INDEX NUMBER THEORY AND MEASUREMENT ECONOMICS

By W.E. Diewert, February, 2015.

CHAPTER 4: The Theory of the Cost of Living Index: The Single Consumer Case

1. The Konüs Cost of Living Index for a Single Consumer

In this section, we will outline the theory of the cost of living index for a single consumer (or household) that was first developed by the Russian economist, A. A. Konüs (1924). This theory relies on the assumption of *optimizing behavior* on the part of economic agents (consumers or producers). Thus given a vector of commodity or input prices p^t that the agent faces in a given time period t, it is assumed that the corresponding observed quantity vector q^t is the solution to a cost minimization problem that involves either the consumer's preference or utility function f or the producer's production function f.[1] Thus in contrast to the axiomatic approach to index number theory, the economic approach does *not* assume that the two quantity vectors q^0 and q^1 are independent of the two price vectors p^0 and p^1. In the economic approach, the period 0 quantity vector q^0 is determined by the consumer's preference function f and the period 0 vector of prices p^0 that the consumer faces and the period 1 quantity vector q^1 is determined by the consumer's preference function f and the period 1 vector of prices p^1.

We assume that "the" consumer has well defined *preferences* over different combinations of the N consumer commodities or items.[2] Each combination of items can be represented by a positive vector $q \equiv [q_1,\ldots,q_N]$. The consumer's preferences over alternative possible consumption vectors q are assumed to be representable by a continuous, increasing and concave[3] utility function f. Thus if $f(q^1) > f(q^0)$, then the consumer prefers the consumption vector q^1 to q^0. We further assume that the consumer minimizes the cost of achieving the period t utility level $u^t \equiv f(q^t)$ for periods t = 0,1. Thus we assume that the observed period t consumption vector q^t solves the following *period t cost minimization problem*:

$$(1)\quad C(u^t,p^t) \equiv \min_q \{\textstyle\sum_{i=1}^N p_i^t q_i : f(q) = u^t \equiv f(q^t)\} = \textstyle\sum_{i=1}^N p_i^t q_i^t ; \quad t = 0,1.$$

The period t price vector for the n commodities under consideration that the consumer faces is p^t. Note that the solution to the cost or expenditure minimization problem (1) for a general utility level u and general vector of commodity prices p defines the *consumer's cost function*, C(u,p). We shall use the cost function in order to define the consumer's cost of living price index.

[1] For a description of the economic theory of the input and output price indexes, see Balk (1998). In the economic theory of the output price index, q^t is assumed to be the solution to a revenue maximization problem involving the output price vector p^t.

[2] In this section, these preferences are assumed to be invariant over time. In chapter 5 when we introduce environmental variables, this assumption will be relaxed (one of the environmental variables could be a time variable that shifts tastes).

[3] f is concave if and only if $f(\lambda q^1 + (1-\lambda)q^2) \geq \lambda f(q^1) + (1-\lambda)f(q^2)$ for all $0 \leq \lambda \leq 1$ and all $q^1 >> 0_N$ and $q^2 >> 0_N$. Note that $q \geq 0_N$ means that each component of the N dimensional vector q is nonnegative, $q >> 0_N$ means that each component of q is positive and $q > 0_N$ means that $q \geq 0_N$ but $q \neq 0_N$; i.e., q is nonnegative but at least one component is positive.

Problem 1: Assume that $f(q)$ is continuous, increasing in the components of q and defined for all $q \geq 0_N$ with $f(0_N) = u_0$. The cost function $C(u,p)$ is defined as $\min_{q\geq 0} \{p \bullet q : f(q) \geq u\}$ for all strictly positive price vectors $p >> 0_N$ and all u such that $u_0 \leq u \leq u_{max}$ where u_{max} is the maximum value of utility that $f(q)$ can attain for $q \geq 0_N$ Show that the cost function C has the following properties:

(i) $C(u,p) > 0$ if $u > u_0$ and $p >> 0_N$ (positivity);
(ii) $C(u,p^2) > C(u,p^1)$ if $u > u_0$ and $p^2 >> p^1 >> 0_N$ (increasing in the components of p);
(iii) $C(u_2,p) > C(u_1,p)$ if $u_2 > u_1 \geq u_0$ and $p >> 0_N$ (increasing in the utility level u);
(iv) $C(u,\lambda p) = \lambda C(u,p)$ for $u \geq u_0$, $p >> 0_N$ and $\lambda > 0$ (C is linearly homogeneous in the components of p);
(v) $C(u,\lambda p^1 + (1-\lambda)p^2) \geq \lambda C(u,p^1) + (1-\lambda)C(u,p^2)$ (C is concave in the components of p).
Note that we do not have to assume that $f(q)$ is concave in q in order to derive the above properties of C.

The Konüs (1924) family of *true cost of living indexes* pertaining to two periods where the consumer faces the strictly positive price vectors $p^0 \equiv (p_1^0,\ldots,p_N^0)$ and $p^1 \equiv (p_1^1,\ldots,p_N^1)$ in periods 0 and 1 respectively is defined as the ratio of the minimum costs of achieving the same utility level $u \equiv f(q)$ where $q \equiv (q_1,\ldots,q_N)$ is a positive reference quantity vector; i.e., we have

$$(2)\quad P_K(p^0,p^1,q) \equiv C[f(q),p^1]/C[f(q),p^0].$$

We say that definition (2) defines a *family* of price indexes because there is one such index for each reference quantity vector q chosen.

It is natural to choose two specific reference quantity vectors q in definition (2): the observed base period quantity vector q^0 and the current period quantity vector q^1. The first of these two choices leads to the *following Laspeyres-Konüs true cost of living index*:

$$\begin{aligned}
(3)\quad P_K(p^0,p^1,q^0) &\equiv C[f(q^0),p^1] / C[f(q^0),p^0] \\
&= C[f(q^0),p^1] / \textstyle\sum_{i=1}^N p_i^0 q_i^0 \qquad\qquad \text{using (1) for } t = 0 \\
&= \min_q \{\textstyle\sum_{i=1}^N p_i^1 q_i : f(q) = f(q^0)\} / \textstyle\sum_{i=1}^N p_i^0 q_i^0 \\
&\quad \text{using the definition of the cost minimization problem that defines } C[f(q^0),p^1] \\
&\leq \textstyle\sum_{i=1}^N p_i^1 q_i^0 / \textstyle\sum_{i=1}^N p_i^0 q_i^0 \\
&\quad \text{since } q^0 \equiv (q_1^0,\ldots,q_N^0) \text{ is feasible for the minimization problem} \\
&= P_L(p^0,p^1,q^0,q^1)
\end{aligned}$$

where P_L is the Laspeyres price index defined in earlier chapters. *Thus the (unobservable) Laspeyres-Konüs true cost of living index is bounded from above by the observable Laspeyres price index.*[4]

The second of the two natural choices for a reference quantity vector q in definition (2) leads to the *following Paasche-Konüs true cost of living index*:

$$(4)\quad P_K(p^0,p^1,q^1) \equiv C[f(q^1),p^1] / C[f(q^1),p^0]$$

[4] This inequality was first obtained by Konüs (1924) (1939; 17). See also Pollak (1983).

$= \sum_{i=1}^{N} p_i^1 q_i^1 / C[f(q^1),p^0]$ using (1) for t = 1

$= \sum_{i=1}^{n} p_i^1 q_i^1 / \min_q \{\sum_{i=1}^{N} p_i^0 q_i : f(q) = f(q^1)\}$

using the definition of the cost minimization problem that defines $C[f(q^1),p^0]$

$\geq \sum_{i=1}^{N} p_i^1 q_i^1 / \sum_{i=1}^{N} p_i^0 q_i^1$

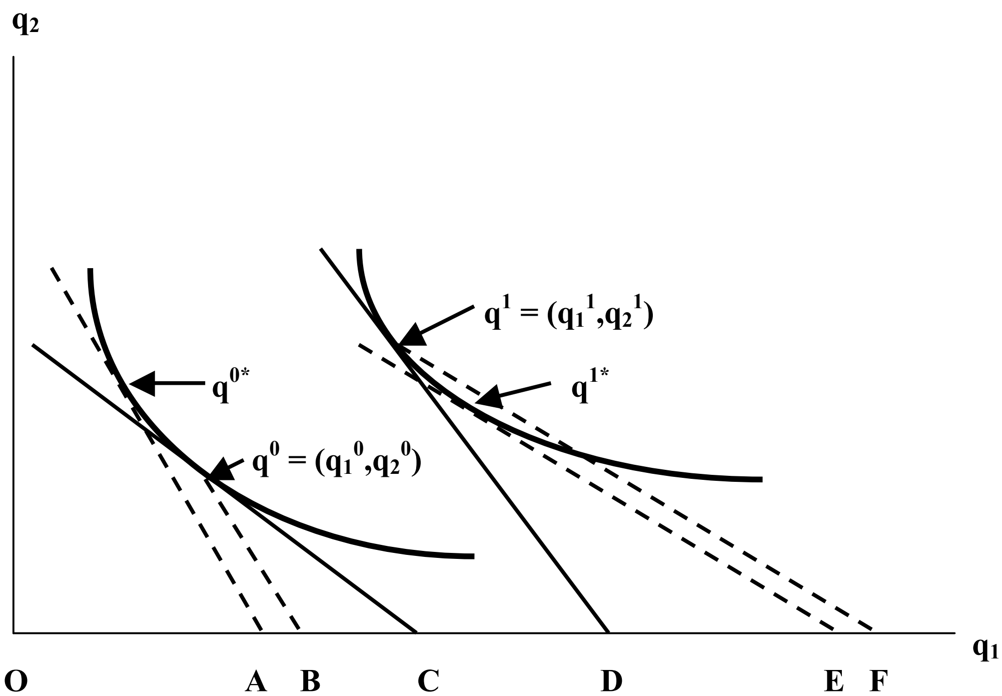
since $q^1 \equiv (q_1^1,\ldots,q_N^1)$ is feasible for the minimization problem and thus $C[f(q^1),p^0] \leq \sum_{i=1}^{N} p_i^0 q_i^1$ and hence $1/C[f(q^1),p^0] \geq 1/ \sum_{i=1}^{N} p_i^0 q_i^1$

$= P_P(p^0,p^1,q^0,q^1)$

where P_P is the Paasche price index defined in earlier chapters. *Thus the (unobservable) Paasche-Konüs true cost of living index is bounded from below by the observable Paasche price index.*[5]

It is possible to illustrate the two inequalities (3) and (4) if there are only two commodities; see Figure 1 below.

