# THE MEASUREMENT OF CHANGES IN QUALITY

BY KELVIN LANCASTER

*Columbia University*

This paper is concerned with the theoretical problems of devising indexes of quality change and with some of the practical problems of constructing such indexes from market data, relating these to the various attempts to construct such indexes in the past. The general conclusion is that, while quality is inherently ordinal, there are three different indexes which might be taken as "measures" of quality change. If changes are sufficiently small, the values of all three indexes will coincide and then, only, can we consider any one of them to be an unambiguous measure of the change.

Although there had been some study of the problem of measuring quality change prior to 1961,[1] the modern work on the problem can be considered to date from the papers by Adelman and Griliches, and by Griliches, published in that year. In the succeeding fifteen years, a considerable body of work on the subject has been developed.[2] An important share of this work was carried out, or commissioned, by the Price Statistics Committee of the Federal Reserve Board, and the general emphasis of much of the work has been towards developing methods for *adjusting for* quality change in cost-of-living indexes, rather than towards measuring quality change as such. Most of the earlier work was concerned with practical problems of econometrics within a relatively simple theoretical framework, but more recent contributions, especially Fisher and Shell (1967) and Muellbauer (1974), have probed more deeply into the theoretical foundations.

The purpose of the present paper is to investigate the problem of measuring quality change as an objective in itself, not specifically related to such other problems as that of developing a "true" cost-of-living index. The author proposes to pursue his normal role as a theorist, concerning himself with keeping everyone honest by pointing out all the difficulties while trying to give some judgment as to the practicality of those techniques which appear to have a firm conceptual basis.

## 1. WHAT IS MEANT BY "MEASURING" QUALITY?

We shall take the initial step of assuming that the ordering of comparable varieties of goods in terms of quality is a well-defined operation for an individual consumer, who can unambiguously rank one variety as of higher or lower quality than another. However, we can also presume than an individual would be unable to answer the question, "By how much is the quality of one variety greater than that of the other?", and thus must regard perceived quality as inherently ordinal, providing no natural basis for a measure. Any numbers we choose to associate

[1]For example, Court (1939), Houthakker (1952), Stone (1956).

[2]A survey of the first ten years is given in Griliches (1971) and a bibliography in Griliches (ed.) (1971). See also a review of this volume by Muellbauer (1972).

with quality or quality change must correctly represent the quality ordering, whenever that ordering is well defined, but the cardinal significance of the numbers can be chosen to be a measure of anything that has the same ordering as quality. It is the purpose of this paper to examine the possibilities for such a measure, some obvious candidates being the prices that would be paid for goods of different quality when all were available at the same time, the amount of a lower quality good considered to be "as good as" a unit of the higher quality good, and some index number of the quantities of the various characteristics of the goods.

What about quality as quantity, the simple repackaging case? If a box of crackers or a candy bar contains 10 percent less content this year than last, it has been the practice among economists to refer to this as a quality change. If the term is appropriate, the quality change would be truly measurable. But it would seem likely that a hypothetical consumer would regard the change, correctly enough, as a mere change in the quantity units having nothing to do with quality in the proper sense.[3]

Thus we conclude that quality is an ordinal concept with no inherent cardinality, and the terms "measuring quality" or "measuring quality change" refer to the construction of some numerical index with the following properties:

(1) that it correctly *ranks* goods varieties in terms of quality whenever an unambiguous ranking exists,

(2) that it *measures* something well-defined and useful in the context to which the index is applied.

From these considerations it is clear that there may be many indexes which rank quality, each of which measures something different and is applicable to its own special use.

## 2. Indexes of Relative Quality I: Divisible Goods

Before attempting to tackle quality *change* which most often occurs in some context in which goods available in one period are replaced by goods of different specification in the next period, we shall first investigate the conceptually simpler problem of assigning quality indexes to different varieties of a good when all varieties are available simultaneously. In the simultaneous case, the market forces consumers (and producers) to compare the varieties with each other under circumstances which directly affect choice and do not involve hypothetical comparisons or possible preference changes.

We shall carry out all analysis in this paper in terms of characteristics,[4] assuming that the objects of consumer preference are collections of characteristics and that goods are purchased only to provide characteristics and not for their own sake. We shall also assume that the goods between which quality comparisons are made are members of a separable group,[5] sharing none of their characteristics

[3]Attempts to treat quality change in indivisible goods by a repackaging approach lack even the merit of a well-defined quantitative variable.

[4]See Lancaster (1966, 1971).

[5]See Lancaster (1971), Chapter 8.

with members of any other group, and that the relevant utility functions are separable in the characteristics of the group.

If consumers have identical preferences and identical income levels, and if the goods in the group have a clearly identifiable natural unit common to all the goods (such as an ounce or a pint), then we can take identical qualities of all the goods and see how the characteristics collections associated with various goods are ranked in terms of preference. This ranking gives the quality ranking.

Even in this highly simplified context, problems arise. One is that the quality rankings need not be independent of the quantities chosen for the comparison, unless either preferences are homothetic or higher quality goods do not ever have less of any characteristic per unit than do lower quality goods.

