

Quality Change a Statistical Approach

(An elaboration of the pure price index framework and its relevance to quality adjustment)

Paper by Don Sellwood for the Fifth Meeting of the Ottawa Group, Iceland 1999

Abstract.

A consumer price index is no more than the observations and practices of which it is composed. The concept of the index is embodied in those observations and practices. This paper seeks to elaborate the issues of quality adjustment in CPI construction by reference to the observable universe of transactions in consumer goods and services and the limited number of actual practices. The paper argues that the practices of stratification and matching, which are central to "the pure price index" concept, may be seen to have correlates in two sub-universes, the static and the dynamic. However, this conceptual framework does not provide a clear operational basis for the determining when quality has changed.

The pure price index concept is taken as defined by the actual sample and operational practices followed in the construction of all actual CPIs including those said to follow other concepts. The elaboration of the measurement objective at the level of the universe in terms of actual observation and practice highlights the role of stratification or matching in quality adjustment. Though stratification can achieve a difference between a pure price index and the simple change in the average transaction price the distinctions between price and quality and between quantity and quality are not well defined. In practice such distinctions must be made according the observable properties of the transactions. Existing procedures for quality adjustment, such as "overlap pricing", "linking", "imputing base prices" or "hedonic" regression are purely mechanical. They have no parallel at the level of the universe where there are no missing prices and no one-to-one replacements. They have yet to be justified as devices for achieving a specific sample representation of price change for a given universe of consumer prices.

The paper concludes with specific requirements for conceptual development or agreement on conventions for quality adjustment practice within the pure price index framework. It is not evident that the "cost of living" concept, based on unobservable utility, can provide the operational guidance necessary for appropriate quality adjustment. The case against utility theory seems so overwhelming that one might ask why a rational and fully informed consumer might yet show a preference for it. Perhaps because they value elegance over utility or perhaps because the relevance of alternatives to quality adjustment has not been examined.

Key words: Consumer Price Index, pure price index, cost of living index, sample representativity, quality adjustment, static and dynamic universes of transaction prices, matched and non-matched prices, stratification.

Introduction

- 1 This paper arises out of work, in the Eurostat project to Harmonize CPIs in the European Union, to formulate guidelines laying down the requirements for the representativity of samples and for making appropriate quality adjustments in the CPI. The aim is to address the problem of quality adjustment from within the "pure price index" framework that is accepted in the design and construction of the majority of actual CPIs. A common framework maybe found in the choices that have been made over the years in deciding, in particular contexts, what to observe and what procedures to follow in constructing an index from what is observed. The CPIs are entirely determined by these, notwithstanding any references made to higher principles or concepts. The framework is elaborated by reference to what is in principle observable in an actual universe of transactions and the limited procedural options then available for compiling an index from what is observed. For convenience the rationale and background of the analysis are given in some detail in the endnotes. The paper is the author's contribution to the ongoing debate. At this stage the limited conclusions are his alone. Whether they lead to changes in actual practice in the European Union is a matter for majority endorsement by Member States (MSs).
- 2 The paper takes a strictly pragmatic view of the role of the conceptual framework, unlike those taken by Triplett (1999) and Diewert (1999) in their papers to the Group. The framework is required to resolve the operational dilemmas facing the index compiler the most urgent of which, by common consensus among European index compilers, is the treatment of quality change. The author believes that the total failure to make appropriate allowance for quality change is in part due to the emphasis of research on the development of conceptual frameworks with no evident practical relevance to this specific problem. Progress is to be achieved by taking new concepts into the "pure price index"

concept, where these resolve practical issues. Consensus on concept is central to Eurostat's search for improvement. While most CPIs are currently Laspeyres type indices, there is no reason why other index forms, such as the Fisher, cannot be adopted where this becomes a practical option. Likewise there is no reason to reject concepts such as "utility" where these provide a solution to the problem of quality adjustment. It is the author's contention, however, that as yet such concepts cannot provide a guide to the choice of practice until their relation to what is observable is demonstrated.

- 3 The author accepts the argument (Triplett 1999) that the position taken herein is only part of the rhetoric. Unlike Triplett he believes that it is by rhetoric we progress. There is, at least in the case of quality adjustment, no truth to be found "beyond the rhetoric" in micro-economic theory or elsewhere. We must seek only agreement on particular courses of action. (endnotes 15 to 21)

The Measurement Objective

- 4 As a starting point this paper takes the concept of a "pure price index" to be that exemplified by the Laspeyres-type index used in the European legal framework. The central idea here is the "changing cost of a fixed basket of goods and services". This differs in concept (but not in practice) from the cost of living index central to which is the notion of "constant utility". The aim of constructing a CPI is taken to be the measurement of the impact of changing prices on a given population of consumers. The impact is the direct effect of changing prices through the payments made by those consumers for the goods and services they buy. In practice, however, the CPI is a measure that is the index compiler's interpretation of the changing cost of the components of the fixed basket when many of the products therein have in fact changed. That interpretation involves various procedural devices to address the problem of quality change.

The Measure in Practice

- 5 A CPI is only strictly defined when it is constructed (Sellwood 1997). That is when all the operational issues have been settled. The index is determined by the prices observed and actual practices followed in its construction. Despite claims that different concepts are followed, actual CPIs are constructed following closely similar practices. Likewise, much debate with users over uses of the CPI has had little impact on CPI practice in general and none on quality adjustment practice.. Actual CPIs are necessarily constructed by the procedural adjustment of selected price observations. The process may be guided by principles such as consistency, coherence, relevance and efficiency but these too can only impact on the measure where they determine the choices of actual practice. Any notion or concept of what is, or is not, inflation is only relevant if it determines the choice of practice.
- 6 The role of the conceptual framework is to provide a basis for choosing between alternative practices. The arguments used for the choices made constitute the existing conceptual framework. The "pure price index" concept is given not simply by the Laspeyres formula but by the sum total of practices followed and their supporting rationale. The focus of this paper is on the arguments used to justify the choices of quality adjustment practices. This is examined by reference to the universe of actual transactions in consumer goods and services. Though such a universe is implicit in CPI practice it is seldom stated explicitly (endnotes 22 to 27). It is elaborated here as a necessary part of the index concept. The view from the universe is that of the "fully informed index compiler". It allows us to say what measure we would construct if every feature of any transaction that can be observed were actually observed.
- 7 In the interests of simplicity the arguments used are for the most part expressed with a simple universe in mind. One equivalent to a stratum comprising a single category of good, say for washing machines, may be taken as typifying any good or service. We first define a universe that is relevant to the fixed basket in order to expose the issues of quality change. We then look at the current practices as a guide to the limits that are imposed on CPI operational design. We consider what choices the conceptual framework is required to guide and whether it does so where quality change is involved. Finally we consider the implications for CPI sample design.

The Universe (endnotes 22 to 27)

- 8 A sample statistic always seeks to represent a universe (or population) statistic (or measure). In order to know how well a sample CPI does this we must first define the universe that is to be represented. This we take as the universe comprising all transactions over a period of years. However, without loss of generality we can take two periods of time $t=1$ and $t=2$ for which we wish to measure price change.

In order to relate the analysis to the "fixed basket" approaches to CPI construction it is convenient to sub-divide the universe into two distinct parts, the static (fixed) and the dynamic (non-fixed).

