IMF Working Paper

Do Unit Value Export, Import, and Terms of Trade Indices Represent or Misrepresent Price Indices?

Mick Silver

IMF Working Paper

Do Unit Value Export, Import, and Terms of Trade Indices Represent or Misrepresent Price Indices?

Prepared by Mick Silver

Authorized for distribution by Kimberly Zieschang

May 2007

Abstract

This Working Paper should not be reported as representing the views of the IMF.
The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Unit value export and import indices compiled from returns to customs authorities are often used as surrogates for price indices in the measurement of inflation transmission, terms of trade (effects), and to deflate import and export value series to derive volume series. Their widespread use is mainly due to their relatively low cost compared with establishment price surveys. This paper provides evidence of substantial bias in their representation of such price changes. Their continued use would mislead economic analysis. The paper considers the efficacy of alternative strategies for their improvement, and argues for a move to establishment-based price surveys.

Contents   Page

Figures

Tables

I. INTRODUCTION

The main purpose of this paper is to consider whether export and import unit value indices derived from customs data, and commonly used as surrogates for export and import price indices, represent or misrepresent such price changes. Unit value indices as measures of price changes of imported and exported goods serve economic analysis in many important ways. They are used as short-term indicators of inflation transmission, to measure changes in a country's terms of trade (effect), and as deflators of export and import values to yield measures of changes in export and import volumes. Yet in spite of their widespread use they are subject to well–recognized bias. The issue of concern is whether such bias misleads economists in their economic analysis to the extent that their compilation and use should not be recommended. Also of importance, is to consider what might be done by statistical agencies if unit value indices are found wanting.

Bias in unit value indices is mainly attributed to changes in the mix of the heterogeneous items recorded in customs documents, but may also arise from the poor quality of recorded data on quantities. The former is particularly important given the increasing differentiation of products and turnover of differentiated products that is a feature of modern markets. Unit value indices may suffer further due to an increasing irrelevance of the source data with first, increasing proportions of trade being in services and by e-commerce, and hence not covered by customs data and second, a constraint on the coverage of such data for countries in customs and monetary unions, for which intra-union trade date may no longer be regularly collected.

Few deny, including United Nations (1981), that narrow specification price indexes provide the best measures of relative price change and that, *a priori*, there are potentially significant biases in using customs unit values to measure price developments in international goods trade. Yet, unit value proxies for narrow specification price data (on specific products collected from the establishments with transactions in them) are still used because they are by-products of existing customs administration systems and have relatively low incremental cost compared with the price surveys of establishments needed for narrow specification prices. In the absence of a systematic examination of the evidence and in view of the low cost of the data, the bias in unit value has been judged tolerable enough that countries are advised to continue compiling them if they do not produce narrow specification price indices. Notwithstanding the putative low cost of obtaining unit values, this paper adds to the weight of advice recommending against their use as proxies for well-defined price indexes; specifically:

- Indexes of unit values have no well-defined relationship over time to the desired narrow specification price indexes; indeed there are substantial discrepancies in direction and magnitude.
- Unit value indexes have no predictive power for narrow specification price indexes.
- For terms of trade indices the discrepancies are worse.
- Measures of the terms of trade effect (for real national income) and deflated volume changes are vastly different when measured using unit value indices as opposed to price indices.

- There is no evidence of homogeneous product classes for which unit value indices may be reliably used.
- Significant unit value bias arises within strata defined at levels of detail well beyond that available in customs systems.
- Unit values are only applicable to trade subject to customs administration, and thus cannot be compiled for trade within economic unions.

The concern over bias in unit value indices is not new. Early critical studies of unit value bias as measures of import and export price changes and terms of trade include Kravis and Lipsey (1971 and 1985). The United States discontinued publication of unit value trade indices in 1989 due to concern over bias and introduced trade price indices based on establishment surveys. More recently, in a speech on statistical challenges raised by globalization, González-Páramo (2006) noted that the European Central Bank is looking forward to the moment that appropriate import and export price indices, instead of unit value changes, become available for the Euro area as a whole, in order that central banks have key information to anticipate domestic inflationary pressures.[1]

As background, Section II of this paper examines recommendations on export and import price measurement given in United Nations statistical manuals and handbooks. In Section III possible sources of bias are outlined in terms of the compliance of unit value indices with desirable axiomatic properties of prices indices, and then summarized. Section IV considers the important matter of evidence. Such evidence is by its nature limited to countries that compile both price indices and unit value indices. It is also limited to the fact that the deficiencies in unit value indices are not measured against a perfect benchmark, the price indices themselves having deficiencies. Yet, as will be outlined, unit value indices suffer mainly from not comparing the prices of like with like, while establishment–based price indices do so. In this important regard this and other studies ask: how well do unit value indices stand up against price indices designed to overcome one of their major failings. The study provides in Section IV some new results using recent data from Germany and Japan. In Section V we consider alternative strategies for countries to measure trade price indices and in Section VI draw conclusions.

II. INTERNATIONAL RECOMMENDATIONS

This is a background paper to the draft Export and Import Price Index (XMPI) Manual[2] developed under the aegis of the United Nations Inter-Secretariat Working Group on Price Statistics (IWGPS) to update the existing United Nations (1981) guidelines. In a major departure from United Nations (1981) the focus of the draft Manual was largely on the establishment survey pricing approach: to obtain the prices of well specified products from

---

[1] The European Commission Short-Term Statistics Council Regulation 1165/98 amended by 1158/2005 introduced requirements for the compilation of import and export price indices based on price surveys.

[2] Available at: http://www.imf.org/external/np/sta/tegeipi/index.htm.

establishments and monitor their prices over time. The use of unit value indices from customs data as surrogates for price changes was considered to be limited and was to be confined to very strictly homogeneous products. The limitations of unit-value indices were recognized in both the United Nations (1979) manual on producer price indices and United Nations (1981). Yet the sole use of unit-value indices was proposed in the latter publication as being an appropriate strategy for statistical authorities with resource constraints. Indeed, United Nations (1983) was composed of case studies on the development and implementation of the two main approaches—the survey pricing approach as used by the Federal Republic of Germany and the unit-value approach as used by Norway—to assist countries in initiating and developing their trade price change measures. In practice the vast majority of countries currently produce unit value indices as the only available information on trade price changes and thus as surrogates for price indices.

The main rationale for unit value indices was the limited resources required to compile them. United Nations (1981) laid down a strategy for countries with a tight budget; they should only use unit value indices with disaggregation by county of origin/destination where deemed appropriate. Further, items with unit value changes outside specified bounds were advised to be excluded unless the exclusions amounted at least half the value of the classification category; the groups "machinery and transport equipment" and "miscellaneous manufacturing" were recommended to require special treatment. The advice for countries on an average budget was to similarly rely on customs documentation, but use product analysts to detect abnormal changes, and undertake a more intensive analysis of the distribution of unit values within a category. Extreme values/changes were to be questioned. Well-endowed countries were advised to use establishment-based price surveys, possibly jointly with unit value indices.

The *System of National Accounts 1993* notes that especially in the field of foreign trade statistics information on prices and volumes are not adequate. Due to the problem of compositional change in non-homogeneous units: "Unit value indices cannot therefore be expected to provide good measures of average price changes over time." (Para. 16.13).

III. UNIT VALUE INDICES AND THEIR BIAS

In this section the nature of the bias in a unit value index arising from changes in the compositional product mix is first outlined in Section A, and then again considered more formally in Section B by means of the properties of the unit value index in relation to the main axiomatic tests used in index number theory to justify formulas. Section C provides a summary of issues of concern in relation to unit value indices.

A unit value index, P_U, for period 1 relative to a reference period 0 is given for comparison over $m=1, ...,M$ prices, p_m^1, and quantities, q_m^1, in period 1 and over $n=1,...,N$ prices, p_n^0, and quantities, q_n^0, in period 0 by:

$$P_U(p^0,p^1,q^0,q^1) \equiv \left(\frac{\sum_{m=1}^{M} p_m^1 q_m^1}{\sum_{m=1}^{M} q_m^1} \right) \Bigg/ \left(\frac{\sum_{n=1}^{N} p_n^0 q_n^0}{\sum_{n=1}^{N} q_n^0} \right).$$ (1)

A. Unit Value Bias Illustrated

United Nations (1981) usefully provides an illustration of unit value bias which we include here. Consider trade in refrigerators. With the exception of the "size" of the refrigerator, assume the mix of all price-determining characteristics—including the brand, frost-free, color, energy-efficiency, possession of ice-making feature, drink dispenser, and so forth—remains constant over the periods compared, or are proxied by "size." Assume further that there is a meaningful division of "size" into the three groups of "small", "medium" and "large" and a change in purchasing patterns towards larger refrigerators. In the illustrative example below, from United Nations (1981, page 15), refrigerator prices, p, double for each size group and there is a shift to the quantities, q, sold now, in proportion to 2, 3, and 5 going from smallest to largest, from what was then 5, 3, 2, though total quantity remains the same over time. The value, v, is given as $p \times q$.

	Size of refrigerator											
	Small			Medium			Large			All sizes		
Period	q	p	v	q	p	v	q	p	v	q	p	v
Now	2	2	4	3	4	12	5	6	30	10	4.6	46
Then	5	1	5	3	2	6	2	3	6	10	1.7	17

Since prices in each size group have doubled, a weighted average of these price changes, $\sum_i w_i \left(p_i^{Now} / p_i^{Then} \right)$, over the i size groups is 2.0. But the change in the unit value is 4.6/1.7=2.71. There is an upwards bias in the UVI due to the change in the product mix towards more expensive refrigerators.

