Title	Theoretical foundations of the "Geary method" for international comparisons of purchasing power and real incomes
Authors(s)	Neary, J. Peter
Publication date	1996-01-18
Publication information	Neary, J. Peter. “Theoretical Foundations of the ‘Geary Method’ for International Comparisons of Purchasing Power and Real Incomes.” University College Dublin. School of Economics, January 18, 1996.
Series	UCD Centre for Economic Research Working Paper Series, WP96/03
Publisher	University College Dublin. School of Economics
Item record/more information	http://hdl.handle.net/10197/3047

THEORETICAL FOUNDATIONS OF THE "GEARY METHOD" FOR INTERNATIONAL COMPARISONS OF PURCHASING POWER AND REAL INCOMES*

J. Peter Neary
University College Dublin

18 January 1996

* I am very grateful to Gerry Boyle for inviting me to contribute on this topic to the centenary symposium in honour of Roy Geary; to Patrick Honohan and Robert Summers for detailed suggestions; and to Brid Gleeson, Will Martin, Brendan Walsh and participants at a meeting of the Dublin Economics Workshop for helpful 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 [illegible]me of the Centre for Economic Performance at the London School of Economics.

Kevin –

for CEP

Pete

J. Peter Neary
Department of Economics
University College Dublin
Belfield, Dublin 4, Ireland
Tel.: +353-1-706 8334
FAX: +353-1-283 0068
nail: JPNEARY@OLLAMH.UCD.IE

THEORETICAL FOUNDATIONS OF THE "GEARY METHOD" FOR INTERNATIONAL COMPARISONS OF PURCHASING POWER AND REAL INCOMES

ABSTRACT

This paper provides a centenary review of the method of calculating real incomes and purchasing power parities proposed by Roy Geary. This method is the most widely used in major international comparisons, but it is often criticised for its lack of theoretical foundations. I discuss the properties of the method and its competitors in the light of both practical and theoretical considerations. I also propose a new method of computing "true" or, as I call them, "Geary-Konüs" exchange rates and world prices and I argue that the Geary method provides the best available approximation to the true values.

Keywords: International Comparisons of Real Incomes; Exchange Rates; Purchasing Power Parities; Price Indexes; R.C. Geary
JEL Classification Nos.: D1, C8, F0

TABLE OF CONTENTS

1. **Introduction** 1
2. **The Geary Method and Alternatives** 2
3. **Review of Index Number Theory** 7
4. **True Indexes of Exchange Rates and Real Incomes** 13
5. **Conclusions** 16

References 17

THEORETICAL FOUNDATIONS OF THE "GEARY METHOD" FOR INTERNATIONAL COMPARISONS OF PURCHASING POWER AND REAL INCOMES

1. Introduction

The late Roy Geary, whose legacy to statistics ranks with those of Edgeworth and Gossen, made at least two major contributions to economic theory. Probably best known is his derivation (in Geary, 1950) of the utility function underlying the linear expenditure system, which is still known as the Stone-Geary utility function. But almost equally widely used is his method of computing exchange rates for international comparisons of real income and purchasing power. This method was originally developed for the U.N.'s Food and Agriculture Organisation and described by Geary in a three-page note (Geary, 1958). It came into wider use when it was adopted as the basic methodology of the U.N.'s International Comparison Project (ICP), which produced cross-section comparisons of real incomes for groups of countries varying in number between 16 and 60 at five-yearly intervals from 1970 to 1985. These comparisons in turn form the basis of the "Penn World Tables".[1] This vast data set, which in its 1991 version gives detailed comparable national income data on 139 countries over 39 years, has made possible the recent explosion of interest in comparative studies of economic growth (see Mankiw et al., 1992 and Quah, 1993).

The basic problem which the Geary method addresses is how to compare real incomes across countries. Such comparisons can be made straightforwardly, of course, using official exchange rates. But these often change dramatically over short periods, without any compensating changes in prices. Even over long periods, prices of many identical commodities measured in a common currency differ persistently, reflecting barriers to arbitrage such as transport costs and imperfect competition.[2] The Geary method does not address the question of *why* prices of identical goods and services differ between countries. Instead, it takes these differences as given and proposes a straightforward procedure for calculating "true" exchange rates which can then be used to calculate the "real" incomes of each country.

However, while the Geary method has had enormous practical success, it has not met with as much approval from economic theorists. Two leading theorists in particular, Paul Samuelson and Erwin Diewert, have frequently criticised its lack of theoretical foundations.[3] Moreover, Diewert has persistently advocated the use of alternative methods, of which the best-known is the so-called EKS method, which has been used by the OECD and by the Statistical Office of the European Union.[4] It would appear that, to quote Harry Johnson's famous comment on Keynes's economics, the Geary method has at best "won the policy battle but lost the theoretical war"!

The purpose of this paper is to reevaluate the Geary method and to suggest that its theoretical foundations are not as weak as have been thought. I begin by reviewing the method and its competitors in Section 2 and then summarise some relevant results from the theory of index numbers in Section 3. Section 4 then outlines the theoretical case for the Geary method. Technical details are relegated to a companion paper, Neary (1995).

2. The Geary Method and Alternatives

The problem addressed by Geary is that of comparing prices and standards of

[1] Kravis (1984) provides an introduction; Kravis et al. (1975, 1978, 1982) are the basic references; Marris (1984) gives a critical though broadly sympathetic review; Summers and Heston (1991) extend the results to 1988; and more recent updates are available free on the Internet via anonymous ftp from nber.harvard.edu.

[2] Such persistent differences are often described as failures of the theory of "purchasing power parity", but they are more appropriately described as rejections of the "law of one price". See Purvis (1982) and Neary (1988) for overviews and Kravis et al. (1975) and Isard (1977) for empirical evidence.

[3] See, for example, the dismissive remarks by Samuelson and Swamy (1974, p. 591), Caves, Christensen and Diewert (1982, p. 83), Samuelson (1984, p. 277 and 1994, p. 212), Diewert (1987, pp. 776-8) and Officer (1989). They and others often refer to the method as the "Geary-Khamis" method. However, this may be over-generous to Khamis (1970, 1972), who derived conditions for the Geary method to give a unique, positive solution and advocated the use of index numbers based on the Geary method. See my comments in Section 3 below.