- 9 The two sub-universes we define are:-
- a) The "static variety universe" is that part of the total universe in which the quality of available products is static, the sub-set of goods and services available at both $t=1$ and $t=2$. In this sub-universe all goods available and purchased (transacted) at the beginning of a period are available and purchased (rather than simply available and hypothetically purchasable by the base period consumers) at the end of the period and no new varieties are included. *This sub-universe is the "actually fixed basket". It corresponds, loosely, to the "matched" part of actual CPI samples where "one to one" matches have been found and where there are no "replacements".*
 - b) The "dynamic variety universe" is the sub-universe comprising the set of transactions in goods and services available at $t=1$ or at $t=2$ but not both (those goods that disappear after $t=1$ and those that appear before $t=2$). None of the models at $t=2$ correspond to those transacted at $t=1$. They will have some characteristics in common and will be sold in the same outlets (mostly). The consumers involved will, for the most part, be different. *This sub-universe is the "actual non-fixed basket". It corresponds, loosely, to the non-matched part of the sample where "replacements" are involved and where quality adjustment is generally considered appropriate.*
- 10 The distinction is between that part of the universe where varieties or models of a product remain effectively unchanged and that part where the varieties disappear or appear over the period. For many goods and services or for short periods the static universe may cover practically all transactions. It is arguably a large part of whole universe for many goods and services not subject to technical innovation where small changes in characteristics are overlooked in current index practice. However, new models are introduced in CPI samples and these can make some claim to represent the dynamic universe. It is, nevertheless, the contention of this paper that there are key issues of quality change and sample representativity that can be illuminated by consideration of the static variety universe. Problems here must be resolved consistent with the resolution of the larger issues raised by the dynamic aspects of the universe.

Practices Determining CPIs

- 11 At the level of the universe the CPI must be expressed as a combination of the prices that are involved in transactions. (Weights are the numbers of transactions). The index is entirely determined by these prices and the limited number of ways in which they can be segregated and aggregated according to any observable characteristics of the products or the transactions. The options can be seen in the actual practices currently involved in index construction at the sample level. These are the enumeration of goods and services, the selection and observation of prices and their associated variables e.g. outlets and characteristics, the matching of observations through time and the replacement when matching fails, re-selection or re-sampling. There are options for stratification or classification of the prices observed on the basis of associated observations. There are routines, or formulae, followed to in combining prices to produce a stratum price index, simple combinations (averages) of the observed prices or of the observed prices after some adjustment. Finally there is the procedures for combining stratum indices to give an overall index.
- 12 The practice of "matching" prices or transactions, and hence that of making a "replacement", does not seem relevant at the level of the universe. There are no one to one correspondences and there are no "missing" prices or products. These practices seem to serve the same purpose for the sample as the distinction between the static and dynamic achieves for the universe.
- 13 The adjustments to observed prices include editing procedures and those for "quality adjustment". Current practices for quality adjustment aims to partition observed price differences between price and quality. For the most part, however, there is a mechanical adjustment, as in "overlap pricing", "linking", "imputing base prices" or "hedonic" procedures. Only where adjustment is achieved subjectively by reference to "price collectors" or "commodity experts" is new information added to that available in the defined universe. Then the new information is not a matter of observation. Hedonic adjustments in practice may involve data sets other than those sampled directly for the index but they are observable data from the universe of transactions. The procedures may involve subjective judgements of the modeller or "experts" but are otherwise mechanical.

- 14 Note that quality adjustment of observed prices is not a practice that is relevant the static universe where there is no change in the quality of any product. Quality change arises from the introduction of new models and the disappearance of old models and is confined the dynamic variety sub-universe. It is in this sub-universe where the problems of quality change arise and must be solved. The universe comprises the set of transactions in goods and services available at t=1 and/or t=2 but not both (those goods that disappear after t=1 and those that appear before t=2). None of the models at t=2 correspond to those transacted at t=1. They will have some characteristics in common and will be sold in the same outlets (mostly). The consumers involved will, for the most part, be different. There is no one to one correspondence in this universe and therefore no obvious correlate to the practice of replacement at the sample level nor to the selective adjustment of individual prices.
- 15 The formulae used for combining prices within the sample strata can be used for combination in the universe strata but the availability of data on the numbers of transactions provides more options. The formulae used for combining sample strata may likewise be used to combine universe strata. The supposition of complete information allows us to define index constructions with whatever weighting we may choose.
- 16 At the level of the universe, the key factors determining CPIs of relevance to quality adjustment are what can be observed, the stratification and the procedures for combining prices within a given stratum (elementary aggregate). The role of concept, theory or framework, is to guide the choice of practice. To do so there must be a distinction based on observables and clear options for subsequent action. "When X is observed do Y". We now look at what can be observed and what distinguished and the association of these with the procedural choices identified above. We then seek to establish what issues conceptual development is required to resolve. Finally we consider the implications for sample design.

What can be observed and what distinctions can be made?

- 17 In the universe the observables are the quoted prices, the transaction prices, the characteristics or specification of the good, the context of the transaction (outlet, service, etc) and the consumer. The universe comprises all transactions at t=1 and t=2 and the prices, characteristics, contexts and consumers involved in these transactions are observable in principle. In practice much of what is observable may be actually captured in scanner data. The consumers at t=2 will differ from those at t=1, and while much may be observed about consumers their motivations or judgements on their purchases are not. In particular cases, such as major durables, there may be no consumers common between the two periods. We, therefore, take this as the norm in this paper.
- 18 Products may be distinguished by their type and/or by their characteristics. Goods and services are divided into categories for various reasons such as an interest in the different purposes they serve. Such classification is a matter of convention rather than principle. The effect of standard expenditure categories in stratification is to impose a certain order derived from purposes other than the design of a CPI. Such categories do, however, reduce price differences (variances) within strata and increase price differences between strata. Product characteristics allow us to make distinctions between products within classes and between products across time. There are neither conventions nor principles for deciding what characteristics distinguish products by quality. Hedonic regression allows us to detect associations between price and characteristics (endnotes 55 & 56).
- 19 Distinction can be made between prices quoted prices and those actually paid in transactions. For some quoted prices there may be no transactions in a given period. Price change, in order to impact on what consumers pay, must involve transactions. This is not to say that a quoted price may not affect the consumer. It may deter him/her from buying.
- 20 Distinction can be made between change in prices through time and differences between prices at a given time with the specification or location of the transaction. That is, we distinguish between $\frac{\partial P}{\partial T}$ – and $-\frac{\partial P}{\partial S}$, between inter-temporal and inter-spatial price differences. It might be argued that only changes in transaction prices with time should represent price change for the purpose of the CPI. However, price differences affect the average price paid where there are changes through time in the numbers of transactions at different price levels. Prices differ between outlets for the same model because of differences in costs or profits. They may differ between closely similar models for

similar reasons or because the consumer is prepared to pay more for one model over another. To the extent that a price difference represents a quality difference the average quality will change with changes in the mix of transactions.

