have different cost-of-living changes than poorer consumers.

77. To see how this works, consider the effect of a change in \( p_i^1 \) in period 1 on the cost-of-living index \( c(u^0, p^1)/c(u^0, p^0) \). In proportional terms

\[
\frac{\partial \log P(u^1, p^0; u^0)}{\partial \log p_i} = \epsilon_i^0
\]

Hence, differentiating again with respect to \( \log x^0 \),

\[
\frac{\partial}{\partial \log x^0} \frac{\partial \log P}{\partial \log p_i} = \epsilon_i^0 - 1
\]

where \( \epsilon_i^0 \) is the income elasticity of good \( i \) evaluated at \( u^0 \) and \( p^0 \).

Hence the effect of a price increase is larger with higher expenditure if the good is a luxury and is smaller with higher \( x \) if the good is a necessity. Thus if, as prices change, there are significant relative price shifts between necessities and luxuries, the distribution of welfare between households will change, even if there has been no change in expenditure levels. Such relative price changes have been important historically especially between food, durables and services,\(^{1/}\)

\(^{1/}\) See, in particular, Kuznets (1962).
while in the short-run, at least recently, fluctuations in the world price of food have been responsible for considerable differential price effects. Note that the dependence of \( P(p^1, p^0; u) \) on \( u \) (or \( x \)) is captured by the Laspeyres, Paasche or Tornquist approximations. Using the last to illustrate, and writing \( w^* \) for \( h(x^0 + x^1) \),

\[
\log P_T = \sum w^*_k \log \left( \frac{p^1_k}{p^0_k} \right)
\]

If, for example, Engel curves take the form

\[
v_i = \alpha_i + \beta_i \log x
\]

with \( \Sigma \alpha_i = 1 \), \( \Sigma \beta_i = 0 \), the Tornquist index is

\[
\log P_T = \alpha \log \left( \frac{p^1}{p^0} \right) + \log x^* (\beta \log \left( \frac{p^1}{p^0} \right))
\]

where \( x^* \) is the geometric mean of \( x^1 \) and \( x^0 \). The first index in (69) is a fixed-weight geometric index with weights adding to unity. This is the base index which is incremented by the marginal index \( \beta \log \left( \frac{p^1}{p^0} \right) \) multiplied by \( \log x^* \). Since \( \beta_i \) is \( > \) or \( < 0 \) as \( i \) is a luxury or a necessity, the marginal index is positive if, between 1 and 0, luxuries have become relatively more expensive, while it is negative if price increases for necessities have been larger. In the former case, \( \log P_T \) increases with \( \log x^* \), in the latter it decreases. Since the parameters of (68) are easily estimated and since the Engel

---

1/ See, for example, Muellbauer (1974) and the other references in Deaton and Muellbauer (1980), Chapter 7.
curve form has extremely nice theoretical and empirical properties, this methodology allows a simple way of incorporating behavioral responses into measurement. See Muellbauer (1978) for the original application of the technique.

78. A precisely similar technique can be used to incorporate differences in family composition and other household characteristics. If the relevant variables are collected in a vector $\mathbf{a}$, then (68) can be modified to yield

$$w_i = c_i + \beta_i \log x + \gamma \cdot a$$

(70)

for parameters $\gamma$ so that the effects of $a$ on the consumption pattern are built-in to the index. I shall discuss various theoretical approaches to the way of entering the $a$'s in the next subsection.

79. I wish finally to note the results of a recent paper by Michael (1979). Michael evaluated Laspeyres indices for a large number of U.S. households using the BLS survey tapes for 1971-2 and then assessed how much of the variability could be explained by variations in observable characteristics. Although such variables were always highly significant, the proportion of variance explained was very low. Michael concludes from this that it is therefore not very useful to prepare different indices for different groups since the intragroup variation is so large relative to the intergroup variation. This conclusion does not seem to me to follow. Presumably, most of the taste differences which are not being explained are those which we do not wish to compensate for or which arise through sampling variability or because of differences in timing of purchases. The real question is whether the household characteristics which are socially regarded as affecting welfare are or are not significant in the analysis. This question is answered in the affirmative by Michael's results.

