The Economic and Social Review, Vol. 28, No. 4, October, 1997, pp. 401-421

Comparing the Wealth of Nations: Reference Prices and Multilateral Real Income Indexes*

J. PETER NEARY

and

BRÍD GLEESON
University College Dublin

Abstract: This paper considers the problem of comparing real incomes across countries. The available methods are reviewed and their performance is compared using the raw data underlying the Penn World Table. We propose and implement a test of the "Gerschenkron Effect": the downward bias in a country's measured real income when its own prices are used as weights. The test confirms that this Effect obtains with fixed-weight indexes but not with true bilateral indexes based on empirical demand parameters. We also demonstrate the feasibility of estimating the GAIA ("Geary-Allen International Accounts") System which yields true multilateral real income indexes.

I INTRODUCTION

It is well known that international comparisons of real income are sensitive to the reference prices used. One hypothesis concerning the nature of the sensitivity involved is the so-called "Gerschenkron Effect": a country's measured real income is higher the more the reference prices differ

Paper presented at the Eleventh Annual Conference of the Irish Economic Association.

*This research 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. We are very grateful to Bettina Aten at the University of Illinois for making the data available to us; to Gavan Conlon and Dave Madden for providing the econometric estimates used in Section IV; and to Patrick Honohan, Colm McCarthy, an anonymous referee and participants at a seminar in University College Dublin and at the 1997 Irish Economic Association Annual Conference for helpful comments.

from its own prices. This effect is usually thought to reflect the bias arising from the lack of substitutability assumed in standard real income indexes. Less widely appreciated is the fact that even "true" indexes, which are based on utility theory and so allow for substitution, are in general sensitive to the reference price vector used. In this paper we set out to test the importance of these sensitivities, using the raw data which underlie the Penn World Table.

The other issue we address in the paper is the choice of index number for use in multilateral comparisons of real incomes between countries. Such comparisons have become increasingly common in recent years, reflecting the greater availability of comparative data, of which the most extensive source is the Penn World Table which draws on the United Nations sponsored International Comparisons Project (ICP).[1] Two principal methods are used in practice in such comparisons. The ICP and Penn World Table use a method originated by Geary (1958), which calculates world prices and exchange rates that correct for deviations from purchasing power parity. However, it has been extensively criticised for its lack of theoretical foundations, notably by Samuelson and Swamy (1974) and Diewert (1981 and 1987). Diewert in particular has argued instead in favour of versions of the so-called EKS method, due to Eltetö and Köves (1964) and Szulc (1964), which is used by the OECD and by Eurostat.

Recent work by Neary (1996a, 1996b) argues that the claims made for the theoretical superiority of the EKS and related indexes do not hold up if tastes are not homothetic. Moreover, in that case, it is necessary to select from an infinity of "true" indexes to use as an ideal or benchmark. Neary proposes for such a benchmark the "GAIA" or "Geary-Allen International Accounts" System, which has the ease of interpretation of the Geary system but is consistent with utility theory in the context of multilateral comparisons. However, the ease of computation of such GAIA indexes and their relationship in practice with the Geary and EKS indexes remain open questions. These issues are addressed below. First, we present a brief overview of some relevant results from the theory of index numbers.

II INDEX NUMBER THEORY[2]

2.1 *Fixed-Weight Bilateral Comparisons and the Gerschenkron Effect*

Figure 1 illustrates the problems which arise in making bilateral comparisons of real incomes across countries. Points J and K represent the quantities consumed of goods 1 and 2 while the slopes of the lines through those points represent the domestic relative prices of good 1 in each of two

1. See Kravis (1984) and Summers and Heston (1991) for overviews of the ICP and the Penn World Table respectively.

2. See Diewert (1981, 1987), Pollak (1971) and Neary (1996a) for more detailed accounts.

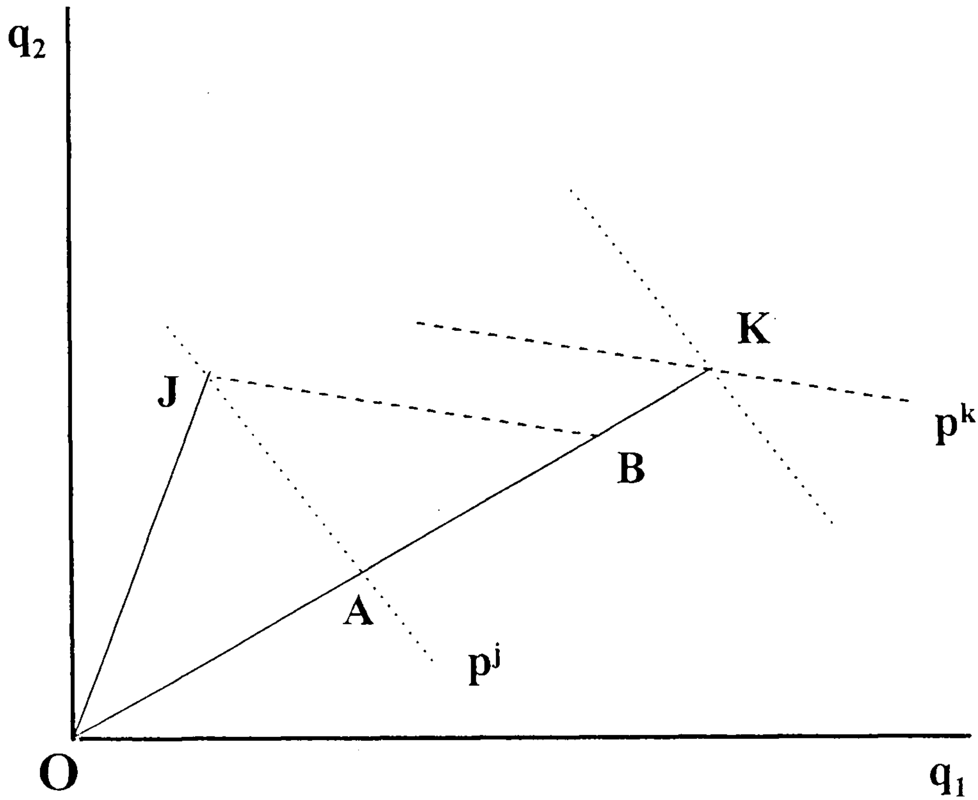

Figure 1: *Fixed-Weight Indexes of Country j's Real Income Relative to Country k: Laspeyres = OB / OK; Paasche = OA / OK*

countries j and k. It is convenient to choose one country as reference and, following ICP conventions, we select the high-income country k. (This choice is arbitrary and does not affect the results in a substantive way.) Hence we seek a measure of the real income of country j relative to that of country k and we measure expenditures in country k's currency, converting at current exchange rates. By this measuring rod, the two country's total expenditures in domestic prices equal OA and OK respectively.

