Volume Title: Household Production and Consumption

Volume Author/Editor: Nestor E. Terleckyj

Volume Publisher: NBER

Volume ISBN: 0-870-14515-0

Volume URL: http://www.nber.org/books/terl76-1

Publication Date: 1976

Chapter Title: The Concept and Measurement of Product Quality

Chapter Author: E. Scott Maynes

Chapter URL: http://www.nber.org/chapters/c3970

Chapter pages in book: (p. 529 - 584)

The Concept and Measurement of Product Quality *

E. SCOTT MAYNES

UNIVERSITY OF MINNESOTA

I. INTRODUCTION

QUALITY is a topic which economists, by and large, have swept under the rug.[1] This paper faces up to the problem by proposing, and arguing for, a concept of product quality, and by showing how it might be measured empirically.

The plan of this paper is first to excite your interest by presenting a sample of actual measurements of product quality in local consumer markets. A discussion of the uses to which quality may be put follows, along with an effort to distinguish my work from that of other toilers in the same vineyard – namely, those working on hedonic price indicators, and Kelvin Lancaster, with his characteristics approach to consumer demand theory. The heart of the paper will deal with the conceptualiza-

* The long gestation period of this paper has given birth to a large number of critics/creditors. My debt is heaviest to Edward M. Foster, and W. Keith Bryant of the University of Minnesota. But I am also deeply indebted to Marcel K. Richter, Leonid Hurwicz, Herbert Mohring, and Neil Wallace of Minnesota; Richard D. Emmerson, J. Edward Russo, Harold Nelson, and Lucille Thompson of the University of California, San Diego; Arthur J. Rolnick of the Federal Reserve Bank of Minneapolis; and Christopher Babb of Cornell. The residual responsibility for accepting or rejecting criticisms so generously given rests with the author.

[1] A survey of the major economic journals during the postwar period yields about twenty articles dealing directly with quality. Most of these belong to the hedonic price index literature, discussed in Section IV of this paper. Several deal with the selection of quality level as a factor in product differentiation. See [10] as an example. Several deal with the proposition of accepting price as an indicator of quality [e.g. 27, 12]. One, Theil's [31], by accepting the mean price as an index of quality, seeks to estimate the demand for quality.

However none essays the task of this paper – conceptualizing and measuring quality per se. Nor has any word about quality reached economic undergraduates if Samuelson's text [26] is representative. One will look in vain to find "quality" in the index. And the closest one comes in substance is a discussion of product differentiation and

tion and measurement of product quality in full detail. A final section will suggest some research possibilities opened up by this essay.

This paper is infused with the view that economic theory and concepts should be the servants of intelligent consumers and handmaidens to empirical research. Accordingly, the concepts presented here represent the simplest possible formulations and have been shaped with empirical measurement in mind.

II. PRICES, QUALITY, AND QUALITY-DEFLATED PRICES IN LOCAL CONSUMER MARKETS: SAMPLE RESULTS

The results reported here are typical of those obtained from a series of investigations made under my direction by upperclassmen and graduate students at the University of Minnesota since the spring of 1970. In all these investigations, students were instructed to obtain estimates of variations in money prices, quality, and quality-deflated prices for a variety of products on the local markets in which consumers would normally purchase. (These markets included mail-order outlets.) The samples of brands/models and retail outlets were selected purposively rather than by probability methods. However, students were urged to take special pains to cover those brands and/or retail outlets whose prices or qualities were expected to be unusually high or low.

Quality may be assessed for either a *variety*—a product/brand/model combination—or, when the characteristics of a seller are also taken into account, for a *specimen* as well. In the sample results reported here, quality was assessed for both varieties and specimens of sofa beds.

The *quality* of a *specimen* was defined as "the subjectively weighted

the factors accounting for it. The identical observation applies to graduate theory texts if [15] is taken as representative.

There is evidence of much earlier interest in the problem of quality. Stigler reported [28, page 315] that John Bates Clark "became preoccupied with a problem to which he could not find a useful solution: how to apply marginal analysis to variations in the quality of goods."

Evidently, the difficulty of the subject has put off many. In his classical article on "The Economics of Information," George Stigler commented: "The search for knowledge on the quality of goods, which has been studiously avoided in this paper, is perhaps no more important but, certainly, analytically more difficult. Quality has not yet been successfully specified by economics, and this elusiveness extends to all problems in which it enters" [29, page 224].

Georgescu-Roegen [13, pages 97–105] asserts gloomily that despite the quantification of some aspects of reality, there will always remain a "qualitative residual" which will defy measurement.

Psychologists and marketers have devoted considerable attention to quality. But their focus has been upon *perceptions* of quality rather than on the conceptualization and measurement of quality. [17] is representative.

average of characteristics," *characteristics* being in turn defined as the "services giving rise to utility," such as safety, durability, and beauty.

The sofa-bed investigation reported here presents data for ten specimens on sale at four different retail outlets. This particular study was selected for presentation on the basis of its representativeness and the care invested in its execution.

In assessing the quality of *varieties* of sofa beds, the investigator identified and weighted the characteristics of the good itself as follows:

Characteristic		*Weight*
Aesthetics		.50
Comfort:		.15
As sofa	.10	
As bed	.05	
Durability:		.15
As sofa	.10	
As bed	.05	
Convenience		.10
Warranty		.06
Safety		.04
	Total:	1.00

As for the retail outlets or sellers, their characteristics were identified and weighted as follows:

Characteristic	*Weight*
Sensitivity, pleasantness	.35
Convenience	.25
Knowledge of product	.20
Reliability	.10
Warranty (retail seller's expected performance with respect to his responsibility)	.10
Total:	1.00

In combining the characteristics of the good itself and of the seller to obtain the quality of a specimen, the investigator assigned a weight of .85 to characteristics of the good and .15 to characteristics of the seller.

Results from this sample investigation are presented in Table 1. Three outcomes are salient: the wide range of money prices, the wide range of quality,[2] and perhaps even more significant, the fact that the

[2] Since the boundaries of the quality index are defined precisely, 0.00 representing a total absence of desirable characteristics and 1.00 representing the quality of the ideal specimen, differences in this scale are meaningful and not arbitrary. See Section V for further discussion of the concept of quality.

TABLE 1

Variations in Money Prices, Quality, and Quality-Deflated Prices in a Local Market: Sofa Beds in Minneapolis [a]

		Varieties [b]			Specimens [c]	
Retail Outlet (1)	Brand/Model (2)	Money Price (3)	Quality (4)	Quality-Deflated Price [d] (5)	Quality (6)	Quality-Deflated Price [d] (7)
1	*A*	$230	.49	$ 469	.52	$ 440
	B – Sale price	250	.61	410	.63	398
	– Regular price	339	.61	556	.63	540
	C	260	.63	411	.65	402
2	*D*	450	.70	643	.92	488
	E – Sale price	399	.98	407	.98	408
	– Regular price	460	.98	469	.98	470
	F – Sale price	269	.64	420	.68	395
	– Regular price	380	.64	594	.68	558
3	*G*	499	.84	594	.86	580
	H – Sale price	560	.63	889	.68	828
	– Regular price	702	.63	1,114	.68	1,033
4	*I*	385	.60	641	.62	623
	J	269	.59	456	.62	434
Means:		$389				
The range:		$230 to $702	.49 to .98	$ 407 to $1,114	.52 to .98	$ 395 to $1,033
Ratio of highest to lowest:		3.06	2.00	2.76	1.88	2.62

[a] These data were collected by Kathryn S. Hochsprung, a senior at the University of Minnesota. The work was done in the spring of 1973. All dollar figures are rounded to the nearest one dollar.

[b] A *variety* of a product is a product/brand/model combination. The assessment of the quality of a variety does *not* take account of characteristics of the seller.

[c] A *specimen* is a product/brand/model/seller combination. Its quality assessment does take account of characteristics of the seller.

[d] The quality-deflated price P^* equals P/G where P = money price and G = quality.

deflation of money prices to take account of quality did not appreciably reduce the range in quality-deflated prices.

These results are not singular but, rather, are typical of the results obtained in more than thirty such investigations. It should be noted that these thirty or so investigations embrace a most diverse set of products, ranging from color TV sets to bicycles to food blenders to auto insurance to wigs, but not including foods.

The reactions of the investigators to their own efforts, though informal in character and necessarily inconclusive, are suggestive. All found the task of assessing quality both tractable and meaningful. When queried, all felt strongly that their quality assessments were both valid and reliable.

The skeptic can find much in the results above and in results not presented here about which he may properly be skeptical:

1. The investigations were carried out by students, not by professional social scientists. (The assessment of quality is, we shall argue, a matter for individuals. However, the students may err in their handling of data as compared with knowledgeable professionals.)
2. Were all the relevant characteristics considered in arriving at the quality measure? (The skeptic may be able to think of other characteristics he would seek in that product.)
3. Is the "product" class on which measurements were performed too broad or too narrow? For example, are there some sofa beds included in the sofa bed comparison which the critic would not consider an acceptable substitute for others in the class?[3]

Despite these doubts, I suggest that the sample results above have two "messages" for us:

1. The measurement of product quality along these lines is feasible.
2. More definitive measurements are likely to confirm the existence of wide ranges of money prices and of the quality-deflated price in local consumer markets.[4]

It is now time to state clearly the uses to which quality may be put.

[3] At the Conference, several discussants voiced doubts that Specimen *H* should be included in the "product" class for Table 1. When this issue was posed to Kathryn Hochsprung after the Conference, she reaffirmed her judgment that Specimen *H* was sufficiently similar in characteristics to the other sofa beds viewed as to warrant "serious consideration" for possible purchase. And there was no question that the seller of Specimen *H* was easily accessible and hence a part of the relevant "market."

[4] The number of brand/models sampled in the sofa bed example was small. Evidence from other investigations with larger samples of models suggests that the wide ranges found in money prices and in quality-deflated prices is not attributable to outliers.

III. THE USES OF QUALITY

An acceptable, empirically measurable concept of quality should have three major uses: (1) as an *index of possible payoffs* to searching (shopping) by consumers, (2) as an *index of the "informational effectiveness"* of markets, (3) as a *building block in economic theory,* facilitating the development of market demand/supply relationships for differentiated products. Each of these requires further explanation. We now turn to this task.

An Index of Possible Consumer Payoffs

Should an intelligent consumer ask an economist how much consumers should shop (or "search").[5] The economist would naturally, correctly, and perhaps fervently (!) recite the marginal rule: "Keep searching (shopping) as long as the expected gross payoff from searching exceeds the expected cost of an additional search."[6] This "answer" at once begets two other questions: (1) What is the nature of a "payoff"? (2) How does one estimate the value of the expected gross payoff?

A *payoff* may be defined as "any gain resulting from searching." A moment's reflection will disclose that gains may take the form of a lower money price (quality equal) or better quality (money price equal) or both. What the marginally calculating consumer would find highly useful is a measure of quality which renders money prices and quality commensurable. The instrument which will perform this task is a cardinal measure of quality. And it is just such a measure which is proposed in this paper.

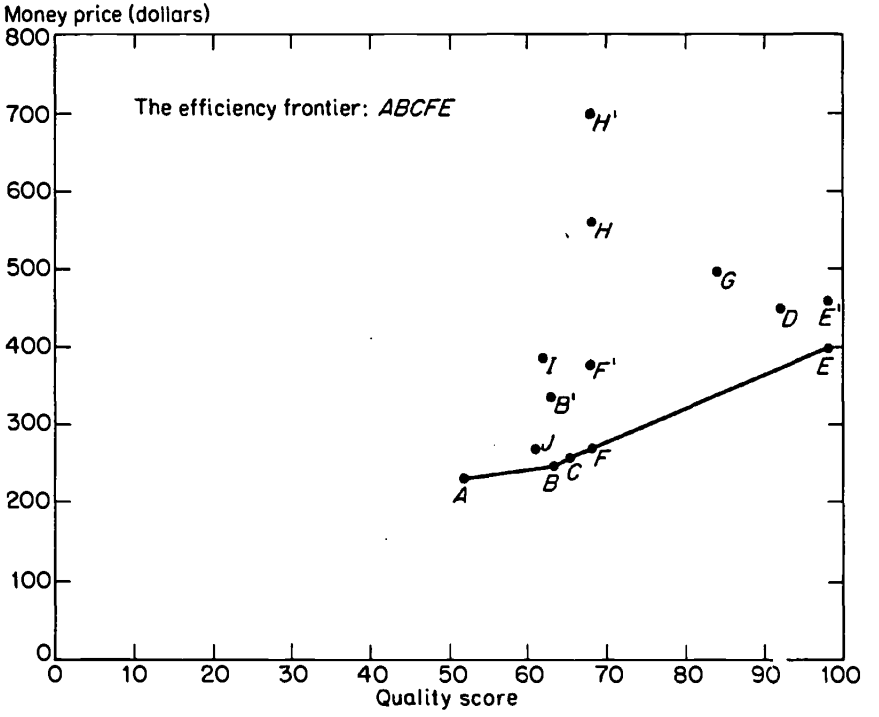
Figure 1, depicting the data of Table 1, conveys price-quality information relating to sofa beds in a choice-facilitating manner. Assume for the moment that the consumer has access to Figure 1 and also to the identity of Specimens A, B, C, and so on. Assume further that all consumers would make uniform assessments of the quality of various

[5] A *search* is defined as "each attempt to secure and to act on information regarding the price and/or quality of a product." Thus, search, as used here, would include all of the following: "shopping," consultation of publications giving product information, telephoning retailers to ask if they carry a given brand, getting information from the "yellow pages," consulting a mail-order catalogue, and—very important—bargaining.