[4] The EKS method is named after its originators Eltetö and Köves (1964) and Szulc (1964). Diewert (1993, p. 57) notes that the EKS method was proposed by Gini in 1931, reflecting the frequent phenomenon that index numbers are not always known by the names of their originators. The same article by Diewert shows that the Laspeyres index was anticipated by Bishop Fleetwood of Ely in 1707 and by Joseph Lowe in 1823 (pp. 34-35); that the Paasche index was proposed by Willard Phillips in 1828 (p. 35); and that the Fisher Ideal index was first developed by Bowley and Pigou (p. 36).

living between a group of countries.[5] 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.[6] Each commodity is assumed to be identical in quality worldwide but because of transport costs and other trade barriers 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.

In seeking to make such comparisons, it is tempting to postulate the existence of "true" exchange rates which give for each country the "purchasing power" of its currency relative to some external standard. The problem is that, in order to calculate such true exchange rates, expenditure on each commodity should be measured in terms of common "world" prices; but such world prices cannot be calculated without a knowledge of the true exchange rates. Geary's resolution of this circular dilemma was to propose an algorithm whereby *both* prices and exchange rates are computed simultaneously.

Let π, with typical element π_i, represent the n-by-one vector of world prices, expressed in units of a hypothetical international unit of currency; and let ϵ, with typical element ϵ_j, represent the m-by-one vector of exchange rates, each giving the cost in international currency units of one unit of domestic currency.[7] The Geary true exchange rates are now defined as follows:

[5] Writing in the 1950s, Geary motivated this problem by referring to the establishment of EFTA and the Common Market, and (perhaps prophetically!) labelled his international currency unit the "europa".

[6] In practice, the raw data used in the ICP were not p_{ij} and q_{ij} but expenditures in domestic currency, $p_{ij}q_{ij}$, and prices relative to the U.S., $p_{ij}/p_{i,US}$. Of course, assembling such a comparable data set is a mammoth task in itself and poses enormous conceptual and technical problems, which are discussed at length in the ICP volumes. While their approach to some of these problems has attracted criticism (Isenman, 1980, for example, criticises their treatment of services), none of them has generated as much controversy as the choice of algorithm for calculating exchange rates.

[7] I follow Geary here in using the "U.K." convention to define exchange rates. Kravis and his co-workers use instead the "U.S." convention, defining the exchange rate as the price in domestic currency of a unit of international currency. Hence, their "purchasing power parities" are the inverse of my true exchange rates: see Kravis (1984), p. 11, equation (3).

$$\epsilon_j = \frac{\sum_i \pi_i q_{ij}}{\sum_i p_{ij} q_{ij}}, \qquad j = 1, \ldots m. \tag{1}$$

Multiplying across by the right-hand-side denominator shows that total expenditure in country j is the same, whether valued at world prices π or at domestic prices p^j converted to international currency units by the exchange rate ϵ_j. The world prices themselves are defined in turn as:

$$\pi_i = \frac{\sum_j \epsilon_j p_{ij} q_{ij}}{\sum_j q_{ij}}, \qquad i = 1, \ldots n. \tag{2}$$

In this case, multiplying across by the right-hand-side denominator shows that total world expenditure on commodity i is the same, whether valued at its world price π_i or at domestic prices p_{ij} converted to international currency units by the exchange rates ϵ_j.

What justification can be provided for these definitions? Geary's own was neither lengthy nor modest. He merely commented: "The special feature which informs equations (1) and (2) is that, if the entities π_i and ϵ_j exist, they could scarcely be defined reasonably in any other terms." As we shall see, not all writers have agreed with this verdict. Nevertheless, the Geary method has many virtues, of which a key one is a property which Kravis (1984) has called "matrix consistency". A reflection of this is that, if we *either* sum (1) over all m countries *or* sum (2) over all n commodities, we obtain the same aggregate condition:

$$\sum_i \pi_i \left(\sum_j q_{ij}\right) = \sum_j \epsilon_j \left(\sum_i p_{ij} q_{ij}\right). \tag{3}$$

This shows that total world expenditure is the same whether valued at world prices π_i or at domestic prices p_{ij} converted to international currency units by the exchange rates ϵ_j. Because of matrix consistency, complete systems of consumer expenditure (or, more generally, of national accounts) measured in Geary world prices may be consistently disaggregated by country or by spending category.

The practical calculation of ϵ and π is relatively straightforward. The systems (1)

and (2) together comprise $m+n$ equations in $m+n$ unknowns. However, the dimensionality of the problem can be reduced in two steps. Firstly, since the equations are mutually dependent from (3), we can drop one equation and normalise one unknown by setting it equal to unity. In practice, it makes sense to choose one country as reference, setting its exchange rate at unity and dropping the corresponding equation from (1). In the ICP, the U.S. was chosen as reference country, but this normalisation is purely arbitrary and does not affect the outcome in any substantive way.

The second step in reducing the dimensionality of the problem is to combine (1) and (2) and solve first for either ϵ or π, depending on whether m or n is smaller. Since in typical applications m is much less than n,[8] it is easiest to eliminate π and solve first for the exchange rates ϵ. The world prices may then be calculated by substituting into (2). Finally, in most applications we are not primarily interested in the prices or exchange rates themselves, but in the implied measures of real income, which in this static framework is the same as real expenditure. These are obtained by multiplying each country's nominal expenditure by its exchange rate. From (1), this gives income at world prices for country j as:

$$Z_j^* = \sum_i \pi_i q_{ij} = \pi.q^j, \qquad j=1, \dots m, \tag{4}$$

where a dot denotes a vector product.

