

Ordinal Utility and the Conceptual Basis of Cost-of-Living Indexes: How History Can Inform Contemporary Debates

Thomas A. Stapleford*
University of Notre Dame
March 7, 2013

ABSTRACT

This essay argues that closer study of the history of neoclassical price index theory can clarify conceptual issues that still hinder contemporary discussions of cost-of-living indexes. I demonstrate that neoclassical price index theory has been marked by important conceptual ruptures, linked to the ordinal revolution in economics, that have transformed the core objective for cost-of-living indexes. By tracking these shifts and their connections to the reception of neoclassical approaches among statisticians and to the applications for price indexes in economic analysis and government policy, we can better understand key issues in contemporary debates.

The publication of the Boskin report on the US Consumer Price Index, CPI (Boskin, Dulberger, et. al., 1996), sparked a lengthy debate about the measurement objectives of consumer price indexes that has yet to be resolved (e.g., Baker 1998; Deaton 1998; National Research Council 2002; Berndt 2006; Triplett 2006). In my view, this intransigence reflects three factors. First, consumer price indexes often serve multiple functions (e.g., deflating components of national income accounts, adjusting inflation-protected bond payments, adjusting wages, etc.), and it is by no means obvious that the same measurement objective will serve each application equally well (cf. Triplett 1983). Second, even for a single task, such as adjusting Social Security benefits, there may be disagreement about the relevant policy goals (what kinds of economic changes

* Associate Professor, Program of Liberal Studies, 215 O'Shaughnessy Hall, University of Notre Dame, Notre Dame, IN 46556, USA. E-mail: tstaplef@nd.edu. This paper is a revised version of an earlier essay: "Aftershocks from a Revolution: Ordinal Utility and Cost-of-Living Indexes," *Journal of the History of Economic Thought* 33, no. 2 (June 2011): 187-222. Those interested in additional historical examples and evidence should consult this earlier version. Research for this project was conducted while I was a Visiting Scholar at the American Academy of Arts & Sciences. I am grateful for the helpful comments on earlier versions of this argument from Franklin M. Fisher, H. Spencer Banzhaf, and two anonymous *JHET* reviewers, as well as subsequent comments from Bert Balk.

should lead to altered benefit payments?) and hence inevitable conflict about measurement goals. (I have written extensively about this issue in Stapleford 2009.) Finally, current price index theory imposes important constraints on the interpretation of cost-of-living indexes, but the consequences of these constraints are often overlooked or misunderstood, even by many economists. As a result, conversations about the first two issues have been further muddled.

Although historical analysis cannot resolve debates about consumer price indexes, I believe that it can clarify the discussion, especially on this third point. By examining the development of neoclassical price index theory (the dominant form today), we can see the challenges that it faced, how its conceptual foundations were modified to meet those challenges, and what those shifts meant for the interpretation of cost-of-living indexes. Furthermore, we can understand how the development of neoclassical price index theory became entangled in the emergence of new applications for consumer price indexes.

The core of my argument is that the history of neoclassical price index theory cannot be understood apart from the history of utility and its role in economic theory. Among contemporary economists, cost-of-living indexes are routinely defined in utility-based terms -- e.g., the minimum change in expenditure required to keep a consumer at the same utility level (or indifference surface) in two different situations. Such unanimity is a recent development; at least through the 1940s, numerous economists denied that a cost-of-living index could or should be predicated on concepts such as utility, welfare, or satisfaction (Banzhaf 2001; Stapleford 2009). Still, the utility-based definition of a cost-of-living index -- what we can call a “constant-utility index” -- can boast a long history of support among neoclassical economists and the early marginalists (e.g., Sidgwick 1887, p. 63; Edgeworth 1925 [1887—1889], pp. 208, 210; Frisch 1936, pp.11-12).

Despite this seeming continuity, however, historians recognize that the conception of utility in economic theory underwent dramatic changes during the early twentieth century. Many early neoclassical economists, especially those in the Marshallian tradition at Cambridge, had treated utility functions as producing cardinal values and had assumed that utility could (in principle) be compared objectively across individuals. By the 1950s, most neoclassical economists would dissent from these propositions, arguing that utility functions should be understood in ordinal terms and that interpersonal comparisons of utility required normative judgments (Schumpeter 1954, pp. 1057-1066; Cooter and Rappoport 1984; Hennisman 1987;

Backhouse 2003). Given that prominent contributors to the theory of constant-utility indexes -- e.g., Edgeworth, A. C. Pigou, Alexander A. Konüs, R. G. D. Allen, John R. Hicks, and Paul Samuelson -- worked on opposite sides of this divide, we might suspect that the “ordinal revolution” (as these collective changes in the treatment of utility are often known) would have important consequences for views on constant-utility indexes. Indeed, a closer look at the history of neoclassical price index theory reveals important conceptual ruptures linked to the ordinal revolution, including transformations in the core objective of cost-of-living indexes.

AN OVERVIEW

The changes to neoclassical price index theory have been subtle, albeit important. Characterizing these shifts therefore requires close attention to the details of different texts, especially since individual authors did not always offer coherent expositions of the fundamental concepts behind their methodological arguments. Before immersing ourselves in these twisting pathways, it may be helpful to gather a sense of the larger terrain, the general ways in which changing ideas about utility affected price index theory over the twentieth century.

For cost-of-living indexes, the importance of the ordinal revolution lay in its implications for interpersonal and intertemporal comparisons of utility levels. In the abstract, ordinal values and noncomparability are two distinct concepts (Sen 1979, p. 13; Hennisman 1987). Cardinal utility functions, for instance, can also be noncomparable, and in fact some neoclassical economists began to adopt a limited form of noncomparability without committing themselves to ordinal utility values. Nevertheless, the critiques that led to the adoption of ordinal utility also implied either full noncomparability or extreme limitations on any comparisons. In this respect, the historical link between the adoption of ordinal utility and the acceptance of noncomparability is not accidental.

For the sake of analytic clarity (not always consistently present in the historical texts themselves) we can distinguish two basic kinds of revisions to neoclassical price index theory that were prompted by the ordinal revolution. I will label these the weak and strong revisions, respectively. The weak revision, which emerged rapidly among early progenitors of the ordinal revolution such as R. G. D. Allen (1933) and John R. Hicks (1940), contended that intertemporal comparisons of welfare had theoretical meaning only when the relevant consumer tastes (represented by indifference surfaces) were identical in the two situations. The strong version,

which was not fully articulated until a paper by Franklin M. Fisher and Karl Shell (1968), argued that intertemporal comparisons of welfare were impossible under any conditions whatsoever, a conclusion that Fisher and Shell described as a “testament to our ordinalist purity” (Fisher and Shell 1968, p. 134 n. 2).

The ordinal revolution was (in the main) epistemological, not ontological. In other words, recourse to ordinal utility did not reflect underlying claims about the nature of human welfare or human psychology but instead reflected epistemological judgments about the limits of knowledge and the necessary preconditions for a working theory of consumer demand. Critics insisted that observations of consumer behavior could only reveal preference rankings (ordinal values) and not the relative magnitudes of those desires (cardinal values). Furthermore, they demonstrated that an ordinal utility function (let us call it “ U_R ”, for its basis in ordinal rankings) was sufficient to formulate a complete theory of consumer demand. Given a collection of goods and services, $X = (x_1, \dots, x_n)$, U_R could then define a particular indifference curve, μ_r :

$$\mu_r = U_R(X)$$

Despite these economists’ commitment to ordinal utility in consumer demand theory, however, many of them also understood that the psychological satisfaction or “welfare” provided by a set of goods -- let us call this μ_w -- was determined by more than simply relative preferences, U_R :

$$\mu_w = U_W [U_R(X); E]$$

Here, E represents various exogenous factors that (according to different economists) might affect the form or result of U_W , such as mental or physical health, other aspects of the social or economic environment beyond the goods purchased (X), and so forth. The catch was that neither the form of U_W nor μ_w itself could be determined by observing consumer purchases.

For consumer demand theory, the distinction between μ_r and μ_w made little difference: consumer choices always depended on the evaluation of goods at a fixed moment in time (when both E and U_W would be fixed), and hence relative preferences were all that mattered. (That was why, of course, it had been possible to eliminate cardinal magnitudes from consumer demand theory in the first place.) At any given instant (for a specific individual), economists could say that bundles of goods on the same indifference surface, μ_r , provided the same “welfare,” μ_w , without getting into trouble. But comparisons over time were another matter. The factors beyond U_R that determined μ_w might have changed in the interim. Thus maintaining constant relative preferences (a constant form for U_R) and even purchasing the same set of goods (X) was not

sufficient to imply constant welfare. As Fisher and Shell put it, a “man’s efficiency as a pleasure-machine may have changed without changing his tastes [i.e., relative preferences, U_R]” (1972, 3). Accordingly, even the strictures of the weak ordinal revision (constant U_R) were insufficient; intertemporal comparisons of welfare were simply impossible without subjective, normative judgments.

The strong ordinal revision -- which was tacitly adopted by leading theorists in foundational texts such as Diewert (1976) and Pollak (1989) -- necessarily required redefining the goal of a cost-of-living index. Under both the cardinal and weak ordinal versions, neoclassical economists had assumed that a “true” cost-of-living index would show the minimum change in expenditure needed to maintain an individual at the same level of economic welfare in two different time periods or geographical locations. The strong ordinal critique denied the possibility of intertemporal comparisons of welfare, and hence the possibility of determining the “same level of welfare” in two real situations. Instead, strong ordinalists substituted a different objective: finding the minimum change in expenditure needed to make a given individual (fixed in time and space) indifferent between two systems of prices. This subtle and seemingly minor shift in fact had important implications, particularly when utility functions changed over time.

The consequences become evident when we compare how the three major approaches to constant-utility indexes (cardinal, weak ordinal, and strong ordinal) responded to shifts in utility functions (see Table 1). Among the early cardinal utility theorists, who sought (in principle) to maintain a constant μ_w across the comparison periods, a constant-utility index could be affected not only by changes in prices but also by changes in the form of the functions U_R or U_W . Since neither of these functions could be directly observed, those economists who sought to create such an index by observing prices and quantities had to assume the stability of E , U_R , and U_W in order to make their calculations operationally viable.

For the proponents of the weak ordinalist revision, where the goal was to maintain a consumer on the same indifference curve, μ_r , the problems were at once narrower and more extreme. They could ignore E and U_W , but differences in the form of U_R presented a thorny dilemma: What could it mean to keep a consumer on the “same indifference curve” when the shape of the entire utility function had been altered? Weak ordinalists tended to respond that this question was meaningless, and hence that comparisons of economic welfare were impossible (without introducing normative judgments) if relative preferences (U_R) had changed. The

assumption of stable utility functions -- which had been an operational necessity for cardinal utility theorists -- now became an essential precondition for even the theoretical validity of a neoclassical cost-of-living index. If relative preferences had changed, calculating a true cost-of-living index for those time periods was simply impossible.

The variability of utility functions was therefore the great bugbear of neoclassical price index theory; indeed, through the 1950s it remained the most serious and damming obstacle to the acceptance of the constant-utility ideal as an objective for actual cost-of-living indexes. That view was particularly prevalent among American institutional economists. Steeped in Veblen and interwar American psychology, institutional economists viewed consumer tastes, desires, and consumption-based satisfaction as being constantly in flux, shaped by advertising, emulation, and the demands of a dynamic environment. In the face of such variation, how could one create a constant-utility index, and indeed, what would be its practical use? Such presuppositions shaped the institutionalist view of neoclassical price index theory and help to explain why U.S. government economists famously rejected the recommendation from a 1959--1961 review committee of the National Bureau of Economic Research to make the constant-utility index into a working ideal for the U.S. Consumer Price Index (National Bureau of Economic Research 1961).

