Application of supply and use tables: centralized deflator compilation system Voltti for production indices

Ville Lindroos
Senior Statistician
Statistics Finland

CONTENTS

1. The need for a new deflator system

1.1. Coherency

The greatest motivation for developing a new system for deflator management, was the need for better coherency between the different volume statistics produced in Statistics Finland. Different statistics producing volume figures:

- National accounts
  - Annual national accounts
  - Quarterly national accounts
  - Trend indicator of output
- Business statistics
  - Industrial Output Indicator
  - Index of Service Production

Two statistics often compared are the Trend Indicator of Output (TIO), produced under national accounts, and the Industrial Output Indicator (IOI), produced under the business statistics. Trend Indicator of Output's volume information concerning manufacturing industries covers the B-F industries, while Industrial Output Indicator covers B-D industries, difference being thus E (*Water supply; sewerage, waste management and remediation activities)* and F (*Construction*), all the other industries are overlapping.

The two statistics tended to show slightly, or sometimes even greatly, different image of the economic outlook for the same industry.

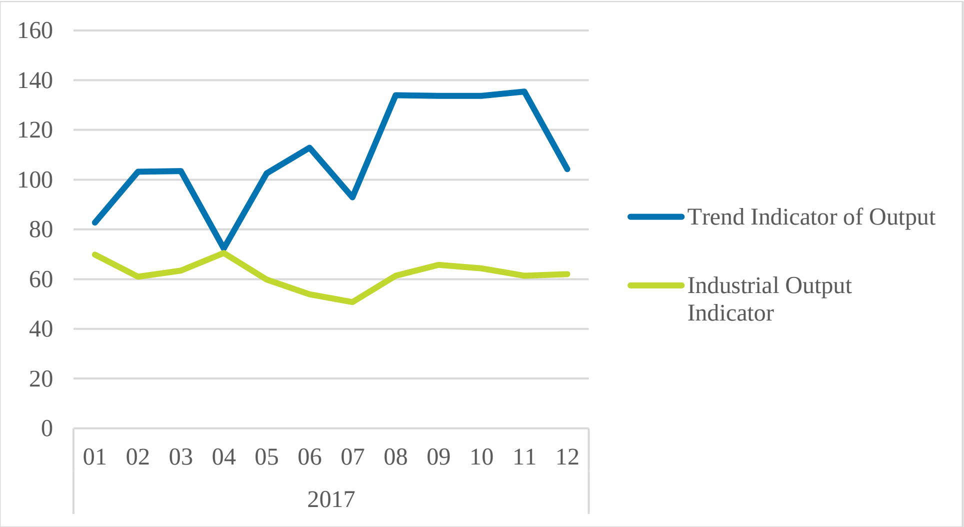

Figure 1 TIO and IOI volume comparison of the NACE industry 14 Manufacture of wearing apparel, before the reform project.

Multiple causes could be determined for the discrepancies between the two statistics:

- The underlying output data in current prices
- The deflation methods:
    - The weights
    - Price sources
    - Including side production
- Extrapolation and interpolation
- Volume calculation methods:
    - Chain-linking
    - Adjustment for seasonal variation

A project for a new volume calculation system Voltti was established. All of the above-mentioned topics were tackled in the project, but on this paper focuses mainly on the deflation part as it utilizes supply and use tables.

1.2. Common methods, different regulation

One challenge for unifying the methods, was the different legislations behind the business statistics and the national accounts. The volume indices, including IOI, produced under business statistics were previously complying with the requirements set by the Short-term statistics (STS) regulation. STS regulation will be however integrated into the new *Framework Regulation Integrating Business Statistics* (FRIBS) and to be able to comply with the new regulation was another reason to reform the volume index production systems. The legislation dictating the methodological choices for volume calculations related to national accounts comes from the European System of Regional and National Accounts (ESA 2010). In practice the preferred methods of ESA 2010 are not contradicting with FRIBS when it comes to the deflators.

The compilation process for deflators differed massively between the two statistics of our interest. The trend indicator of output is produced by the National Accounts unit. Both annual national accounts and quarterly national accounts are using the supply and use tables for the products weights for industry level deflator compilation. Thus also the trend indicator of output has been using already before the Voltti project the SUT product weights. In addition to the common weights, all the volume statistics under national accounts also used the same exact price index series for the product level deflation.

