International comparisons of output and productivity

Ark, Bart van

Chapter 3 - Unit Value Ratios for Industry of Origin Comparisons

Unit Value Comparisons by Industry of Origin

The Matching of Sample Products

International comparisons of unit values are the key element of the comparisons of real output and productivity in this thesis. As mentioned in the previous chapters, exchange rates cannot be used to convert output to a common currency. Neither are purchasing power parities (PPPs) derived from the expenditure side suitable for comparing value added by industry.

Average ratios of the ex-factory unit value were therefore compiled for sample products. Except for adjustments for quality discussed below, no use was made of specification prices such as those used in the ICP. As discussed at length in chapter 2, the main advantage of using unit values instead of specification prices is that the quantities and unit values are consistent with the total value of output.

My unit value ratios are of a binary nature. In most cases the United States is the \`numéraire' country, though in three cases comparisons were made between European countries with the United Kingdom as the base country. These detailed cross-country comparisons were only made for selected benchmark years. In some countries, for example in the United States, full-scale censuses which include product information are only available once in five years. The benchmark years for the comparisons included here are 1975 for Brazil, India and Mexico, 1984 for France and the Netherlands and 1987 for Germany, Korea, Japan and the UK.

The term \`unit value ratio' (UVR) is preferable to the more familiar expression \`purchasing power parity' (PPP) used elsewhere. The two are interchangeable, but for output comparisons the former identifies more clearly the nature of the prices I use. My \`prices' are unit values obtained by dividing the ex-factory sales value by the corresponding quantities obtained from each country's production census or survey.

The first step in estimating the unit value ratios was to match products between countries. The description of the products in the production censuses do not always make such comparisons straigthforward. For example, the production of bricks in one country may be specified in terms of cubic metres and for the other country in tons. In some cases, expert information from industry sources provided a way out of this problem, but in other instances the product match could not be made.

Table 3.1
Coverage of Unit Value Ratios in terms of Total Manufacturing Sales for Benchmark Years (1975, 1984, 1987), in %

Binary Comparison with United States	Own Country(%) (1)	United States (%) (2)	Number of UVRs (3)
1975			
Brazil/USA	27.9	22.9[a]	129
India/USA	19.4[b]	9.6[a]	108
Mexico/USA	31.8	22.8[a]	130
1987			
Korea/USA	36.7	21.0	192
Germany/USA	24.4	24.8	277
Japan/USA	20.0	19.9	193
UK/USA	15.7	14.3	170
Binary Comparison with United Kingdom	**Own Country (%)**	**United Kingdom(%)**	**Number of UVRs**
1984			
France/UK	13.1	9.4	102
Netherlands/UK	17.5	14.5	106
1987			
Germany/UK	21.4	21.9	236

[a] Original product data for the USA are for 1977.
[b] Original product data for India are for 1973/74.
Source: See appendix II. Korea/USA from Pilat (1991b); Japan/USA by Pilat from Pilat and van Ark (1992); Germany/UK from O'Mahony (1992a)

Secondly, for some products no information on sales values or quantities is reported by the census, generally because to do so would breach confidentiality. Thirdly, certain products have a unique character and are produced only in one country and not in the other (for example, super-tankers or guided missiles). Finally, a problem which will be dealt with in more detail later in this chapter, is that many products cannot be matched because they represent different qualities in terms of product content or performance.

Table 3.1 shows the coverage ratios in terms of a percentage of the total sales value and the total number of matches for each binary comparison in this study. Coverage ratios varied from 9.6 to 36.7 per cent of total sales, and was just over 20 per cent on average. There is quite some variation among branches. In some manufacturing branches, close to 50 per cent of sales or even more could be matched, but in other branches coverage was much lower, in particular in the machinery and transport equipment industries.[1]

The Aggregation Procedure

As it is not possible to match all product items in manufacturing, a method is required to fill the holes for the on average 80 per cent of output which could not covered by unit value ratios (UVRs). The aggregation procedure up to the level of total manufacturing was carried out in a number of stages.

The manufacturing sector was divided up in 16 branches, which roughly correspond to the 2-digit level of the International Standard Industry Classification (ISIC) of the United Nations.[2] For each binary comparison, a maximum number of industries within each branch were distinguished which produced the same products in both countries.

Product matches were made for as many products as possible within each industry. The average unit value ratio for the industry was obtained by weighting the unit values by the corresponding quantity weights of one of the two countries:

1 See the tables in appendix II for coverage ratios by branch.

2 See appendix I for the branch and industry classification used for this study.

$$UVR_{j(m)}^{XU(X)} = \frac{\sum_{i=1}^{s} P_{ij}^{X} * Q_{ij}^{X}}{\sum_{i=1}^{s} P_{ij}^{U} * Q_{ij}^{X}} \tag{3.1a}$$

at quantity weights of country X, and:

$$UVR_{j(m)}^{XU(U)} = \frac{\sum_{i=1}^{s} P_{ij}^{X} * Q_{ij}^{U}}{\sum_{i=1}^{s} P_{ij}^{U} * Q_{ij}^{U}} \tag{3.1b}$$

at quantity weights of country U (the USA or the UK).

i =1...s is the sample of matched items in matched industry j(m).

In some cases, the coverage percentage in terms of total sales within the industry was so low, that one could not reasonably assume that the UVRs were representative for the whole industry. On average, there were some 30 industries in each binary comparison for which at least 25 per cent of total sales were matched. These industries represented approximately 40 to 50 per cent of total value added in manufacturing.

For industries for which less than 25 per cent of output was matched, or for which no matches were made at all, the quantity weighted unit value ratio of all matched items in a branch were assumed to be representative for the unknown unit value ratio of a non-matched industry `j(n)' in that branch ‘k’:

$$UVR_{j(n)}^{XU(X)} = \frac{\sum_{i=1}^{s} P_{ik}^{X} * Q_{ik}^{X}}{\sum_{i=1}^{s} P_{ik}^{U} * Q_{ik}^{X}} \tag{3.2a}$$

at quantity weights of country X, and:

$$UVR_{j(n)}^{XU(U)} = \frac{\sum_{i=1}^{s} P_{ik}^{X} * Q_{ik}^{U}}{\sum_{i=1}^{s} P_{ik}^{U} * Q_{ik}^{U}} \tag{3.2b}$$

at quantity weights of country U (the USA or the UK).

