

Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?

Jack E. Triplett¹
Brookings Institution

Paper presented at:
The fifth meeting of the International Working Group on Price Indices (The Ottawa Group)
Reykjavik, Iceland, August 25-27, 1999

I. Introduction and Summary

Does a consumer price index (CPI) need an underlying conceptual framework? If so, should it be the theory of the cost-of-living (COL) index?

With respect to the second question, there is no international consensus among statistical agencies. In several countries (the United States, Netherlands, Sweden), COL index theory provides the conceptual framework for the country's consumer price index (CPI). In some other countries, national statistical offices reject the cost-of-living index framework, in some cases explicitly and in others implicitly--examples are the European Harmonized Indexes of Consumer Prices (HICP) and Australia. Rejection of the COL index is also explicit in an advisory committee recommendation in U.K. In still other countries, Canada being an example, official documents are ambivalent or ambiguous about the role of the COL index as a conceptual framework for the CPI. The international guideline for consumer price indexes—the ILO manual (Turvey, 1989)—does not even mention the COL index.²

The question has taken on heightened interest recently. Part of the stimulus comes from the European Union's relatively new Harmonized Indexes of Consumer Prices (HICP), and the associated European Index of Consumer Prices (EICP) and Monetary Union Index of Consumer

¹ Support for this paper was provided by a contract with Statistics Canada. However, the paper does not represent an official position of Statistics Canada nor does it necessarily reflect the views of any member of its staff. This paper was also presented at The Measurement of Inflation Conference, Cardiff University, Cardiff, Wales, August 31- September 1 1999. For comments on an earlier draft, I thank the following: B.K. Atrostic, Ernst R. Berndt, Jorgen Dahlen, Angus Deaton, Louis Marc Ducharme, Charles L. Schultze, and participants at the Cardiff conference.

² Sources for the statements in this paragraph: U.S.—U.S. Department of Labor (1997); Netherlands—Balk (1994); Sweden—Dalén (1999); Australia—Australian Bureau of Statistics (1997); Canada—Statistics Canada (1995).

Prices (MUICP), which are built up from country HICP indexes. For the HICPs, Eurostat has rejected the COL framework, in favor of an index of “inflation” in “the prices of goods and services available for purchase...for the purposes of directly satisfying consumer needs” (Eurostat, 1999; also, Hill, 1997).

Another part of the stimulus for the current international discussion is an event in North America—the report of The Advisory Commission to Study the Consumer Price Index, known as the Boskin Commission. The Boskin Commission recommended that “The BLS [the U.S. Bureau of Labor Statistics] should establish a cost-of-living index as its objective in measuring consumer prices” (Boskin et al, 1996, page iii).

In the United States, the Boskin Commission recommendation had the effect of confirming the position that the Bureau of Labor Statistics adopted in the 1970s (after considerable debate); however, the Commission’s recommendation was not without controversy in the U.S., even in the 1990’s. Outside the United States, the report probably had the opposite effect, because the Commission used the COL framework to criticize the U.S. CPI, and to justify its estimate of measurement bias. Some statistical agencies distanced themselves from the COL concept, which permitted them to contend that the Commission’s bias estimates pertained to a concept that did not describe their CPI’s.³

The international discussion over the use of the COL index framework for a CPI has a large number of strands which are difficult to partition into tidy headings. This paper is arranged around several major themes, which are summarized immediately below. Each of these themes is developed at greater length in a subsequent numbered section. Each numbered section is written, so far as possible, as a self-contained essay, so the sections can be read independently of the order in which they appear in the paper.

(1) In the current debate, as the one in North America 30 years ago, there is sometimes confusion about what the theory of the COL index says and what it does not say. In section II, I define COL index theory as building the concept of consumption into reasoning about CPI issues, and discuss the content of COL index theory.

(2) In section II, I also ask: What is the alternative to the COL framework? The “not COL” alternative lacks a clear measurement concept, partly because it has been defined negatively, rather than positively.

(3) Measuring owner-occupied housing inevitably enters the debate about the conceptual framework for the CPI. No method for measuring owner-occupied housing is

³ In this, the statistical agencies may have also distanced themselves from their analytical users: Most reviews of CPI’s originating outside statistical agencies (in central banks, for example) have accepted the COL index framework, even if the relevant country statistical agency did not. Examples are Oulton (1995) and Cunningham (1996) for the U.K., Shiratsuka (1998) for Japan, and Hoffman (1998) for Germany.

without serious empirical problems. Section III distinguishes COL index (flow of services) and alternative approaches to owner-occupied housing.

(4) It is sometimes thought that the COL index implies including variables (such as environmental amenities, for example) that are never included in practical CPIs. I refer to this as the “domain” of consumption, or the domain of the standard of living (section IV). There is nothing inherent in the theory of the COL index that necessarily implies a broader domain than the domains that typically appear in CPI indexes. For this reason, the domain is not really an issue for the choice of conceptual framework (section IV).

(5) In section V, I discuss some cases where the COL index and the usual Laspeyres index formulation seemingly suggest different approaches. The theory of the COL “subindex,” presented in section IV, reconciles the two.

(6) Another theme concerns the old relationship between the design of an index and its intended use. The current debate mirrors fairly well what appears in some of the price index literature: In the conventional view, the COL index is linked with escalation uses, and by inference the COL index may not appropriate for other uses, such as measuring “inflation.” In this case, I contend that the price index literature itself is faulty--the COL index does not necessarily provide the theoretical escalator for income payments, but it does provide one appropriate, welfare-oriented measure of inflation (section VI).

(7) In section VII, I discuss the question of forming elementary aggregates, the micro price indexes, such as a price index for bananas, which are aggregated to make the overall CPI. Even though COL theory does not resolve all the questions, it does offer principles for the choice between arithmetic and geometric means. I also show that Eurostat’s rules for computing elementary aggregates in the HICP’s can better be rationalized with COL theory than appeal to the simple Laspeyres concept.

(8) Section VIII responds to some points recently raised by Ralph Turvey (1999).

(9) In the debate on the COL index framework, the rhetoric of the debate dominates in some ways its substance. Words often have associations that go beyond their strict meanings. The choice of words in the debate has influenced the acceptability of ideas, and also sometimes obscured the ideas themselves. For this reason, I turn to the rhetoric of the COL index debate in section II of the paper, before discussing any of the substantive issues.

II. The Cost-of-living Index: Its Definition, Its Rhetoric and Its Content

A. The definition

The cost-of-living index is a price index that measures the change in consumption costs required to maintain a constant standard of living. The index may be unconditional, including

*costs of all variables that affect the standard of living, or it may be conditional on some variables that are held constant, or assumed constant, for the construction of the index.*⁴

The theory of the cost-of-living index originated in the 1920s with Konus (1939). It seems to have been invented independently by others (for example, Allen, 1933).

COL index theory, as developed by Konus, showed that the Laspeyres index number then (and now) in common use has what is now known as substitution bias. Substitution bias has become well known, even outside economics. Substitution bias in the CPI was discussed extensively in the U.S. Congress and the press after the publication of the Boskin Commission report (Boskin et al, 1996), and for the most part the press discussion reflected a surprisingly sophisticated understanding of this index number problem.

The theory of the cost-of-living index applies to an individual consumer. Empirical applications typically employ aggregated data, and not data for individual consumer units. Actual CPI's are also calculated at the aggregate level: Aggregate weights are applied to component price indexes. In turn, each of the component price indexes measures the average change in prices faced by members of the group for which the CPI is defined.

Constructing a COL index (or a CPI) from aggregated data implies that some average standard of living across a group of consumers is held constant in the index. This is often rationalized by the analytic fiction of the "representative" consumer. The representative consumer might be defined democratically (an equally weighted average across households) or plutocratically (where the average across households is weighted by household expenditures).

Theory suggests that unrealistic assumptions are involved in forming any supposedly "representative" consumer (see section II.C). Although the representative consumer language is convenient shorthand, the aggregate COL index is perhaps better thought of as the average of the COL indexes for each member or household of the index population. For this case, it is natural to think of equal weighting for the aggregate index (a democratic index), even though neither empirical COL index estimates nor actual CPI's are calculated in ways that are consistent with a democratic index.

The theory is worked out for two periods, usually called the base or reference period (normally, but not necessarily, the earlier of the two periods), and the comparison period. Even though the theory is defined on two periods, CPIs are produced for many periods. No index

⁴ The word "minimum" conventionally appears in the definition, but is redundant. Hill (1999) contends that my definition of the COL index differs from that of Pollak (1989), apparently because my definition contains the words "price index." I do not intend a different definition, and do not see that my definition differs from Pollak's, except in choice of words. I employ the words "price index" in my definition to make clear that the COL index is indeed a price index, in part because of Hill's (1997) assertion that it is not.

number theory satisfactorily explains a time series of CPIs, rather than two-period comparisons.⁵

Whether aspects of the representative consumer's environment (that affect the standard of living) should be incorporated into the COL index, or should be held constant, is also a part of COL index theory (see section IV, below). Additionally, COL index theory is erected on the assumption that the consumer's tastes can be held constant over the interval of the comparison; Fisher and Shell (1972) consider the COL index when the representative consumer's tastes change. Some of these COL index conceptual difficulties are discussed in subsequent sections.

B. The rhetoric

In the controversy over the application of COL index theory to the CPI, the rhetorical content sometimes matters as much as the substantive questions. In the definition of the COL index (above), I purposefully used the term "standard of living." One could also say that the COL index holds real consumption constant, which would tie the definition into language employed for national accounts. Other words more often appear in definitions of the COL index. Some of those other words are: "constant utility index" or "an index that measures the cost of remaining on the same indifference curve" or sometimes "constant satisfaction index."

In the language usage of most economists, these terms are all synonyms. Thus, an economist will often say that the cost-of-living index is a price index that holds utility constant, or it is a price index that measures the cost of remaining on the same indifference curve. These expressions are economists' jargon for saying that the standard of living is being held constant. Conversely and consistently, when Blackorby and Russell (1978) consider an index of the standard of living, they mean an index that measures the expenditure necessary to move from one indifference curve to another, prices constant, an index that might also be described as a quantity index of real consumption.

For non-economists, and even for some economists, the words "constant utility" and "remaining on the same indifference curve" have a theoretical or ethereal or perhaps even unrealistic sound about them; they are not words that one would encounter in, say, an ordinary journalistic account of what the CPI is and what it measures. The economists' jargon is sometimes parodied by using the term "constant satisfaction index" (often pronounced with a meaningful vocal inflection). Indeed, the term "constant satisfaction index" is most often used by opponents of the COL index, which suggests its pejorative connotation.

Neither "satisfaction" nor "utility" nor "indifference curve" are objectively observable. None of these concepts has any very understandable meaning outside the world of economic theory, nor any currency in ordinary discussion. In ordinary conversation, the terms "constant utility index" and "constant satisfaction index" *appear* not meaningful and they *appear*

⁵ This statement explicitly applies to the so-called "Divisia index," which is not discussed further in the present paper. See Hulten (1973) for the relation between the COL index and the Divisia index.

esoteric. For the nonspecialist (and for some who are) such terms convey a concept filled with unmanageable problems. If some newspaper article on the CPI referred to a “constant satisfaction” index, it is unlikely to be one that conveys a favorable impression of the idea of a COL index.

The term “standard of living,” on the other hand, is used by economists and non-economists alike. There is no discernable difference in the technical and nontechnical employments of the term “standard of living.” Both economists and non-economists behave as if “standard of living” conveys something that is meaningful, realistic, and concrete. And I have never heard “standard of living” used with any pejorative association implied or intended.

Economists know that the *concept* of the standard of living is just as abstract, just as nonobservable, just as problematic as the concept of utility. The two are, after all, wholly equivalent. All of the conceptual and theoretical problems that have been discussed in the *rhetoric* on the “constant utility” index--lacking objective reality, being non-observable, and so forth--have exact counterparts that apply equally to the concept of the standard of living.

To any researcher who seeks to measure changes in the standard of living, economists will bring forth many problems that limit practical measurement, or--more importantly--raise questions about the interpretation of any practical measure, or perhaps raise questions about whether a practical measure of the change in living standards is achievable (in theory) at all. But those are the economist’s technical problems; they do not adhere to the term “standard of living” when that term is used in ordinary communication. One can readily envision a situation where a nontechnical user might request that an economist (or a statistical agency) measure the change in living standards over time (or across countries), where the economist (or the statistical agency) might begin to explain all the problems that adhere to such a measure--to the astonishment and dismay of the user, who would probably dismiss the technical objections as mere quibbles, of concern only to narrow specialists.

