THE ECONOMIC APPROACH TO INDEX NUMBER THEORY: THE MANY-HOUSEHOLD CASE 18

Introduction

18.1 In the previous chapter on the economic approach to index numbers, it was implicitly assumed that the economy behaved as if there were a single representative consumer. In the present chapter, the economic approach is extended to an economy with many household groups or many regions. In the algebra below, an arbitrary number of households, H say, is considered. In principle, each household in the economy under consideration could have its own consumer price index. In practice, however, it will be necessary to group households into various classes. Within each class, it will be necessary to assume that the group of households in the class behaves as if it were a single household in order to apply the economic approach to index number theory. The partition of the economy into H household classes can also be given a regional interpretation: each household class could be interpreted as a group of households within a region of the country under consideration.

18.2 The concepts of a *plutocratic index* and a *conditional index* are introduced in paragraphs 18.3 to 18.13. Using the plutocratic concept, each household in the economy is given a weight in the national index that is proportional to the household's expenditures on commodities for the two periods under consideration. A conditional index is an index that depends on environmental variables that might affect household expenditures on commodities. One example of an environmental variable is the weather: if the weather is cold, then households will spend more on heating fuel. In paragraphs 18.14 to 18.22 it is shown how a national Fisher price index can approximate a plutocratic cost of living index. Finally, paragraphs 18.23 to 18.35 consider an alternative conceptual framework for a national index, the *democratic index*. Using this index concept, each household in the economy is given an equal weight in the national index (as opposed to the plutocratic concept where households that spend more get a higher weight in the national index).

Plutocratic cost of living indices and observable bounds

18.3 In this section, an economic approach to the consumer price index (CPI) is considered that is based on the *plutocratic cost of living index* that was originally defined by Prais (1959). This concept was further refined by Pollak (1980, p. 276; 1981, p. 328), who defined his *Scitovsky–Laspeyres cost of living index* as the ratio of total expenditure required to enable each household in the economy under consideration to attain its base period indifference surface at period 1 prices to the corresponding expenditure required to attain the same standard of living using period 0 prices. In the following paragraph, this concept will be explained more fully.

18.4 Suppose that there are H households (or regions) in the economy and suppose further that there are n commodities in the economy in periods 0 and 1 that households consume *and* that we wish to include in our definition of the cost of living. Denote an n-dimensional vector of commodity consumption in a given period by $q \equiv (q_1, q_2, \ldots, q_n)$ as usual. Denote the vector of period t market prices faced by household h by $p_h^t \equiv (p_{h1}^t, p_{h2}^t, \ldots, p_{hN}^t)$ for $t = 0,1$. Note that it is *not* assumed that each household faces the same vector of commodity prices. In addition to the market commodities that are in the vector q, it is assumed that each household is affected by an M-dimensional vector of *environmental*[1] *or demographic*[2] *variables or public goods*, $e \equiv (e_1, e_2, \ldots, e_M)$. It is supposed that there are H households (or regions) in the economy during periods 0 and 1, and the preferences of household h over different combinations of market commodities q and environmental variables e can be represented by the continuous utility function $f^h(q,e)$ for $h = 1, 2, \ldots, H$.[3] For periods $t = 0,1$ and for households $h = 1, 2, \ldots, H$, it is assumed that the observed household h consumption vector $q_h^t \equiv (q_{h1}^t, \ldots, q_{hN}^t)$ is a solution to the following household h expenditure minimization problem:

$$\min_q \{p_h^t q : f^h(q, e_h^t) \geq u_h^t\} \equiv C^h(u_h^t, e_h^t, p_h^t);$$
$$t = 0,1;\ h = 1, 2, \ldots, H \qquad (18.1)$$

where e_h^t is the environmental vector facing household h in period t, $u_h^t \equiv f^h(q_h^t, e_h^t)$ is the utility level achieved by household h during period t and C^h is the cost or expenditure function that is dual to the utility function f^h.[4] Basically, these assumptions mean that

[1] This is the terminology used by Pollak (1989, p. 181) in his model of the conditional cost of living concept.

[2] Caves, Christensen and Diewert (1982a, p. 1409) used the terms demographic variables or public goods to describe the vector of conditioning variables e in their generalized model of the Konüs price index or cost of living index, while Diewert (2001) used the term environmental variables.

[3] It is assumed that each $f^h(q,e)$ is continuous and increasing in the components of q and e, and is concave in the components of q.

[4] In order to simplify notation, in this section the notation $pq = \sum_{i=1}^n p_i q_i$ as the inner product between the vectors p and q is used rather than the usual summation notation.

each household has stable preferences over the same list of commodities during the two periods under consideration, the same households appear in each period and each household chooses its consumption bundle in the most cost-efficient way during each period, conditional on the environmental vector that it faces during each period. Note again that the household (or regional) prices are in general different across households (or regions).