The second stage of aggregation from industry to branch level is made by weighting the unit value ratios for gross output (UVR_{go}) as derived above by the value added of each industry in country X or country U, i.e.:

$$UVR_k^{XU(U)} = \frac{\sum_{j=1}^{r}[UVR_{j(go)}^{XU(U)} * VA_j^U]}{VA_k^U} \quad (3.3a)$$

for the UVR of branch k at quantity weights of country U, and:

$$UVR_k^{XU(X)} = \frac{VA_k^X}{\sum_{j=1}^{r}[VA_j^X / UVR_{j(go)}^{XU(X)}]} \quad (3.3b)$$

for the UVR of branch k at country X's quantity weights. In the final stage, branch UVRs were weighted at branch value added to obtain a unit value ratio for total manufacturing.[3]

The stage-wise aggregation using either quantities (in the first stage from product to industry level) or value added (in the following stages) has the advantage that the original product UVRs are successively reweighted according to their relative importance in the aggregate. At the end of this chapter the results of sensitivity tests with regard to different aggregation procedures will be presented. It appears that the difference between a stage-wise aggregated UVR and one which is directly build up from the product level using quantity weights of matched products, is largest for comparisons between countries with substantial structural differences. But even for these comparisons (for example, India versus the USA, and the Netherlands versus the UK), the difference in the UVRs is only just over 10 per cent.

Table 3.2 shows the UVRs for total manufacturing for each binary comparison in this study. The own country weighted UVRs are indexes of the Paasche type, whereas the base country weighted UVRs are Laspeyres indexes. Unit value ratios for countries with a similar structure

---

3 The treatment of 'non-matched' industries was slightly different in the earlier ICOP studies, including the 1975 comparisons for Brazil/USA, Mexico/USA and India/USA (see Maddison and van Ark, 1988; van Ark, 1991; see also Szirmai and Pilat, 1990). In these studies the value added-weighted UVR for matched industries was applied to non-matched industries. By using a larger sample of products for the non-matched industies, the average unit value ratio becomes less sensitive to individual matches.

of manufacturing output and employment are not very sensitive to these different weighting systems. However, in comparisons between, for example, India and the United States, the UVRs at US quantity weights are substantially higher than those at Indian weights. Because of the negative relationship between prices and quantities, an item with a relatively high price will be associated with relatively small quantities in the own country. The quantity weights of the other country (in this case the US) are therefore relatively large. As a result, if one weights a country's prices at US quantities, the unit value ratio will be higher than with quantities of the own country. This index number phenomenon is sometimes called the 'Gerschenkron effect', as Alexander Gerschenkron (1962) described it in detail in his analysis of relative backwardness in historical perspective.

The Fisher index, which is a geometric average of the Paasche and Laspeyres indexes, is mostly used in the remainder of this study. Compared to other biliateral index numbers, the Fisher index stands out relatively well in terms of certain index number properties. For example, in contrast to the Paasche and Laspeyres, the Fisher index satisfies the country reversal test (i.e. changing the denominator and numerator does not alter the results) and the factor reversal test (i.e. a Fisher price index times a Fisher quantity index gives a Fisher value index).[4] In addition, Diewert (1981) stressed some economic theoretic properties of the Fisher index, one of them being that it is a 'superlative' index number.[5] Another attractive property of the Fisher index compared to the Paasche or Laspeyres indexes is that when used for extrapolation of price indexes it tends to show a smaller margin of error from the 'true' measure in the year of extrapolation (Krijnse Locker and Faerber, 1984).[6]

Table 3.2 also shows the market exchange rate of the currencies. The ratio of the unit value ratio to the exchange rate gives an indication of relative price levels in each country. For the lower income countries, relative price levels in 1975 are clearly above those of the United States when the

4 The Paasche and Laspeyres indexes satisfy the 'weak' factor reversal test, which is that a Paasche price index times a Laspeyres quantity index give a value index which is identical to a Laspeyres price index times a Paasche quantity index.

5 Superlativity means that the index is exact for a flexible functional form, i.e. a function which closely approximates a class of other functions without having to know, or estimate, the parameters of the latter. For a relatively non-technical discussion, see Hill (1988). Recently Diewert (1992) also emphasised the usefullness of Fisher indices in relation to productivity studies.

6 See chapter 4, p. 81-82, for a more detailed discussion of this point. Multilateral weighting systems are discussed in more detail below.

UVR is weighted by US weights, whereas it is close to or below the US price level when based on own country weights. For 1987 relative price levels of the advanced countries are above those of the USA irrespective of the weighting system, apart from Korea. This reflects the relatively low exchange value of the US dollar in that year. For the European comparisons relative price levels are lowest in France and the Netherlands and highest in Germany.

Table 3.2
Unit Value Ratios for Benchmark Years, Total Manufacturing national currency to numéraire currency (1975, 1984, 1987)

Binary Comparison with United States	US Quantity Weights (1)	Own Quantity Weights (2)	Geometric Average (3)	Exchange rate (4)	Relative Price Level (US=100.0) (3)/(4)
1975					
Brazil/USA[a] (Cr/US$)	8.77	6.91	7.79	8.13	95.8
India/USA[b] (Rs/US$)	12.77	6.70	9.25	8.65	106.9
Mexico/USA[a] (Ps/US$)	15.60	11.97	13.67	12.50	109.4
1987					
Korea/USA (Won/US$)	848.73	576.80	699.60	822.60	85.0
Germany/USA (DM/US$)	2.25	2.16	2.21	1.80	122.8
Japan/USA (Yen/US$)	218.80	150.59	181.52	144.64	125.5
UK/USA (£/US$)	0.748	0.670	0.708	0.612	115.7

Binary Comparison with United Kingdom	UK Quantity Weights (1)	Own Quantity Weights (2)	Geometric Average (3)	Exchange rate (4)	Relative Price Level (UK=100.0) (3)/(4)
1984					
France/UK (FF/£)	11.29	10.70	10.99	11.68	94.1
Netherlands/UK (Dfl/£)	4.23	3.79	4.01	4.29	93.5
1987					
Germany/UK (DM/£)	3.56	3.44	3.50	2.94	119.0

[a] Original product data for the USA are for 1977, and were adjusted to 1975 at the industry level. See Maddison and van Ark (1988).

