

TWINNING CONTRACT

BA 15 IPA SR 01 17

Support to the reform of the statistics system in Bosnia and Herzegovina



MISSION REPORT

Activity 2.C.1:
IT application development for producer prices I

Component 2: Business Statistics

Mission carried out by
Søren Netterstrøm, Statistics Denmark
Martin Birger Larsen, Statistics Denmark
Janni Stavvad, Statistics Denmark

05-09 March 2018

Version: Draft

Expert contact information

Søren Netterstrøm
Statistics Denmark
Copenhagen, Denmark
Email: sne@viv.dk

Martin Birger Larsen
Statistics Denmark
Copenhagen, Denmark
Email: mbl@dst.dk

Janni Stavad
Statistics Denmark
Copenhagen, Denmark
Email: sta@dst.dk

Table of contents

Executive Summary	4
1. General comments	4
2. Assessment and results	4
Requirements for a common IT-application across the SPPI and CPPI in BiH	5
Conceptual model	9
Plan	10
Development tool	10
Pros and cons for using a common system	11
3. Conclusions and recommendations	11
Annex 1. Terms of Reference	12
Annex 2. Persons met	13
Annex 3: Similarities and differences across the five price statistics at Statistics Denmark ...	14
Annex 4: Similarities and differences across the two price statistics at the tre institutions FIS, RSIS and BHAS	16
Annex 5: Quality adjustments	18
Annex 6: Conceptual model	19
Annex 7: Pros and cons of developing a common it system	20

List of Abbreviations

BHAS	Agency for Statistics of Bosnia and Herzegovina
BiH	Bosnia and Herzegovina
CBBH	Central Bank of Bosnia and Herzegovina
EC	European Commission
EU	European Union
FBiH	Federation of Bosnia and Herzegovina
FIS	Institute for Statistics of Federation of Bosnia and Herzegovina
MS	EU Member State
RSIS	Institute for Statistics of Republika Srpska
RTA	Resident Twinning Adviser
ToR	Terms of Reference
CPI	Consumer Price Index
HICP	Harmonised Index of Consumer Prices
PPI	Producer Price Index for Goods

SPPI	Producer Price Index for Services
CPPI	Producer Price Index for Construction

Executive Summary

During a previous Twinning project running from 2014 to 2016 the three institutions in BiH set up pilot projects to start data collections and index calculations for a SPPI and a CPPI in BiH. The pilot indices have been compiled and calculated in different IT-applications across the two price statistics (CPPI and SPPI) and the three institutions (BHAS, FIS and RSIS) resulting in six different compilation-systems.

The purpose of this component was to investigate the possibilities and criteria for setting up an IT-application that could be used across the two price statistics and three institutions. Furthermore a plan for development of a possible common IT-application should be drawn up and agreed upon.

During this mission a number of conclusions and agreements were reached:

- It was concluded that the prerequisites and criteria for developing a common IT-application, for use in both SPPI and CPPI across all three institutions are fulfilled.
- It was hence decided to continue the project of developing a common IT-application.
- The set of functionalities was discussed and agreed upon.
- A conceptual model was discussed and agreed upon.
- A project- and timetable was presented and agreed upon.
- It was agreed to base the system on SQL Server database using Visual Basic for the user interface.
- Given the very tight schedule of the project it will only be possible to develop a basic system, but it should anyway be a substantial improvement compared to compiling the statistics in Excel.
- It was strongly recommended to use the same index calculation methods (index formulas) across all sectors/industries within the same statistic (e.g. in the SPPI).

1. General comments

This mission report was prepared within the EU Twinning Project ” Support to the reform of the statistics system in Bosnia and Herzegovina”. It was the first mission to be devoted to Activity 2.C: an IT application development for producer prices within Component 2: Business Statistics of the project.

The purposes of the mission were:

- Discussion on criteria for development of common IT application for CPPI and SPPI
- Plan for development of common IT application for CPPI and SPPI

The consultants would like to express their thanks to all officials and individuals met for the kind support and valuable information which they received during the stay in Bosnia-Herzegovina and which highly facilitated the work of the consultants.

