What impact does the choice of formula have on international comparisons?

Keir G. Armstrong *Department of Economics, Carleton University*

Abstract. In this paper an empirical comparison of a number of alternative multilateral index-number formulae is undertaken. The magnitude of the effect of choosing one formula over another is ascertained using an appropriate cross-sectional data set constructed under the auspices of the Eurostat-OECD Purchasing Power Parity Programme. To this end, a new indicator is proposed that facilitates the measurement of the difference between two sets of bloc consumption shares, each computed using a different multilateral comparison method. JEL Classification: C31, C43, C81, E31, F31, O57

Quel impact est-ce que le choix des formules a sur les comparaisons internationales? Ce mémoire propose une comparaison empirique d'un certain nombre de formules de nombres-indices multilatéraux. On tente de jauger la magnitude de l'impact du choix d'une formule plutôt qu'une autre en utilisant une base de données transversales appropriée construite sous l'égide du Programme de parité du pouvoir d'achat Eurostat-OCDE. On propose un nouvel indicateur qui facilite la mesure de la différence entre deux ensembles de patterns de consommation, chacun calculé à l'aide d'une méthode de comparaison multilatérale différente.

1. Introduction

Many different methods for aggregating microeconomic price and quantity data into a multilateral index of real output (or a component thereof like consumption) have been proposed during the past ninety-odd years. All along, some of these methods have been put into practice by various statistical agencies and international organizations for the purposes of economic analysis and/or public policy. More recently, attempts have been made to find a theoretical justification for the choice of

The author wishes to thank Erwin Diewert, Ken White, Chuck Blackorby, John Helliwell, and three anonymous referees for helpful comments on earlier versions of this paper. Email: karmstro@ccs.carleton.ca

Canadian Journal of Economics / Revue canadienne d'Economique, Vol. 34, No. 3
August / août 2001. Printed in Canada / Imprimé au Canada

0008-4085 / 01 / 697–718 / © Canadian Economics Association

one method over all others. The most fruitful path taken thus far in attempting to provide such a justification is arguably the test (or axiomatic) approach. Using this approach, the investigator specifies a set of ‘reasonable’ tests and then uses it as the basis for assessing the relative merits of alternative independently motivated methods.[1] Diewert (1986) proposed what has come to be viewed as the benchmark set of multilateral tests and used it to assert the superiority of three methods[2] in relation to five others. Using a modified version of Diewert’s set, Balk (1996) asserted the superiority of two different methods[3] in relation to eight others (including two of the three methods Diewert deemed superior). Setting aside the debate over which of these methods are ‘best,’ there is a clear consensus of opinion among experts in the field of international comparisons that almost any method is better than simply converting economic aggregates into a numéraire currency by means of exchange rates. If almost any method is better than the exchange-rate approach, does the choice among these superior methods matter very much? In the present paper we aim to show that it does: that the choice of one method (or formula) over another can have a substantial impact on the resulting international comparisons.

The question of how to compare multilateral real-output or purchasing-power-parity formulae from an empirical standpoint has received scant attention in the literature. If the formulae under consideration satisfy a certain minimal set of requirements, then the application of any one of them to a bloc consisting of n countries yields a vector of $n - 1$ numbers that can serve as a basis for all possible binary comparisons within the bloc. The universal means by which two such vectors have been compared in the past has been an assessment of the component-wise percentage differences between them.[4] This approach is unsatisfactory for a couple of reasons. First, the percentage difference between two numbers is an asymmetric indicator of the relative difference between them because it depends on which number is used as the point of comparison. To paraphrase an example from Törnqvist, Vartia, and Vartia (1985, 43), 250 is 25 per cent more than 200, or 200 is 20 per cent less than 250. Second, component-wise comparisons between two vectors are unlikely to give rise to a very accurate assessment of the overall difference between them unless the components are few in number or the calculated differences exhibit little variation in size.

In section 2 we propose a new index of the difference between the results of two multilateral comparison methods applied to the same data set. Based on the normed, symmetric, and additive log(arithmic) difference indicator, this index overcomes the problems mentioned above to provide an appropriate summary measure of the

1 The ultimate objective of showing that a set of reasonable tests completely characterizes a particular multilateral comparison formula has never been realized.
2 Viz., the ‘star,’ EKS, and own-share methods.
3 Viz., the van Ijzeren balanced and GK methods.
4 See, for example, Kravis, Kenessey, Heston, and Summers (1975, chaps. 1 and 5) and Ruggles (1967). Note that ‘similarity indexes’ such as those calculated by Kravis, Heston, and Summers (1982, chap. 9) measure the similarity between two vectors of prices or quantities with reference to a *single* multilateral comparison formula.

differences between the purchasing power parities (PPPs) or output shares associated with the two methods. In section 4 we describe the data used in section 5 to undertake an empirical comparison of the twelve specific methods – two new ones and ten proposed elsewhere – described in section 3. An explanation of why different sources provide different values for the same PPPs is given in the conclusion, section 6.

2. A summary measure of the differences between alternative formulae

The maintained domain of comparison consists of a bloc of countries $\mathcal{N} := \{1, \ldots, n\}$ with $\mathbf{h} := (h_1, \ldots, h_n)^\top \in \mathbb{R}^n_{++}$ resident households, a set of consumer goods and services $\mathcal{M} := \{1, \ldots, m\}$ with country-specific national-currency–denominated prices

$$\mathbf{P} := (\mathbf{p}_1, \ldots, \mathbf{p}_n)^\top = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nm} \end{pmatrix} \in \mathbb{R}^{nm}_{++}$$

and a vector of per household consumption bundles

$$\mathbf{X} := (\mathbf{x}_1, \ldots, \mathbf{x}_n)^\top = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \in \mathbb{R}^{nm}_{+}.$$

Thus, each of the h_k households in country $k \in \mathcal{N}$ is considered to be the purchaser of $x_{kl} \geq 0$ units of commodity $l \in \mathcal{M}$ at a price of $p_{kl} > 0$ country-k currency units. Following the conventions of the test approach, the underlying preferences that generate $\mathbf{X}$ are ignored and the elements of $\mathbf{P}$, $\mathbf{X}$, and $\mathbf{h}$ are treated as independent variables.

An axiomatic PPP index for country i relative to country j is a function ρ^{ij}: $\mathbb{R}^{nm}_{++} \times \mathbb{R}^{n(m+1)}_{+} \to \mathbb{R}$ with image $\rho^{ij}(\mathbf{P}, \mathbf{X}, \mathbf{h})$. It is assumed that, at the very least, this index is positive and transitive with respect to i and j. The *positivity* requirement enables the usual interpretation of $\rho^{ij}(\mathbf{P}, \mathbf{X}, \mathbf{h})$ as the number of country-i currency units needed to buy a commodity bundle equivalent to one that can be bought with a single country-j currency unit.

P. *Positivity*: For all $i, j \in \mathcal{N}$, $\rho^{ij}(\mathbf{P}, \mathbf{X}, \mathbf{h}) > 0$.

The *transitivity* requirement guarantees that the results of applying ρ^{ij} to a bloc comprising three or more countries are self-consistent.

T. *Transitivity*: For all $i, k, j \in \mathcal{N}$, $\rho^{ik}(\mathbf{P}, \mathbf{X}, \mathbf{h})\rho^{kj}(\mathbf{P}, \mathbf{X}, \mathbf{h}) = \rho^{ij}(\mathbf{P}, \mathbf{X}, \mathbf{h})$.

In addition to satisfying T, a self-consistent set of PPPs has two further properties. The first, called *identity*, requires the value of ρ^{ij} to be unity when $i = j$.

I. *Identity*: For all $j \in \mathcal{N}$, $\rho^{jj}(\mathbf{P}, \mathbf{X}, \mathbf{h}) = 1$.

The second, *country reversal*, asserts that the value of ρ^{ji} is the reciprocal of the value of ρ^{ij}.

CR. *Country Reversal*: For all $i, j \in \mathcal{N}$, $\rho^{ji}(\mathbf{P}, \mathbf{X}, \mathbf{h}) = 1/\rho^{ij}(\mathbf{P}, \mathbf{X}, \mathbf{h})$.

THEOREM 1. *If ρ^{ij} satisfies P and T, then it also satisfies I and CR.*

The proof of this and all subsequent theorems can be found in the appendix.

