CHAPTER 5

PRICE AND QUANTITY MEASUREMENTS: THEORETICAL BIASES IN EMPIRICAL PROCEDURES

So far in this work, we have dwelt in some length over the causes of inflation and unemployment. Ultimately, the debate revolves around the prices and quantities of commodities. The main question is *why* these attributes tend to change. The issue of *how* they change, on the other hand, seems far less contentious. Of course, there is some disagreement on this latter question, but that mostly pertains to the adequacy of various measurements for different theoretical problems. When considered in isolation, the measurement of prices and quantities is commonly viewed as an empirical, relatively objective procedure. This conviction is not inadvertent. According to Einstein (1931, p. 66), 'The belief in an external world independent of the percipient subject is the foundation of all science.'[1] If economics is to be considered a science, it, too, must be dealing with factual matter: its theories should be tested against objective data and hence the measurement of such data must, whenever possible, be sufficiently independent of the theoretical debate itself. For instance, an increase in the list price of passenger cars can be explained by changes in tastes or technology, which affect demand and supply in a perfectly competitive market. It could also be rationalized by resorting to changes in producer mark-ups under conditions of oligopoly. However, both theoretical approaches are trying to explain changes in the same statistical price series. Similarly, a theory emphasizing rational expectations can argue that the business cycle results from a particular time-series process, while a theory that stresses institutional aspects can explain the cycle by changes in the rate of profit -- and here, too, proponents of both theories will probably use the same statistical series for real GNP as their principal variable of interest. List prices of passenger cars in the first example and real GNP figures in the second illustration are deemed adequate precisely because their measurement is believed to be sufficiently independent of the corresponding theoretical debates.

---

[1] Cited in Feuer (1974, p. 352).

The presumption of theory-neutral measurement seems adequate in simple abstract cases when the 'commodity' being considered remains unaltered. In such cases, it is then sufficient to count how many units were produced in each period and to observe the prices at which they were sold. For instance, suppose Ford Motors produced 100,000 Mustang cars at a unit price of $10,000 in 1975 and manufactured 150,000 units at a price of $14,000 per car in 1985. If we can presume that the Mustang of 1975 was identical to the one produced in 1985, we can, without ever defining what a Mustang is, conclude that there was a 50 percent increase in quantity and a 40 percent rise in price. On the other hand, if we acknowledge that the two models are different, such a direct comparison has little meaning and we must now both define the 'commodity' and describe how it changes over time. The two Mustang models may vary in aspects of production -- such as the technology with which they were manufactured, the labour involved in their assembly, and their material composition. They could also vary in their so-called 'consumption attributes' -- such as weight, size, power, shape, speed, comfort, colour, fuel efficiency, noise and chemical pollution. Under such circumstances, we must somehow denominate all such 'quality' differences in universal, quantitative terms and adjust our computations accordingly. For instance, if because of such changes, a 1985 model contained twice as much 'automobile quality' as the 1975 model, we would have a 200 percent rise in quantity produced and a 30 percent decrease -- not increase -- in unit price! On the other hand, if quality was found to be 50 percent lower in the 1985 model than in the 1975 one, we would end up with a 180 percent rise in price and a 25 percent reduction in quantity![2]

Clearly, whenever the nature of the commodity changes, the *measurement* of such changes in 'quality' is crucial for price and quantity calculations. But then there arises the question of how to measure quality and whether such measurements can remain objective and free of theoretical considerations. This problem is clearly exacerbated as we move from a single commodity to wider

[2] In the first case, when 'automobile quality' is doubled, a purchase of one 1985 Mustang for $14,000 is equivalent to buying two 1975 models for a unit price of $7,000. This imputed price is 30 percent lower than the 1975 price of $10,000. The doubling of quality also implies that output (denominated in 1975 units) rose from 100,000 in 1975 to 300,000 (150,000 · 2) in 1985, or a 200 percent increase. In the second case, when 'automobile quality' is halved, the imputed price for a 1985 equivalent of one 1975 Mustang rises to $28,000 ($14,000 · 2), an increase of 180 percent over the original 1975 unit price. Quantity (denominated in 1975 units) falls to 75,000 (150,000 / 2), or a decrease of 25 percent from the 1975 level.

aggregates. In devising output and price indices for the entire car industry, for instance, we must also account for the addition of new models, deletion of old ones and temporal changes in the industry's product mix. If we move to even broader indices such as real investment and its price deflator, or real GNP and the GNP deflator, our difficulties propagate since the concept of 'aggregate quality' is even more elusive.

The complex issue of comparing different commodities in time and space has occupied economists since the days of Adam Smith. In fact, the need to convert qualitative aspects into quantitative magnitudes of 'invariant' nature relates to epistemological problems of cognition, consciousness, subjectivity and objectivity, which have troubled some of the greatest philosophers since Plato and Aristotle. The predicament only intensified with the rise of capitalism, and prominent thinkers (like Hume, Kant, Hegel and Marx) dealt with them extensively. However, such difficulties remained largely *theoretical* until the present century, when the volume of quantitative statistical data grew substantially. With the evolution of national accounts after the 1930s, there was a pressing need for methods of estimation and measurement that would overcome the problem of historical change. As statistical bureaucracies expanded, techniques were formalized and the older theoretical debates gradually gave way to 'objective' *procedures*, presumably untainted by pseudo-scientific or philosophical disputes.

