On Aggregation Methods of Purchasing Power Parities

J. R. Cuthbert,

M. Cuthbert

OECD

OECD

DEPARTMENT
OF ECONOMICS AND STATISTICS

WORKING PAPERS

No. 56: ON AGGREGATION METHODS OF PURCHASING POWER PARITIES

by

J. R. Cuthbert and M. Cuthbert

November 1988

OCDE

OECD

ECONOMICS AND STATISTICS DEPARTMENT

WORKING PAPERS

This series of Working Papers is designed to
make available, to a wider readership,
selected studies which the Department has
prepared for use within OECD. Authorship is
generally collective, but main individual
authors are named. The Papers are generally
available in their original language, English
or French, with a summary in the other.

Comment on the Papers is invited, and may be
sent to OECD, Department of Economics and
Statistics, 2 rue André Pascal, 75775 Paris
Cedex 16, France. Additional copies of the
Papers on a limited basis can be forwarded on
request.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The development of purchasing power parities as converters of national accounts aggregates to comparable volume figures is important for international economic comparisons. This study is primarily concerned with the aggregation of price relativities to basic heading level: that is, the level below which there are no expenditure weights available across all of a given group of countries. Eight possible methods of aggregation to basic heading level are identified and appropriate summary statistics developed to assist in the subsequent practical investigation of these methods. This is undertaken using price data for 37 basic headings in ten OECD countries.

* * * * *

Pour les comparaisons économiques internationales il est important de développer les calculs de parités de pouvoir d'achat. En effet, ces parités permettent de convertir les agrégats de comptabilité nationale et donc de disposer de chiffres en termes réels qui sont comparables de pays à pays. Cette étude porte principalement sur les problèmes d'agrégation des prix relatifs des biens ou services individuels jusqu'au niveau des rubriques de base, c.a.d. le niveau le plus fin pour lequel une pondération par la dépense est disponible, pour tous les pays comparés. L'étude montre que huit méthodes différentes d'agrégation peuvent être employées et propose des mesures statistiques pour examiner les avantages et inconvénients de chacune de ces méthodes en pratique. Les données sur les prix de 37 rubriques de base dans dix pays de l'OCDE sont utilisées pour tester les différentes méthodes d'agrégation.

NO. 56: ON AGGREGATION METHODS OF PURCHASING POWER PARITIES

by

J. R. Cuthbert and M. Cuthbert

ON AGGREGATION METHODS OF PURCHASING POWER PARITIES

TABLE OF CONTENTS

SUMMARY AND CONCLUSIONS.

A major problem in international comparison work is the development
of purchasing power parities in their role as convertors of national
accounts aggregates from nominal values to comparable real volume
figures. This study is primarily concerned with the aggregation of
price relativities to basic heading level: here, a basic heading is
defined as the lowest level for which expenditure figures are
available for a given group of countries. Aggregation to basic
heading level must therefore take place in the context of limited
information; typically, price data will be available while
expenditure or quantity information will be limited or unavailable.

The study contains four main strands:

(i)   an introductory discussion of some of the main aggregation
      methods :- the Geary Khamis method, normally used from basic
      heading level to GDP level; the Fisher / EKS approach and the
      Country Product Dummy (CPD) method;

(ii) the development of new work on the theory of the three methods.
      This theoretical work contributes to a general understanding of
      the workings of these techniques and some of their weaknesses:
      in addition, it leads to the suggestion of eight possible
      methods of aggregation to basic heading level for further
      practical investigation.

(iii) the development of appropriate summary statistics to help in
      the data analysis. Specifically, there is developed (a) a

summary statistic to provide a measure of the similarity.
between PPPs and (b) a statistic to measure how multiplicative
the structure of a given set of prices is; that is, how closely
the structure of a set of prices approximates to the simple
product of item and country effects.

(iv) an analysis of price and expenditure data for 37 basic headings
in 10 OECD countries. The data is aggregated for each basic
heading by each of the eight methods under investigation, and
the resulting PPP structures analysed.

In more detail, the contents of the individual chapters are as
follows :

In chapter 1, the theory of the Geary Khamis approach [1,2] is
introduced and then developed. The Geary Khamis equations are stated
and explicit solutions of these equations for the 2, 3, and general
n country cases are determined. Important special properties of the
this approach are derived. In particular, it is shown that when the
structure of prices is multiplicative, the Geary Khamis Purchasing
Power Parity solution is independent of the quantities of goods
traded.

Typically, in international comparisons, the Geary Khamis, (GK),
approach is used to aggregate price parities from basic heading
level to GDP level.

Chapter 2 continues the discussion of the GK theory, and

develops options for extending this approach so that it can be applied in aggregating from individual commodities to basic headings. Three possible approaches are suggested :-

The first is based on developing an appropriate sampling design. This approach is most relevant in the two country case - and is set out in Appendix B.

The second is based on assigning imputed quantities, and solving the full Geary Khamis equations. Three methods of imputing quantities are suggested.

The final approach is based on exploiting special characteristics of the price structure - in particular, where the price structure is nearly multiplicative.

Chapter 3 develops the theory of the Fisher index [3], and the EKS approach [4,5] from a statistical viewpoint. By considering a parametric formulation of the PPP estimation problem, alternative estimators of PPPs, based on a regression analysis of the logs of the Fisher indices, are derived. It is shown that the standard EKS approach will normally be sub optimal relative to these alternative approaches. Two versions of the log Fisher regression technique are developed - one of which allows explicitly for the possibility of a consistent pricing differential between characteristic and non characteristic goods: that is, between goods which are in some sense typical and those which are untypical in the country concerned.

In chapter 4, the theory behind the Country Product Dummy (CPD) approach [7,8] is considered. It is shown how this technique can be modified to incorporate the possibility of a consistent pricing differential between characteristic and non characteristic goods. It is also shown that, even where such a systematic pricing effect exists, the decision of whether or not to correct for this effect depends on the circumstances of each case.

Chapter 5 contains the necessary preliminaries to the detailed analysis of actual price data described in the later chapters of the report.

The data analysed is described: (37 product groupings over ten countries).

The PPPs to be calculated from item level to basic heading level are defined. There are 8 of these, falling into three main families: three in the EKS family; three versions of the GK; and two versions of the CPD.

Further, to assist in the data analysis summary measures are developed: in particular, three such statistics have been constructed: (a) a metric, to measure the distance between different PPPs; (b) a measure of the multiplicativity of the price structure; (c) measures of within group homogeneity and between group distances, for groups of PPPs.

In chapter 6, there is a detailed analysis of the question of

whether a systematic price differential exists between
characteristic and non characteristic goods. Using the modified
version of the CPD approach, it is shown that a statistically
significant price differential does exist - and that its size is
such as to have a potentially important operational effect.
The conclusion is drawn that appropriate techniques should be used
in practice, to enable these effects to be detected, and if
necessary, corrected for. To do this adequately, would involve
having some broad estimates of expenditure or quantities at item
level.

Chapter 7 continues the detailed data analysis, by considering the
internal homogeneity of the basic PPP families - and the distances
between these families. It is shown that, when the price structure
is highly multiplicative, the different approaches to calculating
PPPs tend to converge, as the preceding theory suggests.
But, unless the price structure is highly multiplicative, the
different PPP versions considered can diverge very significantly. It
is also shown that, for the data analysed, the CPD family of PPPs
tends, in a sense, to occupy an intermediate position between the
EKS and the GK families.

The conclusion that emerges is that, unless the price structure is
highly multiplicative, the choice of the appropriate method of
calculating PPPs is by no means clear cut. Three possible approaches
are suggested, not necessarily mutually exclusive;

(a) other things being equal, there might be some preference for a CPD approach, as a possible compromise, given its intermediate position relative to the other approaches.

(b) effort could be devoted to obtaining expenditure or quantity data at item level and then using a GK approach.

(c) effort could be put into collecting better expenditure data, so as to reduce the size of the basic headings. As basic headings become smaller, they are likely to become more multiplicative - and hence the choice of PPP method less critical. There are limits to this approach, however - particularly if using the CPD method, since the CPD method involves fitting a model with a relatively large number of parameters, and so cannot be applied reliably if there are too few items in a given basic heading.

Finally, chapter 8 considers the results of aggregating the basic heading level PPPs over the basic headings studied. Two methods of aggregation, the GK and Gerardi (UCW method), are used at this stage [9,10]. Combined with the eight methods used at the first stage, this means that sixteen aggregate PPPs in all are calculated. It is shown that there is a high degree of homogeneity within each of the six groups defined by applying either GK or Gerardi aggregation to one of the three basic PPP families. There is also a fairly high degree of homogeneity among all the GK aggregated indices - and also among all the Gerardi aggregated indices. However, there tend to be consistent differences between GK and

Gerardi aggregated versions, even of the the same first stage indices. It is shown that these differences relate mainly to underlying US price levels being consistently higher for GK aggregated indices compared with Gerardi indices.

CHAPTER 1.

THE GEARY KHAMIS APPROACH

1. Introduction.

1.1 The Geary Khamis approach is one of the most commonly used methods of aggregating purchasing power comparisons from basic heading levels to GDP level. In this section the basic equations of the Geary Khamis approach are introduced. Subsequent sections of this chapter explore some of the theoretical aspects of this approach, including the explicit solution in the 2 and 3 country case, and the form of the general solution in the n country case. In the next chapter, some different possible approaches to the problem of extending the GK approach to aggregation from individual item up to main heading level are discussed.

The Geary Khamis method determines international prices for commodities simultaneously with country PPPs at GDP level. The theory, as usually stated, proceeds in terms of there being data available on prices and quantities of individual commodities. However, the PPP solution can also be obtained from information on price relativities and expenditures by the trick of defining "pseudo" prices in terms of the price relativities at basic heading level and "pseudo" quantities as the quantity of a good which can be

bought for one unit of the numeraire currency in the base country.

The general multilateral problem is that of determining a set of PPPs for a group of countries. Without loss of generality or the introduction of any bias to the results, one of the countries in the group is chosen as having the numeraire currency: that is, the currency to be used in the calculations as the unit of account. For each country expenditure data is expressed in national currencies for national account basic headings. In particular, GDP estimates are available:

$$\text{GDP} = \Sigma\, p_{ij}q_{ij} \qquad\qquad i = 1,..,m;\ j = 0,1,..,n$$

$$\text{where} \qquad p_{ij} = \text{price of item } i \text{ in country } j$$

$$q_{ij} = \text{quantity of item } i \text{ in country } j$$

If estimates of international prices, π_i, $i = 1,...m$, were available, expressed in terms of the numeraire currency, then a volume measure for each country could be defined as

$$\Sigma\, \pi_i\, q_{ij}$$

Note that the quantities of different items cannot simply be added together. For example, lengths of cable and tonnes of wheat cannot be added together. Thus the quantities have to be expressed in a common unit of account; that is, in international prices.

Following on from this, since

$$\text{price} = \frac{\text{expenditure}}{\text{volume}}$$

an average price measure for country j could be derived. This is
defined as

$$\frac{\Sigma \ p_{ij} q_{ij}}{\Sigma \ \pi_i \ q_{ij}} \qquad j = 0,1,\ldots,n$$

The purchasing power parity between countries j and k is defined as
the ratio of the two price measures.

$$_j PPP_k = \frac{\Sigma \ p_{ij} \ q_{ij}}{\Sigma \ \pi_i q_{ij}} \bigg/ \frac{\Sigma \ p_{ik} q_{ik}}{\Sigma \ \pi_i \ q_{ik}} \qquad ----(1)$$

Geary considered that a suitable definition of the international
prices of an item was the sum total of expenditures on the item
throughout the countries being considered in the exercise, divided
by the quantity. Again this requires a means of converting
expenditures in national currencies into the numeraire currency. One
way of doing this is to use the values $_j PPP_0$, where country 0
is that of the numeraire currency. Thus the international price of
commodity i is defined as

$$\pi_i = \frac{\Sigma \ p_{ij} q_{ij} \ / \ (_j PPP_0)}{\Sigma \ q_{ij}} \qquad ----(2)$$

$$i = 1,\ldots,m$$

Equations (1) and (2) define a set of international prices π_i and the PPPs for each country at GDP level. On noting the identity, for the numeraire country, that,

$$\Sigma\ p_{i0}\ q_{i0}\ =\ \Sigma\ \pi_i\ q_{i0}$$

equation (1) in fact simplifies to

$$_jPPP_0\ =\ \frac{\Sigma\ p_{ij}\ q_{ij}}{\Sigma\ \pi_i q_{ij}}$$

2. The Geary Khamis Solution in the 2 and 3 Country Cases.

2.1 In the special cases where only two or three countries are involved, it is a fairly straightforward matter to solve the basic Geary Khamis equations, Equations (1) and (2) above, explicitly. In the two country case, suppose that the countries are denoted by 0 and 1, and that the currency of country 0 is the numeraire currency. Then the purchasing power parity, $_1PPP_0$, which is denoted here by R_1, is given by the formula:

$$R_1\ =\ \frac{\sum \dfrac{p_{i1}\ q_{i0}\ q_{i1}}{q_{i0}\ +\ q_{i1}}}{\sum \dfrac{p_{i0}\ q_{i0}\ q_{i1}}{q_{i0}\ +\ q_{i1}}}\qquad ----(3)$$

In the three country case, if the countries are denoted by 0, 1 and 2, and country 0 is again taken as providing the numeraire currency, then the PPP, R_1, is given by

$$R_1 = \frac{\left(\sum \frac{p_{i1}q_{i0}q_{i1}}{(q_{i0}+q_{i1}+q_{i2})}\right)\left(\sum \frac{p_{i2}q_{i2}(q_{i0}+q_{i1})}{(q_{i0}+q_{i1}+q_{i2})}\right) + \left(\sum \frac{p_{i1}q_{i1}q_{i2}}{(q_{i0}+q_{i1}+q_{i2})}\right)\left(\sum \frac{p_{i2}q_{i0}q_{i2}}{q_{i0}+q_{i1}+q_{i2}}\right)}{\left(\sum \frac{p_{i0}q_{i0}q_{i1}}{q_{i0}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i2}(q_{i0}+q_{i1})}{(q_{i0}+q_{i1}+q_{i2})}\right) + \left(\sum \frac{p_{i0}q_{i0}q_{i2}}{q_{i0}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{i0}+q_{i1}+q_{i2}}\right)}$$

$$\text{---- (4)}$$

with a corresponding expression for R_2

A derivation of equation (4) is given in Appendix A.

2.2   It is instructive to consider the forms of these solutions in the two and three country cases. First of all, equation (3) shows that the two country PPP is in fact a ratio of an arithmetic average of prices in country 1 to a corresponding average of prices in country 0. The weights in these arithmetic averages are functions of the quantities in both countries: and the form of these weights,

$$\frac{q_{i0} \; q_{i1}}{(q_{i0} + q_{i1})}$$

is such that the greatest weight will be attached in the price comparisons to items which have relatively large quantities in both countries: conversely, items which have small quantities in either

country will be given small weight.

The form of the three country solution, as given in equation (4), is rather more complex. R_1 is the ratio of two sums of products of price terms, where the prices are, again, weighted by functions of the quantities. The numerator of equation (4) includes prices in countries 1 and 2; while the denominator involves prices in countries 0 and 2. In fact the numerator and denominator of equation (4) are symmetric expressions, in the sense that the two expressions are identical if the subscripts 0 and 1 are switched with each other.

It is also interesting to consider the limiting behaviour of the three country solution, as the quantities in country 2, the q_{i2} 's, become either very small or very large.

Suppose, for convenience, that equation (4) is written in the form

$$R_1 = \frac{ab + cd}{ef + gh} \qquad \text{-------(5)}$$

Consider, first of all, what happens if $q_{i2} \longrightarrow 0$ for all i. In this case, R_1 behaves asymptotically like a/e, which, in turn, converges simply to the two country solution given in equation (3) as would be expected.

Conversely, if $q_{i2} \longrightarrow \infty$ for all i, then R_1 behaves asymptotically as cd/gh, in the notation of equation 5. It is readily seen that the limiting value of this expression is simply,

$$R_1 = \frac{\Sigma p_{i1} q_{i1}}{\Sigma p_{i0} q_{i0}} \cdot \frac{\Sigma p_{i2}\, q_{i0}}{\Sigma p_{i2}\, q_{i1}}$$

or, equivalently,

$$R_1 = \frac{\dfrac{\Sigma p_{i1} q_{i1}}{\Sigma p_{i2} q_{i1}}}{\Bigg/}\ \frac{\Sigma p_{i0} q_{i0}}{\Sigma p_{i2} q_{i0}} \qquad\text{------(6)}$$

The numerator of this expression is a price level for country 1, defined as the ratio of country 1 expenditures, to a measure of volume for country 1 determined at country 2 prices. The denominator is the corresponding price measure for country 0. In other words, equation (6), the limiting value for R_1 as the quantities in country 2 become large, is simply the PPP between country 1 and 0 using country 2's prices as international prices. This, again, is as expected.

3. The General Geary Khamis Solution.

3.1 As noted earlier, Appendix A contains a derivation of equation (4), the explicit Geary Khamis solution in the three country case. The argument in Appendix A proceeds by showing that the country 1 and country 2 PPP's, R_1 and R_2, are solutions of the equations:

$$\sum \frac{p_{i1} q_{i1}(q_{i0} + q_{i2})}{(q_{i0} + q_{i1} + q_{i2})} \cdot R_1^{-1} - \sum \frac{p_{i2} q_{i2} q_{i1} \cdot R_2^{-1}}{(q_{i0} + q_{i1} + q_{i2})} = \sum \frac{p_{i0} q_{i0} q_{i1}}{(q_{i0} + q_{i1} + q_{i2})}$$

$$-\sum \frac{p_{i1}q_{i1}q_{i2} \cdot R_1^{-1}}{(q_{i0}+q_{i1}+q_{i2})} \quad + \sum \frac{p_{i2}q_{i2}(q_{i0}+q_{i1}) \cdot R_2^{-1}}{(q_{i0}+q_{i1}+q_{i2})} \quad - \sum \frac{p_{i0}q_{i0}q_{i2}}{(q_{i0}+q_{i1}+q_{i2})}$$

It is not difficult to see that a similar argument can be applied in the (n+1) country case, to show that $R_1, R_2, R_3, \ldots, R_n$ must be solutions of the $\underline{n}$ equations:

$$\sum \frac{p_{i1}q_{i1}(q_{i0}+q_{i2}+\ldots+q_{in})R_1^{-1}}{(q_{i0}+q_{i1}+\ldots+q_{in})} - \sum \frac{p_{i2}q_{i2}q_{i1}R_2^{-1}}{(q_{i0}+\ldots+q_{in})} - \ldots - \sum \frac{p_{in}q_{in}q_{i1}R_n^{-1}}{(q_{i0}+\ldots q_{in})}$$

$$- \sum \frac{p_{i0}q_{i0}q_{i1}}{(q_{i0}+\ldots+q_{in})}$$

$$-\sum \frac{p_{i1}q_{i1}q_{i2}R_1^{-1}}{(q_{i0}+q_{i1}\ldots+q_{in})} + \sum \frac{p_{i2}q_{i2}(q_{i0}+q_{i1}+q_{i3}\ldots+q_{in})R_2^{-1}}{(q_{i0}+\ldots\ldots+q_{in})} - \ldots - \sum \frac{p_{in}q_{in}q_{i2}R_n^{-1}}{(q_{i0}+\ldots+q_{in})}$$

$$- \sum \frac{p_{i0}q_{i0}q_{i2}}{(q_{i0}+\ldots+q_{in})}$$

etc.

$$-\sum \frac{p_{i1}q_{i1}q_{in}R_1^{-1}}{(q_{i0}+\ldots+q_{in})} - \ldots\ldots + \sum \frac{p_{in}q_{in}(q_{i0}+\ldots+q_{in-1})R_n^{-1}}{(q_{i0}+\ldots+q_{in})}$$

$$- \sum \frac{p_{i0}q_{i0}q_{in}}{(q_{i0}+\ldots+q_{in})}$$

In order to simplify the way in which these equations are written, it is convenient to adopt the following special notation - which is used for the remainder of this section only.

Denote,

$$\sum \frac{p_{i1}q_{i1}(q_{i0}+q_{i2}+\ldots+q_{in})}{(q_{i0}+\ldots+q_{in})}$$

as

$$p_1q_1(\ q_0+q_2+\ldots+q_n)$$

and

$$\sum \frac{p_{i1}q_{i1}q_{i2}}{(q_{i0}+\ldots+q_{in})}$$

as

$$p_1q_1q_2 \qquad ; \text{ etc};$$

Then, in this special notation, the above equations can be written more simply as :-

$$p_1q_1(q_0+q_2+\ldots+q_n)R_1^{-1} - p_2q_2q_1R_2^{-1}-\ldots -p_nq_nq_1R_n^{-1} = p_0q_0q_1$$

$$-p_1q_1q_2R_1^{-1}+p_2q_2(q_0+q_1+q_3+\ldots+q_n)R_2^{-1} -\ldots-p_nq_nq_2R_n^{-1} = p_0q_0q_2$$

$$-p_1q_1q_nR_1^{-1} -p_2q_2q_nR_2^{-1} -\ldots+p_nq_n(q_0+\ldots+q_{n-1})R_n^{-1} = p_0q_0q_n$$

This set of equations can be expressed in matrix form, as follows:

If A is defined to be the nxn matrix

$$
\begin{bmatrix}
p_1 q_1 (q_0 + q_2 + \ldots q_n) & -p_2 q_2 q_1 & \cdots & -p_n q_n q_1 \\
\quad - p_1 q_1 q_2 & +p_2 q_2 (q_0 + q_1 + q_3 + \ldots q_n) & \cdots & -p_n q_n q_2 \\
\\
\quad - p_1 q_1 q_n & & & +p_n q_n (q_0 + \ldots + q_{n-1})
\end{bmatrix}
$$

If R^{-1} is defined to be the vector

$$
\begin{bmatrix}
R_1^{-1} \\
\cdot \\
R_n^{-1}
\end{bmatrix}
$$

and b is defined to be the vector

$$
\begin{bmatrix}
p_0 q_0 q_1 \\
p_0 q_0 q_2 \\
\cdot \\
\cdot \\
p_0 q_0 q_n
\end{bmatrix}
$$

then the vector R^{-1} must satisfy

$$
A\, R^{-1} = b
$$

and hence, in general,

$$
R^{-1} = A^{-1} b \qquad \text{-------(7)}
$$

This is one form of the general Geary Khamis solution in the (n+1)
country case.

It is possible to express equation (7) in an alternative form, which
is in many ways more informative about the structure of the Geary
Khamis' PPPs.

Recalling that, for any non singular vector A,

$$A^{-1} = \frac{1}{|A|} \text{adj}(A)$$

where $|A|$ is the determinant of A, and adj(A) is the adjunct of A,
equation (7) can be written equivalently as

$$R^{-1} = \frac{\text{adj}(A) \cdot b}{|A|}$$

It is not difficult to see that the first element in the vector

$$\text{adj}(A) \cdot b$$

is, in fact, just the determinant of the matrix obtained by
replacing the first column of A by b.