[b] Original product data for India are for 1973/74, and were adjusted to 1975 at the industry level. See van Ark (1991). For USA see footnote a).

Sources: See appendix II. Matchings for Korea/USA from Pilat (1991b); Japan/USA by Pilat from Pilat and van Ark (1992); Germany/UK from O'Mahony (1992a); exchange rates from IMF, *International Financial Statistics*.

An Assessment of the Unit Value Method

A Comparison with Alternative Converters

Table 3.3 compares my unit value ratios with PPPs for total expenditure from ICP and with 'proxy ICP PPPs' for expenditure on manufacturing products. The ICP PPPs in table 3.3 which are expressed in terms of national currencies to the US dollar are based on direct binary comparisons with USA.[7] For the European countries the ICP PPPs are weighted by multilateral European weights. The proxy PPPs are compiled on the basis of a set of PPPs for expenditure categories which mainly consist of manufacturing products, including food products, beverages, and tobacco, clothing and footwear, transport equipment and producer durables. These PPPs were weighted by value added derived from the production censuses to obtain proxy PPPs for total manufacturing.

For the lower income countries the expenditure PPPs are substantially below my manufacturing UVRs. This is caused by the fact that expenditure PPPs include comparisons of prices for services, which are relatively low in lower income countries. The manufacturing proxy PPPs for these countries are much closer to the UVRs.

Unit value ratios are a more appropriate indicator for price comparisons in manufacturing than the purchasing power parities from ICP which cover total expenditure. The latter are designed for expenditure comparisons, and most scholars actively involved in compiling these estimates refrained from using them for sectoral productivity comparisons.

Proxy PPPs serve no purpose and can easily lead to misleading results. Firstly, expenditure by category adds up to national income and not to domestic output. Although ICP makes an adjustment at the economy-wide level to arrive at GDP, expenditure prices for individual items include prices of imported products and exclude prices of exported items. Secondly, the PPPs include relative transport and distribution margins which are more difficult to take out. For example, one reason for the high ICP proxy PPP in Japan might be the relatively high distribution margins in Japan. Thirdly, PPPs are usually expressed at market prices, which may explain the relatively high proxy PPP for Germany, as it includes value added tax and excise duties. For comparisons of the

[7] Binary PPPs for 1975 are from Kravis, Heston and Summers (1982). Since ICP III binary PPPs have not been published anymore, but they were kindly provided by Eurostat.

Table 3.3
Comparisons of Unit Value Ratios, ICP Purchasing Power Parities and Proxy Purchasing Power Parities for Manufacturing

Binary Comparison with United States	Unit Value Ratios for Manufactu -ring (1)	ICP PPPs for Total Economy (2)	Proxy ICP PPPs for Manufac-turing (3)	Exchange rate (4)
1975				
Brazil/USA(Cr/US$)	7.79	5.40	7.77	8.13
India/USA (Rs/US$)	9.25	2.82	7.28	8.65
Mexico/USA (Ps/US$)	13.67	7.17	12.46	12.50
1987[a]				
Germany/USA (DM/US$)	2.21	2.57	2.64	1.80
Japan/USA (Yen/US$)	811.52	235.65	250.53	144.64
UK/USA (£/US$)	0.708	0.604	0.663	0.612

Binary Comparison with United Kingdom	Unit Value Ratios for Manufac-turing (1)	ICP PPPs for Total Economy (2)	Proxy PPPs for Manufac-turing (3)	Exchange Rate (4)
1984				
France/UK (FF/£)	10.91	12.77	11.83	11.68
Netherlands/UK (Dfl/£)	3.99	4.66	4.30	4.29
1987				
Germany/UK (DM/£)	3.50	4.23	3.63	2.94

[a] ICP PPPs for Korea versus the USA are not available.

Note: Proxy PPPs for manufactured products were obtained from the Fisher or multilateral average PPPs for the following categories: food, beverages and tobacco; clothing and footwear; furniture; household textiles and appliances; personal transport equipment and machinery and equipment. The PPPs were weighted at value added weights derived from each country's production statistics.

Sources: UVRs are geometric averages taken from table 3.2. PPPs for 1975 are 'augmented' binary PPPs derived from Kravis, Heston and Summers (1982). PPPs for 1987 (apart from Germany/UK) are Fisher binary PPPs for 1985 kindly provided by Eurostat, updated to 1987 on the basis of national deflators. 1987 Germany/UK PPPs from O'Mahony (1992a). PPPs for 1984 were obtained from multilateral PPPs at European weights from Eurostat (1988) for 1985, and backdated to 1984 on the basis of national deflators.

performance of production factors, value added should ideally be expressed at factor cost, i.e. excluding indirect taxes and including subsidies. This implies that output prices should be exclusive of indirect taxes as well. The fourth argument against the use of proxy PPPs is that these price ratios only refer to final expenditure items, and exclude price comparisons of intermediate goods. Finally, below the basic heading level (of which there are 151), ICP PPPs are unweighted and at basic heading level they are weighted by expenditure per capita. This may lead to quite different results from the output weights required for the purpose of this study.

The Quality Problem

The accuracy of the unit values used for the converters in this thesis depends to an important extent on the detail of product descriptions given in the censuses of each country. In practice unit values mostly represent an average price for a mix of product varieties which may be available in different proportions in two countries.

The expenditure approach uses specification prices for narrowly defined product items, which to some extent meets the product mix-problem. Despite this advantage of expenditure comparisons on the whole, the quality problem is not necessarily more serious in industry of origin studies. Firstly, quality differences are most important in consumer durables and investment goods, but less so for basic goods which represent intermediate stages of production. This latter group, which includes relatively homogeneous products such as paper, steel, cement, planed wood, etc., makes up a large share of manufacturing output but is by definition excluded from final expenditure comparisons.[8] Secondly, compared to specification prices unit values relate to a relatively large share of output and they cover the production of a whole year. In particular for comparisons between countries with a different structure in manufacturing, it is questionable how representative specification prices are of the total output in the countries.