The purpose of this chapter is to assess some of the literature that sought to resolve problems in commodity measurement, specifically those posed by quality change. Explicit points of controversy in this literature have received considerable attention and we do not attempt to provide yet another summary of these debates. Instead, we direct our examinations toward some fundamental aspects which remained largely unexplored in the literature. Our objective is two-fold. First, we seek to demonstrate that, despite its implicit promise, the literature has failed to produce theory-neutral methods of measurement. Instead of being independent of theoretical debates, the measurements of prices and quantities are in fact deeply embedded in the particular world-view of neoclassical economics. Existing measurement procedures seem to require a society of free, utility maximizing individuals, a perfectly competitive organization of markets and continuous equilibrium. Our second goal is hence to suggest

that current methods may be partly or wholly inadequate when these conditions are not met. In other words, whenever individual preferences are open to coercion and persuasion, whenever collective action and differential power override voluntary atomistic interaction, whenever conflict and dynamic change replaces equilibrium and stability -- our measurements for quantity and price may be telling us a very biased story.

In a certain fundamental sense, then, our data on *how* prices and output change may not be sufficiently independent from our views on *why* they change and that inherent subjectivity must be recognized. Within the present historical epoch, the predisposition of price and quantity data toward the neoclassical economic outlook means that these data may not be altogether suitable to test the neoclassical outlook against competing frameworks. Furthermore, the problem is not really soluble since there is no practical way to encompass *conflicting* explanations into the *same* category of perception. For example, if we adjust price changes as if they were the consequences of quality improvements, we are already assuming that the change in price was not the result of variations in market power. On the other hand, if we were to adjust our price and quantity indices so that they reflect changes in social organization, we would in fact bias those categories against hedonic-based theories. It may hence be better to follow Myrdal (1956, p. 336) and accept that 'our very concepts are value-loaded' and that they 'cannot be defined except in terms of political valuations.' Indeed, according to Robinson (1962, p. 27), the whole subject matter of economics is immersed in political and ideological convictions. 'A unit of measurement,' she observes (p. 66), 'implies and agreed convention that is the same for everybody.' Yet, to the extent that such unit is '[l]ocked in the individual's subjective consciousness, it is not a unit at all.' Ultimately, the neoclassicist's 'unit of happiness is the same kind of mirage as Ricardo's *absolute value* or Marx's *abstract labour*.' Thus, instead of trying to devise some universally accepted indices for prices and quantities, it is perhaps better to accept from the outset that *any* scientific method of measuring these categories must, to some extent, be anchored in our initial values. Indeed, it is these initial values which make our analysis worthy in the first place, so they must be clearly identified for that analysis to carry any weight.

We begin our discussion by introducing, in the first section, the notion of commodity 'characteristics' as the basic building bloc for quality adjustment. Then, in the second section, we turn to examine criteria for discriminating relevant from irrelevant characteristics. In the third section, we explore how this framework is used to periodically adjust the standard price and quantity indices. In the fourth section, we examine an alternative and increasingly fashionable procedure, which uses hedonic regressions as a means of achieving continuous quality adjustment. In the final section, we argue that the evident limitations of existing methods require that we develop alternative indices which may better suit our own theoretical framework.

5.1 Commodities and Their Characteristics: Search for Objective Criteria

Because they seem to change, commodities themselves cannot be used as a uniform standard for temporal comparison and alternative units must be sought. One solution for this difficulty has been to define commodities as collections of smaller building blocks, or 'characteristics,' which are readily measurable. Lancaster (1971) for instance, formalized a taxonomy of characteristics as a basis for a new approach to consumer demand (although this could be equally used in reference to non-consumer goods). He proposed to view the relationship between people and things as a two-staged affair consisting of an *objective* relationship between things (commodities) and their characteristics and a *subjective* relationship between characteristics and people:

> All goods possess objective characteristics . . . The relationship between a given quantity of a good (or a collection of goods) and the characteristics which it possesses is essentially a technical relationship, depending on the *objective* properties of the goods and, sometimes, a context of technological 'know-how' as to what the goods can do, and how. Individuals differ in their reactions to different characteristics, rather than in their *assessment* of the characteristics content of various good collections. (p. 7, emphases added)

Clearly, in order to identify objective characteristics, we must effectively demarcate them from subjective interpretations. This separation is of paramount significance but then we may ask on what practical criteria it should be based? For example, how do we know that an apparent attribute of a medicine is an objective characteristic and not simply our personal conviction? Lancaster's solution is straightforward:

> Operationally speaking, it is *universality* that is important. If everyone *believes* that snake oil has special medical properties, we would analyze behaviour as though this were indeed true even if, in some objective sense, it could not be shown to be true; or even if it could be shown to be false, provided the negative proof was unknown or unaccepted in the society being studied. (p. 18-9, emphases added)

This method is seriously flawed in a number of ways. First, it embodies a double standard for objectivity which may lead to logical inconsistency: if a characteristic was deemed to be objectively true by a criterion of unanimous consent, how could such a characteristic be shown to be objectively false at the same time? Second, the need for 'universal acceptance' is a highly demanding because even minor dissent will render it non-operational. Third and most important, if 'acceptance' and 'belief' are criteria for objectivity what should be the criteria for subjectivity? This latter point deserves further elaboration.