The views and observations stated in this report are those of the consultants and do not necessarily correspond to the views of EU, BHAS, FIS, RSIS and Statistics Denmark.

2. Assessment and results

During a previous Twinning project running from 2014 to 2016 the three institutions in BiH set up pilot projects to start data collections and index calculations for a SPPI and a CPPI in BiH. The project goals of that project were fully met and pilot price indices for both the SPPI and CPPI have been

calculated. Furthermore the CPPI from the FIS has already been published. However the pilot indices have been compiled and calculated in different IT-applications across the two price statistics (CPPI and SPPI) and the three institutions (BHAS, FIS and RSIS) resulting in six different compilation-systems. The purpose of this component is to investigate the possibilities and criteria for setting up an IT-application that could be used across the two price statistics and three institutions. Furthermore a time plan for development of a possible common IT-application should be drawn up and agreed upon.

A common IT-application for price indices in Denmark

Statistics Denmark is already using a common IT-application across five different price statistics (CPI/HICP, PPI, SPPI, Construction cost index and CPPI). The table in annex 3 presented to the local experts shows some of the similarities and differences across these five price statistics that are being compiled and calculated in the same IT-application. The Danish case clearly highlights that it is possible to construct a common compilation-system across different price statistics.

Differences and similarities between the CPPI and SPPI in BiH

In annex 4 the Danish experts have together with the experts from BHAS, FIS and RSIS filled in a table similar to the table in annex 3 but for the SPPI and CPPI in BiH. This shows that there are many similarities between the two price statistics. The process also led to the identification of requirements, wishes and suggestions regarding a common IT-application that could be used for the compilation and calculation of the SPPI and CPPI in all three institutions in BiH.

Requirements for a common IT-application across the SPPI and CPPI in BiH

The Danish price index experts have together with the experts from BHAS, FIS and RSIS discussed the requirements needed for a common IT-application that can handle the compilation and calculation of a CPPI and SPPI in all three institutions in BiH. These necessary criteria and requirements have been drawn up under the different headings below, that can be grouped in four (six) head groups; Data entry and editing, Calculations, System Control and Output/Input. Besides that there is also some requirements regarding some general criteria for the system and hierarchies.

GENERAL

Frequency:

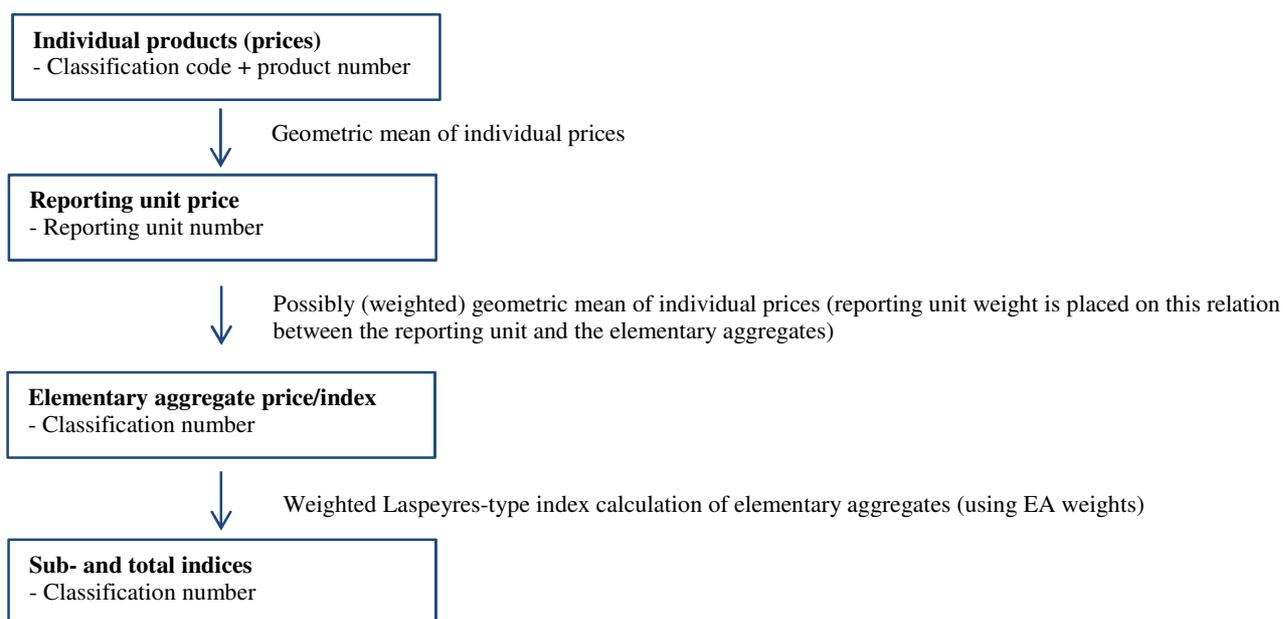
- All SPPI and CPPI indices are produced quarterly.
- The system should be able to handle different sampling and calculation periods.
- Before every new sampling period, the system should be set up with a clear sheet for typing and editing new prices (change of period).