The terms “same indifference curve,” “constant utility” and, especially, “constant satisfaction” have emotive connotations that the term “standard of living” does not have. Whether these emotive connotations are misleading or misplaced is beside the point. Expressing an idea with one set of words (constant standard of living) conveys something that is different from expressing it with another set of words (constant utility or constant satisfaction), even though the objective meaning in economics is the same.

Paying attention to the rhetoric in which a discussion is conducted is an essential part of communication. Why economists should abjure use of terminology, like standard of living, that communicates to non-economists (and indeed to economists), in favor of synonyms that do not communicate, is a topic that is beyond this paper.

However, many of the statistical agency rationales for rejecting the cost-of-living index mention, at some point, the ambiguity or etherealness of the idea of a constant utility price index, with the implication (and sometimes an explicit claim) that the idea of estimating a COL is both unrealistic empirically and ill-defined conceptually. There is merit to the contention that

the COL index is difficult to estimate. I suspect, however, that the same people would not say anything similar about the idea of a *constant standard of living*, or at least they would not say it for popular consumption. The reason is that the one (constant utility) brings forth negative connotations, and the other (standard of living) does not.

As an experiment, consider the following questions, which are addressed to those who say they do not accept a “constant utility” price index as the conceptual basis for a CPI⁶: Would you also be willing to say publicly that your CPI does not, by design, attempt to hold constant the standard of living? Would you say to users that it is inappropriate, conceptually, to compare the change in consumer expenditures with your CPI and to infer whether real consumption has gone up or gone down? Would you say that your CPI, by design, would rise or fall with at least some changes in the standard of living?⁷

Probably most statistical agencies would not be willing to say those things. But most users will not understand that when an agency says it is not, *by design*, trying to approximate a COL index in its CPI, its also saying that the CPI is not intended for questions concerning changes in the standard of living. Moreover, I suspect that if this relation were explained to them, some (at least) of the public opposition to the concept of the COL index would evaporate.

For all of the above reasons, I prefer to say that the cost-of-living index measures the cost of holding the standard of living constant. First, that is technically correct. Second, it communicates an idea to the largest number of people in the plainest possible way. To the extent that economists’ jargon causes communications problems, I opt for other language that communicates more effectively. And third, it reduces the level of rhetoric that has surrounded this debate. Reduction in the rhetoric and the emotive tones in which these issues have been discussed recently among the world statistical community would be salutary.

C. The content

The substantive content of the theory of the COL index can be stated in a simple axiom: *The economic concept of consumption drives reasoning about consumer price index number issues.* The unique intellectual advance provided by Konus (1939) was building the concept of consumption into reasoning about consumer price indexes. Previously, reasoning about index numbers revolved around a small set of more or less mathematical properties of different index number formulas (see, for example, the exchanges between Fisher, 1921 and Walsh, 1921).

Although substitution bias is the best known part of the content of COL index theory, the substitution bias question is only one issue for which COL theory can be employed in reasoning about CPIs. Substitution bias is not so much the content of COL theory as a concrete application of the theory to a particular (and in a sense, rather narrow) problem—determining

⁶ Here, I am debating a bit, but the purpose is to highlight the importance of language.

⁷ The key words are “by design.” Obviously, one lacks the ability to control exactly for changes in the standard of living. That is also precisely equivalent to saying that one can not estimate exactly a COL index.

the aggregate CPI index number formula. The substitution bias question is an important one in price indexes. But it is not the only important topic in constructing price index numbers, and it is probably not the most important one.

Constructing a CPI is not just a matter choosing a formula that combines the detailed component indexes—price indexes for coats and carrots and computers and cars, indexes that are sometimes called “elementary aggregates.” Hundreds and perhaps thousands of decisions must be made in measuring those detailed component indexes. Those decisions are not solely statistical or sampling or collection and processing decisions. Many of them involve economic questions—they are “what do we want to measure?” questions.

Applying the theory of the COL index to the CPI means that those “What do we want to measure?” decisions are guided by a consistent, overall decision-making framework, which is the economic theory of consumption. Use of the theory of consumption in constructing the CPI means that constructing price indexes is an exercise in what economists call “applied micro theory,” comparable to the use of consumer theory for estimating consumer demand, or for the analysis of regulations, or for tax policy analysis, and so forth. Use of the economic theory of consumption for the CPI raises few additional questions beyond those that arise in the use of consumer theory for other purposes for which it is routinely employed by economists.

Any use of the theory of consumption, including its application to the CPI, raises some real issues, which are well known to economists. One set of questions concerns whether the theory is realistic—do consumers behave as consumer theory suggests? Do they actually respond to relative price changes by substituting among commodities? A second set of questions, raised most often by consumer demand theorists than by others, concerns aggregation: Even if individual consumers do behave as the theory suggests, does this theory adequately describe aggregate consumer expenditures, and therefore the aggregate cost-of-living index?

1. The aggregate level. The literature on the second set of questions—the interpretation of aggregate consumer expenditures—emanates mainly from consumer demand theorists. The topic is known as aggregation theory (actually, one branch of aggregation theory). The theoretical question is whether it is possible to aggregate individual-level consumer demands for commodities to obtain observed aggregate consumer expenditures, or, equivalently, whether one can estimate consumer demand behavior using aggregate consumer expenditure data. For the COL index, the question is whether one can calculate an aggregate index directly from aggregated data.

Theory says that the conditions for aggregation of individual demands (that is, for aggregation over consumers) are very restrictive, and not realistic. This means that one cannot necessarily interpret aggregate consumption data as if the aggregate were generated by the behavior of individual consuming units. In particular, changes in aggregate consumption shares by commodity may be generated by changes in the income distribution, and not by changes in relative prices alone. Angus Deaton (1998, pages 37-38), one of the most prominent contemporary analysts of consumer demand, has recently written:

“That the Bureau of Labor Statistics should establish a cost-of-living index as its objective in measuring consumer prices, taken by them [the Boskin Commission] as essentially obvious, is a contentious proposition that requires serious argument. In fact, it is unclear that a quality-corrected cost-of-living index *in a world with many heterogeneous agents* is an operational concept” (emphasis supplied).

Yet, one must always balance any theoretical shortcomings with the theory’s usefulness. Someone once remarked that the major substantive content of consumer demand theory is that the demand curves slope down—consumers demand more of a commodity when its price falls.⁸ Empirically, the evidence is overwhelming that demand curves do slope down. An extensive review of studies of commodity demand was conducted by the Bureau of Labor Statistics (United States Department of Labor, Bureau of Labor Statistics, 1997).

If demand curves slope down, the equivalent proposition for the cost-of-living index is that consumers substitute in response to relative price changes. At the level of the roughly 200-600 commodities that receive consumption weights in the typical CPI,⁹ the empirical evidence is overwhelming that the commodities that have the fastest growth rates are those whose relative prices fall. At this level of detail (that is, roughly 200-600 commodities), commodity substitution exists, and the substitutions observed in the aggregate data are consistent in direction with the predictions of the theory. Furthermore, estimates of substitution bias in consumer price indexes are remarkably similar, no matter what estimation methods are used, what time period is explored, or what countries’ data are being examined.

One could debate whether the commodity substitution that is in fact observed at the aggregate level is substitution that holds a consumer’s standard of living constant. Indeed, this is precisely what is debated in the consumer demand literature. Even here, there is at least some empirical evidence in favor of the aggregate interpretation of the theory. Manser and McDonald (1988) tested whether the conditions were met for interpreting aggregate movements in consumer demand as if they were generated by a representative consumer, and could not reject this hypothesis—their data contained around 100 commodities.

⁸ That has also been described as a theory that predicts little that is not obvious. But its implications are apparently not so obvious to non-economists. To take one example, U.S. energy policy in the 1970's was predicated on the presumption that demand for gasoline was not responsive to its price. Politicians and policy makers are often subsequently surprised when price effects actually influence consumers’ behaviors in ways not considered when the policies were adopted. Some entertaining and instructive examples appear in Krugman (1998).

⁹ The US CPI identifies approximately 200 commodities for which component price indexes are calculated and weights are assigned, the French CPI approximately 600. The Australian CPI has something on the order of 1,500 commodities.

The commodity substitution in COL index theory concerns the level at which weights are held fixed in the CPI. At that level, 200-600 commodities, aggregate consumption behavior conforms to the theory's predictions. Accordingly, the commodity substitution portion of COL index theory is consistent with empirical evidence, for the limited purpose of analyzing fixed-weight indexes with 200 or so commodities.

However, commodity substitution is not that large a source of bias in consumer price indexes. The Boskin Commission, which estimated bias in the US CPI at 1.1 percentage points per year, obtained only 0.1 of that estimate from aggregate commodity substitution. This is consistent with what has always been my own view: Although the commodity substitution bias is present in fixed-weight indexes, and it goes in the direction that COL index theory predicts, the substitution bias is small enough to be neglected, for practical purposes, when an agency computes a chained Laspeyres formula (with up-to-date weights). The amount of intellectual attention given to aggregation formulas for price indexes has been huge and its results have been instructive, but the effort has been disproportionate to any realistic estimate of the sources of potential errors in consumer price indexes.

A second qualification concerns the level of detail. Even though the empirical evidence at this level of detail is overwhelmingly in support of the predictions of the theory, any theory of consumer behavior is not pushed very hard when tested at the level of 200 or so commodities. At this level, these are already aggregations of commodities, and not the detailed commodities on which consumers make choices (and of course, the data are also aggregated over consumers). The theory proves empirically valid at this aggregated level of detail because one cannot see in the data the behavior that might be inconsistent with the theory, so the power of any test that one could conduct on the theory's predictions will be low. Too little of consumers' real behaviors are apparent in such aggregated data.

In summary, the aggregation objections to COL index theory, though in theory important, seem secondary considerations empirically, with respect to forming an aggregate index from component indexes that are already aggregated (over commodities to an extent, and also over consumers). This is partly because substitution bias at this level of aggregation is empirically fairly small, and partly because, as noted, much aggregation is already in the data before they are aggregated further with Laspeyres index weights.

2. The micro level. The really important and relevant issues--quantitatively--in measuring CPIs concern how one measures the detailed indexes, and not those issues that concern the aggregation formula. I believe the theory of the cost-of-living index is much more pertinent and relevant for making decisions about measuring the individual commodity price indexes that make up a CPI—deciding how to handle quality changes, new products, changes in transaction terms, multipart pricing, and so forth. Pollak (1998, page 75) remarked: “In the case of the CPI, economists and others often appeal to the theory of the cost-of-living index for a principled resolution of technical issues.”

Although COL index theory is useful--indeed I would say irreplaceable--at the micro level, it is at the detailed level where the theory of the cost-of-living index has its greatest

ambiguities and greatest need for development. This takes us back to the first question from the beginning of this section: Do individual consumers (or, more relevantly, households) behave as the theory suggests? Do they in fact optimize? Do not micro-studies of individual behavior provide evidence that is inconsistent with consumer theory?

No economist can be unaware of criticisms of consumer theory. Indeed, anomalies in consumer behavior, anomalies that are not consistent with theory of consumption, are well known. Thaler (1991) is prominently associated with this strand of research. My impression is that the most telling anomalies arise in choices involving risk, uncertainty, and the valuation of options; perhaps people process information about risk incorrectly, or perhaps the theory of how they should behave in risky situations is wrong. Behavior toward risk affects consumer choice outside of such obviously risky activities as gambling and insurance (buying a used car, for example), but anomalous behavior toward risk does not seem central to the application of COL index theory to the CPI.

Looking at anomalies can help focus effort on improving and extending the theory to make it useful for dealing with behaviors that the simple theory does not confront satisfactorily. It was certainly true that the specification of consumer behavior that Konus (1939) introduced into the index number literature was a very simple specification of the theory of consumption. It was advanced for its day, but it is not advanced now. The simple theory does not say much, for example, about how the consumer gets information to make choices. Information has been incorporated into economic theory since at least Stigler (1961); attempts have been made to bring the acquisition of information (consumer search to find the lowest prices, for example) into the theory of the COL index, but it is not easy.¹⁰

The analysis of quality change is another topic that is not handled well by the simple theory of consumer behavior that was known in Konus' day. A theory of consumer behavior toward the characteristics of goods, rather than the goods themselves, is extremely complicated. An early example is Court (1941a, 1941b--not the same Court as the man who invented hedonic methods, incidentally). Other contributions include Ironmonger (1972) and Lancaster (1971). A more recent contribution (which is directly relevant for analyzing quality change in price indexes) is Berry, Levinson and Pakes (1995). Pakes (1997) presents a nontechnical summary of some of this work, discusses its application to the CPI, and unresolved problems.

