

Text

Johanna Jonsson  
Richard Nilsson  
Data

<b>28<sup>th</sup> Meeting of the Wiesbaden Group on Business Registers The Hague, 2 - 6 October 2023</b>
Johanna Jonsson (Statistics Sweden) Richard Nilsson (Statistics Sweden)
Session 2 – Profiling complex Statistical Units (SUs)
<b>Implementing the results of profiling in the Swedish Statistical Business Register</b>

### Introduction

For several years there has been a large ongoing project at Statistics Sweden: Coordinated economic statistics with “new” statistical units. The ongoing project and the parallel implementation of the results is done through close cooperation between the Business Register and the Large Cases Unit at Statistics Sweden, as well as a dedicated IT-team.

One of the project goals is to fully implement the demands of the Regulation (EU) 2019/2152 on European business statistics, with special focus on the creation and updating of complex enterprises. This is achieved through a combination of automatic and intensive profiling.

Profiling is an ongoing process, and we keep learning as we keep profiling. Throughout this project a lot of the initial ideas have been modified, refined or even discarded. We are always limited by the available data and in the end, national legislation and our budget set hard boundaries for what we are able to do.

We have now, however, implemented the profiled structures in the live Business Register and are developing the routines for regularly updating the register as the annual profiling process progresses.

The majority of enterprise groups are subject to automatic profiling. Intensive profiling is carried out on the largest and/or most diversified enterprise groups. These only constitute about 5% of all ENTs, but they hold about 57% of all persons employed.

## **Live register**

The Swedish Business Register consists of two parts, the live register and the annual frame. Profiling is done in the annual frame and new structures are updated in the live register.

The live register is updated every week with information from administrative sources, mainly the Swedish Tax Office, and register maintenance updates the register daily.

The live register is frozen around the end of the year. This annual frame is then used for profiling, both intensive and automatic. Changes in relationships between enterprise groups, enterprise units and legal units as well as changes in coding of ancillary units are inserted back into the live register at regular intervals as the profiling process proceeds.

Register maintenance of the live register can discover changes that otherwise might not be included in the administrative sources that update the annual frame. A priority system is in place, which prevents old information from the annual frame from over-writing newer information in the live register.

Also, the register maintenance investigates the continuity of the identity of several statistical units, including LOU, KAU and ENT. This is especially important for the latter from the perspective of profiling, since it means that complex demographic events might be handled as they happen instead of at the end of the year.

## **Annual frame**

The live register is frozen at the end of the year. Since we want to include all events of the year, we wait for weekly delivery from the Swedish Tax Office and create our frozen frame at the start of the new year. While we receive weekly updates of administrative data every week and register maintenance is continuous throughout the year, some variables are less than ideal for our frozen annual frame.

The number of employees is updated monthly, but with a time lag of two months. What we want is for the frozen frame to include the number of employees for December rather than October. Number of employees is especially important for two reasons. Firstly, it is used to set the NACE-codes on several statistical units. The number of employees is registered on the LOU and is then multiplied with the percentages of the NACE-codes on the LOU to aggregate the NACE-codes on the LEU, KAU and ENT. Secondly, it is used in the continuity of the ENT in the automatic profiling as a proxy for the survival of the factors of production.

The turnover is based on VAT and show the previous year, but we want the turnover to reflect the year of profiling rather than the previous year. During the production year of profiling, the turnover is updated and is first replaced with turnover from the financial statements, as they are made available, and later with SBS-data for those companies that are part of that survey. The turnover is used when coding ancillary units. Also, the turnover is especially important in profiling when the turnover for an ENT is consolidated as part of intensive profiling.

### **Group structures**

During the year, we do not have access to information on changes in ownership for legal units. Neither The Swedish Companies Registration Office nor the Swedish Tax Office request information on changes in ownership from the legal units. For this information, we need to wait for the yearly annual reports from the legal units themselves. However, the only legal requirement for information on group structure is the closest relations of a legal unit.

