Disaggregated international trade prices

Dawit Sisay Temere*
Modelgroup, Statistics Denmark, mail: *dsi@dst.dk*

Summary: This paper presents new annual price indices for imports and exports of OECD countries classified by SITC system. The paper uses detailed SITC data from the OECD international trade by commodity statistics database. The price indices at 1-digit SITC are constructed by aggregating unit values at the most detailed SITC level. Before aggregation, the unit values at the detailed level are filtered for outliers using the Hidiroglou-Berthelot method. The final price indices depend on the choice of index formula: Paasche, Laspeyres or Fisher chain index. The Paasche and Laspeyres chain price indices have downward and upward biases, respectively. There is a lot of volatility in the detailed unit values which makes it natural to prefer the Fisher index that is a geometric mean of the Paasche and Laspeyres indices.

Keywords: Unit value; trade price; international trade

JEL: C43; F10

1. Introduction

The OECD *International Trade and Competitiveness Indicators* (ITCI) database provides prices and volume of imports and exports for member countries. Trade flows excludeing services are grouped into four SITC components: agricultures, raw materials, energy and manufactures, cf. Durand, Madaschi and Terribile (1998). Unfortunately, the OECD ceased publishing these data in 2002. Since then only aggregated data for total goods is available. A disaggregated dataset can provide a more informative picture since the response of imports and exports to price changes is different in the different SITC components. Hence, detailed trade data improves the

* The author has benefited greatly from the comments and conversations with staffs of the economic modeling unit at Statistics Denmark. All errors are the author's own responsibility.

scope for esti-mating trade price elasticities that are often crucial in economic analysis. For example, the lack of disaggregated trade prices has been a challenge for Danish model builders, and the new data presented in this paper has already been applied to create export market variables for the macro models – ADAM (Annual Danish Aggregate Model) and SMEC (Simulation Model of the Economic Council).

This paper presents new annual price indices for disaggregated imports and exports of OECD countries based on unit values. The raw data is obtained from the OECD International Trade by Commodity Statistics (ITCS) database and contains both values and quantities on a detailed SITC level. Unit values defined as annual value divided by annual quantity are used to construct price indices. The unit values, at the most detailed SITC level, are first filtered for outliers using the Hidiroglou-Berthelot (1986) method.

Trade prices approximated by unit values have well-known shortcomings. Bias can arise from compositional changes in the product mix and it is often difficult or impossible to identify quality changes and distinguish these from price changes. Unit value indices rely to a larger extent on outlier detection and deletion, given price stickiness such deletion run the risk of missing price catch-ups and risks understating inflation, cf. Silver (2007). By using the most detailed unit values at 5-digit SITC level some of the shortcomings can be minimized. Furthermore, by constructing the trade prices yourself you may reduce bias from methodological differences and facilitate cross-country comparison, Gaulier et al. (2008).

Comparison of the constructed unit-value-based price indices with the official ITCI-database and with the OECD national accounts statistics shows a great deal of similarity between the constructed prices and official data. The sensitivity analysis with alternative filtering techniques, such as winsorizing, presents no strong argument against the Hidiroglou-Berthelot method for removing outliers. The rest of the paper is organized as follows. Section 2 describes the dataset, section 3 outlines the methodology for filtering outliers, section 4 reviews index formulas widely used in statistical institutes, section 5 presents price indices for selected OECD countries, section 6 presents sensitivity analysis, and section 7 concludes.

2. The raw data

The OECD *International Trade by Commodity Statistics* provide detailed quantity and value trade data for OECD countries. Data is available annually up to the most detailed 5-digit SITC level. Trade values are reported in current US dollar and quantities are reported in different units of measurement, e.g. area in square meters, electrical energy in kilowatt-hours, volume in cubic meters, etc. This paper uses SITC revision-3 which is available from 1989 onwards.

The raw data presents a number of challenges. Quantities are reported in different units of measurement. Missing values at the detailed level create problems at each step of the data construction, and outliers can introduce bias in the final trade prices. Different countries have different reporting thresholds and not all countries have data for the full sample 1989-2012. It is important to address these potential problems before constructing price indices.

Firstly, the focus is on member countries' import/export vis-à-vis the rest of the world, and consequently bilateral trade flows are dropped from the database. Secondly, totals, subtotals and zero trade flows are removed from the database to make sure price indices are built from the most detailed SITC level. Different reporting thresholds can introduce heterogeneity in the dataset. To deal with this, a reporting threshold of 10000 US dollar is imposed on trade flows at 5-digit level. Whenever the threshold is not met, the observation is dropped. Observations with missing quantity or missing dollar value are also dropped for the price computations. By using the growth rates of unit values instead of levels, the problem of different units of measurement can be minimized. An alternative is to start by converting different units into a single unit, which requires a great deal of time and more information on quantities than found in the ITCS-database. Measurement errors are likely to be correlated between reported values and quantities, so that working directly on unit values helps to reduce measurement bias. Price indices based on unit values rely on outlier detection and deletion, and the following section explains the methodology for filtering outliers.