The strong ordinal revision, however, bridged the impasse between theoretical discussion and practical application by redefining the objective of constant-utility indexes, and thereby transforming the challenges that variations in utility functions posed to these calculations. A strong ordinal cost-of-living index avoided intertemporal comparisons of real situations and instead evaluated two price systems (abstracted from their larger economic and social context) from the perspective of a hypothetical consumer fixed in time and space. Accordingly any given calculation necessarily depended on only a single ordinal utility function, U_R . The puzzle created by different utility functions now reduced to this: When drawing prices from two different time periods (0 and 1) in which the relevant consumer might have different utility functions (U_R^0 and U_R^1), which function should be used for the calculation?¹ That decision proved to be far more tractable than the empirical problems posed by the cardinal constant-utility index, and it simultaneously avoided the need to assume the constancy of either U_R or U_W .

Redefining the objectives of cost-of-living indexes, however, also raised an additional issue: How would the strong ordinal revision affect the functions that cost-of-living indexes

served in economics? Significantly, for at least two important applications -- estimates of consumer price change and deflation of consumer data in the national accounts -- the change to a strong ordinal formulation eased acceptance of the neoclassical perspective. Since the 1920s, many influential economists had argued that, in principle, changes in utility functions ought not to affect comparative estimates of national “economic welfare.” The strong ordinal approach fit neatly with this tradition, and thus few economists engaged in empirical macroeconomic analysis recognized the significance of eschewing intertemporal comparisons. In turn, it was easy to overlook the transformations in the basic objectives of neoclassical cost-of-living indexes that had occurred over the previous half-century.

Yet for one of the original applications for constant-utility indexes -- the calculation of proper compensation to ensure constant economic welfare -- the shift to a strong ordinal approach was more consequential. Unfortunately, many economists have not been quick to recognize these implications nor to appreciate the variety of objectives that can be (and have been) ascribed to a constant-utility index. In turn, public discussion of cost-of-living adjustments for wages or government income payments has been deeply muddled, and important conceptual questions obscured.

CARDINAL UTILITY AND COST-OF-LIVING INDEXES, 1880s -- 1930s

That well-known cardinal utility theorists such as Edgeworth or Pigou extended cardinality to their treatment of constant-utility indexes may seem like a *prima facie* conclusion. Nevertheless, exploring that connection will help to illustrate the consequences of adopting a cardinal approach. Along the way, we will consider several other examples of utility-based cost-of-living indexes from this period, including oft-cited authors such as Arthur Bowley, Ragnar Frisch, and Alexander A. Konüs.

Cardinal utility treatments of constant-utility indexes are generally characterized by four features. First, there is an explicit or implied presumption that the “utility” in question is some kind of psychic phenomenon. Second, there is a clear belief that the objective of cost-of-living indexes is to make intertemporal or interlocal comparisons of economic welfare (typically described as “utility” or “satisfaction”). Third, authors recognize that the utility which concerns them, μ_w , may be affected by more than merely a change in relative preferences. Fourth, variations in utility functions (whether U_R or U_W) present operational challenges to what

nevertheless remains a meaningful comparison. Of these four points, all but number two are unique to the cardinal utility approach (the second being shared with the weak ordinal position).

The most systematic treatment of index numbers in the nineteenth century came from Francis Edgeworth in a series of papers written in 1887--1889 for the British Association for the Advancement of Science (Edgeworth 1925 [1887--1889]). Edgeworth, as was his wont, developed an elaborate classification scheme in which “Measurements of Change in Value of Money” were categorized by their objectives and assumptions. A major section dealt with constructing a “standard for deferred payments” which might “afford to the consumer a constant-value-in-use” (pp. 207-208) In the 1880s Edgeworth drew on William Stanley Jevons to characterize “value-in-use” as “final utility”; by the 1920s, he regularly cited Henry Sidgwick as the formal progenitor of this standard and typically described the goal as constant “satisfaction,” perhaps to make the psychological dimensions of the measurement more clear (e.g., Edgeworth 1923, p. 344).

Edgeworth cheerfully acknowledged the uncertainty that must necessarily accompany such “subjective” factors and conceded “that index numbers for the cost of living...have not the objective character of nautical tables or reports of yesterday’s weather” (Edgeworth 1925b, 381-382). Even at the end of his career, though, he remained committed to this psychological basis for intertemporal comparisons of the cost of living. When the American economist Correa M. Walsh criticized Edgeworth’s equanimity in the face of the “indefiniteness and vagueness” that suffused his writings on index numbers (including the use of utility: Walsh 1921; Walsh 1924, esp. 500), Edgeworth defended the scientific character of “mental measurements” and hence the legitimacy of an economic statistic based (at least in principle) upon them (Edgeworth 1925a, 558-559).

For Edgeworth, the psychological nature of utility had one important practical consequence for cost-of-living comparisons: the marginal utility supplied by a collection of goods and services would depend on the surrounding economic context. As he put it in 1887, “where the total wealth of a people has increased, an equal quantity of utility is represented by a larger quantity of wealth.” Accordingly, Edgeworth distinguished two different standards for deferred payments: one in which “National Wealth” might be assumed to be constant and a second in which it might vary (Edgeworth 1925 [1887--1889], pp. 207, 210, cf. 222). In 1894, he approvingly quoted Edward A. Ross’s view that “the total well-being we derive from goods

depends not only on the positive satisfaction experiences in use or consumption,' but also 'on the social satisfactions that flow to us in consequence, the latter largely determined by the relation of our consumption to that of our neighbours.'" As a result, Edgeworth concluded, "in a progressive society," the ongoing increase in productivity "tends to depreciate goods with respect to utility" (Edgeworth 1894, p. 162).

Edgeworth's reasoning provides a classic illustration of how a cardinal interpretation affects a constant-utility index: relative preferences, U_R , might remain the same, but the financial improvement of the larger society means that the same goods now provide less satisfaction, μ_w . In effect, Edgeworth's analysis left the distribution of income as an important part of a cost-of-living calculation: if average prosperity improved over time, then the same bundle of goods would actually define a lower standard of living. Similar concerns would later shape the debate over the constant-utility ideal in the United States until the strong ordinal approach eliminated the possibility of intertemporal comparisons of welfare.

Although Edgeworth was a staunch proponent of the constant-utility index, he did little to advance its practical calculation; his greatest attention went towards the applications of probability to index number theory. The most influential early contribution came instead from Pigou, who necessarily wrestled with price indexes while trying to define the "national dividend" over time.

Pigou's *Wealth and Welfare* (1912) is the canonical example of the cardinal utility approach. (His later treatments in the many editions of *The Economics of Welfare* are more complicated; we will return to them at the end of this essay.) The book opened with Pigou's definition of economic welfare as "the psychic return of *satisfaction*" [emphasis in original] arising from "those parts of the community's net income that enter easily into relation with the measuring rod of money" (Pigou 1912, pp. 4, 3). Any attempt to analyze welfare over time in terms of money must necessarily deal with changes to the value of money, which in turn led Pigou to the basic problem of Chapter III: "to devise, if we can, some measure of price changes, ...the reciprocal of which shall vary in the same sense as the economic welfare of the group of persons whose fortunes we are at the time considering" (p. 41).

He began by making what would subsequently become a fundamental distinction, namely between cases in which "the tastes and temperament of the group remain constant" and those "in which they vary" (p. 41). The former was far more amenable to analysis, and Pigou devoted

several pages to establishing what he felt to be the most practical and illuminating measure (pp. 42-47, esp. 46). Along the way, he implicitly noted a crucial point. If the welfare delivered by a set of goods remained constant (a condition that Pigou believed he had established by assuming constant tastes and temperaments), then consumers could always maintain a constant level of welfare by purchasing the same goods. Of course, there might be less expensive avenues to the same end; consumers might be able to redistribute their expenditures in more efficient ways under a new set of prices. In any event, however, a price index based on the original set of goods (what is today known as a Lapeyres index) would form an upper limit to the change in expenditure needed to maintain the original level of welfare, I_w^0 :

$$I_w^0 \leq \frac{\sum p^1 q^0}{\sum p^0 q^0}$$

where p and q are prices and quantities, respectively. The task which Pigou set himself was to find a lower bound to I_w^0 , potentially by relating it to a price index that utilized current quantities (today known as a Paasche index):

$$\frac{\sum p^1 q^1}{\sum p^0 q^1}$$

In fact, devising a suitable lower bound for a constant-utility index (based on the original level of welfare) proved quite difficult, even under the assumption of constant “tastes and temperament.” Abandon those, and everything seemed to fall apart entirely, for a given set of goods could no longer be guaranteed to deliver the same level of welfare and hence changes in its cost could not provide an upper limit to I_w^0 . Pigou recognized two conditions that might violate his basic assumption: “Some of the representative man’s tastes may change *relatively* to others; or the level of his tastes in general may change absolutely” (p. 48). As these comments suggest, Pigou saw economic welfare, μ_w , as a function of at least three factors: the goods purchased (X), relative preferences (U_R , albeit understood in a cardinal fashion), and the “level of tastes in general”, which we will designate as T . Thus for Pigou:

$$\mu_w = U_W [U_R (X); T]$$

Unfortunately, Pigou had no straightforward way to account for either changes in U_R or changes in T . The latter, in fact, was especially troubling because it seemed unresolvable; as Pigou conceded in a footnote, “The existence of differences in the absolute level of tastes in general

between two points of time or space can never be revealed by statistics” (p. 50, n. 1). Despite the operational inaccessibility of “differences in the absolute level of tastes,” however, Pigou continued to believe that they ought properly to be included in constant-utility index. Accordingly, he regarded his inability to account for changes in “tastes and temperament” as a “failure” (p. 49). A fully accurate constant-utility index would find the minimum cost of maintaining μ_w even amidst changes in “tastes,” relative or absolute.

Pigou’s work established a basic pattern that neoclassical economists would follow when they resumed active work on cost-of-living indexes after the war: (a) the distinction between two cases (changing tastes and constant tastes), (b) the focus on situations with constant tastes, and (c) the attempt to find some means of bounding I_w^0 (given the assumption of constant tastes), typically in relation to both the Laspeyres and Paasche indexes. This project was taken up most persistently by the mathematical economist Arthur Bowley, who led British attempts to produce a working estimate of a neoclassical cost-of-living index during the 1920s (Bowley 1919; 1920; 1928).²

Outside of Britain, a similar dynamic prevailed. The Italian statistician Corrado Gini highlighted the “subjective” nature of utility and the problems that changing utility functions could cause for the practical measurement of constant-utility indexes (Gini 1924, pp. 139-141). Likewise, efforts to bound a constant-utility index under an assumption of constant tastes appeared in other texts -- notably Haberler (1927) and Bortkiewicz (1932) -- so that by the mid-1930s, economists could recognize these projects as comprising a more or less coherent research field (e.g., Staehle 1934, pp. 74-92; Frisch 1936, pp. 11-12).

The most famous contribution to index number theory from the 1920s, Alexander A. Konüs’s 1924 paper in Russian (translated into English in 1939) on “The Problem of the True Index of the Cost of Living,” is a more complicated case. Both Konüs’s clear intention to make intertemporal comparisons and his description of a “standard of living” as the “general state of want-satisfaction” are characteristic of the cardinal utility approach. Not surprisingly, as Henry Schultz later noted, a large portion of Konüs’s paper (namely his attempt to find two standards of living that were “approximately equal”) rested on an implicit assumption of cardinal utility (Schultz 1939, pp. 8, n. 8; Konüs 1939 [1924], pp. 21-29). Like other neoclassical theorists, Konüs also assumed stable utility functions, stipulating that his analysis was “valid only when the habits of consumers, their family compositions, and their environments (exclusive of relative

prices and total expenditures) do not change” (p. 13). With this condition in place, Konüs demonstrated that the Laspeyres and Paasche indexes were upper and lower bounds, respectively, on two different “true” indexes (from the base period and current period), and that a third “true” index -- which defined a standard somewhere between the base period and current standards -- would always lie between the Laspeyres and Paasche boundaries:

$$I_w^0 \leq \frac{\sum p^1 q^0}{\sum p^0 q^0}$$

$$I_w^1 \geq \frac{\sum p^1 q^1}{\sum p^0 q^1}$$

$$\frac{\sum p^1 q^0}{\sum p^0 q^0} > I_w^E > \frac{\sum p^1 q^1}{\sum p^0 q^1} \quad \text{for } U_w^0 > U_w^E > U_w^1$$

In all these respects, Konüs’s paper fell squarely within the cardinal utility tradition. Still, his brief comments about intertemporal comparisons in the absence of stable utility functions potentially suggest a stronger view of noncomparability, one that would later be associated with the weak ordinal revision. Specifically, Konüs warned that “there is no possibility of comparing standards of living in the summer and winter months of any year, since conditions of life differ as between summer and winter” (p. 13). It is unclear whether these remarks referred to the operational limits of his own analysis (which followed “the method of the economists of the so-called mathematical school”), or whether Konüs was stating a more general stricture about the impossibility of welfare comparisons (in principle) when environmental conditions had altered utility functions. The published record is ambiguous, though by the late 1960s, Konüs had clearly adopted the weak ordinal view (Konüs 1968, pp. 94-97, 105).