So much for the Geary method. What alternatives are available for making international comparisons of real income? A much simpler approach would be to take one country as reference and to compute the real income of every other country in terms of the reference country's prices. This is an example of a "star" method, since the reference country can be thought of as the centre of a star, which has as many points as there are comparison countries. It has an advantage over the Geary method in that it maximises what Dreschler (1973) calls the *characteristicity* of bilateral comparisons: a comparison between any country and the reference country involves the prices and

[8] For example, in Phase IV of the ICP, the Geary exchange rates and world prices for 1980 were calculated from a data set covering 60 countries and 150 commodity categories. In practice, such large systems are solved by iterative methods rather than by inverting a submatrix of the partitioned full matrix.

quantities of those two countries only. Characteristicity is often taken for granted in time-series applications: it is often assumed that comparisons between two successive years should ideally depend on data pertaining to those two years alone. However, in cross-section applications, the price that must be paid for characteristicity is high: the star method does not exhibit *base-country invariance*, meaning that the choice of reference country can make a great difference to the results.

Though the star method is unsatisfactory, its starting point in bilateral comparisons is the origin of the EKS method. Note first that in principle a single bilateral comparison could be carried out using the prices of *either* country: this is the cross-section analogue of the choice which arises in time-series applications between Laspeyres and Paasche indexes. A widely-used compromise between these two is the so-called "Fisher Ideal" index, which equals the geometric mean of the Laspeyres and Paasche indexes. Thus, the Fisher Ideal index giving the real income of country j relative to that of country k is (expressed in logarithms):

$$\ln Q_F^{jk} = \frac{1}{2}\left\{\ln \frac{p^j.q^j}{p^j.q^k} + \ln \frac{p^k.q^j}{p^k.q^k}\right\}. \tag{5}$$

The EKS method replicates this bilateral comparison over all possible pairs. Thus, the EKS index giving the real income of country j relative to that of country k equals the geometric mean of the ratios of Fisher Ideal indexes arising from all m bilateral comparisons between the two countries using each of the m countries in turn as base:

$$\ln Q_{EKS}^{jk} = \frac{1}{m}\sum_{l=1}^{m}\left\{\ln Q_F^{jl} - \ln Q_F^{kl}\right\}. \tag{6}$$

Since the Fisher index is reflexive ($Q_F^{jj}=1$) and symmetric ($Q_F^{jk}.Q_F^{kj}=1$), this formula may be written in an alternative way:

$$\ln Q_{EKS}^{jk} = \frac{1}{m}\left[2\ln Q_F^{jk} + \sum_{l=1, l\neq j,k}^{m}\left\{\ln Q_F^{jl} - \ln Q_F^{kl}\right\}\right]. \tag{7}$$

When m equals two the term preceded by a summation sign disappears and the EKS index reduces to the Fisher index.

What are the advantages of the EKS index? Firstly, since it treats all countries symmetrically, it satisfies base-country invariance. Secondly, it exhibits *transitivity*,

meaning that it leads to a unique cardinal ranking of the real incomes of all m countries: $Q^{jk}_{EKS}=Q^{jl}_{EKS}.Q^{lk}_{EKS}$. In both these respects, the EKS index is clearly superior to both the star system and the Fisher Ideal index. (For example, it is *not* true that $Q^{jk}_F=Q^{jl}_F.Q^{lk}_F$). Finally, equation (7) shows that the EKS index exhibits a high degree of characteristicity, albeit less than either of the bilateral indexes. This follows by construction, since it can be shown that the EKS index 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.)

It is also instructive to compare the EKS index with the Geary method (which also exhibits transitivity and base-country invariance, though less characteristicity). Unlike the Geary method, the EKS approach calculates real incomes directly. If desired, true exchange rates may then be inferred by normalising the nominal exchange rate of an arbitrary country at unity (exactly as in the Geary method). However, it does not seem to be possible to derive estimates of "world" prices from the EKS method.[9] As a result, the EKS method does not exhibit matrix consistency: it is not possible to disaggregate consistently the calculated real incomes across commodities. It was largely for this reason that the Geary method was chosen for use in the ICP. However, the choice should also be guided by considerations from economic theory. In order to examine those, we must first review briefly the economic theory of index numbers. While this topic is usually presented in terms of intertemporal comparisons, identical considerations arise in international comparisons.

3. Review of Index Number Theory[10]

In practice, most empirical index numbers are based on weighted averages of relative prices or quantities between two periods or regions. The Laspeyres, Paasche and

[9] Other methods have been suggested which give almost identical results to the EKS method but which ensure matrix consistency. One such example, due to Gerardi, is to define world prices as unweighted geometric means of the national prices. However, like the EKS method, this suffers from a serious deficiency: the estimated world prices are not invariant to the ways in which the boundaries which make up the group of countries are drawn. See Hill (1984) for discussion and references.

[10] For more details on this topic, see Pollak (1971) and Diewert (1981 and 1987).

Fisher Ideal formulae, which have already been mentioned, are the best-known of these, but there are many alternatives. Indeed, Geary's algorithm for computing exchange rates can be interpreted as implying yet another index number formula. As Geary himself noted, with only two countries ($m=2$), the world prices may be eliminated from (1) and (2) to give an expression for the ratio of the two exchange rates as:

$$\frac{e_k}{e_j} = \frac{\sum_i p_{ij}\,\omega_i}{\sum_i p_{ik}\,\omega_i}, \qquad \text{where: } \omega_i = \frac{2q_{ij}q_{ik}}{q_{ij}+q_{ik}}. \tag{8}$$

This is a price index with fixed weights equal to the harmonic means of the quantities, and its use has been advocated strongly by Khamis (1972).[11] It has the same form as the Laspeyres and Paasche price indexes, whose weights ω_i equal q_{ik} and q_{ij}, respectively; and as the price indexes of Edgeworth and C.M. Walsh, whose weights equal the arithmetic and geometric means of the quantities, respectively. (See Diewert (1993, p. 36).)

However, the difficulty with all fixed-weight indexes has been bluntly expressed in the remark of Afriat (1977) that they provide "answers without questions". For all their "commonsensical" appeal, unless we specify an underlying theoretical framework, there is no well-defined concept corresponding to the "real income" or "average price level" which fixed-weight indexes purport to measure.[12] In the absence of such concepts, one approach to choosing between different indexes is the "test" or axiomatic approach of Irving Fisher (1922). This was illustrated in the last section, where we discussed a number of plausible tests or criteria which satisfactory multilateral real-income indexes

[11] Khamis (1972, p. 105) notes that this Geary price index, like "the so-called 'true' index" can lie outside the Laspeyres and Paasche bounds. For this reason he claims that it is superior to the Edgeworth and Fisher Ideal indexes, which must lie between these bounds.