[6] *Costs* include "anything which is undergone or forgone in order to attain a given end." Thus, the dollar equivalent of someone's distaste for shopping may be a "cost" of shopping. Alternatively, the dollar equivalent of the activity forgone in order to undertake shopping, e.g., an afternoon's sailing, represents another type of cost. The more common varieties of costs, such as transportation costs, are naturally included.

FIGURE 1

Price and Quality in a Local Market: Sofa Beds in Minneapolis



NOTE: Letters represent various specimens. Where an identical letter appears twice, the primed designation, e.g., H', represents a "regular" price while the unprimed designation, e.g., H, represents a sale price.

SOURCE: Table 1.

specimens. (The assumption of uniform assessments of quality will be discussed later.) For the choice-maker the dominant and most useful feature of Figure 1 is the *efficiency frontier, ABCFE*. The efficiency frontier consists of the set of points and the line segments connecting them, for which a given level of quality may be purchased at the lowest price.

In choosing among the specimens available,[7] the rational consumer would proceed along the efficiency frontier from left to right. His first inclination would be to consider Specimen A with a money price of \$230. Should he purchase more quality? Figure 1 (and the supporting data from Table 1) would reveal that by purchasing Specimen B for a sale price of \$250 he could achieve a 21 per cent increase in quality (.11/.52) for only a 9 per cent increase in money price paid (\$20/230). Or, to consider another alternative which would be attractive to someone with a taste for the "best," by purchasing Specimen E at the sale price of \$399, he could achieve an 88 per cent increase in quality as compared with A (.46/.52) for a 74 per cent increase in price (\$169/230). Which should he do? Only the consumer himself can say whether the marginal improvement in a particular comparison is "worth" the additional money outlay. But one thing is certain: the choice of the rational consumer would lie on the efficiency frontier.

But, supposing our prospective purchaser had access to Figure 1, but *not* to the identification of the brand/model/seller represented by A, B, C, etc. Under these conditions four pieces of information would be particularly useful: (1) the range in quality, given on Figure 1 by the distance on the horizontal axis from A to E; (2) the visual correlation, if "high," between price and quality; (3) the average slope of the efficiency frontier, $\Delta P/\Delta G$, where G = quality (a mnemonic for "goodness"); (4) the ratio of the maximum price to the frontier price, here approximately H'/F. But why should these four be singled out as important? The question is answered below in terms of the data of Figure 1.

The large range in quality, E being approximately twice as "good" as A, poses a major choice for the purchaser: Should he opt for more or less quality? If the correlation between price and quality is "high" meaning "visually obvious," then price becomes a proxy for quality and the purchaser, with a knowledge of the average slope of the efficiency frontier, $\Delta P/\Delta G$, can obtain the quality he desires by focusing on the money price asked.

When the correlation is low, the intelligent consumer should shift his attention to the discovery of the efficiency frontier. At this point, he will be interested, too, in the ratio of the maximum price to the corresponding frontier price. The greater this ratio, the greater the amount he should search for a lower money price.

[7] Careful definitions of *product*—needed to identify the specimens properly counted as part of the product class—and *market*—needed to determine which sellers and which buyers are properly part of the class—are given in Sections VII and VIII. For the moment, we assume that satisfactory definitions of product and market are at hand.

But suppose our prospective purchaser had access to only a single statistic from Table 1. What would be most useful? My candidate would be the range in quality-deflated prices P^*. If this range was small, it would suggest that price was indeed closely associated with quality, and that the searcher should concentrate on price. If this range was large, considerable searching would be appropriate and considerable payoffs might be expected. For specimens of sofa beds, this range was \$395 to \$1,033, suggesting very large payoffs, indeed.

It is useful to recall that, in this discussion, we have assumed that consumers make uniform assessments of quality. If, and as, quality assessments by different individuals become increasingly less uniform, then different consumers are likely to identify different efficiency frontiers, and variations in P^* as assessed by one individual would become increasingly meaningless to other individuals. In mentioning the issue of uniformity versus nonuniformity in quality assessments, it is worth recalling the informal judgments of the student investigators: they felt that others undertaking the same quality assessments would achieve highly similar results. Obviously, a test of the congruity of quality assessments and of efficiency frontiers is an important topic for future research.

A Measure of the "Informational Effectiveness" of Markets

Economists will naturally be interested in the extent to which different buyers achieve the best possible terms offered in a given market. As a first approximation, we will designate as *informationally effective* any market in which all buyers purchase specimens situated on the efficiency frontier, assuming uniform assessments of quality. And we will propose variations in the vertical distance of prices from the efficiency frontier for any given level of quality (the excess of the price actually asked over the price on the frontier) as a measure of the informational effectiveness of markets: the greater the variation in distances from the frontier, the less informationally effective a market.

Note that this concept applies only to *informational* effectiveness and not to overall effectiveness. The single price charged in a market which is perfect from the informational viewpoint might be the "high" price of the perfect monopolist.

We might check the working of the proposed measure against the sofa bed data presented in Figure 1. Here the most distant price, H', exceeds the corresponding price on the efficiency frontier, F, by \$433 (\$702 minus \$269). Other points are also considerably removed from the efficiency frontier, though not as extremely as H'. It is clear that the

market for sofa beds is a long way removed from the zero variation from the frontier which we would expect to find in an informationally perfect market. The culprit in this case is undoubtedly the fact that efforts to obtain information regarding price and quality are costly and imperfect, as are efforts to act on the basis of such information, e.g., bargaining, or a visit to another seller.

Why do we propose the variation in the distance from the efficiency frontier only as a first approximation measure? First, the validity of the measure rests on the assumption that, within product classes, different consumers make identical quality assessments. If consumers differ in their assessments of quality, then it may be perfectly reasonable for Consumer I to prefer Specimen *A* over Specimen *B* at the same money price, while Consumer II prefers *B* over *A*. This event, generalized, would yield variations in the distance from the frontier, as measured by either I or II, which would be consistent with an informationally *effective* market.[8] For this reason, too, the extent of congruity in quality assessments is an important topic for future research.

A second qualification arises from the prevalence of price discrimination based upon objective factors. Suppose—and this is an actual case—American Airlines offers a two-thirds reduction in fares between certain cities to dependents traveling with a family head and returning within a specified period. This example of price discrimination will increase the distance from the efficiency frontier, but need not imply informational ineffectiveness. Such would be the case if advertising and/or airline clerks succeeded in informing all eligible persons of this favorable fare. If, however, the efforts of advertising and airline clerks failed to inform all of those eligible, then the observed variation in the distance from the efficiency frontier would be attributable, in part, to an informationally ineffective market.

It should be noted that much price discrimination arises from the very fact that consumer ignorance makes possible the separation of markets. An example would be the different prices paid, as a result of bargaining, by different purchasers of identical new cars. To the extent that it results from differences in information and *not* from differences in bargaining skills (knowledge equal), this kind of price discrimination contributes to the informational ineffectiveness of the market in which it occurs.

Summing up, economists should be interested in the informational

[8] The task will be simpler if it turns out that individuals, similarly situated, tend to make uniform judgments regarding the delineation of "markets," "products," and quality. If individuals' judgments on these matters differ, the analysis will become more complex and less general but not impossible.

effectiveness of markets. The assessment of quality is a necessary ingredient in the evaluation of this facet of market performance. Just how much variation in the distance from the efficiency frontier is required to justify a verdict of "informational ineffectiveness" is rendered uncertain, due to the possibility of differences in quality assessments and also to the existence of price discrimination based on objective factors.

A Building Block in Economic Theory

With respect to the theoretical treatment of quality, Edward Chamberlin's legacy, Chamberlin [5, 6], is at once laudable and lamentable. It is laudable because—although he rarely used the word *quality*—Chamberlin successfully explained why producers/sellers seek to incorporate real or imagined quality differences into their products and thereby "differentiate" their products. The successful product differentiator, as all of us know, is not a "price taker," but is, in some degree, a "price maker."

Chamberlin's legacy is lamentable, however, because instead of developing the concept of quality, he moved in the dead-end direction of defining each product variant as a separate "product." And this ineluctably led to the well-accepted conclusion in economic theory[9] that the development of *market* demand/supply relations for sets of closely related "products" is inadmissible.

It is my contention that the deflation of money prices by an acceptable measure of quality—the P^* cited earlier—plus appropriate definitions of "product" and "market" will provide us with the necessary building blocks from which market demand and supply relationships for differentiated products may be constructed.

It is my hope that the concepts of quality, product, and market presented in this paper will prove up to the task. And it is my belief that the task is both important and feasible. Most of the purchase decisions consumers make involve the assessment of quality as an important ingredient. If economics is to be helpful and relevant in this sphere, there must be developed models which incorporate quality differences. If the concepts proposed in this paper—or their successors—are found acceptable, then the task of moving to market demand/supply relationships for differentiated products should be relatively straightforward.

[9] This conclusion is "well accepted" in economic theory but not in the domain of applied economics and econometrics, as witnessed by various studies of the demand for automobiles, refrigerators, houses, and food.

IV. THIS PAPER AND THE EXISTING LITERATURE

The central task of this paper is to conceptualize and measure product quality. It is this emphasis on quality per se which distinguishes this effort from Kelvin Lancaster's pioneering work on the characteristics approach to consumer demand and from the hedonic price literature. For those with a taste for taxonomy, this paper probably fits best in the economics-of-information literature. We shall consider each in turn.

For the contributors to the hedonic price index literature, quality [10] is a "pain in the neck." Their primary objective is to purge time-series price data of quality effect so that they can obtain valid measures of "pure" price changes.

Their view of quality is highly similar to that proposed in this paper. Adelman and Griliches [1, page 539] propose that "the quality of a commodity be regarded as a composite of different characteristics." However, since they wish to obtain measures of *average* money prices over time, the hedonists confine their attention to the *average* effects of *only* those characteristics which have a measurable influence on money prices. By contrast, our formulations will embrace *all* characteristics entering the utility function and will be concerned with individual, as well as with average, assessments of quality.

It is ironic, considering the length [11] and strength [12] of the hedonic literature to note that none of its practitioners has sought to generalize the measurement of quality as this paper does.

Though my work could have been inspired by the hedonic approach, it was, in fact, stimulated by Kelvin Lancaster's seminal work on the characteristics approach to consumer demand.[13] But the approach here differs from Lancaster's in several important ways.

First, in two articles and one book, Lancaster never felt it necessary or desirable to develop a measure of quality. By contrast, it is my view that the concept of quality is essential for the pursuit of the three objectives set forth above—the estimation of consumer payoffs, the assessment of the informational effectiveness of markets, and as a building block in economic theory. In addition, it is my expectation that the concept of quality will be helpful, both normatively and positively,

[10] For a succinct summary of the hedonic approach, cf. [32]. Landmarks in the development of the hedonic approach include [7, 1, 11] 1971. See Griliches [14], 1971, for a set of papers covering most aspects of the literature.

[11] The hedonic approach dates back at least to Court's paper in 1939 [7].

[12] The quality of both the papers and the contributors to the volume edited by Griliches [14] testify to the strength of the hedonic literature.

[13] The basic paper is [19]. But many of Lancaster's ideas are spelled out more fully in his 1971 book [21]. For a simpler version of the paper, see [20], and for trenchant criticism, see [4].

in dealing with the problem of consumer choice among products whose quality differs. The stress which subscribers to *Consumer Reports* place on the quality ratings supports this view.[14]

A second major difference between this paper and Lancaster's work concerns the assumptions which each makes regarding the properties of characteristics. Lancaster assumes that characteristics are both *objective* and *universal,* whereas this paper rejects both assumptions, and instead takes the view that characteristics may be subjective and nonuniversal. The significance of these different assumptions will be discussed later. Here, it is our intention to alert the reader to the fact that this paper departs from, rather than follows, Lancaster in major respects.

As has been said, this paper probably fits best into the economics-of-information literature. Indeed, it deals with a problem *not* taken up by Stigler in the classical and seminal 1961 paper which launched the economics-of-information literature. In that paper Stigler wrote [29]:

> The search for knowledge on the quality of goods, which has studiously been avoided in this paper, is perhaps more important but, certainly, analytically more difficult. Quality has not yet been successfully specified by economics, and this elusiveness extends to all problems in which it enters.

The central task of this paper, to which we now turn, is to take up the challenge which Stigler eschewed: the successful specification of quality.

V. THE CONCEPT OF QUALITY

Some Preliminary Definitions

Before proceeding to the concept of quality per se, we must pause to define again, or to define in at least a preliminary way, the closely related concepts of *specimen, variety, characteristic, product,* and *market.*

A *specimen* is a "product/brand/model/seller combination," for

[14] Of the subscribers to *Consumer Reports* who purchased a given product, the following proportions reported that they had bought a "top-rated model":

Hair shampoo	81%
Color TV set	75
AM/FM radio	48
TV or FM antenna	51
Coffee maker	77
Sewing machine	64
Record changer	62

Source: Benson and Benson, Inc. "Survey of Present and Former Subscribers to *Consumer Reports*" (Consumers Union, mimeographed, 1970), Tables T21 and T22.

example, a 1971 Buick Sports (i.e., station) Wagon purchased from Fairways Buick. (A more careful specification of a specimen would require the listing of the motor type, the accessories, and other features which distinguish this class of Buick from other Buicks.) If we dropped the seller designation ("Fairways Buick") and spoke of *any* 1971 Buick Sports Wagon, we would be dealing with a *variety*. Though it is appropriate to conceptualize and measure the quality of either a specimen or a variety, we shall deal below with the quality of a specimen.