The industrial output indicator however did not prior to the Volume reform project use product level weights in the compilation of the deflators. The more detailed level industries were directly deflated with appropriate producer price indices. There was no control or consistency examination whether the price indices chosen for each industry matched between the two statistics.

The need for having a common way to define the deflators kindled the Deflator subsystem part of the Voltti project.

One challenge for unifying the methods, was the different legislation behind the business statistics and the national accounts. The legislation dictating the methodological choices for volume calculations related to national accounts comes from the European System of Regional and National Accounts (ESA 2010). The volume indices produced under business statistics are aiming to comply with the Framework Regulation Integrating Business Statistics (FRIBS). In practice though the preferred methods of ESA 2010 are not contradicting with FRIBS when it comes to the deflators.

1.3. Common tools – Deflators software

As every volume statistic needed a way to define which prices to use for which products and ultimately industries, a new tool was needed. The principal idea being that eventually each statistic would have exactly the same deflators by definition. It was decided that each volume statistic would accept the Supply and Use tables for the weights, and over time, all the calculation methods set by SUTs as well. As now everything was the same product level as supply and use tables, circa 800 CPA products, the prices could also be defined precisely the same way for each product and four supply and use classes. From the supply and use classes two were supply classes: domestic output and imports, and two were use classes: household consumption and exports.

The Volume project thus got its *Deflators* –subsystem. The graphical interface of this system is used to control all the metadata that defines the compilation of deflators. The most important part is defining for each product the price index series it uses for the product level price change information. Another common definition controlled with the new tool, were the extrapolation and interpolation methods. Extrapolation methods include four and the interpolation methods three different alternatives.

2. Supply and use tables in Finland

2.1. General info

Annual supply and use tables as well as input-output tables have been published since 2002. Balanced annual supply and use tables in current prices as well as input-output tables has been compiled for the years 1995-2015. The supply and use table framework has been used to compile the price and volume calculations in the Finnish national accounts since 2005. The input-output tables have been compiled previously for years 1956, 1959, 1970, 1980, 1982, 1985, 1992, 1993, 1995. The compilation is done in the level of detail of 182 industries (NACE rev2) and 836 products (CPA). The publishing level is 64 industries and 64 products. Other relevant classifications related to the compilation include the household consumption classification COICOP, 182 classes, and institutional sector classification, 23 classes.

2.2. Product data sources

The basis for the current priced compilation of the supply and use tables is the preliminary product transaction data, published in t+19 months, of various sub systems of national accounts. These product transactions data categories are:

- output and intermediate consumption
- imports and exports
- gross fixed capital formation
- private consumption, public consumption expenditure and consumption expenditure of non-profit institutions
- changes in inventories
- taxes on products and subsidies on products

In the compilation of the supply and use tables, the product transaction data concerning the supply and use of the national accounts is divided into 836 products. The data is further classified into 19 supply data categories and 51 use data categories.

Various data sources are used when obtaining the product structures for the different accounts:

- Data variables of the output and intermediate consumption of production account (based of the turnover and expenditure items of SBS)
- Commodity statistics' production and raw material data
- Imports and exports of goods (Finnish Customs)
- Imports and exports of services Statistics Finland's Balance of payment statistics
- Business services statistics (NACE 58, 62_63,69, 702, 71, 73, 78)
- Data processed by NA experts (NACE 01-03)
- Product data based on the previous year's product data corrected with price changes (NPISHs)

The source data (SBS) for the intermediate consumption contains the following data variables:

Data variable	*Heading*
P22_131	Acquisition of materials and supplies
P22_132	Acquisition of packaging materials
P22_133	Acquisition of fuels
P22_134	Contracted repair, maintenance and installation work
P22_135	Contracted paid labor
P22_136	Subcontracting
P22_137	Renting of labor force
P22_141	Acquisition of electricity for own use
P22_1410	Representational expenses
P22_1411	Other expenses not mentioned above
P22_142	Acquisition of heat for own use
P22_1431	Research and development expenses
P22_1432	Research and development expenses

P22_144	Transport and storage expenses
P22_1451	Advertising, sale and marketing expenses
P22_1452	Activated marketing expenses
P22_1453	Marketing costs (agency etc.)
P22_146	Computer, design and programming expenses
P22_147	Expenses from patents and licenses
P22_148	Leasing expenses
P22_149	Other rents