18.5 With the above assumptions in mind, the example of Pollak (1980; 1981) and Diewert (1983a, p. 190)[5] is followed. The class of *conditional plutocratic cost of living indices*, $P^*(p^0, p^1, u, e_1, e_2, \ldots, e_H)$, pertaining to periods 0 and 1 for the arbitrary utility vector of household utilities $u \equiv (u_1, u_2, \ldots, u_H)$ and for the arbitrary vectors of household environmental variables e_h for $h = 1, 2, \ldots, H$ is defined as follows:

$$P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u, e_1, e_2, \ldots, e_H) \equiv \frac{\sum_{h=1}^{H} C^h(u_h, e_h, p_h^1)}{\sum_{h=1}^{H} C^h(u_h, e_h, p_h^0)} \quad (18.2)$$

The numerator on the right-hand side of equation (18.2) is the sum over households of the minimum cost, $C^h(u_h, e_h, p_h^1)$, for household h to achieve the arbitrary utility level u_h, given that the household h faces the arbitrary vector of household h environmental variables e_h and also faces the period 1 vector of prices p_h^1. The denominator on the right-hand side of equation (18.2) is the sum over households of the minimum cost, $C^h(u_h, e_h, p_h^0)$, for household h to achieve the *same* arbitrary utility level u_h, given that the household faces the *same* arbitrary vector of household h environmental variables e_h and also faces the period 0 vector of prices p_h^0. Thus in the numerator and denominator of equation (18.2), only the price variables are different, which is precisely what is wanted in a theoretical definition of a consumer price index.

18.6 The general definition (18.2) is now specialized by replacing the general utility vector u by either the period 0 vector of household utilities $u^0 \equiv (u_1^0, u_2^0, \ldots u_H^0)$ or the period 1 vector of household utilities $u^1 \equiv (u_1^1, u_2^1, \ldots u_H^1)$. The general definition is also specialized by replacing the general household environmental vectors $(e_1, e_2, \ldots e_H) \equiv e$ by either the period 0 vector of household environmental variables $e^0 \equiv (e_1^0, e_2^0, \ldots e_H^0)$ or the period 1 vector of household environmental variables $e^1 \equiv (e_1^1, e_2^1, \ldots, e_H^1)$. The choice of the base period vector of utility levels and base period environmental variables leads to the *Laspeyres conditional plutocratic cost of living index*, $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^0, e^0)$.[6] The choice of the period 1 vector of utility levels and period 1 environmental variables leads to the *Paasche conditional plutocratic cost of living index*, $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^1, e^1)$. It turns out that these last two indices satisfy some interesting inequalities, which are derived below.

18.7 Using definition (18.2), the Laspeyres conditional plutocratic cost of living index, $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^0, e^0)$, may be written as follows:

$$P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^0, e_1^0, e_2^0, \ldots, e_H^0)$$

$$\equiv \frac{\sum_{h=1}^{H} C^h(u_h^0, e_h^0, p_h^1)}{\sum_{h=1}^{H} C^h(u_h^0, e_h^0, p_h^0)}$$

$$= \frac{\sum_{h=1}^{H} C^h(u_h^0, e_h^0, p_h^1)}{\sum_{h=1}^{H} p_h^0 q_h^0} \quad \text{using equation (18.1) for } t = 0$$

$$\leq \frac{\sum_{h=1}^{H} p_h^1 q_h^0}{\sum_{h=1}^{H} p_h^0 q_h^0} \quad (18.3)$$

since $C^h(u_h^0, e_h^0, p_h^1) \equiv \min_q \{p_h^1 q : f^h(q, e_h^0) \geq u_h^0\} \leq p^1 q_h^0$ and q_h^0 is feasible for the cost minimization problem for $h = 1, 2, \ldots, H$

$$\equiv P_{PL}$$

where P_{PL} is defined to be the observable (in principle) *plutocratic Laspeyres price index*, $\sum_{h=1}^{H} p_h^1 q_h^0 / \sum_{h=1}^{H} p_h^0 q_h^0$, which uses the individual vectors of household or regional quantities for period 0, $(q_1^0, \ldots, q_H^0)$, as quantity weights.[7]

18.8 If prices are equal across households (or regions), so that

$$p_h^t = p^t \quad \text{for } t = 0, 1 \text{ and } h = 1, 2, \ldots, H, \quad (18.4)$$

then the plutocratic (or disaggregated) Laspeyres price index, P_{PL}, collapses down to the usual aggregate Laspeyres index, P_L; i.e., then P_{PL} becomes

[5] These authors provided generalizations of the plutocratic cost of living index attributable to Prais (1959). Pollak and Diewert did not include the environmental variables in their definitions of a group cost of living index.

[6] This is the concept of a cost of living index that Triplett (2001) found most useful for measuring inflation: "One might want to produce a COL *conditional* on the base period's weather experience. . . . In this case, the unusually cold winter does not affect the *conditional* COL subindex that holds the environment constant. . . . the COL subindex that holds the environment constant is probably the COL concept that is most useful for an anti-inflation policy." Hill (1999, p. 4) endorsed this point of view.