The two countries' real incomes can now be compared at either set of domestic prices. Using country k's prices leads to the Laspeyres measure, $p^k.q^j / p^k.q^k$, given by OB/OK in Figure 1.[3] By contrast, using country j's prices leads to the Paasche measure, $p^j.q^j / p^j.q^k$, given by OA/OK. The latter is clearly much lower, reflecting the "Gerschenkron Effect" due to

3. With m commodities, we let p^h and q^h denote the m-by-1 vectors of prices and quantities respectively in country h. A dot denotes a vector inner product.

Gerschenkron (1951). To quote Samuelson, "It is always better to ride the other man's horse"; more precisely, each country's relative real income is lower when the comparison is made at its own prices and higher when it is made at the other's. However, as we shall see, this outcome is not inevitable on theoretical grounds and so its empirical relevance is an open question which deserves investigation.

2.2 *Reference Prices and True Bilateral Comparisons*

The fact that fixed-weight indexes of real income are sensitive to the reference prices used suggests that we should consider a "true" or utility-based index instead. The most natural of these is the Allen (1949) quantity index, which equals the ratio of the expenditure functions of the two countries evaluated at a common reference price vector p^r:

$$\frac{e(p^r, u^j)}{e(p^r, u^k)}. \tag{1}$$

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. However, it is not independent of reference prices in general. An exception is the special case of homothetic tastes, when the expenditure function takes the form $e(p^r,u^j) = u^j.e(p^r)$, and so (1) reduces to the ratio of utilities, u^j/u^k, which is independent of p^r. More realistically, when tastes are not homothetic, the Allen index is sensitive to the choice of reference prices. Figure 2 illustrates the difference between what we call the "Laspeyres-Allen" and "Paasche-Allen" indexes, which use country k and country j prices as reference respectively.

While the Allen index does not (except when tastes are homothetic) avoid the dependence of our measure of real income on the reference prices used, we can use it to throw light on the Gerschenkron Effect. Since the expenditure function gives the minimum cost of attaining a given utility level facing particular prices, it follows that $e(p^k, u^j) \le p^k.q^j$. Hence the Laspeyres-Allen index cannot exceed the corresponding fixed-weight Laspeyres index:

$$\frac{e(p^k, u^j)}{e(p^k, u^k)} \le \frac{p^k.q^j}{p^k.q^k}. \tag{2}$$

By similar reasoning, the Paasche-Allen index cannot be less than the corresponding fixed-weight Paasche index:

$$\frac{e(p^j, u^j)}{e(p^j, u^k)} \ge \frac{p^j.q^j}{p^j.q^k}. \tag{3}$$

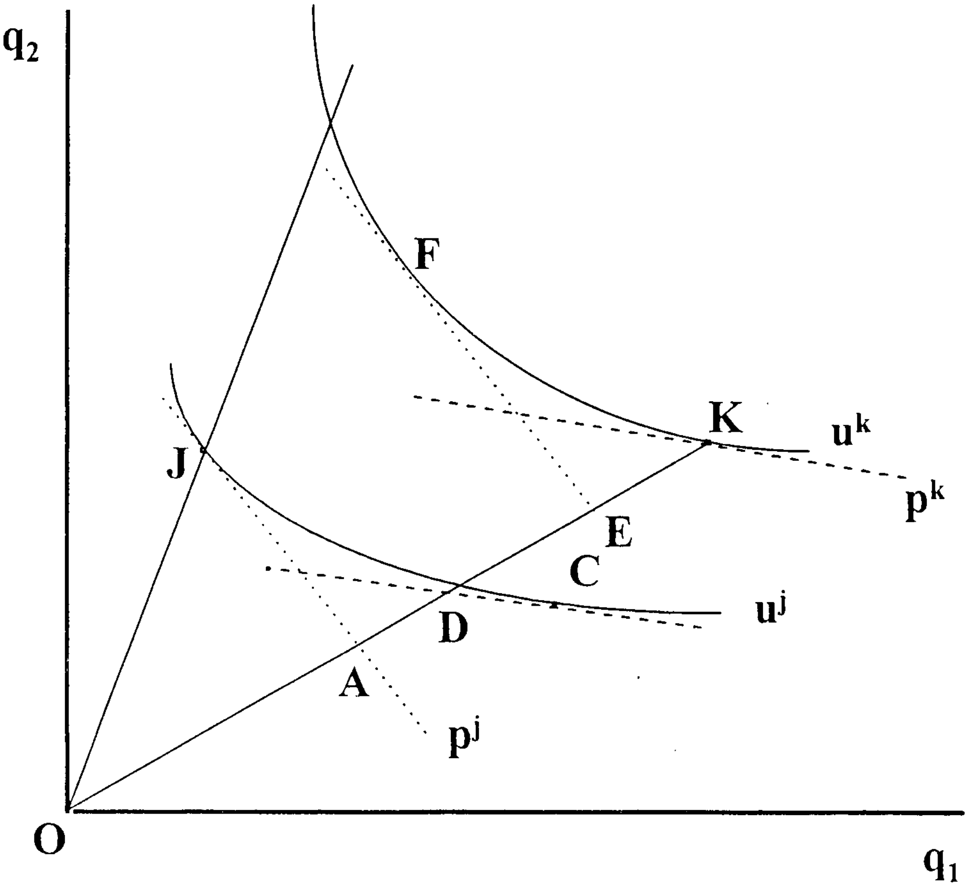

Figure 2: *True (Allen) Indexes of Country j's Real Income Relative to Country k: Laspeyres = OD / OK; Paasche = OA / OE*

These two inequalities can be combined in the special case of homothetic tastes. Substituting for the special form of the expenditure function in this case, the left-hand sides of (2) and (3) are equal to one another and the Laspeyres index necessarily exceeds the Paasche index. (More precisely, the Laspeyres index cannot be less than the Paasche index and must strictly exceed it if any inter-commodity substitution occurs.) This gives a key result: the Gerschenkron Effect is a necessary consequence of utility maximisation by a single consumer whose tastes are homothetic.[4]

Figure 3 illustrates the result. Points J and K represent countries j and k as before. Since tastes are homothetic, all indifference curves have the same

4. Nuxoll (1994) states that the Gerschenkron Effect is implied by the Weak Axiom of Revealed Preference when tastes are homothetic. For bilateral comparisons, the Weak Axiom is equivalent to utility maximisation.

