CENTRE
for
ECONOMIC PERFORMANCE

DISCUSSION PAPER NO.  329

February 1997

# TRUE MULTILATERAL INDEXES FOR INTERNATIONAL COMPARISONS OF PURCHASING POWER AND REAL INCOME

J. P. NEARY

## ABSTRACT

I consider the problem of choosing index numbers of purchasing power and real income for international comparisons. I show that the desirable properties of methods based on the Fisher "Ideal" index do not extend to *multilateral* comparisons, except when tastes are homothetic. By contrast, the Geary method, which underlies the Penn World Tables, provides an approximation to a set of "true" exchange rate indexes which have many desirable properties. In particular, if demands exhibit generalized linearity, the true indexes measure real incomes relative to a hypothetical country whose income is an appropriate average of individual countries' incomes.



This paper was produced as part of the Centre's
Programme on International Economic Performance

# TRUE MULTILATERAL INDEXES FOR INTERNATIONAL COMPARISONS OF PURCHASING POWER AND REAL INCOME

J. Peter Neary

FEBRUARY 1997

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE



ISBN 0 7530 0075 X

# TRUE MULTILATERAL INDEXES FOR INTERNATIONAL COMPARISONS OF PURCHASING POWER AND REAL INCOME

J. Peter Neary



The Centre for Economic Performance is financed by the Economic and Social Research Council. The Centre also gratefully acknowledges financial support from the Senior Business Forum.

**ACKNOWLEDGEMENTS**

I am very grateful to Jim Anderson, Eric Fisher, Brid Gleeson, Patrick Honohan, Morgan Kelly, Dave Madden and Kevin O'Rourke for helpful discussions and to participants in seminars at ESEM 96 in Istanbul and in Florence (EUI), Hong Kong (CUHK), Keele and LSE for comments. My work on this paper has been supported by the Social Science Research Council of the Royal Irish Academy and forms part of the International Economic Performance Programme of the Centre for Economic Performance at the London School of Economics.

# TRUE MULTILATERAL INDEXES FOR INTERNATIONAL COMPARISONS OF PURCHASING POWER AND REAL INCOME

J. Peter Neary

How should we compare price levels and real incomes between countries? The question is of great importance, since the demand for such comparisons is enormous. Apart from their intrinsic interest, they are essential for testing hypotheses about comparative growth performance. Indeed, such tests have themselves become a major growth industry in recent years. This reflects in part the revival of interest in growth theory and the development of models of endogenous growth. It also reflects the relatively recent availability of comparative data on real incomes for a wide range of countries and years, of which the major source is the Penn World Tables, an enormous data set which originates from the United Nations International Comparison Project (ICP). (See Kravis (1984) and Summers and Heston (1991).)

However, a paradox lies at the heart of the Penn Tables. The basic method they use to construct internationally comparable data on real incomes relies on a method for computing "true" or "purchasing-power-corrected" exchange rates, devised by the Irish statistician Roy Geary (1958). This method has many practical advantages, most notably that it leads to a consistent set of world accounts which can be disaggregated by country and commodity. However, the method lacks a secure theoretical foundation and has been heavily criticised by theorists, most notably by Erwin Diewert, who argues for alternative approaches in his authoritative survey (1987, pp. 776-8).[1] The best-known of these alternative methods is the "EKS" index, named after its originators Eltetö and Köves (1964) and Szulc (1964), which has been used by the OECD and by Eurostat (the Statistical Office of the European Union) to produce purchasing-power-corrected real income data for their member countries.[2]

This paper re-examines the theoretical foundations for international comparisons of purchasing power and real incomes. I suggest that the claims of Diewert and others for methods based on the Fisher Ideal index do not hold up. Essentially, the Fisher-type indexes have desirable properties for *bilateral* comparisons which do not extend to the *multilateral* case. More positively, I propose a new set of "ideal" indexes for international comparisons which combine the desirable aggregation property of the Geary method with a firm foundation in economic theory. I also show that, under a wide class of assumptions about demand behaviour, my "ideal" indexes yield international comparisons of real income relative to a hypothetical country whose income is an average of world incomes in an appropriate sense. Finally, I argue that, for practical purposes, the Geary method is to be preferred to the EKS method and its variants, since it gives an approximation, though not necessarily a very good one, to an appropriate ideal procedure, whereas the EKS method yields a set of inconsistent multilateral comparisons.

Section 1 sets up the problem and introduces the three multilateral indexes which will be compared in the paper, the EKS index, the CCD index of Caves, Christensen and Diewert (1982) and the Geary index. Section 2 reviews some relevant results from the theory of index numbers, paying particular attention to the specific issues which arise in multilateral cross-section comparisons. Section 3 considers the results of Konüs and Byushgens (1926) and Diewert (1976) which provide a theoretical justification for Fisher-type indexes and shows that they do not extend to multilateral comparisons. Section 4 introduces my own proposed indexes, which I call the "Geary-Konüs" method. I note its theoretical properties, show how it relates to the Geary method and, in Section 5, draw on the theory of linear aggregation to explain the world prices which underlie the method.

## 1. PRELIMINARIES

### 1.1 The Problem

Suppose that, for each of $m$ countries, labelled $j = 1, \ldots m$, we have observations on the prices (expressed in national currencies) and the quantities (expressed in international units) of $n$ commodities, labelled $i = 1, \ldots n$. Price and quantity vectors in country $j$ are denoted $p^j$ and $q^j$, with typical elements $p_{ij}$ and $q_{ij}$, respectively. Each commodity is assumed to be identical in quality worldwide but because of transport costs, imperfect competition or other barriers to arbitrage there is no tendency for prices to be equalised internationally. Hence, official exchange rates are not appropriate for comparing price levels or real incomes between countries. What we seek is a set of index numbers which express the real income of each country $j$ relative to every other country $k$: $\{Q_{jk}, \forall j,k\}$.