We have been using a commercial provider (Dun & Bradstreet) to obtain group structures. Every month, we receive updates on the structures based on information from annual reports and this means that enterprise groups are continuously updated during the production process. When a group has been updated, the automatic profiling is started.

The provider does not send a date of cessation on relationships between legal units, but rather no update at all of the group structures if no new information exists. However, in some instances groups are missing from the monthly updates during the time that personnel at the commercial provider are manually updating the groups and sometimes groups are missing for a few months, only to return again as active. This has led to entire groups, consisting of many legal units, to have been interpreted as having ceased when it has not been included in an update from the commercial provider. As a consequence, the automatic profiling created new ENTs out of every LEU, since the group had ceased.

This gap has been closed with implementing routines from the EGR and the group is no longer interpreted as ceased if it is missing in an update. Instead, larger groups are investigated to see if there is any information available on their status.

## **Ancillary units**

We have had several projects on automatic coding of other NACE codes using the list of NACE codes recommended by Eurostat, which include such diverse activities as freight transport and cargo handling to IT consulting and hosting. While many of the proposed NACE codes could possibly be ancillary activity, not very many of them turned out to be. At least not in the Swedish BR and at least not more than a very small portion.

In a project designed to investigate the outcome of a proposed model for automatic coding of ancillary activity, we investigated the result of our model that included all NACE-codes proposed by Eurostat.

When we compared the outcome with what the companies had stated in their annual reports, we found that a large portion had been wrongfully identified as ancillary units. This was in part due to the fact that when we do automatic profiling and interpret an entire enterprise group as one enterprise, the ENT becomes quite large and diversified. The whole that we compare the part to, becomes so big that even a quite large activity with many employees and large turnover often was mistaken as an ancillary activity.

Also, while many of these activities may be ancillary in some cases, they are also very often market activity. For example, it is not unheard of that some large companies put their IT-departments in one legal unit that serves as an ancillary unit for the group. However, this does not seem to be very common and most legal units with IT-activity seem to be market orientated rather than for internal use.

When creating the framework for automatic coding of ancillary units, we narrowed it down to only three types of activities:

- holding companies (64.20)
- business management (70)
- real estate (68.10)

Because the Swedish national version of NACE (called SNI) has a fifth digit, we are able to differentiate between apartment buildings and non-residential buildings which simplify the automatic coding of real estate as an ancillary unit.

In the automatic profiling the principle is that if the turnover is less than 10% of the total turnover of the ENT and the principal activity is included in the list, then the unit is automatically coded as an ancillary unit.

For groups that are part of the intensive profiling, the coding of ancillary units is manual and the list of possible NACE codes to be considered as ancillary is longer.

Another outcome from the discussions on ancillary units is that Statistics Sweden has changed the handling of large head offices compared to previous years. It used to be that we considered them not as separate ENTs but rather as ancillary units to the rest of the group even if they had a lot of market activity and very large turnover. This view has changed in order to better harmonise with other European countries. As of 2023 this has only been implemented halfway. Head offices with large proportions of turnover originated from outside the ENT are no longer marked as ancillary units, but are still part of a larger ENT.

### **Note on small groups**

There is a special problem with very small ENTs since number of employees are used as a proxy for the continuity of the ENT. Instead of using the usual criteria of control, economic activity and location, this method follows the majority of employees. If the larger portion of the employees remain in the group and the ENT, then the ENT-id stays in the group. If the larger portion of the employees leave the group along with one of several LEUs, then the ENT-id follows and the original group receives a new ENT-id.

This causes some problems for very small ENTs. Every employee receives proportionally larger weight which means that a change of one or two employees can determine if the ENT-id survives or not.

There is also a problem with creating an ENT that makes sense out of these very small groups. It is not unusual that all the employees are in one LEU and they are rented out to other LEUs in the same group. Those other LEUs are the ones who are actually selling their goods and services outside of the group. But this largest LEU have all the persons employed and a relatively high turnover and will therefore not be automatically coded as an ancillary unit. Instead it will have the highest weight when setting the NACE-code for the ENT which often will be some form of employment activity (NACE 78).