In a modern industrial society, there is abundant information about commodities, some of which appears to be 'generally accepted.' Yet, one may still ponder on how such general acceptance makes the information 'objective.' To illustrate some of the attendant difficulty, consider for instance Lancaster's simple diet example (p. 17) where he tabulates the 'consumption technology' that relates different food items such as milk, eggs, and meat, with some nutrient characteristics like calories, protein and vitamins. The amount of each nutrient per unit of food is said to be objective because it was measured by 'official agencies,' but what happens if such agencies err in their measurements or intentionally falsify the data? Given this possibility, should we not conclude that the 'objective' characteristics of food in fact depend on the social organization of information? Let us carry this example one step further and consider the possibility of scientific progress, where new theories or methods of measurements may produce different numbers of nutrient contents. Could we permit the objective method for measuring protein in milk, for instance, to alter over time, or should we adopt a single method of measurement for all periods? Lancaster provides no clues as for how to address such dynamic historical questions.

While the accuracy of official data and stability of measurement methods in natural sciences may seem to present little practical difficulty, they nevertheless point to the potential hazard in relying on 'universality' as a criterion for 'objectivity.' The difficulties are no longer hypothetical, however, when we consider information that is made universally accepted not by 'official agencies' but *via* means of mass persuasion. In Huxley's *Brave New World*, genetic engineering and brain washing are used to create

legions of twin consumers that share not only the same *preferences* but also an identical *perception* of reality. Clearly, one does not have to go as far as Huxley's black utopia to identify the co-influence stemming from various media of persuasion. When advertisement argues that vitamin pills enhance our long term health, for instance, it concurrently affects our perception of what these vitamins can do and our preference for them. Similarly, when an advertisement encourages us to consume artificial sweeteners but fails to identify their possible link to cancer, it may influence both our preferences and knowledge at the same time. The existence of such ambiguities greatly reduce our ability to use 'universal acceptability' as a criterion for separating objective characteristics from subjective preferences.

All of these questions belong to an age-long and unresolved philosophical debate on the nature of subject and object which, surprisingly, is not acknowledged by Lancaster.[3] He simply takes it for granted that characteristics are objective and hence that their measurement is 'simply a technical matter' which presents no 'operational problems' and should be resolved by engineers (p. 115).[4] The difficulty in his opinion lies elsewhere, namely in determining the *relevant* characteristics.

5.2 The Choice of Relevant Characteristics

Complete characterization of commodities is often impractical according to Lancaster, simply because the number of characteristics may be too large. However, we can still find an adequate approximation with a partial description that ignores immaterial aspects of a commodity, provided such description encompasses all of the *relevant* facets of quality. The natural question arising in this context is how can we distinguish between 'relevant' and 'irrelevant' characteristics? Lancaster stipulates that a

---

[3] Ever since the Greek epistemology separated the earthly *appearance* of things from the universal *truth* of the world spirit, philosophers have attempted to bridge this gap between the subjective and the objective. British empiricists like Locke and Hume rejected the notion of universality and argued that apparent uniformity in human thinking was rooted only in customs and habits. German idealists, like Kant, contested this view and argued that perception was universal -- namely common to all humans -- because it required the *a priori* intuition of space and time and various categories such as unity, substantiality and causality. Nevertheless, the Platonic spirit continued to fly unarrested for Kant agreed that objects of experience were indeed never given by themselves and could only be perceived through human experience. Even Hegel's later attempt to unify object and subject into one 'whole' failed to resolve the riddle.

[4] This view is widely held. For similar expressions, see Adelman and Griliches (1961, p. 539), Court (1939, p. 107-8), Rosen (1974, pp. 75-6) and Triplett (1983, pp. 277-8; 1986, p. 37).

characteristic is relevant to a situation if ignoring its existence would lead to different *predictions* about people preferences and choice. But the bases of predictions are subjective. They must rely on a choice among theories and so Lancaster's recipe for the concrete description of commodities is no longer theory-neutral.

An example of the intrusions of such theoretical considerations is provided by Triplett (1983) as part of a broad distinction between input and output characteristics:[5]

> . . . quality variation in an input exists if substitution of different varieties or examples of this input creates variations in output or cost that are not explained by the factors included in the production or cost function. A quality is an input characteristic if it reduces that unexplained variation. (p. 279)

Similarly,

> Something is an output characteristic if it accounts for, or partly accounts for, the unexplained variation in resource usage occasioned by changes in the varieties of nonhomogenous goods produced. (p. 294)

To illustrate Triplett's approach, consider the case of personal computers. 'Computer speed' should be regarded as a relevant input characteristic if variations in speed help to explain changes in the output or 'user-value' generated by the computer. On the other hand, computer speed should not be viewed as a relevant output characteristics if increased speed does not seem to entail higher production cost or an additional use of resources in producing the computer. Following this same logic, 'car size' is not an input characteristic if it does not appear to affect consumer preferences or utility, but it is an output characteristic if it has an apparent impact on production cost.

Two questions arise in this context. First, one may ask which methods -- 'user value' for input, 'resource-cost' for output, or both -- should be used to identify relevant characteristics for any particular commodity? According to Triplett (p. 305) it does not matter. Under conditions of perfect competition and equilibrium, the two methods yield the same numbers because 'the marginal cost of producing a quality change must approximate the incremental value of it to the user.' This must be so, writes Triplett, because otherwise 'a reallocation of resources would take place.'

---

[5] This distinction was also central to the earlier work by Fisher and Shell (1972).