## 1.2 The EKS Index

The simplest way of making multilateral comparisons of real incomes is the so-called "star" method, which revalues each country's consumption vector in terms of the prices of a single reference country. This amounts to constructing a set of Laspeyres quantity indexes, with the country at the centre of the star as base. However, there seems little justification for privileging one country in this way. Even in bilateral comparisons some compromise between the base-weighted Laspeyres index and the current-weighted Paasche index is normally preferred, of which the most widely-used is their geometric mean, the Fisher Ideal index:

$$\ln Q_{jk}^F = \tfrac{1}{2}\left\{\ln\frac{p^k.q^j}{p^k.q^k} + \ln\frac{p^j.q^j}{p^j.q^k}\right\}. \qquad (1)$$

The Fisher Ideal index has many desirable properties but it is not suited to multilateral comparisons. The EKS index extends it to the multilateral context since it equals the geometric mean of the ratios of all $m$ bilateral Fisher Ideal indexes, taking each of the $m$ countries in turn as base:

$$\ln Q_{jk}^{EKS} = \frac{1}{m}\sum_{l=1}^{m}\left\{\ln Q_{jl}^{F} - \ln Q_{kl}^{F}\right\}. \qquad (2)$$

Since the Fisher index is reflexive ($Q^F_{jj}=1$) and symmetric ($Q^F_{jk}.Q^F_{kj}=1$), this may be rewritten as:

$$\ln Q_{jk}^{EKS} = \frac{1}{m}\left[2\ln Q_{jk}^{F} + \sum_{l=1,l\neq j,k}^{m}\left\{\ln Q_{jl}^{F} - \ln Q_{kl}^{F}\right\}\right]. \qquad (3)$$

which reduces to the Fisher index when $m=2$. Thus the EKS index is indeed an appropriate multilateral generalisation of the Fisher Ideal.

## 1.3 The CCD Index

Caves, Christensen and Diewert (1982) have proposed an alternative to the EKS index which resembles it in many respects but has superior theoretical properties. Its starting point is the bilateral Törnqvist index, defined as:

$$\ln Q_{jk}^{T} = \frac{1}{2}\sum_{i}(\omega_{ij}+\omega_{ik})\ln(q_{ij}/q_{ik}), \qquad (4)$$

where $\omega_{ij}$ is the budget share of good $i$ in country $j$.[3] The CCD index extends the Törnqvist index to multilateral comparisons in the same way
as the EKS index extends the Fisher Ideal index:

$$\ln Q_{jk}^{CCD} = \frac{1}{m}\sum_{l=1}^{m}\left\{\ln Q_{jl}^{T} - \ln Q_{kl}^{T}\right\}. \qquad (5)$$

Caves, Christensen and Diewert have applied this index to international comparisons of output and productivity and Prasada Rao, Selvanathan and Pilat (1995) have applied the corresponding price index to international comparisons of consumer prices.

## 1.4 The Geary Method

The Geary method proceeds in a very different way to the other two indexes. It first postulates the existence of "world" prices p and "true" exchange rates e. The true exchange rates are Laspeyres price indexes, which compare the world prices with the prices of each country in turn:[4]

$$e_j = \frac{\sum_i p_i q_{ij}}{\sum_i p_{ij} q_{ij}}, \qquad j = 1, \ldots m. \tag{6}$$

As for the world prices themselves, they satisfy the property that total world spending on commodity *i* is the same whether valued at its world price or at domestic prices converted at the true exchange rates:

$$p_i = \frac{\sum_j e_j p_{ij} q_{ij}}{\sum_j q_{ij}}, \qquad i = 1, \ldots n. \tag{7}$$

Solving for e and p, it is then straightforward to calculate the income of each country at world prices:

$$z_j^G = e_j z_j = \sum_i p_i q_{ij}, \qquad i = 1, \ldots m. \tag{8}$$

These real income measures in turn imply a set of indexes, $Q^G_{jk} = z^G_j / z^G_k$, $\forall\, j,k$. Thus the Geary method is a star system with the hypothetical country ("the world") whose prices are p as centre.

## 2. CRITERIA FOR CHOOSING BETWEEN INDEX NUMBERS[5]

How can we choose between the different real-income indexes which have been introduced in the last section? There are two distinct approaches which can be taken to this problem. The “test” approach, following Fisher (1922), treats prices and quantities as independent variables and assesses the extent to which different indexes satisfy certain desirable, though not necessarily mutually consistent, properties. By contrast, the “economic” approach assumes that prices and quantities arise from optimising behaviour in each country and explores how closely empirical indexes approximate to the “true” indexes based on economic theory.

Consider first the test approach. While there are a great many tests which a satisfactory index number formula might be expected to satisfy, four in particular are especially relevant to multilateral comparisons:[6]

***i. Base-Country Invariance*:** It is intuitively desirable that indexes of real income should not be sensitive to the choice of base or reference country.

***ii. Transitivity* or *Circularity*:** A satisfactory index number formula should provide a unique cardinal ranking of the real incomes of the countries considered. Thus, the real income of country *j* relative to country *k* should be the same whether the two are compared directly or via an arbitrary intermediate country *l*: $Q_{jk} = Q_{jl}.Q_{lk}$.

***iii. Characteristicity* or *Independence of Irrelevant Countries*:** The comparison between two countries should as far as possible depend only on variables which characterise them and not on variables characteristic of other countries. Thus, country *j*’s real income relative to country *k*’s should ideally be unaffected by changes in third countries.

***iv. Matrix Consistency*:** Finally, the usefulness of a set of real income indexes is much enhanced if they can be consistently disaggregated by commodity as well as by country.