8 This also explains why in the UVR comparisons, relatively high matching percentages of output were achieved in countries such as Brazil, Mexico and Korea (see table 3.1). In these countries homogeneous items are relatively more important than in the more advanced capitalist countries. On the other hand, vague descriptions of many product items in the censuses of lower income countries and the lack of a suitable product classification system seriously hampers comparisons for some industries, notably for investment industries. See also Beckerman (1966).

Even if the product mix-problem can be tackled, both approaches still face the second aspect of the quality problem, namely ‘product content’. This is related to the capacity of the product to perform certain functions, which are not easily observed from even the most detailed product description. For example, for a passenger car one can specify its physical characteristics, such as cylinder capacity, the type of fuel it uses, the number of gears and doors, whether it is equipped with a sunroof or not, etc.. It is more difficult to indicate the durability of its parts, the degree of safety of the car and its actual performance in terms of speed, braking distance, etc.. It should be emphasised that, from a conceptual point of view, ‘product-mix’ and ‘product content’ are not different. The distinction lies in the fact that even the most detailed product description will not pick up quality aspects related to product content.

There is an extensive post-war literature on the problem of quality differences in comparisons of real output and income, most of it in relation to time series, such as the retail price index and the producer price index. In recent decades the quality problem has shown a new dimension. Previously quality improvements were mostly reflected in a price rise, and the debate revolved around the question which part of the price rise should be interpreted as a quality increase and which part as a price increase. Presently, one of the major items in manufacturing, namely computers, has shown a continuous and very substantial price fall over the past two decades, which was largely caused by a continued supply of cheaper components with a higher performance.[9]

The problem of adjusting for quality differences is even more difficult for cross-country comparisons than for time series. Over time the quality of most products can be expected to increase along with real output, but

9 The early postwar literature on the quality problem was concerned with the debate what to view as quality change. Stone (1956), Denison (1957) and Gilbert (1961) suggested measuring only quality differences, which are proportional to the change in resource costs (or the price) of a product. However, Griliches (1964, 1971) argued that there are also quality differences which are non-proportional to the price of the product. According to Jaszi (1964) and Denison (1964a) many of these quality differences are related to the ‘user value’ rather than to the resource costs of the product, and should not be taken into account in real product comparisons. Nowadays there appears to be consensus that non-proportional quality changes should also be taken into account (Baily and Gordon, 1988). Gordon (1990) shows that the measurement of the fall in resource cost per ‘computer box’ results in a deflator which shows a much slower price decline than the change of the computer price per unit of ‘calculating power’.

quality differences between countries are not gradual. In particular when countries have traded off comparative advantages, relative quality advantages in one area of production may go together with quality backwardness in other areas.

There are basically two approaches to handle the quality problem. The first is the conventional method of comparing prices of 'matched' models, i.e. products which possess similar quality characteristics. The second is the hedonic pricing technique. Here a product is not matched directly, but considered as a bundle of quality characteristics. Each quality characteristic is considered as a premium on the price. This premium is derived by way of regression analysis. The hedonic technique has been applied in the US producer price index for computers since 1986 (Sinclair and Catron, 1990). It was also used by ICP for the estimation of PPPs of dwellings and cars (Kravis, Heston and Summers, 1982). A strong point of the hedonic technique is that it can pick up 'product mix' and certain 'product content' characteristics as described above. Its main disadvantage is that the results depend strongly on which quality characteristics are specified. Gordon (1990) pointed also at the problem of multicollinearity and the unclear relation between the characteristics within and outside the hedonic pricing model.

For the comparisons in this thesis, the conventional approach was adopted. In this respect one can distinguish between 'identical products', which have the same specifications and characteristics in both countries, 'common products', which serve the same purpose and have the same product name but with different specifications, and 'unique products' which are products available in one country and not to be found in the other country (Gilbert and Kravis, 1954, p. 79). For example, a steel product of a particular size or thickness and a specified carbon content is typically an identical product. Similarly, for cement one can assume quality differentials to be insignificant. A textile yarn made of a particular fibre may not be identical in terms of thickness compared to the yarn in the other country but it can still be taken as a common product. In our approach we included identical and common products, but in the latter case only when the product mix was judged to have a negligible effect on the unit value ratio.

The crucial assumption in the conventional approach is that the unit value ratio for the matched products is representative for that of the non-matched products, and it needs to be considered whether or not a bias may have occurred. For example, one could assume that the identical and common products included in the matches have a relatively 'low-quality content'. As a result relative prices of the matched products in the country with relatively low productivity levels may be too low, because the non-matched high quality items in the latter country are relatively scarce and are therefore produced at a relatively higher price. This downward bias in the UVR of the low-productivity country compared to the high-productivity country is reinforced by the fact that in case the matched items are not entirely free from quality differences in terms of product content, the relative price in the high productivity country is too high as it embodies an uncaptured quality premium on the price. The assump-tions of the conventional approach therefore imply that the productivity gaps between low-productivity and high-productivity countries which are presented in this thesis, are more likely to slightly understate rather than overstate the 'actual' productivity gap.[10]

Table 3.4
Quality Adjustment of Unit Value Ratios for Passenger Cars

Binary Comparison	Before Quality Adjustment (1)	After Quality Adjustment (2)	Ratio (2):(1) (3)
1975			
Brazil/USA[a] (Cr/US$)	3.97	4.97	125
India/USA[ab] (Rs/US$)	3.20	4.13	129
Mexico/USA[a] (Ps/US$)	9.13	10.94	120
1987[c]			
UK/USA (£/US$)	0.510	0.604	118
1984			
France/UK (FF/£)	8.16	9.02	111
Germany/UK (DM/£)	4.28	4.05	95

[a] Original product data for the USA are for 1977, and were adjusted to 1975 at the industry level. See Maddison and van Ark (1988).
[b] Original product data for India are for 1973/74, and were adjusted to 1975 at the industry level. See van Ark (1991).
[c] No quality adjustments were made for the comparisons between Germany and the USA and between Japan and the USA.
Source: As for table 3.2.

[10] Alternatively one can put forward the argument that due to the availability of high quality products in the high productivity-country, low quality products will be lower priced than in the low productivity-country, because they are regarded as old fashioned. However, this argument primarily relates to the consumer price of the products and not to their ex-factory cost price.

