OECD OCDE

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

A World Free of Poverty

THE WORLD BANK GROUP

Using Scanner Data to Estimate Country Price Parities: An Exploratory Study

Cardiff Business School, Cardiff University – UK
Department of Economics, University of Pennsylvania- USA
Bureau of Labor Statistics - USA

Agenda item n° 5

JOINT WORLD BANK – OECD SEMINAR ON PURCHASING POWER PARITIES

Recent Advances in Methods and Applications

WASHINGTON D.C.

30 January – 2 February 2001

USING SCANNER DATA TO ESTIMATE COUNTRY PRICE PARITIES: AN EXPLORATORY STUDY

Saeed Heravi, Alan Heston**, and Mick Silver*•*
*Cardiff Business School, Cardiff University - UK
** Department of Economics, University of Pennsylvania - USA
• Visiting Senior Research Fellow, Bureau of Labor Statistics - USA

**We acknowledge the support from GfK Marketing UK Ltd for the provision of the data and Pat Barr in particular for his help. Support has been received from ** the National Science Foundation (NSF) and † the NSF/American Statistical Association and BLS. The usual disclaimers apply.*

Introduction

1. Experiments with scanner data assess their suitability for time to time indexes of consumer prices, estimate price differences across outlet types or by location of outlet and test theories of consumer and market behavior. It is a natural extension of this research to use scanner data to estimate the purchasing powers of currencies within or between countries, the focus of this paper. Many of the problems that arise in temporal price comparisons, for example relating to choice index number formula, have their counterpart in spatial price analysis.

2. The plan of the paper is to provide in Part A some background on how parities are typically estimated in traditional purchasing power studies such as the International Comparison Programme of the United Nations Statistical Division and by Eurostat and the OECD. The methodology for such estimates will be shown to be restricted by the limited nature of the data available. In Part B the estimation of parities using hedonic regressions is discussed as is a framework which uses superlative index number formulae which are exact for (correspond to) flexible functional forms for the underlying aggregator (Diewert, 1976 and 1978). Fixler and Zieschang (1992) and Feenstra (1995) have extended the framework to provide for hedonic quality-adjustments to such comparisons allowing for the compilation of superlative/exact hedonic indices, hereafter referred to as SEHI. In a recent, seminal paper Kokoski *et al.* (1999) have applied this framework to interarea price comparisons within the U.S. In this study we adapt the framework for use with scanner data to compare prices across countries.

3. The framework is important for two reasons. First, scanner data is a rich source with an extensive coverage. Because the coding of the same or similar items may vary across countries, matching is unsuitable for such spatial comparisons since much of the data will be lost through mismatched comparisons. However, because scanner data includes information on the characteristics of the items, an hedonic formulation is most suitable. Second, scanner data includes information on the unit values, volumes and thus values to provide weights at a detailed level below basic headings. The theoretical frameworks referred to above provide the means for suitable superlative aggregators to be used to compile such indices. This framework, as will be discussed, extends to multilateral as well as bilateral comparisons. The purpose of this paper is to explore alternative approaches to inter-country price comparisons using scanner data and provide a suitable framework that preserves the integrity of the comparisons, without undue loss of the richness of the data source.

4. Part C presents an application using scanner data from different types of outlets on a variety of models of television sets in France, the Netherlands and the UK to estimate price parities for television sets (TVs). The data set is rich in the coverage and scope of transactions as well as the extensive detail available on unit values, volumes (and thus weights) and characteristics for each individual model of TV sold. It thus permits us to examine alternative formulae and approaches to the estimation of such parities. While the study is limited to one consumer durable, such data are available for a wide range of products, the results here being illustrative of a proposed methodology for a rich data source. Part D concludes.

Part A. BACKGROUND ON SPATIAL PARITIES

1. Some Terminology

5. We will follow a convention that reserves the term PPP for an aggregate purchasing power parity, say over consumption or GDP, and use the term ‘detailed heading parity’ for the first level comparison made from price ratios for specified items. This usage is roughly parallel in reference to elementary aggregates of CPIs. As in temporal indexes, where usually the ratio of item prices in t and t+1 are the starting point, spatial comparisons usually begin with the ratio of item prices in different countries to a numeraire country or an average of the countries as in the EU. A basic heading parity is an average of these ratios, or the ratio itself if only one item represents a heading, as might happen in the case of eggs. In the

International Comparison Programme (ICP) expenditure weights exist for a basic heading, and below this level weights are not available or used.

6. Thus in the empirical application our concern is with the estimation of the basic heading parity for TVs across countries. In scanner data there are of course quantity and, through unit values, value weights for individual models of TVs. However, we maintain for clarity the terminology of a detailed heading parity below the basic heading level for each individual model or variety of TV in spite of the existence of information for weighting. Before discussing hedonic methods let us quickly review the usual ways parities have been generated at the basic heading level.

2. The Two Country Case

7. The binary case is important for at least two reasons. First, binary comparisons have been taken as the model for both two country comparisons as well as multilateral comparisons using the star system of linking a country to one or several nodes. The accepted method of binary comparisons has been to compute a geometric mean of the price ratios within a basic heading parity. This geometric mean is used as the estimate of the basic heading parity, $pp_{jk,}$ between country *j* and *k*. Usually no weights are attached to the individual price ratios, but weights can be used and Fisher's formula is typically applied.

8. A second important reason to look at the binary case is that it has been used by the EU as a basis for developing multilateral estimates of basic heading parities. The EU practice is for countries to distinguish over a basket of several thousand specified goods and services between items regarded as *representative* (also formerly denoted as '*' items), other less important that are *priced*, and items *not priced* by a country. An item not priced is clearly even less representative for a country than items that are priced in the basic heading. Between any pair of countries, only representative items for country *j* will enter into $pp_{jk,}$ and for country *k*, only its representative items will enter into pp_{kj}. So an item may be representative in country *j*, but not in country *k*, where it will be priced and used for the pp_{jk} comparison, but not used for the pp_{kj} comparison. These become the inputs for the EKS estimation that generate a transitive set of parities for all the countries at the basic heading level. Should country *k* and *j* not price items that are representative in the other country, their estimated pp_{kj} will be based upon their indirect comparisons through third countries. Both the 0,1 weighting system and the willingness to not use some price information (that is if an item is not representative in both countries, it will not enter into their binary comparison) have been questioned. However, the EU method is an interesting way to combine binary and multilateral techniques. As will be outlined in Part C, scanner data provides a much richer data source for such comparisons.

3. Multilateral Comparisons: The Usual Suspects, CPD and EKS

9. A variety of approaches to multilateral comparisons exist as outlined and discussed in Balk (1996), Diewert (1999) and Hill (1999). If each country prices each item within a basic heading and the geometric means of the ratios are calculated, these will be transitive across the countries. In the early comparisons in Latin America this condition was imposed on the countries, often producing strained comparisons, particularly between Caribbean rim countries and the Mercosur countries. The more usual case is where there is an incomplete price tableaux; for example, consider a heading there are 8 items, and 20 countries, but only 100 of the possible 160 prices are supplied by the countries. Then a directly estimated pp_{kj} will not in general equal an indirect pp_{kj} derived from pp_{kl}/pp_{jl}. This led to the development of multilateral approaches that would take account of the fact that not all prices would be available for all items within a basic heading. The two approaches most commonly used to estimate heading parities in ICP benchmark comparisons have been the CPD and EKS methods. They both have the property that if each country prices all items, their estimated pp_{jk} will be the same and equal to that obtained in the binary case.

a) The EKS Method

10. The EKS method permits transitivity to be attained taking into account the indirect and direct comparisons by :

$$(1)\quad pp_{jk} = \prod_{l=1}^{n} [pp_{jl} / pp_{kl}]^{1/n}$$

11. In EKS the direct parities (pp_{jk}) where j=k) and (pp_{kj}, where k=j), i.e. the price in country j is also available in country k are both counted, while each indirect parity is counted once. Clearly the weight of the indirect comparisons increases very rapidly as the number of countries increases, a problematic property.

12. The EKS method produces transitivity making use of both direct price comparisons between each pair of countries, as well as all indirect price relationships between each pair of countries and the remaining countries. The EKS method is derived from a minimization procedure that is basically mathematical in formulation. Because it builds up from Fisher comparisons it is often thought these properties somehow produce desirable multilateral properties in multilateral comparisons.[1] The CPD method that follows is derived from an explicit model of how the price tableau is generated.[2]

b) The Country-Product-Dummy (CPD) approach

13. An alternative way of dealing with an incomplete matrix of prices is the Country-Product-Dummy (CPD) procedure developed by Robert Summers [3]. It was employed in the ICP calculations for the initial studies, although in recent years most regions have used the EKS method. It is a basic hedonic regression model and it is a simpler version of what has been used for temporal price indexes using scanner data. In (2) below j = 1,2,...,m for the

$$(2)\ \ln p_{ij} = \sum_{i=1}^{n} \beta_i z_{ij} + \sum_{j=2}^{m} \alpha_j D_j + \varepsilon_{ij}$$

1 It has also been argued that if EKS is used at the basic heading level it should also be used to aggregate the basic headings to larger aggregates, though the basis for this assertion, other than symmetry, is not clear. Diewert (1999) provides a very valuable discussion of the properties of EKS and other aggregation methods.

2 In the EKS method, the importance assigned to individual price observations is variable and is not self evident. Even when the "*" system is used the importance assigned to individual prices will depend on the number of observations and on whether they are starred. Further, the weight given to indirect price ratios will normally not depend on the number of price observations involved unless explicit weighting is employed. This last problem is also present in the CPD method where it has been explicitly treated by assigning the same weight to each country so that each price for a country will receive a weight inversely related to the total number of prices for a basic heading for the country. While this could be done in EKS, in applications of EKS it has not been carried out. Also both systems can explicitly handle weights, and in some applications this has been done.

3 See R. Summers (1973). In Summer’s formulation the missing prices in the tableau are assumed to be randomly distributed. Other assumptions are also possible.

14. m countries, $i = 1,2,\ldots n$ for the number of items in a basic heading, and p_{ij} is the price of item i in country j, and $\bullet_{ij}$ is the error term. The prices are regressed against the two sets of dummy-variables: one set contains a dummy for each country, $\bullet_j$, other than the numeraire country, and the second set with a dummy for each item specification, z_i.

15. The transitive parities, pp_j are derived from the coefficients of the country dummies. The country coefficients, the β_js, are the logarithm of the estimated country parity for the heading relative to the numeraire country which is taken as 1. The item coefficients, the β_js are the logarithms of the estimates of the average item price in the currency of the numeraire country.

16. The CPD method in effect produces an estimated complete price matrix for a basic heading, and the country parities are the GEOMEAN of the ratios of country prices to those estimated for the numeraire country. The estimated average prices given above, the $\bullet_j$s, are of considerable interest in themselves since they are an estimate of the average price for each specification in the currency of the numeraire country across the group of countries. In regional comparisons, for example, a bi-product of the application of the CPD method in the ESCAP region in 1985 was a set of Asian item prices. These CPD prices provide a basis of comparison for any country in a region of their item prices for matching specifications with the average prices from a CPD regression for the region. They may also be used to link a country that did not participate in the benchmark comparison to an existing ICP study. This exercise was carried out as a non-official research exercise for Taiwan based on the CDP average ESCAP prices for 1985. (Yotopolous and Lin, 1993). The item prices can further be used when prices only become available at a date later than the initial CPD for a region. This occurred in the 1993 ESCAP comparison in the case of Laos and Malaysia.[4] (UNESCAP, 1999) One other example is the U.S. interarea comparisons by Kokoski, Moulton and Zieschang (1999). The BLS entry level item sampling technique for the consumer price index lends itself to CPD because many of the characteristics of an item are yes/no type entries. City parities for aggregates are then built up using a weighted Törnqvist index. When comparisons of the CPD and the EKS have been carried out they are quite close together when the price tableau has few holes. Unfortunately the two indexes are further apart when most needed, that is when there are many holes in the price tableau of item and country prices for a basic heading.

c) Hedonic Methods and Multilateral SEHI

17. For most headings the selection of items to be compared has involved a great deal of judgement because quantity weights are not available and because there is really no satisfactory theory of how to sample for spatial price comparisons. The identity of items has usually been defined in terms of brand name, model number and the like. But as the following quotation makes clear, there is more to comparing like with like than just trademarks.