Figure 1: The Laspeyres and Paasche bounds to the true cost of living



The solution to the period 0 cost minimization problem is the vector q^0 and the straight line through C represents the consumer's period 0 budget constraint, the set of quantity points q_1,q_2 such that $p_1^0 q_1 + p_2^0 q_2 = p_1^0 q_1^0 + p_2^0 q_2^0$. The curved line through q^0 is the consumer's period 0 indifference curve, the set of points q_1,q_2 such that $f(q_1,q_2) = f(q_1^0,q_2^0)$; i.e., it is the set of consumption vectors that give the same utility as the observed period 0 consumption vector q^0. The solution to the period 1 cost minimization problem is the vector q^1 and the straight line through D represents the consumer's period 1 budget constraint, the set of quantity points q_1,q_2 such that $p_1^1 q_1 + p_2^1 q_2 = p_1^1 q_1^1 + p_2^1 q_2^1$. The curved line through q^1 is the consumer's period 1 indifference curve, the set of points q_1,q_2 such that $f(q_1,q_2) = f(q_1^1,q_2^1)$; i.e., it is the set of consumption vectors that give the same

[5] This inequality is also due to Konüs (1924) (1939; 19). See also Pollak (1983).

utility as the observed period 1 consumption vector q^1. The point q^{0*} solves the hypothetical cost minimization problem of minimizing the cost of achieving the base period utility level $u^0 \equiv f(q^0)$ when facing the period 1 price vector $p^1 = (p_1{}^1, p_2{}^1)$. Thus we have $C[u^0, p^1] = p_1{}^1 q_1{}^{0*} + p_2{}^1 q_2{}^{0*}$ and the dashed line through A is the corresponding isocost line $p_1{}^1 q_1 + p_2{}^1 q_2 = C[u^0, p^1]$. Note that the hypothetical cost line through A is parallel to the actual period 1 cost line through D. From (3), the Laspeyres-Konüs true index is $C[u^0, p^1] / [p_1{}^0 q_1{}^0 + p_2{}^0 q_2{}^0]$ while the ordinary Laspeyres index is $[p_1{}^1 q_1{}^0 + p_2{}^1 q_2{}^0] / [p_1{}^0 q_1{}^0 + p_2{}^0 q_2{}^0]$. Since the denominators for these two indexes are the same, the difference between the indexes is due to the differences in their numerators. In Figure 1, this difference in the numerators is expressed by the fact that the cost line through A lies *below* the parallel cost line through B. Now if the consumer's indifference curve through the observed period 0 consumption vector q^0 were L shaped with vertex at q^0, then the consumer would not change his or her consumption pattern in response to a change in the relative prices of the two commodities while keeping a fixed standard of living. In this case, the hypothetical vector q^{0*} would coincide with q^0, the dashed line through A would coincide with the dashed line through B and the true Laspeyres-Konüs index would *coincide* with the ordinary Laspeyres index. However, L shaped indifference curves are not generally consistent with consumer behavior; i.e., when the price of a commodity decreases, consumers generally demand more of it. Thus in the general case, there will be a gap between the points A and B. The magnitude of this gap represents the amount of *substitution bias* between the true index and the corresponding Laspeyres index; i.e., the Laspeyres index will generally be *greater* than the corresponding true cost of living index, $P_K(p^0, p^1, q^0)$.

Figure 1 can also be used to illustrate the inequality (4). First note that the dashed lines through E and F are parallel to the period 0 isocost line through C. The point q^{1*} solves the hypothetical cost minimization problem of minimizing the cost of achieving the current period utility level $u^1 \equiv f(q^1)$ when facing the period 0 price vector $p^0 = (p_1{}^0, p_2{}^0)$. Thus we have $C[u^1, p^0] = p_1{}^0 q_1{}^{1*} + p_2{}^0 q_2{}^{1*}$ and the dashed line through E is the corresponding isocost line $p_1{}^1 q_1 + p_2{}^1 q_2 = C[u^0, p^1]$. From (4), the Paasche-Konüs true index is $[p_1{}^1 q_1{}^1 + p_2{}^1 q_2{}^1] / C[u^1, p^0]$ while the ordinary Paasche index is $[p_1{}^1 q_1{}^1 + p_2{}^1 q_2{}^1] / [p_1{}^0 q_1{}^1 + p_2{}^0 q_2{}^1]$. Since the numerators for these two indexes are the same, the difference between the indexes is due to the differences in their denominators. In Figure 1, this difference in the denominators is expressed by the fact that the cost line through E lies *below* the parallel cost line through F. The magnitude of this difference represents the amount of *substitution bias* between the true index and the corresponding Paasche index; i.e., the Paasche index will generally be *less* than the corresponding true cost of living index, $P_K(p^0, p^1, q^1)$. Note that this inequality goes in the opposite direction to the previous inequality between the two Laspeyres indexes. The reason for this change in direction is due to the fact that one set of differences between the two indexes takes place in the numerators of the two indexes (the Laspeyres inequalities) while the other set takes place in the denominators of the two indexes (the Paasche inequalities).

The bound (3) on the Laspeyres-Konüs true cost of living $P_K(p^0, p^1, q^0)$ using the base period level of utility as the living standard is *one sided* as is the bound (4) on the Paasche-Konüs true cost of living $P_K(p^0, p^1, q^1)$ using the *current period* level of utility as the living standard. In a remarkable result, Konüs (1924; 20) showed that there exists an intermediate consumption vector q^* that is on the straight line joining the base period consumption vector q^0 and the current period consumption vector q^1 such that the corresponding (unobservable) true cost of living index $P_K(p^0,p^1,q^*)$ is between the observable Laspeyres and Paasche indexes, P_L and P_P.[6] Thus we have:

Proposition 1: There exists a number λ^* between 0 and 1 such that

(5) $P_L \leq P_K(p^0,p^1, \lambda^* q^0 + (1-\lambda^*)q^1) \leq P_P$ or $P_P \leq P_K(p^0,p^1, \lambda^* q^0 + (1-\lambda^*)q^1) \leq P_L$.

Proof: Define $g(\lambda)$ for $0 \leq \lambda \leq 1$ by $g(\lambda) \equiv P_K(p^0,p^1,(1-\lambda)q^0 + \lambda q^1)$. Note that $g(0) = P_K(p^0,p^1,q^0)$ and $g(1) = P_K(p^0,p^1,q^1)$. There are 24 = (4)(3)(2)(1) possible a priori inequality relations that are possible between the four numbers $g(0)$, $g(1)$, P_L and P_P. However, the inequalities (3) and (4) above imply that $g(0) \leq P_L$ and $P_P \leq g(1)$. This means that there are only six possible inequalities between the four numbers:

(6) $g(0) \leq P_L \leq P_P \leq g(1)$;
(7) $g(0) \leq P_P \leq P_L \leq g(1)$;
(8) $g(0) \leq P_P \leq g(1) \leq P_L$;
(9) $P_P \leq g(0) \leq P_L \leq g(1)$;
(10) $P_P \leq g(1) \leq g(0) \leq P_L$;
(11) $P_P \leq g(0) \leq g(1) \leq P_L$.

Using the assumptions that: (a) the consumer's utility function f is continuous over its domain of definition; (b) the utility function is increasing in the components of q and hence is subject to local nonsatiation and (c) the price vectors p^t have strictly positive components, it is possible to use Debreu's (1959; 19) Maximum Theorem (see also Diewert (1993; 112-113) for a statement of the Theorem) to show that the consumer's cost function $C(f(q),p^t)$ will be continuous in the components of q. Thus using definition (2), it can be seen that $P_K(p^0,p^1,q)$ will also be continuous in the components of the vector q. Hence $g(\lambda)$ is a continuous function of λ and assumes all intermediate values between $g(0)$ and $g(1)$. By inspecting the inequalities (6)-(11) above, it can be seen that we can choose λ between 0 and 1, λ^* say, such that $P_L \leq g(\lambda^*) \leq P_P$ for case (6) or such that $P_P \leq g(\lambda^*) \leq P_L$ for cases (7) to (11). Thus at least one of the two inequalities in (5) holds. Q.E.D.

The above inequalities are of some practical importance. If the observable (in principle) Paasche and Laspeyres indexes are not too far apart, then taking a symmetric average of these indexes should provide a good approximation to a true cost of living index where the reference standard of living is somewhere between the base and current period living standards. To determine the precise symmetric average of the Paasche and Laspeyres indexes, we can appeal to the results in Chapter 1 above and take the geometric mean, which is the Fisher price index. Thus the Fisher ideal price index receives a fairly strong justification as a good approximation to an unobservable theoretical cost of living index.

The bounds (3)-(5) are the best bounds that we can obtain on true cost of living indexes without making further assumptions. In a subsequent section, we will make further assumptions on the

[6] For more recent applications of the Konüs method of proof, see Diewert (1983a;191) for an application to the consumer context and Diewert (1983b; 1059-1061) for an application to the producer context.

class of utility functions that describe the consumer's tastes for the N commodities under consideration. With these extra assumptions, we are able to determine the consumer's true cost of living exactly.

2. The True Cost of Living Index when Preferences are Homothetic

Up to now, the consumer's preference function f did not have to satisfy any particular homogeneity assumption. In this section, we assume that f is (positively*) linearly homogeneous*[7]; i.e., we assume that the consumer's utility function has the following property:

(12) $f(\lambda q) = \lambda f(q)$ for all $\lambda > 0$ and all $q >> 0_N$.

Given the continuity of f, it can be seen that property (12) implies that $f(0_N) = 0$ so that our old u_0 is now equal to 0. Furthermore, f also satisfies $f(q) > 0$ if $q > 0_N$.

In the economics literature, assumption (12) is known as the assumption of *homothetic preferences*.[8] This assumption is not strictly justified from the viewpoint of actual economic behavior, but it leads to economic price indexes that are independent from the consumer's standard of living.[9] Under this assumption, the consumer's expenditure or cost function, C(u,p) defined by (1) above, decomposes as follows. For positive commodity prices $p >> 0_N$ and a positive utility level u, we have by the definition of C as the minimum cost of achieving the given utility level u:

$$
\begin{aligned}
(13)\quad C(u,p) &\equiv \min_q \{\textstyle\sum_{i=1}^N p_iq_i : f(q_1,\ldots,q_N) \geq u\} \\
&= \min_q \{\textstyle\sum_{i=1}^N p_iq_i : (1/u)f(q_1,\ldots,q_N) \geq 1\} && \text{dividing by } u > 0 \\
&= \min_q \{\textstyle\sum_{i=1}^N p_iq_i : f(q_1/u,\ldots,q_N/u) \geq 1\} && \text{using the linear homogeneity of f} \\
&= u \min_q \{\textstyle\sum_{i=1}^N p_iq_i/u : f(q_1/u,\ldots,q_N/u) \geq 1\} \\
&= u \min_z \{\{\textstyle\sum_{i=1}^N p_iz_i : f(z_1,\ldots,z_N) \geq 1\} && \text{letting } z_i = q_i/u \\
&= u\, C(1,p) && \text{using definition (1) with } u = 1 \\
&= u\, c(p)
\end{aligned}
$$

where $c(p) \equiv C(1,p)$ is the *unit cost function* that is corresponds to f.[10] Using Problem 1, it can be shown that the unit cost function c(p) satisfies the same regularity conditions that f satisfied; i.e.,

[7] This assumption is fairly restrictive in the consumer context. It implies that each indifference curve is a radial projection of the unit utility indifference curve. It also implies that all income elasticities of demand are unity, which is contradicted by empirical evidence.