How well do the index numbers considered in the last section meet these criteria? It is clear that the "star" and bilateral Fisher Ideal index numbers do not satisfy either base-country invariance or transitivity, whereas all three multilateral systems do. As for characteristicity, the EKS index exhibits this to a high degree by construction, since it is the solution to the problem of finding a transitive index which minimises the sum of squared deviations from the bilateral (and non-transitive) Fisher Ideal indexes. (See Dreschler, 1973, p. 28.) However, both the EKS and CCD indexes fail to satisfy matrix consistency, whereas, because of its linear structure, the Geary system does satisfy this test. It was primarily for this reason that the Geary system was used in the ICP and subsequently as the foundation for the Penn World Tables.

The test approach is a useful starting point in choosing between competing index numbers. However, ever since Frisch (1936), it has been criticised by economists on a number of grounds. At a practical level, different tests often turn out to be mutually inconsistent. For example, we have seen above that some trade-off is necessary in practice between the criteria of transitivity and characteristicity. At a theoretical level, the test approach does not require that the indexes have any basis in economic theory; in particular, little or no intercommodity substitution may be allowed. Finally, at a conceptual level, all empirical index number formulae are open to the devastating criticism of Afriat (1977) that they provide no more than "answers without questions": without a clear conceptual framework no meaning can be attached to the concept of "real income" which empirical indexes purport to measure.

The "economic" approach to index numbers avoids these difficulties by explicitly starting from maximising behaviour. In the context of international comparisons, this implies that the data are assumed to be generated by the utility maximising behaviour of a representative consumer in each country, with identical tastes worldwide. Even with these assumptions, the economic approach applies strictly only to comparisons of expenditure levels (in the absence of a theoretical model of investment and government behaviour) and only if the problem of intra-national aggregation over households can be resolved (or safely ignored). Nevertheless, its superior theoretical foundation makes it a natural starting point in evaluating real income indexes.

One immediate difficulty with the economic approach is that it does not imply a unique ideal real income index. Even confining attention for the present to bilateral comparisons, at least three distinct measures of real income have been proposed:

***i. The Allen Quantity Index,*** $Q^A_{jk}$: This equals the ratio of the expenditure functions of the two countries evaluated at a common reference price vector $p^r$:

$$\ln Q^A_{jk} = \ln e(p^r,u_j) - \ln e(p^r,u_k). \tag{9}$$

Since the expenditure function gives the minimum cost of attaining a given utility level facing given prices, this index allows for intercommodity substitution and so avoids the biases of fixed-weight indexes.

***ii. The Konüs Quantity Index,*** $Q^K_{jk}$: A problem with the Allen index is that it is not in general consistent with the true price or cost-of-living index due to Konüs. An alternative index which meets this criterion by construction is the Konüs quantity index:

$$\ln Q^K_{jk} = [\ln(p^j.q^j) - \ln(p^k.q^k)] - [\ln e(p^j,u_r) - \ln e(p^k,u_r)]. \tag{10}$$

This index equals the ratio of actual expenditures in the two countries divided by the Konüs price index, evaluated at a reference utility level $u_r$.

***iii. The Malmquist Quantity Index,*** $Q^M_{jk}$: Finally, a difficulty with both the Allen and Konüs indexes is that they are not homogeneous of degree one in quantities.  An index which meets this desirable criterion is that of Malmquist:

$$\ln Q^{M}_{jk} = \ln d(q^{j},u_r) - \ln d(q^{k},u_r) \; . \tag{11}$$

This is defined not in terms of the expenditure function but of the distance function, $d(q,u_0) \equiv Max_d \{d: u(q/d) \geq u_0\}$; and, like the Konüs index, it is evaluated at a reference utility level $u_r$.

If tastes are homothetic, all three indexes reduce to the ratio of utility levels, $u_j/u_k$ (since the expenditure function and distance function become $e(p,u)=u.e(p)$ and $d(q,u_0)=u(q)/u_0$ respectively). But otherwise the three indexes differ among themselves and the value of each one depends on the reference price vector or utility level chosen. This underlines the fact that there is no such thing as a *unique* measure of real income.

Finally, what can be said about the different empirical index numbers introduced in the last section in the light of the economic approach to index numbers? As far as the Geary method is concerned, the consensus appears to be that it has no basis in economic theory.[7] However, the EKS and CCD indexes have obtained considerable support from results of Konüs and Byushgens (1926) and Diewert (1976) which relate the bilateral Fisher and Törnqvist indexes to particular specifications of preferences. In the next section I review these results and consider their relevance to multilateral comparisons.

## 3. SUPERLATIVE INDEXES AND MULTILATERAL COMPARISONS OF REAL INCOME

The first result relating true to empirical indexes is the following:

***Result 1*** [Konüs and Byushgens (1926)]: *The Fisher Ideal index is exact when the utility function is a homogeneous quadratic:* $u=(q'Aq)^{1/2}$, *A symmetric.*

Since tastes are homothetic in this case, saying that the Fisher Ideal index is exact means simply that it equals the ratio of the utility levels in the two countries:

$$\ln Q^{F}_{jk} = \ln u_j - \ln u_k . \tag{12}$$

As Diewert (1976) has noted, the quadratic utility function is a *flexible* functional form, *i.e*, it provides a second-order approximation to an arbitrary twice-differentiable linearly homogeneous utility function. He argues strongly for the use of index numbers which are *superlative*, in the sense that they are exact for flexible functional forms, and Result 1 shows that the Fisher Ideal index is superlative.