[7] Thus the plutocratic Laspeyres index can be regarded as an ordinary Laspeyres index except that each commodity consumed by each household (or in each region) is regarded as a separate commodity.

$$P_{PL} \equiv \frac{\sum_{h=1}^{H} p_h^1 q_h^0}{\sum_{h=1}^{H} p_h^0 q_h^0} = \frac{p^1 \sum_{h=1}^{H} q_h^0}{p^0 \sum_{h=1}^{H} q_h^0} = \frac{p^1 q^0}{p^0 q^0} \equiv P_L \tag{18.5}$$

where the total quantity vector in period t is defined as

$$q^t \equiv \sum_{h=1}^{H} q_h^t \quad \text{for } t=0,1 \tag{18.6}$$

18.9 The inequality (18.3) says that the theoretical Laspeyres plutocratic conditional cost of living index, $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^0, e^0)$, is bounded from above by the observable (in principle) plutocratic or disaggregated Laspeyres price index, P_{PL}. The special case of inequality (18.3) when the equal prices assumption (18.4) holds[8] was first obtained by Pollak (1989, p. 182) for the case of one household with environmental variables and by Pollak (1980, p. 276) for the many-household case, but where the environmental variables are absent from the household utility and cost functions.

18.10 In a similar manner, specializing definition (18.2), the Paasche conditional plutocratic cost of living index, $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^1, e^1)$, may be written as follows:

$$P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^1, e_1^1, e_2^1, \ldots, e_H^1)$$
$$\equiv \frac{\sum_{h=1}^{H} C^h(u_h^1, e_h^1, p_h^1)}{\sum_{h=1}^{H} C^h(u_h^1, e_h^1, p_h^0)}$$
$$= \frac{\sum_{h=1}^{H} p_h^1 q_h^1}{\sum_{h=1}^{H} C^h(u_h^1, e_h^1, p_h^0)} \quad \text{using equation (18.1) for } t=1$$
$$\geq \frac{\sum_{h=1}^{H} p_h^1 q_h^1}{\sum_{h=1}^{H} p_h^0 q_h^1} \quad \text{using a feasibility argument}$$
$$\equiv P_{PP} \tag{18.7}$$

where P_{PP} is defined to be the plutocratic or disaggregated (over households) Paasche price index, $\sum_{h=1}^{H} p_h^1 q_h^1 / \sum_{h=1}^{H} p_h^0 q_h^1$, which uses the individual vectors of household quantities for period 1, $(q_1^1, \ldots, q_H^1)$, as quantity weights.

[8] The general case was obtained by Diewert (2001, p. 222).

18.11 If prices are equal across households (or regions), so that assumptions (18.4) hold, then the disaggregated Paasche price index P_{PP} collapses down to the usual aggregate Paasche index, P_P; i.e., then P_{PP} becomes

$$P_{PP} \equiv \frac{\sum_{h=1}^{H} p_h^1 q_h^1}{\sum_{h=1}^{H} p_h^0 q_h^1} = \frac{p^1 \sum_{h=1}^{H} q_h^1}{p^0 \sum_{h=1}^{H} q_h^1} = \frac{p^1 q^1}{p^0 q^1} \equiv P_P \tag{18.8}$$

18.12 Returning to the inequality (18.7), it can be seen that the theoretical Paasche conditional plutocratic cost of living index, $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^1, e^1)$, is bounded from below by the observable plutocratic or disaggregated Paasche price index, P_{PP}. Diewert (1983a, p. 191) first obtained the inequality (18.7) for the case where the environmental variables are absent from the household utility and cost functions, and prices are equal across households. The general case is attributable to Diewert (2001, p. 223).

18.13 In the following section, it will be shown how to obtain a theoretical plutocratic cost of living index that is bounded from above and below rather than the theoretical indices in inequalities (18.3) and (18.7) that just have the one-sided bounds.

The Fisher plutocratic price index

18.14 Using the inequalities (18.3) and (18.7) and the continuity properties of the conditional plutocratic cost of living $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u, e)$ defined by equation (18.2), it is possible to modify the method of proof used by Konüs (1924) and Diewert (1983a, p. 191) and establish the following result:[9]

There exists a reference utility vector $u^* \equiv (u_1^*, u_2^*, \ldots, u_H^*)$ such that the household h reference utility level u_h^* lies between the household h period 0 and 1 utility levels, u_h^0 and u_h^1 respectively for $h=1, \ldots, H$, and there exist household environmental vectors $e_h^* \equiv (e_{h1}^*, e_{h2}^*, \ldots, e_{hM}^*)$ such that the household h reference mth environmental variable e_{hm}^* lies between the household h period 0 and 1 levels for the mth environmental variable, e_{hm}^0 and e_{hm}^1 respectively for $m=1,2,\ldots,M$ and $h=1,\ldots,H$, and the conditional plutocratic cost of living index $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^*, e^*)$ evaluated at this intermediate reference utility vector u^* and the intermediate reference vector of household environmental variables

[9] See Diewert (2001, p. 223). Note that the household cost functions must be continuous in the environmental variables; this is a real restriction on the types of environmental variables which can be accommodated by the result.

