

Integrated audit aspects in the Hungarian GDP compilation

Bálint Murai¹

¹ *Hungarian Central Statistical Office, Budapest, Hungary, balint.murai@ksh.hu*

Abstract

The paper describes the multi-year modernisation project of the National account's compilation system in Hungary. One of the main driving forces behind the re-engineering of the annual statistical process was to fully harmonise it with the EU's audit standard. As a result of the redesigned system data are produced, stored and validated in the framework of the Process Table which is the tool for the direct verification exercise of the GDP/GNI data in the EU. The paper presents the challenges and the implementation of this giant modernisation project, summarises the main achievements and the outline of the new compilation methodology.

Keywords: National Accounts, GDP compilation, Process table, GNI audits

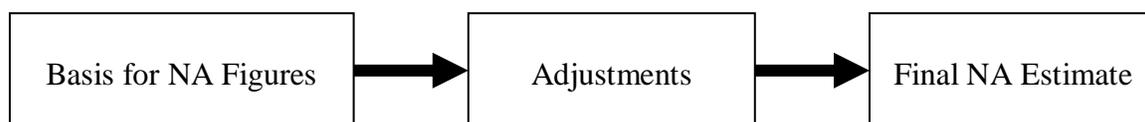
1. The Process Table

The Process Table (PT) is a standardized table set developed by EUROSTAT to standardize and facilitate the GNI audits in the member states. While compiling the GNI inventory, which is the detailed methodology of the National accounts compilation, member states have to compile the PT for the reference year of the GNI inventory.

The compilation of the National Accounts is a very complex process which draws together many different types of data and balances them against one another to produce macroeconomic aggregates, such as Gross Domestic Product (GDP) or Gross National Income (GNI). The NA compilation process starts with different types of data sources as input. The process includes a series of adjustments to the basic data to make them consistent with each other and bring them in line with the National Accounts concepts and methods (see Figure 1). (EUROSTAT, 2014a)

The aim of PT is to give an overview on the characteristics of the national account, as well as on data sources and estimation methods for the GDP and GNI calculations.

Figure 1: Structure of the National Accounts compilation process



Based on the recommendations of the European Court of Auditors (ECA) for the comprehensive quality check of the data a special, so called “Direct Verification” exercise has been introduced in the GNI audit missions. During the “Direct verification” exercise the PT is used as the main tool by the auditors to check the quality of the data and its compliance with the Inventory. The role of the PT in the upcoming round of verification of GNI for own resource purposes is likely to increase (EUROSTAT, 2014a p. 4.).

Figure 2. The structure of the PT

GDP & GNI components	Basis for NA Figures				Adjustments				Final estimate
	Surveys and Censuses	Administrative Records	Combined Data	Extrapolations and Models	Data Validation	Conceptual	Exhaustiveness	Balancing	
Production									
Expenditure									
Income									
GNI									

2. Driving forces

Until recently the Hungarian GDP data, especially the largest, and most complex sector, the non-financial sector (S11) had not been compiled in the framework of the PT. The compilation process for S11 was a result of development started mainly in the mid 90's when the PT had not yet been implemented. In the old system for S11 there used to be two types of adjustment categories in the transition from bookkeeping data to the published national account figures: There was a SAS software based compilation algorithm, which made the corrections at individual (company) level and the results were uploaded in ORACLE table. The other types of adjustments were calculated directly in Excel files based on different data sources and expert estimates. At the end the results of these two sources were put together and stored in excel files. The structurally different Excel files and the several links within them made the system difficult to overview. Automatic conversion of National Account data to the categories and logic of the PT was not possible, the transformations had to be done manually, which was extremely time consuming, demanding significant human resources.

As the importance of the GNI Audits and the role of PT had been gaining importance Hungarian National Account had to decide how to adapt to the new conditions. After careful consideration the decision was made to modernize, re-engineer and redesign the compilation methodology for S11 in a way which fully incorporates the PT in the compilation process. Besides the methodological work it involved the development of a new IT system (architecture, data processing scheme and multidimensional database). The restructuring of the data process will not have any effect on the data themselves. The new system has to be capable of further extension, so later on the other sectors (S12-S15), and other accounts of non-financial accounts can also enter the system. But, in the meantime, to the estimation of the GDP by production approach it was necessary that the other sectors' production accounts figures, and taxes less subsidies on products could be uploaded into the system from Excel files.

The development has had two main advantages. First, no PT-transformation mechanism will be necessary in the future. Second, with a built-in audit function the PT categories could be

used for validation, right in the compilation and finalization process of GDP. This will help national accountants to better control and supervise its data production system internally, as well as to facilitate Eurostat's audits of GNI data.

Besides the aims above, there were a couple of other goals to be achieved with the modernization.