These conclusions, of course, are valid only to the extent one accepts the theoretical notions of efficient allocation and equilibrium together with a presumption of perfect competition. Indeed, Triplett acknowledges that in 'reality' the two methods would usually yield different measures for quality because there are 'shifts in functions, interference with competitive allocation, or wrong data' (*ibid.*). The prerequisite of 'competitive equilibrium' in this context has been repeatedly emphasized by most commentators in the quality change debate. Unfortunately, these commentators have generally failed to provide guidance for identifying characteristics when commodities are produced and traded in alternative market structures or in disequilibrium.[6]

A second problem arises because Triplett's recipe is simply incomplete and cannot be readily applied. The procedure requires that we identify the way in which characteristics affect output or input. Unfortunately, such causal relationships cannot simply be 'observed' and must be based, at least in part, on *economic theory*. Furthermore, given a variety of competing theoretical perspectives, the identification of relevant characteristics hinges on the *particular* choice of theory. In our earlier computer illustration, for instance, we need to chose a specific production theory of computer services (to identify input characteristics) and a cost theory of computer manufacturing (to identify output characteristics). Unfortunately, Triplett's procedure provide no guidance on how to select the 'appropriate' theory that should be used in each case.

This latter point is significant for measurement because if distinct theories generate different sets of relevant characteristics they also lead to different measurements of quality. For instance, consider the relation between the required duration of a bachelor programme in economics and the overall 'quality' of the programme in producing graduates. If years of schooling are believed to enhances the 'human capital' of programme participants (in other words, if 'duration' is taken as a relevant input characteristic), an increase from three to four years in the duration of the programme should be considered as quality improvement. On the other hand, if programme duration is regarded as an irrelevant input characteristics, such an increase obviously does not affect the programme's quality. Yet

---

[6] For recognition of these prerequisites, see Early and Sinclair (1983, p. 108), Hofsten (1949, pp. 285-86), Rosen (1974), Ulmer (1949, pp. 67-8).

a third perspective may hold that programme duration is a relevant input characteristic because it *undermines* the production of human capital in graduates. From this perspective, increase in duration is definitely a quality deterioration! Evidently, each theoretical perspective in this context implies a different measure for quality change. The choice of one particular measure must then depend on our *theoretical* preferences and hence cannot be considered 'objective.'[7]

Triplett, like other participants in the quality change literature, does not acknowledge the potential impact on quality measurement of such theoretical diversity. As noted above, this literature generally accepts the neoclassical foundations of perfect competition and equilibrium, while alternative views are simply not considered. Furthermore, even within this limited framework, it is customary to talk about *the* production, or *the* utility function which must be empirically 'discovered' as if competing formulations cannot coexist.[8] Unfortunately, even under this imposed theoretical uniformity, the precise functions for consumption, production, cost and utility are admittedly unknown and practical methods of measurement always rely on *ad hoc* procedures.

5.3 Specification Pricing and Quality Change

In practice, most price indices (and associated quantity measures) are based on what is known as 'specification pricing,' where each commodity in the index must fit into some predetermined range of product attributes or 'specifications.' In this way, comparisons are restricted to only those commodities whose specifications remain unchanged over the period and a 'pure' price change can be thus assessed. An important shortcoming of this method stems from its inability to properly account for quality change. Indeed, the Price Statistics Review Committee (headed by Stigler in 1961) was concerned that adherence to strict specifications, in the presence of continuous changes in product mix and commodity quality, would greatly reduce the relevance of the index. As a partial remedy, the Committee

---

[7] There is another reason for concern here. For consumer goods, for instance, the choice of relevant characteristics in this method depends on utility functions and consumer preferences. But how can one rely on subjective consumer preferences to describe 'objective' characteristics of commodities? We return to this important question is the fourth section.

[8] For example, see Triplett (1983, p. 274, 302).

recommended we adopt 'more general rules for recognition of comparable qualities' and restructure specifications along 'more flexible lines' (p. 34). Following this suggestion, U.S. agencies, such as the Bureau of Labor Statistics, developed 'subspecifications' within main specifications. These extensions provided greater 'flexibility' in fitting products into existing descriptions but they still did not resolve the pending problem of quality change. Two points of difficulty deserve attention here.

First, the rules governing the development of specifications are obscure. The United Nations' *Guidelines on Principles of a System of Price and Quantity Statistics* do not explain how to distinguish between objective variations in quality and differences in subjective interpretations. In fact, the *Guidelines* allow national statisticians considerable freedom in their choice of characteristics:

> In the case of goods, *all* differences in physical composition, components, size style, packaging and operating characteristics (for example, capacity, power, speed, durability etc.) should be considered quality differences . . . In the case of services, quality characteristics relate to such attributes as the activities constituting the services, the conditions under which the services are rendered, the level of skill and training of the persons rendering the services and, if feasible, the benefit generally expected from the services. (p. 9, section 45, emphasis added)

This latitude in choosing relevant characteristics is somewhat restricted along lines discussed in our previous section:

> As an exception to the rule stated above, differences in the circumstances of production that do not lead to differences in utilization or function should not generally be considered quality differences. (p. 9, section 46)

Unfortunately, the United Nations' manual does not provide guidance for the practical implementation of this latter exception. In the absence of clear criteria, the choice of 'relevant' characteristics (for input or output) to be included in specifications is left to the discretion of statistical-service officials and naturally, arbitrary definitions of quality lead to arbitrary conclusions about *changes* in quality. The ambiguity is further enhanced because most statistical services do not provide adequate narratives for their commodity specifications.[9]

A second difficulty arises in bridging the operational gap between fixed specifications and temporal changes in quality. Even with many layers of 'objective' subspecification, the overall description

[9] On the lack of such narratives, see comments made in the Price Statistics Review Committee (1961) and by Griliches (1967; 1971).

of commodities in a price index is ultimately given and hence does not allow for variations arising from quality change in existing commodities, introduction of new commodities and disappearance of old ones. Most statistical services attempt to overcome these problems and preserve the continuity and validity of specification indices, but procedures employed for this purpose, such as ones described in the United Nations' *Guidelines*, contain strong theoretical biases, particularly toward conventional neoclassical price theory and the presumption of equilibrium.