Hence, it follows that,

$$R_1^{-1} = \frac{\begin{vmatrix} p_0 q_0 q_1 & 1 & \\ \cdot & 1 & \\ \cdot & 1 & [\text{remaining } (n-1) \text{ columns of } A] \\ \cdot & 1 & \\ p_0 q_0 q_n & 1 & \end{vmatrix}}{|A|} \quad \text{--(8)}$$

It is also possible to show, although the proof is not given here,
that the determinant in the numerator of equation (8) is equal to
the determinant of the matrix obtained by switching the subscripts 0

and 1 wherever they occur in writing down the matrix A. In other words,

$$
R_1^{-1} = \frac{\begin{vmatrix} p_0 q_0 (q_1+q_2\cdots+q_n) & -p_2 q_2 q_0 & -p_n q_n q_0 \\ -p_0 q_0 q_2 & +p_2 q_2 (q_0+q_1+q_3\cdots+q_n) & -p_n q_n q_2 \\ \vdots & & \\ -p_0 q_0 q_n & & p_n q_n (q_0+\cdots q_{n-1}) \end{vmatrix}}{\begin{vmatrix} p_1 q_1 (q_0+q_2+\cdots q_n) & -p_2 q_2 q_1 & -p_n q_n q_1 \\ -p_1 q_1 q_2 & +p_2 q_2 (q_0+q_1+q_3\cdots+q_n) & -p_n q_n q_2 \\ -p_1 q_1 q_n & & p_n q_n (q_0+\cdots q_{n-1}) \end{vmatrix}}
$$

$$--(9)$$

Thus it follows that R_1 consists of the ratio of sums of products of weighted sums of prices; the numerator ratio involving the prices in country 1 and countries 2 to n; the denominator being a similar expression in the prices of country 0 and countries 2 to n. The numerator and denominator expressions are symmetric, in the sense that the two expressions are identical on switching country subscripts 0 and 1 throughout. Similar expressions, of course, hold for the other PPPs , R_2 .... R_n .

4. Two Important Special Cases.

4.1 It is useful to conclude this section on the theory of the Geary Khamis solution by considering two important special cases.

a). Exactly Multiplicative Prices.

The price structure, p_{ij} , is defined to be exactly multiplicative if there exist parameters α_i and β_j such that

$$p_{ij} = \alpha_i \cdot \beta_j$$

In other words, the price of commodity i in country j can be expressed exactly as the product of a commodity price parameter (α_i) and a country price effect (β_j). When the price structure is exactly multiplicative, it can be shown that the Geary Khamis international prices are proportional to the parameters α_i, and the Geary Khamis PPPs, $R_1, \ldots, R_n$, are simply proportional to the β_j's. This follows on noting that:

$$\frac{\Sigma p_{ij} \, q_{ij}}{\Sigma \alpha_i \, q_{ij}} = \beta_j \left[\frac{\Sigma \alpha_i \, q_{ij}}{\Sigma \alpha_i \, q_{ij}} \right] = \beta_j$$

and that

$$\frac{\Sigma p_{ij} q_{ij} (\beta_j)^{-1}}{\Sigma q_{ij}} = \alpha_i \left[\frac{\Sigma \beta_j \, q_{ij} \beta_j^{-1}}{\Sigma q_{ij}} \right] = \alpha_i$$

so that
$$\pi_i = \alpha_i \, \beta_0$$
and
$$R_j = \beta_j \, / \, \beta_0$$

are solutions of the Geary Khamis equations, taking country
0 as the base country. One important implication of this
result is that, under exact multiplicativity, <u>the Geary
Khamis PPPs are independent of the quantities q.</u>

b).<u>Proportional Quantities Between Countries.</u>

Suppose that, for two particular countries, say country A
and country B, the quantities of products are proportional
between the two countries. By this is meant that there
exists some constant τ, such that $q_{iA} = \tau\, q_{iB}$ for all i

It follows, from the Geary Khamis defining equations, that
the Geary Khamis PPP between countries A and B, $_A PPP_B$,
has the following expression:

$$_A PPP_B = \frac{\sum p_{iA} q_{iA}}{\sum \pi_i q_{iA}} \bigg/ \frac{\sum p_{iB} q_{iB}}{\sum \pi_i q_{iB}}$$

$$= \frac{\sum p_{iA} q_{iA}}{\tau \sum \pi_i q_{iB}} \bigg/ \frac{\sum p_{iB} q_{iB}}{\sum \pi_i \, q_{iB}}$$

$$= \frac{1}{\tau} \cdot \frac{\sum p_{iA} q_{iA}}{\sum p_{iB} q_{iB}}$$

In other words, when the quantities of products are proportional between two countries, the Geary Khamis PPP between these countries does not depend on international prices, or on the prices and quantities in any other countries, but only on the prices and quantities in the two countries in question.

CHAPTER TWO

OPTIONS FOR ESTIMATING THE GEARY KHAMIS APPROACH AT ITEM LEVEL.

1. Introduction

1.1 The Geary Khamis approach to calculating PPPs, the theory of
which was discussed in the preceding section, has mainly been
applied at the stage of aggregation from basic headings up to GDP.
It is not difficult to see why this is so: calculation of Geary
Khamis PPPs depends on knowledge not just about the prices of
commodities, the p_{ij}s, but also on the quantities, the q_{ij}s, - or,
alternatively, on knowledge of relative prices and expenditures.
Basic headings are, by definition, normally the lowest level of
aggregation for which quantity or expenditure data are available.
Straightforward application of the Geary Khamis approach below basic
heading level is, therefore, not normally possible, because of lack
of the necessary data.

In principle, however, the Geary Khamis method could be used for
aggregation from individual items up to basic heading, as well as
for aggregation from basic heading up. It is therefore considered
here, whether it is possible to develop techniques of estimating the
Geary Khamis method up from individual item level, making use of

such limited information as may be available about quantities or
expenditures in order to do so. This chapter discusses three
possible approaches to the problem:-

(a) by devising an appropriate sampling scheme;

(b) by solution of the Geary Khamis equations, using imputed
   quantities;

(c) by exploring special features of the price structures.

2.   The Sampling Approach.

2.1 Calculation of PPPs at basic heading level involves an
additional problem - that of sampling. There is the problem of
choosing those individual items to be priced, as well as the problem
of arriving at a suitable method of combining the item prices into a
PPP. In certain circumstances, it may be possible to derive a
suitable probability sampling scheme for selecting individual items
- in such a way that efficient estimates of a desired PPP can be
derived as relatively simple expressions of the sampled prices.

This approach is developed in Appendix B, for the two country or
binary expression. In Appendix B, it is shown how,

   (a) if prior knowledge is available in terms of rough estimates
       of item level expenditures, and an initial estimate of the
       inter country purchasing power relativity; and

   (b) if it can be assumed that the price structure within the

basic heading in question was approximately multiplicative;

then a probability sampling scheme for selecting items can be

derived, which has the property that an efficient estimate of the

inter country Geary Khamis PPP can be obtained as a relatively

simple function of the sampled prices.

This approach is interesting in the binary comparison case.

However, in the three country case, or the more general multilateral

case, the complexity of the formula for the general Geary Khamis

solution, as shown in formulae 4 and 9 in Chapter 1, is such that

there does not appear to be any obvious extension of the sampling

approach to these cases. This sampling approach, therefore, must be

regarded as a useful suggestion for the binary comparison case, but

not one that has implications for multilateral comparisons.

3.   Solving the Geary Khamis Equations Directly Using Imputed Quantities

3.1 Suppose that price data is available for a selection of items:

that is, the price data $, p_{ij}$, is known. Then one approach which

could be adopted to calculate a Geary Khamis PPP could be to impute,

from whatever information is available, a set of notional

quantities, q_{ij}. This price and notional quantity data could then

be used to solve the defining Geary Khamis equations; that is,

equations (1) and (2) in Chapter 1.

There are a number of different ways in which notional quantities

could be imputed, depending on the available information. Three possible ways are illustrated here.

(a) Suppose that all the information that is available is whether item i is priced in country j or not. In this situation of very limited information, a possible approach would be to define

$$q_{ij} = 1 \quad \text{if item i is priced in country j;}$$
$$= 0 \quad \text{otherwise.}$$

In fact, after assigning initial q values in this fashion, it might be sensible to normalise the qs, to sum to a constant total within each country. This normalisation step would prevent any individual country being given greater volume weight in the calculation just because more items happened to have been priced in that country.

(b) Alternatively, it might be the case that additional information is available about the individual items, in the sense that it is known for each item, i, whether or not that item is "characteristic" or not in country j. The definition of characteristic taken here is that suggested by Drechsler, [6]: namely, a product exhibits characteristicity in a country if its specifications are typical for that country. Given information on characteristicity, a possible approach would be to define

$$q_{ij} = 2 \quad \text{if item i is characteristic in country j;}$$
$$= 1 \quad \text{if i priced, but not characteristic in j;}$$
$$= 0 \quad \text{if i is not priced in j.}$$

Again, after assigning initial q values in this fashion, a within country normalisation would be appropriate, as in method (a).

(c) Finally, it might be the case that, in addition to the characteristic/ non characteristic status of individual items, data is available for broad volume magnitudes of the countries concerned - probably from prior GDP estimates. If it appears plausible to assume, to a first order of approximation, that these same relativities could hold at basic heading level, then the normalised qs, as derived for each country in (b), could be multiplied by overall factors describing the inter country volume relativities.

Of course, the three techniques illustrated above for ascribing notional quantities, q, are only three examples of a large number of possibilities. Other approaches might be much more sensible, depending on the form of whatever prior information is available.

The fact that the techniques are crude is not, in itself, a strong argument against applying them. It may be better to take account, however crudely, of prior information, rather than to discount this information entirely. The important point is that, by imputing notional quantities, versions of the Geary Khamis PPP can be formally calculated. It is an empirical question how sensitive the results are to the particular techniques employed for imputing quantities. This question will be addressed in the later chapters which consider the analysis of actual data.

4.   Exploiting Special Features of the Price Structure.

4.1 The third approach to obtaining an approximation to the Geary Khamis PPP at basic heading level is one which is only feasible in exceptional circumstances, since it involves exploiting special features which the price structure, the p_{ij}s, may possess.

It has already been seen, in section 4 of Chapter 1, that if the price structure, p_{ij}, is exactly multiplicative, then the Geary Khamis PPPs are simply functions of the prices alone, and do not depend on the quantities, q_{ij}. In this case, therefore, the Geary Khamis PPPs could be obtained from knowledge of the p_{ij}s alone.

Of course, exact multiplicativity is unlikely ever to be encountered in the real world. But it is reasonable to hypothesise that, if a price structure was encountered, which was, in some sense, close to being exactly multiplicative, then an approach which breaks down the realised prices into "item effects" and "country effects", could well give PPPs which are close to the Geary Khamis PPPs.

In fact, one of the standard alternative approaches to estimating PPPs, the Country Product Dummy (CPD) approach, consists exactly of decomposing prices into item and country effects. The theory of the CPD approach is discussed in Chapter 4.

The hypothesis, therefore, is that when the observed price structure

is "close" to multiplicativity, then the CPD approach to PPPs will be a good approximation to the Geary Khamis PPP. As the CPD depends only on knowledge of prices, and not quantities, this means that an approximation to the Geary Khamis can be derived without bringing in quantity data.

Of course, this discussion leaves unsettled the question of how "closeness" to multiplicativity is to be measured - and how close one has to be before the CPD is a good approximation to the Geary Khamis. These are questions which will be considered in later chapters.

5. Implications of the Gerschenkron Effect

It is a well known feature of the GK method, when applied to aggregation up to GDP level that, if one country or a group of countries are dominant in overall volume terms, then the GK international prices will tend to be close to the prices of the dominant country or group. This fact, together with the commonly observed inverse relationship between relative prices and quantities, means that the Geary Khamis approach will tend to give lower aggregate volumes for the dominant country or group, compared to aggregation methods which define international prices as equi-weighted averages of country prices. The converse of this is that, in a sense, the aggregate price level for the dominant country or group will be higher under the GK method than under a method based

on equi - weighted international prices. This is the so called
Gerschenkron effect. This feature of the Geary Khamis method has
been the root of considerable debate between proponents and
opponents of the approach. It is not the purpose here to argue the
pros and cons of this aspect of the Geary Khamis method: however it
is important to note that the effect exists, and to consider some of
the implications.

The question then arises: suppose that the Geary Khamis approach, or
a version of it of the kind suggested in this chapter, is being used
to aggregate up to basic heading level, and that the standard Geary
Khamis method is then being applied again to aggregate up to GDP
level. Is there a danger that the Gerschenkron effect will enter
into the calculation of the basic heading level Geary Khamis PPPs,
and then be compounded in the second stage Geary Khamis aggregation
to GDP level, leading to a kind of overall "super" Gerschenkron
effect?

It is relevant to note, however, that the Gerschenkron effect cannot
arise at the level of first stage aggregation in the following
circumstances:-

(a) if either of the first two methods suggested above for imputing
quantities is being used: this is because the imputed quantities for
each country are scaled to the same total;

(b) if the structure of prices within the basic heading is
multiplicative, or almost multiplicative: this is because the

inverse relationship between prices and quantities cannot hold when prices are multiplicative.

Therefore, in the context of the " imputed quantities" suggested in section 3 as possible approaches to estimating Geary Khamis parities at basic heading level, the Gerschenkron effect is only potentially a problem when the third possible method is being employed - and only then if the structure of prices within the basic heading is non multiplicative. Whether there is evidence of a compounded Gerschenkron effect associated with the third imputed quantities method is one of the questions to be considered later on in this report, in the analysis of actual price data.

CHAPTER THREE

THE EKS AND RELATED METHODS

1.   Introduction.

1.1 In this chapter, the EKS method of aggregating purchasing power

comparisons up to a basic heading level is described. It is shown

how adopting a parametric formulation of the PPP estimation problem

can be used to give insights into the properties of the EKS

approach.   Finally, some alternative approaches, broadly similar to

the EKS approach, but likely to have superior estimating properties,

are suggested.

2.   The EKS Method : Description.

2.1 Suppose that price information, p_{ij}s, is available on items

within a basic heading. In addition, for each item, i, which is

priced in country j, it is known whether the item is

"characteristic" in that country.

This information on the characteristic status of items is used to define weights, r_{ij}, as follows:-

$$r_{ij} \quad = \quad 1 \text{ if item } i \text{ is characteristic in } j$$
$$= \quad 0 \text{ otherwise}$$

2.2 Given this information, the "Laspeyres" and "Paasche" indices are defined as follows, for any pair of countries j and k.

$$L_{jk} = \left[\prod_{i} \left(\frac{p_{ij}}{p_{ik}} \right)^{r_{ik}} \right]^{\frac{1}{\Sigma r_{ik}}}$$

$$P_{jk} = \left[\prod_{i} \left(\frac{p_{ij}}{p_{ik}} \right)^{r_{ij}} \right]^{\frac{1}{\Sigma r_{ij}}}$$

where, in these definitions, only items which are priced in both countries j and k are included in the construction of the indices.

Note that the Laspeyres index, L, is therefore defined if there exists at least one item which is priced in both countries, and is characteristic in country k. Similarly, the Paasche index, P, is only defined if there is at least one item which is priced in both countries, and is characteristic in country j.

2.3 When both the Laspeyres and Paasche indices exist, the Fisher index, F_{jk}, is defined to be the geometric mean of the Laspeyres and the Paasche : that is;

$$F_{jk} = [L_{jk} \cdot P_{jk}]^{1/2}$$

$$= \left[\prod_i \left[\frac{p_{ij}}{p_{ik}} \right]^{\left[\frac{r_{ik}}{\Sigma r_{ik}} + \frac{r_{ij}}{\Sigma r_{ij}} \right]} \right]^{1/2} \quad ..(1)$$

The Fisher index is thus <u>invertible</u>, in the sense that,

$$F_{jk} = [F_{kj}]^{-1}$$

but is not, in general, <u>transitive</u>; that is, for countries j, k and c

$$F_{jk} \cdot F_{kc} \neq F_{jc}$$

2.4 Suppose that, for a given group of countries, all possible Fisher indices have been formed, linking pairs of individual countries.

Then,

(a) this matrix will not in general be complete, in the sense that there may be some pairs of countries for which the Fisher index, as defined above, cannot be formed.

(b) The matrix of Fisher indices will not, as noted above, be transitive.

These problems can be overcome as follows:

Suppose that for country j and k, the Fisher index is not directly defined. A value for the Fisher index between j and k can be estimated by taking the geometric mean

$$\left[\prod_c [F_{jc} \cdot F_{ck}] \right]^{1/m}$$

over all countries c where F_{jc} and F_{ck} exist, where m is the total number of such countries. By applying this procedure, repeated several times if necessary, the matrix of Fisher indices can usually be completed, but the problem of non transitivity still remains.

2.6 A transitive index, the EKS index, can be obtained by applying the EKS formula to the complete Fisher matrix. The EKS purchasing power relativity between countries j and k, and denoted by E_{jk}, is defined as

$$E_{jk} = \left[\prod_c [F_{jc} \cdot F_{ck}] \right]^{1/n}$$

where n is the total number of countries. Here, F_{jj} and F_{kk} are taken to be equal to 1.

The EKS derived PPPs, defined in this way, then have the properties of invertibility and transitivity.

3.   The Properties of the EKS in the Light of a Parametric

      Approach to the Problem.

3.1 The EKS formulation, as described in the previous section, is

an example of the "index number" approach to purchasing power

comparisons.

A proponent of this kind of approach might well regard the formulae

as being reasonable ways of defining relative price levels, if

perfect and complete information were available on the prices of all

items. Viewed thus, the formulae stand in their own right,

independent of consideration of sampling and measurement error.

In the real world, however, consideration of the random elements of

the process of estimating PPPs is inescapable: items will normally

be sampled, and prices will be subject to random measurement errors

and within country variation.

It is therefore relevant to consider the statistical properties of

any given estimator — even if the ultimate rationale of that

estimator is based on an "index number" type approach.

One way in which insight can be gained into the statistical

properties of an estimator is to formulate the estimation problem in

parametric terms.

3.2 This approach is applied first of all to the Fisher index, defined in the previous section. Suppose that, for countries j and k, the price ratios

$$\frac{p_{ij}}{p_{ik}} \qquad \text{for items i priced in both countries}$$

are random variables with the properties that

$$\log\left[\frac{p_{ij}}{p_{ik}}\right] \qquad \text{are independent with mean } \mu_{jk} \text{ and variance } \sigma^2.$$

Another way of writing this is as follows:-

$$\log\left[\frac{p_{ij}}{p_{ik}}\right] = \mu_{jk} + \varepsilon_{ijk},$$

where the ε_{ijk} are independent error terms, with common variance σ^2. This is equivalent to

$$p_{ij} = \exp[\mu_{jk} + \varepsilon_{ijk}] \cdot p_{ik}$$

In other terms, the parameters, $\exp[\mu_{jk}]$, determine, in a sense, an average relativity between the prices in country j and country k – so motivating this particular choice of parametric model.

Assuming that this parametric model holds, the Fisher PPP, F_{jk}, between countries j and k, (equation 1) above, can be written

$$\log F_{jk} = \tfrac{1}{2}\Sigma\left[\frac{r_{ij}}{\Sigma r_{ij}} + \frac{r_{ik}}{\Sigma r_{ik}}\right]\log\left[\frac{p_{ij}}{p_{ik}}\right]$$

$$= \frac{1}{2} \Sigma \left[\frac{r_{ij}}{\Sigma r_{ij}} + \frac{r_{ik}}{\Sigma r_{ik}} \right] (\mu_{jk} + \epsilon_{ijk})$$

Hence,

$$E (\log F_{jk}) = \frac{1}{2} \mu_{jk} \Sigma \left[\frac{r_{ij}}{\Sigma r_{ij}} + \frac{r_{ik}}{\Sigma r_{ik}} \right]$$

$$= \mu_{jk} \qquad\qquad ...(2)$$

Thus, in the statistical sense, $\log F_{jk}$, is an unbiased estimator of μ_{jk}. Moreover, it is a straightforward matter to calculate the variance of $\log (F_{jk})$. We note that,

$$V [\log (F_{jk})] = V \left[\frac{1}{2} \Sigma \left[\frac{r_{ij}}{\Sigma r_{ij}} + \frac{r_{ik}}{\Sigma r_{ik}} \right] (\mu_{jk} + \epsilon_{ijk}) \right]$$

$$= \frac{\sigma^2}{4} \Sigma \left[\frac{r_{ij}}{\Sigma r_{ij}} + \frac{r_{ik}}{\Sigma r_{ik}} \right]^2 \qquad\qquad ...(3)$$

on using the fact that the ϵ_{ijk} are assumed to be independent.

Let s_1 denote the number of goods which are characteristic in j, and
    priced in k ;

  s_2 denote the number of goods which are characteristic in k, and
    priced in j;

  s_3 denote the number of goods which are characteristic in both j
    and k,

so that $s_1 = \Sigma r_{ij}$; $s_2 = \Sigma r_{ik}$; and $s_3 = \Sigma r_{ij} \cdot r_{ik}$

Then, equation 3 for the variance of the log Fisher index can be written as

$$V [\log(F_{jk})] = \frac{\sigma^2}{4} \left[\frac{1}{s_1} + \frac{1}{s_2} + \frac{2s_3}{s_1 \cdot s_2} \right] \quad \text{---(4)}$$

This variance expression therefore becomes smaller, that is, $\log(F_{jk})$ becomes a more precise estimator of μ_{jk}, the larger s_1 and s_2 are - in other words, the more characteristic items are involved in the comparison between country j and country k.

The fact that the variances of the log Fisher indices between different pairs of countries will, in general, be different, means, in effect, that different Fisher indices contain differing amounts of information about the underlying price level comparisons. This has very important theoretical implications for the EKS method - as will be seen later in this chapter.

3.3 The parametric approach can also be used to explore another topic of considerable potential importance - namely, the question of characteristic / non characteristic bias in the Fisher index.

It is a well known feature in price comparison work that there tends to be an inverse relationship between the price of a commodity, and the volume of that commodity traded: other things being equal, high volume commodities will tend to be cheaper, relative to low volume commodities. But since, by definition, characteristic goods will

generally be high volume goods, and non characteristic goods low volume, this suggests that there might be a systematic effect on prices, depending on the characteristic / non characteristic status of an item.

This suggests the following generalisation of the above model, to allow for such an effect; namely,

$$\log\left[\frac{p_{ij}}{p_{ik}}\right] = \begin{cases} \mu_{jk} + \alpha + \epsilon_{ijk}, & \text{if item i characteristic in k, not in j} \\ \mu_{jk} + \epsilon_{ijk}, & \text{if item i characteristic in both j and k} \\ \mu_{jk} - \alpha + \epsilon_{ijk}, & \text{if item i characteristic in j, not in k} \end{cases}$$

A systematic characteristic / non characteristic price effect of the kind postulated above would correspond to a positive value of α in the formulation of the model.

Substituting this expression for

$$\log\left[\frac{p_{ij}}{p_{ik}}\right]$$

into the equation for $\log [F_{jk}]$, it follows that,

$$\log [F_{jk}] = \frac{1}{2} \Sigma \left[\frac{r_{ij}}{\Sigma r_{ij}} + \frac{r_{ik}}{\Sigma r_{ik}}\right] \log\left[\frac{p_{ij}}{p_{ik}}\right]$$

and hence,

$$E[\log(F_{jk})] = \mu_{jk} + \frac{\alpha}{2}\left[\frac{\text{no. of items characteristic in k but not in j}}{\Sigma r_{ik}}\right. $$
$$\left. - \frac{\text{no. of items characteristic in j but not in k}}{\Sigma r_{ij}}\right]$$

$$= \mu_{jk} + \alpha/2 \left[\frac{s_2 - s_3}{s_2} - \frac{s_1 - s_3}{s_1} \right] \qquad ---(5)$$

in the notation introduced earlier in this section.

Equation (5) shows how, if there is a systematic characteristic/ non-characteristic effect on prices, then this works through into the log Fisher index in quite a complicated fashion, depending on the relative numbers of goods characteristic in k but not in j, and vice- versa.