Passenger cars were the one product item included in these comparisons for which a quality adjustment was made using information from secondary sources. The production censuses of most countries only provide figures for the total quantity and sales value of passenger cars. Only the censuses in Germany, Japan and Mexico make a crude distinction between passenger cars on the basis of cylinder capacity. Information from industry and trade sources was therefore used to allocate the passenger cars in each country to four or five size categories on the basis of their cylinder capacity. It was not possible to obtain ex-factory prices for different cylinder categories, but trade sources were consulted to obtain retail prices for domestically manufactured models representing 'typical' models for each size group. On average 3 to 4 typical prices were collected for each size group. The average unit value for each group was then inferred from the average retail prices by category and the actual unit value for all passenger cars which was taken from the production census.

Table 3.4 compares the original unit value ratio for passenger cars with the unit value ratio after adjustment for quality differences. In the binary comparisons with the USA, the unit value ratio after adjustment for quality differences goes up, because of the relatively larger cylinder capacity of cars in the United States. For the France/UK comparisons the UVR also increases as France produces relatively more small cars than the United Kingdom.[11]

The Problem of Double Deflation

Industry of origin comparisons of real output and productivity face a major problem not encountered in comparisons from the expenditure side. This concerns the need to get UVRs for both the value of gross output (GO) and intermediate inputs (I). The UVR for value added of branch 'k' is then obtained as:

[11] No quality adjustments were made at this stage for the other binary comparisons. The procedure for the adjustments in the Brazil/USA and the Mexico/ USA comparisons was slightly different from that described above. See Maddison and van Ark (1988), Statistical Appendix (Notes).

$$UVR_{k(va)}^{XU(U)} = \frac{\sum_{j=1}^{r} [UVR_{j(go)}^{XU(U)} * GO_j^U] - [UVR_{j(i)}^{XU(U)} * I_j^U]}{VA_k^U} \tag{3.4a}$$

at quantity weights of country U, and:

$$UVR_{k(va)}^{XU(X)} = \frac{VA_k^X}{\sum_{j=1}^{r} [GO_j^X / UVR_{j(go)}^{XU(X)}] - [I_j^X / UVR_{j(i)}^{XU(X)}]} \tag{3.4b}$$

at country X's quantity weights.

The double deflation-method has been used in a number of output and productivity comparisons for agriculture, which is a sector characterised by a relatively simple input structure.[12] So far no cross-country comparisons of manufacturing output systematically applied a full-scale double deflation procedure. To convert intermediate inputs to a common currency, one needs separate UVRs for raw materials, fuels, electricity, and for industrial and non-industrial inputs. The coverage of inputs by UVRs needs to be high in particular for raw materials. In contrast to output prices one cannot assume that the UVRs for a few main inputs are representative for the other 'non-matched' inputs in an industry.

Some countries publish information on the value of the main inputs by industry, but quantity information is often lacking. For the United Kingdom and the United States figures on physical quantities of raw materials, packaging materials and energy inputs are provided at the (four-digit) industry level but only for a few main items. Paige and Bombach (1959) and van Ark (1990a) adjusted output UVRs for price differences of elec-tricity and fuel input, but these adjustments made only little difference to the results at the level of branches and for manufacturing as a whole.

Table 3.5 shows the results of an experiment with double deflation on the basis of input-output tables in the Netherlands and the United Kingdom for 1984. For domestic raw materials, output UVRs were used for the branches from which the inputs were used. Service inputs were converted with ICP PPPs and imported inputs at the official exchange rate. The double deflated UVRs for value added show very large fluctua-

[12] See, for example, FAO (1986), van der Meer and Yamada (1990) and Maddison and van Ooststroom (1993).

Table 3.5
Conversion Factors in Double Deflation Procedure of Manufacturing Output, Netherlands/UK, 1984, DFL/£

	Gross Output (DFL/£)		Intermediate (DFL/£)		Imports (DFL/£)	Value Added (DFL/£)	
	Neth quantity weights (1)	UK quantity weights (2)	Neth quantity weights (3)	UK quantity weights (4)	(5)	Neth quantity weights (6)	UK quantity weights (7)
Food Products and Beverages	3.72	3.94	3.69	3.79	4.27	3.14	4.16
Tobacco Products	2.50	2.93	3.84	3.82	4.27	0.82	1.59
Textiles	3.81	4.19	3.95	4.19	4.27	3.13	4.14
Wearing Apparel	4.78	5.14	4.28	4.32	4.27	6.46	6.37
Leather and Footwear	5.42	5.67	4.27	4.46	4.27	11.28	7.67
Wood Products	3.79	4.23	4.40	4.13	4.27	3.08	4.36
Paper Products	2.36	2.34	3.51	3.49	4.27	1.18	-1.08
Printing and Publishing	3.79	4.23	3.75	3.98	4.27	3.69	4.42
Chemicals	3.74	3.90	4.17	4.07	4.27	2.67	3.43
Rubber and Plastic Products	3.79	4.23	4.24	4.07	4.27	3.18	4.38
Stone, Clay and Glass Products	2.45	2.39	3.86	3.81	4.27	1.57	0.27
Basic Metals and Metal Products	4.40	4.46	5.85	4.32	4.27	3.75	4.88
Electric Engineering	3.79	4.23	4.22	4.29	4.27	3.32	4.17
Machinery and Transport Equipment	4.85	4.96	5.15	4.46	4.27	5.25	5.80
Instruments and Other Manufacturing	3.79	4.23	4.68	4.22	4.27	3.09	4.23
Total Manufacturing	3.79	4.23	4.06	4.11	4.27	3.07	4.31

Note: Gross output UVRs from appendix table II.10. Domestic raw materials were converted by gross output UVRs for branches from which inputs were obtained; imported raw materials were converted at exchange rate. Services were converted at ICP PPPs for specific services categories. The results shown here are only on Fisher-basis.
Source: UK from BSO (1988), *Input-Output Tables for the United Kingdom 1984*, London; Netherlands from CBS (1987), *Nationale Rekeningen 1986*, The Hague.

tions at branch level, though errors appear to cancel out at the level of total manufacturing.[13]

For better results with double deflation at a more disaggregated level one requires much larger input-output tables, and more specific information on the prices of intermediate inputs by industry. Such information cannot be obtained without separate surveys at firm level for individual product items.