> "In particular, identity means that (1) the size of the good/service should be the same. As mentioned above, this requirement has a double meaning: not only that the unit price reported relates to the same size (for example the price of one kg of potato) but also that the price originally observed should relate to approximately the same purchased quantity in each country. It would not be appropriate, for instance, to observe the price of sugar in 1 kg package in country A while observing it in a 10 kg package (and dividing it by ten) in country B, since the price of a larger size package covers relatively less distributive services (per kilo) and is, typically, relatively lower. (2) the physical and functional properties should be the same (for instance, the thread count of fabrics, the capacity of machines, the life of electric bulbs). This relates to all such properties that may have significant influence on the price of the given product. Shape and colour may also be

4 It is possible to link the country to a regional or world comparison in the following way. First the item prices within a basic heading for such a country would be divided by the CPD estimates of the same item prices in a numeraire currency; then the geometric mean of these price ratios would be calculated to provide the basic heading parity that would allow the country to be linked to the regional and world comparison.

> relevant for some products, although not for others. Types of outlet should ideally be the same when matching items across countries (this is one of the more important conditions that are discussed further in the following section.(3) delivery conditions (for example package, warranty, transportation cost included or excluded) should also be the same. This is particularly important for producers and consumers durables (4) other circumstantial factors (for example type of restaurants for consumption of specified dishes, access to repair services) should be the same insofar as they have a significant effect on the price." (United Nations, 1992)

18. All of these guidelines suggest the desirability of using as many objective criteria as possible in comparing items. Clearly, hedonic regressions are appropriate when there is a strong relationship between objective quality characteristics and the market price. In practice hedonic estimation has been carried out for appliances, producer durables like pumps or motors, automobiles and housing for a number of ICP benchmarks.

19. The use of a characteristics' approach is not without critics. Pieper (1999) for example, commenting on Kokoski *et al.* (1999), argues that:

> "Hedonic quality adjustments are necessary in the authors' data set not because the goods themselves differ across regions but because the product definitions are imprecise. Thus it would be possible to control for quality differences directly, without the use of hedonic regressions, by narrowly defining the good, for example, Red Delicious apple, unpackaged and sold at a chain grocery store, versus just apples. The advantages of a narrow definition are a better control for quality differences and lower computational costs, whereas the hedonic method has the advantage of allowing a greater coverage of data. It is probably not possible to construct area price indexes using narrow product definitions with the author's data set because there are likely to be only a few observations in a given area at the narrowest level of product definition." (Pieper, 1999:168).

20. It is the last sentence that is particularly pertinent to this paper. If, for example, the price of a specific model of television set (TV) could be priced across all areas, then a case could be argued for the comparing of like with like via computational matching, as opposed to partialling out differences in a regression equation. However, this would restrict the sample to a set of narrowly defined items whose existence could be guaranteed in each area. For intercountry comparisons the restriction would be quite severe for many products, and unduly so when an hedonic framework would adjust for any such differences. Moreover, scanner data, with its extensive coverage of transactions, requires a framework in which data are not thrown out because corresponding items cannot be computationally matched due to their being not sold in other countries, or the idiosyncrasies of the model codes used across countries. The SEHI framework allows scanner data's information on weights and characteristics to be used in the measurement of intercountry price comparisons for bilateral and, following Kokoski *et al.* (1999), multilateral comparisons both benefiting from the impressive coverage of the data as, explained in the next section.

21. Before considering these approaches we highlight some of the benefits from using the characteristics of goods for quality-adjusted price comparisons by reviewing some of the past practice, particularly for housing and automobiles.

d) The Characteristics Approach (I) House Rents

22. Household rents are a significant percent of consumer expenditures in most countries and appropriate quantity comparisons across countries, or even between regions of a country are difficult. There are direct quantity comparisons within and between countries that look at persons per room or per unit area of housing. But variation in housing quality within and between countries is so large that such a comparison on the basis of housing census data would not be comparing like with like. It is possible to use

hedonic regression estimates for countries that routinely conduct rental surveys, and this has been done for some countries in the ICP benchmarks.[5]

23. For the ICP benchmarks price determining characteristics for rental housing are size of dwelling, availability of amenities like water, electricity, central heating, bathroom facilities, and usually age of dwelling. There are many such studies *within* a country including Ramussen and Zuehlke (1980) and Mills and Simenauer (1986). The monthly rent is regressed against the above amenities, size and age of dwelling, and usually additional survey variables that may be particular to each country, including any rent control information. Pooled regressions are not possible for all countries because many countries do not have surveys, but only rental rates for certain types of accommodation. Even for those countries with surveys, important variables were often not comparable, like floor area versus number of rooms. Where available, additional variables, like location, if dwelling is furnished, rent controlled or subsidized, availability of garage, and air-conditioning were included in the estimating equations for individual countries to sharpen estimates on the two most important variables, floor area and number of bathrooms.

24. This points to a general problem in spatial comparisons, namely that there is much more variation in country provision of price information than there is in year to year information to a country to produce its CPI. In the rental case, for the ICP benchmarks a second stage was added to the hedonic regressions. To obtain parities for all countries, some 60 or so rent cells were created that were combinations of the age of dwelling, availability of water, electricity, plumbing, and central heating, and the area of an unfurnished dwelling space. The weights for these cells came from national housing censuses and if they were at least 3% of the housing stock, countries usually provided rental estimate either directly obtained or estimated from the hedonic country regressions. The overall parity for rental housing was then obtained from a weighted CPD regression, though weighted EKS could also be used.

25. What should be noted about the handling of rents and automobiles is that quantity weights are explicitly introduced below the basic heading level. And in terms of the discussion below, this procedure that has been carried out in two stages, is analogous to the Superlative Exact Hedonic Indexes (SEHI) developed below when the aggregation of the cells is based upon a superlative index like a Fisher or Törnqvist index.

(ii) The Characteristics Approach: Automobiles, Other Durables, and Price-slope adjustments

26. The methods used in the ICP for automobiles were very similar to those for house rents. Generally a grid of horsepower and weight cells defines a number of automobile models and within each cell from one to four actual models are described. Countries would either provide prices for those models based on their importance using sales in the benchmark year or the prices might be derived from hedonic regression equations or other survey data. Sometimes list prices would be used for hedonic regressions and an overall adjustment would then be made for the average national discount factor for various classes of cars, for example, the percentage markdown from list prices is typically larger for more expensive cars. As in the case of house rents, there are weights associated with each automobile price that are taken into account in building up the basic heading parity. Again to determine an overall parity a weighted CPD or EKS, or other procedure could be used.

27. It was found for most appliances that usually one or two factors were sufficient to explain price variation, such as total liters and freezer size in the case of refrigerators. But in a number of other cases 'price-slope adjustments' were adopted to modify the actual prices to meet a specification, these adjustments being based on coefficients from an hedonic regression. Such procedures can be grouped into two types, those arising from the size of package or purchase, and those arising from the technical capacity of the item.

28. Size adjustments arise because the price per unit of weight usually declines with the size of purchase. Often choosing the correct unit of comparison for an item is in effect a price-slope adjustment.

[5]. These procedures are described more fully in Kravis, Heston and Summers, World Product and Income, pp. 54-9.

Much of price variation for household textiles and floor coverings may be explained by specifying price per square meter for a given material of a given weight. For example, cotton towels come in a bewildering variety of sizes and weights, but the main element in price is the physical quantity of toweling. This is often standardized by using an approximate range of towel sizes of a given weight per square meter and converting the collected price to a price per square meter.[6]

29. Price-slope adjustments have been frequently used for consumer durables such as refrigerators, air-conditioners, heating appliances, fans, TVs and the like. Where the same brand and model are widely sold in different countries simple comparisons of the same item are appropriate. However, a company uses often the same model numbers in different countries while the technical specifications differ, or almost identical models are given different model numbers in different countries precluding direct matching. When price variation can be readily explained by the characteristics of the items, hedonic estimation, or other such adjustment may be appropriate. For example, most price information about room air conditioners is contained in the number of BTUs of capacity. This is a case where it may take more resources to match air conditioner across countries than to use the relationship of price to BTUs. Price per BTU declines with capacity, but the gradient allows for the adjustment of actual prices to a normalized price for a standard size.

30. This discussion points to the obvious advantage of having objective adjustments, such as those provided by hedonic regressions, for many items. And this leads to the question of whether the use of scanner data across countries would provide a feasible and efficient way to estimate parities for basic headings.

Part B. SCANNER DATA AND PARITIES

31. Scanner data benefit from having an extensive coverage of transactions over all items sold as opposed to a selection of varieties of selected representative items in a sample of outlets. The coverage extends to all transactions in the period as opposed to a survey date and takes transaction as opposed to display prices. For individual varieties or models within a basic heading data are available on the unit value, volume and thus values for weighting at this lower level of aggregation. Data are also available on the characteristics of each variety or model allowing hedonic regressions to be estimated using this extensive data. The use of scanner data for the basic headings of CPIs has moved from the advocacy stage to the stage of serious research about costs and benefits.[7] In the first section of this part we take up the some general questions about the applicability scanner to purchasing power comparisons compared to its use for the CPI. In the second section we discuss some models for using scanner data for estimating basic heading parities.

1. Similarities and differences between CPI and ICP applications of scanner data

32. One general point of difference between CPI and PPP applications of scanner data should be noted, namely, statistical standards are much higher for the former than the latter.[8] There are practical

---

6 Differences in packaging for the same unit of purchase of the same commodity often do not affect relative prices across countries. Frequently wage rates, recycling legislation, or relative costs of paper will lead one country to adopt one form of packaging as less expensive, while another form of packaging may be cheaper in another country. If volume sellers use different packaging in the two countries it is a difference that can usually be ignored.

7 An early advocacy paper was by Silver (1995). Subsequently, much more work was done, much of it reported at a Conference on Income and Wealth organized by Robert Feenstra and Robert Shapiro held in Washington, September 15-16, 2000. This conference was devoted to use of scanner data for CPIs and included a number of experiments being carried out by statistical offices of Canada, the U.K., and the U.S.

8 The point was brought home forcefully to Irving Kravis when he was engaged with Milton Gilbert and others in the lst OEEC purchasing power comparison. Kravis was familiar with the Paasche-Laspeyere spreads common

reasons for this since billions of dollars, for example, ride on a small change in the CPI in the United States, and in other countries too. There are also payroll implications to spatial differences in purchasing power, for example, in the U.S. the Federal government pays over $300 million a year in cost of living allowances to employees outside the continental 48 states. However, the error involved in getting the spatial price index between Washington and Anchorage wrong by say 2% in a given year, is trivial compared to a similar error in the CPI. The same is generally true for most other spatial versus temporal uses of price indexes. Two important considerations flow from this observation.

33. First, consider one limitation of scanner data, namely that it may not be available from all retail chains that sell a given item. This might well be a matter of concern for CPI estimation. However, for comparisons of parities for the same heading between countries this is likely to be a second order consideration, especially if information is available on where the source of scanner data fits into the average level of prices for the item. Furthermore market research agencies that compile such data include data collected by price collectors from a sample of outlets without bar-code readers.

34. Second, considerable attention has been paid to alternative formula that might be used to construct time to time indexes for various commodity groups, like cereals, canned tuna or TVs. To those familiar with how the indexes for such headings move from month to month based on existing CPI methods, it is often alarming how much different scanner results based on different index formula differ from each other and existing CPI indexes. In the case of cereals, Richardson (2000) reports the CPI as showing a rise of the cereal price index in New York City over the 28 months ending June, 2000 from 100.0 to 101.1. Using scanner data the indexes were higher, 103.6 (Geomean), 103.9 (Törnqvist) and 104.9 (Sato-Vartia). Suppose we accept for now that the scanner data was much closer to the truth because of its very large number of monthly observations (80,000) compared to the CPI (55) in New York. Should one be concerned about the differences in the three indexes that Richardson calculated? Clearly one might want to choose between them based upon say monthly variance or similar criteria. But from the standpoint of using scanner data for place to place comparisons, these differences between the methods are very small compared to other sources of error.

35. Similarly, another CPI issue in using scanner data is the length of period analyzed. Scanner data are frequently available weekly and for some types of research (for example, consumer response to price changes) this is useful. For CPI purposes, monthly data are quite adequate. Usually basic heading parities have been estimated on an annual basis, though typically the price observations have been taken in one month. So for purposes of PPP work, one could clearly make use of scanner data for as often in a year as available, but all that is really needed are one set of quantities and unit values, be they averages of monthly data, or those of a single month.