It is, however, too simplistic to say that the term involving α in equation (5) represents a "bias" to the Fisher index. The term "bias" has connotations of error; there may be circumstances in which it would be desirable to correct for this effect, but, likewise, there will be other circumstances in which it is not appropriate to make any correction, even when the effect exists. This is a point which will be discussed in greater depth in the following chapter.

3.4 The next step is to consider the EKS formula itself, and to show how a parametric formulation of the problem can give additional insight into the characteristics of the formula. In fact, it can easily be shown that a transitive set of PPPs among n countries is formally equivalent to establishing n individual country price levels.

For example, given a set of transitive PPPs, written as P_{jk}, the price level for country 1 can be arbitrarily taken as 1, and price levels for country 2, ...,n as P_{21},.....,P_{n1} respectively. Then, any P_{jk} can be defined in terms of these price levels, since,

$$P_{jk} = \frac{P_{j1}}{P_{k1}} \quad , \qquad \text{by transitivity.}$$

Conversely, a set of n price levels, P_1, P_2, ...,P_n, obviously defines a transitive set of PPPs, by the relationship,

$$P_{jk} = P_j / P_k$$

Given, therefore, that the existence of a set of n underlying country price levels is implicit in the transitive set of PPPs produced by the EKS formula, it seems natural to adopt the following parametric formulation to describe the steps of obtaining a transitive set of PPPs from a matrix of Fisher indices. Namely,

$$\log (F_{jk}) = \mu_j - \mu_k + \epsilon_{jk}, \quad (j=1,n; k=1,n; j \neq k) \qquad ...(6)$$

where the μ_j and the μ_k are the logs of the desired country price levels.

The standard statistical approach to estimating the (μ) parameters

in a relation like (6) would be by the application of the technique of regression.

If it is assumed that the matrix of Fisher indices is complete, that is, F_{jk} exists for all comparisons of j and k, then it can readily be shown that the regression estimator of

$$\mu_j - \mu_k$$

is simply

$$\log\left[\prod_c F_{jc} \cdot F_{ck}\right]^{1/n}$$

Thus the regression estimator is identical with the log of the EKS estimator.

The proof is as follows:-

Proof

$$\text{Define} \quad \mu = 1/n \sum_{c=1}^{n} \mu_c$$

Then on averaging equation (6), holding j constant, it follows that,

$$1/n \sum_c \log F_{jc} = \mu_j - \mu + 1/n \sum_c \varepsilon_{jc} \quad, \text{ for all } j$$

Thus the regression estimator of $\mu_j - \mu$ is simply

$$1/n \sum_c \log F_{jc}$$

Thus the regression estimator of

$$\mu_j - \mu_k = 1/n \sum \log F_{jc} - 1/n \sum \log F_{kc}$$

$$- \log \left[\prod_c F_{jc} \cdot F_{ck} \right]^{1/n}$$

which is the desired result.

Starting from a complete matrix of Fisher indices, therefore, and under the stated assumptions on the error term, the EKS and regression approaches yield the same PPP estimates. Since, under the stated assumptions, regression estimates are optimal, in the sense of being minimum variance unbiased linear, this is an encouraging finding from the EKS viewpoint.

However, if the matrix of Fisher indices, as initially derived before filling in the gaps, is incomplete, a different picture emerges. In this situation, it will normally be possible, unless there are too many gaps in the matrix, to apply the regression techniques directly to the incomplete matrix - and this will again under the stated conditions, give optimal estimates. Further, it can be shown that these estimates will in general differ from the estimates obtained by filling in the gaps, as described in section 2.5, and then applying EKS. The exception is when only one gap exists, where the EKS and regression estimates will still coincide.

Since regression estimates are optimal under the stated conditions, the implication is that this technique of filling in the gaps and then applying EKS will in general give less precise estimates than the straightforward application of the regression model (6) to the incomplete matrix of log Fisher indices.

3.5 The position is, however, in general, even more complicated than this. It has been shown above that the terms log $[F_{jk}]$ will usually have unequal variance, as given by the expression in formula (4) above. Optimal application of regression to estimate the parameters in equation (6), therefore, would take account of this by fitting, not an unweighted regression, but a weighted Generalised Least Squares regression, with weights determined by equation (4). Thus, even when the matrix of Fisher indices is complete, the EKS formula, which is identical to an unweighted regression, is likely to be sub optimal relative to a weighted regression of the log Fisher indices. In these circumstances, the EKS estimator will still be unbiased, but will have sub optimal precision.

3.6 It is appropriate to summarise what has been a complicated argument :-

(a) when the matrix of first step Fisher links is complete, or almost complete, then the EKS approach is identical to an unweighted regression of the log Fisher indices.

(b) where the matrix of first step Fisher links is incomplete, the method of filling in gaps and applying EKS will not in general be equivalent to an unweighted regression of the log Fisher values. So if the assumptions for an unweighted regression hold, the EKS will be an unbiased, but sub optimal estimator.

(c) In any event, because the log Fisher variances are in general unequal, (equation 4), it will generally be appropriate to apply a weighted log Fisher regression.

Points (b) and (c) both suggest that the techniques of estimating PPPs by applying a weighted log Fisher regression is likely to give results which are close to, but more precise than, the straightforward EKS.

In addition, the weighted log Fisher regression model can be specified either in the straightforward form of equation (6), or, in a form which allows for the possible existence of characteristic / non characteristic bias, utilising the expression for the bias term in equation (5) above. The regression model allowing for bias would be

$$\log F_{jk} = \mu_j - \mu_k + a/2 \left[\frac{s_2 - s_3}{s_2} - \frac{s_1 - s_3}{s_1} \right] + \epsilon_{jk}$$

where s_1, s_2, and s_3 are defined in terms of the numbers of items characteristic in countries j and k, as in section 3.2 above.

3.7 One final cautionary remark should be made. Consider two direct Fisher indices, F_{jk} and F_{jc} say, both involving country j. It is possible, indeed it is likely, that some of the items appearing in

the numerator of these indices (priced in country j), are the same in both indices. In this case, there will be correlation between these two indices. Thus, in applying regression to estimate the parameters in equations (6) or (7) allowance should strictly be made not just for unequal error variances, but for potential correlation between the errors. This is a problem which applies not just to the log Fisher regression methods, but also to the EKS itself, given the equivalence between the EKS and the log Fisher regression under given circumstances. This suggests the desirability of considering other methods which get round the error correlation problem. One such method is considered in the next chapter.

CHAPTER FOUR

THE COUNTRY PRODUCT DUMMY TECHNIQUE.

1.   Introduction.

1.1 The Country Product Dummy (CPD) approach is the third of the standard PPP techniques discussed. The next section introduces the technique. Following that, there is a description of how the technique can be generalised to allow for characteristic / non characteristic product bias. The theoretical question of when characteristic / non characteristic bias ought to be corrected for is also discussed.

2.   The CPD Technique.

2.1 The CPD is a parametric technique, which depends only on knowledge of price information, p_{ij}s, without any information on quantities, or the characteristic/non characteristic classification of items. For reference, see Kravis et al; [7,8] and Hill, [11].

Suppose that there are n countries, and m items, and that the price of item i in country j is p_{ij}. It is not necessary that every item be priced in every country.

The basic assumption of the technique is that $\log p_{ij}$ can be expressed in the form

$$\log p_{ij} = \alpha_i + \beta_j + \epsilon_{ij} , \quad ----(1)$$

where α_i is a price effect parameter for item i;

and β_j is a country effect parameter for country j.

Under the assumption that the errors ϵ are uncorrelated, with common variance σ^2, the parameters are estimated by the application of standard, unweighted regression. To avoid parameter redundancy, a single constraint has to be imposed on the parameters: without loss of generality, this could be taken to be the condition that $\beta_1 = 0$. This would have the effect of expressing the price effect terms, the α parameters, in the currency of country 1.

Once the β parameters have been estimated, the PPP as determined by the CPD method between any two countries, j and k say, is given by

$$C_{jk} = \exp \{ \beta_j - \beta_k \}$$

These CPD derived PPPs are thus clearly invertible and transitive.

The CPD technique has major attractions:
* it is simple, and straightforward to apply;
* being a parametric approach, it has a good underlying rationale;
* unlike the EKS and log Fisher regression techniques, there is

no a priori reason to suppose that correlation between the errors in the regression equations is likely to be a problem.

However, the CPD approach also has limitations. In particular:-

(a) implicit in the approach is the assumption that the price structure being analysed is reasonably approximated by the product of item and country price effects: i.e., that the prices are basically multiplicative. This may not be a reasonable assumption in particular cases.

(b) the model fitted is relatively parameter intensive. That is, it involves $(n+m-1)$ price or country effect parameters plus one variance parameter. Unless at least $(n+m)$ price observations are available, the model cannot be fitted at all: and even where more than $(n+m)$ prices are available, precision may be poor if the number of observed prices is close to $(n+m)$.

(c) The approach makes no use of either available quantity information, or of any information which may be available on the characteristic / non characteristic classification of items. In fact, this statement is not quite true : versions of the technique have been developed which apply overall quantity weights, per item, in estimating the basic equations. But this involves weighting all countries similarly for a given item - and so represents only a very partial use of quantity data.

3. **Extension of the CPD Technique to allow for Characteristic / Non Characteristic Bias.**

3.1 As noted in the preceding section, the standard CPD technique makes no allowance for characteristic / non characteristic bias. It is not difficult, however, to see how the basic CPD model could be extended to allow for the possibility of a differential price between characteristic and non characteristic products, if information on the characteristic / non characteristic classification of items is available. The required adjustment to the basic CPD equation is

$$\log (p_{ij}) = \alpha_i + \beta_j + \tau z_{ij} + \epsilon_{ij} \qquad ----(2)$$

where $z_{ij} = 1$ if item i is non characteristic in country j;
$= 0$ if item i is characteristic in country j.

If a characteristic / non characteristic pricing differential exists, as postulated in an earlier chapter, then this would show up as a positive coefficient of z in this equation.
There are no technical difficulties about estimating the coefficients in equation (2) by standard regression - although occasionally, in practice, circumstances may occur when the τ parameter is not identified: for example, if all the items in country j are non characteristic, and no items are non

characteristic in other countries, then the τ parameter would be confounded with the β_j country parameter.

Two models have thus been developed - the log Fisher regression with bias term, and the CPD with non characteristic dummy, which could be applied to real life data to test whether significant non characteristic bias exists in practice. Which model should be preferred?

As noted above, there is a potential problem of error correlation in the log Fisher regression; this is liable to distort tests of statistical significance. Because of this, the detailed analysis of non characteristic bias reported in a later chapter is based on the CPD approach.

The CPD model with non characteristic dummy offers a convenient framework for detecting the existence of a differential price effect for non characteristic items. However, the question still remains of whether, in calculating the final inter country PPPs any correction should in principle be made for non characteristic bias. This is not a simple question to answer - indeed, each case probably has to be considered on its own merits. This is illustrated by considering two hypothetical examples.

Example 1.

Consider two countries, j and k, with five items priced in each country. Suppose that all items are characteristic in country j, but four are non characteristic in country k. In country k, however, these

non characteristic items are of minor importance, accounting for
only 10% of the total volume of trade in that basic heading.

Example 2.

Suppose that the situation is the same as in example 1, except that
now the four non characteristic items each account for 15% of trade
in country k, that is 60% of trade in all.
If there is a significant non characteristic price differential, how
should this be allowed for in these two cases?

It is fairly clear that, in example 1, the effect of the non
characteristic price differential should be removed, or largely
removed, in calculating country k's price level. 90% of the trade
in country k is in the characteristic product - so the average price
level in country k will be largely determined by the price of that
product. It would give a distorted estimate of the average level of
prices actually paid in country k to give equal weight to four
prices including the non characteristic price effect, and one price
excluding it. Yet this is what the standard CPD regression would do.

On the other hand, in the case of example 2, 60% of the volume
traded in the country includes the non characteristic differential.
So the average price of goods actually traded will be substantially
influenced by the non characteristic effect. In this case
application of the CPD model without the non characteristic dummy
would probably give an estimated country effect for country k which
would be closer to the true average price.

The optimal answer to the problem would presumably involve running the CPD model with the non characteristic dummy, and then, for each country, adding to the estimated country effect a fraction of the estimated non characteristic coefficient – that fraction depending on a separate assessment of what proportion of trade in the country is accounted for by non characteristic products.

CHAPTER FIVE

DATA ANALYSIS

1.   Data Analysis : Preliminaries.

1.1 For those OECD countries not within the EEC, price and
expenditure data were collated by the OECD. In the first instance,
this data was collected by individual member government statistical
departments. In all, price and expenditure data were analysed for
the following ten countries; Austria, Finland, Norway, Sweden,
Australia, New Zealand, Japan, Canada, United States and Turkey,
numbered in the analysis, 1 to 10 respectively.

Price Data

Price information is collected at individual item level. Each
country in the group supplied information on the specification of
the commodity, for example, name of product and how packaged;
together with its price, and detail on whether the good was
characteristic or not. In general, several items would be priced
within each basic heading. For the purposes of this study, the
price collection methods are taken as given.

Expenditure Data

Briefly, expenditure data was collected by the OECD, at Gross
Domestic Product level, for each of the countries in the study.
Expenditure figures are in national currencies (millions). Sub
divisions of expenditure data were also provided, both at a fairly
aggregate level, such as Consumer Expenditure, and also at basic
heading level. The latter is more important for the current study.
Basic heading level is defined as the smallest grouping for which
individual country expenditure data is available. For this study,
the price data for 37 product groupings, falling into 34 basic
headings, were examined. The following is a list of the products
studied:

111.1 Rice
111.3 Bread
111.41 Rusks
111.51 Noodles
111.6 Other Cereals
112.1 Beef
112.4 Lamb and Goat
112.6 Sausages
112.7 Other Processed Meat
113.1 Fish
113.4 Preserved Fish
114.1 Milk
114.2 Preserved Milk
114.3 Cream
114.4 Cheese
116.11 Citrus Fruit
116.12 Apples   ] Fresh Fruit
116.2 Dried Fruit
116.3 Frozen Fruit
116.41 Vegetables
116.42 Salad Veg.   ] Fresh Vegetables
116.43 Root Veg.
116.7 Preserved Vegetables
117.2 Potato Products
118.1 Sugar
119.1 Coffee
119.2 Tea
1110.1 Jams
1110.2 Chocolate
1110.3 Sweets
1110.5 Condiments

121.2 Soft Drinks
131.1 Alcohol
131.2 Wine
131.3 Beer
131.4 Aperatifs
141.1 Cigarettes

2. Outline of the PPPs Calculated for this Study.

2.1 For each of the 37 product groupings selected for detailed study, 8 different methods were applied for aggregating the item level price data up to a PPP for each basic heading. The methods employed fall into three main groups, as follows :-

(a) The Geary Khamis Group.

Three methods of calculating Geary Khamis indices were used, each involving imputing quantities, as described in Chapter 1 , and then solving the resulting full Geary Khamis equations. The three techniques differ in the methods used for imputing quantities : the procedures adopted are those outlined in Chapter 1. The resulting PPPs have been denoted GK1, GK2 and GK3.

GK1 : this involved imputing the quantities q_{ij} as

q_{ij} = 1   if item i priced in country j

     = 0   otherwise,

and then scaling the resulting q 's for each country so that the country totals are the same.

GK2 : for this, the quantities q_{ij} are imputed initially as

q_{ij} = 2   if item i characteristic in country j:

     = 1   if item i priced but non characteristic:

     = 0   otherwise.

As for GK1, the resulting q 's are then scaled to the same total within each country.

GK3 : as for GK2, but country totals are scaled to a prior estimate of the overall GDP volume relativities between the countries.

The volume relativities assumed were :-

Austria :   1.76

Finland :   1.18

Norway: 1.18

Sweden: 2.94

Australia: 3.53

New Zealand: .59

Japan: 22.9

Canada: 5.89

United States: 58.82

Turkey: 1.18

These percentages were the approximate shares of each country in the total GDP of the 10 countries determined from Purchasing Power Parities and Real Expenditures, 1985, as published by OECD.

(b) The EKS Group

The three methods used were :-

E   : the standard EKS approach.

F1 : the log Fisher regression approach, without bias term; the regression was weighted by the inverse of the square root of the theoretical log Fisher variance, to correct for the problem of heteroscedasticity.

F2 : the log Fisher regression with bias term to allow for a possible characteristic / non characteristic price differential: again, this regression is weighted to correct for heteroscedasticity.

See Chapter 3 for a description of the underlying theory for these approaches.

(c) The CPD Group.

C1 : CPD regression without bias term;

C2 : CPD regression with bias term;

See Chapter 4 for a discussion of the relevant theory.

As noted above, each of these 8 methods for calculating PPPs was applied for each of the 37 product groupings selected for detailed study. In fact, in a few cases it was not technically possible to calculate the versions of the log Fisher and CPD PPPs involving bias terms – for example, because all of the items priced were characteristic in every country.

3. The Results.

Table 5.1 gives examples of the PPPs obtained by these different

methods for a number of basic headings. In each case the PPPs are

relative to the United States (Country 9): that is, for example for

preserved fish the price relative of Austria to the US using the EKS

method is 9.6 Austrian schillings to the US dollar.

4. Summary Statistics Used in the Analysis.

4.1 Before embarking on the description of the data analyses, in

the subsequent chapters, it is necessary to develop some summary and

descriptive statistics, which will be used as tools in the data

analyses.

(a) A Measure of the Difference between the PPPs.

The first requirement is to have a simple and convenient means to

assess how close one PPP is to another. Suppose, for example, that,

8 different PPPs are calculated for a particular basic heading,

using the methods outlined in the previous section. It is then

natural to ask questions like - is the F1 index or the F2 index

closer to the EKS index ? - and so on.

To answer this kind of question, a measure of the distance between

two PPPs is required. But since each PPP consists of a set of 9

TABLE 5.1 EXAMPLES OF PPPs AT BASIC HEADING LEVELS

All PPPs in this table are shown relative to the United States

A.　PRESERVED FISH

Country Method	Aus	Fin	Nor	Swe	Australia	NZ	Jap	Can	Turkey
EKS	9.60	10.16	11.28	9.69	0.88	1.26	312.10	1.13	328.30
LOG F with bias	8.92	9.65	11.33	9.77	0.90	1.23	302.62	1.10	325.46
LOG F w'out bias	9.71	10.00	11.38	9.76	0.80	1.24	313.01	1.15	332.67
GK1	12.32	11.27	11.45	11.49	0.91	1.30	343.26	1.37	341.53
GK2	11.76	11.63	11.65	11.58	0.92	1.30	330.98	1.26	345.51
GK3	10.92	10.64	10.44	10.69	1.03	1.26	326.34	1.20	294.98
CPD with bias	9.83	10.20	11.48	10.54	0.93	1.22	309.90	1.19	336.30
CPDw'out bias	11.70	10.63	11.49	10.40	0.88	1.24	329.27	1.27	350.45

B.　CHEESE

Country Method	Aus	Fin	Nor	Swe	Australia	NZ	Jap	Can	Turkey
EKS	15.91	5.01	7.42	8.65	0.72	1.14	236.67	1.31	411.34
LOG F with bias	16.28	5.16	7.67	8.70	0.73	1.20	234.06	1.37	424.83
LOG F w'out bias	15.80	5.06	7.52	8.68	0.71	1.16	237.79	1.33	416.49
GK1	16.12	4.56	7.12	8.98	0.70	1.25	341.13	1.31	411.62
GK2	16.41	4.68	7.13	8.66	0.70	1.26	334.46	1.33	411.07
GK3	14.88	3.56	5.91	7.28	0.68	1.17	265.29	1.42	363.28
CPD with bias	17.19	5.28	8.06	9.36	0.73	1.29	249.48	1.44	446.06
CPDw'out bias	15.85	4.79	7.45	9.07	0.66	1.16	263.68	1.24	407.96

C.　JAM

Country Method	Aus	Fin	Nor	Swe	Australia	NZ	Jap	Can	Turkey
EKS	15.99	7.60	12.06	7.56	0.75	1.06	345.29	1.47	296.48
LOG F with bias	15.70	7.59	12.01	7.62	0.75	1.07	346.88	1.47	295.99
LOG F w'out bias	16.02	7.59	11.96	7.59	0.75	1.06	344.92	1.47	295.99
GK1	16.24	7.40	11.33	7.56	0.73	1.03	335.66	1.40	288.65
GK2	15.93	7.40	11.34	7.56	0.73	1.03	336.63	1.40	288.65
GK3	16.22	7.60	11.54	7.70	0.75	1.05	343.96	1.40	296.33
CPD with bias	15.48	7.44	11.37	7.63	0.74	1.05	341.96	1.41	290.13
CPDw'out bias	16.30	7.44	11.28	7.57	0.74	1.03	335.87	1.41	290.13

price levels relative to a base country, (or, equivalently, of the set of 45 possible inter country comparisons), it is not obvious how such a distance measure between PPPs should be defined.

The following argument suggests a possible approach. Consider two PPPs - denoted F and G say. For each PPP, construct the vector consisting of all 45 possible inter country comparisons : (there are 45 comparisons between 10 countries). F and G will be identical PPPs if the two vectors are the same - or equivalently, if the vector obtained by dividing each term in the F vector by the corresponding term in the G vector is a vector of ones. This suggests that a measure of the distance between F and G might be based on how great the dispersion is between the terms in the vector F/G .

A standard measure of the dispersion of a set of numbers is the sample variance. Thus an initial distance measure might be taken as the variance of F/G.

Note: here the sample variance of a vector ($x_1, \ldots x_k$) is defined as

$$\frac{1}{(k-1)} \Sigma (x_i - \bar{x})^2$$

However, a little thought indicates that this measure of distance is deficient, in that the variance of (F/G) is not in general equal to the variance of (G/F); and clearly it is desirable in defining a distance measure that the distance from F to G should be the same as the distance from G to F.

A simple modification to the definition gets round the problem -
namely, taking the variance of log(F/G).
This follows since

$$\log(F/G) = - \log(G/F)$$

and hence

$$\text{var } [\log(F/G)] = \text{var } [\log(G/F)]$$

Thus the distance measure D between F and G is defined as

$$D(F,G) = \text{var } [\log(F/G)]$$

or equivalently,

$$D(F,G) = \text{var}[\log F - \log G]$$

It would be possible to define other distance measures between PPPs
which fulfill a similar function - however, the distance measure
defined above is certainly adequate for the purpose intended.

It is important to have a feel for the magnitude of the differences
between PPPs implied by different values of the measure. Such a feel
can be supplied by the following heuristic argument.

Roughly speaking, about 95% of a given set of observations are
likely to fall within a range of $\pm 2 \sqrt{\text{(sample variance)}}$ of the mean
of the observations. This suggests that, for two PPPs, F and G,
about 95% of the terms in the vector log(F/G) will fall within a

range of $\pm 2 \sqrt{[D(F,G)]}$.

Or, in other words, about 95% of the terms in the vector F/G can be expected to fall in the range

$$[\exp\{-2\sqrt{D}\} , \exp\{+2\sqrt{D} \}].$$

The right hand limit of the approximate 95% confidence interval lies, in percentage terms, at $[\exp(2\sqrt{D})-1] * 100\%$ above G. For example, suppose that, for particular PPP vectors F and G, the corresponding D value is 0.01, then the above approximate 95% confidence interval for the terms in F/G is

$$[\exp\{-2\sqrt{(0.01)}\} , \exp \{+2\sqrt{(0.01)}\}]$$

that is,

$$.819 \leq \frac{F}{G} \leq 1.22$$

with approximately 95% confidence.

Expressing this another way, approximately 95% of the terms in F will be expected to lie in the interval

$$.819G \leq F \leq 1.22G$$

In other words, the right hand bound of the confidence interval is 22% above G.