HIERARCHIES

Classification (incl. code lists):

- The system should be able to handle different classifications and code lists.
- The system should be able to handle classification codes of at least five digits (any number of digits).
- The classifications are used to identify elementary aggregates and sub-indices.
- Individual products (services and construction works), below the elementary aggregate level, are identified by classification codes or codes from the code lists and a simple (consecutively) product number (1,2,3,...,n). In this way it is easy to insert a new individual product or terminate one.

The system must be able to handle the following relations through the hierarchies:



DATA ENTRY AND DATA EDITING

Reporting units (RU):

- The system should have a table of reporting units.
- The reporting unit should be identified by a reporting unit code (e.g. the one from the business register)
- For each reporting unit the system should have information regarding name, address, phone number, e-mail address and contact person.
- It should be possible for the statisticians to insert new reporting units and to terminate existing reporting unit from the table of reporting units (the sample).
- When a reporting unit is terminated and removed from the table/sample, the system should save information about the enterprise, and it should be possible for the Statistician to make a comment about why the reporting unit is terminated.
- When a reporting unit is terminated and removed from the sample, the individual products/prices belonging to this reporting unit should also be terminated (no longer be part of the calculations).
- It should be possible for the statistician at all time to see a status about the current data collection, i.e. number of edited and accepted prices, number of prices in process of being edited and number of missing price reports.

Weights for reporting units:

- The system should be able to handle weights for reporting units. This includes securing that weights sum to 100 within an elementary aggregate, also when a reporting unit is terminated and removed from the sample or when a new one is added to the sample. The weights are registered by the statistician. If there are not registered any weights, the system just makes an equal weighting.
- These should be placed on the relation between the reporting unit and the elementary aggregate (the reporting unit price).

Individual products (prices):

- The system should have a table of individual products (and their prices).

- It should be possible for the statisticians to see which individual products (/prices) that belongs to which reporting units.
- Individual products (as mentioned under classifications) are identified by classification codes or codes from the lists and a simple (consecutively) product number (1,2,3,...,n). In this way it is easy to insert a new individual product or terminate one.
- It should be possible for the statisticians to terminate and insert new individual products.
- Individual products (and their prices) should always belong to a reporting unit and an elementary aggregate. The system must insure this (e.g. through the consistency control).
- It should be possible for the statisticians to type in prices for actual and previous periods and already at this stage to do some editing of prices (automatically – see “Outlier/inlier detection” and manually).
- It should be possible (desirable) for the statisticians to register the type of quality adjustments that are made for the individual products/prices every period (the actual quality adjustment decision has to be made manually outside the system by the statistician¹).

Types of prices:

- It should be possible for the statistician to mark which kind of price that is reported in a given period for the individual products; transaction price, unit value price etc.

Price collection methods:

Prices are collected through paper questionnaires or excel sheets. The prices should therefore be typed in the system. See also ”No of price observations”.

- The system should be able to export information required to make the questionnaire, that is
 - Which individual products (classification code, product code and product description) that each reporting unit should report prices for.
 - Information about the reporting unit.
 - Etc.

