An Approach to Pooled Time and Space Comparisons

Ana Aizcorbe and Bettina H. Aten[*]

SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity
Vancouver, BC
June 30-July 3, 2004

Introduction

The problem of reconciling aggregate Gross Domestic Product (GDP) price indexes across space and time has received more attention as international statistical offices and institutions such as the European Union, the OECD (Organization for Economic Cooperation and Development) and the World Bank, increase the frequency and scope of their GDP comparisons at purchasing power parities (PPPs). The problem has also been present in the construction of the Penn World Tables (PWT), where successive benchmark studies of the International Comparison Programme (ICP) from 1970, 1975, 1980 and 1985 had to be reconciled with more recent comparisons. Summers and Heston (1984) termed it "consistentization" and others, including Heston, Summers & Aten (2001), Aten and Heston (2002), Hill (2002), and the OECD have examined a number of approaches to obtaining various degrees of temporal and spatial consistency. Varjonen (2001) summarizes the problem succinctly as quoted below:

> *GDP comparisons are organised annually for countries participating in the Eurostat PPP Programme. In the OECD Programme, the interval between comparisons is three years and results for intermediate years are derived using implicit price indices of GDP. However, comparison results are in strict sense valid only for those years when comparisons are carried out, that is in benchmark years. They are based on prices and price structures prevailing in those years, and updating results for other years using implicit price indices of GDP provides only a proxy solution because changes in price structures are not taken into account.*

Hill (2004) provides a thorough discussion of the history and taxonomy of panel methods, including an extension of the minimum-spanning tree approach that links countries successively across space and time. He applies seven different varieties of index number methods [Geary-Khamis, EKS (Elteto, Koves and Szulc) and Minimum Spanning Trees] to a disaggregated panel data set, merging the Harmonized Index of Consumer Prices of the European Union over the period 1995-2000 with OECD cross-section data for fifteen countries.

[*] Aizcorbe (Ana.Aizcorbe@bea.gov) and Aten (Bettina.Aten@bea.gov) are Economists at the Bureau of Economic Analysis, 1441 L St., NW. Washington, DC, 20230.

Our objective in this paper is to add a time dimension to a multilateral price comparison method called the country product dummy (CPD) method that has gained some attention in the 2004 ICP, and to illustrate its application on panel data with a simple example. This time-extended CPD specification has the property that in its elementary form, and when the data are complete, cross-country comparisons reduce to those obtained from single-equation CPD estimates, and price comparisons over time reduce to those obtained from a variant of the hedonic regressions used in the price measurement literature. When expenditure weights are available, this correspondence no longer holds, but arguably, the desirable properties of the dummy regression approach remain, and would provide a baseline check on other multilateral methods that attempt to provide spatial and temporal consistency.

The Extended Country Product Dummy Approach

The CPD was originally derived in the context of incomplete price data when multilateral comparisons were needed across a number of categories for many countries. It was not possible then, nor is it now, to price all items within each category for every country. Summers (1973) felt it inefficient to discard all price data on items where country prices were missing and proposed an alternative. The CPD "uses all available data in an economical way" [p.2, Summers (1973)]. A similar problem arises in studies comparing the price of goods over time, as it is not always possible to 'match' specific products over long periods of time, particularly in rapidly-changing industries. In the time series literature, hedonic regression techniques are used to construct price measures when data sets are incomplete, and in particular we look at the specification proposed by Aizcorbe, Corrado and Doms (2000) that is algebraically identical to the CPD, but uses time instead of country as a dummy variable (referred to as TPD in this paper). We propose two combinations of these single equation models, the first a Time-interaction-Country Product Dummy approach (TiCPD) and the second an even simpler model, without the interaction terms, labeled Time-Country-Product Dummy or TCPD approach.

An important difference in CPD and hedonic techniques lies in their theoretical underpinnings. The CPD method is based on a statistical model where regression coefficients are interpreted as means of underlying distributions. In contrast, the theoretical basis for hedonic regressions is a structural model of consumer behavior [Rosen (1974)]. In that view, regression coefficients are interpreted either as reduced-form parameters reflecting changes in demand and supply conditions [Pakes (2003)] or as parameters corresponding to consumer preferences [Moch and Triplett (2002)]. As Sergeev (2004) notes, if we do not "combine the CPD method with hedonics then it is better to use the presentation of the CPD as an index number method", which is the path followed below.

a. The Country Product Dummy (CPD) Method

The CPD method has been used in the ICP benchmark comparisons below the basic heading level, and in interarea work by the Bureau of Labor Statistics [Kokoski, Moulton & Zieschang (1999)]. More recently it has been discussed in a weighted form for use at the more aggregate level in research on the new round of the ICP [Aten & Heston (2002), Cuthbert (1997). Deaton (2004), Diewert (2002), Rao (2002), Sergeev (2004)]. In extensions of this framework, the ICP handbook[1] introduces the useful distinction between CPD and CPRD (Country Product Representative Dummy). The CPRD adds a third dimension, "representativity", for products that are representative or common in their countries [see also Sergeev (2003) and (2004)]. This representativity dummy can be thought of as a proxy for weights and could also be applied in the time-product-dummy approach discussed below. This would allow for better[2] comparisons of price changes over time between products that are still common in one period for some countries, for example, older computers and peripherals, but less common in other countries. Of course, actual weights would be better than a dummy variable, but first we consider the unweighted models.