Table 1 Intermediate consumption data variables

2.3. The compilation of the balanced SUT

The compilation of the balanced supply and use tables can be split to roughly five steps. The first four are related to compiling the unbalanced tables and the fifth is the balancing:

1. Compilation of supply data at basic prices and use data at purchaser's prices by product
2. Compilation of price formation items
3. Converting use data at purchases prices into basic prices
4. Compilation of the unbalanced supply and use tables
5. Balancing the supply and use tables
    i. Manual balancing
    ii. Automatic balancing of price formation items
    iii. Elimination of the statistical discrepancy
    iv. Automatic balancing of the remaining small differences

Manual Balancing

The most significant product imbalances between supply and use at basic prices are corrected within manual balancing. Manual balancing focuses on products with an imbalance at basic prices greater than ten percent and the absolute value of the imbalance over EUR 30 million. Manual balancing can touch easily even half of the 836 products.

Automatic balancing of price formation items

The automatic balancing of price formation items consists of two functions. The subsidies and taxes on products in use data are adjusted with the paid and collected subsidies and taxes on products. The trade and transport margins are adjusted with the supply of service products that produce trade and transport margins.

Elimination of the statistical discrepancy

Statistical discrepancy of total supply and total use at basic prices is eliminated balancing the products whose differences between supply and use are biggest and similar to the statistical discrepancy. In the elimination of the statistical discrepancy, the preliminary levels of industry-specific output and intermediate consumption data, imports and final use change.

Automatic balancing with RAS algorithm

All the remaining, at this point relatively small, imbalances are eliminated using the RAS algorithm. In this method the product values may change ***within use*** excluding the use category specific and product values that have been decided to fix in advance (e.g. fuels in household consumption expenditure). As a result of automatic balancing, the balanced supply and use tables are generated, where supply and use at basic prices are in balance by product and output type.

3. Organization of the system management

3.1. Roles and responsibilities

As the new system was something that was used cross the major business units, it also required some new ways of working. The software as a product is owned by the business statistics. Senior statisticians of the National accounts unit, more specifically the supply and use table –team's price and volume experts, have the responsibility over the product-level decisions. This includes the supply and use tables for product weights, and the deflator choice for each product for each supply and use class. However a cross-departmental expert group called the *Deflator group* has been formed in Statistics Finland to discuss price and volume issues. This group has representatives from each price and cost statistics available, representatives from all the volume statistics, and representatives of national accounts. The chair is held by national accounts.

The project itself is in the change management –stage, and regular development is still done. The major decisions, especially regarding IT resources, are done with the approval of a steering group consisting of heads of statistics.

4. How the system works

4.1. User interfaces

As the one of the main ideas of the new system was to unify the deflator definitions, a graphical user interface was needed to be able to manage all the definitions. The software has a browser based user interface. Through the interface, all the necessary metadata definitions can be managed. The different functionalities of the Deflator software are:

1. Product basket management
2. Price source linking to products
3. Other price sources - management
4. Methods
5. Reports
6. Tax management