Konüs aside, it seems clear that many neoclassical economists sided with Edgeworth or Pigou (1912) on this issue -- that is, they viewed different utility functions as providing operational rather than theoretical obstacles to intertemporal comparisons of welfare. The clearest evidence on this point comes from the attempts to find some empirical test that could identify two expenditure levels that provided similar “satisfaction” (μ_w) irrespective of such differences. For example, at the end of Ragnar Frisch’s subsequently famous survey article on index numbers, he developed a cost-of-living comparison based on “monetary flexibility,” which he suggested might be applicable to comparisons across different utility functions (Frisch 1936,

pp. 31-38, esp. 38). In a presentation to U.S. government economists a year later, Frisch pursued this theme more confidently, now couched in terms of “spending elasticity.” (The latter, conceived as the inverse of monetary flexibility, can be loosely defined as a consumer’s willingness to defer part of his or her expenditures to a later date).³ “This elasticity,” Frisch told the Americans, “can be compared from one group to another even though the markets are structurally different and the want constitutions are different. And it may under certain very general conditions be taken as an indicator of well-being, thus solving the problem of computing cost of living indices in these more complicated cases” (Frisch 1937, p. 5). In a similar vein, Hans Staehle (an economist at the International Labour Office in Geneva) likewise sought to identify “an objective criterion” for “determining money incomes which yield equivalent satisfaction” but without imposing an “assumption of identical tastes and milieu” (Staehle 1934; Staehle 1937, p. 205).

Unfortunately, neither of these proposals garnered widespread support. The only theoretically unassailable solution remained bounding a “true” index through standard price indexes such as the Laspeyres or Paasche indexes. Establishing these limits, however, required what Staehle characterized as “the very serious assumption” of stable utility functions; absent that condition, he concluded glumly, “the method of price index numbers must break down altogether” (Staehle 1935, p. 188). Beyond this narrow field economists lacked an operational technique to find what neoclassical theory told them to seek: expenditures that provided equivalent utility, μ_w , in two different situations.

NONCOMPARABILITY AND THE WEAK ORDINAL REVISION, 1933-1950

The viability and necessity of cardinal utility as a fundamental concept for economics had been under question since turn-of-the-century. Yet it was contributions from three British economists (Robbins 1932; Hicks and Allen 1934) that brought these critiques to wider notice in the early 1930s and gave them a clear focus (Schumpeter 1954, pp. 1057-1066; Cooter and Rappoport 1984; Hennipman 1987; Weber 2001). By mid-century, most young economists and the majority of their elders had accepted that utility need only (and, scientifically-speaking, could only) be discussed in ordinal terms and that interpersonal comparisons of utility fell outside of science, and hence outside of economics proper (Backhouse 2003).

For some neoclassical economists, the conceptual upheavals surrounding utility had little effect on price index theory beyond a shift in language. When Ragnar Frisch compiled his 1936 survey, for example, he cobbled together a host of authors who discussed index numbers based upon equal “utility” or “well-being,” but did not worry about the different meanings that they attached to such terms. In a second survey nearly two decades later, he explicitly adopted an ordinal view of utility, but the change still had no effect on his analysis; he continued, for instance, to suggest that equivalent “money flexibility” could indicate equivalent real income even in the face of differing utility functions (Frisch 1954, pp. 409, 421).

Among other neoclassical economists, however, the shift from μ_w to μ_r as the “scientific” version of utility meant that welfare comparisons could now occur only under strict conditions. If utility merely referred to a multi-dimensional surface on a complex utility function, in what sense could you “compare” utility levels when the structure of the underlying function was entirely different? Among early ordinal theorists such as R. G. D. Allen or John Hicks, the answer was simple: you could not; such comparisons were simply meaningless. From being an operational constraint, the assumption of stable utility functions had shifted into a theoretical precondition for the possibility of intertemporal comparisons of welfare, and hence for constant-utility indexes.

Demand Theory, the Ordinal Revolution, and Welfare Comparisons: Allen, Hicks, and Samuelson

R.G.D. Allen provided the first, and characteristically incisive, intervention from an ordinal theorist in the midst of a larger article “On the Marginal Utility of Money and Its Application” (1933). Having constructed a theory of consumer demand based upon the use of indifference curves, Allen turned to the problem of what he (following an earlier book by Frisch) called a “price of living” index. The task of such an index, in Allen’s interpretation, was to find the “change in expenditure” that would leave an individual on the same indifference curve under a new set of prices, and thus as “well off” as he or she was before (Allen 1933, pp. 200-202). What meaning could such a comparison have, however, if tastes or market conditions had changed?

Allen tackled that question most explicitly near the end of his article, when he turned to “geographical” comparisons across locations, “where individual tastes are different and where the sets of consumers' goods quoted are of entirely different composition” (207). Here, Allen

suggested, “it would be possible to extend the notion of the price of living” only if economists could find “a one-one correspondence...between the indifference *loci* [indifference curves] of the two systems. Price lines (and the expenditures which go with them), which were previously associated if they touched the same indifference *locus* in a single diagram, would now be associated if they touched corresponding indifference *loci* in two distinct diagrams” (207). But, Allen argued, it was impossible to establish such a correspondence, and so price of living comparisons could only be made within a stable system of indifference surfaces. “There is, therefore, no theoretical support for the comparison of the price of living in countries or on markets with different modes of living” (208).

In general, Allen repeatedly emphasized the stringent constraints on any price index derived from an ordinal, neoclassical analysis of consumer demand:

The conditions under which such a price of living can be defined must be stressed. *P* [the index] refers necessarily to one individual only, and this individual is faced with various price situations on a given consumers’ goods market. The concept is in no way a collective one, nor can it refer to the same, or different, individuals at different times, or on different markets. In particular the consumers’ goods themselves are a given and invariable set (201).

These constraints are remarkably binding, and indeed they verge on the strong ordinal position. The difference, though, was that Allen believed in the legitimacy of welfare comparisons across time and space; the task of the economic theorist was to determine the conditions under which such comparisons could be justified (just as he had in the case of “geographical comparisons”).

Not surprisingly, Allen found these conditions to be extremely rare. “The practical application of time index numbers to the price of living problem,” he declared, “is confined to the case where tastes, habits and general market conditions vary little, so that changes in relative prices are the only variations of importance” (206). Such stasis was an unusual occurrence, however, since “no matter how short the period of time may be, individual tastes and (more important still) the number, kinds, and qualities of the goods quoted on the market are bound to change” (205). Thus, as he concluded, “there is...a definite but limited connection between the practical price of living index numbers (in their ideal forms) and the theoretical concept of the price of living” (206).

As these excerpts make clear, Allen depicted neoclassical, cost-of-living index theory as having only limited relevance to real-world situations. The justification of price indexes by “price of living” theory could be undercut both by changes in goods (the familiar problem of

quality change) but also by changes in tastes (U_R), which altered the form of indifference curves and made it impossible to establish meaningful correspondence between two situations. Since these changes occurred almost constantly, the constant-utility index appeared more as a theoreticians' toy than a practical objective for statisticians.

Other leading figures in the reformulation of consumer demand theory recognized similar constraints. In a subsequently famous 1940 paper on "The Valuation of Social Income," for example, John R. Hicks declared that "comparisons of economic welfare must proceed under the hypothesis of *constant wants*....If this assumption cannot be granted, the question whether [an individual] is better off in one situation or in the other loses all economic meaning" (Hicks 1940, p. 107). Like Allen, Hicks interpreted "constant wants" to mean a constant set of relative preferences (U_R): "We say that tastes are unchanged between two situations if any preference for one collection of goods over another, which holds in one situation, holds also in the other. This assumption can be expressed by assuming a given utility function or sufficiently by assuming given indifference surfaces" (107). Likewise, working from a slightly different theoretical foundation, Paul Samuelson also concluded that the objective behind index number comparisons ("determining which of two situations is higher up on an individual's preference scale") "admits of a partial answer if certain rigid assumptions are fulfilled." The latter included a restriction "to one individual, whose tastes do not change in the period under consideration" (Samuelson 1947, pp. 146-147).

In light of these requirements, the validity of applying neoclassical theory to real-world situations depended heavily on the likelihood of taste stability during the relevant period. As Hicks described, welfare comparisons "are significant so far, and only so far, as we *judge* the assumption of unchanged tastes to be a tolerable assumption with reference to any particular actual comparison" (Hicks 1940, p. 107). Whether from the cardinal or weak ordinal perspective, economists through the mid-twentieth century treated constant tastes as an essential precondition for the application of neoclassical, cost-of-living index theory. Unfortunately for the advocates of a utility-based approach, many of their colleagues, especially those in the United States, were as (or more) skeptical than R. G. D. Allen about the existence of such stability, and hence about the practical value of neoclassical price index theory as well.

TASTE CHANGE, INSTITUTIONAL ECONOMICS, AND THE REJECTION OF THE CONSTANT-UTILITY IDEAL, 1930-1965

Although Europeans led the development of neoclassical price index theory during the first half of the twentieth century, U.S. economists became the first to apply that theory vigorously to official statistics, a process that began with a major report from the National Bureau of Economic Research (NBER) in 1961 and culminated in 1996 report from a committee led by Michael Boskin of Stanford University (National Bureau of Economic Research 1961; Boskin, Dulberger et al. 1996). Accordingly, the U.S. is a valuable site for illustrating how close attention to the conceptual shifts in the definition of a constant-utility index can explain much about the reception and use of that theory for practical calculations.

Surprisingly, in light of later developments, the U.S. was not a hospitable environment for neoclassical price index theory at mid-century. Indeed, when the 1961 NBER committee first recommended that the U.S. Bureau of Labor Statistics (BLS) use a constant-utility index as a model for the Consumer Price Index, most government officials and their external advisers reacted with skepticism, if not derision. To some degree, that rejection reflected the much weaker tradition of a utility-based approach to price index theory in the United States. To be sure, several American economists in the late nineteenth century had suggested adjusting debts in accordance with the changes in the utility-value of money (e.g., Merriam 1893; Fetter 1895; Adams 1901; Adams 1902). Yet the most influential American text on price indexes prior to the First World War, Correa M. Walsh's *The Measurement of General Exchange Value* (1901), deliberately eschewed subjective concepts of utility in favor of what Walsh deemed a more objective phenomenon: the change in "exchange-value." Likewise, even the leading American neoclassical economist, Irving Fisher, doubted the viability of a statistic constructed on utility. Not only was "the marginal utility of the same goods...different for different persons," it was also "different for the same person at different periods of his life" (Fisher 1911, 220). Like Edgeworth, Fisher agreed that a rise in prosperity might reduce the marginal utility of a fixed set of goods, yet he was highly skeptical that one could measure the extent of this shift; such an objective, he declared, "seems as fatuous a quest as the search for the philosopher's stone." Overall, he concluded, "the 'utility standard' [for measuring the change in the value of money] is...impracticable, even if the theory of such a standard were tenable" (222). Accordingly, Fisher

ignored the constant-utility approach in his *magnum opus* on price indexes, *The Making of Index Numbers* (1922).