Let us assume homotheticity and identical preferences (identical incomes are not required) so that there is a universally agreed quality ranking, and concentrate on the construction of a suitable numerical index. Since the goods are divisible, the most obvious index is one based on the relative quantities of the various goods required to attain the same utility level as some specified amount of any one of the goods chosen as numeraire. The inverse of these quantities would give an appropriate index, and we can call such a measure a *quality-equivalent* index.

For uniform homothetic preferences, a quantity-equivalent index of quality has the following properties, all desirable:

(1) Its quality ranking will coincide with the true and unambiguous ranking.

(2) It will be independent of the amount of the numeraire good chosen.

(3) Recomputation of the index to make the value for a non-numeraire good equal to 100 will give the same results as would making this good the new numeraire.

(4) The value of the index depends only on the technical properties of the goods and the shape of any one indifference curve, and is independent of the properties of the market.

(5) For goods having all characteristics in the same proportions, the value of the index will coincide with that obtained by treating the difference between the goods as due to simple repackaging.

Note that, although the quantity-equivalent index is independent of market conditions, the market will reflect the content of the index. Under the circumstances assumed above (uniform homothetic preferences), the relative prices of the goods must be directly proportional to their quality indices[6] if all goods are available simultaneously and are all actually purchased, provided the goods are not used in combination. If the goods are combined, the relative costs of the combinations must reflect the qualities of the combinations but the prices of individual goods need not reflect their quality indices.

If preferences are not uniform or homothetic, a quantity-equivalent index can be constructed for a reference group of consumers at a specific income level. The index will not be relevant to other consumer groups, and even the quality ranking may not hold outside the reference group.

[6] A special problem arises if there are no natural units, such as detergents which may be powdered or liquid, or of different degrees of concentration. One should convert to equivalent units, such as the "dose" required per load of washing.

As a practical matter, we cannot really expect to possess information as to the exact shape of the typical indifference curve (or surface) of the reference group. It is more reasonable to expect that we can isolate and measure the leading characteristics of the goods, so an obvious alternative to a true quantity-equivalent index would be some kind of fixed weight index number of quantities of characteristics, constructed in the same way a quantity index of a bundle of goods is constructed. If we were to construct such an index, how would its properties compare with those of the quantity-equivalent index?

The essential difference between the true quantity-equivalent index and a fixed weight index number is that the latter takes as equivalent all combinations of characteristics corresponding to points on the same hyperplane, the former combinations corresponding to points on the same indifference surface. It is obvious that, if preferences are strictly convex, the fixed weight number will *overstate* the quality of all goods except that corresponding to the point of tangency between the index hyperplane and an indifference surface, as compared with the quantity-equivalent index. It is easy to construct examples in which the *ranking* of goods in terms of quality is not even the same in the two cases. The degree of divergence between the two measures will increase with increased convexity of the indifference surface and with increased dispersion of the characteristics proportions of the goods in the group.

The slope of the index hyperplane is, of course, determined by the choice of weights for the index, these in turn being shadow prices allocated to the various characteristics.[7] If the characteristics proportions of the various goods are clustered together and/or the indifference surface has little curvature, the fixed weight index will be relatively close to the true quantity-equivalent index if the slope of the index hyperplane is close to the slope of the indifference surface for a good somewhat near the center of the cluster. In this case, it is clearly worthwhile attempting to obtain the best possible estimates for the true shadow prices on characteristics, in order to use these for the weights. But if there is likely to be considerable divergence between the fixed weight index and the true index, we can only regard the former as a roughly approximate measure and undue concern with estimating shadow prices is not only a waste of resources but tends to give a spurious impression of accuracy in the final index.

## 3. Indexes of Relative Quality II: Nondivisible Goods

Historically, most of the attempts to measure or adjust for quality differences have been carried out on such essentially indivisible goods as automobiles and refrigerators. There are no problems in determining appropriate units for such goods, but we cannot compare different quantities of different goods as in constructing a quantity-equivalent index.

Banking goods in terms of quality presents no more and no less of a problem than in the divisible goods case—if anything, perhaps less, since we will always compare single units and problems of combinability do not arise. A good ranks higher in quality than another if the characteristics collection asscoiated with a

[7]See Section 4 for the theoretical determination of the relevant shadow prices, and Section 6 for the practical problems of estimating these.

unit of the good is preferred to the collection associated with the other, at least by the reference consumer group. We can presume that a Mercedes woud be ranked higher in quality than a Volkswagen, but can we associate a numerical index with the quality that conveys any useful information other than the relative ranking?