The specification of the original, unweighted CPD is shown in (1.1).

(1.1)

$$\begin{array}{l} CPD: \\ \ln P_{ij} = \sum_{j=1}^{M} \alpha_j G_{ij} + \sum_{i=1}^{N} \beta_i C_{ij} + \varepsilon_{ij} \ , \exists \text{ time period } t = 1,...T \\ \beta_i = 0 \text{ (for any } i = 1,..., N) \end{array}$$

Where $\ln P_{ij}$ is the log of the price of good j in country i, (G_{ij}, C_{ij}) are dummy variables for good j and country i, respectively, with j=1,…,M and i=1,…,N. The antilog of the country dummy coefficients (the antilog of the β_is) are the estimated parities for each country i (P_i) at the aggregate level, with one country as the base (P_i=1). Summers showed that when there are no missing data, the differences between the β_is reduce to the geometric means of the relative prices[3]. For countries c and d, this difference is:

(1.2)

$$\beta_c - \beta_d = \sum_{j=1}^{M} (\ln P_{cj} - \ln P_{dj}) / M$$

Where M is the number of goods in each country.

[1] Chapter 10 'The Extended CPD method or CPRD method' in the ICP Handbook 2004 (v.21/05/04). Online at http://siteresources.worldbank.org/ICPINT/Resources/Ch10.doc.

[2] Better in the sense that it would reduce the bias in coefficient estimates that may result from products that are unusual – either because they are rarely available or because they are highly priced, yet included in a comparison because they 'match' a specification.

[3]The ratios of the aggregate GDP parities are termed price levels when normalized on a common currency. This is exactly equal to the EKS (Elteto, Koves and Szulc) method when there are no missing data. The EKS equals the geometric mean of the Fisher price ratio averages across all possible pairs of countries.

b. The Time Product Dummy (TPD) Method

The specification used by Aizcorbe, Corrado and Doms (2000) that is applied to time-series data is shown in (1.3).

(1.3)

$$\begin{array}{|l|}\hline TPD: \\ \ln P_{jt} = \sum_{j=1}^{M} \alpha_j{}' G_{jt} + \sum_{t=1}^{T} \gamma_t T_{jt} + \mu_{jt} \ , \exists \text{ country } i = 1,...N \\ \gamma_t = 0 \text{ (for any } t = 1,...,T) \\ \hline \end{array}$$

Where $\ln P_{jt}$ is the log of the price of good j in time period t, (G_{jt}, T_{jt}) are dummy variables for good j and time period t, respectively, with j=1,…,M and t=1,…,T. This specification—algebraically identical to the CPD specification—may be viewed as a special case of a traditional hedonic regression[4]. Differences in the coefficients on the time dummies are interpreted as measures of price change over time. Again, when there are no missing data, the differences between the γ_ts reduce to the geometric means of the relative prices:

(1.4)

$$\gamma_t - \gamma_{t-1} = \sum_{j=1}^{M} (\ln P_{jt} - \ln P_{jt-1}) \big/ M$$

Despite the similarities of the two specifications, the use of CPD regressions from two time periods to infer something about price change in one country over time will not give the same answer as a direct TPD regression using time series data for that country. In Hill's (2004) nomenclature, the CPD is an "average price-method", and results are based on the price structure of the countries included in each time period. Similarly, the separate TPDs are based on the price structure of each country over time. However, it is possible to specify a single regression that provides direct estimates of price levels across space and time, and we turn to two elementary versions below.

c. The Time Country Product Dummy (TCPD) Method

The simple Time Country Product Dummy shown in Equation (1.5) assumes an average price structure across both space and time. The time change is constant across countries, and the price levels are constant over time. The purpose is to provide an

[4] See Triplett and MacDonald (1977), Berndt and Griliches (1993), Diewert (2003) and Heravi, Heston and Silver (2003)

aggregate measure of price change, either by partitioning countries into regions, Southeast Asia versus Latin America, or grouping time-periods into decades, for examples.