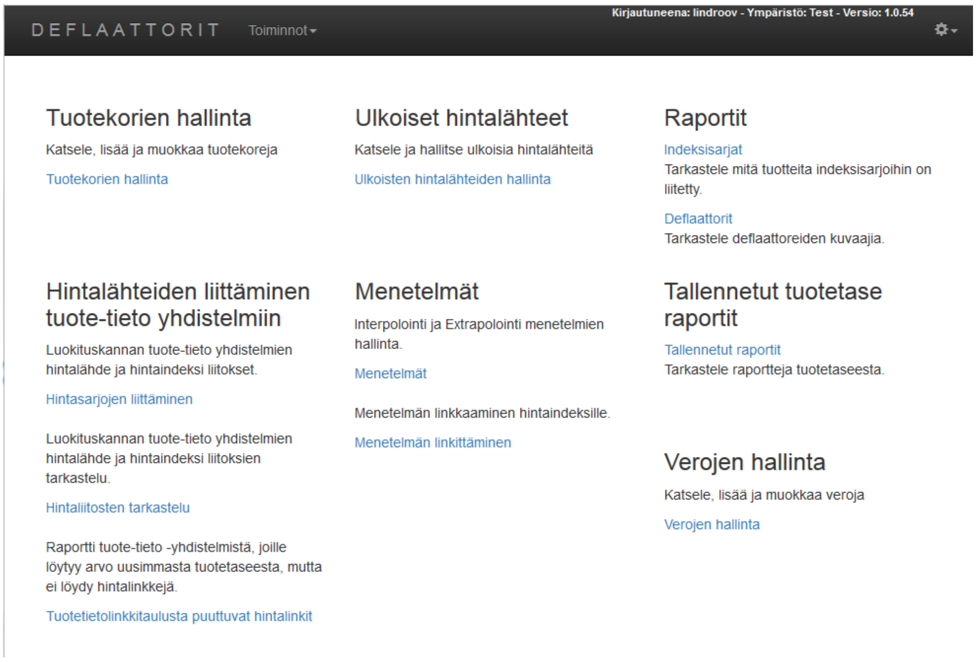

Figure 2 Deflator software user interface

4.2. Deflator query

The deflators can be called by the main volume calculation software *Voltti* or by anybody via web browser. What this means in practice, is typing a standard form URL in the web browser, where multiple parameters are defined for the desired deflator. These parameters include notably:

- the industry
- the sector
- time window

Additionally many other parameters can be defined if the user wants to deviate from the default values. These are for example:

- the index formula for the deflator aggregation (default is Paasche)
- different product link versions
- different product basket versions (balanced vs unbalanced for example)
- clearing cache (index series are cached to make the queries faster)

Here is an example of a url for a query with all the parameters separated to its own line:

http://apptest1.ad.stat.fi:51461/ Deflaattorit2/api/deflator/	*address for the server*

deflatorvalues?name=14	*NACE industry, only one at a time can be specified*
&sector=S11	*Institutional sector, multiple can be specified*
&datatypes=1010	*1010= domestic market production* *NTP=domestic production including non-market producers* *2010=intermediate use*
&productBasketVersion=201901	*Supply and use table version, can be left undefined in which case the default version is used. The default is always defined as the most recent version.*
&beginDate=2017-01-01	*The first observation month for which the deflators are desired*
&endDate=2017-12-31	*The last observation month for which the deflators are desired*
&calculationMethod=2	*1= base index, 2 = chain index, for most of the users the chain index is more practical*
&baseYear=2015	*The desired base year, if the deflators are wanted as a base index. If chain index is wanted, this parameter is obsolete*

As a part of a production process the national accounts users post a query for the deflators as an url via SAS software. This means that also the results also appear in SAS and fits smoothly in the process flow.

4.3. Extrapolation and interpolation of prices

4.3.1. Extrapolation

The volume indices of business statistics as well as the trend indicator of output by national accounts, are produced in such a quick cycle, that not all the price sources are available at the time of compilation. For this purpose, the prices that are unavailable need to be forecasted, through extrapolation. Three main models are used for extrapolation: Previous observation carried forward, Autoregressive model (AR) and Autoregressive integrated moving average (ARIMA). In addition a special method is available for the producer price index for services, which utilizes the consumer price index to forecast the SPPI. The default method can be managed with the main user interface of the Deflators –software. It is also possible to specify each price series their own extrapolation method. The examination of the different methods showed that most often the normal AR model produced best results. In addition to the afore-mentioned four methods, a fifth option exists for the default method: the previous month's best method. This means fitting all the different methods and seeing which performed best in the previous month. This is the standard default option, as it automates the process of monitoring which method works best in different situations.

4.3.2. Interpolation

Some of the important price sources are published quarterly instead of monthly. This means that for the use of monthly volume statistics, the monthly prices need to be interpolated from the quarterly data. The interpolation methods are defined in the same tab of the software as the extrapolation. The three different options available: Same value for each month of the quarter, linear interpolation, and spline interpolation. One of the difficulties of determining the method that yields best results is the lack of monthly comparison data. The decisions have been thus made mainly visually examining the results of the different methods.