Fig 5.1 Plot of $\left\{Exp(2\sqrt{D})-1\right\} \times 100$ as a function of D
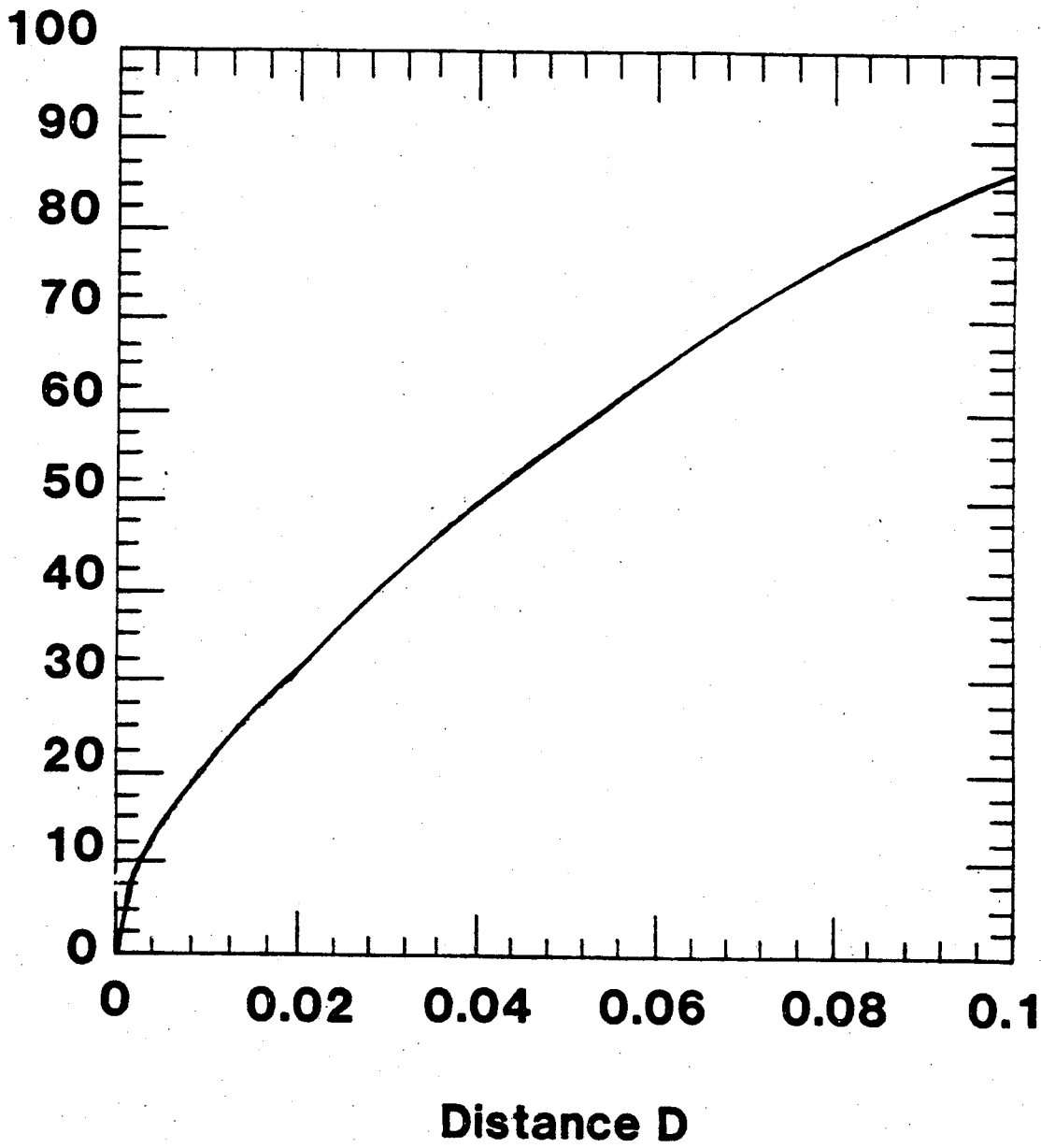

Distance D

Figure 5.1 plots [exp{2√D} - 1] * 100 as a function of D.

From the chart, it can be seen that, for a distance D of .05 between

the PPPs, approximately 95% of inter country comparisons between the

different PPPs are likely to lie within 60% of each other. Putting

this another way, for a D value of .05, it might be expected that 2

or 3 inter country comparisons out of the 45 possible for 10

countries, would differ by 60% or more, between the different

versions of the PPPs. This is certainly not very close. It can be

seen that, roughly speaking, for inter country comparisons to differ

by less than 10%, D values of .003 or smaller are required.

(b) A Measure of the Degree of Multiplicativity for a Set of Prices.

It was hypothesised, in Chapter 1, that the closer the set of prices

(p_{ij}) corresponds to being multiplicative, then the less would the

Geary Khamis PPP depend on the particular quantities (q_{ij}), and the

closer the Geary Khamis PPP would converge to what is, in effect,

the CPD PPP. This begs the question, however, of how closeness to

multiplicativity is to be defined, for a given set of prices.

Consideration of the CPD regression model suggests a natural measure

of closeness to multiplicativity. If the price structure is exactly

multiplicative, then there will be an exact fit in the CPD

regression, and the coefficient R^2, (the coefficient of multiple

correlation for that regression) will be 1. This suggests taking

the R^2 value from the CPD regression as the measure of closeness to

multiplicativity.

R^2 itself, however, suffers from one potential drawback. In a CPD regression with n countries, and m items, an apparently very good fit (high R^2) will tend to be observed, if there are too few price observations in fitting the regression. That is, if the number of observed prices is not much greater than m+n. To attempt to get round the problem of spuriously good fit, when observations are too few, a better measure of multipicativity is R^2 adjusted for degrees of freedom, conventionally denoted by $\bar{R}^2$. Formally, if a regression is being fitted involving r observations and s parameters: and if the total sum of squares of the dependent variable about its mean is denoted TSS, and the regression residual sum of squares is denoted RSS, then

$\bar{R}^2$ is defined as

$$\bar{R}^2 = 1 - \left[\frac{RSS}{r-s} \right] / \left[\frac{TSS}{r-1} \right]$$

Like R^2 , $\bar{R}^2$ from the basic CPD regression will equal 1 if the price structure is exactly multiplicative. The closeness of $\bar{R}^2$ to 1 can be taken as a reasonable measure of how multiplicative the price structure is.

$\bar{R}^2$ cannot be used as an absolute measure of multiplicativity, to compare price structures between different sets of countries, since it will be dependent, for example, on the absolute price levels in the different countries - which in turn are dependent on the

particular currencies in use in the different countries. But for a
fixed set of countries, $\bar{R}^2$, is a satisfactory means of comparing
relative closeness to multiplicativity, for different sets of
prices.

(c). Internal Homogeneity of a Family of PPPs.

As noted above, the 8 PPPs which have been calculated fall into
three basic families - the Geary Khamis family, the EKS family, and
the CPD family. It is useful to have a summary statistic to describe
how internally homogeneous each family is. For a given basic
heading. the internal homogeneity of a particular family of PPPs is
taken to be the largest distance between any two PPPs in the family,
using the D measure of distance defined at (a) above.

(d). Closeness of Different PPP Families.

Similarly, for two different families of PPPs, it is useful to have
a summary statistic measuring how close the families are to each
other. A measure of the closeness of two families is taken to be the
smallest distance, in terms of the D measure, between any two PPPs
in the different families.

CHAPTER SIX

ANALYSIS OF CHARACTERISTIC / NON CHARACTERISTIC BIAS.

1.   Introduction

1.1 In previous chapters, it was noted that there may be a

consistent price differential in the prices of non characteristic

goods; that is, a characteristic / non characteristic bias.

In the discussion of the CPD approach in Chapter 4, it was suggested

that the CPD regression model, with non characteristic dummy, would

be an appropriate model to investigate this phenomenom. This chapter

reports on the results of applying the CPD model to the price data

described in the previous chapter for the purposes of studying the

characteristic / non characteristic pricing differential. Two main

questions are considered:

(a) Is there statistically significant evidence that a differential

pricing effect exists, of the kind hypothesised?

(b) If so, is the magnitude of the effect such that it has a

potentially important operational effect: that is, is it liable to

make any real difference to calculated PPPs?

1.2 Table 6.1 shows the following information for the 33 out of the

37 basic headings for which it was possible to estimate the CPD
model with non characteristic dummy.

(a) the coefficient of the non characteristic dummy estimated from
the regression. It will be recalled that a positive coefficient
corresponds to a differential price effect of the direction
hypothesised.

(b) the "t" values associated with the coefficient.

(c) the values of [exp(dummy coefficient) - 1]*100 : this is the
estimated percentage effect on the price of an item through
being non characteristic.

(d) the distance D, as defined in Chapter 5, between the CPD PPPs,
calculated with and without the non characteristic dummy variable in
the regression.

1.3 It will be seen, from the first two columns of the table, that,
out of the 33 product groupings, the coefficient of the non
characteristic dummy variable is positive in 26 cases, negative in
7. On the basis of the "t" statistic, five of the coefficients are
significantly different from 0 at the 5% level- and in each of these
cases the coefficient is positive. These five commodities are
cheese, preserved vegetables, condiments, alcohol and beef. The
disproportionate number of positive coefficients, both overall, and
among the individually significant items, suggests that there may
well be an effect at work of the form hypothesised - although the
reader may well not find the evidence yet conclusive. However, a
more refined analysis of the data shows that this conclusion can
indeed be strengthened.

TABLE 6.1   CHARACTERISTIC/NON CHARACTERISTIC BIAS

ITEM	COEFFICIENT OF NON CHAR DUMMY	t VALUE	% EFFECT ON ITEM PRICE	D MEASURE BETW. CPD WITH & W'OUT BIAS
RICE	.127	.44	13.5417	.0005
BREAD	.064	.42	6.6092	.0006
OTHER CEREALS	.017	.11	1.7145	0.
LAMB	.088	.73	9.1988	.0013
SAUSAGES	-.107	-.9	-10.1474	.0005
PRES MEAT	.068	.3	7.0365	.0007
FISH	.061	.43	6.2899	.0005
PRES FISH	.303	1.59	35.3915	.0074
CHEESE	.282	3.59	32.5779	.007
CITRUS FRUIT	.572	1.33	77.1807	.0037
APPLES	.088	.21	9.1988	0
DRIED FRUIT	.195	.78	21.5311	.0056
FROZ FRUIT	.144	1.14	15.4884	.0011
VEGETABLES	-.025	-.11	-2.469	0
SALAD VEG	-.117	-.28	-11.0415	.0002
ROOT VEG	.485	1.77	62.4175	.0075
PRES VEG	.257	2.1	29.3045	.0047
COFFEE	-.247	-1.02	-21.8859	.004
TEA	.089	.59	9.308	.0005
JAM	.18	2.41	19.7217	.0007
CHOCOLATE	-.014	-.07	-1.3902	0
SWEETS	.104	.51	10.96	.0011
CONDIMENTS	.562	2.27	75.4177	.0112
ALCOHOL	.323	2.49	38.1265	.0161
WINE	-.318	-1	-27.2397	.001
BEER	.184	.75	20.2016	.0132
RUSKS	.142	.81	15.2577	.0017
NOODLES	.034	.16	3.4585	.0001
BEEF	.403	3.57	49.6307	.0161
CREAM	-.376	-1.18	-31.3398	.0085
SOFT DRINKS	.062	.25	6.3692	.0006
APERITIFS	.149	.5	16.0673	.0019
CIGARETTES	.148	1.05	15.9513	.0036

For all of the regression results recorded in the table, a probability value (P) can be associated with the observed t - statistic, as follows:

Let T be the value of the "t" statistic actually observed, and let z be a random variable with the appropriate "t" distribution possessed by T, under the null hypothesis that the coefficient of the non characteristic dummy is actually zero.

Then define

$$P = P\{z \leq T\}$$

In other words, P is the probability of obtaining a value of the "t" statistic $\leq$ the value actually observed, under the null hypothesis that the coefficient of the non characteristic dummy is actually zero.

Under this hypothesis, the 33 P values associated with Table 6.1 would be equivalent to 33 independent observations on the uniform distribution over the interval [0,1]. It is appropriate, therefore, to apply the Kolmogorov Smirnov test to this set of P values, to see whether they could indeed reasonably be regarded as being drawn from this particular distribution. Figure 6.1 shows the comparison of the two distributions. It will be seen that the empirical distribution is shifted uniformly to the right compared with the theoretical; the value of the Kolmogorov Smirnov statistic is indeed significant, at higher than the 0.3% level. Viewed thus, therefore, there is strong evidence of the existence of a positive differential effect on the prices of non characteristic items.

Fig 6.1 Empirical and Fitted Cumulative Distribution Functions
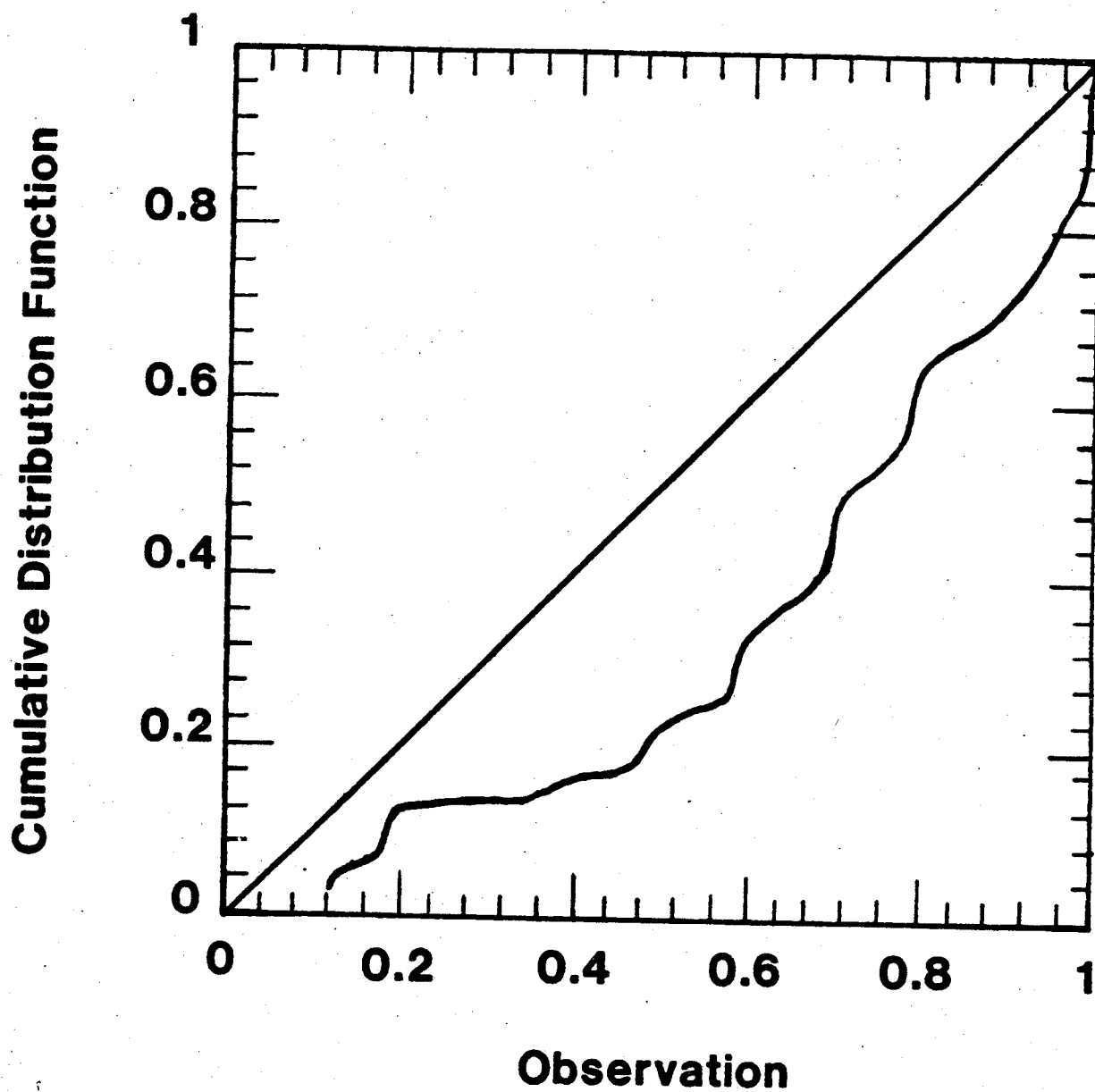
1.4 Having established that there is strong evidence for the
existence of characteristic / non characteristic bias - the next
question is - does it potentially make an important difference to the
PPP comparisons? Column 3 in Table 6.1 shows the estimated effect on
individual prices of being characteristic as a multiplicative
factor. The largest effects, (for condiments, and citrus fruits) are
indeed large - of the order of 75%. Most of the effects are much
smaller - typically around 10% on individual prices. Nevertheless,
this is large enough to be of potential concern.

1.5 Column 4 of Table 6.1 illustrates what is perhaps a more
relevant measure of the size of this effect. This is the distance,
in terms of the distance measure D, between the CPD PPPs calculated
with and without the non characteristic dummy. It will be noted that
the large coefficients in column 1 do not always correspond to large
D values: this, in fact, is not unexpected and can arise if there
are, proportionately, only a small number of non characteristic to
characteristic products in the basic heading.

The largest D values occur for condiments, alcohol, beef and beer,
where the D value is greater than .01 in each case. For D values of
this magnitude, it would be anticipated (see Fig.5.1 in Chapter 5)
that some inter country PPP comparisons would be affected by 20% or
more - and this is indeed the case. For example for condiments, the
estimated Turkey / Austria PPP for the C2 index ,that is the CPD
with bias, is 14.8, but the CPD without bias, C1, is only 11.4, - a

difference of 30%. In fact, three out of the five condiments priced in Austria are non characteristic - so it is not surprising that correcting for non characteristic bias naturally lowers the Austrian price level relative to other countries.

Similarly, for beer, there is an effect approaching 20% on some of the inter country comparisons involving Austria (five out of the seven beers priced in Austria are non characteristic); and over 15% on some of the comparisons involving Japan, (three out of five beers priced in Japan are non characteristic).

These examples illustrate that the effect of differential prices for non characteristic items can indeed be of potentially major importance for particular basic headings. For most of the basic headings considered, however, the apparent effect is much smaller: that is the D values observed would typically imply differences of a few percentage points in inter country PPPs.

1.6 To summarize therefore:-

(a) There is evidence of a significant positive differential effect on the prices of non characteristic items.

(b) The magnitude of this is such that, for some basic headings, it could potentially have a very serious effect in distorting inter country PPP comparisons.

The problem of non characteristic bias is shown here to be one that has to be taken very seriously. It is relevant, however, to recall the theoretical discussion at the end of Chapter 4. In particular, each case really has to be considered on its merits: there are no a priori grounds for regarding either of the CPD PPPs as being intrinsically superior to the other.

CHAPTER SEVEN

INTERNAL HOMOGENEITY OF PPP GROUPINGS AND BETWEEN GROUP DISTANCES.

1. Introduction.

1.1 This chapter continues the detailed analysis of the data for 37

product groupings, looking at the internal homogeneity of the three

basic PPP groupings, and the inter group distances. It will be

recalled that the necessary concepts for this analysis, for example,

inter PPP distances etc., were defined in Chapter 5.

2.   The Internal Homogeneity of the Geary Khamis Family.

2.1   When the theory of the Geary Khamis method was discussed in

Chapter 1, it was noted that, where the price structure is exactly

multiplicative, then the Geary Khamis PPP is independent of the

quantities. This suggests that, in real life, as the price structure

becomes more multiplicative, the different versions of the Geary

Khamis defined in Chapter 2, which differ only in their imputed

quantities, should come closer together. In other words, as $\overline{R}^2$, the

measure of price structure mulitplicativity defined in chapter 5,

increases, the internal homogeneity of the Geary Khamis grouping should increase.

Figure 7.1 shows $\bar{R}^2$ on the horizontal axis plotted against the measure of internal homogeneity for the Geary Khamis group, defined in Chapter 5. Recall that internal homogeneity is measured by the largest within group distance : so the smaller this measure is, the greater the degree of internal homogeneity.

Examination of Figure 7.1 suggests that there is a relationship of the form suggested - but that it is fairly weak. All basic headings where $\bar{R}^2$ is > .99 have a homogeneity measure of .01 or less: and below an $\bar{R}^2$ value of .99, there is, apart from one notable outlier, discussed in more detail below, the suggestion of a downward slope to the upper bound of the scatter of plotted points. On the other hand, the maximum observed value of the homogeneity measure rapidly becomes very large as $\bar{R}^2$ falls below .99. Thus, it is clear that, for the bulk of the basic headings in the sample, the degree of price multiplicativity ($\bar{R}^2$) is not in itself such as to place any very stringent bound on the expected homogeneity of the Geary Khamis family.

Further examination of the marked outlier towards the top right of Figure 7.1 is of interest. This product grouping is, in fact, noodles: and within this group, there are only four items which are priced within the ten countries considered. For three countries, Australia, Canada and Turkey, only one item is priced in each country. The large value of the Geary Khamis internal homogeneity

Fig 7.1 GK Internal Homogeneity versus R (BAR) Squared

R(BAR) Squared (X 0.01)
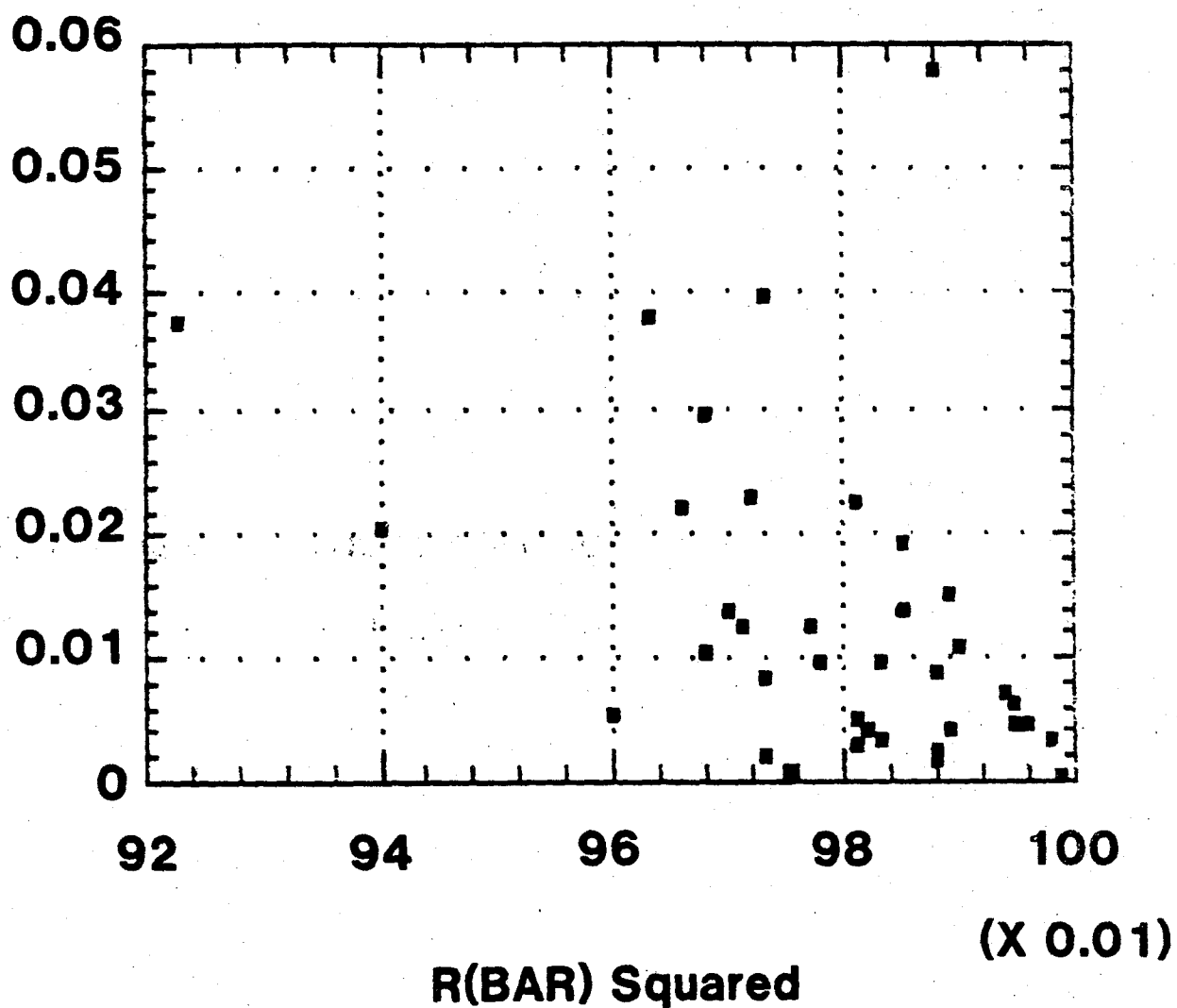
measure for this group is largely accounted for by the Turkey / Canada price comparison which has a value of 117.7 for GK1 and 196.7 for GK3 - a difference of 67%. This large difference arises because of two contributory factors :-

(a) First of all, the price structure for items 4 and 5 is very non multiplicative between Japan and the USA :

	JAPAN	USA
ITEM 4	601	.83
ITEM 5	250	.83

Largely because of this, the international price of item 5 is much higher under GK3 (which attaches relatively more weight to the USA) than under GK1 . The GK1 and GK3 international prices of item 5 are 12.85 and 18.33 respectively, in Austrian schillings.