- 21 A distinction is often made between price levels and price changes. The emphasis here is on differences rather than levels because price differences can be easily mistaken for price changes. The issue is whether any price differences should count as price change in certain circumstances and if so how they should be counted in the measure. *Where no quoted price is changed is there inflation? If there are no transactions should a change in a quoted price be counted as inflation?*

Distinctions made by Stratification (Matching)

- 22 Stratification of the universe for the purpose of constructing a CPI it allows us to refine the measure. It allows us to make a distinction between one set of transactions and another and hence between their transaction prices. It allows us to treat prices within one stratum separate from those in another stratum. Stratification must be according to some observable feature(s) of the transaction. Without stratification (separation) we cannot treat some prices differently from others..
- 23 The universe at a given time t may be stratified down to (or up from) the level of individual transactions. Price is not used per se but features such as model and outlet are often associated with price and to this extent stratification tends to aggregate transaction with similar price levels. From a sampling point of view stratification may improve efficiency and reduce the overall sampling error where it maximises between strata variance in price change and minimises within stratum variance of price change. The aim must be to choose strata appropriate for the resolution of the quality change problem. The choice of stratification must distinguish prices affected by quality change from those that are no. The distinction between the static and the dynamic is thus an obvious case so long as there are procedural choices available.
- 24 For the static variety universe embracing both $t=1$ and $t=2$ stratification is restricted to features that are common to transactions in both periods. This will depend on the degree of detail specified in the stratification criteria. This may include outlet type or individual outlet. It may define specific models or brands or it may take several varieties together. The stratum includes the transactions for the particular model set at $t=1$ and those at $t=2$ that occur in the same outlet set. There is no question of consumers repeating the exact same transactions at $t=2$ as they made at $t=1$. So that correspondence is limited to the content and context of the transaction. The limit of stratification is the transaction price (or quoted price) for a given model within a set of models in a given outlet within a set outlets at $t=1$ with the observable prices quoted for the same model in the same outlets at $t=2$. These are “matched prices” and matching may thus be seen as an extended form of stratification of the universe. Conversely stratification is a limited form of matching. Strict and detailed stratum criteria or matching reduce the size of the sub-universe.
- 25 For each stratum encompassing $t=1$ and $t=2$ in the static or fixed basket universe we can define a corresponding stratum in the dynamic or non-fixed universe. These strata comprise products serving similar purposes but differing in certain characteristics. They do not meet the strict criteria of the static stratum and would not count as matched or comparable in sample constructs. The dynamic universe therefore comprises at least as many strata as the static universe. (New categories of product might increase the number of strata.)
- 26 Hedonic regression offers the possibility of further stratification according to the characteristics of products. In the static universe products will have the same characteristics at $t=1$ and $t=2$ and there will be no new characteristics. The dynamic universe will comprise products with characteristics at $t=2$ that were not observed in products at $t=1$.

Choice of Procedures for combining prices within a stratum (endnotes 40 to 64)

- 27 The measurement objective for the sample based CPI can and should be defined at the level of the universe. It is at this level that the relevance of any notion of the pure price index or fixed basket index must be evident. In the endnotes we look at the more obvious ways in which prices may be averaged to produce a stratum index. Between time $t=1$ and $t=2$, in the static variety universe, prices of the individual goods can change but there is, by definition, no change in the quality of any available good since no new model is added or old model lost. There may, however, be changes in the quantities of a good with given characteristics transacted. The mix of characteristics in the transactions changes only

in as much as the mix of transactions changes, not with the addition of new characteristics. Even where no single price changes between $t=1$ and $t=2$ the average transaction price may change because the numbers of transactions at different prices change.

- 28 This problem is not new. The problem might be resolved by taking a Laspeyres formula for combining prices or, since the Paasche formula would give a different index we might, in a fixed basket approach agree on a Fisher index. However, the point here is that there is an issue of quality change even though the quality of no single available product has changed because of changes in the numbers of transactions in different products. The problem arises not because there is quality change but because there is quality/price variation within the strata. If we accept this as a quality adjustment problem then we must ensure the consistency of the procedures used to resolve it with those used to deal with quality change in the dynamic universe.
- 29 For the combination of prices within the dynamic strata we do not have the same choice of formulae available as for the static strata. We have a set of transaction and prices at $t=1$ and a non-corresponding set for $t=2$. The quality is changed. Fixed basket constructs are not possible. This perspective contrasts with that found in sample practice where allowance for quality change is made on the basis of quality differences in one for one replacements. At the level of the universe there is a many to many product mapping with both price and quality change.
- 30 The distinction between the static and the dynamic is not clear-cut. The extent to which a product specification must change before it is judged to be different is a matter of choice. The resulting universe CPI depends crucially on the choice that is made. The static variety universe will never be entirely static. It is a universe in which product change is ignored for the purpose of measuring price change. To this extent it is a universe of pure price change.

Hedonic Methods

- 31 As noted above hedonic regression can extend stratification to product characteristics. It thus allows the mix of characteristics to be controlled in the aggregation of prices within static strata thus achieving a fixed basket solution. It will, however, produce a different stratum index from those produced by other fixed basket approaches thus implying a different allowance for quality. There is no reason in practice why associations between price and characteristic found in the static universe should not be used for adjustment in the dynamic universe but the case for doing so needs to be established and/or agreed. If agreed, then criteria as to when and where it should be applied must be stated. The fact that some significant coefficients have been found does not validate the assumptions and tells us nothing about the limits of applicability.(endnotes 56&57)

Choice of procedure for combining strata.

- 32 In principle an overall stratum index may be constructed by weighting together separate stratum indices, one for the static variety sub-universe and one for the changing variety universe and aggregating these across strata. There is an issue of what the criteria should be for choosing strata corresponding the distinction static v dynamic. There is no principled argument for any particular choice of strata. Why separate meat from bread? Why separate one outlet type from another? Clearly one might argue that certain markets are segregated but there will remain some arbitrariness in the categories used with potential implications for the treatment of quality change.
- 33 The universe CPI is defined for single periods, $t=1$ to $t=2$, but operationally it must be defined for continuous periods. Between any two points in time there is a distinguishable universe of transactions where there are no new models. The included models depend on the time period taken. Alongside this universe there is a universe of changing varieties the dynamic universe where old varieties disappear and new varieties appear. The static variety universe exists only from $t=1$ to $t=2$ we have considered only the construction of a measure for such single periods. Extending the measure to cover several periods raises separate issues that, in part, correspond to practices of re-sampling. New models may be accommodated in the assumed universe at $t=1$ but not between $t=1$ and $t=2$. That is to say that new models might be included when there are significant numbers of transactions.

Implications for Sample Design (endnotes 65 to 71)

- 34 It is evident that the sample practices of matching and replacement attempt to capture both the static and dynamic universes. The matching rules (practices) define the size of the static universe that is represented, the more strict the rules the smaller the universe. The replacement rules define the

dynamic universe and a requirement for "close replacements" will place an unduly restrictive limit on what is represented. Together these rules have crucial implications for amount of "quality change" that is detected and the number of quality adjustment that are made.

Conclusions

- 35 Even if we could observe all that there is to be observed about the universe of transactions in consumer goods and services there are strictly limited options for the CPI that we can construct. Price data can be segregated or stratified only according to the observed features of transactions or products thereafter there are few options for their combination and aggregation. Conceptual frameworks can only determine the measure through the few choices that exist between actual practices.
- 36 The pure price or fixed basket framework dictates certain practices that are followed in all CPIs but does not provide adequate criteria for the choice of strata necessary for distinguishing quality change from price change. As a result CPIs are determined by arbitrary decisions on what counts as evidence for quality change. Even with appropriate adjustments for quality change two actual CPIs for the same universe would differ more on account of operational practice than on account of sample error. Given that most of the actual quality adjustment is mechanical the potential divergence between two CPIs on this count will be determined by decisions on sample matching and the choice of replacement.
- 37 The pure price index framework must be developed so as to provide guidance to actual CPI practice in deciding where quality should be deemed to have changed. A principled basis for the stratification of transaction is required. This should say where an observed price difference reflects a quality difference and what must be observed as evidence of this. In the absence of new concepts there is an urgent need for internationally agreed conventions.
- 38 Only where the concepts of the cost of living theory can be demonstrably related to actual practice will they be accepted as providing an advance on the pure price index framework. That the theory provides a warrant for hedonic quality adjustment applies only under unrealistic assumptions about the universe of transactions. Empirical evidence of informed rational consumers in some markets would be a start. Meanwhile hedonic regression may be the most promising of the few options for practical quality adjustment providing that conventions can be agreed for its use.
- 39 CPI sample design should be reviewed by reference to some defined universe of transactions. The actual practices of stratification, matching, replacement and re-sampling require an appropriate framework of statistical inference. Otherwise there is no consensus on what sample CPIs are intended to represent. The analysis herein suggests that there are a number of related issues to be resolved (endnotes 72 to 76).