Under specification pricing, a quality change is said to occur when the product or service no longer fits into the predetermined specification. The extent of the change can vary from a minor modification affecting a single aspect in the specification (such as a 2 percent increase in 'car speed') to a multifaceted metamorphosis (such as a replacement of an old automobile model by a new one). The problem, of course, is how to quantify such changes in some universal units of 'quality.' For this, we need to know the relative contribution of each characteristic to the overall quality of the commodity; that is, we must know the 'consumption technology' or the 'production function,' whatever the case may be. Unfortunately, these structures are unknown so methods of indirect estimation are substituted for direct measurements of quality change.

The common adjustment procedure of statisticians uses observed changes in cost or price to estimate non-observable changes in quality. The U.S. Bureau of Labor Statistics, for instance, relies on cost information furnished by producers to estimate quality changes in the Producer Price Index. For instance, if Westinghouse increases the BTU output of an air-conditioning model and informs the Bureau that this modification requires a 15 percent rise in cost (under the *original* technology and factor prices), the change is taken to represent a 15 percent improvement in quality.[10] The same principle is applied in 'splicing,' a method used when a new commodity replaces an old one. As a hypothetical illustration for splicing, consider the introduction of Sony compact disc (CD) players priced at $400, as potential replacement for turntables of an earlier technology which cost $200 per unit. When it is decided to replace the turntable by the CD player in the Consumer Price Index, the observed price difference is interpreted as a 100 percent increase in quality (namely, in quantity of 'music machines'

[10] For further details on quality adjustment to the PPI, see Early and Sinclair (1983).

produced), while the CPI itself remains unaffected.[11] The rationale behind these methods is, again, very simple. It is presumed that in equilibrium, under conditions of perfect competition, the ratio of qualities between two varieties in time or place is just equal to the corresponding ratio of prices.[12] But then this explanation raises two disturbing questions.

The first difficulty stems from the possibility of disequilibrium. When markets are out of equilibrium, there is no unique mapping between qualities and prices and the popular method simply breaks down. To overcome this obstacle in the case of splicing for instance, the United Nations' *Guidelines* (p. 10) recommend we replace products in a price index when the assumption that price differences between the two products are proportional to quality differences is 'most likely to be true.' In simple words, statisticians are advised to perform splicing when markets are in equilibrium. Unfortunately, criteria for identifying occurrences of equilibrium are yet to be developed and the *Guidelines* concede there is a 'difficulty' here. So how should the quality of air conditioners, 'music machines' or any other commodity be adjusted in the meantime? The United Nations have no solution and admit that such corrections must be 'essentially pragmatic' (p. 10).

The second complication emerges when commodities are produced and exchanged in 'imperfect' markets. Incidence of government intervention, oligopolistic practices, or non arms-length transactions could distort or completely destroy the functional relations between price and quality which are presumed to exist under 'perfect' competition. Again the method breaks down. The United Nations' *Guidelines* note these potential hazards on in passing. They discuss two cases of market 'imperfections' but, far from resolving the problem, their proposed solutions only serve to accentuate it.

One case involves the mandatory installation of anti-pollution equipment on automobiles. Here the *Guidelines* recommend to treat such changes as quality improvements and evaluate them by their

---

[11] For an explanation of 'splicing', see The United Nations (1977, p. 9).

[12] The process is not as simple if the older turntable disappears from the market before the new CD player is introduced into the CPI. When this happens, the price for the non-existing turntable that 'would have prevailed' in the time of the swap must be somehow imputed. A similar difficulty arises when a totally new product is introduced into the index.

cost of production. The explanation for this recommendation, however, is highly confused. The United Nations agree that because consumers and producers are given no choice in this matter, changes in cost may not provide an accurate measure for quality improvement. Nevertheless,

> . . . it is still appropriate to treat the required improvements as increases in quality, not price. Price and quantity statistics are intended to provide ***objective*** measures (in so far as these are possible) of what has happened to output and its price. Measuring the utility or welfare that output yield is beyond their scope. (p. 10, emphasis added)

But then,

> It is, of course, true that the very concept of "output" involves some assumption about *utility*; however, certain conventions have been adopted for use in national accounting and it seems preferable to maintain the same conventions here. (*ibid.*, emphasis added)

So despite elaborate theoretical attempts to resolve the issue 'once and for all,' it appears that the philosophical questions of objectivity and subjectivity persist unabated. Within the framework of equilibrium and perfect competition, these issues are conveniently ignored but outside this framework one must resort to 'conventions.'

The *Guidelines* also identify another 'special case of difficulty' involving internal transactions between related enterprises or branches of the same company. Since prices set under these conditions may be 'quite arbitrary,' the United Nations (p. 12) suggest we 'abandon value as one of the primary measures' and replace it with 'a measure of physical quantity,' combined with an estimate of 'what the equivalent market price would have been.' This recommendation is puzzling for two reasons. First, how could we obtain direct measures of physical quantity when, in the presence of quality changes, such measures were deemed infeasible to begin with? Second, what formula should be used to impute hypothetical market prices and what is the validity of such imaginary prices? The *Guidelines* do not explore these questions.