- It is expected that the new system will make the processing of national accounts data more structured, transparent and auditable.
- The old calculation process required several weeks of work that was necessary to be shortened to let more time for quality checks.
- To retrieve data a fast and user friendly query system is needed which can deal with multi-dimensional datasets.
- As it is also aimed to cover other sectors and accounts by the new system, the whole IT application would be developed (capacity, dimensions, special characteristics of institutional sectors etc.) in a way that makes it capable for future extensions.
- For the SUT (Supply and Use table) balancing transparent GDP compilation is needed, where the input sources and the different types of adjustments are explicitly presented. This approach is equivalent to the idea behind the PT.
- It was also required from the system to harmonize with the current IT and META infrastructure of the HCSO.

3. Modernisation steps

3.1. Concept and preparation

The basic concept had long been in existence before the IT implementation could be started, because the preparation needed extra amount of highly trained human capacity, while ensuring the day-by-day duties. The preparation phase included several difficult tasks:

- Rethink the data processing
- Develop the design of the new structure

- Rationalize and streamline the calculations and reassure their internal consistency
- Create dimensions
- Define the new indicators, their computation algorithms and hierarchy
- Complete documentation for programming
- Harmonize classification of nomenclatures and name convention for input databases

3.2. Architecture

The board of HCSO decided that all the actions planned in the framework of this project should be implemented within the office using existing IT resources integrated into the HCSO's IT structure. The system has to be developed in house by the national accountants and IT experts. After discussions with database architecture experts it was decided that the system would be built on two pillars: data processing and data access.

3.2.1 Data processing: Integrated Data Processing System (EAR)

The HCSO considers the EAR as a key important system in the business processes with a lot of functionality to support statisticians in the design and management of the data processing tasks. The purpose of the EAR is to support the data processing activities in general using standard protocols for any statistical domain. One of the main outcomes of the EAR is the increase in quality of data processing tasks due to standardization, full integration with the HCSO metainformation system and the automatically generated documentation of data processing activities.

The functionality of the EAR is twofold:

- framework – it supports the design and the management of the data processing activities.
- set of standard elements – EAR procedures/functions like Lego cubes underlie the whole processing.

3.2.2 Data access: Data warehouse

The HCSO Data warehouse is a collection of statistical data in a multidimensional format for fast data access, with built-in OLAP functions. The Data warehouse has an external part, called Dissemination database where HCSO disseminates its statistical database.

3.2.3. The input database

A harmonised and consistent database, the Common Database (KAB) is used as the main data source of the GVA compilation in non-financial corporations sector. This database is operated by the Business Statistics Department, where data from integrated economic statistical surveys are supplemented with corporate tax returns data at individual (enterprise) level. The integrated economic statistical surveys contain every basic variable to compile the value added at individual level. Tax returns are used to validate, impute and gross up data. The KAB is the basis of Structural Business Statistics (SBS) and the non-financial corporations sector data of national accounts. Data from KAB are classified into ‘Combined data’ in Process Table.

3.3. The implementation

Colleagues of three departments of HCSO took part in the implementation: National accounts department, IT department (EAR team, Data Warehouse team) and Methodological department (META experts). After creating the project team, the work started by training of the team members. First, NA compilers delivered presentations to the IT staff about the calculation process and the expectations towards the system. Concepts, definitions, nomenclatures, methodology and steps of implementation had to be clarified. The creation of the IT environment and the adoption of the new system into the existing system of HCSO was managed by the IT department. Next, IT experts trained the statisticians to prepare them to program the whole compilation process in a user friendly IT environment.

3.3.1. Dimensions

In the concept of multidimensionality the following five dimensions have been established:

3.3.1.2 Time periods

Due to the revisions, the figures of a given year can differ from each other, depending on the year when they were calculated. The following table demonstrates the relation between the reference year (the period of time the data refers to) and the calculation year (the year when the calculation is done):

Table 1. Dimensions according to time periods

Calculation year	Reference year		
	2012	2013	2014
2013	2012 preliminary		
2014	2012 semi-final	2013 preliminary	
2015	2012 final	2013 semi-final	2014 preliminary

3.3.1.2 NACE Rev2 codes

The current price calculations are made according to the NACE Rev. 2 nomenclature in 4 digit level, but there is a possibility to present figures on more aggregated levels.

3.3.1.3 PT Indicators

There are more than one hundred fifty PT indicators, of which two-thirds are primary indicators, the others are derived indicators, which are calculated from the primary indicators. In developing the PT indicator codes, the requirement to the code was to unambiguously identify the sector, the sequence of the accounts, and the PT category.

3.3.1.4 Types of enterprises

Eleven types of enterprises are distinguished according to their characteristics, information sources and available individual data for S11 sector in Hungary. (e.g. Enterprises using

double-entry bookkeeping, Enterprises, whose business year differs from calendar year, Non-profit institutions classified into the non-financial corporations sector, etc.)