(1.5)

$$\begin{aligned}&TCPD:\\&\ln P_{ijt} = \sum_{j=1}^{M} \alpha''_j G_{ijt} + \sum_{i=1}^{N} \beta'_i C_{ijt} + \sum_{t=1}^{T} \gamma'_t T_{ijt} + \upsilon_{ijt}\\&\beta'_i = 0 \text{ (for any } i = 1,\dots,N\text{) and } \gamma'_t = 0 \text{ (for any } t = 1,\dots,T\text{)}\end{aligned}$$

The country coefficients (β'_i) equal the average of the separate country coefficients from the CPDs across all years [β_i in Equation (1.1)], and the time coefficient (γ'_t) equals the average of the separate time coefficients from the TPDs across all countries [γ_t in Equation (1.3)]. That is:

$$\begin{aligned}\beta'_i &= \sum_{t=1}^{T} \beta_{i(t)} \Big/ T \quad (\forall i = 1,\dots,N)\\ \gamma'_t &= \sum_{i=1}^{N} \gamma_{(i)t} \Big/ N \quad (\forall t = 1,\dots,T)\end{aligned}$$

d. The Time-interaction-Country Product Dummy (TiCPD) Method

The second combination of the CPDs and TPDs is the Time-interaction-CPD (TiCPD) model, shown in Equation (1.6)[5].

(1.6)

$$\begin{aligned}&TiCPD:\\&\ln P_{ijt} = \sum_{j=1}^{M} \alpha'''_j G_{ijt} + \sum_{i=1}^{N}\sum_{t=1}^{T} \delta_{it} C_{ij} T_{jt} + \upsilon_{ijt}\\&\delta_{it} = 0 \text{ (for any } i = 1,\dots,N \text{ and for any } t = 1,\dots,T\text{)}\end{aligned}$$

Where G_{ijt}, are dummy variables for good j, in country i and in time period t, and $C_{ij}T_{jt}$ are dummy variables for each combination of country and time period with j=1,…,M, i=1,…,N and t=1,…,T.

[5]Note that it would be possible to add separate country and/or time dummies to Equation (1.6), but the results would be identical to the TiCPD. This is because additional country or time dummies would over-identify the model, forcing more parameters in the interaction term of Equation (1.6) to be zero.

The relative price levels across countries for a time period in the TiCPD (1.6) are equal to those in the CPDs (1.1), and price changes in the TiCPD (1.6) are equal to those in (1.3) for one country. These relationships are analogous to structuring a 'Chow' test using dummy interaction terms[6]. That is:

$\delta_{it} - \delta_{kt} = \beta_i - \beta_k$ (differences between countries i and k for time t)

and

$\delta_{it} - \delta_{is} = \gamma_t - \gamma_s$ (differences betwen time periods t and s for country i)

In their unweighted or elementary forms, the indexes reduce to simple differences of means with no consistency problems. One possible application of the TCPD or the TiCPD in the ICP context is to the data that must be collected by individual countries and groups of regions over a period of several months, as some form of aggregation and consistency must be obtained at the basic heading level[7] before further processing takes place.

Complete Data

We illustrate this point using selected data from the OECD benchmark studies for 1990, 1993 and 1996[7,8]. *Table 1a* provides parameter estimates corresponding to the three regression specifications. The top panel provides results from single-equation CPD regressions, where a separate regression was done for each time period. As may be seen, each regression is normalized to the UK[9], so that the estimates from each regression provide price differences relative to the UK for that time period. Similarly, the second panel provides parameter estimates from single-equation TPD regressions, where a separate regression was done for each country and the estimates are expressed relative to 1996. Finally, the third panel provides estimates from the combined specification, where the normalization is on prices in the UK in 1996. We make several observations.

---

[6] See for example, Kennedy (1985) p. 186.

[7] Items below the basic heading level generally have no weights associated with them.

[7] All data, including the 1999 round, were kindly provided by Seppo Varjonen, but 1999 was dropped because it uses a different set of basic heading classifications. The basic headings consist of the components of GDP used in the OECD PPP studies. More recent PPP data, revised from 1995 onwards to meet new SNA guidelines, are available from the OECD website.

[8] There are twenty four overlapping countries in the three years, but for the Tables below, only five countries (Australia, Canada, Japan, the Netherlands and the United Kingdom) were chosen. These five countries have no missing data for the available headings. Thus i=1,…,5 countries, j=1,…199 goods, t=90, 93, 96 time periods. The 199 goods and services correspond to expenditure headings of GDP. The total number of observations is thus 2985.

[9] The choice of base country does not affect the results. An average of the countries could also be used, and is the preferred method in Eurostat and the OECD.

First, note that some of these parameters are identical. The CPD estimates for 1996 (in the last column in the top panel) are identical to the corresponding estimates from the combined specification (in the last column in the bottom panel): both provide price comparisons for each country's prices in 1996 relative to prices in the UK in 1996. Similarly, the TPD estimates for the UK (last row in the 2nd panel) are identical to those in the combined specification (last row in the 3rd panel): they both give price changes in the UK in each period relative to 1996.