End notes

- 1) The paper arises out of work in the Eurostat project to Harmonize CPIs in the European Union to formulate guidelines laying down the requirements for the representativity of samples and for making appropriate quality adjustments. It follows previous papers by the author (Sellwood 1994, 1995, 1997 and 1998) seeking to relate discussions of concepts to the practical issues faced by index compilers. The present argument is part of an ongoing debate and is far from complete so there can therefore be no firm conclusions. Where the arguments require further elaboration it is given here in note form. There are few matters of fact with regard to quality adjustment or representativity in the CPI. *Text in italic is used to highlight propositions involving particular conceptual ambiguity.*

The Rationale

- 2) The paper attempts to tackle the problem of quality adjustment in CPIs from the perspective of sampling theory within the context of actual operational practice and free from presupposition about the nature of the consumer. It starts from the proposition that a break from established dogma is needed if a consensus is to be reached. The main barriers are seen as sample thinking, the failure to define what is to be represented, and argumentation from economic concepts that provide no clear distinctions as a guide to possible procedural action. The paper seeks to establish a basis for improved in index construction by re-thinking practices within the long established and accepted framework of statistical sampling theory and by introducing concepts only where they affect practice.
- 3) Emphasis is placed on what can be actually observed and on actual procedures the effects of which are known and warranted. The author recognises that any attempt to measure inflation involves a theoretical perspective but it is suggested that concepts that relate to unobservable entities confuse discussion without providing such warrant. The aim is nevertheless to remove as much of the conceptual presupposition surrounding CPI construction as possible.

Guiding Premises

- 4) All action, however ostensibly practical, including observation, description and re-thinking, involves some theory, concept or assumption on the part of its advocates. Some simple assumptions such as; “a good measure for part of the universe is better than a poor measure for all” are not examined. Various heuristics such as “don’t let the problems you can’t solve get in the way of solving those you can” also go unquestioned. Otherwise the main premises on which the re-thinking is based are given below. They have a better pedigree. It is accepted that arguments for changing CPI construction must be justified by whether or not they lead to practices that are by common agreement “better. The argument should lead to specific conclusions as to practice that will guarantee improved measures.
- 5) **Sampling Theory.** Sampling theory has been empirically tested and widely accepted. It is a theory concerning what legitimate inferences valid can be made about the universe from which we have drawn a random sample on the basis of that sample. It follows by deduction from what must be the case when we take random samples from a given universe. Price statisticians must accept the tenets of this theory. In particular, the theory
 - requires the population or universe that is to be represented by the sample to be defined
 - provides rules of inference from the sample to that universe
 - provides a framework of method by which to design efficient samples with a focus on the variables of interest.
- 6) All CPIs are based on sample data but the basis of sample selection for most price indices is “purposive” rather than “random”. We must therefore examine the consequences for our inferences when we do not have random samples. In this context the role of stratification achieving adequate representativity would seem critical.
- 7) **Simplicity of Concepts.** Following Ockham’s razor (the principle of parsimony), which dictates a bias towards simplicity, as few concepts as possible should be used. Unnecessary concepts lead to ambiguity in agreeing the common practices that are required for the international comparability of CPIs. Concepts should only be introduced where they are necessary to make a distinction from which action, with an effect on the actual measure, follows. Historical concepts that are irrelevant to practice or unwarranted should be discarded in favour of new concepts as required. We must know how, and to what extent, particular concepts affect results.

- 8) **Distinction.** Reyes & Zrama (1998) argue that “Making a new distinction allows us to bring forth something we did not recognise before”. The approach asserts that the actual measure of inflation produced is determined by the actual practices followed and thus only those distinctions that affect actual practice are relevant. Distinctions in the form of agreed definitions or concepts should only be made where they make a difference to practice and results. Definitions and concepts are relevant to the distinctions made between particular observations or specific actions (computations) within the CPI operations. Definitions or concepts outside this domain should have clear consequences for observation and action within the domain.
- 9) **Observables.** Boylan and O’Gorman argue that “Although one must be careful to distinguish between observing and detecting, science, by means of theory-laden instrumentation, makes an indispensable contribution to human knowledge by extending the bounds of the observable... The hidden, however, which we can reveal, is limited to what is in principle observable”. The discussion of quality adjustment in CPIs has become somewhat removed from the observation. CPIs are determined solely by observations and by the actions performed on those observations. Notwithstanding the a long standing argument that no observation can be independent of the observer is implicit in the quotation we take observables to be entities that are matters of fact independent of the observer. They replicable, the same result is obtained whoever observes be he/she expert or not. The possibilities offered by scanner data are used in the paper as a guide to whether something is observable. The key observables are prices, quoted and paid, quantities purchased, product characteristics and consumers. Observables are to be distinguished from non-observables such as utility and quality, which as yet remain "hidden".
- 10) **Testability.** The approach requires that wherever possible tests should be made to ascertain that rules governing observation or practice lead to a measure giving appropriate results in all probable circumstances. For example, the conditional “if no prices change then the measure of inflation should record no change” might be one such test. Supplementary statistics can help to highlight where procedures give poor results or where the measure is insensitive to procedural choice. Judicious use of tests can give credibility to the CPI. Few statistical offices can say what allowance has been made for quality change and many do not calculate sampling errors. They have been vulnerable to, often poorly founded, criticism of bias or inaccuracy.
- 11) **Progressive Approximation.** The argument finally accepts the premise that the perfect CPI is not definable. A "better" CPI must be discovered by a process of progressive improvement in practice as proposed by Demming (see for example Ishikawa 1985) in his Total Quality Management approach. Any improvement is to be welcomed and it is inappropriate to impose standards on new approaches that far exceed those that are met by existing practice, which has major limitations.

Background Remarks

Precedent -Mind sets -Barriers to progress.-Why a sampling theoretic approach?

- 12) Early papers by the author argued that the conceptual frameworks within which CPIs have been developed can not be reconciled (Sellwood 1994). Existing practice is a function of its historic development with different emphasis in different countries and flowing from different practitioners and theorists. The two dominant paradigms, characterised by microeconomic theory and the sample index construct or “economic think” and “sample think” (see below), have failed to resolve the central problems of bias and quality change. The former has introduced concepts that confuse rather than guide actual practice and the latter has resulted in a failure to relate the sample measures to the dynamic universe for which the measure is required. Notwithstanding the requirements of sampling theory CPIs are without exception defined by the samples used in their construction. These samples are said to be representative but little attempt is made to say “representative of what”. That this is so is testament to the difficulty of the problem of defining a universe that gives an appropriate perspective on the issue of changes in the goods and services available (quality change in the broadest sense)
- 13) Like all thinkers CPI statisticians and their critics are constrained by precedent (Sellwood 1994). Their current arguments and actions follow the thinking of their predecessors The choices between COLIs, fixed or chain Laspeyres indices reflect prejudices of those involved in, or with an influence on, the construction of CPIs over a number of decades. This is both a strength and a weakness. A strength in that the consensus that has been achieved lends credence to the CPIs but a weakness in the failure to develop further consensus on the way to overcome the most intractable of CPI problems, quality

change. Here, while common practices are followed they are unwarranted and, without exception, unsatisfactory. There are differences in the conceptual frameworks used to support the attempted rationalisations for particular practices that cannot be resolved but in fact these differences have rather less impact on actual practice than might be expected from the strengths of conviction expressed.