[12] Here, as elsewhere, practitioners doing what seems to them "obvious" are, in Keynes's memorable phrase, "merely the slaves of some defunct economist." For example, prior to Keynes (1930, chap. 6) a common approach to calculating price indexes was to base them on properties of the distribution of the *unweighted* price relatives p_{ij}/p_{ik}. This approach seemed "commonsensical" to its originators, Jevons and Edgeworth, since according to the quantity theory of money all prices should move proportionately except for random fluctuations.

might satisfy (such as transitivity and base-country invariance). However, the test approach was fatally undermined by Frisch (1936), among others, who showed that many *a priori* reasonable properties of index numbers are mutually inconsistent.[13] For this reason, the attention of economists has switched to the so-called "economic" approach to index numbers. This seeks first to identify "true" indexes which provide answers to well-defined questions assuming optimising behaviour on the part of economic agents; and it then investigates which empirical indexes come closest to the unobservable true indexes.

The economic approach can be illustrated in the context of international comparisons of real incomes using Figure 1 (based on Samuelson, 1974) and Table 1. Assume that we know the prices and consumption levels of two goods in each of two countries, j and k. (I follow the ICP convention of taking the higher-income country, in this case country k, as the base.) Each country is inhabited by a single utility-maximising individual and tastes are identical internationally. With consumption levels given by points J and K respectively, we wish to establish how poor is country j relative to country k.

Consider first the empirically-based fixed-weight indexes which value the consumption bundles of the two countries at a common reference price vector p^r. (The measure of real incomes implied by the Geary method, given by equation (4), is an example of this index, where the reference price vector is the world price vector π.) The value of each index can be read in Figure 1 by projecting the relevant lines onto the (arbitrary) ray from the origin OQ. Thus the Laspeyres index Q_L^{jk} (using country k's prices) evaluates country j's real income relative to k's as OA/OB; while the Paasche index Q_P^{jk} (using country j's prices) evaluates it at OC/OD. The latter is smaller, reflecting a common (though not inevitable) phenomenon known as the "Gerschenkron (1951) effect": because relative prices are inversely related to relative quantities consumed, a country's measured real income is lower the closer are the reference prices to its own prices.

The first economic index of real income is due to Allen (1949) and equals the ratio

[13] Of course, the test approach remains useful and even in its extreme form (which ignores the implications of optimising behaviour) it has its modern adherents. For example, Diewert (1988) and Eichhorn and Voeller (1990) apply the approach to the choice between indexes for international comparisons, though neither of these papers mentions the Geary method.

of expenditures evaluated at the utility levels (*not* the consumption bundles) of the two countries and at a common reference price vector p^r: $Q_{Ar}^{jk} = e(p^r,u^j)/e(p^r,u^k)$. (Here $e(p,u)$ is the expenditure function, giving the minimum cost of attaining the utility level u facing prices p.) The Allen index thus answers the question "what fraction of the cost of attaining country k's utility level would be needed to attain country j's utility level at the reference prices?" Using k's prices as reference involves locating country j at the hypothetical consumption point H and gives a value for the Laspeyres-Allen index (i.e., the Allen index using base-country weights) of OE/OB. This is less than the fixed-weight Laspeyres index OA/OB, reflecting the substitution possibilities available to the consumer in country j when faced with country k's prices. Similarly, the Paasche-Allen index (which evaluates country k at the hypothetical point G) equals OC/OF, which is greater than the fixed-weight Paasche index OC/OD. Thus, we can see that the Gerschenkron effect is offset or even eliminated when the possibility of substitution is recognised. The Gerschenkron effect is an empirical artefact reflecting in cross-section comparisons the same phenomenon that in time-series applications causes the fixed-weight Laspeyres index to "overestimate" and the fixed-weight Paasche index to "underestimate" the "true" (i.e., utility-based) rate of inflation.

The Allen index is intuitively appealing. However, it is very desirable that the product of a price and a real-income index should equal the ratio of actual expenditures between the two countries. (Fisher called this requirement the *weak factor reversal test*.) The Allen index does not generally satisfy this test if the true price or cost-of-living index is (as is usually assumed) that of Konüs (1924): $e(p^j,u^r)/e(p^k,u^r)$, which answers the question "what fraction of the cost of attaining the reference utility level at k's prices is needed to attain it at j's prices?" Hence, an alternative real-income index may be preferred, which satisfies the weak factor reversal test by construction. This equals the expenditure ratio OC/OB divided by the Konüs price index and may be called the Konüs real-income index. As shown in Table 1, the Laspeyres-Konüs real-income index equals the Paasche-Allen index and the Paasche-Konüs real-income index equals the Laspeyres-Allen index.

Finally, a defect of both the Allen and Konüs real-income indexes is that they are not in general linearly homogeneous in prices: if the quantities of all goods consumed in country j are exactly λ times those consumed in country k, neither index need equal λ, as

intuitively it should. A real-income index which satisfies this desirable property is that of Malmquist (1953). (See Deaton (1979) for an exposition.) This index is not defined in terms of the expenditure function but in terms of an alternative representation of preferences, the distance function, $d(q,u)$. This function gives the factor of proportionality by which we have to deflate a given consumption vector q in order to attain a given utility level u; i.e., $d(q,u^r) \equiv Max\{\delta: u(q/\delta)=u^r\}$. The Malmquist index equals $d(q^j,u^r)/d(q^k,u^r)$, where u^r is a reference utility level. It thus answers the question "what is the ratio of the scaling factors which must be applied to country j's and country k's consumption bundles to attain the reference utility level?" And, since the distance function is linearly homogeneous in q, it follows that the Malmquist index equals λ when $q_j=\lambda q_k$. As shown in Figure 1, the Laspeyres-Malmquist index equals *OJ/OR* and the Paasche-Malmquist index equals *OS/OK*.