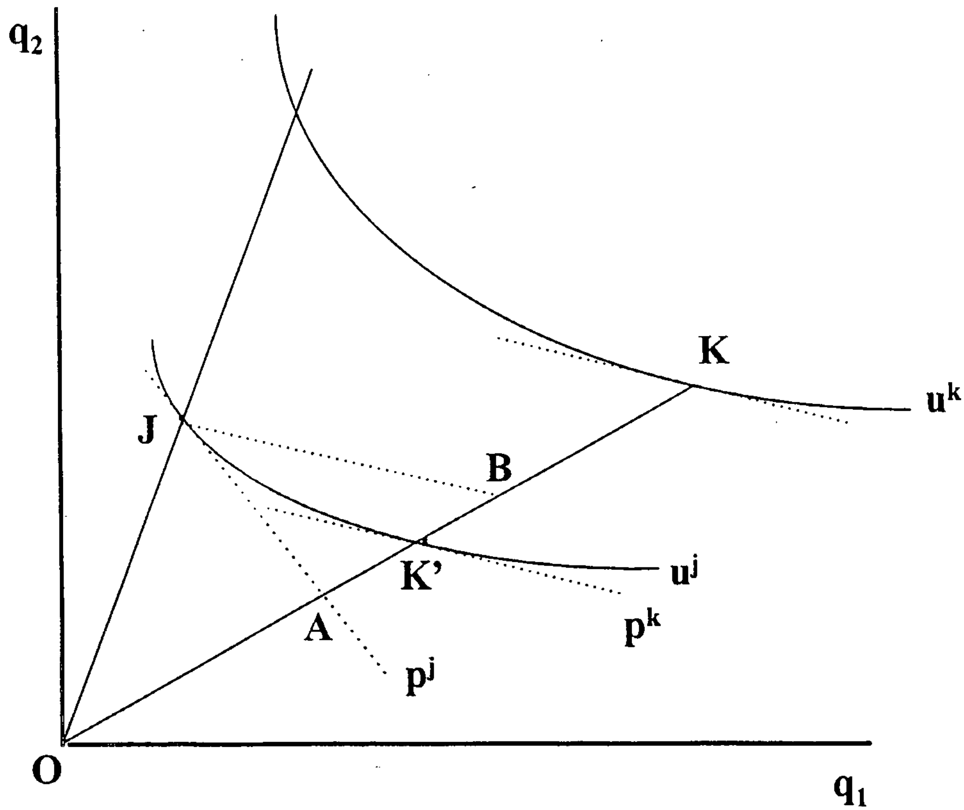

Figure 3: *Homothetic Preferences Imply the Gerschenkron Effect:* $OB/OK > OA/OK$

slope along the ray OK. In particular, the slope of the indifference curve through J at point K′ is the same as the slope of the price line at K. The Laspeyres-Allen index therefore equals OK′/OK. It follows immediately that the fixed-weight Laspeyres index, which equals OB/OK, must exceed the fixed-weight Paasche index, which equals OA/OK.

However, a similar geometric approach may be used to show that, if tastes are not homothetic, the Gerschenkron Effect may not arise. In Figure 4 good 1 is income-inelastic but nevertheless has a higher relative price in k than in j. As a result, the Laspeyres index OB/OK is less than the Paasche index OA/OK. The Gerschenkron Effect is violated, even though the data relate to a single utility-maximising individual. We may conclude from these examples that substitutability tends to encourage the Gerschenkron Effect and that non-homotheticity may, but need not, work against it. Of course, when we compare data for whole countries these results derived from individual behaviour need not apply.

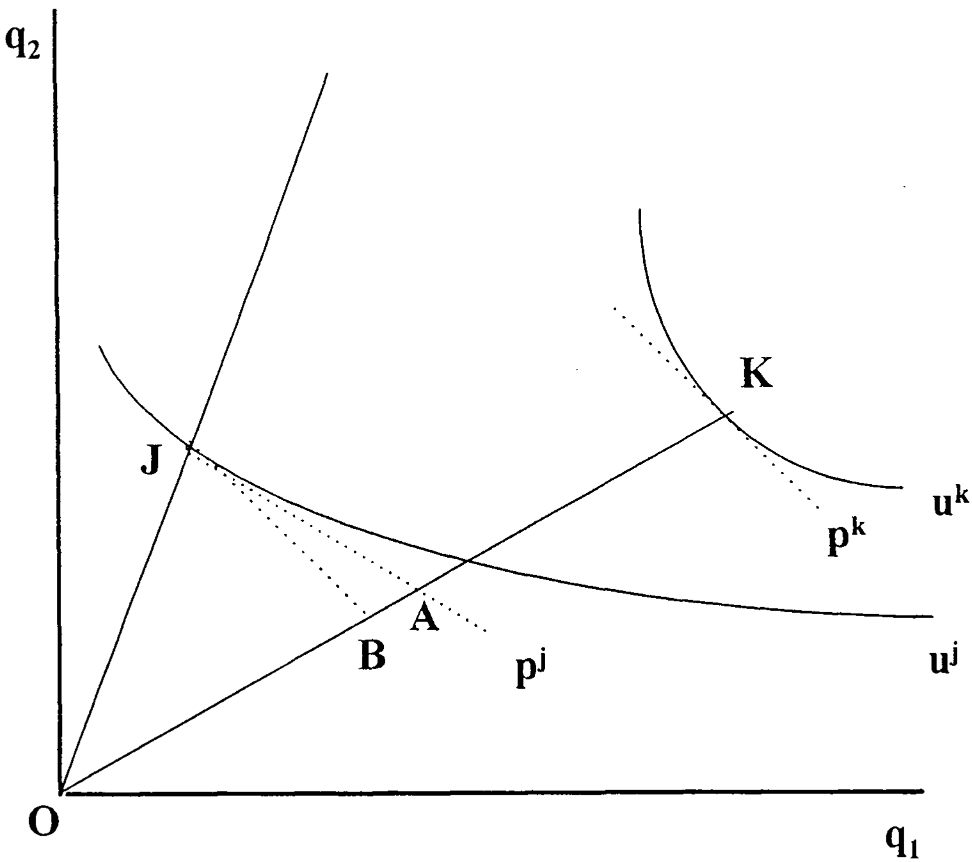

Figure 4: *Non-Homothetic Preferences may Violate the Gerschenkron Effect:* $OB/OK < OA/OK$

2.3 *Multilateral Comparisons*

So far, we have considered only bilateral international comparisons. To compare the real incomes of a group of countries the issues already considered are still relevant and many others also arise. Rather than giving a full account, we summarise the three multilateral indexes whose empirical performance is considered in the remainder of the paper.