1/ See again Deaton and Muellbauer (1980), Chapters 1, 3 and 6.
C. Welfare and Family Composition

80. If we follow the general approach of attempting to model taste differences, then the utility function for all households is written \( u = y(q; a) \) for a vector of conditioning household characteristics \( a \). Corresponding to this there is a cost function \( c(u, p, a) \), also conditioned on \( a \). Given the money metric approach, the quantity

\[
CV = c(u^R, p^R, a^h) - c(u^R, p^R, a^R)
\]

is the compensating variation the reference household (with utility \( u^R \), characteristics \( a^R \) and facing prices \( p^R \)) would require in order to restore its original welfare having been transformed into household \( h \). The ratio form of this index number, i.e.

\[
\frac{c(u^R, p^R, a^h)}{c(u^R, p^R, a^R)}
\]

is called an equivalence scale. If, for example, household \( h \) consists of two identical adults, while the reference household has one, (72) might give the answer 2 (although economies of scale within the household might reduce this somewhat). If \( h \) has two adults and a child, the scale would presumably be some fraction between two and three. Hence, when the reference household is a single adult, the equivalence scale can be thought of as a number of equivalent adults. (Of course, there is no reason why a childless couple or any other combination cannot be taken as reference.)

81. There are a number of different ways of specifying the conditional cost function \( c(u, p, a) \) and each gives rise to its own equivalence scale and its own model of household behavior. Much of the recent work in this area has been
done by Muellbauer (1977) and (1980) and the following discussion is based on this as presented in Deaton and Muellbauer (1980, Chapter 8), but see also Pollak & Wales (1978).

82. The simplest form of \( c(u, p, a) \) is a multiplicative one, i.e.

\[
c(u, p, a) = m(a) \gamma(u, p) \tag{73}
\]

for some functions \( m \) and \( \gamma \). This might be called "compositional homotheticity" since, given \( u \) and \( p \), the pattern of demand \( (v_i = \partial \log c / \partial \log p_i) \) is independent of household composition. Hence, changes in \( a \) affect the demand pattern only insofar as they affect welfare. In such circumstances, there is a simple way of evaluating \( m(a) \), which given (73) and the normalization \( m(a^R) = 1 \), is the equivalence scale (72). Two households, say \( h \) and \( R \), which have the same demand pattern must have the same welfare (since we assume that prices are the same for all). Hence, by (73), \( x_h / m(a^h) \) must equal \( x_R / m(a^R) = x_R \). This immediately gives the equivalence scale as \( x_h / x_R \); i.e. given the same demand pattern, differences in income reflect the exact compensation for the differences in characteristics.

83. In practice, this theory might be used as justification for the oldest and most popular technique of determining equivalence scales, that due to Engel. This uses simply the food share as an indicator of welfare rather than the whole budget pattern. Engel himself not only noticed that the food share typically decreased with welfare level but also that it increased with family size. Thus arose the idea of using the food share to indicate welfare, and it is still a widely-used concept, e.g. the food share is widely used as a dimensionless indicator of welfare in international comparisons, while the U.S. government implicitly defines its official poverty level by the expenditure corresponding to a food share of 1/3.
Although it is true that the food share shows the most dramatic variation with
the welfare level, the theory of (73) applies to all goods and if the model were
correct, the same equivalence scale should be obtained in whatever good is chosen
as indicator. In practice this is not the case implying, therefore, that the data
apparently do not conform to the model (73).

84. It is not hard to see why (73) should be too restrictive. The equation shows
clearly that the only effect of changes in composition is to reduce the purchas-
ing power of a given total expenditure. While it is reasonable to suppose that
this is one of the effects of having more children it is unlikely to be the only
one. In particular the special needs of children for specific types of goods are
ignored. This is remedied in the model of Barten (1964) who suggests a "scale"
for each good $m_i(a), i = 1, \ldots, n$, so that the utility function becomes

$$u = u(q; a) = u(q_1/m_1, q_2/m_2, \ldots q_n/m_n)$$

(74)

with, once again, $m = 1$ for all $i$ for the reference household. It is clear
that the cost function corresponding to (74) must take form

$$c(u, p, a) = c^*(u, p_1 m_1, p_2 m_2, \ldots, p_n m_n)$$

(75)

in order that changes in a simply change the "effective" prices $p_i m_i$ for the
"effective" consumption levels $q_i/m_i$. Note that changes in family composition will
thus have both income effects (like (73) to which (75) reduces if $m_1 = m_2 =
\ldots = m_n$) and substitution effects (whiskey becomes cheap relative to soft
drinks when you acquire a family). In a single cross-section with prices con-
stant, (75) is like a normal cost function with the $m$'s replacing the $p$'s so
normal demand functions for $q$'s in terms of $x$, and the $m$'s (themselves some
functions of the $a$'s) can be fitted to give estimates of the parameters of the
cost function (75). Armed with these, equivalence scales are given by
However, it turns out that without some price variation (or some prior information) only \((n - 1)\) of the \(m\)'s can be identified. It is also true that with price variation, estimating the model is a considerable task. Muellbauer has also found that the one-for-one analogy with prices is not fully supported by the evidence, so that it is doubtful whether the extra complexity of the calculations can be fully justified for practical measurement.