Outlier/inlier detection (editing):

- The system should be able to do some automatic outlier detection. This could be confidence intervals calculated on the basis of all price changes for the last year within each elementary aggregate. Price changes outside these intervals must be flagged as possible errors that should be handled by the statisticians.
- The system should be able to do some automatic inlier detection that is, a check for unchanged prices during a given number of periods.
- The statistician must be able to set the following status during the process of editing a price observation;
 - “In process”
 - “Finished” when the editing is done and the price change (or lack of price change) is accepted

The status will be placed on the level of the reporting unit.

¹ See also annex 5

CALCULATIONS

Elementary aggregates (EA):

- The system should have a table of elementary aggregates.
- It should be possible for the statisticians to terminate and insert new elementary aggregates in this table (this requires a weight update).
- The system must ensure that (new) elementary aggregates have a starting index value (e.g. through the consistency control), to be able to calculate the first index value. This could just be set to 100 (This will only occur when updating the weights).
- It should be possible for the statisticians to connect elementary aggregates to sub-indices.
- It should be possible for the statisticians to see which enterprises and individual products that belong to each elementary aggregate.
- It should also be possible for the statisticians to connect reporting units and individual products to an elementary aggregate.
- The system must not allow for an elementary aggregate to be empty (e.g. through the consistency control).
- If an elementary aggregate no longer has any individual products (and therefore no prices) connected (it is empty), and it is not possible to find other individual products, maybe from another reporting unit, the system should be able to handle this, by making it possible to follow the price development from another elementary aggregate (until the next weight update where the elementary aggregate can be deleted).
- The system must ensure that all elementary aggregates are connected to a sub-index, e.g. just the total index (e.g. through the consistency control).
- The system should be able to calculate elementary aggregates for each institution, that can be weighted together (another set of weights?)

Elementary aggregate weights:

- The system must be able to handle elementary aggregate weights. These are also used for sub-indices. The weights should be in turnover values or similar. In this case the system should convert them to percentage shares, depending on which sub-indices the elementary aggregates belongs to.
- The system must ensure that every elementary aggregate has a weight (e.g. through the consistency control).
- The system must ensure that all weight shares sum to 100 within the total index (e.g. through the consistency control).

Weight updates (EA weights):

- The system must be able to handle weight updates at different time intervals.

Index formula elementary aggregates:

- From individual price observations to elementary aggregates the system is suggested to use the Jevons (geometric mean) index formula:

$$I_{0t}^{Jv} = \prod \left(\frac{p_t^i}{p_0^i} \right)^{w^i} = \frac{\prod (p_t^i)^{w^i}}{\prod (p_0^i)^{w^i}}, \sum w^i = 1$$

Index formula higher level index:

- From elementary aggregates to higher level indices the system must use the Laspeyres-type index (with fixed weights):

$$I_{0t} = \sum w_b^j \cdot I_{0t}^j, \quad \sum w_b^j = 1$$

Index reference period:

- The Statisticians must be able to define the base year in the system.
- The system must be able to handle changes in base year (i.e. correct chaining and correct index reference period).

SYSTEM CONTROL

Consistency control:

- The system should be able to make a number of system controls to check that everything is ready for calculations.

OUTPUT/INPUT

Output tables:

- Export of individual products and prices
- Export of EA index tables for visual inspection
- Export of sub- and total indices for visual inspection and publication
- Export/import of data from all institutions
- A report of quality indicators (response rate etc.)

The above mentioned requirement can be grouped as follows:

Most important features:

It is the assessment of the experts that the following groups are the most important:

- Data entry and editing
- System Control (consistency control)
- Calculations
- Output

The following features are also assessed to have a high priority:

- It should be easy to establish and maintain hierarchies (controlling the flow of data)
- It should be easy to maintain individual products by reporting units (adding, terminating and changing) and to maintain reporting unit weights.

Other features are:

- It should be easy to maintain response units (adding, terminating and changing e.g. contact person etc.)
- It should be easy to maintain elementary aggregate weights

Conceptual model

The Danish IT expert has translated the criteria described above into a conceptual model.

The purpose of this conceptual model is to identify the different types of data objects needed in a system to calculate price indices. It is not a data base model, but kept on the conceptual level. The primary purpose of the model is to be able to demonstrate how different tasks related to the use of the system could be managed, mainly by identifying the data objects that would be used and/or changed for each task.