All three real-income indexes take particularly simple forms when preferences are homothetic (so all income-elasticities of demand are unity). Formally, homotheticity implies that $e(p,u)=u.\epsilon(p)$ and $d(q,u)=u.\delta(q)$). Hence, it follows that all three indexes are independent of reference prices or utility levels. Moreover, all three indexes are equal to one another and in turn equal the ratio of the two country's utility levels, $u(q^j)/u(q^k)$. Unfortunately, homothetic tastes are overwhelmingly rejected in general, and so all three indexes are inherently ambiguous. One country's true real income relative to another depends on which true index is used and on the reference price vector or utility level at which the comparison is made. When multilateral comparisons between many countries are being made, the degree of ambiguity is correspondingly greater. However, it seems intuitively clear that all countries should ideally be compared using the *same* reference prices or utility level, even though this means abandoning bilateral characteristicity.

While each of the true real-income indexes provides a clear answer to a well-defined question, they can only be calculated precisely if we know the true structure of preferences (whether expressed in terms of the expenditure or the distance function). In practice, such knowledge is never available, of course, and in its absence we can only approximate the true indexes if we first estimate a complete system of demand equations. This is an expensive and difficult operation, especially if there are many commodity groups. Hence, most of the time we are forced to use some empirically-based index. A huge research program is thus opened up, exploring the relationship between various

index number formulae and the true indexes under different assumptions about consumer preferences. Among the many results of this kind (surveyed in Diewert, 1981), I give here three which are particularly relevant to international comparisons:

1. The fixed-weight Laspeyres real-income index provides a first-order approximation to the Laspeyres-Allen index (and thus to the Paasche-Konüs index); moreover, the Laspeyres index is exact (i.e., it equals the true index) under Leontief preferences. This is easily seen by taking a Taylor's Series approximation to $e(p^k,u^j)$ in the neighbourhood of p^j:

$$e(p^k,u^j) \approx e(p^j,u^j) + (p^k-p^j)'q^j + (p^k-p^j)'e^{jj}_{pp}(p^k-p^j)/2. \qquad (9)$$

   Cancelling terms and dividing by base-country expenditure gives the desired result:

$$Q^{jk}_{Ak} \approx Q^{jk}_{L} + [(p^k-p^j)'e^{jj}_{pp}(p^k-p^j)]/(2p^k.q^k). \qquad (10)$$

   With Leontief preferences, consumers have no substitution possibilities and so e^{jj}_{pp} and all higher-order derivatives vanish and the fixed-weight Laspeyres index is exact.

2. The Fisher Ideal real-income index is exact when the utility function is a homogeneous quadratic: $u=(q'Aq)^{1/2}$, where A is a symmetric matrix. This importance of this result, due to Konüs and Byushgens (1926), has been emphasised by Diewert (1976). He notes that the quadratic utility function is a *flexible* functional form, meaning that it provides a second-order approximation to an arbitrary twice differentiable linearly homogeneous utility function. Since the Fisher Ideal index is exact for a flexible functional form, there is a presumption that it should provide a good approximation to the true index for *any* homogeneous utility function. Diewert argues strongly for the use of what he calls "superlative" indexes, i.e., indexes which are exact for a flexible functional form.

3. The Theil-Törnqvist real-income index is exact when the utility function is a homogeneous translog, i.e., $\ln u(q) = \alpha_0 + \Sigma_i \alpha_i \ln q_i + ½\Sigma_i\Sigma_h \alpha_{ih} \ln q_i \ln q_h$. This result, proved by Diewert (1976), makes use of another index, the Theil-Törnqvist real-income index defined by $\ln Q_T = ½\Sigma_i (\omega_{ij}+\omega_{ik}) \ln (q_{ij}/q_{ik})$, where ω_{ij} is the budget share of good i in country j. Diewert also proved an important extension: if the distance function is a general (nonhomogeneous) translog, then

the Theil-Törnqvist real-income index equals the Malmquist index (and so is exact) provided the latter is evaluated at u^*, where $u^* = (u^j u^k)^{0.5}$. Since the translog is also a flexible functional form, the Theil-Törnqvist index is also superlative. And, since the translog is more general than the quadratic, these results suggest that the Theil-Törnqvist real-income index is likely to provide a good approximation to the true index in very general circumstances, though Diewert concedes that the Fisher Ideal is very close to it in most practical applications.

These results underlie Diewert's advocacy of the EKS method. However, it has not been sufficiently stressed hitherto that the results are appropriate to *bilateral* comparisons only. It can be shown (see Neary, 1996) that when the utility function is a homogeneous quadratic the EKS index is exact. But, from the Konüs-Byushgens result, this implies that the EKS index is *identical* to the Fisher Ideal index in this case. In other words, the EKS index is *redundant* whenever it is exact. Similar results hold for the CCD (Caves, Christensen and Diewert, 1982) index, which extends the Theil-Törnqvist index to multilateral comparisons in the same way as the EKS index extends the Fisher Ideal index.

These results suggest that the economic theory of index numbers does not provide a rationale for using the EKS or CCD indexes. When they are exact they are identical to bilateral indexes; and when they are not exact their performance relative to bilateral indexes (which are also superlative in Diewert's sense) is an empirical question which needs to be investigated. Moreover, except when preferences are homothetic, these indexes do not compare the real incomes of different countries at a common reference price or utility level.

4. True Indexes of Exchange Rates and Real Incomes

Where do the essentially negative results reported at the end of the last section leave the Geary method? While we have seen that the economic theory of index numbers provides little justification for the EKS method, it does not endorse the Geary method either. In this section, I want to suggest how the two approaches may be reconciled.[14]

[14] Marris (1984) also relates the Geary exchange rates to true quantity indexes but he does not discuss how corresponding true world prices may be calculated.