An alternative procedure, superficially similar to that of Barten was proposed much earlier by Prais and Houthakker (1955). Like the Barten model, each good has a specific scale \(m_i\), and the demand functions are specified as

\[
q_i = \frac{m_i}{m_0} g_i(x/m_0)
\]

for some general scale \(m_0\). Although Prais and Houthakker did not base their model on utility theory, Muellbauer has shown that, provided \(m_0\) is interpreted as the general scale \((72)\), the Prais-Houthakker model is only consistent with the general formulation \((74)\) if preferences are non-homothetic Leontief, i.e. if all substitution effects are zero. This is essentially because \((77)\), unlike the demands from \((75)\), does not allow for the substitution effects which occur in the more general model as changes in family composition alter the effective prices of goods and services. Although there is no compelling reason why \((77)\) has to be consistent with utility theory, the substitution effects of the Barten model would be hard to deny in principle and it is undoubtedly to the disadvantage of the Prais-Houthakker formulation that it cannot accommodate them.
The Prais-Houthakker model has been relatively widely estimated. However, like the Barten model on a single cross-section, the Prais-Houthakker model is not fully identified (even with price variability). Nevertheless, a number of authors have still "estimated" the model, essentially by incorporating untested prior information; these estimates are worth very little. Muellbauer's experience with the model, incorporating the Leontief preferences, is not encouraging, and although Pollak and Wales (1978) obtain a much better showing, the \( m_0 \) in their model is not equal to the ratio of the two cost functions (72); hence, there are two distinct general scales implicit in their model rendering the interpretation unclear.

A number of other possibilities exist. For example Gorman (1976) has suggested that some fixed costs for each good \( d_i(a) \), be allowed, so that (75) is modified to

\[
c(u, p, a) = \sum_p p_k d_k(a) + c^*(u, p_{-1}, \ldots, p_n). \tag{78}
\]

Without the \( m \)'s in (78), this cost function is what Pollack and Wales call "demographic translating"; instead of, as in the Barten model, scaling the price for demographic change, the composition variables translate effective consumption from \( q_i \) to \( q_i - d_i(a) \). These models (and others) are far from having been fully worked-out as, yet and there is undoubtedly much more work still to be done. For example, there is also the issue of the functional form relating the \( m \)'s to the \( a \)'s and this has much effect on whether there are "economies of scale" within the household. These are crucial issues in welfare measurement and it is unfortunate that so few economists are working on them.

1/ Singh (1972), Singh and Nagar (1973) and McClements (1977).

2/ See the summary and discussion in Deaton and Muellbauer (1980), Chapter 8.
88. Finally, I must return to the general issue of whether this whole exercise is soundly based and whether it is possible, by observing the effects of compositional variables on behavior, to construct valid equivalence scales. There are two separate issues. First, children are usually viewed as resulting from a conscious and deliberate decision by the parents. Hence, although the equivalence scale approach might well be valid for measuring the consequences of a disease, for example, or even country of residence, it is not valid to implicitly compensate people for a deliberate choice (and one which presumably increased their welfare.) There are a number of difficulties with this argument. First, the world is an uncertain place and it is not possible to "dispose of" children as one would "dispose of" a mistakenly purchased durable good. Second, social welfare is concerned with the welfare of the children as well as with that of their parents and this is not considered in the argument. But I think the conclusive reply is that the welfare we are concerned to measure is short-term and economic; hence, by the same arguments I have used above, the non-economic benefits of parenthood ought not to be included in welfare measurement, while the economic costs of keeping children ought to be included if only because the state sees these as overriding committed costs even if the individual household does not. The second challenge to the validity of equivalence scales, similar to that discussed in the context of quantity constraints in paragraph 64, is that there is no reason to suppose that all the welfare consequences of changes in household composition will show up in the commodity demands, for example if there is a

1/ See for example Pollak and Wales (1979).
separable "children" branch in the utility function, so that no amount of demand analysis can construct valid welfare comparisons. An old man, living alone, may be overjoyed to discover that he has become a grandfather, but it may nevertheless not change his consumption pattern. Formally, if direct utility $u(q, a)$ is given by $f(a, u^*(q, a))$ so that the cost function takes the general form $c \{ g(u, a), a, \}$ then neither the parameters of $f(\cdot)$ nor $g(u, a)$ are empirically identified from the demands for $q$ (although they could conceivably be from fertility studies.)