B. Unit Value Indices: the Test Approach

The unit value index fails the *Proportionality Test*: $P(p,\lambda p,q^0,q^1) = \lambda$ for $\lambda > 0$; that is, if all prices are multiplied by the positive number λ, then the new price index is λ. The unit value index only satisfies the proportionality test in the unlikely event that relative quantities do not change (preferences are linearly homogeneous and identical—Bradley, 2005). Changes in the index can thus reflect both changes in price and changes in the product mix over the two periods compared. Since the index should only measure price changes, the index number formula has a potential bias.

It follows from the failure of the proportionality test that the unit value index also fails the *Identity or Constant Prices Test*: $P(p,p,q^0,q^1) = 1$; that is, if the price of every good is identical during the two periods, then the price index should equal unity, no matter what the quantity vectors are. It only satisfies the identity test if relative quantities, that is the composition of the products compared, do not change.

The unit value index, however, satisfies the proportionality in current period prices test, $P(p^0,\lambda p^1,q^0,q^1) = \lambda P(p^0,p^1,q^0,q^1)$ for $\lambda > 0$; that is, if all period 1 prices are multiplied by the positive number λ, then the new price index is λ times the old price index.

The unit value index fails the *Invariance to Changes in the Units of Measurement (commensurability) Test*: $P(\alpha_1 p_1^0,...,\alpha_n p_n^0; \alpha_1 p_1^1,...,\alpha_n p_n^1; \alpha_1^{-1} q_1^0,...,\alpha_n^{-1} q_n^0; \alpha_1^{-1} q_1^1,...,\alpha_n^{-1} q_n^1)$ $= P(p_1^0,...,p_n^0; p_1^1,...,p_n^1; q_1^0,...,q_n^0; q_1^1,...,q_n^1)$ for all $\alpha_1 > 0, \dots, \alpha_n > 0$; that is, the price index does not change if the units of measurement for each product are changed. For example, if the measurement of one of the products changed from pounds weight to kilograms, the index should not change.

Unit value indices pass the other main index number tests, including the time reversal test, the circularity test, and the product test. However, that it is affected by changes in the composition of products, and (changes in) its units of measurement—that is, it fails the proportionality and commensurability tests—is critical to concluding that it is an inappropriate measure.

That a price index can be affected by changes in relative quantities is a serious deficiency. The essence of the fixed basket concept of price indices is the need to hold quantities constant over time. There is a very real sense in which a unit value index should not be properly described as a price index unless applied to transactions for homogeneous products and thus, by definition, the composition of products cannot change.

In addition, where there is more than one product, there would need to be some natural units of measurement for the unit value index. For example, say an index covered two products, product A measured in terms of the number of items imported/exported and product B in terms of the weight of the items. Then a change in the units of product B, *for both periods*, from say pounds to kilograms would affect the results of the unit value index. That is, different results would arise if pounds were used to measure total quantity in each period as opposed to kilograms. That the results of the index depend on the units adopted results in a quandary as to which units are correct.

A particular instance of the effect of the failure in the commensurability test impacting adversely on the unit value index is one in which the quality of products imported/exported changes. When this occurs, the actual units of measurement may not change, but the implicit unit of productive service or utility would change, and bias the index. Accounting for the effects of quality changes on prices is difficult enough for index number work based on price surveys (ILO *et al.*, 2004a). Customs data on quality characteristics is likely to fall well short of the corresponding information that would normally be available from establishments producing for export, or purchasing as imports.

Balk (1998) draws the following conclusion: the unit value bias will be equal to zero for a comparison between periods 0 and 1 if one or more of the following conditions are met:

(1) all base period prices p_m^0 are equal to each other and all current period 1 prices p_n^1 are equal to each other;

(2) all quantity relatives q_m^1 / q_m^0 are equal to each other; and

(3) there is no correlation between p_m^0 and q_m^1 / q_m^0, and also no correlation between p_n^1 and q_m^1 / q_m^0.

These are all highly restrictive conditions. The first condition requires product homogeneity to an extent that defeats the purpose of a price index, in that if all prices were equal in each period, then there is no index number problem; the price change of a single product would suffice. The second condition, is the assumption required above for satisfaction of the identity and proportionality tests. If all quantity relatives were equal, and this were known, the price index number problem would be solved by dividing the total outlays by this single quantity relative. At the heart of the index number problem is that such quantity changes cannot be assumed to be the same. The third condition arises from the fact that if price relatives and quantity relatives are uncorrelated, then a change in prices would not affect quantity relatives, and vice versa, and there will be no change in the composition mix due to relative price changes. There may be some markets in which there is market failure or temporal inconsistencies, but for the large part the laws of economics cannot be assumed away.

Bradley (2005) compares unit value indices with theoretical counterparts derived using economic theory, for which price and quantities are related. He concludes that a unit value index will be an appropriate aggregate price index only under extreme conditions.

C. Unit Value Indices: the Cause for Concern

Unit value indices derived from data collected by customs authorities are mainly used by some countries as surrogates for price changes at the elementary level of aggregation. The following are grounds upon which unit value indices can be deemed unreliable:

- Bias arises from compositional changes in quantities and quality mix of what is exported and imported. Even with best practice stratification the scope for reducing such bias is limited due to the sparse variable list—class of (quantity) size of the order and country of origin/destination)—available on customs documents. Indeed Párniczky (1974) shows that it does not follow that such breakdowns are always beneficial to a UVI;
- For unique and complex goods, model pricing can be used in establishment-based surveys where the respondent is asked to price each period a product, say a machine with fixed specified characteristics. This possibility is not open to unit value indices.

- Methods for appropriately dealing with quality change,[3] temporarily missing values, and seasonal goods can be employed with establishment-based surveys to an extent that is not possible with unit value indices;
- The information on quantities in customs returns, and the related matter of choice of units in which the quantities are measured, has been found in practice to be seriously problematic;
- With customs unions countries may simply have limited intra-area trade data to use;
- An increasing proportion of trade is in services and by e-trade and not subject to customs documentation;
- Unit value indices rely to a large extent on outlier detection and deletion. Given the stickiness of many price changes, such deletions run the risk of missing the large price catch-ups when they take place and understating inflation.

A main advantage of the use of unit value indices is held to be their coverage and relatively low resource cost. However, the unit values used are drawn as non-random samples and exclude: products traded irregularly; that have no quantity reported (especially for parts and machinery); have low-value shipments; and erratic month-to-month changes. The extent of such exclusions is substantial, as illustrated in Section VE below. Establishment-based surveys can be quite representative. Often a small number of wholesalers or establishments are responsible for much of the total value of imports or exports and, assuming cooperation, will be a cost-effective source of reliable data. Further, good sampling, can, by definition, realize accurate price change measures and finally, the *value* shares of exports and imports, obtained from customs data, will form the basis of information for weights for establishment-based surveys.

Errors and biases are recognized in price index measurement. Silver (2006) in Chapter 11 of the draft export and import price index Manual outlines such sources of error and bias and refers to the chapters of the draft manual in which methods to mitigate them are outlined in some detail—methods that statistical authorities are well versed in for compiling consumer and producer price indices. These include obtaining detailed specifications of representative goods and services so the prices each month of like are compared with like, something customs data does not allow for. The consumer price index (CPI) and producer price index (PPI) compilation practices have benefited from much research and experience since the publication of United Nations (1981) and extensive guidelines on good practice are available in the CPI and PPI Manuals (ILO *et al.*, 2004a and 2004b respectively), the benefits of which should carry over to XMPI practice—see Trewin (2006) for an example of country practice.

IV. EVIDENCE

We adopt for brevity the terminology of PI to refer to an establishment-survey based price index as distinguished from a customs-data based unit value index, UVI.

---

[3] Von der Lippe (2007) shows that adjustments for quality change is one reason why price indices are less volatile than unit value indices.

Given the above reasons to expect that UVIs will not be suitable surrogates for PIs, it is necessary to consider the empirical evidence available on the nature and extent of any differences. The evidence is presented in this section first at the aggregate level for some existing studies and then for Germany and Japan as new results. Results at a more disaggregated level will be considered in Section V for Germany and some other European countries.

A. Some Existing Studies

Alterman (1991) compared price changes between March 1985 and June 1989 for the United States (U.S.) as measured by UVIs and by PIs based on establishment surveys that replaced them.[4] For imports, over this period, the PI increased by 20.8 percent and the UVI increased by 13.7 percent. For exports, the figures were much closer, 13.0 and 12.2 percent for the price survey and UVI respectively. Some of the difference between the two series may be attributed to their use of different periods for weights. However, when price survey indices were recalculated using the same weights as the UVIs, the differences were exacerbated: a 20.6 percent and 16.4 percent increase for the import and export price indices respectively. The average (absolute quarter-on-quarter) UVI change for imports and exports respectively were 27 and 70 percent larger than the corresponding PI changes. One method of considering whether such differences matter is to evaluate the implications of such discrepancies for deflation of the foreign trade component of the national accounts. Alterman (1991) found that the annualized second-quarter 1989 "real" trade deficit in March 1985 dollars would have been $128.4 billion if deflated by UVIs, but just $98.8 billion, 23 percent less, if deflated by the PIs.

Such findings are not new. Kravis and Lipsey (1971) found that the prices of manufactured goods exported by developed countries to developing countries had risen over about twenty years by 75 percent, as compared to the l4 percent shown UVIs. Kravis and Lipsey (1985) found a decrease in the terms of trade of manufactures relative to all primary products between 1953 and 1976 of over 36 percent, using price indices, almost a quarter greater than that suggested by the UVIs (28 percent). With a further correction for quality change the price data suggested a fall in manufactures terms of trade of over 45 percent, more than 50 percent greater than UVIs.