The conceptual model can be seen in annex 6 and was accepted by the local experts.

Plan

The actual development work after this mission will not involve Danish price index experts. The development work will from hereon be led by the Danish IT expert, who has suggested the following plan for the experts from BHAS, FIS and RSIS.

When	What	Who
April	Detail planning Data base model designed User Interface designed Functionality described	Methodologists
May	Programming <ul style="list-style-type: none"> - Base menu - Data Entry screen(s) - Consistency controls - Calculations 	Methodologists Ad hoc meetings
September	Programming <ul style="list-style-type: none"> - Export to Excel - Creation and maintenance of hierarchies - Maintenance of individual products - Maintenance of Response Units Maintenance of Weights - Data Exchange 	Methodologists Ad hoc meetings
October	Final testing Handing over to ... (for translation) Workshop on DB structure etc.	Methodologists IT-staff

The experts from BHAS, FIS and RSIS accepted this first draft.

Development tool

It has been discussed and decided that the system will be based on a SQL Server database. The user interface will be created using Visual Studio in Visual Basic.

The application (user interface) will be created in English, but VB makes it easy to translate to any language in a way, that language is selected on start up.

Excel may be used for parts of the system.

It should be noted, that it is intended to create a database for each survey and for each institution. They will all have the same structure, but the content will be different, but there will only be one application.

To facilitate the work it would be very helpful, if it is possible to reuse as much as possible from the work done by FIS and RSIS in terms of having started the development of an IT system. The BiH IT expert responsible for these systems expressed their willingness to deliver the necessary materials.

It was suggested by the Danish experts that when the common system has been developed, it will be useful to write a cookbook on how to handle different options in the system. This should mainly be written by statisticians at BHAS.

Pros and cons for using a common system

At the end of the mission pros and cons for using a common it system was discussed. These can be seen in Annex 7. This discussion led to the decision, that a common system for SPPI and CPPI for all three institutions should be developed.

With the resources available, it will only be possible to develop a basic IT system with the most necessary features. In addition to this, it was noted by one of the local experts that a primary goal for the system, must be that it should be a substantial improvement to an Excel system.

3. Conclusions and recommendations

The following conclusions were reached:

- It was concluded that the prerequisites and criteria for developing a common IT-application for use in both in SPPI and CPPI across all three institutions are fulfilled.
- It was hence decided to continue the project of developing a common IT-application.
- The set of functionalities was discussed and agreed upon.
- A conceptual model was discussed and agreed upon.
- A project- and timetable was presented and agreed upon.
- It was agreed to base the system on SQL Server database using Visual Basic for the user interface.
- Given the very tight schedule of the project it will only be possible to develop a basic system, but it should anyway be a substantial improvement compared to compiling the statistics in Excel.

The following recommendation was made:

- It was strongly recommended to use the same index calculation methods (index formulas) across all sectors/industries within the same statistic (e.g. in the SPPI)

The following actions are needed for moving forward and for preparation of the next mission:

Action	Deadline	Responsible person
Decision on the exact index calculation methods (formulas) at the level below elementary aggregates needs to be decided	Before next mission in April 2018	Local experts
Re-examine the conceptual model and the required functionality stated in the final mission report checking for completeness and errors	Before next mission in April 2018	Local experts
Consider potential data validation rules (editing) to be included in the IT-application	Before next mission in April 2018	Local experts
Prepare first draft of DB-model	Before next mission in April 2018	Søren Netterstrøm
Prepare first drafts of the forms for the IT-application	Before next mission in April 2018	Søren Netterstrøm

Annex 1. Terms of Reference

Terms of Reference

EU Twinning Project BA 15 IPA ST 01 17

Component 2: Business Statistics

05-09 March 2018

Activity 2.C.1: IT application development for producer prices I

1. Result for the activity

- Plan for development of common IT application for CPPI and SPPI agreed

2. Purpose of the activity

- Discussion on criteria for development of common IT application for CPPI and SPPI
- Plan for development of common IT application for CPPI and SPPI