As in the earlier analysis, my belief is that it is correct to exclude such effects from the measurement of economic welfare. Neither the function $f(a)$ or the function $g(u, a)$ have any effect on purchases of goods, and although they may well reflect the joys (or miseries) of parenthood, they are not part of the economic costs of children as reflected in the rest of welfare. One is, of course, free to disagree with the conception of welfare implied by such assumptions, but nevertheless I do not believe there is anything incorrect about basing welfare equivalence-scales on observable behavior.

D. Summary: Money Metric Utility Again

89. I wish to conclude this part by restating the form of the basic unit of welfare and its relationship to price indices and equivalence scales. Money metric utility, the minimum cost for a reference household of attaining a specified standard of living at reference prices is represented by welfare $c(u, p^R, a^R)$ (79) for reference prices $p^R$ and household characteristics $a^R$. If $u$ is the utility level of household $h$ which in fact has a total budget $x^h$, faces prices $p^h$ (in another country or time period) and composition $a^h$, then $u$ is given by the indirect utility function $*(x^h, p^h, a^h)$ so that welfare of $h$ is welfare of $h = c \{ *(x^h, p^h, a^h), p^R, a^R \}$ (80)
30. To illustrate, assume that preferences are once again those of the linear expenditure system and that household composition effects follow the Engel model so that \( m \) is the number of adult equivalents in \( h \). Then direct utility is

\[ u^h = \frac{(x^h - p^h \cdot y)}{m} \prod_k h^8_k \]  

(81)

i.e. the real value per equivalent adult of supernumerary expenditure. If the reference household has a single adult, money metric utility for household \( h \) would be

\[ p^R \cdot y + \left( \frac{\prod_k R^8_k}{\prod_k h^8_k} \right) \left( \frac{x^h - p^h \cdot y}{m^h} \right) \]  

(82)

which is subsistence expenditure at reference prices plus per capita supernumerary expenditure revalued at reference prices (using the supernumerary expenditure price index). Note that (82) involves only observable or estimatable magnitudes.

91. Note again the point made in paragraph 73, that if tastes between \( h \) and \( R \) differ in a way not allowed for by the characteristics \( a^h \) and \( a^R \), then (80) is the measure of the money needed by \( R \) to maintain itself at the welfare level it would actually reach given \( h \)'s circumstances. To make this a useful measure of welfare it is important that the reference household have an acceptably "typical" composition.

92. Since the expression (80) is equal to \( x^h \) if \( p^h = p^R \) and \( a^R = a^h \), it is useful to use this as a base for evaluation. Indeed, writing \( x^h \) for money metric utility, (80) can be identically rewritten as
\[ x^h_* = x^h - \frac{c(u^h, p^R, a^h)}{c(u^h, p^a, a^h)} \times \frac{c(u^h, p^R, a^R)}{c(u^h, p^a, a^h)} \]  

Now \( c(u^h, p^h, a^h)/c(u^h, p^R, a^h) \) is the true cost of living index at \( u^h \) for a household with characteristics \( a^h \), while \( c(u^h, p^R, a^h)/c(u^h, p^R, a^R) \) is the equivalence scale or number of equivalent adults evaluated at \( u^h \) and \( p^R \). Hence (83) gives money metric utility as total expenditure divided by a price index and an equivalence scale (both appropriately defined). Hence, the deflation idea for dealing with prices can be extended to deal with household composition.

93. The consumer surplus analogue of (83) is also straightforward; employing differences rather than ratios,

\[ x^h_* = x^h - \left[ \frac{c(u^h, p^h, a^h)}{c(u^h, p^R, a^h)} - \frac{c(u^h, p^R, a^h)}{c(u^h, p^R, a^h)} \right] \]

which is total expenditure less two compensating variations, first, the compensation needed to keep \( h \) on \( u^h \) given a price change from \( p^R \) to \( p^h \), and second, the compensation needed at \( p^R \) to restore \( R \) to \( u^h \) on its family composition having changed to \( a^h \). These formulas emphasise the "corrected expenditure" aspect of money metric utility once suitable price indices and equivalence scales have been evaluated.