Second, it is easy to verify that differences between coefficients are also identical. For example, the differences in CPD coefficients for any two countries in a time period **($\beta_i - \beta_k$)** equal those of the TiCPD coefficients **($\delta_{it} - \delta_{kt}$)**. That is, the difference between rows for any column in the CPD regressions equals those for the TiCPD regression. For example, in 1990, the difference between Australia and the UK in the CPD is 0.777, and from the TiCPD it is also 0.777 (0.690-(-0.087). Similarly, the difference between any two time periods **($\gamma_t - \gamma_s$)** in the TPD regressions equals the corresponding difference **($\delta_{it} - \delta_{is}$),** in the TiCPD . For example, Australia in 1996 – Australia in 1990 is -0.002 in the TPD, and 0.0688-0.0690 = -0.002 in the TiCPD.

Thus, if we are looking at relative levels, either across countries or between time periods, the TiCPD provides the same answers as separate CPD or TPD models, with the advantage that it normalizes the relationships on a single country and time period.

Table 1b looks at the estimated parities of the TCPD for two regional groupings: Asia (Australia and Japan), Europe (the Netherlands and the UK) and for the group of five countries. They can be seen to equal the average of the relevant countries from the TiCPD specification.

Missing Data

When there are missing data, the CPD estimates still reduce to a geometric mean, where the missing data are replaced by predicted values from the regression. Unfortunately, different regressions will likely generate different predicted values so that the correspondences discussed above no longer hold. This is because increases in the number of missing observations reduce the part of the estimate that is shared across equations—i.e., the geometric mean and increase the weight on the part of the equation that is specific to each specification—the residuals[10].

---

[10] To see this, consider the case where good *g* is missing from country *d*. In that case, it can be shown that the CPD comparison of prices in the two countries is given by:

(i) $(\beta_c - \beta_d) = (M_c / M_d)\,[\Sigma_{j=1,Mc}(\ln P_{ij} - \ln P_{dj}) \;/\; M_c\,] + (1 / M_d)\;\;[(\ln P_{cg} - \ln P^{hat}_{dg}) \;/\; M_c\,]$

where M_c and M_d denote the number of prices observed for countries *c* and *d* and the imputed price for good *g* is: $\ln P^{hat}_{dg} = \alpha_g + \beta_d$

note that the difference in the second term ($\ln P_{cg} - \ln P^{hat}_{dg}$) reduces to a difference of parameters and the residual: $\ln P_{cg} - \ln P^{hat}_{dg} = (\alpha_g + \beta_c + \varepsilon_{cg} - (\alpha_g + \beta_d) = \beta_c - \beta_d + \varepsilon_{cg}$. Substituting this expression into (i) and simplifying yields a sum of two terms: one with observed prices and the other with the residual.

Below, we illustrate this point by generating successive datasets with varying degrees of completeness, beginning with approximately 4% of its observations missing and increasing to 74% missing. *Table 2a* shows the correspondence of the missing observations and their representative share-weights. In Group 1, there are 199 observations per country per time period, but by Group 7, the number drops to approximately 50 observations, with fewer overlaps between countries

The characteristics of the missing data may also introduce bias to the estimated coefficients. If, for example, all the missing data in country A are 'important' or representative, but those in country B are not, the predicted parity between A and B will have an upward bias (if the price of a product is expected to be higher when it is not representative)[11]. In the examples above, price observations were deleted for observations where the weights were less than 0.01 percent of GPD (for each country), and incrementally up to observations less than 0.50 percent of GDP. The former removed 4% of total observations and the latter 74% of total observations. Since the share weights of the products vary by country, these deletions would correspond to removing progressively more representative observations from all countries, but not necessarily the same products over time or across countries, hence it is *not* introducing the type of bias that one might expect if only representative products were priced in one country but not the other.

Tables 2b shows the percentage difference between the estimated parities for the TiCPD and the separate CPDs when there are missing observations (Group 1 with 4% missing observations). The differences in percentage points are calculated as percentage differences relative to the TiCPD of each country relative to the UK, in each time period. *Table 2c* is identical, for Group 7, with 74% missing observations.

The percentage differences (third panel in *Table 2b and 2c*) are small, below one percent in the data set with only 4% of its observations missing, but rise to over three percent (Australia's parity using TiCPD is 3.4% higher than the corresponding CPD for 1993, relative to the UK) in the data set with 74% missing observations. Although it is not our objective to suggest a cut-off for using pooled regressions versus separate regressions, the exercise points to the need for a more formal variance test when comparing specifications with many missing observations.

If we examine only one specification, the TiCPD, and look at the differences in estimated parities as we increase the number of observations, a similar pattern emerges. That is, the average absolute difference increases relative to the complete data set, beginning with an average of less than one percent, and rising to 5.6% for Group 7. These

---

[11] See for example, Case 2, ICP Handbook, Chapter 10 at http://siteresources.worldbank.org/ICPINT/Resources/Ch10.doc.

are shown in *Figure 1.* The general pattern is that with unweighted data, differences in specification are somewhat less than differences due to large amounts of missing data.