B. Unit Value Indices and Price Indices for Germany and Japan Compared

This study compares UVIs and PIs for both Germany and Japan for exports and imports using monthly data for1996:7 to 2006:9 from the IMF's *International Financial Statistics* (*IFS*). Results are presented to ascertain the magnitude of the discrepancies between UVIs and PIs for measures of export and imports price changes as: short-run indicators (month-on-month and month-on-12 month comparisons); long-run cointegration; and predictive ability (leading indicators). Export and import UVIs are also used for the measurement of changes in terms of

---

[4] The official U.S. UVIs were discontinued by the Bureau of the Census in 1989 so these figures are the latest available estimates.

trade and discrepancies between UVIs and PIs used for this purpose are also considered, as is the use of UVIs for the measurement of the terms of trade effect and as a long-run deflator.

Short-run indicator

Figures 1 to 4 compare month-on-month changes between UVIs and PIs for exports and imports for Germany and Japan. While for Germany UVIs are clearly much more volatile than PIs, the relative volatility of UVIs for Japan is much less marked. For exports and imports in both countries substantial discrepancies between PIs and UVIs can be identified. Note that for Japan in some periods, especially the earlier years, UVIs and PIs appear to track each other, but this cannot be relied upon and breaks down in later periods.

Table 1 provides summary statistics on the magnitude of the (absolute value of the) discrepancies: the ratio of UVIs to PIs for export and imports in both countries along with the root mean squared deviation between the UVIs and PIs. The mean month-on-month discrepancy is calculated as : $\sum_{m=2}^{M} [|((UVI_m / UVI_{m-1})/(PI_m / PI_{m-1})) - 1|]/(M - 1)$ (where $|\ |$ denotes the modulus—absolute value) and other summary measures accordingly.[5] The mean discrepancy for imports to Germany was 1.1 percent. A need exists to draw a line as to the extent to which a discrepancy is acceptable, in the sense that on empirical grounds the matter of choice between a UVI and PI is of little consequence. A discrepancy of 0.011 implies that if the month-on-month change in the PI was unity, no change, then the UVI would take a value of a 1.1 percent change on average; or if the PI was a 1 percent change, the UVI would be $1.01 \times 1.011 = 1.021$, a 2.1 percent change. Such discrepancies can be regarded as seriously misleading for economists. The discrepancy for individual months can of course be much larger than this mean, as reflected by the standard deviation of 1.0 percent and maximum of 7.3 percentage points for these month–on-month changes. The month–on-12 month changes benefit from some of the positive and negative discrepancy over the 12 months compared cancelling. Yet with a mean 12 month PI change for German imports of 4.75 percent, a discrepancy of 1.8 percent on average and standard deviation of 1.6 percent (Table 1) provides no cause for complacency.

Such discrepancies might be argued to be a problem of magnitude rather than direction. Table 2 shows the extent to which positive (and negative) changes in UVI indices are mirrored by positive (and negative) ones in PIs. For about 25 percent of month-on-month comparisons the signs differ; that is in one quarter of comparisons the economist would read prices were rising (falling) when they were falling (rising). The results are better for month-on-12 month comparisons, but this cannot be relied upon, as German exports demonstrate.

---

[5] Von der Lippe (2007) in a study of German data uses $\sum_{m=1}^{M} (UVI_m - PI_m)/M$, i.e. differences in the index number levels which understates the mean differences as positive and negative differences to some extent cancel. With inflation, it also gives more importance to later period data (higher index levels) than data from earlier periods. Yet it contains interesting information on the higher levels of volatility of unit value indices compared with that of price indices.

Cointegration

It may be argued that the concern should be with the long-run equilibrium between the alternative measures and the extent of the short-run error correction. The results of unit root tests for Germany and Japan's import and export UVIs, and corresponding PIs, are given in Table 3 for month-on-month changes, month–on-12 month changes, and the index. The null hypothesis of a unit root is rejected at a 5 percent level for all month-on-month comparisons, and for all month-on-12 month comparisons, with the exception of German exports. The UVIs and PIs are generally not I(1) and thus it is not possible to establish cointegrating relationships. Although it is the changes that are the concern of economic analysts, we consider the series themselves. The index series, that have not been subject to differencing, are all I(1) and the cointegration test statistics all have *p-values* that exceed 0.05; the null hypothesis of a unit root in the cointegrating regression cannot be rejected at this level and we conclude at this level that the linear combination of the unit value and price index is not I(0), so they are not cointegrated.

Prediction

A further question is whether UVIs have any information content useful to *predict* next month's PI. Changes in past UVIs may be used as indicators of future trade price changes. We estimated for each PI series:

$$PI_t = \alpha_0 + \sum_{j=1}^{n} \beta_{1j} UVI_{t-j} + \varepsilon_{1t} \tag{2}$$

and tested the null hypothesis that $\beta_{1j} = 0$ for all j and observed the sign if the null hypothesis was rejected (the signs were all positive when significant). Table 4 shows the F-test for this null hypothesis to be rejected in three out four cases for the month-on month indices and in all cases for month-on- 12 month changes. Thus for most cases UVIs have some predictive power in relation to PIs. However, when they have, it is of little substance. Table 4 provides the means of the PIs and standard errors of the regression. It can be seen that the predictive intervals are quite wide: for example the 95 percent interval for German imports is ± 1.8 percent. While lagged UVIs have some predictive power regarding PIs, there is the question as to whether lagged UVIs have any contribution to predictive power *over and above* that of lagged values of the PIs themselves: that the UVIs Granger-cause (GC) the PIs; that lagged UVIs better predict the PIs than lagged PIs would themselves. The test requires ordinary least squares (OLS) (given the stationarity) estimates of :

$$PI_t = \alpha_{10} + \sum_{j=1}^{n} \alpha_{1j} PI_{t-j} + \sum_{j=1}^{n} \beta_{1j} UVI_{t-j} + \varepsilon_{1t} \qquad \text{and}$$

$$UVI_t = \alpha_{20} + \sum_{j=1}^{n} \alpha_{2j} PI_{t-j} + \sum_{j=1}^{n} \beta_{2j} UVI\pi_{t-j} + \varepsilon_{2t} \tag{3}$$

and tests for the joint hypothesis that $\beta_{1j} = 0$ and $\alpha_{2j} = 0$ for all j. The Granger-causality (GC) tests in Table 4 find that in half the cases lagged UVIs contain no predictive power over and

above lagged PIs, but this is not to demonstrate that in the cases where there is some such power UVIs GC PIs, since the GC tests reject the null hypothesis that PIs GC UVIs.

The above evidence is that UVIs are misleading proxies for PIs: they mislead in the sense that the relative and absolute errors can be substantial and that in many cases the signs of changes are wrong. There is no evidence of long-run (cointegrating) relationships and UVIs are of little further help for predicting PIs.

Terms of trade indices

The concern above was with bias in UVIs as indicators of import and export price inflation, as measured by PIs. Yet another use of UVIs is in the measurement of changes in the terms of trade (ToT) of a country, determined as the ratio of the price index of exports to the price index of imports. If export and import UVIs are used as surrogates for export and import PIs, and export and import UVIs are biased to the same extent and direction, the UVIs will provide a correct indication of changes in the terms of trade as the bias cancels. However, if the export and import UVIs are biased in different directions, then the ToT UVI bias will compound. Our analysis is similar to that used above, but for ToT measured using UVIs as against PIs. Table 5 shows the average discrepancy between the UVI and PI measures of ToT. The discrepancies are generally *larger* than the substantial discrepancies found in Table 1 for the export and import indices. For example, the mean month-on–month discrepancy for ToT changes for Germany was 1.3 percent compared with 1.1 and 0.9 percent respectively for imports and exports. For month-on-12 month changes the ToT discrepancy for Japan was 3.7 percent compared to 2.4 and 2.5 percent for imports and exports respectively. The ToT discrepancy for Japan implies that if the TOT PI change was unity, the ToT UVI index would on average show a month–on-month change of 3.7 percent with, given its standard deviation over time of 10 percent (0.10) and maximum of 70 percent (0.70), the possibility of very misleading results.

Table 6 presents the results on the percentage of months in which TOT UVIs have the same sign to their change as ToT PIs. ToT indices perform worse on average than export and import indices, Table 2, in this respect. The month-on-month ToT indices had the wrong sign in over one-third of the month-on-month comparisons. Japan's month-on-12 month series: had the wrong sign in 22 percent of cases while the export and import series had the wrong sign in 15 and 4 percent respectively (Table 2).

Table 7 finds that ToT indices measured by both UVIs and PIs do not have unit roots and thus are not cointegrated. The ToT indices measured as lagged UVIs are shown in Table 8 to have some predictive information in relation to ToT PIs but it is very weak, as demonstrated by the standard errors of the regression in relation to the means of the TOT PIs. In the case of Germany, lagged ToT UVIs have no predictive ability over ToT PIs, while in Japan, lagged UVIs have some such ability; PIs are found to have a similar predictive effect for UVIs so we cannot establish that ToT UVIs Granger-Cause ToT PIs.

Terms of trade effect

Terms of trade effect, or trading gain (loss), is a measure of the effect on income of changes in the terms of trade of a country, the relative price change of imports against exports. The *SNA 1993* (Eurostat and other, 1993) outlines the method of calculation as:

$$T = \frac{X-M}{P} - \left(\frac{X}{P_X} - \frac{M}{P_M} \right) ,$$ (4)

where the first term is a measure of the goods and services balance (exports of goods and services (X) less imports of goods and services (M)) using a single deflator, P and the second term is the goods and services balance by taking the difference between a volume (say constant price) measure of exports and a volume measure of imports, that is after X and M have been deflated by respective price indices for exports and imports, P_X and P_M. Note in the second term how, for example, as export prices increase more slowly than import prices, the larger the sum deducted from the first term is, and hence the smaller the terms of trade effect is. Note also that the magnitude of the terms of trade effect is contingent on the deflator in the first term. There is no agreement as to the best deflator to use for this component (Silver and Madhavy, 1989). The interpretation of the trading gain would be in terms of the gain in purchasing power with regard to the bundle of such goods and services to which P relates.