$e^* \equiv (e_1^*, e_2^*, \ldots, e_H^*)$ lies between the observable (in principle) plutocratic Laspeyres and Paasche price indices, P_{PL} and P_{PP}, defined above by the last equalities in (18.3) and (18.7).

18.15 The above result says that the *theoretical national plutocratic conditional consumer price index* $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^*, e^*)$ lies between the plutocratic or disaggregated Laspeyres index P_{PL} and the plutocratic or disaggregated Paasche index P_{PP}. Hence if P_{PL} and P_{PP} are not too different, a good point approximation to the theoretical national plutocratic consumer price index will be the *plutocratic or disaggregated Fisher index* P_{PF} defined as:

$$P_{PF} \equiv \sqrt{P_{PL} P_{PP}} \tag{18.9}$$

The plutocratic Fisher price index P_{PF} is computed just like the usual Fisher price index, except that each commodity in each region (or for each household) is regarded as a separate commodity. Of course, this index will satisfy the time reversal test.

18.16 Since statistical agencies do not calculate Laspeyres, Paasche and Fisher price indices by taking inner products of price and quantity vectors, as was done in equation (18.9) and the previous definitions, it will be useful to obtain formulae for the Laspeyres and Paasche indices that depend only on price relatives and expenditure shares. In order to do this, it is necessary to introduce some notation. Define the expenditure share of household h on commodity i in period t as

$$S_{hi}^t \equiv \frac{p_{hi}^t q_{hi}^t}{\sum_{k=1}^n p_{hk}^t q_{hk}^t}; \quad t=0,1; \quad h=1,2,\ldots,H; \quad i=1,2,\ldots,n \tag{18.10}$$

Define the expenditure share of household h in total period t consumption as:

$$S_h^t \equiv \frac{\sum_{i=1}^n p_{hi}^t q_{hi}^t}{\sum_{k=1}^H \sum_{i=1}^n p_{ik}^t q_{ik}^t} = \frac{p_h^t q_h^t}{\sum_{k=1}^H p_k^t q_k^t} \quad t=0,1;\ h=1,2,\ldots,H \tag{18.11}$$

Finally, define the national expenditure share of commodity i in period t as:

$$\sigma_i^t \equiv \frac{\sum_{h=1}^H p_{hi}^t q_{hi}^t}{\sum_{k=1}^H p_k^t q_k^t} \quad t=0,1; \quad i=1,2,\ldots,n$$

$$= \sum_{h=1}^H \left(\frac{p_{hi}^t q_{hi}^t}{p_h^t q_h^t}\right)\left(\frac{p_h^t q_h^t}{\sum_{k=1}^H p_k^t q_k^t}\right)$$

$$= \frac{\sum_{h=1}^H s_{hi}^t p_h^t q_h^t}{\sum_{k=1}^H p_k^t q_k^t}$$

$$= \sum_{h=1}^H s_{hi}^t S_h^t \tag{18.12}$$

The Laspeyres price index for region h (or household h) is defined as:

$$P_{Lh} \equiv \frac{p_h^1 q_h^0}{p_h^0 q_h^0} \quad h=1,2,\ldots,H$$

$$= \frac{\sum_{i=1}^n \left(\frac{p_{hi}^1}{p_{hi}^0}\right) p_{hi}^0 q_{hi}^0}{p_h^0 q_h^0}$$

$$= \sum_{i=1}^n s_{hi}^0 \left(\frac{p_{hi}^1}{p_{hi}^0}\right) \tag{18.13}$$

18.17 Referring back to equation (18.3), the plutocratic national Laspeyres price index, P_{PL}, can be rewritten as follows:

$$P_{PL} \equiv \frac{\sum_{h=1}^H p_h^1 q_h^0}{\sum_{h=1}^H p_h^0 q_h^0} \tag{18.14}$$

$$= \sum_{h=1}^H \left(\frac{p_h^1 q_h^0}{p_h^0 q_h^0}\right)\left(\frac{p_h^0 q_h^0}{\sum_{h=1}^H p_h^0 q_h^0}\right) = \sum_{h=1}^H \left(\frac{p_h^1 q_h^0}{p_h^0 q_h^0}\right) S_h^0$$

$$= \sum_{h=1}^H S_h^0 P_{Lh} \tag{18.15}$$

$$= \sum_{h=1}^H S_h^0 \sum_{i=1}^n s_{hi}^0 \left(\frac{p_{hi}^1}{p_{hi}^0}\right)$$

$$= \sum_{h=1}^H \sum_{i=1}^n S_h^0 s_{hi}^0 \left(\frac{p_{hi}^1}{p_{hi}^0}\right) \tag{18.16}$$

Equation (18.15) shows that the plutocratic national Laspeyres price index is equal to a (period 0) regional expenditure share-weighted average of the regional Laspeyres price indices. Equation (18.16) shows that the national Laspeyres price index is equal to a (period 0) expenditure share-weighted average of the regional price relatives, (p_{hi}^1/p_{hi}^0), where the corresponding weight, $S_h^0 s_{hi}^0$, is the period 0 national expenditure share of commodity i in region h.