Since one consumer can use only one car at a time, the characteristics of six Volkswagens are not markedly different from those of one except for increased reliability (they will surely not all fail to start at the same time) and durability, although the ability to provide cars for the family and friends may provide increased utility. Thus someone who really likes the characteristics mix of a Mercedes may prefer one to twenty Volkswagens without implying that he considered the Mercedes to have at least twenty times as much "quality".[8]

One way of attempting to solve the indivisibility problem is to introduce hypothetical divisibility. In the Mercedes–Volkswagen example, we take the Volkswagen as quality numeraire and consider a hypothetical good having its characteristics in the same *proportions* as the Mercedes, but at a scale that makes the consumer indifferent between this hypothetical good (a "Merc-wagen") and the Volkswagen. The ratio of the characteristics of the Mercedes to those of the Merc-wagen is then a potential quality index, which we shall call a *hypothetical quantity-equivalent index.*[9]

Since the comparison on which the hypothetical quantity-equivalent index is based does not correspond to any real choice situation, no real meaning can be ascribed to the numerical value of the index, as opposed to the implied ordering. In particular, there is no reason why the numerical values should be reflected by the market under any circumstances. Let us suppose, purely for the sake of illustration, that a standard Mercedes contains 1.8 times as much of every characteristic as a hypothetical "Merc-wagen" with characteristics in the same proportions as the Mercedes but which would be ranked equivalent to the Volkswagen. The only market predictions we could make are that *if* the Merc-wagen actually existed, it would have to sell at the same price as the Volkswagen (and then only if there were uniform preferences over all consumers), and that the Mercedes must sell for a higher price than the Volkswagen. The Mercedes might actually sell at a price 20 percent higher than that of a Volkswagen or it might sell at ten times the price, either being consistent with the sale of both cars and with uniform preferences. There is absolutely no reason for expecting the price ratio to be 1.8, the numerical value of the quality index.

The lack of relationship between market price ratios and the value of the hypothetical quantity-equivalent index is, of course, due to the necessity of comparing goods at different expenditure levels for the group under study (e.g., automobiles) and thus involving comparison with expenditure alternatives outside the group. A consumer will be indifferent between a Volkswagen at $3,000 and a Mercedes at $12,000 if and only if he is indifferent between a Volkswagen plus $(Y-3,000) of other things and Mercedes plus $(Y-12,000) of other things, where Y is his total budget. He may well be indifferent between these two collections even though the quality index of the Mercedes is only 1.8 times that of

[8]The customer is, of course, prohibited from trading any of these Volkswagens, directly or indirectly.

[9]This is the index actually sought in most applications of the so-called hedonic technique.

the Volkswagen on a hypothetical quantity-of-characteristics comparison. It all depends on how he rates the enjoyment of other things as against more quality in an automobile. Separability of his utility function is not sufficient to avoid this comparison with expenditures outside the group.

We might try to construct an *expenditure-equivalent* index[10] of quality by conducting the following imaginary experiment. Take a particular consumer with a given budget, choose a specific automobile as numeraire (let us say a Volkswagen), and take the price of the Volkswagen and of all goods except other automobiles as given. The consumer is now asked to buy a Volkswagen, spend the rest of his budget optimally, and not his preference level. We then ask him to return to scratch and calculate how much he would be willing to pay for a Mercedes (or any other variety of car) in place of the Volkswagen under the following conditions:

(1) His total budget is to be the same as before.

(2) He is to achieve the same level of preference as before after optimal allocation of the remainder of his budget.

With normal assumptions on preferences there will be, in principle, a unique solution for every type of automobile which, when divided by the price of the numeraire, will give our expenditure-equivalent index. This index will obviously conform to the true quality ranking and the numbers convey information relevant to a real choice situation. As with the quantity-equivalent index, the numbers are applicable only to a particular consumer or group of consumers with identical preferences and may depend on the choice of numeraire. In this case the index depends, in addition, on the budget level chosen, the price of the numeraire good and the prices of all other goods except nonnumeraire members of the group of goods being indexed.

## 4. Indexes of Quality Change

Constructing an index of quality change involves the same problems as constructing an index of relative quality, with the addition of problems from two new areas of difficulty. The first is that the new and old goods may not be (and typically are not) available simultaneously, so that comparisons are hypothetical and not real. The second is that preferences, incomes and prices may change at the same time as the specifications of the goods.

Complications due to preference changes will not be taken up in this paper,[11] since these present the same kind of problem whether there is quality change or not, so we will assume that each consumer maintains the same preferences over the period in which quality change is being investigated. Since we are taking preferences to be defined over characteristics collections, the introduction of new goods varieties does not require new preferences.

In the case of divisible goods, quality change with constant preferences can be handled in essentially the same way as relative quality. Even if the old variety of a

[10]This is the index implicitly being sought in Cagan (1965) and other studies based on used-goods markets; and also the index implicit in Fisher and Shell (1968). See also Rosen (1974).

[11]Discussion of quality change along with preference change is given in Fisher and Shell (1968) and Muellbauer (1975).

good is replaced by a new variety, so that the two varieties are not available at the same time, we can reconstruct the characteristics collection that would have been provided by the old good at the time the new good is available and construct a quantity-equivalent index of the new good in terms of the old, from which we directly obtain the quality change as measured in a well-defined way. In most contexts in which we are likely to wish to measure quality change over a short period, we will be comparing new and old models or varieties of what would be identified as essentially the same good. This implies that we are making comparisons between new and old varieties having similar characteristics proportions and thus that a fixed-weight index of characteristics with appropriate weights will be a close approximation to the true quantity-equivalent index.

