

# Chapter 5

## Conclusions and Further Research

### 5.1 Conclusions

There are two features that this dissertation aims to study. The first and most apparent target is to describe and compare theoretically some of the estimation procedures discussed or used by Eurostat and to make proposals, wherever this is possible, for the adoption of better practices. The second target is less evident but also of great importance. This target concerns the over space harmonisation of the estimation procedures. The idea of harmonisation mainly refers to those procedures that are being applied at the Member State level.

As far as the first target is concerned, in this dissertation we have described three estimation domains. These are: the micro-aggregation techniques for producing confidential data, the backward calculation methods for constructing homogeneous time series and some aspects of the sampling techniques discussed by Eurostat and applied by the Member States.

For the micro-aggregation techniques we have concluded that there is a need for the adoption of more complex methods like the improvement of the Hanani's algorithm and the modified Ward's criterion. These methods must be incorporated into the micro-aggregation software that Eurostat develops at this time point. Furthermore, we have argued that an enrichment of the evaluation criteria is required. This means that more criteria aim at assessing the confidentiality aspect of the modified data set as well as the violation of the structure of the original data set must also be incorporated into the new micro-aggregation software. Criteria of both types are the indicator of data perturbation, the examination of the correlation coefficient of the initial and the modified data set and the further processing ability. Finally, we have proposed the adoption of a model choice approach for those that are being involved into the micro-aggregation process. This

means that they have to apply different micro-aggregation techniques and different evaluation criteria in order to arrive at an “optimal” solution. Using the term “optimal” we mean a solution that is acceptable, from the confidentiality point of view, by the Member States and useful, from quality point of view, for the independent analysts.

The second domain of estimation concerns the backward calculation techniques. In this chapter we have described generally the backward calculation techniques and we have mainly focused in the models applied by the Netherlands and France. We recognised that both countries apply the benchmark/interpolation technique. The difference among the two approaches is in the backward calculation of the intermediate years. While the Netherlands use a linear pattern for between years interpolation, France uses the Kalman filter algorithm. It is common belief that the best method for the backward calculation is the annual backward calculation. However, given the time and staff constraints, the Member States can resort to the benchmark years and interpolation method but only under certain quality requirements.

The third domain of estimation concerns some aspects of the sampling techniques used by the Member States. The complexity of this domain is connected with its the multinational character. As a result, the three surveys that we have chosen to describe make easier the study of this complex and multivariate domain. Two aspects of the surveys have been selected for description i.e. the sample designs and the weighting procedures. In other words, we have tried to identify the most frequently used sample designs and weighting procedures. With respect to the sample designs we have found that in general the Member States are in favour of clustered samples. More specifically, the Member States usually use multistage stratified cluster samples with probability proportional to size selection for the primary units. However, there are cases where one-stage element samples are used. These designs involve stratified samples with random or systematic selection of the units within each stratum. As far as it concerns the weighting procedures, we have found that the most frequently used are the post-stratification and the raking ratio adjustment procedures.

As it has already been stated, the second feature that this dissertation aims to describe, is the progress in the domain of the over space harmonisation of the estimation procedures. First of all, we have to stress the importance of harmonisation as an indicator

of quality. This means that the more harmonised the methods used by the Member States are, the easier the comparisons between the different results and consequently the better the evaluation of the quality of the results is.

The role of Eurostat regarding the harmonisation procedure can be viewed in two dimensions. The first dimension is this under which Eurostat revises the methodology on a specific topic, makes proposals to the Member States that have no methodology to develop one on that topic and at the same time sets internally the bases for applying the methodology. An example is the role of Eurostat in the domain of backward calculation methods. We have already said that Eurostat revises the backward calculation techniques and the existing models (Dutch and French backward calculation models) in order to be in a position to make proposals to countries that have no methodology to develop one. At the same time Eurostat acts so as to develop internally the same methodology. It is apparent that the role of Eurostat aims at the better harmonisation of the backward calculation procedures. It is also apparent that the internal development of the same methodology aims at the better evaluation of the quality of the micro results that are transmitted to Eurostat.

Under the second dimension, every Member State has already developed an estimation procedure on a topic. In this case Eurostat receives micro (individual) results which have to be assessed for their quality, and to be composed in order to produce the macro (aggregated results). It is apparent that the evaluation of data quality as well as the composition of the individual results become more difficult due to the multinational character of these estimation domains. Consequently, the harmonisation in terms of using common definitions, tools and measurement procedures is important. The role of Eurostat in this case is to study and describe the different definitions, tools and methods and to be able to propose the adoption of better and more harmonised practices. Some examples of this role of Eurostat have been studied in the chapter concerning the sampling techniques. For example, for the household budget survey and for the European Community household panel Eurostat has proposed step by step weighting procedures both for the first wave as well as for the subsequent waves (longitudinal weighting procedures). Furthermore, progress has been made in the use of common definitions and tools. For example in the labour force survey the use of common definitions for employment,

unemployment and underemployment that have been introduced by the thirteenth international conference of labour statisticians at Geneva. Also harmonisation attempts of the same type have been made for the household budget survey and for the European Community household panel. These attempts involve the adoption of common definitions for the consumption expenditures, the income of the households and for the use of a common set of variables.

Under the first dimension the role of Eurostat has more degrees of freedom. Eurostat should identify estimation domains where Member States have not developed yet a methodological approach, study the existed estimation procedures and make proposals in the direction of harmonisation.

Under the second dimension the role of Eurostat is more restricted. In this case, the study of the different methodologies plays a central role. Only by analysing the different methodologies Eurostat will be in the position to identify those cases where the heterogeneity problem is severe and to make proposals in the direction of the better harmonisation.

The present dissertation analyses the procedures in three domains of estimation. This study aims at the evaluation of the effectiveness of the estimation procedures, the comparison of the different procedures, the development of quality criteria for the evaluation of the results in each estimation domain as well as the investigation of the progress concerning the harmonisation of the estimation methods. Setting the same scopes in the future, this work must be also expanded into the other domains of estimation in Eurostat.