Even so, the application of characteristic-space consumer theory to COL indexes is inadequately worked out. My own work (Triplett, 1983, 1987) makes use of simplifying assumptions that make it a special case. Pollak (1989) contends that there are many special cases, and that it will be difficult to find a general case that is tractable for empirical work.

Aggregation issues, discussed in the previous section, arise here also. There are many consumers, and they evaluate quality change differently. Yet, in the CPI it is necessary to make

¹⁰ See the discussion in Pollak (1998). Baye (1985) proposed adding search behavior to COL index theory. A discussion of the consumer "shopping" problem in designing CPI basic components is in Triplett (1998).

one quality adjustment when quality change is encountered, essentially because only one price index for, say, automobiles is calculated, and not specific automobile price indexes for each consumer or each consumer type. This means that quality change is handled in practice as if there is a representative consumer, whose valuations of quality change are incorporated into the CPI, and the representative consumer is a very questionable concept.

It would be wrong to say that modern theory resolves all the problems with which we must deal, but it is also quite wrong to say that there is nothing in the modern theory that confronts the problem of quality change in price indexes, as I interpret Turvey (1999) as asserting.¹¹ Turvey (1999) and I might agree that we need an improved theory, but he overstates his case, by a large margin, when he implies that the theory contains nothing that is relevant.

It is certainly true that, in many other ways, the theory has not advanced enough, and in many ways COL index theory does not confront, fully, some of the problems for which CPI compilers most need guidance. Pollak (1998, pages 69-70) wrote: "To deal with current concerns about the CPI within the framework of economic theory requires developing the theory of the cost-of-living index under more general assumptions than have thus far been standard."

There is also the perennial question of the realism of the theory's assumptions. Economists know that the theory of the cost-of-living index, like any theory, rests on assumptions that are often (or, more accurately, often seem) unrealistic. Realism objections are quite old. For example, when Friedman (1938) raised the substitution bias issue with respect to the fixed-weight indexes of output and productivity constructed by Copeland and Martin (1938), the authors dismissed this as more or less an academic curiosity. Theory always abstracts. Krugman (1998, p. 19) has noted that "Economic theory is...a menagerie of thought experiments--parables, if you will--that are intended to capture the logic of economic processes in a simplified way." That is often lost sight of in criticisms that confuse descriptive usefulness with analytic usefulness. The issue is the analytic usefulness of consumer theory, and not whether one can find some consumer behavior that the theory (as presently developed) cannot explain.

Moreover, the debate on COL index theory often reads as if the statistical agency contributors thought that the only practical application of consumption theory was to the COL index. It is important to recognize that economists use the theory of consumer behavior in all kinds of practical applications, demand analysis, tax analysis and other issues of public policy analysis. It would be hard to describe the life of an economist working in those areas who was told--as Turvey (1999) instructs CPI compilers--to ignore the theory of consumer behavior. Those economists also know of the shortcomings of the theory for applied problems. But it is

¹¹ Turvey (1999) says that quality change is among five problems "where the fruitfulness of cost-of-living theory is not apparent" because, he says, "the theory that we have now... tells us only that demand curves slope downwards because of substitution effects, which we knew anyway."

useful, extremely so, and so it is used. The same thing can be said of the theory of the COL index: It is useful, and it is used (it is even used sometimes when the agency officially says it is not computing a COL index). And why not? Fully adequate or not, it is the only thing we have.

D. What is the alternative to the COL index?

At the beginning of section II.C, I stated that in my view “COL index theory” means mainly that the economic concept of consumption drives reasoning about consumer price index issues. I understand that mine is a broader definition than the usual understanding of COL index theory, and that some might debate it (this conference is one place for that).

One problem with discussing alternatives is that the opponents of COL index theory have not described cogently the conceptual framework they advocate as an alternative. In much of the recent discussion, the protagonists more or less construct an argument against the COL index, implicitly on the logic that the alternative to the COL index then wins, by default.

To resolve this debate, one needs to understand whether the alternative conceptual frameworks to the COL index are less ambiguous or less problematic, or are more consistent with empirical knowledge or better-defined conceptually. Rather than specifying alternative concepts to the COL index, the detractors from the COL index take, mostly, “not COL index” positions. Perhaps the flaws of the COL conceptual framework are less serious than the flaws of the “not COL” conceptual framework, if only that “not COL” framework were carefully written down.

An internal Office for National Statistics (U.K.) paper by Sue Holloway (1999) makes an interesting contribution to the discussion. She sets the COL index at one end of a continuum. In her interpretation, the essence of the COL index is suggested by: “Anything which relates to consumer preferences [and] behavior . . . moves a pure price index toward a cost of living index.” Holloway’s definition is consistent with the definition of COL index theory I have given above.

She puts at the other end of the continuum “not COL” (or alternatively, “pure price index”¹²). With this continuum idea, she deduces that “not COL” at the other end of the continuum must be a price index in which consumer preferences and behavior *have no place* in the measurement. For example, consumer expenditures are the end product of consumer behavior and consumer preferences. Thus, in Holloway’s definition of the continuum, even

¹² The name “pure price index” has been around for a very long time. It was once employed in the US. The BLS, for example, used this term extensively in response to the Stigler Committee’s recommendations of 1961. But no one has ever written down very clearly what a pure price index was supposed to mean, except when it was identified with the Laspeyres weighting system; I never understood exactly what the old BLS meant by it (except, very clearly, it meant “not COL index”!).

using consumer expenditure weights introduces an element of consumer behavior, and therefore an index that uses consumption weights is not at the far end of the “not COL” continuum.

I suspect that most of the statistical agencies that advocate “not COL” for their CPIs will not accept Holloway’s description of the continuum. For most of them “not COL” probably means a fixed-basket, Laspeyres index. Others may say they favor a “pure price index” (but this has never, to my knowledge, been defined). However, Holloway’s paper suggests that people of that persuasion have already let the camel’s nose into the tent. Once you accept consumer behavior with respect to the weights as appropriate for the CPI, why stop there? Why not admit some other form of consumer behavior into reasoning about the index? What indeed is the stopping rule on Holloway’s continuum if one does not want to be at either the COL or the “no consumer behavior at all” poles?

Asking what is the appropriate stopping rule along Holloway’s continuum seems a better way to think about the problem than simply to say “not COL.” I suspect, as I noted earlier, that most statistical agencies who advocate “not COL” will respond to Holloway by saying: “We will admit consumer behavior up to the point of deriving weights for the Laspeyres index, but not one step further.” Whether that is a tenable intellectual or analytical position is something to be debated. Putting the question that way would engage a debate with substance.

In summary, whether Holloway’s continuum ultimately proves useful in facilitating the debate on the appropriate concept for the CPI depends on whether other people accept her definition of the “not COL” pole represented by a pure price index. A useful dialog on the question of the economic concept underlying the CPI requires considering alternative concepts, not just finding weaknesses in one. Before one rejects the COL index as the economic concept for the CPI, there must be some assurance that we understand the implications of the alternative, “not COL” index.

In section IV, I consider the COL index end of Holloway’s continuum.

III. Owner Occupied Housing and the CPI

Beyond the rhetoric, the issue that drives much statistical agency uneasiness over the concept of the COL is the treatment of owner-occupied housing. COL index theory suggests pricing the flow of monthly housing services—the monthly cost of living in the house.

Statistical agencies in many countries have deemed the empirical problems with the flow of services approach so distasteful that it has been rejected. It is perhaps an oversimplification to say that the unpalatable empirical implications of the theory of the cost-of-living index for owner-occupied housing has induced rejection of the COL index framework, but there is nevertheless considerable truth in the oversimplification.

It is probably worth restating at the outset the logic of the flow of services approach to owner occupied housing. The concept of consumption implies that the standard of living

depends on the consumption of housing services, and not on the purchase of houses. I think that is not controversial, even among detractors of COL index theory.

For rental housing, the price and the quantity are quite clear. The monthly quantity is the use of an apartment (flat) or house of a particular specification (location, size, amenities, and so forth).¹³ The price is the monthly rent. Measuring rent change is certainly not free from empirical problems—for example, Randolph (1988) shows that price indexes for rental housing have a substantial downward bias because of unobserved aging effects.

In the case of owner-occupied housing, the quantity is in principle the same as for the rental housing case, the use of the dwelling for a month. But there is no transaction between the owner of an owner-occupied house and the tenant. There is no directly observable price, and also no directly observable monthly or annual expenditure weight.

Two empirical methods exist. One is to estimate owner's equivalent rent. The second is to estimate a user cost function for housing (see below).

A. Rental equivalence

In the rental equivalence approach to owner-occupied housing, one estimates the *change* in monthly cost for owner occupied housing by the *change* in monthly rents for housing of similar types that are in fact rented. Three strong objections to the rental equivalence procedure arise.

In the first place, in many countries rent control and publicly subsidized rents are prominent. For this reason, changes in rents may not measure very well monthly dwelling costs for owner-occupied dwellings. This is a serious problem.

Second, it is sometimes said that owner-occupied and rental housing are different markets and their prices do not move together. As a general proposition I would not place a lot of weight on this objection. In the absence of rent controls and similar regulatory distortions or nonmarket determined rents, evidence suggests that within a particular urban area rents—though they differ substantially in levels—tend toward similar rates of change, whether single family or multifamily, and for inexpensive and expensive housing. There are indeed differences in levels of rents, but market forces tend to push up all the levels at more or less the same rates, except for short-run shortages or surpluses that might be caused by some rapid shift in the composition of the rental housing market (a sudden influx into an area of managers and professionals, for example, that creates excess demand for luxury housing). The second objection is valid mainly when the first one (rent control and so forth) creates problems.

¹³ So long as only one dwelling is used by a household, this implies that a change in the quantity of housing consumed is identical with a change in the quality of the dwelling unit. It is not really true, as sometimes said, that the quantity of housing consumed is always one unit.

A third objection is that the rental market is “thin.” Usually, this actually means that the rental market is thin for the *exact type* of housing that is owner occupied. If the rental market is thin for the *exact type* of housing that is owner occupied, then the observation that rents in a particular urban area tend to move together also suggests that, thin or not, an index of rents, suitably partitioned or segmented, will provide a useful measure for owner’s equivalent rent. If the objection just means that rental sample sizes are too small, then they should be too small to compute a rent index; nearly every country to my knowledge includes rent in its CPI. In the absence of rent controls and publicly subsidized housing, this objection has less force.

It is sometimes said that pricing owner-occupied housing with the rental equivalence method is an imputation, and that imputations should not appear in a consumer price index. It probably is an imputation, but this contention seems largely rhetorical. For example, in the HICP the “no imputations” rule has been used to exclude “imputing” rent change for owner-occupied housing from rent changes in the rental housing market. But the HICP contains an imputation in the case of insurance, where the pricing concept and the weight is the imputed services of *administering* an insurance policy, and not the service of insuring the policy holder. This imputation is acceptable because it is not called an imputation, but it is an imputation nonetheless. There is no price for an insurance administration service so defined, no market transaction in such services, and no separate way to estimate consumer spending on such a service. In contrast, a rental equivalence price is a market price, it is not an imputed *price*; it is only imputed to a closely related commodity to that for which the price is collected.

Whether or not the price is imputed should not be the question. The question is whether the occupants of owner-occupied housing are affected by inflation in the housing market, and if they are how should we measure inflation for this portion of the CPI population.¹⁴

B. User cost functions

Another way to estimate the cost of owner-occupied housing is to make use of a relationship, developed in capital theory, that relates the price of houses, the cost of providing housing services, and the market rent:

$$(1) c_t = (d + i) P_t - (P_t - P_{t-1})$$

In this equation, c_t is the monthly *cost* of providing housing of specified characteristics, P_t is the price of the house itself, d is the rate of depreciation and i is the appropriate interest rate for housing investment. The term $P_t - P_{t-1}$ is the capital gain (or loss) from holding the property for one period. For simplicity, I have ignored property taxes and maintenance and repair expenses in equation (1) because their effects on cost are obvious, and because they are normally measured directly in CPI’s.