36. New products are a perennial problem in CPI work for which scanner data provides two attractive possibilities for improvement. First the introduction of the item and quantities are recorded immediately so that there need be no lag in their introduction. And second, the availability of quantity data allows for the possibility of approximation of virtual prices of the new good. (Hausman, 1997) Happily the new good problem does not arise in estimating spatial parities when a new good appears in all the markets of all countries at the same time. While this is the case with the data set with which we are working in this paper, it may not be the case in general, so in thinking about how to use scanner data, this should remain a consideration.

---

in CPI data. He was due to fly from Paris to Philadelphia with the first sets of binary results to begin writing up the results. But he made the mistake of glancing at the Italy/US spread, which was 20 times larger than anything he had encountered in temporal price indexes. So he delayed his return until he had in fact verified the initial Italy/US result was, in fact, correct.

2. Methodology

37. We consider three methods for measuring quality-adjusted price changes using scanner data: the dummy variable hedonic method, a matching technique and a SEHI approach, whose application for comparisons over time has been considered by Silver and Heravi (2000).

a) Dummy Variable Hedonic Method

38. This is akin to that given by equation (2) except that it includes a detailed set of variables on the quality characteristics, including the makes, of the items. Thus price variation, in local currencies, within and between countries is explained by variation in the quality characteristics of each item as well as the country dummy variables to pick up the parities. The variation in quality across (and within) countries is not controlled for by the selection of a sample of items, so that like is compared with like. It is controlled for in the regression equation by partialling out any quality differences. As such there is no need in principle to sample from the population of items, since matching is not used, thus making the most of the rich coverage of scanner data and its detailed information on quality characteristics. The theoretical basis for the method *within a country* has been derived by Rosen (1974) where a market in characteristic space is established (see also Triplett, 1988 and Arguea *et al*., 1994). Empirical studies and econometric issues are surveyed in Griliches (1990), Triplett (1990) and Gordon (1990); but see also Berndt *et al.* (1995), Moulton *et al.* (1998), Hoffmann (1998) and Silver (1999). The hedonic regression is given by:

$$(3)\ \ln p_{ij} = \beta_0 + \sum_{k=1}^{K_1} \beta_k z_{ikj} + \sum_{k=K_1}^{K} \beta_k z_{ikj} + \sum_{j=2}^{m} \alpha_j D_j + \varepsilon_{ij}$$

39. As in the discussion of the CPD equation, the index *j* represents the *m* countries, and the index *i*, the individual products. However, in this case *i* refers to an item that has *K* characteristics. In the CPD equation there was no intercept, and each p_i was the log of the price of the item. In (3), β_0 is the log of the price of the base item, in our application the simplest Sony 14" TV set, in the currency of the base country, the UK, sold in multiples. The remaining β_k s simply modifies that coefficient and α_j are the heading parity that would convert the price to say guilders.

40. The β_k s have been broken into two groups, those for $k = 1,2,\ldots K_1$, and those for $k = K_1 + 1,\ldots K$. The distinction being made here is between what we will term core characteristics of an item, in our case screen size, a feature that will be present for all *i*. The other group includes characteristics, like widescreen, which is a feature of less than 20% of the models. This division of the characteristics is not essential to the estimation but it is useful for thinking about the general problem of spatial parities, and we will return to this in the next section.[9]

41. The dummy variable hedonic approach as conventionally used is not without problems. First, it implicitly treats each model as being of equal importance, when some models will have quite substantial sales, while for others sales will be minimal. Scanner data includes data on sales and a weighted least squares estimator may be employed, (Ioannidis and Silver, 1999), this being considered in the empirical section. A second problem relates to the averaging procedure necessary for multilateral comparisons. It is apparent from equation (3) that comparisons are unaffected by the size or value of consumption in each country. If, for example, the regression was constrained so that all α_j, the coefficients on the *country specific* dummy variables, were constrained to be zero, the intercept β_0 would be the unweighted average of all of the country's intercepts, not a weighted international average. Kennedy (1986) explains how country weighted averages of the *z* characteristic set must be used if international averages are required as

9 The essential advantage of the SEHI formulation in (3) is that it permits use of observations on characteristics that may not match those in other countries so long as at least some other characteristics of the item do overlap with observations in other countries.

benchmarks. Scanner data with its associated unit vales and volumes allows for such estimates. Finally, a problem arises with the manner in which the direct method takes account of changing marginal values (coefficients) over time, or in this context, across countries. It is the usual practice that the coefficients are held constant and thus not allowed to reflect changes in the marginal worth of the characteristics. Dummy slope coefficients on each characteristic for each period would relax the constraint. Yet this renders the estimate of quality-adjusted price changes, the coefficient on the dummy (time) intercept, dependent on the values of the performance characteristics (Silver, 1999 and Kokoski *et al.*, 1999). In an international context the restriction amounts to assuming preferences for characteristics are constant across countries, though we will consider the implication of relaxing this in the empirical section. We will also see that these problems are dealt with in the SEHI formulation, the dummy variable hedonic method being a restricted version of the SEHI hedonic approach.

b) Matched Models Method

42. The aim is to devise a method using scanner data which allows for quality adjustment by comparing only 'like' with 'like'. This is akin to the process used by price collectors for statistical offices in the compilation of CPIs, but the matching is electronic using scanner data. Scanner data have a code to describe each model of a good. The code can be extended to include the type of outlet in which it is sold, in order that a particular model of a good in a particular type of outlet is matched against its counterpart in successive periods. Since individual retailers often have unique codes for the same model, the matching is in practice closer than by 'model and outlet type.' A problem with such matching is missing observations. For scanner data they arise when there is no transaction in that outlet (type) in a period, possibly because the item is no longer being sold, or is on display but no one has bought it. This contrasts with the practice of price collectors who may record a display price whether or not the item has been sold. More importantly in this context is the validity of the codes being used. Providers of scanner data use codes to describe the models and it can be the case that the same model sold in a different country is given a different code. While it may well be the case that the market research providers of such data can overcome this problem, initial empirical investigations were problematic. The study thus moves to a third approach

c) Superlative Exact Hedonic Indices (SEHI)

43. Konüs (1939) and Diewert (1976) define a theoretical cost-of-living index (COLI), P_c as the ratio of the minimum expenditure required to achieve a given level of utility, U, when the consumer faces period *t* prices compared with period *t*-1 price, p_t and p_{t-1}; i.e.

$$P_c(p_t, p_{t-1}, U) = E(p_t, U)/E(p_{t-1,} U) \qquad (4)$$

44. Konüs (1939) demonstrated that there is a substitution bias in the use of fixed base index number formulae, the Laspeyres index, for example, being unable to accommodate shifts in purchases away from goods with relatively high price changes. Laspeyres and Paasche were shown to provide upper and lower bounds respectively on a true COLI given by (4). Index number formulae can be shown from economic theory to correspond to, or be *exact* for, COLI for specific functional forms of the expenditure function. For example if preferences themselves are fixed, Leontief, then Laspeyres is exact. The base and current period weighted geometric means indices (equation (6a)) are exact for (correspond to) utility maximising (representative) consumers with constant elasticity of substitution. Diewert (1976 and 1978) found that symmetric averages of these bounds provide index number formulae that correspond to flexible functional forms for the expenditure function, which are much less restrictive. He defined such index number formulae as being *superlative*. Fisher's index is the geometric mean of Laspeyres and Paasche and is superlative. The Törnqvist index uses a symmetric mean of the weights of the bounds in equation (6a) and is superlative and exact for (corresponds to) a flexible translog utility function. Fisher's and Törnqvist indices are thus quite special in that they are *superlative,* though Diewert (1995 and 1999) has also shown the two formulae to be superior to many others from an axiomatic approach, with Fisher's in particular satisfying more 'reasonable' tests than its competitors.

45. The above does not recognise that changes may occur in the quality mix of the items compared. Fixler and Zieschang (1992) and Feenstra (1995) define an analogous *hedonic* COLI:

$$P_c(p_t, p_{t-1}, z_t, z_{t-1}, U) = E(p_t, z_t, U) \,/\, E(p_{t-1}, z_{t-1}, U) \qquad (5)$$

i.e. the ratio of the minimum expenditure required to maintain a given level of utility when the consumer faces p_t and p_{t-1} prices and quality characteristics z_t and z_{t-1}. Feenstra (1995) developed bounds for COLI that take account of quality variations over time and Ioannidis and Silver (1999) have provided some empirical results.

46. There is a natural extension of equation (5) to bilateral price comparisons where countries *A* and *B* replace periods *t* and *t*-1. The extension of the framework in equation (5) to multilateral international comparisons from both an axiomatic and economic theoretic approach has been undertaken by Diewert (1999), who identifies the relative merits of a range of formulae. Optimising behaviour in each country by economic agents is assumed, as is a common linearly homogeneous aggregator function – each household in each country is assumed to maximise the increasing concave and linearly homogeneous utility function subject to a budget constraint.

47. We first consider bilateral comparisons and borrow the framework for equation (5) developed by Feenstra(1995), substituting spatial for temporal comparisons. We will then move to a multilateral framework based on these bilateral building blocks. The country *A* (base) and country *B* weighted quality-adjusted bounds for a COLI, when $ln\ p_i$ is concave in z_i, i.e. for a semi-logarithmic specification of the hedonic equation, are given by (Feenstra, 1995, *proposition* 7):

$$\prod_{i=1}^{N}\left(\frac{p_{iB}}{\hat{p}_{iA}}\right)^{s_{iB}} \leq \frac{E(p_B, z_B, U)}{E(p_A, z_A, U)} \leq \prod_{i=1}^{N}\left(\frac{\hat{p}_{iB}}{p_{iA}}\right)^{s_{iA}} \qquad (6a)$$

where E(.) denote the expenditure function, in countries *A* and *B*, evaluated at a fixed level of utility and the arguments in the index are given by :

$$\hat{p}_{iA} \equiv p_{iA}\, exp[\sum \beta_{Ak}(z_{ikB} - z_{ikA})]$$

$$\hat{p}_{iB} \equiv p_{iB}\, exp[-\sum \beta_{Bk}(z_{ikB} - z_{ikA})] \qquad (6b)$$

which are prices in countries *A* and *B* respectively adjusted for the sum of the differences in each quality characteristic weighted by its respective marginal value derived from a semi-log hedonic regression, s_{iB} and s_{iA} are the shares in total value of sales of product *i* in countries *A* and *B* respectively.

48. An arithmetic aggregation for a *linear* hedonic equation is given by:

$$\left[\frac{\sum_{i=1}^{N} x_{iB}\, p_{iB}}{\sum_{i=1}^{N} x_{iB}\, \hat{p}_{iA}}\right] \leq \frac{E(p_B, z_B, U)}{E(p_A, z_A, U)} \leq \left[\frac{\sum_{i=1}^{N} x_{iA}\, \hat{p}_{iB}}{\sum_{i=1}^{N} x_{iA}\, p_A}\right] \qquad (7a)$$

where $\hat{p}_{iB} \equiv p_{iB} - \sum \beta_{kB}(z_{ikB} - z_{ikA})$

$$\hat{p}_{iA} \equiv p_{iA} + \sum \beta_{kA}(z_{ikB} - z_{ikA}) \qquad (7b)$$

where Laspeyres and Paasche are upper and lower bounds on 'true', economic theoretic COLIs: x is quantity sold, p is price, and z a vector of characteristics with associated marginal values β_{kiB} derived from linear hedonic regressions over i=1…N product varieties (models) for each characteristic k. Differences in the quality of models are picked up via differences in their characteristics ($z_{kB} - z_{kA}$) which are multiplied by estimates of their associated marginal values β_{kB}. With sales data available, the vector z can be the sales-weighted average usage or mix of each characteristic in each country. Note that $\hat{p}_{iB}$ corrects the observed prices p_{iB} for differences in the characteristics between the two countries, corresponding to the "explicit quality adjustment" described by Triplett (1990: 39).