The Geary System: The method proposed by Geary (1958) postulates the existence of "world" prices π and "true" exchange rates ε. The true exchange rates are Laspeyres price indexes, which compare the world prices with the prices of each country in turn:

$$\varepsilon_j = \frac{\sum_i \pi_i q_{ij}}{\sum_i p_{ij} q_{ij}}, \quad j = 1,\ldots m. \tag{4}$$

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:

$$\pi_i = \frac{\sum_j \varepsilon_j p_{ij} q_{ij}}{\sum_j q_{ij}}, \quad i = 1,\ldots n. \tag{5}$$

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

$$z_j^G = \varepsilon_j z_j = \sum_i \pi_i q_{ij}, \quad i = 1,\ldots m. \tag{6}$$

These real income measures in turn imply a set of indexes, $Q_{jk}^G = z_j^G / z_k^G, \forall\, j,k$. Thus the Geary method yields fixed-weight index numbers of real income which use as reference the prices π of a hypothetical "world" country.

The EKS Index: The starting point for the EKS system is the Fisher Ideal index, which is the geometric mean of the base-weighted Laspeyres index and the current-weighted Paasche index:

$$\ln Q_{jk}^F = \tfrac{1}{2}\left\{\ln \frac{p^k.q^j}{p^k.q^k} + \ln \frac{p^j.q^j}{p^j.q^k}\right\}. \tag{7}$$

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:

$$\ln Q_{jk}^{EKS} = \tfrac{1}{m} \sum_{l=1}^{m} \left\{\ln Q_{jl}^F - \ln Q_{kl}^F\right\}. \tag{8}$$

This index, unlike the Fisher index, yields a transitive ranking across countries and is not sensitive to the choice of base country. It also reduces to the Fisher index when m=2. Thus the EKS index is indeed an appropriate multilateral generalisation of the Fisher Ideal. Note that it is not possible to give a reference price interpretation to the EKS index.

The GAIA ("Geary-Allen International Accounts") System: This is an "ideal" counterpart to the Geary system, proposed by Neary (1996a and 1996b). Its starting point is to replace the fixed-weight Laspeyres formula in the Geary exchange rates with their true equivalents, which may be called Geary-Konüs exchange rates:

$$E_j = \frac{e(\Pi, u^j)}{e(p^j, u^j)} = \frac{\sum_i \Pi_i q_{ij}^*}{\sum_i p_{ij} q_{ij}}, \quad j = 1, \ldots m. \tag{9}$$

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

$$q_{ij}^* = e_i(\Pi, u^j). \tag{10}$$

As for the world prices Π_i, they must satisfy aggregation conditions of the Geary type. They cannot do so in terms of actual quantities consumed but they can in terms of virtual quantities. This leads to a set of Geary-Konüs world prices:

$$\Pi_i = \frac{\sum_j E_j p_{ij} q_{ij}}{\sum_j q_{ij}^*}, \quad i = 1, \ldots n. \tag{11}$$

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

$$z_j^* = E_j z_j = \sum_i^n \Pi_i q_{ij}^* = e(\Pi, u_j), \quad j = 1, \ldots m. \tag{12}$$

As noted in Neary (1996b), this system combines the theoretical consistency of Allen indexes with the ease of interpretation of the Geary system. It also allows an interpretation of the world prices: if preferences are characterised by "generalised linearity" as proposed by Muellbauer (1975), the world prices correspond to the prices of a hypothetical country whose income is an appropriate average of world incomes. Of course, notwithstanding these theoretical advantages, the GAIA system is unobservable and so its empirical relevance depends on how easily and plausibly it can be estimated using available information on demand parameters. We investigate this issue in Section V below. First we consider the sensitivity of both fixed-weight and true indexes to the choice of reference prices.

III LASPEYRES INDEXES AND THE GERSCHENKRON EFFECT

In this section and the next we examine empirically the significance of the issues raised in the theoretical discussion above. The data we use are taken from the International Comparison Project (ICP) which underlies the Penn World Table and are described in detail in the Appendix. In this pilot study we use data on 16 countries in 1970, giving the prices and quantities consumed of 11 categories of personal consumption. All prices are measured in current US dollars, converted at *market* exchange rates.

The first step is to calculate the real income of each country relative to the US, using the prices of each country in turn as reference, and the results of this are given in Table 1. The countries are ranked by z_j, their per capita expenditure in dollars, converted at current exchange rates. For reference, this is given in levels and relative to the US in the first two rows of the table. The next sixteen rows, constituting the bulk of the table, give the relative Laspeyres indexes of real income, using the prices of each country in turn as reference. Reading down each of the country columns gives alternative estimates of the real income of each country relative to the US. Reading across each row gives the indexes corresponding to a particular country's prices; i.e., the incomes of each country relative to the US using a particular country's prices as weights. Thus the entry in row k and column j equals $p^k.q^j / p^k.q^k$ over $p^k.q^1 / p^k.q^k$; i.e., $p^k.q^j / p^k.q^1$, the real income of country j relative to country 1 (the US) using the prices of country k as reference. This matrix may be called the Laspeyres "star" matrix, since each row gives a star index where the reference country is the centre and the other countries are the points of a star. The next three rows and the final three columns of the table give the average, standard deviation and coefficient of variation of the different Laspeyres indexes in each row and column. The final three rows of the table give the EKS, Geary and GAIA indexes of real expenditure, to be discussed in Section V.

The first major feature of the Laspeyres star matrix is that all the different indexes show less dispersion in real incomes than the data on nominal expenditure. Thus the coefficient of variation of nominal expenditure across the 16 countries is 98.2 per cent, whereas the coefficient of variation of the Laspeyres indexes ranges from 81.3 per cent to 85.8 per cent, with an average

Table 1: *Laspeyres, EKS, Geary and GAIA Indexes of Real Consumption Expenditure per capita, 1970*