Consider two sets of PPPs, A and B, each computed using a different multilateral index-number formula satisfying P and T. For ease of exposition, let the n-dimensional square matrices (ρ_A^{ij}) and (ρ_B^{ij}) represent the elements of A and B, respectively. Since P together with T implies I, so that $\rho_A^{jj} = \rho_B^{jj} = 1$ for all $j \in \mathcal{N}$, there are up to $n^2 - n$ possible differences between these matrices. One way to construct a summary measure of these differences is to calculate their mean. For such an index to be meaningful, however, the elementary difference indicator must be unit-independent; that is, it must measure the *relative* difference between ρ_A^{ij} and ρ_B^{ij}.[5]

Vartia (1974, 5) defined an indicator of relative difference as a function d: $\mathbb{R}^2_{++} \rightarrow \mathbb{R}$ that is continuous, increasing in ρ_A^{ij} when ρ_B^{ij} is fixed, homogeneous of degree zero, and has the additional property that

$$d(\rho_A^{ij}, \rho_B^{ij}) \geq 0 \text{ if and only if } \rho_A^{ij} \geq \rho_B^{ij}. \tag{1}$$

The homogeneity property implies that only the ratio ρ_A^{ij}/ρ_B^{ij} matters; that is, there exists a function $D: \mathbb{R}_{++} \rightarrow \mathbb{R}$ such that $D(\rho_A^{ij}/\rho_B^{ij}) \equiv d(\rho_A^{ij}/\rho_B^{ij}, 1)$. Obviously, D is continuous, increasing, and satisfies

$$D(\rho_A^{ij}/\rho_B^{ij}) \geq 0 \text{ if and only if } \rho_A^{ij}/\rho_B^{ij} \geq 1. \tag{2}$$

Examples of this function can be found in Törnqvist, Vartia, and Vartia (1985, 44).

As explained in section 1, the problem with using percentage difference as an indicator of relative difference is that it is asymmetric. Formally, the indicator D is said to be symmetric if

$$D(\rho_B^{ij}/\rho_A^{ij}) = -D(\rho_A^{ij}/\rho_B^{ij}). \tag{3}$$

It is easy to show that the percentage difference indicator $\rho_A^{ij}/\rho_B^{ij} - 1$ does not satisfy this requirement.

5 The *absolute* difference $\rho_A^{ij} - \rho_B^{ij}$ is of the dimensionality $i\$/j\$$ – the number of units of country i's currency per unit of that of country j.

Since the index-number formulae generating (ρ_A^{ij}) and (ρ_B^{ij}) are assumed to be transitive, it would be desirable if the relative difference with respect to any two countries were equal to the relative difference with respect to the first country and some third country plus the relative difference with respect to the same third country and the second country; that is, for any $k \in \mathcal{N}$,

$$D\left(\frac{\rho_A^{ij}}{\rho_B^{ij}}\right) = D\left(\frac{\rho_A^{ik}}{\rho_B^{ik}}\right) + D\left(\frac{\rho_A^{kj}}{\rho_B^{kj}}\right). \tag{4}$$

This property is equivalent to *D*'s being additive:

$$D\left(\frac{\rho_A^{ik}}{\rho_B^{ik}}\frac{\rho_A^{kj}}{\rho_B^{kj}}\right) = D\left(\frac{\rho_A^{ik}}{\rho_B^{ik}}\right) + D\left(\frac{\rho_A^{kj}}{\rho_B^{kj}}\right). \tag{5}$$

Using T and then CR, equation (4) can be rewritten as

$$D\left(\frac{\rho_A^{ij}}{\rho_B^{ij}}\right) = D\left(\frac{\rho_A^{ik}}{\rho_B^{ik}}\right) + D\left(\frac{\rho_B^{ik}}{\rho_A^{ik}}\frac{\rho_A^{ij}}{\rho_B^{ij}}\right). \tag{6}$$

Setting $j := i$ and then invoking I shows that an additive indicator of relative difference is symmetric.

Another desirable property for the indicator *D* would be that it behave approximately as the percentage difference indicator $D_1(\rho_A^{ij}/\rho_B^{ij}) := \rho_A^{ij}/\rho_B^{ij} - 1$ when ρ_A^{ij}/ρ_B^{ij} is close to one. Törnqvist, Vartia, and Vartia (1985, 45) formalized this property as

$$\lim_{u\to 1}\frac{D(u)}{D_1(u)} = 1. \tag{7}$$

THEOREM 2. *Let* $D: \mathbb{R}_{++} \to \mathbb{R}$ *be a continuous, additive function that satisfies (7), and let* $(\rho_A^{ij}, \rho_B^{ij}) \in \mathbb{R}_{++}^2$. *Then* $D(\rho_A^{ij}/\rho_B^{ij}) = \ln(\rho_A^{ij}/\rho_B^{ij})$, *an increasing function that satisfies (2).*

Thus, continuity, additivity, and approximation of D_1 in the neighbourhood of unity are sufficient conditions for the log difference indicator $\ln(\rho_A^{ij}/\rho_B^{ij})$.

The symmetry property of $\ln(\rho_A^{ij}/\rho_B^{ij})$ in conjunction with CR implies that the mean of the $n^2 - n$ (off-diagonal) log differences between (ρ_A^{ij}) and (ρ_B^{ij}) is zero. The most natural way to eliminate this offsetting of positive and negative log differences is to take the absolute value of each and then calculate the *mean absolute log difference* (MALD):

$$\Delta_{A,B} = \frac{\sum_i \sum_{j\neq i}\left|\ln\left(\frac{\rho_A^{ij}}{\rho_B^{ij}}\right)\right|}{n(n-1)}. \tag{8}$$

Another technique that accomplishes the same end is to square each log difference, add the results, divide by their number, and then extract the square root to get the root mean square log difference (RMSLD). Since this measure is more sensitive to outliers than the MALD, it is less representative of the 'typical' log difference between (ρ_A^{ij}) and (ρ_B^{ij}). Furthermore, the interpretation of the value of the RMSLD is much less straightforward than the interpretation of the MALD. For these reasons, the former is rejected in favour of the latter as the appropriate summary measure for undertaking the empirical comparisons below.

Using T, for any $k \in \mathcal{N}$, (8) can be rewritten as

$$\Delta_{A,B} = \frac{\sum_i \sum_{j \neq i} \left| \ln \left(\frac{\rho_A^{ik} \rho_A^{kj}}{\rho_B^{ik} \rho_B^{kj}} \right) \right|}{n(n-1)} \tag{9}$$

$$= \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left| \ln \left(\frac{\rho_A^{ik}}{\rho_B^{ik}} \right) - \ln \left(\frac{\rho_A^{jk}}{\rho_B^{jk}} \right) \right|, \text{ by CR.} \tag{10}$$

For a particular bloc of countries, let $\mathcal{P} := \{A, B, \ldots\}$ denote the set of all sets of PPPs that satisfy P and T.

THEOREM 3. *Δ defined by the right-hand side of (10) is a metric on $\mathcal{P}$; that is, $\Delta_{\cdot,\cdot}$ is a real-valued function on $\mathcal{P} \times \mathcal{P}$ that satisfies (i) $\Delta_{A,B} \geq 0$; (ii) $\Delta_{A,B} = 0$ if and only if $A = B$; (iii) $\Delta_{A,B} = \Delta_{B,A}$; and (iv) $\Delta_{A,B} \leq \Delta_{A,C} + \Delta_{C,B}$ (triangle inequality).*

Thus, Δ possesses the most important properties of ordinary distance, making it a reasonable and intuitive measure of the difference between alternative sets of PPPs.