49. The advantages of the SEHI approach are threefold. First, it utilises the coefficients on the characteristics to adjust observed prices for quality differences. Second, it incorporates a weighting system using data on the sales of each model and their characteristics, rather than treating each model as equally important. This also has important implications for the extension of the analysis to a multilateral framework where relative sales in each country forms part of the weighting system. Finally, it has a direct correspondence to a constant utility index number formulation defined from theory.

50. The interpretation of the N models or cells in these equations is essential to understanding our approach. In traditional place to place comparisons, these N models would be matched in both countries, in which case the actual prices in A (B) would enter the left (right) hand side of (6a) and (7a) and there would be no need for equations (6b) and (7b). However, implementation of the traditional matched approach requires meetings and agreement between countries A and B. And if scanner data are used, the number of matched observations is reduced significantly because models with all characteristics are not available in all countries. This is particularly problematic with makes, where models may not be sold in all countries or sold in different countries under different brand names. Furthermore the electronic matching of models is confused by model codes for similar models of a TV not always being the same for each country.

51. Note how there is no loss of information in the SEHI formulation given by equations (6a) and (7a) since each model has a corresponding model in both countries, any minor differences being corrected by equations (6b) and (7b). However, if we aggregated over all characteristics in equation (7a), with no adjustments in (7b) we would have a matching process with some models having no price data for either country B or A in any two-way comparison. These would be excluded. However, by allowing the aggregation over a limited number of characteristics, which include models available in both countries, no information is lost. The adjustment for the variables not included in this weighted aggregation takes place in (6b).[10]

52. We believe that it is especially important to use as much scanner or other price information as possible in spatial comparisons. Given the costs and difficulty of matching, and the uncertainty of whether matches really match, we have adopted an approach that uses as much information as possible. This in turn leads to our formulation based upon core and secondary characteristics of products. The N models in our formulation are a set of core characteristics, say makes, outlet types and screen sizes available in both A and B. There will be many more N actual transactions because there will be TV sets sold that have secondary characteristics that vary for the same model. What the formulation in equations (6b) and (7b) does is adjust the price for each model to the common set of characteristics in A or B. Where the market shares of models of country B are the weight, the actual price in B for a given model that embodies B's mix of secondary characteristics, is compared to the price that would occur in A for that model with B's mix of secondary characteristics. And when A's weights are used, the actual price in A is compared to an estimated price for B for that model, with A's mix of secondary characteristics. An alternative to this

---

10 If a secondary characteristic is not available in country A, then only β_B can be estimated and it would have to be used in both sides of (6b) and (7b).

formulation also discussed below would be to use the average mix of secondary characteristics in A and B to estimate a price of each model in A and B.[11]

53. Equations (6b) and (7b) are the general case where separate marginal values of the characteristics, the β_{Ak}s and β_{Bk}s are assumed, which would be generated from separate country in hedonic equations. It is also possible to estimate a pooled hedonic regression for the two countries, where the β_ks are common for both countries, or are common for some, but not all characteristics, be they core or secondary. Further, equations (6) and (7) may be implemented, using β_ks that are estimated from hedonic regressions using pooled data from more than just two countries.[12] We explore some of these issues in the empirical section.

3. The Multilateral Case

54. Having developed a methodology for estimating superlative bilateral parity of a basic heading parity, where weights are available, how do we move to the multilateral case? At least four approaches suggest themselves. The first, which would be easier on the reader, is to simply run a pooled hedonic regression across countries and use the country parities that emerge, full stop. We present this approach, but also examine others in part to gain insight into the sensitivity of the results. The second would be to use a classic Gini-EKS approach, the third would be the purposeful binary chain, a variation of Robert Hill's spanning tree approach (Hill, 1999) and, finally, the fourth is to use a method outlined in Kokoski *et al.* (1999), a 'superclass' of the EKS/CCD approach.

55. From our perspective, which tries to use as much price information as possible, we would propose two ways to obtain a common set of binary price comparisons to implement the Gini/EKS and Hill's approach. Following our distinction between core and secondary characteristics, we would generate a set of cells for core characteristics, say makes, outlet types, and screen sizes. The first way is to generate price ratios for each cell as in (6a) or (7a) above. This will be termed the country mix method.

56. The second way is to obtain a price for each cell for each country where the sales share was significant, using a mix of secondary characteristics common to all the countries. This will be termed the average mix method. The price would be estimated from a hedonic regression equation, weighted and pooled or not, that question being left to the empirical section. For countries without scanner data, but with some information on weights, prices could be directly collected.

57. This set of cells or models could then be used to develop binary comparisons for each pair of countries. As noted we could construct a set of binary comparisons based upon (6a) or (7a). The Fisher index could be used and transitivity imposed on the set of binary comparisons using Gini-EKS as in (1). This would be a way to combine the various binary comparisons

58. A third approach is to use purposeful binaries along the lines of the spanning tree approach of Robert Hill (1999). In this approach, binary comparisons would be made between pairs of countries that would be likely to have similar mixes of secondary characteristics and similar weights in the various cells. In the Hill approach this would correspond to a path that would minimize the Paasche-Laspeyre spread implicit in (6a) or (7a). It is not proposed that the Hill algorithim actually be carried out for each basic heading, but rather this consideration is used to generate a set of binaries linking all of the countries. In our three-country example, this is fairly easy to do, so we will illustrate it in the empirical section.

---

[11] If the β_{Ak}s and β_{Bk}s are identical, the estimates of the price ratio for a model entering both sides of 6a and 7a would be the same.

[12] One reason to do this might be that a given characteristic of model is found in country C and B, but not often in A. In this case the β_k estimate will be sharper if the price observations in country C are used to supplement that in A and B.

59. The multilateral SEHI approach is outlined in Kokoski *et al.* (1999) for a Tornqvist formula. The approach is based on their first noting the transitivity condition and then estimating a suitable reference charactersitic set, mix of characteristics, regression coefficients and estimated average prices that if used will meet this condition. Thus the first step is to estimate parameters to form the reference set for the comparisons and meet the transitivity conditions. The second stage is then to calculate for each country (multilateral indices) in relation to this common reference set.

4. Further Considerations

60. The criteria for selection of core and secondary characteristics require some mention. First, the core variables should have some explanatory power with regard to price variation both within and between countries and the results from the hedonic regressions in Tables 5 and 6 are helpful in this respect. Second, they should be of some substantive importance, Table 3 showing, for example, little to be gained from including satellite (tuner) as a core feature. Finally they should have counterparts in all countries. The use of ‘makes’ in this study fails in this regard in the sense that makes not available in both countries were aggregated in a category defined as ‘others’. Some makes and screen sizes also had relatively small sales. However, the computational ease of extending the matching with scanner data is so slight, that the use of a finer disaggregation of, for example, screen size or the extension of core characteristics to a richer set is a relatively trivial matter as long as such comparisons exist. The evidence from Tables 4, 5 and 6 is that the extension of core characteristics to include flat screen technology might be appropriate for a France/UK comparison.

61. Let us close this section with some more pragmatic reasons for adopting the framework in (3) and modifications of (6a) and (7a) that distinguishes between core and secondary characteristics of items. First, the framework is not special to consumer durables and can be extended to other products. For example, consider a non-durable basic heading like non-tropical fruits. If scanner data were available then it would be possible to implement (6a) where the interpretation of N might be varieties of fruit, like apples and pears from different outlet types, convenience outlets and supermarkets. The secondary characteristics might Fuji, golden delicious, and Granny Smith apples, and Bartlett and Bosc pears. Similarly, consider an item like rice, which is also a basic heading, where size of package is crucial to the price per kg. Then the interpretation of N in (6a) will be core characteristics, perhaps long grain and short grain, outlet type, and size of purchase as a continuous variable or as dummies, where other characteristics might include whether purchased in bulk, bag or box, brands and the like.

62. Another practical advantage of formulating the relationship in terms of core and secondary characteristics is that it is a convenient framework for combining price and sales information from countries that have scanner data (or weights below the basic heading level) with price data from countries that do not have such rich data. The core characteristics can become the set of cells across which comparisons are made. Countries with scanner data can both provide estimates for the cells as well as weights. Countries without weights or scanner data can provide prices from their most typical outlets, including whatever information on secondary characteristics is available. The latter can be used with the $\bullet_k$s to generate more accurate price comparisons than would be possible, unless direct matches are available.

63. Before reporting the empirical results, two additional comments about this formulation. First, if the cost of harmonizing scanner data across countries is prohibitive, it is still possible to estimate an hedonic regression for each country separately and to utilize this as described above and carried out in the empirical section. Second, it is very time consuming but very important to match specifications across countries in traditional spatial comparisons. In the approach suggested here, more resources would be invested in the data processing, and less in matching, which has frequently involved expensive meetings.[13]

[13] One problem with the matching process over the years has been the fact that country involvement is too small after the initial item prices have been provided to a central organization for a regional or country group. This point has been make in the Rhyten Report (1998). Use of estimating equations might offer some respite here too.

Part C. AN APPLICATION OF SCANNER DATA TO TELEVISION SETS

1. Variables and data

64. The empirical work utilises scanner, bar-code data for television sets for June and July 1998 in three countries: the U.K., France and the Netherlands. Scanner data provides, for each item, information on price (unit value), sales (through aggregation of transactions), characteristics of the product (linked to the item or model number), retailer and time of purchase. Since data are available on the characteristics of each item the potential exists to allow for quality-adjustment in the measurement of prices between countries. The scanner data provide transaction prices for all transactions as opposed to a price collector collecting display prices for a sample of branded varieties on what are intended to be representative items taken from a sample of outlets for a single day in each month. As will be seen, the coverage of the scanner data for an item extends to millions of transactions each year. Supplementary data were also collected from outlets without bar-code readers. The relative merits of the practical use of such data for CPI compilation have been discussed in Silver (1995), Fenwick *et al*. (2000) and Richardson (2000).

65. The observations are for a model of the product for which there was a transaction in the country in a particular outlet type. For example, an observation in the data set for June and July 1998 includes the unit value (£284.52), volume (3,686 transactions) and quality characteristics (including possession of Nicam stereo and fastext text retrieval facilities) of the a (Panasonic TX21MD3 21 inch screen) television set sold in electrical multiples only in the U.K. For June and July 1998 there were 4,827 observations: 1,186 for the Netherlands, 2,146 for France and 1,495 for the U.K. representing over a million transactions over June and July 1998: about 0.2, 0.6 and 0.3 million transactions in each of the Netherlands, France and the U.K. respectively. While observations may be for this same model in different outlet types, models were, on average sold in 1.37 different outlet types in the Netherlands, but more so for France at 1.97 and the U.K. at 2.2. There were 2,425 different models and 4,827 when also differentiated by the outlet type in which they were sold: models being sold in 1.84 outlet types on average. Their model numbers of course distinguishes models here and these may change over time with fairly small differences in styling and coexist in the market with their previous incarnation. However, as Table 1 demonstrates, the variation in prices between models/outlet types within countries is not trivial and requires explanation in terms of their quality characteristics.

For one thing, estimates from an equation like (3) do not have any of the confidentiality considerations surrounding national collection of prices so countries could be provided a grid of estimated cell prices, and a list of other characteristics from which they could then provide information.

Table 1: Summary statistics of price in local currencies

	N	Mean	Standard deviation	Coefficient of variation	0.05 percentile	0.95 percentile
The Netherlands	1,186	1,309.33	1008.23	0.77	353.60	3,397.60
France	2,146	4,176.79	3,328.87	0.80	1,182.10	10,788.2 5
U.K.	1,495	466.58	432.24	0.93	118.68	1,310.31

66. The variable set is quite extensive and includes:

a. Price is the unit value of a model in a month/outlet type across all transactions, i.e. sales value/sales volume (see Balk, 1996 for the statistical properties of unit values).

b. Volume is the sum of the transactions during the period. As will be shown below many of the models sold in any month have relatively low sales.

c. c.Vintage is the year in which the first transaction of the model took place. New models can coexist with old models, both as a result of an inability to dump the old model before the launch of the new one and as an appreciation that different sub-markets exist for models of different vintages.

d. The characteristics set includes: (i) manufacturer (make) – dummy variables for about 36 makes; (ii) size of screen; dummy variables for possession of: (iii) Nicam stereo; (iv) tube type - flat screen tube/Trinitron; (v) tuner - Pal, Pal/Secam, Pal/Secam/NTSC, Pal plus varieties; (vi) satellite; (vii) text retrieval system – fastext/TOP, teletext; (viii) dolby system; (ix) wide screen; (x) s-vhs socket; (xi) digital.

e. .Outlet type[14]

67. Table 2 shows the volume of transactions for models/outlet types to be highly negatively skewed with a relatively large number of models having with low sales volumes. Models of TVs in an outlet type with sales of 30 or less in the two months were ignored, we would be left with only 45, 60 and 58 percent of observations for the Netherlands, France and U. K. respectively. Yet these remaining observations would account for 97, 99 and 98 % of transactions respective countries. The top 71 or 5 percent of makes in an outlet type in the U.K by sales can be seen from Table 2 to account for 126, 229 or 41 percent of transactions.