The first step is to note that equation (1) defining the Geary exchange rates has a straightforward Laspeyres price-index form. It is therefore subject to the usual criticism that is fails to allow for substitution possibilities. This suggests using instead the Konüs alternative:

$$\mathrm{E}_j = \frac{e(\Pi, u^j)}{\sum_i p_{ij} q_{ij}}, \qquad j = 1, \ldots m. \tag{11}$$

This compares the cost of attaining country j's utility at some world prices Π with the actual cost of attaining it at domestic prices. However, the world prices Π have not yet been defined. The crucial next step is to follow Geary and to require that Π satisfy some desirable aggregation properties. They cannot satisfy (2), the commodity balance equations expressed in terms of *actual* quantities consumed. However, they can satisfy analogous equations expressed in terms of *imputed* consumption bundles, which by Shephard's Lemma equal the price derivatives of the numerator of (11):

$$q^{*j} = e_{\Pi}(\Pi, u^j). \tag{12}$$

Now, it is straightforward, by analogy with (2), to define world prices as follows:

$$\Pi_i = \frac{\sum_j \mathrm{E}_j p_{ij} q_{ij}}{\sum_j q^*_{ij}}, \qquad i = 1, \ldots n. \tag{13}$$

It seems natural to describe (11) and (13) as defining *Geary-Konüs* true exchange rates and world prices respectively. Finally, the implied measures of each country's real income may be obtained from (11) by multiplying each country's nominal income $p^j.q^j$ by its true exchange rate E_j. This leads to *Geary-Allen* indexes of the real income of each country at world prices:[15]

[15] Extending the approaches reviewed in Section 3, it should also be possible to devise Geary-Malmquist true exchange rates, world prices and real incomes. However, their relationship with the original fixed-weight Geary approach is more tenuous.

$$Z^j_{GA} = \sum_i^n \Pi_i q^*_{ij} = e(\Pi, u^j), \qquad j=1, \dots m. \tag{14}$$

I claim that the system given by equations (11), (13) and (14) 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 intercommodity substitution. Like the latter, it leads to a matrix of expenditure levels expressed in a common world currency which can be consistently aggregated across countries and across commodities. Admittedly, this property of matrix consistency applies to imputed rather than to actual consumption levels. Nevertheless, this is appropriate when we wish to allow for intercommodity substitution and so avoid the bias of the Gerschenkron effect in international comparisons.

The final question is which of the available empirical methods best approximates the Geary-Konüs exchange rates, world prices and real incomes. In principle, this is an empirical question and the answer to it must await extensive applied work, examining the implications of alternative assumptions about consumer preferences. However, in the present stage of our knowledge, it seems safe to take the original Geary method as providing the best available approximation to the true E and Π. The Geary method provides a Laspeyres-type approximation to the true indexes, whereas the EKS index and its extensions provide a second-order approximation to an inconsistent set of international comparisons. (See Neary, 1996.)

This approach also throws light on a question which is obscure in the Geary approach: to which country, if any, do the calculated "world" prices correspond? In the ICP, these prices were found to lie closest to the prices in middle-income countries such as Italy. In the Geary-Konüs case, we can see why this must be so, if we make the additional assumption that preferences are quasi-homothetic, so that the expenditure function exhibits the "Gorman (1953) polar form":

$$e(p,u) = a(p) + u.b(p), \tag{15}$$

where $a(p)$ and $b(p)$ are arbitrary linearly homogeneous functions of prices.[16] This demand specification is the most general which allows exact aggregation across individuals. Using this in (13) yields:

$$\sum_j E_j p_{ij} q_{ij} = \Pi_i \sum_j q^*_{ij} = m\left[a_i(\Pi) + \bar{u} b_i(\Pi)\right], \tag{16}$$

where $\bar{u} = \Sigma_j u^j/m$. This shows that, when demand behaviour is characterised by the Gorman polar form, total world expenditure on each commodity (both imputed world expenditure valued at world prices and actual world expenditure aggregated at true exchange rates) would be voluntarily consumed by m hypothetical individuals, each of whose utilities equalled a simple average of the utilities in all m countries. This shows that the Geary-Konüs world prices (and presumptively the Geary world prices which approximate them) correspond to the domestic prices of an "average" country in an appropriate sense. Putting this differently, the Geary-Konüs world prices would generate the actual world demand pattern if world utility was equally distributed.[17]

5. Conclusions

In this paper, I have reviewed alternative methods of making international comparisons of purchasing power and real incomes. For major multilateral comparative studies, the choice lies between two methods, that of Roy Geary and that of Eltetö, Köves and Szulc. Both are widely used in practice, the former by the U.N.'s International Comparison Project associated with the University of Pennsylvania and the latter by the Statistical Office of the European Union. The Geary method has the great advantage of allowing consistent disaggregation of real spending across countries and commodities.

---

[16] This can be thought of a generalisation of the Stone-Geary linear expenditure system, whose expenditure function is: $e(p,u) = \Sigma\, \gamma_i p_i + u.\Pi(p_i/\beta_i)^{\beta i}$, $\Sigma\beta_i = 1$.

[17] I am grateful to Patrick Honohan for this interpretation and for encouraging me to make explicit that (16) applies to world expenditure on each commodity and not just to aggregate expenditure. It is straightforward to extend this to allow for different numbers of individuals in each country, provided all have identical quasi-homothetic preferences.

However, the consensus in the literature, until now, has been that the EKS method is more firmly grounded in the economic theory of index numbers.

I have argued in this paper that the Geary method should be preferred to the EKS method on theoretical grounds, both negative and positive. Negatively, I have suggested that the desirable properties of the EKS method apply only in bilateral contexts and that when this method is exact it is also redundant. More positively, I have argued that the "ideal" approach to making international comparisons leads to true indexes of exchange rates, world prices and real incomes which combine the micro foundations of the economic approach to index numbers with the matrix consistency of the Geary method. Of course, unless we know the underlying demand system, true indexes of this type cannot be calculated exactly. Of the available empirical methods, the Geary method is to be preferred since it provides an approximate answer to the correct question.