In order to mitigate this effect, there is a manual check of the larger of these groups and a possibility to do manual adjustments if necessary.

### **Profiling in the live register**

Profiling has been implemented in the annual frame since the reporting year of 2018, it was not introduced in the live register until 2022.

While profiling in the annual frame is about refining a frozen frame as more and more information becomes available, maintaining the

profiled structures in the live register is about reacting to limited information. The profiled structures were put to a completely different kind of test when they were exposed to weekly updates.

### **Different time periods**

Much can happen in a year. A legal unit that was a holding company without any persons employed can take over the main production of a group and the next year have several hundreds of persons employed and millions in turnover.

When first implementing the profiled ENTs in the live register, we had problems with historical information overwriting newer information in the live register. While register maintenance in the live register might code a LEU as an ancillary unit that has NACE 70.10, the automatic profiling might code the same unit as primary production since that is what they had been during the profiling year.

There were also issues in the population of intensive profiling, in the cases when the profiler had not checked the latest information in the live register. This led to the creation of ENTs that make sense given the information that was available for the specific year, but made no sense in the live register.

We are continuously working to find ways to mend the gaps caused by time lags. In the case of ancillary units, we plan to implement automatic coding of ancillary units in the live register using the current values of variables instead of the historical values from the annual frame. Also, ENTs that in the live register “make no sense”, large ENTs that have died, large ENTs that are new or ENTs that have changed drastically in number of employees or turnover are checked manually.

### **Register maintenance**

The register maintenance is not limited to maintaining the survival of the enterprise unit, but also of the local units and kind-of-activity units. For many surveys, the maintenance of the KAU is more important than the ENT. Various demographic events need to be handled in the live register as they happen in order to assure the quality of the KAU and short-term statistics. A restructuring that involves multiple legal units must be handled as soon as possible. Does it involve one ENT or several ENTs? The same group or are several groups involved?

This means that it is necessary to make decisions on the survival of the ENT in the live register when information is limited. Later, when the annual profiling is done and more information is available, we might

discover that some of the choices made in the live register might have been wrong.

But register maintenance can also aid profiling. One big problem is that we have no information on ownership of newly registered legal units. Not until the first annual report do we know which group it belongs to, if any. Consequently, a newly registered LEU is not part of an enterprise group and will become a new enterprise unit. It is not until the next year that we can identify which enterprise group they belong to and which (if any) existing ENT they should be part of. Depending on how late in the year the LEU was registered it can take up to two years before there is any information on ownership.

### **PAYE – a source for continuity of statistical units**

Before 2019, the main source for number of employees in the Swedish BR was the total amount of wages per LEU every month. A method using an average for every NACE-code was used to calculate the number of employees for every LEU. Once a year in April, we got information on which individuals that had received pay from which LEUs, but because of the time lag the use of this information was limited.

However, since 2019 it is mandatory for Swedish employers to every month declare PAYE (pay-as-you-earn tax) to the Swedish Tax Office. This gives us monthly updates on which individuals that have received income from which LEUs and the total amount of the compensation. And this information can be used in maintaining the continuity of statistical units in the Business Register.

Some newly registered LEUs are flagged in the BR and are investigated. In connection with the monthly delivery of PAYE, there is an automatic check to identify clusters of individuals that have changed legal units. Business transitions when a cluster of individuals receives payments from the same legal unit and then the same cluster of individuals instead receive payments from another legal unit. This is mainly used for the continuity of the LOU and the KAU but this could also mean that the legal unit should become part of an existing ENT and/or enterprise group.

Of course, this works both ways. When a large LEU closes down its activity, we are able to track where the employees go. Through this we are able to conclude if the activity has truly ceased or if it continues under another legal unit and which structure it belongs to.

## **Conclusions**

Completing the milestone of implementing the profiled units in the live register has led to new ways of working. Limited information and time lags lead to discrepancies that need to be handled in a consistent way. As we move to the next phase of our profiling activities we will surely encounter other phenomena that we need to examine and figure out how to deal with.

Creating new routines for updating the live register with information from the annual frame has been, and continues to be, a challenge.