Table 9 shows the annual terms of trade effects for Germany and Japan in each case measured in terms of the change in prices for the preceding year and in terms of the purchasing power of imports, $P = P_X$, though a similar conclusion arises from using exports or some average of the two. Data are also provided for each year on the country's trade balance. The effect of using UVIs to calculate the ToT effects as against PIs is most marked . Note how, for example, in 2005 Japan's trade balance of 6,956 billion Yen is eliminated by the adverse change in its terms of trade when using PIs, but only halved when using UVIs.

Long run changes and deflation

Table 10 is concerned with comparing long-run changes between UVIs and PIs. One way of considering this is in terms of the use of such indices as deflators. In Table 10 the values of exports and imports of Germany and Japan are deflated over the period from 1999 to 2005 by corresponding UVIs and PIs and the results are compared. The volume of exports by Japan can be seen to have increased by 50 percent over this period when a UVI deflator is used, but the increase is halved when a PI is used. The volume of imports by Germany was about constant over this period when a UVI deflator is used, but fell by about 10 percent when a PI deflator is used. Such discrepancies can be misleading for economic analysis and policy decision-making.

A framework for testing whether there is a difference in long-run changes is provided in Annex 1. This uses simple *t*-tests to test the null hypothesis whether the differences between the mean log inflation rates for UVIs and PIs are equal to zero. The null hypothesis could not be rejected

for exports for Germany (t=0.828; p-value=0.41); imports for Germany (t=0.219; p-value=0.83); exports for Japan (t=-1.554; p-value=0.12); imports for Japan (t=0.742; p-value=0.46); ToT for Germany (t=0.463; p-value=0.64); and ToT for Japan (t=-1.946; p-value=0.052). There is no evidence of long-run differences by this approach or any of the preceding analysis.

The evidence in Section B is that export and import UVIs are inadequate surrogates for their PI counterparts when used in economic analysis. Such analysis includes their use in the measurement of short- and long-run inflation, prediction, terms of trade, terms of trade effects, and as deflators. Indeed, the evidence is that they are seriously misleading.

V. WHAT IS TO BE DONE?

Given what should be grave concern over the use of UVIs, there is the practical matter of what should be done. UVIs are used by most countries and a move to PIs has resource consequences.

One possibility is to identify whether there are particular products more prone to UV bias and utilize UVIs for these sub aggregates in a hybrid overall index. This has the resource advantage of undertaking price surveys only for products for which they are necessary. The efficacy of such advice depend on the extent to which reliable UVIs will be available at a disaggregated level.

A. Use Unit Value Sub-indices for Homogeneous Product Groups: the Reliability of Sub-indices

Disaggregated export and import UVIs and PIs for Germany

We extend the study to disaggregated monthly data for Germany. Such data are for export and import UVIs and PIs for the period from January 2000 to November 2006 at the 4-digit level of the Statistical Classification of Products by Activity in the European Economic Community, 2002 version, (CPA). The relative discrepancies between month-on-month changes in the UVIs and PIs were calculated in the same manner as in Table 1 for each of the 150 class series available for both UVIs and PIs at this level. The 15 classes in the lowest percentile were then identified and their mean, maximum, minimum, and standard deviation month-on-month change over the period January 2000 to November 2006 calculated; the results are presented in Table 11. These are the best product classes in the sense that they have the least average discrepancies. The results in Table 11 identify the magnitude of the discrepancies of what should be the most homogeneous product classes, to enable us to assess whether they are usable in a reliable way, and to characterize the product groupings in some meaningful way.

The best two product classes, the manufacture of motor vehicles and the manufacture of pulp, have mean month-on-month discrepancies of 2.00 percent; that is, if the PI showed no change the UVI would show a 2 percent change. At the bottom of the best percentile range is the manufacture of fertilizer and nitrogen compounds with a discrepancy on average of 4.00 percent. These averages of course have standard deviations that at 6 and 7 percent for the best

two classes and 15 percent for the worst end of the best percentile, demonstrate along with the maximum values in Table 11, serious cause for concern.

The results in Table 11 reflect an average discrepancy and volatility in excess of that for the weighted aggregate import price index found for Germany above. On aggregation there must be some smoothing of these fluctuations, though not to an extent, as revealed in the previous section, that renders them as suitable surrogates for PIs. Had the results been more favorable it would have been useful to attempt to characterize these "best" product classes for use in the compilation of UVIs in hybrid UVI/PI indices. Unexpectedly they include three heterogeneous classes composed of "other" and "n.e.c." products. There is also some concentration around plastic products and motor vehicle related activities. But given the size of the discrepancies these are not useful groupings.

In informal discussions with UVI compilers as to which product groups give reliable series, the anecdotal information confirmed the above results. For example, for Denmark, coal was initially considered to be homogeneous, but they discovered that the UVIs were completely unpredictable and uncorrelated with price changes. The explanation was that coal is not just coal; it varied in quality in terms of the amount of energy it produced, the extent of cleaning/filtering rquired, and residual use, say for road construction.

The PLANISTAT Report

Also of particular help in examining UVI and PI discrepancies at a disaggregated level is the extensive study PLANISTAT Europe Reports, Decoster (2003a and 2003b) commissioned by Eurostat for European Member States. In particular, the Second Report provided a comparative analysis of import PIs and UVIs for Finland, Germany,[6] Netherlands, and Sweden. The monthly import indices used are those provided by these countries to Eurostat. The UVIs were extracted from the Comext database.[7] The series are available at a 3-digit level CPA and while results at aggregated levels are provided, they are unweighted and not useful for our purposes. The series are monthly from January 1995 (=100) to September 2001. Some of the results are provided in Table 12.[8]

Table 12 shows that, for example for Finland, of the 77 product groups at the 3-digit level CPA for which import data were available, 17 percent had an average discrepancy between UVIs and XMPIs of less than 2.5 percent, and about another 40 percent between 2.5 and

---

[6] Data were not available at a detailed level of aggregation for Germany.

[7] Unit value indices are subject to outlier detection and revision and the series available in Comtext may differ from those available from the individual countries in this regard.

[8] The PLANISTAT report was undertaken by Renaud Decoster. The results provided here are based on the worksheets of summary measures for the individual series, provided to the author of this chapter by Mr. Decoster. The author acknowledges Mr. Decoster's help and advice. The above tables do not appear in the report, but are based on the data series used for the report. The conclusions drawn here and in the report are very similar and differ only in that a less favorable consideration is given in the report to unit value indices than here.

5 percent. There was less difference between UVIs and PIs for Sweden with about one-third of 3-digit product groups having a discrepancy of less than 2.5 percent. Bear in mind that a discrepancy of 0.025 implies that if the month-on-month change in the PI was zero, no change, then the UV index would take a value of a 2.5 percent change on average. Such discrepancies can be regarded as misleading for economists and 83, 87, and 66 percent of 3-digit groups in Finland, Netherland, and Sweden respectively had discrepancies that on average exceeded this value. Table 12 shows the average discrepancies for the three countries to be 5.3, 5.4, and 4.1 respectively and the minimum discrepancy for a 3-digit group to be between 1.5 and 2 percent—at best still significant potential to mislead economists.

The figures cited in Table 12 are for the mean discrepancy over the 68 month-on-month comparisons for the period January 1995 (=100) to September 2001. The standard deviation of each discrepancy for each group was calculated for the 68 comparisons over time to quantify the volatility of the discrepancies. The minimum, maximum, mean, and median of these standard deviations are given in Table 12, along with their standard deviation across product groups. It is apparent that there is substantial volatility in the discrepancy over time. The average dispersion is high: for each country the mean standard deviation over the groups exceeds the mean of the groups. The very lowest dispersion over time of the month-on-month discrepancy for a 3-digit group is for Finland at 1.9 percent allowing an approximately 95 percent plus or minus range of 2 x 1.9=3.8 percentage points around the mean discrepancy, and this is the lowest dispersion.

The results of Table 12 clearly show that some product groups have a lower discrepancy, but they also allow us to conclude, as does Decoster (2003a and 2003b), that for no product group is the average monthly discrepancy acceptable. Decoster (2003a) analyzed the data in considerably more detail. He found that PIs are more stable over time than UVIs, and that UVIs often display erratic behavior that PIs do not; he therefore concluded in the Final Report (2003b, p.9) that:

> "Any list of CPA categories for which UVIs are a priori acceptable as proxies for SPIs [import price indices] would be very short, especially as regards monthly data. It would include almost only aggregates and raw materials, even if sizable discrepancies between UVIs and SPIs are deemed acceptable. Apparently, any list of product categories for which short-term UVIs are acceptable proxies for SPIs seems country specific. For a few low-tech products, for which quality changes are slow, UVI changes over the long term (several years) may be acceptable proxies for SPIs."