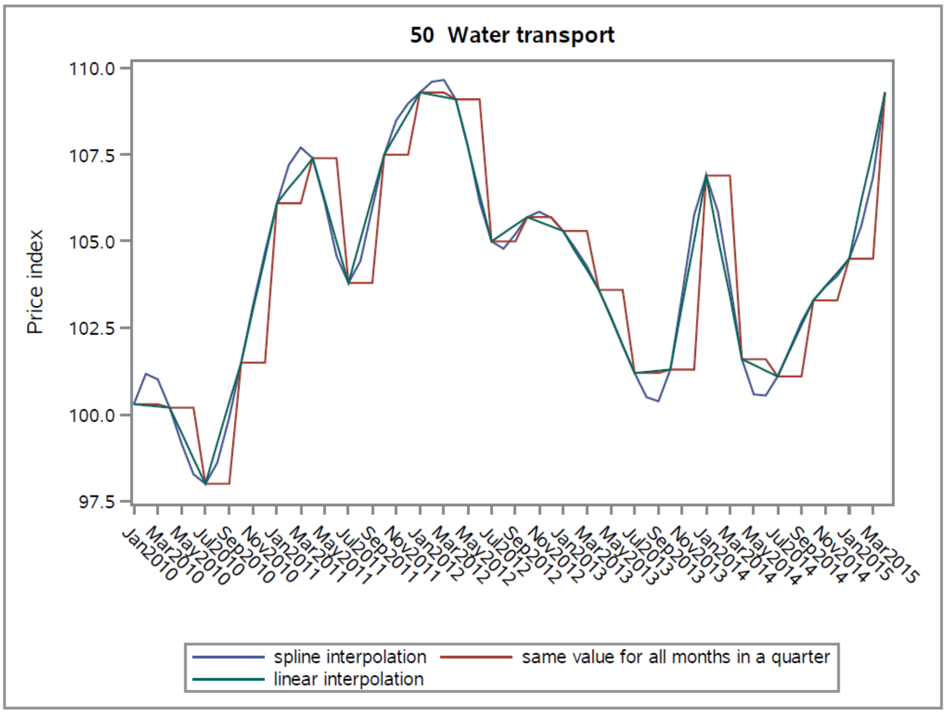

Figure 3 Comparison of the three interpolation methods of the price index of water transport

Conclusion of the examination was to use linear interpolation as the default method. Spline interpolation produces very similar results to the linear interpolation. Same value for each month seems to be the worst method.

4.4. Calculation process flow

4.4.1. SUT-based product baskets

One of the most important methods to unify the volume calculations, was to get everyone the same way of compiling deflators, and this method was decided to be the SUT way. For this purpose the SUTs need to be provided on the Volume-system's use. To ensure the identical product weights for each industry between different users, the precisely same SUTs are being used for all the Volume users. This means that the most recent, despite being an unbalanced provisional SUT, is being used for all the deflator aggregation purposes.

When the user calls for a deflator of a certain industry and sector, the first part of the process is looking up this industry-sector combination in the product basket. Inside the programs process flow, this query now returns the products on that industry-sector with their weights.

4.4.2. Price index series

Part of the process is connecting these product weights with prices, to obtain the product level deflators. Each product-SUT class combination can have multiple price series linked to it.

The decisions on the product level deflators, i.e. the price index series for each product by supply and use class, are made based on the ESA2010 and more precisely the additional information provided by the Handbook on price and volume measures. Most important price sources are:

– Producer price indices
– Producer price indices for services
– Consumer price index
– Building cost index
– Cost index of civil engineering works
– Index of real estate maintenance costs
– Rents of dwellings

For the deflators of the domestic output, naturally the producer price indices are in the most important role. They are however complemented by multiple smaller specialized price and cost indices, for example regarding construction costs or forestry prices. In practice this work can be managed in the *Price source linking to products* –section of the software's user interface. A relation database is used to manage the metadata for the selections

4.4.3. Aggregation order of compilation

The next step in the process after obtaining the product weights for the industry-sector combination from the SUT based product basket, and linking them to price indices to get the product level deflators, is to aggregate the product level weights and product level deflators to the industry-sector level deflator. The order of the aggregation from the most detailed level price index series to the most aggregated industry-sector level deflators can make a difference to the end results, and thus needs to be considered carefully.

A link between CPA products and price index series for each supply and use category/class is used to define which product is deflated with what prices. Each product-SUT class combination can have multiple index series as its deflator, but only from one price source. The reason why multiple series might be needed, is the fact that the classifications do not always match completely, e.g. CPA compared COICOP. Another reason is that the price index classifications can be more detailed than the SUT product classification, e.g. 8 digit CPA in the producer price index compared to the 6 digit CPA of the supply and use tables.

4.4.3.1. Splitting the domestic production by use

For the domestic production, a separation is made for the domestic and export prices, depending on the use. For this purpose the ratio between the domestic use and exportation is obtained from the use table. This is the only information obtained from the use table for the purpose of compiling the deflators for domestic production.