---

[14] The country & outlet classification are:
NL Departmental and Catalogue 11
NL Electrical multiples 12

a) ***NL Photographers 13***

Fr Departmental
Fr Electrical multiples 22
Fr Hypermarkets 23
Fr Specialist (independents) 24
Fr Catalogue 25
UK Mass merchandisers (departmental) 31
UK Electrical multiples 32
UK Renters and others n.e.c 33
UK Independents 34
UK Catalogue 35

These were combined in four groups for each country: multiples, mass merchandisers, catalogue and independents with NL=11 being allocated to mass merchandisers.

Table 2

Monthly volume of sales	Country					
	The Netherlands		France		U.K.	
	Number	*Sales Volume*	*Number*	*Sales Volume*	*Number*	*Sales Volume*
1-10	426	1,665	553	2,543	384	1,427
11-20	137	2,109	222	3,536	145	2,239
21-30	87	2,166	111	2,844	106	2,676
31-100	242	14,143	446	26,844	329	19,857
101-500	185	43,644	461	110,293	364	86,445
501-1000	63	43,064	190	132,146	96	69,674
1,001-5,000	44	79,549	156	311,822	71	126,229
5,001-10,000	2	10,857	7	45,192	-	-
Total	1,186	197,197	2,146	635,220	1,495	308,547
Active sample†	540	191,377	1,281	626,927	873	302,595
as % of population	45	97	60	99	58	98

† sales of 30 or more per make in an outlet type (limited sample used for hedonic regression only).

68. Table 3 provides information on the sample according to the characteristics of the TV's sold**.** The relative importance of different features of the sets is provided for each country both by number of sets and by percent of sales. The information on the various makes is provided for only the top dozen brands, but more than thirty brands are retained in the hedonic regressions. The top three makes can be seen to dominate the market accounting for between a third and a half of the volume of sales. Some makes achieve relatively high sales with relatively few models, for example Philips and Sony in the U.K. sold a similar number of varieties, though Sony achieved very much higher sales. Though, as shown by Table 2, this is more likely to be a reflection of their ability to clear redundant models than any actual strategy on market offerings. Table 3 also shows how the desirability of different features of TVs varies between countries. For example, digital TV at this time was lees important in France, while flat screen/Trinitron technology was taken up to a similar degree between countries. Purchases of wide screen TVs were more prevalent in Netherlands and larger screen sizes in France.

Table 3: Characteristics of television sets

	Country					
	The Netherlands		*France*		*U.K.*	
	Percent of models	*Percent of sales volume*	*Percent of models*	*Percent of sales volume*	*Percent of models*	*Percent of sales volume*
Characteristic						
Digital*	24.3	21.8	11.1	4.0	40.2	28.1
Dolby	7.4	4.6	8.4	3.0	24.3	14.8
Flat screen Trinitron	73.2	65.5	80.0	66.9	70.0	66.1
Nicam	22.6	19.8	59.6	47.5	60.0	49.2
Satellite	0.4	0.0	2.6	1.2	1.0	0.1
SVHS	47.1	41.3	58.2	47.0	50.8	55.6
Teletext	28.5	32.6	43.9	40.4	11.2	9.0
Fastext/TOP	59.0	52.0	19.9	14.0	71.9	69.6
Widescreen	17.6	17.8	15.8	7.8	13.3	8.6
Reception (Tuner)						
PAL	47.2	58.7	0.2	1.2	51.8	56.8
PAL/SECAM	36.9	31.5	79.4	86.3	18.9	11.4
SECAM	-	-	0.2	0.2	0.1	0.1
PAL/SECAM/NTSC	10.5	5.1	20.2	12.2	28.2	31.5
PAL PLUS**	5.5	4.8	-	-	0.4	0.0
Screen size						
Under 14”	1.6	0.3	2.0	2.1	0.9	1.6
14”	18.8	26.8	14.8	25.8	21.1	30.2
15” to 19”	1.6	1.8	2.3	0.9	1.4	1.4
20”	7.6	10.0	5.2	9.0	5.6	5.5
21”	16.5	18.3	20.4	21.1	20.7	24.5
24”	2.8	2.9	1.3	0.3	1.3	0.6
25”	15.3	13.7	11.2	5.6	17.4	15.0
26”	0.1	0.0	0.1	0.0	0.0	0.0
28”	20.8	18.5	24.8	24.6	17.1	11.8
29”	8.9	5.1	8.7	6.8	6.5	6.0
32” and more	6.3	2.4	9.1	3.8	8.3	3.5
Makes (with more than 1.5% on average of total international market share)						
Sony	13.5	15.2	12.5	10.7	10.8	18.1
Panasonic	4.6	4.4	4.0	2.0	10.0	17.9
Grundig	7.3	6.5	7.3	5.2	1.9	1.7
JVC	3.7	3.6	0.9	0.6	6.2	4.3
Sharp	4.0	5.1	2.3	2.5	3.4	2.4
Toshiba	0.7	0.0	3.3	3.2	7.5	10.2
Phillips	15.0	29.0	12.2	15.5	11.0	6.8
Daewoo	4.0	2.9	2.7	3.3	2.8	4.1
Thornson	2.6	0.4	11.5	10.4	2.1	2.0
Bush	-	-	-	-	2.4	5.4
Mitsubishi	1.6	0.3	1.4	1.5	6.0	4.3
Akai	3.8	2.3	1.6	2.2	0.9	0.2

* Digital includes any such facility: Digital memory, PIP (Picture in Picture), mult-PIP and Video
** Includes PAL+; PAL+/SECAM;PAL + SNJSC.

69. Table 4 provides a breakdown of outlet type with sales represented in numbers sold and value of sales. The classifications vary across countries and comparisons are less reliable. However, multiples appear to be the dominant outlet type for TV sales in the Netherlands while the U.K. market makes more use of catalogue sales.

Table 4: Sales patterns by outlet types

	Country					
	The Netherlands		*France*		*U.K.*	
	Number	Sales Volume	Number	Sales Volume	Number	Sales Volume
Mass merchandisers }	14.0	5.6	7.7	0.7	23.0	19.0
and catalogue			7.5	3.8	14.3	12.0
Electrical multiples	70.3	92.1	31.4	35.4	25.3	32.2
Independents	-	-	25.9	26.8	33.8	35.9
Hypermarkets	-	-	27.5	33.2	-	-
Renters and other niec	-	-	-	-	3.6	0.8
Photographers	15.7	0.7	-	-	-	-

Part C – RESULTS

1. The Dummy Variable Hedonic Method

70. The OLS regression results are given in Model I of Table 5. The sample of 2,633 excludes models with country sales of less than 30 in any outlet type. Such models with limited sales may have unusual pricing policies. The $\overline{R}^2$ can be seen to be relatively high at 0.965 and the magnitude and signs are of the appropriate order and nature. For example, a windscreen TV has a price margin of (exp (0.300) –1)*100 = 35%, or more properly, with the adjustment of half the standard error (Goldberger, 1968), i.e. (exp (0.300 + 0.00769) – 1) * 100 = 36%. All quality features have a positive sign, with the exception of the provision of a satellite tuner, probably as a result of some multi-collinearity. The dummy variables for makes were benchmarked on a Sony, which in Europe carries an up-market premium, most other makes having negative coefficients. Exceptions include Bang & Olufsen who serve a niche, relatively high-priced market. The outlet types were benchmarked on multiples with, for example, purchases in independent outlets having an (exp (0.100073 + 0.5 (0.012407)) –1) * 100 = 11% price premium. The nature of the purchase experience is regarded as a quality characteristic of the transaction and thus expected service affects price. The country dummy variables are benchmarked on the UK with an estimate of parities at this basic heading for the Netherlands against the UK of 2.83 fl/£ or 33% above the exchange rate, and France against the UK of 8.75 fr/£ or 45% above the exchange rate. The models suffer from some heteroskedasticity through heteroskedastic consistent standard errors were used for the tests following a procedure by White (1980).

Table 5
Regression results

	Model I - OLS, n>30, country & outlet		**Model I - OLS, n>30, country * outlet**		**Model I - WLS, n>30, country & outlet**	
Variable	**Estimated coefficient**	**Standard Error**	**Estimated coefficient**	**Standard error**	**Estimated coefficient**	**Standard error**
Characteristics						
constant	6.532***	0.529	6.268***	0.512	5.968***	1.040
Flatscreen	0.143***	0.014	0.152***	0.014	0.196***	0.026
Teletext	0.119***	0.014	0.116***	0.013	0.124***	0.019
Fastext	0.177***	0.016	0.173***	0.015	0.208***	0.021

Digital	0.089***	0.008	0.086***	0.008	0.084***	0.013
Satellite	-0.176**	0.057	-0.186***	0.053	-0.242**	0.075
S-VHS	0.033**	0.012	0.033***	0.012	0.011	0.020
Widescreen	0.300***	0.015	0.300***	0.015	0.336***	0.023
Dolby	0.179***	0.017	0.187***	0.017	0.170***	0.023
Nicam stereo	0.111***	0.014	0.114***	0.014	0.116***	0.018
screen size	-0.125***	0.025	-0.122***	0.024	-0.999	0.051
$(\text{screen size})^2$	0.0060***	0.0011	0.0058***	0.0011	0.0046	0.0023
$(\text{screen size})^3$	-.000083***	0.000016	-.000051***	0.000015	-.000031	0.000033
Vintage	-0.009	0.005	-0.007	0.005	-0.005	0.009
Tuner (PAL1 omitted)						
PAL/SECAM	0.059***	0.015	0.060***	0.014	0.076**	0.026
PAL/SECAM/NTSC	0.072***	0.016	0.068***	0.016	0.059*	0.025
Makes (Sony omitted)						
Akai	-0.086*	0.030	-0.099***	0.029	-0.120*	0.051
Alba	-0.319	0.172	-0.346***	0.207	-0.174*	0.072
Amstrad	-0.148	0.082	-0.203	0.117	0.006	0.065
Beko	-0.317***	0.038	-0.345	0.030	-0.185**	0.063
Binari	-0.264***	0.017	-0.276***	0.017	-0.211***	0.028
Bang & Olufsen	1.346***	0.083	1.334***	0.086	1.303***	0.110
Blaupunkt	0.109***	0.019	0.090***	0.017	0.180***	0.038
Bush	-0.074	0.040	-0.123***	0.037	0.015	0.048
Crown	-0.279***	0.019	-0.228**	0.038	-0.212***	0.028
Daewoo	-0.122***	0.023	-0.127***	0.023	-0.903*	0.038
Decca	0.329*	0.160	0.337	0.219	0.288	0.178
Dual	-0.134	0.093	-0.096	0.093	-0.047	0.083
Ferguson	-0.202***	0.043	-0.189***	0.043	-0.222***	0.054
Goodman	-0.159**	0.053	-0.180***	0.077	-0.097	0.053
Grundig	0.096***	0.023	0.091***	0.023	0.091**	0.033
Hitachi	-0.782***	0.020	-0.077***	0.018	-0.019	0.028
JVC	-0.046	0.023	-0.051*	0.022	-0.016	0.030
LG	-0.141***	0.028	-0.151	0.029	-0.770	0.048
Loewe	0.512***	0.034	0.505***	0.035	0.534***	0.048
Mitsubishi	-0.148***	0.033	-0.131***	0.033	-0.063*	0.031
NEI	-0.306***	0.046	-0.286***	0.049	-0.240***	0.036
Nokia	0.012	0.043	-0.003	0.044	0.011	0.063
Nordmen	0.226*	0.110	0.220	0.120	0.276*	0.131
Orion	-0.293*	0.117	-0.257*	0.126	-0.182*	0.073
Panasonic	0.082***	0.018	0.073***	0.017	0.077**	0.026
Phillips	0.129***	0.016068	0.122***	0.015	0.175***	0.025
Pye	-0.061	0.040	-0.100***	0.036	0.031	0.039
Saba	-0.082*	0.032	-0.079***	0.032	-0.011	0.049
Samsung	-0.128***	0.039	-0.142***	0.035	-0.040	0.065
Sanyo	-0.098***	0.028	-0.117***	0.026	-0.041	0.035
Schneider	0.015	0.037	0.012	0.036	0.050	0.050
Sharp	-0.081***	0.021	-0.076***	0.021	-0.012	0.028
Tatung	0.002	0.032	-0.005	0.033	0.093*	0.046
Teletun	0.045	0.072	0.033	0.071	0.107	0.058
Thomson	0.053*	0.022	0.048***	0.021	0.147***	0.031
Toshiba	0.009	0.020	-0.001	0.020	0.049	0.027
Others	-0.259***	0.025	-0.229	0.024	-0.283***	0.041
Country and outlet (UK multiples omitted)						
NL Department	-	-	1.135***	0.026	-	-
NL Multiple	-	-	1.140***	0.020	-	-
NL Photographer	-	-	1.084***	0.036	-	-