Table 1 contains three sets of PPPs covering the same twenty-four countries: the first two were calculated by the Organization for Economic Cooperation and Development (OECD) using the Eltetö-Köves-Szulc (EKS) method and the Geary-Khamis (GK) method,[6] respectively; the third was calculated for the Penn World Table (PWT) using the GK method. For comparison, the corresponding exchange rates (ER) are also included in table 1. The differences among these four sets of numbers can be summarized by computing the associated Δ values using equation (10): $\Delta_{\text{EKS,GK}} = 0.04773$, $\Delta_{\text{EKS,PWT}} = 0.10425$, $\Delta_{\text{GK,PWT}} = 0.08354$, $\Delta_{\text{EKS,ER}} = 0.28832$, $\Delta_{\text{GK,ER}} = 0.30469$ and $\Delta_{\text{PWT,ER}} = 0.34659$. Thus the OECD PPPs differ from one another by about 4.8 per cent and from the exchange rates by roughly 30 per cent, and the GK PPPs differ from one another by about 8.4 per cent[7] and from the exchange rates by over 30 per cent. Note that the conventional component-wise approach to making any of these comparisons would involve the sequential evaluation of twenty-three percentage differences. The size variation among these differences is sufficient to make an accurate overall assessment of them (based on inspection

6 These methods are described in the next section.
7 The reason for this difference is provided at the end of section 5.

TABLE 1
OECD- and PWT-calculated PPPs for private final consumption expenditure in 1990, national currency per U.S. dollar

Country	OECD-EKS	OECD-GK	PWT-GK	Exchange rate
Belgium	40.4	39.1	39.	33.3
Denmark	9.92	9.08	9.06	6.17
France	6.69	6.48	6.92	5.43
Germany	2.06	2.	2.1	1.61
Greece	140.	132.	115.	158.
Ireland	0.688	0.679	0.649	0.603
Italy	1380.	1332.	1422.	1195.
Luxembourg	36.6	36.	37.4	33.3
Netherlands	2.15	2.02	2.06	1.82
Portugal	105.5	93.7	90.	142.2
Spain	113.1	109.5	109.3	101.6
United Kingdom	0.597	0.586	0.565	0.561
Austria	14.2	14.	13.9	11.3
Switzerland	2.23	2.2	2.28	1.38
Finland	6.87	6.66	5.89	3.83
Iceland	90.9	85.4	90.4	58.3
Norway	10.68	10.13	10.41	6.26
Sweden	9.5	9.02	8.52	5.92
Turkey	1597.	1232.	1076.	2613.
Australia	1.44	1.39	1.22	1.28
New Zealand	1.65	1.58	1.52	1.68
Japan	207.	186.	211.	145.
Canada	1.34	1.31	1.16	1.17
United States	1.	1.	1.	1.

SOURCES: OECD (1992, table 2.5; 1993b, table 2.8), Penn World Table (Mark 5.6a)

alone) highly unlikely. In the case of the EKS-GK comparison, for example, the relevant percentage differences range from 0.013 to 0.296 (or from −0.229 to −0.013) and have a median of 0.036 (or −0.035), a mean of 0.057 (or −0.051) and a standard deviation of 0.059 (or 0.046).

A system of bloc-specific (real) consumption indexes for countries $1,\ldots,n$ is a function $\boldsymbol{\sigma} : \mathbb{R}^{nm}_{++} \times \mathbb{R}^{n(m+1)}_{+} \rightarrow \mathbb{R}$ with image $\boldsymbol{\sigma}(\mathbf{P},\mathbf{X},\mathbf{h}) := [\sigma_1(\mathbf{P},\mathbf{X},\mathbf{h}),\ldots,\sigma_n(\mathbf{P},\mathbf{X},\mathbf{h})]^\top$. To enable the ith element ($i \in \mathcal{N}$) of this system to be interpreted as country i's share of total bloc consumption, $\boldsymbol{\sigma}$ is required to satisfy

S1. *Fundamental Share Test*: $\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h}) > 0$ for all $i \in \mathcal{N}$ and $\sum\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h}) = 1$.

THEOREM 4. *If $\boldsymbol{\sigma}$ satisfies S1, then ρ^{ij} defined implicitly by*

$$\rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h})\,\frac{\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h})}{\sigma_j(\mathbf{P},\mathbf{X},\mathbf{h})} = \frac{h_i\,\mathbf{p}_i^\top\,\mathbf{x}_i}{h_j\,\mathbf{p}_j^\top\,\mathbf{x}_j} \tag{11}$$

satisfies P and T, and

$$\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h}) = \left\{ \sum_j \frac{h_j \mathbf{p}_j^\top \mathbf{x}_j}{h_i \mathbf{p}_i^\top \mathbf{x}_i} \middle/ \rho^{ji}(\mathbf{P},\mathbf{X},\mathbf{h}) \right\}^{-1}. \tag{12}$$

Under the assumptions of this theorem, the number $\rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h})$ is the amount by which the total bloc expenditure of country-*i* households relative to those of country *j* must be deflated in order to make it equal to the corresponding total consumption ratio.

Substituting for ρ^{tk} in (10) using (11) yields an equivalent expression for the mean absolute log difference between multilateral comparison methods *A* and *B*:

$$\Delta_{A,B} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left| \ln\left(\frac{\sigma_{A,i}}{\sigma_{B,i}}\right) - \ln\left(\frac{\sigma_{A,j}}{\sigma_{B,j}}\right) \right|. \tag{13}$$

Thus $\Delta_{A,B}$ can be calculated from associated basis sets of PPPs using (10) or from the associated consumption-share systems using (13).

3. Some specific methods

As noted above, Diewert (1986) evaluated eight specific multilateral comparison formulae in the light of his test approach. In the empirical work of section 5, results from seven of these formulae, two variants of the eighth, and three others, two of which are new, are compared. A brief description of each is given in this section.

All multilateral comparison formulae proposed in the literature are averages of some sort made on the basis of prices $\mathbf{P}$ and *total* quantities $\hat{\mathbf{h}}\mathbf{X}$, where $\hat{\mathbf{h}}$ is the $n \times n$ diagonal matrix with $\hat{h}_{kk} = h_k$. In other words, they do not treat per-household quantities $\mathbf{X}$ and household numbers $\mathbf{h}$ as distinct variables. The two new formulae proposed here are motivated by the notion that each binary national-price–level comparison should be an average of the corresponding comparisons made from the perspectives of the *n* 'average households.' Given the available data $(\mathbf{P},\mathbf{X},\mathbf{h})$, household *k*'s best price-level comparison between country *i* and country *j* is obtained by taking the ratio of the results of pricing $\mathbf{x}_k$ at both $\mathbf{p}_i$ and $\mathbf{p}_j$. This index will be approximately exact if the household has preferences that admit very little substitution among the *m* commodities, or if the *n* price vectors are not very different from one another.

Let $\bar{h}_k := h_k / \mathbf{1}_n^\top \mathbf{h}$ denote the fraction of bloc households living in country *k*, $\mathbf{1}_n$ being the *n*-dimensional (column) vector of ones. The household-share–weighted geometric mean of the *n* average-household PPP indexes is called the *household democratic PPP index for country i relative to country j*:

$$\rho_{\mathrm{HD}}^{ij}(\mathbf{P},\mathbf{X},\mathbf{h}) := \prod_k \left[\frac{\mathbf{p}_i^\top \mathbf{x}_k}{\mathbf{p}_j^\top \mathbf{x}_k} \right]^{\bar{h}_k}. \tag{14}$$

By assigning each average-household index a weight that is proportional to the corresponding total number of households, ρ_{HD}^{ij} affords equal treatment to all households in the bloc.

A weaker democratic aggregation rule would treat countries rather than households as equals. Accordingly, the *country democratic PPP index for country i relative to country j* is defined as the unweighted geometric mean of the *n* average-household PPP indexes:

$$\rho_{\mathrm{CD}}^{ij}(\mathbf{P},\mathbf{X},\mathbf{h}) := \prod_k \left[\frac{\mathbf{p}_i^\top \mathbf{x}_k}{\mathbf{p}_j^\top \mathbf{x}_k} \right]^{1/n}. \tag{15}$$

A formal evaluation of the properties of ρ_{CD}^{ij} and ρ_{HD}^{ij} can be found in Armstrong (1999).

The proximate alternative to averaging over the relative costs of consumption bundles $\mathbf{X}$ using prices $\mathbf{p}_i$ and $\mathbf{p}_j$ is to compute the relative cost of the bloc average consumption bundle $\mathbf{X}^\top \bar{\mathbf{h}}$ using $\mathbf{p}_i$ and $\mathbf{p}_j$. This method, first used by the United Nations Economic Commission for Latin America (ECLA) in the 1960s, is called the *ECLA* or *average basket PPP index for country i relative to country j*:[8]

$$\rho_{\mathrm{AB}}^{ij}(\mathbf{P},\mathbf{X},\mathbf{h}) := \frac{\mathbf{p}_i^\top (\mathbf{X}^\top \bar{\mathbf{h}})}{\mathbf{p}_j^\top (\mathbf{X}^\top \bar{\mathbf{h}})}. \tag{16}$$

Early multilateral comparison methods were based on bilateral index-number formulae. The simplest and most popular of these methods involved the use of the Laspeyres formula in making binary quantity comparisons between a pre-selected base country and each of the other countries in the bloc. In general, such a ‘star system’ can be constructed using any index-number formula of the form $\phi(\mathbf{p}_i, \mathbf{p}_j, \mathbf{x}_i, \mathbf{x}_j)$. Accordingly, for a given base country $k \in \mathcal{N}$, the *country-k star system of consumption shares* is defined by

$$\sigma_{k*,i}(\mathbf{P},\mathbf{X},\mathbf{h}) := \frac{\phi(\mathbf{p}_i, \mathbf{p}_k, h_i \mathbf{x}_i, h_k \mathbf{x}_k)}{\sum_j \phi(\mathbf{p}_j, \mathbf{p}_k, h_j \mathbf{x}_j, h_k \mathbf{x}_k)}. \tag{17}$$