The lack of attention to the constant-utility index in the United States (even among neoclassical economists) was reinforced during the interwar years by a heightened cognizance of the dynamic and transient nature of consumer tastes. Given the fatal problems that changing tastes caused for both the cardinal and weak-ordinal versions of neoclassical price index theory, attention to the dynamic qualities of consumer desire necessarily created serious doubts about the practical value of the constant-utility approach, just as had happened for Irving Fisher. Nowhere was this perspective more evident, though, than among American institutional economists.

Taste Change and Institutional Economists

European economists saw differences in tastes as a serious problem primarily when making comparisons that crossed social classes or geographical locations with different customs (e.g., Pigou 1912, pp. 48-49, 51; Keynes 1930, pp. 96-98, 120; Staehle 1934, p. 88; Frisch 1954, p. 420). By contrast, they believed consumer tastes would remain relatively stable over time; Bowley, for example, argued that “an interval of time in which customs and desires have not changed” would correspond to “say, 60 years in one country,” a conclusion that he justified by (uncited) “general statistics of consumption, which show for England that the changes in habit were slight between 1904 and 1927” (Bowley 1928, pp. 223-224).

To my knowledge, no American institutional economists during the interwar period reviewed Bowley’s assertion that “customs and desires” would be largely stable over a sixty-year period, but undoubtedly they would have been bewildered by his claim. Whereas Bowley found stasis in the purchasing habits of British consumers between 1904 and 1926, government statisticians in the U.S. were impressed by the dramatic changes in American consumption over a shorter (but later) time period: 1919 - 1935. Major BLS surveys of urban, working-class families revealed that proportional expenditure on staples such as food and clothing had fallen by 4 percentage points and 6 percentage points, respectively, indicating (in the common interpretation) a significant rise in material standards of living. More importantly, BLS officials described a range of new products that now appeared in family budgets: radios, washing machines, vacuum cleaners, electric and gas refrigerators, light bulbs, and automobiles (U.S. Bureau of Labor Statistics 1940).

Alongside this empirical evidence of changing consumption patterns, the institutional economists who led most American statistical agencies from the 1930s into the 1960s had a theoretical perspective that linked shifts in consumer behavior to the creation and alteration of consumer desires. Impressed by Thorstein Veblen's analysis of consumer culture (1899) and immersed in interwar American psychology (with its attention to instinct, habit, and behavior modification), institutional economists emphasized the malleability of human desires (Rutherford 2000, pp. 295-297). As John M. Clark explained in an early article, economists had once "thought of the self as a sovereign will, in some sense independent of the universe. Men had their wants, and the universe granted or denied their gratification." But "modern psychology" taught that "our wants...are molded by our environment" (Clark 1918, pp. 7-8). In contrast to the stasis of traditional economic analysis, Clark touted the "dynamic nature of consciousness" and the consequentially dynamic qualities of consumer demand (12-16). In contemporary society, Clark found one overwhelming source of transformation in demand: advertising. "In a single business establishment," he remarked dryly, "one department furnishes the desires which the other departments are to satisfy" (8).

Institutional economists' view of consumer psychology had important implications for how they would perceive the consequences of structural changes in the economy -- especially rising productivity and innovation -- on consumer desires and psychological satisfaction. Again, a comparison with the British neoclassical tradition may be helpful. British economists such as Pigou recognized that changes in productivity and novel goods could alter demand. Significantly, though, Pigou viewed these changes as enhancements whereby consumers learned to appreciate novel items and hence improved their ability to derive pleasure from them:

It is true that a cause which augments the supply of some article will often thereby indirectly alter people's taste for that article. But the alteration will practically always be an enhancement. Thus, when machines are sent out on trial, or articles presented in sample-packets, or pictures exhibited free to the public, the popular desire for these objects tends to be augmented (Pigou 1924, p. 73).

Accordingly, Pigou remained convinced that the increased production of goods would (generally speaking) "carry with it an increase in economic welfare, ultimately as well as immediately" (74). For that reason, he was also unconcerned by the hypothetical conundrum that had given Edgeworth pause in his treatment of a constant-utility index, namely that an increase in "National Wealth" or productivity might cause the utility delivered by a given set of goods to decrease. Pigou conceded that "if, at a time when production was becoming easier, the level of

general tastes was falling, it might happen in reality that economic welfare was growing less, though the reciprocal of our index number would necessarily be increasing.” But he found such a coincidence “not probable, at all events between closely adjacent times” (Pigou 1912, p. 50).

By contrast, institutional economists were more skeptical about the ability of rising productivity and innovation to enhance the psychological satisfaction of consumers who had already reached a level of basic subsistence, even though such changes might boost their material comfort. Whereas Pigou had assumed that increased supply of a good would augment desire, any reader of Veblen would recall that, for the leisure class, it was scarcity, not abundance, that drove desire (Veblen 1899). From that vantage point, the increased availability of a specific good could actually reduce the amount of pleasure derived from its consumption, and the satisfaction provided by novelty would be diminished in proportion to its ubiquity. In general, Veblen’s analysis insisted that consumer desires for particular goods (and any satisfaction that they derived from their purchases) would be highly correlated to their own relative position within a local economic and social environment.

Economists in American statistical agencies applied this logic to the analysis of price indexes from the 1930s through the 1950s. For example, in 1938 (perhaps prompted by Ragnar Frisch’s presentation to government economists the previous summer; Frisch 1937), a BLS staff member concluded that for the bureau’s cost-of-living index, “what we really want is the changing cost of purchasing the same economic well-being.” Yet she felt it was “impossible” to create such a measure, in part because “standards change” and thereby altered the well-being delivered by specific goods. As she put it, “a model-T Ford is definitely less satisfactory now, even when new, than it was in 1917.”⁴ The economist Louis Bean from the Department of Agriculture took a similar view. Responding to Andrew Court’s 1939 attempt to use a utility framework to adjust the BLS automobile price index for quality change (the first example of what Court dubbed a “hedonic” price index), Bean complained that Court had ignored how “consumers’ appraisal of values and the nature of their satisfactions tend to change simultaneously with improvements,” so that a physically identical car actually delivered less satisfaction over time (General Motors Corporation 1939, p. 118). A postwar article on quality change from BLS economists took a similar view (Williams and Hoover 1946, p. 361); likewise, a 1951 memo describing the conceptual basis for the CPI reported that “in theory a fixed standard of living is interpreted as ‘constant want satisfaction’ or ‘equal well-being,’” but “in

practice, it virtually defies measurement,” in part because “the kinds of goods that...make up a fixed level [of living] change as new products are placed on the markets and influence the tastes and preferences of the consumer.”⁵

Taste Change and Reactions to the 1961 NBER Report on Federal Price Statistics

These examples suggest that BLS economists were likely to regard a neoclassical, constant-utility index as a dubious ideal for working calculations since it required the assumption of constant consumer tastes as either an operational (cardinal utility) or theoretical (weak ordinal) precondition. As it happened, their skepticism -- and that of most economists concerned with the practical creation of official statistics -- was heightened by the manner in which that proposal was first made.

In 1959, the Office of Statistical Standards in the U.S. Bureau of the Budget commissioned the NBER to review the federal government’s price statistics. The NBER asked University of Chicago economist George Stigler to assemble a committee, and in the spring of 1961, Stigler and his colleagues presented their findings to the Joint Economic Committee of the U.S. Congress (U.S. Congress. Joint Economic Committee 1961).⁶ The report (National Bureau of Economic Research 1961) retrospectively served as a watershed moment in the history of federal price statistics and the history of applications for neoclassical price index theory. In part, this reputation grew from the committee’s four hundred pages of detailed staff research papers, which included a subsequently famous paper on hedonic price indexes by Zvi Griliches. In part, it derived from the committee’s concerted efforts to use the constant-utility ideal as a framework to guide a range of operational decisions for the U.S. Consumer Price Index, including procedures for introducing novel goods, techniques for translating quality changes into price changes, and methods for handling durable goods and housing.

Despite the sophistication of many of its applied analyses, however, the report’s discussion of the theory behind its recommendations was startlingly, even shockingly, simplistic. Although the report repeatedly referred to “welfare” as a basis for cost-of-living measurements, readers heard nothing about the nearly three decades of debate among economists about the nature of “welfare” and its role in empirical economics; indeed, there was no serious discussion about what the committee meant by that term. Furthermore, even though every major, neoclassical treatment of cost-of-living indexes since Konüs (1939 [1924]) -- including George

Stigler's own textbook on price theory (Stigler 1952, pp. 64, 88-91) -- had discussed the need to assume constant consumer tastes in order to estimate a constant-utility index, the report never mentioned these strictures. It is unsurprising, therefore, that many economists regarded the committee's theoretical perspective as naïve and unworkable.

BLS commissioner Ewan Clague spoke for many critics when he reported that "it is a little hard for us to get quite clear in mind what this welfare index would be and how you would measure this concept" (U.S. Congress. Joint Economic Committee 1961, p. 637). Lacking guidance from the committee, the BLS understood the report to be calling for a cardinal comparison of the minimum change in expenditure needed to secure a constant level of consumer satisfaction (μ_w). Understandably, Clague was mystified about how to measure what he called the "cost of living in this psychological sense" (564). In private comments to the Stigler committee, bureau staff members complained that "any attempt to construct an index of this nature would involve such a complex of subjective decisions, assumptions reflecting personal judgments with respect to satisfaction levels, and other abstractions, that the index would lack objectivity...."⁷ The degree to which the bureau interpreted the report in cardinal terms is evident in the summary of the report that one high-level staff member, Sidney Jaffe, offered to union economists on the bureau's Labor Research Advisory Council: "A price index attempts to measure the pure price change in a fixed basket of goods. A cost of living index would attempt to measure the price or cost of equivalent levels of satisfaction. The major difference between a price index and a cost of living index is the fact that a cost of living index would take into account change of preference patterns by consumers."⁸

The report's emphasis on neoclassical price index theory did find some enthusiastic supporters (e.g., U.S. Congress. Joint Economic Committee 1961, pp. 742-745). Nonetheless, most economists and statisticians who were closely involved with official statistics either rejected the constant-utility ideal outright or were unsure how it could be implemented in operational terms.⁹ The University of Chicago's economist-turned-senator, Paul Douglas, provides a good example. Douglas, who was a member of the Joint Economic Committee, unfortunately had to miss most of the hearings on the day that Stigler and his colleagues presented their findings. Several days later, however, Douglas expressed his skepticism about the NBER committee's proposed constant-utility index in a brief monologue:

It is bad enough to try to do this [compare utilities] with external objects, common units of utility, but when you have to take into account the fact that people themselves change, it often becomes more difficult. It has been said that one never steps in the same stream twice. It is also true that the same person never steps in the same stream twice. The stream is not only different but the person is different. In 1900 the average education of an American was I think the sixth grade. Today the average education of an American is high school. Now, does this make for greater enjoyment? Do you get more enjoyment? Does a college graduate get more enjoyment out of a hotdog than a sixth grader? On the other hand, he may get more enjoyment from a painting--and probably does....How can you compare the decrease in satisfaction with the hotdog with the increase in satisfaction on the painting? How can you measure either one to begin with, or how can you equate them after you have measured them, which is, I think, practically impossible... (639).

For Douglas, who had spent his own time wrestling with price indexes back in the 1920s (Douglas 1930), the dynamic quality of utility functions vitiated efforts to construct a constant-utility index.

Following a recommendation of the Joint Economic Committee, the BLS assigned several staff members to pursue further research on the concept of a constant-utility index, but they continued to stumble over the problem of taste change. In 1962, for example, a union economist balked at the use of “hedonic” price indexes to account for quality change (a method recommended by the NBER committee) because such analysis “does not take into account the value of an item as a status symbol.” The BLS concurred with the spirit of this objection, which followed the same logic as Louis Bean’s 1939 critique of hedonic methods for adjusting automobile prices (see above).¹⁰ In a 1965 report on concepts for the CPI, BLS staff members reiterated that “the calculation of [a constant-utility] index across periods of time which experience changes in tastes and in real income (a determinant of tastes and preferences) conflicts with one of the basic assumptions--that tastes and preferences are fixed.”¹¹ Of course, those constraints meant that a constant-utility index would only be valid for an economically and culturally static society, a vision that did not match typical perceptions of the United States during the 1960s. BLS economist Sidney Jaffe felt that utility functions were unlikely to remain stable “over long periods” (such as a decade) and that changes in tastes therefore presented major obstacles to implementing a constant-utility index.¹² Ewan Clague’s successor as BLS commissioner, Arthur Ross, voiced the same doubts in his testimony before the Joint Economic Committee in the spring of 1966. Ross noted that “some very important restrictions” governed the calculation of a constant-utility index, including the requirement that the consumer’s “environment and tastes” remained unchanged. “Unfortunately,” he continued dryly in a

parenthetical commentary, “the persons responsible for calculating actual price index numbers cannot bypass such situations” (U.S. Congress. Joint Economic Committee 1966, p. 72).