3. Outlier filtering: Hidiroglou-Berthelot method

Outliers are defined as unrealistic unit value growth rates, measured as p_t/p_{t-1}, for a given product over time. It is not straightforward to set a threshold for price variations to define outliers. The simplest method is to set a simple threshold, say a price growth above 100% is unrealistic. This can, however, be misleading as the true price changes can be very high for some commodities. Another technique is to winsorize outliers, i.e. define a confidence interval and set data values outside the confidence thresholds to the nearest threshold value. However, this method produces unrealistic unit values, as section 6 below demonstrates. Various other methods have been tried before settling with the H-B method.[1]

The H-B method is used by, among others, Statistics Sweden, U.S. Census Bureau, Statistics Canada and Statistics Denmark. The method is developed to detect

1. One of the alternative methodologies is that of CEPII's, the French institute for research in international economics. CEPII publishes unit value indices for different countries, using UN comtrade database for commodities classified under the harmonized system. The idea is to first compute a median price change for each product and country over time and a median price change for each product across countries at each point in time. Based on the calculated medians, price changes above 5 or below 1/5 times the median are dropped.

outlying observations for a periodic business survey conducted on a regular basis from the same sampling unit, cf. Hidiroglou and Berthelot (1986).

The H-B method uses a composite score function that includes a measure for the *suspicion* that a data value is erroneous and a measure for the *impact* the suspicious record has on publication totals, cf. Jäder and Norberg (2006) and Bartha and García (2012). The method requires accurate and relevant medians and quartiles calculated for *homogenous groups*. Hence, the first challenge is to define homogeneous groups. Jäder and Norberg (2006) apply the H-B method to Swedish trade data that is classified according to the Combined Nomenclature at the detailed 8-digit level and available by country, enterprise, year and month. These data offer various options to define homogeneous groups – based on SITC classification, trade flow, enterprise, country of origin or destination, and there are long monthly time series based on several years of historical data. This is the type of data the H-B method requires.

The OECD database, in contrast, provides only annual data for each member country's trade with the world. This limits the options for classifying traded items into homogeneous groups and calculating relevant medians and quartiles. Consequently, the H-B method has to be adapted to the OECD database. It is not possible to copy the approach in Jäder and Norberg.

One option is to form a group consisting of similar commodities across member countries and calculate quartiles accordingly. But this can be misleading if the trade volume of a specific commodity differs among different countries. In addition, a particular commodity may be reported in one country and not in another, or reported for some part of the sample period only. An alternative is to consider a particular commodity-country combination over time and calculate appropriate time medians and quartiles. One possible shortcoming of this approach is that the sample period (1989-2012) might not be sufficient. On the other hand, it has the advantage of treating each product-country combination separately. Thus, it is chosen to use this approach. Originally, the composite H-B index is computed for the level of unit values. The OECD database reports quantities, but for some products the unit of measurement is changed within the same period, e.g. from metric ton to kilo, producing a clear shift in the quantity level. Working with growth rates instead of levels can make it easier to address the change in units of measurement. Thus, the H-B formula is applied to growth rates of unit values. Furthermore, the growth rates are log-transformed, which makes the distribution of ratios more symmetric. With these considerations suspicion is defined as:

$$suspicion = \begin{cases} \dfrac{q1_t - \log\left(\frac{p_t}{p_{t-1}}\right)}{q3_t - q1_t} & if \log\left(\dfrac{p_t}{p_{t-1}}\right) < q1_t \\ \dfrac{\log\left(\frac{p_t}{p_{t-1}}\right) - q3_t}{q3_t - q1_t} & if \log\left(\dfrac{p_t}{p_{t-1}}\right) > q3_t \\ 0 & otherwise \end{cases} \tag{1}$$

Where p_t is unit value, and $q1_t$ and $q3_t$ are first and third quartiles of unit value growth rates over time, $log(p_t/p_{t-1})$. The suspicion formula flags growth rates outside the lower and upper quartile based on their relative distance to the nearest quartile. A preliminary scrutiny of the data indicates that the SITC components differ in terms of noise level. In general, the manufactured components SITC5-9 contain more outliers than other categories, and the United Nations (1981) manual for compiling unit values recommends paying special attention to machinery and transport equipment (SITC7) and to miscellaneous manufacturing (SITC8). Accordingly, it makes sense to use different suspicion thresholds for different SITC components. An observation is considered an outlier if its suspicion value is above 1.5 for SITC 0, 1, 2, 3, 4, 5, 6 & 9 and above 0.5 for SITC 7 & 8. In this way the amount of data thrown out as outlier is minimized. The lower the suspicion threshold the larger will be the amount of data thrown out as outlier.