With indivisibility, we can base an index of quality change on either the hypothetical quality-equivalent index or the expenditure-equivalent index. In discussing relative quality, we showed that these indexes will be different in general, and that only the expenditure-equivalent index corresponds to a real choice situation or is likely to be reflected by the market. For measuring quality *change*, however, the potential gap between the two indexes in both numerical values and operational foundations is considerably reduced for several reasons.

In the standard context of model replacement, the comparisons between old and new models will be hypothetical in any case, reducing the operational realism of the expenditure-equivalent index. In addition, the dependence of this index on prices in general requires that adjustments be made for price changes among goods outside the group, giving another hypothetical comparison. On the other hand, if the quality change being assessed is that between successive models of the "same" automobile (or other typical nondivisible good), the changes in both the characteristics proportions *and* the absolute quantities of characteristics will be relatively small. Under these circumstances, the marginal rate of substitution between a fixed-weight aggregate of characteristics of goods in the quality-changing group and an aggregate of other characteristics can be taken to be constant. Thus for small quality changes of this kind, the expenditure-equivalent index, the hypothetical quantity-equivalent index, and the fixed-weight characteristics index will be approximately equal, all three converging as the quality change becomes sufficiently small.

To prove this last statement, consider a model of consumption that satisfies the conditions of our analysis, namely that the set of goods for which we are indexing quality forms a separable group. This requires that the characteristics of the group are not possessed by any goods outside the group and that the utility function is separable. We shall write the latter in the form

$$U = U[u(z), v(Z)]$$

where $z$ is the $m$-vector of characteristics possessed by goods within the group and $Z$ is the scalar representing the aggregate of characteristics from other goods, relative prices being constant outside the group. We shall suppose other goods to be normalized with respect to prices, so that the cost of the unit aggregate is unity.

First consider the expenditure-equivalent index. Initially, the consumer is constrained to consume one unit of the indivisible group good with specification

$z = b$, for which he must pay $P$. His utility is then given by

$$U = U[u(b), v(Y-P)]$$

where $Y$ is his total budget. This is fully determined, given that the aggregate $Z$ is already optimal over the prices of non-group goods, and no optimization is involved.

Now suppose the specification of the group good is changed to $z = b + db$, with the expenditure on the good changed to $P + dP$, where $dP$ is to be chosen so leave utility unchanged. We have

$$dU = U_1 \Sigma u_i \, db_i - U_2 v' \, dP = 0$$

giving

$$dP = \frac{U_1}{U_2 v'} \Sigma u_i \, db_i.$$

The measure of quality change in terms of the expenditure-equivalent index is given by $dP/P$. To complete our analysis, we need to express $P$ in terms of the same parameters as $dP$. This can be done by finding the set of shadow prices on characteristics such that, if the consumer was charged these shadow prices and given a budget $P$ for buying characteristics within the group, he would buy exactly the collection of characteristics given by $z = b$. If we denote the vector of these shadow prices by $y$, they must be such as to give the solution $z = b$, $yz = P$, to the problem:

$$\text{Max } U[u(z), v(Z)], \quad \text{subject to} \quad yz + Z = Y.$$

The first order conditions of the optimization give

$$U_1 u_i = \lambda y_i \qquad i = 1, \ldots, m$$
$$U_2 v' = \lambda$$

so that $y_i = U_1 u_i / U_2 v'$. Substituting in $yb = P$, we obtain

$$P = \frac{U_1}{U_2 v'} \Sigma u_i b_i$$

and the expenditure-equivalent index of quality change is given by:

$$\frac{dP}{P} = \frac{\Sigma u_i \, db_i}{\Sigma u_i b_i}.$$

Now let us turn to calculation of the hypothetical quantity-equivalent index. We shall make the comparison between the old variety with specification $z = b$ and a hypothetical variety with specification $z = (1 + d\beta)b$, where $d\beta$ is a scalar chosen to that the utility is the same with the hypothetical good as the actual new variety of specification $z = b + db$. If $u$ is the original utility (in the group) and $u + du$ the utility after the quality change, we must have,

$$du \text{ (actual)} = du \text{ (hypothetical)}$$

that is

$$\Sigma u_i \, db_i = d\beta \Sigma u_i b_i.$$

The measure of quality change in terms of the hypothetical quantity-equivalent index is given by the value of $d\beta$, that is, by the expression

$$\frac{\Sigma u_i \, db_i}{\Sigma u_i b_i}$$

which is identical with the value given by the expenditure-equivalent index.

Finally, we note that the expression for the measure of quality change is identical with that given by a fixed-weight index of characteristics with weights proportional to the $u_i$'s, giving the third equivalent measure.

Thus an unambiguous measure of quality change, in the sense of a number which has the same value in all three potential indexes, ranks the qualities correctly, and can be given an operational meaning, under the following conditions:

(1) The change in characteristics per unit of the good should be small (but characteristics need not change in the same ratio).

(2) We are concerned with a single good which changes in specification and which is consumed by itself without cooperating inputs or combinations with other goods.