For purposes of understanding our concept of quality, the preliminary definition of a characteristic used earlier will suffice: A *characteristic* is "any service giving rise to utility," such as safety, durability, beauty. A full discussion of characteristics, their specification and the measurement problems they pose, will follow our delineation of the concept of quality.

Bear in mind the uses to which the concept of quality is to be put. We will wish, for example, to estimate the range of quality of a given "product" in a given "market." Or we will wish to draw a demand curve, with the money price deflated by quality, for some differentiated "product." Thus, we need a definition of *product* to decide *which specimens* will be appropriately included in the class for which quality evaluations and comparisons are to be made. And, analogously, we will need a definition of *market* to decide *which sellers and which buyers* (consumers) are to be appropriately included in the evaluation/comparison class. For now, we will provisionally assume that satisfactory definitions of product and market are at hand. The actual definitions proposed and argued for in this paper will be taken up after we have dealt with quality per se.

The Concept of Quality

Verbally *quality*, by the definition proposed here, consists of:

> "the extent to which a specimen possesses the service characteristics you desire."

Formally, the *quality* of the kth specimen, G_k^{ij}, is given by

$$G_k^{ij} = \frac{\sum_{l=1}^{L} (W_l^{ij} \cdot Ch_{kl}^{ij})}{\sum_{l=1}^{L} W_l^{ij}} \tag{1.0}$$

where

G_k^{ij} = the quality of the kth specimen of the jth product class as assessed by the ith individual.

W_l^{ij} = the weight assigned to the lth characteristic in the jth product class by the ith individual.
Note that for different specimens of the same product class the weights assigned to a particular characteristic are identical.

Ch_{kl}^{ij} = the characteristic score assigned to the lth characteristic of the kth specimen in the jth product class by the ith individual.

Characteristic scores range from 0.00 to 1.00 with two alternative interpretations, giving rise to two alternative models for quality:

Model A: The Proportionality Model. Here 0.00 denotes a total absence of the characteristic and 1.00 denotes the amount of this characteristic in the "ideal" specimen, k^o. Values between 0.00 and 1.00 denote the ratio of the amount of the characteristic possessed in the kth specimen to the amount possessed in the ideal specimen.

Model B: The Direct Measurement Model. Now the characteristic score Ch_{kl}^{ij} refers to the *marginal utility* conferred by whatever amount is possessed by the specimen under consideration. A score of 0.00 denotes zero marginal utility, 1.00 the marginal utility of the ideal specimen (with respect to that characteristic), and other real numbers the ratio of MU_k to MU_{k^o} where k^o is the ideal specimen.

For convenience, weights will be assigned so that they sum to 1.00 for each specimen. This assumption, coupled with the convention adopted for characteristic scores, implies that G_{k^o}, the quality of the ideal specimen, will equal 1.00.

For some purposes it may be helpful to normalize the quality score of a specimen, G_k^{ij}, by expressing it as a ratio to the mean quality for that product. Then, $\hat{G}_k^{ij}$, the normalized quality index would be

$$\hat{G}_k^{ij} = G_k^{ij}/\bar{G}^{ij} \tag{2.0}$$

where

$\bar{G}^{ij}$ = the mean score for all specimens in the jth product group as assessed by the ith individual:

And, of course, we can utilize G_k^{ij}, the normalized quality index,

along with the money price of a specimen, $P_k{}^j$, to obtain the P_k^{*ij}, which we encountered earlier

$$P_k^{*\,ij} = P_k{}^j / G_k{}^{ij} \tag{3.0}$$

Variation in P_k^{*ij} among the various specimens in a product class provides a convenient index to possible payoffs to consumer searches.

And, as we shall now show, P_k^{*ij} is exactly equivalent to $P_k{}^j / MU_k{}^{ij}$, or the ratio of price to marginal utility for that specimen.[15]

Quality and the Utility Function: Two Models

We turn first to the Direct Measurement Model and then to the Proportionality Model.

Direct Measurement Model. The proof (dropping the i and j superscripts as unnecessary here) that G_k is equal to marginal utility follows immediately from our assumptions:

a. As specified above, the characteristic score, Ch_{k1} measures the marginal utility conferred by characteristic number l.
b. We assume that the total increment to utility provided by specimen k (MU_k) is a weighted sum of the utility provided by each characteristic

$$U(k) = \alpha_1 Ch_{k1} + \ldots + \alpha_L Ch_{kL} \tag{4.1}$$

c. It is apparent that if the weights, W_l, are appropriately chosen to be equal to the α_l, the quality index $G_k = W_1 Ch_{k1} + \cdots + W_L Ch_{kL}$ is identical to the utility specified in (4.1).

The Proportionality Model. Here the proof is the same, but requires a somewhat stronger assumption concerning the utility provided by each class:

a. We once again need to assume that

$$U(k) = \alpha_1 Ch_{k1} + \ldots + \alpha_L Ch_{kL} \tag{4.1}$$

Given that assumption, the proof is again immediate.
b. Since in this model the characteristic score measures the quantity possessed rather than the utility provided, (4.1) now says that utility provided is directly proportional to the quantity possessed of each characteristic. That is, if we let $U(Ch_{kl})$ represent the amount of utility provided by characteristic l in specimen k, we are assuming that

[15] In this paper "marginal utility" refers to the utility summed over a discrete step, namely the increment resulting from going from zero to a single specimen of a product.

$$U(Ch_{kl}) = \alpha_l \cdot Ch_{kl} \tag{4.2}$$

This assumption is not required for the Direct Measurement Model.

In the Direct Measurement Model it is assumed of course that each individual will make the judgment required.

The essential difference between the two models is that the Direct Measurement Model allows for declining marginal utility as more of a characteristic is found, or is thought to be found, in a particular specimen.

A property the two models share is the assumption of cardinality.

Knowledge and Quality

Quality assessments are knowledge-dependent in several respects. In the first place, the weights assigned to a characteristic will depend upon the knowledgeability of the assessor. For example, in an area where sanitation practices are poor, the assessor who understands bacterial theory is likely to assign greater importance (a larger weight) to cleanliness of food preparation in assessing the quality of restaurant meals than is an assessor who does not understand bacterial theory.

In the second place, the knowledgeability of the assessor is likely to affect the assignment of characteristic scores. In both the Direct Measurement Model and the Proportionality Model, characteristic scores are assigned with reference to "ideal" specimens. The assessor's perception of the "ideal" will depend upon his general knowledgeability. The assessor of the characteristic, automobile "performance," who is familiar with the performance of the Wankel motor, is likely to have a higher reference standard for "performance" than an assessor who is not. Hence, the base on which the characteristic score depends will differ between these two assessors.

Sometimes assessors may omit a characteristic which they "really" believe relevant. That is, if someone asked them—after a quality assessment had been completed—what weight they would have assigned to a particular omitted characteristic, they may answer with a nonzero value.[16] The effect of omitting a characteristic which "should" have been included is to produce errors in the quality score.

Finally, assessors may be unaware of the existence of particular

[16] Morris Kaplan, the late longtime Technical Director of Consumers Union, stated—on the basis of many episodes—that ordinary consumers identify fewer relevant characteristics than test engineers and, further, that, when prompted, acknowledge that they would accept "as their own" most of the missed characteristics. (Personal conversation with Morris Kaplan, August 12, 1971.)

brands, models, or sellers who are properly part of a given product/market set. Such omissions would produce errors in estimates of parameters of quality over a given product/market set, e.g., estimates of variation in P^*.

Since the state of knowledge varies over time, operational use of quality measurements may require the use of a time subscript. A reading of *Consumer Reports* or the driving of a new car may change one's quality index for that model.

Quality: Objective or Subjective?

If my reading of the "message" of several seminars on "quality" is correct, our fraternity yearns for an *objective* measure of quality. Unfortunately, such cannot be: quality is intrinsically subjective.

To see why, let us assume provisionally that (1) the characteristic scores in equation 1.0, the definition of quality, are provided "objectively" by (say) Consumers Union, and that (2) these scores are accepted as correct by all consumers. Then, assuming the Proportionality Model, the only source of variation in quality assessments of identical specimens comes from the weights assigned by different individuals, the $W_{j1}{}^{ij}$, to the various characteristics.

Whence cometh these weights? They come from a time analysis a la Becker [3] of alternative utility-maximizing activities, which give rise to a derived demand for characteristics. The latter then lead to the planned purchase from a particular product set. Understandably, different individuals will prefer different activities, and this may lead to different preferences for characteristics, even within a product set.

All of which makes sense, as an example may make clear. Consider two individuals, A and B, both contemplating the purchase of an FM radio receiver. For convenience, let us assume that the product/market set of specimen is identical for both. We will further assume that the quality of these FM sets is a weighted average of but two characteristics, sensitivity and selectivity. Sensitivity denotes the ability of a receiver to pick up distant signals while selectivity refers to the capacity of a receiver to reproduce a signal without interference from other signals in adjacent bands. It makes sense and should come as no surprise that A, who lives at the very fringe of a metropolitan area should assign a heavy weight to sensitivity and hence high "quality" to receivers judged high with respect to sensitivity. By the same token, B who lives in the heart of the metropolitan area beset by many competing FM signals assigns a heavy weight to selectivity. Thus, their quality assessments are at once different and subjective.

But what of Consumers Union and its weights? Are they not "ob-

jective"? The answer is, "No, of course not!" CU's product testers, whose quality-scoring system closely corresponds to that proposed here, seek to estimate weights appropriate to an "average" user of a product. But this part of the process is unavoidably judgmental and subjective. The reason: the taste for characteristics comes only from the utility functions of individuals.[17]

Subjectivity can arise also from the *characteristic score* component of the quality measurement. Even if Consumers Union or other testing agencies can devise satisfactory objective tests of characteristics,[18] the individual is free to accept in whole, in part, or not at all CU's assessments. This freedom to choose for himself endows the characteristic score with subjectivity.

The discussion of characteristic scores above pertains only to the Proportionality Model. In the Direct Measurement Model, the individual is required to make an intrinsically subjective assessment of the rate at which marginal utility declines as the amount of a characteristic produced by a specimen increases. So, here, too, the objectivist is thwarted.

The only possible consolation for the objectivistic is the hedonic approach, which, in the view of some of its exponents, converts subjective notions of "quality" into a combination of objectively measurable or "rankable" traits.[19] The hedonist, using a statistically reproducible method, regresses time-series or cross-section data on a set of objective characteristics in order to ascertain the implicit prices of the characteristics. By summing these implicit prices, he obtains an estimate of the *average* quality of a given *variety* (but not a specimen). However, the objective measure of quality here exists only as a mean estimated over an entire product/market set.

Uniformity versus Nonuniformity in Quality Assessments

The formulation of quality proposed in this paper permits, but does not impose, nonuniformity in quality assessments. This is the reason

[17] Recognizing this essentially subjective element, many people argue—like a University of California colleague of mine—that Consumers Union should not publish quality ratings per se.

[18] As footnote 14 indicates, about 50 per cent to 75 per cent of subscribers to *Consumer Reports* reported that they had purchased "top-rated" models for a varied set of products. The remaining 25 per cent to 50 per cent of subscribers either were unable to follow *Consumer Reports*' ratings (due perhaps to the nonavailability of the top-rated model) or chose not to do so.

[19] Adelman and Griliches [1, page 539] claim the conversion of "subjective" notions of quality to something objective as a property of the hedonic approach. Using the "quality" of milk as an example, they assert: "Note that, in this process, the subjective notion of the 'quality' of a particular type of milk has been quantified by a specific combination of objectively measurable or rankable traits."

for including an "*i*" superscript in the formula for quality. Three reasons support this procedure. First, as argued above, knowledgeability affects quality assessments, and knowledgeability will probably differ among individuals. Second, as just argued, the weights attached to characteristics come from the utility function, and individuals' tastes for various characteristics may differ sufficiently to produce different quality assessments of identical specimens. Finally, even if these two reasons were not compelling, the spirit of consumer sovereignty would support the notion that individuals should be permitted to make different quality assessments.

The possibility of different assessments raises a most important empirical question: To what extent, in fact, do different individuals make different quality assessments over identical sets of specimens? We may recall that the student assessors cited at the outset of this paper felt that the quality assessments of other individuals of the identical set of specimens would be closely similar to their own. On the other hand, we have substantial evidence that many readers of *Consumer Reports* buy models other than those top-rated by *Consumer Reports*.[20] Some acting in this way may do so because their weights, or tastes, differ from those adopted by CU's testers.

Additivity

The definition of quality in equation 1.0 embodies additivity on the grounds that (1) additivity is intuitively attractive, and (2) a simpler formulation is preferable to a more complex one. Only after strong evidence is offered regarding the unacceptability of the simpler formulation should one move toward a more complex formulation.

VI. RELATED CONCEPTS: (1) CHARACTERISTICS

Some of the points made below have been hinted at earlier. Clarity, however, demands that they be dealt with explicitly, even at the cost of redundancy.

Characteristics As Services

Characteristics, repeating our earlier definition, are the "*services* giving rise to utility."[21] But why should they be defined in terms of

[20] See footnote 18.