IV. FROM INDIVIDUAL TO AGGREGATE WELFARE

A. Social Welfare Functions

94. The essential assumptions necessary for the construction of social welfare have already been made in the previous section and all that is required here is
their assembly into a suitable measure. The key assumption is that interpersonal comparisons can be made in terms of money metric utility so that social welfare is simply some aggregate or index based on the money metric utilities of each household. As in the case of individual welfare defined over a vector of goods, we can only define social welfare over the vector of individual welfares if we have some consistent ordering over possible allocations. As we shall see, this need not be explicit, but it is convenient to begin as if it were.

95. Social preferences are represented by a social welfare function \( V(\cdot) \), the arguments of which are the individual levels of money metric utility, \( x^1, x^2, \ldots, x^H \) for the \( H \) households concerned. Hence, if the level of welfare is \( W \), we write

\[
W = V(x^1, x^2, \ldots, x^H). \tag{85}
\]

I shall assume that the function \( V \) has the following properties:

(i) **linear homogeneity**: doubling each household's welfare doubles social welfare. This seems to be relatively harmless but it implies a separability in attitudes over the total and its allocation, so that inequality is viewed in the same light independently of the level of per capita expenditure. In practice, this is largely accepted for want of any better specific assumption.

(ii) **quasi-concavity**: a weighted average of any two allocations is always at least as good as either of the original allocations. This embodies the idea that greater equality is always good **per se**. Note carefully that greater equality may well have a cost in terms of less available resources for allocation and that quasi-concavity does not imply that perfect equality is even attainable, let alone desirable. Quasi-concavity is closely
related to Dalton's (1920) principle of transfers, that reallocations from better-off to worse-off consumers (which do not reverse their positions) are desirable.

(iii) Symmetry: Social welfare is unchanged by permutations of the (superscript) indices in (85). The same set of money metric utilities gives the same welfare, independently of which households have which welfare levels. Note the crucial dependence of this on the fact that we have \( x_+ \)'s in (85) and not \( x \)'s; the former are corrected for variations in household characteristics while the latter are not.

A fourth assumption additivity is sometimes made. This occurs if judgments between any two households can be made independently of the welfare position of any third household. As its name suggests, the assumption implies that (85) is a monotone increasing transform of a function additive in the \( x_+ \)'s. One useful additive social welfare function is the CES form

\[
W = \left\{ \frac{1}{H} \sum \frac{x_+^{1-\varepsilon}}{1-\varepsilon} \right\}^{1/1-\varepsilon}
\]

(86)
in which the parameter \( \varepsilon \) controls the degree of quasi-concavity, e.g. for \( \varepsilon = 0 \), the indifference curves are straight lines of slope \(-1\) while for \( \varepsilon = \infty \), \( W = \min x_+ \) and the indifference curves are right angles. Since, in this context, quasi-concavity reflects beliefs about the desirability of equality, \( \varepsilon \) is known as the degree of inequality aversion, see Atkinson (1970).
B. The Unit of Measurement

97. A suitable measure of social welfare based on (85) and the three assumptions can now be constructed in exactly the same way as was done for individual welfare in Part II. However, the money metric approach is not particularly attractive in this context. The individual $x_*$'s are not available at fixed prices subject to any overall constraint. In many discussions of income inequality, there is an implicit allocation model in the background wherein a fixed total of GDP (or consumption) is to be allocated to the individual households. Of course, no such constraint exists and attempts to alter distribution (e.g. by implementing a tax system) will alter behavior in such a way as to change the constraints. Indeed, the optimal tax literature contains quite general theorems on the impossibility of obtaining complete equality.

98. An alternative approach is to use the quantity-metric utility concepts of Part II; this presentation has recently been worked out in a paper by Blackorby and Donaldson (1978), although the basic concepts go back at least as far as Debreu's (1951) paper on the coefficient of resource utilization. The results themselves are essentially those first derived by Atkinson (1970) and Kolm (1969) (1976). The procedure is basically to choose the point of complete equality as reference and scale the social welfare function such that social welfare equals average per equivalent capita real income at this point; indeed if everyone possessed $x_*$, $x_*$ is the obvious measure of welfare. Actual welfare is measured relative to this by the point at which the actually attained social indifference curve cuts the vector of complete equality. If we write this vector $\bar{x}_*$ (for unit vector $1$), then according to paragraph 8
above, quantity metric utility referenced on \( \overline{x}_* \) is \( d \left\{ V(x_*), \overline{x}_* \right\}^{-1} \) for actual vector \( x_* \). Since this is social welfare taking \( \overline{x}_* \) as unity, and since this latter generates a welfare level of \( \overline{x}_* \), our proposed measure of social welfare is

\[
W^* = \overline{x}_* d \left\{ V(x_*), \overline{x}_* \right\}^{-1}.
\]

(87)

The next paragraph expands and interprets (87) algebraically, the subsequent one does so geometrically.