Jäder and Norberg (2006) propose a score function consisting of equation (1) and an additional equation for the potential *impact* that the suspicion value could have on the final price index. To calculate potential impact a measure of expected trade value for a given quantity is required. This calculation is not possible with the available trade data, and equation (1) is used without adding an impact equation. Once a growth rate at the detailed level is declared an outlier, it is set to zero, i.e. the level of unit value a year before becomes the best forecast for the current period with outlying growth rate. Setting the outlying growth rate to zero prevents the unit value from shifting level when the units of measurement is changed. In the final aggregation of the detailed unit values to 1-digit SITC, the unit values with outlying growth rates are excluded. A particular commodity is dropped for the whole sample period if more than half of the observations are declared outliers. This test preserves about 90-95 percent of the original data.

4. Price Index: Theory

Once the data have been treated for outliers and other statistical issues, price indices at 1-digit SITC level can be calculated. There are various ways of doing this. The general idea is to relate the growth rate of unit values at 1-digit SITC to a weighted sum of the growth rates of the detailed unit values within each 1-digit SITC. Weighting the detailed unit values requires an index formula, and the most popular indices used by national statistical institutions are Laspeyres, Paasche, Fisher and Törnqvist indices, either fixed-base or chained. For example, member states of the European Union use chain Paasche price indices in their national accounts and the United States uses chain Fisher price indices. The chained Paasche price index is given as:

$$P_t = \left(\frac{\sum_k p_{k,t} q_{k,t}}{\sum_k p_{k,t-1} q_{k,t}}\right) P_{t-1} = \left(\frac{1}{\sum_k w_{k,t} \frac{p_{k,t-1}}{p_{k,t}}}\right) P_{t-1} \quad (2)$$

Where k denotes a 5-digit SITC good, $p_{k,t}$ is unit value of good k in period t and $q_{k,t}$ is quantity of good k in period t, $w_{k,t} = \frac{p_{k,t} q_{k,t}}{\sum_k p_{k,t} q_{k,t}}$ is the weight of good k in period t and P_t is the Paasche unit value index for 1-digit SITC good. The Paasche chain index is known to underestimate volatile price changes by attributing the new larger weight to products that are in larger demand following a relative price drop, Gaulier et al. (2008). Alternatively, the Laspeyres index can be used, given as:

$$L_t = \left(\frac{\sum_k p_{k,t} q_{k,t-1}}{\sum_k p_{k,t-1} q_{k,t-1}}\right) L_{t-1} = \left(\sum_k w_{k,t-1} \frac{p_{k,t}}{p_{k,t-1}}\right) L_{t-1} \quad (3)$$

Where $w_{k,t-1} = \frac{p_{k,t-1} q_{k,t-1}}{\sum_k p_{k,t-1} q_{k,t-1}}$ is the weight of good k in the previous year and L_t is the Laspeyres unit value index for 1-digit SITC good. The Laspeyres index, conversely, tends to overestimate volatile price changes as it uses the weight of the previous year, ignoring that consumers can shift to goods that have become relatively cheaper.

The more noise the data contains the larger will be the upward and downward bias in the Laspeyres and Paasche indices, respectively. These indices also suffer from a measurement problem because they use a single time period for constructing the weights of the index formula. In this regard, the Fisher and Törnqvist indices are preferable as their weights cover two time periods. The Fisher index (F_t) and the Törnqvist index (T_t) are given as:

$$F_t = (P_t \cdot L_t)^{1/2} \quad (4)$$

$$T_t = \Pi_k \left(\frac{P_{k,t}}{P_{k,t-1}}\right)^{(w_{k,t-1} + w_{k,t})/2} \quad (5)$$

The downward and upward biases can be illustrated by a simple example, as table 1 demonstrates. Consider two goods, in period one the price of good-1 drops leading to a higher quantity consumption of good-1, and in period two the price of good-1 and the quantity consumed go back to their values in the base period zero. The chained Paasche index systematically understates the aggregate price for good-1 and good-2 and the Laspeyres index does the opposite. The irreversibility of the chain indices implies that neither the Paasche nor the Laspeyres price index return to their base period value of 1 in period two, although good-1 and good-2 have both returned to their base period price and quantity. This illustrates the permanent bias that price fluctuations can create in chained Paasche and Laspeyres indices. In other words, the Paasche and Laspeyres indices give the lower and upper bounds of the real price change. The Fisher index lies between the two indices and is one way to approach the true but unobserved aggregate price, see also Feenstra (1997) and Gaulier et al. (2008).