4.4.3.2. Aggregating prices at the use class level

When it comes to the order of compilation, the domestic and export parts must be first aggregated separately to the product level. This is to use as little as possible Laspeyres type of aggregation. This step needs to be done with Laspeyres formula, as there is no time T weight information yet available for the different price index series when the product level aggregation is done, thus making it impossible to use Paasche formula yet.

4.4.3.2.1.1. The case of price index being a base index

If the price index is a base index, the aggregation of the multiple series goes in a simple straightforward Laspeyres fashion where the series' current period price and previous year's average are calculated by aggregating the series-weights obtained from the price statistics along with the index number. The current period price aggregate is divided by the previous year's average price aggregate to derive the product level deflator.

4.4.3.2.1.2. The case of price index being a chain index

In the case of the price index series that is linked to the product being a chain index, another step needs to be taken in the process. To correctly aggregate the product deflators from the price index series, the index series needs to have their chain-linking reversed first. When the links of all the series have been reversed to the point where the price statistic in question started chain-linking, the aggregation can be done. Next the new aggregate has to be chained again to the time period of original interest. In this fashion the current period and previous year's average price aggregates are being formed, and with those the product level deflator can be calculated.

4.4.3.3. Aggregating to product level

After having aggregated the price series separately for the domestic use and the export proportions, the two can be aggregated to obtain the complete product level deflator. The aggregation from the domestic and export disaggregation to the product level aggregate deflator does have the most recent use table information at its disposal, and thus can be done in Paasche fashion. Same process is done for both the current period T and for the previous period T-1 (usually previous year's average) prices. From the two periods, the price change, i.e. the deflator, can be derived.

This step is not completely necessary. Aggregating directly from the product-use level to the industry-sector level, if done correctly, would yield the exactly same results. However having this intermediate stage done is helpful for debugging potential problem situations. It gives us the possibility to compare directly to the product level deflators of the supply and use tables.

4.4.3.4. Aggregating to industry-sector level

The third aggregation step is to compile the industry-sector level aggregates from the product level weights and deflators, using Paasche index formula.

4.4.4. Index formulae

The index formula used in each step is decided based on the desired end result, and data availability. The final product according to both of the governing regulations is a Laspeyres type volume index (ESA2010, FRIBS).

$$L_Q = \sum_{i=1}^{n} \left(\frac{q_i^t}{q_i^0} \right) \frac{v_i^0}{\sum v_i^0}$$

Formula 1 Laspeyres volume index

This means that we need to compile our deflators to be Paasche price index which matched with a Laspeyres volume index, will decompose the value change into a price index times a volume index. Reversely, matching the value change with a Paasche price index, produces us the desired Laspeyres volume index.

$$P_P = \left[\sum_{i=1}^{n} \left(\frac{p_i^t}{p_i^0} \right)^{-1} \frac{v_i^t}{\sum v_i} \right]^{-1}$$

Formula 2 Paasche price index

4.5. Delivery of the end results

After the user has posted a query, the above process will be executed by the program. The end result, the industry-level deflator, will be returned. The values can be shown in JSON in a more human-readable form as shown in figure 5, or in raw data lines as shown in figure 6.

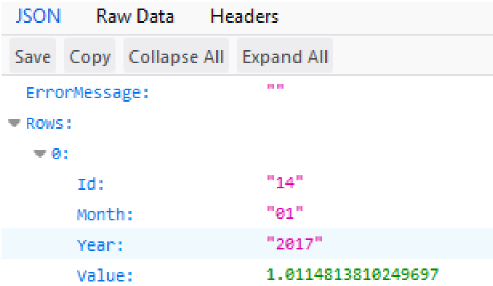

Figure 5 01/2017 NACE 14 deflator value returned in JSON

Figure 6 01/2017 – 12/2017 NACE 14 deflator values as raw data lines

5. Advantages gained with the new system

5.1. Coherence

As it was in the beginning stated, the spark for the whole Volume –project was the lack of coherency between the economic outlook provided by business statistics and national accounts. The new system has made sure, that everything that can be done with the same methods, same definitions, and same underlying data, is indeed done identically. The economic outlook that each volume statistic shows tells now a coherent story of the Finnish economy as we can see from the figure 7.