This commentary is significant because the practical estimation of quality change appears to rely exclusively on the assumption of equilibrium and perfect competition while, in reality, perfectly competitive markets are not very common and experiences of rapid dynamic change seem contrary to the notions of stability and equilibrium. Consider, for instance, quality changes in the nature of military hardware purchased by governments. In order to adjust the price deflator for military spending to such changes, the U.S. Bureau of Economic Analysis adopts the usual methodology of 'production cost' and

'splicing.' However, weapon systems are not produced and sold in a perfectly competitive market but, rather, in an environment of monopsony and oligopoly from the buyer and seller sides, respectively. Consequently, the meaning of 'equilibrium' in such a market is unclear. It is also not clear what 'utility function' the government attempts to maximize. The 'cost function' introduces further ambiguities because, in the reality of military procurement, it is often hard to establish the meaning of 'normal profits' and even to separate genuine cost from concealed earnings.[13]

Questions raised by these difficulties are by no means limited to military commodities. Can markets for automobiles, consumer electronics, civilian aircraft, medical equipment, machine tools and ships, for instance, be considered perfectly competitive? What is the level of 'normal profit' to be included in cost functions for industries producing for these markets? What are the 'utility functions' for consumers in these markets? How should we interpret attempts to influence consumer preferences in each of these cases? What is the impact of government intervention (through regulation, taxes and subsidies for instance) on prices in such markets? These questions must be clearly answered before we can interpret the validity of quality adjustments to commodities produced in each market. To our knowledge, however, these issues have not been sufficiently addressed in the quality-change literature. Rather than deal with such basic questions, many researchers have chosen to cultivate and refine existing approaches. We deal with one of these attempts in the following section.

5.4 Hedonic Regressions for Price Indices

The most systematic method of adjustment for quality change can be found in the application of so-called 'hedonic regressions.' The method was first suggested and applied by Court (1939) and was latter used in studies by Stone (1956), Griliches (1961) and others.[14] During the 1980s, hedonic regressions were employed, for the first time, to adjust official price indices for computer equipment and

---

[13] For further details on the construction of price indices for military spending, see Ziemer and Galbraith (1983).

[14] For a detail bibliography on 'hedonic regressions', see Griliches (ed.) (1971) and Triplett (1975). References to recent studies are cited in Cole *et al.* (1986) and Triplett (1986).

new one-family houses in the national income and product accounts of the United States.[15] Despite frequent suggestions to the contrary, hedonic regressions do not provide any new insight into the issue of quality adjustment. The hedonic technique is nevertheless interesting because its elaborate nature heightens difficulties which are also implicit in other adjustment procedures.

Like other existing approaches for quality adjustment, the hedonic regression is firmly rooted in the assumption that quality is correlated with price. Indeed, Griliches (1961, p. 57) is resolute when he asserts that

> The reason why [at any one time] different varieties or models sell at different prices *must be* due to some differences in their properties, dimensions, or other "qualities," real or imaginary. (emphasis added)

Under the common system of specification pricing, the primary focus is on a commodity and adjustment for differences in quality are performed only irregularly. The hedonic approach, in contrast, focuses directly on characteristics to enable a more or less continuous adjustment for quality differences. In this approach, the familiar functional relationship between the commodity's quantity (Q) and price (P) is recast in terms of quantity $\{q_i\}$ and 'implicit' prices $\{p_i\}$ for its n different quality dimensions, or characteristics. The general expression for such relation can be summarized by equations (1) and (2):

$$(1) \qquad P = f_1 (Q, t) ,$$

where quantity (Q) can be written as some aggregation of n different qualities:

$$(2) \qquad Q = f_2 (q_1, \ldots, q_n) .$$

In Equation (1), the parameter associated with t is the 'pure' price change that occurs over time, while the parameter associated with Q reflects the implicit impact of overall quantity on the commodity price. In Equation (2), each characteristic q_i is associated with a corresponding parameter p_i which could be interpreted both as the 'implicit' price per unit of this characteristic, as well as the weight, or 'contribution' of that characteristic to the overall quantity of the commodity. This simple framework

---

[15] See Cole *et al.* (1986) and Triplett (1986) for more information.

seems to provide a systematic alternative to other, apparently more erratic, quality adjustments discussed in the previous section. A closer examination reveals, however, that this framework does not resolve any of the difficulties haunting the other methods. First, we still do not know which quality variables are 'relevant' for each situation, nor do we know the 'functional form' through which such characteristics presumably affect the price. Second, the emphasis on the seemingly 'technical' nature of hedonic regressions is highly misleading and, in fact, the neoclassical paradigm is as dominant here as elsewhere in the quality-change literature. An illustration of these issues is provided by the work of Griliches (1961) on automobiles.

Griliches examined data for U.S. passenger four-door sedans for the years 1937, 1950 and 1954 through 1960. He related list prices for different models to three numerical quality variables (horsepower, weight and length), as well as to six dummy variables which indicated the presence of other quality attributes as 'standard' features (V-8 engine, hardtop, automatic transmission, power steering, power brakes and whether the model was considered a 'compact' car). Econometric estimations were based on the following linear semi-logarithmic form:

$$(3) \qquad log\ P = p_0 + \Sigma_i\, p_i\, q_i + u\ ,$$

where P is the list price for the model, q_i is the value for the ith quality characteristic, p_i is the implicit price for that characteristic, p_0 is the 'pure' price (equal for all different models) and u is an error term. This functional form was applied to (1) cross-sectional data for each individual year, and (2) cross-sectional data with observations drawn from two or more adjacent years. (In the second formulation, shift variables were added to the regression in order to capture the pure price change that occurred between adjacent years.) Hence, the first scheme allows implicit price estimates to vary from year to year, while the second assumes that these implicit prices are fixed for the entire period.