B. Use A More Detailed Stratification Of Unit Values

A second possibility is to improve UVIs by more detailed stratification of the customs data. United Nations (1981) emphasized the need to stratify unit values to the (limited) extent possible and drew attention to doing so where possible by country of destination and size of batch, though see Párniczky (1974) on the limitations to this. Stratification is also possible for shipments by/to (major) establishments to/from given countries. The author is unaware of studies as to how UVIs derived from poorly disaggregated customs data stand up to UVIs derived from highly disaggregated customs data. Indeed, the absence of highly detailed criteria by which to stratify unit values precludes any benchmark as to what is a reliable UVI. However, such experiments can be undertaken for consumer goods using highly detailed bar-code scanner

data. Bradley (2006) examined the issue in some detail and found that even for detailed data of sales of cereal in 169 selected stores by 1,369 brands, aggregating unit values that distinguish a brand of Tuna according to the week of purchase and store in which it is sold, as against simply aggregating unit values for the self-same brand and item, leads to substantial differences in the results. Silver and Webb (2002) took (brand and) model numbers for washing machines, dishwashers, vacuum cleaners, television sets, cameras and personal computers (PCs) and compared unit value changes for the same models over different store types, finding quite different result when aggregating with and without store type as a variable. Haan and Opperdoes (1999) undertook a similar study on coffee, further apportioning their data according to the week of the month the data relates to, again finding unit value bias. Given such bias at this fine level of detail for aggregating identical items it is hard to imagine disaggregated unit values based on customs returns being robust to unit value bias.

C. Use Other Country Data or Global Product Price Indices

An alternative to using UVIs is to use corresponding series from other countries, for example an export price index of personal computers from the U.S. to proxy an import price index, or global commodity price indices to proxy exports or imports. The assumption is that there is a global market in which countries are price takers with little to no price discrimination between countries. In advocating stratification by country of origin/destination United Nations (1981) implicitly argued against this as a general strategy. However, there may well be product areas for which this is useful. It will not of course be a panacea for the measurement of trade price indices.

D. Different Formulas

PIs and UVIs are compiled in two stages. The first stage is the price relative (PIs) or unit value change (UVIs) at the elementary level of aggregation, to form elementary indices. The second stage is the weighted aggregation of these elementary indices. PIs and UVIs may be compiled using different formulas at this second stage, so differences in the results may in part be due to formula differences. Data were not available to recompile the indices to identify the effect of such formulas differences. Some insights are available for Germany.[9] Germany is in the fortunate position of having import price indices, import deflators of the national accounts and unit value indices.[10] The import price indices are of the Laspeyres type and refer to the year 2000. The Laspeyres principle is applied, however, only to the basket of goods, but not to the

---

[9] Unit value indices are compiled in Japan using a Fisher index and the trade price indices (from their Corporate Goods Price Index) using a chained Laspeyres index, though there is some lag in the adoption of the most recent period's weights.

[10] Data and this account on Germany is from private correspondence with Johannes Hoffmann and Hans-Albert Leifer, Deutsche Bundesbank, December, 2006.

countries of origin, meaning that any shifts to low cost producers will be captured by the import price index. The national accounts deflator are annually chained Paasche indices, and the unit value indices are Paasche indices referring to the year 2000. The product specific price indices used for the compilation of the national accounts deflators are taken from the price statistics. Hence, the main difference between the import price index and the import deflator is to be found in the index formula.

In the years 2000 to 2005, the UVI displayed a decline of 1.8% pa, whereas the import price index increased slightly and the import deflator decreased less strongly (+0.3 and -0.8, respectively). Taking the geometric average of the change in the import price deflator and the import price index gives an estimate of -0.2% as the "true" annual change in import prices, implying that the German UVI is significantly distorted downwards.

Bear in mind we are comparing Paasche unit value indices with Laspeyres PIs. Von der Lippe (2007) demonstrates how the components of such formula discrepancy may cancel and any differences would be the result of unit value bias.

E. Lack of Customs Data and Unit Value Indices within Monetary Unions and, in General, for Services

There remains a potential problem of customs data itself becoming unsuitable for measuring trade flows for some countries. Countries with customs/monetary unions may abandon or limit the requirements on trade within the union to be documented. Furthermore, with services and e-commerce making up an increasing share of trade, customs data on merchandise trade will be unsuitable as the sole data source. Establishment-based sources for external trade price data have become the only practical option in these cases (though trade within customs unions may well be measured as a by-product of administering, for example, value added taxes).

F. Use Deletion Routines for Unusual Price Changes

Of widespread use in the compilation of UVIs are deletion routines. This is because much of the data from customs records on unit value changes are extreme outliers and have to be discarded. Some of this arises from absent or poor quantity data. In other cases it is due to unit value bias. Alterman (1991) estimated that the U.S. UVIs produced in 1985 were calculated for only 56 percent of the value of imports and 46 percent of the value of exports. For capital goods the respective figures for imports and exports fell to 30.3 and 26.1 percent. The problem of such deletions are two-fold. First, the implicit effect on the sample representativity and coverage. PIs are based on selected items from selected establishments with the purpose in mind that they are representative. Second, is that the deletion removes signal as opposed to noise. There is much evidence in CPI compilation, for example Hoffmann and Kurz-Kim (2006), that price changes can be substantial, and irregular, with long periods of constant prices followed by relatively large catch-up price changes. These large price increases may be deleted by outlier detection routines, resulting in UVIs that are unduly stable.

G. The Resource Constraint

A main reason why countries do not compile PIs is the cost of doing so. United Nations (1981) recognized the superiority of PIs by recommending well-endowed countries compile them, while advising countries with limited resources to compile UVIs. Countries not only require price indices for trade flows, but for the deflation of output, intermediate and final consumption goods and services by resident units. In particular, an output producer price index (PPI) is required that measures the changes in the prices of output of resident establishments. PPIs are compiled by selecting representative items from major/representative establishments and comparing the prices of like with like over time. Such output covers the domestic and export market (ILO *et al.*, 2004b). For a self-standing export price index there would be a need to identify price changes from such establishments for foreign markets as well as overall output and, as necessary, expand the sample size to ensure those establishments serving foreign markets are included in a representative manner. In some instances specialist import/export wholesalers may be an efficient contact. Poorer countries have fewer establishments serving foreign markets with large proportions of exports usually being the responsibility of a relatively small number of establishments. Similar arguments apply to imports. Establishment-based trade price indices are but an extension of establishment-based price surveys for producer prices. There are resource costs to both PPIs and, by extension, to trade PIs. But they have their benefits which are the proper measurement of the major economic flows affecting the country, to allow for appropriate policy responses when necessary.

VI. CONCLUSIONS

There has been a long-held view that unit value indices based on customs data can seriously misrepresent price changes as measured by price indices. The evidence in Section IV of this paper supports that view: UVIs were found to seriously mislead in the sense that discrepancies between UVIs and PIs were substantial; changes could not be relied upon to have the same sign; there was no evidence of long-run (cointegrating) relationships between PIs and UVIs; and UVIs were of little help in predicting PIs. The findings held both for month-on-month and month-on-12 month changes. The marked unreliability of UVIs as measures of export and import price inflation was surpassed by the unreliability of the terms of trade indices based upon them. Terms of trade indices based on UVIs failed with regard to the substantial magnitude of the discrepancy with PI-based ones, the wrong sign, absence of long-run relationship, and poor predictive value.

The results from using UVIs to measure the terms of trade effect, as part of a measure of real national income, and to deflate import and export current period values to derive volume measure, were seriously misleading when compared to those from using price indices.

We reiterate the caveat to these findings at the start of this paper. The evidence presented here is limited to two countries that compile both price indices and unit value indices, though other studies have similar conclusions. It is also limited by the fact that the deficiencies in unit value indices are not measured against a perfect benchmark, the price indices themselves having deficiencies. Yet, as outlined above, unit value indices suffer mainly from not comparing the prices of like with like, while establishment–based price indices do so. Furthermore, the

coverage of PIs is by design representative, while the coverage of UVIs results from a substantial discarding of outliers. This and other studies asked how well unit value indices stand up against price indices designed to overcome their major failings and the answer is that that they do not.

We then turned to the question of what can be done? The answer is to commence as soon as possible a program of establishment-based survey price collection. In Section VA we demonstrated that a strategy of hybrid indices, using UVIs where feasible, was unproductive, and was, indeed, misleading. In Section VB we argued that customs data was by its nature limited to the extent that it could benefit from further stratification. On a positive note we stated in Section VC that other country indices or global product prices may play a useful role, but this was not a panacea. The fact that our comparisons between UVIs and PIs were not pure was reiterated in Section VD. It was argued in Section VE that customs and monetary unions and the increasing role of services in world trade give rise to further cause of concern over a reliance on customs data. The main advantage of customs data had been argued to be their superior coverage of transactions and relatively low resource cost. In Section VF we argued that the extent of deletions gives rise to concern over the representativity of UVIs and potential bias in deletion of some of the signal. As regards the resource constraint, in Section VG, the development of establishment-based surveys was identified as a natural part of the development of a system of price indices, with a smaller resource demand on countries with less developed import and export markets. Indeed it seems apparent that a disservice is being done to countries by advocating the cheaper alternative of unit value indices.