[8] More precisely, Shephard (1953) defined a homothetic function to be a monotonic transformation of a linearly homogeneous function. However, if a consumer's utility function is homothetic, we can always rescale it to be linearly homogeneous without changing consumer behavior. Hence, we simply identify the homothetic preferences assumption with the linear homogeneity assumption.

[9] This particular branch of the economic approach to index number theory is due to Shephard (1953) (1970) and Samuelson and Swamy (1974). Shephard in particular realized the importance of the homotheticity assumption in conjunction with separability assumptions in justifying the existence of subindexes of the overall cost of living index. It should be noted that if the consumer's change in real income or utility between the two periods under consideration is not too large, then assuming that the consumer has homothetic preferences will lead to a true cost of living index which is very close to Laspeyres-Konüs and Paasche-Konüs true cost of living indexes defined above by (3) and (4).

[10] Economists will recognize the producer theory counterpart to the result $C(u,p) = uc(p)$: if a producer's production function f is subject to constant returns to scale, then the corresponding total cost function C(u,p) is equal to the product of the output level u times the unit cost c(p).

c(p) is positive, concave and (positively) linearly homogeneous for positive price vectors.[11] Substituting (13) into (1) and using $u^t = f(q^t)$ leads to the following equations:

(14) $\sum_{i=1}^{N} p_i^t q_i^t = c(p^t)f(q^t)$ for t = 0,1.

Thus under the linear homogeneity assumption on the utility function f, observed period t expenditure on the n commodities (the left hand side of (14) above) is equal to the period t unit cost $c(p^t)$ of achieving one unit of utility times the period t utility level, $f(q^t)$, (the right hand side of (14) above). Obviously, we can identify the period t unit cost, $c(p^t)$, as the period t price level P^t and the period t level of utility, $f(q^t)$, as the period t quantity level Q^t.[12]

The linear homogeneity assumption on the consumer's preference function f leads to a simplification for the family of Konüs true cost of living indices, $P_K(p^0,p^1,q)$, defined by (2) above. Using this definition for an arbitrary reference quantity vector q, we have:

(15) $P_K(p^0,p^1,q) \equiv C[f(q),p^1]/C[f(q),p^0]$
$= c(p^1)f(q)/c(p^0)f(q)$ using (14) twice
$= c(p^1)/c(p^0)$.

Thus under the homothetic preferences assumption, the entire family of Konüs true cost of living indexes collapses to a single index, $c(p^1)/c(p^0)$, the ratio of the minimum costs of achieving unit utility level when the consumer faces period 1 and 0 prices respectively. Put another way, *under the homothetic preferences assumption, $P_K(p^0,p^1,q)$ is independent of the reference quantity vector q.*

If we use the Konüs true cost of living index defined by the right hand side of (15) as our price index concept, then the corresponding implicit quantity index defined using the product test has the following form:

(16) $Q(p^0,p^1,q^0,q^1,q) \equiv \sum_{i=1}^{N} p_i^1 q_i^1 / \{\sum_{i=1}^{N} p_i^0 q_i^0 P_K(p^0,p^1,q)\}$ using the product test and P_K as the price index
$= c(p^1)f(q^1) / \{c(p^0)f(q^0) P_K(p^0,p^1,q)\}$ using (14) twice
$= c(p^1)f(q^1) / \{c(p^0)f(q^0)[c(p^1)/c(p^0)]\}$ using (15)
$= f(q^1)/f(q^0)$.

[11] Obviously, the utility function f determines the consumer's cost function C(u,p) as the solution to the cost minimization problem in the first line of (13). Then the unit cost function c(p) is defined as C(1,p). Thus f determines c. But we can also use c to determine f under appropriate regularity conditions. In the economics literature, this is known as *duality theory.* For additional material on duality theory and the properties of f and c, see Samuelson (1953), Shephard (1953) and Diewert (1974) (1993; 107-123).

[12] There is also a producer theory interpretation of the above theory; i.e., let f be the producer's (constant returns to scale) production function, let p be a vector of input prices that the producer faces, let q be an input vector and let u = f(q) be the maximum output that can be produced using the input vector q. $C(u,p) \equiv \min_q \{ \sum_{i=1}^{N} p_i q_i : f(q) \geq u \}$ is the producer's cost function in this case and $c(p^t)$ can be identified as the period t input price level while $f(q^t)$ is the period t aggregate input level.

Thus under the homothetic preferences assumption, the *implicit quantity index* that corresponds to the true cost of living price index $c(p^1)/c(p^0)$ is the *utility ratio* $f(q^1)/f(q^0)$. Since the utility function is assumed to be homogeneous of degree one, this is the natural definition for a quantity index.

Problem 2: Assume that the consumer has homothetic preferences. Show that for any reference quantity vector $q >> 0_N$, we have:

(i) $P_P(p^0,p^1,q^0,q^1) \equiv p^1\bullet q^1/p^0\bullet q^1 \leq P_K(p^0,p^1,q) \leq p^1\bullet q^0/p^0\bullet q^0 \equiv P_L(p^0,p^1,q^0,q^1)$

where P_K is the true cost of living index defined by (2) above and P_P and P_L are the ordinary Paasche and Laspeyres price indexes. Thus under the assumption of homothetic preferences, all true cost of living indexes lie between P_P and P_L and we can also deduce that $P_P \leq P_L$.

3. Wold's Identity and Shephard's Lemma

In subsequent sections, we will need two additional results from economic theory: Wold's Identity and Shephard's Lemma.

Wold's (1944; 69-71) (1953; 145) *Identity* is the following result. Assuming that the consumer satisfies the cost minimization assumptions (1) for periods 0 and 1 and that the utility function f is differentiable at the observed quantity vectors $q^0 >> 0_N$ and $q^1 >> 0_N$ it can be shown[13] that the following equations hold:

(17) $p_i^t/ \sum_{k=1}^N p_k^t q_k^t = [\partial f(q^t)/\partial q_i]/ \sum_{k=1}^N q_k^t\, \partial f(q^t)/\partial q_k$; $t = 0,1$; $k = 1,\ldots,N$

where $\partial f(q^t)/\partial q_i$ denotes the partial derivative of the utility function f with respect to the ith quantity q_i evaluated at the period t quantity vector q^t.

If we make the homothetic preferences assumption and assume that the utility function is linearly homogeneous, then Wold's Identity (17) simplifies into the following equations which will prove to be very useful:[14]

(18) $p_i^t/ \sum_{k=1}^N p_k^t q_k^t = [\partial f(q^t)/\partial q_i]/f(q^t)$; $t = 0,1$; $k = 1,\ldots,N$.

Using vector notation, (18) can be rewritten as follows:

(19) $p^t/p^t\bullet q^t = \nabla f(q^t)/f(q^t)$; $t = 0,1$.

[13] To prove this, consider the first order necessary conditions for the strictly positive vector q^t to solve the period t cost minimization problem. The conditions of Lagrange with respect to the vector of q variables are: $p^t = \lambda^t\, \nabla f(q^t)$ where λ^t is the optimal Lagrange multiplier and $\nabla f(q^t)$ is the vector of first order partial derivatives of f evaluated at q^t. Note that this system of equations is the price equals a constant times marginal utility equations that are familiar to economists. Now take the inner product of both sides of this equation with respect to the period t quantity vector q^t and solve the resulting equation for λ^t. Substitute this solution back into the vector equation $p^t = \lambda^t\, \nabla f(q^t)$ and we obtain (17).

[14] Differentiate both sides of the equation $f(\lambda q) = \lambda f(q)$ with respect to λ and then evaluate the resulting equation at $\lambda = 1$. We obtain the equation $\sum_{i=1}^N f_i(q)q_i = f(q)$ where $f_i(q) \equiv \partial f(q)/\partial q_i$.

Shephard's (1953; 11) *Lemma* is the following result. Consider the period t cost minimization problem defined by (1) above. If the cost function $C(u^t,p^t)$ is differentiable with respect to the components of the price vector p, then the period t quantity vector q^t is equal to the vector of first order partial derivatives of the cost function with respect to the components of p; i.e., we have

(20) $q_i^t = \partial C(u^t,p^t)/\partial p_i$; $\quad$ $i = 1,\ldots,N$; $t = 0,1$.

To explain why (20) holds, consider the following argument. Because we are assuming that the observed period t quantity vector q^t solves the cost minimization problem defined by $C(u^t,p^t)$, then q^t must be feasible for this problem so we must have $f(q^t) = u^t$. Thus q^t is a feasible solution for the following cost minimization problem where the general price vector p has replaced the specific period t price vector p^t:

$$(21) \quad C(u^t,p) \equiv \min_q \{\textstyle\sum_{i=1}^N p_iq_i : f(q_1,\ldots,q_n) \geq u^t\} \\ \leq \textstyle\sum_{i=1}^N p_iq_i^t$$

where the inequality follows from the fact that $q^t \equiv (q_1^t,\ldots,q_N^t)$ is a feasible (but usually not optimal) solution for the cost minimization problem in (21). Now define for each strictly positive price vector p the function g(p) as follows:

$$(22) \quad g(p) \equiv \textstyle\sum_{i=1}^N p_iq_i^t - C(u^t,p)$$

where as usual, $p \equiv (p_1,\ldots,p_n)$. Using (1) and (21), it can be seen that g(p) is minimized (over all strictly positive price vectors p) at $p = p^t$. Thus the first order necessary conditions for minimizing a differentiable function of N variables hold, which simplify to equations (20).

If we make the homothetic preferences assumption and assume that the utility function is linearly homogeneous, then using (13), Shephard's Lemma (20) becomes:

(23) $q_i^t = u^t\, \partial c(p^t)/\partial p_i$; $\quad$ $i = 1,\ldots,n$; $t = 0,1$.

Equations (14) can be rewritten as follows:

(24) $\sum_{i=1}^N p_i^tq_i^t = c(p^t)f(q^t) = c(p^t)u^t$ $\quad$ for $t = 0,1$.

Combining equations (23) and (24), we obtain the following system of equations:

(25) $q_i^t / \sum_{k=1}^N p_k^tq_k^t = [\partial c(p^t)/\partial p_i] /c(p^t)$; $\quad$ $i = 1,\ldots,N$; $t = 0,1$.

Using vector notation, we can rewrite (25) as follows:

(26) $q^t/p^t\bullet q^t = \nabla c(p^t)/c(p^t)$; $\quad$ $t = 0,1$.

Note the symmetry of equations (19) with equations (26). It is these two sets of equations that we shall use in subsequent material.

Problem 3: Suppose that the consumer's cost function C(u,p) is differentiable with respect to the components of the commodity price vector p. Then the consumer's system of Hicksian (1946; 331) demand functions is defined as:

(i) $d_i(u,p) \equiv \partial C(u,p)/\partial p_i$; $\qquad$ i = 1,...,N.