[21] One of the leading hedonists, Triplett, also opts for the definition of characteristics as services [32, pages 9–10].

Lancaster does not state whether characteristics are services or things. But an extensive example in which the characteristics of an automobile are analyzed deals mostly

services rather than the more objective factors—inputs—which "produce" them? There are several reasons.

First, it is the services, we believe, which consumers really want, not the means which produce them. They want the cool air and the comfort resulting from the operation of an air conditioner, not the air conditioning apparatus itself. Similarly, consumers want safety, not disc brakes per se.

Second, because (sometimes) many different inputs are required to produce a given service, the definition of characteristics as services will reduce the number of characteristics to be considered and thus render measurement easier. For example, a characteristic such as the "safety" of an automobile might have the following inputs (and more!): brakes (stopping distance, fade resistance, lack of "pull"), tires, interior layout (effect on crash survival), locks, emergency handling, layout of controls, visibility (location, size and reflection properties of glass areas). As we shall see a few pages hence, sixty-one or so inputs (or subcharacteristics) in a modern automobile reduce to eleven service characteristics. If individuals assign characteristic scores for the characteristic as a whole without separately evaluating each input or each subcomponent of a characteristic, then a reduction in the number of characteristics will render measurement easier.

A third reason for preferring to view characteristics as services is that the process of producing a desired service may be complex, marked by nonlinearities or perhaps multiplicative relationships. Under such circumstances it may be easier and more accurate to measure the output (the service rendered) than the inputs.

In some cases, there may exist a rather simple and dependable relationship between inputs and the service outputs. In such cases, it becomes a matter of convenience as to whether the assessor seeks to measure the characteristics directly, or by the inputs which produce them. The assessment of the quality of a mattress provides a convenient example. One assessor identified four "characteristics": (1) spring construction, (2) handles, (3) firmness and edge support, (4) surface conformity to body. It seems clear that "spring construction" was thought to be closely related to the service characteristic, "durability"; that "handles" were proxies for "portability"; that "firmness" and "surface conformity" were both viewed as components of "comfort."

It should be noted that the product-testing organizations such as

with characteristics which are services [21, page 170]. However, Lancaster is adamant in asserting that characteristics are objective and that, therefore, intrinsically subjective characteristics, such as beauty lie outside his domain of analysis [21, page 114].

Consumers Union frequently utilize the measurement of inputs as a proxy for service characteristics. But it should be remembered that their taste is necessarily different from that of the ordinary consumer. In order to protect the organization's reputation and financial integrity, they must confine themselves to products for which reproducible information provides an acceptable approximation to quality. Their reproducible information may include the results of product tests, or reports of experience (frequency of repair data, satisfaction with insurance settlements), or judgments (taste tests for beers or foods) from either panels of experts or from probability samples of a relevant population.

Characteristics Subjective, Nonuniversal

We will assume that the amount of a characteristic associated with a particular specimen is subjective and nonuniversal. By this, we mean that each individual will decide for himself, on the basis of whatever information he possesses, the extent to which a particular specimen possesses a given characteristic. This assumption does not preclude the possibility that for some characteristics of some specimens, objective evidence regarding the amount of the characteristic exists and is so compelling that it is universally accepted.

What we do argue is that objective data regarding characteristics are unusual and that for most characteristics of most products, individuals will not possess such objective evidence. Or, even if they do possess it, they may choose to accept it wholly, in part, or not at all.

Consumer Reports, with a current circulation of 2,100,000 and a maximum estimated audience of approximately ten million, is the most important producer and distributor of objective data on characteristics and quality.[22] Still, it reaches at most but one out of six families. And, in a year, it tests about 120 products.

The Optimal Level of Abstraction

A vexing problem in the measurement of characteristics is the determination of the optimal level of abstraction. Consider the performance of a soprano. Should we consider "beauty" as the characteristic? Or, assuming that beauty in a vocal performance has as its components such things as "color" and "range," should we consider *each compo-*

[22] *Consumer Research Magazine* also provides objective data on a wide variety of products. Its circulation is approximately 100,000. Other publications such as *Motor Trend* (circulation of 610,000), *Car and Driver* (circulation of 619,000), and *U.S. Camera* distribute both subjective and objective information to specialized audiences.

nent of "beauty" as a separate characteristic? My recommendation is that each assessor use that level of abstraction or aggregation at which he feels he can make the most accurate judgment, overall. My guess would be that more knowledgeable assessors would be aware of more components and would feel that greater accuracy would be achieved by judging each component separately and then combining assessments on the various components.

Negative Characteristics

Some characteristics, which we denote as *negative characteristics,* are undesirable. Ugliness, which goes well beyond the absence of beauty, would be such a characteristic. In the quality formula, we would attach to such a negative characteristic a negative weight, but a positive characteristic score.

The Characteristics of an Automobile

Perhaps nothing communicates as effectively as a concrete example. In this spirit, an illustrative set of service characteristics for an automobile are presented in Table 2. Several comments are in order.

Even for a good as complex as an automobile, the number of characteristics is "small" enough to be easily manageable. However, the example does not include the characteristics of the seller (reliability, convenience, etc.) although our theory calls for this. As a rough guess, seller characteristics might double the number of characteristics to be dealt with.

The information required to score each subcomponent separately is considerable. Little wonder that so many better-educated consumers turn to *Consumer Reports!* Given these considerable information requirements, it is only realistic to expect that most consumers will assess quality in a very rough, nonquantitative way.

Table 2 represents one person's supposedly exhaustive set of characteristics for an auto. It seems highly likely that the reader might be inclined to partition the characteristic space differently: to use different names, or assign different components to the same major component. Additionally, it seems likely that the reader may conceive of major characteristics or components which he feels relevant, but which were omitted from this set and hence given an implicit weight of zero. In practice, the partitioning of the characteristic space by one individual will necessarily be arbitrary and omissions of relevant characteristics will be inevitable.

The clear implication is that most consumers do not have access to

TABLE 2

The Service Characteristics of an Automobile:
A Tentative Taxonomy [a]

Characteristic	Characteristic Component or Input
1. Economy-Durability [b]	Operating Costs: Gas-oil use Repairs – Frequency and average bill (parts and labor) Parts cost and accessibility Geographic access to servicing Insurance costs Capital Costs: Depreciation – Expected loss of market value due to obsolescence Wear Deterioration due solely to passage of time (not use) Warranty Security against theft Time off the road for repairs
2. Comfort	Temperature control and ventilation (heating, air conditioning) Noise levels Ride Space (legroom, etc.) Seating: Comfort (padding, shape, angle, adjustability) Height
3. Performance	Maneuverability (size) On the road versus local Acceleration: From stop Passing Control Shifting Handling Speed

TABLE 2 (continued)

Characteristic	Characteristic Component or Input
4. Convenience	Steering (ease and precision)
	Ease of entry and exit
	Access to controls
	Storage:
	Baggage
	Glove compartment
	Ease of cleaning
	Ease of starting
	Application of power shifting
	Opening and closing of windows (manual versus electrical operation)
	Ease of parking:
	Steering
	Size
5. Safety	Collision absorption
	Visibility:
	Forward
	Side
	Backward (rear window defogger)
	Day/night
	Brakes:
	Stopping
	Fade resistance
	Emergency characteristics
	Tires
	Layout—effect on crash survival
	Locks
	Emergency handling
	Layout of controls
	Structural strength and rigidity
	Size
6. Aesthetics	Lines
	Color
	Finish
7. Status	Technical virtuosity
	Opulence
	Scarcity
8. Carrying Capacity: People	Number
	Effect on handling characteristics (an interaction)

(continued)

TABLE 2 (concluded)

Characteristic	Characteristic Component or Input
9. Carrying Capacity: Things	Usable cubic footage
	Flexibility for carrying odd shapes
	Durability of carrying surfaces
10. Pollution Effect	Noise
	Exhaust fumes

[a] Compare this set of characteristics with those used by Lancaster [21, page 106] in an illustrative computation of his technical efficiency approach: accommodation (size and comfort of seating, etc.), ride qualities, handling and steering, engine (quietness and performance), brakes, frequency of repair record (based on model of previous year), manufacturer's suggested retail price. Lancaster abstracted his characteristics from *Consumer Reports*.

[b] From a more sophisticated view, the durability of a specimen might be more carefully defined as "the weighted average of the rate at which characteristic scores do *not* decline over time" or, equivalently, as "the subjectively weighted average of quality over time." That is, if a specimen continued to provide service characteristics at the initial rate, it would be perfectly durable. By this interpretation each characteristic would have its own weight which might or might not have a time subscript. Each characteristic score—see equation 1—would have its own time subscript.

In the author's view, this more sophisticated definition is too complex to be operational. For this reason, economy-durability is defined more judgmentally, as above.

objective data on characteristics even if they desired it. They must perforce make subjective judgments regarding the extent to which specimens possess a given characteristic. Until empirical research suggests otherwise, it would be fatuous to *assume* that characteristics are objective and uniform.[23]

VII. RELATED CONCEPTS: (2) PRODUCT

It may be helpful to repeat background assumptions made earlier. We assumed that the consumer has already undertaken a personal "survey" of the activities in which he might engage, and that this "survey" has resulted in a commitment to purchase some specimen in a given product class, subject to a maximum outlay. The quality assessment will help determine *which specimen* to purchase, *not whether to spend* on this product category.

[23] Nonetheless, this is the assumption adopted by Lancaster in his technical efficiency approach. Lancaster declares [21, page 114]: "It is essential that the characteristics be an objective, universal property of the good (or activity)." This assumption was criticized by John S. Chipman [4, page 46] who asserted "that this assumption is a postulate and not a consequence of Lancaster's scheme."

The product class consists of *all* of the goods from which the consumer might choose for a particular purchase. Estimates of P^* (the quality-deflated price) will be made over this entire population or, more likely, a sample which is representative of this population. Realistically, the purchasing consumer is likely to be aware of and to purchase from only a small subset of this population.

The concept of *product* is needed to decide which specimens are appropriately included in the population for which quality evaluations and comparisons of P^* are to be made.

We define a *product* as "the set of goods which, assuming perfect information regarding their characteristics and money prices, would in the consumer's judgment serve the same general purpose for some maximum outlay." Again, clarifying comments are in order.

This concept is personal and subjective. Each individual must decide for himself which specimens are sufficiently similar to be seriously considered for possible purchase.

The perfect-information assumption enables us—the economist or the empirical researcher—to identify *all* specimens which are *potentially* relevant to this consumer's purchase decision, even though, realistically, this consumer may remain unaware of some of them.

Presumably, different specimens within a product class would possess rather similar characteristics.

The maximum-outlay specification effectively keeps the consumer from possibly violating a self-imposed budget restraint which he has adopted at an earlier stage of the purchase process. For example, on grounds of size, maneuverability, and other characteristics, a Mercedes sedan might qualify as a "compact sedan." For some consumers, however, the $7,000 plus price of the Mercedes would exceed their budget constraint and thus effectively purge the Mercedes from the "compact sedan" product group.[24]

In practice, it would appear that the delineation of product classes has posed few problems. None of the thirty or more student investigators undertaking quality assessments in 1970 reported any difficulty in determining which specimens were appropriately included, and which were appropriately excluded, from the relevant product classes.

[24] It is interesting to see how the related literature deals with the concept of "product." Lancaster [19, 21] does not define a concept of "product."

Not surprisingly, in the hedonic literature, the *product* is market determined. Triplett places different varieties in the same product class if (1) they share the same set of characteristics, and (2) the implicit characteristics prices are the same [31, page 14].

VIII. RELATED CONCEPTS: (3) MARKET [25]

While the concept of *product* enables us to say *which specimens* are appropriately included in the evaluation/comparison class, a concept of *market* is needed to say *which sellers* and *which buyers* are appropriately included.

In terms of its functioning, a *market* may be defined as "the area within which the price of a commodity tends to uniformity, allowance being made for transportation costs" [30, page 85]. Substitute "quality" for "transportation" and acknowledge that prices must be adjusted for quality, and we have a *functional* definition which is suitable for our purposes.

Our problem is to operationalize the above conception by developing a further definition which enables us to decide which sellers and which buyers are appropriately included in the same "market."

What we need is the population of sellers from whom a particular consumer *might* have purchased, not the set which he actually considered. In seeking to identify the "from-whom-might-buy" population, we assume that our sample consumer possessed perfect information regarding the probability distribution of the net payoff, i.e., the gross payoff less the cost of the search. We assume further that he followed the marginal rule: "Undertake additional searches as long as the expected gross payoff from a given search exceeds the expected cost of that search." Any seller "discovered" by an infinite number of repetitions of this rational search process would be appropriately included in the market. This procedure gives us the set of sellers in the market defined for a particular consumer.

Now we seek to identify the buyers appropriately included in this market. Any consumer who, by applying the same marginal search rule, might have "discovered" *any* of the above-listed set of sellers would be counted in the same "market."

Summing up, a *market* may be operationally defined as "the set of sellers which might be 'discovered' by Consumer A in applying the rational search rule *plus* the set of buyers who, pursuing the same rule, might have purchased from A's set of sellers. In identifying relevant sellers and buyers, it is assumed that all consumers possessed perfect information regarding the probability distribution of net payoffs to search."

[25] Neither the hedonists nor Lancaster define a *market*. In practice, however, the hedonic approach has been applied to removing quality effects from price indexes for entire economies. In effect, therefore, means are estimated over the aggregate of all local markets.