Table 1. Upward and downward biases, t-0 = 1

		Period		
Goods		t-0	t-1	t-2
good-1	price	1	0,5	1
	quantity	1	2	1
good-2	price	1	1	1
	quantity	1	1	1
		t-0	t-1	t-2
aggregate price	Paasche	1	0,667	0,889
index	Laspeyres	1	0,75	1,125
	Fisher	1	0,707	1

It should be noted that the Fisher and Törnqvist indices produce almost identical results, hence only the Fisher index is reported. Fixed-base price indices are also not reported, as they are not used anymore in the national accounts.

After unit value indices at 1-digit SITC are constructed for each country using index formula (2), (3) or (4), quantities of imports and exports at 1-digit SITC are constructed by deflating values of imports and exports with their respective unit value indices.

5. Evaluating the price indices

The lack of official disaggregated trade prices implies that there is only little to check the calculated unit values indices against. The old OECD's ITCI-database contains official figures up to 2002 and this can be one benchmark. Another possible benchmark is the OECD's national accounts statistics that have price indices for total goods for the full sample. This section presents unit values for selected countries[2] and compares to existing figures whenever possible.

The use of unit values as surrogates for price indices is more reliable the more homogeneous the products are (Silver, 2007). Energy products are relatively homogeneous compared to other SITC components. Figure 1 presents price indices for energy products (SITC3) for selected countries. The Laspeyres, Paasche and

2. To save space, only import prices for a few OECD countries are presented, additional data can be obtained from ADAM's homepage, or from the link: http://www.dst.dk/ext/adamzone/Unit_value_indices.

Fisher indices are almost identical, i.e. the result is not sensitive to the index formula applied, which may reflect that energy products follow a largely similar international price development. Note also that the unit values closely resemble the official data from the OECD for the available sample period. The difference between the official data and the unit value in US and Japan is due to a difference in price increase in a single year, and in both cases the official data indicates the biggest price change.