Figures and Tables

Figure 1, Export unit value and price indices for Germany

Germany UVI — Germany PI

Figure 2, Export unit value and price indices for Japan

Japan UVI — Japan PI

Figure 3, Import unit value and price indices for Germany

Germany PI — Germany UVI

Figure 4, Import unit value and price indices for Japan
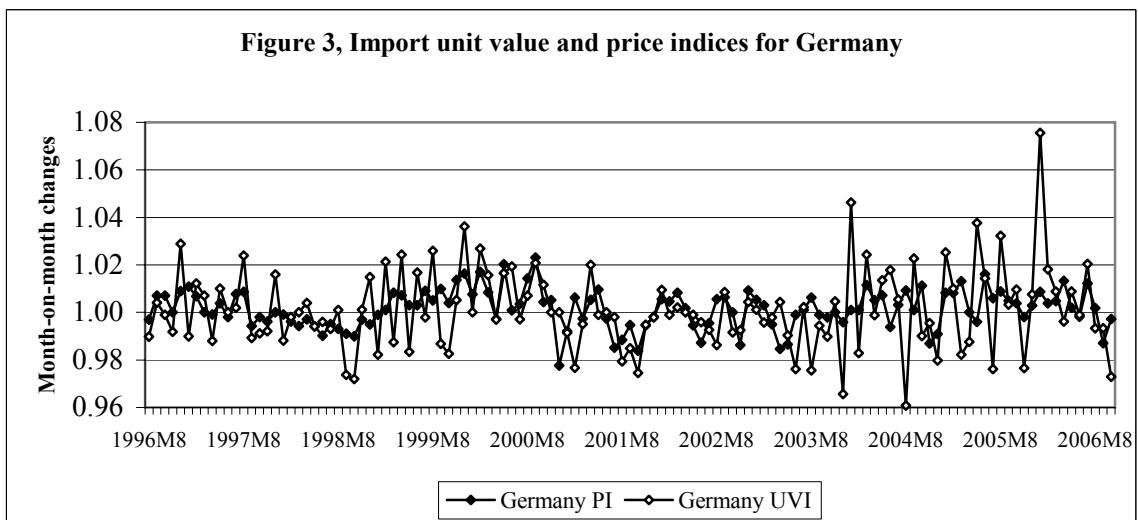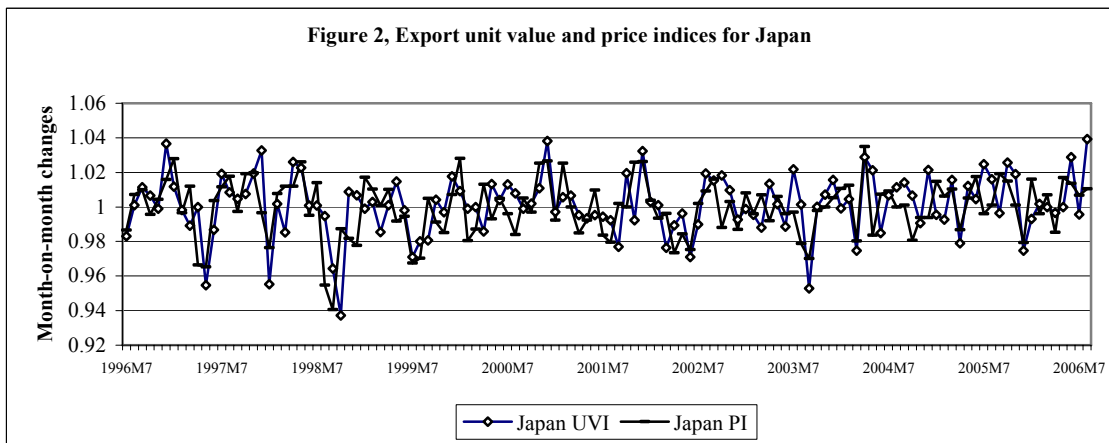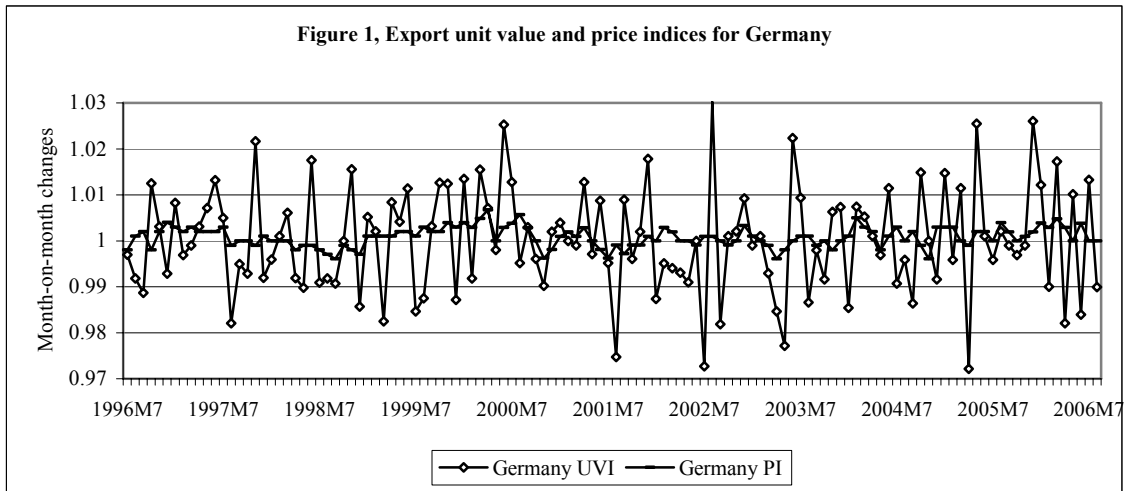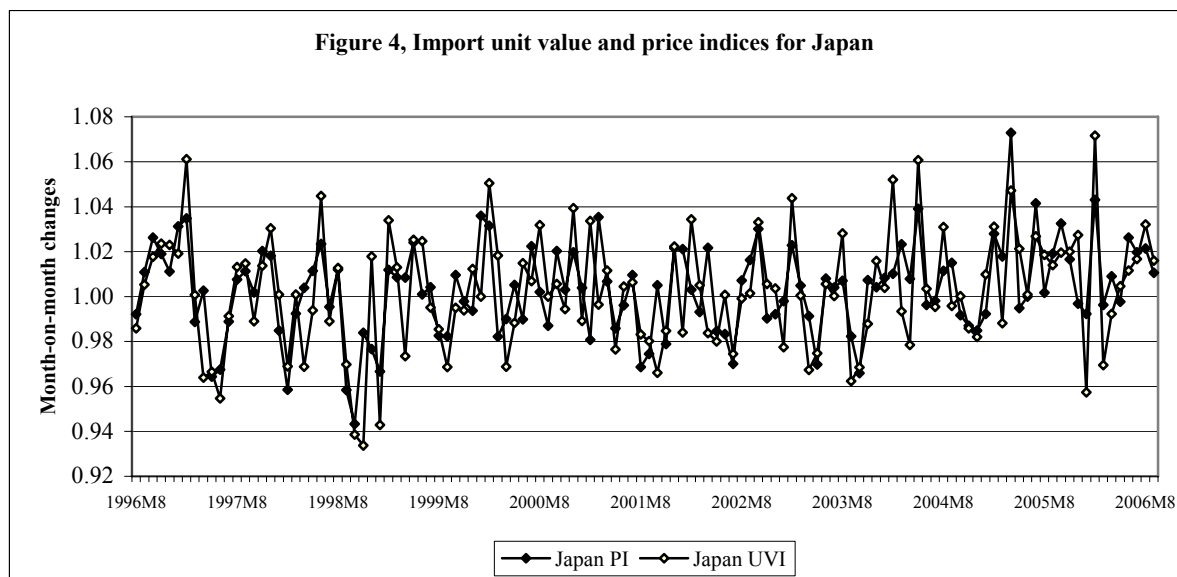
Table 1, Average discrepancy between import unit value and price indices

	Germany Month-on-month		Japan Month-on-month		Germany Month-on-12 month		Japan Month-on-12 mo	
	Import	Export	Import	Export	Import	Export	Import	Export
Absolute value of ratio of UVI to PI:[†]								
Mean	0.011	0.009	0.015	0.013	0.018	0.016	0.024	0.025
Standard deviation	0.010	0.007	0.015	0.010	0.016	0.012	0.019	0.019
Maximum	0.073	0.045	0.060	0.055	0.096	0.068	0.089	0.099
Root mean squared error	0.014	0.011	0.018	0.016	0.024	0.020	0.032	0.031
Mean absolute deviation	0.011	0.009	0.015	0.013	0.018	0.016	0.025	0.025

†Summary measures of the absolute value of the discrepancy [(UV/PI)-1]

Table 2, UVI and PI: percentage of changes of same and different sign

	Month-on-month				Month-on-12 month			
	Germany		Japan		Germany		Japan	
	Exports	Imports	Exports	Imports	Exports	Imports	Exports	Imports
Same sign								
both positive	41.7	39.9	36.2	41.8	56.4	49.1	39.1	45.7
both negative	16.6	32.1	33.0	35.2	19.4	37.4	45.9	50.0
Different signs								
UV positive : PI negative	12.5	11.1	15.7	11.2	0.6	4.3	13.0	0.6
UV negative : PI positive	29.2	16.8	15.1	11.8	23.7	9.1	2.0	3.7
Total[†]	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

† Total excludes cases where either change is exactly unity.