99. Recalling the definition of the distance function, i.e. \( V \left[ x_* / d(\overline{W}, x_*) \right] = \overline{W} \), the linear homogeneity of \( V \), assumption (i), implies that \( d(\overline{W}, x_*) = V(x_*) / \overline{W} \).

Hence, from (87), social welfare \( W^* \) can be written as

\[
W^* = \overline{x}_* \left\{ V(x_*), \overline{x}_* \right\} \left\{ V(\overline{x}_*) \right\}^{-1}.
\]

(88)

By the quasi-concavity and symmetry of \( V \), assumptions (ii) and (iii),

\( V(\overline{x}_*) \geq V(x_*) \) with the equality holding if, and only if, \( \overline{x}_* = x_* \), so that \( W^* \) is less than \( \overline{x}_* \) unless there is perfect equality at which point it attains its maximum of \( \overline{x}_* \). It is thus natural to define equality by the ratio

\[
E = \frac{V(x_*)}{V(\overline{x}_*)} = d \left\{ V(x_*), \overline{x}_* \right\}^{-1}
\]

(89)

This magnitude clearly satisfies \( 0 \leq E \leq 1 \) so that inequality, \( I \), may be defined by

\[
I = 1 - E
\]

(90)

Note that our earlier measure of quantity metric utility \( d(u(q), q)^{-1} \) is now serving as an equality measure. Equality is measured by how actual welfare compares with what it would be if perfect equality were attained (even if this state is a purely imaginary one).
100. Figure 5 illustrates the (probably) more familiar geometry. The actual distribution is \((x^1_*, x^2_*)\) at B; by symmetry this lies on the same indifference curve as C which is \((x^2_*, x^1_*)\): By quasi-concavity, points of more equal distribution, i.e. along BC, are preferred. The distance function value \(d\{V(x^1_*, x^2_*), x^*_*, x^*_*\}\) is the amount by which CE has to be shrunk to get it on to the actual welfare curve; this is CE/OA. Hence, equality is the ratio \(OA/CE\) (NB \(0 \leq OA/CE \leq 1\)) and inequality is \(1 - OA/CE = AE/OA\) in accord with Atkinson's original definition. Note that the equality and inequality measures are determined not only by the position of the vector \(x_*\), but also by the shape of the social indifference curve. Some writers feel that this dependence of "objective" measurement on "subjective" criteria (such as the shape of social welfare functions) is inappropriate, but it is surely no less appropriate than basing the measurement of the welfare of an individual consumer on his preference ordering. The reverse position is much closer to the truth, that the use of apparently "objective" measures of equality or inequality conceals hidden welfare judgments.

![Figure 5](image_url)

**Figure 5**

*Equality and inequality*
101. The preceding analysis can easily be reversed so that inequality measures can be used to generate social welfare functions. For example, the gini coefficient \( \gamma \), defined over real per equivalent capita expenditures \( x^h_\ast \) takes the form

\[
\gamma = 1 + \frac{1}{H} - \frac{2}{H} \sum_{x^h_\ast} \left( x^1_\ast + 2x^2_\ast + \ldots + Hx^H_\ast \right)
\]

(91)

where the \( x^1_\ast \)'s are arrayed in descending order of magnitude, i.e. \( x^1_\ast \geq x^2_\ast \geq \ldots \geq x^H_\ast \) (it does not matter how ties are resolved.) Checking first on its limits, ultimate inequality is reached when \( x^1_\ast = Hx^H_\ast \) so the \( \gamma = 1 - 1/H \) which for large \( H \) is approximately 1. Equality is \( x^h_\ast = x^\ast_\ast \) which gives \( 1 + H^{-1} = 2 \frac{H(H + 1)}{H^2} = 0 \), so that the coefficient lies within the correct theoretical limits without further adjustment. We can thus take \( E = 1 - \gamma \) and, by (87), the social welfare function is

\[
W = \bar{x}_\ast(1 - \gamma) = \frac{H}{1} \left[ 2\rho(h) - 1 \right] x^h_\ast
\]

(92)

where \( \rho(h) \) is the rank of \( x^h_\ast \) in the distribution (richest first). Such rank-order social welfare functions have the (perhaps unattractive property) that the social marginal utility of giving an extra dollar to \( h \) is a discontinuous function of \( x^h_\ast \). Other inequality indices can be given a social welfare interpretation in exactly this way.