TRANSFORMING THE CONSTANT-UTILITY INDEX: THE STRONG ORDINAL REVISION

By the mid-1960s, the status of constant-utility index remained much as it had several decades before: attractive because of its coherence with neoclassical consumer demand theory, but impractical as a working ideal for actual statistics because of its reliance on clearly false presuppositions about the stability of consumer utility functions. Within a few short years, however, the BLS would aggressively but unobtrusively adopt the constant-utility index as the basic theoretical model for the Consumer Price Index. Meanwhile, theoretical and empirical research into the construction of working versions of neoclassical cost-of-living indexes would enjoy a small renaissance. What accounts for this transition?

To some degree, it reflected the larger shift in the economics profession in the United States, as institutional economics faded after the Second World War and neoclassical microeconomic theory became the basic conceptual framework for most American economists (Morgan and Rutherford 1998). That shift was sharper in the BLS, where a wave of retirements hit the agency in the mid-1960s and created a large outflow of older, institutional economists who were replaced by younger scholars trained in the increasingly ubiquitous neoclassical tradition. The staff turnover had a particularly important effect on the bureau’s Division of Price and Index Number Research, which was created in the wake of the 1961 NBER report and would (under the guidance of a host of new, younger staff members) lead the agency in its acceptance of the constant-utility ideal.¹³

Nonetheless, the growing dominance of neoclassical economists does not, in itself, provide a sufficient explanation. Prior to the 1970s, neoclassical economists had always worried about how changes in utility functions might render the calculation of a constant-utility index impossible or meaningless. The widespread acceptance of the constant-utility ideal as a practical model for actual statistics therefore required overcoming this fatal weakness. The solution emerged from surprising sources: a reconsideration of how an ordinal approach to utility affected temporal comparisons of welfare and a related redefinition of the constant-utility index itself.

Formulating the Strong Ordinal Revision

The first steps in this direction came from R. G. D. Allen, who redefined the neoclassical cost-of-living index into a sharply narrower measure which avoided the hazards posed by changes in utility functions but who did not take up the larger reconsideration of welfare comparisons in general. Recall that in 1933, Allen had insisted on the need for stringent assumptions about the stability of consumer tastes and had been the British economist who worried most about the viability of those assumptions over time. When Allen (1949) returned to the topic, however, these concerns had vanished. Though Allen described his paper as a “synthesis” (197) of existing views (giving particular attention to Hicks 1940, and Samuelson 1947), he in fact (consciously or not) offered a new definition of cost-of-living indexes.

Previous discussions of cost-of-living indexes had always assumed that the ultimate goal was to calculate the cost of achieving the same level of welfare in two different, real situations (geographical locations or time periods). The task of economic theory had been to elucidate the conceptual basis for such a comparison and describe the conditions under which it might be operationally possible and theoretically valid. By contrast, Allen described the task as calculating the change in expenditures needed to keep a given consumer (selected from a real situation) at the same “indifference level” when facing a set of “hypothetical” prices (197; “hypothetical” is Allen’s term). Though these hypothetical prices might be (and usually were) drawn from a later time period, the cost-of-living index did not involve that subsequent time period directly; it made no pretense to calculating what expenditures might be needed to bring equal “satisfaction” or “welfare” to that individual at a later date.

As Allen put it, “there is no reference to the behavior of the consumer in situation 2; the index is independent of the preference map of situation 2 and of the purchases...made in this situation” (197). Such independence implied that Allen’s previous worries changing consumer tastes were irrelevant to this new version of a cost-of-living index. For those who missed the point, Allen returned to it again more bluntly one page later: “We get no further by making the additional assumption that the preference map remains unchanged from situation 1 to situation 2....This hypothesis, indeed, has no relevance to price index numbers which (by definition) relate to one indifference map alone” (198). In Allen’s hands, the neoclassical cost-of-living index no longer involved a comparison of real consumer welfare across time periods or places; it

was a hypothetical calculation in which an existing consumer (fixed in time and space) suddenly faced a new set of prices (which, implicitly, were not allowed to affect his tastes).

Allen did not derive his redefinition from a deeper consideration of how ordinal utility affected constant-utility indexes (cf. 198-199) or welfare comparisons in general. Indeed, when he turned his attention from “price indexes” to “volume indexes” (quantity indexes), he assumed that the latter *did* involve a direct comparison between two different time periods. Moreover, he returned to the weak-ordinal view of such comparisons: they required constant utility functions. Thus, in contrast to the independence of price indexes from changes in utility functions, Allen concluded that “the [volume] index has no meaning unless we make the assumption that the preference map is the same in the two situations” (199). As this differential analysis implies, Allen effectively broke the symmetry often ascribed to price indexes and quantity indexes: under his definitions, deflating a change in expenditures through a price index would not necessarily yield a matched quantity index; the former was a hypothetical comparison while the latter involved real situations and could only be applied under limited circumstances (cf. 202).

Perhaps because Allen (for whatever reasons) failed to highlight the novelty of his definition for price indexes and did not link that redefinition to theoretical concerns, his article appears to have had little effect on American approaches to price index theory. A more influential analysis had to wait until the late-1960s, when Franklin M. Fisher and Karl Shell (then both at the Massachusetts Institute of Technology) began to work on the neoclassical foundations of price index theory courtesy of support from the Federal Reserve Board Committee on Prices and Price Measurement. The fruit of this research, a paper on “Taste and Quality Change in the Pure Theory of the True Cost-of-Living Index,” appeared as Fisher and Shell (1968) and was reprinted along with a second essay on a national output deflator in an oft-cited collection (Fisher and Shell 1972).

Unlike Allen, Fisher and Shell began their analysis by highlighting the “rather uncomfortable treatment” of “taste and quality changes” in existing theory (Fisher and Shell 1968, p. 97). “The assumption of an unchanging indifference map,” they explained, “even defined over nonexistent goods is apparently crucial for a theory which is often erroneously thought to answer the question: How much would it cost in today’s prices to make the consumer just as well off as he was yesterday?” Such a formulation was both common (they cited the

NBER report as one example) and problematic, since “taste changes do occur.” Behind this dilemma lay what they called “misinterpretation” of the “true cost-of-living index” (97-98).

Correcting the problem required reconsidering the viability of welfare comparisons in light of an ordinal conception of utility. First, Franklin and Shell highlighted an issue recognized by weak ordinal theorists: “What is meant by [the phrase] ‘just as well off as he was yesterday’ if the indifference map has shifted?” (98). Like the weak ordinalists, they suggested that such a question was meaningless. However, they further insisted that the “same difficulty appears even if tastes do not change....In both periods, the man’s utility function is determined only up to a monotonic transformation; how can we possibly know whether the level of true utility (whatever that may mean) corresponding to a given indifference curve is the same in both periods?” (99). In their view, intertemporal comparisons were analogous to the interpersonal comparisons criticized by early ordinalists: “One never steps into the same river twice and the comparison between a man’s utility now and his utility yesterday stands on precisely the same lack of footing as the comparison of the utilities of two different men.” Both intertemporal and interpersonal welfare comparisons were “completely without any operational content whatsoever” (99). In essence, Paul Douglas had been right: the cost-of-living index described by the 1961 NBER report was indeed “impossible.”

In place of the unsatisfactory, standard definition, Fisher and Shell offered an alternative that (independently) echoed Allen’s formulation: “‘Given an indifference map, we compare two *hypothetical* situations, A and B. We ask how much income the consumer in B would require to make him just indifferent between facing B’s prices and facing A’s prices with a stated income.’ Note that the question of whether the consumer has the same utility in A as in B never arises” (99).

Even under this definition, taste change still created certain difficulties for economists. Notably, one had to choose which “indifference map” to use for a calculation: that from A, from B, or from somewhere else entirely. But these issues proved more amenable to theoretical analysis (e.g., Diewert 1976), and most importantly, they were no longer fatal. In the worst scenario, economists would have two equally valid indexes for a temporal comparison -- one using current tastes and one using base period tastes -- but under no circumstances would taste change make it impossible to calculate an index (theoretically or operationally). The problems posed by taste change had been tamed.

During the 1970s and 1980s, most leading neoclassical theorists began describing cost-of-living indexes in terms compatible with the strong (rather than weak) ordinal revision (e.g., Samuelson and Swamy 1974, pp. 567, 585-587; Allen 1975, pp. 265-266; Diewert 1983, pp. 167-168; Pollak 1989, p. 6). Other economists quickly realized that Fisher and Shell's analysis had implications beyond price indexes *per se*. In 1979, Amartya Sen restated their arguments in more formal and general terms during a survey of "The Welfare Basis of Real Income Comparisons," in which he also noted that even a cardinal utility framework did not necessarily imply the possibility of intertemporal comparisons of welfare (Sen 1979, pp. 12-14). Still, though scrupulous price index theorists adopted the strong ordinal formulation, the extent to which they explained or justified that definition varied, and Fisher and Shell's (1968, 1972) discussion remained the most detailed.

Attitudes within the BLS followed a similar trajectory. In 1967, BLS economist Thomas Gavett argued that changing tastes *ought* to be excluded from price measurements, even when based on "utility or consumer satisfaction." One year later, in a paper on "A Constant Utility Cost of Living Index," staff member Joel Popkin (who would later lead the BLS Office of Prices and Living Conditions) offered a strong ordinal definition of the concept.¹⁴ Finally, a 1974 article by BLS economist Robert Gillingham, which subsequently became the bureau's most commonly-cited description of the theoretical framework for the Consumer Price Index, firmly established the strong ordinal revision as the working model for BLS calculations (Gillingham 1974, pp. 246-247).

The bureau's rapid assimilation of the strong ordinal perspective is partially explained by the presence of Robert A. Pollak, who completed his dissertation under Franklin Fisher in 1964. Pollak worked briefly for the BLS from 1968-1969 and then served as a consultant to the bureau over the next decade while a faculty member at the University of Pennsylvania. During that time, Pollak gave a series of seminars on price index theory at the BLS which formed the basis for published articles and a later book (Pollak 1989), which is now the standard theoretical reference for neoclassical cost-of-living indexes. Pollak's influence was extensive; Gillingham, for example, described three of Pollak's essays (two of which were unpublished seminar papers) as the "foundations" for his 1974 article (Gillingham 1974, p. 246).

By the 1990s, the strong ordinal version of the neoclassical constant-utility index had developed extensive support both within American academic circles and within the BLS. Thus

when the 1996 Boskin commission recommended that the bureau adopt the constant-utility index as the official objective for the Consumer Price Index, the bureau could rightly claim that it had already been working within that framework for several decades (Boskin, Dulberger et al. 1996; Greenlees 2001). Theory and practice had merged, and the neoclassical cost-of-living index, long confined to academic papers alone, now had a home in official statistics.

PRICE CHANGE, THE STRONG ORDINAL REVISION, AND MACROECONOMIC ANALYSIS

Of course, finding that home for neoclassical price index theory had come at a cost: redefining the objectives of a constant-utility index. No longer would economists seek to compare the cost of obtaining a fixed level of satisfaction or welfare in two time periods or locations; instead, they calculated the change in expenditures needed to keep a given consumer (fixed in time and space) on the same indifference curve when faced with a different set of prices. How did this shift in objectives affect the use of cost-of-living indexes? To answer that question, we need to return to the early history of the constant-utility ideal.