Table 3, Unit root tests and coointegrating relationships between month-on-month and 12-month percentage changes

	Germany Import		Germany Export		Japan Import		Japan Export	
	Unit value	Price index	Unit value	Price index	Unit value	Price index	Unit value	Price index
Month-on-month								
Unit root tests:								
Weighted symmetric	-4.422	-7.421	-4.248	-4.194	-6.389	-6.351	-11.237	-9.693
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Dickey-Fuller	-4.295	-5.937	-4.107	-4.159	-6.284	-6.247	-7.361	-9.621
p-value	*0.00*	*0.00*	*0.01*	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*
Phillips-Peron	-574.441	-176.442	-520.963	-271.929	-267.931	-171.440	-295.361	-252.684
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Month-on-12 month								
Unit root tests:								
Weighted symmetric	-3.210	-3.531	-3.300	-3.006	-3.163	-3.358	-3.647	-3.785
p-value	*0.05*	*0.02*	*0.04*	*0.08*	*0.05*	*0.03*	*0.01*	*0.01*
Dickey-Fuller	-3.070	-3.445	-3.171	-3.003	-3.127	-3.273	-3.489	-3.704
p-value	*0.11*	*0.05*	*0.09*	*0.13*	*0.10*	*0.07*	*0.04*	*0.02*
Phillips-Peron	-24.130	-22.523	-34.532	-20.800	-26.205	-24.882	-32.997	-28.599
p-value	*0.03*	*0.04*	*0.00*	*0.06*	*0.02*	*0.03*	*0.00*	*0.01*
p-value	*0.08*	*0.07*	*0.01*	*0.04*	*0.03*	*0.07*	*0.01*	*0.00*
Index series								
Unit root tests:								
Weighted symmetric	-2.147	-2.377	-1.075	-0.334	-2.353	-2.048	-1.940	-2.383
p-value	*0.54*	*0.37*	*0.97*	*1.00*	*0.39*	*0.61*	*0.69*	*0.37*
Dickey-Fuller	-2.884	-2.738	-2.465	-2.597	-2.272	-2.092	0.686	0.389
p-value	*0.17*	*0.22*	*0.35*	*0.28*	*0.45*	*0.55*	*0.69*	*0.39*
Phillips-Peron	-9.531	-7.347	-5.227	-4.732	-6.735	-4.776	0.521	0.474
p-value	*0.47*	*0.64*	*0.80*	*0.84*	*0.69*	*0.84*	*0.52*	*0.47*
E-G cointegration test	-2.464	-2.149	-3.311	-3.384	-3.306	-3.141	-1.690	-2.265

Table 4, Predictive ability of UVIs in relation to PIs

	Germany Month-on-month		Japan Month-on-month		Germany Month-on-12 month		Japan Month-on-12 month	
	Import	Export	Import	Export	Import	Export	Import	Export
Prediction:								
F-statistic	5.911	7.549	6.263	1.475	503.024	248.851	1149.070	220.214
p-value	*0.00*	*0.00*	*0.00*	*0.13*	*0.00*	*0.00*	*0.00*	*0.00*
Std err of reg	0.009	0.003	0.023	0.016	0.025	0.011	0.046	0.036
Mean of PI change	0.001	0.001	0.000	-0.001	0.015	0.016	0.012	-0.011
Granger causality:								
F-statistic UVI GC PI	0.339	2.491	2.571	0.781	0.405	0.57341	13.486	3.160
p-value	*0.98*	*0.00*	*0.00*	*0.67*	*0.81*	*0.68*	*0.00*	*0.01*
F-statistic PI GC UVI	9.133	5.969	17.505	18.258	27.304	19.824	57.566	62.057
p-value	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*

Table 5, Terms of trade indices: discrepancy between UVIs and PIs

	Month-on-month		Month-on-12 month	
	Germany	**Japan**	**Germany**	**Japan**
Absolute value of ratio				
of UVI to PI:				
Mean	0.013	0.016	0.016	0.037
Standard deviation	0.010	0.013	0.014	0.027
Maximum	0.070	0.062	0.100	0.183
Root mean squared error	0.016	0.020	0.021	0.046

Table 6, Terms of trade: Percentage of changes of same and different sign

	Month-on-month		Month-on-12 month	
	Germany	**Japan**	**Germany**	**Japan**
Same sign				
both PI and UVI positive	32.1	29.5	40.6	38.9
both PI and UVI negative	30.7	34.0	40.6	39.2
Different signs				
UV positive : PI negative	18.8	22.2	4.8	16.9
UV negative : PI positive	18.3	14.3	14.1	5.1
Total[†]	100.0	100.0	100.0	100.0

† Total excludes cases where either change is exactly unity.

Table 7, Terms of trade indices: unit root and cointegration tests

	Month-on-month Germany Terms of trade		Month-on-12 month Germany Terms of trade		Month-on-month Japan Terms of trade		Month-on-12 month Japan Terms of trade	
	Unit value	Price index	Unit value	Price index	Unit value	Price index	Unit value	Price index
Weighted symmetric	-4.685	-7.793	-3.250	-3.600	-5.178	-6.295	-3.653	-3.404
p-value	*0.00*	*0.00*	*0.04*	*0.02*	*0.00*	*0.00*	*0.01*	*0.03*
Dickey-Fuller	-4.562	-6.309	-3.074	-3.482	-5.069	-6.198	-3.962	-3.261
p-value	*0.00*	*0.00*	*0.11*	*0.04*	*0.00*	*0.00*	*0.01*	*0.07*
Phillips-Peron	-477.879	-197.403	-36.274	-24.510	-467.081	-131.064	-29.482	-27.380
p-value	*0.00*	*0.00*	*0.00*	*0.03*	*0.00*	*0.00*	*0.01*	*0.02*
E-G coint. test	-4.842	-4.507	-3.471	-3.348	-5.100	-7.651	-4.393	-3.809
p-value	*0.00*	*0.01*	*0.11*	*0.14*	*0.00*	*0.00*	*0.01*	*0.00*

Table 8, Terms of trade: predictive ability of UVI in relation to PI

	Month-on-month		Month-on-12 month	
	Germany	**Japan**	**Germany**	**Japan**
Prediction:				
F-statistic	2.988	9.495	338.566	620.952
p-value	*0.00*	*0.00*	*0.00*	*0.00*
Std err of reg	0.007	0.015	0.022	0.040
Mean of PI	0.000	-0.001	0.003	-0.006
Granger causality:				
F-statistic UVI GC PI	0.517	2.051	0.992	3.545
p-value	*0.90*	*0.02*	*0.41*	*0.01*
F-statistic PI GC UVI	10.330	9.041	28.310	26.926
p-value	*0.00*	*0.00*	*0.00*	*0.00*

Table 9, Terms of trade effect: previous year's prices

	Germany (Euros bill.) Terms of trade effect			Japan (Yen bill.) Terms of trade effect		
	Unit value indices	Price indices	Trade balance[†]	Unit value indices	Price indices	Trade balance[†]
1999	2.1	-0.2	17	2,663.6	-518.9	7,893
2000	-43.7	-46.0	7	-2,866.6	-5,174.5	7,316
2001	13.7	2.5	43	3,743.1	289.7	3,174
2002	19.4	18.9	98	-2,016.8	183.2	6,412
2003	20.8	14.8	86	960.1	-1,998.8	7,975
2004	0.2	-4.5	111	-1,420.9	-3,548.3	9,626
2005	-16.4	-26.5	116	-3,614.7	-7,001.2	6,956

† National accounts estimates of exports minus imports of goods and services

Table 10, Comparison of deflated exports and imports by UVIs and PIs

	Germany			Japan		
	Billion Euros at constant 1999 prices		Percentage	Billion Euros at constant 1999 prices		Percentage
	1999	2005	change	1999	2005	change
Exports						
Unit values	591.5	799.0	35.1	51144.0	65009.2	27.1
Price indices	591.5	857.9	45.0	51144.0	76741.8	50.1
Imports						
Unit values	796.3	785.5	-1.4	64957.0	54353.9	-16.3
Price indices	796.3	705.2	-11.4	64957.0	52574.6	-19.1

Table 11, CPA 4-digit classes in percentile with the least discrepancy for month-on-month UVIs and PIs

		Month-on-month discrepancy: UVI and PI			
		Mean	Maximum	Minimum	Standard deviation
Manufacture of motor vehicles	CPA-3410	0.0197	0.0601	0.0009	0.0153
Manufacture of pulp	CPA-2111	0.0197	0.0664	0.0003	0.0148
Manufacture of motor vehicles parts and accessories	CPA-3430	0.0260	0.1266	0.0001	0.0212
Manufacture of plastic plates, sheets, tubes and profiles	CPA-2521	0.0278	0.0920	0.0000	0.0202
Manufacture of plastics in primary forms	CPA-2416	0.0282	0.0898	0.0001	0.0209
Manufacture of paper and paper board	CPA-2112	0.0296	0.0916	0.0000	0.0217
Manufacture of other fabricated metal products n.e.c†	CPA-2875	0.0325	0.1289	0.0005	0.0257
Manufacture of plastic packing goods	CPA-2522	0.0344	0.1139	0.0001	0.0245
Processing and preserving of fruit and vegetables n.e.c†	CPA-1533	0.0350	0.1281	0.0004	0.0280
Manufacture of made up textile articles, except apparel	CPA-1740	0.0368	0.1048	0.0012	0.0269
Manufacture of basic iron and steel and ferrous-alloys (ECSC)*	CPA-2710	0.0370	0.1591	0.0002	0.0312
Manufacture of rubber tyres and tubes	CPA-2511	0.0371	0.1490	0.0000	0.0305
Manufacture of other plastic products	CPA-2524	0.0380	0.1193	0.0021	0.0282
Manufacture of underwear	CPA-1823	0.0387	0.1115	0.0009	0.0284
Manufacture of fertizers and nitrogen compounds	CPA-2415	0.0395	0.1524	0.0001	0.0330

† Not elsewhere classified; * European Coal and Steel Community.

Table 12, Average discrepancy between import unit value and price indices: absolute values of month-on-month percentage changes

Absolute values of ratios of UVIs toPIs: abs(UVm,m-1/PIm,m-1)-1	Finland 3-digit group	Netherlands 3-digit group	Sweden 3-digit group
0 to under 0.025	13	13	23
0.025 to under 0.05	32	52	31
0.05 to under 0.075	19	17	7
0.075 top under 0.1	8	10	3
0.1 to under 0.15	3	3	2
0.15 to under 0.2	1	3	0
0.0 to under 0.3	1	2	1
	77	100	67
Mean	0.053	0.054	0.041
Standard deviation	0.034	0.041	0.030
Minimum	0.015	0.016	0.019
Maximum	0.203	0.247	0.213
Standard deviation of UV/PI ratio			
Mean	0.069	0.073	0.059
Standard deviation	0.046	0.058	0.075
Minimum	0.019	0.022	0.024
Maximum	0.285	0.323	0.594

References

Alterman, William, 1991, Price Trends in U.S. Trade: New Data, New Insights in *International Economic Transactions*, ed. by Peter Hooper and J. David Richardson, pp. 109-143, (Chicago: University of Chicago Press).