Result 1 would appear to justify the use of the EKS index for multilateral comparisons, since it is an appropriate generalisation of the bilateral Fisher Ideal index. However, my first proposition throws doubt on this:

***Proposition 1****: The EKS index is exact when the utility function is a homogeneous quadratic.*

***Proof***: The proposition follows immediately on substituting from (12) into the expression for the EKS index (2)

$$\ln Q^{EKS}_{jk} = \frac{1}{m} \sum_{l=1}^{m} \left\{ (\ln u_j - \ln u_l) - (\ln u_l - \ln u_k) \right\}, \tag{13}$$

$$= \ln u_j - \ln u_k . \qquad \sim \tag{14}$$

At first sight, Proposition 1 appears to justify the use of the EKS method. Since the quadratic utility function is a flexible functional form, Proposition 1 implies that the EKS method is superlative. The difficulty with Proposition 1 is that it goes too far. It shows that, as far as economic theory is concerned, there is nothing to be gained by using the EKS procedure over the bilateral Fisher index. While the EKS index is exact for the quadratic utility function, it is also redundant, since it actually equals the bilateral Fisher index in that case. Of course, the EKS index by construction yields a transitive ranking of income levels, unlike the Fisher Ideal. However, this is a statistical property and does not imply that the EKS method approximates an underlying transitive preference ordering.

What about the non-homogeneous case? The quadratic utility function does not generalise to this case. However, its logarithmic equivalent, the translog, does. Consider first a result of Diewert's which deals with the homogeneous translog:

***Result 2*** [Diewert (1976)]: *The Törnqvist index, $Q_{jk}^T$, is exact when the utility function is a homogeneous translog*:

$$\ln u(q) = a_0 + a'\ln q + \tfrac{1}{2}(\ln q)'A\ln q\,. \qquad (15)$$

The translog is also a flexible functional form but is more general than the quadratic. Result 2 therefore suggests that the Törnqvist index is even more "superlative" than the Fisher Ideal. This in turn has been interpreted to justify the use of the CCD Index, which as we saw in Section 1 is the appropriate multilateral extension of the Törnqvist index. However, my next proposition shows that it fares no better than the EKS index:

***Proposition 2***: *The CCD index is exact when the utility function is a homogeneous translog.*

The proof is identical to that of Proposition 1. Hence, like the EKS, the CCD index is redundant when it is exact.

Consider next the extension of the translog to the non-homogeneous case, which leads to the translog distance function:[8]

$$\ln d(q,u) = a_0 + a'\ln q + \tfrac{1}{2}(\ln q)'A\ln q + b_0\ln u + (\ln u)b'\ln q + \tfrac{1}{2}c_0(\ln u)^2. \qquad (16)$$

The appropriate index number corresponding to this specification of preferences is given by another result of Diewert's:

***Result 3*** [Diewert (1976)]: *The Törnqvist index $Q^T_{jk}$ equals the Malmquist index $Q^M_{jk}$ (and so is exact) if the distance function is a general (non-homogeneous) translog and $Q^M_{jk}$ is evaluated at the geometric mean of the two countries' utilities, $u^{jk} = (u^j u^k)^{0.5}$.*

However, the difficulty with this result is that the Malmquist index is evaluated at a particular utility level which is specific to the two countries being compared. This suggests that the corresponding multilateral index, the CCD index, aggregates in general over *m* inconsistent bilateral comparisons. This is confirmed by the next proposition:

***Proposition 3****: The CCD index deviates systematically from the Malmquist index evaluated at the geometric mean of the m countries' utilities, if the distance function is a general (non-homogeneous) translog.*

It is straightforward to calculate the bias of the CCD index explicitly:

$$\ln Q^{CCD}_{jk} - \ln Q^{M}_{jk} = \tfrac{1}{2}(\ln u_j - \ln u^*)b'(\ln q^j - \ln q^*) - \tfrac{1}{2}(\ln u_k - \ln u^*)b'(\ln q^k - \ln q^*) \,. \tag{17}$$

where:

$$\ln u^* \equiv \tfrac{1}{m}\sum_h \ln u_h \qquad \text{and} \qquad \ln q^* \equiv \tfrac{1}{m}\sum_h \ln q^h \,.$$

Equation (17) shows that the CCD index is exact only when tastes are homothetic (*i.e.* $b=0$), when from Proposition 2 it is redundant, or when the two countries compared deviate symmetrically from average. Proponents of the CCD index can perhaps draw some consolation from the fact that, in any sample of countries, the bias of a given bilateral comparison depends only on the deviations of utilities and quantities in the two countries being compared from the corresponding worldwide averages. However, it should be noted that $u^*$ and $q^*$ are not mutually consistent in general.

## 4. THE GEARY-KONÜS METHOD

The propositions in the last section throw doubt on the claims that the EKS and CCD indexes have a firm basis in economic theory when applied to international comparisons of expenditure patterns. By contrast, the Geary method at least uses a consistent set of world prices to compare real incomes. However, it suffers from the drawback of all fixed-weight indexes in that it does not allow any substitution in consumption. In this section I propose a new set of true indexes which overcome this drawback while preserving the spirit of the Geary method.

The first step is to replace the fixed-weight Laspeyres formula in the Geary exchange rates (6) with their true equivalents, which I call Geary-Konüs exchange rates:

$$\Pi_j = \frac{e(\pi, u_j)}{e(p^j, u_j)} = \frac{\sum_i \pi_i q_{ij}^*}{\sum_i p_{ij} q_{ij}}, \qquad j = 1, \ldots m. \tag{18}$$

Here the $q_{ij}^*$ denote the "virtual" or imputed quantities which country $j$ would choose if it were faced with world prices $\pi$:

$$q_{ij}^* = e_i(\pi, u_j). \tag{19}$$

Comparing the prices of all countries with a common world price vector is unremarkable in itself. The next step is to require that the world prices satisfy aggregation conditions of the Geary type. They cannot do so in terms of actual quantities consumed but they can in terms of imputed quantities. This leads to a set of Geary-Konüs world prices:

$$? _i = \frac{\sum_j ?_j p_{ij} q_{ij}}{\sum_j q^*_{ij}}, \qquad i = 1, \ldots n. \tag{20}$$

Finally, the implied measures of income at world prices are defined as follows:

$$z^*_j = E_j z_j = \sum_i^n ?_i q^*_{ij} = e(?, u_j), \qquad j = 1, \ldots m. \tag{21}$$

The advantages of this proposed system are that it combines the best features of the economic approach to index numbers and the Geary method. Like the former, it is firmly based on the microeconomic theory of the consumer and allows for the possibility of inter-commodity substitution. Like the latter, it satisfies matrix consistency, albeit in terms of imputed rather than actual consumption levels: it leads to a matrix of expenditure levels expressed in a common world currency which can be consistently aggregated across countries and across commodities. Finally, the system presented here avoids the conflict between bilateral Allen and Konüs quantity indexes noted in Section 2: each exchange rate $E_j$ is a Konüs true price index, while the real income measures (21) imply Allen true indexes of real income, $Q^*_{jk} = z^*_j/z^*_k$, using ? as the reference prices.