These functions trace out the consumer's demand for goods and services as prices p vary but the standard of living u is held fixed. Now suppose that the consumer has homothetic preferences so that:

(ii) $C(u,p) = uC(1,p)$.

Show that under this assumption of homothetic preferences that all N real income elasticities of demand are one; i.e., prove that

(iii) $[\partial d_i(u,p)/\partial u][u/d_i(u,p)] = 1$; $\qquad$ i = 1,....,N.

4. Superlative Indexes: The Fisher Ideal Index

Suppose the consumer has the following utility function:

$$(27)\quad f(q_1,\ldots,q_N) \equiv [\textstyle\sum_{i=1}^N\sum_{k=1}^N a_{ik}\, q_iq_k]^{1/2}\ ;\ \ a_{ik} = a_{ki}\quad \text{for all i and k.}$$

Differentiating f(q) defined by (27) with respect to q_i yields the following equations:

$$(28)\quad f_i(q) = (1/2)[\textstyle\sum_{j=1}^N\sum_{k=1}^N a_{jk}\, q_jq_k]^{-1/2}\, 2\sum_{k=1}^N a_{ik}\, q_k\ ; \qquad i = 1,\ldots,N$$
$$= \textstyle\sum_{k=1}^N a_{ik}\, q_k / f(q) \qquad \text{using (27)}$$

where $f_i(q) \equiv \partial f(q^t)/\partial q_i$. In order to obtain the first equation in (28), we need to use the symmetry conditions, $a_{ik} = a_{ki}$. Now evaluate the second equation in (28) at the observed period t quantity vector $q^t \equiv (q_1^t,\ldots,q_N^t)$ and divide both sides of the resulting equation by $f(q^t)$. We obtain the following equations:

$$(29)\quad f_i(q^t)/f(q^t) = \textstyle\sum_{k=1}^N a_{ik}\, q_k^t / [f(q^t)]^2 \qquad t = 0,1\ ;\ i = 1,\ldots,N.$$

Assume cost minimizing behavior for the consumer in periods 0 and 1. Since the utility function f defined by (27) is linearly homogeneous and differentiable, equations (18) (Wold's Identity) will hold. Now recall the definition of the Fisher ideal quantity index, Q_F defined by the first line of (30) below:

$$(30)\quad Q_F(p^0,p^1,q^0,q^1) \equiv [\textstyle\sum_{i=1}^N p_i^0q_i^1/\sum_{k=1}^N p_k^0q_k^0]^{1/2}\, [\sum_{i=1}^N p_i^1q_i^1/\sum_{k=1}^N p_k^1q_k^0]^{1/2}$$
$$= [\textstyle\sum_{i=1}^N f_i(q^0)q_i^1/f(q^0)]^{1/2}\, [\sum_{i=1}^N p_i^1q_i^1/\sum_{k=1}^N p_k^1q_k^0]^{1/2} \qquad \text{using (18) for } t = 0$$
$$= [\textstyle\sum_{i=1}^N f_i(q^0)q_i^1/f(q^0)]^{1/2}\, /[\sum_{k=1}^N p_k^1q_k^0 /\sum_{i=1}^N p_i^1q_i^1]^{1/2}$$
$$= [\textstyle\sum_{i=1}^N f_i(q^0)q_i^1/f(q^0)]^{1/2}\, /[\sum_{i=1}^N f_i(q^1)q_i^0 /f(q^1)]^{1/2} \qquad \text{using (18) for } t = 1$$
$$= [\textstyle\sum_{i=1}^N\sum_{k=1}^N a_{ik}\, q_k^0q_i^1/[f(q^0)]^2]^{1/2}\, /[\sum_{i=1}^N \sum_{k=1}^N a_{ik}\, q_k^1q_i^0 /[f(q^1)]^2]^{1/2} \qquad \text{using (29)}$$

$= [1/[f(q^0)]^2]^{1/2} / [1/[f(q^1)]^2]^{1/2}$ using $a_{ik} = a_{ki}$ and canceling terms
$= f(q^1)/f(q^0)$.

Thus under the assumption that the consumer engages in cost minimizing behavior during periods 0 and 1 and has preferences over the N commodities that correspond to the utility function defined by (27), the Fisher ideal quantity index Q_F is *exactly* equal to the true quantity index, $f(q^1)/f(q^0)$.[15]

As was noted in Chapter 3 above, the price index that corresponds to the Fisher quantity index Q_F using the product test is the Fisher price index P_F. Let c(p) be the unit cost function that corresponds to the homogeneous quadratic utility function f defined by (27). Then using (24), the product test and (30), it can be seen that

(31) $P_F(p^0,p^1,q^0,q^1) = c(p^1)/c(p^0)$.

Thus under the assumption that the consumer engages in cost minimizing behavior during periods 0 and 1 and has preferences over the N commodities that correspond to the utility function defined by (27), the Fisher ideal price index P_F is exactly equal to the true price index, $c(p^1)/c(p^0)$.

A twice continuously differentiable function f(q) of N variables $q \equiv (q_1,\ldots,q_N)$ can provide a *second order approximation* to another such function f*(q) around the point q* if the level and all of the first and second order partial derivatives of the two functions coincide at q*. It can be shown[16] that the homogeneous quadratic function f defined by (27) can provide a second order approximation to an arbitrary f* around any (strictly positive) point q* in the class of linearly homogeneous functions. Thus the homogeneous quadratic functional form defined by (27) is a *flexible functional form*.[17] Diewert (1976; 117) termed an index number formula $Q_F(p^0,p^1,q^0,q^1)$ that was *exactly* equal to the true quantity index $f(q^1)/f(q^0)$ (where f is a flexible functional form) *a superlative index number formula*.[18] Equation (30) and the fact that the homogeneous quadratic function f defined by (27) is a flexible functional form shows that the Fisher ideal quantity index Q_F defined by the first line of (30) is a superlative index number formula. Since the Fisher ideal price index P_F also satisfies (31) where c(p) is the unit cost function that is generated by the homogeneous quadratic utility function, we also call P_F a superlative index number formula.

It is possible to show that the Fisher ideal price index is a superlative index number formula by a different route. Instead of starting with the assumption that the consumer's utility function is the homogeneous quadratic function defined by (27), we can start with the assumption that the consumer's unit cost function is a homogeneous quadratic. Thus we suppose that the consumer has the following unit cost function:

(32) $c(p_1,\ldots,p_N) \equiv [\sum_{i=1}^N \sum_{k=1}^N b_{ik}\, p_i p_k]^{1/2}$

---

[15] For the early history of this result, see Diewert (1976; 184).
[16] See Diewert (1976; 130) and let the parameter r equal 2.
[17] Diewert (1974; 133) introduced this term to the economics literature.
[18] Fisher (1922; 247) used the term superlative to describe the Fisher ideal price index. Thus Diewert adopted Fisher's terminology but attempted to give some precision to Fisher's definition of superlativeness. Fisher defined an index number formula to be superlative if it approximated the corresponding Fisher ideal results using his data set.

where the parameters b_{ik} satisfy the following symmetry conditions:

(33) $b_{ik} = b_{ki}$ for all i and k.

Differentiating c(p) defined by (32) with respect to p_i yields the following equations:

$$(34)\quad c_i(p) = (1/2)[\sum_{j=1}^N\sum_{k=1}^N b_{jk}\, p_jp_k]^{-1/2}\, 2\sum_{k=1}^N b_{ik}\, p_k\,; \qquad i = 1,\ldots,N$$
$$= \sum_{k=1}^N b_{ik}\, p_k / c(p) \qquad \text{using (32)}$$

where $c_i(p) \equiv \partial c(p^t)/\partial p_i$. In order to obtain the first equation in (34), we need to use the symmetry conditions, (33). Now evaluate the second equation in (34) at the observed period t price vector $p^t \equiv (p_1^t,\ldots,p_N^t)$ and divide both sides of the resulting equation by $c(p^t)$. We obtain the following equations:

$$(35)\quad c_i(p^t)/c(p^t) = \sum_{k=1}^N b_{ik}\, p_k^t / [c(p^t)]^2 \qquad t = 0,1\,;\ i = 1,\ldots,N.$$

As we are assuming cost minimizing behavior for the consumer in periods 0 and 1 and since the unit cost function c defined by (32) is differentiable, equations (25) (Shephard's Lemma) will hold. Now recall the definition of the Fisher ideal price index, P_F given by the first line of (36) below:

$$(36)\quad P_F(p^0,p^1,q^0,q^1) = [\sum_{i=1}^N p_i^1q_i^0/\sum_{k=1}^N p_k^0q_k^0]^{1/2}\, [\sum_{i=1}^N p_i^1q_i^1/\sum_{k=1}^N p_k^0q_k^1]^{1/2}$$
$$= [\sum_{i=1}^N p_i^1c_i(p^0)/c(p^0)]^{1/2}\, [\sum_{i=1}^N p_i^1q_i^1/\sum_{k=1}^N p_k^0q_k^1]^{1/2} \qquad \text{using (25) for } t = 0$$
$$= [\sum_{i=1}^N p_i^1c_i(p^0)/c(p^0)]^{1/2} /[\sum_{k=1}^N p_k^0q_k^1 /\sum_{i=1}^N p_i^1q_i^1]^{1/2}$$
$$= [\sum_{i=1}^N p_i^1c_i(p^0)/c(p^0)]^{1/2} /[\sum_{i=1}^N p_i^0c_i(p^1)/c(p^1)]^{1/2} \qquad \text{using (25) for } t = 1$$
$$= [\sum_{i=1}^N\sum_{k=1}^N b_{ik}\, p_k^0p_i^1/[c(p^0)]^2]^{1/2} /[\sum_{i=1}^N \sum_{k=1}^N b_{ik}\, p_k^1p_i^0/[c(p^1)]^2]^{1/2} \qquad \text{using (35)}$$
$$= [1/[c(p^0)]^2]^{1/2} / [1/[c(p^1)]^2]^{1/2} \qquad \text{using (33) and canceling terms}$$
$$= c(p^1)/c(p^0).$$

Thus under the assumption that the consumer engages in cost minimizing behavior during periods 0 and 1 and has preferences over the n commodities that correspond to the unit cost function defined by (32), the Fisher ideal price index P_F is *exactly* equal to the true price index, $c(p^1)/c(p^0)$.[19]

Problem 4: Suppose the consumer's utility function is defined as $f(q) \equiv [q^TAq]^{1/2}$ where $A = A^T$ and A^{-1} exists. Let p be a strictly positive vector of commodity prices and use calculus to solve the following constrained minimization problem:

(i) $\min_q \{p^Tq : [q^TAq]^{1/2} = 1\} \equiv c(p)$.

Show that $c(p) = [p^TA^{-1}p]^{1/2}$.