¹⁴ Thus, Stott (1998) reports that the New Zealand advisory committee for the CPI felt that “the credibility of the CPI is enhanced by the use of prices paid in actual market transactions and is reduced by the use of imputed prices and notional transactions,” yet “there was a strong view by a minority that housing costs should be measured using imputed rents....”

The first term on the right-hand side represents the depreciation and interest cost of investment in housing. Note that a capital gain subtracts from the cost of providing rental housing, a relationship that is readily apparent from economic data on rental housing markets. In equilibrium, c_t , the cost of housing, as defined in equation (1), is equal to the competitive market rent, r_t .

The user cost equation offers an attractive option if one believes that the rental housing market is too unrepresentative to provide a good estimate of the cost of owner occupied housing. Unfortunately, existing estimates of user cost for housing are often far more volatile than market rents suggest (the best analysis with respect to the consumer price index is Gillingham, 1983). The basic reason is that capital gains, which in equation (1) reduce the cost of providing housing, are volatile, and when capital gains are high (which will happen when house prices are accelerating), they may be large enough to create negative user cost. CPI users are not likely to accept a cost function that shows negative monthly costs of housing precisely when inflation in house prices is the highest.

However, there is sense in this result, properly understood. Equation (1) measures user *cost*, not rent. In the short run, rents may exceed or fall short of user cost. When the prices of rental housing units are accelerating, user cost may be negative (see equation 1); if so, rent greatly exceeds the cost, and housing is very profitable, which accords with common sense. In the long run, rent will equal the cost of providing housing, but in the short run, housing may be a very profitable or an unprofitable investment.

Another objection to user cost functions comes from central banks, who seldom like to see the price they control (interest rates) included in the CPI. When interest rates are a main instrument of anti-inflation policy, this concern is understandable, even though it remains true that interest is a major part of the cost of consuming the services of durable goods, especially those that are long lived. Moreover, it is quite clear from housing market data that the short term movement of rents is nowhere nearly so volatile as a user cost function (with current mortgage rates) suggests, which validates from a different perspective the dislike of central bankers to see the price they control increasing the CPI.

Turvey (1999) opposes the flow of services approach for owner-occupied housing. His argument contains two surprising errors. “To ignore the actual prices of [durables] in favor of their imputed rental values...would not appear sensible to most economists.” Equation (1) shows that the selling price of a durable good is included in a user cost expression; the actual selling price is hardly ignored. The sentence quoted sets up a false relationship that misconstrues the issue.

Another slip in Turvey’s argument is even more surprising: “Most people will judge it as absurd to regulate Social Security benefits, pensions and taxes according to...the amount consumers are deemed paying themselves for the use of their owner-occupied dwellings....” Here, Turvey confuses the *measure* of inflation with the important issue of which income flow should be adjusted to compensate for inflation.

Owner-occupiers have an implicit flow of income from their owner-occupied housing (this is the treatment in national accounts). Thus, a pensioner's total income includes pension and nonpension income. To adjust pension income by a COL index that includes the cost of owner-occupied housing--ignoring, that is, the fact that pensioners have implicit income from housing--will of course leave those individuals at a higher standard of living, whenever the price of housing is growing faster than other prices. This point was made in my old paper (Triplett, 1983). It is not an argument against measuring inflation by the COL index, or against measuring housing by a flow of services approach; it is, rather, one of the many arguments that can be levied against the use of 100% escalation in income payments. See the cogent discussion of this problem by Griliches (1996).

C. Other alternatives

The ILO manual on consumer price indexes (Turvey, 1989), in an otherwise excellent discussion of the measurement of CPI owner occupied housing, suggests banishing interest rates from the user cost function. Because interest is such a major part of using any long lived durable good, this is not a very good alternative. If interest cost is simply dropped, it underweights owner-occupied housing relative to rental housing (because in equilibrium, rent is equal to cost, so rent includes interest cost--indeed, interest cost is a large proportion of the cost of housing, because houses last so long); it also overweights, relatively, the other components of the user cost function. For example, if the weight of interest cost is reallocated to depreciation, it overweights depreciation in the user cost function (see equation 1).

A "not COL" solution, sometimes suggested, is to include in the CPI only the price change for houses. Equation (1) shows some of the problems with that. For owners, a rise in the price of houses has two effects on the cost of housing. The direct effect raises the cost of housing, because the house price is multiplied by depreciation and interest rates (see the first term on the right-hand side of equation (1)). But the capital gains effect lowers the cost of housing, through the second term in equation (1). Equation (1) rationalizes the widely-observed fact that owners like to see house prices go up, but prospective owners do not (for prospective owners, there is only the direct effect, they do not benefit from the capital gain). Including house prices only in the CPI overstates the cost of housing to owners during a period of rising house prices, and understates it if house prices fall.

D. Conclusions

In summary, then, the urge to reject the COL index formulation for owner-occupied housing is driven by practical considerations. On the one hand, use of owner equivalent rent is thought to be unacceptable to CPI users.¹⁵ On the other hand, putting mortgage interest into the

¹⁵ Perhaps one should distinguish the short-run, in which any change is bound to create some controversy, from the long run impact. Moving to a rental equivalence basis for housing in the U.S. CPI was indeed very controversial when it was first proposed in the 1970's. Once it was put in place, however (in the 1980's), and users became familiar with it, the controversy

CPI, even in the form of a user cost function for housing, is often disliked by central banks who use interest rates as a tool of monetary policy.

There is no fully satisfactory solution to measuring the price of owner-occupied housing. The problem arises because housing is a large proportion of consumption, so it cannot be ignored in the CPI, and because there is no transaction that can be observed for the monthly cost of housing. The problem of owner-occupied housing is not caused by the COL index concept. It cannot be solved by ignoring the concept of the COL index.

IV. The Domain of the COL and of a CPI

It is sometimes said that a COL index would include variables, such as nonmarket goods and services and environmental amenities, that are included in no country's CPI. The implication is that moving to a COL index concept would force adding to the CPI components whose measurement is contentious and possibly not germane to, e.g., measuring consumer inflationary experience. But it is misleading to think that the CPI domain is necessarily inconsistent with what is specified by COL index theory.

I define the "domain" of the CPI as the list of the goods and services that are contained in a typical country's CPI. The n goods and services in the domain can be represented as:

$$(2): \text{domain of the CPI} = [x_1, x_2, \dots, x_n]$$

Note that this list of n commodities is generally much larger than the number of component indexes (sometimes called "elementary aggregates") in any country's CPI, because agencies always base their CPI's on samples of the n goods and services in the domain.

As defined in section II, the COL index holds constant the reference period's standard of living. The domain of the COL index depends therefore on the list of commodities that determine the standard of living.

There cannot be much doubt that the list of commodities (perhaps I should say the list of variables) that determine the standard of living is broader than the list included in the domain of any country's CPI. For example, there are government provided goods and services. It is also hard to see how the standard of living can be defined independently of the level of pollution, say, or of crime and safety. The Boskin Commission had a long list of "broader considerations on the quality of life" which the Commission included in its notion of what the CPI should measure.¹⁶

completely evaporated. Where any country has moved to rental equivalence, and there are no major problems with rent control or other government operations on the rental housing market, little controversy seems to remain after the change has been put into effect.

¹⁶ "Overall, we find that the presumed negatives (pollution, crime, suicide, divorce), the worsening of which may have increasingly detracted from the quality of life at one time, have reached a plateau...but seem to us to have been more than offset by increased quality and

Accordingly, the domain of a *comprehensive* COL index might be written:

$$(3) \text{ Domain of the COL index} = [x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_m]$$

This is just a formal statement that there are some variables, $[y]$, that would be included in a comprehensive measure of a COL index but that are not normally included in a CPI. If one tried to include *all* of the determinants of the standard of living in an estimate of a COL index, I think it would generally be accepted that this is a measurement project that lacks, even remote, feasibility.

There is no reason, however, to be so negative. Pollak's concept of the COL "subindex" (Pollak, 1989) provides the way to think about the COL index domain and how the domain of the COL index connects to a CPI.

Pollak began by noting that CPI's always include published indexes such as, for example, food or clothing. He called these "subindexes" of the cost-of-living index. Part of his paper is concerned with deriving theoretical statements about the construction of familiar subindexes of this kind.

Pollak (1989) went on to note that the entire CPI, as every country publishes it, is also a subindex of the cost-of-living index. That is evident from the definitions of the domains of the CPI and of the COL index, given above. Suppose that the $[x]$ variables in equations (2) and (3) include all the market purchased consumption goods and services, so that the domain of the CPI includes all market-purchased consumption goods and services. Then, following Pollak, we can say that such a CPI is an approximation to the COL *subindex* on market-purchased goods and services. Indeed, similar language has been used to describe the domain of the COL index that provides a conceptual framework for the U.S. CPI.

The idea of a subindex does not solve all the problems, because the theoretical conditions for producing subindexes are quite stringent: The expenditures on the goods and services that are included in the subindex must depend only on their prices and not on the level of commodities that are outside the subindex. For example, a subindex of the COL index that is restricted to market-purchased goods and services requires that trade-offs between, say, expenditures on door locks and burglar alarms compared with, say, garden parties, do not depend on the level of police protection (which is outside the subindex).

variety of goods, services, and choice of outlets along dimensions that are partly but not entirely captured by our measures of bias, but most importantly by the major increase in longevity which perhaps swamps everything else. Accordingly, our estimate of the current bias in the CPI, is, if anything, probably understated" (Boskin et al, 1996, pg. 76; the Commission's recommendation #10 made this view more explicit). This very broad concept of consumption made professionals in statistical agencies (and indeed, many economists outside them) uncomfortable.

The subindex idea makes it clearer what we are doing when we compute a CPI that depends only on market purchased goods and services. Surely no one believes that the displacement of expenditures on garden parties when police protection declines in favor of expenditures on locks and burglar alarms leaves the standard of living unchanged. The CPI on market purchased goods and services must maintain such an assumption. The theory of the COL subindex makes the limitations of an actual CPI index clearer by making its assumptions more transparent.

The conclusion: There is nothing inherently inconsistent in saying that one wants to adopt a COL index concept and also wants to restrict the CPI to market-purchased goods and services. For this case, one says that the CPI is an approximation to the COL subindex on market-purchased goods and services. Or, to put it another way, the CPI is a price index for that part of the standard of living that arises from market-purchased goods and services.

Saying it this way, defining the domain of the COL subindex this way, makes more clear some of the uncomfortable boundary issues that arise in a COL subindex defined on market-purchased goods and services. The alternative, however, is setting an arbitrary boundary for the CPI. An arbitrary boundary leaves the same boundary issues equally unresolved, but it covers them up. An arbitrary boundary tends to hide the measurement shortcomings inherent in defining the living standard to include only market-purchased goods and services. Moreover, such a restricted definition of the living standard has implications for the use and interpretation of the CPI.

Another valuable use of the subindex idea is closely related. Suppose we were to agree that, ultimately, we want a COL on all of the variables, x and y . But it will always be the case that measuring the prices of some of the y variables is not feasible, or that the methods to measure them are too “soft” to be defended except as a research exercise. In other cases, the variables themselves may be too controversial for widespread public acceptance. If for any of these feasibility or acceptability reasons we decide against including, say, measures of y_r and y_s , this does not preclude doing a COL subindex on the others. That is, we can compute a COL subindex where the domain is:

$$(4) \text{ domain COL subindex} = [x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_m - (y_r, y_s)]$$

Again, use of this subindex idea does not solve any problem that arises from excluding y_r and y_s . The subindex idea just gives a consistent way to characterize what is done and to consider, properly, the limitations of the subindex that arises from excluding costs of variables y_r and y_s . It is tremendously useful to have such a conceptual way to both describe the measurement and to characterize its strengths and limitations.

In summary, adopting the COL concept as the framework for the CPI does not necessarily require that a statistical agency estimate costs for all variables that might conceivably be put into the standard of living. The Boskin Commission Report has been a bit misinterpreted in this respect, and probably its failure to discuss the subindex idea frightened some users and some statistical agencies away from the idea of the COL index. Use of Pollak’s

concept of the subindex can permit introducing feasibility and user acceptability as considerations in defining the domain of the COL subindex that one wants to approximate in the CPI, or indeed in a research price index.

V. Some Concrete Issues

Does the COL index framework give different answers in practical situations from the “not COL” framework? One great difficulty in answering this question is the lack of a concrete and explicit “not COL” framework. If one had alternative conceptual frameworks which gave clear alternative prescriptions for the CPI, then one might assess the alternatives by whether one of the two indexes were better for some purpose.