18.18 The Paasche price index for region h (or household h) is defined as:

$$P_{Ph} \equiv \frac{p_h^1 q_h^1}{p_h^0 q_h^1} \quad h=1,2,\ldots,H$$

$$= \frac{1}{\sum_{i=1}^n \left(\frac{p_{hi}^0}{p_{hi}^1}\right) p_{hi}^1 q_{hi}^1 / p_h^1 q_h^1}$$

$$= \frac{1}{\sum_{i=1}^{n} s_{hi}^1 \left(\frac{p_{hi}^1}{p_{hi}^0}\right)^{-1}}$$

$$= \left\{ \sum_{i=1}^{n} s_{hi}^1 \left(\frac{p_{hi}^1}{p_{hi}^0}\right)^{-1} \right\}^{-1} \quad (18.17)$$

18.19 Referring back to equation (18.7), the plutocratic national Paasche price index, P_{PP}, can be rewritten as follows:

$$P_{PP} \equiv \frac{\sum_{h=1}^{H} p_h^1 q_h^1}{\sum_{h=1}^{H} p_h^0 q_h^1} \quad (18.18)$$

$$= \frac{1}{\left\{ \sum_{h=1}^{H} \left(\frac{p_h^0 q_h^1}{p_h^1 q_h^1}\right) \left(\frac{p_h^1 q_h^1}{\sum_{h=1}^{H} p_h^1 q_h^1}\right) \right\}}$$

$$= \frac{1}{\sum_{h=1}^{H} \left(\frac{p_h^1 q_h^0}{p_h^0 q_h^0}\right)^{-1} S_h^1}$$

$$= \left(\sum_{h=1}^{H} S_h^1 P_{Ph}^{-1} \right)^{-1} \quad (18.19)$$

$$= \left\{ \sum_{h=1}^{H} S_h^1 \sum_{i=1}^{n} s_{hi}^1 \left(\frac{p_{hi}^1}{p_{hi}^0}\right)^{-1} \right\}^{-1}$$

$$= \left\{ \sum_{h=1}^{H} \sum_{i=1}^{n} S_h^1 s_{hi}^1 \left(\frac{p_{hi}^1}{p_{hi}^0}\right)^{-1} \right\}^{-1} \quad (18.20)$$

Equation (18.19) shows that the national plutocratic Paasche price index is equal to a (period 1) regional expenditure share-weighted harmonic mean of the regional Paasche price indices. Equation (18.20) shows that the national Paasche price index is equal to a (period 1) expenditure share-weighted harmonic average of the regional price relatives, (p_{hi}^1/p_{hi}^0), where the weight for this price relative, $S_h^1 s_{hi}^1$, is the period 1 national expenditure share of commodity i in region h.

18.20 Of course, the share formulae for the plutocratic Paasche and Laspeyres indices, P_{PP} and P_{PL}, given by equations (18.20) and (18.16), can now be used to calculate the plutocratic Fisher index, $P_{PF} \equiv [P_{pp} P_{PL}]^{1/2}$.

18.21 If prices are equal across regions, the formulae (18.16) and (18.20) simplify. The formula for the plutocratic Laspeyres index becomes:

$$P_{PL} = \sum_{h=1}^{H} \sum_{i=1}^{n} S_h^0 s_{hi}^0 \left(\frac{p_{hi}^1}{p_{hi}^0}\right)$$

$$= \sum_{h=1}^{H} \sum_{i=1}^{n} S_h^0 s_{hi}^0 \left(\frac{p_i^1}{p_i^0}\right) \quad \text{using assumptions (18.4)}$$

$$= \sum_{i=1}^{n} \sigma_i^0 \left(\frac{p_i^1}{p_i^0}\right) \quad \text{using equation (18.12) for } t=0$$

$$= P_L \quad (18.21)$$

where P_L is the usual aggregate Laspeyres price index based on the assumption that each household faces the same vector of commodity prices; see equation (18.5) for the definition of P_L. Under the equal prices across households assumption, the formula for the plutocratic Paasche index becomes:

$$P_{PP} = \left\{ \sum_{h=1}^{H} \sum_{i=1}^{n} S_h^1 s_{hi}^1 \left(\frac{p_{hi}^1}{p_{hi}^0}\right)^{-1} \right\}^{-1}$$