A second multilateral comparison method based on a bilateral formula is due to Gini (1931, 12). Known by the initials of its three independent re-discoverers, Eltetö and Köves (1964) and Szulc (1964), the (*generalized*) *EKS system of consumption shares* is defined by[9]

$$\sigma_{\mathrm{EKS},i}(\mathbf{P},\mathbf{X},\mathbf{h}) := \frac{\prod_k [\phi(\mathbf{p}_i, \mathbf{p}_k, h_i \mathbf{x}_i, h_k \mathbf{x}_k)]^{1/n}}{\sum_j \prod_t [\phi(\mathbf{p}_j, \mathbf{p}_t, h_j \mathbf{x}_j, h_t \mathbf{x}_t)]^{1/n}}. \tag{18}$$

8 This index can be expressed in terms of total quantities $\hat{\mathbf{h}}\mathbf{X}$ because $\bar{\mathbf{h}} = \hat{\mathbf{h}}\mathbf{1}_n / \mathbf{1}_n^\top \mathbf{h}$.

9 In the version of this index advanced by Eltetö and Köves (1964) and Szulc (1964), the Fisher formula was used in place of ϕ.

A third bilateral-formula–based multilateral comparison method is due to Diewert (1986, 25). His *own-share system of consumption indexes* is defined by

$$\sigma_{\mathrm{OS},i}(\mathbf{P},\mathbf{X},\mathbf{h}) := \frac{\left\{\sum_k [\phi(\mathbf{p}_i,\mathbf{p}_k,h_i\mathbf{x}_i,h_k\mathbf{x}_k)]^{-1}\right\}^{-1}}{\sum_j\left\{\sum_t [\phi(\mathbf{p}_j,\mathbf{p}_t,h_j\mathbf{x}_j,h_t\mathbf{x}_t)]^{-1}\right\}^{-1}}. \tag{19}$$

Note that in the own-share and EKS systems, respectively, a harmonic mean of bilateral comparisons and a geometric mean of bilateral comparisons substitute for the base-country bilateral comparison of the star system.

The next three multilateral methods are based on weighted averages of the country-k star systems. Respectively, the *democratic weights*, *plutocratic weights*, and *quantity weights consumption-share systems* are defined by

$$\sigma_{\mathrm{DW},i}(\mathbf{P},\mathbf{X},\mathbf{h}) := \sum_k \frac{1}{n}\sigma_{k*,i}(\mathbf{P},\mathbf{X},\mathbf{h}) \tag{20}$$

$$\sigma_{\mathrm{PW},i}(\hat{\boldsymbol{\gamma}}\mathbf{P},\mathbf{X},\mathbf{h}) := \sum_k s_k(\hat{\boldsymbol{\gamma}}\mathbf{P},\mathbf{X},\mathbf{h})\sigma_{k*,i}(\hat{\boldsymbol{\gamma}}\mathbf{P},\mathbf{X},\mathbf{h}) \tag{21}$$

and

$$\sigma_{\mathrm{QW},i}(\mathbf{P},\mathbf{X},\mathbf{h}) := \sum_k \sigma_{OS,k}(\mathbf{P},\mathbf{X},\mathbf{h})\sigma_{k*,i}(\mathbf{P},\mathbf{X},\mathbf{h}), \tag{22}$$

where

$$s_k(\hat{\boldsymbol{\gamma}}\mathbf{P},\mathbf{X},\mathbf{h}) := \frac{(\gamma_k\mathbf{p}_k)^\top(h_k\mathbf{x}_k)}{\sum_t(\gamma_t\mathbf{p}_t)^\top(h_t\mathbf{x}_t)} \tag{23}$$

is country k's share of (nominal) bloc expenditure, $\boldsymbol{\gamma} := (\gamma_1,\ldots,\gamma_n)^\top$ is a vector of exchange rates and $\hat{\boldsymbol{\gamma}}$ is the $n \times n$ diagonal matrix with $\hat{\gamma}_{kk} = \gamma_k$ for all $k \in \mathcal{N}$.[10]

None of the three remaining multilateral methods is based on a bilateral formula. The first is a proposal by Geary (1958) that was later amplified by Khamis (1970, 1972); the second and third are variants of van Ijzeren's (1956) weighted balanced method.

10 Since γ_k is the price of a unit of country k's currency in terms of some numéraire currency, $\hat{\boldsymbol{\gamma}}\mathbf{P}$ is the matrix of numéraire-denominated bloc commodity prices.

The Geary-Khamis or GK consumption shares are found by solving the following system of equations:

$$\sigma_i = \sum_l \pi_l [h_i x_{il}], \qquad i = 1, \ldots, n \tag{24a}$$

$$\pi_l = \frac{\sum_i \omega_{il} \sigma_i}{\sum_k h_k x_{kl}}, \qquad l = 1, \ldots, m, \tag{24b}$$

where $\omega_{il} := p_{il}[h_i x_{il}]/\mathbf{p}_i^\top (h_i \mathbf{x}_i)$ is the lth country-i expenditure share. Equations (24b) define the 'international price' of each commodity as the ratio of the expenditure-share–weighted sum of the n consumption shares to the total quantity consumed. Equations (24a) define the share of bloc consumption for each country as the cost of its national basket at international prices.

The $n + m$ equations (24a,b) are not independent, since each constituent set implies

$$\sum_l \pi_l \sum_i h_i x_{il} = \sum \sigma_i, \tag{25}$$

and, consequently, at least one non-trivial solution exists. Khamis (1970, sect. 3) showed that, subject to any normalization on the σ_is,[11] the system consisting of any $n + m - 1$ of the equations (24a,b) has a unique positive solution.

The consumption shares associated with van Ijzeren's weighted balanced method are found by solving the following system of equations:

$$\sum_{k \neq i} a_k \frac{\mathbf{p}_i^\top (h_k \mathbf{x}_k)}{\mathbf{p}_i^\top (h_i \mathbf{x}_i)} \frac{\sigma_i}{\sigma_k} = \sum_{k \neq i} a_k \frac{\mathbf{p}_k^\top (h_i \mathbf{x}_i)}{\mathbf{p}_k^\top (h_k \mathbf{x}_k)} \frac{\sigma_k}{\sigma_i}, \qquad i = 1, \ldots, n, \tag{26}$$

where a_k is the country-k 'weighting coefficient.' If $\xi_1 \equiv \mathbf{p}_1^\top (h_1 \mathbf{x}_1)/\sigma_1, \ldots, \xi_n \equiv \mathbf{p}_n^\top (h_n \mathbf{x}_n)/\sigma_n$ are called 'equivalents,' the left-hand side of (26) is the number of equivalents that would be required to buy, in country i, the quantities in the weighted national baskets that can be bought for one equivalent in countries $1, \ldots, i - 1, i + 1, \ldots, n$. The right-hand side is the number of equivalents that would be required to buy, in each of countries $1, \ldots, i - 1, i + 1, \ldots, n$, the weighted quantities purchased in country i for one equivalent. The balanced method asserts that, for $i = 1, \ldots, n$, these two quantities of money are equal.

Van Ijzeren (1956, 25–7) showed that, subject to any normalization on the σ_is, the system consisting of any $n - 1$ of equations (26) has a unique positive solution. Under the normalization $\sum \sigma_i = 1$, this system is referred to as the household-weighted balanced (VH) method if $a_k := h_k$ and the quantity-weighted balanced (VQ) method if $a_k := \sigma_k / h_k$. The former weighting scheme originates with van Ijzeren (1956, 3–5); the latter with van Ijzeren (1983, 45).