Since the homogeneous quadratic unit cost function c(p) defined by (32) is also a flexible functional form, the fact that the Fisher ideal price index P_F exactly equals the true price index $c(p^1)/c(p^0)$ means that P_F is a *superlative index number formula*.[20]

[19] This result was obtained by Diewert (1976; 133-134).

Suppose that the b_{ik} coefficients in (32) satisfy the following restrictions:

$$(37)\quad b_{ik} = b_i b_k \qquad \text{for } i,k = i,\ldots,N$$

where the N numbers b_i are nonnegative. In this special case of (32), it can be seen that the unit cost function simplifies as follows:

$$\begin{aligned}(38)\quad c(p_1,\ldots,p_n) &\equiv [\textstyle\sum_{i=1}^N \sum_{k=1}^N b_i b_k\, p_i p_k]^{1/2} \\ &= [\textstyle\sum_{i=1}^N b_i\, p_i \sum_{k=1}^N b_k\, p_k]^{1/2} \\ &= \textstyle\sum_{i=1}^N b_i\, p_i\, .\end{aligned}$$

Substituting (38) into Shephard's Lemma (20) yields the following expressions for the period t quantity vectors, q^t:

$$(39)\quad q_i^{\,t} = u^t\, \partial c(p^t)/\partial p_i = b_i\, u^t \qquad i = 1,\ldots,N\, ;\, t = 0,1.$$

Thus if the consumer has the preferences that correspond to the unit cost function defined by (32) where the b_{ik} satisfy the restrictions (37), then the period 0 and 1 quantity vectors are equal to a multiple of the vector $b \equiv (b_1,\ldots,b_N)$; i.e., $q^0 = b\, u^0$ and $q^1 = b\, u^1$. Under these assumptions, the Fisher, Paasche and Laspeyres indices, P_F, P_P and P_L, *all coincide.* However, the preferences which correspond to the unit cost function defined by (38) are not consistent with normal consumer behavior since they imply that the consumer will not substitute away from more expensive commodities to cheaper commodities if relative prices change going from period 0 to 1.

Problem 5: (a) Show that the linear utility function f* defined as

(i) $f^*(q) \equiv a^T q = \sum_{n=1}^N a_n q_n$ where $a_n > 0$ for each n

is a special case of the homogeneous quadratic function f(q) defined by (27).
(b) Consider the unit cost minimization problem that corresponds to the utility function defined by (i) above: i.e., for $p^* >> 0_N$, define

(ii) $c(p^*) \equiv \min_q \{p^{*T}q: f^*(q) = 1\}$.

Show that if $q^* >> 0_N$ is a solution to (ii), then p* is proportional to the vector a which occurs in (i).

5. Quadratic Mean of Order r Superlative Indexes

It turns out that there are many other superlative index number formulae; i.e., there exist many quantity indexes $Q(p^0,p^1,q^0,q^1)$ that are exactly equal to $f(q^1)/f(q^0)$ and many price indexes

[20] Note that we have shown that the Fisher index P_F is exact for the preferences defined by (27) as well as the preferences that are dual to the unit cost function defined by (32). These two classes of preferences do not coincide in general. However, if the N by N symmetric matrix A of the a_{ik} has an inverse, then it can readily be shown that the N by N matrix B of the b_{ik} will equal A^{-1}. See Problem 4 above.

$P(p^0,p^1,q^0,q^1)$ that are exactly equal to $c(p^1)/c(p^0)$ where the aggregator function f or the unit cost function c is a flexible functional form. We will define two families of superlative indexes below.

Suppose the consumer has the *following quadratic mean of order r utility function*:[21]

(40) $f^r(q_1,\ldots,q_N) \equiv [\sum_{i=1}^N\sum_{k=1}^N a_{ik}\, q_i^{r/2}\, q_k^{r/2}\,]^{1/r}$

where the parameters a_{ik} satisfy the symmetry conditions $a_{ik} = a_{ki}$ for all i and k and the parameter r satisfies the restriction $r \neq 0$. Diewert (1976; 130) showed that the utility function f^r defined by (40) is a flexible functional form; i.e., it can approximate an arbitrary twice continuously differentiable linearly homogeneous functional form to the second order. Note that when $r = 2$, f^r equals the homogeneous quadratic function defined by (27) above.

Define the *quadratic mean of order r quantity index* Q^r by:

(41) $Q^r(p^0,p^1,q^0,q^1) \equiv \{\sum_{i=1}^N s_i^0\, (q_i^1/q_i^0)^{r/2}\}^{1/r}\, \{\sum_{i=1}^N s_i^1\, (q_i^1/q_i^0)^{-r/2}\}^{-1/r}$

where $s_i^t \equiv p_i^tq_i^t/\sum_{k=1}^N p_k^tq_k^t$ is the period t expenditure share for commodity i as usual. It can be verified that when $r = 2$, Q^r simplifies into Q_F, the Fisher ideal quantity index.

Using exactly the same techniques as were used in section 4 above, it can be shown that Q^r is exact for the aggregator function f^r defined by (40); i.e., we have

(42) $Q^r(p^0,p^1,q^0,q^1) = f^r(q^1)/f^r(q^0)$.

Problem 6: Show that (42) is true.

Thus under the assumption that the consumer engages in cost minimizing behavior during periods 0 and 1 and has preferences over the n commodities that correspond to the utility function defined by (40), the quadratic mean of order r quantity index Q_F is *exactly* equal to the true quantity index, $f^r(q^1)/f^r(q^0)$.[22] Since Q^r is exact for f^r and f^r is a flexible functional form, we see that the quadratic mean of order r quantity index Q^r is a *superlative index* for each $r \neq 0$. Thus there are an infinite number of superlative quantity indexes.

For each quantity index Q^r, we can use the product test in order to define the corresponding *implicit quadratic mean of order r price index* P^{r*}:

(43) $$\begin{aligned} P^{r*}(p^0,p^1,q^0,q^1) &\equiv \sum_{i=1}^N p_i^1q_i^1/\{\sum_{i=1}^N p_i^0q_i^0\, Q^r(p^0,p^1,q^0,q^1)\} \\ &= c^{r*}(p^1)/c^{r*}(p^0) \end{aligned}$$

where c^{r*} is the unit cost function that corresponds to the aggregator function f^r defined by (40) above. For each $r \neq 0$, the implicit quadratic mean of order r price index P^{r*} is also a superlative index.

---

[21] This terminology is due to Diewert (1976; 129).

[22] See Diewert (1976; 130).

When r = 2, Q^r defined by (41) simplifies to Q_F, the Fisher ideal quantity index and P^{r*} defined by 43) simplifies to P_F, the Fisher ideal price index. When r = 1, Q^r defined by (41) simplifies to:

$$(44)\quad Q^1(p^0,p^1,q^0,q^1) \equiv \{\sum_{i=1}^N s_i^0\,(q_i^1/q_i^0)^{1/2}\}/\{\sum_{i=1}^N s_i^1\,(q_i^1/q_i^0)^{-1/2}\}$$
$$=\{[\sum_{i=1}^N p_i^0q_i^0/\sum_{i=1}^N p_i^0q_i^0](q_i^1/q_i^0)^{1/2}\}/\{[\sum_{i=1}^N p_i^1q_i^1/\sum_{i=1}^N p_i^1q_i^1](q_i^1/q_i^0)^{-1/2}\}$$
$$=\{\sum_{i=1}^N p_i^0(q_i^0q_i^1)^{1/2}/\sum_{i=1}^N p_i^0q_i^0\}/\{\sum_{i=1}^N p_i^1(q_i^0q_i^1)^{1/2}/\sum_{i=1}^N p_i^1q_i^1\}$$
$$=[\sum_{i=1}^N p_i^1q_i^1/\sum_{i=1}^N p_i^0q_i^0]/P_W(p^0,p^1,q^0,q^1)$$

where P_W is the *Walsh* (1901) (1921) *price index* defined in previous chapters. Thus P^{1*} is equal to P_W, the *Walsh price index*, and hence it is also a superlative price index.

Suppose the consumer has the *following quadratic mean of order r unit cost function*:[23]

$$(45)\quad c^r(p_1,\ldots,p_N) \equiv [\sum_{i=1}^N\sum_{k=1}^N b_{ik}\,p_i^{r/2}\,p_k^{r/2}\,]^{1/r}$$

where the parameters b_{ik} satisfy the symmetry conditions $b_{ik} = b_{ki}$ for all i and k and the parameter r satisfies the restriction $r \neq 0$. Diewert (1976; 130) showed that the unit cost function c^r defined by (45) is a flexible functional form; i.e., it can approximate an arbitrary twice continuously differentiable linearly homogeneous functional form to the second order. Note that when r = 2, c^r equals the homogeneous quadratic unit cost function defined by (32) above.

Define the *quadratic mean of order r price index* P^r by:

$$(46)\quad P^r(p^0,p^1,q^0,q^1) \equiv \{\sum_{i=1}^N s_i^0\,(p_i^1/p_i^0)^{r/2}\}^{1/r}\,\{\sum_{i=1}^N s_i^1\,(p_i^1/p_i^0)^{-r/2}\}^{-1/r}$$

where $s_i^t \equiv p_i^tq_i^t/\sum_{k=1}^N p_k^tq_k^t$ is the period t expenditure share for commodity i as usual. It can be verified that when r = 2, P^r simplifies into P_F, the Fisher ideal price index.

Using exactly the same techniques as were used in section 4 above, it can be shown that P^r is exact for the unit cost function c^r defined by (45); i.e., we have

$$(47)\quad P^r(p^0,p^1,q^0,q^1) = c^r(p^1)/c^r(p^0).$$

Problem 7: Show that (47) is true.

Thus under the assumption that the consumer engages in cost minimizing behavior during periods 0 and 1 and has preferences over the n commodities that correspond to the unit cost function defined by (45), the quadratic mean of order r price index P_F is *exactly* equal to the true price index, $c^r(p^1)/c^r(p^0)$.[24] Since P^r is exact for c^r and c^r is a flexible functional form, we see that the quadratic mean of order r price index P^r is a *superlative index* for each $r \neq 0$. Thus there are an infinite number of superlative price indexes.

[23] This terminology is due to Diewert (1976; 130). This unit cost function was first defined by Denny (1974).
[24] See Diewert (1976; 133-134).

For each price index P^r, we can use the product test in order to define the corresponding *implicit quadratic mean of order r quantity index* Q^{r*}:

$$(48)\quad Q^{r*}(p^0,p^1,q^0,q^1) \equiv \sum_{i=1}^N p_i^1q_i^1/\{\sum_{i=1}^N p_i^0q_i^0\, P^r(p^0,p^1,q^0,q^1)\} = f^{r*}(q^1)/f^{r*}(q^0)$$

where f^{r*} is the aggregator function that corresponds to the unit cost function c^r defined by (45) above.[25] For each $r \neq 0$, the implicit quadratic mean of order r quantity index Q^{r*} is also a superlative index.