- 14) Following the notion of “group think” used by psychologists researching the “Bay of Pigs” and similar cases of human error the paper characterises two types of thinking that inhibit progress on CPI problems, “economic think” and “sample think”

“Economic Think”

- 15) CPIs are only determined by observation and practice (the procedures that are actually followed). Concepts and theory are there to help, to give consistency and reduce arbitrariness in decisions. Different concepts and theories may justify a single practice. If different theories do not indicate different practice the resolution of the theoretical difference has no consequence for the measure. The test of a concept or theory is the its descriptive adequacy, its relation to observables and practice. The accusation here is not that economic theories, such as "utility theory", applied to CPI thinking are wrong (the jury is still out on whether they are true, approximately true or false). It is that there is a lack of shared understanding as to the relevance of theory to actual practice and this has meant that consensus has not been reached and is unlikely to be. The belief that the solution to the problem of quality lies in concepts such as utility is a barrier to progress towards improved practice. (Sellwood 1997).
- 16) The author's response to the Diewert/Triplett case is that it contributes nothing immediately to the practical problems of quality change. Triplett (1999) suggests that micro-economics can take us past the rhetoric to the truth. McCloskey, who he quotes, seems to argue (McCloskey 1994) that we cannot escape from the rhetoric, noting that "The main figure of economic rhetoric has become the conspicuous use of mathematics". As to the rigour mathematics brings McCloskey says "If we are going to be rigorous we should be rigorous, not rigorous about the proof and extremely sloppy about its range of application." The author's line reflects more the analysis of Boylan and O'Gorman (1995) and seeks consensus rather than truth. Hence the attempt to distinguish observables from concepts or non-observables. It is surely only a matter of rhetoric, not "beyond rhetoric", where the line between income and price is to be drawn for owner-occupiers' housing costs. Purity of concept has ultimately to be judged on its practical impact.
- 17) The relevance of the hypothetically fully informed consumer and his/her utility and preferences to actual CPI practice has not been demonstrated. Results of hedonic methods, which assume the theory to be true, cannot provide such a demonstration. What is needed is some empirical tests of the assumptions on consumer behaviour or the consequence of departure from those assumptions on the quality change estimates. In the absence of such evidence there is only rhetoric. Debate is a perfectly legitimate means by which to resolve problems, indeed it is the normal means. However, it should not be allowed to limit the options for wider debate. "Utility theory" may have been the only game in town for the past twenty or so years but, as many of us don't see any point in playing it, it is time we invented alternatives. Microeconomics is theory in search of an application. It has yet to find that in quality adjustment, notwithstanding the contribution that it has made to the debate so far.
- 18) The case against utility theory seems overwhelming (Sellwood1997). The theory cannot differentiate "utility" and "quality". Utility is a truism, a useful tautology. It says only that consumers always make what they consider to be the best choice, only a fool would do otherwise. It does not follow that consumers aim to maximise one thing that is “utility”. There is no “thing” that is “utility” and the concept can embrace any basis for preference (including colour or style) including those that are not what are generally understood as “quality change”, such as technical advances in motor vehicles and computers. Utility of the theory neither explains nor predicts actual consumer choices. The assumptions (fully informed and market clearing) need not be valid for the theory to provide a useful perspective on market processes but they must surely be substantially true for the empirical estimates to be an appropriate measure of "quality change" in the sense used by both Boskin (1999) and index compilers. That micro-economic theory provides a good description of the market process, as Kirzner (1992) argues, is surely only all that Triplett's empirical evidence for consumer preference shows. That a consumer expresses a preference in taking a decision to buy says nothing for the basis or stability of preferences within or between consumers

- 19) .If COL theory does give a good approximation to reality then surely the rational consumer should take price as the best (least search cost) indicator of quality? Elizabeth Anderson(1997) argues that “When we evaluate something as wonderful, awesome, delightful, and charming, we are judging that it makes sense for us to be wondered, awed, delighted, and charmed by them.Kantians reduce ‘x is good’ roughly to ‘x embodies or satisfies a principle of practical reason.’ I reduce ‘x is good’ roughly to ‘it is rational to value x,’ where to value something is to adopt toward it a favourable attitude susceptible to rational reflection.” However, when we talk of quality change and CPIs we generally have in mind technical change and some aspect of the performance of the good or service. A theory must separate this kind of valuation from other kinds.
- 20) The theory must also enable us to differentiate fashion from quality. There is not only women's dresses but also "this years' model" of durable. The tendency of consumers to be influenced by advertising and by other consumers which Ormerod (1998) describes as ant-like behaviour that can influence prices and lead to the preference for technically inferior goods (the much quoted case of QWERTY key boards). We all know the quality of PCs is improving but we do not agree on what this statement means. Some believe that “the consumer” or “the market” has sorted out the value of the quality improvement and are prepared to take the “speed” or “memory” of the PC as the basis of a measure of quality improvement ignoring the role of software. GM foods and food additives show how consumers can come to give overwhelming weight to particular features of a product
- 21) That micro-economics has contributed to a fascinating debate is beyond question. Diewert's (1999) propositions as to alternative index formulae are not questioned here. The challenge is to his support for Triplett's (1999) claim that a COLI framework is somehow superior to that of the fixed basket approach, particularly when it comes to the issues of quality adjustment. Both frameworks are an amalgam of disjoint concepts lacking an overall coherence. Both borrow ideas from each other and other frameworks such as national accounting. It maybe that COLI theory currently favours a user cost approach to housing but there is no theoretical necessity in this and no necessity that the pure price approach should preclude a "user cost" measure. There are simply agreements in practice. The CPI concept is embodied the practice until there is a COLI there is no relevant alternative concept.

“Sample think”

- 22) The question of what the CPI sample represents seems to have been abandoned as unanswerable. Since sampling universes have proved particularly difficult to define the practice has been to avoid any reference to a universe in discussion of the practical problem of CPI construction. This paper contends that this is a pervasive way of thinking that has stood in the way of solving some problems.
- 23) .It is in fact meaningless (lacking a theoretic frame and therefore incoherent) to talk of any aspect of a sample other than in relation to what it represents for some universe. The sample per se is of no interest. It is what one can infer about the universe that matters. Whether the sample represents all the changes that occur in the universe is an empirical question but, by the rules of inference, each and every change in the sample represents some change in the universe. There is no change in the sample that is purely a matter of the sample without implication for the price change that is being estimated for the universe.
- 24) If a particular price can no longer be observed this cannot be solely a sampling issue. There must, in theory, be a relevant issue of inference to the population CPI (there are, it must be inferred, similar events in the non-sampled universe). This is not to say that there are not considerable conceptual and practical difficulties in specifying the universe, nor that the failure to do so invalidates a particular CPI, only that representativity necessarily requires reference to the universe. Even if a universe is not specified one is implied by any given sample. The questions to be answered are “What is it about the universe that must be represented in the sample?” What are the most appropriate possible universes? What difference does the choice of universe make?
- 25) Evidence for “sample think” is:-
- i. The literature of price indices ignores the problem of defining a universe. The Laspeyres and other index concepts are not defined for the universe of prices but only for the sample of prices
 - ii. The requirement to match prices from month to month has become an end in itself rather than simply a procedural device to deal with a changing universe.
 - iii. A missing observation is treated as a “failure to match” rather than evidence of a change in the universe.