Fr Department	-	-	2.341***	0.032	-	-
Fr Multiple	-	-	2.292***	0.023	-	-
Fr Hypermarket	-	-	2.147***	0.024	-	-
Fr Specialist	-	-	2.387***	0.023	-	-
Fr Catalogue	-	-	2.257***	0.027	-	-
UK Merchandiser	-	-	0.129***	0.019	-	-
UK Renter	-	-	0.129***	0.032	-	-
UK Independent	-	-	0.182***	0.018	-	-
UK Catalogue	-	-	0.286***	0.020	-	-
Outlet type (multiples omitted)						
Mass merchandisers	-0.041***	0.012			-0.064**	0.023
Independents	0.100***	0.012			0.130***	0.018
Catalogues	0.112***	0.016			0.108***	0.023
Country (UK omitted)						
The Netherlands	1.033***	0.016			1.054***	0.023
France	2.161***	0.017			2.132***	0.027
Adjusted R^2	0.965		0.968		0.969	
Sum of squared residuals	125.545		116.637		99.658	
n	2633		2633		2633	
F statistic, zero coefs.	1288		1196		927	

Standard errors are adjusted for heteroskedasticity
***, **, * denotes statistically significant at a 0.1, 1 and 5% level respectively for a two-tailed test.

71. The relatively crude classification of outlet types into multiples, independents, mass merchandisers and catalogue is an attempt to provide a consistent categorisation of the richer classification available as outlined in the previous section. Estimates in Model II of Table 5 use this finer classification benchmarked on UK multiples. The results for individual outlet types in individual countries are estimates of how price comparisons differ across purchases in different outlet types within and between countries. For example, in France the hypermarkets are 16% cheaper than the mulitples, while there is no difference in the U.K. Further, the independents in the U.K. are only 5% cheaper than the multiples while the differential is 10% in France. In the context of estimating heading parities, knowledge of outlets and their likely influence on price is clearly important for making appropriate price comparisons. here is of note.

72. Examination of Model III in Table 5 indicates very little difference between the results from OLS and WLS estimators. The estimates of parities for TVs in particular are very close at 1.054 and 2.132 for NL and France using WLS and 1.033 and 2.161 using OLS. Both of these models restricted the sample to the volumes of transactions greater than 30 for an individual observation. However, when the whole sample was used for WLS, the coefficients were still very similar Finally, the model was re-estimated having converted prices to their dollar equivalent using the average exchange rates for June and July (IMF, *International Financial Statistics*, December 1998, Washington D.C.). Again the model performed well in terms of $\overline{R}^2$ = 0.90, the coefficients being of the expected nature (signs) and

73. The above framework is quite restrictive since it assumes the coefficients attached to each quality variable are constant across countries. By including interaction effects for each of France and the Netherlands benchmarked on the UK, an unconstrained model in which coefficients can vary across countries can be estimated. An F-test for the constrained versus unconstrained model test the null hypothesis that:

$$\beta_k = \beta_{jk} \text{ for all } k = 1,...K \text{ in (3)}$$

i.e. that the coefficients are the same across countries. Not all makes, though all characteristics were available in all countries, the unconstrained model having 128 variables, compared with 57 in the constrained model. The F-test statistic was 0.624, with a critical $F_{71,2505,0.05}$ = 1.297, thus rejecting the null hypothesis at a 5% level. Having rejected the null, it is of interest to identify which characteristics are subject to inter-country variability in their estimated preferences. These are given in Table 6. The omitted, benchmarks for the *t*-tests for these interaction terms are based on the UK, for example, the test in Table 6 for the Blaupunkt make being that $(\beta_{FR,Blaupunt} - \beta_{UK,Blaupunt}) = 0$, and similarly for other characteristics. The Phillips brand, for example, is estimated to have a marginal value of about (exp (0.15 + 0.5 (0.04) – 1) * 100 = 18.5% above that of the UK for both the Netherlands and France and the estimated marginal values of digital facilities exceeds the UK in both countries. It should be noted that the 32 differences listed in Table 6 are only those coefficients where the differences are statistically significant, there being a further 96 bilateral comparisons where the differences were not statistically significant.

Table 6, Constancy of coefficient across countries

	$(\beta_{France} - \beta_{UK})$	Standard error	$(\beta_{NL} - \beta_{UK})$	Standard error
Makes (Sony omitted)				
Blaupkunt	-0.099*	0.043	-	-
Daewoo	0.222***	0.051	0.254***	0.048
Grundig	0.202**	0.074	0.207*	0.080
Mitsubishi	-0.183*	0.086	-	-
Nokia	-0.413***	0.111	-	-
Panasonic	0.106*	0.041	-	-
Phillips	0.156***	0.039	0.154***	0.041
Telelun	-0.235**	0.078	-	-
Thomson	0.190*	0.092	-	-
Amstrad	-	-	0.450***	0.067
Hitachi	-	-	0.259***	0.059
JVC	-	-	0.135*	0.056
Tatung	-	-	0.130*	0.054
Characteristics				
Flat screen technology	0.097**	0.034	-	-
Digital	0.075**	0.023	0.055*	0.022
Satellite	-0.249**	0.080	-	-
S-VHS	0.059*	0.027	0.117***	0.038
Fastext	-	-	0.115**	0.038
PAL/SECAM	-	-	0.068*	0.033
PAL/SECAM/NTSC	-	-	0.132**	0.047
Vintage	-	-	-0.028*	0.012
Outlet (multiples omitted)				
Catalogue	-0.313***	0.029	-	-
Independents	-0.074**	0.024	-0.189***	0.037
Mass merchandisers	-0.245***	0.025	-0.125***	0.030

Standard errors are heteroskedastic consistent ***,**, * denote statistically significant at 0.1, 1 and 5% level respectively for 2 tailed tests

Comparisons are only given when the differences are statistically significant at a 5% level or less.

2. Superlative Exact Hedonic Indices: Results

74. In presenting our SEHI results we have used an unconstrained equation that employs 10 different screen size dummy variables instead of treating size continuously as in the regression equations in Table 5. Table 7 shows how we build up the SEHI results. While this is clearly stated in the equations, Table 7 is also useful because it provides a more general framework for using hedonic equations to generate detailed heading parities. We believe the framework of Table 7 is general enough to include basic headings for services and non-durables and flexible enough to accommodate price information from individual country

regressions or prices obtained otherwise with information on core and secondary characteristics. Table 7 also provides the basis for the presentation in Tables 8 and 9.

Table 7. Illustration of SEHI Calculation for 12 Screen Sizes

Size of Screen Inches	*FR Mean Price*	*FR-NL Adjust Price*	*FR-UK Adjust Price*	*FR Share Sales*	*NL Mean Price*	*NL-FR Adjust Price*	*NL Share Sales*	*UK Mean Price*	*UK-FR Adjust Price*	*UK Share Sales*
1.14:>	1616	1780	2193	0111	533	498	0018	210	239	0111
2.14	1293	1071	1010	1097	395	401	094	152	156	1241
3.15-19	2440	1961	1743	0081	706	632	0115	270	244	0099
4. 20	1388	1230	1378	0413	534	505	0473	219	219	0328
5. 21	2299	1972	2256	1594	806	743	1311	296	260	1956
6. 24	5797	5074	6012	0050	1822	1631	0471	554	467	0085
7. 25	4229	3204	3736	0776	1312	1315	1596	447	409	1819
8. 26	3965	4478	4498	3208	2076	1502	3426	631	507	2016
9. 29	5988	4996	5522	1345	1744	1641	0794	643	562	1051
10. >29	10667	11913	11075	1325	3979	3125	0855	1398	1202	1319

75. Table 7 gives the simplest breakdown of core characteristics, namely 10 screen sizes. For each screen size the mean price in each country is provided in national currencies, without holding constant the other quality characteristics of the set. The average price of a 14 inch set is 1283 fr, 395 fl and 152 £, corresponding to ratios to the £ of 8.29 fr and 2.60 fl. The adjusted prices take account of the secondary characteristics that differ between the countries for the typical TV set in the cell. The FR-NL adjusted price of 1010 fr is obtained by applying the French coefficients to the average quality of 14 inch sets in the Netherlands. Similarly the FR-UK, NL-FR, and UK-FR adjusted prices apply the coefficients of the lst country to the characteristics of the 2nd, the NL-UK, and UK-NL being implicit in the others. The following illustrates the process for 14-inch sets:

	French/UK(fr/£)		NL/UK (fl/£)		France/NL(fr/fl)	
Mean Prices 14”	8.51	(1293/152)	2.60	(395/152)	3.27	(1293/395)
Fr Adjusted	6.64	(1010/152)	n.a.		2.71	(1071/395)
UK Adjusted	8.29	(1293/156)	3.06	(395/129)	n.a.	
NL Adjusted	n.a.		2.06	(313/152)	3.22	(1293/401)

76. The two ratios for each pair of countries are derived from equations 6-8. In Table 7, we have country weights, sales shares, for each of the screen size groups, so we can aggregate these price ratios in the usual ways. To illustrate, in Table 8A, the weighted average of the FR/NL ratios (2.45) is given as the first entry using NL weights, and in the second entry (3.06) using French weights. The column labeled Törnqvist (2.74) is in the geometric formulation also the GM of the NL and FR weighted averages.

77. For many basic headings one will not have the detail available to us for Table 7. However, we would argue that to the extent one can replicate a framework that involves core characteristics of items in a heading, with any adjustment factors to control for the quality of secondary characteristics, one would have a consistent basis for comparison. We would also argue that the type of framework in Table 7, which would differ by basic headings, might well take the place of written specifications. All of this said, it is also clearly desirable to have available weights or country indicators of the representativeness of cells.

78. The aggregation in the first core component of equations 7(a) and 8(a) was according to ten screen sizes. The correction for quality differences within each of these categories across countries as given by 7(b) and 8(b) utilized makes, outlet types and all other characteristics. The core characteristic(s) should be selected on the grounds that the categories exist in both countries so there will be no missing cells, and that this characterization is salient in exploring price variation. The former was chosen for this very reason and the latter is evidenced from hedonic studies including Lowe (1998), Moulton *et al.* (1998) and Ioannidis and Silver (1999).

79. An initial analysis using makes found a loss of 20% by sales volume of comparisons with different makes not being available in both countries, possibly appearing under different brand names. An analysis over time found the same makes to be available in the subsequent time period, the overall index when chained, involving little loss of data due to missing comparisons (Silver and Heravi, 2000). For spatial comparisons we can extend the categories to include possession of other quality features, such as Nicam stereo and widescreen. The weighted aggregation would then be for particular screen sizes, with or without Nicam stereo and widescreens, substantially extending the aggregation in equations 7(a) and 8(a) with little loss of data. This extension is considered later. We start with aggregation by screen size using weighted geometric means for each country defined by equations 7(a) and 7(b), a Törnqvist index given by:- $\exp\left[0.5\left(s_{KA}+s_{KB}\right)\ln\left(\bar{\hat{p}}_{KA}/\bar{p}_{KB}\right)\right]$ for country A compared with country B: weighted arithmetic means given by Laspeyres and Paasche in equations 8(a) and 8(b) and Fisher's index given by the geometric mean of Laspeyres and Paasche.