11 For example, $\sum \sigma_i = 1$.

4. The data

The raw price and expenditure data used in the empirical work of the next section are those of the Eurostat-OECD PPP Programme. These data cover the bloc comprising the twenty-four OECD countries of 1990 and the general commodity list made up of the 158 basic headings[12] of the major aggregate called 'Final Consumption of Resident Households.' Let $\mathbf{V} := (v_{kl})$ denote the (24×158) matrix of national expenditures (in national currency units) at the basic heading level, and let $\bar{\mathbf{P}} := (\bar{p}_{kl})$ denote the corresponding matrix of basic-heading PPPs in national currency units per U.S. dollar. Hence, for all $k \in \mathcal{N}$ and for all $l \in \mathcal{M}$,

$$v_{kl} \equiv h_k p_{kl} x_{kl} \tag{27}$$

and

$$\bar{p}_{kl} \equiv \frac{p_{kl}}{p_{\mathrm{US},l}}. \tag{28}$$

Several different sources were employed in the determination of the household numbers ($\mathbf{h}$) presented in table 2. For the United States and Japan, Turkey, each of the Nordic countries excluding Iceland,[13] and each of the European Union countries excluding Denmark and Germany, the corresponding datum was furnished by, respectively, the United Nations (1993, table 3), the State Institute of Statistics (1993, table 65), the Nordic Statistical Secretariat (1994, table 105), and Eurostat (1992, table 3.13). Estimates of h_k for Austria, Switzerland, Canada, Australia, and New Zealand were made by linear interpolation using, respectively, the corresponding 1983 and 1993 figures reported by Eurostat (1992, table 3.13; 1995, table 3.13), the 1980 figure reported by the United Nations (1993, table 3) and the 1993 figure reported by Eurostat (1995, table 3.13), the 1986 figure reported by the United Nations (1993, table 3) and the 1991 figure reported by Statistics Canada (1992, table 8), the 1986 and 1991 figures reported by the Australian Bureau of Statistics (1995, table 5.11), and the 1986 and 1991 figures reported by the United Nations (1993, table 3). Linear extrapolation was used to estimate h_k for Germany[14] based on the corresponding 1986 and 1989 figures reported by Eurostat (1989, table 3.13; 1992, table 3.13). For Iceland, which has not undertaken a national census in over thirty-five years, it was assumed that the population-household ratio in 1990 was the same in relation to the range of values exhibited by Norway, Sweden, and Fin-

12 In principle, a basic heading consists of a small group of similar well-defined goods or services. In practice, it is the lowest level of classification for which expenditures can be estimated. Consequently, an actual basic heading can cover a broader range of commodities than is theoretically desirable.

13 Viz., Denmark, Finland, Norway, and Sweden.

14 More precisely, the Federal Republic of Germany, including West Berlin, as constituted prior to 3 October 1990.

TABLE 2
Population and household data for the Member countries of the OECD in 1990

Country	Persons (1,000)	Households (1,000)	Ratio
Belgium	9,967	3,786	2.6
Denmark	5,140	2,229	2.3
France	56,735	21,644	2.6
Germany	63,253	27,497	2.3
Greece	10,161	3,449	2.9
Ireland	3,503	1,060	3.3
Italy	56,661	20,766	2.7
Luxembourg	382	143	2.7
Netherlands	14,952	6,011	2.5
Portugal	9,896	3,301	3.0
Spain	38,959	11,444	3.4
United Kingdom	57,561	22,902	2.5
Austria	7,718	2,958	2.6
Switzerland	6,712	2,782	2.4
Finland	4,986	2,037	2.4
Iceland	255	97	2.6
Norway	4,241	1,751	2.4
Sweden	8,559	3,830	2.2
Turkey	56,098	11,189	5.0
Australia	17,065	5,698	3.0
New Zealand	3,363	1,155	2.9
Japan	123,537	40,670	3.0
Canada	26,584	9,813	2.7
United States	249,911	93,347	2.7
OECD total	836,199	299,559	2.8

SOURCE: Population data: United Nations (1996, table 5)

land as it was in 1960.[15] The number of households was computed by using this estimated ratio to deflate the corresponding population value.

5. Empirical results

The matrix $\bar{\mathbf{X}} := (\bar{x}_{kl})$ of per-household quantities consistent with $(\mathbf{V}, \bar{\mathbf{P}}, \mathbf{h})$ is defined as

$$\bar{x}_{kl} := \frac{v_{kl}}{\bar{p}_{kl} h_k} \tag{29}$$

$$\equiv p_{\mathrm{US},l}\, x_{kl}, \text{ by (27) and (28).} \tag{30}$$

15 The 1960 population and household numbers for Iceland, Norway, Sweden, and Finland were furnished by the OECD (1993a, 156) and the United Nations (1993, table 3).

The calculation of PPPs based on the data set $(\bar{\mathbf{P}}, \bar{\mathbf{X}}, \mathbf{h}) = (\mathbf{P}(\hat{\mathbf{p}}_{\text{US}})^{-1}, \mathbf{X}\hat{\mathbf{p}}_{\text{US}}, \mathbf{h})$ can be accomplished only by means of formulae that, in addition to satisfying P and T, satisfy *commensurability*, the requirement that a change in the unit of measure of each commodity has no effect on the value of ρ^{ij}.

C. *Commensurability*: For all $i, j \in \mathcal{N}$ and for all $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_m)^\top \in \mathbb{R}^m_{++}$, $\rho^{ij}(\mathbf{P}\hat{\boldsymbol{\lambda}}^{-1}, \mathbf{X}\hat{\boldsymbol{\lambda}}, \mathbf{h}) = \rho^{ij}(\mathbf{P}, \mathbf{X}, \mathbf{h})$, where $\hat{\boldsymbol{\lambda}}$ is the $m \times m$ diagonal matrix with $\hat{\lambda}_{ll} = \lambda_l$ for all $l \in \mathcal{M}$.

Similarly, the calculation of consumption shares based on the data set $(\bar{\mathbf{P}}, \bar{\mathbf{X}}, \mathbf{h})$ can be facilitated only by formulae that satisfy S1 and *share commensurability*, the requirement that the consumption shares be invariant to changes in the units of measure of commodities.

S5. *Share Commensurability*: For all $i \in \mathcal{N}$ and for all $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_m)^\top \in \mathbb{R}^m_{++}$, $\sigma_i(\mathbf{P}\hat{\boldsymbol{\lambda}}^{-1}, \mathbf{X}\hat{\boldsymbol{\lambda}}, \mathbf{h}) = \sigma_i(\mathbf{P}, \mathbf{X}, \mathbf{h})$.

All twelve of the multilateral comparison methods presented in section 3 satisfy S1 and S5 or, equivalently, by theorem 4, P, T, and C. Consequently, the computation of consumption shares was a straightforward exercise involving simple substitutions into the defining formulae. Table 3 contains a selection of the results of this exercise. Included are the two new methods (HD and CD) three of the four methods not based on a bilateral formula (AB, GK and VH) and three of the six bilateral-formula–based methods (EKS, OS and the k-star with $k :=$ US (US*)). The GK and VH consumption-share systems were each calculated iteratively using the household-democratic consumption shares as initial values. The Fisher 'ideal' consumption index ϕ_F defined by

$$\phi_F(\mathbf{p}_i, \mathbf{p}_j, \mathbf{x}_i, \mathbf{x}_j) := \left[\frac{\mathbf{p}_j^\top \mathbf{x}_i}{\mathbf{p}_j^\top \mathbf{x}_j} \frac{\mathbf{p}_i^\top \mathbf{x}_i}{\mathbf{p}_i^\top \mathbf{x}_j} \right]^{1/2} \tag{31}$$

was used as the basis for each of the bilateral-formula–based methods.

The mean absolute log differences among the eight methods of table 3 and the exchange-rate approach[16] are expressed as percentages in table 4. If the cut-off between 'substantial' and 'insubstantial' is set at 2 per cent, this table partitions the considered methods into five groups based on whether or not they are substantially different from one another. HD, AB, and CD are grouped together, since all of the differences among them lie below the cut-off while all of the differences involving just one of them lie above. Similarly, VH, EKS, and OS[17] form a group[18] as do each

16 The exchange-rate–based consumption-share system was calculated by substituting the country-i exchange rate with respect to country j, as reported in OECD (1992, table 2.5), for $\rho^{ij}(\mathbf{P}, \mathbf{X}, \mathbf{h})$ in equation (12).

17 Since it is so close to 2, $100\Delta_{\text{US}^*,\text{OS}} = 1.992$ is treated as a substantial difference.

18 An extended version of table 4 would show that the quantity-weighted van Ijzeren, democratic weights, plutocratic weights, and quantity weights methods also belong to this group.