A. Questions already considered

It is sometimes asserted that if a county’s CPI were shifted over to a COL index, it would imply a broader measure, with added variables that go beyond what the CPI can be expected to measure. I have contended, in section IV, that this contention is a misunderstanding.

The international statistical agency discussion contains other examples that reflect misunderstandings of what COL index theory says. It might be useful to catalog such errors, essentially to eliminate confusions and to clarify the real issues, but this is not the place for that.

It is sometimes asserted that shifting to a COL index framework would imply a flow-of-services approach to owner-occupied housing, and that a flow-of-services approach is unacceptable to users, or (sometimes) to economic policy uses of the CPI, such as its use as an inflation index for macroeconomic or monetary policy. Although there is basis for the first part of this assertion, as discussed in section III, it is not clear how a “not COL” index would measure owner-occupied housing (at present, the HICP indexes omit owner-occupied housing because no satisfactory method for measuring it has been found). It is not clear that an alternative to the flow-of-services method for estimating the cost of owner-occupied housing is more satisfactory. And particularly, it is probably not true that a “not COL” approach to owner-occupied housing is necessarily preferred by economists in central banks, treasury ministries and other macroeconomic policy-making authorities. Again, this issue is discussed in section III.

B. An example

One interesting case where COL index and Laspeyres index formulations seem, under some circumstances, to give different answers involves home heating and cooling.¹⁷ Suppose

¹⁷ The issue discussed in this section is a very old one, and will be familiar to almost anyone who has worked on price indexes. I am not sure, however, just where this matter is written down, if anywhere. It was brought back to my attention in a recent conversation with Angus Deaton, Robert Pollak, and Charles Schultze.

an unusually cold winter (or an unusually hot summer). Suppose there is no unusual increase in the price of home heating fuel (or of electricity for air conditioning). The unusual weather increases the cost of heating (or cooling) one's home because the quantity consumed increased, not because prices rise. How should we think about what would appear, to most homeowners, as an increase in their cost of living?

In the Laspeyres' index formulation, one simply says: Only price change matters, quantity changes do not matter, the weights for home heating fuel (electricity) are held constant at the base period levels, and no inflation has taken place. If the objective is to produce an inflation index for monetary policy, one might put this even stronger: One would not want to mistake the effects of a severe winter for underlying inflationary forces.

Yet, if it is not counted as consumer inflation, then the increased consumption of home heating oil must be a rise in the standard of living (it would show up in this way in national accounts, unless offset by declines in consumption elsewhere). This seems questionable, which suggests that the Laspeyres view is not totally satisfactory.

On the COL index view, the situation is more complicated, because the COL view forces one to ask: Exactly what is being held constant? No one wants home heating fuel for its own sake, so what matters to the dweller is the cost of keeping one's home to a comfortable temperature in the winter (which is what the heating fuel was purchased to provide). On this concept of consumption, we might specify that the COL index should measure the cost of holding constant the winter-time temperature in the living quarters of the house; the COL index will then rise with cold winters, and fall with unusually mild ones.

There is another way to look at this. One might want to produce a COL subindex *conditional* on the base period's weather experience (the concept of the conditional COL subindex is attributable to Pollak, 1989). In this case, the unusually cold winter does not affect the *conditional* COL subindex that holds the environment constant. Even though the unusual weather conditions raise, in some sense, the cost of living, they do not raise the COL subindex that we want to measure, which is a conditional COL index. The COL subindex that holds the environment constant is probably the COL concept that is most useful for an anti-inflation policy (see section VI).

This example shows that the purpose for which one wants a price measure is essential in specifying the nature of a COL index. The COL index framework is a very flexible one. It can be applied in different ways, depending on the purpose.

The real value of the COL index framework is to make us specify more precisely what it is that we want to measure, and to make us state more precisely the question for which we want the CPI as an answer. I am not saying that one could not work out those questions precisely in a non-COL framework. However, having a conceptual framework based on the COL index helps, because asking what conditions "outside" the index are held constant is not a very natural question in the Laspeyres framework. And if the COL index framework helps, why not have it?

VI. The COL Index as an Escalator and as an Inflation Measure

One theme that has emerged in the current discussion of the COL index concerns the relation between the design of an index and its intended use. The CPI is used as an escalator for wages or income payments. The CPI is also a measure of inflation for households. It has frequently been asserted that the COL index is the appropriate measure for escalation, or compensation for inflation, but that the COL index is not appropriate if the objective is measuring inflation.¹⁸

In this section, I will contend that the existing price index literature has it backward--the COL index is not *necessarily* the measure that suits the escalation or compensation problem that is faced in public and private escalation arrangements, and, conversely, the COL index is the concept that one wants for measuring inflation.

A. The COL index as an inflation measure

This is a complicated issue, on which different economists might reasonably have different views. It has become a newly complicated by some confusions, which I will address in subsections 2-4.

1. The COL index as a measure of inflation. A division of opinion exists on whether a consumer inflation index should be identified with a COL index. A paper by Hill has been interpreted as giving a negative answer to this question: “A cost of living...index does not measure the change in the value of a fixed basket of goods and services so that its meaning as a measure of price change is not self evident (Hill, 1997).¹⁹ The European HICP indexes have been designated explicitly as indexes of consumer inflation, and not COL indexes, following the statement in Hill’s paper.

¹⁸ However, Stott (1998) remarks that, whatever the link between purpose and measurement in the price index literature, little agreement emerged from users of the New Zealand CPI on either the purposes that the index should serve, or the type of index that should be constructed.

¹⁹ Hill also writes: “...As an increase in the ‘level of prices’ suggests measuring inflation by the increase in the total monetary value of a specified, fixed set, or ‘basket,’ of goods and services....” And: “As an inflation index [is] a price index that measures the change between two periods in the total value of a fixed basket of goods and services. On the other hand...a COL index measures the change in value between two baskets of goods and services whose quantities...are just sufficient to keep the consumer on the same indifference curve.” Additionally: “Neither objective, having sufficient resources to maintain a constant level of utility [the COL index] as against having sufficient resources to purchase a constant basket of goods and services [Hill’s inflation index] seems intrinsically superior to the other from a theoretical or scientific point of view.”

However, one can make a good case that the COL index is exactly what is wanted as a measure of consumer inflation. The COL index is a welfare-oriented measure, it is the price index that holds constant the standard of living between two periods.

If the “not COL” index deviates from the COL index, it must not hold the standard of living constant, so it must increase when some (though not all) component of the standard of living increases. Why monetary authorities should want to stabilize an inflation measure that incorporates increases in the standard of living (as does the “not COL” index) is not clear. Stabilizing such an inflation measure would then imply a falling standard of living, which would, I believe, correspond to no central banker’s objective.

Thus, if a central bank sets a zero inflation goal, the COL index serves as an appropriate *standard* for determining whether or not the goal has been met.²⁰

There is support from central banks for the use of the COL index as the standard for monetary policy. In the U.S., the central bank publicized shortcomings of the CPI as a COL index. Reviews by central and regional Federal Reserve staffs adopted the COL framework (Wynne and Sigalla, 1993; Lebow, Roberts, and Stockton, 1994; Shapiro and Wilcox, 1996), without in any way suggesting that the COL framework was inappropriate for monetary policy. Reviews of CPI’s carried out by central banks in other countries have generally adopted a COL framework, even when country statistical agencies did not (see references in note 3). A paper prepared by the Bank of New Zealand for that country’s review of its CPI states: “The theoretical roots of the CPI are found in the literature on cost-of-living indexes,” and goes on to conclude that for both inflation measurement and money policy uses of the CPI in New Zealand: “The implied measure of CPI inflation is an increase in the cost-of-living or cost of consumption, henceforth referred to as consumer price inflation” (Connolly, 1996).²¹ Hoffmann (1998) documents that the German Bundesbank stated many years ago that it considered probable biases in the CPI is setting its inflation target, which implicitly accepts the COL index as an inflation standard (because no other standard exists for estimating biases in CPI indexes). In personal conversations, economists in the Bank of Australia described removing mortgage interest from the CPI as an important issue (so the inclusion of mortgage interest in a user cost measure for housing was undesirable for a CPI for monetary purposes), but whether the CPI, otherwise, should approximate a COL index was not objectionable.

²⁰ I distinguish below the difference between the noninflationary standard and the information that the bank needs for monitoring inflation.

²¹ Connolly (1996) distinguishes between the use of the CPI as an inflation measure for monetary policy and as an escalator of wages and pensions. However, she concludes that for both purposes, durable goods should be measured by the flow of services approach suggested by COL index theory. See the discussion of housing in section III, above. It is important to note that in New Zealand the head of the central bank has as an inflation target written into his “job description,” so the measure of consumer inflation that is appropriate for monetary policy--of concern to central banks everywhere--has immediate and substantial relevance to the administration of the New Zealand central bank.

In a survey, the Organisation for Economic Co-operation and Development (1997), noting disagreement on whether monetary policy demands a COL index, lists eight countries where policy makers responded affirmatively (U.S., Germany, France, U.K., Netherlands, Norway, and Sweden). Disagreement was registered “most often [by] national statistical offices” including France, Finland, Sweden, Austria, Belgium, Greece, Ireland, New Zealand, Switzerland (note the inclusion of some countries on both sides of the question).

There are issues yet to be resolved, and policy makers do not always have the same views. For example, Blinder (1997) suggests that consumer inflation may not be the appropriate objective of an anti-inflation policy if the major costs of inflation arise from business decision-making.²²

However, I conclude from the above that the appropriateness of the COL index concept for monetary policy is questioned more vigorously by statistical agencies than by central banks and economic policy makers. To be specific, I believe that Hill’s (1997) views on an inflation index are not shared by central bankers, economic policy makers, and administrators of anti-inflation policies.²³

The following addresses points that have been raised in the discussions among statistical agencies.

2. Causes of inflation and the measurement of it. One often sees statements such as: “Because inflation is a monetary phenomenon, certain prices that would be included in a COL index do not belong in as an inflation index.” Often this assertion is used to justify excluding owner-occupied housing, or as an argument against measuring owner-occupied housing with a rental equivalence measure, or against (some?) imputations.

The statement that inflation is a monetary phenomenon is a statement about its causes. I am inclined to agree with the hypothesis. But the analysis of inflation, and the determination of its causes, is a different topic from the measurement of inflation, even though there are clearly interactions between them. Bringing in a particular hypothesis about the causes of inflation to justify a particular decision about the measurement of inflation confuses and commingles these two topics in a way that is not at all useful. There may be instances where the implications of the COL index *might* create problems for the analysis or the monitoring of inflation or for

²² He notes that economists do not have very good estimates of who bears the cost of inflation, but that much current thinking about it suggests that inflation produces inefficiencies in business decision-making. If that is so, Blinder contends, one should stabilize the price index that influences business decision-making.

²³ I think it relevant to note that I have experience in an anti-inflation program (I was Assistant Director for Price Monitoring in the U.S. Carter Administration’s Council on Wage and Price Stability—CWPS). None of the top economists at CWPS questioned the concept of the COL index, and indeed two of us (myself and Deputy Director R. Robert Russell) had published on the topic.

determining whether a central bank has met its zero inflation goal, but we should be unwilling to accept any ukase that runs from a hypothesis about the cause of inflation to a rule about measuring it.

3. The difference between information that the monetary authorities might need to monitor, predict or forecast future inflation and the standard that indicates whether a “zero inflation” policy has been successful. This point requires distinguishing two different kinds of price information. On the one hand there is the *standard* for determining whether or not a central bank has met its zero inflation goal. As I elaborated above, I believe that the *standard* for zero inflation is provided by the COL index.

On the other hand, the COL index most certainly does not provide all the information one would need to determine future inflation. Monitoring, forecasting, and analyzing inflation requires much more information about price movements and determinants than one could get out of an aggregate COL index. One might need a set of “leading indicators” for inflation. Wage measures often serve this function. Conversely, the idea of “core inflation” suggests a narrower focus.

The need for additional information for monitoring inflation, or for forecasting it, does not invalidate the COL index as the standard for anti-inflationary policy. It is a bit surprising that these two ideas have become confused.