Price Change and the Cost-of-Living Index

When Edgeworth had explored various means of measuring the “change in the value of money” for the British Association for the Advancement of Science in the 1880s, he developed the constant-utility index as a “standard for deferred payments” which would allow the adjustment of payments in long-term contracts to provide “constant value-in-use” to the recipient (Edgeworth 1925 [1887--1889], p. 208). Prices were obviously extremely important to such a standard, but Edgeworth recognized that other kinds of shifts could also alter the amount of money needed to sustain constant value-in-use, hence his concern with changes in National Wealth or productivity.

By contrast, for the purposes of what we might anachronistically call monetary policy, Edgeworth offered a different conceptual model (one which he drew from William Stanley Jevons): a measure of average price change intended to reveal the effect of “‘causes which operate upon all goods whatever’, or at least upon a considerable group of goods” (Edgeworth 1925 [1887--1889], p. 233). In practice, that meant looking for the effects of changes in the money supply. In Edgeworth’s depiction, such a calculation did not directly involve subjective

valuations of goods; it was a stochastic measure of price changes rather than an attempt to quantify the shifting cost of “constant value-in-use”. Its theoretical foundations lay in the science of “probabilities” rather than neoclassical microeconomics.

Later economists often blurred or ignored the distinctions that Edgeworth drew between a constant-utility index and a measure of price changes, in part perhaps because Edgeworth’s stochastic approach to index numbers came under sharp criticism and hence disappeared as a popular alternative (Aldrich 1992, pp. esp. 679-684). By the late 1920s, many of Edgeworth’s younger colleagues regarded the constant-utility index primarily as a gauge of how price changes might affect consumers. These attitudes manifested themselves in subtle but telling ways. Bowley, for example, described a cost-of-living index as answering the question, “What change in expenditure is necessary *after a change of prices* to obtain the same satisfaction as before?” (Bowley 1928, p. 233; emphasis added). A similar emphasis on price change appeared in definitions from Allen (1933, 200) and Konüs (1939 [1924], 12) among others, and it partially explains Ragnar Frisch’s insistence on calling these statistics “price of living” indexes rather than cost-of-living indexes (e.g., Frisch 1954, p. 407).

Despite this attention to price change, early authors remained committed to finding the changing cost of reaching the same level of utility in two actual situations (time periods or locations), hence their concern about the consequences of differences in utility functions or consumer tastes. But it was surely just a small step to Allen’s (1949) redefinition, to a subtle but crucial conclusion that if the effects of price changes (and price changes alone) were the real variables of interest, then there was no need to worry about changes in utility functions at all. From this perspective, the advent of the strong ordinal version of the constant-utility index -- which compared price systems from a fixed perspective rather than comparing two real situations -- represented a resolution towards which many neoclassical price index theorists had long (though seemingly unconsciously) been heading. The final link in the history of neoclassical price index theory during the twentieth century, therefore, requires us to consider the reasons for that trend.

Empirical Macroeconomics and Measures of Price Change

Economists’ attention to price changes (and their corresponding desire to ignore shifts in utility functions within cost-of-living indexes) had its roots in practical concerns, especially concerns

arising from empirical macroeconomics. One subsequently important function was the deflation of consumer expenditures (and, through mid-century, the deflation of output for consumer use as well). This application connected cost-of-living index theory with a long-time preoccupation of economists: calculating the real change in national income.

Pigou provides an excellent example of how attention to national income could affect economists' attitudes towards changes in utility functions. In 1920, Pigou published an almost completely rewritten version of *Wealth and Welfare* (1912), now titled *The Economics of Welfare*, which was revised in 1924 and subsequently went through multiple editions. As in the 1912 book, Pigou's proposed technique for measuring "changes in the size of the national dividend" was only applicable to "a community with constant tastes" (Pigou 1924, p. 52). Yet Pigou's attitude towards taste change and its role in evaluating the national dividend had undergone a radical transformation.¹⁵

Recall that in 1912, Pigou had insisted that a measure of economic welfare ought to account for differences in absolute and relative tastes, though he was often unsure how to accomplish that task (cf. Pigou 1912, pp. 48-50). By 1920, however, he had begun to qualify that judgment (Pigou 1920, pp. 70-72), and by 1924 he had abandoned it entirely:

The dividend is an objective thing, consisting in any period of such and such a collection of goods and services....Since it is an objective thing, we should naturally wish, if we were able, to define changes in the size of it by reference to some objective physical unit, and without any regard to people's attitude of mind towards the several items contained in it. I do not mean that changes in public tastes would be thought of as incapable of affecting the size of the national dividend. They are obviously capable of affecting it by causing changes in the objective constituents of the dividend. I mean that, *given those objective constituents*, the size of the dividend should depend on them alone, and not at all on the state of people's tastes. This is the point of view which everybody intuitively wishes to take. (Pigou 1924, p. 44)

To paraphrase Pigou, changes in tastes might indirectly affect the national dividend (by altering demand and hence production and consumption) but they should not directly affect national dividend calculations (for example, by reducing the value of a given collection of goods because it was less favored under new consumer tastes). For Pigou in 1924, therefore, the adoption of constant tastes when calculating changes in the national dividend was not an assumption required to construct an imperfect operational measure (as it had been in 1912); it was a stipulation about the kinds of factors that ought to affect those calculations.

Pigou recognized that his exclusion of taste change could create interpretive problems: conceivably a comparison might show a decrease in the national dividend when using period 1 tastes but an increase when using period 2 tastes. He concluded that “the only escape from this is to admit that, in these circumstances, there is no meaning in speaking of an increase or decrease in the national dividend in an absolute sense.” There were two different measurements, from two different perspectives, “and there is nothing more to say” (46).

Pigou’s ideal measure is obviously compatible with the strong ordinalist version of a constant-utility index. Note, however, that whereas strong ordinalists such as Fisher and Shell denied the theoretical possibility of intertemporal comparisons of welfare, Pigou’s position reflected his view of what “everybody intuitively” wanted from national dividend calculations. The two positions thus derived from different foundations; but nevertheless, once the constant-utility index was reformulated on strong ordinalist terms, it fit neatly with Pigou’s goals.

A similar compatibility would develop between the strong ordinal cost-of-living index and the tradition of national accounts that emerged in the U.S. after the Second World War. Department of Commerce economists (led by Milton Gilbert) argued staunchly that national income calculations did not encompass “comparisons of economic welfare,” which would necessitate accounting for changes in “wants” or “tastes” (Gilbert, Jaszi et. al. 1948, p. 189; cf. Perlman 1987). Not surprisingly, that stance led Gilbert to criticize utility-based price indexes, which he interpreted in cardinal terms (Gilbert 1961). The strong ordinal revision, however, made constant-utility indexes compatible with the Gilbert’s view and hence with the approached engrained in the Department of Commerce. Moreover, it had a similar effect on another longstanding application for price indexes: measuring aggregate consumer inflation.

The BLS had long treated its “Cost-of-Living Index” (as the CPI was called from its inception in 1919 through 1945) as a measure of retail price change for urban, working-class families. Indeed, with the help of Wesley Mitchell, it had justified the use of a fixed market basket for the index on precisely these grounds: with the basket fixed, the only variables in the index were prices (Banzhaf 2001; Stapleford 2009, pp. 110-113). When labor unions criticized the index during the Second World War because it did not reflect all of the increased expenses faced by workers (such as commuting costs to new, war production facilities), the bureau emphasized the index’s isolated focus on changes in retail prices and eventually (after the war) agreed to rename it the Consumer Price Index. Though staff members (and other experts, cf.

Mills, Bakke et al. 1943, p. 402) conceded that the index's narrow scope limited its suitability for assessing wartime wage adjustments, they noted its utility for other purposes, notably gauging the success of the government's wartime price control program (Stapleford 2009, pp. 208-211).

The ties between the index and the U.S. Office of Price Administration during the war foreshadowed the prominence that the CPI would increasingly achieve in macroeconomic analysis and public assessment of the economy. As the federal government took greater responsibility for managing the national economy in the wake of the Second World War (often in a Keynesian framework), consumer inflation developed into a potent political topic in the United States and a crucial variable for debates over monetary and fiscal policy (cf. Collins 2000; Jacobs 2005, pp. 221-261). Naturally, that prominence drove attention to concepts and methods behind the CPI: the 1961 NBER review, for example, was prompted by previous complaints before the Joint Economic Committee that U.S. price indexes were overestimating inflation by failing to account for consumer substitution and by undervaluing quality improvements (cf. U.S. Congress. Joint Economic Committee 1959, p. 109). The NBER committee suggested that a constant-utility approach could overcome these alleged weaknesses, but that suggestion (interpreted in cardinal utility terms) also introduced a host of other factors into index number calculation, including changes in consumer tastes. The strong ordinal revision therefore not only made it possible to calculate a constant-utility index even amidst changes in utility functions; it also apparently brought that concept into alignment with what the bureau had been claiming as its objective all along: the calculation of "the change in the cost of living which is the result of changes in market prices," as Ewan Clague put it in 1961 (U.S. Congress. Joint Economic Committee 1961, p. 587).

In truth, the strong ordinal cost-of-living index does not provide exactly what Pigou, Gilbert, or even Clague had sought; indeed, interpreting its economic significance requires careful thought. For example, it is tempting (but misleading) to suggest that, when used with base period prices and current prices, such an index answers the question, "How would consumer expenditures have changed between A and B if prices had changed as they actually did, but consumers themselves (tastes, real income, etc.) had remained the same?" This formulation makes it easy to assume that we are running a historical thought-experiment: hold consumers constant, let the rest of the economy follow its actual course, and see how expenditures change. In this way, we can isolate the changes in expenditures driven by non-consumer, macroeconomic factors (monetary policy, productivity, etc.) from those driven by changes in real income or

changes in utility functions. Such a thought experiment would appear to have many applications. It purports, for example, to isolate changes in expenditures that are “forced” on consumers (by factors beyond their control) from those which derive from shifts in consumer desires or real income. Likewise, it might seem the most suitable measure of consumer inflation for macroeconomic analysis of the sort advocated by Pigou or Gilbert.

Unfortunately, this interpretation of cost-of-living index theory is erroneous; a strong ordinal cost-of-living index does not offer such a thought experiment. The problem is that the market prices used in the index are not independent of shifts in real income or utility functions that occurred between A and B. On the contrary, changes in utility functions will affect demand, and hence affect prices. Accordingly, every strong ordinal cost-of-living index, even though it postulates only a single utility function, may be indirectly affected by changes in utility functions (or real income) insofar as those changes have altered the market prices being used in the index. Likewise, there remains the real possibility (suggested by Veblen) that price changes may in turn affect utility functions. In short, there is no way of running the thought experiment described above, no way of cleanly separating price changes (and thus changes in expenditures) that are due to altered utility functions from those price changes (or changes in expenditures) that are due to other factors such as monetary policy or the supply of consumer goods.

What the strong ordinal index can tell us is the answer to the original question: how a rational consumer (fixed in time and space) would have altered his or her expenditures if he had faced B’s prices rather than A’s prices. This question is clear and answerable. However, the relationship between it and the interpretation of economic changes from A to B is more obscure.

Despite these limitations of the strong ordinal cost-of-living index, it was undoubtedly closer to the objectives that Pigou or Gilbert had defined than was the original cardinal utility version. Likewise, despite the important operational differences between “constant goods” market basket pursued by the BLS through early 1960s and the strong ordinal version, there was nonetheless a common emphasis on isolating price changes. In that respect, it was easy to envision the strong ordinal cost-of-living index as a more sophisticated version of what the BLS had been pursuing for many years -- the calculation of a price index for consumers -- but now made coherent with neoclassical, microeconomic theory.

THE HISTORY OF THE CONSTANT-UTILITY INDEX IN RETROSPECT

As we have seen, the history of the neoclassical cost-of-living index is a history of heterogeneous projects. Edgeworth or Pigou (1912) set themselves a very different task than Fisher and Shell (1968), or even Allen (1933). On the level of economic theory, the major transitions in that history were driven by successive reconsiderations of the concept of utility and how welfare could (or could not) be compared over time and space, even for a single individual. From another perspective, however, the history of the constant-utility index is a tale of its realignment from one set of applications to another, from Edgeworth's advocacy of the constant-utility ideal as a "standard for deferred payments" to its later enshrinement as a mechanism for measuring aggregate consumer inflation and deflating consumer expenditures for national accounts. These two histories did not move in lockstep: Pigou, for example, excluded changes in utility functions from his ideal calculations without having critiqued the possibility of intertemporal comparisons of welfare, much less adopted a concept of ordinal utility.