Share-Weighted Dummy Models

Researchers have considered weighted versions of the CPD method at the aggregate GDP level [Cuthbert (1997), Diewert (2002), Sergeev (1982), Rao (2002)] , where the weights are nominal expenditure shares. Will the consistency between separate CPDs, separate TPDs and the TiCPD hold? We examine this by looking at the weighted versions of equations (1.1) through (1.6), where the weights (w_{ijt}) are nominal expenditure shares that add to one in each country and year, termed share-weights:

(1.7)

$$w_{ijt} = \frac{(pq)_{ijt}}{\sum_{j=1}^{M}(pq)_{ijt}} \quad (\exists i = 1,\ldots,N \; and \; \exists t = 1,\ldots,T)$$

Table 3a compares the coefficient estimates from the share-weighted CPD, TPD and TiCPD specifications. Although the differences are not large, in the tenths of a percent range, the resulting parities and price changes will depend on which specification is used. One can alternatively use a Laspeyres-type weight in the TiCPD that brings the coefficients closer to the 1990 CPD, or a Paasche-type TiCPD, that is closer to the separate 1996 CPD, but neither will exactly equal the CPD parities or the TPD time changes.

Figure 2 shows the graph of differences between the complete and missing data sets for the TiCPD parity estimates. These are noticeably smaller than the differences in *Figure 1* for the unweighted parities. The largest value in *Figure 2* is for Group 7, with an average absolute difference of only 2.5%, whereas in the unweighted case, the Group 7 average absolute difference is 5.6%. For 50% or fewer missing observations, the differences among TiCPD parities are half a percent or less.

These examples reinforce the notion that when no weights are available, a representative version of the dummy approach may have more of an impact than the separate versus pooled specification, as the parities appear very sensitive to the number of missing observations. It is only in the extreme case, when nearly three-quarters of the observations are missing, that the parities in the weighted TiCPD show such sensitivity. The same is true for the weighted case, namely that missing observations create larger differences among the same specification than among different specifications (TiCPD versus CPD) for the same data set.

Conclusions

One of the criticisms of the unweighted CPD and related methods is that they are average price methods and likely to underestimate the price differentials across countries (or over time). The advantages of the various dummy-method approaches shown here lie in their simplicity and transparency, as the properties of regression models are well understood and easily reproduced. Additionally, goodness-of-fit comparisons can be made, and variations and extensions of the variables and their coefficients can provide more flexibility to the framework, for example, by adding representativity in the unweighted specifications.

In this paper, we illustrate how the unweighted combined model, the TiCPD, exactly equals separate benchmark comparisons or separate time-series deflators from hedonic regressions, and how they may be useful in the new round of ICP. The ICP is collecting price observations at a detailed level for a number of countries over an extended period of time (approximately two years). The TiCPD could be used for obtaining parities at the basic heading level below which weights are not available. Once weights are introduced, the equality among separate regressions and the combined specification no longer holds but the differences among estimated parities are in the one percent range. Although these results are particular to this sample, share-weighted CPD estimates have been shown to approximate other index number formulae [Rao (2002), Diewert (2002), Sergeev (2004)] and the share-weighted TiCPDs may serve a useful purpose as a baseline check on other methods that attempt to move GDP PPPs over time.

REFERENCES

Aizcorbe, Ana, Carol Corrado and Mark Doms (2000), 'Constructing Price and Quantity Indexes for High Technology Goods', Paper presented at NBER/CRIW Summer Institute, July.

Aten, Bettina and Alan Heston, (2002), 'Benchmark Reconciliations Revisited', International Association for Research in Income and Wealth, Stockholm, Sweden, August 18-24.

Berndt, Ernst R. and Zvi Griliches, 'Price Indexes for Microcomputers: An Exploratory Study', in *Price Measurements and Their Uses*, ed. Murray Foss, Marilyn Manser, and Allan Young. (Chicago: The University of Chicago Press, 1993), pp. 63-93.

Cuthbert, J. and M. Cuthbert (1988), 'On Aggregation Methods of Purchasing Power Parities', OECD working paper.

Cuthbert, J (1997), 'Aggregation of Price Relatives to Basic Heading Level: Review and Comparison', ISI meeting, Istanbul, 18-26 August.

Deaton, Angus with Jed Friedman, Vivi Alatas (2004), 'Purchasing Power Parity Exchange Rates from household survey data: India and Indonesia', Research Program in Development Studies, Princeton University.

Diewert, W. E. (2002), 'Weighted Country Product Dummy Variable Regressions and Index Number Formulae', Department of Economics, Discussion paper 02-15, University of British Columbia, Vancouver, BC, Canada.

Diewert, W. E. (2003) 'The Treatment of Owner Occupied Housing and Other Durables in a Consumer Price Index, Department of Economics, Discussion paper 03-08, University of British Columbia, Vancouver, BC, Canada.

Heravi, Saeed, Alan Heston and Mick Silver (2003) 'Using Scanner Data to Estimate Country Price Parities: An Exploratory Study," Review of Income and Wealth, March.