I have also shown how one great Irish invention - the Gorman polar form - can be used to throw further light on another - the Geary method of calculating purchasing power parities. When demands are generated by the Gorman polar form (of which the linear expenditure system is a special case), the true Geary world prices can be interpreted as the prices of a hypothetical country whose utility level is the average of individual countries' utilities. This result draws an interesting link between Roy Geary's two principal contributions to economic theory; it also points towards extensions to investigate the properties of the true Geary indexes when demands are generated by more general specifications, such as those proposed by Muellbauer (1975).

Of course, these arguments do not settle the case for the Geary method. At a theoretical level, the results of consumer theory only provide a justification for the Geary method when applied to international comparisons of household expenditure patterns. The theoretically appropriate treatment of investment and government spending in international comparisons must await further work. At the empirical level, many questions remain concerning the relative performance of the different indexes we have considered and a rich empirical agenda is opened up by the need to answer them. However, the desirability of consistent disaggregation of the estimates of real consumer spending in different countries suggests that the Geary approach to calculating world prices is unlikely to be surpassed. In that sense, the last word may indeed go to Geary: if world prices exist, "they could scarcely be defined reasonably in any other terms" than

those given by (2) and (13).

References

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

Allen, R.G.D. (1949): "The economic theory of index numbers," *Economica*, 16, 197-203.

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, 73-86.

Deaton, A.S. (1979): "The distance function in consumer behaviour with applications to index numbers and optimal taxation," *Review of Economic Studies*, 46, 391-405.

Diewert, W.E. (1976): "Exact and superlative index numbers," *Journal of Econometrics*, 4, 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, 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, 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, 67-86; reprinted in W.E. Diewert and A.O. Nakamura (eds.): *op. cit.*

Diewert, W.E. (1993): "The early history of price index research," in W.E. Diewert and A.O. Nakamura (eds.): *op. cit.*, 33-65.

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, 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, 321-358.

Eltetö, O. and P. Köves (1964): "On a problem of index number computation relating to international comparisons" [in Hungarian], *Statisztikai Szemle*, 42, 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, 1-39.

Geary, R.C. (1950): "A note on 'A constant-utility index of the cost of living'," *Review of Economic Studies*, 18, 65-66.

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, 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*, 21, 63-80.

Hill, T.P. (1984): "Introduction: The special conference on purchasing power parities," *Review of Income and Wealth*, 30, 125-133.

Isard, P. (1977): "How far can we push the law of one price?," *American Economic Review*, 67, 942-948.

Isenman, P. (1980): "Inter-country comparison of 'real' (PPP) incomes: Revised estimates and unresolved questions," *World Development*, 8, 61-72.

Keynes, J.M. (1930): *A Treatise on Money, Volume 1: The Pure Theory of Money*, London: Macmillan.

Khamis, S.H. (1970): "Properties and conditions for the existence of a new type of index numbers," *Sankhyā*, B, 32, 81-98.

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

Konüs, A.A. (1924): "The problem of the true index of the cost of living" [in Russian], *The Economic Bulletin of the Institute of Economic Conjuncture* (Moscow), 9-10, 64-71; English translation in *Econometrica*, 7, 1939, 10-29.

Konüs, A.A., and S.S. Byushgens (1926): "On the problem of the purchasing power of money" [in Russian], *Voprosi Konyunkturi* II(1), 151-172.

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

Kravis, I.B., A.W. Heston and R. Summers (1978): *International Comparisons of Real Product and Purchasing Power*, Baltimore: Johns Hopkins Press.

Kravis, I.B., A.W. Heston and R. Summers (1982): *World Product and Income: International Comparison of Real Gross Domestic Product*, Baltimore: Johns Hopkins Press.

Kravis, I.B., Z. Kennessy, A.W. Heston and R. Summers (1975): *A System of International Comparisons of Gross Product and Purchasing Power*, Baltimore: Johns Hopkins Press.

Malmquist, S. (1953): "Index numbers and indifference surfaces," *Trabajos de Estadistica*, 4, 209-242.

Mankiw, N.G., D. Romer and D.N. Weil (1992): "A contribution to the empirics of economic growth," *Quarterly Journal of Economics*, 107, 407-437.

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

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

Neary, J.P. (1988): "Determinants of the equilibrium real exchange rate," *American Economic Review*, 78, 210-215.

Neary, J.P. (1996): "True multilateral indexes for international comparisons of purchasing power and real incomes" (in preparation).

Officer, L.H. (1989): "The national price level: Theory and estimation," *Journal of Macroeconomics*, 11, 351-373.

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, 5-77.

Purvis, D.D. (1982): "Exchange rates: Real and monetary factors," *Economic and Social Review*, 13, 303-314.

Quah, D. (1993): "Empirical cross section dynamics in economic growth," *European Economic Review*, 37, 426-434.

Samuelson, P.A. (1974): "Analytical notes on international real-income measures," *Economic Journal*, 84, 595-608.

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

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

Samuelson, P.A. and S. Swamy (1974): "Invariant economic index numbers and canonical duality: Survey and synthesis," *American Economic Review*, 64, 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, 327-336.

Szulc, B.J. (1964): "Index numbers for multilateral regional comparisons" [in Polish], *Przeglad Statystyczny*, 3, 239-254.

Index Number	Formula	Base-Weighted (Laspeyres): $r=k$	Current-Weighted (Paasche): $r=j$
Fixed-Weight	$\frac{p^r.q^j}{p^r.q^k}$	$\frac{OA}{OB}$	$\frac{OC}{OD}$
Allen	$\frac{e(p^r,u^j)}{e(p^r,u^k)}$	$\frac{OE}{OB}$	$\frac{OC}{OF}$
Konüs	$\frac{p^j.q^j}{p^k.q^k} \div \frac{e(p^j,u^r)}{e(p^k,u^r)}$	$\frac{OC}{OF}$	$\frac{OE}{OB}$
Malmquist	$\frac{d(q^j,u^r)}{d(q^k,u^r)}$	$\frac{OJ}{OR}$	$\frac{OS}{OK}$

Table 1: Alternative Quantity Index Numbers, Each Measuring the Real Income of Country *j* Relative to that of Country *k*, using Country *r* as Reference