When r = 2, P^r defined by (46) simplifies to P_F, the Fisher ideal price index and Q^{r*} defined by (48) simplifies to Q_F, the Fisher ideal quantity index. When r = 1, P^r defined by (46) simplifies to:

$$\begin{aligned}(49)\quad P^1(p^0,p^1,q^0,q^1) &\equiv \{\sum_{i=1}^N s_i^0\,(p_i^1/p_i^0)^{1/2}\}/\{\sum_{i=1}^N s_i^1\,(p_i^1/p_i^0)^{-1/2}\}\\ &=\{[\sum_{i=1}^N p_i^0q_i^0/\sum_{i=1}^N p_i^0q_i^0](p_i^1/p_i^0)^{1/2}\}/\{[\sum_{i=1}^N p_i^1q_i^1/\sum_{i=1}^N p_i^1q_i^1](p_i^1/p_i^0)^{-1/2}\}\\ &=\{\sum_{i=1}^N q_i^0(p_i^0p_i^1)^{1/2}/\sum_{i=1}^N p_i^0q_i^0\}/\{\sum_{i=1}^N q_i^1(p_i^0p_i^1)^{1/2}/\sum_{i=1}^N p_i^1q_i^1\}\\ &= [\sum_{i=1}^N p_i^1q_i^1/\sum_{i=1}^N p_i^0q_i^0]/Q_W(p^0,p^1,q^0,q^1)\end{aligned}$$

where Q_W is the *Walsh quantity index* defined in previous chapters. Thus Q^{1*} is equal to Q_W, the *Walsh quantity index*, and hence it is also a superlative quantity index.

6. Superlative Indexes: The Törnqvist index

In this section, we will revert to the assumptions made on the consumer in section 1 above. In particular, we do not assume that the consumer's utility function f is necessarily linearly homogeneous as in sections 2-5 above.

Before we derive our main result, we require a preliminary result. Suppose the function of N variables, $f(z_1,\ldots,z_N) \equiv f(z)$, is quadratic; i.e.,

$$(50)\quad f(z_1,\ldots,z_N) \equiv a_0 + \sum_{i=1}^N a_i z_i + (1/2)\sum_{i=1}^N \sum_{k=1}^N a_{ik}\, z_iz_k\ ;\ a_{ik} = a_{ki} \text{ for all i and k,}$$

where the a_i and the a_{ik} are constants. Let $f_i(z)$ denote the first order partial derivative of f evaluated at z with respect to the ith component of z, z_i. Let $f_{ik}(z)$ denote the second order partial derivative of f with respect to z_i and z_k. Then it is well known that the second order Taylor series approximation to a quadratic function is *exact*; i.e., if f is defined by (50) above, then for any two points, z^0 and z^1, we have

$$(51)\quad f(z^1) - f(z^0) = \sum_{i=1}^N f_i(z^0)[z_i^1-z_i^0] + (1/2)\sum_{i=1}^N \sum_{k=1}^N f_{ik}(z^0)[z_i^1-z_i^0][z_k^1-z_k^0].$$

It is less well known that *an average of two first order Taylor series approximations* to a quadratic function is also *exact*; i.e., if f is defined by (50) above, then for any two points, z^0 and z^1, we have[26]

[25] The function f^{r*} can be defined by using c^r as follows: $f^{r*}(q) \equiv 1/\max_p \{\sum_{i=1}^N p_iq_i : c^r(p) = 1\}$.

(52) $f(z^1) - f(z^0) = (1/2)\sum_{i=1}^N [f_i(z^0) + f_i(z^1)][z_i^1 - z_i^0]$.

Diewert (1976; 118) and Lau (1979) showed that equation (52) characterized a quadratic function and called the equation the *quadratic approximation lemma*. We will be more brief and refer to (52) as the *quadratic identity*.

We now suppose that the consumer's *cost function*,[27] C(u,p), has the following *translog functional form*:[28]

$$(53)\quad \ln C(u,p) \equiv a_0 + \sum_{i=1}^N a_i \ln p_i + (1/2) \sum_{i=1}^N \sum_{k=1}^N a_{ik} \ln p_i \ln p_k + b_0 \ln u + \sum_{i=1}^N b_i \ln p_i \ln u + (1/2)\, b_{00} [\ln u]^2$$

where ln is the natural logarithm function and the parameters a_i, a_{ik}, and b_i satisfy the following restrictions:

(54) $a_{ik} = a_{ki}$; $i,k = 1,\ldots,N$;
(55) $\sum_{i=1}^N a_i = 1$;
(56) $\sum_{i=1}^N b_i = 0$;
(57) $\sum_{k=1}^N a_{ik} = 0$; $i = 1,\ldots,N$.

The parameter restrictions (55)-(57) ensure that C(u,p) defined by (53) is linearly homogeneous in p, a property that a cost function must have. It can be shown that the translog cost function defined by (53)-(57) can provide a second order Taylor series approximation to an arbitrary cost function.[29]

We assume that the consumer has preferences that correspond to the translog cost function and that the consumer engages in cost minimizing behavior during periods 0 and 1. Let p^0 and p^1 be the period 0 and 1 observed price vectors and let q^0 and q^1 be the period 0 and 1 observed quantity vectors. Thus we have:

(58) $C(u^0,p^0) = \sum_{i=1}^N p_i^0 q_i^0$ and $C(u^1,p^1) = \sum_{i=1}^N p_i^1 q_i^1$

where C is the translog cost function defined above. We can also apply Shephard's lemma, (20) above:

$$(59)\quad q_i^t = \partial C(u^t,p^t)/\partial p_i = [C(u^t,p^t)/p_i^t]\, \partial \ln C(u^t,p^t)/\partial \ln p_i \; ; \qquad i = 1,\ldots,N \; ; \; t = 0,1$$

---

[26] To prove that (51) and (52) are true, use (50) and substitute into the left hand sides of (51) and (52). Then calculate the partial derivatives of the quadratic function defined by (50) and substitute these derivatives into the right hand side of (51) and (52).
[27] The consumer's cost function was defined by (1) above.
[28] Christensen, Jorgenson and Lau (1971) introduced this function into the economics literature.
[29] It can also be shown that if $b_0 = 1$ and all of the $b_i = 0$ for $i = 1,\ldots,N$ and $b_{00} = 0$, then $C(u,p) = uC(1,p) \equiv uc(p)$; i.e., with these additional restrictions on the parameters of the general translog cost function, we have homothetic preferences. Note that we also assume that utility u is scaled so that u is always positive.

Now use (58) to replace $C(u^t,p^t)$ in (59). After some cross multiplication, equations (59) become the following system of equations:

$$(60)\quad p_i^t q_i^t / \sum_{k=1}^N p_k^1 q_k^1 \equiv s_i^t = \partial \ln C(u^t,p^t)/\partial \ln p_i \; ; \qquad i = 1,\ldots,n \; ; t = 0,1 \text{ or}$$

$$(61)\quad s_i^t = a_i + \sum_{k=1}^N a_{ik} \ln p_k^t + b_i \ln u^t \; ; \qquad i = 1,\ldots,N \; ; t = 0,1$$

where s_i^t is the period t expenditure share on commodity i and (61) follows from (60) by differentiating (53) with respect to $\ln p_i$.

Define the geometric average of the period 0 and 1 utility levels as u*; i.e., define

$$(62)\quad u^* \equiv [u^0 u^1]^{1/2} .$$

Now observe that the right hand side of the equation that defines the natural logarithm of the translog cost function, equation (53), is a quadratic function of the variables $z_i \equiv \ln p_i$ if we hold utility constant at the level u*. Hence we can apply the quadratic identity, (52), and get the following equation:

$$(63)\quad \ln C(u^*,p^1) - \ln C(u^*,p^0)$$
$$= (1/2)\sum_{i=1}^N [\partial \ln C(u^*,p^0)/\partial \ln p_i + \partial \ln C(u^*,p^1)/\partial \ln p_i][\ln p_i^1 - \ln p_i^0]$$
$$= (1/2)\sum_{i=1}^N [a_i + \sum_{k=1}^N a_{ik} \ln p_k^0 + b_i \ln u^* + a_i + \sum_{k=1}^N a_{ik} \ln p_k^1 + b_i \ln u^*][\ln p_i^1 - \ln p_i^0]$$
differentiating (53) at the points (u^*,p^0) and (u^*,p^1)
$$= (1/2)\sum_{i=1}^N [a_i + \sum_{k=1}^N a_{ik} \ln p_k^0 + b_i \ln [u^0u^1]^{1/2} + a_i + \sum_{k=1}^N a_{ik} \ln p_k^1 + b_i \ln [u^0u^1]^{1/2}][\ln p_i^1 - \ln p_i^0]$$
using definition (62) for u*
$$= (1/2)\sum_{i=1}^N [a_i + \sum_{k=1}^N a_{ik} \ln p_k^0 + b_i \ln u^0 + a_i + \sum_{k=1}^N a_{ik} \ln p_k^1 + b_i \ln u^1][\ln p_i^1 - \ln p_i^0]$$
rearranging terms
$$= (1/2)\sum_{i=1}^N [\partial \ln C(u^0,p^0)/\partial \ln p_i + \partial \ln C(u^1,p^1)/\partial \ln p_i][\ln p_i^1 - \ln p_i^0]$$
differentiating (53) at the points (u^0,p^0) and (u^1,p^1)
$$= (1/2)\sum_{i=1}^N [s_i^0 + s_i^1][\ln p_i^1 - \ln p_i^0] \qquad \text{using (60).}$$

The last equation in (63) can be recognized as the logarithm of the Törnqvist-Theil index number formula P_T defined in earlier chapters. Hence exponentiating both sides of (63) yields the following equality between the true cost of living between periods 0 and 1, evaluated at the intermediate utility level u* and the observable Törnqvist (1936) (1937) Theil (1967) index P_T:[30]

$$(64)\quad C(u^*,p^1)/C(u^*,p^0) = P_T(p^0,p^1,q^0,q^1).$$

Since the translog cost function which appears on the left hand side of (64) is a flexible functional form, the Törnqvist-Theil price index P_T is also a *superlative index*.

It is somewhat mysterious how a ratio of *unobservable* cost functions of the form appearing on the left hand side of the above equation can be *exactly* estimated by an *observable* index number formula but the key to this mystery is the assumption of cost minimizing behavior and the quadratic

[30] This result is due to Diewert (1976; 122).

identity (52) along with the fact that derivatives of cost functions are equal to quantities, as specified by Shephard's lemma, (20). In fact, all of the exact index number results derived in sections 4 and 5 can be derived using transformations of the quadratic identity along with Shephard's lemma (or Wold's identity, (18) above).[31] Fortunately, for most empirical applications, assuming that the consumer has (transformed) quadratic preferences will be an adequate assumption so the results presented in sections 3 to 6 are quite useful to index number practitioners who are willing to adopt the economic approach to index number theory.[32] Essentially, the economic approach to index number theory provides a strong justification for the use of the Fisher price index P_F, the Törnqvist-Theil price index P_T, the implicit quadratic mean of order r price indexes P^{r*} defined by (43) (when r = 1, this index is the Walsh price index P_W) and the quadratic mean of order r price indexes P^r defined by (46).