80. Table 8 has 12 panels; the first 6 use unconstrained semi log formulations and the last 6 use arithmetic formulations for the hedonic equations. Each set of panels begins with simple breakdowns of core characteristics, starting with the 10 screen sizes, adding to that, whether or not flat screened or not, and finally in 8F and 8L, involving 20 brands, 10 screen sizes and 3 types of outlets, a total of 600 cells. The left side of each panel of Table 8 has been discussed above. The right hand side contains two sets of information. The columns labeled *Obs del/Used* and *Volume del/used* illustrate what happens as we increase the number of core cells. In panel 8A all the observations are used, representing the volume of transactions indicated. There are potentially 3,332 price ratios between France and the Netherlands representing 832,417 transactions. In the lst row of Table 8A it is indicated that it was possible to use all potential price ratios. When we move to 8B we begin to lose observations because there are empty cells. The cost of having a large number of cells is that we lose more and more observations, so by 8F more than 20% of the observations cannot be used. If this technique were extended to a larger number of countries, then it is likely that one would want to adopt a set of core characteristics that combined to produce a number of combinations closer to 100 than 500 or more.

Table 8A: Bilateral SEHI Results, by sizes(10)

	Geometric							
	Base country weights		**Current country weights**					
Bilateral Comparison	**(base GM)**		**(current GM)**		**Törnqvist**	**Obs del/used**	**Volume del/used**	**EKS/ AEKS**
France/NL	2.4494	(NL)	3.0627	(France)	2.7389	0/3331	0/832401	2.9251/ 2.7978
NL/UK	2.9162	(UK)	3.3803	(NL)	3.1397	0/2680	0/505728	2.7855/ 2.9590
UK/France	0.1157	(France)	0.1309	(UK)	0.1230	0/3641	0/943767	0.1227/ 0.1248

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8B: Bilateral SEHI Results, by sizes&flat screen(10*2)

	Geometric					Geometric		
	Base country weights		Current country weights					
BilateralComparison	(base GM)		(current GM)		Törnqvist	Obs del/used	Volume del/used	EKS
FRANCE/NL	2.4459	(NL)	3.0608	(France)	2.7361	4/3327	172/832229	2.6558/ 2.6616
NL/UK	2.8676	(UK)	3.3907	(NL)	3.1182	51/2629	6360/499368	2.5828/ 2.8318
UK/France	0.1160	(France)	0.1406	(UK)	0.1277	54/3587	7722/936045	0.1312/ 0.1297

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8C: Bilateral SEHI Results, by sizes&flat screen & nicam

	Geometric					Geometric		
	Base country weights		Current country weights					
Bilateral Comparison	(base GM)		(current GM)		Torqvist	Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.4205	(NL)	3.0547	(France)	2.7192	39/825162	6824/825162	2.6934/ 2.8076
NL/UK	2.7236	(NL)	3.0160	(UK)	2.8661	201/2464	25344/479969	2.8778/ 2.8882
UK/France	0.1181	(France)	0.1469	(UK)	0.1317	117/3524	15633/928134	0.1174/ 0.1264

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8D: Bilateral SEHI Results, by makes (20)

	Geometric							
	Base country weights		Current country weights					
Bilateral Comparison	(base GM)		(current GM)		Törnqvist	Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.5241	(NL)	3.0337	(France)	2.7672	0/3332	0/832417	2.6773/ 2.5898
NL/UK	2.9150	(UK)	3.3779	(NL)	3.1379	0/2681	0/505744	2.6870/ 2.9234
UK/France	0.1080	(France)	0.1335	(UK)	0.1201	0/3641	0/943767	0.1390/ 0.1383

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8E: Bilateral SEHI Results, by makes&sizes(20*10)

	Geometric							
	Base country weights		Current country weights					
Bilateral Comparison	(base GM)		(current GM)		Törnqvist	Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.4707	(NL)	2.9678	(France)	2.7079	161/3170	34527/797874	2.8913/ 2.8036
NL/UK	2.7145	(UK)	3.4230	(NL)	3.0482	203/2477	30307/475421	2.8495/ 3.1955
UK/France	0.1192	(France)	0.1310	(UK)	0.1250	185/3456	41343/902424	0.1224/ 0.1204

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8F: Bilateral SEHI Results, by makes&sizes&outlets(20*10*3)

	Geometric							
	Base country weights		Current country weights					
Bilateral Comparison	(base GM)		(current GM)		Törnqvist	Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.3285	(NL)	2.6019	(France)	2.4614	610/2721	120047/712354	3.0812/ 2.8746
NL/UK	2.1358	(UK)	3.2184	(NL)	2.6218	594/2086	90704/415024	2.8635/ 3.0539
UK/France	0.1399	(France)	0.1469	(UK)	0.1434	459/3182	85598/858169	0.1195/ 0.1215

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8G: Bilateral SEHI Results, by sizes(10)

	Arithmetic							
	Base country weights		Current country weights					
Bilateral Comparison	(Laspeyres)		(Paasche)		Fisher	Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.9628	(NL)	3.0128	(France)	2.9877	0/3331	0/832401	2.9251/ 2.9764
NL/UK	3.1410	(UK)	3.1587	(NL)	3.1499	0/2680	0/505728	2.7855/ 2.9833
UK/France	0.1176	(France)	0.1136	(UK)	0.1156	0/3641	0/943767	0.1227/ 0.1206

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8H: Bilateral SEHI Results, by sizes&flat screen(10*2)

Bilateral Comparison	Arithmetic: Base country weights (Laspeyres)		Current country weights (Paasche)		Fisher	Geometric: Obs del/used	Volume del/used	EKS
FRANCE/NL	2.9589	(NL)	3.0159	(France)	2.9873	4/3327	172/832229	2.6558/ 2.8430
NL/UK	3.1740	(UK)	3.1620	(NL)	3.1680	51/2629	6360/499368	2.5828/ 2.9897
UK/France	0.1177	(France)	0.1136	(UK)	0.1156	54/3587	7722/936045	0.1312/ 0.1245

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8I: Bilateral SEHI Results, by sizes&flat screen & nicam

Bilateral Comparison	Arithmetic: Base country weights (Laspeyres)		Current country weights (Paasche)		Fisher	Geometric: Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.9089	(NL)	2.9635	(France)	2.9360	39/825162	6824/825162	2.6934/ 3.0584
NL/UK	3.1649	(NL)	3.1321	(UK)	3.1485	201/2464	25344/479969	2.8778/ 2.9157
UK/France	0.1175	(France)	0.1133	(UK)	0.1154	117/3524	15633/928134	0.1174/ 0.1224

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8J: Bilateral SEHI Results, by makes (20)

Bilateral comparison	Arithmetic: Base country weights (Laspeyres)		Current country weights (Paasche)		Fisher	Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.9222	(NL)	2.9600	(France)	2.9410	0/3332	0/832417	2.6773/ 2.8686
NL/UK	3.2097	(UK)	3.4380	(NL)	3.3219	0/2681	0/505744	2.6870/ 3.1442
UK/France	0.1128	(France)	0.11237	(UK)	0.1126	0/3641	0/943767	0.1390/ 0.1285

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8K: Bilateral SEHI Results, by makes&sizes(20*10)

Bilateral Comparison	Arithmetic: Base country weights (Laspeyres)		Current country weights (Paasche)		Fisher	Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.9790	(NL)	2.97967	(France)	2.9793	161/3170	34527/797874	2.8913/ 2.8425
NL/UK	3.1895	(UK)	3.4678	(NL)	3.3258	203/2477	30307/475421	2.8495/ 3.3379
UK/France	0.1092	(France)	0.1127	(UK)	0.1109	185/3456	41343/902424	0.1224/ 0.1144

Countries in parentheses are used for the weights and the hedonic coefficients.

Table 8L: Bilateral SEHI Results, by makes&sizes&outlets(20*10*3)

Bilateral Comparison	Arithmetic: Base country weights (Laspeyres)		Current country weights (Paasche)		Fisher	Obs del/used	Volume del/used	EKS/ AEKS
France/NL	2.9919	(NL)	3.0364	(France)	3.0141	610/2721	120047/712354	3.0812/ 2.8982
NL/UK	3.1854	(UK)	3.3798	(NL)	3.2811	594/2086	90704/415024	2.8635/ 3.1449
UK/France	0.1137	(France)	0.1136	(UK)	0.1137	459/3182	85598/858169	0.1195/ 0.1159

Countries in parentheses are used for the weights and the hedonic coefficients.

81. Table 9 and 10 summarize some of the results from Tables 5 and 8. First, Table 9 shows that the selection of core characteristics for the aggregation can matter. The results from aggregating over screen size; screen size and flat screen; and screen size, flat screen and Nicam stereo are similar for each of Törnqvist and Fisher parities. This is to be expected from the loss of relatively little data as the list of core variables for aggregation increases, as shown in Tables 8G to 8I and 8A to 8C respectively. However when aggregating over makes; makes and screen sizes; and makes, screen sizes and outlets, there is much more variability in the results as is consistent with the increased loss of data as the aggregation expands as shown in Tables 8 D to 8F and 8J to 8L. For example, the Törnqvist NL/UK parities were 3.05 and 2.62 respectively when extending the aggregation from aggregating from screen size/make to screen size/make and outlet The number and volume of observations lost increased from 5.3 and 4.6 percent respectively to 14.4 and 10.0 per cent (Tables 8E and 8F) with this extension of the aggregation to outlets. We thus focus on the results from using screen sizes, flat screen and Nicam stereo as the core variables for aggregation with its more limited loss of data.

Table 9: Fisher and Törnqvist SEHI indices aggregated by different core variables*

	Törnqvist			Törnqvist		
	Screen size	*Screen size/Flat Screen*	*Screen size/Flat Screen/Nicam*	*Screen size*	*Screen size/Makes*	*Screen size/Makes/Outlets*
France/UK	8.13	7.83	7.59	8.33	8.00	6.97
France/NL	2.74	2.74	2.72	2.76	2.71	2.46
NL/UK	3.14	3.12	2.87	3.14	3.05	2.62

	Fisher			Fisher		
	Screen size	*Screen size/Flat Screen*	*Screen size/Flat Screen/Nicam*	*Screen size*	*Screen size/Makes*	*Screen size/Makes/Outlets*
France/UK	8.65	8.65	8.67	8.88	9.02	8.80
France/NL	2.99	2.99	2.94	2.94	2.98	3.01
NL/UK	3.14	3.17	3.15	3.32	3.33	3.28

***For screen sizes ten non-overlapping groups were used each of which had observations in each country**

Table 10: Summary Table of Binary and Multilateral Results

	1	2	3	4	5	6	7
Heading Parities for TV Heading Price Levels							
	FR/UK	FR/NL	NL/UK	Node	U.K.=100	U.K.=100	NL= 100
	fr/£	fr/fl	fl/£	fl/£	FR/UK	NL/UK	FR/NL
A. Pooled Regression Common Slopes	8.68	3.09	2.81		87.1	80.4	108.3
B. Pooled Regression Weighted	8.43	2.94	2.87		84.6	82.1	103.0
C.Törnqvist:Size/FlatScr./Nicam SEHI	7.59	2.72	2.87	2.65	76.1	82.1	95.4
EKS	7.66	2.70	2.84		76.8	81.2	94.7
D.Fisher: Size/FlatScr./Nicam SEHI	8.67	2.94	3.15	2.95	87.0	103.1	90.1
EKS							
Average June/July Exchange Rates	9.97	2.85	3.50				

82. Second, Törnqvist with its geometric aggregation and Fisher with its arithmetic basis can give very different results. Table 9 shows that for the aforementioned aggregation the Törnqvist France/UK parity was 7.59 compared with 8.67 for Fisher. This is an unusual result since both index number formulae are superlative and empirical work on a time series basis has found little difference between the two. Diewert (1995) however has advocated Fisher against Törnqvist from an axiomatic stance. The results in Table 9 find a greater consistency at different levels of aggregation for Fisher than for Törnqvist, thus recommending Fisher.