It would have been most gratifying to report that we succeeded in specifying an acceptable set of search procedures which readily identified the particular sellers and buyers to be included in a market.

Unhappily, such was not the case. Despite much racking of brains and consulting of colleagues, we were unable to specify an acceptable set of search procedures beyond the general statement of the marginal rule.

Ironically, as in the case of the concept *product*, practice has proved easier than theory. Student investigators reported no problems in determining which sellers should be included or excluded from a particular market. And they were strongly confident that other investigators would decide as they had.

Some "qualitative" observations may prove useful. Like the concept of product, this concept, too, is personal and subjective. However, it is our guess that the set of sellers and buyers assigned to the same market by different, but similarly situated, consumers will turn out to be highly similar.

As an example of the market concept, it may be useful to examine the illustrative market for sofa beds cited at the beginning of the paper. The sellers included four retail outlets in Saint Paul (where the investigator lives): the dominant department store, a local mail-order outlet for a national chain, a local discount-type store, a local "high-class" home furnishings store. Since each establishment or branch has the authority to modify prices and choose the brand/models it offers, it is the establishment or branch's customers who make up the buyer side of the market. In this case, this outlet of the dominant department store may attract customers from several hundred miles away. Thus, the consumer side of this market would be more geographically dispersed than the seller side. Note, however, that the market would not include *all* people within a given radius of Saint Paul. Instead, it includes only those who might actually have purchased a sofa bed from the four retail outlets *while following the marginal rule.*

A final observation concerns the application of the search rule. One must differentiate between purchases in which the product in question bears the full marginal cost of the search versus a purchase which rides piggyback on some other "more important" purchase(s). An example of a full marginal-cost-bearing purchase might be the toothpaste purchased by someone who had just arrived at a hotel or motel late in the evening, toothpasteless. By contrast, a partial-marginal-cost purchase would occur when a person engaged in a major shopping trip added toothpaste to his shopping list. Which of these situations yielded the

larger "market" (more sellers) would depend upon the number and geographic clustering of sellers relative to the location of the consumer for whom the market is defined.

IX. AN AGENDA FOR FUTURE RESEARCH

The concepts proposed in this paper were formulated explicitly with an eye to future empirical measurement. Set forth below are my priorities regarding alternative research projects which might stem from the conceptual framework proposed in this paper. This listing is made without testing the reader's patience by specifying how the proposed research might be carried out.

My first-priority candidate would be a large number of "demonstration" measurements, seeking to assess the quality of widely variegated products. The purpose would be to demonstrate the feasibility of measuring quality along the lines proposed here and to ferret out problems associated with it. Any such measurements should include follow-up interviews with the quality assessors to ensure that they understood the relevant concepts and that the measurements taken conformed to the model specified.

In case a quality assessor purchased a specimen of the product under investigation, one would want to know whether he purchased a specimen on the efficiency frontier. An off-frontier choice would seem to call for an explanation.

Of almost equal importance would be research designed to ascertain the extent to which quality assessments of an identical set of specimens are "uniform." In its strongest form, "uniformity" of quality assessments would call for the designation of identically sited frontiers by different assessors. Next strongest (but still highly satisfactory) would be the designation of efficiency frontiers consisting of identical specimens in the same left-to-right sequence, but with different widths. Less pleasing, but still interesting, would be an efficiency frontier consisting of the same set of specimens, but containing some reversals of order.

Even if different individuals differ in their quality assessments, it would still be interesting and useful to obtain some measure of the extent of variation of prices from one individual's frontier.

A fallback position on the uniformity-of-assessment front might center on the distinction between "primarily objective" and "primarily subjective" characteristics. *Objective* characteristics would be those which are subject to reproducible tests, e.g., the durability of a sofa bed. Objective characteristics might also include characteristics for

which judgments by panels result in "little" variance. Subjective characteristics would include all others. In the sofa bed example, the "aesthetic" aspects of a sofa bed might represent a subjective characteristic. The objective of this fallback research would be to ascertain through correlation techniques the extent (if any) to which reproducible measurements of objective characteristics dominate overall quality judgments. (Research in this direction was suggested to me by Richard Emmerson of the University of California, San Diego.)

Two other potential projects depend upon a finding of considerable uniformity in the making of quality assessments. *If* extensive research shows a considerable degree of uniformity in quality assessments for informed consumers, then these two tasks assume relevance: (1) an attempt to assess the "efficiency" of consumption of low-income versus high-income households; (2) the development of market demand/supply relationships taking account of quality.

Embodied in the first is the notion that low-income households, being less amply endowed—both genetically and culturally—on the average, may, in purchasing, use their income less effectively than high-income households. More specifically, it might be hypothesized that low-income households while spending in a given product category may obtain less quality per dollar of outlay. If such a finding should be confirmed, it would suggest that current data on the distribution of money income understate the "true" degree of inequality.

REFERENCES

[1] Adelman, Irma, and Griliches, Zvi. "On an Index of Quality Change." *Journal of the American Statistical Association* 56 (September 1961): 535–548.

[2] Akerlof, George A. "The Market for Lemons: Quality, Uncertainty, and the Market Mechanism." *Quarterly Journal of Economics* 84 (August 1970): 488–500.

[3] Becker, Gary S. "A Theory of the Allocation of Time." *Economic Journal* 75 (September 1965): 493–516.

[4] Chipman, John S. "Discussion of Lancaster's 'Change and Innovation in the Technology of Consumption.'" *American Economic Review* 56 (May 1966): 45–49.

[5] Chamberlin, Edward H. "The Product as an Economic Variable." *Quarterly Journal of Economics* 67 (February 1953): 1–29.

[6] ———. *The Theory of Monopolistic Competition.* Cambridge: Harvard University Press, 1933.

[7] Court, A. T. "Hedonic Price Indexes with Automotive Examples." In *The Dynamics of Automobile Demand.* New York: General Motors Corporation, 1939, pages 99–117.

[8] Dawes, Robyn M.; Coombs, Clyde H.; and Tversky, Amos. *Mathematical Psychology: An Elementary Introduction.* Englewood Cliffs, N.J.: Prentice-Hall, 1970.

[9] Dhrymes, Phoebus J. "Price and Quality Changes, in Consumer Capital Goods: An Empirical Study." In Zvi Griliches, editor, *Prices Indexes and Quality Change*. Cambridge: Harvard University Press, 1971, pages 88–175.
[10] Dorfman, Robert, and Steiner, Peter O. "Optimal Advertising and Optimal Quality." *American Economic Review* 44 (December 1954): 826–836.
[11] Fisher, Franklin M., and Shell, Karl. "Taste and Quality Change in the Pure Theory of the True Cost-of-Living Index." In Zvi Griliches, editor, *Price Indexes and Quality Change*. Cambridge: Harvard University Press, 1971, pages 16–54.
[12] Gabor, André, and Granger, C. W. J. "Price as an Indicator of Quality: Report on an Enquiry." *Economica* 33 (February 1966): 43–70.
[13] Georgescu-Roegen, Nicholas. *Entropy as an Economic Process*, Cambridge: Harvard University Press, 1971.
[14] Griliches, Zvi (ed.). *Price Indexes and Quality Change*, Cambridge: Harvard University Press, 1971.
[15] Henderson, James M., and Quandt, Richard E. *Microeconomic Theory: A Mathematical Approach*, 2nd ed., New York: McGraw-Hill, 1971.
[16] Hirshleifer, Jack. "Where Are We in the Theory of Information?" *American Economic Review* 63 (May 1973): 31–39.
[17] Jacoby, John: Olson. Jerry C.; and Haddock, Rafael A. "Price, Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality." *Journal of Applied Psychology* 55 (December 1971): 570–579.
[18] Krantz, David H. "Measurement Structures and Psychological Laws." *Science* 175 (March 31, 1972): 1427–1435.
[19] Lancaster, Kelvin J. "A New Approach to Consumer Theory." *Journal of Political Economy* 74 (April 1966): 132–157.
[20] ———. "Change and Innovation in the Technology of Consumption." *American Economic Review* 56 (May 1966): 14–23.
[21] ———. *Consumer Demand*, New York: Columbia University Press, 1971.
[22] Maynes, E. Scott. "Consumerism: Origins and Research Implications." In Eleanor Sheldon, editor, *Understanding Family Behavior*. Philadelphia: J. B. Lippincott, 1973, pages 270–294.
[23] ———. *Decision-Making for Consumers*. New York: Macmillan, 1976.
[24] ———. "The Power of the Consumer." In Burkhard Strümpel, James N. Morgan, and Ernest Zahn, editors, *Human Behavior in Economic Affairs*, Essays in Honor of George Katona. Amsterdam: Elsevier, 1972, pages 399–420.
[25] Morris, Ruby Turner, and Bronson, Claire Sekulski. "The Chaos of Competition Indicated by Consumer Reports." *Journal of Marketing* 33 (July 1969): 26–34.
[26] Samuelson, Paul A. *Economics*. 9th ed. New York: McGraw-Hill, 1973.
[27] Scitovsky, Tibor. "Consequences of Judging Quality by Price." *Review of Economic Studies* 12 (1944–1945): 100–105.
[28] Stigler, George J. "The Development of Utility Theory, I." *Journal of Political Economy* 58 (August 1950): 95–113.
[29] ———. "The Economics of Information." *Journal of Political Economy* 69 (June 1961): 213–225.
[30] ———. *The Theory of Price*. 3rd ed. New York: Macmillan, 1966.
[31] Theil, Henri. "Qualities, Prices and Budget Enquiries." *Review of Economic Studies* 19 (1951–1952): 129–147.
[32] Triplett, Jack E. *The Theory of Hedonic Quality Measurement and Its Use in Price Indexes*. Bureau of Labor Statistics, Staff Paper 6. Washington, D.C.: Government Printing Office, 1971.

Comments on "The Concept and Measurement of Product Quality"

F. THOMAS JUSTER

NATIONAL BUREAU OF ECONOMIC RESEARCH
AND UNIVERSITY OF MICHIGAN

SUMMARY OF THE MODEL

IN this paper, Maynes essentially argues the case for a cardinal scale of overall quality measurements, designed to achieve these objectives:

1. to measure the informational effectiveness of markets;
2. to measure the payoff to search; and
3. to create a building block for demand theory.

The cardinal measurement of quality is derived from essentially ad hoc judgments of individuals, using a fixed weighting scheme to determine the relative importance of product characteristics and individual consumer judgments about the extent to which different products embody specified characteristics. Quality scores for characteristics are then combined into an index, defined as ranging between 0 and 1. Finally, the quality scores are used to estimate deflated or quality-adjusted prices.

The basic philosophy behind Maynes's concern with quality measurement and with the construction of quality indexes is an essentially normative presumption that markets function very imperfectly, and that development of quantitative information on quality would permit consumers to make different choices that would come closer to optimizing their welfare, given the budget constraint. But to prove that case, it is necessary to show that the observed variation in money prices cannot be attributed to variation in quality, or alternatively, that there is significant variation across products in quality-adjusted prices. That is, in order to demonstrate that quality measurement is important, it is essential to show that existing market decisions reflect significant imperfections resulting from lack of knowledge about quality.

The alternative model of consumer markets, which is widely accepted by economists, is, of course, that differences in prices for functionally comparable products at a single point in time simply reflect differences in quality as perceived by buyers. Hence, we usually proceed to use price differences in cross section to measure changes in average quality over time, as in the hedonic price index literature.

To me, there seem to be two basic flaws in the measurements that Maynes provides, and in the inherent structure of the way he proposes to go about inquiring whether markets function with serious imperfections. Suppose that, as the data in his Table 1 show, there is a significant positive association between the quality rankings and the money prices. Then, whether the quality-adjusted prices show a significant amount of variation across products or only small random variation will depend largely on the scaling used to measure quality. Taking the extreme case, suppose that the rank correlation between price and quality were perfect. Then, there must be some way of scaling the quality variable so that the observed variation across products would be zero, and any other way of scaling would provide either a positive or negative association between the original money prices and the quality-adjusted prices.

If price and quality scores are not perfectly correlated, then no utility transformation of the quality scale will provide the result that quality-adjusted prices are equal across different specimens. Instead, it must be true that some consumers are buying less quality than they could get at the same price, or more generally, paying a higher price for less quality. That provides prima facie evidence of imperfect functioning of markets and suggests that gains in social welfare would accrue as a result of an index of product quality.

However, even a less than perfect ranked correlation between price and quality indexes cannot be interpreted as necessarily reflecting market imperfections if one assumes that the relative importance (weights) of certain characteristics vary across households. In terms of invariance among households it is far more plausible to suppose that judgments about the quantitative differences among products in quality or characteristic dimensions is invariant among households than to suppose that the relative importance attached to certain characteristics is also invariant. As Maynes himself points out, there are solid reasons for supposing that the circumstances of particular individuals (location, age) are very likely to result in differential importance for particular product characteristics and, hence, in a weighting scheme which is not universally applicable to all groups of individuals or households. For

example, relatively high-income households are quite apt to place quite heavy weight on product characteristics which minimize shopping time and provide insurance against malfunction, rapid and convenient delivery and/or installation, and so on. Specimens which are relatively high priced because they include substantial components of these types of sellers' services would not, in Maynes's scheme, show up with a higher quality rating than a specimen with more of the desirable product characteristics but very little desirable seller characteristics. And in that case, the rank correlation between price and quality would be far from perfect.