Apart from practical data limitations, there are also certain methodological objections against double deflation. Firstly, value added UVRs at Paasche or Laspeyres weights can be far apart in particular if the share of intermediate inputs in gross output differs strongly between countries. Secondly, relatively small measurement errors in the price ratios of output or inputs tend to become magnified in the UVR for value added, in particular when intermediate inputs make up a large part of gross output.

Instead of applying an incomplete and unsatisfactory double deflation procedure, I followed the practice of earlier industry of origin studies, which derive the UVR for value added from the UVR for gross output weighted by the value added of the corresponding industry, as shown by equations (3.3a) and (3.3b) above. This method is called the 'adjusted single indicator' method, because although the product UVRs refer to the gross output level, it is adjusted for value added weights.[14] The method is based on the following assumptions:

1) at the product level, the value share of intermediate inputs in each unit of output is the same for all products within that industry and across countries.
2) the UVRs for inputs of industries and branches equal the corresponding UVRs for gross output.

Paige and Bombach (1959) defended the superiority of the adjusted single indicator method which 'although not so tidy and conceptually less satisfying' (p. 82) tends to provide more robust results than the double deflation method.

---

13 Szirmai and Pilat (1990) experimented with a similar kind of double deflation procedure for their Japan/USA and Korea/USA comparison for 1975, which also showed rather volatile results at branch level. See Frank (1977) for a partial double deflation procedure, which included fuels, electricity and raw material inputs.

14 This method is similar to what has been common practice in compiling wholesale- or producer price indexes in many countries, namely to weight the indexes of producer prices at the value added of specific industries (see, for example, Carter, Reddaway and Stone, 1948).

Binary versus Multilateral Weighting Systems

The unit value ratios presented in this thesis are all based on binary comparisons, with either the United States or the United Kingdom as the 'numéraire' or base country. In fact these binary comparisons take the form of a star comparison with the base country as the centre of the star. Comparisons between two or more countries representing points of the star can be made when using unique weights, for example the weights of the star country. However, as discussed above, the use of single country weights creates biases in one or the other direction. In the present study the binary results are expressed in terms of the Fisher index.[15]

Binary comparisons are characterised by some major index number problems of which the three most important are discussed here. Firstly, binary indexes are not transitive. In the present context this means that the unit value ratio between two countries does not equal the ratio of the UVRs between each of those two countries and a third country.

Secondly, binary indexes lack base country invariance, which implies that the results depend on the base country with which each country is compared. Base country invariance can only be achieved if the weights represent an average of all countries in the sample.

Finally, a binary index does not generate additivity (or matrix consistency). The requirements for additivity are twofold. If one conceives of an international comparison of output as a matrix with the columns representing the countries in the sample and the rows representing the products or industries, then each row should add up to the total value of output of all countries for one particular product or industry, and each column should add up to the total value of output in a country.[16]

The problems of transitivity, base country variance and additivity can be tackled by multilateral weighting systems. Multilateralisation is now

15 The comparisons in the OEEC studies (Gilbert and Kravis, 1954; Gilbert and Associates, 1958; Paige and Bombach, 1959) are also of a binary nature comparing each country on an individual basis with the United States. Gilbert and Kravis (1954) and Gilbert and Associates (1958) employed a rather primitive multilateral weighting system to obtain average European price weights. In each European country, products were priced in terms of US dollars. The average European dollar price for each product was then obtained weighting the dollar prices for each country at the national product in US prices.

16 Other index properties such as the factor reversal test and transaction equality are discussed in Kravis, Heston and Summers (1982). See also Pilat and Prasada Rao (1991).

common practice for all ICP studies.[17] Recently Pilat and Prasada Rao (1991) calculated multilateral indexes on the basis of industry of origin estimates from ICOP for the benchmark year 1975. Their study covers six originally binary comparisons with the United States, which include Brazil, Mexico, India, the United Kingdom, Korea and Japan.[18]

The first index variant used by Pilat and Prasada Rao is the Geary-Khamis method, which is also mostly applied by ICP. It derives average prices at a disaggregated level simultaneously with a PPP for the aggregate on the basis of two interdependent equations. In ICOP-terminology this implies that the average 'international' unit value, P_k, for each branch 'k' and the Geary-Khamis unit value ratio, $UVR_m^{Z(GK)}$, for total manufacturing 'm' of any country Z are derived on the basis of two interdependent equations:[19]

$$P_k = \sum_{Z=1}^{N} \frac{P_k^Z}{UVR_m^{Z(GK)}} [Q_k^Z / \sum_{Z=1}^{N} Q_k^Z] \qquad (3.5a)$$

and

$$UVR_m^{Z(GK)} = \frac{\sum_{k=1}^{t} P_k^Z Q_k^Z}{\sum_{k=1}^{t} P_k Q_k^Z} \qquad (3.5b)$$

where P_k^Z and Q_k^Z are the unit value and quantity of branch k in country Z.

For their sample of seven countries Pilat and Prasada Rao found that the Geary-Khamis index moves into the direction of or even beyond the Paasche VR. This is caused by the fact that the Geary-Khamis index is

17 A range of methodological studies on multilateralisation methods for ICP has appeared over the past decade, including Hill (1981), Kravis, Heston and Summers (1982), Ward (1985), Salazar-Carillo and Prasada Rao (1988) and Kurabayashi and Sakuma (1990).

18 Multilateral indexes were calculated at three different aggregation levels, i.e. at branch level, at industry level and at product level (for food products and chemicals). In the remainder of this section I will only deal with multilateralisation at branch level, which implies that the results below that level are still of a binary nature.

19 The terminology and sub-scripts of our equations are adjusted to that used for this study and different from the original ICP terminology.

dominated by the largest country in the sample, which affects the results in particular if the distance between the Paasche and the Laspeyres index is wide.