$$= \left\{ \sum_{h=1}^{H} \sum_{i=1}^{n} S_h^1 s_{hi}^1 \left(\frac{p_i^1}{p_i^0}\right)^{-1} \right\}^{-1} \quad \text{using assumptions (18.4)}$$

$$= \left\{ \sum_{i=1}^{n} \sigma_i^1 \left(\frac{p_i^1}{p_i^0}\right)^{-1} \right\}^{-1} \quad \text{using equation (18.12) for } t=1$$

$$= P_P \quad (18.22)$$

where P_P is the usual aggregate Paasche price index based on the assumption that each household faces the same vector of commodity prices; see equation (18.8) for the definition of P_P.

18.22 Thus with the assumption that commodity prices are the same across regions, in order to calculate national Laspeyres and Paasche indices, only "national" price relatives and national commodity expenditure shares are required for the two periods under consideration. If there is regional variation in prices, however, then the simplified formulae (18.21) and (18.22) are not valid and it is necessary to use the earlier formulae (18.16) and (18.20), which require the use of regional price relatives and regional expenditure shares.

Democratic versus plutocratic cost of living indices

18.23 The plutocratic indices considered above weight each household in the economy according to the size of its expenditures in the two periods under consideration. Instead of weighting in this way, it is possible to define theoretical indices (and "practical" approximations to them) that give each household or household group in the economy an equal weight. Following Prais (1959), such an index will be called a *democratic index*. In this section, the plutocratic index number theory developed in paragraphs 18.3 to 18.22 will be reworked into the democratic framework.

18.24 Making the same assumptions as in paragraph 18.4, define the class of *conditional democratic cost of living indices*, $P_D^*(p^0, p^1, u, e_1, e_2, \ldots, e_H)$, pertaining to periods 0 and 1 for the arbitrary utility vector of household utilities $u \equiv (u_1, u_2, \ldots, u_H)$ and for the arbitrary vectors of household environmental variables e_h for $h = 1, 2, \ldots, H$ as follows:

$$P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u, e_1, e_2, \ldots, e_H) \equiv \sum_{h=1}^{H} \left(\frac{1}{H}\right) \frac{C^h(u_h, e_h, p_h^1)}{C^h(u_h, e_h, p_h^0)} \quad (18.23)$$

Thus P_D^* is a simple unweighted arithmetic average of the individual household conditional cost of living indices, $C^h(u_h, e_h, p_h^1)/C^h(u_h, e_h, p_h^0)$. In the numerator and denominator of these conditional indices, only the price variables are different, which is precisely what is wanted in a theoretical definition of a consumer price index. If the vector of environmental variables, e_h, is not present in the cost function of household h, then the conditional index $C^h(u_h, e_h, p_h^1)/C^h(u_h, e_h, p_h^0)$ becomes an ordinary Konüs true cost of living index of the type defined earlier in Chapter 17.

18.25 Now specialize the general definition (18.23) by replacing the general utility vector u by either the period 0 vector of household utilities $u^0 \equiv (u_1^0, u_2^0, \ldots u_H^0)$ or the period 1 vector of household utilities $u^1 \equiv (u_1^1, u_2^1, \ldots u_H^1)$. Further specialize the general definition by replacing the general household environmental vectors $(e_1, e_2, \ldots e_H) \equiv e$ by either the period 0 vector of household environmental variables $e^0 \equiv (e_1^0, e_2^0, \ldots e_H^0)$ or the period 1 vector of household environmental variables $e^1 \equiv (e_1^1, e_2^1, \ldots, e_H^1)$. The choice of the base period vector of utility levels and base period environmental variables leads to the *Laspeyres conditional democratic cost of living index*, $P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^0, e^0)$, while the choice of the period 1 vector of utility levels and period 1 environmental variables leads to the *Paasche conditional democratic cost of living index*, $P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^1, e^1)$. It turns out that these two democratic indices satisfy some interesting inequalities, which are derived below.

18.26 Specializing definition (18.23), the Laspeyres conditional democratic cost of living index, $P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^0, e^0)$, may be written as follows:

$$P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^0, e_1^0, e_2^0, \ldots, e_H^0) \equiv \sum_{h=1}^{H} \left(\frac{1}{H}\right) \frac{C^h(u_h^0, e_h^0, p_h^1)}{C^h(u_h^0, e_h^0, p_h^0)} = \sum_{h=1}^{H} \left(\frac{1}{H}\right) \frac{C^h(u_h^0, e_h^0, p_h^1)}{p_h^0 q_h^0}$$

using equation (18.1) for $t=0$

$$\leq \sum_{h=1}^{H} \left(\frac{1}{H}\right) \frac{p_h^1 q_h^0}{p_h^0 q_h^0} \qquad (18.24)$$

since $C^h(u_h^0, e_h^0, p_h^1) \equiv \min_q \{p_h^1 q : f^h(q, e_h^0) \geq u_h^0\} \leq p^1 q_h^0$ and q_h^0 is feasible for the cost minimization problem for $h = 1, 2, \ldots, H$

$$\equiv P_{DL}$$

where P_{DL} is defined to be the observable (in principle) democratic Laspeyres price index, $\sum_{h=1}^{H} \left(\frac{1}{H}\right) p_h^1 q_h^0 / p_h^0 q_h^0$, which uses the individual vectors of household or regional quantities for period 0, $(q_1^0, \ldots, q_H^0)$, as quantity weights.