Clearly, parameter estimates derived in this context depend crucially on the *particular* choice of included variables, functional form and cross-sectional method of estimation being employed. Griliches (p. 53) admits that '[T]here is no a priori reason to expect price and quality to be related in

any particular fixed fashion,' so he cannot explain why any specific scheme is to be preferred over alternative ones. According to Griliches, his own decisions were based on 'empirical' considerations. A large number of different regressions were computed for different combinations of years and independent variables, and the final choice was based on 'an inspection of the data and the convenience of this particular formulation.' In a later article (Griliches, 1971, p. 15), it was suggested that one should chose the framework that provided the 'most *concise* and stable explanation of reality.' But reliance on 'empiricism' here could be quite precarious for two reasons. First, there are no objective criteria for choosing the most 'convenient,' 'concise,' or 'stable' explanation of reality among competing econometric formulations. Second, the *measurement* of quality cannot depend on the *explanation* of price changes and remain theory-neutral at the same time. It is fairly clear that despite its apparent rigour, the decomposition of price into 'quality' and 'pure price' elements in the hedonic regression is far from being 'objective.' Like simpler methods, the hedonic regression also involves a mixture of theoretical and arbitrary decisions.

To study these difficulties further, let us consider the estimates for implicit prices obtain by Griliches (reported in Table 3.4, p. 66). The most striking feature here is the marked variability of coefficient estimates between the different periods. For instance, in 1954-55, the 'implicit price' of 10 horsepower units amounted to 2.4 percent of the overall automobile value. In 1957-58, however, this fell to 0.4 percent, only to rise again in 1959-60 to 1.1 percent. A similar variability is evident for most other parameter estimates. Griliches (p. 64) suggests that this instability in implicit price estimates for various 'quality' attributes is an 'empirical' problem, which stems from a high correlation between the different quality variables and, also, from a lack of variability in some quality attributes between various models. Of course, this response is not the only possible explanation since, in the hedonic framework, implicit prices should be also sensitive to changes in the 'supply' and 'demand' functions for characteristics. Griliches (p. 79) is aware of this possibility, which he equates with the 'classical index number problem of changing weights,' but declares that 'Not much can be done about this in practice'[!] For our purpose however, the issue here is not so much the solution but the proper identification of the problem.

In the hedonic method, 'quality' is perceived to be embedded in an array of characteristics and the hedonic regression purports to provide us with an estimated weight, or contribution of each of these characteristics to 'overall quality' (or quantity). The trouble is that these weights are not fixed, in other words, the 'quality contents' of each characteristic changes over time. But then if the contribution of 10 horsepowers, 1 pound of weight, or 10 inches of length to 'automobile quality' can change from year to year, this means that quality is not solely intrinsic to the commodity and also depends on other, 'external,' factors. What are these other factors which make quality unstable? Griliches points to changes in 'tastes' and 'supply factors' but, in doing so, he confirms the view that quality is essentially a *subjective* matter. The issue is crystallised when we note how his weights for some characteristics fluctuate between positive and negative values (p. 66). The weight for a 100 inches of automobile length, for instance, was 0.07 percent in 1957-58, but, in 1958-59, fell to -0.18 percent. Is it possible for an 'objective' characteristic to represent a positive quality in one year and an 'anti-quality' in the next? In light of Lancaster's treatise (1971) on objective characteristics and subjective preferences, we suspect the answer should be negative.[16] Fisher and Shell (1968, p. 24) are also troubled by this conceptual riddle when they raise the basic question of 'just what we mean by taste change as opposed to a quality change.' They provide a separate theoretical treatment for each of these concepts but, unfortunately, they too fail to indicate how we should distinguish between quality and taste in practice.

Estimates from hedonic regressions are open to reinterpretation for other reasons. Griliches (1961, p. 76) writes that

> One of the problems associated with the use of list prices in this study is the extent to which they may just represent *pricing mistakes* by manufacturers at some point in time. A manufacturer may *overprice* or *underprice* a particular innovation, and there is nothing in our method that would catch it. (emphases added)

Again, before we turn to Griliches' proposed solutions, let us contemplate the problem further. What should we understand from Griliches' reference to 'pricing mistakes,' 'overpricing' and 'underpricing?' Clearly, these concepts are meaningful only when we have a yardstick for 'correct' pricing, but it is common knowledge that such an unequivocal benchmark is provided only by equilibrium in a perfectly competitive market. The presence of this prerequisite leads to three dilemmas. (1) If pricing 'mistakes'

---

[16] As we argued, Lancaster's taxonomy contains some serious inconsistencies and ambiguities so the answer here cannot be definite.

are the result of disequilibrium, we can no longer accept the basic assumption made in Griliches (1971, p. 4) on the existence of a 'reasonably well-fitting' relation between the prices of different models and the level of their characteristics. (2) In a perfectly competitive market, prices are determined by the 'invisible hand,' not by individual producers. Under these circumstances, it is not clear how pricing 'mistakes' by manufacturers are possible. (3) If markets are oligopolistic or monopolistic rather than perfectly competitive, the meaning of a 'correct' price is unclear. Should we consider a price to be the 'correct' one when it reflects the 'true quality' of the product, when it gives rise to the maximum profit, or when company managers view it as the 'proper' one under the circumstances? How could we know whether or not the price satisfies any one of these requirements? Each one of these three quandaries casts heavy doubt on the validity of hedonic regressions. Griliches prefers to ignore these questions because they are 'general:'

> [T]he doubt whether the evidence of the marketplace reflects adequately, if at all, the "true" marginal utility of different items or qualities to the consumer can be turned against any other price or commodity. It is not a problem peculiar to the measurement of "quality." (pp. 60-1)

Once this is said, Griliches feels free to proceeds with suggested solutions. His remedies are interesting because they, again, attest to strong neoclassical theoretical biases.