Heston, Alan, Robert Summers and Bettina Aten (2001), 'Price Structures, the Quality Factor and Chaining', *Statistical Journal of the United Nations Economic Commission for Europe*, Vol.18, 2001.

Hill, Robert (2002), 'Measuring Price Differences Across Space and Time: The Case of the European Union's Harmonized Index of Consumer Prices', International Association for Research in Income and Wealth, Stockholm, Sweden, August 18-24.

Hill, Robert (2004), 'Constructing Price Indexes Across Space and Time: The Case of the European Union', SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver, BC, June 30-July 3.

Kennedy, Peter (1985), '*A Guide to Econometrics*', 2nd edition, The MIT Press, Cambridge, Massachusetts.

Kokoski, Mary and Brent Moulton and Kim Zieschang (1999), 'Interarea Price Comparisons for Heterogenous Goods and Several Levels of Commodity Aggregation', in *International and Interarea Comparisons of Income, Output and Prices*, ed.,by Alan Heston and Robert Lipsey, 123-66. University of Chicago Press.

Moch, Dietmar and Jack E. Triplett, (2002), International Comparisons of Hedonic Price Indexes for Computers: A Preliminary Examination'. Presented at the 27th General Conference of the International Association for Research in Income and Wealth. Djurhamn, Sweden, August.

Pakes, Ariel, (2003), 'A Reconsideration of Hedonic Price Indices with an Application to PC's', *American Economic Review*, Vol: 93 (5) pp. 1578 – 1596.

Rao, D.S. Prasada (2002), 'On the equivalence of Weighted Country Product Dummy Method and the Rao System for Multilateral Price Comparisons', School of Economics, University of Queensland, Brisbane, Australia.

Rosen, S. (1974), 'Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition', *Journal of Political Economy*, Pp. 34-55.

Selvanathan, E., and D. S. Prasada Rao (1994), '*Index Numbers: a Stochastic Approach*', Ann Arbor, the University of Michigan Press.

Sergeev, Serguei (1982), 'Multilateral Methods for International Comparisons', Ph.D. dissertation, Central Statistical Committee of the Soviet Union, Moscow (in Russian).

Sergeev, Sergei (2003), 'Equi-representativity and some Modifications of the EKS Method at the Basic Heading Level', Working Paper No. 8,UN ECE, Geneva, March 31-April 2. (http://www.unece.org/stats/documents/2003/03/ecp/wp.8.e.pdf)

Sergeev, Sergei (2004), 'The Use of Weights within the CPD and EKS Methods at the basic heading level', Statistics Austria, mimeo.

Summers, Robert (1973), 'International Price Comparisons based upon Incomplete Data', *The Review of Income and Wealth*, Volume 19, Issue 1, March.

Triplett, Jack E. and Richard J. McDonald (1977) 'Assessing the Quality Error in Output Measures: The Case of Refrigerators.' *The Review of Income and Wealth* 23(2): 137-56.

Varjonen, Seppo (2001), 'Consistency between GDP based on PPPs and National Accounts Time Series', OECD Paris, France, October 2001.

Table 1a. Comparison of Parameter Estimates: Unweighted Regressions

	1990	1993	1996
CPD Regressions: β_i			
Australia	0.777	0.765	0.688
Canada	0.723	0.651	0.585
Japan	5.786	5.759	5.597
Netherlands	1.240	1.187	1.071
UK	0	0	0
TPD Regressions: γ_t			
Australia	0.002	0.027	0
Canada	0.051	0.016	0
Japan	0.102	0.112	0
Netherlands	0.081	0.066	0
UK	-0.087	-0.050	0
TiCPD Regression: δ_{it}			
Australia	0.690	0.715	0.688
Canada	0.635	0.601	0.585
Japan	5.699	5.709	5.597
Netherlands	1.152	1.137	1.071
UK	-0.087	-0.050	0

Note: Number of observations in each regression are as follows: CPD regressions contain 995 observations per time period; TPD regressions contain 597 observations per country and the TiCPD regression contains 2985 observations.

Table 1b. Estimated Parities : TiCPD vs. TCPD

	1996-1990	1996-1993
TiCPD Regressions: antilog(δ_{i1996}- δ_{it})		
Australia	-0.002	-0.027
Canada	-0.051	-0.016
Japan	-0.102	-0.112
Netherlands	-0.081	-0.066
UK	0.087	0.050
TCPD Regressions: antilog(γ_{1996}- γ_t)		
Asia	-0.052	-0.069
Europe	0.003	-0.008
World	-0.030	-0.034

Note: Number of observations in each regression is as follows: Asia (Australia, Japan) 1194 observations, Europe (Netherlands, UK) 1194 observations, World (5 countries) 2985 observations.

Table 1b (cont.). Estimated Parities : TiCPD vs TCPD

	1990	1993	1996	All Years
	TiCPD Regressions: antilog(δ_{it}- δ_{UKt})			TCPD Regressions: antilog(β_i-β_{UK})
Australia	2.18	2.15	1.99	2.10
Canada	2.06	1.92	1.79	1.92
Japan	325.8	317.1	269.7	303.2
Netherlands	3.45	3.28	2.92	3.21
UK	1	1	1	1

Note: TiCPD and TCPD regressions contain 2985 observations.