3.3.1.5 Main aggregates of the NA by production approach

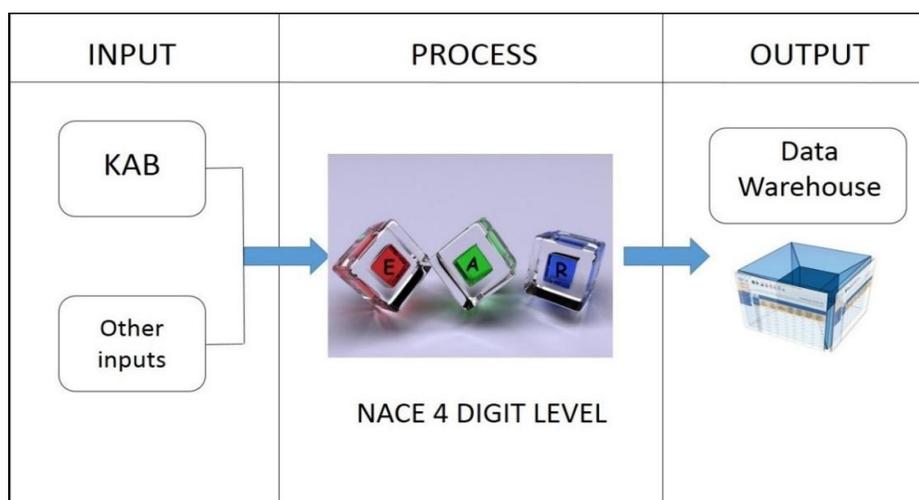
This dimension consists of the three main figures of the Production Account. These values show the contribution of a given indicator to the value of Output, IC and GVA.

4. The new system

To create a new architecture, a new indicator set and hierarchy structure had to be established with the corresponding nomenclatures. The restructuring of the data process did not have any effect on the data themselves. Classification of nomenclatures and name convention for input databases were also harmonized. Correlations, formulas had been updated and rationalised. The whole process had been documented for the programming procedure. The documentation includes the inputs used in the estimation, the filtering options and the calculation algorithms.

The following chart shows the data estimation process in the EAR-Data warehouse environment.

Figure 2. GVA compilation process for S11



From the input data sets EAR computes the PT indicators at 4-digit level of NACE, which are transferred to the Data warehouse. From the Data warehouse there are two options to derive data for validation and publication. There is a built-in query module in the Data warehouse or the data can be accessed directly from Excel using the ORACLE Smartview add-in extension.

5. Results and future activities

All of our objectives undertaken for this project have been achieved. Calculations have become faster, transparent and the procedure lasts days instead of weeks. Faster calculations and user-friendly data access gives more time for analysis and data check. The system has become safer, as data are stored in database. The run of GVA during the compilation is easier to follow. PT framework is used in the validation process and is generated automatically every year for sector S11. The modernisation had an extra advantage, as the compilation process shortened, we could apply the new compilation method not only to the final but also for the semi-final estimates. This was a huge step forward, because in the past the semi-final estimate had to be based on quarterly benchmarking technics, which resulted in larger revisions between the semi-final and final estimates. Thanks to the modernisation, as both the final and the semi-final estimates are based on individual data the amount of revision has become significantly lower.

To get familiar with the know-how of the system, for the colleagues being involved hands on training courses have been organized in smaller and larger groups.

Since the establishment of the new, EAR-Data Warehouse system the following features have also been included: (between the time period from 2013-2016)

- The whole time series for the production account were back casted to 1995.
- The previous year calculations were also streamlined and re-designed in order to be put in the system, so the constant price estimations could also be incorporated in EAR since 2013 back to 1995.

- The next sector to be incorporated into EAR fully in line with the PT was Household sectors (S14). In order to reach this goal, as for sector S11, the complete calculation methodology of the production account for S14 has been deeply analysed, restructured and re-designed by the categories of the PT. The redesigned estimation of the Household sector has been included in the system back to 1995.
- The current development is the introduction of the sixth dimension for S11 in the system, i.e. the subsector of ownership (public, national private or foreign) for Gross Value Added. The completion of the task is planned by the end of this year.
- The next planned development is the generation of income account for S11 to be entered in the system.

References

EUROPEAN COMMISSION [2013]: *European System of Accounts – ESA 2010*. Luxembourg.

EUROSTAT [2004]: *GNI Process Tables Analysis PT/01*. Luxembourg.

EUROSTAT [2005]: *Eurostat's Tabular Approach to Exhaustiveness Guidelines GNIC050*. Luxembourg.

EUROSTAT [2009]: *Guidelines for the direct verification exercise in the framework of the 2007-2010 verification of member states GNI inventory GNIC/137Rev.1*. Luxembourg.

EUROSTAT [2014a]: *Process Tables Compilation Guide GNIC/281*. Luxembourg.

EUROSTAT [2014b]: *ESA 2010 GNI Inventory Guide GNIC/280*. Luxembourg.