7. The Lloyd Moulton Index Number Formula

The index number formula that will be discussed in this section on the single household economic approach to index number theory is a potentially very useful one for statistical agencies that are faced with the problem of producing a CPI in a timely manner. The Lloyd Moulton formula that will be discussed in this section makes use of the same information that is required in order to implement a Laspeyres index except that one additional piece of information is required.

In this section, the same assumptions on the consumer are made that were made in section 2 above. In particular, it is assumed that the consumer's utility function f(q) is linearly homogeneous[33] and the corresponding unit cost function is c(p). It is supposed that the unit cost function has the following functional form:

$$(65)\ c(p) \equiv \alpha_0 \left[\sum_{n=1}^N \alpha_n p_n^{1-\sigma}\right]^{1/(1-\sigma)} \quad \text{if } \sigma \neq 1;$$
$$\equiv \alpha_0 \prod_{n=1}^N p_n^{\alpha_n} \quad \text{if } \sigma = 1$$

where the α_i and σ are nonnegative parameters with $\sum_{i=1}^N \alpha_i = 1$. The unit cost function defined by (65) corresponds to *a Constant Elasticity of Substitution (CES) aggregator function* which was introduced into the economics literature by Arrow, Chenery, Minhas and Solow (1961)[34]. The parameter σ is the *elasticity of substitution*; when $\sigma = 0$, the unit cost function defined by (65) becomes linear in prices and hence corresponds to a fixed coefficients aggregator function which exhibits 0 substitutability between all commodities. When $\sigma = 1$,the corresponding aggregator or utility function is a Cobb-Douglas function. When σ approaches $+\infty$, the corresponding aggregator function f approaches a linear aggregator function which exhibits infinite substitutability between each pair of inputs. The CES unit cost function defined by (65) is *not* a fully flexible functional

[31] See Diewert (2000). Wold's identity says that derivatives of the utility function are proportional to prices.

[32] However, if consumer preferences are nonhomothetic and the change in utility is substantial between the two situations being compared, then we may want to compute separately the Laspeyres-Konüs and Paasche-Konüs true cost of living indexes defined above by (3) and (4), $C(u^0,p^1)/C(u^0,p^0)$ and $C(u^1,p^1)/C(u^1,p^0)$ respectively. In order to do this, we would have to use econometrics and estimate empirically the consumer's cost or expenditure function. However, if we are willing to make the assumption that the consumer's cost function can be adequately represented by a general translog cost function, then we can use (64) to calculate the true index $C(u^*,p^1)/C(u^*,p^0)$ where $u^* \equiv [u^0u^1]^{1/2}$.

[33] Thus homothetic preferences are assumed in this section.

[34] In the mathematics literature, this aggregator function or utility function is known as a mean of order r; see Hardy, Littlewood and Polyá (1934; 12-13).

form (unless the number of commodities N being aggregated is 2) but it is considerably more flexible than the zero substitutability aggregator function (this is the special case of (65) where σ is set equal to zero) that is exact for the Laspeyres and Paasche price indexes.

Under the assumption of cost minimizing behavior in period 0, Shephard's Lemma, (20) above, tells us that the observed first period consumption of commodity i, q_i^0, will be equal to $u^0\ \partial c(p^0)/\partial p_i$ where $\partial c(p^0)/\partial p_i$ is the first order partial derivative of the unit cost function with respect to the ith commodity price evaluated at the period 0 prices and $u^0 = f(q^0)$ is the aggregate (unobservable) level of period 0 utility. Using the CES functional form defined by (65) and assuming that $\sigma \neq 1$, the following equations are obtained:

$$(66)\ q_i^0 = u^0\alpha_0\,[\textstyle\sum_{n=1}^N \alpha_n\,(p_n^0)^r]^{(1/r)-1}\alpha_i\,(p_i^0)^{r-1}; \qquad r \equiv 1-\sigma \neq 0\ ;\ i = 1,\ldots,N$$
$$= u^0\ c(p^0)\ \alpha_i\,(p_i^0)^{r-1}/\textstyle\sum_{n=1}^N \alpha_n\,(p_n^0)^r\ .$$

These equations can be rewritten as

$$(67)\ p_i^0q_i^0/u^0c(p^0) = \alpha_i\,(p_i^0)^r/\textstyle\sum_{n=1}^N \alpha_n\,(p_n^0)^r\ ; \qquad i = 1,\ldots,N.$$

Now consider the following *Lloyd* (1975) *Moulton* (1996) *index number formula*:

$$(68)\ P_{LM}(p^0,p^1,q^0,q^1) \equiv [\textstyle\sum_{i=1}^N s_i^0\ (p_i^1/p_i^0)^{1-\sigma}]^{1/(1-\sigma)}\ ; \qquad \sigma \neq 1$$

where s_i^0 is the period 0 expenditure share of commodity i as usual:

$$(69)\ s_i^0 \equiv p_i^0q_i^0/\textstyle\sum_{k=1}^N p_k^0q_k^0\ ; \qquad i = 1,\ldots,N$$
$$= p_i^0q_i^0/u^0c(p^0) \qquad \text{using the assumption of cost minimizing behavior}$$
$$= \alpha_i\,(p_i^0)^r/\textstyle\sum_{n=1}^N \alpha_n\,(p_n^0)^r \qquad \text{using (67).}$$

If (69) is substituted into (68), it is found that:

$$(70)\ P_{LM}(p^0,p^1,q^0,q^1) \equiv [\textstyle\sum_{i=1}^N s_i^0\ (p_i^1/p_i^0)^r]^{1/r} \qquad \text{using } r \equiv 1-\sigma$$
$$= [\textstyle\sum_{i=1}^N \{\alpha_i\,(p_i^0)^r/\sum_{n=1}^N \alpha_n\,(p_n^0)^r\}(p_i^1/p_i^0)^r]^{1/r} \qquad \text{using (69)}$$
$$= [\textstyle\sum_{i=1}^N \alpha_i\,(p_i^1)^r/\sum_{n=1}^N \alpha_n\,(p_n^0)^r]^{1/r}$$
$$= [\textstyle\sum_{i=1}^N \alpha_i\,(p_i^1)^r]^{1/r}\ /\ [\sum_{n=1}^N \alpha_n\,(p_n^0)^r]^{1/r}$$
$$= \alpha_0[\textstyle\sum_{i=1}^N \alpha_i\,(p_i^1)^r]^{1/r}\ /\ \alpha_0[\sum_{n=1}^N \alpha_n\,(p_n^0)^r]^{1/r}$$
$$= c(p^1)/c(p^0) \qquad \text{using } r \equiv 1-\sigma \text{ and (65).}$$

Equation (70) shows that the Lloyd Moulton index number formula P_{LM} is *exact* for CES preferences. Lloyd (1975) and Moulton (1996) independently derived this result but it was Moulton who appreciated the significance of the formula (70) for statistical agency purposes. Note that in order to evaluate (68) numerically, it is necessary to have information on:

- base period expenditure shares s_i^0 ;

- the price relatives p_i^1/p_i^0 between the base period and the current period and
- an estimate of the elasticity of substitution between the commodities in the aggregate, σ.

The first two pieces of information are the standard information sets that statistical agencies use to evaluate the Laspeyres price index P_L (note that P_{LM} reduces to P_L if $\sigma = 0$). Hence, if the statistical agency is able to estimate the elasticity of substitution σ based on past experience[35], *then the Lloyd Moulton price index can be evaluated using essentially the same information set that is used in order to evaluate the traditional Laspeyres index*. Moreover, the resulting consumer price index will be free of *substitution bias* to a reasonable degree of approximation.[36] Of course, the practical problem with implementing this methodology is that estimates of the elasticity of substitution parameter σ are bound to be somewhat uncertain and hence the resulting Lloyd Moulton index may be subject to charges that it is not *objective* or *reproducible*. The statistical agency will have to balance the benefits of reducing substitution bias with these possible costs.

The final section in this chapter looks at some econometric approaches to the estimation of the elasticity of substitution parameter σ.

8. Econometric Approaches to the Estimation of the Elasticity of Substitution

Recall equations (67) above. The utility level u^0 on the left hand side of equations (67) can be eliminated if we set period 0 expenditures, $u^0c(p^0)$, equal to the observed expenditure, say Y^0, on the N commodities during period 0; i.e.,

$$(71)\ u^0c(p^0) = Y^0 \equiv \sum_{n=1}^N p_n^0q_n^0.$$

Substituting (71) into (67) leads to the following equation for the consumer's ith expenditure share in period 0, s_i^0:

$$(72)\ s_i^0 \equiv p_i^0q_i^0/ Y^0 = \alpha_i (p_i^0)^r / \Sigma_{k=1}^N \alpha_k (p_k^0)^r\ ; \qquad i = 1,2,\ldots,N.$$

Letting s_n^t be the expenditure share on commodity n in period t, we can derive the counterpart to equations (72) for period t as well:

$$(73)\ s_n^t \equiv p_n^tq_n^t/ Y^t = \alpha_n (p_n^t)^r / \Sigma_{k=1}^N \alpha_k (p_k^t)^r\ ; \qquad n = 1,2,\ldots,N\ ;\ t = 0,1,\ldots,T.$$

---

[35] For the first application of this methodology (in the context of the consumer price index), see Shapiro and Wilcox (1997; 121-123). They calculated superlative Törnqvist indexes for the U.S. for the years 1986-1995 and then calculated the Lloyd Moulton CES index for the same period using various values of σ. They then chose the value of σ (which was .7) which caused the CES index to most closely approximate the Törnqvist index. Essentially the same methodology was used by Alterman, Diewert and Feenstra (1999) in their study of U.S. import and export price indexes. For alternative methods for estimating σ, see Balk (2000) and the following section.

[36] What is a "reasonable" degree of approximation depends on the context. Assuming that consumers have CES preferences is not a reasonable assumption in the context of estimating elasticities of demand: at least a second order approximation to the consumer's preferences is required in this context. However, in the context of approximating changes in a consumer's expenditures on the N commodities under consideration, it is usually adequate to assume a CES approximation.

Note that the right hand side expressions in equations (73) are homogeneous of degree 0 in the positive parameters, $\alpha_1, \alpha_2, \ldots, \alpha_N$, so that we will require a normalization on these parameters in order to identify them (such as $\alpha_1 + \alpha_2 + \ldots + \alpha_N = 1$). Other than this normalization problem, equations (73) could have error terms appended to them and we could estimate the unknown α_n and r parameters using nonlinear regression techniques. However, in the standard CES estimation literature, this is not what is done. Usually, each equation in (73) is divided by say the 1st equation in period t in order to obtain the following system of equations:

(74) $s_n^t/s_1^t = \alpha_n (p_n^t)^r/\alpha_1 (p_1^t)^r$; $n = 2,\ldots,N$; $t = 0,1,\ldots,T$.