(b) The second factor at work is that only one item is priced in each of Turkey and Canada: item 1 in Turkey, item 5 in Canada. So if Turkey is denoted by the suffix T and Canada by the suffix C, it follows that

$$GK_{TC} = \frac{\sum p_{iT} \, q_{iT}}{\sum \pi_i q_{iT}} \bigg/ \frac{\sum p_{iC} \, q_{iC}}{\sum \pi_i q_{iC}}$$

$$= \frac{p_{1T} \, q_{1T}}{\pi_1 q_{1T}} \bigg/ \frac{p_{5C} q_{5C}}{\pi_5 q_{5C}}$$

$$\bullet \quad \frac{P_{1T}}{P_{5C}} \cdot \frac{\pi_5}{\pi_1}$$

Thus the change in the Turkey / Canada comparison between the GK1 and GK3 versions depends on the change in the ratio π_5/π_1 from GK1 estimated prices to GK3 estimated prices – and this is in fact very large mainly due to the instability of π_5 already noted at (a).

Study of the noodles outlier, therefore, suggests two important points.

(i) Despite an apparently high $\bar{R}^2$ (.988 in this case), the structure of international prices may be very non multiplicative – and this can have a significant effect on the Geary Khamis methods proposed.

(ii) When small numbers of items are priced in individual countries, then the Geary Khamis PPPs between individual countries can reduce to ratios of individual prices : thus the averaging effects of adding prices are lost, and the resulting price comparisons can be fairly unstable.

The overall conclusion of this section are, therefore, fairly sobering for the Geary Khamis methods considered. In particular,

(a) unless the price structure involved is very multiplicative the

different Geary Khamis methods can be sensitive to the choice of imputed quantities;

(b) there is a suspicion that the Geary Khamis methods may tend to be fairly unstable in cases where only a small number of items are priced in particular countries.

3. Distance between Geary Khamis and CPD Families: Relationship with $\overline{R}^2$.

3.1 The same theoretical argument that suggested that the internal homogeneity of the Geary Khamis would increase as $\overline{R}^2$ tends to 1, also suggests that the distance between the Geary Khamis and CPD families is likely to decrease with increasing multiplicativity. Fig 7.2 examines this, plotting $\overline{R}^2$ on the horizontal axis against the Geary Khamis / CPD distance on the vertical. For very high values of $\overline{R}^2$, the Geary Khamis and CPD families are indeed very close together. As is the case for the Geary Khamis internal distance, however, the GK /CPD distance can rapidly become quite large when $\overline{R}^2$ falls below about .99.

The main outlier in Figure 7.2, with a GK / CPD distance of about .04, is processed meat. The characteristics of this basic heading are described in more detail below.

Fig 7.2 GK / CPD Distance
versus R(BAR) Squared

R(BAR) Squared

(X 0.01)
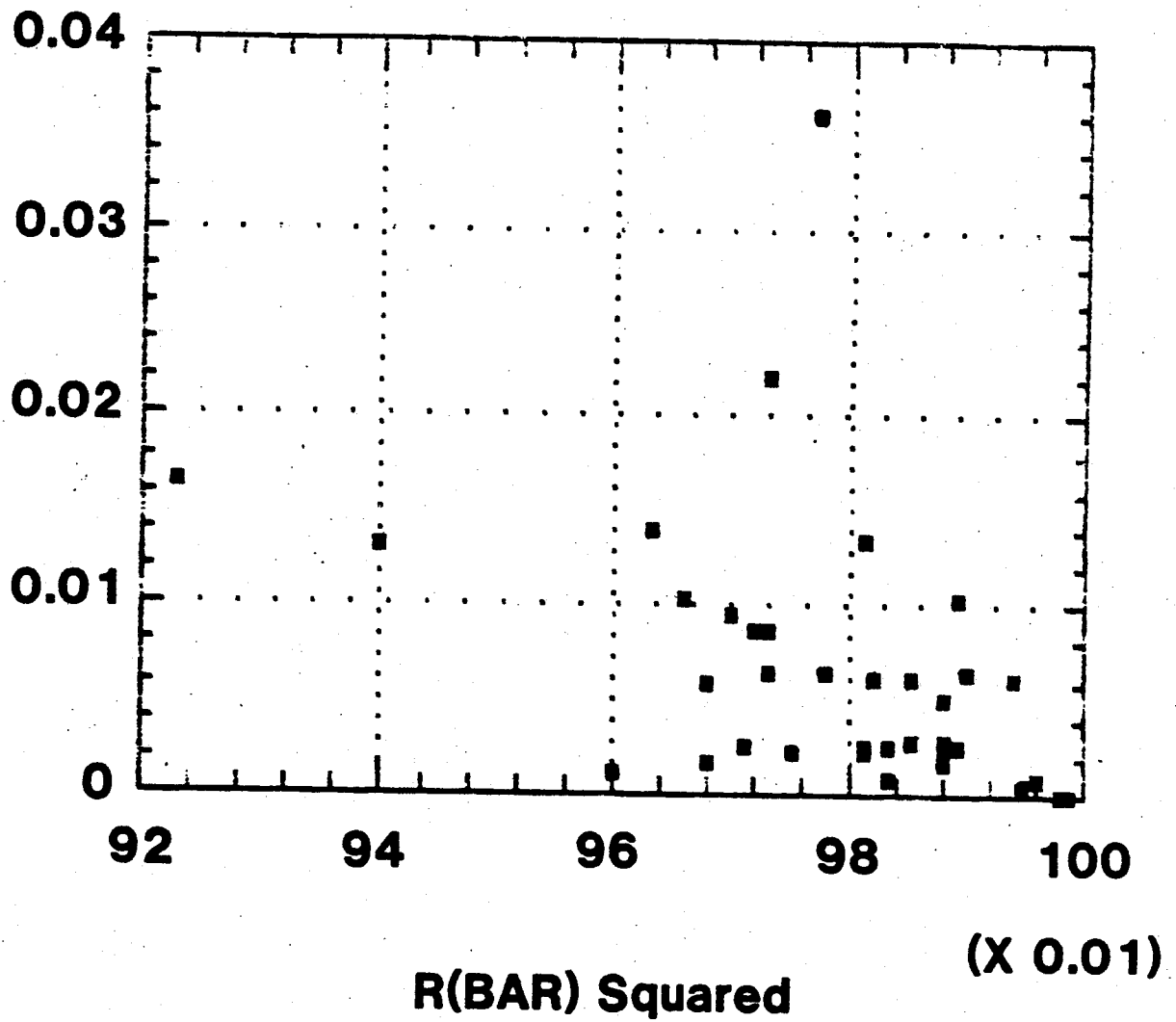
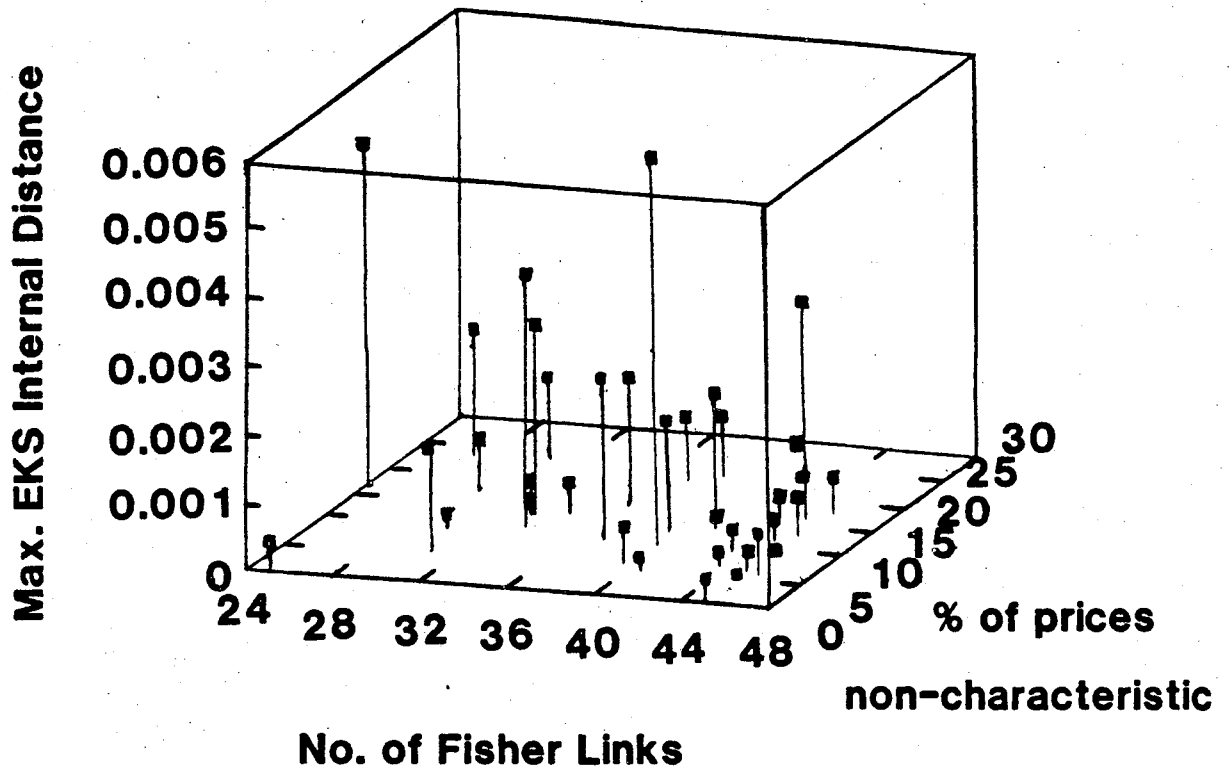4.   <u>Internal Homogeneity of the EKS Family.</u>

4.1 It will be recalled that the term "EKS" has been used to denote the group of three PPPs consisting of the EKS itself, the log Fisher regression with bias term, and the log Fisher without bias. This section examines the internal homogeneity of this family.

4.2 With one exception, Dried Fruit, the EKS family is very internally homogeneous for the 37 product groupings considered. For dried fruit, 4 out of the 6 items priced in the USA are non characteristic - and the US in fact forms only two Fisher links. the effect is that the estimation of the bias coefficient in the "log Fisher with bias" regression for the US is confounded with the estimate of the US price effect, leading to imprecise estimates - and accounting for the very large internal homogeneity measure of 1.5 for the EKS family. For this reason, dried fruit has been excluded from the further analysis of the EKS family.

4.3 Apart from Dried Fruit, the largest internal distance within the EKS family occurs for Sweets (.0056) and Beer (.005) : in fact, for most other basic headings, the EKS internal distances are considerably smaller than this. Figure 7.3 shows a three dimensional plot of the maximum EKS internal distance (vertical axis) against the number of Fisher links found at the first stage in calculating the EKS, and the percentage of prices which relate to non characteristic products. It can be seen that, when the number of first stage Fisher links is close to its theoretical maximum, (45),

Fig 7.3

and where there is a low percentage of non characteristic prices, then the maximum internal distance within the EKS family tends to be very small indeed.

4.4 It should be noted, however, that, despite the high degree of internal homogeneity within the EKS family, some of the distances observed are still large enough to be operationally important. Thus a distance measure of .005, (which was the largest internal distance observed within the EKS family, apart from the dried fruit anomoly) corresponds to an approximate range of +/- 15% on inter country price comparisons. This is certainly large enough to be of some concern in practice.

5.   Distance between EKS and Other Families : Relation with $\bar{R}^2$.

5.1 If the price structure is exactly multiplicative, then the EKS purchasing power parities will coincide with the GK and CPD PPPs. It is therefore appropriate to examine whether there is evidence of convergence between the EKS PPPs and those derived by the other methods, as $\bar{R}^2$ increases. Figure 7.4 shows the EKS /CPD distance plotted against $\bar{R}^2$: and Figure 7.5 shows the corresponding plot for the EKS / GK distance,

5.2 Examination of Figure 7.4 suggests that there is some relationship, albeit fairly weak, for the EKS /CPD distance: all basic headings with a value of $\bar{R}^2$ greater than .99 have an EKS / CPD

Fig 7.4 EKS / CPD Distance versus R(BAR) Squared
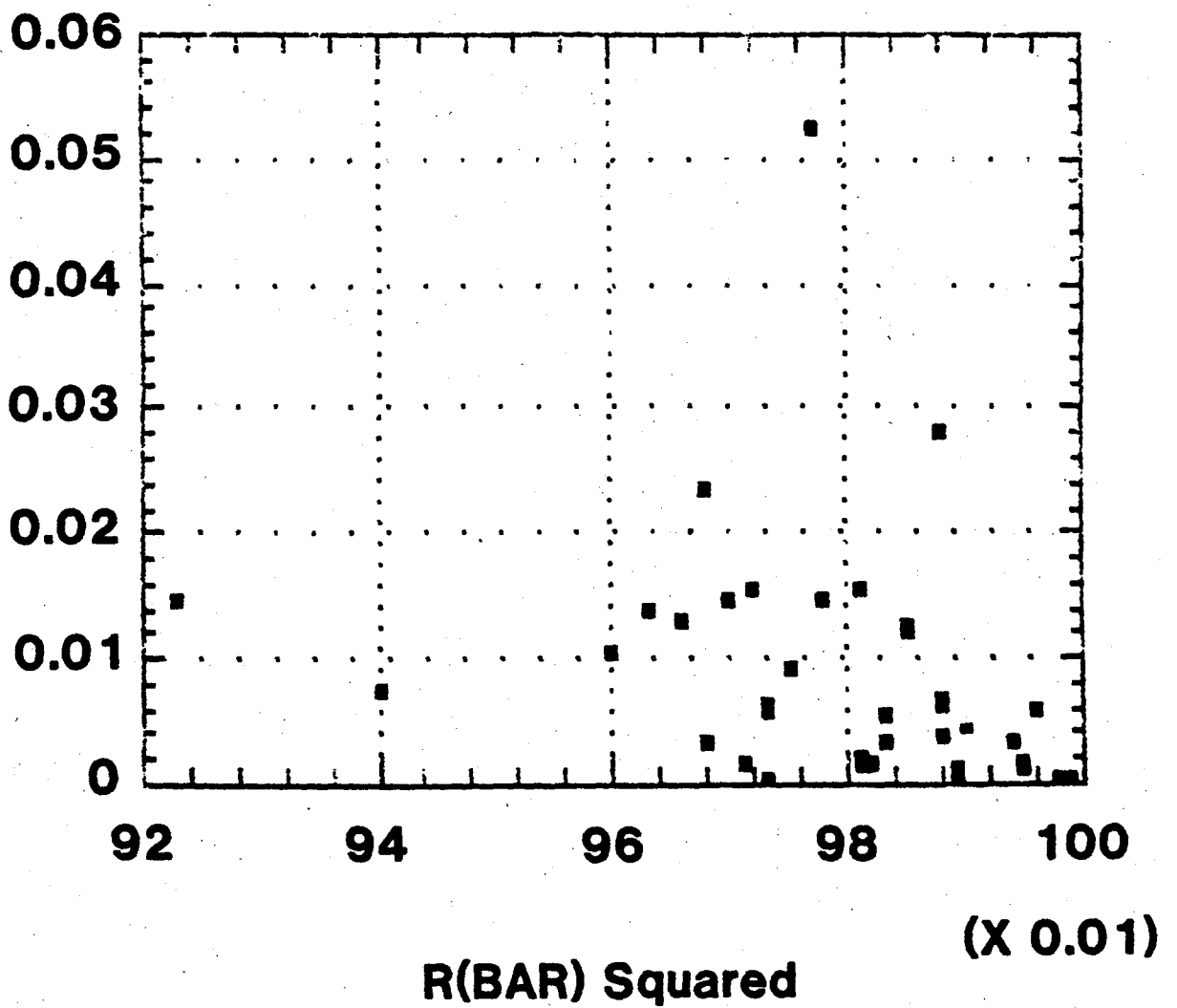
Fig 7.5 EKS / GK Distance
versus R(BAR) Squared

R(BAR) Squared

(X 0.01)
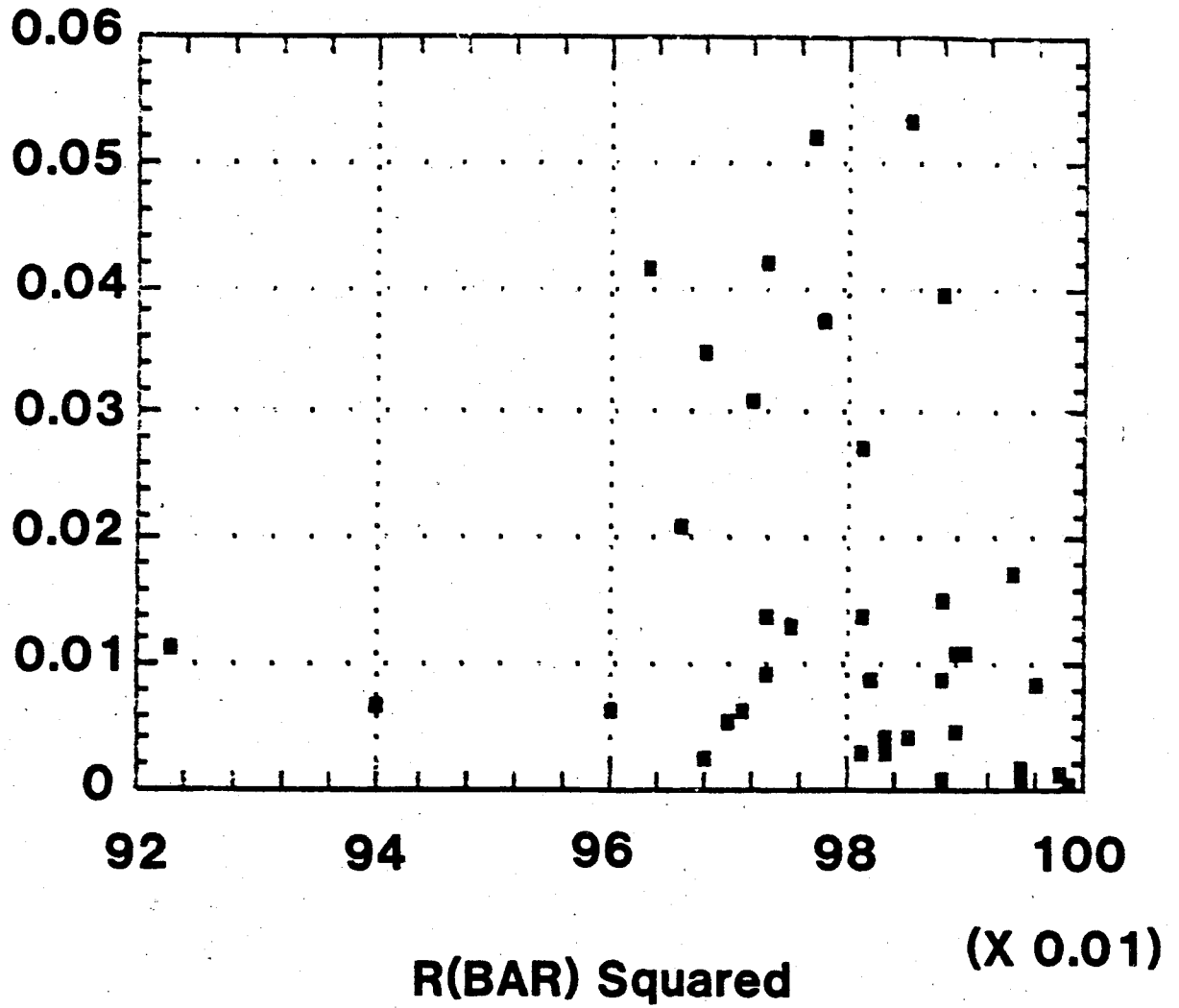
distance of less than .01. For $\overline{R}^2$ values of less than .99, the
maximum observed value of the EKS / CPD distance increases quite
rapidly. The main outlier in this plot, with an EKS / CPD distance
of greater than .05, is again, processed meat.

5.3 Figure 7.5, on the EKS /GK distance, suggests that any
relationship is much weaker than for the other distance comparisons
considered above. With the overall greater dispersion of points in
Figure 4 there are less well defined outliers - the largest values
of the EKS /GK distance, however, which are both larger than .05,
occur for processed meat and beef.

5.4 It is of interest that processed meat shows up as an outlier in
each of Figures 7.2, 7.4 and 7.5 : in other words, for processed
meat there is a large distance between each of the three basic PPP
families. More detailed analysis of the price data for the processed
meat basic heading shows that it has the following features:

(a) 23% of the observed prices for processed meat are non
characteristic: this is the second largest percentage of any basic
heading. However, the estimated bias coefficient from the CPD
regression, at .068, is not particularly large.

(b) In the first stage of the EKS procedure, only 27 direct Fisher
links can be formed. This is the third smallest number of any of the
basic headings considered.

(c) Three countries (Finland, Turkey and Canada) have only one item

priced in each of them: and in fact, the distribution of the number
of priced items over countries is the second most skewed of all the
basic headings considered.

The above factors suggest that the observed price structure for
processed meat is particularly ill conditioned : it is little
surprise, therefore, that it emerges as a consistent outlier.

6. <u>Further Examination of Distances between the Basic PPP Groupings.</u>

6.1 This section continues analysis of the distances between the
basic PPP groupings: the main question considered is - is there any
consistent pattern or relationship among the inter group distances.

Figures 7.6, 7.7, and 7.8 illustrate respectively:

Figure 7.6 : the EKS /CPD distance plotted against the EKS / GK
        distance;

Figure 7.7 : the GK / CPD distance plotted against the EKS / GK
        distance;

Figure 7.8 : the GK / CPD distance plotted against the EKS / CPD.