C. Inequality Measure

102. Hence, although the approach of this section mirrors the earlier analysis in allowing social welfare to be measured by modifying real per capita expenditure, now by an equality index, we have little direct guidance on what functional form the index should take. All indices based on a social welfare

1/ See, for example, Blackorby and Donaldson (1978) or Deaton and Muellbauer (1980), Chapter 9, for further examples.
function satisfying the three assumptions (or any index satisfying the principle of transfers) will rank distributions in the same way if the corresponding Lorenz curves do not cross. If the Lorenz curves do cross however, different indices will give different rankings depending on the implicit weights in the underlying social welfare function. There is a very considerable literature discussing the merits and demerits of specific inequality indices and specific properties of indices. It is not clear to me that this discussion has been very decisive, certainly not to the point of demonstrating the superiority of any one specific index or even of having suggested new axioms which have a general appeal. 1/ Kolm's (1969) (1976) contributions are at the root of much subsequent work (although many of his ideas seem far-fetched). Among the issues discussed the following are perhaps important enough to be mentioned.

(a) Specific measures can be ruled-out because they are inconsistent with the basic assumptions (i.e.) (i) - (iii) or the principle of transfers). The relative mean deviation and the variance of logarithms can be disposed of in this way.

(b) Does linear homogeneity impart a "rightist" bias to inequality measurement? Perhaps we should say that the distribution is unchanged if everyone gets an equal addition to their income. (See Kolm for discussion).

(c) How do the various measures stand up to various practical problems, e.g. in dealing with negative incomes for a specific period? Note that this problem does not arise with an expenditure based index.

(d) Some measures can be decomposed so that overall inequality can be ascribed,

---

for example, to intra- and inter-group inequalities. This is particularly useful in underdeveloped countries where the population may fall into quite distinct groups, e.g. racial groups, or distinct urban/rural groups. We also may wish to disaggregate by individual or household characteristics, for example in assessing the impact on inequality of changes in demographic composition.

Papers by Bhattacharya and Mahalanobis (1967) and by Pyatt (1976) show how the (1967, Chapter 4) inequality measure also has attractive decomposition properties. 1/

103. One theoretical property of inequality indices which does seem to me to be a desirable one is absence of what Blackorby and Donaldson (1978), call distributional homotheticity. A social welfare function is said to the distributionally homothetic if its contours of intersection with the fixed sum hyperplane \( \sum x_i = \text{constant} \) are concentric around the point of complete equality. This concept is only operative if there are at least three consumers; note that with two consumers, Lorenz curves cannot cross and there are few interesting questions about inequality measures.

Figure 6 illustrates the fixed sum hyperplane for three consumers together with its midpoint \( E \), the point of complete equality; this diagram corresponds to a three-dimensional version of Figure 5 viewed from the northeast corner. An initial allocation is the point \( Z \); by symmetry, PQRX and Y must all be on the same welfare contour. Apart from this, different welfare functions will trace out different equal-welfare contours around the point \( E \) (where a social welfare contour just touches the simplex from above). The case of sole concern with the poorest (a maximin or Rawlsian social welfare function) gives contours concentric to the

---

1/ See also the paper by Paglin (1975), and the criticisms in Danziger et al (1977).
triangle ABC. Along each of the sides of the triangle the welfare of the poorest is constant; at a vertex, the name of the poorest changes. Other social welfare functions, which accept some loss to the poorest if the gain is large enough elsewhere, will lie outside the triangle along ZY, XR and PQ and inside it on YX, RQ and PZ. Blackorby and Donaldson argue, in my view convincingly, that the social welfare function should be such that, as we move further from E, the contours should become more triangular so that the further society is from equality, the more it should be concerned with the welfare of the poorest. Now the contours of the gini coefficient are concentric hexagons around E so that the social welfare function (92) is distributionally homothetic and the gini does not have the required property. Similarly for the coefficient of variation. However, inequality measures based on the CES form (86) have the desired property as does Theil's entropy coefficient.
I note finally that the idea of writing individual welfare as money expenditure deflated by a price index and an equivalence scale can be extended to the various components of social welfare as discussed in this part. From (87) and (89), social welfare is written

$$W^* = \bar{x}_* E$$

and we can use a decomposition corresponding to (83) to break out the contributions of price and demographic change from each of these components.