Unless one is willing to make strong assumptions about the cardinality of characteristic scores (or utility derived from characteristics), and about invariance among the population in weights assigned to characteristics and in the assessment of characteristic scores, the model in Maynes's paper will not have any empirically verifiable content. The paper actually contains two models: one described as the Direct Measurement Model, the other described as the Proportionality Model. The only difference between them is that the Proportionality Model says that quality is proportional to the weighted quantity of characteristics of different specimens, while the Direct Measurement Model says that quality is proportional to the weighted sum of the utilities conveyed by these same quantities of characteristics. The Proportionality Model, being more objective, is less subject to differences in individual judgments about the value of additional units of a given quality characteristic. For example, the Proportionality Model might have as an ingredient the horsepower of automobiles—or perhaps more appropriately, the time it takes to accelerate to a given speed during a fixed time span. However, the Direct Measurement Model would require consumers to judge the gain in utility from having more acceleration, not just the objective facts of differences in acceleration among vehicles. While it is no doubt true that proportional differences in a characteristic like acceleration do not provide proportional differences in utility, it is also true that there must be substantial differences among people in the relative importance ascribed to acceleration—the middle-aged shopper is not likely to value that characteristic very highly, whereas the teen-age hot rodder might regard it as the dominant characteristic of automobile quality. There seems to be no solution to this problem and, hence, as Maynes himself recognizes, no uniform solution to the quality measurement problem.

Despite these difficulties, Maynes argues that it is useful and important to construct comprehensive quality indexes, and moreover, that

they can be constructed via weighting with marginal utilities rather than by quantities of different characteristics. However, on the basis of the evidence and argument presented in his paper, I am not persuaded that there is a great deal to be gained from pursuing the work of producing and refining a broad-gauge index of average product quality.

SOME SPECIFIC COMMENTS AND QUIBBLES

1. In the data shown in Table 1, Maynes notes that the range of quality-deflated prices is about as large as in original prices. However, the rank correlation between money price and quality is significantly positive, and the coefficient of variation in quality-adjusted prices is less than that in original prices. If one clear outlier is ignored, the coefficient of variation is cut by more than a third.

2. In the weighting schemes presented for both product and seller characteristics, the inherently subjective nature of the characteristics is their dominant characteristic. For product characteristics, aesthetics gets half the weight. In seller characteristics, sensitivity and pleasantness gets a third of the weight, while convenience gets another quarter. No weight is apparently given to such significant dimensions of seller characteristics as the range of models available for comparison; the time it takes to be served; whether or not delivery is available, and if so, at what cost; how long it takes to obtain delivery for a purchase; and so on. Presumably some of these characteristics are buried in the convenience terms, but that is not clear from the discussion.

3. The importance of Maynes's cardinal measurement assumption should be underlined—ordinal ranking will not suffice for his purposes. Only if all consumers assess characteristic quantities as the same across all products, have identical marginal utility functions for characteristics, possess perfect knowledge of market alternatives, and have uniform weights for all characteristics will it be true that the variations in p^* will be zero. While one can easily conceive of uniform judgments about the relative importance of various quality dimensions or about the marginal utility of increased amounts of characteristics, a consumer with a low wage rate and consequently a low value of search and shopping time is not going to place a high premium on retail establishments which minimize both, but at substantial cost. Nor does it seem plausible that aesthetics will receive the same relative weight as durability and serviceability for consumers with different levels of income.

4. A related point is the meaning of the end points on the Maynes quality scale. Is there really a zero point where the characteristic has

zero quantity or zero utility? Or a unity point where the characteristic reaches a limit that cannot be exceeded?

If not, and if the absolute *differences* in the observed quality scale are more accurate than the levels or the proportional differences, the observed results could be reversed: in the data, the maximum and minimum quality observations are in a 2:1 ratio on the present scale if seller characteristics are ignored, and on a 3:2 ratio if seller characteristics are included. But the price range has a maximum to minimum ratio of 3:1, hence one obtains a substantial range of quality-adjusted prices. But adding or subtracting a constant from the quality scale is capable of reducing the variation in p^* so long as there is a positive rank correlation between price and quality.

5. Re: market functioning and knowledge. Efficient functioning doesn't depend on complete knowledge for all units: it requires only sufficient knowledge to make inferior products unprofitable.

6. Re: other forms of empirical evidence that could be used. Suppose that buyers were shown to have typically visited x outlets in the process of acquiring a given product, or to shop typically at y different outlets for purchases of sundries? Would that not mean that any observed differences in p^* are more likely due to differences in judgments about quantitative characteristics or differences in the weights given to characteristics rather than to lack of knowledge about alternatives? In short, what does the empirical evidence look like regarding actual consumer search and experimentation?

7. Re: structure. Is additivity of characteristic scores plausible a priori? I submit not, since there must be threshold effects of essential characteristics without which large amounts of other characteristics are useless: a well-designed piece of furniture that won't last more than a month under normal usage isn't worth anything, whereas a well-designed article that has better durability characteristics may be worth a great deal.

8. Re: research strategy using the quality scale. How are quality characteristics correlated across commodities? If the correlations are high enough and positive, the whole measurement problem becomes enormously simplified. Maynes presents no data here, and does not even discuss the problem.

As a way of determining whether the notion of aggregate quality measurement is meaningful, why not test it by application to areas where one expects differences in the informational effectiveness of markets a priori? For example, there may be a lot or a little informational effectiveness *generally*, but there should be less than average

where: (a) the market area has a high degree of mobility among consumer units, and word-of-mouth judgments are therefore less important. (b) the market area has a relatively large fraction of poorly educated consumers, whose ability to process information is on the average less than that of consumers in other market areas. In these situations, the variance in quality-adjusted P should be higher than average.

An alternative strategy which would reach one of Maynes's objectives—increasing the informational effectiveness of markets—although it would not contribute to either theoretical developments or to measuring the payoff to search—would be to concentrate on the development of a better information system by which consumers could assess relevant data as an input into rationalizing purchase decisions and, indirectly, improving the functioning of markets. As I see it, the basic difficulty with existing consumer information systems is that, ordinarily, they are not relevant to the decisions that consumers actually have to make. To the extent that information about product characteristics and performance is available, it is subject to a substantial time lag between the time the data are collected and the time they can be used. Such data are quite apt to be product- and area-specific and of limited relevance to many potential users, and they are only available in a form which entails going through substantial amounts of nonessential information in order to find out what one would like to know.

CONCLUDING COMMENTS

Although the general tenor of most of these comments is critical of the notion that quality change can be measured in an effective way using the techniques described by Maynes, it should be noted that the presence of formidable conceptual difficulties does not necessarily warrant the conclusion that the procedures and measurements lack economic or social utility. It is one thing to argue, as I have above, that differences in the characteristic weights among the population make it difficult or perhaps impossible to draw conclusions about the efficient functioning of markets from the sort of data that Maynes has provided, and another to say that these differences are quantitatively important enough to make the information useless for the purpose of improving consumers' decision making. Most attempts to measure conceptually complex phenomena can be subject to the same type of critical comment, but it is not until someone has attempted to make the measurements, and then tried to use them, that the issue can be fully resolved. Hence, although my own judgment is that aggregate measures of prod-

uct quality along the lines described by Maynes will not really prove to be useful, it should be recognized that judgments of this sort are apt to be biased on the conservative side, and that there is some detectable proportion of cases in which they will prove to be wrong.

Comments on "The Concept and Measurement of Product Quality"

JACK E. TRIPLETT

BUREAU OF LABOR STATISTICS *

MAYNES commends his proposal for measuring quality for a number of purposes and asserts that it is different from alternative procedures, such as the hedonic method. His argument for the usefulness of his proposal for economic analysis consists merely of a recitation of problem areas where *some* way of dealing with the quality problem is needed; moreover, he makes no attempt to show how his proposal is related to, or exactly how it differs from, Lancaster's "New Theory of Demand," on the one hand, and the hedonic technique on the other.

I intend to compare Maynes and the alternatives, and will in the process also consider the usefulness—to economists—of his proposal. I agree fully with Juster's points, which deal mainly with the question of whether Maynes's proposals are likely to be of use for constructing better consumer information systems, and will therefore ignore that aspect of Maynes's paper.

Maynes assumes that "quality" is a concept that is inherently measurable, so that a suitable approach will yield a *scalar, nonmonetary* measure of quality. He also asserts that his measure is cardinal.

My own view of "quality"—and this view is explicitly advanced as the concept that underlies the hedonic method—is that quality itself is not a measurable concept, in the sense of obtaining a *scalar, non-*

* The views expressed are those of the author and do not represent an official position of the Bureau of Labor Statistics.

monetary quality indicator. Under this view (which I will henceforth refer to as the "hedonic" viewpoint), when we use the word *quality*, what we really are doing is making a kind of shorthand reference to the several quantities in a *vector* of characteristics. Under the hedonic view, there is no measure of quality, as such, because there is no way to combine directly the various elements of the characteristics vector—the problem is exactly analogous to adding apples and oranges. However, provided we can find an appropriate (implicit) price for each of the characteristics, it is possible to obtain a measure of the *value* of the vector, by valuing the quantity of each characteristic by its implicit price, and then combining the results (the most natural way is simply to add them up).

To reiterate the distinction in another form, under Maynes's view there is, in principle, a measure of quality comparable to the quantity measures which we are accustomed to use in measuring inputs and outputs—a scalar measure whose formation requires no monetary valuation. Under the hedonic view, a measure of "quality" can only be obtained by a process comparable to the construction of GNP – we can combine the apples and oranges (characteristics) into an aggregate called "fruit" (quality) only by valuing quantities of the individual fruits (characteristics) by market prices. We may speak of the resulting aggregate (be it "fruit," real GNP, or quality) *as if* it were a physical measure, and we often do so; but this should not disguise the fact that it has properties which are different from the properties of a purely physical measure of quantity.

Maynes's view and the hedonic view of the economic concept of "quality" are two very different ways of looking at the matter. It is of no value to argue which one is a better or more accurate perception of reality—possibly neither is very good.

But there is a straightforward answer to Maynes's plaint that those of us who have been working within the hedonic framework have failed to put forth a concept of measurable quality along the lines he proposes: for purposes of economic analysis and measurement, a scalar notion of measurable quality simply is not necessary. Even if we conceded that quality is intrinsically representable as a scalar, and even if we believed that Maynes has found the tool for measuring it (which I do not believe), we would have to find a way to attach a value to the quality index anyway. What we require—and, really, *all* we require—is a measure of the *value* of quality, and that, in the hedonic view, is feasible *without* a scalar quality index.

Moreover, the measurable quality view has some built-in limita-

tions that are avoided by the approach inherent in the hedonic or characteristics analysis. One of these is the tastes problem.

Consider two individuals (I and II) who are asked to choose between two different varieties (call them A and B) of a product which has only two characteristics (*r* and *s*). Suppose further that the proportions of characteristics *r* and *s* in variety A are different from the combination found in variety B, though both A and B sell at the same price, and that individual I prefers A to B, while individual II prefers B to A.

Maynes explicitly states that under his proposal, A is judged higher quality by I's assessment, but by II's, variety B is higher quality than variety A. Hence, nothing can be said, unless interpersonal comparisons can be made. Differences of taste of this kind make consistent quality measurement impossible under his system, and, moreover, no analysis of the situation can be carried out. This is a serious defect for economic uses, for the situation he describes must be regarded as an empirically pervasive one.

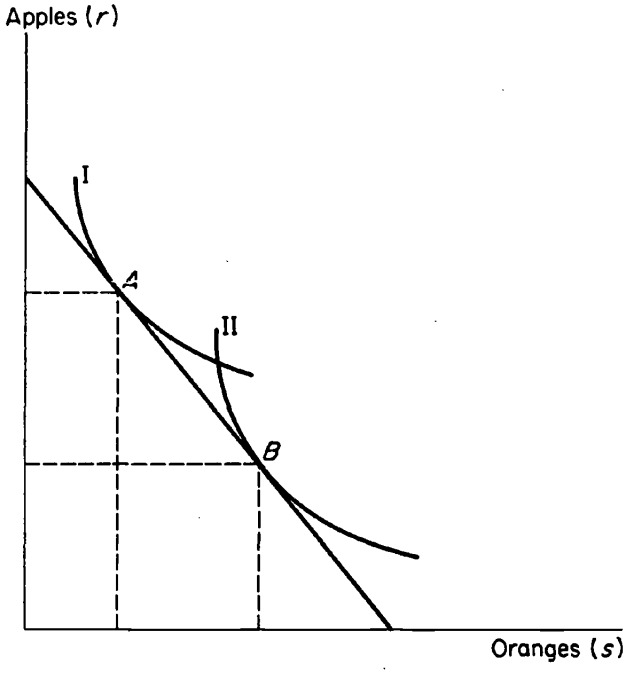
In contrast, the hedonic view does not require that tastes be anywhere near so uniform across consumers. Under the hedonic way of looking at quality, the situation described above could result from the preferences mapped in Figure 1. Given the relative implicit prices of *r* and *s*, and their respective indifference curves (I and II), individuals I and II locate at points A and B, which points correspond to the proportions of characteristics *r* and *s* contained in varieties A and B. There is nothing particularly surprising or perplexing about the situation diagramed. Budget allocations for goods and services as conventionally defined are frequently found to differ, and pose no massive problems for the analysis of consumer behavior. No greater problems are posed when the same thing occurs with respect to budget allocations for characteristics.