Balk, B. M., 1998, On the Use of Unit Value Indices as Consumer Price Subindices, in *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, pp. 112-120, (Washington DC: U.S. Bureau of Labor Statistics). Available at: http://www.ottawagroup.org.

Bradley, Ralph, 2005, "Pitfalls of Using Unit Values as a Price Measure or Price Index," *Journal of Economic and Social Measurement*, Vol. 30, pp. 39-61.

Commission of the European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank, 1993, *System of National Accounts 1993* (Brussels/Luxembourg, New York, Paris, Washington: EC, IMF, OECD, UN, World Bank).

Decoster, Renaud, 2003a, *Import and Export Price Indices: A Comparison of Import Unit Value Indices and Import Specific Price Indices for Four EU Member States*, *Second Report*, Study Performed for Eurostat Unit C-4: Methodology, Nomenclature and Statistics of External and Intra-Community Trade (Paris: PLANISTAT Europe).

Decoster, Renaud, 2003b, *Import and Export Price Indices: Final Report*; Study Performed for Eurostat Unit 0074 C-4: Methodology, Nomenclature and Statistics of External and Intra-Community Trade (Paris: PLANISTAT Europe).

González-Páramo, José Manuel, 2006, "Statistical Challenges Raised by Globalization," European Central Bank (ECB) (24 October), (Frankfurt: ECB).

Eurostat, IMF, OECD, United Nations, and World Bank, 1993, *System of National Accounts 1993,* (Brussels/Luxembourg, New York, Paris, Washington: Eurostat, IMF, OECD, United Nations, and World Bank).

Haan De, Jan. and Eddy Opperdoes, (1999), Estimation of the Coffee Price Index Using Scanner Data: Simulation of Official Practices in *Proceedings of the Third International Conference on Price Indices*, ed. by B.M. Balk, (Voorburg: Statistics Netherlands). Revised version in J. De Haan, 1999, "*Empirical Studies on Consumer Price Index Construction,*" pp. 59-74 (Voorburg: Statistics Netherlands).

Hoffmann, Johannes and Jeong-Ryeol Kurz-Kim, 2006, "Consumer Price Adjustment Under the Microscope: Germany in a Period of Low Inflation," *Deutsche Bundesbank, Discussion Paper Series* 1: *Economic Studies*, No. 16/2006.

International Labour Office (ILO), IMF, OECD, Eurostat, United Nations, World Bank, 2004a, *Consumer Price Index Manual: Theory and Practice*, (Geneva: ILO). http://www.ilo.org/public/english/bureau/stat/guides/cpi/index.htm.

International Labour Office (ILO), IMF, OECD, UN ECE, World Bank, 2004b, *Producer Price Index Manual: Theory and Practice* (Washington: International Monetary Fund). http://www.imf.org/np/sta/tegppi/index.htm.

Kravis, Irving B. and Robert E. Lipsey (1971), *Price Competitiveness in World Trade*, Studies in International Economic Relations, Vol. 6 (New York: National Bureau of Economic Research).

Kravis, Irving B. and Robert E. Lipsey, 1985, "Prices and Terms of Trade for Developed Country Exports of Manufactured Goods," *National Bureau of Economic Research* (*NBER*) *Working Papers*: No. 0774.

Von der Lippe, Peter, 2007, " Price Indices and Unit Value Indices in German Foreign Trade Statistics," Unpublished Paper, University of Duisburg-Essen.

Párniczky, G., 1974, "Some Problems of Price Measurement in External Trade Statistics," *Acta Oeconomica*, Vol. 12, No. 2, pp. 229-240.

Silver, Mick, 2006, Errors and Biases in Export and Import Price Indices, Chapter 11 in *Draft Export and Import Price Index Manual*, (International Monetary Fund, Washington) http://www.imf.org/external/np/sta/tegeipi/index.htm.

Silver, Mick and Khashi Mahdavy, 1989, "The Measurement of a Nation's Terms of Trade Effect and Real National Disposable Income Within a National Accounting Framework," *Journal of the Royal Statistical Society*, Series A, Vol. 152, Part 1 (March). Reprinted in *The IMF's Statistical Systems in Context of the Revision of the United Nations' A System of National Accounts* ed. by Vicente Galbis (Washington DC, International Monetary Fund, 1991).

Silver, Mick and Bruce Webb, 2002, "The Measurement of Inflation: Aggregation at the Basic Level," *Journal of Economic and Social Measurement*, Vol. 28, No. 1-2, pp. 21-36.

Trewin, Dennis, 2006, Australian Bureau of Statistics Information Paper, *Producer and International Trade Price Indexes: Concepts, Sources and Methods* (Canberra, Australian Bureau of Statistics).

United Nations, Department of International Economic and Social Affairs, *Strategies for Price and Quantity Measurement in External Trad*e, Statistical Papers, Series M, No. 69 (United Nations, 1981).

United Nations, Department of International Economic and Social Affairs,, *Manual on Producers' Price Indices for Industrial Goods*, Statistical Papers, Series M, No. 66 (United Nations, 1979).

United Nations, Department of International Economic and Social Affairs, *Price and Quantity Measurement in External Trade: Two Studies of National Practice*, Statistical Papers, Series M, No. 76 (United Nations, 1983).

Derivation of test of long-run differences between UVIs and PIs and TOT indices[11]

We have the two prices series, p^t and P^t, for $t = 0,1,...,T$ with both price series starting at unity in period 0:

(1) $p^0 = P^0 \equiv 1$.

If one series grows at different rates compared to the other series, then we have the following model:

$$(2)\ p^t = \alpha_1\alpha_2...\alpha_t\,P^t\,; \hspace{4cm} t = 1,...,T$$

where α_t is one plus the differential rate of growth between the two series going from period $t-1$ to period t. Assumptions (1) and (2) imply that

$$(3)\ p^t/p^{t-1} = \alpha_t\,(P^t/P^{t-1})\,; \hspace{4cm} t = 1,...,T.$$

Define the logarithmic rates of growth of both series as follows:

$$(4)\ y^t \equiv \ln(p^t/p^{t-1})\,;\ Y^t \equiv \ln(P^t/P^{t-1})\,; \hspace{3cm} t = 1,...,T.$$

Now take logarithms of both sides of (3), use definitions (4) and rearrange terms slightly in order to obtain the following equations:

$$(5)\ Z^t \equiv y^t - Y^t = \beta_t\,; \hspace{4cm} t = 1,...,T$$

where β_t is defined as the logarithm of α_t:

$$(6)\ \beta_t \equiv \ln\alpha_t\,; \hspace{4cm} t = 1,...,T.$$

The exact equations (5) can be converted into the following simple stochastic model:

$$(7)\ Z^t = \beta + \varepsilon^t\,; \hspace{4cm} t = 1,...,T$$

where the ε^t are independently distributed normal variables with mean 0 and constant variance. The least squares and maximum likelihood estimator for β is of course the average of the Z^t which in turn is equal to the average of the logarithmic growth rates of the p^t series, the average of the y^t, less the average of the logarithmic growth rates of the P^t series, the average of the Y^t.

---

[11] I am grateful to Erwin Diewert for this derivation.

The hypothesis of interest is:

(8) $\beta = 0$

which in turn corresponds to the hypothesis

(9) $\alpha = 1$.

If the hypothesis (8) or (9) is accepted, then the two measures of price change give the same answer in the long run.

The above setup also simplifies testing for no changes in the terms of trade, using a traditional definition for the terms of trade as the export price index divided by the import price index. Thus let the unit value measures of export and import prices in period t be P_X^t and P_M^t respectively and let the price index measures of export and import prices in period t be p_X^t and p_M^t respectively. Repeat the above algebra for export prices and compare the rate of growth of p_X^t with the rate of growth of P_X^t and obtain the counterpart to (7), the estimator for the difference in the logarithmic rates of growth, say β_X^*. Similarly, repeat the above algebra for import prices and compare the rate of growth of p_X^t with the rate of growth of P_X^t and obtain the estimator for the difference in the logarithmic rates of growth, say β_M^*. The two series for the logarithmic rates of growth of the terms of trade are defined as follows:

(10) $y_{TT}^t \equiv \ln[(p_X^t/p_M^t)/(p_X^{t-1}/p_M^{t-1})]$; $Y_{TT}^t \equiv \ln[(P_X^t/P_M^t)/(P_X^{t-1}/P_M^{t-1})]$; $\qquad t = 1,...,T.$

The terms of trade regression counterpart to (7) is

(11) $Z_{TT}^t \equiv y_{TT}^t - Y_{TT}^t = \beta_{TT} + \varepsilon_{TT}^t$; $\qquad t = 1,...,T.$

The two sets of price indexes will give the same answer with respect to measuring changes in the terms of trade if β_{TT} equals 0. The least squares or maximum likelihood estimator for β_{TT} is the average of the y_{TT}^t less the average of the Y_{TT}^t. It can also be shown that this estimator, β_{TT}^*, is equal to β_X^* less β_M^* where β_X^* and β_M^* are the estimators for β_X and β_M defined above (10).

All of the above suggested tests are simple and consistent with one another.