Taking $\bar{x}_*$ first, and with period 1 as current and period 0 (or reference characteristics $a^0$) as base, we have

$$\bar{x}_* = \bar{c}(u^{1h}, p^0, a^0)$$

$$= \frac{1}{H} \sum_h \left[ \bar{c}(u^{1h}, p^0, a^h) \right]$$

$$= \bar{x}_1 / \left\{ \sum_h \bar{c}(p^0; u^{1h}, a^h) \right\}$$

where

$$\bar{c}(u^{1h}, p^0, a^0) = \frac{c(u^{1h}, p^0, a^h)}{\bar{c}(u^{1h}, p^0, a^0)}$$

Hence, (94) gives $\bar{x}_*$ as average money expenditure $\bar{x}_1$ deflated by a weighted average of individual current-weighted true cost-of-living indices and also by a weighted average of individual equivalence scales. The weights in each case add to unity and correspond to expenditure shares for each household in some hypothetical configuration of expenditures and prices. In practice, one
could hope that these weights can adequately be approximated by actual distributional shares. A similar device can be used to analyze changes in E from 0 to 1. From (89)

\[ \frac{E^1}{E^0} = \frac{V\left(\frac{x^1}{x^*_0}\right)}{V\left(\frac{x^0}{x^*_0}\right)} = \frac{v(x^1_*)}{v(x^0_*)} \quad (95) \]

in the case where \( x^1_* = x^0_* \) (for convenience of presentation) and where, since 0 is the base period, \( x^h_0 = x^h \). Using (83) to write \( x^1_* = x^h / m^h \cdot P^h \), (95) can be written

\[ \frac{E^1}{E^0} = \frac{v(x^1)}{v(x^0)} \cdot \frac{v(x^1/m^h P^h)}{v(x^0/m^h P^0)} \quad (96) \]

The first term is the contribution due to changes in money expenditures, the second that due to changes in demographic structure and the third that due to changes in prices. Note however, that these decompositions are rarely unique (e.g. one can 'reprice' for prices first or for demographics first) and considerable more experience is needed with them in practical application. Again, see Muellbauer (1978) for an example.
List of Works Cited


__________, (1976), "What Can We Deduce About Welfare from Behaviour?", mimeo.


The World Bank
Publications Order Form

SEND TO: YOUR LOCAL DISTRIBUTOR OR TO WORLD BANK PUBLICATIONS
\(\text{See the other side of this form.}\)
P.O. BOX 37525
WASHINGTON, D.C. 20013 U.S.A.

Date ____________________________

Name ____________________________________

Title ____________________________________

Firm ____________________________________

Address __________________________________

City _______ State _______ Postal Code ______

Country __________ Telephone (_____) __________

Ship to: (Enter if different from purchaser)

Name ____________________________________

Title ____________________________________

Firm ____________________________________

Address __________________________________

City _______ State _______ Postal Code ______

Country __________ Telephone (_____) __________

Check your method of payment.
Enclosed is my ☐ Check ☐ International Money Order ☐ Unesco Coupons ☐ International Postal Coupon.
Make payable to World Bank Publications for U.S. dollars unless you are ordering from your local distributor.

Charge my ☐ VISA ☐ MasterCard ☐ American Express ☐ Choice. (Credit cards accepted only for orders addressed to World Bank Publications.)

Credit Card Account Number ____________________________
Expiration Date ____________________________
Signature ____________________________

☐ Invoice me and please reference my Purchase Order No. ________________

Please ship me the items listed below.

<table>
<thead>
<tr>
<th>Stock Number</th>
<th>Author/Title</th>
<th>Customer Internal Routing Code</th>
<th>Quantity</th>
<th>Unit Price</th>
<th>Total Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All prices subject to change. Prices may vary by country. Allow 6-8 weeks for delivery.

Subtotal Cost $________

Total copies _______, Air mail surcharge if desired ($2.00 each) $________

Postage and handling for more than two complimentary items ($2.00 each) $________

Total $________

Thank you for your order.