Other multilateralisation methods have been developed, some of them aiming to obtain results which are independent of country size. For example, the Gerardi-method derives the international unit value P_k on the basis of a simple unweighted geometric average of each country's unit value, adjusted for purchasing power. The Gerardi international price for each country is then derived as follows:[20]

$$P_k = \prod_{Z=1}^{N} [\frac{P_k^Z}{PPP}]^{1/N} \tag{3.6}$$

Pilat and Prasada Rao also show results for a multilateral version of the binary Theil-Tornqvist indexes. The binary Theil-Tornqvist UVR for manufacturing between two countries X and U, $UVR_m^{XU(TT)}$, is a geometric average of binary branch (Fisher) UVRs weighted at the average value share of the two countries in each branch:

$$UVR_m^{XU(TT)} = \prod_{k=1}^{t} [UVR_k^{XU(F)}]^{\frac{v_k^X + v_k^U}{2}} \tag{3.7a}$$

where $UVR^{XU(F)}$ is the Fisher UVR between countries X and U, and v_k^X and v_k^U are the value of branch 'k' in countries X and U

These binary index are multilateralised (and therefore made transitive) on the basis of a procedure developed by Eltetö, Köves and Szulc (EKS). This index aims to minimise the distortion between the original binary index between country X and U and the multilateral version, which can be seen from the following equation:

$$UVR_m^{XU(EKS)} = [UVR_m^{XU(TT)^2} * \prod_{Z \neq X,U}^{N} UVR_m^{XZ(TT)} * UVR_m^{ZU(TT)}]^{\frac{1}{N}} \tag{3.7b}$$

The main problem with the EKS procedure is that it does not provide full additivity, so that no UVRs for the sub-aggregates can be obtained.[21]

20 See Hill (1981, pp. 54-61) for a critical analysis of the Gerardi procedure. Compared to the Geary-Khamis method, one disadvantage is that the PPP is not simultaneously derived with the international price.

21 See Prasada Rao and Pilat (1991) for attempts to achieve additivity in the EKS system, but so far this has not produced satisfactory results.

I calculated the Geary-Khamis and the TT-EKS indexes to obtain a transitive unit value ratio between the Germany/USA, the Germany/UK and the UK/USA comparison for 1987. Table 3.6 compares these multilateral UVRs with the original binary UVRs taken from table 3.2. In contrast to the binary indexes, one can see that the Geary-Khamis index and the TT-EKS indexes produce a transitive result, as the actual and implicit UK/USA UVRs coincide (see the last two entries in the third and fourth row).

Table 3.6
Comparison of Binary UVRs and Multilateral UVRs for Manufacturing in Germany, the United Kingdom and the United States, 1987

	Binary UVRs			Geary-Khamis UVR	EKS Theil-Tornqvist
	Paasche	Laspeyres	Fisher		
(1) Germany/USA (DM/US$)	2.16	2.25	2.21	2.24	2.21
(2) Germany/UK (DM/£)	3.42	3.59	3.50	3.29	3.23
(3) UK/USA (£/US$)	0.670	0.748	0.708	0.680	0.684
(4) UK/USA - implicitly derived from (1)/(2) (£/US$)	0.619	0.627	0.631	0.680	0.684

Source: Binary UVRs from table 3.2; multilateral UVRs were calculated from binary branch results.

Despite the attractive properties of multilateral methods for comparisons between more than two countries, I have reservations about multilateralising the complete price system for the purpose of this study. As shown above there is no index number which can possess all desirable properties. The most important shortcoming of all multilateral methods is the loss of a very important property which binary index numbers possess, i.e. country characteristicity.[22] For a comparison between any pair of countries, the weights of the two countries themselves most adequately reflect the relative price structures. In particular if one is primarily interested in how each country's productivity compares to and catches up with the leading country, a comparison based on weights of third countries is less valid. Among the binary indexes, the Fisher index stands out relatively well in terms of its index number characteristics and economic theoretic properties, and it does not produce the biases which are inherent of the Paasche and the Laspeyres indices.

[22] The term was first coined by Laszlo Drechsler (1973).

Testing the Unit Value Ratios

One can of course question the realism of some of the assumptions and adjustments made above to derive unit value ratios. It is therefore necessary to analyse carefully the sensitivity of the unit value ratios to the various assumptions and adjustments. These tests were carried out for five of the ten binary comparisons included in this thesis.

My first sensitivity tests were aimed at checking the robustness of the average UVRs for the inclusion of UVRs for small products or for outlier UVRs. As can be seen from the UVRs in the country tables in appendix II the unit value ratios varied substantially between the branches. This appears also also from the coefficients of variation for the product UVRs in column (1) of table 3.7, which range from 0.26 in the France/UK comparison to 0.77 in the India/USA comparison.

One might infer that this large variation in product UVRs is caused by ‘outlier’ UVRs for relatively small products. However, it appears from columns (2) and (3) in table 3.7 that the coefficient of variation does not change much if one drops from the sample the relatively small items with a value of less than 0.1 per cent of total sales. This implies that ‘outlier’ UVRs, i.e. UVRs which are very high or very low compared to the average, are not just those of the smaller items.

In column (4) of table 3.7, ‘outlier’ UVRs which are more than 0.5 times the standard deviation below the mean of the full sample or more than one time the standard deviation above the mean are excluded from the sample.[23] Naturally, the coefficient of variation fell, but there was no statistically significant difference between the averages. So even if one is suspicious of ‘outlier’ UVRs it does not make much difference to the overall results. Of course these tests can be repeated for each of the 16 manufacturing branches. The UVRs will then be slightly more sensitive to the exclusion of outliers.

The second test is related to the aggregation procedure explained above. In column (1) of table 3.8 the product UVRs are directly aggregated to the level of total manufacturing weighted by their quantities. In column (2) an intermediate stage of value added-weights at industry level is included, whereas column (3) shows my preferred unit value ratios which are reweighted by industry- and branch value added.

---

23 The exclusion criteria are skewed, as the UVRs can never fall below zero, whereas at least in theory they can become many times higher than the mean.