(3) The good belongs to a separable group, one that is separable both in terms of the consumption technology (having no characteristics also possessed by goods outside the group) and the utility function.[12]

(4) The quality change is measured from the point of view of a single consumer (or a group of consumers with identical preferences) whose tastes and income do not change.

(5) The prices of goods outside the group do not change, nor is there quality change anywhere except for the good under investigation.

Condition (1) is not unduly restrictive for measuring typical year-to-year changes, but implies that long-period changes should be estimated by chaining and not by constant base-period weights. Condition (5) is the most certain to be violated, since it will be normal to expect price and/or quality changes over many goods in a typical year-to-year comparison. The problems are similar to those encountered in many other contexts in which one change is to be measured when many are taking place. If all changes are small and cross-effects can be neglected, the individual contributions to the total change can be treated as additive—otherwise, as in mixed price-quantity changes, there is no "true" measure of the individual components.

The remaining conditions are technical. Condition (3) is simply a tighter version of the structure we implicitly assume in our ordinary consumption decisions while (2) and (4) are concerned with definitional context of our measurement.

[12]See Fisher and Shell (1968) for discussion of the non-separable case along rather different lines from this paper.

## 5. The Quality of What?

So far, we have concerned ourselves with the measurement of quality change for a well-defined single good with reference to a well-defined consumer or consumer single group, and implicitly assumed that the good was not used in combination with other goods or inputs. We shall now explore some of the problems which arise when we relax these assumptions.

Consider the case of two divisible and linearly combinable goods in the same group which are consumed in combination by the reference consumer, the vectors of characteristics per unit for the goods being $b_1$, $b_2$. These are consumed in a linear combination with quantities $km_1$, $(1-k)m_2$, where $m_i$ is the budget for the group divided by the price of the $i$th good, so that the characteristics combination is given by

$$z = km_1b_1 + (1-k)m_2b_2.$$

Now suppose that there is a quality change in the first good only, its unit specification becoming $b_1 + db_1$. The measure of quality change for the good in isolation will be given by $wdb_1/wb_1$, where $w$ is the vector of appropriate shadow prices, while the measure of quality change for the combination of goods will be given by $wdb_1/w[b_1 + [[(1-k)m_2]/km_1]b_2]$, smaller than the former expression so long as $k < 1$.

Which is the appropriate measure, the quality change for the good treated as if it were consumed in isolation (a hypothetical comparison) or considered in combination (a real comparison)? It will depend on the context in which the measure of quality change is to be used. If for example, the measure is to be used to adjust a price index for quality change, the single-good value will be appropriate if the price index is based on single-good prices, the combination value if the price index is based on composite prices of goods combinations.

A kind of inverse problem arises in the common practice of setting out to measure the quality change of, say, automobiles as a group, rather than the quality change in a particular automobile.[13] In this case we are making comparisons between combinations of goods, although they are consumed separately. The only justification for aggregation would seem to be on an implicit assumption that the "true" quality change measures are really the same for all automobiles, and that averaging is used to minimize purely statistical errors.

A deeper problem arises when the good in which we are interested requires cooperating inputs of a kind that breaks down the assumption of separability. One of the more interesting cases falling into this classification is the need for time as an input from the consumer, so that use of the particular good is doubly linked to the rest of his consumption, through both time expenditure and money expenditure.

If time is needed for the use of an indivisible good, like an automobile, and if a change in specifications does not affect either the time per unit of "use" or the intensity of utilization,[14] the preceding analysis holds unchanged and we can

[13]The classic studies, such as Griliches (1961), and official statistics attempt to measure this average change.

[14]Note that we have been short-cutting here by taking the characteristics of the durables themselves, rather than the characteristics of the services of the durables, as the arguments of utility functions. The general conclusions are not affected by the more correct analysis in terms of service flows, provided the same separability conditions hold. See, however, Fisher and Shell (1968) for examples suggesting that separability is somewhat less realistic when the analysis is cast in terms of services, and Muellbauer (1974) for a discussion of quality change in a household production context.

construct an expenditure-equivalent index in the standard way. But if the change affects time per unit of utilization (greater reliability may reduce the ratio of total time devoted to the car to time spent in use of the car for its desired purpose), or the degree of use (the car may be easier to drive, leading to more driving per year), there is a reduction in time available for other consumption activities. This may make the original collection of characteristics from goods other than automobiles no longer optimal, and perhaps even technically infeasible, contrary to the assumptions underlying the expenditure-equivalent index.

Let us simplify the problem somewhat by assuming that the primary variable use of time, other than time consumed by activities involving goods in the special group, is in pure leisure which is a single-characteristic good requiring no cooperating inputs. Then any variation in time required for goods in the group is reflected in changes in leisure time only and not in the mix of other characteristics. We shall suppose, as before, that the group is separable with respect to all characteristics other than time. Not let us consider the expenditure-equivalent index for a quality change involving time as well as other characteristics.

