

ENBES Workshop “The Unit Problem in Business Statistics Methodology” held in Geneva, Switzerland, on November 10th 2014

Summary

Introduction

The unit problem arises due to errors in creation of units that are used for producing statistical estimates: errors in their identification, assignment of characteristics, or establishment of relations between them. This is of particular concern for big and complex businesses that tend to contribute most to the statistical estimates. Administrative units, created for administrative (e.g. tax or legal) reasons, are not necessarily optimal for direct use for producing statistical estimates, implying the need to create statistical units, intended to be better suited for the purpose of producing official statistics.

The creation and maintenance of the statistical units, including profiling, is a core task often given to a separate business register department. Whatever the methods used, potential errors and effects are far from transparent and are not included in the uncertainty evaluation of the statistics produced. Misidentification of statistical units causes unit error, which will propagate to all statistics that involve the erroneously profiled or classified units. The unit delineation and related issues require a statistical approach in order to bring them into the total error framework.

Immediate goals of the workshop were to support forming a common understanding of the unit problem, to share the related expertise across the disciplines, and to outline the important issues for development and research.

In order to structure and direct the discussions the following **fundamental questions** were presented at the beginning of the workshop:

What is the unit problem?

Where does it appear?

Is it one problem, or an umbrella term covering several problems?

Is the problem real, or an artefact that can be ignored?

How can it be measured/estimated?

What theoretical approaches are available to tackle it?

How can its effects be reduced?

Way forward in the short, medium and long run?

The workshop started with presentations focusing more on conceptual aspects followed by more methodological presentations. Finally, it ended with a discussion session summarizing the day's findings and allowing the exchange of ideas about possible further steps.

In the following you may find an overview of the presentations. The overview is structured by clusters of interest that became evident during the workshop and does not always respect the order in the workshop itself.

Where to use the expensive profiling instrument?

[Harrie van der Ven, CBS: How the Dutch statistical business register is dealing with unit problem challenges and the role of the Complexity Statistical Impact factor](#)

- Use of a complexity factor score (CSI - score) in order to define groups that have to be put in profiling.
- Initially automatic profiling of groups below this threshold. The result of that appeared to be very instable over time. For the short run it was decided to treat this category as one enterprise (enterprise group = enterprise). The impact of this measure still has to be analysed.

[Emmanuel Gros, Philippe Brion, Thomas Deroyon, INSEE: A first assessment of profiling already made, and evaluation of the impact of profiling on statistics](#)

- Impact of profiling (Legal unit vs Profiled units (groups)) on structural business statistics.
- Turnover may decrease if looking at groups due to neutralisation of internal flows (up to factor 3): profiling has a significant quality impact.
- Consolidation rules needed to derive turnover, delivered by administrative sources on legal entity level, on the enterprise level.
- Wrong NACE / non-consolidation of non-additive variables cause bias in estimates.
- Idea: estimate the bias as a function of size and sector.
- Distinguish face-to-face profiled groups, automatic profiled groups and legal entities.
- Categorisation between face-to-face and automatic profiling based on a cost function to be minimised.
- Suggestion from discussion: extend profiling partly to validate the automatic profiling results into domains with smaller units.

Sharing of experiences on automatic profiling

Suggestion from discussion: Compare results on automatic profiling for instance between the studies in France and the Netherlands (previous paragraph).

[Frank Verschaeren, Statistics Belgium: Creating a set of known good data for detecting wrongly delineated statistical units with sufficiently low true error rate](#)

- Check reaction of groups when presenting them the results of automatic profiling methods.

Modelling the errors

[Li-Chun Zhang, University of Southampton and SSB: Elements of a theoretical framework for dealing with the unit problem](#)

- Total error framework should also cover unit error.
- Mathematical expression of unit errors using allocation matrices. Facilitating the basic operations of aggregation and apportion.

[Arnout van Delden, CBS: Classification errors and how they affect population estimates](#)

- Estimation of the effect of classification errors (bias and variance) based on assumptions of probabilities of false classifications represented by a matrix.
- Sensitivity analysis based on uniform probabilities showed that misallocation by economic activity results in large bias when the different economic activities have very different values for target variables (like turnover) A more realistic approach would be with probabilities based on expert knowledge.
- Outcome of analyses of audited wrong NACE-Codes: Factors explaining the probabilities are size and NACE-Code of firms.
- Questions: If one observes a bias – then why not just adjust the results accordingly? But then, how reliable must this bias estimate be?