What is the relationship between the original Geary system and the "ideal" system proposed here? Firstly, it is immediate though worthy of note that the two systems coincide when there are no substitution possibilities in consumption:

***Proposition 4**: The Geary-Konüs world prices and real incomes coincide with those from the Geary system when preferences are* *of the fixed coefficients kind.*

More generally, the relationship between the two systems is most easily seen by considering the explicit solution for the Geary world prices:

$$\mathrm{p} = [\hat{q} - Z_{-m}\hat{z}_{-m}^{-1}Q'_{-m}]^{-1} z^{m}. \tag{22}$$

where $\hat{q}$ and $Q_{-m}$ are quantity matrices, $\hat{z}_{-m}$ and $Z_{-m}$ are expenditure matrices, and the exchange rate of an arbitrary country (indexed by *m*) has been set equal to one. (Details are given in the Appendix.) From (18) and (20), the Geary-Konüs world prices are given by an equation identical to (22), except with $Q^{*}$ and $\hat{q}^{*}$ instead of $Q$ and $\hat{q}$. Hence the explicit relationship between the Geary-Konüs and Geary prices is as follows:[9]

$$\pi = \mathrm{p} - R^{-1}\{R^{-1} + (\Delta R)^{-1}\}^{-1}R^{-1}z^{m}, \tag{23}$$

where:

$$R \equiv \hat{q} - Z_{-m}\hat{z}_{-m}^{-1}Q'_{-m} \quad \text{and} \quad \Delta R \equiv R - \hat{q}^{*} + Z_{-m}\hat{z}_{-m}^{-1}Q^{*\prime}_{-m}. \tag{24}$$

Note that $\Delta R$ is the only unobservable term in this expression.

Equation (22) also suggests an algorithm for estimating the Geary-Konüs prices. First, calculate Geary prices in the normal way, then use p and the estimated functional form of the demand functions to calculate first-round estimates of $Q^{*}$ and $\hat{q}^{*}$. Then, reapply the Geary formula using these estimates and repeat until the process converges. While this may not be the most efficient algorithm from a computational point of view, it shows clearly that the Geary-Konüs prices can be viewed as the outcome of a tâtonnement process which adjusts prices at each stage to ensure worldwide imputed commodity balance. Moreover, it gives a further justification for the Geary method, as providing a first-round approximation to the true but unobservable

Geary-Konüs prices and real incomes.

## 5. INTERPRETATION OF THE WORLD PRICES

The Geary-Konüs system satisfies base-country invariance and matrix consistency only because it chooses a particular set of world prices. This raises the question: to which countries, if any, do the world prices correspond? In the ICP, the Geary prices have been found to come closest to the prices of middle-income countries such as Italy. In this section, I show why this must be so for the Geary-Konüs prices, under a very wide class of preferences.

This issue is most easily addressed when the equations defining the world prices and true exchange rates are re-expressed in terms of budget shares. In the Geary case, equations (6) and (7) yield:

$$S_j\ ß_j^G\ ?_{ij} \quad ' \quad S_j\ ß_j^G ?_{ij}^G, \qquad i\ '\ 1, \ldots .\ n, \tag{25}$$

where the $?_{ij}$ are actual budget shares, $p_{ij}q_{ij}/z_j$; the $?_{ij}^G$ are budget shares in world prices, $p_{ij}q_{ij}/z_j^G$; and the $ß_j^G$ are the shares of each country in world income, $z_j^G/S_k z_k^G$. Expressing the aggregation condition in this way is not so useful because the budget shares $?_{ij}^G$ have no behavioral significance, since they are chosen facing prices $p^j$ but aggregated using prices $p$. The same is not true, however, of the corresponding equation for the Geary-Konüs system:

$$S_j\ ß_j^(\ ?_{ij} \quad ' \quad S_j\ ß_j^( ?_{ij}^(, \qquad i\ '\ 1, \ldots .\ n, \tag{26}$$

where $?_{ij}^* = ?_{ij} q_{ij}^*/z_j^*$ and $ß_j^* = z_j^*/S_k z_k^*$. Now the budget shares at world prices $?_{ij}^*$ are both generated by and aggregated by the same world prices and so they have a straightforward behavioral interpretation which links our results with the theory of linear aggregation developed by Gorman (1953) and Muellbauer (1975). The key result is the following:

***Proposition 5**: If preferences exhibit Generalized Linearity, then world demand patterns would be generated by a hypothetical country facing the Geary-Konüs world prices and with an income equal to a weighted quasi-linear mean of the individual countries' incomes.*

Generalized Linearity is a specification of preferences introduced by Muellbauer (1975) which implies an expenditure function of the following form:

$$e(?\,,u) \quad ' \quad ?[\,a(?\,),b(?\,),u]\;, \tag{27}$$

where the functions *a* and *b* are linearly homogeneous in ? and the function ? is linearly homogeneous in *(a,b)*. Muellbauer also shows that the budget shares in this case are:

$$?^{(}_{ij} \quad ' \quad ?^{GL}_{i}(?\,,z^{(}_{j}) \quad ' \quad f\,(?\,,z^{(}_{j})\,.\,A_{i}(?\,)\;\%\;B_{i}(?\,)\;. \tag{28}$$

where $A_i$ and $B_i$ are commodity-specific functions which are independent of income. Substituting into (26):

$$\overline{?}^{(}_{i} \quad / \quad \underset{j}{S}\;ß^{(}_{j}?_{ij} \quad ' \quad ?^{GL}_{i}(?\,,\tilde{z}^{(})\;, \tag{29}$$

where:

$$\tilde{z}^{(} \quad ' \quad f^{\&1}\left(?\,,\underset{j}{S}\;ß^{(}_{j}\,f\,(z^{(}_{j},?\,)\right)\;. \tag{30}$$