For an initial specification $z = b$, $t = t$ of a good sold at price $P$, utility is given by $U = U[u(z, t), v(Y-P, T-t)]$ where $Y$ is income and $T$ the time available for use of the group good plus leisure. If the specification of the good changes to $b + db$, $t + dt$ and the price of the good is adjusted to $P + dP$, chosen to leave $U$ unchanged, we obtain:

$$dP = \frac{U_1}{U_2 v_1}(\Sigma u_i\, db_i + u_t\, dt) - \frac{v_2}{v_1}\, dt.$$

Note that $dP$ differs from a shadow-price weighted sum of characteristics (including time use) changes by the term $-(v_2/v_1)\, dt$. This last term represents the marginal valuation of leisure in terms of goods outside the group and can be taken to be the wage rate in typical situations. This suggests that the appropriate index for measuring quality change is not the plain expenditure-equivalent index but an index based on an augmented expenditure concept defined by $\bar{P} = P + wt$, where $w$ is the wage rate and $t$ the time devoted to activities using goods in the group. ($\bar{P}$ corresponds to the "full price" in the sense of Becker (1965)).

So long as we can assume that $w = v_2/v_1$, it is easy to proceed along the lines of the preceding section analysis and show that

$$\frac{d\bar{P}}{\bar{P}} = \frac{\Sigma u_i\, db_1 + u_t\, dt}{\Sigma u_i b_i + u_t t}.$$

Thus the augmented expenditure-equivalent index gives the same measure of quality change as the hypothetical quantity-equivalent index and the fixed-weight characteristics index, both with time use included among the characteristics. Clearly it is the augmented expenditure-equivalent index and not the standard expenditure-equivalent index which is the appropriate measure of quality change.

Note that, although time appears in the augmented expenditure with a weight that can be taken to be given by the relevant wage rate, the weight to be given to time in the index of characteristics is different—it is the marginal utility of time *in using the car* relative to the marginal utility in generalized leisure. The former may be greater than the latter (if driving is considered very pleasurable) or less, and the marginal utility of time in use of the car may even be negative if driving is considered an unpleasant chore.

What happens if the assumption of separability must be discarded altogether, there being no characteristics obtainable from goods in the group which cannot also be obtained from other goods? In this case, we cannot find an operational concept corresponding to any number which purports to measure the quality change in a single good. We can measure only the overall "quality of consumption" change which results from the change in specifications of the various goods in the consumption bundle.[15]

## 6. Using Market Data

The various potentially acceptable measures of quality change require information with respect to one or more of the following:

(1) Market data showing the relative prices at which the original and quality-changed varieties are being sold to the same consumers at the same time. (For expenditure-equivalent indexes.)

(2) Technical data on the quantitative content of relevant characteristics in the original and quality-changed varieties, plus the shape of the relevant portion of the reference indifference curve. (For quantity-equivalent and hypothetical quantity-equivalent indexes.)

(3) The same technical data, plus true shadow prices for characteristics rather than indifference curve shapes. (For fixed-weight indexes of characteristics.)

Since it is separate from the others, let us first consider the use of market data to obtain expenditure equivalents. To read such equivalents directly from the market, we require that the original and changed version of the good be available simultaneously and that there be evidence that the reference consumer is indifferent between the two versions at the going prices, which can then be treated as expenditure equivalents.

In most typical cases (such as automobiles), the new and old models are not really available at the same time. Although there may be some overlap at the commencement of the model year, the new models will not yet have attained the equilibrium price and the old models are usually being discounted heavily to clear inventory, so that the relative prices are not representative of the equilibrium relationship. In addition, depreciation depends on the nominal year of the model rather than the actual data of first use, further distorting relative prices during the changeover period.

One way of trying to avoid the problem of non-simultaneity simultaneity is the Burstein–Cagan–Hall[16] technique of looking at the market for used goods in cases (such as automobiles and houses) in which this market has more than fringe significance. In 1975, the used market will give relative prices for 1973 and 1974 models. These relative prices are presumed to depend on two separable factors, the quality difference between the two models which is presumed to be independent of time, and the difference in age between the two models, the effect of which

[15]See also Fisher and Shell (1968), Muellbauer (1974, 1975).

[16]Burstein (1961), Cagan (1965), Hall (1971).

is presumed to be independent of the characteristics mix. The assumption of separability enables us to determine the effect of pure time depreciation in a variety of ways, adjust for the contribution of this to the price difference between 1973 and 1974 models, and assume that the residual represents the price difference due to quality change. The used-good approach has the advantage of being based on an effective comparison between successive model years, but depends crucially on the separability of time depreciation from other characteristics (implying that the characteristics mix does not change as the good grows older, very dubious) and on the estimation of the effect of this time depreciation. In addition, the quality change measure is not available except with a lag and this quality change is being assessed in terms of the preferences of a later year and not either of the two years in which at least one of the models was actually available new. Nevertheless, this approach may be the best we have for developing historical series of quality change measures for indivisible goods.[17]

The alternatives to expenditure equivalents both require that we possess information on the characteristics of the good both before and after the quality change. This information is technological rather than market information, although we may use market data to satisfy ourselves as to the set of characteristics which is really relevant to consumer choice. In any case the presumptive set of relevant characteristics and their quantification is a non-market problem, to be obtained from engineering or testing service data.