	USA	*Germ.*	*Belg.*	*France*	*Neth.*	*UK*	*Italy*	*Japan*	*Hung.*	*Colum.*	*Korea*	*Iran*	*Malay.*	*Phil.*	*India*	*Kenya*	*Mean*	*SD*	*CV*
z ($):	3,362	1,777	1,753	1,709	1,654	1,511	1,203	1,099	286	258	215	207	201	140	76	5	966	949	98.2
z:	1,000	529	521	508	492	449	358	327	85	77	64	62	60	42	23	1	287	282	98.2
USA	1,000	536	598	560	577	512	431	450	140	137	101	133	102	90	44	2	338	283	83.5
Germany	1,000	530	598	550	578	520	435	438	139	145	111	130	106	93	49	2	339	280	82.7
Belgium	1,000	529	561	542	563	500	428	432	136	135	103	120	100	83	44	2	330	279	84.5
France	1,000	526	583	530	566	505	412	429	136	137	103	121	100	84	43	2	330	279	84.6
Netherlands	1,000	527	583	537	564	515	423	426	138	137	106	119	104	87	46	2	332	279	84.0
UK	1,000	538	582	560	580	508	428	448	142	135	96	122	99	84	41	2	335	283	84.4
Italy	1,000	523	598	537	574	517	418	428	137	144	108	125	103	90	47	2	334	280	83.6
Japan	1,000	534	601	557	585	534	446	417	131	140	109	138	109	99	50	2	341	282	82.7
Hungary	1,000	531	608	555	585	538	447	435	137	155	125	131	111	101	56	3	345	280	81.3
Colombia	1,000	544	584	568	572	517	443	438	134	133	99	136	104	93	45	2	338	282	83.5
Korea	1,000	553	612	597	610	551	465	445	133	138	98	149	107	103	46	2	351	289	82.4
Iran	1,000	530	569	553	575	528	448	426	137	146	116	117	107	92	51	2	337	279	82.3
Malaysia	1,000	514	584	541	567	510	416	427	135	142	104	121	99	87	44	2	331	279	84.3
Philippines	1,000	549	528	572	579	460	444	488	152	133	100	110	93	72	41	2	333	282	84.9
India	1,000	553	603	594	628	547	464	475	146	154	114	127	103	98	49	2	354	289	81.8
Kenya	1,000	524	565	546	565	517	420	408	125	129	91	123	97	85	39	2	327	281	85.8
Mean	1,000	534	585	556	579	517	435	438	137	140	105	126	103	90	46	2	337	281	83.5
SD	0	11	21	19	17	21	16	20	6	7	9	10	5	8	4	0	12*	7*	59.3*
CV (%)	0.0	2.1	3.6	3.5	3.0	4.1	3.8	4.6	4.5	5.2	8.2	7.6	4.6	8.9	9.2	7.8	5.4*	2.3*	43.0*
EKS	1,000	526	580	551	575	520	427	421	133	136	100	121	102	82	46	2	333	281	84.6
Geary	1,000	532	592	551	576	518	432	436	138	141	107	127	104	91	47	2	337	281	83.3
GAIA	1,000	532	580	545	578	519	428	428	126	138	97	112	104	84	31	9	332	283	85.3

SD: Standard Deviation; CV: Coefficient of Variation (%).
*Exceptionally, these entries are based on 15 countries only (i.e., excluding the US).

across countries of 83.5 per cent.[5] The different indexes also give broadly similar rankings across countries which may differ from the rankings based on nominal expenditure. For example, almost all the different indexes of real expenditure rank Germany below each of Belgium, France and The Netherlands, and rank Korea below Iran, in all cases reversing the corresponding rankings of nominal expenditure.

While there are major similarities between the indexes there are also important differences. These are highlighted by the summary statistics below the star matrix, in particular by the row giving the coefficient of variation of the indexes for each country. The average of these for the 15 countries excluding the US is 5.4 per cent, indicating a significant variation in the estimate of real income depending on the choice of weights. This effect is considerably greater for poorer countries: the correlation coefficient (over 15 countries) between the coefficient of variation and the level of nominal expenditure is –0.824.

We next wish to examine the magnitude of the Gerschenkron Effect in the sample. This means considering all possible bilateral comparisons. With m countries there are m(m–1)/2 such comparisons; 120 in our 1970 sample of 16 countries. For each of these comparisons we treat the ratio of the Laspeyres to the Paasche index, $p^k.q^j / p^k.q^k$ divided by $p^j.q^j / p^j.q^k$, as an independent observation and we examine the sample distribution of these observations. The results are given in the first row of Table 2. (We postpone consideration of the second row until Section IV.)

Table 2: *Tests of the Gerschenkron Effect*

Index		*Distribution of Laspeyres to Paasche Ratios*							
	Mean	*Std. Dev.*	*<0.9*	*0.9-1.0*	*1.0-1.1*	*1.1-1.2*	*1.2-1.3*	*>1.3*	*% > 1.0*
Fixed-Weight	1.114	0.118	1	7	59	33	12	8	93.3
Allen	1.528	1.953	25	50	30	1	0	14	37.5

Index	*Regression of LP Ratio on PD*		
	R^2	*a*	*b*
Fixed-Weight	0.139	1.027	0.047
		(0.022)	(0.011)
Allen	0.004	1.780	–0.136
		(0.397)	(0.191)

All comparisons are based on the 120 LP ratios; i.e., the ratios of bilateral Laspeyres to Paasche indexes for the 16 countries in 1970. PD: Measure of price dispersion, equal to the sum of squared deviations of prices between the two countries. Regression coefficient estimates have standard errors in parentheses.

5. These coefficients of variation are independent of the choice of base country. With the US as base, the lower dispersion shows up as a higher average real income with a near-identical standard deviation.

Considering first the properties of the distribution, the table shows that the average of the 120 ratios was 1.114, with a standard deviation of 0.118. Although this is not significantly greater than unity, the distribution is highly skewed, with 112, or 93.3 per cent, of the ratios exceeding unity. This suggests that the Laspeyres index is indeed systematically greater than the corresponding Paasche index. A more formal test of the Gerschenkron Effect requires that we investigate how their ratio varies with the difference in prices between the two countries. To do this we estimate a simple regression equation which regresses the Laspeyres to Paasche ratio on the sum of squared deviations between prices of individual goods in the two countries; i.e., a measure of the extent of price *dispersion* between the two countries. As the table shows, the Gerschenkron Effect is overwhelmingly confirmed. A unit increase in this measure of price dispersion is significantly associated with a rise in the Laspeyres to Paasche ratio of 0.047.

IV ALLEN INDEXES BASED ON ESTIMATED STONE-GEARY PREFERENCES

The previous section has confirmed that the choice of reference prices makes a significant difference to the estimated levels of real expenditure and that the differences involved reflect the Gerschenkron Effect. This suggests the desirability of estimating alternative indexes which allow for substitutability between commodities in response to price differences. Of course, there are many ways in which such substitutability might be parameterised. We have chosen to do so in a simple fashion, using the linear expenditure system, which corresponds to the Stone-Geary utility function. The demand function for commodity group i in this system (suppressing country subscripts for simplicity) is:

$$p_i q_i = p_i \gamma_i + \beta_i (z - \sum_h p_h \gamma_h) \qquad (13)$$

The γ_i coefficients are usually interpreted as subsistence parameters, while the marginal budget shares β_i (which must sum to unity) determine the allocation of "supernumerary" income $z - \Sigma_h p_h \gamma_h$ among commodity groups.