One solution is to relate the temporal behaviour of residuals from the hedonic regression to the market share for each individual model. The residuals are taken as proxies for 'pricing mistakes' as compared to the 'true' implicit prices predicted by the regression. With this interpretation, 'overpricing' (or 'underpricing') should lead to a reduction (or an increase) in the market share of the particular model, as consumers rearrange their purchases to maximize the quality return on their dollar. The problem with this cure is that the proof is already embedded in the hypothesis: if producers are oligopolies, how could the hedonic regression estimates be taken to represent the 'true' implicit price for characteristics? Furthermore, market shares in oligopolistic markets can be influenced by a host of factors in addition to the price of marginal utility. In particular, with massive advertising campaigns by manufacturers, one should be careful before attributing changes in consumption patterns to 'rational' consumer decisions.

Another way of verifying the validity of coefficient estimates derived from data on new automobiles is to compare them with quality estimates based on markets for used cars because, according to Griliches (p. 77), 'prices of used cars are not tied any more to the manufacturers' list prices and are set, presumably, more directly by the "market".' This alternative is deficient in two main ways. First, one may suspect that the presence of a large oligopolistic market for new cars affects prices set in the used-car market. Second, even if we believe that prices in this market are equilibrium outcomes of the interplay of supply and demand in a perfectly competitive environment, we must still be able to distinguish between the price associated with 'quality' embedded in new automobiles and the discount allowed for 'depreciation' of that quality in second hand cars. Cagan (1965) attempted to resolve this difficulty by first estimating a 'depreciation coefficient' from market data and then adjusting quality estimates accordingly. Unfortunately, the depletion of quality over time is no more observable than quality itself so Cagan's estimates of depreciation are only as good as the arbitrary assumptions on which they are based.

These are only two ways by which one can amend reality to fit the preconceived neoclassical framework. With sufficient resolve, many more can be devised. An alternative path is to recognize the reality of an oligopolistic market structure and evaluate the consequences for the hedonic-regression method. Dhrymes (1971), for instance, estimated implicit prices for characteristics of automobiles and refrigerators and found that these coefficients varied between the different manufacturers. This led him (p. 104) to conclude that the estimated coefficients represented

> the manufacturer's own evaluation of the model's features in the context of his own price-quantity selling strategy -- remembering that we deal essentially with an oligopolistic market and that the oligopolist may well be "satisficing" his profits on a cost-plus-markup basis.

Under these conditions, writes Dhrymes (p. 93), 'we cannot, strictly speaking, construct "quality corrected" price indices routinely in the manner suggested by Court (1939) and Griliches (1961).'

5.5 Alternatives to Utilitarianism: Beyond Perfect Competition and Equilibrium

We have argued that the measurement of price and quantity indices is neither objective nor free of theoretical biases. The attempt to develop 'objective' commodity measures in the presence of quality

change is besieged by a constant resort -- explicit or implicit -- to 'subjective' considerations. Both the idea that quality can be measured (objectively or not) and the methods developed for that purpose are closely tied with the neoclassical paradigm. The evidence supporting these conclusions seems overwhelming.

This view may prove disheartening for those who prefer to have 'reliable' and 'objective' data, with which they can describe the world and rigorously test their theories. One may hope that, although there are serious methodological difficulties, they present only minor *practical* problems. Triplett (1975), for instance, surveyed numerous studies on quality change and concluded that there are no clear evidence that price indices are systematically biased in one way or the other. This conclusion, in itself, reflects a misunderstanding of the problem. Such inference is possible only if we already assumed that price indices can, in fact, be denominated in some neutral, *a*-historic units of 'quality.' The nature of this problem is best illustrated in reference to long-term historical comparisons. Economists often examine price and quantity series that extend over a century or more but the meaning of such comparisons is unclear. For instance, how should we interpret the measure of real GNP in 1882 when denominated in '1982 prices'? Most commodities produced in 1882 were simply unavailable in 1982 and hence could not have '1982 prices.' Furthermore, every generation of goods and services introduces new features and eliminates older ones, so even the imputation of implicit prices for characteristics is infeasible. Finally and perhaps most significantly, the 'desirability' of any commodity attribute and the 'satisfaction' it may provide change drastically with social conditions. In this context of vast changes, comparisons of uniform quality units may often seem absurd but the United Nations' *System of National Accounts* (1968, p. 61) nevertheless hopes to circumscribe the problem by arresting history in a straightjacket: it suggests to formulate characteristics so that 'their validity is as far as possible *timeless*.'

Is it at all possible to have theory-neutral, objective indices for price and quantity in the presence of historical change? In our view, the answer is negative but this need not destroy the prospect for empirical economic research. Like Carr (1961), we also believe that any study of society necessarily reflects subjective biases of the observer and that such biases must be recognized, not ignored. In the next part of this work, we propose an alternative framework which seeks to analyze inflation as an

antagonistic process of dynamic restructuring. We emphasize not the individual, but the group; we accentuate not equilibrium, but continuous historical change; we focus not on universal utilitarianism, but on the conflicting duality of creativity and power. From that vantage point, 'timeless,' hedonic-based indices for prices and quantities are not only impossible to construct, but also quite unhelpful. Instead, we will suggest (in Chapter 7) a new family of indices which are biased -- in a clearly defined way, we hope -- toward our own theoretical dispositions.