6.2 Drawn in on each plot are the 45° lines, to help assess which
distances tend to be the smaller.
It is clear from examination of the plots that

Fig 7.6 EKS / CPD Distance
versus EKS / GK Distance
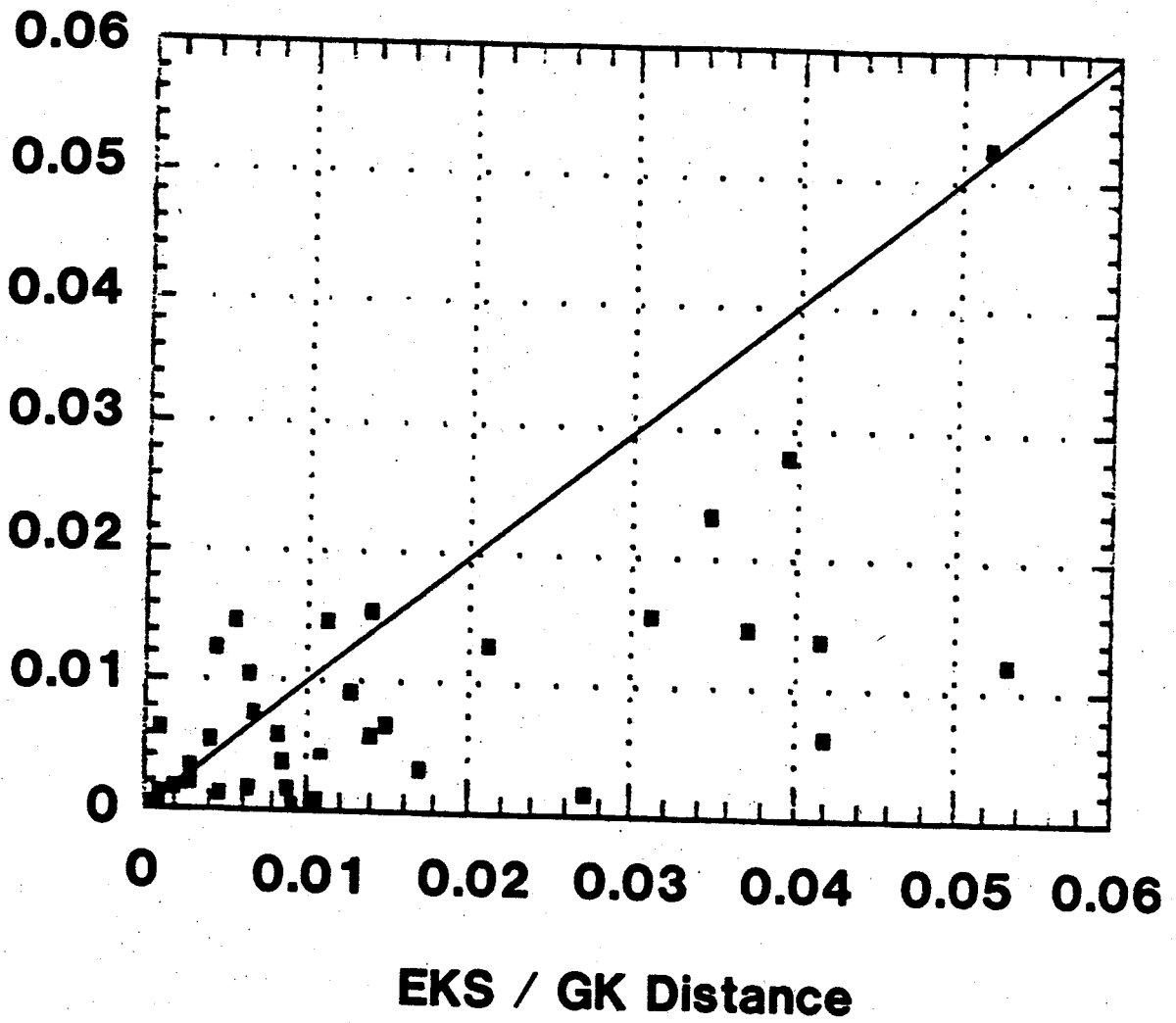
EKS / GK Distance

Fig 7.7 GK / CPD Distance
versus EKS / GK Distance
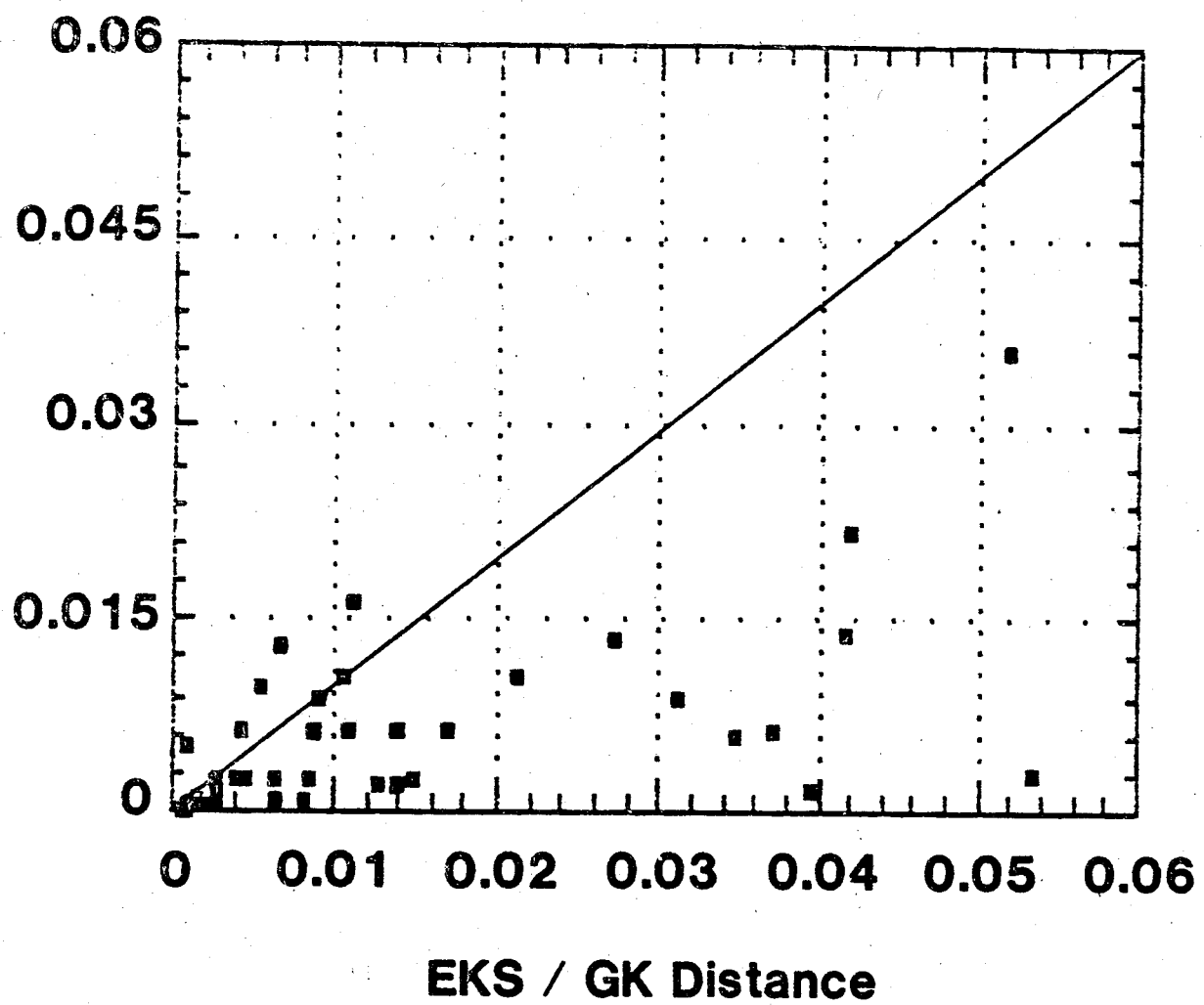
EKS / GK Distance

Fig 7.8 GK / CPD Distance
EKS / CPD Distance
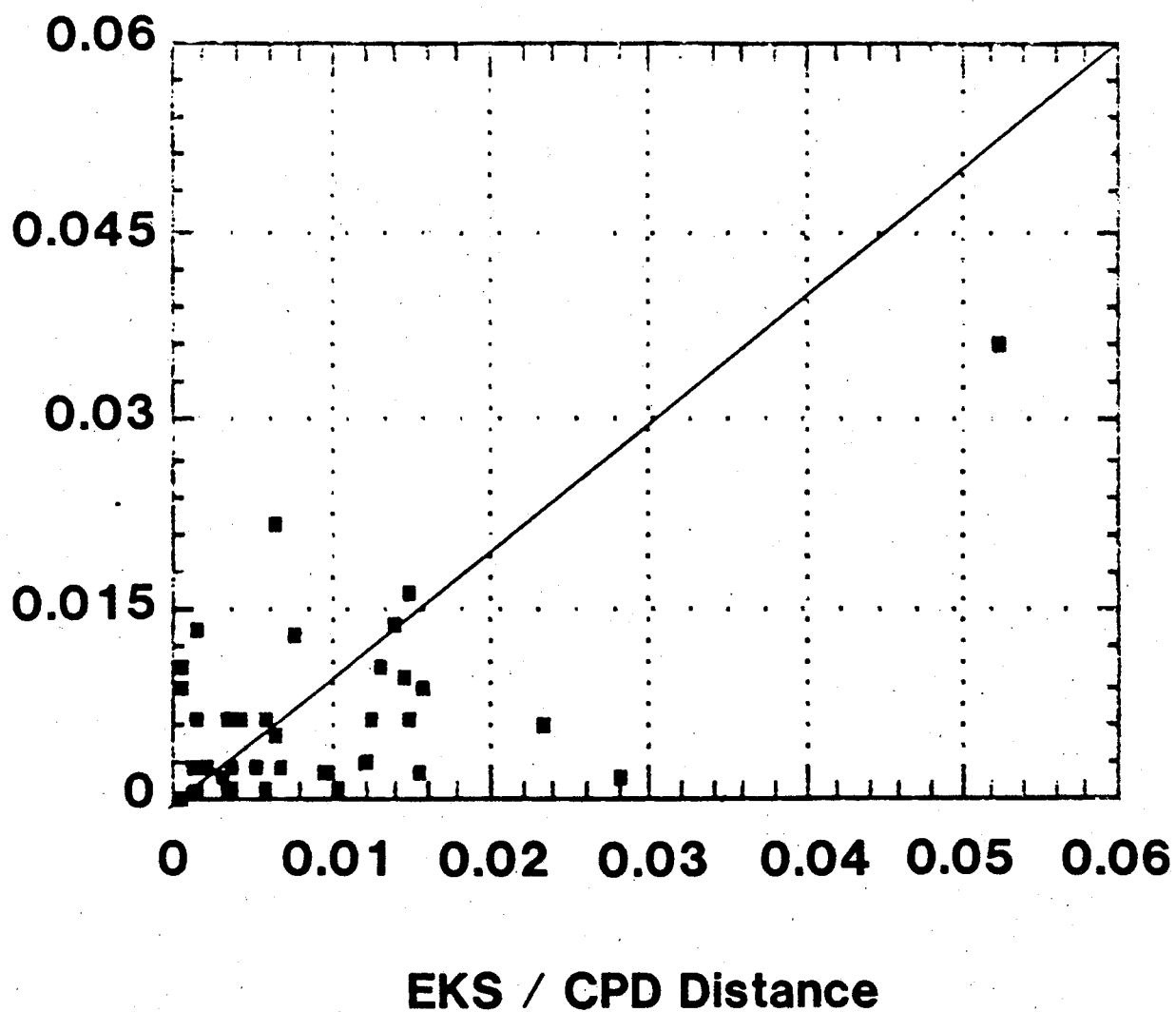
EKS / CPD Distance

(a) for most basic headings, the EKS / CPD distance is smaller than the EKS / GK distance : (Figure 7.6)

(b) similarly, for most basic headings, the GK / CPD distance is smaller than the EKS / GK distances : (Figure 7.7)

(c) the striking factor about Figure 7.8 is that both of the GK / CPD and EKS / CPD distances tend to be fairly small. For 20 out of the 37 product groupings, both distances are less than .01 : and for all but 4 of the groups, both distances are less than .02. There is little or nothing to suggest that the CPD tends to be consistently closer to either the EKS or the GK. In fact, in 25 out of the 37 basic headings, the CPD is closer to the GK than to the EKS. This is weak evidence : however, taken together with the pattern of outliers in Fig 7.8, there is a slight suggestion that the CPD may tend to be closer to the GK than to the EKS.

6.3 To summarise the information in Figs 7.6, 7.7 and 7.8 : from the point of view of the Geary Khamis, the distance from the GK to the CPD tends to be smaller than the distance from the GK to the EKS. Conversely, from the point of view of the EKS, the distance from the EKS to the CPD tends to be smaller than the distance from the EKS to the GK. We could illustrate this state of affairs in a heuristic diagram as follows, by saying that the CPD tends to lie, in distance terms, somewhere between the EKS and the GK:

<pre>
        EKS   GK

         *   *

                        CPD

                         *
</pre>

Furthermore, there is a weak suggestion that the CPD may tend to lie closer to the GK than to the EKS.

Of course, this kind of heuristic diagram is not to be interpreted too literally : but it provides a useful way of summarising the information in Figures 7.7 and 7.8.

7. Precision of CPD Estimates.

7.1 Since the CPD technique involves fitting a regression equation, the standard properties of regression provide a direct estimate of the precision with which the CPD coefficients are estimated – in terms of the estimated standard errors of the regression coefficients. This is a bonus in the CPD approach: no such estimates of precision are readily available for the EKS or GK approaches.

7.2 In fitting the CPD regressions for the 37 product groupings studied, country 1 (Austria) was arbitrarily taken as the base country. This means that the estimated country coefficients from the

Fig 7.9

AV. Standard Errors (y-axis, 0 to 0.4)

Residual degrees of freedom in CPD regression (x-axis, 0 to 60)
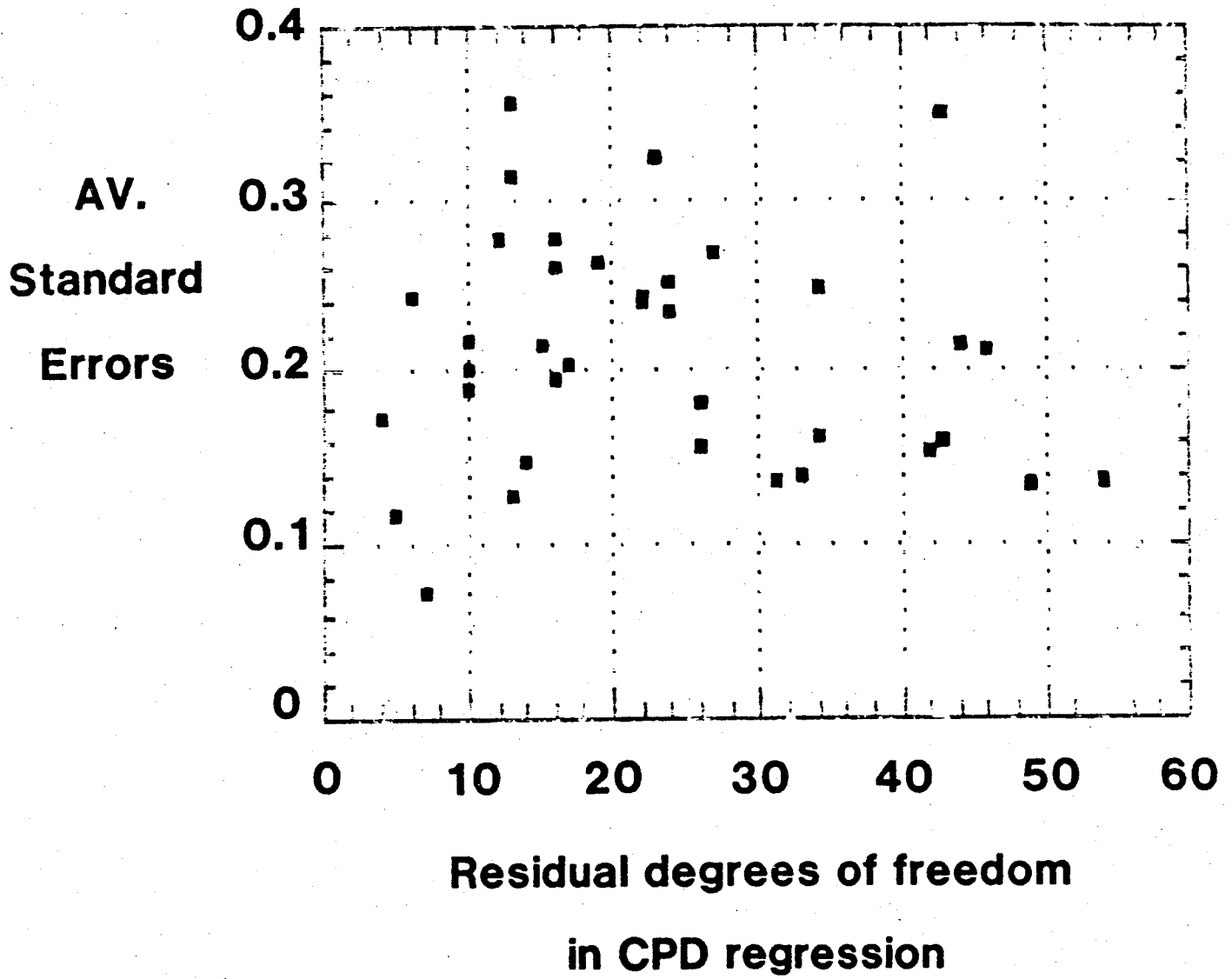
CPD regression represent the logs of the purchasing power relativities to Austria. In order to provide a convenient summary measure, the average of the standard errors of the 9 country coefficients was calculated within each basic heading.

7.3 Figure 7.9 shows the average product grouping standard errors, plotted against the residual degrees of freedom from the CPD regression. The version of the CPD regression model without a non characteristic dummy was used in this analysis. Since the number of coefficients estimated in this regression is

> 1 constant
>
> plus 9 country effects
>
> plus (number of items - 1) item effects,

the residual number of degrees of freedom is equal to

> (number of observed prices) - (number of items) - 9.

7.4 Two factors are of potential interest in Figure 9.

(a) The first is the size of the average standard errors: only one standard error is smaller than 0.1, and most are around 0.2. A 95% confidence interval for a given log price relativity is approximately +/- 2 standard errors. Hence, for a given price relativity for which the standard error of the coefficient was 0.2, the half width of a 95% confidence interval around the rice relativity p would be given by a multiplicative factor of $\exp(2*0.2) = 1.5$; that is, the confidence interval would be given by the range

{ p/1.5 , p*1.5 }

This suggests that the CPD technique, on the data studied, will typically give very imprecise estimates of purchasing power relativities. This is sobering - but is not, in itself, a criticism of the CPD technique, relative to the EKS or GK, since corresponding estimates of precision for the EKS and GK methods are not readily available. There is therefore no reason to suppose that the CPD is any less precise than the EKS or GK approaches. Indeed, since, as has been noted earlier, the CPD tends to occupy a central position between the EKS and the GK, there may be some circumstantial grounds to suggest that the CPD may be more precise than the other techniques.

(b) The second interesting feature in Figure 7.9 is the suggestion that the vertical spread of average standard errors is greater towards the left of the diagram - that is, where the residual degrees of freedom in the regression are small. This is as one would expect - since there will be greater imprecision in estimating the standard error, the smaller the number of residual degrees of freedom. This feature points to a difficulty with the CPD approach, if the number of observed prices is only slightly greater than the number of countries plus items. In this situation, there is a danger of getting a severely misleading impression of the precision of the CPD estimates.

8. <u>Conclusions</u>

8.1 Finally, what conclusions can be drawn from the data analysis presented in this and the previous chapter?

(a) First of all, on the question of characteristic / non characteristic bias, the conclusions are fairly positive. Significant evidence has been found that this kind of bias exists - and of a magnitude which could have an important bearing on real life PPP comparisons. No uniform a-priori rule can be laid down as to whether this effect should be corrected for in any specific instance : this really depends on how large a proportion of total trade within the country is accounted for by non characteristic products.
The implication, therefore, is that , unless there is to be a risk of significant distortion arising from this cause, effort has to be put into more refined estimates of commodity volumes below basic heading level : and a technique of calculating PPPs should be employed which would enable the size of the potential bias to be assessed , and any appropriate adjustment made, if necessary.

This points to the use of the CPD with non characteristics dummy, or the Log Fisher with dummy. In addition, with refined enough quantity data, the Geary Khamis method will automatically make the appropriate adjustment.

(b) More generally , the choice between different PPP methods is by no means clear cut. Where the price structure is very

multiplicative, there does tend to be a general convergence between the different measures considered. In this happy circumstance it is unlikely to make much difference which measure is chosen. But for this to hold, the degree of multiplicativity, as measured by $\bar{R}^2$ from the CPD regression, has to be very high indeed. Unless the price structure is very multiplicative, there can be large differences between the PPPs assessed by different methods, both within the Geary Khamis family, and between the different families. In the absence of any objective reasons for preferring one particular approach to another, this points to a major problem of choice.

(c) What can be done about this?

There are perhaps three possible approaches, not necessarily mutually exclusive:

(i) as has been noted earlier, the CPD approach tends to produce a PPP which lies "between" the EKS and Geary Khamis families. This may suggest a preference for the CPD as a reasonable compromise approach.

(ii) alternatively, effort could be put into deriving more realistic quantity data within basic headings. In practice, this is unlikely to be a feasible option, for cost and other reasons. But if fairly accurate quantity data could be derived, the theoretical arguments for the Geary Khamis approach might suggest adopting the Geary Khamis.

(iii) A third approach would involve putting effort into deriving more refined quantity or expenditure data - but this time directed

at reducing the size of the basic headings, that is, splitting them down into smaller aggregates, rather than attempting to obtain quantities at individual item level.

As basic headings become smaller, there is a good chance that the internal price structure may become more multiplicative - so making the choice of method less critical. There are, however, limits to this approach - and these relate only partly to cost and data availability. As the number of items in the basic heading gets smaller, the number of price observations upon which to base estimates of the required parameters is therefore also reduced. This could lead to imprecision in parameter estimates. In particular, since the CPD approach involves the simultaneous estimation of a relatively large number of parameters, (equal to the number of countries plus the number of items), this might be a more serious problem for the CPD approach compared to the other techniques.

CHAPTER EIGHT

AGGREGATION OF BASIC HEADINGS

1.   Introduction

1.1 The final stage in the analysis was to aggregate the PPPs derived at product group level over the 37 product groups. Two methods were used in this aggregation - the Geary Khamis and the Gerardi (UCW) method. Both of these methods were applied to each of the eight PPPs calculated at product group level - yielding, therefore, sixteen different ways of calculating the final PPP at the aggregate level. These sixteen methods are denoted by numbers, as follows :-

1 : GK	aggregation of EKS
2 :Gerardi	aggregation of EKS
3 : GK	aggregation of (log Fisher with bias)
4 :Gerardi	aggregation of (log Fisher with bias)
5 : GK	aggregation of (log Fisher without bias)
6 :Gerardi	aggregation of (log Fisher without bias)
7 : GK	aggregation of GK1
8 :Gerardi	aggregation of GK1

9 : GK   aggregation of GK2

10:Gerardi aggregation of GK2

11: GK   aggregation of GK3

12:Gerardi aggregation of GK3

13: GK   aggregation of CPD with bias

14:Gerardi aggregation of CPD with bias

15: GK   aggregation of CPD without bias

16:Gerardi aggregation of CPD without bias

In the few cases where it was not possible to calculate the log
Fisher or CPD with bias, the corresponding indices without bias were
used. For Dried Fruit, where the log Fisher with bias can be
calculated but is suspect, because of an identifiability problem,
the without bias version was also used.