Calculating the Allen indexes is straightforward. The expenditure function corresponding to the Stone-Geary utility function, evaluated at the prices of country k and the utility level of country j is:

$$e(p^k, u^j) = \gamma^k + \beta^k u^j, \quad \text{where: } \gamma^k \equiv \sum_i \gamma_i p_{ik} \text{ and } \beta^k \equiv \prod_i \left(\frac{p_{ik}}{\beta_i} \right)^{\beta_i}. \qquad (14)$$

The γ^k and β^k terms can be calculated using the parameter estimates from Table A1 in Appendix 2, while the level of utility in each country is found by solving (14) for j=k (when the left-hand side equals z^j and so all the terms in the equation except u^j are observable). All 16-by-16 terms implied by (14) can then be calculated and the results are shown in the central matrix in Table 4. The entry in row k and column j of the matrix gives the real income of country j relative to the reference country (the US), evaluated at the prices of country k; that is, $e(p^k,u^j)/e(p^k,u^k)$ over $e(p^k,u^1)/e(p^k,u^k)$, or $e(p^k,u^j)/e(p^k,u^1)$. The special demand system used imposes considerable structure on the entries in the matrix.[6] From (14), the difference between any two entries j and h (not necessarily adjacent) in row k is:

$$\Delta^{jh} = \frac{u^j - u^h}{u^1} H^k, \quad \text{where:} \quad H^k \equiv \frac{u^1\beta^k}{\gamma^k + u^1\beta^k}. \tag{15}$$

In words, this is the utility difference between countries j and h (relative to the US) times H^k, which measures the ratio of supernumerary to total income for the US at country k prices. Thus, any entry in the table can be calculated from the values of u and H only (given in the first row and last column respectively).[7]

Since all the indexes are linear in utility, they are perfectly correlated across columns. Nevertheless, reading down columns there is considerable variation between the different indexes for each country. This is less true of the high-income countries in the sample, for which the results are relatively insensitive to the reference prices used. However, for the low-income countries the opposite is true. In fact, the average coefficient of variation (over 15 countries) for these indexes is 10.4 per cent, almost twice as much as for the Laspeyres indexes in Table 1.[8] Of course, some of this variation may reflect the failure of the linear expenditure system to capture the variations

6. Our discussion in the remainder of this section has benefited from a stimulating correspondence with Patrick Honohan.

7. Equation (14) also applies to any member of the Gorman Polar Form family of demand systems. This generalises the linear expenditure system to allow for arbitrary γ^k and β^k functions which are linearly homogeneous in prices. Hence the comments in this paragraph apply to any member of this family.

8. As with the Laspeyres indexes, this effect is greater for poorer countries though rather less so than in Table 1: the correlation coefficient (over 15 countries) between the coefficient of variation and the level of nominal expenditure is –0.532. These computations are not significantly affected if India and Kenya are omitted.

Table 3: *Allen Indexes of Real Consumption Expenditure per capita, 1970*

	USA	*Germ.*	*Belg.*	*France*	*Neth.*	*UK*	*Italy*	*Japan*	*Hung.*	*Colum.*	*Korea*	*Iran*	*Malay.*	*Phil.*	*India*	*Kenya*	*Mean*	*SD*	*CV*	*H*
Utility	1,000	490	543	505	541	476	378	377	49	63	17	34	25	3	−54	−78	273	308	112.7	
USA	1,000	529	578	543	576	516	426	425	121	134	92	108	100	79	26	4	329	284	86.5	0.924
Germany	1,000	534	582	547	580	521	431	430	130	142	101	116	108	88	36	14	335	281	84.0	0.915
Belgium	1,000	530	578	544	577	517	427	426	123	136	94	109	101	81	28	6	330	284	86.0	0.922
France	1,000	530	578	543	577	517	426	426	123	136	93	109	101	81	28	6	330	284	86.0	0.922
Netherlands	1,000	532	580	545	578	519	429	428	126	139	97	113	105	84	32	10	332	282	85.0	0.919
UK	1,000	529	577	542	575	515	425	424	120	133	90	106	98	78	25	2	327	285	86.9	0.925
Italy	1,000	533	581	547	579	520	430	430	129	142	100	115	107	87	34	12	334	282	84.3	0.916
Japan	1,000	536	584	549	582	523	434	433	134	147	105	121	113	93	40	18	338	280	82.8	0.910
Hungary	1,000	543	590	556	588	530	442	441	146	159	118	133	125	105	54	32	348	276	79.4	0.897
Colombia	1,000	530	578	543	577	517	426	426	123	135	93	109	101	81	28	6	330	284	86.1	0.922
Korea	1,000	536	584	550	582	524	434	434	134	147	106	121	113	93	41	19	339	280	82.7	0.910
Iran	1,000	539	586	552	585	526	437	437	139	152	110	126	118	98	46	24	342	278	81.4	0.905
Malaysia	1,000	532	580	545	578	518	428	428	125	138	96	112	104	84	31	9	332	283	85.2	0.919
Philippines	1,000	533	581	546	579	519	429	429	127	140	98	114	106	85	33	11	333	282	84.7	0.917
India	1,000	542	589	555	587	529	441	440	145	157	116	132	124	104	52	31	347	276	79.8	0.899
Kenya	1,000	528	577	542	575	515	424	424	119	132	90	106	98	77	24	2	327	285	87.1	0.925
Mean	1,000	534	581	547	580	520	431	430	129	142	100	116	108	87	35	13	335	282	84.2	0.915
SD	0	5	4	4	4	5	6	6	9	8	9	9	9	9	9	10	7*	2*	31.9	0.009
CV (%)	0.0	0.9	0.7	0.8	0.7	0.9	1.3	1.3	6.6	5.9	8.8	7.5	8.1	10.2	27.0	75.2	10.4*	19.2*	184.6	1.0

SD: Standard Deviation; CV: Coefficient of Variation (%).
*Exceptionally, these entries are based on 15 countries only (i.e., excluding the US).
H: Ratio of supernumerary to total income for US at country k prices.

in expenditure patterns between the countries in the sample.[9] Nevertheless it suggests that the choice of reference prices makes at least as much difference to the true indexes as it does to the fixed-weight Laspeyres indexes. Finally, applying to the Allen indexes the same tests for the Gerschenkron Effect as we applied to the fixed-weight indexes in Section III, we find that the presence of the effect is rejected. (See the second row of Table 2.) This shows that the sensitivity of the Allen indexes to reference prices is not systematically related to the degree of price dispersion between the countries compared.

V GROPING TOWARDS GAIA

The last set of indexes we calculate are those corresponding to the GAIA system. To do this we implement the iteration procedure proposed in Neary (1996b). This is a tâtonnement-type algorithm, which takes as starting point the standard Geary world prices and true exchange rates and uses them to calculate initial estimates of the virtual quantities defined in (10). The Geary method is then applied to the estimated virtual quantities and the process is continued until it converges. In practice, it proved possible to implement this procedure fairly straightforwardly and convergence was reasonably rapid: very much so for real incomes, rather less so for exchange rates and slowest of all for world prices.[10] The results are shown in the last row of Figure 1.