TABLE 3
Indexes of private final consumption in 1990 (OECD = 100)

Country	HD	AB	CD	GK	VH	EKS	OS	US*
Belgium	1.12	1.12	1.13	1.11	1.12	1.11	1.12	1.14
Denmark	0.482	0.478	0.492	0.505	0.488	0.484	0.488	0.483
France	6.66	6.61	6.74	6.58	6.61	6.58	6.61	6.63
Germany	8.02	8.03	8.04	7.73	7.82	7.75	7.85	8.04
Greece	0.575	0.568	0.59	0.615	0.592	0.596	0.59	0.579
Ireland	0.239	0.237	0.243	0.229	0.233	0.236	0.23	0.221
Italy	6.5	6.42	6.63	6.62	6.54	6.58	6.51	6.42
Luxembourg	0.05	0.05	0.051	0.051	0.051	0.051	0.051	0.051
Netherlands	1.6	1.6	1.6	1.61	1.59	1.58	1.6	1.62
Portugal	0.535	0.526	0.546	0.628	0.578	0.577	0.575	0.561
Spain	3.06	3.03	3.12	3.13	3.11	3.13	3.09	3.
U.K.	6.46	6.42	6.49	6.28	6.34	6.37	6.31	6.14
Austria	0.802	0.799	0.806	0.777	0.787	0.787	0.784	0.772
Switzerland	0.938	0.938	0.944	0.889	0.915	0.908	0.916	0.923
Finland	0.441	0.438	0.449	0.436	0.436	0.438	0.435	0.427
Iceland	0.026	0.026	0.027	0.027	0.027	0.027	0.027	0.026
Norway	0.351	0.35	0.353	0.36	0.353	0.354	0.351	0.344
Sweden	0.795	0.79	0.808	0.819	0.803	0.806	0.801	0.789
Turkey	1.46	1.43	1.5	2.14	1.74	1.71	1.74	1.73
Australia	1.82	1.81	1.83	1.8	1.79	1.79	1.8	1.8
N.Z.	0.315	0.315	0.318	0.314	0.312	0.311	0.313	0.315
Japan	13.1	13.1	13.	14.3	13.5	13.3	13.6	13.7
Canada	3.4	3.42	3.37	3.28	3.32	3.31	3.33	3.38
U.S.	41.2	41.5	40.9	39.8	41.	41.2	40.9	40.9

of GK, US*, and ER.[19] Thus the choice of one method over another can have a substantial impact on international comparisons of consumption.

To get a feel for what ‘substantial impact’ means at the margin in practical terms, consider a hypothetical international project to be financed by reference to 1990 OECD consumption shares. Using the own-share system instead of the US* system ($100\Delta_{US^*,OS} \approx 2$) would change the average national contribution by 1.58 per cent. For some countries, however, this switch in methods would change their contribution by as much as 3.99 per cent. In an era of government fiscal restraint, $4.00 per hundred can easily be viewed as a substantial difference, that is, one that is large enough to merit justification of the choice of methods.

Table 5 presents eight per household consumption indexes derived from the results in table 3 using the household numbers in table 2. Each of these indexes measures the consumption of the average household in each OECD country as a percentage of that in the United States. Figure 1 is a graphical representation of selected results

19 Preliminary support for the hypothesis that this partition is not merely an artefact of the (1990) data is given by the fact that the same partition obtained from the corresponding 1993 data.

TABLE 4
Mean absolute log-percentage differences

	HD	AB	CD	GK	VH	EKS	OS	US*
ER	26.7	26.5	26.8	30.6	28.7	28.7	28.6	28.6
US*	4.18	4.39	4.4	5.37	2.39	3.07	1.99	
OS	3.7	4.18	3.42	4.31	0.428	1.23		
EKS	3.71	4.26	3.29	4.61	0.85			
VH	3.74	4.26	3.39	4.27				
GK	7.89	8.38	7.37					
CD	1.14	1.83						
AB	0.754							

in table 5 along with those of the exchange-rate approach. Therein, the relevant countries[20] are arranged from left to right along the horizontal axis in order of decreasing per household consumption calculated via the household democratic method.

In table 6, each entry is a PPP associated with the corresponding consumption share in table 3 by means of equation (11). Comparison of tables 6 and 1 shows that there are relatively large differences between the PWT(-GK) PPPs and the GK PPPs calculated by both the author and the OECD. For instance, the PPP for Canada is 1.31 according to the former and 1.16 according to the latter – a difference of about 12 per cent. Given that both sets of numbers were calculated using the same formula, why do they differ so much? The answer to this question is provided by Hill (1982, 7): 'It is inherent in all multilateral aggregation methods . . . that the characteristics of the group of countries as a whole impose themselves on measurements made within the group. [The PPP] of France [relative to] Germany, for example, cannot be exactly the same within the context of the [European Union] as it appears within the context of the world economy as a whole, even if exactly the same basic[-heading] PPPs, [national] expenditures and aggregation method are used.' With respect to the GK PPPs of the present paper, then, the noted differences are due to the fact that one set was calculated in the context of the 152-country PWT, whereas the other two were calculated in the context of the 24-country OECD.[21] Consequently, it seems sensible to regard the OECD-GK PPPs and the corresponding PWT PPPs as distinct concepts rather than as different measures of the same concept.

20 The United States (US), Luxembourg (LU), Canada (CA), Switzerland (CH), Japan (JP), Australia (AU), Italy (IT), France (FR), Belgium (BE), Germany (DE), the United Kingdom (UK), New Zealand (NZ), Iceland (IS), Austria (AT), Spain (ES), the Netherlands (NL), Ireland (IE), Finland (FI), Denmark (DK), Sweden (SE), Norway (NO), Greece (GR), Portugal (PT), and Turkey (TR).

21 The 'consistentization' procedure used in the construction of the PWT (PWT appendix, Dec. 1994, n. /4/) had no effect on the component-level PPPs used in the present paper. Under this procedure, the 'adjustment factor' for each 1990 benchmark country was multiplied by the *value* of real consumption for that country – *not* the PPP.

TABLE 5
Indexes of per-household private final consumption in 1990

Country	HD	AB	CD	GK	VH	EKS	OS	US*
Belgium	67.2	66.7	67.9	69.0	67.1	66.4	67.3	68.5
Denmark	49.0	48.2	50.3	53.2	49.8	49.2	49.9	49.4
France	69.6	68.6	71.1	71.3	69.5	68.9	69.6	69.8
Germany	66.1	65.6	66.7	65.9	64.8	63.9	65.1	66.6
Greece	37.7	37.0	39.1	41.8	39.1	39.1	39.0	38.3
Ireland	51.1	50.3	52.2	50.7	50.0	50.4	49.6	47.5
Italy	70.9	69.5	72.9	74.8	71.7	71.8	71.5	70.4
Luxembourg	79.8	79.2	81.1	83.2	80.8	80.7	80.5	80.6
Netherlands	60.1	59.7	60.8	62.9	60.4	59.4	60.7	61.6
Portugal	36.7	35.8	37.8	44.6	39.9	39.6	39.7	38.8
Spain	60.6	59.6	62.1	64.2	61.9	62.0	61.6	59.7
U.K.	63.9	63.0	64.6	64.3	63.0	63.0	62.8	61.1
Austria	61.4	60.8	62.2	61.6	60.6	60.3	60.5	59.5
Switzerland	76.3	75.8	77.4	74.9	74.9	73.9	75.1	75.6
Finland	49.0	48.4	50.3	50.2	48.8	48.8	48.7	47.8
Iceland	61.7	61.2	62.6	66.0	62.7	62.3	62.6	61.8
Norway	45.3	44.9	46.0	48.2	45.9	45.8	45.7	44.8
Sweden	47.0	46.4	48.1	50.1	47.8	47.7	47.7	47.0
Turkey	29.6	28.8	30.5	44.8	35.5	34.6	35.5	35.3
Australia	72.1	71.5	73.4	74.1	71.7	71.1	71.8	72.1
N.Z.	61.8	61.3	62.8	63.8	61.6	60.9	61.8	62.2
Japan	72.9	72.3	72.9	82.2	75.5	74.3	76.0	76.8
Canada	78.4	78.4	78.4	78.4	77.1	76.3	77.3	78.6
U.S.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Comparison of tables 6 and 1 also reveals small differences between the (non-PWT-) GK and EKS PPPs.[22] For both methods, these differences have arisen because the definition of private final consumption expenditure employed herein excludes expenditures by private non-profit institutions serving households, whereas that of the OECD does not. For the EKS method, the differences are also due to the imposition of 'fixity' by the OECD. Under this requirement, the 'official' PPPs for the European Union (EU) must remain unchanged in any comparison involving a larger group of countries. The achievement of fixity is a two-step process.[23] First, each OECD-specific PPP comparing two EU countries is replaced with the corresponding EU-specific PPP. Second, each OECD-specific PPP comparing an EU country and a non-EU country is adjusted to restore transitivity. Thus, there are three distinct PPP concepts embedded within the OECD-EKS results.