As a practical matter, we can rule out the possibility of obtaining the shape of a reference indifference curve from any kinds of data likely to be available in a real situation. Thus the real alternative to an expenditure-equivalent index is a fixed-weight index of characteristics, with weights proportional to the appropriate shadow prices. It is these shadow prices for which we require the market data.

In Section 4, it was shown that the appropriate shadow prices are those proportional to the marginal utilities of the characteristics (or relative shadow prices should equal the marginal rates of substitution between characteristics), and can thus be obtained from the slope of the hyperplane which is tangent to the indifference surface at the point corresponding to the characteristics bundle of the good.

Historically, the slope of this hyperplane has been estimated by regressing price (adjusted for the "pure" price change) on characteristics content over cross-section and/or time series data for the individual varieties of goods.[18] This technique assumes that the shadow prices on characteristics are the same for all varieties, and it is this assumption that needs close scrutiny.

Although almost all the applications of the above "hedonic technique" have been to goods which are essentially indivisible and non-combinable, let us examine its foundations by first considering the combinable case. Since combinability implies either divisibility or indivisibility with consumption of many units, we can eliminate direct consideration of price by taking the set of all characteristics bundles than can be obtained by spending a given budget on all possible combinations of goods (in the group). Now if all consumers have identical preferences and incomes or identical homothetic preferences, the only cir-

[17]See Hall (1971) for discussion of the econometric problems.
[18]See Griliches (1961, 1971).

cumstance under which all goods could simultaneously find a place in the market would be when the relative prices of the goods were such as to make all points corresponding to characteristics obtainable from the budget lie on a hyperplane. Under these circumstances, there is a unique hyperplane which is tangent to the indifference surface and the regression method gives the true shadow prices.

It is surely much more realistic to suppose that the existence of many varieties of goods on the market implies the existence of varied preferences over the population. In this case, goods can be sold at prices which do not lead to points on a hyperplane, provided the points lie on the boundary of a convex feasible set of characteristics, with different consumers buying different goods combinations. This is shown in Figure 1, where there are three different consumers buying

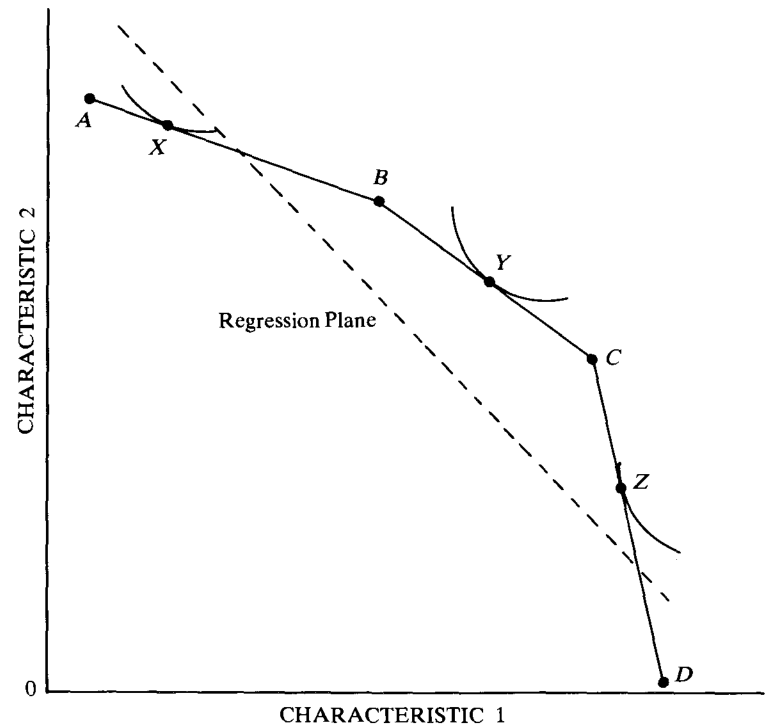

Figure 1

different combinations of goods (two at a time) to achieve their respective optimal characteristics combinations X, Y, Z. If our reference consumer is the one with optimal point X, the relevant shadow prices are given by the slope of BC. A regression will fit a line to all four points, giving shadow prices determined by the broken regression line in the diagram. These are not the appropriate shadow prices for the reference consumer, nor for any of the three consumers.

Given combinability, we can determine the true shadow prices from the slope of the facet BC.[19] Note that the facet method uses information about the

[19]See Lancaster (1967, 1971), Muellbauer (1974), Klevmarken (1973).

characteristics of goods B and C only—it measures quality change only over those goods actually purchased by the reference group of consumers. In many cases we might expect a cluster of goods with very similar characteristics proportions near that part of the spectrum in which our reference consumer is interested, suggesting that a regression through this cluster might be an acceptable estimate of shadow prices with some error-diminishing properties. In any case, the statistically most useful observations (the most divergent from the cluster) would not be used.

The problems outlined above are relatively minor ones compared with those that arise with indivisible and non-combinable goods. Here we are in real trouble because the true shadow prices may not become manifest in any form at all. Let us consider a group of, say, automobiles which differ in characteristics proportions and which, for simplicity, are all sold on the market simultaneously at the same price. If all consumers are identical, these facets are consistent only if the characteristics bundles corresponding to the various cars correspond to points on the (unique) indifference surface. In principle, we could then draw the indifference surface from the data and obtain the shadow prices relevant to quality change measurement for any one of the automobiles from the slopes at the proper point.