Table 4 summarises the relationships between the principal indexes we have calculated (including the EKS and Geary indexes) as well as some of the statistical properties of each. As can be seen the correlations between all the indexes are extremely high, in both levels and first differences. (The average of the Allen indexes is perfectly correlated with the GAIA index, reflecting the fact that, with Stone-Geary preferences, both these indexes are linear functions of the data.) The nominal expenditure index is the least correlated with the other indexes in levels while the GAIA index is slightly more

9. The two negative utility levels for India and Kenya pose special difficulties of interpretation. Since utility is ordinal rather than cardinal, negative values do not matter in principle, but with Stone-Geary preferences they imply that the country in question has total expenditure less than the "subsistence" level (i.e., $z^j < \Sigma_h p_{hj} \gamma_{hj}$). This in turn implies that the substitution matrix is not negative semi-definite when evaluated at those countries' consumption bundles. Such a violation of concavity means that the hypothesis of utility maximisation is not valid at these points. See Irvine and McCarthy (1980). This finding is common with the linear expenditure system. Brown and Deaton in their 1972 survey (Section IV.3) note that even in time-series studies supernumerary income is often negative for the first few observations.

10. Copies of the GAUSS program which calculates the GAIA system are available on request from the authors. Implementation on a Pentium 90 notebook with 16MB RAM running under Windows 95 took 2.69 seconds of calculation time. The convergence criterion selected was that the sum of squared deviations between the estimated Π vectors in successive iterations be less than 10^{-14} and this was satisfied after 316 iterations. Further details are given in Neary and Gleeson (1997).

Table 4: *Correlations Between Different Indexes*

Index:	*Expenditure*	*Av. Laspeyres*	*Av. Allen*
Mean:	287.3	337.2	334.5
Standard Deviation:	282.2	281.5	281.6
Coeff. of Var (%):	98.2	83.5	84.2
Correlation Coefficients:			
Expenditure	1.0000000000	0.9962482043	0.9965736256
Average Laspeyres	0.9986399639	1.0000000000	0.9998062474
Average Allen	0.7700815976	0.7406721740	1.0000000000
EKS	0.9984205271	0.9999458452	0.7343514622
Geary	0.9985890156	0.9999982507	0.7397424918
GAIA	0.8468864765	0.8247088435	0.9924840417
Index:	*EKS*	*Geary*	*GAIA*
Mean:	332.6	337.2	331.5
Standard Deviation:	281.4	280.9	282.9
Coeff. of Var (%):	84.6	83.3	85.3
Correlation Coefficients:			
Expenditure	0.9966830510	0.9962349641	0.9965736256
Average Laspeyres	0.9999030691	0.9999612073	0.9998062474
Average Allen	0.9998299199	0.9997905392	1.0000000000
EKS	1.0000000000	0.9998947261	0.9998299199
Geary	0.9999555033	1.0000000000	0.9997905392
GAIA	0.8195704744	0.8239766478	1.0000000000

Correlation coefficients above the diagonal relate to the levels of the indexes.
Those below the diagonal relate to the percentage first differences of the indexes.

correlated with the EKS index than with the Geary index in levels and conversely in first differences. Of course, the high correlations between the different indexes does not mean that they have identical implications for the levels of real consumption expenditure in different countries. Note finally that the GAIA index is the most dispersed of all the indexes, except for that based on nominal expenditure.

VI CONCLUSIONS

This paper has used a specially constructed data set drawn from the International Comparisons Project which underlies the Penn World Table to investigate a number of related issues which arise in making comparisons of real incomes and expenditures between countries.

The first issue we addressed was the empirical importance of the Gerschenkron Effect, which postulates that a country's relative real income is higher the more the reference prices used to make the comparison deviate from its own prices. We noted that, on theoretical grounds, this phenomenon may be expected to occur to the extent that utility-maximising consumers substitute away from commodities which are relatively more expensive. However, when tastes are not homothetic and income levels are sufficiently different, it may not occur even if we had observations on the behaviour of individual utility-maximisers.

Given these theoretical ambiguities concerning the likelihood of the Gerschenkron Effect, we proposed and implemented two tests of its severity, based on the distribution of the ratios of the Laspeyres to Paasche indexes for all possible bilateral comparisons in the sample. Our results confirmed that the Gerschenkron Effect was present in the sample. 93.3 per cent of the ratios exceeded unity and they were significantly and positively related to the degree of price dispersion between the two countries compared. The same tests were applied to the Allen indexes, calculated using the estimated parameters of the Stone-Geary utility function. By contrast with the results for the fixed-weight indexes, there was no significant relationship between a country's Allen real income index and the deviation between its own prices and the reference prices used. This suggests that taking account of inter-commodity substitution substantially eliminates the Gerschenkron Effect.

While the Allen indexes eliminate the Gerschenkron Effect, they do not reduce the dependence of real income measures on reference prices. On the contrary, a comparison of Tables 1 and 4 shows that the sensitivity of real incomes to reference prices is considerably greater for the Allen than for the Laspeyres indexes, with the increased sensitivity especially pronounced for low-income countries. This strengthens the case for a system such as the GAIA which provides an explicit rationale for the choice of reference prices.

Finally, the actual estimation of the GAIA system was shown to be feasible, with convergence occurring rapidly. The resulting indexes are more dispersed than any of the other real income indexes, though they are slightly more strongly correlated with the EKS than with the Geary indexes.

Naturally, this paper has raised more questions than it has answered. It would be especially desirable to extend the methods of calculating real expenditures to other categories of aggregate demand, with a view to calculating true indexes of national output. As for the true indexes which we have estimated, they are clearly conditional on the hypothesis that demands in all the countries of the sample can be represented by the linear expenditure system. Since the Stone-Geary utility function which underlies this exhibits additive separability, it is open to the criticism of Deaton (1974) that it imposes on the results an inverse relationship between income and price elasticities. (Deaton calls this "Pigou's Law".) This may explain why the EKS system does not fare badly relative to the GAIA system: deviations from homotheticity work against the EKS system but, because of Pigou's Law, they also tend to impose greater price substitutability, which also works against the Geary system. These considerations suggest that it would be very desirable to repeat the calculations of this paper using demand systems which do not impose additive separability.

REFERENCES

ALLEN, R.G.D., 1949. "The Economic Theory of Index Numbers", *Economica*, Vol. 16, pp. 197-203.