Table 3.7
Testing the Sensitivity of the Unit Value Ratios to the Exclusion of Outliers

	All Unit Value Ratios (1)	UVRs more than 0.1% of total matched value own country (2)	UVRs more than 0.1% of total matched value base country (3)	UVRs less than 0.5*STD below mean or 1*STD above mean (4)
Germany/USA (1987)				
number of UVRs	273	131	141	153
arithmetic mean UVR	2.48	2.52	2.47	2.54
standard deviation (STD)	1.03	0.99	0.91	0.41
coefficient of variation	0.42	0.39	0.37	0.16
UK/USA (1987)				
number of UVRs	170	107	77	92
arithmetic mean UVR	0.755	0.737	0.737	0.767
standard deviation (STD)	0.28	0.24	0.24	0.11
coefficient of variation	0.38	0.33	0.33	0.14
Netherlands/UK (1984)				
number of UVRs	106	89	92	61
arithmetic mean UVR	3.984	3.866	3.914	4.097
standard deviation (STD)	1.22	1.21	1.26	0.47
coefficient of variation	0.31	0.31	0.32	0.12
France/UK (1984)				
number of UVRs	102	80	102	60
arithmetic mean UVR	11.457	11.337	11.457	11.613
standard deviation (STD)	3.01	3.05	3.01	1.02
coefficient of variation	0.26	0.27	0.26	0.09
India (1973/74)/USA (1977)				
number of UVRs	108	87	83	81
arithmetic mean UVR	6.379	6.485	6.138	5.958
standard deviation (STD)	4.90	5.15	3.88	1.88
coefficient of variation	0.77	0.79	0.63	0.32

Sources: see tables in appendix II.

Table 3.8
Comparison of Quantity-Weighted UVRs for Total Manufacturing with Value Added-Weighted UVRs for Benchmark Years

	Quantity-Weighted UVR (1)	Reweighted at Industry Level (2)	Reweighted at Industry and Branch Level (3)
Germany/United States (DM/US$) - 1987			
own quantity weights	2.06	2.10	2.16
US quantity weights	2.16	2.19	2.25
geometric average	2.11	2.15	2.21
United Kingdom/United States (£/US$) - 1987			
own quantity weights	0.643	0.664	0.670
US quantity weights	0.703	0.718	0.748
geometric average	0.675	0.690	0.708
France/United Kingdom (FF/£) - 1984			
own quantity weights	10.26	10.83	10.70
US quantity weights	11.21	11.27	11.29
geometric average	10.73	11.05	10.99
Netherlands/United Kingdom (Dfl/£) - 1984			
own quantity weights	3.42	3.61	3.79
US quantity weights	3.82	3.95	4.23
geometric average	3.62	3.78	4.01
India/United States (Rs/US$) - 1975			
own quantity weights	5.57	5.98	6.70
US quantity weights	11.99	12.46	12.77
geometric average	8.17	8.63	9.25

Source: see tables in appendix II

Table 3.9
Sensitivity Tests of Unit Value Ratio by Product Category

	All Unit Value Ratios (1)	UVRs Consumer Goods (2)	UVRs Basic Goods (3)	UVRs Investment Goods (4)
Germany/USA (1987)				
number of UVRs	273	187	69	21
arithmetic mean UVR	2.48	2.61	2.32	1.91
standard deviation (STD)	1.03	1.06	0.95	0.72
coefficient of variation	0.42	0.41	0.41	0.38
UK/USA (1987)				
number of UVRs	170	119	42	9
arithmetic mean UVR	0.755	0.743	0.814	0.640
standard deviation (STD)	0.28	0.26	0.34	0.29
coefficient of variation	0.38	0.35	0.41	0.45
Netherlands/UK (1984)				
number of UVRs	106	82	23	1
arithmetic mean UVR	3.984	4.058	3.719	4.038
standard deviation (STD)	1.22	1.29	0.95	0.00
coefficient of variation	0.31	0.32	0.25	0.00
France/UK (1984)				
number of UVRs	102	56	39	7
arithmetic mean UVR	11.457	11.682	10.879	12.876
standard deviation (STD)	3.01	3.28	2.42	3.08
coefficient of variation	0.26	0.28	0.22	0.24
India (1973/74)/USA (1977)				
number of UVRs	108	52	55	1
arithmetic mean UVR	6.397	6.365	6.490	2.962
standard deviation (STD)	4.90	3.56	5.91	0.00
coefficient of variation	0.77	0.56	0.91	0.00

Sources: see appendix II.

The table shows that the difference in UVRs according to the alternative weighting procedures is largest in the case of the India/USA comparison. The structure of the Indian and US industry is very different, which makes reweighting necessary in order to correct for products which are important in one country but unimportant in the other country. The geometric average UVR on the basis of the stage-wise aggregation procedure is more than 13 per cent above the product-weighted UVR. This confirms the observations made above concerning the quality problem, namely that the product sample in the low-productivity country is characterised by relatively low unit values. By reweighting this bias is reduced and the UVR increases. For the other countries, the unit value ratios also turn out to be slightly higher when based on the stage-wise aggregation procedure, but the differences are less than for the India/US case.

The conventional approach to the quality problem in this thesis has led to a relative overrepresentation of UVRs for durable and non-durable consumer goods in the product sample. Table 3.9 shows that, on average, some three-quarters of the sample consists of this kind of products, with the remainder covering basic goods and a limited number of investment goods. Although the average UVRs show substantial differences between the three sub-samples, there was only a statistically significant difference between the average UVR for investment goods and the overall manufacturing UVR for Germany versus the USA and for India versus the USA. This implies that one cannot speak of a systematic bias in our sample due to a relatively large number of consumer goods in the sample. In any event because of the stage-wise aggregation procedure described above, the impact of consumer good UVRs on industries which mainly consist of basic and investment goods is substantially reduced.

Conclusion

The conclusion of this assessment is that, at least for aggregates such as for branches and for manufacturing as a whole, the unit value method as applied here is sufficiently robust for obtaining appropriate indicators to convert output to a common currency. For comparisons at more disaggregated levels, such as for industries and products, a careful assessment is required in every case to assess quality differences and to evaluate the impact of different unit value relationships for inputs compared to output. This requires consultation of experts and trade sources. Some adjustments of this nature, in particular for passenger cars, have been included in this thesis, but further research is necessary to cover other goods as well, in particular for investment goods.

Unit value ratios, which are derived from the quantities and ex-factory sales value of products, are more suitable for industry of origin comparisons than ICP purchasing power parities. The latter are designed for expenditure comparisons, and as far as productivity comparisons are concerned only applicable for the economy as a whole. Our UVRs for manufacturing are clearly superior to proxy PPPs for expenditure on manufactured products. The latter include transport and distribution margins, reflect prices of imported goods, and are largely exclusive of information on intermediate products.

A particular strong point of the unit value method compared to the expenditure PPP method is the direct relationship between values and quantities. The unit values and quantities are also directly related to the concepts of gross output and value added which are used in real output and productivity comparisons which follow in the next chapter.