3. Expected output of the activity

- Possibilities for development of common IT application for CPPIs and SPPI discussed
- Criteria defined
- Plan for development of common IT application for CPPI and SPPI agreed

Annex 2. Persons met

Agency for Statistics of BiH (BHAS)

- Fahir Kanlić, BHAS
- Dženita Babić, BHAS
- Anita Brković, BHAS
- Alen Bajramović, BHAS
- Lejla Borovina, BHAS
- Ševala Abaz, BHAS, IT
- Belma Mesihović, BHAS, IT

Institute for Statistics of Federation of BiH (FIS)

- Edina Dulic, FIS
- Nusreta Imamovic-Kaljanac, FIS
- Merima Hadzalic, FIS
- Ahmet Fazlic, FIS, IT
- Amel Sikiric, FIS, IT

Institute for Statistics of Republika Srpska (RSIS)

- Jelena Glamočika, RSIS
- Aleksandra Đonlaga, RSIS
- Želimir Radišić, RSIS
- Nataša Gojković, RSIS, IT

MS Experts

- Janni Stavad, Statistics Denmark
- Martin Birger Larsen, Statistics Denmark
- Søren Netterstrøm, Statistics Denmark

Twinning Project Administration

- Katja Møller Hjeltvang, RTA
- Đemka Šahinpašić, RTA Assistant
- Haris Imamovic, Interpreter

Annex 3: Similarities and differences across the five price statistics at Statistics Denmark

Denmark:	CPI/HICP	PPI	SPPI	Construction cost index	CPPI (renovations)
Frequency	Monthly	Monthly	Quarterly	Quarterly	Yearly
Classification	COICOP	CN and NACE	NACE and CPA	Nace and own classification	NACE and CPA
No of outlets/firms (RU)	1.600	1.000	350	NA	350
No of price observations	25.000	6.300	1.700	NA	3.000
Types of prices	Consumer prices	- Producer prices - Import prices	- Producer prices - Unit values (Turnover value/quantity) - Profit share * wage	Indices from PPI and wage indices.	Producer prices
Price collection methods	Questionnaires (digital), internet page collections, price collectors and scanner data	Questionnaires (digital), internet page collections	Questionnaires (digital), internet page collections	From databases	Questionnaires (digital)
No of elementary aggregates	450	1.050	45	1.000	55
Index formula elementary aggregates	Jevons (geometric mean)	Jevons (geometric mean)	Jevons (geometric mean)	Jevons (geometric mean)	Jevons (geometric mean)
Index formula higher level index	Laspeyres-type	Laspeyres-type	Laspeyres-type	Laspeyres-type	Laspeyres-type
Index reference period	2015=100	2010=100	2010=100	2015=100	2015=100
Weight sources	National accounts and	National accounts	National accounts	Survey data	Expert data

	HBS				
Frequency of weight updates	Every year	Every five years	Every five years	Not regularly	Every five years?
Outlier detection (editing)	Extremes (HB-method) and inlier detection				
System controls	Consistency control				
IT program used	Oracle and Oracle forms				

Annex 4: Similarities and differences across the two price statistics at the three institutions FIS, RSIS and BHAS

BiH:	CPPI - BHAS	CPPI - RSIS	CPPI-FIS	SPPI-BHAS	SPPI-RSIS	SPPI-FIS
Frequency	Quarterly	Quarterly	Quarterly	Quarterly	Quarterly	Quarterly
Classifications and code lists used	Own classification and NACE rev2	Own classification and NACE rev2	Own classification and NACE rev2	NACE rev2 and CPA	NACE rev2 and CPA	NACE rev2 and CPA
No of outlets/firms (RU)	10	93	112	11	45	71
No of price observations	419	2476	2564	20	80	120
Types of prices	Market prices excl. VAT	Market prices excl. VAT	Market prices excl. VAT	Transaction prices (B2B), contract prices and unit values	Transaction prices (B2B), contract prices and unit values	Transaction prices (B2B), contract prices and unit values
Price collection methods	Questionnaires	Questionnaires	Questionnaires	Questionnaires	Questionnaires	Questionnaires
No of elementary aggregates	30	30	30	100	100	100
Index formula elementary aggregates	Jevons (geometric mean)	Jevons (geometric mean)	Jevons (geometric mean)	Jevons (geometric mean)	Jevons (geometric mean)	Jevons (geometric mean)
Index formula higher level indices	Laspeyres-type	Laspeyres-type	Laspeyres-type	Laspeyres-type	Laspeyres-type	Laspeyres-type
Index reference period	2015q4=100	2015q4=100	2015q4=100	2015=100	2015=100	2015=100
Weight sources	Bill of quantities	Bill of quantities	Bill of quantities	Survey data	Survey data	Survey data
Frequency of weight updates	Every 5 years	Every 5 years	Every 5 years	Annually	Annually	Annually
Outlier detection	Outliers (percentage)	Outliers (percentage)	Outliers (percentage)	Outliers (percentage)	Outliers (percentage)	Outliers (percentage)