Given ? , $\tilde{z}^{*}$ is a *symmetric mean*, or more specifically a *weighted quasi-linear mean* of the $z_j^*$.[10] Equation (29) thus states that world expenditure patterns, in the sense of the world budget shares at world prices, would be generated by a hypothetical country which faces the same prices and whose income is an appropriate average of the individual countries' incomes.

Generalized linearity is an extremely general specification of preferences, which nests many of the most widely-used demand systems. Proposition 5 is significantly strengthened when we specialise to some of these sub-cases:

### *i. Price-Independent GL ("PIGL") Preferences*

In this case, also due to Muellbauer (1975), the expenditure function specialises to a CES form:

$$e(?\,,u) \quad ' \quad [(1\&u)\,a(?\,)^{a} \;\%\; u\,.\,b(?\,)^{a}\,]^{\frac{1}{a}}\,, \tag{31}$$

and the income function in the budget shares is independent of prices (whence the name):

$$?_{ij}^{(} \quad ' \quad ?_{i}^{PIGL}(?\,,z_j^{(}) \quad ' \quad f\,(z_j^{(})\,.\,A_i(?\,) \;\%\; B_i(?\,)\,. \tag{32}$$

It follows that the average income level which generates world spending patterns at the Geary-Konüs prices is also independent of prices and equals a CES mean of individual countries' incomes:

$$\tilde{z}^{(} \quad ' \quad \left\{ \underset{j}{S}\; ß_j^{(}\,(z_j^{(})^{\&a} \right\}^{\&\frac{1}{a}}. \tag{33}$$

### *ii. Price-Independent Generalized Logarithmic ("PIGLOG") Preferences*

This system, which nests the AIDS model of Deaton and Muellbauer (1980), is the limit of the PIGL system as a approaches zero. Hence the expenditure function takes a Cobb-Douglas form:

$$\ln e(\pi, u) = (1-u)\ln a(\pi) + u \ln b(\pi)\,; \tag{34}$$

the budget shares depend on ln $z_j^*$; and average world income is a weighted geometric mean of individual countries' incomes:

$$\ln \tilde{z}^* = \sum_j \beta_j^* \ln z_j^*\,. \tag{35}$$

### *iii. The Gorman Polar Form*

A different special case of PIGL, obtained by setting a equal to one, is the Gorman polar form, which nests the Linear Expenditure System corresponding to the Stone-Geary utility function. The expenditure function is now:

$$e(\pi, u) = (1-u)\, a(\pi) + u . b(\pi)\,, \tag{36}$$

and the budget shares are:

$$\omega_{ij}^* = \omega_i^{GPF}(\pi, z_j^*) = (z_j^*)^{-1} . A_i(\pi) + B_i(\pi)\,. \tag{37}$$

In this case, the Geary-Konüs prices generate demand patterns which aggregate not just in the sense of yielding the same budget shares but in the much stronger sense of yielding the same *levels* of world expenditure on each commodity:

$$\sum_j E_j p_{ij} q_{ij} = \sum_j z_j^* \omega_{ij} = m[A_i(\pi) + \bar{z}^* B_i(\pi)] = m\bar{z}^* \bar{\omega}_i^*\,, \tag{38}$$

where:

$$\bar{?}_i^{*} = ?_i^{GPF}(?, \bar{z}^{*}) \quad \text{and} \quad \bar{z}^{*} = \frac{1}{m} \sum_j z_j^{*} . \tag{39}$$

Thus, when preferences exhibit the Gorman polar form, the Geary-Konüs world prices would generate actual world demands if world income was equally distributed (or, since expenditure is linear in utility from (36), if world utility was equally distributed).

## 6. CONCLUSIONS

While many researchers have worked with the Penn World Tables, not many have asked what exactly the numbers mean, and those who have considered the question have mostly advocated very different methods for calculating real or purchasing-power-corrected incomes. In this paper I have proposed a new standard for such international comparisons which combines the best features of the economic approach to index numbers and of the existing Geary method. The method I propose meets all of the tests discussed in Section 2, except that of "characteristicity"; it allows for inter-commodity substitution; and it relates directly to the theory of linear aggregation.

I have also emphasised that there is no unique "true" measure of real incomes. On the contrary, even confining attention to methods based on the expenditure function, there are as many methods as there are candidate reference price vectors. For different purposes, different reference vectors may be preferred: for example, a Swiss multinational wishing to calculate local allowances for its executives should ideally use a "star" system based on Swiss prices rather than the method proposed here. However, for researchers interested in world growth patterns it makes sense to have a system which disaggregates consistently across commodities and countries. By contrast, the EKS and CCD methods provide good (second-order) approximations to an *inconsistent* set of multilateral comparisons and can only be recommended if tastes are close to homothetic.

As far as the Penn World Tables themselves are concerned, the bottom line of this paper is that the Geary method which underlies them is an acceptable, though not necessarily a particularly good, approximation to the ideal system. Of course, an appreciation of the theoretical underpinnings of the Geary method draws attention to potential pitfalls in applying it. For example, since the Geary method is exact only when preferences are of the fixed-coefficient type, it would not be appropriate to use the Penn World Tables to *test* hypotheses concerning the degree of inter-commodity substitutability.