Note that in the above numbering, odd numbered methods refer to GK
aggregation; even numbered to Gerardi.

2. The Results

2.1 Table 8.1 shows the complete set of resulting PPPs aggregate level
for the 16 methods outlined above. For each PPP, all 45 possible

TABLE 8.1 PURCHASING POWER PARITIES, AGGREGATE LEVEL AMONG 10 COUNTRIES

AGGREGATION SECOND STAGE	GK	GER'DI	GK	GER'DI	GK	GER'DI	GK	GER'DI
FIRST STAGE	EKS	EKS	LOG F (with bias)	LOG F	LOG F (without bias)	LOG F	GK1	GK1
COUNTRY COMPARISON								
2 TO 1	.4738	.4730	.4770	.4742	.4721	.4705	.4501	.4559
3 TO 1	.7084	.7229	.7224	.7329	.7075	.7212	.6738	.7047
4 TO 1	.6665	.6626	.6742	.6682	.6630	.6585	.6357	.6406
5 TO 1	.0659	.0663	.0657	.0659	.0654	.0658	.0632	.0647
6 TO 1	.0831	.0818	.0828	.0813	.0822	.0809	.0771	.0775
7 TO 1	18.75	18.45	18.77	18.37	18.61	18.32	18.92	18.96
8 TO 1	.0856	.0814	.0863	.0819	.0853	.0812	.0816	.0787
9 TO 1	.0626	.0581	.0626	.0581	.0620	.0577	.0599	.0566
10 TO 1	10.80	10.94	10.89	10.95	10.73	10.83	10.06	10.48
3 TO 2	1.495	1.528	1.514	1.546	1.499	1.533	1.497	1.546
4 TO 2	1.407	1.401	1.413	1.409	1.404	1.399	1.412	1.405
5 TO 2	.1390	.1402	.1377	.1390	.1384	.1399	.1404	.1419
6 TO 2	.1754	.1730	.1735	.1715	.1742	.1720	.1714	.1700
7 TO 2	39.58	39.00	39.36	38.75	39.43	38.93	42.04	41.59
8 TO 2	.1807	.1720	.1810	.1727	.1806	.1723	.1813	.1727
9 TO 2	.1321	.1228	.1313	.1225	.1314	.1226	.1331	.1242
10 TO 2	22.79	23.12	22.84	23.10	22.72	23.01	22.35	22.99
4 TO 3	.9408	.9166	.9334	.9117	.9370	.9130	.9435	.9090
5 TO 3	.0930	.0917	.0909	.0899	.0924	.0913	.0938	.0918
6 TO 3	.1174	.1132	.1146	.1110	.1162	.1122	.1144	.1100
7 TO 3	26.47	25.52	25.99	25.07	26.31	25.40	28.08	26.90
8 TO 3	.1208	.1126	.1195	.1117	.1205	.1124	.1211	.1117
9 TO 3	.0883	.0803	.0867	.0792	.0877	.0800	.0889	.0803
10 TO 3	15.24	15.13	15.08	14.95	15.16	15.01	14.93	14.87
5 TO 4	.0988	.1000	.0974	.0987	.0986	.0999	.0994	.1009
6 TO 4	.1247	.1235	.1227	.1217	.1240	.1229	.1213	.1210
7 TO 4	28.14	27.84	27.85	27.50	28.08	27.82	29.76	29.59
8 TO 4	.1284	.1228	.1280	.1226	.1286	.1231	.1284	.1229
9 TO 4	.0939	.0877	.0929	.0869	.0936	.0876	.0942	.0884
10 TO 4	10.20	16.51	16.16	16.39	16.18	16.44	15.82	16.36
6 TO 5	1.262	1.234	1.260	1.234	1.26	1.230	1.221	1.199
7 TO 5	284.7	278.2	285.8	278.7	284.8	278.3	299.5	293.2
8 TO 5	1.300	1.227	1.314	1.242	1.305	1.232	1.292	1.217
9 TO 5	.9500	.8764	.9535	.8811	.9493	.8767	.9483	.8754
10 TO 5	163.9	165.0	165.8	166.2	164.1	164.5	159.2	162.0
7 TO 6	225.6	225.4	226.9	225.9	226.4	226.3	245.3	244.6
8 TO 6	1.030	.9944	1.043	1.007	1.037	1.002	1.058	1.016
9 TO 6	.7528	.7100	.7568	.7143	.7545	.7129	.7767	.7303
10 TO 6	129.9	133.7	131.6	134.7	130.4	133.8	130.4	135.2
8 TO 7	.0046	.0044	.0046	.0045	.0046	.0043	.0043	.0042
9 TO 7	.0033	.0031	.0033	.0032	.0033	.0032	.0032	.0030
10 TO 7	.5759	.5930	.5803	.5962	.5782	.5911	.5316	.5528
9 TO 8	.7310	.7140	.7256	.7092	.7276	.7116	.7341	.7191
10 TO 8	126.2	134.4	126.2	133.8	125.8	133.6	123.2	133.1
10 TO 9	172.6	188.2	173.9	188.6	172.9	187.7	167.9	185.1

TABLE 8.1 PURCHASING POWER PARITIES, AGGREGATE LEVEL AMONG 10 COUNTRIES

AGGREGATION SECOND STAGE	GK	GER'DI	GK	GER'DI	GK	GER'DI	GK	GER'DI
FIRST STAGE	GK2	GK2	GK3	GK3	CPD (with bias)	CPD (with bias)	CPD (without bias)	CPD (without bias)
COUNTRY COMPARISON								
2 TO 1	.4517	.4571	.4456	.4528	.4795	.4797	.4567	.4564
3 TO 1	.6782	.7069	.6734	.7018	.7137	.7360	.6793	.7006
4 TO 1	.6363	.6405	.6305	.6350	.6653	.6632	.6392	.6376
5 TO 1	.0634	.0648	.0636	.0656	.0659	.0669	.0644	.0655
6 TO 1	.0780	.0782	.0770	.0779	.0818	.0817	.0784	.0780
7 TO 1	18.72	18.67	19.03	18.95	18.89	18.78	18.39	18.38
8 TO 1	.0817	.0789	.0846	.0814	.0864	.0829	.0824	.0788
9 TO 1	.0600	.0568	.0605	.0573	.0616	.0580	.0602	.0565
10 TO 1	10.10	10.47	10.44	10.78	10.54	10.86	10.11	10.40
3 TO 2	1.501	1.546	1.511	1.550	1.488	1.534	1.487	.1535
4 TO 2	1.408	1.401	1.415	1.402	1.387	1.383	1.400	1.397
5 TO 2	.1403	.1416	.1426	.1448	.1375	.1395	.1411	.1434
6 TO 2	.1727	.1710	.1729	.1721	.1705	.1702	.1717	.1709
7 TO 2	41.43	40.84	42.71	41.83	39.39	39.15	40.27	40.27
8 TO 2	.1809	.1725	.1897	.1798	.1802	.1729	.1803	.1727
9 TO 2	.1329	.1243	.1358	.1264	.1284	.1209	.1317	.1238
10 TO 2	22.36	22.91	23.42	23.82	21.97	22.64	22.14	22.78
4 TO 3	.9381	.9061	.9364	.9048	.9323	.9011	.9410	.9101
5 TO 3	.0935	.0916	.0944	.0935	.0924	.0909	.0948	.0934
6 TO 3	.1150	.1106	.1144	.1110	.1146	.1109	.1154	.1114
7 TO 3	27.60	26.41	28.26	27.00	26.47	25.52	27.08	26.23
8 TO 3	.1205	.1116	.1256	.1160	.1211	.1127	.1212	.1125
9 TO 3	.0885	.0804	.0899	.0816	.0863	.0788	.0886	.0806
10 TO 3	14.90	14.82	15.50	15.37	14.76	14.76	14.88	14.84
5 TO 4	.0996	.1011	.1008	.1032	.0990	.1009	.1008	.1026
6 TO 4	.1226	.1220	.1222	.1227	.1229	.1231	.1227	.1224
7 TO 4	29.42	29.15	30.18	29.84	28.39	28.32	28.78	28.83
8 TO 4	.1285	.1231	.1341	.1282	.1299	.1250	.1288	.1236
9 TO 4	.0944	.0887	.0960	.0902	.0926	.0874	.0941	.0886
10 TO 4	15.88	16.35	16.55	16.98	15.84	16.38	15.82	16.31
6 TO 5	1.231	1.207	1.212	1.188	1.240	1.221	1.217	1.192
7 TO 5	295.3	288.3	299.4	288.9	286.5	280.7	285.5	280.8
8 TO 5	1.290	1.218	1.330	1.241	1.311	1.239	1.278	1.204
9 TO 5	.9473	.8773	.9523	.8729	.9341	.8665	.9338	.8633
10 TO 5	159.4	161.7	164.2	164.4	159.8	162.3	156.9	158.9
7 TO 6	239.9	238.8	247.1	243.1	231.0	230.0	234.6	235.6
8 TO 6	1.048	1.009	1.098	1.045	1.057	1.015	1.050	1.010
9 TO 6	.7697	.7266	.7859	.7347	.7530	.7099	.7673	.7242
10 TO 6	129.5	134.0	135.5	138.4	128.8	133.0	129.0	133.3
8 TO 7	.0044	.0042	.0044	.0043	.0046	.0041	.0045	.0043
9 TO 7	.0032	.0030	.0032	.0030	.0033	.0031	.0033	.0031
10 TO 7	.5397	.5609	.5485	.5692	.5578	.5783	.5497	.5658
9 TO 8	.7346	.7202	.7159	.7032	.7125	.6991	.7305	.7167
10 TO 8	123.6	132.8	123.4	132.5	121.9	131.0	122.8	131.9
10 TO 9	168.2	184.4	172.4	188.4	171.1	187.4	168.1	184.1

inter country comparisons among the 10 countries are shown.

Visual inspection of Table 8.1 suggests that there is a fairly high degree of uniformity among the different PPPs. But for a more informative analysis, it is useful to consider the inter PPP distance measure, D, defined earlier. Table 8.2 shows the complete 16x16 matrix of the distances, D, between every possible pair of PPPs. For ease of reference, the GK PPPs are grouped together first of all, followed by the Gerardis: also, lines have been drawn on Table 8.2 to indicate the main PPP groupings.

2.2 The main points to be noted from Table 8.2 are as follows:-

(a) First of all, consider the group defined by PPPs numbered 1, 3 and 5: that is, the group consisting of GK aggregates of the EKS family. The within group distances for this group form the top left hand 3x3 block of the matrix in Table 8.2. It can be seen that this group is, in fact, very internally homogeneous, with a maximum distance within the group of .0001. Examination of the other diagonal blocks of the matrix shows that a high degree of internal homogeneity is also a feature of the other groups which are defined by a common method of aggregation of the basic PPP families: the GK - GK, GK - CPD, Gerardi - EKS, Gerardi - GK, and Gerardi - CPD groups.

(b) It is also noticeable that there is a fair degree of internal

TABLE 8.2   INTER PPP DISTANCES

		GEARY KHAMIS AGGREGATION								GERARDI AGGREGATION							
		EKS FAMILY			GK FAMILY			CPD FAMILY		EKS FAMILY			GK FAMILY			CPD FAMILY	
		1	3	5	7	9	11	13	15	2	4	6	8	10	12	14	16
GEARY / EKS FAMILY	1	.0000	.0001	.0000	.0014	.0009	.0016	.0002	.0006	.0014	.0016	.0014	.0023	.0018	.0022	.0016	.0017
	3	.0001	.0000	.0001	.0017	.0012	.0019	.0003	.0009	.0013	.0013	.0012	.0023	.0018	.0022	.0013	.0017
	5	.0000	.0001	.0000	.0013	.0009	.0015	.0002	.0005	.0014	.0015	.0013	.0022	.0017	.0021	.0015	.0016
KHAMIS / GK FAMILY	7	.0014	.0017	.0013	.0000	.0001	.0003	.0012	.0004	.0033	.0036	.0031	.0019	.0019	.0020	.0032	.0020
	9	.0009	.0012	.0009	.0001	.0000	.0004	.0008	.0002	.0028	.0030	.0026	.0018	.0016	.0019	.0027	.0017
	11	.0016	.0019	.0015	.0003	.0004	.0000	.0014	.0006	.0035	.0038	.0033	.0021	.0021	.0017	.0033	.0022
CPD FAMILY	13	.0002	.0003	.0002	.0012	.0008	.0014	.0000	.0005	.0018	.0019	.0017	.0023	.0019	.0022	.0015	.0018
	15	.0006	.0009	.0005	.0004	.0002	.0006	.0005	.0000	.0024	.0027	.0023	.0020	.0018	.0020	.0024	.0016
GERARDI / EKS FAMILY	2	.0014	.0013	.0014	.0033	.0028	.0035	.0018	.0024	.0000	.0001	.0000	.0011	.0006	.0011	.0002	.0006
	4	.0016	.0013	.0015	.0036	.0030	.0038	.0019	.0027	.0001	.0000	.0001	.0013	.0008	.0013	.0002	.0008
	6	.0014	.0012	.0013	.0031	.0026	.0033	.0017	.0023	.0000	.0001	.0000	.0010	.0006	.0010	.0002	.0005
GK FAMILY	8	.0023	.0023	.0022	.0019	.0018	.0021	.0023	.0020	.0011	.0013	.0010	.0000	.0001	.0003	.0009	.0003
	10	.0018	.0018	.0017	.0019	.0016	.0021	.0019	.0018	.0006	.0008	.0006	.0001	.0000	.0003	.0006	.0001
	12	.0022	.0022	.0021	.0020	.0019	.0017	.0022	.0020	.0011	.0013	.0010	.0003	.0003	.0000	.0009	.0003
CPD FAMILY	14	.0016	.0013	.0015	.0032	.0027	.0033	.0015	.0024	.0002	.0002	.0002	.0009	.0006	.0009	.0000	.0006
	16	.0017	.0017	.0016	.0020	.0017	.0022	.0018	.0016	.0006	.0008	.0005	.0003	.0001	.0003	.0008	.0000

homogeneity within the 8x8 block of PPPs defined by GK aggregation; the top left 8x8 block in the matrix. Within this block, the largest internal distance is .0019, between PPPs numbered 3 and 11, and most of the distances within this group are a good deal smaller than this. .

Similarly, there is also a high degree of homogeneity within the Gerardi group, which is shown in the bottom Right hand 8x8 block in the matrix. In fact, within the Gerardi block, the largest internal distance is .0013, between PPPs numbered 8 and 4, and also between PPPs 12 and 4.

(c) One of the most interesting features in Table 8.2, however, is the relatively high distances between the GK and Gerardi groupings. These inter group distances are those occuring in the top right or bottom left 8x8 block of the table. The smallest of these is .0012. In other words, the GK aggregates are uniformly fairly distant from the Gerardi aggregates.

Examination of Table 8.1 suggests at least one of the main factors accounting for this feature. From inspection, it appears that the major differences between the GK and Gerardi aggregates of a particular PPP tend to occur in country comparisons involving the US. In each of rows 8, 16, 23, 29, 34, 38, 41, and 43, the US is in the numerator of the inter country comparison, and the price ratio is <u>higher</u> in the GK aggregated version compared with the corresponding Gerardi aggregated version. The opposite is true for row 45; there the US is in the denominator of the comparisons. This

pattern is consistent with the underlying US price level being

higher in the GK aggregate method than for the corresponding Gerardi

aggregate method.

This consistent difference between the price levels for the US in

the GK and Gerardi aggregation methods is not unexpected. The

Gerardi method computes international prices as the unweighted

geometric average of basic heading prices. On the other hand, the GK

approach computes international prices essentially as a weighted

average, where the weights defined are the quantities in the

individual countries. Since the US has by far the largest volumes of

the 10 countries, it is not unexpected that the largest difference

between the GK and the Gerardi approach shows up primarily for the

US. In other words, this observed pattern is consistent with the

Gerschenkron effect.

2.3 In chapter 2, section 5, the question was raised as to whether

the combined effect of applying the Geary Khamis approach in

aggregating first to basic heading level and then to GDP level might

lead to a compounded Gerschenkron effect - with a consequent over

estimation of the price level of the dominant country or group. It

was noted that this could not happen with the GK1 and GK2 methods

for imputing quantities - since, under both these methods the

imputed quantities are scaled to the same totals for all countries.

In the context of the present data analysis, therefore, the

compounded Gerschenkron effect could only arise when the GK3 method

is used in the first stage of aggregation, combined with the Geary

Khamis method in the second.

Examination of the aggregate PPPs considered in this chapter is not entirely conclusive on this point. In Table 8.2, the largest distances, D, occur between the PPP derived from combining the GK3 with the Geary Khamis (PPP number 11) and the PPPs obtained by combining the EKS family with the Gerardi, and the CPD (with bias) and the Gerardi. This is consistent with the existence of a compounded Gerschenkron effect. However, both the GK1 - GK and GK2 - GK PPPs , also have relatively large distances from the EKS and CPD families combined with the Gerardi. Thus, the evidence for the existence of a compounded Gerschenkron effect is still equivocal.

Direct examination of Table 8.1 is also relevant. On considering the price relativities in the Table between the US (country 9), and each of the other countries, it can be seen that for four of these countries the GK3 - GK price relativity is higher than for all of the other aggregation methods. The countries for which this is true are 2 (Finland), 3 (Norway), 4 (Sweden), and 6 (New Zealand) - all small countries in overall terms. This suggests that a compounded GK effect might indeed be at work.

In summary, it would not be appropriate to reach any firm conclusion, given the rather inconclusive nature of the evidence: nevertheless, the evidence does tend to suggest that a compounded Gerschenkron effect could be a possibility if the GK3 method of aggregating up to basic heading level is combined with Geary Khamis aggregation up to GDP level.

2.4 It is appropriate to conclude with the following cautionary remark. The emphasis of this report is on problems and techniques arising in aggregating purchasing power parities up to basic heading level. This final chapter has involved moving from basic heading level to a higher level of aggregation. However, the discussion in this chapter is intended to throw more light on the techniques used in aggregating up to basic heading level - it is not intended as an account of aggregation beyond basic headings. In particular, the two higher aggregation techniques used, the Geary Khamis and Gerardi (UCW) methods, have been selected merely as two useful illustrative examples of the kind of techniques which can be applied in aggregating at higher level. It would not be appropriate to draw any conclusions, from the discussion in this chapter, on the relative merits of different aggregation techniques up to GDP level.

APPENDIX A

GEARY KHAMIS SOLUTION IN THE THREE COUNTRY CASE.

Let the three countries be designated 0, 1, and 2. Further, with no loss of generality, let us choose country 0 as the country with the numeraire currency and which is also the base country. Thus,

$$R_o = 1$$

where R_j is the ratio of prices in the numerator country to the base country 0.

$$\pi_i = \frac{\Sigma p_{ij} R_j^{-1} q_{ij}}{\Sigma q_{ij}}$$

$$= \frac{[p_{io} q_{io} + p_{i1} q_{i1} R_1^{-1} + p_{i2} q_{i2} R_2^{-1}]}{(q_{io} + q_{i1} + q_{i2})} \qquad \text{---(a)}$$

$$R_j = \frac{\Sigma p_{ij} q_{ij}}{\Sigma \pi_i q_{ij}} \qquad \text{---(b)}$$

Therefore,

$$R_1^{-1} = \frac{[\Sigma \pi_i q_{i1}]}{(\Sigma p_{i1} q_{i1})}$$

$$= \frac{1}{(\Sigma p_{i1} q_{i1})} \left[\sum \frac{p_{io} q_{io} q_{i1} + p_{i1} q_{i1}^2 R_1^{-1} + p_{i2} q_{i1} q_{i2} R_2^{-1}}{(q_{io} + q_{i1} + q_{i2})} \right]$$

Thus,

$$(\Sigma p_{i1}q_{i1})R_1^{-1} = \sum \frac{p_{io}q_{io}q_{i1}}{(q_{io}+q_{i1}+q_{i2})} + R_1^{-1}\left[\sum \frac{p_{i1}q_{i1}^2}{(q_{io}+q_{i1}+q_{i2})}\right] + R_2^{-1}\left[\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right]$$

That is,

$$\left[\sum \frac{p_{i1}q_{i1}(q_{io}+q_{i1}+q_{i2})-p_{i1}q_{i1}^2}{q_{io}+q_{i1}+q_{i2}}\right]R_1^{-1} - \left[\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right]R_2^{-1}$$

$$= \sum \frac{p_{io}q_{io}q_{i1}}{q_{io}+q_{i1}+q_{i2}}$$

That is,

$$\left[\sum \frac{p_{i1}q_{i1}(q_{io}+q_{i2})}{q_{io}+q_{i1}+q_{i2}}\right]R_1^{-1} - \left[\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right]R_2^{-1}$$

$$= \sum \frac{p_{io}q_{io}q_{i1}}{q_{io}+q_{i1}+q_{i2}} \qquad ---(c)$$

Similarly,

$$-\left[\sum \frac{p_{i1}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right]R_1^{-1} + \left[\sum \frac{p_{i2}q_{i2}(q_{io}+q_{i1})}{q_{io}+q_{i1}+q_{i2}}\right]R_2^{-1} = \sum \frac{p_{io}q_{io}q_{i2}}{q_{io}+q_{i1}+q_{i2}} \qquad ---(d)$$

Thus,

$$R_1^{-1}\left[\sum \frac{p_{i1}q_{i1}(q_{io}+q_{i2})}{q_{io}+q_{i1}+q_{i2}} - \frac{\left(\sum \frac{p_{i1}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)}{\sum \frac{p_{i2}q_{i2}(q_{io}+q_{i1})}{q_{io}+q_{i1}+q_{i2}}}\right] =$$

$$\sum \frac{p_{io}q_{io}q_{i1}}{q_{io}+q_{i1}+q_{i2}} + \frac{\left(\sum \frac{p_{io}q_{io}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)}{\sum \frac{p_{i2}q_{i2}(q_{io}+q_{i1})}{q_{io}+q_{i1}+q_{i2}}}$$

Thus,

$$R_1 = \frac{\left(\sum \frac{p_{i1}q_{i1}(q_{io}+q_{i2})}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i2}(q_{io}+q_{i1})}{q_{io}+q_{i1}+q_{i2}}\right) - \left(\sum \frac{p_{i1}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)}{\left(\sum \frac{p_{io}q_{io}q_{i1}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i2}(q_{io}+q_{i1})}{q_{io}+q_{i1}+q_{i2}}\right) + \left(\sum \frac{p_{io}q_{io}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)}$$

$$= \frac{\left(\sum \frac{p_{i1}q_{io}q_{i1}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i2}(q_{io}+q_{i1})}{q_{io}+q_{i1}+q_{i2}}\right) + \left(\sum \frac{p_{i1}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{io}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)}{\left(\sum \frac{p_{io}q_{io}q_{i1}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i2}(q_{io}+q_{i1})}{q_{io}+q_{i1}+q_{i2}}\right) + \left(\sum \frac{p_{io}q_{io}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)\left(\sum \frac{p_{i2}q_{i1}q_{i2}}{q_{io}+q_{i1}+q_{i2}}\right)}$$

APPENDIX B
_____

SAMPLING CONSIDERATIONS AND THE USE OF THE GEARY KHAMIS METHOD IN ESTIMATING BILATERAL PURCHASING POWER PARITIES AT THE BASIC HEADING LEVEL.