If preferences vary and consumers are not identical, the market tells us very little. Since each consumer chooses only one good, the fact that all are sold implies only that the points corresponding to the characteristics bundles of the goods *not* bought by each particular consumer must all lie on the origin side of the indifference curve through the point corresponding to the good he actually did buy. This imposes some limits on the possible values of the shadow prices, but these limits may be very broad unless there is independent evidence that the indifference curves show very high elasticity of substitution.

The best we can do in the indivisible case—which is the main case in terms of the usefulness of measuring quality change—is to assume relatively flat indifference curves (high elasticity of substitution) and choose a reference consumer whose optimal characteristics collection is in a cluster of goods with similar characteristics proportions. The slope of the indifference curve cannot then differ greatly from the slope of the line joining the points corresponding to the goods closest in specification to that actually chosen by the reference consumer. Since there is simply no way of determining the *exact* shadow prices, even in principle, it would probably be appropriate to take the regression through the cluster as a safe approximation. As in the divisible case, the appropriate regression would ignore points that were not close to the cluster, increasing the theoretical validity at the expense of statistical reliability.

## 7. Conclusion

We have shown that, if the quality changes are "small", the numbers derived from all three of the indexes which we might choose as criteria for quality change will coincide, and we can then consider these numbers to give an unambiguous "measure" of quality change. The choice of index will then be a matter of practical convenience, the two realistic options being the expenditure-equivalent index (especially applicable to vintage and used-good data) and a fixed-weight index of characteristics with properly derived shadow price weights and not regression weights. The major complications which may arise are non-separability, taste changes, and relative price changes among goods not being indexed for quality change.

## References


Adelman, I. and Griliches, Z., "On an Index of Quality Change," *J. Amer. Stat. Asscn.*, 56 (1961), pp. 535–48.

Becker, G. S., "A Theory of the Allocation of Time," *Econ. J.*, 75 (1965), pp. 493–516.

Burstein, M., "Measurement of Quality Change in Consumer Durables," *Manchester Sch. Econ. and Soc. Stud.*, 29 (1961), 267–79.

Cagan, P., "Measuring Quality Change and the Purchasing Power of Money: An Exploratory Study of Automobiles," *Nat. Banking Rev.*, 3 (1965), 217–36. (Reprinted in Griliches (ed.) (1971).)

Court, A. T., "Hedonic Price Indexes with Automotive Examples." In *The Dynamics of Automobile Demand*, New York, General Motors Corporation, 1939.

Dhrymes, P. J., "Price and Quality Changes in Consumer Capital Goods: An Empirical Study." In Griliches (ed.) (1971).

Fisher, F. M. and Shell, K., "Taste and Quality Change in the Pure Theory of the True-Cost-of-Living Index." In J. N. Wolfe (ed.), *Value, Capital, and Growth: Essays in Honour of Sir John Hicks*, Edinburgh University Press, 1968. (Reprinted in Griliches (ed.) (1971).)

Griliches, Z., "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change." In *Price Statistics of the Federal Government, 1961*, General Series, No. 73. New York, National Bureau of Economic Research. (Reprinted in Griliches (ed.) (1971).)

———, "Introduction: Hedonic Price Indexes Revisited." In Griliches (ed.) (1971).

——— (ed.), *Price Indexes and Quality Change: Studies in New Methods of Measurement*, Cambridge, Mass., Harvard Univ. Press, 1971.

Hall, R. E., "The Measurement of Quality Change from Vintage Price Data." In Griliches (ed.) (1971).

Houthakker, H., "Compensated Changes in Quantities and Qualities Consumed," *Rev. Econ. Stud.*, 19 (1952), 155–64.

Klevmarken, N. A., "A Note of New Goods and Quality Changes in the True Cost-of-Living Index in View of Lancaster's Model of Consumer Behavior," *Research Report*, Institute of Statistics, University of Stockholm, 1973.

Lancaster, K., "A New Approach to Consumer Theory," *J. Polit. Econ.*, 74 (1966), 132–57.

———, "New Methods of Price Measurement: Discussion," *Am. Stat. Assocn., Proc. Econ. and Bus. Sect.*, 1967, 346–47. (Comments on an earlier version of Griliches (1971.)

———, *Consumer Demand: A New Approach*, New York, Columbia Univ. Press, 1971.

Muellbauer, J., Review of Griliches (ed.) (1971), *J. Econ. Lit.*, 10 (1972), 830–32.

———, "Household Production Theory, Quality, and the 'Hedonic Technique'," *Amer. Econ. Rev.*, 64 (1974), 977–94.

———, "The Cost of Living and Taste and Quality Change," *J. Econ. Theory*, 10 (1975).

Rosen, S., "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *J. Polit. Econ.*, 82 (1974), 34–55.

Stone, R., *Quality and Price Indexes in National Accounts*. Paris, OECD, 1956.