Two drawbacks of the system proposed here, one genuine and one illusory, should be mentioned. A genuine objection to the Geary-Konüs system is that it draws on consumer theory only and hence relates only to comparisons of household expenditure patterns. Appropriate theoretical foundations for making international comparisons of investment and government spending must await further research. Of course, an alternative approach is to make international comparisons of real *output* only but such comparisons have no implications for *income* or *living standards*. A second but in my view spurious objection to the approach adopted here is that it assumes that tastes are identical worldwide. Taken literally, this is clearly an implausible assumption, but it should be seen as a necessary requirement for making international comparisons in the first place. Insofar as data on real income have any meaning, it is that they provide an answer to the question "how well-off would the same reference consumer be in different countries?" Of the multitude of candidate reference consumers, it seems sensible for economists to take the hypothetical consumer whose consumption patterns mimic world consumption behaviour as closely as possible.

Finally, the results of this paper have implications for other issues. At an empirical level, they point to the need for more international studies of demand patterns which may provide the basis for true measures of real income based on the Geary-Konüs rather than the raw Geary method. The paper's results also have implications for time-series comparisons. The requirement of "characteristicity" is often taken for granted in a time-series context: why should the growth rate of real income between 1970 and 1971 change when data on 1995 become available? But once it is recognised that there is no such thing as the "true" growth rate of real income, this question loses its paradoxical character. For some purposes, a consistent time series expressed in terms of the prices of a central year or even of a consistent average of different years, as in the Geary-Konüs system, may be more appropriate. As the results of Section 3 have shown, superlative indexes have desirable properties only in bilateral contexts. Since almost all comparisons in economics are multilateral, this calls into question the standard practice of "chaining" together a set of bilateral annual comparisons to produce a multi-period time series.[11]

## ENDNOTES

1. See also the dismissive remarks by Samuelson and Swamy (1974, p. 591), Caves, Christensen and Diewert (1982, p. 83) and Samuelson (1984, p. 277 and 1994, p. 212).

2. These two organisations convened a conference in 1989, at which their expert advisors failed to agree on whether the Geary or EKS methods should be adopted. As a result, the OECD now publishes annually two complete tables of real income indexes for its member countries. However, "Eurostat requires that only one set of results be recognised as the official results of the Community," so that based on the EKS method is released a year before that based on the Geary method. See OECD (1990).

3. The relationship between the Törnqvist and Fisher indexes may be seen more clearly by rewriting the latter as: $\ln Q^F_{jk} = ½(S_i \ln ?_{ij} q_{ij}/q_{ik} \; ! \; S_i \ln ?_{ik} q_{ik}/q_{ij})$.

4. The ICP defines true exchange rates or "purchasing power parities" as the inverse of (6), following the U.S. convention of measuring exchange rates. I follow Geary in using the U.K. convention, since it facilitates the matrix derivations.

5. Overviews of the vast literature on index numbers may be found in Pollak (1971) and Diewert (1981) and (1987). A more extended but non-technical treatment of the issues specific to international comparisons of real income may be found in Neary (1996).

6. For alternative perspectives, see Diewert (1988) and Eichhorn and Voeller (1990), who apply the test approach to the choice of indexes for international comparisons without mentioning the Geary method. See also Diewert (1987, Section 9).

7. An exception to this rule is Marris (1984), who compares the Geary method with a set of multilateral Allen indexes, though without discussing how the world prices may be calculated. Geary himself did not provide any theoretical justification for his system, other than remarking in passing: "if the entities $p_i$ and $e_j$ exist, they could scarcely be defined reasonably in any other terms."

8. The vector b is the source of non-homogeneity, which may be seen from the equations for the budget shares: ? = a + Alnq + blnu. When b is zero, (16) reduces to (15) by setting d(q,u)=1 and (without loss of generality) normalizing $b_0$=!1 and $c_0$=0.

9. This makes use of a standard result in matrix algebra: $(A+B)^{!1}=A^{!1}!A^{!1}(A^{!1}+B^{!1})^{!1}A^{!1}$.

10. A symmetric mean is a function which is symmetric in the $z_j^*$ and which equals $z_0$ when $z_j^*=z_0$, all j. See Chew (1983) and Diewert and Nakamura (1993, chap. 14).

11. This rationale for using methods such as the Geary method in contexts other than international comparisons is very different from that proposed by Khamis (1972).

# APPENDIX: CALCULATING THE GEARY AND GEARY-KONÜS PRICES

The $m+n$ equations (6) and (7) are not independent since both imply the same aggregate equation:

$$\sum_i \pi_i \left( \sum_j q_{ij} \right) = \sum_j e_j \left( \sum_i p_{ij} q_{ij} \right). \qquad (40)$$

Hence, we need to drop one equation and normalise one unknown. Following the ICP convention, we set the exchange rate for one country, indexed $m$, equal to unity and drop the equation from (6) corresponding to that country. Then, to solve explicitly for e and π, rewrite (6) and (7) as a single matrix equation as follows:

$$\begin{bmatrix} \hat{z}_{-m} & -Q'_{-m} \\ Z_{-m} & -\hat{q} \end{bmatrix} \begin{bmatrix} e_{-m} \\ \pi \end{bmatrix} = \begin{bmatrix} 0 \\ -z^m \end{bmatrix}. \qquad (41)$$

Here, $z$ denotes the $m$-by-one vector of total expenditures by countries, with typical element $z^j = p^j.q^j$; $Z$ denotes the $n$-by-$m$ matrix of expenditures by commodity and country, with typical element $z_{ij} = p_{ij}q_{ij}$; $z^m$ denotes the $n$-by-one vector of expenditures by commodity in country $m$ (i.e., the final column of $Z$); $Q$ denotes the $n$-by-$m$ matrix of quantities by commodity and country, with typical element $q_{ij}$; and $q$ denotes the $n$-by-one vector of world consumption levels of each commodity, with typical element $q^i = \Sigma_j \; q_{ij}$. Finally, a prime (′) denotes a transpose; a circumflex (^) over a vector denotes a diagonal matrix formed by placing on the principal diagonal the corresponding elements of the vector; and the subscript ($_{-m}$) denotes a vector or matrix from which the entries corresponding to the reference country (numbered $m$) have been deleted. Note that $z$ can be written as $Z'i_n$ and $q$ as $Qi_m$, where $i_h$ is a $h$-by-one vector of ones; and that $Z_{-m}\hat{z}^{-1}_{-m}$ is the matrix of world budget shares, with typical element $\omega_{ij} = z_{ij}/z^j$.