- iv. Replacements are chosen so as to re-establish or maintain the sample rather than to represent the change in the universe.
 - v. The procedures for treating replacements are defended, where they are defended rather than just accepted, by reference to the sample. E.g. imputing a new reference price might be justified as being the result that would have been obtained had the replacement been chosen at the outset.
 - vi. Forced replacement is seen as confined to the outlet chosen in the sample rather than an event occurring in the many outlets represented by the sample outlet.
 - vii. Fruitless discussion as to whether “substitution” is what statisticians do or what consumers do.
- 26) Sample thinking has led to a situation in MSs where the methods of constructing CPIs are justified by reference only to the samples used. Problems of replacement are treated as problems of the sample and not as problems of estimating an index of changing prices in the universe. As a consequence a variety of procedures have been developed for resolving operational difficulties and it is impossible to agree, when they produce different results, whether one procedure is to be preferred over another.
- 27) Thinking from the CPI sample implicitly accepts many assumptions about the nature of the universe from which it is drawn. These hidden assumptions should be exposed to critical appraisal.

Choice of Procedures for combining prices within a stratum

40 The measurement objective for the sample based CPI can and should be defined at the level of the universe. It is at this level that the relevance of any notion of the pure price index or fixed basket index must be evident. We look first at the more obvious ways in which prices may be averaged in the elementary aggregate index, and how these raise issues of quality change, starting with the simple ratio of the average prices paid, the unit value index.

41 **The Unit Value Index for the Static Variety Universe.** For this universe we can define, in terms of the observables, an average transaction price and changes in this average. The unit value index for the universe is the change in the average transaction price

$$U = \frac{\sum P_{2i} / N_2}{\sum P_{1i} / N_1}$$

Where every transaction (*i*) is regarded as unique.

42 Where the contexts of the transactions (outlets or regions) remain static we can stratify the universe by context. Likewise we can stratify according to available models (or brands). Stratification of the universe simply produces sub-aggregates according to observables. At the level of the universe it has no effect on the unit value index. All the transactions are used. They are simply aggregated in sets according to the stratification criteria.

43 Where there are N_s transactions in each context at prices P_{is} differing only in that different consumers and different models are involved we may write

$$U_a = \frac{\sum_{s2} N_{2s} \sum_i P_{2is} / N_{2s}}{\sum_{s1} N_{1s} \sum_i P_{1is} / N_{1s}} = U$$

44 Where there are no new and no disappearing contexts (outlets) the number of strata will be the same in the numerator as in the denominator. U_a can represent a stratified aggregate where either “different models of the same good are purchased in the same context by different consumers”, e.g. by outlet or location, or where “the same model of the good is bought in different contexts by different consumers”, that is stratification by product specification e.g. by brand. Both U and U_a are in principle calculable since all transactions are observable and in some cases may be found in scanner data. U is equal to U_a , which is simply obtained by performing the aggregation in stages.

45 The unit value index is not affected by changes in the total number of transactions but it is affected by changes in the mix of transactions. If no prices change, U can rise simply as a result of more

transactions being conducted at higher prices. That is, U and U_a reflect changes in characteristics and contexts of transactions in as much as these reflect price differences. Changes in consumers do not in themselves reflect price differences. The effects come through their choice of context and characteristics. The effects of changes in the mix of transactions associated with characteristics or the specification of the goods are generally taken as changes in average quality. The effects of changes in context (e.g. outlet) are sometimes regarded as changes in the average price. However, the distinction is not clear where the transaction is more broadly defined to include the context as involving a service. These effects are avoided by stratification in the traditional Laspeyres type indices.

- 46 **Laspeyres and Similar Index Constructs.** Traditionally changes with time in the CPI necessarily require that some price or prices are changed between $t=1$ and $t=2$ (and that some transactions take place at the new prices). The argument here is that, whatever else may change if no price changes and no model changes then the CPI should not change (there is no inflation, notwithstanding any changes in the consumers' view of the quality of goods available). Price change necessarily involves change through time. At one point in time t one cannot observe price change. Price differences do not represent price change. At time t price differences must reflect only non-price change factors.

- 47 If changes in models and outlets are not to be regarded as price change it is necessary to discount their effects in the measure. That is to remove the price differences associated with model substitution and outlet substitution. The Laspeyres index of the form

$$L = \frac{\sum P_2 N_1}{\sum P_1 N_1}$$

and similar indices are constructs for holding the "non-price factors" (the numbers of transactions) in U constant between $t=1$ and $t=2$.

- 48 The "Laspeyres index" poses the question "What if the consumers making transactions at $t=1$ made the same transaction at $t=2$?" It introduces a concept that involves a non-observable "fixed consumer". The index cannot be strictly defined for transaction prices in the static variety universe because full correspondence of the transactions at $t=2$ with those at $t=1$ is required and this does not exist, consumer behaviour is not static and nor are consumers. As noted above there may be some correspondence between quoted prices at $t=2$ and quoted or transaction prices at $t=1$. These are the observable matched prices sampled for the construction of most CPIs. The fact that the index can be defined for matched samples should not be accepted as providing an automatic warrant for this construct as many seem to think it does.

- 49 By stratification a Laspeyres type index can be implemented for the transactions in a static variety universe. Where transactions involve unchanging characteristics, for example outlet type or model specification, we obtain

$$L_a = \frac{\sum_s N_{1s} \sum_i P_{2is} / N_{2s}}{\sum_s N_{1s} \sum_i P_{1is} / N_{1s}}$$

Where the weight given to the strata are the numbers of transactions at $t=1$. N_{1s} is the number of transactions at $t=1$ in outlet s (or model s). The averages of within stratum prices are averages over all transactions the prices are not "matched" beyond meeting the criteria defining the stratum.

- 50 Matching in the sense used in sample estimates might be seen as involving very detailed stratum criteria, a particular model specification for a single outlet. Such stratification will reduce the size of the static variety sub-universe since slight changes between $t=1$ and $t=2$ will be taken as part of the dynamic sub-universe. An index based on strictly matched samples represents a limited universe. On the other hand loose stratum criteria allows more change to be disregarded and increases the static sub-universe and reduces the dynamic sub-universe.
- 51 The Laspeyres construct removes the effect of outlet (or model) substitution in a static variety universe. In as much as the universe may be static in respect of context it can stratified (e.g. by outlet, or outlet type and/or location). The formula is not defined where there are new outlets or new models. It is a weighted average of unit value indexes for outlets (or models). Price differences between outlets

(or models) are not counted as price changes when there is a shift in transactions between outlets (or models). The formula can be extended to cover both models and outlet type together.

- 52 Note that the stratum index in the above is

$$I_s = \frac{\sum_i P_{2is} / N_{2s}}{\sum_i P_{1is} / N_{1s}}$$

the ratio of the averages of the actual transaction prices where the

numbers of transactions may differ. This is necessarily(?) a unit value index at the level of the universe. It is affected by any change in the mix of models in the stratum. If this change in mix is a change in the average quality then some adjustment is necessary. Sample estimates based on matched transactions might remove such effects. What is the argument for doing so? Matching with replacement and quality adjustment is another option.