1. This appendix considers the implications of applying sampling theory to the problem of estimating price relativities by the Geary Khamis method at the basic heading level in the bilateral case and suggests:

(a) how individual commodities should be sampled within basic headings;

(b) an alternative to existing methods for combining the selected price ratios to determine basic heading price relativities.

There are resource and time constraints to the numbers of items which can be selected for pricing within any one basic heading. The method of choosing the items to represent a basic heading and the manner of combining the prices of these items to estimate a price relativity between two countries are interdependent.

2. Suppose there are two countries, 0 and 1, and that within the basic heading being considered there are n commodities; the prices and quantities of commodity i are denoted by p_{i0} and q_{i0} in country 0 and p_{i1} and q_{i1} in country 1. The Geary Khamis purchasing power parity $_1PPP_0$, of country 1 relative to country 0, is denoted R_1, and the international prices determined simultaneously with R_1 are denoted π_i. Without loss of generality, country 0's currency is taken as the numeraire currency in which international prices are denominated.

It has been seen that

$$R_1 = \frac{\sum \dfrac{p_{i1}q_{i0}q_{i1}}{q_{i0}+q_{i1}}}{\sum \dfrac{p_{i0}q_{i0}q_{i1}}{q_{i0}+q_{i1}}} \qquad \text{------[1]}$$

3. The sampling problems associated with the estimation of the expression for R_1 are now considered. When the collection of information is restricted to a sample of commodities, the following problems have to be considered:

(a) how should the sample be drawn?
(b) what estimator should be used to determine the expression for R_1 ?

It is assumed here, to begin with, that the quantities q_{i0} and q_{i1} of the items in both countries are known. The unknowns,which are observed for the sample of commodities, are the prices p_{i0} and p_{i1}. In formulating the problem, country 1 has been taken arbitrarily as the numerator country and country 0 as the denominator. But the ordering of the two countries is arbitrary: and it seems a reasonable first condition to impose on the technique to be derived that it should satisfy the country reversal test; that is, the resulting price relativities should be independent of which country is used as the base country. Thus, if we want to find $_0PPP_1$, then we require the sampling and estimating techniques to be such that we have the same rules for drawing the sample of items as we would have in estimating $_1PPP_0$. Further, we require that the estimate of $_0PPP_1$ should be the reciprocal of $_1PPP_0$. If this criterion is accepted, then it has implications for the approach to estimating the expression for R_1, which is in the form of a ratio.

Now, if t is an unbiased estimator of a quantity τ; that is,

$$E[t] = \tau$$

then, in general, it will not be true that $1/t$ is unbiased for $1/\tau$. That is,

$$E[t^{-1}] \neq 1/\tau$$

This suggests that if we look for an estimator which is an unbiased estimator of the ratio in the expression for R_1, we will be breaking the principle of country reversal — since the reciprocal of the estimator will not be unbiased for the reciprocal of the expression.

This in turn suggests that rather than looking for an unbiased estimator of the ratio, we should look for a technique which will give, in some sense, equally good estimates of the numerator and denominator in the expression. In other words, in order to secure the country reversal principle, we should consider the problem of estimating the expressions:

$$\Sigma \ \frac{q_{i0}q_{i1}p_{i1}}{q_{i0}+q_{i1}} \qquad \text{and} \qquad \Sigma \ \frac{q_{i0}q_{i1}p_{i0}}{q_{i0}+q_{i1}}$$

separately.

4. Consider then, the problem of estimating the quantity

$$\Sigma \; \frac{q_{io} q_{i1} p_{i1}}{q_{io} + q_{i1}} \qquad\qquad \text{------[2]}$$

For simplicity, consider first of all the case in which we draw a sample of a single commodity, selecting commodity i with probability τ_i, and, having established the price of the commodity p_{i1}, we take as the estimate of [2] the quantity $w_i p_{i1}$, where the w_i are weights to be determined.

How should the probabilities, τ_i and the weights w_i be determined to give a "good" estimate of [2] ?

The problem can be approached as follows:

$$E[wp] \; = \; \Sigma \; w_i p_{i1} \tau_i$$

For the estimator to be an unbiased estimator of expression [2], therefore, the equality

$$\Sigma \; w_i p_{i1} \tau_i \quad = \quad \Sigma \; \frac{q_{io} q_{i1} p_{i1}}{q_{io} + q_{i1}}$$

must hold for all possible p_{i1}.

This implies that, for unbiasedness, the w_i and τ_i, should satisfy the relationship

$$w_i \tau_i \quad = \quad \frac{q_{io} q_{i1}}{q_{io} + q_{i1}} \qquad \text{for all i} \qquad \text{----- [3]}$$

Now consider the variance of the estimator wp :

$$V[wp] \quad = \quad \Sigma \; w_i^2 p_{i1}^2 \tau_i - (E[wp])^2$$

$$= \quad \sum \frac{(q_{io} q_{i1})^2 p_{i1}^2}{(q_{io} + q_{i1})^2 \; \tau_i} - E^2 \qquad \text{----- [4]}$$

on using the relationship [3] to remove w_i from the expression.

Expression [4] for the variance of the estimator wp would be minimised if we were able to take the τ_i, the probability of selecting commodity i, as proportional to

$$\frac{q_{io} q_{i1} p_{i1}}{(q_{io} + q_{i1})}$$

We cannot of course achieve this as the $\{p_{i1}\}$'s are unknown in advance. However, if the $\{p_{i1}\}$'s are relatively constant, in the sense that they do not vary greatly between different items within the basic heading, then we would expect to achieve an efficient estimator, that is, one with a small sample variance, if we took τ_i proportional to

$$\frac{q_{io}q_{i1}}{(q_{io}+q_{i1})}$$

In this case, it is easy to determine what the weights w_i should be. Since, from equation [3],

$$w_i\tau_i = \frac{q_{io}q_{i1}}{(q_{io} + q_{i1})}$$

the weights must be $w_i = 1$, for all i.

To summarise, the suggested procedure would be,

(a) select a commodity at random, with selection probability proportional to

$$\frac{q_{io}q_{i1}}{(q_{io} + q_{i1})}$$

(b) take as the estimator of the numerator of the expression in equation [1] the observed price of the selected commodity p_{i1}. Correspondingly, the estimator of the denominator is p_{io}.

The above argument has been developed in the simplest case, dealing with a sample of size one. When dealing with larger samples, the expression for the variance of the estimator, [4] above, becomes more complicated and will depend on the particular sampling scheme used, not just on the probabilities τ_i.

But similar principles apply, and this suggests that a reasonable approach when dealing with a sample size of n could be as follows :

(a') select commodities with probabilities proportional to

$$\frac{q_{io}q_{i1}}{(q_{io} + q_{i1})}$$

(b') estimate the numerator and denominator of [1] by Σp_{i1} and Σp_{io} respectively. That is, the estimator of R_1 based on the sample is

$$R_1 = \frac{\Sigma\, p_{i1}}{\Sigma\, p_{io}} \qquad\qquad ----- [5]$$

5. This approach certainly has some intuitive appeal. Both the sample selection stage and the form of the estimator are symmetric between country 0 and cou try 1, so the desirable property of country reversal is achieved. Moreover, since the sample selection probabilities are proportional to

$$\frac{q_{io}q_{i1}}{(q_{io} + q_{i1})}$$

it can be seen that an item which has high volumes in both countries has a high probability of being selected in the sample; an item which has a low volume in either country has a reduced probability; and a commodity which does not appear in one or other country has zero probability. These features appear emminently reasonable. However, further consideration suggests that the above approach has the following two basic weaknesses:

(a) the term

$$\frac{q_{io}q_{i1}}{(q_{io} + q_{i1})}$$

is denominated in the unit in which the quantity of commodity i happens to be measured. In other words, if commodity i happens to be measured in tons, then the scale of the term is measured in

$$\frac{tons^2}{tons} = tons$$

If the unit of measurement is then changed, for example to kilos, the value of the above term will change, and hence the probability of selecting commodity i in the sample.
Since, even within a basic heading, there need be no common unit for measuring the quantities of all the commodities involved, this means that the sampling scheme will vary depending on the arbitrary units for measuring quantities which happen to have been chosen.

(b) The second problem with the approach is that, in selecting a particular sampling scheme to reduce the variance of the estimator, (equation [4]), we make the assumption that the prices $\{p_{ij}\}$ did not vary greatly between different commodities within this basic heading. But prices themselves are a function of the arbitrary unit of volume which is being used. For example, if we change from expressing the quantity of a given commodity from units of tons to kilos, we will reduce the price per unit of volume by a factor of 1000. But since the choice of units for measuring the volumes of different commodities is essentially arbitrary, the assumption of relatively constant prices is certainly untenable. It has to be considered therefore, whether the procedure outlined above can be modified to overcome both of these difficulties.

6. We approach the problem by going back to the basic expression for the solution of the bilateral Geary Khamis PPPs, and expressing the formula in a more general form. The original formula is

$$R_1 = \frac{\sum \dfrac{q_{io}q_{i1}p_{i1}}{(q_{io} + q_{i1})}}{\sum \dfrac{q_{io}q_{i1}p_{io}}{(q_{io} + q_{i1})}}$$

Suppose that we take a set k_i of arbitrary non zero constants; then we can rewrite the expression for R_1 in the following identically equivalent form, on noting that the k_i terms cancel out.

$$R_1 = \frac{\sum \left(\dfrac{(k_i q_{io})(k_i q_{i1})}{k_i(q_{io} + q_{i1})} \cdot \dfrac{p_{i1}}{k_i} \right)}{\sum \left(\dfrac{(k_i q_{io})(k_i q_{i1})}{k_i(q_{io} + q_{i1})} \cdot \dfrac{p_{io}}{k_i} \right)} \qquad \text{---- [1']}$$

To overcome problem (5a) above, we want to be able to choose k_i in such a way that k_i varies inversely with q_{i1} and q_{io}, if the units in which the quantities are expressed are altered. To overcome problem (5b), we would like the quantities (p_i/k_i), and (p_{io}/k_i) to be relatively stable; and, in particular, to be invariant under changes in the units in which we measure the quantities of goods. Finally, to retain country reversal, we require the definition of k_i to be in some sense symmetric between country 0 and country 1.

One possible way of defining the k_i to satisfy these properties would be to take

$$k_i = 1/2 \, (p_{io} + \alpha p_{i1}) \qquad \text{------ [6]}$$

where α is an arbitrary constant.

It can readily be checked that the first two properties above are satisfied, as follows.
First of all, if there is a change in the unit in which commodity i is denominated, then there will be a corresponding reciprocal change in the prices p_{io} and p_{i1}, so that the quantities

$$1/2(p_{io} + \alpha p_{i1})q_{io} \, , \quad 1/2(p_{io} + \alpha p_{i1})q_{i1}$$

will be invariant.

Similarly, the quantities,

$$\frac{p_{i1}}{1/2(p_{io} + \alpha p_{i1})} \quad \text{and} \quad \frac{p_{io}}{1/2(p_{io} + \alpha p_{i1})}$$

are invariant under changes in the unit in which the volumes are measured.

We would also expect these latter quantities to be relatively stable: certainly a good deal more stable than the original $\{p_{io}\}$'s and $\{p_{i1}\}$'s.

7. An argument analogous to that used in paragraph [4], therefore, applied to equation [1'], suggests that we might approach the problem of Geary Khamis estimation in the following way:

(a) select a sample of commodities with probabilities proportional to

$$\frac{(k_i q_{io})(k_i q_{i1})}{k_i(q_{io} + q_{i1})}$$

where $k_i = 1/2 (p_{io} + \alpha p_{i1})$
and α is a constant, yet to be determined.

(b) take as the estimator of the PPP, the expression

$$R_1 = \frac{\Sigma \left(\dfrac{p_{i1}}{1/2(p_{i0} + \alpha p_{i1})} \right)}{\Sigma \left(\dfrac{p_{io}}{1/2(p_{io} + \alpha p_{i1})} \right)}$$

There are however two problems we need to solve before this becomes a practicable procedure. First of all, what value should we choose for the constant α in the definition of the k_i, in equation [6]. Secondly, the sampling weights depend both on the quantities q_{io} and q_{i1} and the unknown prices p_{io} and p_{i1}. How can these weights be estimated?

The following argument shows how one might be able to solve these problems if there were available an initial estimate of the price relativity between the two currencies, and of commodity level expenditures.

Suppose that α is an initial approximation to the price relativity between country 0 and country 1 for the basic heading being considered; and that to a first approximation

$$p_{io} = \alpha p_{i1} \qquad \text{----- [7]}$$

for each commodity, then,

$$1/2(p_{io} + \alpha p_{i1})q_{io} \doteq e_{io}$$

where e_{io} is expenditure on commodity i in country 0, and

$$1/2(p_{io} + \alpha p_{i1})q_{i1} \doteq \alpha e_{i1}$$

where e_{i1} is expenditure on commodity i in country 1.
In this case the sampling weights at (a) above are approximately

$$\frac{e_{io}\alpha e_{i1}}{(e_{io} + \alpha e_{i1})}$$

Equivalently, since the constant α in the numerator does not affect the relative proportions of these weights, we would sample commodities with probability proportional to

$$\frac{e_{io}e_{i1}}{(e_{io} + \alpha e_{i1})}$$

8. We have thus arrived at the following suggested procedure for estimating the Geary Khamis PPP. Suppose we have an initial estimate α of the country 0 to country 1 price relativity, and estimates e_{io} and e_{i1} of the expenditures of the commodities in the two countries. Then,

(a) one should select commodities with probabilities proportional to

$$\frac{e_{io}e_{i1}}{(e_{io} + \alpha e_{i1})}$$

(b) the estimate of $_1PPP_0$ should be taken as

$$R_1 = \frac{\Sigma\left(\dfrac{p_{i1}}{1/2(p_{io} + \alpha p_{i1})}\right)}{\Sigma\left(\dfrac{p_{io}}{1/2(p_{io} + \alpha p_{i1})}\right)}$$

As has been seen from the discussion in paragraph [6] above, both the sampling weights and the estimator are invariant under arbitrary changes in the units in which individual items are measured. Further, it is readily seen that the above procedure also satisfies the country reversal test, both as regards the determination of sampling weights, and the form of the estimator R_1.

9. The procedure outlined in paragraph [7] depends on the availability of initial, but possibly crude, estimates of the price relativity a, and of the expenditures on the individual items. In general some initial estimate of a will be available; for want of any more refined information we could take as an estimate of a the market exchange rate between the two currencies. If, however, some initial information is available on prices at item level, then a could be estimated directly from this.

The availability of estimates of expenditures at item level is likely to be much more problematical. However, in many cases, estimates will be available which will be tantamount to rough estimates of expenditure share. For example, if a classification of items into characteristic and non characteristic items has been attempted, this is likely to be based on some crude assessment of the overall expenditure shares of individual items - even if this only takes the form of an intuitive assessment of whether the expenditure share of individual items falls above or below some particular threshhold. Such crude estimates of broad shares of expenditure, together with the available data on total expenditure at basic heading level, can be converted into rough estimates of item level expenditures, which can then be utilised in the above technique.

Note that the condition in equation [7] means that the price structure within the basic heading has to be approximately multiplicative: (recall that the concept of multiplicativity was defined in chapter 1, section 4). We have not investigated how efficiently the sampling scheme suggested in this appendix performs when the price structure is very far from being multiplicative: but it is likely that its performance would be robust for moderate departures from multiplicativity.

References

1. Geary, R.C. A Note on Comparisons of Exchange Rates and Purchasing Power Between Countries. Journal of the Royal Statistical Society, 1958,Series A 121, pp 97-99.

2. Khamis, S.H. "A New System of Index Numbers for National and International Purposes." Journal of the Royal Statistical Society, 1972 Series A 135, pp 96-121.

3. Fisher, I. "The Making of Index Numbers", 1922, Boston, Houghton Mifflin.

4. Eltetö, O. and Köves, P. On a Problem of Index Number Computation Relating to International Comparison. Statisztikai Szemple, 1964 pp 507-518.

5. Szulc, B. Indices for Multiregional Comparisons, Przeglad Statystyczny, 1964 pp 239-254.

6. Drechsler, L. Regionalisation of the ICP and its consequences, Presented at the Expert Group on ICP Methodology, Luxembourg June 1988.

7. Kravis, Kenessey, Heston and Summers, A System of International Comparisons of Cross Product and Purchasing Power, 1975, Baltimore : The John Hopkins University Press.

8. Kravis, Heston, Summers, World Product and Income, International Comparisons of Real Gross Product, 1982, John Hopkins University Press.

9. Gerardi, G. for example, see "Selected Problems of Inter Country Comparisons on the Basis of the Experience of the EEC, Dec. 1982, Review of Income and Wealth.

10. "Comparison in Real Values of the Aggregates of ESA, 1980" 1983, Eurostat.

11. Hill, P.,Multilateral Measurements of Purchasing Power and Real GDP, Office for Official Publications of the European Communities, 1982.

ECONOMICS AND STATISTICS DEPARTMENT
WORKING PAPERS

**

DEPARTEMENT DES AFFAIRES ECONOMIQUES ET STATISTIQUES
DOCUMENTS DE TRAVAIL

**

In April 1983, the Economics and Statistics Department initiated a new series of economic studies entitled ESD Working Papers.

The following titles have been circulated:

1. Use of Demand Elasticities in Estimating Energy Demand (out of print)
   Utilisation des Elasticités de la Demande dans l'Estimation de la Demande de l'Energie

   Axel Mittelstadt

2. Capital, Energy and Labour Substitution: the supply block in OECD medium-term models
   Substitution du Capital, de l'Energie et du Travail : le bloc de l'offre dans les modèls à moyen terme de l'OCDE (épuisé)

   Patrick Artus

3. Wage Formation in France: sectoral aspects (out of print)
   Formation des Salaires en France : aspects sectoriels

   Patrick Artus

4. Service Lives of Fixed Assets (out of print)
   Durée de Vie Utile des Actifs Fixes

   Derek Blades

5. Resource Prices and Macroeconomic Policies: Lessons from Two Oil Price Shocks
   Prix des Ressources Naturelles et Politique Macro-Economique : Les Enseignements de Deux Chocs Pétroliers (épuisé)

   John Llewellyn

6. Output Responsiveness and Inflation: An Aggregate Study
   Souplesse de la Production et Inflation : Etude Globale

   David T. Coe and Gerald Holtham

7. The Determinants of Exchange Rate Movements (out of print)
   Les Déterminants des Mouvements des Taux de Change (épuisé)

   Graham Hacche

8. Simulated Macroeconomic Effects of a Large Fall in Oil Prices (out of print)
   Simulation des Effets Macro-économiques d'une Forte Baisse des Prix Pétroliers

   Flemming Larsen and John Llewellyn

9. Medium-Term Financial Strategy: The Co-ordination of Fiscal Monetary Policy (out of print)
   Stratégie Financière à Moyen Terme : la Coordination des Politiques Monétaire et Budgétaire (épuisé)

   Jean-Claude Chouraqui and Robert Price

10. Price Dynamics and Industrial Structure: A Theoretical and Econometric Analysis
    Dynamique des Prix et Structure Industrielle : Une analyse théorique économétrique (épuisé)

    David Encaoua (with collaboration from Paul Geroski and Riel Miller

11. Evidence on Income Distribution by Governments (out of print)
    L'Action Exercée par l'Etat sur la Redistribution du Revenu

    Peter Saunders

12. Labour Force Participation: An Analysis with Projections
    Taux d'Activité : Analyse et Projections

    James H. Chan-Lee

13. The Demand for Money and Velocity in Major OECD Countries (out of print)
    La Demande de Monnaie et la Vitesse de Circulation dans les Grands Pays de l'OCDE

    A. Blundell-Wignall, M. Rondoni and H. Ziegelschmidt

14. The Conduct of Monetary Policy in the Current Recovery
    La Conduite de la Politique Monétaire dans la Phase Actuelle de Reprise Economique

    Paul Atkinson and Jean-Claude Chouraqui

15. Structural Budget Deficits and Fiscal Stance
    Déficits Budgétaires Structurels et Orientation de la Politique Budgétaire

    Patrice Muller and Robert W.R. Price

16. Monetary Policy in the OECD Interlink Model
    La Politique Monétaire dans le Modèle Interlink

    A. Blundell-Wignall, M. Rondoni, H. Ziegelschmidt and J. Morgan

17. Real Gross Product in OECD Countries and Associated Purchasing Power Parities (out of print)
    Produit Brut Réel et Parités de Pouvoir d'Achat dans les pays de l'OCDE (épuisé)

    Peter Hill

18. The OECD Compatible Trade and Production Data Base (out of print)
    Base de Données Compatibles sur le Commerce et la Production de l'OCDE

    Derek Blades and Wendy Simpson

19. Nominal Wage Determination in Ten OECD Economies
    Détermination des Salaires Nominaux dans Dix Economies de l'OCDE

    David T. Coe and Francesco Gagliardi

20. Profits and Rates of Return in OECD Countries
    Profits et Taux de Rendement dans les Pays Membres de l'OCDE

    James H. Chan-Lee and Helen Sutch

21. Real Interest Rates and the Prospects for Durable Growth
    Taux d'Intérêt Réels et Perspectives de Croissance Durable

    Paul Atkinson and Jean-Claude Chouraqui

22. Energy Prices: Trends and Prospects
    Les Prix de l'Energie : Evolution et Perspectives

    Axel Mittelstadt

23. Changes in the Composition of Output and Employment
    Changements dans la Composition de la Production et de l'Emploi

    Axel Mittelstadt and Françoise Correia

24. Labour Market Flexibility and External Price Shocks
    Flexibilité du Marché du Travail et Chocs Extérieurs sur les Prix

    F. Klau and A. Mittelstadt

25. Discrepancies Between Imports and Exports in OECD Foreign Trade Statistics (out of print)
    Ecart entre les Importations et les Exportations dans les Statistiques du Commerce Extérieur de l'OCDE

    Derek Blades and Marina Ivanov

26. Aggregate Supply in Interlink: Model Specification and Empirical Results

    John Helliwell, Peter Sturm, Peter Jarrett and Gérard Salou

27. Commodity Prices in Interlink

    Gerry Holtham, Tapio Saavalainen, Paul Saunders and Helen Sutch

28. Exchange Rates and Real Long-Term Interest Rate Differentials: Evidence for Eighteen OECD Countries

David T. Coe and Stephen S. Golub

29. Method of Calculating Effective Exchange Rates and Indicators of Competitiveness

Martine Durand

30. Public Debt in a Medium-Term Context and its Implications for Fiscal Policy

Jean-Claude Chouraqui, Brian Jones and Robert Bruce Montador

31. The OECD Compatible Trade and Production Data Base 1970-1983

Anders Brodin and Derek Blades

32. The Formulation of Monetary Policy: A Reassessment in the Light of Recent Experience

Paul Atkinson and Jean-Claude Chouraqui

33. Mécanismes de Transmission et Effets Macro-Economiques de la Politique Monétaire en France : les Principaux Enseignements Econométriques

Marc-Olivier Strauss-Kahn

34. Pure Profit Rates and Tobin's q in Nine OECD Countries

James H. Chan-Lee

35. Wealth and Inflation Effects in the Aggregate Consumption Function

G.H. Holtham and H. Kato

36. The Government Household Transfer Data Base

Rita Varley

37. Internationalisation of Financial Markets: Some Implications for Macroeconomic Policy and for the Allocation of Capital

Mitsuhiro Fukao and Masaharu Hanazaki

38. Tracking the US External Deficit, 1980-1985: Experience with the OECD Interlink Model

Pete Richardson

39. Monetary Policy in the Second Half of the 1980s: How Much Room For Manoeuvre?

Kevin Clinton and Jean-Claude Chouraqui