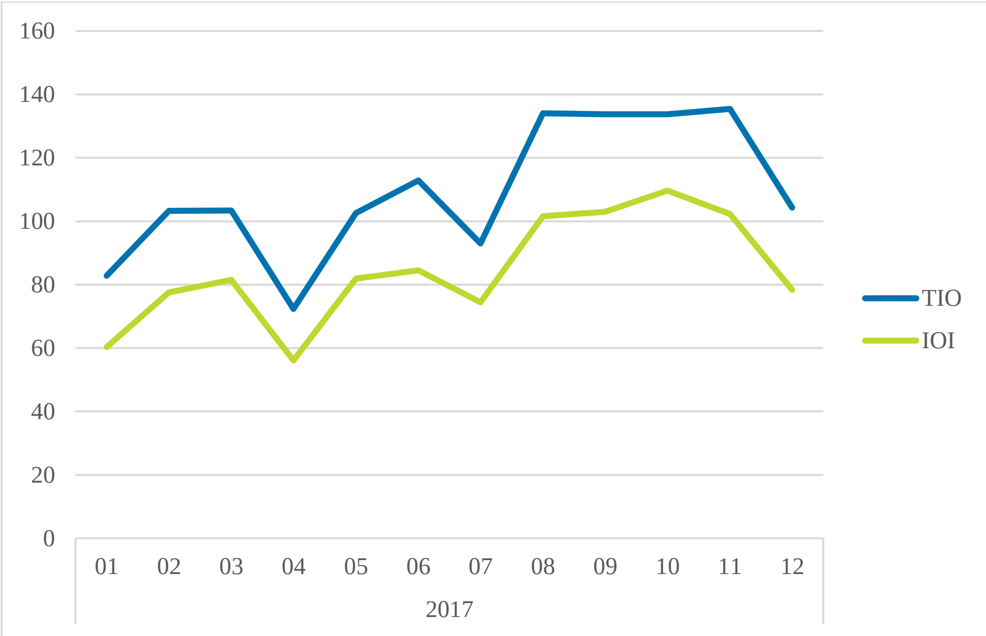

Figure 7 TIO and IOI volume comparison of the NACE industry 14 Manufacture of wearing apparel, after the reform project.

5.2. Quality of the deflators

Not only are the deflators now similar between the different statistics, but they are also overall better. The quality comes from theoretically correct methods, and from a developed process that reduces human errors. With the first part I am referring to the fact that all the volume statistics that can, use now supply and use tables for obtaining the best possible weights for aggregating deflators. Also the supply and use table team continues to spend time and effort to ensure that the price sources are utilized as well as possible.

Previously a situation where different users chose a different version (different year, balanced or imbalanced, different round of preliminary tables) of the supply and use tables for product weights. Now the correct table is always prepared and delivered directly for the new Voltti system's disposal whenever a new version is ready.

The price links previously were maintained in a multitude of excel files, thus prone to human errors. They were often out of date which lead to less than ideal price indices being used for deflation purposes. Now there is one link for all the products and price sources, and for all volume purposes. Technically the link is now a sql database table, which can be very practically and securely accessed for updating or examination purposes. New price index series are developed constantly in the different price statistics, and each new series will help to better deflate the industries through the supply and use tables.

5.3. Workload redistribution

A major benefit gained from the project is the redistributed workload. One part of it is simply the developed new software and chosen methods, that reduce the amount of manual tasks needed. Another part is being able to use same definitions and solutions for all the statistics, compared to the pre-Voltti situation where each statistics did their own, often overlapping process steps. For IOI this reduced workload has been realized by not collecting as much quantity data now that deflating turnover data has replaced it. In addition to the lower total workload for all the volume statistics, the remaining workload is now organized more efficiently. The share of labor between national accounts and business statistics is clear, and help is gained from the price source statistics easily via aforementioned deflator-group. In general the use of the cross-departmental deflator-group has made methodological discussion, development, and decision making easier and faster. When all the relevant statistics are present and aware what is happening, we avoid having to go back and forth with different solutions. At the same time we have always the maximum amount of expertise at our disposal.

5.4. Faster production

When each step of the process has been rethought and the workload redistributed, the overall process is faster. One important contributor for this is avoiding doing any work more times than necessary, another one is more automated processes. An example of the former one is IOI now using for some industries exactly the same methods as TIO would use. IOI is slightly faster statistic and is published roughly one week before TIO. This means that instead of TIO repeating the same process as IOI for certain industries, TIO can actually use the volume information produced by IOI and use it as a volume indicator. Same applies for the index of service production, its volumes can be used for TIO’s service industries.