4. Should the standard for “zero inflation” be an index that is broader than the COL index (to include investment goods, internationally traded goods, or asset prices)? It has been asserted that the success of a central bank’s anti-inflation policy should be judged by inflation in the entire economy, and not just in the consumer sector. Among asset prices, inflation in house prices is often mentioned.

In part, this assertion confuses the information necessary for monitoring, forecasting, and predicting inflation with the standard for assessing inflation. Suppose that price indexes for investment goods, correctly measured to account for changes in the productivity of those goods, increased forever at 10% per year, but that consumer prices, measured by a COL index that corrected for quality change, showed zero inflation. Why should the monetary authority care, for its anti-inflation program? If there were never any feedback in the consumer sector, there is no reason to consider inflation in the investment goods part of the economy.

The reason the monetary authority would in fact care about investment goods is different: Price increases in investment goods will eventually feed back into future consumer inflation. Ignoring the warning signs in the non-consumer part of the economy would not be prudent. But that does not make investment goods prices part of the stability criterion; instead, they are being used as forecasters of the future change in consumption prices.

Alternatively, one might believe that a rise in investment goods prices has implications for investment, and in the long run, for economic growth. Growth is an economic policy concern, but it is not anti-inflation policy. What matters for anti-inflation policy, in the end, are

consumer prices. And in the end, other prices only matter for anti-inflation policy if they ultimately feed back into consumer prices (which of course normally they will).

5. Summary for this section. The COL index is a welfare-type measure, as has often been noted. It is very hard to understand why a monetary authority should be interested in stabilizing something that is not a welfare measure. As the standard for anti-inflation policy, then, I believe the COL index is exactly what is wanted. However, the COL index is not the only price information needed for economic policy making in an anti-inflationary regime.

B. The COL index as an escalation measure

In Triplett (1983), I suggested that escalation uses of the CPI do not necessarily imply a COL index. The reason has little to do with any defects in the COL index concept, but depends instead on the objectives of escalation policies. Briefly, the argument goes as follows.

Escalation is generally applied to income payments, such as wages or pensions, so it is natural to think of the COL index as the escalator that would hold the living standard constant. An “income defined” COL index gives the total income necessary in the comparison period to maintain the standard of living of the base period.

But escalation is never applied to *total* income. Many recipients of wages or pensions have other sources of income. Escalation of one component of income by the COL index does not necessarily hold the living standard constant, it depends on what happens to other components. I think this is the reason why no parties to escalation, so far as I can tell, pay much attention to the methodology of the index they build into their agreements, and it probably explains as well why 100% escalation is so rare in private sector agreements.

Alternatively, we might define the purpose of escalation as leaving the income *payment* with the same command over goods and services as it had in the base period. This is not the same objective as holding constant the standard of living of pension recipients. Same command over goods and services implies a measure of inflation, or a deflator for consumer purchases, which in turn implies the expenditure-defined COL index. But that is because the COL index is a measure of *inflation*, not because the COL index provides an escalator that holds constant the incomes of pension, or wage, recipients. It is a subtle distinction, perhaps, but as an important one. The objectives of an escalation policy (for pensions, for example) are seldom thought through.

More information is in my original article, Triplett (1983). A very thoughtful and insightful discussion of the purposes of escalation of pensions is Griliches (1996).

VII. Choosing between Arithmetic and Geometric Mean Formulas for Elementary Price Indexes

It has long been established (see, for example, Carruthers, Sellwood, and Ward, 1980) that an equally-weighted geometric mean of price relatives $(\vartheta_i (p_{i,t} / p_{i0})^{1/n})$ yields a price index

that is lower than one computed using an equally-weighted arithmetic mean of price relatives ($\sum_i (p_{i,t} / p_{i0}) (1/n)$). Two questions arise:

- (a) What determines the size of the difference?
- (b) What price index interpretation should be given to the difference?

After all, if the difference between geometric and arithmetic means is simply a mathematical property, the one that gives the lower inflation rate is not necessarily the “correct” measure.

A. The size of the difference

In the U.K., the difference between geometric and arithmetic means for elementary aggregates (the detailed commodity price indexes that are aggregated into the RPI) amounts to approximately 0.5 index points per year for the aggregate RPI (Fenwick, 1999). This difference is higher than what has been reported for some other European countries, but it roughly equals the comparable calculation for the U.S. (Moulton, 1996).

What accounts for the size of the difference between geometric and arithmetic means? The answer is not entirely clear, in the U.K. and also in other countries. One can cite more or less mechanical reasons for the difference.

For example, elementary aggregates may be defined on broader or narrower definitions of commodities. Both the U.K. and the U.S. have sample selection procedures that explicitly or implicitly include a wide spectrum of commodity varieties within some of the elementary aggregates. With broader definitions of commodities, the elementary price indexes will have more dispersion in the base period prices and—possibly—more dispersion in the price relatives over which the geometric means and arithmetic means are calculated, thus producing a greater difference between them.

A more inclusive sampling procedure must result in a more representative index. Thus, with a given sample size changing the sampling procedure to minimize the difference between arithmetic and geometric means compromises the index.

Alternatively, the difference between geometric and arithmetic means may be influenced by whether the base month for the index is one in which an abnormally high number of special sales is encountered (the U.K. uses January as a base month, some other European countries use December). A base month with more special sales in it will have more dispersion in the base period prices, and elementary price indexes may therefore be more sensitive to the arithmetic-geometric mean calculation.

Obviously, there is no uniquely correct base month. However, Schultz (1994) implicitly points to special sale prices, and to the effects of commodities coming off sale prices, in his demonstration of the effects of alternative formulas (Schultz used soft drink prices collected for the Canadian CPI). Part of the analysis of “formula bias” in the U.S. CPI also points to over-weighting or over-emphasis of price increases on observations returning to normal from sales (Moulton, 1996; United States Department of Labor, Bureau of Labor Statistics, 1998). In the U.K., RPI components that exhibit the greatest arithmetic-geometric mean difference include

furniture and clothing indexes, for which January sales are a prominent aspect of marketing. Thus, sale prices are a reasonable focus in understanding the arithmetic-geometric mean problem in elementary price indexes, and the proportion of sales prices in the base month is also an important consideration.

Both of these possible causes suggest the importance of the conceptual question: Which calculation method for the elementary index is preferred, conceptually?

B. The interpretation of the difference

The Boskin Commission interpreted the difference between geometric means and arithmetic means in the U.S. CPI as “lower level substitution” in consumption. This lower-level substitution interpretation was accepted by the BLS (Abraham, Greenlees, and Moulton, 1998; United States Department of Labor, Bureau of Labor Statistics, 1998). This substitution interpretation makes sense for elementary aggregates that cover a spectrum of varieties, and for which divergent relative price trends occur. However, the commodity substitution interpretation does not make sense for homogeneous elementary aggregates (this is discussed below). Additionally, Moulton (1996) points out that the arithmetic-geometric mean difference will exist whether substitution occurs or not.

The Boskin Commission’s lower-level substitution interpretation of the arithmetic-geometric mean difference depends explicitly on the idea that the U.S. CPI should be an approximation to a cost of living (COL) index. However, many statistical agencies do not view their CPI’s as cost of living indexes. For example, the UK Retail Price Index (RPI) has not in the past been given a COL index interpretation. A consumer substitution interpretation of the geometric mean again turns the discussion back to the conceptual question: In an aggregate price index for consumption or retail sales to households, is the COL index the appropriate underlying conceptual framework?

Figure 1 presents a “scorecard” for summarizing interactions between index number concept and economic interpretation of the difference between geometric mean and arithmetic mean price indexes. The scorecard incorporates the fact that two alternative formulations of arithmetic mean indexes have been discussed in the price index literature, the arithmetic mean of price relatives (the average of the relatives, or AR, defined above) and the ratio of arithmetic averages (RA), defined as: $\sum_i p_{it} (1/n) / \sum_i p_{i0} (1/n)$.

1. Laspeyres index. Statistical agencies in many countries maintain that the Laspeyres index provides the underlying concept for their consumer price indexes or retail price indexes. Eurostat (see Hill, 1999) maintains that the harmonized indexes of consumer prices (HICP) are “fixed basket index numbers,” conceptually, without reference to the concept of the cost of living index.

If the price index is Laspeyres, conceptually, then it is hard to see why the elementary price indexes should not also be Laspeyres. As the scorecard indicates, when the underlying concept is Laspeyres, the arithmetic mean “wins” over the geometric mean, by definition, and

this “score” is tallied in the scorecard. Additionally, there is no conceptual reason for preferring the RA over the AR form of arithmetic mean, when the concept is Laspeyres; in the scorecard, these are entered as a “tie” in the Laspeyres index row.²⁴

2. Cost of living index. Unlike the Laspeyres case, which gives the arithmetic mean as an unambiguous solution, the cost of living index does not fully resolve the issue. The cost of living index framework does, however, help us analyze the question and helps determine the factors on which the decision depends.

There are several cases. To illustrate, suppose a CPI component that contains non-homogeneous observations—a price index for chairs, for example in the furniture indexes of the CPI. This chair price index might include leather, cloth, wooden and plastic chairs, which might have different price movements, even in the short run, either because the underlying supply conditions differ, or because more special sales took place on one kind of chair in the base period.

Commodity substitution behavior. Consumers respond to changes in relative prices by substituting toward the relatively cheaper chairs. In the standard COL index formulation, this means that the arithmetic mean price index for chairs has a substitution bias, relative to the cost of living subindex for chairs.

The geometric mean index is the correct COL index when substitution elasticities are unity. Conversely, the arithmetic mean index is the correct COL index when substitution elasticities are zero.

Thus, if the substitution elasticity is greater than 0.5 (including cases where the substitution elasticity is greater than 1) the geometric mean is a better measure of the COL index for chairs than is the arithmetic mean, even if the geometric mean is not exact. This is shown in the scorecard as a win for the geometric mean.

Conversely, if the true substitution elasticity among different kinds of chairs is closer to zero than to unity (that is, if the substitution elasticity is less than 0.5), then some form of arithmetic mean provides a better measure of the cost of living index than does the geometric mean. But which arithmetic mean? There are two, RA and AR. Both are equally-weighted arithmetic means, but they often give different results, because the implicit weighting structure implied by the two formulas differs.

To resolve this question, a distinction made Fenwick (1999) is useful. Suppose that some of the stores were conducting sales on some kinds of chairs in the base period. RA implies that the quantities are equal across all of the observations in the base period; this is a

²⁴ Eurostat permits either the ratio of average prices (RA) or the geometric mean, partly on the grounds that these two methods for computing basic price indexes do not differ very much. It excludes only the average of relatives (AR) method, which tends to differ empirically from the other two.

reasonable assumption if substitution elasticities across chairs *and across stores* are zero “within the year,” to use Fenwick’s language—the sale does not work. AR implies, instead, that expenditures on all observations are the same in the base period; this is a reasonable assumption if the elasticity of substitution equals unity across chairs and across stores within the year—consumers adjust their purchases to the sales prices, but each store earns the same revenue, sale or no sale.

However, this suggests an inconsistency in AR. Its implicit equal expenditure weights in the base period imply that the elasticity of substitution equals unity within the year. However, because the AR is a form of Laspeyres index, the AR implies that the elasticity of substitution equals zero “between the years,” using again Fenwick’s language. Thus, from a COL index perspective, AR implies inconsistent assumptions within the year and between the years.

RA, on the other hand, is consistent: it relies on the zero substitution assumption both within the year and between the years. In the scorecard, this is shown as a “win” for the RA form of the arithmetic mean, compared with either AR or the geometric mean, when the substitution elasticity is less than 0.5.

Thus, considering only consumer substitution among the different commodities included in the basic component index, the decision between the RA version of the arithmetic mean and the geometric mean depends on the size of the substitution of elasticity within that component. The Bureau of Labor Statistics (1998) carried out an extensive review of available information about substitution elasticities at the lower levels of its consumer price index (CPI). In 39 of 48 components investigated, the BLS report concluded that one could not reject the hypothesis that the substitution elasticity equals unity. For those cases, the BLS decided to compute elementary price indexes using the geometric mean. For the other components, the BLS continued to use a form of the arithmetic mean.²⁵

Other consumer behavior. The Boskin Commission considered that substitution behavior in response to changes in relative prices was the only economic behavior that was relevant to constructing lower level cost of living indexes. That, however, is clearly not the case.