Equation (41) can now be solved to give (22) using the formula for inverting a partitioned matrix. Replacing $Q$ and $\hat{q}$ with $Q^*$ and $\hat{q}^*$ gives the Geary-Konüs world prices, though since $Q^*$ and $\hat{q}^*$ themselves depend on ? this is not a closed form.

## REFERENCES


Afriat, S.N. (1977), *The Price Index*, Cambridge: Cambridge University Press.

Caves, D.W., L.R. Christensen and W.E. Diewert (1982), 'Multilateral comparisons of output, input and productivity using superlative index numbers', Economic Journal, 92, pp.73-86.

Chew, S.H. (1983), 'A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox', Econometrica, 51, pp.1065-1092.

Deaton, A.S. and J. Muellbauer (1980), 'An almost ideal demand system', American Economic Review, 70, pp.312-326.

Diewert, W.E. (1976), 'Exact and superlative index numbers', Journal of Econometrics, 4, pp.115-145; reprinted in W.E. Diewert and A.O. Nakamura (eds.), *op. cit.*

Diewert, W.E. (1981), 'The economic theory of index numbers: A survey', in A. Deaton (ed.): *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, London: Cambridge University Press, pp.163-208; reprinted in W.E. Diewert and A.O. Nakamura (eds.), *op. cit.*

Diewert, W.E. (1987), 'Index numbers', in J. Eatwell, M. Milgate and P. Newman (eds.), *The New Palgrave: A Dictionary of Economics*, Volume 2, London: Macmillan, pp.767-780; reprinted in W.E. Diewert and A.O. Nakamura (eds.), *op. cit.*

Diewert, W.E. (1988), 'Test approaches to international comparisons', in W. Eichhorn (ed.), *Measurement in Economics*, Heidelberg: Physica-Verlag, pp.67-86; reprinted in W.E. Diewert and A.O.

Nakamura (eds.), *op. cit.*

Diewert, W.E. and A.O. Nakamura (eds.) (1993), *Essays in Index Number Theory*, Volume 1, Amsterdam: North-Holland.

Dreschler, L. (1973), 'Weighting of index numbers in multilateral international comparisons', Review of Income and Wealth, 19, pp.17-34.

Eichhorn, W. and J. Voeller (1990), 'Axiomatic foundation of price indexes and purchasing power parities', in W.E. Diewert (ed.): *Price Level Measurement*, Amsterdam: North-Holland, pp.321-358.

Eltetö, O. and P. Köves (1964), 'On a problem of index number computation relating to international comparisons' (in Hungarian), *Statisztikai Szemle*, 42, pp.507-518.

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

Frisch, R. (1936), 'Annual survey of general economic theory: The problem of index numbers', Econometrica, 4, pp.1-39.

Geary, R.C. (1958), 'A note on the comparison of exchange rates and purchasing power between countries', Journal of the Royal Statistical Society, Series A, 121, pp.97-99.

Gorman, W.M. (1953), 'Community preference fields', Econometrica, 21, pp.63-80.

Khamis, S.H. (1972), 'A new system of index numbers for national and international purposes', Journal of the Royal Statistical Society, Series A, 135, pp.96-121.

Konüs, A.A., and S.S. Byushgens (1926), 'On the problem of the purchasing power of money' (in Russian), Voprosi Konyunkturi II(1), pp.151-172.

Kravis, I.B. (1984), 'Comparative studies of national incomes and prices', Journal of Economic Literature, 22, pp.1-39.

Marris, R. (1984), 'Comparing the incomes of nations: A critique of the international comparison project', Journal of Economic Literature, 22, pp.40-57.

Muellbauer, J. (1975), 'Aggregation, income distribution and consumer demand', Review of Economic Studies, 62, pp.525-543.

Neary, J.P. (1996), 'Theoretical foundations of the 'Geary method' for international comparisons of purchasing power and real incomes', Economic and Social Review, 27:2 (January), pp.161-179.

OECD (1990), *Purchasing Power Parities and Real Expenditures: EKS Results, Volume 1*, Paris: Organisation for Economic Co-Operation and Development.

Pollak, R.A. (1971), 'The theory of the cost-of-living index', U.S. Bureau of Labor Statistics Working Paper 11; revised version in W.E. Diewert (ed.) (1990): *Price Level Measurement*, Amsterdam: North-Holland, pp.5-77.

Prasada Rao, D.S., E.A. Selvanathan and D. Pilat (1995), 'Generalized Theil-Törnqvist indices with applications to international comparisons of prices and real output', Review of Economics and Statistics, 77, pp.352-360.

Samuelson, P.A. (1984), 'Second thoughts on analytical income comparisons', Economic Journal, 94, pp.267-278.

Samuelson, P.A. (1994), 'Facets of Balassa-Samuelson thirty years later', Review of International Economics, 2, pp.201-226.

Samuelson, P.A. and S. Swamy (1974), 'Invariant economic index numbers and canonical duality: Survey and synthesis', American Economic Review, 64, pp.566-593.

Summers, R. and A. Heston (1991), 'The Penn World Tables (Mark 5): An expanded set of international comparisons, 1950-1988, *Quarterly* Journal of Economics, 106, pp.327-336.

Szulc, B.J. (1964), 'Index numbers for multilateral regional comparisons' (in Polish), Przeglad Statystyczny, 3, pp.239-254.