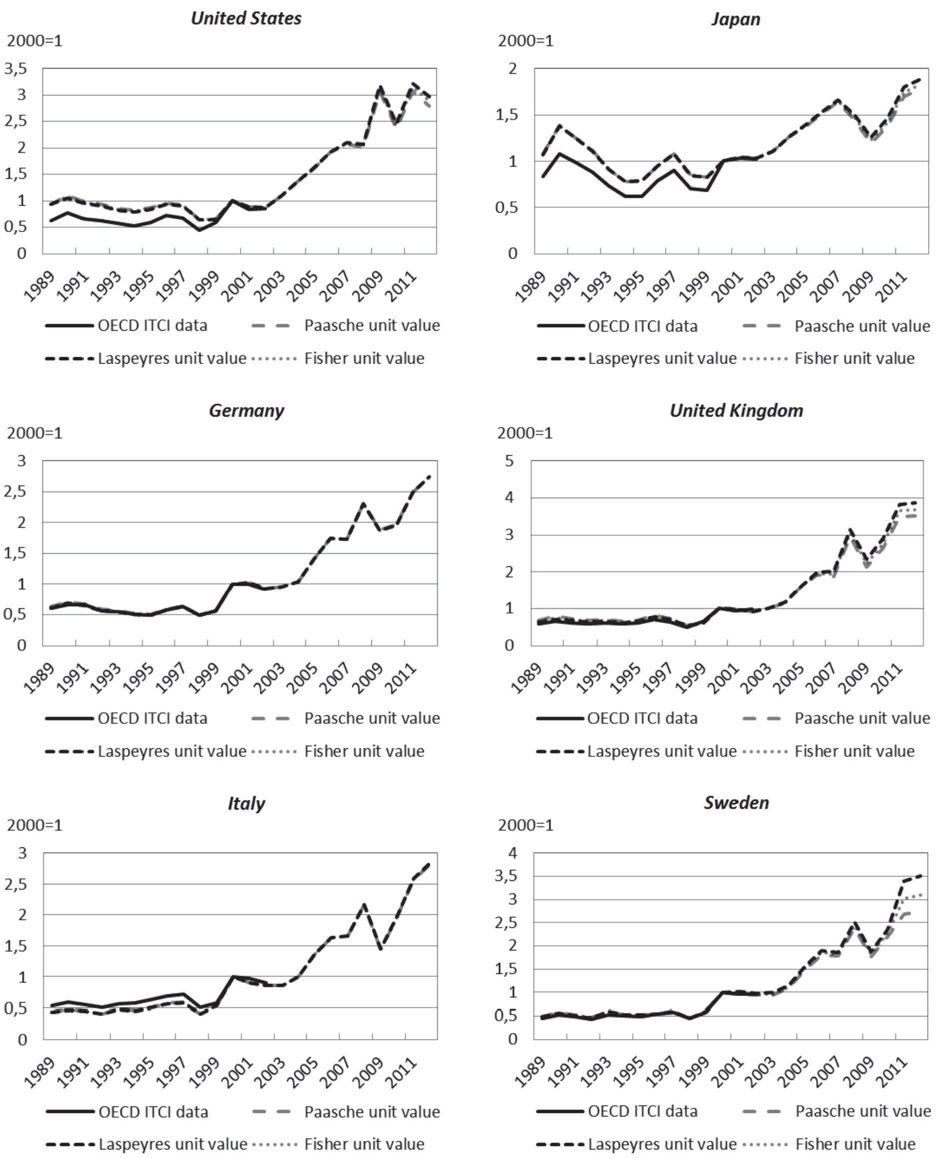

Figure 1. Price indices for energy imports, SITC3, national currency

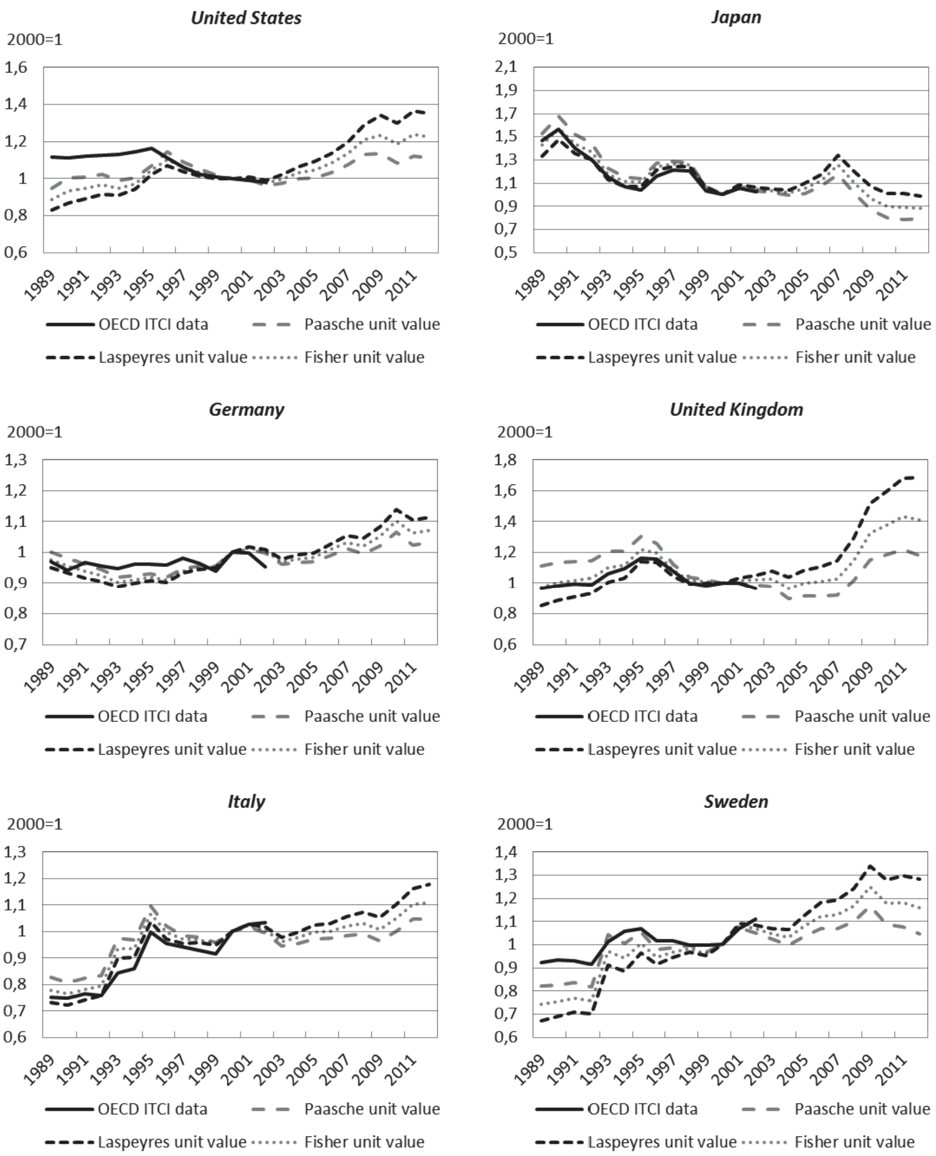

Figure 2. Price indices for manufactured imports, SITC5-9, national currency

Figure 2 presents price indices for manufactured imports for the same set of countries. These price indices are more sensitive to the index formula. The upward bias in Laspeyres indices and the downward bias in Paasche indices seem pronounced. The Laspeyres index sets the upper bound and the Paasche index sets the lower bound, while the Fisher index lies in the middle as expected. The bias margin increases with the amount of noise in the data and the degree of heterogeneity contributes positively to the bias. With a significant bias problem, the Fisher index is the best index to use. In spite of the bias problem, the unit value indices are clearly related to the official data, and the unit value indices for Japan, UK and Italy closely resemble the official figures. In the case of Germany, the growth rates of the official data differ from the unit value indices only in the beginning and in the end of the overlapping periods. A closer look at US prices reveals that the differences are pronounced only in the beginning of the sample. As previously mentioned, the United Nations manual for trade prices recommends the use of unit values in the case of homogeneous products. However, despite the heterogeneity in manufactures, the unit values seem to do a good job in approximating the official prices.

The OECD's national accounts statistics provides prices for total imports and exports of goods for the whole sample period. It can be used to compare the unit value indices for recent years, cf. figure 3. Member states of the European Union use Paasche price indices, and this could explain why the Paasche unit value is close to the national accounts price. However, it is important to note that the national account prices are quality adjusted, which reduces the price growth of IT products. This may be the real reason or part of the reason why the Paasche chain index (with a down-ward bias) resembles the national accounts price. For Japan, the Laspeyres unit value index resembles the national accounts data better than the other indices, it may be the case that the national account prices are Laspeyres chain indices.