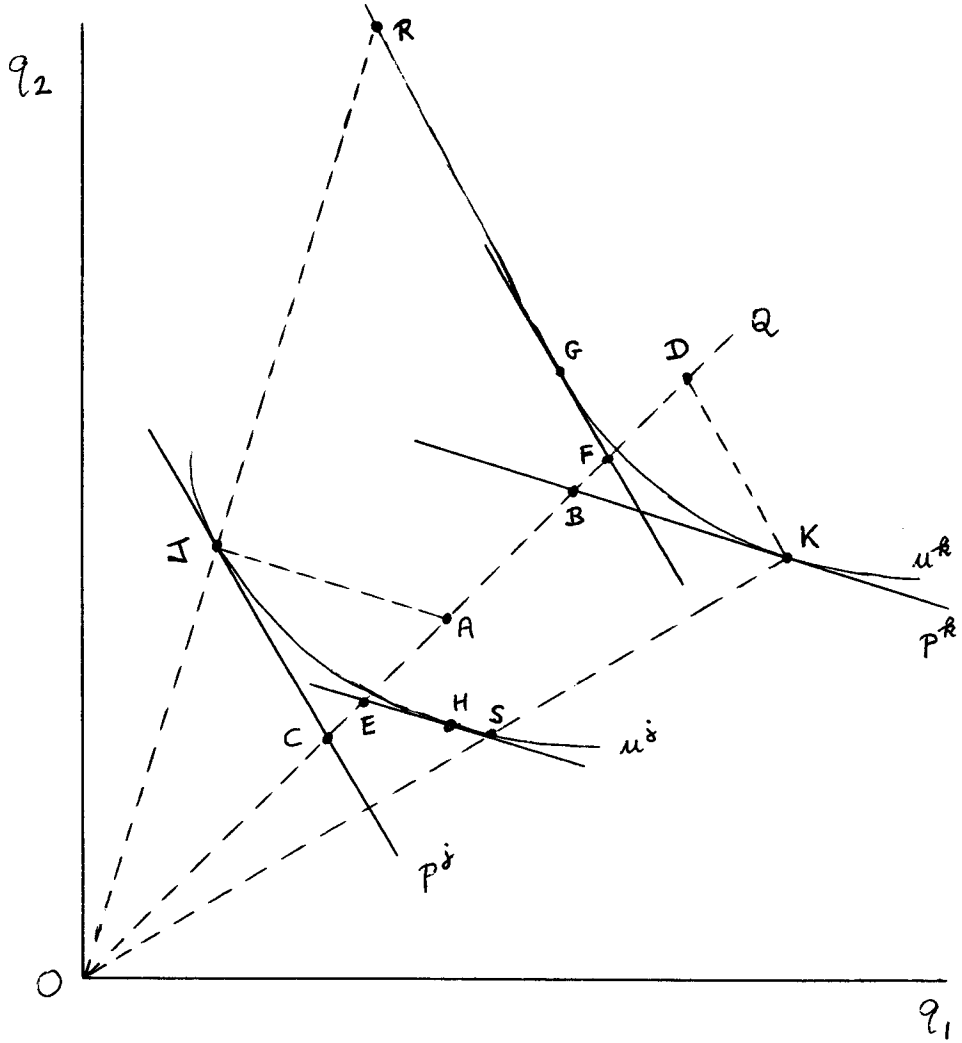

Figure 1: Comparing the Real Incomes of Countries *j* and *k*

Centre for Economic Research Papers

Working Papers represent preliminary reports on research in progress and should not be cited without permission from the authors. Policy Papers represent preliminary reports on policy-oriented research carried out by members or associates of the Department of Economics, University College Dublin. The following is a comprehensive list of the last year's working and policy papers issued by the Centre A complete list of other publications of the Centre is available on request.

Working Papers 1994

WP94/1: Dermot Leahy: "Export Enhancing Tariff Protection with Strategic Precommitment". March 1994.

WP94/2: Rodney Thom: "A (Weak) Exogeneity Test of the German Leadership Hypothesis." March 1994

WP94/3: Frank Barry, John Bradley, Kieran Kennedy and Nuala O'Donnell: "Labour Market Performance in the EU Periphery: Lessons and Implications" March 1994.

WP94/4: David Madden: "How does Unemployment Affect Marginal Direct and Indirect Tax Reform?." April 1994

WP94/5: Dermot Leahy and J. Peter Neary: "Learning by Doing, Precommitment and Infant-Industry Protection." April 1994

WP94/6: Brendan Walsh: "Wage Convergence and Integrated Labour Markets: Ireland and Britain, 1841-1991. May 1994.

WP94/7: Cormac Ó Gráda "Anthropometric History What's in it for Ireland?" June 1994.

WP94/8: Joel Mokyr and Cormac Ó Gráda: "Height and Health in the United Kingdom 1815-1860: Evidence from the East India Company Army" June 1994.

WP94/9: J. Peter Neary: "Factor Mobility and International Trade" June 1994.

WP94/10: Cormac Ó Gráda and Brendan Walsh: "Recent Trends in Fertility and Population in Ireland, North and South" July 1994.

WP94/11: James Bergin and B L Lipman, Queen's University Ontario: "Evolution with State - Dependent Mutations" August 1994.

WP94/12: David Madden: "Omitted Variables, Dynamic Specification and Tests for Homogeneity" August 1994.

WP94/13: Frank Barry: "Peripherality in Economic Geography and Modern Growth Theory: Evidence from Ireland's Adjustment to Free Trade". August 1994.

WP94/14: Frank Barry: "Exchange Rate Policy when the Labour Market Exhibits Hysteresis" September 1994

WP94/15: Kevin O'Rourke: "Industrial Policy, Employment Policy and the Non-Traded Sector". October 1994.

WP94/16: Anthony Murphy: "Artificial Regression Based Mis-Specification Tests for Discrete Choice Models" October 1994.

WP94/17: Anthony Murphy: "A Simple Artificial Regression Based LM Test of Asymmetry in the Logit Model". October 1994.

WP94/18: Anthony Murphy: "A Simple Artificial Regression Based Test of the fit of