- 53 **The effective adjustment for changes in the mix of transactions.** Each of the stratification options gives a different picture of price change. The ratio of U_a to L_a shows the adjustment for removing the mix effects of outlets or models. That is

$$U_a / L_a \cong \frac{\sum_s N_{2s} \sum_i P_{2is} / N_{2s}}{\sum_s N_{1s} \sum_i P_{2is} / N_{2s}}$$

- 54 This is similar to the implicit quality index (IQI) and shows the change in price resulting from substitution between t=1 and t=2. The numerator includes all transactions at t=2 of models available at both t=1 and 2, whereas the denominator includes models available in the same outlets at both t=1 and t=2. New and closed outlets are excluded. Stratification by outlet-type would be less restrictive.
- 55 **Combining Prices in the Strata of the Dynamic Universe.** In this universe we have the same set of strata with slightly looser rules for inclusion. We can clearly construct a unit value index but the fixed basket constructs do not exist. Nothing is fixed beyond the stratum criteria, such as laptop PC. These may for example be distinguishable from other product strata only the purpose of product. The unit value index is affected by both price and quality as was the corresponding index for the static stratum. The difference lies in the degree of quality change. Note that there is no one to one correspondence, no replacements there are different sets of transactions and a many to many mapping.

Hedonic Constructs

- 56 The CPI can be regarded as the changing cost of a bundle of products and it is legitimate to argue that the products in this bundle may each be seen as a bundle of characteristics. In this case an index based on the regression of prices on the characteristics of the goods should provide an alternative index form. At the level of the universe there are no sampling errors but there will be associations (or correlations) between the prices and model characteristics. The extent to which characteristics appear to account for price differences is not in itself justification for hedonic procedures. The regression does not “explain” variations in price levels it is purely descriptive. It does not establish a separate price for particular characteristics corresponding to the separate prices for the models. Nevertheless, to the extent that it extends stratification to characteristics and allows the mix of characteristics to be held constant, hedonic regression is a means of avoiding counting changes in the mix of models transacted as price change.
- 57 The argument that the hedonic regression can establish the consumers' valuation of the characteristics involves unwarranted assumptions about non-observable consumer preferences etc. The associations between price levels and outlets, models or model characteristics are observable patterns in the data. Whether they reflect underlying mechanisms is a matter of conjecture. Whether these patterns may be used in the hedonic model to prevent price differences (model substitutions) being taken as price change requires an appropriate rationalisation. This might involve empirical comparison of hedonic measures with unit value indices and Laspeyres type indices using various degrees of stratification or matching.

Choice of Formula

- 58 Given full information on all transactions the straightforward options for measuring price change in the static variety universe are
- i. If changes in the mix of both models and outlets are to be counted as price changes then the unit value index is appropriate. Note that to the extent that price differences reflect quality difference the unit value index will reflect changes in the average quality of transactions.
 - ii. If changes in the mix of models are not to be counted as price change then the Laspeyres formula with stratification by model is appropriate. Where price differences between models reflect quality differences such differences do not impact on the index. (The numbers of transactions of particular models at $t=1$ should be used to calculate stratum weights.)
 - iii. If changes in the mix of outlet types are not to be counted as price change then the Laspeyres formula with stratification by outlet types is appropriate. Where price differences between outlets represent price reductions rather than quality differences such reductions will not be reflected in the index. (The numbers of transactions of the various models in particular outlet types at $t=1$ should be used to calculate stratum weights.)
 - iv. If changes in the mix of neither models nor the mix of outlet types are to be counted as price change then the Laspeyres formula with stratification by model and outlet type is appropriate. (The numbers of transactions of the various models in particular outlet types at $t=1$ should be used to calculate outlet stratum weights and the numbers of transactions of particular models in particular outlet types at $t=1$ should be used to calculate models within outlet stratum weights.)
- 59 The "pure price index" concept followed in the Laspeyres type CPI compiled by most countries takes option iv) above. In so doing they claim only that neither changes in the mix of outlets nor of models are in themselves treated as changes in price.
- 60 Whilst this option does not allow certain reductions in the average price paid the alternatives allow changes in the average quality to affect the measure. Unless there is some basis for partitioning price differences between model or between outlets into price and quality these are the only options. That certain price differences are generally accepted as reflecting quality difference is of little value. What is required is to know to what extent any price difference represents quality difference. Or in aggregate how much price change is accounted for by quality change.
- 61 A framework is built by posing hypothetical situations or asking "What if" questions. These are arguments *ceteris paribus* or argument by counterfactual conditionals. They cannot be conclusive because their assumptions are not met. They suppose "other things in the universe" remain constant, which they do not, or suppose that the universe is other than it is. The supposition that consumers repeatedly buy the same goods is the Laspeyres approach. The supposition that consumers are fully informed and maximise utility is the cost of living approach. Nevertheless they are a useful way of developing and exploring possibilities. The danger comes in believing that they are more than part of the rhetoric (see Triplett 1999).
- 62 **Quality difference and quality change.** Changes in transactions associated with changes in the consumers involved are addressed in the above analysis without reference to quality. If the concept of quality is introduced as a consumer dependent phenomenon then changes in consumers may be directly associated with quality change and hence price change. Price differences between brands or outlets at a given point in time might be said to reflect quality differences. Even so they are distinct from price change or quality change. If there is no inflationary element to price difference then either the difference may be explained by commensurate difference in quality or some other factor must be introduced to explain them. Price differences may be associated with different characteristics (models or brands) or different contexts (outlets or locations). If quality is introduced there is a parallel distinction to that made for "price" between quality change and quality difference. That is between $\frac{\partial Q}{\partial T}$ - and - $\frac{\partial Q}{\partial S}$. However, the relationship that matters is that between the differences $\{P_{it} - P_{jt}\}$ and $\{Q_{it} - Q_{jt}\}$. That is between the prices and qualities of different models or outlets (i and j) at a given time t .
- 63 The conceptual framework must give reason to treat some price differences differently from others. For example, why should lower prices associated with outlets be taken as lower prices while lower prices associated with models or brands are taken to reflect quality? Criteria, based on observation, are required to distinguish pure price difference from quality dependent price difference. Neither, micro-economic theory nor the pure price approach has yet provided such a rationale or criteria.

- 64 Existing procedures for quality adjustment, which are applied to sample data, are not relevant to the static variety universe where there is no issue of replacement. In the static variety universe the “quality problem”, if there is one, is solely the question changes in the mix of transactions. "Overlap pricing", "linking", "imputing base prices" and "hedonic" procedures take price differences as either all price or all quality or some arbitrary division between these. Such procedures must be judged by whether they achieve appropriate representation of the universe by the sample. Subjective quality adjustment procedures raise different issues.