83. Third, the results from Table 8I for arithmetic aggregation by screen sizes, flat screen and Nicam stereo show minimal spread between the Laspeyres and Paasche indices. Indeed this finding is consistent across all levels of aggregation for the arithmetic indices (Tables 8G to 8I), but not so for the geometric

indices in Tables 8A to 8C. While such geometric bounds are often thought to be narrower than their arithmetic counterparts, this is not always the case. Shapiro and Wilcox note that:

84. Fourth, Table 10 compares the Fisher results aggregated with three core variables and minimal loss of data, with the regression estimates. As explained above, in the pooled equation with common coefficients the parities simply come out of the hedonic regression. For completeness we have also given in Table 9 the parities from the weighted regression in Table 5. These are rows [A] and [B] of Table 9; the parities are presented in two forms, as national currencies per £ and francs per guilder; and as a price level with the U.K. or Netherlands as 100. The TV parities generated from the equations in Table 5 are transitive, so that if one has francs and guilders per £, it is redundant to present fr/fl ratios. However, the SEHI results in Table 10 are not transitive, so there is an additional column for parities and price levels for the FR/NL comparison. Finally there is a column for an added NL/UK comparison based on a binary linking explained further below. The price levels are the parities divided by the exchange rate expressed as a percentage of the base country.

85. The results are quite close to the currency exchange rates for all three countries concerned. That they may differ is not questionable. Goldberg and Knetter (1997) comment on the 34% appreciation of the yen against the dollar between January 1994 and April 1995 which, after accounting for differential productivity growth and wage inflation implied a fall in labour costs in the U.S. relative to Japan of about 30%, yet:

> "..the suggested retail price of a large-screen SONY Trinitron actually *fell* [their emphasis] by 15% in the U.S. between 1994 and 1995. In the jargon of the international economists, there was incomplete (or in the case of the Trinitron, negative) “pass-through” of exchange rates to imported goods prices.” (Goldberg and Knetter, 1997 page 1244)."

86. They also discuss evidence of substantial differential pass-through rates for different countries for the same product, something not evidenced here, though quite possible when treating different countries as market segments.

87. The lower entries for Columns 1-3 of C are the EKS parities that are transitive. As described earlier, this is one way to produce multilateral parities from the SEHI results. Column 4 presents an alternative method of generating a transitive multilateral result. It follows the procedure that Robert Hill has advocated of using a minimal spanning tree that chooses a chain of binary comparisons the minimizes the Paasche-Laspeyere spread over all possible chains. With only 3 countries, the possible chains are few. In fact, there is rather little to choose between France and the U.K., but we believe France has a small claim to be the node to link the U.K. and the Netherlands. Thus the entry in column (4) is the derived fl/£ parity that is transitive with the columns 1 and 2 entries. Thus from comparing multilateral results, columns 1-3 are relevant for case A and B, columns 1,2, and 4 for the cases in the upper row of C-I, and columns 1-3 of the EKS rows for cases C-I.

88. Finally, multilateral SEHI can be calculated by estimating a reference parameter set that meets transitivity conditions and comparing each country’s prices with their counterparts from derived from this reference point Kokoski *et al.* (1999). The results are pending and should be available in a later version.

Part D CONCLUSION

89. This paper has explored the use of scanner data to estimate country parities. Hedonic regression equations were estimated and demonstrated the feasibility of the exercise. In addition a framework developed by Fixler and Zieschang (1992) and Feenstra (1995), and applied to interarea comparisons by Kokoski *et al.* (1999), was used to exploit the data on weights below the basic heading level to develop both binary and multilateral estimates. This framework was adapted to meet the special needs of scanner data. By the standards of these things the regression models fitted well. The successful application of the hedonic dummy variable framework, coupled to its application to a data source, which benefits from a

quite comprehensive coverage of transactions, is encouraging. The extension of the work to the SEHI framework which takes advantage of the information on weights available at this detailed level allowed an analysis of the results to be undertaken using different forms of aggregation and aggregation formulae. The importance of using core variables, which can be matched across countries, and thus not lose much information, was identified as were the closeness of Laspeyres and Paasche-type bounds for the preferred Fisher index. The extension to multilateral comparisons was also undertaken, as were comparisons of the EKS and Fisher results.

90. As might be expected from an exploratory study, these results raise questions and possibilities. One of the issues that seem crucial is outlet classification across countries, whether from scanner data or regular data question. It is one thing to base CPIs in a country on an outlet classification for temporal comparisons, but it is clearly important for spatial comparisons to make sure that the outlet classifications are referring to same aspects of service and price across countries. The fairly crude classification used here did not fully exploit scanner data's potential whereby the bar-code records details of the actual store in which each transaction took place, though such data were not available in the information set used in this study.

91. Another direction that this work might be extended to is the use of an hedonic framework for a broader range of basic headings than consumer durables. We believe it would be feasible to apply both scanner and price data from other sources into a framework such as the one attempted for TV sets. Obviously scanner data has the advantage of providing weights within a basic heading. However, we believe that a framework such as the one used here could also be used alongside one based on product matching representative items, but with a greater emphasis combining the core characteristics of a heading, in each country.

92. Finally, scanner data of the sort discussed are available on a cross-country and pooled time series basis and a natural extension of the work is the integration of consumer price indices and international price indices.

REFERENCES

Arguea N.M., Haseo C and Taylor G.A., (1994) Estimating Consumer Preferences using Market Data – An Application to US Automobile Demand, *Journal of Applied Econometrics*, 9, 1-18.

Balk B. (1995), On the Use of Unit Value Indices as Consumer Subindices. In B. Balk (editor), *Proceedings* of the International Working Group on Price Indices, Voorburg: Statistics Netherlands.

Balk (1996), A Comparison of ten methods for multilateral International price and Volume Comparisosns, *Journal of Official Statistics*, 12, 199-222.

Berndt, E.R., Grilches, Z., and Rappaport N.J. (1995), Econometric estimates of price indices for personal computers in the 1990s, *Journal of Econometrics*, 68, 243-68.

Diewert, W.E. (1976) Exact and Superlative Index Numbers, *Journal of Econometrics*, 4 115-45.

Diewert,W.E., (1978), Superlative Index Numbers and Consistency in Aggregation, *Econometrica*, 46, 883-900.

Diewert, W.E. (1995). 'Axiomatic and Economic Approaches to Elementary Price Indices'. Discussion Paper No. 95-01, Department of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W. Erwin (1999) "Axiomatic and Economic Approaches to International Comparisons," in Heston and Lipsey, eds. *International and Interarea Comparisons of Prices, Output and Productivity*, Committee on Research on Income and Wealth, NBER, Chicago: University of Chicago Press.

Feenstra RC (1995), Exact Hedonic Price Indices, *Review of Economics and Statistics*, LXXVII, 634-54.

Feenstra, R.C. and Diewert, E.W.(2000), Imputation and Price Indices: Theory and Evidence from the International Price Program, *National Bureau of Economic Research Working Paper*, forthcoming.

Fenwick, D., Ball, A., Silver, M. Morgan, P. (2000), Price collection and Quality Assurance of Item Selection in the Retail Prices Index: How can Scanner Data Help? Conference on the Use of Scanner Data in CPIs, NBER and CRIW, Washington, Sept 15-16, 2000, In Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 61, Chicago: University Press, forthcoming 2001.

Fixler, D and Zieschang, K (1992) Incorporating Ancillary Measures of Processes and Quality Change into a Superlative Productivity Index, *Journal of Productivity Analysis* 2: 245-267.

Goldberg, P.K. and Knetter, M.M. (1997), Goods Prices and Exchange Rates, *Journal of Economic Literature*, 35,3,pp. 1243-1272.

Goldberger, A.A. (1968), The Interpretation and Estimation of Cobb-Douglass Functions, *Econometrica,* 35, 3-4, pp.464-472.

Gordon, RL (1990) *The Measurement of Durable Goods Prices*, Chicago: University of Chicago Press.

Griliches, Z., (1990). 'Hedonic Price Indices and the Measurement of Capital and Productivity: Some Historical Reflections'. In *Fifty Years of Economic Measurement: The Jubilee Conference of Research in Income and Wealth, Studies in Income and Wealth*, Vol.54, NBER (eds. E.R. Berndt and J.E. Triplett), pp.185-206. Chicago University of Chicago Press.

Hausman J.R. (1997). 'Valuation of New Goods Under Perfect and Imperfect Conditions'. In *The Economics of New Goods*, Studies in Income and Wealth 58, NBER (eds. T. Bresnahan and R.J. Gordon), pp.209-248, Chicago and London: University of Chicago.

Heston, A. and Lipsey, R. (1999), *International and Interarea Comparisons of Prices, Output and Productivity*, Committee on Research on Income and Wealth, NBER, Chicago: University of Chicago Press.

Hill, R.J. (1999), International Comparisons Using Spanning Trees. In Alan Heston, and Robert E. Lipsey, R., (editors) *International and Interarea Comparisons of Prices, Output and Productivity*, Committee on Research on Income and Wealth, NBER, Chicago: University of Chicago Press, pp109-122.

Hoffmann, J (1998) Problems of Inflation Measurement in Germany, *Economic Research Group of Deutsche Bundesbank*, Discussion Paper 1/98, Frankfurt: Deutsche Bundesbank.

Ioannidis C and Silver, M (1999) Estimating Hedonic Indices: an Application to UK Television Sets, *Journal of Economics, Zeitschrift fur Nationalokonomie*, 69,1, 71-94.

Kennedy P. (1986) Interpreting Dummy Variables, *Review of Economics and Statistics,* vol. 68, 174-175.

Kokoski, Mary F., Brent R. Moulton and Kimberly D. Zieschang, (1999) "Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation," in Heston and Lipsey ed., *op. cit.*

Moulton, B., LaFleur, T.J., Moses, K.E. (1998). 'Research on improved quality adjustments in the CPI: the case televisions'. In Proceedings of the Fourth Meeting of the International Working Group on Price Indices (ed. W. Lane), pp.77-99, Washington DC: Bureau of Labor Statistics.

Rasmussen, DW and Zuehlke, TW (1990), On the Choice of Functional Form for Hedonic Price Functions, *The Review of Economics and Statistics*, November, 668-675.

Reinsdorf M (1996) Constructing Basic Components for the U.S. CPI from Scanner Data: A Test Using Data on Coffee, Working Paper 277, Bureau of Labor Statistics, Washington, D.C.

Richardson, D. (2000), Scanner Indexes for the CPI, Conference on use of Scanner Data in CPI, NBER and Conference on Research in Income and Wealth, Washington, Sept 15-16, 2000. In Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 61, Chicago: University Press.

Rosen S (1974) Hedonic Prices and Implicit Markets: Product Differentiation and Pure Competition, *Journal of Political Economy*, 82 34-49.

Silver, M (1995) Elementary Aggregates, Micro-Indices and Scanner Data: Some Issues in the Compilation of Consumer Prices, *Review of Income and Wealth*, 41, 427-38.

Silver, M (1999) An Evaluation of the Use of Hedonic Regressions for Basic Components of Consumer Price Indices, *Review of Income and Wealth*, 45, 1, 41-56.

Silver, M. and Heravi, S. (2000), The Measurement of Quality-Adjusted Price Changes, Conference on use of Scanner Data in CPI, NBER and Conference on Research in Income and Wealth, Washington, Sept 15-16, 2000. In Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 61, Chicago: University Press.

Summers, R. (1973), International Comparisons with Incomplete Data, *The Review of Income and Wealth*, March.

Triplett JE, (1987), Hedonic Functions and Hedonic Indices, in *The New Palgraves Dictionary of Economics,* 630-634.

Triplett, J.E. (1990). 'Hedonic Methods in a Statistical Agency Environments: An Intellectual Biopsy'. In *Fifty Years of Economic Measurement: The Jubilee Conference on Research in Income and Wealth, Studies in Income and Wealth*, Vol 56, NBER (eds. E.R. Berndt and J.E Triplett), pp.207-238. Chicago: University of Chicago Press.

United Nations Economic and Social Commission for Asia and the Pacific (1999) *ESCAP Comparisons of Real Gross Domestic Product and Purchasing Power Parities*, 1993, (ST/ESCAP/1918), Bangkok.

United Nations Statistical Division (1992) *Handbook of the United Nations International Comparison Programme*, United Nations, New York.

White, H, (1980) A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity, *Econometrica*, 48, pp.721-746.

Yotopoulos, Pan A. and Jenn-Yih Lin (1993), "Purchasing Power Parities for Taiwan: The Basic Data for 1985 and International Comparisons," *Journal of Economic Development*, (Korea) No.1.).