I would certainly agree that if we *have* to confront the issue of taste differences, little is really gained by evading the question. But that does not alter my preference for systems—when the choice presents itself—in which taste differences play a lesser, rather than a larger, role. Under the hedonic view, taste differences present major difficulties only for a small number of situations (for example, if individual I were to regard characteristic *r* as desirable, while II would pay to get rid of it).

I turn now to the question of whether Maynes is doing hedonics (or Lancastrian analysis) by another name. I think not, but a clear resolution of the question is impeded by the lack of precision in Maynes's specification.

FIGURE 1

Budget Allocations in a Goods World and in a Characteristics World



Maynes's quality definition is set forth in his equation 1.0, on p. 542. This can be written (dropping the superscripts)

$$G_k = \frac{\sum_l (w_l \cdot ch_{kl})}{\sum_l w_l}$$

where G_k is the quality of variety k of some product, w_l is the weight assigned to characteristic l (the weights "for convenience" sum to unity), and ch_{kl} is the "characteristic score assigned to the lth characteristic of the kth specimen" (Maynes, p. 543). Interpretation of what is meant by these weights and by the term "characteristics scores" is greatly hampered by persistent ambiguity in Maynes's text.

Maynes gives two specifications (pp. 543–544) for the *Ch*'s: (a) the "characteristic score" is the *quantity* of characteristic l (expressed as an index); (b) the "characteristic score" is the *marginal utility* of the

total quantity of characteristic l found in variety k. The weights are spoken of in the paper as "subjective" or "assigned," without a clear specification of what is meant. In discussion, Maynes has suggested two alternatives: (a) the weights are marginal utilities for characteristics (which interpretation is consistent with the example of radio receivers, discussed on p. 546); (b) the weights are the proportions of total expenditure on characteristics allocated to the various characteristics (the latter definition is probably the proper interpretation to place on survey responses, if one were to establish the weights by asking consumers to rate the importance of various characteristics in the purchase of a particular product).[1]

Some combinations of these various possible definitions of weights and characteristic scores are unreasonable. For example, characteristics definition (b) and weight definition (a) require us to weight marginal utilities by marginal utilities. Moreover, I cannot see any meaning to weighting characteristic scores (however defined) by expenditures on characteristics, so I rule out combining weight definition (b) with anything.

This leaves a relatively simple and familiar interpretation for Maynes's equation 1.0. Quality (at one point he writes "utility") is defined as the summation of quantities of characteristics, each of which is assigned an appropriate weight. Equation 1.0 is thus equivalent to a "branch" utility function, assumed to be additive, defined on characteristics as arguments—and, of course, with the simple additive form, the w_l's are marginal utilities.

In this form, Maynes's proposal may be contrasted with Lancaster's "New Theory of Demand" (1971), and with the hedonic quality measurement technique. Lancaster proposed that consumer theory be reformulated by entering characteristics, rather than goods, in the utility function (the utility function could be a very general one), and

[1] To revert to the grocery-cart simile advanced elsewhere, suppose a cart contains X pounds of corned beef and Y pounds of cabbage, and the total price of the grocery cart of food consists of the sum of expenditures (made at that purchase) on corned beef and on cabbage—i.e., $p_bX + p_cY$ (where the p's are implicit prices of corned beef and cabbage, respectively). If one asked the purchaser of this cart to assign a "subjective" weight to the importance of the two characteristics corned beef and cabbage in determining the value (or "quality") of the cart of food, the most likely answer would probably be based on the relative sizes of expenditures (i.e., on p_bX and p_cY). Some consumers, however, might have the number of pounds (X and Y) in mind, some the relative prices p_b and p_c and some may well be referring to something else. This underscores the problem of trying to evaluate consumer survey or opinion data of this particular type. One wonders exactly how the students in Professor Maynes's classes interpreted their task of determining weights.

explored the differences this would make to theory, and some of the implications to be drawn for empirical work. What Maynes has done is similar to Lancaster's approach in that both have written a utility function defined on characteristics. But where Lancaster correctly and properly wrote the utility function in a very general form (in the absence of any information that could specify its form), Maynes assumes a very specific utility function defined on characteristics—a utility function which, moreover, has severely restrictive properties.[2] And he proposes it, not as a theoretical or analytical tool, but as a measurement device! Hence, the connection with Lancaster's work is, as Maynes says, rather remote, but hardly for the reasons Maynes gives.

The hedonic technique is employed as a measurement tool, so in terms of function, it is directly comparable with Maynes's proposal. In the hedonic technique, the price of a product variety (or "specimen" to use Maynes's term[3]) is viewed as simply a sum of expenditures on characteristics. That is, given the implicit price of characteristic l (p_l), and the quantity of characteristic l embodied in a particular product variety (Q_{lk}), then (assuming for simplicity a strictly linear specification of the hedonic function) the product ($p_l Q_{lk}$) represents the expenditure on l when variety k is purchased, and

$$p_k = \sum_l p_l Q_{lk} \qquad \text{(I)}$$

equals total expenditure on all characteristics when variety k is purchased.[4]

It is a mistake to assume, from the superficial similarity of the form of equation I and Maynes's equation 1.0 that they are indeed similar. Characteristics (Q_{lk} in equation I, or Ch_{lk} in Maynes's equation 1.0) are measurable, in principle, if the right kind of data are available. The

[2] Such as additivity, which Maynes finds "intuitively appealing"; economists who have worked on consumer demand models regard the implications of additivity as a serious liability. As Goldberger (1967, p. 31) remarked: "Direct additivity may be a plausible specification when goods are defined broadly, but not when one works with a fine classification of consumer expenditures."

[3] Many hedonic studies have not considered store services as an attribute or characteristic, mainly because they have been carried out using list prices or some other data source which contains no information on store services, or on store-to-store price variation. In principle, however, the hedonic technique would encompass such factors where appropriate, and where data were available.

[4] Other forms of the hedonic function require modification of this statement, but the basic nature of the disaggregation described in it holds for most forms which have been used in the hedonic literature.

hedonic method and Maynes's proposal both require this. But in addition to measurable characteristics, the hedonic technique requires only that prices for different varieties be available, and from this the implicit prices—the p_l's—are estimated. Maynes, on the other hand, requires information on the *utilities* of characteristics.

In other words, if cardinal utility exists, and if someone invents a "utilometer" to measure it, we could implement Maynes's proposal. Of course, then we wouldn't have to, because we could also measure the utility of variety k directly, without recourse to characteristics at all, and without bothering about establishing a functional form for the branch utility function linking characteristics to goods.

In short, the Maynes proposal contains nothing that has measurable implications, and the "measures" he purports to have made cannot be interpreted as anything but ad hoc judgments, without scientific validity either as numbers or as a procedure.

Finally, I think it important to reiterate the distinction between the idea that quality is intrinsically measurable in a scalar, nonmonetary form (which is the premise that Maynes labors under), and the opposing view—namely, that what we mean by the term "quality" is no more than a nonspecific and nonquantitative shorthand reference to a vector of characteristics. I would like to be as conservative as Juster in eschewing "impossibility theorems." But I really believe that pursuing any notion of scalar, nonmonetary quality measurement is following a will-o'-the-wisp, for the reason that it will inevitably lead right down the road that Maynes has followed—a road which ends in a tollgate labeled "measurable, cardinal utility."

We have, it is true, notions that consumers assess the quality of different product varieties, and decide that one is higher or lower quality than another; moreover, we often think of situations where a consumer decides that, even though one product variety may be higher quality, it is not worth the price differential asked for it. It is therefore tempting to assume that these actions must imply behavior in reaction to some unique, scalar measure of quality inherent in varieties of goods—a scalar measure that can be set against price in the consumer's decision-making process.

These consumer decisions, however, can be given an interpretation within the framework of behavior toward characteristics (using the analytic tools of Lancaster); the characteristics interpretation of consumer behavior dispenses with a measure of quality as such, in favor of an optimization process similar to the standard theory of consumer behavior, but conducted over the characteristics of goods (rather than

on the goods themselves, or the "quality" of goods). If we have data on characteristics of goods (required under Maynes's proposal, as well as for Lancastrian analysis and hedonic estimation), we can study demands for characteristics directly (as King has done, in the paper in this volume), without finding it necessary to "deflate" prices by a "quality index."

One of the major accomplishments of the characteristics approach to consumer behavior is that it has relieved us of the necessity for producing a scalar quality measure in order to analyze the phenomenon we call "quality." The characteristics approach to consumer behavior—and to the interpretation and measurement of quality—has, it is true, ambiguities and inchoate aspects of formidable magnitude. But even though that path is not yet thoroughly explored, and though the measurements it yields are no doubt imperfect and inadequately understood, still it seems to me to have much more promise than attempting to implement any notion of a scalar quality index. Imperfect or not, the characteristics approach to quality has yielded both measurement and analysis, and has advanced our understanding of the quality phenomenon. The scalar quality measure approach has no accomplishments to recommend it, even though it has been around far longer. I doubt if it will yield anything of value even if much more work is expended upon it.

REFERENCES

Goldberger, Arthur S. "Functional Form and Utility: A Review of Consumer Demand Theory." Social Systems Research Institute, University of Wisconsin, Systems Formulation, Methodology, and Policy Workshop Paper 6703 (October, 1967).

Lancaster, Kelvin. *Consumer Demand: A New Approach.* New York and London: Columbia University Press, 1971.

Reply to F. Thomas Juster and Jack E. Triplett

E. SCOTT MAYNES *

SENSING economies of scale, I have framed a single reply to Juster and Triplett. The two of them direct three major criticisms toward my paper:

1. *Juster doubts that local consumer markets are informationally imperfect,* contrary to Maynes. It follows therefore that "differences in prices for functionally comparable products at a single point in time simply reflect differences in quality as perceived by buyers."

2. *Juster expects different individuals to have varying preferences for different characteristics.* Hence, different individuals will make differing assessments of the quality of an identical specimen and, for this reason, one individual's overall measure of "quality" will not be useful to another.

3. *Juster doubts and Triplett denies that the measurement of quality,* according to the cardinal formulation proposed in my paper, is *useful or meaningful.*

I shall deal with each in turn.

THE INFORMATIONAL IMPERFECTIONS OF LOCAL CONSUMER MARKETS

It is in fact my view that, informationally, many local retail markets function very badly indeed. Due to costly price/quality information and imperfect searches by consumers, there coexist high and low money prices, high and low quality, and—worst of all—high and low quality-deflated prices. Price in many of these markets is *not* an indicator of quality. Indeed the difficulty of assessing quality contributes to the informational imperfection of these markets.

The desire to test this view carefully and validly has been a major

* In view of the conspicuous references to Consumers Union, it should be noted that the author, though Treasurer of CU, is speaking as an economist and *not* as an officer of CU.

factor in motivating my efforts to formulate a theoretically acceptable and empirically measurable concept of *quality*.

But, as a challenger to accepted views, hopeful of exciting interest, I would seek to provide at least preliminary evidence of the informational imperfection of markets. Three pieces of evidence are adduced.

The first evidence comes from calculations by Morris and Bronson [8] of rank correlations between price and quality for 48 sets of products, mostly consumer durables. The sample consisted of all the product tests conducted by Consumers Union (CU) in the 1958–1967 period for which defensible price and quality data were available. The results were emphatic: for only 12 out of 48 tests, or 25 per cent, were statistically significant correlations obtained (5 per cent level of significance, one-tailed test). For 10 of the same set of 48, the correlation coefficients were negative, though not statistically significant. The Morris-Bronson analysis is subject to several limitations: (1) It assumes that Consumers Union's placement of brand/models into ordinal ratings groups would be accepted by all consumers; (2) it utilized list prices instead of more realistic "bargained" or "discounted" prices; (3) its quality measure pertains only to varieties and then only to the testable characteristics of these varieties.

In defense of the wide acceptance of Consumers Union quality ratings, it may be noted that subscribers to *Consumer Reports* reported that they purchased "top-rated models" from 48 per cent (AM/FM radios) to 81 per cent (hair shampoo) of the time [1].

Consider, next, five-year term life insurance policies. In the language of my paper, its predominant characteristic is "after-death income protection for survivors." Since consumers tend to purchase life insurance in large multiples and the same purchase price (or set of prices) applies for a long period in the future, consumers should be strongly motivated to search for a low price. Surely, in an informationally effective market, the prices charged by different sellers to an identical purchaser would exhibit only small variations. (As Juster so helpfully reminds us (page 565): "Efficient informational functioning of markets doesn't depend on complete knowledge for all units: it requires only sufficient knowledge to make inferior products unprofitable.")

The facts contrast starkly with the prediction. They come from a careful study, reported in the January, 1974 issue of *Consumer Reports*. CU reported the "interest-adjusted" prices charged by 125 companies, including the 20 largest. Separate prices were obtained for potential insurees classified by three variables: age, size of policy,

whether the policy was participating or nonparticipating. This gave rise to 18 estimates of the ratio of the highest price to the lowest price as follows:

Age	Face Value of Policy	Ratio of Highest to Lowest Price [a]	
		Nonparticipating	Participating
25	$ 10,000	1.89	1.55
	25,000	1.85	1.42
	100,000	1.84	1.42
35	$ 10,000	1.60	1.35
	25,000	1.58	1.37
	100,000	1.54	1.41
45	$ 10,000	1.49	1.35
	25,000	1.49	1.59
	100,000	1.44	1.39

[a] Companies whose policies are available only to specialized clientele, e.g., Teachers Insurance and Annuity Association of America (TIAA), are excluded from these calculations.