22 The mean absolute log-percentage differences are 0.313 and 0.472, respectively.

23 Such a process is necessary, since ρ^{ij}_{EKS} is not invariant to even small changes in the size of the bloc (Diewert 1986, proposition 8).

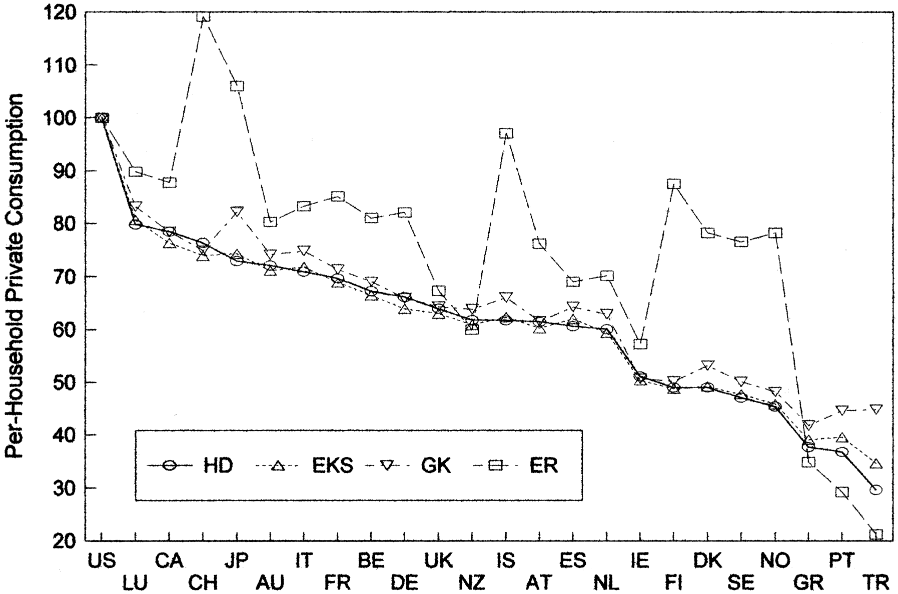

FIGURE 1 1990 Consumption indexes

6. Concluding remarks

Gordon (1996, 292) notes that the use of PPPs from alternative sources can lead to very different assessments of the relative standards of living of countries. In an expression of the widespread confusion that exists among users of PPP data about this seemingly 'fragile state of international . . . comparisons,' he goes on to ask the obvious question: '[W]hy [do] the sources differ so much?' There are three essential reasons.

First, different sources calculate the same PPPs in the context of different blocs of countries. An example of this was given above when the OECD-calculated EKS PPPs comparing two EU countries were contrasted with the corresponding EKS PPPs calculated by the author. The differences between them are due (in part) to the fact that the EKS index, like all other multilateral indexes, is bloc specific: The comparison of two EU countries in the context of the EU is conceptually different from the comparison of the same two countries in the broader context of the OECD. Similarly, the GK PPPs calculated by the author differ from the corresponding (GK) PPPs calculated for the PWT because the comparison of two OECD countries in the context of the OECD is conceptually different from the comparison of the same two countries in the context of the world as a whole.

Second, different sources build the same aggregates from different baskets of goods and services. For example, Final Consumption of Resident Households con-

TABLE 6
PPPs for private final consumption expenditure in 1990, national currency per U.S. dollar

Country	HD	AB	CD	GK	VH	EKS	OS	US*
Belgium	40.1	40.5	39.7	39.1	40.2	40.6	40.1	39.4
Denmark	9.86	10.01	9.59	9.08	9.68	9.8	9.67	9.77
France	6.64	6.73	6.5	6.48	6.64	6.71	6.63	6.62
Germany	2.	2.02	1.98	2.01	2.04	2.07	2.03	1.98
Greece	146.	148.	141.	131.	140.	140.	141.	144.
Ireland	0.674	0.685	0.66	0.681	0.69	0.684	0.696	0.725
Italy	1404.	1433.	1366.	1331.	1387.	1386.	1392.	1413.
Luxembourg	37.5	37.8	36.9	35.9	37.	37.1	37.1	37.1
Netherlands	2.12	2.14	2.1	2.03	2.12	2.15	2.1	2.07
Portugal	113.2	116.1	110.1	93.1	104.2	104.9	104.6	107.2
Spain	115.7	117.7	112.9	109.2	113.2	113.1	113.8	117.4
U.K.	0.591	0.599	0.584	0.587	0.599	0.599	0.601	0.617
Austria	14.	14.2	13.8	14.	14.2	14.3	14.2	14.5
Switzerland	2.16	2.17	2.12	2.19	2.2	2.22	2.19	2.17
Finland	6.84	6.92	6.66	6.67	6.86	6.87	6.88	7.01
Iceland	91.8	92.5	90.4	85.8	90.3	90.9	90.4	91.6
Norway	10.8	10.89	10.64	10.17	10.67	10.69	10.71	10.92
Sweden	9.64	9.76	9.4	9.03	9.48	9.5	9.5	9.64
Turkey	1869.	1920.	1811.	1235.	1559.	1598.	1559.	1565.
Australia	1.43	1.44	1.4	1.39	1.43	1.45	1.43	1.43
N.Z.	1.63	1.64	1.61	1.58	1.64	1.65	1.63	1.62
Japan	211.	213.	211.	187.	203.	207.	202.	200.
Canada	1.31	1.31	1.31	1.31	1.33	1.35	1.33	1.31
U.S.	1.	1.	1.	1.	1.	1.	1.	1.

sists of 159 basic headings under the OECD's classification and 215 under that of Eurostat. This fact points to an additional dimension of conceptual disparity among the OECD-EKS PPPs, since those that compare two EU countries were calculated on the basis of the latter classification, while all the others were calculated on the basis of the former.

Third, different sources calculate the same PPPs using different methods of aggregation. Using a new type of difference indicator, in the preceding section we showed that the choice of one method over another can have a substantial impact on the results obtained. Consequently, those who use PPP data for policy or research purposes should care about the justification for the way in which these data are calculated. Clearly, then, further research on the theoretical basis for international comparisons is needed in order to better inform users about which method is 'best' in some relevant sense.

Appendix

Proof of theorem 1. By T, for any $k \in \mathcal{N}$,

$$\rho^{jj}(\mathbf{P},\mathbf{X},\mathbf{h})\rho^{jk}(\mathbf{P},\mathbf{X},\mathbf{h}) = \rho^{jk}(\mathbf{P},\mathbf{X},\mathbf{h}) .$$

Thus, by P, $\rho^{jj}(\mathbf{P},\mathbf{X},\mathbf{h}) = 1$. Now,

$$\rho^{ji}(\mathbf{P},\mathbf{X},\mathbf{h}) = \frac{\rho^{ji}(\mathbf{P},\mathbf{X},\mathbf{h})\rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h})}{\rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h})}, \text{ by P}$$

$$= \frac{\rho^{jj}(\mathbf{P},\mathbf{X},\mathbf{h})}{\rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h})}, \text{ by T}$$

$$= \frac{1}{\rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h})}, \text{ by I.} \qquad \blacksquare$$

Proof of theorem 2. Since D is additive,

$$D(uv) = D(u) + D(v)$$

for any $(u,v) \in \mathbb{R}^2_{++}$. By Eichhorn (1978, 12–13), the only solution to this Cauchy-type equation is $D(u) = \alpha \ln u$, $\alpha \in \mathbb{R}$. By (7),

$$\lim_{u\to 1} \frac{\alpha \ln u}{u-1} = \lim_{u\to 1} \frac{\alpha \ln u - \ln 1}{u-1} =: \frac{d}{du}(\alpha \ln u)\bigg|_{u=1} = \alpha = 1.$$

Since $d \ln u/du > 0$ for all $u > 0$, D is increasing. Since $\ln u \geq 0$ if and only if $u \geq 1$, D satisfies (2). $\blacksquare$