Now take logarithms on both sides of (74) and we obtain the following system of estimating equations after adding the error terms e_n^t:

(75) $\ln[s_n^t/s_1^t] = \beta_n + r \ln[p_n^t/p_1^t] + e_n^t$; $n = 2,\ldots,N$; $t = 0,1,\ldots,T$

where $\beta_n \equiv \ln\alpha_n - \ln\alpha_1$ for $n = 2,\ldots,N$. The system of estimating equations (75) is linear in the unknown r and β_n parameters and thus linear regression techniques can be used in the estimation procedure.

Now a problem with (75) is that the first commodity plays a asymmetric role in the estimation. In what follows, we attempt to find symmetric estimating equations.[37] Define the geometric mean of the α_n which occur in (73) by:

(76) $\alpha. \equiv [\prod_{n=1}^N \alpha_n]^{1/N}$.

Define the geometric mean of the period t expenditure shares s_n^t by:

(77) $s.^t \equiv [\prod_{n=1}^N s_n^t]^{1/N}$; $t = 0,1,\ldots,T$.

Finally, define the geometric mean of the period t prices p_n^t by:

(78) $p.^t \equiv [\prod_{n=1}^N p_n^t]^{1/N}$; $t = 0,1,\ldots,T$.

Using equations (73) and the above definitions, it can be seen that

(79) $s.^t = \alpha. (p.^t)^r / \Sigma_{k=1}^N \alpha_k (p_k^t)^r$; $t = 0,1,\ldots,T$.

Now take each of the N period t equations in (73) and divide by equation t in (79) in order to obtain the following system of equations:

(80) $s_n^t/s.^t = \alpha_n (p_n^t)^r/\alpha. (p.^t)^r$; $n = 1,\ldots,N$; $t = 0,1,\ldots,T$.

Finally, take logarithms of both sides of (80) and add the error terms e_n^t in order to obtain the following system of estimating equations:

[37] Our analysis is based on an approach that was originally suggested by James Cuthbert.

(81) $\ln[s_n^t/s_\cdot^t] = \gamma_n + r \ln[p_n^t/p_\cdot^t] + e_n^t$; $n = 1,\ldots,N$; $t = 0,1,\ldots,T$

where

(82) $\gamma_n \equiv \ln\alpha_n - \ln\alpha_\cdot$; $n = 1,\ldots,N$.

The system of estimating equations (81) is linear in the unknown r and γ_n parameters and thus linear regression techniques can be used in the estimation procedure.

Using definitions (76) and (82), it can be shown that:

(83) $\sum_{n=1}^{N} \gamma_n = 0$.

Using (79) and (80), we have:

(84) $0 = \sum_{n=1}^{N} \ln[s_n^t/s_\cdot^t]$
$= \sum_{n=1}^{N} \gamma_n + r \sum_{n=1}^{N} \ln[p_n^t/p_\cdot^t] + \sum_{n=1}^{N} e_n^t$; using (81)
$= 0 + r\,[0] + \sum_{n=1}^{N} e_n^t$ using (83) and (78)
$= \sum_{n=1}^{N} e_n^t$; $t = 0,1,\ldots,T$.

Thus for any period t, the errors sum to zero and hence *cannot* be independently distributed. Hence for each period t, we must drop one of the estimating equations in (81). Hence there will be $(N-1)(T+1)$ degrees of freedom to estimate the unknown $r = \sigma-1$ and the $N-1$ independent γ_n parameters.[38] Suppose the first equation is dropped in the estimation procedure. Then we have estimates for $\gamma_2^*, \gamma_3^*,\ldots, \gamma_N^*$ and we define γ_1^* using (83); i.e.,

(85) $\gamma_1^* \equiv - \sum_{n=2}^{N} \gamma_n^*$.

The corresponding estimated α_n^* parameters are now defined as follows:

(86) $\alpha_n^* \equiv \exp[\gamma_n^*] > 0$; $n = 1,\ldots,N$.

It turns out that the estimated α_n^* satisfy the following normalization:

(87) $1 = \alpha_\cdot^* \equiv [\alpha_1^*\alpha_2^*\ldots\alpha_N^*]^{1/N}$.

Of course, the important parameter is r and standard errors for it can readily be obtained.

Let us return to the regression model (81) and discuss the distributional assumptions on the error terms, e_n^t. As mentioned above, we cannot simply assume that the e_n^t are independently distributed over time due to the constraints (84) on these errors. Define (the transpose of) the period t vector of errors, e^t, as

[38] If we use a seemingly unrelated regression package to do the estimation, the estimates will be invariant to the equation that is dropped.

(88) $(e^t)^T \equiv [e_1^t,\ldots,e_N^t]$; $t = 0,1,\ldots,T$.

The simplest stochastic specification for the e^t is the following one:

(89) $Ee^t = 0_N$;[39]
(90) $Ee^t(e^t)^T \equiv \Sigma$

where Σ is a positive semidefinite symmetric matrix of rank $N-1$ that satisfies the restriction

(91) $\Sigma 1_N = 0_N$.

The resulting model can be estimated by maximum likelihood or the method of seemingly unrelated regressions after dropping any one of the N equations.[40]

A useful generalization of (81) is to add a time trend to the r parameter, leading to the following model:

(92) $\ln[s_n^t/s_\cdot^t] = \gamma_n + (r + st) \ln[p_n^t/p_\cdot^t] + e_n^t$; $n = 1,\ldots,N$; $t = 0,1,\ldots,T$

where s is a new time trending parameter. This model allows the elasticity of substitution to trend over time.[41] Again, the restrictions (84) on the error terms e_n^t will hold for this new model. We continue to make the stochastic specification (89)-(91). Again, one of the N equations can be dropped and the resulting model can be estimated by the seemingly unrelated regression model.[42]

References

Arrow, K.J., H.B. Chenery, B.S. Minhas and R.M. Solow (1961), "Capital-Labor Substitution and Economic Efficiency", *Review of Economics and Statistics* 63, 225-250.

Balk, B.M. (1998), *Industrial Price, Quantity and Productivity Indices*, Boston: Kluwer Academic Publishers.

Balk, B.M. (2000), "On Curing the CPI's Substitution and New Goods Bias", Research Paper 0005, Department of Statistical Methods, Statistics Netherlands, P.O. Box 4000, 2270 JM Voorburg, The Netherlands.

Christensen, L.R., D.W. Jorgenson and L.J. Lau (1971), "Conjugate Duality and the Transcendental Logarithmic Production Function," *Econometrica* 39, 255-256.

[39] This means the expectation of each component of e^t is 0. The vector 0_N is an N dimensional vector of 0's.
[40] The seemingly unrelated regression model is a standard option in the econometric programming packages TSP or SHAZAM.
[41] Various further generalizations of (92) could be considered, including spline models.
[42] Again, the estimates will be invariant to the equation that is dropped.

Denny, M. (1974), "The Relationship Between Functional Forms for the Production System", *Canadian Journal of Economics* 7, 21-31.

Debreu, G. (1959), *Theory of Value*, New York: John Wiley and Sons.

Diewert, W.E., 1974. "Applications of Duality Theory," pp. 106-171 in M.D. Intriligator and D.A. Kendrick (ed.), Frontiers of Quantitative Economics, Vol. II, Amsterdam: North-Holland.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 114-145.

Diewert, W.E. (1983a), "The Theory of the Cost of Living Index and the Measurement of Welfare Change", pp. 163-233 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada, reprinted as pp. 79-147 in *Price Level Measurement*, W.E. Diewert (ed.), Amsterdam: North-Holland, 1990.

Diewert, W.E. (1983b), "The Theory of the Output Price Index and the Measurement of Real Output Change", pp. 1049-1113 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

Diewert, W.E. (1993), "Duality Approaches To Microeconomic Theory", in *Essays in Index Number Theory*, pp. 105-175 in Volume I, Contributions to Economic Analysis 217, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North Holland.

Diewert, W.E. (2000), "The Quadratic Approximation Lemma and Decompositions of Superlative Indexes", Discussion Paper 00-15, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1, 25 pp.

Fisher, I. (1922), *The Making of Index Numbers*, Houghton-Mifflin, Boston.

Hardy, G.H., J.E. Littlewood and G. Polyá (1934), *Inequalities*, Cambridge: Cambridge University Press.

Konüs, A.A. (1924), "The Problem of the True Index of the Cost of Living", translated in *Econometrica* 7, (1939), 10-29.

Konüs, A.A. and S.S. Byushgens (1926), "K probleme pokupatelnoi cili deneg", *Voprosi Konyunkturi* 2, 151-172.

Lau, L.J. (1979), "On Exact Index Numbers", *Review of Economics and Statistics* 61, 73-82.

Lloyd, P.J. (1975), "Substitution Effects and Biases in Nontrue Price Indices", *American Economic Review* 65, 301-313.

Moulton, B.R. (1996), "Constant Elasticity Cost-of-Living Index in Share Relative Form", Bureau of Labor Statistics, Washington D.C., December.

Pollak, R.A. (1983), "The Theory of the Cost-of-Living Index", pp. 87-161 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada; reprinted as pp. 3-52 in R.A. Pollak, *The Theory of the Cost-of-Living Index*, Oxford: Oxford University Press, 1989; also reprinted as pp. 5-77 in *Price Level Measurement*, W.E. Diewert (ed.), Amsterdam: North-Holland, 1990.

Samuelson, P.A. (1953), "Prices of Factors and Goods in General Equilibrium', *Review of Economic Studies* 21, 1-20.

Samuelson, P.A. and S. Swamy (1974), "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis", *American Economic Review* 64, 566-593.

Shapiro, M.D. and D.W. Wilcox (1997), "Alternative Strategies for Aggregating Prices in the CPI", *Federal Reserve Bank of St. Louis Review* 79:3, 113-125.

Shephard, R.W. (1953), *Cost and Production Functions*, Princeton: Princeton University Press.

Shephard, R.W. (1970), *Theory of cost and Production Functions*, Princeton: Princeton University Press.

Törnqvist, Leo (1936), "The Bank of Finland's Consumption Price Index," *Bank of Finland Monthly Bulletin* 10: 1-8.

Törnqvist, L. and E. Törnqvist (1937), "Vilket är förhållandet mellan finska markens och svenska kronans köpkraft?", *Ekonomiska Samfundets Tidskrift* 39, 1-39 reprinted as pp. 121-160 in *Collected Scientific Papers of Leo Törnqvist*, Helsinki: The Research Institute of the Finnish Economy, 1981.

Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.

Walsh, C.M. (1921), *The Problem of Estimation*, London: P.S. King & Son.

Wold, H. (1944), "A Synthesis of Pure Demand Analysis, Part 3", *Skandinavisk Aktuarietidskrift* 27, 69-120.

Wold, H. (1953), *Demand Analysis*, New York: John Wiley.