Pollak (1998) and also Triplett (1998) point to consumer search behavior, shopping behavior with respect to sales, switching between stores, stockpiling of more durable consumption commodities (soft drinks, for example) by acquiring them at sale prices, and so forth. Indeed, if a CPI component covers only a single, homogeneous commodity (bananas), then by definition no commodity substitution can take place: The difference between

²⁵ For two cases, medical care and utilities, BLS rejected its own statistical findings on the grounds that substitution between different kinds of medical services was unlikely, as was substitution between, say, electricity and trash collection.

arithmetic and geometric means cannot be given a commodity substitution interpretation for a homogeneous commodity, yet arithmetic and geometric means will still differ.²⁶

At present, the theory of the cost of living index is not well developed with respect to consumer behavior other than commodity substitution. For this reason, the theory does not indicate the proper way to construct a cost of living index that incorporates consumer shopping and search behavior. For the same reason, when the difference between arithmetic and geometric means stems from consumer shopping and search behavior—and therefore the cause of the difference in indexes arises from special sales prices—cost of living index theory does not help resolve the choice between arithmetic and geometric mean indexes. This ambiguity is scored as a tie in the scorecard (even though it is very unlikely that the arithmetic mean will ever prove to be the best measure to incorporate consumer search behavior into a cost of living index).

3. Summary. To summarize, for the COL index case, and considering only commodity substitution behavior, the scorecard shows that the choice between arithmetic mean (RA) and geometric mean depends on whether the substitution within a basic component of the cost of living index is greater than or less than 0.5. In the COL case, use of the AR form of the arithmetic mean involves inconsistent assumptions about commodity substitution.

Consumer behavior other than commodity substitution may not be handled adequately by a geometric mean across price relatives. A better understanding of consumer shopping and search behavior toward sales and so forth might in the future lead to an alternative proposal for calculating elementary price indexes.

VIII. Some Issues Raised by Ralph Turvey

Turvey (1999) concludes:

“The formal concept of a true cost-of-living index can safely be ignored by statistical offices when designing and implementing a consumer price index. It:
As yet provides little assistance to the construction of [the actual CPI];
Results in the inappropriate inclusion of imputed items...
Is congenial to economic theorists but not to most users.”

Much of Turvey’s paper consists of a list of shortcomings of COL index theory. Many of his points are well taken (although some are not). But even where I agree that he has identified a shortcoming, his antipathy toward the COL index concept leads him to extreme positions.

²⁶ It has been proposed that consumer shopping and search behavior across stores could be incorporated into the indexes by treating each product in each retail outlet as a separate commodity, calculating the geometric mean across all of the retail outlets, and interpreting the difference between geometric and arithmetic means as substitution across retail outlets. Triplett (1998) shows with an example that this does not work.

A. Flow of services approach to durable goods

It is one thing to be uncomfortable--as many people are, to some degree--with the rental equivalence approach to owner-occupied housing. All approaches to owner-occupied housing have empirical difficulties.

But Turvey makes a different, and unsupportable (at least by me) point: “To include imputed rentals for owner-occupied housing and for such other durables *where it is feasible*, but not for all those durable goods and services *where it is not feasible*, seems distinctly odd” (emphasis implied).²⁷ He concludes from this that a flow of services approach should not be used even where it is feasible, under the rule of consistency. This argument says that it is better to have a measurement that is universally wrong than one that is right for the major cases where it matters.

I prefer a far different rule. The flow of services from durables is the right measure of consumption, and pricing the flow of services from durables is the right component of a COL index. But the appropriate measure is hard to implement, and in an operational index, pragmatism enters in as well. So my own rule is: Apply the flow of services pricing approach where it matters most, and do not worry about the more complicated approach where it does not matter, subject in all cases to feasibility. Clearly, in housing, it matters, both because housing has a large weight in the CPI and because we know (from the BLS experience in the 1970's, for example) that the flow of services approach creates a different price measure.

Unfortunately, there are too few empirical studies to apply my rule very extensively across the CPI. My rule says that it is important to do research on flow of services measures because improvements can be made gradually as progress is made. Turvey's rule implies that one should never put any resources into it because the last obscure durable good will always remain to be estimated, preventing implementation of any of the others.

B. “Common sense” solutions and the COL index

Turvey (1999) states: “To decide how [the CPI] should treat problems such as those rehearsed above requires a clear formulation of the purpose to be served by the index and the application of only elementary economics and common sense.” The topics he notes include use of the Fisher Ideal index number, the treatment of multipart tariffs, and the use of hedonic functions for adjusting for quality change.

Certainly, the contribution of good economic judgment to the solution of economic problems cannot be denied. There is little worse than a mediocre theorist run amok. And no doubt, as Turvey implies, the economic techniques necessary to comprehend and to solve CPI problems do not require all of the technical apparatus on display in a typical graduate-level economic theory textbook. Yet, Turvey underestimates the contribution of his own substantial

²⁷ Actually, I might quibble about that word “odd.” If something is not feasible, it is not feasible. What is odd about not doing something that is not feasible?

economic skills in deriving his proposed solutions to the problems he lists, and underestimates especially the contribution of economic theory to the work of those economists who originated of those solutions.

1. Fisher indexes. “The common-sense view is simply that the best way of comparing 1998 prices with 1997 prices is to take account of both the 1997 and the 1998 patterns of consumption rather than just one or the other” (Turvey, 1999).

This “common sense” view is not as an unappealing one. (I have sometimes used it myself in explaining Fisher indexes to a nontechnical audience.) But why--if only common sense is required--have statistical agencies argued so forcefully against the Fisher index over so many years, and not infrequently contended that the Fisher index was conceptually inappropriate? Was “common sense” in so short supply? And why is the acceptance of Fisher indexes (and other superlative index numbers) growing today? I suggest it is less that we have more common sense than our predecessors and more that we now have the benefit of Mr. Diewert’s elegant theory of the superlative index number (Diewert, 1976). It is very hard to contend that COL index theory had nothing to do with Diewert’s development of the superlative index idea.²⁸

2. Multipart tariffs. “There is a common sense solution: The fixed charge and the unit charge are, respectively, the price of being able to consume and the price per unit consumed, so should be given separate weights” (Turvey, 1999).

Of course, I agree with Turvey’s solution, which amounts to “unbundling” the commodity into more fundamental units that are the real units that households consume. Perhaps only our joint common sense determines this agreement and economic theory has had no influence on our joint thinking. However, I can recall years ago vigorous exchanges of memos within one statistical agency on exactly this same topic (when I then proposed what Turvey now proposes), suggesting that--in the U.S., at least--ordinary common sense did not always extend to the idea of unbundling the commodity into more basic elements of consumption.

I believe that I got this idea for treating multipart tariffs from Stone (1956), and I suspect that Turvey got it from his own extensive experience in electricity rate-setting. One good thing about the idea is its appeal to common sense. But it did not seem an ordinary idea when Stone (1956) first proposed it, it was controversial when it was first proposed and if more acceptable today may not still be fully appreciated. Did his idea come to Stone because he had as an unusual amount of common sense or did his unusual command of economics also have

²⁸ In the preceding sentence to the one quoted, Turvey makes a common mistake: He asserts that the axiomatic approach to index numbers also yields a Fisher Ideal formula. Since Frisch (1936), we have known that the axiomatic approach yields a Fisher index *only* if we put in the “right” axioms. Additivity is quite a useful “common sense” property for both CPI and national accounts index numbers. If one includes additivity in the set of axioms, the axiomatic approach excludes the Fisher index.

something to do with it? In any event, this now old idea of Stone's does not make a very good example for why we do not "need" the conceptual assistance of COL index theory, which I defined (in section II) as use of the economic concept of consumption in thinking about the problems that arise in computing the CPI.

3. Hedonic functions. "But [use of hedonic coefficients] can also be justified without [theory], by treating some components of consumption as bundles of market-price-relevant characteristics. This means regarding the index as measuring changes in the cost of buying a reference-period collection of bundles of such characteristics rather than a set of consumption items" (Turvey, 1999).

Again, I like Turvey's suggestion. Indeed, I *believe* that I was the first one to propose redefining a (Laspeyres) index number from goods space to characteristics space, in a quite old and now obscure document prepared for the BLS some time before I went there (Triplett, 1971). At the time, I must admit that I thought this solution was novel, and so it must be said did many others (some who accepted it, some who did not).

But even though I developed a rationale for interpreting the hedonic characteristics in a Laspeyres index framework (exactly the same as the one that Turvey now proposes, nearly 30 years on), the approach was hardly devoid of the theory of consumption. A number of economists (Gorman, 1980, but written and circulated in the 1950's; Ironmonger, 1972, but written in the 1960's and known earlier in manuscript; and Lancaster, 1966, 1971) had proposed redefining consumer demand on the characteristics of goods rather than on the goods themselves. It was clear to most of us who were working on hedonic functions that a hedonic function also treated the consumption good as a bundle of utility-generating characteristics (Griliches, 1988, described automobiles this way). What I did, then, was to extend these two ideas into the making of price indexes.

Later, I (Triplett, 1983, 1987), and others, also discussed building into COL index theory the idea that consumption commodities could be unbundled into more basic utility-generating variables (the characteristics). One reason that I began by redefining the Laspeyres index into characteristic space was that I wanted to lessen resistance to hedonic indexes, so for strategic reasons I did not want to couple the then-controversial hedonic technique with the also controversial COL index.

Perhaps, as Turvey asserts, all of this was merely common sense. I agree with him that the approach seems sensible. But only common sense and a minimal amount of economics? It took years for the basic idea of hedonic indexes to become acceptable (even ten years ago, hedonic indexes were still very controversial, even in the U.S.--see my discussion of the reasons for this in Triplett, 1990). On Turvey's interpretation, I suppose that the greater acceptability of hedonic functions today than thirty years ago (or even ten years ago) is because we have somehow accumulated a larger world-wide stock of common sense. I also observe that it is precisely those statistical agencies that have the strongest antipathy to the idea of the COL index that also have historically most questioned the applicability of hedonic indexes in CPIs, which is strong evidence that the two are coupled in most professionals' minds.

4. Conclusion to this section. I am not saying that one has to be a master economic theorist to understand or accept Fisher indexes, the unbundling of goods to simplify pricing of multipart tariffs, or hedonic price indexes. Many of the ideas are quite simple. They can also be given interpretations that are neither exotic nor violate common sense, which in my view is a very good thing.

But Turvey undervalues the economics that generated these innovations. One reason is that the examples he gave are now quite old ones--from 45 in the case of Stone's to 20-25 in the case of Diewert's and Triplett's--and they have moved from the fringe of index numbers into, if not quite the mainstream, at least the familiar. None of the originators was unaware of developments in theoretical consumption economics. Turvey vastly understates the role of the economic concept of consumption--which is the essence of COL index theory--in generating the innovations he uses as examples in his attempt to show that the theory is not useful.

IX. Concluding Remarks

This paper is long. A summary (beyond what has already been provided in section I) is perhaps inappropriate.

Its major themes come from the two questions that opened the paper. The first question (does a CPI *need* an underlying conceptual framework?) was addressed only implicitly. When price index agencies explicitly adopt the COL index, it is obvious that they do so because they also believe that an underlying CPI conceptual framework is necessary. But even when agencies do not adopt an explicit framework, an implicit framework evolves out of the practical decisions that are made in constructing the index, though it is often hard to discern what that implicit framework is. One issue, then, is whether it is better for the underlying framework to be explicit and written down, or implicit.

With respect to the second question, I contend that the theory of the cost-of-living index does provide the underlying conceptual rationale for constructing a practical CPI, for the reasons elaborated in the seven substantive sections of the paper. These sections put much more emphasis on COL index theory as a tool for resolving practical index number issues—and much less emphasis on substitution bias and index number formulas—than has generally been the case in the price index literature. The important difficulties in estimating price indexes concern measuring the components. Constructing detailed price indexes for coats and carrots and computers and cars poses more vital questions, empirically, than aggregating those components into an overall CPI.

No coherent alternative to the COL index framework, what I referred above as the “not COL index,” exists. Neither the Laspeyres index, nor a “fixed-basket” index, nor the idea of a “pure price” index provides an underlying conceptual framework for resolving measurement issues within CPI detailed component indexes.