AFRIAT, S.N., 1977. *The Price Index*, Cambridge: Cambridge University Press.

BROWN, J.A.C., and A.S. DEATON, 1972. "Models of Consumer Behaviour: A Survey", *Economic Journal*, Vol. 82, pp. 1145-1236.

DEATON, A.S., 1974. "A Reconsideration of the Empirical Implications of Additive Preferences", *Economic Journal*, Vol. 84, pp. 338-348.

DIEWERT, W.E., 1976. "Exact and Superlative Index Numbers", *Journal of Econometrics*, Vol. 4, pp. 115-145.

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.

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.

DRESCHLER, L., 1973. "Weighting of Index Numbers in Multilateral International Comparisons", *Review of Income and Wealth*, Vol. 19, pp. 17-34.

ELTETÖ, O., and P. KÖVES, 1964. "On a Problem of Index Number Computation Relating to International Comparisons" (in Hungarian), *Statisztikai Szemle*, Vol. 42, pp. 507-518.

FRISCH, R., 1936. "Annual Survey of General Economic Theory: The Problem of Index Numbers", *Econometrica*, Vol. 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)*, Vol. 121, pp. 97-99.

GERSCHENKRON, A., 1951. *A Dollar Index of Soviet Machinery Output*, Santa Monica: The Rand Corporation.

GORMAN, W.M., 1953. "Community Preference Fields", *Econometrica*, Vol. 21, pp. 63-80.

IRVINE, I., and C. McCARTHY, 1980. "Further Evidence on Inflation and Redistribution in the United Kingdom", *Economic Journal*, Vol. 90, pp. 907-911.

KRAVIS, I.B., 1984. "Comparative Studies of National Incomes and Prices", *Journal of Economic Literature*, Vol. 22, pp. 1-39.

MUELLBAUER, J., 1975. "Aggregation, Income Distribution and Consumer Demand", *Review of Economic Studies*, Vol. 62, pp. 525-543.

NEARY, J.P., 1996a. "Theoretical Foundations of the 'Geary Method' for International Comparisons of Purchasing Power and Real Incomes", *The Economic and Social Review*, Vol. 27, pp. 161-179.

NEARY, J.P., 1996b. "True Multilateral Indexes for International Comparisons of

Purchasing Power and Real Income", Centre for Economic Research, University College Dublin, Working Paper No. 96/22.

NEARY, J.P., 1997. "R.C. Geary's Contributions to Economic Theory", in D. Conniffe (ed.), *Roy Geary, 1896-1983: Irish Statistician. The Centenary Lecture by John Spencer and Associated Papers*, Dublin: Oak Tree Press, pp. 93-118.

NEARY, J.P., and B. GLEESON, 1997. "Comparing the Wealth of Nations: Reference Prices and True Multilateral Real Income Indexes", Centre for Economic Research, University College Dublin, Working Paper No. 97/19.

NUXOLL, D.A., 1994. "Differences in Relative Prices and International Differences in Growth Rates", *American Economic Review*, No. 84, pp. 1423-1436.

POLLAK, R.A., 1971. "The Theory of the Cost-of-Living Index", US Bureau of Labor Statistics Working Paper 11; revised version in W.E. Diewert (ed.), 1990, *Price Level Measurement*, Amsterdam: North-Holland, pp. 5-77.

SAMUELSON, P.A., and S. SWAMY, 1974. "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis", *American Economic Review*, Vol. 64, pp. 566-593.

SUMMERS, R., and A. HESTON, 1991. "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988", *Quarterly Journal of Economics*, Vol. 106, pp. 327-336.

SZULC, B.J., 1964. "Index Numbers for Multilateral Regional Comparisons" (in Polish), *Przeglad Statystyczny*, Vol. 3, pp. 239-254.

APPENDIX 1

The Data

The raw data on 16 countries are taken from Phase II (1970) of the United Nations International Comparisons Project (ICP), and were retrieved, via the Internet, from files at the University of Pennsylvania. Expenditure data $z_{ij} = p_{ij} .q_{ij}$ were given as percentages of GDP while the price data were given relative to the US, p_{ij} /p_{i1}. The data on personal consumption expenditure were broken down into 110 categories, so these had to be aggregated to obtain the 11 categories used in the present project: food; beverages; tobacco; clothing and footwear; gross rents; fuel and power; house furnishings, appliances and operations; medical care; transport and communication; recreation and education; and miscellaneous goods and services. To get expenditure on food, for example, the expenditure shares of all items in the food category were summed and then grossed up by GDP, taken from the IMF statistical yearbooks. The aggregate price level for food was calculated as the weighted sum of all food item prices where the weight was the share of each food item in total food expenditure. Finally, dividing each expenditure entry, $p_{ij}.q_{ij}$, by the corresponding relative price, p_{ij}/p_{i1}, gives quantities in country j measured in US prices, $p_{i1}.q_{ij}$. The price and quantity data are given in Neary and Gleeson (1997).

Table A1: *Estimated Parameters for the Stone-Geary Utility Function, 1970*

Commodity Group	*Average Budget Shares*	*β's: Marginal Budget Shares*		*γ's: Subsistence Parameters*		R^2
		Coeff.	*t-value*	*Coeff.*	*t-value*	
Food	0.219	0.140341	9.80	123.783	5.99	0.896
Beverages	0.033	0.027571	6.45	12.366	3.17	0.637
Tobacco	0.022	0.015973	6.84	13.261	6.49	0.782
Clothing and Footwear	0.079	0.079609	22.59	26.514	5.41	0.974
Gross Rents	0.106	0.130337	17.95	8.939	0.89	0.946
Fuel and Power	0.027	0.030799	37.74	2.398	3.81	0.985
Household	0.092	0.092661	10.92	22.007	1.93	0.894
Medical Care	0.077	0.092666	22.88	2.903	0.88	0.963
Transport and Communications	0.102	0.127407	16.12	3.752	0.50	0.939
Recreation and Education	0.128	0.138453	27.96	26.602	4.01	0.985
Miscellaneous	0.108	0.124183	*	14.622	1.01	*

*β for the Miscellaneous group was estimated residually, so these entries are not relevant.

APPENDIX 2

The Demand System Parameter Estimates

Table A1 gives the parameters of the linear expenditure system estimated from the 1970 data for 16 countries. In this pilot study we did not experiment with alternative specifications. The estimation method is Seemingly Unrelated Regressions, with no correction for heteroscedasticity. The statistical performance of the equations is satisfactory and the results are plausible in economic terms. Comparing average and marginal budget shares, we see that food, beverages and tobacco decline in importance as total expenditure rises, whereas gross rents; fuel and power; medical care; transport and communications; recreation and education and miscellaneous items are luxuries.