Table 2a. Distribution of Missing Data

Group	Share Weight Cut-off	Total number of Observations	Percent missing
1	0	2985	0%
2	0.0001	2867	4%
3	0.0002	2711	9%
4	0.0005	2386	20%
5	0.0010	1970	34%
6	0.0020	1473	51%
7	0.0050	787	74%

Note: Share weights sum to one for each country in each time period. For example, the 0.0001 cut-off corresponds to observations with weights less than 0.01 percent of total expenditures.

Table 2b. Comparison of Estimated Parities from CPDs vs TiCPDs relative to the UK (4% Missing Data)

	1990	1993	1996
CPD: antilog(β_c)-antilog(β_{UK})			
Australia	2.172	2.152	2.000
Canada	2.065	1.923	1.788
Japan	320.2	312.7	264.7
Netherlands	3.456	3.282	2.915
UK	1.000	1.000	1.000
TiCPD: antilog(γ_{ct})-antilog(γ_{UKt})			
Australia	2.176	2.140	1.989
Canada	2.070	1.916	1.772
Japan	320.3	310.7	265.2
Netherlands	3.467	3.275	2.902
UK	1	1	1
Difference (percentage points)			
Australia	0.19%	-0.56%	-0.58%
Canada	0.26%	-0.39%	-0.91%
Japan	0.05%	-0.67%	0.19%
Netherlands	0.34%	-0.22%	-0.45%
UK	0%	0%	0%

Note: Parities are calculated as differences in the regression coefficients: CPD parities are for countries relative to the UK: antilog(β_c) – antilog(β_{UK}) . The TiCPD parities are calculated relative to the UK for each year: antilog(γ_{c1990})- antilog(γ_{d1990}) for 1990, antilo(γ_{c1993})- antilog(γ_{d1993}) for 1993, and so on. Differences are expressed as a percentage of the TiCPD parities: ((antilog(γ_{c1990})- antilog(γ_{d1990}))- (antilog(β_c)- antilog(β_{UK})))/ (antilog(γ_{c1990})- antilog(γ_{d1990}))*100. For example, the Australia-UK parities for 1990 based on the TiCPD estimates are 0.19 percent higher than the parities from the CPD estimates.

Table 2c. Comparison of Estimated Parities from CPDs vs TiCPDs relative to the UK (74% Missing Data)

	1990	1993	1996
CPD: antilog(β_c)-antilog(β_{UK})			
Australia	2.166	1.984	1.997
Canada	1.998	1.832	1.755
Japan	309.9	265.4	249.2
Netherlands	3.381	3.312	3.221
UK	1	1	1
TiCPD: antilog(γ_{ct})-antilog(γ_{UKt})			
Australia	2.126	2.054	2.023
Canada	1.942	1.861	1.738
Japan	303.2	264.4	255.8
Netherlands	3.364	3.395	3.170
UK	1	1	1
Difference (percentage points)			
Australia	-1.9%	3.4%	1.3%
Canada	-2.9%	1.6%	-1.0%
Japan	-2.2%	-0.4%	2.6%
Netherlands	-0.5%	2.5%	-1.6%
UK	0%	0%	0%

Note: Parities are calculated as differences in the regression coefficients: CPD parities are for countries relative to the UK: antilog(β_c) – antilog(β_{UK}) . The TiCPD parities are calculated relative to the UK for each year: antilog(γ_{c1990})- antilog(γ_{d1990}) for 1990, antilo(γ_{c1993})- antilog(γ_{d1993}) for 1993, and so on. Differences are expressed as a percentage of the TiCPD parities: ((antilog(γ_{c1990})- antilog(γ_{d1990}))- (antilog(β_c)- antilog(β_{UK})))/ (antilog(γ_{c1990})- antilog(γ_{d1990}))*100. For example, the Australia-UK parities for 1990 based on the TiCPD estimates are 1.9 percent lower than the parities from the CPD estimates.