Proof of theorem 3. First, by (10), there is a real number $\Delta_{A,B} \in \text{range}\,\Delta$ that is associated with any two elements A and B of $\mathcal{P}$. Clearly, $\Delta_{A,B} > 0$ if $B \neq A$. If $B = A$ then

$$\Delta_{A,A} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left| \ln\left(\frac{\rho_A^{ik}}{\rho_A^{ik}}\right) - \ln\left(\frac{\rho_A^{jk}}{\rho_A^{jk}}\right) \right|$$

$$= 0.$$

Next,

$$\Delta_{A,B} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left| \ln\left(\frac{\rho_A^{ik}}{\rho_B^{ik}}\right) - \ln\left(\frac{\rho_A^{jk}}{\rho_B^{jk}}\right) \right|$$

$$= \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left| \ln\left(\frac{\rho_B^{ik}}{\rho_A^{ik}}\right) - \ln\left(\frac{\rho_B^{jk}}{\rho_A^{jk}}\right) \right|$$

$$= \Delta_{B,A}.$$

Finally, for any $C \in \mathcal{P}$,

$$\begin{aligned}
\Delta_{A,B} &= \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left| \ln\left(\frac{\rho_A^{ik}}{\rho_B^{ik}}\right) - \ln\left(\frac{\rho_A^{jk}}{\rho_B^{jk}}\right) \right| \\
&= \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left| \ln\left(\frac{\rho_A^{ik}}{\rho_C^{ik}}\right) + \ln\left(\frac{\rho_C^{ik}}{\rho_B^{ik}}\right) - \ln\left(\frac{\rho_A^{jk}}{\rho_C^{jk}}\right) - \ln\left(\frac{\rho_C^{jk}}{\rho_B^{jk}}\right) \right| \\
&\leq \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left\{ \left| \ln\left(\frac{\rho_A^{ik}}{\rho_C^{ik}}\right) - \ln\left(\frac{\rho_A^{jk}}{\rho_C^{jk}}\right) \right| + \left| \ln\left(\frac{\rho_C^{ik}}{\rho_B^{ik}}\right) - \ln\left(\frac{\rho_C^{jk}}{\rho_B^{jk}}\right) \right| \right\} \\
&= \Delta_{A,C} + \Delta_{C,B}. \qquad \blacksquare
\end{aligned}$$

Proof of theorem 4. By S1,

$$\rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h}) := \frac{h_i \, \mathbf{p}_i^\top \mathbf{x}_i}{h_j \, \mathbf{p}_j^\top \mathbf{x}_j} \frac{\sigma_j(\mathbf{P},\mathbf{X},\mathbf{h})}{\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h})} > 0.$$

Next,

$$\begin{aligned}
\rho^{ik}(\mathbf{P},\mathbf{X},\mathbf{h})\rho^{kj}(\mathbf{P},\mathbf{X},\mathbf{h}) &:= \frac{h_i \, \mathbf{p}_i^\top \mathbf{x}_i}{h_k \, \mathbf{p}_k^\top \mathbf{x}_k} \frac{\sigma_k(\mathbf{P},\mathbf{X},\mathbf{h})}{\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h})} \frac{h_k \, \mathbf{p}_k^\top \mathbf{x}_k}{h_j \, \mathbf{p}_j^\top \mathbf{x}_j} \frac{\sigma_j(\mathbf{P},\mathbf{X},\mathbf{h})}{\sigma_k(\mathbf{P},\mathbf{X},\mathbf{h})} \\
&= \frac{h_i \, \mathbf{p}_i^\top \mathbf{x}_i}{h_j \, \mathbf{p}_j^\top \mathbf{x}_j} \frac{\sigma_j(\mathbf{P},\mathbf{X},\mathbf{h})}{\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h})} \\
&=: \rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h}).
\end{aligned}$$

Finally, from (11),

$$\begin{aligned}
&\frac{h_j \, \mathbf{p}_j^\top \mathbf{x}_j}{h_i \, \mathbf{p}_i^\top \mathbf{x}_i} \rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h})\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h}) = \sigma_j(\mathbf{P},\mathbf{X},\mathbf{h}) \\
&\quad \Rightarrow \sum_j \frac{h_j \, \mathbf{p}_j^\top \mathbf{x}_j}{h_i \, \mathbf{p}_i^\top \mathbf{x}_i} \rho^{ij}(\mathbf{P},\mathbf{X},\mathbf{h})\sigma_i(\mathbf{P},\mathbf{X},\mathbf{h}) = \sum_j \sigma_j(\mathbf{P},\mathbf{X},\mathbf{h}) \\
&\quad \Leftrightarrow \sigma_i(\mathbf{P},\mathbf{X},\mathbf{h}) = \left\{ \sum_j \frac{h_j \, \mathbf{p}_j^\top \mathbf{x}_j}{h_i \, \mathbf{p}_i^\top \mathbf{x}_i} \Big/ \rho^{ji}(\mathbf{P},\mathbf{X},\mathbf{h}) \right\}^{-1}, \text{ by S1 and CR.} \qquad \blacksquare
\end{aligned}$$

References

Armstrong, Keir G. (1999) 'A restricted-domain multilateral test approach to the theory of international comparisons,' Carleton Economics Papers 99-16, Carleton University

Australian Bureau of Statistics (1995) *Year Book*: *Australia*, Number 77, Cat. No. 1301.0 (Canberra)

Balk, Bert M. (1996) 'A comparison of ten methods for multilateral international price and volume comparison,' *Journal of Official Statistics* 12, 199–222

Diewert, W.E. (1986) 'Microeconomic approaches to the theory of international comparisons,' Discussion Paper No. 86-31, Department of Economics, University of British Columbia
Eichhorn, Wolfgang (1978) *Functional Equations in Economics* (London: Addison-Wesley)
Eltetö, O., and P. Köves (1964) 'On a problem of index number computation relating to international comparisons,' *Statisztikai Szemle* 42, 507–18
Eurostat (1989, 1992, 1993) *Basic Statistics of the Community*, 26th, 29th, and 30th eds (Luxembourg)
— (1995) *Basic Statistics of the European Union*, 32nd ed. (Luxembourg)
Geary, Robert C. (1958) 'A note on the comparison of exchange rates and purchasing power between countries,' *Journal of the Royal Statistical Society A* 121, 97–9
Gini, Corrado (1931) 'On the circular test of index numbers,' *Metron* 9, 3–24
Gordon, Robert J. (1996) 'Discussion of chapter 8,' in *The New System of National Economic Accounts*, ed. John W. Kendrick (Norwell, MA: Kluwer Academic)
Hill, Peter (1982), *Multilateral Measurements of Purchasing Power and Real GDP* (Luxembourg: Eurostat)
Ijzeren (Yzeren), J. van (1956) 'Three methods of comparing the purchasing power of currencies,' *Statistical Studies* 7 (The Hague: Centraal Bureau voor de Statiskiek)
— (1983), 'Index numbers for binary and multilateral comparison,' *Statistical Studies* 34 (The Hague: Centraal Bureau voor de Statiskiek)
Khamis, Salem H. (1970) 'Properties and conditions for the existence of a new type of index number,' *Sankhya B* 32, 81–98
— (1972) 'A new system of index numbers for national and international purposes,' *Journal of the Royal Statistical Society A* 135, 96–121
Kravis, I.B., A. Heston, and R. Summers (1982) *World Product and Income: International Comparisons of Real Gross Product* (Baltimore: Johns Hopkins University Press)
Kravis, I.B., Z. Kenessey, A. Heston, and R. Summers (1975) *A System of International Comparisons of Gross Product and Purchasing Power* (Baltimore: Johns Hopkins University Press)
Nordic Statistical Secretariat (1994) *Yearbook of Nordic Statistics*, vol. 32 (Copenhagen: Nordic Council of Ministers)
OECD (1992) *Purchasing Power Parities and Real Expenditures 1990*: *EKS Results* (Paris)
— (1993a) *National Accounts: Main Aggregates 1960–1991* (Paris)
— (1993b) *Purchasing Power Parities and Real Expenditures 1990: GK Results* (Paris)
Ruggles, Richard (1967) 'Price indexes and international comparisons,' in *Ten Economic Studies in the Tradition of Irving Fisher*, ed. W. Fellner et al. (New York: J. Wiley)
State Institute of Statistics (1993) *Statistical Yearbook of Turkey* (Ankara)
Statistics Canada (1992) *Dwellings and Households*: *The Nation*, Cat. No. 93-311 (Ottawa: Minister of Industry, Science and Technology)
Szulc, B. (1964) 'Indices for multiregional comparisons,' *Przeqlad Statystyczny* 3, 239–54
Törnqvist, Leo, Pentti Vartia, and Yrjö Vartia (1985) 'How should relative changes be measured?' *American Statistician* 39, 43–6
United Nations (1993) *Demographic Yearbook Special Issue: Population Ageing and the Situation of Elderly Persons*, ST/ESA/STAT/SER.R/22 (New York)
— (1996) *Demographic Yearbook 1994*, ST/ESA/STAT/SER.R/25 (New York)
Vartia, Yrjö O. (1974) 'Relative changes and economic indices,' unpublished licentiate thesis, Department of Statistics, University of Helsinki