18.27 In a similar manner, specializing definition (18.23), the Paasche conditional democratic cost of living index, $P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^1, e^1)$, may be written as follows:

$$P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^1, e_1^1, e_2^1, \ldots, e_H^1) \equiv \sum_{h=1}^{H} \left(\frac{1}{H}\right) \frac{C^h(u_h^1, e_h^1, p_h^1)}{C^h(u_h^1, e_h^1, p_h^0)} = \sum_{h=1}^{H} \left(\frac{1}{H}\right) \frac{p_h^1 q_h^1}{C^h(u_h^1, e_h^1, p_h^0)}$$

using equation (18.1) for $t=1$

$$\geq \sum_{h=1}^{H} \left(\frac{1}{H}\right) \frac{p_h^1 q_h^1}{p_h^0 q_h^1} \text{ using a feasibility argument}$$

$$\equiv P_{DP} \qquad (18.25)$$

where P_{DP} is defined to be the *democratic Paasche price index*, $\sum_{h=1}^{H} \left(\frac{1}{H}\right) p_h^1 q_h^1 / p_h^0 q_h^1$, which uses the individual vector of household h quantities for period 1, q_h^1, as quantity weights for term h in the summation of individual household Paasche indices. Thus, it can be seen that the theoretical Paasche conditional democratic cost of living index, $P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^1, e^1)$, is bounded from below by the observable (in principle) democratic Paasche price index P_{DP}. Diewert (1983a, p. 191) first obtained the inequality (18.25) for the case where the environmental variables are absent from the household utility and cost functions, and prices are equal across households.

18.28 It is now shown how to obtain a theoretical democratic cost of living index that is bounded from above and below by observable indices. Using the inequalities (18.24) and (18.25) and the continuity properties of the conditional democratic cost of living $P^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u, e)$ defined by equation (18.23), it is possible to modify the method of proof used by Konüs (1924) and Diewert (1983a, p. 191) and establish the following result:

> There exists a reference utility vector $u^* \equiv (u_1^*, u_2^*, \ldots, u_H^*)$ such that the household h reference utility level u_h^* lies between the household h period 0 and 1 utility levels, u_h^0 and u_h^1 respectively for $h = 1, \ldots, H$. Also, there exist household environmental vectors $e_h^* \equiv (e_{h1}^*, e_{h2}^*, \ldots, e_{hM}^*)$ such that the household h reference mth environmental variable e_{hm}^* lies between the household h period 0 and 1 levels for the mth environmental variable, e_{hm}^0 and e_{hm}^1 respectively for $m = 1, 2, \ldots, M$ and $h = 1, \ldots, H$. The conditional democratic cost of living index $P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^*, e^*)$, evaluated at this intermediate reference utility vector u^* and the intermediate reference vector of household environmental variables $e^* \equiv (e_1^*, e_2^*, \ldots, e_H^*)$, lies between the observable (in principle) democratic Laspeyres and Paasche price indices, P_{DL} and P_{DP}, defined above by the last equalities in (18.24) and (18.25).

18.29 The above result says that the *theoretical national democratic conditional consumer price index* $P_D^*(p_1^0, \ldots, p_H^0, p_1^1, \ldots, p_H^1, u^*, e^*)$ lies between the democratic Laspeyres index P_{DL} and the democratic Paasche index P_{DP}. Hence if P_{DL} and P_{DP} are not too different, a good point approximation to the theoretical national democratic consumer price index will be the *democratic Fisher index* P_{DF}, defined as:

$$P_{DF} \equiv \sqrt{P_{DL} P_{DP}} \tag{18.26}$$

The democratic Fisher price index, P_{DF}, will satisfy the time reversal test.