Figure 1. Unweighted TiCPDs

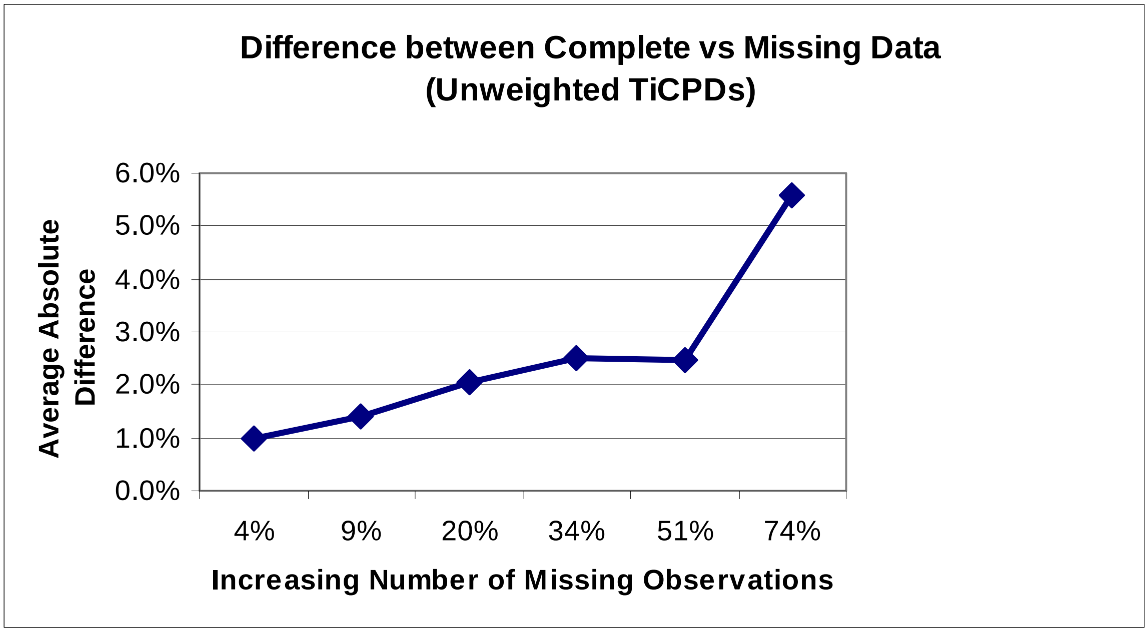

Figure 2. Weighted TiCPDs

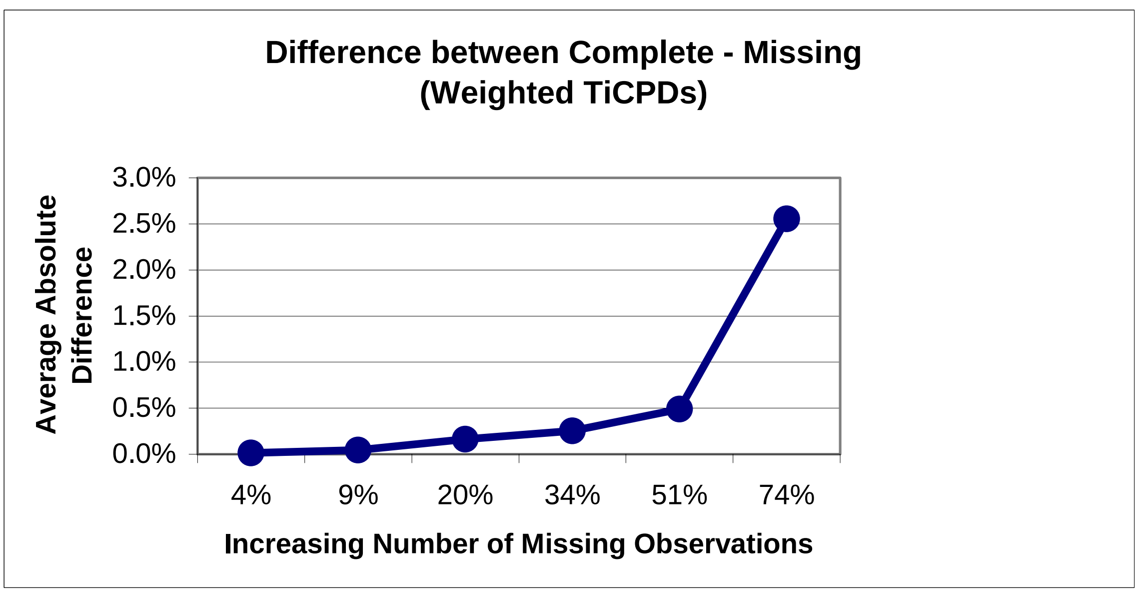

Table 3a. Comparison of Parameter Estimates: Share-weighted Regressions

	1990	1993	1996
CPD Regressions: β_i			
Australia	0.847	0.745	0.703
Canada	0.785	0.670	0.600
Japan	5.783	5.662	5.556
Netherlands	1.289	1.204	1.152
UK	0	0	0
TPD Regressions: γ_t			
Australia	0.041	0.040	0
Canada	0.080	0.052	0
Japan	0.134	0.077	0
Netherlands	0.056	0.036	0
UK	-0.093	-0.025	0
TiCPD Regression: δ_{it}			
Australia	0.731	0.728	0.693
Canada	0.670	0.648	0.595
Japan	5.681	5.628	5.540
Netherlands	1.194	1.171	1.135
UK	-0.102	-0.033	0

Note: Number of observations in each regression are as follows: CPD regressions contain 995 observations per time period; TPD regressions contain 597 observations per country and the TiCPD regression contains 2985 observations.