Sample Estimates

- 65 The implication of the foregoing analysis for CPI sample design and operational procedures for CPI construction needs further work but some points seem relevant. All CPI practice involves a matched/non-matched distinction that determines the line drawn between the static and dynamic. The domain of quality change and quality adjustment is distinguished from that of stable or near stable quality. The distinction is thus determined by the rules for matching products at $t=2$ with those sampled at $t=1$. These rules involve:-
- a) Data, recorded, observed or surmised, on the products (specification)
 - b) Matching requirements and their interpretation in practice, the judgements that there are no "comparable" products.
 - c) Replacement rules and their interpretation in practice, the choice of the replacement product.
- The universe represented by CPI samples depends on justifiable inference from these.
- 66 Clearly if the full universe of transactions is to be represented samples should be designed to guarantee an appropriate balance of the two sub-universes. New models must be identified and incorporated as they appear and then treated appropriately, whatever that means. In scanner data the two sub-universes are distinguishable. With sample data post stratification on the basis of sample observation is necessary. Given a random samples of prices at $t=1$ and $t=2$ stratification requires knowledge of the existence of the varieties represented by sample observation at $t=1$ in the universe at $t=2$ and vice versa. Stratification gives some guarantee of appropriate representation of the static universe and representation of the changing universe from $t=1$ but not at $t=2$. The effect of matching sample observations and subsequent choice of replacements on the universe represented should be explored. The closer the matching the smaller the static variety universe. The closer the replacement the smaller the dynamic universe represented.
- 67 Samples must be designed to represent the introduction of new varieties (replacement and re-sampling). Sample design to ensure replacement and re-sampling properly represent the introduction of new varieties is critical to the accuracy of the CPI
- 68 It is not evident with current practices how CPI estimates based on samples relate to either the unit value index or the Laspeyres index at the level of the universe. Various procedures are followed when there are judged to be changes between existing models, non-matches and replacement. Some of these treat price differences as price change, which is inappropriate for the Laspeyres index, but the adjustments made are not according to changes in the mix of transactions, as would be appropriate for the unit value index. If the aim is a Laspeyres index then price differences between models should not count as price change. There should be no quality adjustments when models in the sample are replaced or, failing that, no net allowance for quality change (on average).
- 69 Stratification at the level of the universe was seen as an aid to the distinction between price difference and price change. At sample level it is usually seen as a device for partitioning variation and reducing sampling error, i.e. improving the efficiency of sample design. In a Laspeyres type construct appropriate stratification may provide some restriction on bias from purposive samples. (“Stratified purposive sampling” should be closer to “stratified random sampling” than “purposive sampling” is to “random sampling”.) Though this may require good estimates of the numbers and distribution of transactions in the base period for the stratum weights. The effects of taking quoted prices in place of transaction prices should be tested empirically.
- 70 There are no missing observations in the universe and it is therefore necessary to establish what missing sample observations represent.

- 71 The difference between the Laspeyres index and the unit value index is the allowance for model and/or outlet substitution but there is no obvious basis for deciding whether such an allowance is appropriate or, if it were, what allowance should be made. The unit value index based on unqualified observed prices and quantities is, therefore, an important reference statistic. The SRI (the standard reference index or change in the crude average of prices collected) and the IQI (implicit quality index, the ratio of the CPI to the SRI which measure the effect of all procedural adjustments to the crude prices) are proposed in the European work as estimates of allowances for outlet or model substitution (for so-called quality change).

Comments

- 72 The main conclusion to be drawn from this analysis is that, in deciding how to construct a pure price index, even a "fully informed" index compiler faces some difficult decisions in differentiating price change from quality change. Existing conceptual frameworks provide no guidance or operational criteria for deciding that quality change has had a price effect or, were there has been such an effect, how much adjustment should be made. Operational practice must be based on observables and reference to non-observables cannot resolve the practical dilemma. The only quantified observables are prices and numbers of transactions (and sometime pack sizes). Qualitative observations on product characteristics may be used to explain price differences in general terms but what is required is a clear partitioning of price and quality in each specific case.
- 73 The paper suggests that viewing the quality adjustment problem from the actual changing universe of available consumer durables, rather than from the information available in CPI samples, might provide an interesting and helpful perspective on the current practice of "matching" prices in CPI construction. The advent of scanner data allows us to think in terms of the alternative CPI measures in a universe where all observables may be observed and to test these empirically.
- 74 We can define a "static variety universe for which two kinds of measure of price change can in principle be computed. One allows changes in the mix of transactions, the "unit value index" and the other removes such effects, the "Laspeyres type index" (and other pure price constructs such as Paasche and Fisher indices). The indices may be constructed from observable data on transactions and the numbers and prices involved. There is no change in quality in this universe and, therefore, no need for "quality adjustment" in the sample. There may be changes in the mix of models transacted and it is necessary to determine how allowance can be made for these. Stratification by model in a Laspeyres type index with constant stratum weights based on the numbers of transactions will ensure that model substitution does not count as price change.
- 75 Reference to the universe to be represented by the CPI and to alternative measures of price change should lead to improved index construction and sample designs. They should also indicate empirical tests for actual CPIs and for assessing the potential bias in these. For the static variety universe there is a straightforward choice of measure of price change between the unit value index and the Laspeyres (or similar) index. Alternative measures of price change based on hedonic regression offer the possibility of further stratification of price differences according selected model characteristics but there are no obvious merits in this. It is yet to be established whether this technique can contribute to CPI construction as a purely statistical device for overcoming operational difficulties. The relevance of conceptual arguments usually given in support of such methods has not been demonstrated.
- 76 It is the introduction of new models and discontinuation of old varieties where the dynamics of the actual universe of transactions raises particular problems for the identification of price change as distinct from quality change. It is here where the pure price framework must be developed with actual practice in mind. It must provide criteria for observation that will distinguish price change from quality change. Any definition of quality change must be reconciled with the various factors that have been associated with price difference in the static variety universe. The resolution of these problems may require new or old conceptual constructs. The argument here has been for a minimum concept approach. Many of the concepts used to underpin existing practice are unwarranted and unsatisfactory

References.]

Anderson, Elizabeth (1997) (See pages 94 and 95) "Incommensurability, Incomparability, and Practical Reason", Edited by Ruth Chang, Harvard.

Boskin, Michael J et al. (1998) "Consumer Prices, the Consumer Price Index, and the Cost of Living" Journal of Economic Perspectives Vol 12, No1.

Boylan and O'Gorman (1995), Beyond rhetoric and realism in economics (p146 & 203), Routledge,.

Diewert, W.E.(1999) "THE CONSUMER PRICE INDEX AND INDEX NUMBER PURPOSE" paper presented at the Fifth Meeting of the International Working Group on Price Indices (The Ottawa Group), Reykjavik, Iceland, August 25-27, 1999; Revised: September, 1999.

Ishikawa, Kaoru. (1985), "What is Total Quality Control?", Prentice Hall, 1985.

Kirzner Israel M. (1992) "The Meaning of Market Process, essays in the development of modern Austrian economics". Routledge (92& 96).

McCloskey. Donald N (1994) "Knowledge and persuasion in economics." Cambridge University Press.

Ormerod. Paul (1998) "Butterfly Economics". Faber,

Reyes & Zrama (1998) "The Process of Embodying Distinctions", Cybernetics & Human Knowing Vol5, No.3,

Sellwood. Don.J.(1994) "Constrain macro effects of the use of different methods at the basic level in the project to harmonize consumer price indices in the European Union". Paper presented at the 1st Meeting of the Ottawa Group on Price Indices Ottawa November 1994

Sellwood (1995) "The Harmonization of Quality Adjustment Practices in the European Union". Paper presented at the 2nd Meeting of the Ottawa Group on Price Indices Stockholm November 1995

Sellwood (1997) "In Search of new approaches to the problem of quality adjustment in the CPI." Paper presented at the 3rd Meeting of the Ottawa Group on Price Indices Voorburg April 1997.

Sellwood (1998) "Constraining unwarranted variations in sampling and quality adjustment practices in the construction of CPI elementary aggregates." Paper presented at the 4th Meeting of the Ottawa Group on Price Indices. Washington April 1998

Triplett, J.E. (1999), "Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?" Paper presented at the 5th Meeting of the Ottawa Group on Price Indices, Reykjavik, Iceland, August 25-27, 1999.