A kind of “two different worlds” syndrome exists in the price index literature. On the one hand, academic contributors are often unfamiliar with the complexities of price index

construction, and may underestimate the degree of difficulty in the decisions that go into an actual price index. Perhaps for this reason, they typically are more concerned with index number formulas, which is the topic, after all, that makes up the bulk of the index number literature. Thus, academic contributors are likely to see the COL index in terms of substitution bias, and not the theory's contribution to estimating the component price indexes that are aggregated with the index number formula. Statistical agency contributors are more likely to understand the difficulties, and are less likely to put questions such as substitution bias at the forefront of their concerns. For this reason—the COL index usually being presented in terms of the price index substitution bias—they are less disposed to the theory, which seems to them (as it does to some of the academics) esoteric or unhelpful. Because I view both the typical academic and the typical agency positions as partly right and partly wrong, this paper is a contribution toward breaking down the barriers between those “two different worlds.”

References

- Abraham, Katherine G., John S. Greenlees, and Brent R. Moulton. 1998. "Working to Improve the Consumer Price Index." *Journal of Economic Perspectives* 12(1): 27-36, Winter.
- Allen, R.G.D. 1933. "On the Marginal Utility of Money and Its Application." *Economica* (36): 186-209, May.
- Australian Bureau of Statistics. 1997. Outcome of the 13th Series Australian Consumer Price Index Review. Australia: Australian Bureau of Statistics.
- Balk, Bert M. 1994. "The New Consumer Price Indices of Statistics Netherlands." *Statistics Journal of the United Nations* 11(2): 119-123.
- Baye, M.R. 1985. "Price Dispersion and Functional Price Indices." *Econometrica* 53:1 (January): 213-223.
- Berry S., J. Levinsohn, and A. Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63(4): 841-90, July.
- Blackorby, Charles and Robert R. Russell. 1978. "Indices and Subindices of the Cost-of-Living and the Standard of Living," *International Economic Review* 19:1 (February): 229-240.
- Blinder, Alan S. 1997. "Measuring Short-Run Inflation for Central Bankers: Commentary." *Federal Reserve Bank of St. Louis Review* 79(3): 157-60, May/June.
- Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgenson. 1996. "Toward a More Accurate Measure of the Cost of Living." Advisory Commission to Study the Consumer Price Index. Interim Report to the Senate Finance Committee, December 4.
- Carruthers, A.G., D.J. Sellwood, and P. Ward. 1980. "Recent Developments in the Retail Prices Index." *The Statistician* 29(1), March.
- Connolly, Catherine. 1996. "The Purpose of the CPI." *The New Zealand Reserve Bank Research Memo* M96/34, July 16.
- Copeland, Morris A. and Edwin M. Martin. 1938. "The Correction of Wealth and Income Estimates for Price Changes." Conference on Research in Income and Wealth, Studies in Income and Wealth 2: 85-119 and 131-35. New York: National Bureau of Economic Research.
- Court, Louis M. 1941a. "Entrepreneurial and Consumer Demand Theories for Commodity Spectra (Part I)." *Econometrica* 9(2): 135-162, April.

- Court, Louis M. 1941b. "Entrepreneurial and Consumer Demand Theories for Commodity Spectra (Concluded)." *Econometrica* 9(3/4): 241-297, July-October.
- Cunningham, Alastair W. F. 1996. "Measurement Bias in Price Indices: An Application to the UK's RPI." Bank of England, Working Paper Series, No. 47, March.
- Dalén, Jörgen. 1999. "Some Issues in CPI Construction." Prepared for the Fifth Meeting of the International Working Group on Price Indices, Reykjavik, Iceland, August 25-27.
- Deaton, Angus. 1998. "Getting Prices Right: What Should Be Done?" *Journal of Economic Perspectives* 12(1): 37-46.
- Diewert, W. Erwin. 1976. "Exact and Superlative Index Numbers." *Journal of Econometrics* 4:2 (May): 115-45.
- Eurostat. 1999. Compendium of HICP Reference Documents: Harmonisation of Price Indices. Eurostat, January.
- Fenwick, David. 1999. "The Impact of Choice of Base Month on the Relative Performance of Different Formulae Used for Aggregation of Consumer Price Index Data at an Elementary Aggregate Level." Prepared for the Fifth Meeting of the International Working Group on Price Indices, Reykjavik, Iceland, August 25-27.
- Fisher, Franklin M. and Karl Shell. 1972. The Economic Theory of Price Indices: Two Essays on the Effects of Taste, Quality, and Technological Change. New York: Academic Press.
- Fisher, I. 1921. "The Best Form of Index Number." *Journal of the American Statistical Association* 17: 535-37, March.
- Friedman, Milton. 1938. "Discussion." Conference on Research in Income and Wealth, Studies in Income and Wealth 2: 123-30. New York: National Bureau of Economic Research.
- Frisch, R. 1936. "Annual Survey of General Economic Theory: The Problem of Index Numbers." *Econometrica* 4: 1-39.
- Gillingham, Robert F. 1983. "Measuring the Cost of Shelter for Homeowners: Theoretical and Empirical Considerations." *The Review of Economics and Statistics* 65(2): 254-65, May.
- Gorman, W.M. 1980. "A Possible Procedure for Analyzing Quality Differentials in the Egg Market." *Review of Economic Studies* 47 (October): 843-56.
- Griliches, Zvi. 1988. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change." (1961 article, reprinted in Zvi Griliches' Technology, Education, and Productivity. Great Britain: Bookcraft Ltd.)

Griliches, Zvi. 1996. Testimony before the Senate Finance Committee.

Hill, Peter. 1997. "The Measurement of Inflation and Changes in the Cost of Living." Prepared for the Conference of European Statisticians, Joint ECE/ILO Meeting on Consumer Price Indices, Geneva, November 24-27.

Hill, Peter. 1999. "COL Indexes and Inflation Indexes." Presented at the Fifth Meeting of the International Working Group on Price Indexes, Reyjavik, Iceland, August 25-27.

Hoffman. 1998. Problems of Inflation Measurement in Germany. Frankfurt, Germany: Deutsche Bundesbank.

Holloway, Sue. 1999. "Where Is the RPI on the Pure Price Index--Cost-of-Living Continuum?" Unpublished paper, January 20.

Hulten, Charles R. 1973. "Divisia Index Numbers." *Econometrica* 41(6): 1017-25, November.

Ironmonger, D.S. 1972. New Commodities and Consumer Behavior. *University of Cambridge, Department of Applied Economics Monographs* 20. Cambridge: Cambridge University Press.

Konus, A.A. 1939. "The Problem of the True Index of the Cost of Living." *Econometrica* 7 (January): 10-29.

Krugman, Paul R. 1998. The Accidental Theorist: and Other Dispatches from the Dismal Science. Norton, 1998.

Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *The Journal of Political Economy* 74(2): 132-157.

Lancaster, Kelvin J. 1971. Consumer Demand: A New Approach. New York: Columbia University Press.

Lebow, David E., John M. Roberts, and David J. Stockton. 1994. "Monetary Policy and 'The Price Level.'" Washington, DC, Board of Governors of the Federal Reserve System.

Manser, Marilyn E. and Richard J. McDonald. 1988. "An Analysis of Substitution Bias in Measuring Inflation, 1959-85." *Econometrica* (July): 909-30.

Moulton, Brent R. 1996. "Bias in the Consumer Price Index: What Is the Evidence?" *Journal of Economic Perspectives* 10(4): 159-77, Fall.

Organisation for Economic Co-operation and Development. 1997. "Synthesis Paper on Shortcomings of the Consumer Price Index Measure of Inflation for Economic Policy Purposes." Prepared at the request of the Economic Policy Committee at its meeting of December 1996.

- Oulton, Nicholas. 1995. "Do UK Price Indexes Overstate Inflation?" *National Institute Economic Review* 0(152): 60-75.
- Pakes, Ariel. 1997. "New Goods, Characteristic Based Models, and the Role of Hedonics." Unpublished paper. Yale University.
- Pollak, Robert A. 1989. The Theory of the Cost of Living Index. New York: Oxford University Press.
- Pollak, Robert A. 1998. "The Consumer Price Index: A Research Agenda and Three Proposals." *Journal of Economic Perspectives* 12(1): 69-78, Winter.
- Randolph, William C. 1988. "Housing Depreciation and Aging Bias in the Consumer Price Index." *Journal of Business and Economic Statistics* (July): 359-71.
- Schultz, Bohdan J. 1994. "Choice of Price Index Formulae at the Micro-Aggregation Level: The Canadian Empirical Evidence." Prepared for the International Conference on Price Indices, Ottawa, Canada, November.
- Shapiro, Matthew D. and David W. Wilcox. 1996. "Measurement in the Consumer Price Index: An Evaluation." *National Bureau of Economic Research Working Paper* 5590 (May).
- Shiratsuka, Shigenori. 1998. "Measurement Errors in Japanese Consumer Price Index." Prepared for the Bank of Japan and the Federal Reserve Bank of Chicago, June.
- Statistics Canada. 1995. The Consumer Price Index Reference Paper. Ottawa, Canada: Minister of Industry, July.
- Stigler, George J. 1961. "The Economics of Information." *The Journal of Political Economy* 69(3): 213-225.
- Stone, Richard. 1956. Quantity and Price Indexes in National Accounts. Paris: Organisation for European Economic Cooperation.
- Stott, Helen. 1998. "The CPI Purpose and Definition--The Australasian Debate." In Proceedings of the Fourth Meeting of the International Working Group on Price Indices, edited by Walter Lane. Washington, DC: Department of Labor.
- Thaler, Richard H. 1991. Quasi Rational Economics. New York: Russell Sage Foundation, 1991.
- Triplett, Jack E. 1971. "The Theory of Hedonic Quality Measurement and Its Use in Price Indexes." *BLS Staff Paper* 6.

Triplett, Jack E. 1983. "Escalation Measures: What is the Answer? What is the Question?" In Price Level Measurement, edited by W. Erwin Diewert and C. Monmarquette. Ottawa, Canada: Minister of Supply and Services Canada, 457-82.

Triplett, Jack E. 1987. "Hedonic Functions and Hedonic Indexes." In John Eatwell, Murray Milgate, and Peter Newman, eds. The New Palgrave: A Dictionary of Economics. New York: Stockton Press, 630-34.

Triplett, Jack E. 1990. "Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy." In Ernst R. Berndt and Jack E. Triplett, eds. Fifty Years of Economic Measurement: The Jubilee. Conference on Research in Income and Wealth, Studies in Income and Wealth 54: 207-33.. University of Chicago Press for the National Bureau of Economic Research.

Triplett, Jack E. 1998. "Elementary Indexes for a Consumer Price Index." In Proceedings of the Fourth Meeting of the International Working Group on Price Indices, edited by Walter Lane. Washington, DC: Department of Labor.

Turvey, Ralph. 1989. Consumer Price Indices. Geneva: International Labour Organisation.

Turvey, Ralph. 1999. "True Cost of Living Indexes." Prepared for the Fifth Meeting of the International Working Group on Price Indices, Reykjavík, August.

United States Department of Labor, Bureau of Labor Statistics. 1997. "The Consumer Price Index." In BLS Handbook of Methods for Surveys and Studies (April): 167-202.

United States Department of Labor, Bureau of Labor Statistics. 1998. "The Use of the Geometric Mean in the Elementary Indexes of the Consumer Price Index." Unpublished draft report.

Walsh, C.M. 1921. "The Best Form of Index Number: Discussion." *Journal of the American Statistical Association* 17: 537-44, March.

Wynne, Mark A. And Fiona Sigalla. 1993. "A Survey of Measurement Biases in Price Indexes." *Federal Reserve Bank of Dallas Research Paper* 9340, October.

Figure 1

Score Card, Arithmetic and Geometric Means

Index Concept	Arithmetic Mean of Relatives (AR) ¹	Ratios of Arithmetic Means (RA) ²	Geometric Mean ³
Laspeyres index	1	1	0
Cost of living index			
Substitution (lower level)	0	0	1
Elasticity>0.5	0	1	0
Elasticity<0.5	0	0	0
Other behaviors			

¹ AR = average of price relatives = $\sum_i (P_{it}/P_{i0})(1/n)$. AR implies that the observations have equal expenditure weights in the base year.

² RA = ratio of average prices = $(\sum_i P_{it}/n)/(\sum_i P_{i0}/n)$. RA implies that the observations have equal quantity weights in the base year.

³ Though it is usually not explicitly stated, calculation of geometric means implies that observations have equal expenditure weights in the base period, which is required for the cost of living index interpretation of the geometric mean index. Equal quantity weights among observations in the base period would yield unequal expenditure weights for the geometric mean (see Triplett, 1998, for the distinction between quantity weighted and expenditure weighted geometric mean indexes).