(editing)	range) and inliers	range) and inliers	range) and inliers	e range) and inliers	e range) and inliers	e range) and inliers
System controls	?	Manual and automated controls	Manual and automated controls	?	?	?
IT program used	Excel	Serbian system	Acces	Excel	Excel	Excel

Annex 5: Quality adjustments

During the discussions a question arose regarding whether it would be possible to (fully) automate quality adjustment procedures in the common IT-application.

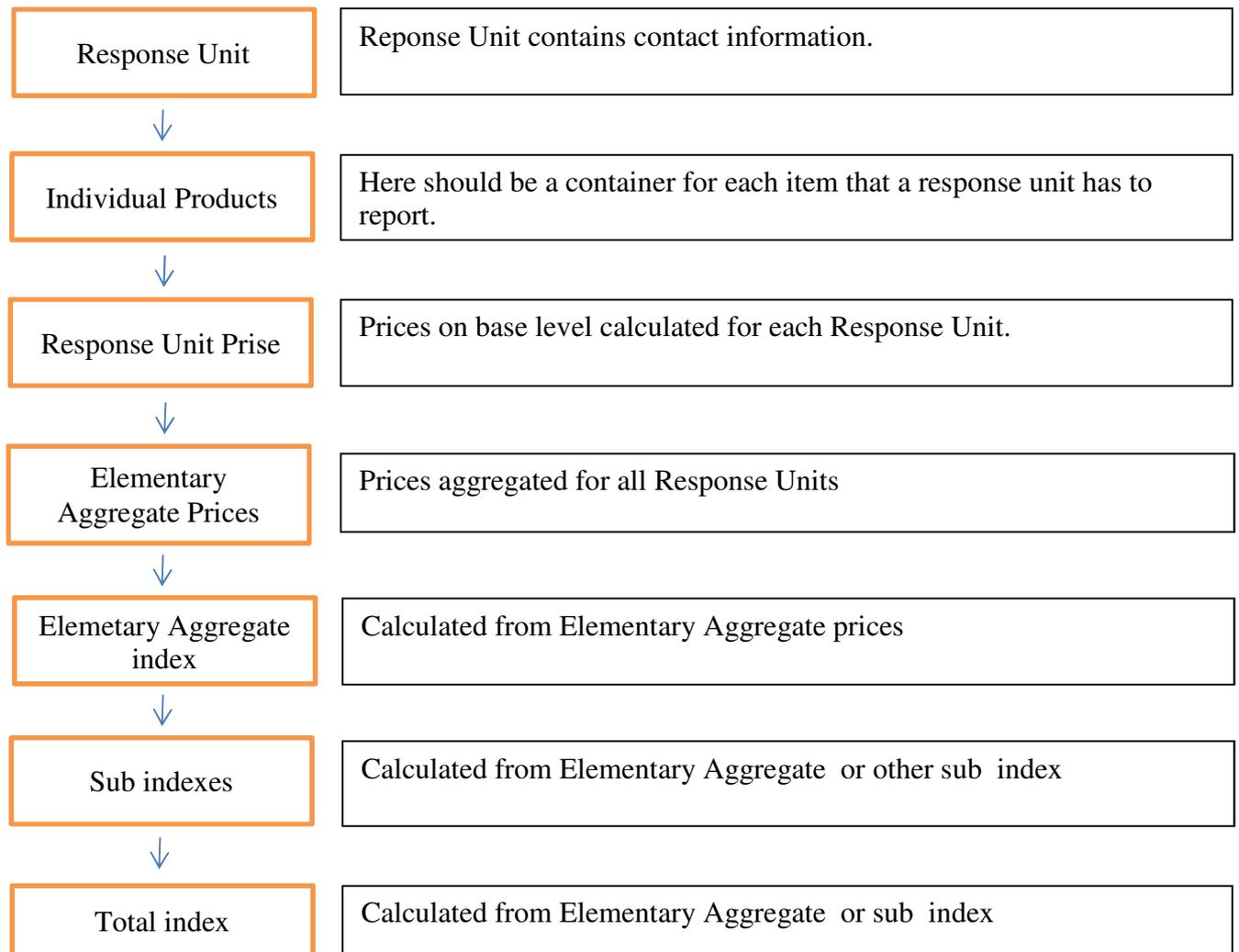
Whenever a product in the sample of a price index is replaced with another product a quality adjustment has to be made between the price of the product leaving the sample and the price of the product entering the sample. The quality adjustment should ensure that price differences reflecting possible quality differences between the two products do not enter the price index calculations so that the indices only reflect “pure” price changes. In general methods for quality adjustment can be divided into two groups; implicit methods (e.g. direct comparison or bridged overlap) and explicit methods (e.g. hedonics or expert evaluation).