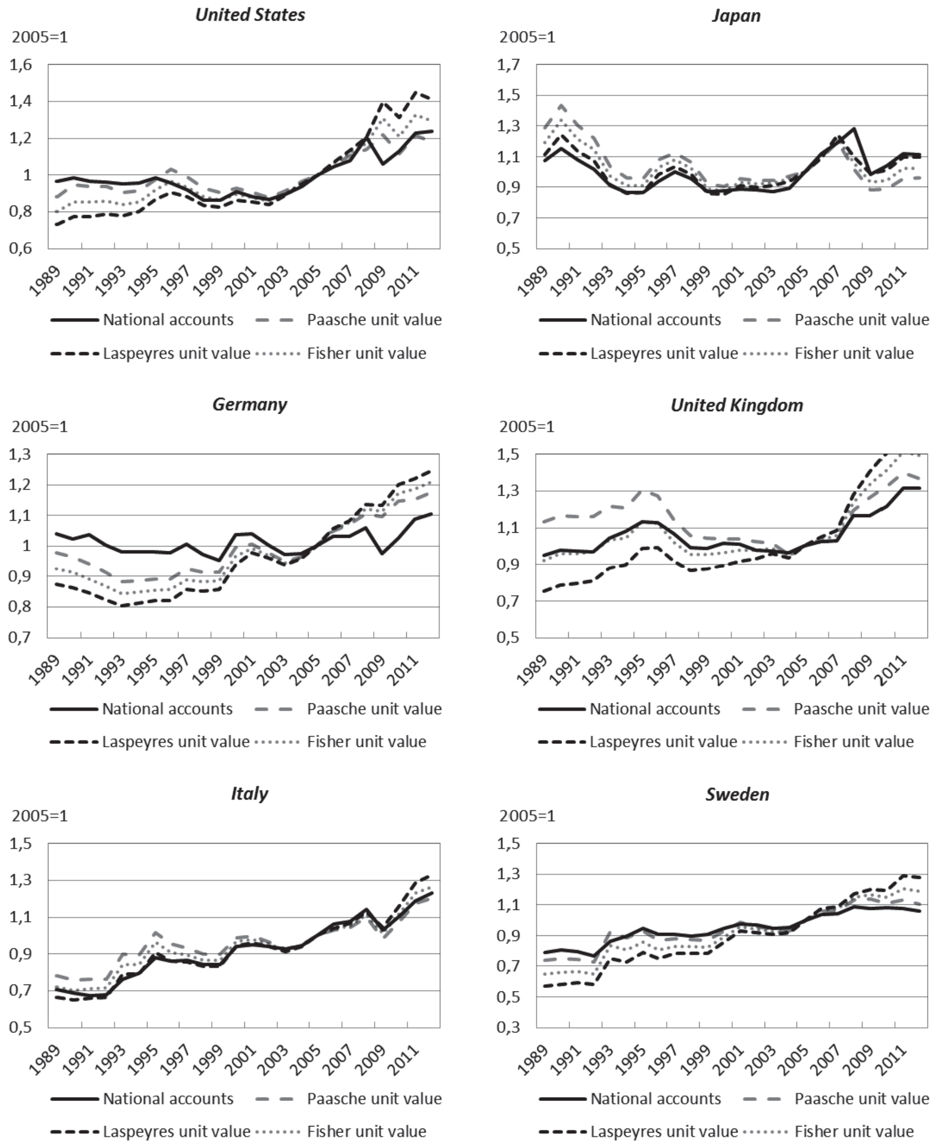

Figure 3. Price indices for total good imports, SITC0-9, national currency

In general, there is a lot of similarity between the national account prices and the unit value indices. Table 2 reports simple correlation statistics between the unit values and the national accounts import price, i.e. three correlation coefficients for each of the six countries.

Table 2. Correlation between national accounts import price and unit values, SITC0-9

	Unit values		
	Paasche index	Laspeyres index	Fisher index
United States	0.879	0.851	0.866
Japan	0.499	0.901	0.724
Germany	0.730	0.643	0.678
United Kingdom	0.839	0.915	0.971
Italy	0.962	0.989	0.984
Sweden	0.985	0.971	0.982

For five countries, one or more of the three unit values has a high correlation coefficient above 0.85 with the official national accounts import price. In the case of Germany, the highest correlation is 0.73 for the Paasche index. A closer look at the German price indices in figure 3 indicates that much of the difference between the national accounts price and the unit value index is attributed to the end of the sample period. Particularly the world-wide recession year 2009 stands out, the correlation coefficient increases to 0.83 when the observation in 2009 is excluded. In this year the unit value index shows little or no price change while the national accounts data shows a significant price drop. This could be due to measurement errors or composition effects in the unit value indices. It is also peculiar that the national accounts import price for Germany has no upward trend since the mid-2000s, while the other countries' import prices trend upward.

Normally, the unit value indices tend to fluctuate more than the national accounts data. This is for example the case for Sweden, where the devaluation in 1993 is more visible in the unit values than in the national accounts data. In general, unit values are not perfect substitutes for price indices, but given the lack of genuine price indices for trade flows at a disaggregated level unit values can be an alternative.

6. Sensitivity analysis

The final price indices have been subjected to several sensitivity tests. First, different suspicion thresholds for the chosen H-B method have been tried. The choice of a threshold of 1.5 (excluding SITC 7 & 8) is ad hoc and is not that sensitive. A threshold of 0.5, 1, 2, 5, 10 and 20 has been tried and the basic result does not change significantly for threshold values above 1.5. As the threshold increases, a few price growth rates show up with extra-large value without significantly changing the year-to-year price variations in 1-digit SITC components. However, the upward and downward biases in Laspeyres and Paasche index cumulate the outlying observations and produce diverging price trends in the long run, specifically if the threshold for SITC 7 & 8 is greater than 0.5. The chosen threshold of 1.5 represents the case where the correlation between the constructed unit values and the official national accounts prices is the highest and at the same time recovers a significant portion of the original data, cf. Sisay (2012).

Alternative filtering techniques have also been tried, for instance winsorizing outliers. The technique defines a specific percentile beyond which values are considered outliers. For example a 95 percent winsorization sets all values below the fifth percentile to the fifth percentile and all values above the ninety-fifth percentile to the ninety-fifth percentile. One drawback of this technique is that the cutoff points originate from the data and can be placed among the outliers. If a unit value has a level shift due to changes in the unit of measurement, setting that particular growth rate to zero level corrects the series, which is not necessarily the case in winsorization. Winsorizing replaces values outside the chosen confidence interval by the nearest confidence threshold, implying that the impact of the level shift is reduced but not removed, especially not if the confidence interval is influenced and widened by outliers. As already mentioned, the H-B method flags outliers according to their relative distance with respect to the interquartile distance. Thus, the method pays attention to the distribution of the individual series and not all outliers outside the chosen interval are deleted.