My question for doubters: Are these the results you would expect for an informationally effective market?

A third example, so hoary that it has become a consumerist cliché, is aspirin. Despite declarations by consumer product testing organizations,[1] the Federal Trade Commission, and eminent pharmacologists [see 2, page 69] that "aspirin is aspirin," the Bayer brand of aspirin sells for $0.78 per 100 tablets, whereas an equivalent amount of unbranded aspirin on the same shelf sells for $0.18 [6, page IV-7]. A second question for doubters: How long does it take for the word to get around?

These three examples suggest—but do not establish—that local retail markets are informationally ineffective. The conceptualization and empirical measurement of quality are necessary for a careful and persuasive test of the informational effectiveness of markets.

[1] *Consumer Bulletin* stated in its February, 1973 issue (page 13): "A government agency [FTC] holds [as Consumers Research did many years before] that one aspirin is about as good as another, in spite of advertising to the contrary for Bayer and other aspirins. . . . there is no persuasive scientific evidence that one brand is more effective for the relief of headaches than another."

UNIFORMITY OF QUALITY ASSESSMENTS AND THE CONCEPTS OF "PRODUCT" AND "MARKET"

Economists have been remarkably cavalier regarding the delineation of "product" and "market" sets. Juster scrupulously honors this tradition. (A possible exception to the generalization is the analytically useful, but nonoperational, device of separating "products" on the basis of cross elasticities.)

And, in so acting, Juster may have rendered one of his major criticisms of my paper partially invalid. Juster asserts, on a priori grounds, that differing preferences for a given set of characteristics on the part of different individuals would tend to make uniform quality assessments unlikely. He cites two examples:

1. ". . . high-income households are quite apt to place quite heavy weight on product characteristics which minimize shopping time. . . ."
2. ". . . the middle-aged shopper is not likely to value that characteristic ["acceleration" in an auto] very highly whereas the teenage hot rodder might regard it as the dominant characteristic of automobile quality."

Juster concludes: "There seems to be no solution to this problem and, hence, . . . no uniform solution to the quality measurement problem." Juster's speculations regarding the two examples seem plausible. And, unquestionably, they give rise to differing assessments of the quality of identical specimens.

But a crucial issue is the delineation of product/market classes within which quality comparisons might appropriately be made. My paper devotes two major sections to problems of delineating product/market sets, an investment which went unremarked in Juster's criticisms. It would be my contention, with references to the above examples that:

1. Otherwise similar high-income and low-income households might include (exclude) different sellers in "their" market set—due to possible differences in search cost, high-income households tending to view shopping as more costly;
2. Middle-aged shoppers would tend to exclude "high-performance" models from their product set while teen-age hot rodders might include only high performance models in theirs.

If my contentions are correct, these examples contain an important lesson: the confinement of quality comparisons to appropriate and carefully delineated market/product sets should eliminate a major source of nonuniformity in quality assessments. Nonetheless, the possibility of nonuniform quality assessments will continue to exist and should, of course, be investigated empirically.

THE USE AND USEFULNESS OF CARDINAL UTILITY

In an apt turn of language, Triplett states (page 573): "I really believe that pursuing any notion of scalar, nonmonetary quality measurement is following a will-o'-the-wisp, and the reason is that it will lead right down the road that Maynes has followed—a road which ends in a toll-gate labeled 'measurable, cardinal utility.'"

This is indeed the road on which I am traveling. But, contrary to Triplett, there exists strong evidence—from the market!—that the road exists and, further, that the benefits of the road exceed the tolls. Millions of American consumers have paid the tolls and have taken repeated trips on this highway since 1936 when it was first opened. Though the toll authority has not usually allowed users to travel at the high speeds for which the road was designed, there is strong evidence that users are pleased with the services provided. Let me explain.

The "highway" is *Consumer Reports* published by Consumers Union, the services provided are quality ratings (and other relevant consumer information), and the toll is presently $11.00 per year. Currently about two and one-quarter million subscribers are paying the toll. "Passengers" in these "vehicles"—using the road, but not paying a separate toll—are estimated at about ten million.

With some qualifications the quality scoring system employed by Consumers Union conforms to my Proportionality Model. The weights, or relative marginal utilities associated with a given characteristic, represent the consensus judgment of CU's testers. The characteristic scores are based on reproducible tests or on the reproducible judgments of panels of users, sometimes "experts." CU's quality assessments take account of only those *characteristics* for which such reproducible measurements may be obtained and its quality assessments are restricted to products for which such "objective" characteristics are judged to be dominant.

In *form,* CU's quality scoring system is cardinal. But in presentation and in textual interpretation, CU has for the most part acted conservatively and taken an ordinal posture. In this way, CU has permitted less than maximum "speeds" on its "highway."

CU does not publish the implicitly cardinal, numerical quality scores. Instead, it divides the varieties tested into "rating groups" in which quality differences are implicitly nonexistent (as reflected in alphabetical listing) or are described as small ("closely ranked models differ little in overall quality"). Differences between groups are assigned verbal labels, e.g., Very Good, Acceptable—Good, Acceptable—Poor, Unacceptable.

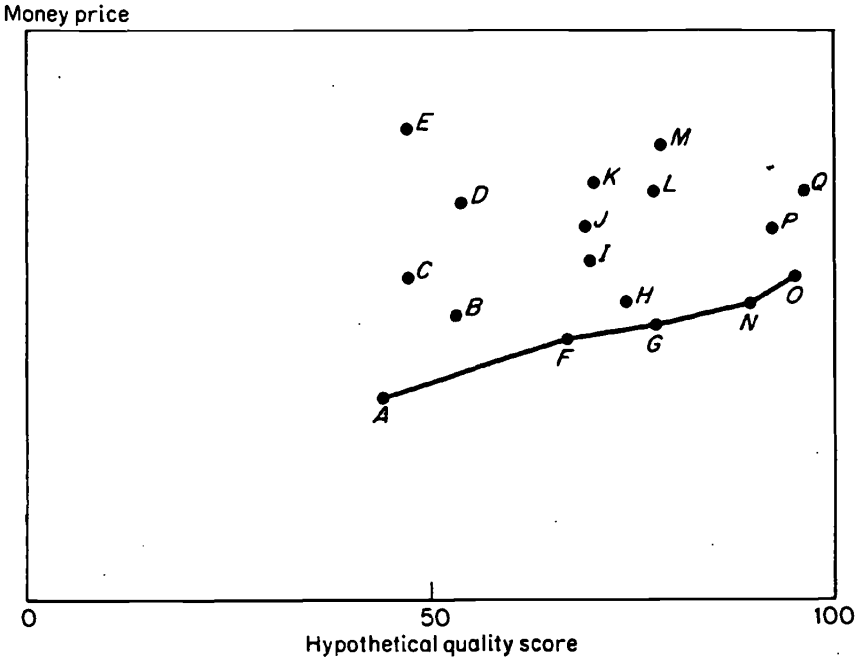
Cardinality manifests itself in CU's ratings in the form of a "Best Buy" designation. "Best Buys" come *only* from the highest quality rating group and should, in CU's words, "provide more quality per dollar." In my usage a "Best Buy" would lie on the efficiency frontier.

To illustrate CU's handling of quality scores, consider the hypothetical data in Figure 1. CU might designate varieties *NOPQ* as "Acceptable–Very Good," *F* through *M* as "Acceptable–Good," and *A* through *E*, as "Acceptable–Fair." Then, cardinally, CU might designate *N* and *O* as "Best Buys."

Let me spell out the limitations of CU's ratings and the differences between what it does and what is proposed in my paper. CU's product tests are restricted to nationally or regionally distributed products and

FIGURE 1

Consumers Union's Largely Ordinal Treatment of Cardinal Quality Scores



NOTE: Letters denote various specimens.

brands and, largely, to products for which "objective" characteristics dominate. My product set, by contrast, would be all-inclusive. CU confines its attention to characteristics for which reproducible measurements may be devised, whereas, again, mine would be all-inclusive. CU's reports of prices are list, or those encountered in a variety of cities, whereas mine would be those found in a single, local market.

CU's basically additive model is sometimes modified by the use of thresholds for certain characteristics. Specifically, a subthreshold value of some desirable characteristic, safety, for example, or an above-threshold value of some undesirable characteristic will automatically map the quality score of that variety to zero or to a verbally equivalent "Unacceptable" rating. This practice of CU's—which corresponds exactly to a suggestion by Juster (page 565)—is one which I should be pleased to follow.

In my empirical investigations stemming from my proposals, I would also plan to follow CU's path in interpreting results ordinally until convincing evidence is obtained in support of the stronger, cardinal interpretation.

Juster and Triplett address themselves to particular problems which they identify with the quality measure proposed. Juster asks (pages 564–565): Is there really a zero point or a unity point [on the scale for the characteristics score]? His doubts seem to rest not with the definitions offered, but rather with the question of whether these points can be identified by the users of the scale. My answer, on the basis of ex post discussions with students using the scale, would be that they understand the scale and make appropriate use of the end points and the cardinal (ratio) interpretation of the characteristic scores.

Juster laments (page 563) the absence of "any empirically verifiable content." While I have not been able to devise any empirical test of the structure of the measure, I would argue that for a quality assessor, whether he purchased a specimen on the efficiency frontier would constitute a strong test of the consistency of this model with observable behavior.

Triplett observes correctly (page 573) that there exists no universally acceptable measure of utility along the lines of, say, a kilometer for measuring distance. Does acceptance of this proposition imply the acceptance of the opposite proposition—that utility and other similar "subjective" variables are not susceptible to useful measurement? I believe not.

Consider, for example, "consumer attitudes," a variable which seems, if anything, more nebulous than utility. Yet Juster's research

career has centered upon largely successful efforts to demonstrate the predictive usefulness of such subjective variables as "consumer attitudes" and "subjective purchase probabilities" [4, 5]. The use of such subjective variables is plagued by the need to calibrate the apparently different implicit scales which each individual uses. (See [7] for a discussion of this problem.) Yet it would be my judgment that calibration is feasible and that success in this endeavor would yield very large payoffs.

A last comment relates to economists' rather schizoid views of what individuals can and cannot do by way of data assimilation and calculation. I am struck by the readiness of many economists to assume that individuals can perform prodigious feats of data assimilation and complex calculation, *as long as this behavior is "well buried" in a theoretical model.* And I am equally struck by the readiness of many economists to doubt individuals' capacities in this respect when the research requires—as in the case under discussion—actual data assimilation and calculation. Let me illustrate the point.

To utilize wage differences as measures of accepted risk, Thaler and Rosen, in a paper presented at this Conference, had to assume that in selecting an occupation and accepting a wage rate, workers in a risky occupation had access to and were able to digest data of the sort contained in Thaler and Rosen's Table 1. Or, alternatively, that "somehow" they could absorb these "facts." This, despite the fact that an audience including many professional data collectors were fascinated by some unexpected data in the table. For example, who would have thought of elevator operators as accepting a risky occupation?

Another example is the life-cycle hypothesis of saving. "All" that this theory requires of an individual is that he: (1) estimates the probability distribution of his employment income for each year up to and including the year of his death (but not beyond!); (2) select an appropriate discount rate; (3) discount the stream of expected income back to the present; (4) adjust his consumption—interpreted as "services enjoyed"—to this present value of labor income, and to net worth as well.

Our profession seems to have accepted the assumptions of data assimilation and calculation implicit in these two examples. By contrast, the task implicit in assessing quality seems like child's play.

THE MEANING OF WEIGHTS AND CHARACTERISTICS

Triplett (page 570) found my definitions of "weights" and "characteristic scores" ambiguous. Let me clarify my intentions.

The weights are intended to represent the relative importance, or relative marginal utilities, assigned to each characteristic in a product. The characteristic scores are scalars, denoting the marginal utility conveyed by this characteristic in a given specimen as a ratio to the marginal utility provided by the "ideal specimen." Thus, for the "ideal specimen," the weight represents the share of the total marginal utility of that specimen which would be conferred by a given characteristic. Thus, too, the overall quality score of the ideal specimen would be 1.0.

REFERENCES

[1] Benson & Benson. *Survey of Present and Former Subscribers to Consumer Reports.* Princeton: Benson & Benson, 1970.

[2] Burack, Richard. *The New Handbook of Prescription Drugs, Revised Edition.* New York: Pantheon, 1970.

[3] Fellner, William. "Operational Utility, The Theoretical Background and A Measurement." In W. Fellner, editor, *Ten Economic Studies in the Tradition of Irving Fisher.* New York: Wiley, 1967.

[4] Juster, F. Thomas, and Wachtel, Paul. "Anticipatory and Objective Models of Durable Goods Demand." *American Economic Review* 62 (September 1972): 564–579.

[5] ———. "Uncertainty, Expectations, and Durable Goods Demand Models." In Burkhard Strumpel et al., editors, *Human Behavior in Economic Affairs: Essays in Honor of George Katona.* New York: Elsevier, 1972.

[6] Maynes, E. Scott. *Decision-Making for Consumers.* New York: Macmillan, 1976.

[7] ———. "The Power of the Consumer." In Strumpel et al., editors, *Human Behavior in Economic Affairs* (see 5 above).

[8] Morris, Ruby Turner, and Bronson, Claire Sekulski. "The Chaos of Competition Indicated by Consumer Reports." *Journal of Marketing* 33 (July 1968): 26–34.