18.30 Again, it will be useful to obtain formulae for the democratic Laspeyres and Paasche indices that depend only on price relatives and expenditure shares. Using definition (18.10) for the household h expenditure share on commodity i during period t, s_{hi}^t, the Laspeyres and Paasche price indices for household h can be written in share form as follows:

$$P_{Lh} \equiv \frac{p_h^1 q_h^0}{p_h^0 q_h^0} = \sum_{i=1}^{n} s_{hi}^0 \left(\frac{p_{hi}^1}{p_{hi}^0} \right); \quad h = 1, \ldots, H \tag{18.27}$$

$$P_{Ph} \equiv \frac{p_h^1 q_h^1}{p_h^0 q_h^1} = \left\{ \sum_{i=1}^{n} s_{hi}^1 \left(\frac{p_{hi}^1}{p_{hi}^0} \right)^{-1} \right\}^{-1}; \quad h = 1, \ldots, H. \tag{18.28}$$

Substituting equation (18.27) into the definition of the democratic Laspeyres index, P_{DL}, leads to the following share type formula:[10]

$$P_{DL} = \sum_{h=1}^{H} \left(\frac{1}{H} \right) \sum_{i=1}^{n} s_{hi}^0 \left(\frac{p_{hi}^1}{p_{hi}^0} \right) \tag{18.29}$$

Similarly, substituting equation (18.28) into the definition of the democratic Paasche index, P_{DP}, leads to the following share type formula:

$$P_{DL} = \sum_{h=1}^{H} \left(\frac{1}{H} \right) \left\{ \sum_{i=1}^{n} s_{hi}^1 \left(\frac{p_{hi}^1}{p_{hi}^0} \right)^{-1} \right\}^{-1} \tag{18.30}$$

18.31 The formula for the democratic Laspeyres index in the previous paragraph simplifies if it can be assumed that each household faces the same vector of prices in each of the two periods under consideration. Under this condition, equation (18.28) can be rewritten as follows:

$$P_{DL} = \sum_{i=1}^{n} s_{di}^0 \left(\frac{p_i^1}{p_i^0} \right) \tag{18.31}$$

where the period 0 democratic expenditure share for commodity i, s_{di}^0, is defined as follows:

$$s_{di}^0 \equiv \sum_{h=1}^{H} \left(\frac{1}{H} \right) s_{hi}^0; \quad i = 1, \ldots, n \tag{18.32}$$

Thus s_{di}^0 is simply the arithmetic average (over all households) of the individual household expenditure shares on commodity i during period 0. The formula for the democratic Paasche index does not simplify in the same way, under the assumption that households face the same prices in each period, because of the harmonic form of averaging in equation (18.30).

18.32 The conclusion at this point is that democratic and plutocratic Laspeyres, Paasche and Fisher indices can be constructed by a statistical agency provided that information on household-specific price relatives, p_{hi}^1/p_{hi}^0, and expenditures is available for both periods under consideration. If expenditure information is available only for the base period, then only the Laspeyres democratic and plutocratic indices can be constructed.

18.33 It is now necessary to discuss a practical problem that statistical agencies face: namely, that existing household consumer expenditure surveys, which are used in order to form estimates of household expenditure shares, are not very accurate. Thus the detailed commodity by region expenditure shares, $S_h^0 s_{hn}^0$ and $S_h^1 s_{hn}^1$, which appear in the formulae for the plutocratic Laspeyres and Paasche indices, are generally measured with very large errors. Similarly, the individual household expenditure shares for the two periods under consideration, s_{hn}^0 and s_{hn}^1, which are required in order to calculate the democratic Laspeyres and Paasche indices defined by equations (18.29) and (18.30) respectively, are also generally measured with substantial errors. Hence, it may lead to less overall error if the regional commodity expenditure shares s_{hn}^t are replaced by the national commodity expenditure shares σ_n^t defined by equation (18.12). Whether this approximation is justified would depend on a detailed analysis of the situation facing the statistical agency. In general, complete and accurate information on household expenditure shares will not be available to the statistical agency, and hence statistical estimation and smoothing techniques will have to be used in order to obtain expenditure weights that will be used to weight the price relatives collected by the agency.

18.34 It should be noted that the conditional index framework used above can be used to model situations where household preferences change (continuously) from the base period to the current period: simply choose the environmental variable to be time t. The theoretical results in paragraphs 18.14 and 18.28 imply the existence of cost of living indices that lie between observable Laspeyres and Paasche bounds, where the preference functions for the households are taken to be some preferences that are intermediate between the preferences pertaining to the two periods under consideration. As usual, if the observable bounds are not too far apart, taking the geometric average of the bounds leads to an adequate approximation to these theoretical cost of living indices.[11]

18.35 For criticisms and some limitations of the economic approach to index number theory, see Turvey (2000) and Diewert (2001).[12]

[10] Comparing the formula for the democratic Laspeyres index, P_{DL}, with the previous formula (18.16) for the plutocratic Laspeyres index, P_{PL}, it can be seen that the plutocratic weight for the ith price relative for household h is $S_h^0 s_{hi}^0$, whereas the corresponding democratic weight is $(1/H)s_{hi}^0$. Thus households that have larger base period expenditures and hence bigger expenditure shares S_h^0 get a larger weight in the plutocratic index as compared to the democratic index.

[11] For a more extensive treatment of cost of living theory in the context of taste change, see Balk (1989a).

[12] For a vigorous defence of the economic approach, see Triplett (2001).