Only in the 1970s did these two histories completely converge. This recent union begs a question, though: How suitable is the strong ordinal constant-utility index for supplying what Edgeworth called a standard of "constant value-in-use"? Even as the CPI has become a basic tool for empirical macroeconomics, it has also served as an (increasingly powerful) mechanism for adjusting payments and income-based parameters in forms such as cost-of-living escalator clauses in union contracts, U.S. Treasury Inflation-Protected Securities (TIPS), and the widespread indexation of a whole host of federal programs and benefits, including tax brackets, the official poverty threshold, and Social Security payments. How relevant is the strong ordinal cost-of-living index, now the official basis for the CPI, to these tasks?

Answering that question requires deciding how much one wishes changes in utility functions to affect specific adjustments. Would a computer with the same physical and performance characteristics as a mid-range, 1995 desktop bring the same satisfaction to a consumer in 2009 as it had fourteen years earlier? If not, should tax brackets be adjusted by changing cost to purchase a (hypothetical) machine with those 1995 characteristics (the strong ordinal approach), or should it track the average cost of a mid-range desktop (irrespective of overall improvements in speed and storage capacity), a calculation which would show a much smaller decline? Or again, would a patient be equally satisfied receiving 1985-style health care in 2009 as he or she would have been in 1985? If not, does that imply that Social Security payments should also be adjusted for changes in patient expectations arising from overall improvement in

treatments and outcomes? What constitutes being equally “well off” (to borrow Frisch’s 1936 phrase) in an environment of rapid technological innovation?

Naturally, those questions cannot be decided by economic theory. They are, as Edgeworth concluded in his own analysis, topics for “*la haute politique*” (Edgeworth 1925 [1887--1889], p. 222). Unfortunately, the tangled history of neoclassical cost-of-living indexes has tended obscure, rather than illuminate, such issues. The distinctions between the three major formulations of the constant-utility index are important, but they are also subtle. Quick readings of earlier papers with an eye towards mathematical theorems rather than conceptual foundations are apt to perceive continuity rather than change. Bowley’s (1928) definition -- taken out of context -- looks much like Allen (1933) and not far from Pollack (1989). Moreover, despite Fisher and Shell (1968), economists have been liable to slip into older formulations, not recognizing that these definitions can imply much more than contemporary authors may have in mind. Thus, even though the 1996 CPI review committee led by Michael Boskin based its analysis on theoretical papers from the strong ordinal tradition, the committee often discussed its ideal index in much looser terms, stating (for example) that a cost-of-living index tells us “how much would we need to increase (or decrease) initial (period 1) expenditure in order to make the consumer as well off as in the subsequent period (period 2)” (Boskin, Dulberger et al. 1996, pp. 22-23).

That definition, of course, implies a broad scope that the theory cannot and does not support. Such slippage tends to muddle public (and expert) discussion about the methods behind price index calculations and their suitability for various purposes. To the extent that disambiguating the multiple meanings that have been ascribed to a constant-utility index can begin to clarify possible interpretations and their implications, then a history of neoclassical price index theory can have more than historical value.

TABLE 1: THE THREE MAJOR VERSIONS OF THE CONSTANT-UTILITY INDEX

	Measures change in expenditures needed to...	Intertemporal comparisons of welfare?	Changes in utility functions create...	Exemplars
<i>Cardinal</i>	Maintain same level of cardinal utility, μ_w	Yes	Operational problems	Edgeworth (1925 [1887-9]) Pigou (1912)
<i>Weak Ordinal</i>	Remain on same indifference curve	Yes	Theoretical problems	Allen (1933) Hicks (1940)
<i>Strong Ordinal</i>	Remain on same indifference curve	No	No fundamental obstacles	Fisher & Shell (1968) Pollak (1989)

TABLE 1. Not all essays on neoclassical price index theory fit neatly into one of these categories, both because authors were ambiguous or perhaps inconsistent and because these categories do not exhaust all logical possibilities. Thus Konüs (1939 [1924]) is primarily an example of the cardinal approach but may be read as advocating the kind of noncomparability that is characteristic of the weak ordinal view.

NOTES

¹ As we will see, after the work of A. A. Konüs (1939 [1924]), neoclassical economists recognized that any cost-of-living comparison had two possible “true” indexes: one that maintained the consumer’s original, base-period level of utility (or, in an alternate formulation, kept the consumer on the base-period indifference curve) and a second that maintained the current level of utility (or indifference curve). The strong ordinal formulation thus extended this division to a parallel choice between the utility function of the base period or the utility function of the current period.

² I am grateful to Bert Balk for pointing out that Bowley’s attention to Laspeyres and Paasche indexes dates back at least to 1899, when he suggested a geometric mean of the two (Fisher’s ideal index) as a measure of the purchasing power of money. Bowley later abandoned this view, however (Balk 2009, p. 16).

³ In Frisch’s terms, S represents a potential transfer of expenditure to a later period and e would be the consumer’s expenditure absent such a transfer. Then, “the spending elasticity, b , expresses the ratio of a small percentage transfer of expenditure, S/e , (against the wishes of the consumer) from one consumption period to another where prices and expenditure otherwise would have been unchanged, and the largest relative premium, s/S , which the person or family is willing to pay to avoid such a transfer”:

$$b = \frac{2S/e}{s/S}$$

(Frisch multiplied the ratio by 2 merely for “convenience in the subsequent computations”). For details, see Frisch 1937, pp. 12-21.

⁴ Ragnar Frisch, “Methods of Measuring the Relative Cost of Living,” 27 July 1937, Consumption, box 410, Subject Files of the Office of the Commissioner, 1934-55, Records of the U.S. Bureau of Labor Statistics (BLS). Miss Rice [first name unknown], “Retail prices in the cost of living index: Their use and significance,” January 1938, as excerpted in Laura Mae Webb and Ethel D. Hoover, “The Consumer Price Index: Issues of Concept and Practice”, 1 October 1965, Div of Prices & COL, 1965-1968, box 3, RCL-PPF 1950-1975, Records of the U.S. Bureau of Labor Statistics (BLS).

⁵ U.S. Bureau of Labor Statistics, “Uses and Limitations of Consumers’ Price Index” (draft), p. 1, 18 January 1951, folder 13, box 43, Papers of Katherine P. Ellickson, Walter Reuther Library. For more on the reception of hedonic price indexes among American economists and its connection to concerns about changing tastes, see Stapleford 2011.

⁶ Other members included Dorothy Brady, Irving Kravis, Philip McCarthy, Albert Rees, Richard Ruggles, and Boris Swerling.

⁷ U.S. Bureau of Labor Statistics, “Comments of the Bureau of Labor Statistics on the Conclusions and Recommendations Contained in the Preliminary Report of the Price Statistics Review Committee,” November 1960, Stigler Committee, box 217, Statistical Records Relating to Particular Subject Areas, 1940-1968, Records of the U.S. Office of Management and Budget (OMB).

⁸ Minutes of the Labor Research Advisory Council [LRAC], Seminar on prices, p. 4, 21 March 1961, box 2, LRAC, Records of the BLS.

⁹ For examples from the committee hearings, see U.S. Congress. Joint Economic Committee 1961, pp. 638-640, 669, 674. Statisticians and economists in the Office of Statistical Standards within the U.S. Bureau of the Budget expressed similar concerns about the practical implementation of the constant-utility ideal. See Raymond T. Bowman to George Stigler, 8 November 1960, Stigler Committee, box 217, Statistical Records Relating to Particular Subject Areas, 1940-1968, Records of the OMB; Thomas Mosiman to Bowman, 18 October 1960, *ibid.*; Stanley Lebergott to Bowman, 21 October 1960, *ibid.*; Mosiman to Bowman, 25 October 1960, *ibid.*

¹⁰ Minutes—LRAC Committee on Consumer and Wholesale Prices, 31 October 1962, p. 5, LRAC, Price Committee, Fiscal 1963, box 2, LRAC, Records of the BLS.

¹¹ Laura Mae Webb and Ethel D. Hoover, “The Consumer Price Index: Issues of Concept and Practice”, 1 October 1965, p. 26, Div. of Prices & Cost of Living, 1965-1968, box 3, Program & Policy Files, 1950-75, Records of the BLS.

¹² Sidney Jaffe to Arnold Chase, 10 March 1966, Addendum, p. 2, *ibid.*

¹³ I have described and documented these changes in more detail in Stapleford 2009, pp. 347-352.

¹⁴ Thomas Gavett, “Research on Quality Adjustments in Price Indexes,” p. 2, unlabeled folder, box 2, LRAC, Records of the BLS. Joel Popkin, “A Constant Utility Cost of Living Index”, October 1968, pp. 1-2, CPI Revision—Cost of living index, box 4, Janet Norwood, A-Z Files, Records of the BLS.

¹⁵ Bert Balk has discussed related changes in the various editions of Pigou’s texts, noting a constant emphasis on attempts to bound a true index by Laspeyres and Paasche indexes. Balk also sees “revealed-preference type arguments” in the “background” to Pigou’s work; in my view, of course, these arguments are confined to the later editions (Balk 2009, pp. 21-22).

ARCHIVAL SOURCES

Papers of Katherine P. Ellickson, Walter Reuther Library, Wayne State University, Detroit, Mich.

Records of the U.S. Bureau of Labor Statistics (BLS), Record Group 257, U.S. National Archives II, College Park, Md.

Records of the Office of Management and Budget (OMB) Record Group 51, U.S. National Archives II, College Park, Md.

REFERENCES

Adams, Thomas S. (1901). "Index Numbers and the Standard of Value." *Journal of Political Economy* 10, no. 1: 1-31.

---. (1902). "Index Numbers and the Standard of Value, II." *Journal of Political Economy* 10, no. 2: 193-213.

Aldrich, John. (1992). "Probability and Depreciation: A History of the Stochastic Approach to Index Numbers." *History of Political Economy* 24, no. 3: 657-687.

Allen, R. G. D. (1933). "On the Marginal Utility of Money and Its Application." *Economica*, no. 40 (May): 186-209.

---. (1949). "The Economic Theory of Index Numbers." *Economica* 16, no. 63: 197-203.

---. (1975). *Index Numbers in Theory and Practice*. Chicago, Aldine.

Backhouse, Roger E. (2003). "The Stabilization of Price Theory, 1920-1955". *A Companion to the History of Economic Thought*. Warren J. Samuels, Jeff Biddle and John B. Davis. Malden, Mass., Blackwell.

Baker, Dean (1998). *Getting Prices Right: The Debate Over the Consumer Price Index*. Armonk, M.E. Sharpe.

Balk, Bert M. (2009). *Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference*. Cambridge, Cambridge University Press.

Banzhaf, H. Spencer. (2001). "Quantifying the Qualitative: Quality-Adjusted Price Indexes in the United States, 1915-1961". *The Age of Economic Measurement*. Judy L. Klein and Mary S. Morgan. Durham, Duke University Press: 345-370.

---. (2004). "The Form and Function of Price Indexes: A Historical Accounting." *History of Political Economy* 36, no. 4: 589-616.

Berndt, Ernst R. (2006). "The Boskin Report after a Decade: After Life or Requiem?" *International Productivity Monitor* no. 12 (Spring): 61-73.

Bortkiewicz, L. von. (1932). "Die Kaufkraft des Geldes und ihre Messung." *Nordic Statistical Journal* 4: 1-68.