Figure 4 demonstrates the possible differences between the H-B method and winsorization for two OECD countries. For winsorization two thresholds are reported, 95% and 90%. In the case of Germany, a 90 percent winsorization produces a unit value that is close to the unit value based on the H-B method. An attempt to increase the data recovery rate by increasing the percentile to 95 percent produces two distinct level shifts. Removing the two level shifts produces a unit value that resembles the other two unit values in the figure. This speaks against winsorizing at 95% threshold and also in favor of setting outlying growth rates to zero. For the case of US, both 95 and 90 percent winsorization produce a unit value that is totally different from both the H-B based unit value and the official national accounts statistic. One can increase the cutoff percentile further, but already 20 percent of the data is trimmed at 90 percent winsorization, hence the method can be costly in terms of data recovery. Based on the H-B method the average data recovery is 97% for Germany and 91% for US for all SITC, these values increase to 98% and 94%, respectively, if we exclude the more volatile SITC 7 & 8; in either case the data recovery rate is better in the H-B method.

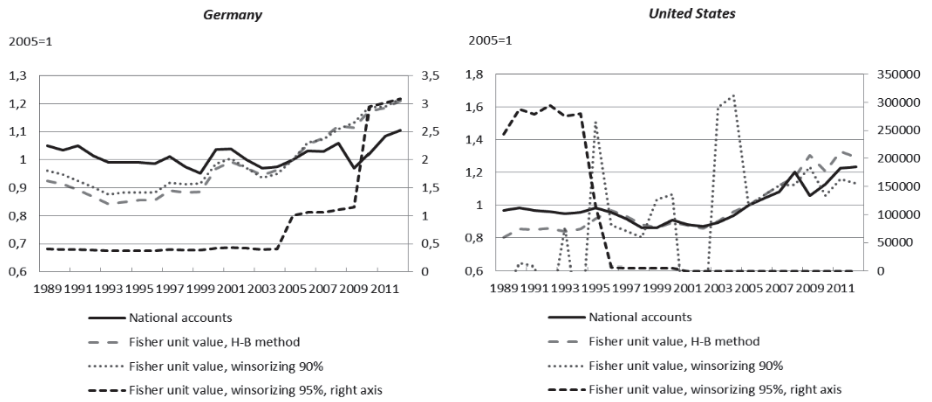

Figure 4. Price indices for total good imports, SITCO-9, national currency – different filtering

7. Summary

This paper has presented disaggregated trade prices for OECD countries. Prices are constructed based on unit values using annual data from OECD's ITCS database. The new data can be used as a substitute for OECD's ITCI database that was discontinued in 2002. The price indices are in general sensitive to the choice of index formula. The Laspeyres index has an upward bias and the Paasche index has a downward bias. The Fisher index is a geometric mean of the two indices and it is the preferred choice. Unit value indices are shown to be highly correlated with the available national account prices for total goods. More specifically, the national account prices are quality adjusted and one would expect a higher price growth in the unit values that include the rising quality in the price development. This seems confirmed as the national account import prices of some OECD countries are closer to the unit-value-based Paasche chain index, which grows slower than the Laspeyres and Fisher indices. It is the volatility of the unit values that creates the downward bias in the Paasche chain index, and this bias may to some extent compensate for the lack of quality adjustment. The new trade prices has already been used for an overdue update of export market prices and volumes in ADAM, and the dataset can be used for further and more disaggregated analysis of the foreign trade performance in Denmark and other OECD countries.

Literature

Bartha, E. and M. García. 2012. Score Function for Selective Editing of the US Census Bureau Trade Data. *U.S. Census Bureau Research Report Series*, 2012-10.

Durand, M., C. Madaschi and F. Terribile. 1998. Trends in OECD Countries' International Competitiveness: The Influence of Emerging Market Economies. *Economic Department Working Paper* no. 195.

Feenstra, R. C. 1997. U.S. Exports, 1972-1994: With State Exports and Other U.S. Data. *NBER Working Paper* no. 5990.

Gaulier, G., J. Martin, I. Méjean and S. Zignago. 2008. International Trade Price Indices. *CEPII Working Paper* 2008-10.

Hidiroglou, M. and J. Berthelot. 1986. Statistical Editing and Imputation for Periodic Business Surveys. *Survey Methodology* vol. 12, no. 1.

Jäder, A. and A. Norberg. 2006. A Selective Editing Method Considering both Suspicion and Potential Impact, Developed and Applied to the Swedish Foreign Trade Statistics. *Economic Statistics* 2006-3.

Silver, M. 2007. Do Unit Value Export, Import and Terms of Trade Indices Represent or Misrepresent Price Indices? *IMF Working Paper* no. 121.

Sisay, D. 2012. Export Market and Market Price Indices: Trade Statistics Data. *Statistics Denmark Model-group Working paper DSI23* nov12.

Statistics Denmark. 2013. ADAM – a Model of the Danish Economy.

United Nations. 1981. Strategies for Price and Quantity Measurement in External Trade. *Department of International Economic and Social Affairs Statistical Paper* no. 69.