[Luca Mancini, Simona Toti, ISTAT: Determinants of administrative registers' coverage error](#)

- Presentation of techniques to measure administrative sources coverage errors.
- Part of the development towards replacing Census with Admin register for detailed census-like inpopulation statistics.

What is happening in global organisation of businesses and what is the role of official statistics?

[Gianluca Spina, Politecnico di Milano: New organizational models in global values chains](#)

- Organizational models generally unrelated to legal entities.
- Value chains have become very international. Very different global strategies (for instance shoppers, octopus, full players) even within one branch.
- The organisational model can change rapidly over time.
- Discussion about the sources of information for research in the field: mainly based on information collected from the groups in collaboration projects with them. Official statistics is not an important source in this type of research.

Need to organise the discussion on units and their revision over the whole chain (input, business registers, basic statistics, national accounts)

[Steven Vale, UNECE: The different dimensions of the unit problem\(s\)](#)

- Lack of connection between statistical and real world units.
- One size fits all solutions? Different countries, times, statistics...
- Beside of legal units and statistical units it is relevant to know how the enterprise sees itself.
- Is it a problem because we try to fit a fixed model to an anarchic changing world?
- Work towards solution: not report to a specific expert group. Should be overseen by chief statisticians to ensure a neutral and strategic view. Should not be driven too much by history. Small groups with limited time should come up with proposals for defining statistical units.
- Discussion: full chain management. No isolated discussions on units in national accounts, business registers, business statistics. Such discussion should be deliberately scheduled and performed in an overarching way.

[Norbert Rainer, Statistik Austria: Business unit\(s\), their identification and significance from a Business Register as well as a wider perspective](#)

- The unit problem arises when statistical units are not given but must be delineated.
- Appears also in other domains than business statistics (e.g. household/family).
- Kind of units:
 - Administrative units
 - Legal units
 - Statistical units (adequate units for statistics, avoid inconsistencies in statistical data)
- Statistical Units are theoretical constructs. The more the conceptual model is different from given real units, the bigger is the unit problem.
- Unit model: conceptual relations between the kinds of units.
- Classification (region, activity etc.) depends on choice of statistical units.
- Rather a unit challenge; the challenge is a continuous task because of the dynamic environment.
- First of ten principles for choosing statistical units.
 - Units should be chosen in agreement with the goals of the statistics. Comment at the other hand: variety of units should be minimised (coherence of the statistical system).

[Roland Sturm, Destatis: Conceptual differences between the current and the Eurostat-proposed re-definitions of "enterprise"](#)

- Differentiate the function of units for instance reporting unit, statistical unit.
- ESS-consistency: only 8 member states use concept of complex statistical units – other use legal units.
- Profiling: analysing structure of enterprise group to define appropriate statistical units (different forms like intensive, light, schematic).

- Profiling has to meet different user needs regarding information about the structure of a complex unit (f.e. geographical structure, activity structure)
- Profiling is an expensive instrument, it should be used wisely.
- Comment in discussion: the concept of the statistical unit must also fit with the possibility of the enterprises to deliver the needed data – else efforts may remain useless.
- No one size fits all solutions.

[Tihomira Dimova, UNECE: International discussions about units in the national accounts](#)

- Discussion of challenges regarding treatment / classification of factoryless good producers and special purpose entities– research on guidance by task force on national accounts.
- Intellectual property as gaining importance.
- The work within the National Accounts on a revision of units.

[Wim Kloek, Eurostat: Longitudinal perspective on units in business statistics](#)

- Criteria for defining continuity vs change of identity based approximatively on administrative information.
- Judgement may not only depend on the changes observed but also on the time interval these changes happened. Such information is in practice not available with the exception for the large take-all enterprises.
- Important not only to the analysis of changes in the aggregates over time, but also as a source of micro-economic statistics. The take-all part in business statistics seems to produce a natural starting point for an enterprise panel. Why are panels no longer in fashion?
- Discussion: wish to capture structural changes / but no evolutions due to corrections. Stability is deemed important.

Final discussion

The final discussion and evaluation of the workshop (*via a participant questionnaire*) indicated that while the participants appreciated the workshop, there was also an expectation to be more concrete about the next steps, including how to work towards a more unified view on units. As a first step in this direction, ENBES, as the organizer, intends to produce a position paper. ENBES hopes also that further discussions and contributions on the unit problem topic will take place during the [5th International Conference on Establishment Statistics](#), 2016 in Geneva and the [European Establishment Statistics Workshop 2015](#) in Poznan, Poland, and invites all interested colleagues to submit their abstracts there.