In the price indices compiled by Statistics Denmark implicit methods are being used on a case by case basis for almost all product groups. The same is true for other European statistical institutes where implicit quality adjustment methods are being used for most product groups across most of the different price indices.

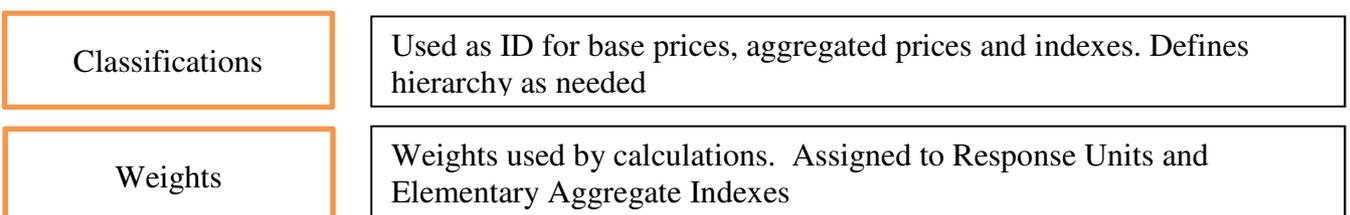
In the view of the Danish experts it is not possible to fully automate the quality adjustment procedures used with the possible exception of product groups where it is possible (in terms of available data and resources) to use hedonics for quality adjustments. But even in the case of hedonics it is very important that statisticians with at least some knowledge of the product group in question very regularly overviews the results of the hedonic quality adjustments.

The decision on implicit quality adjustment methods to be used in case of product replacements should be made or at least be checked on a case by case basis by a statistician. One reason is that it is almost always not possible to quantify all relevant characteristics of a product relevant for the (consumer) evaluation of the quality. In effect this makes some subject evaluation of quality differences unavoidable. There would be a risk that fully automated quality adjustment procedures run of track and in effect creates less reliable price indices if not handled or checked by a statistician on a case by case basis.

Annex 6: Conceptual model



Hierarchical structures!



Annex 7: Pros and cons of developing a common it system

Pros and cons of having a shared IT-application/system for the calculation of SPPI's and CPPI's across the three institutions (BHAS, FIS and RSIS):

Pros:

- Statistics produced by the same system are with certainty fully comparable across the three institutions
- Cheaper to develop one system instead of several
- Cheaper to maintain one system instead of several
- New functions developed for use in one statistic/institution could possibly be used with benefit in other statistics/institutions that uses the system

Cons:

- If IT-knowledge of the system is not present in each institution there could be dependence on IT-staff from other institutions in case of e.g. a system crash
- The three institutions have to agree on the design e.g. in case of development of new functions in the system