Boskin, Michael J., Ellen R. Dulberger, et al. (1996). *Toward a More Accurate Measure of the Cost of Living: Final Report to the Senate Finance Committee from the Advisory*

- Commission to Study the Consumer Price Index*. Washington, D.C., U.S. Government Printing Office.
- Bowley, Arthur L. (1919). "The Measurement of Changes in the Cost of Living." *Journal of the Royal Statistical Society* 82, no. 3 (May): 343-372.
- . (1920). "Cost of Living and Wage Determination." *Economic Journal* 30, no. 117 (March): 114-117.
- . (1928). "Notes on Index Numbers." *Economic Journal* 38, no. 150 (June): 216-237.
- Clark, John M. (1918). "Economics and Modern Psychology I." *Journal of Political Economy* 26, no. 1: 1-30.
- Collins, Robert M. (2000). *More: The Politics of Economic Growth in Postwar America*. New York, Oxford University Press.
- Cooter, Robert and Peter Rappoport. (1984). "Were the Ordinalists Wrong About Welfare Economics?" *Journal of Economic Literature* 22, no. 2: 507-530.
- Deaton, Angus. (1998). "Getting Prices Right: What Should Be Done?" *Journal of Economic Perspectives* 12, no. 1: 37-46.
- Diewert, W. Erwin. (1976). "Exact and Superlative Index Numbers." *Journal of Econometrics* 4, no. 2 (May): 115-145.
- . (1983). "The Theory of the Cost-of-Living Index and the Measurement of Welfare Change". *Price Level Measurement: Proceedings from a Conference Sponsored by Statistics Canada*. W. E. Diewert and C. Montmarquette. Ottawa, Minister of Supply and Services Canada: 163-233.
- . (1993). "The Early History of Price Index Research". *Essays in Index Number Theory, Vol 1*. W. Erwin Diewert and Alice O. Nakamura. Amsterdam, Elsevier Science Publishers: 33-65.
- . (1998). "Index Number Issues in the Consumer Price Index." *Journal of Economic Perspectives* 12, no. 1 (Winter): 47-58.
- Douglas, Paul H. (1930). *Real Wages in the United States, 1890-1926*. Boston, Houghton Mifflin Co.
- Edgeworth, Francis Y. (1894). "Recent Writings on Index Numbers." *Economic Journal* 4, no. 13: 158-165.
- . (1923). "The Doctrine of Index Numbers According to Mr. Correa Walsh." *Economic Journal* 33, no. 131 (September): 343-351.
- . (1925). "The Element of Probability in Index-Numbers." *Journal of the Royal Statistical Society* 88, no. 4 (July): 557-575.
- . (1925). "The Plurality of Index-Numbers." *Economic Journal* 35, no. 139 (September): 379-388.
- . (1925 [1887--1889]). "Measurement of the Change in the Value of Money". *Papers Relating to Political Economy*. Francis Y. Edgeworth. London, MacMillan & Co. 1: 195-297.

- Fetter, Frank. (1895). "The Exploitation of Theories of Value in the Discussion of the Standard of Deferred Payments." *Annals of the American Academy of Political and Social Science* (May): 882-896.
- Fisher, Franklin M. and Karl Shell. (1968). "Taste and Quality Change in the Pure Theory of the True Cost-of-Living Index". *Value, Capital, and Growth: Papers in Honour of Sir John Hicks*. J. N. Wolfe. Edinburgh, University of Edinburgh Press: 97-139.
- . (1972). *The Economic Theory of Price Indices*. New York, Academic Press.
- Fisher, Irving. (1911). *The Purchasing Power of Money: Its Determination and Relation to Credit Interest and Crises*. New York, MacMillan Co.
- . (1922). *The Making of Index Numbers: A Study of Their Variety, Tests, and Reliability*. Boston, Houghton Mifflin Co.
- Frisch, Ragnar. (1936). "Annual Survey of General Economic Theory: The Problem of Index Numbers." *Econometrica* 4, no. 1 (January): 1-38.
- (1937). "Methods of Measuring the Relative Cost of Living." Washington, D.C., July 27, 1937, [mimeograph] Source: Price indexes--technical notes, box 218, Statistical Records Relating to Particular Subject Areas, OMB.
- . (1954). "Some Basic Principles of Price of Living Measurements." *Econometrica* 22, no. 4 (October): 407-421.
- General Motors Corporation. (1939). *Dynamics of Automobile Demand*. New York, General Motors Corporation.
- Gilbert, Milton. (1961). "Quality Changes and Index Numbers." *Economic Development and Cultural Change* 9, no. 3 (April): 287-294.
- Gilbert, Milton, George Jaszi, et al. (1948). "Objectives of National Income Measurement: A Reply to Professor Kuznets." *The Review of Economics and Statistics* 30, no. 3: 179-195.
- Gillingham, Robert. (1974). "A Conceptual Framework for the Consumer Price Index." *Proceedings of the Business and Economic Statistics Section of the American Statistical Association*: 246-252.
- Gini, Corrado. (1924). "Quelques Considérations au Sujet de la Construction des Nombres Indices des Prix et des Questions Analogues." *Metron* 4: 3-162.
- Greenlees, John S. (2001). "The U.S. CPI and the Cost-of-Living Objective." Joint ECE/ILO Meeting on Consumer Price Indexes, Geneva, Switzerland, November 2001.
- Haberler, Gottfried. (1927). *Der Sinn der Indexzahlen*. Tübingen, J. C. B. Mohr.
- Hennipman, Pieter. (1987). "A Tale of Two Schools: Comments on a New View of the Ordinalist Revolution." *De Economist* 135, no. 2: 141-162.
- Hicks, J. R. (1940). "The Valuation of the Social Income." *Economica* 7, no. 26: 105-124.
- Hicks, J. R. and R. G. D. Allen. (1934). "A Reconsideration of the Theory of Value. Parts I & II." *Economica* 1, no. 1: 52-76; no. 2: 196-219.
- Jacobs, Meg. (2005). *Pocketbook Politics: Economic Citizenship in Twentieth-Century America*. Princeton, Princeton University Press.

- Keynes, John Maynard. (1930). *A Treatise on Money*. New York,, Harcourt Brace & Co.
- Konüs, A. A. (1939 [1924]). "The Problem of the True Index of the Cost of Living." *Econometrica* 7, no. 1 (January): 10-29.
- . (1968). "The Theory of the Consumer Price Indexes and the Problem of the Comparison of the Cost of Living in Time and Space". *The Social Sciences: Problems and Orientations*. UNESCO. The Hague, Mouton & Co.: 93-107.
- Kuznets, Simon. (1948). "National Income: A New Version." *The Review of Economics and Statistics* 30, no. 3: 151-179.
- Merriam, Lucius S. (1893). "The Theory of Final Utility in its Relation to Money and the Standard of Deferred Payments." *Annals of the American Academy of Political and Social Science* (January): 483-501.
- Mills, Frederick C., E. Wright Bakke, et al. (1943). "An Appraisal of the U.S. Bureau of Labor Statistics Cost-of-Living Index." *Journal the American Statistical Association* 38, no. 224 (December): 387-405.
- Morgan, Mary S. and Malcolm Rutherford, eds. (1998). *From Interwar Pluralism to Postwar Neoclassicism*. Durham, Duke University Press.
- National Bureau of Economic Research. (1961). *The Price Statistics of the Federal Government: Review, Appraisal, and Recommendations*. New York, National Bureau of Economic Research.
- National Research Council. (2002). *At What Price? Conceptualizing and Measuring Cost-of-Living Indexes*. Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes, Charles Schultze and Christopher Mackie, eds. Washington, D.C., National Academies Press.
- Perlman, Mark. (1987). "Political Purpose and National Accounts". *The Politics of Numbers*. William Alonso and Paul Starr. New York, Russell Sage Foundation: 133-151.
- Pigou, A. C. (1912). *Wealth and Welfare*. London, MacMillan & Co.
- . (1920). *The Economics of Welfare*. London, MacMillan and Co.
- . (1924). *The Economics of Welfare*. 2nd ed. London, MacMillan & Co.
- Pollak, Robert A. (1989). *The Theory of the Cost-of-Living Index*. New York, Oxford University Press.
- Robbins, Lionel. (1932). *An Essay on the Nature and Significance of Economic Science*. London, MacMillan & Co.
- Rutherford, Malcolm. (2000). "Understanding Institutional Economics: 1918-1929." *Journal of the History of Economic Thought* 22, no. 3: 277-308.
- Samuelson, Paul A. (1947). *Foundations of Economic Analysis*. Cambridge, Mass., Harvard University Press.
- Samuelson, Paul. A. and S. Swamy. (1974). "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis." *American Economic Review* 64, no. 4: 566-593.

- Schultz, Henry. (1939). "A Misunderstanding in Index-Number Theory: The True Konüs Condition on Cost-of-Living Index Numbers and Its Limitations." *Econometrica* 7, no. 1 (January): 1-9.
- Schumpeter, Joseph Alois. (1954). *History of Economic Analysis*. New York, Oxford University Press.
- Sen, Amartya. (1979). "The Welfare Basis of Real Income Comparisons: A Survey." *Journal of Economic Literature* 17, no. 1: 1-45.
- Sidgwick, Henry. (1887). *The Principles of Political Economy*. 2nd ed. London, MacMillan and Co.
- Staehle, Hans. (1934). *International Comparison of Food Costs*. Geneva, International Labour Office.
- . (1935). "A Development of the Economic Theory of Price Index Numbers." *Review of Economic Studies* 2, no. 3 (June): 163-188.
- . (1937). "A General Method for the Comparison of the Price of Living." *The Review of Economic Studies* 4, no. 3: 205-214.
- Stapleford, Thomas A. (2003). "'The Most Important Single Statistic': The Consumer Price Index and American Political Economy, 1880-1955." Ph.D. dissertation, Harvard University, Cambridge, MA.
- . (2009). *The Cost of Living in America: A Political History of Economic Statistics, 1880–2000*. New York: Cambridge University Press.
- . (2011). "Reconceiving Quality: Political Economy and the Rise of Hedonic Price Indexes." In *Histories on Econometrics*, annual supplement to *History of Political Economy*, vol. 43, Marcel Boumans, Ariane Dupont-Keiffer, and Duo Qin, eds. Durham, Duke University Press, 309-328.
- Stigler, George J. (1952). *The Theory of Price*. Rev. ed. New York, Macmillan Co.
- Triplett, Jack E. (1983). "Escalation Measures: What is the Answer? What is the Question?" In *Price Level Measurement: Proceedings from a Conference Sponsored by Statistics Canada*, ed. W. E. Diewert and C. Montmarquette. Ottawa, Minister of Supply and Services, 472-476.
- . (2001). "Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?" *Economic Journal* 111, no. 472 (June): F311-F334.
- . (2006). "The Boskin Commission Report After a Decade." *International Productivity Monitor* no. 12 (Spring): 42-60.
- U.S. Bureau of Labor Statistics. (1940). "The Bureau of Labor Statistics' New Index of Cost of Living." *Monthly Labor Review* 51, no. 2 (August): 367-404.
- U.S. Congress. Joint Economic Committee. (1959). *Staff Report on Employment, Growth, and Price Levels*. 86th Cong., 1st sess. Washington, D.C., U. S. Govt. Print. Off.
- . (1961). *Government Price Statistics: Hearings before the Subcommittee on Economic Statistics of the Joint Economic Committee*. 87th Cong., 1st sess. Washington, D.C., GPO.

- . (1966). *Government Price Statistics: Hearings before the Subcommittee on Economic Statistics*. 89th Cong., 2nd sess. Washington, D.C., GPO.
- Veblen, Thorstein. (1899). *The Theory of the Leisure Class*. New York, MacMillan Co.
- Walsh, Correa M. (1901). *The Measurement of General Exchange-Value*. New York, MacMillan Co.
- . (1921). *The Problem of Estimation*. London, P.S. King & Son, Ltd.
- . (1924). "Professor Edgeworth's Views on Index Numbers." *Quarterly Journal of Economics* 38, no. 3 (May): 500-519.
- Weber, Christian E. (2001). "Pareto and the 53 Percent Ordinal Theory of Utility." *History of Political Economy* 33, no. 3: 541-576.
- Williams, Faith M. and Ethel D. Hoover. (1946). "Measuring Price and Quality of Consumers' Goods." *Journal of Marketing* 10, no. 4: 354-369.