Abstract

The Swiss Federal Statistical Office uses administrative data for statistical purposes as well as a tool to control the quality of the register units. The value added tax (VAT) data, as a new available data source of annual turnover in the Swiss business register, is used to improve the quality of the records of the business register as well as an auxiliary variable in the calibration of certain statistics which leads to a gain in precision and quality of business statistics. However, given the conditions of imposition of VAT, individual tax data do not allow full coverage of the universe.

The main challenge in providing annual turnover data is the lag between a fiscal year and the time when enterprises have completed their tax returns. A detailed analysis calculating the best trade-off between timeliness and accuracy determined that a realistic value of the total annual turnover could be inserted in the business register one year after the reference date. However in the laps of time between the end of the fiscal year and the insertion of the turnover data in the business register, structural enterprise changes such as births, merges, splits and deaths can occur.

The business register is also used as a sampling frame. Once a year a frame is set and kept in this frozen state for all business survey samples of the Swiss Federal Statistical Office. The annual turnover will be inserted in the business register for the first time this year. Therefore this new variable will be included in the next survey frame. The difficulty is to combine the needs of the business register and the needs of the survey users. The former needs a record of the “historical” annual turnover data which provides an accurate record of the economic variables of an enterprise as well as a great potential for temporal data analysis. The latter require “current” turnover data that takes into account the enterprise demography, including employment changes. We are currently developing a method to estimate this “current” turnover. The “historical” turnover value will be recorded once a year in the business register. In addition to these data, an estimation of the “current” turnover value will be provided. A recalculation of this value will be triggered every time a sampling frame is created.
Introduction

According to European regulations on business registers, turnover data must be introduced in the Swiss business register at the enterprise level to establish a common framework for business registers for statistical purposes. Under the Swiss Law on Federal Statistics, Switzerland has also committed itself to promote the use of administrative data, including monetary data.

The Swiss Federal Statistical Office uses administrative data not only for statistical purposes but also as a tool to control the quality of the register units. The turnover data offer an additional tool which makes it possible to detect potential errors in the coding of units in the business register and therefore increase the quality of the register.

Administrative data offer real advantages to official statistics. More specifically, turnover data are one of the key indicators used to measure economic performance and growth. Turnover data also offer considerable information on the structure of the economic landscape, as well as the state, needs and potential of the economy.

In this context, the register of the Swiss Federal Tax Administration was identified as the primary source to provide turnover data. The annual turnover was inserted in the business register for the first time this year. Turnover, as a new mandatory stratification and weighting variable, brings a fresh perspective on the universe of enterprises that is currently measured by employment.

Indeed, VAT data are an important source in the field of business statistics and provide a wide coverage of the target population, contrary to surveys. Furthermore it provides the added value of decreasing the burden on businesses. These data are used as auxiliary variables in the calibration of certain statistics which leads to a gain in precision and quality of business statistics. They represent a solid basis to make statistical results more reliable and directly contribute to limit the burden on businesses.

However, given the conditions of imposition of VAT, individual tax returns do not allow full coverage of the universe. The total coverage of the universe by VAT data is higher than 60% and largely varies depending on employment and economic activities. The coverage of medium and large enterprises is very high, contrary to small enterprises (with fewer than 5 employees). This gap in coverage is mainly due to the tax conditions, which exclude from taxation all enterprises with an annual turnover lower than CHF 100,000, as well as enterprises from the public or primary sector and those engaged in financial, real estate, health care, education or insurance activities. Therefore issues related to the coverage of missing data and data availability were analysed. A thorough quality assessment was also performed on the data. In cases where units are not subject to VAT, turnover must be estimated by imputation.
Lag between the reference year and insertion in the business register

The main challenge in providing annual turnover data is the lag between a fiscal year and the time when enterprises have completed their tax returns. A detailed analysis calculating the best trade-off between timeliness and accuracy determined that a realistic value of the total annual turnover could be inserted in the business register one year after the reference date.

A balance had to be found between data delivered too quickly that are of poor quality and data delivered too late that delay the normal course of projects. The data received by the Swiss Federal Tax Administration contains a very useful variable: the date individual tax returns are received. Two criteria were used to measure, at any given date, the quality of the turnover data:

- the cumulated percentage of turnover corresponding to the addition of individual tax returns received during the reference year (graph 2)
- the cumulated percentage of taxpayers who filled in all their tax returns during the reference year (graph 3).

Graph 1 illustrates the number of tax returns received by the tax administration each day for the reference year 2011. Most businesses are taxed quarterly which explains the 4 peaks. A vast majority of tax returns are received by July of the following year, in this case July 2012.

Graph 1: Number of tax returns received per day (reference year 2011)
Graph 2 illustrates the cumulated turnover over time for the reference year 2011. The amount of each individual tax return is added to the other tax returns received previously by the tax administration. As time goes by, the cumulated turnover increases until reaching the total turnover amount. The total amount of declared turnover can be determined for any given date. As shown in Graph 2, the cumulated turnover of the reference year increases very rapidly until end of March of the following year, and much more slowly thereafter. On 31.03.2012, over 97% of the total 2011 turnover had been sent to the tax administration.

Graph 2: Cumulated turnover (reference year 2011)
Graph 3 illustrates the cumulated number of taxpayers who filled in all their tax returns. On 31.03.2012, this only represents 67%. In June it rises to almost 90%. It was concluded that it would pay off to wait until early June of the following year, date beyond which the evolution of the number of taxpayers who have submitted all their tax returns drastically slows down.

Graph 3: Cumulated number of taxpayers who filled in all their tax returns (reference year 2011)

Therefore, June 1st of the year following the reference year is a milestone. But this is a theoretical date: the time elapsed between the reception of the tax return and the date it was checked, corrected and inserted in the database of the tax administration must be added. Taking this into account, analyses and treatment of the turnover data can begin in September of the following year and annual turnover can be inserted in the business register one year after the reference date.
Historical versus current annual turnover

Turnover is inserted at the enterprise level, which is the basis unit for monetary variables. Currently, this is the lowest level for which we can obtain this information, since it is not possible to obtain turnover data at the local unit level. It is also the level of information requested by Eurostat.

When turnover data are not available from the VAT administrative source it must be estimated, but only if certain basic conditions are met. The turnover of the remaining missing units are estimated using a mathematical model based on the number of employees using a productivity per head ratio calculated by economic activity branch. The model is a robust regression conducted according to employment higher than 0 and turnover higher than 0 for each NACE economic branch at level 2 digits. The imputation methods for missing turnover data were developed in collaboration with the Statistical Methods section of the Swiss Federal Statistical Office.

Enterprises can also be taxed as a group under certain conditions. For VAT groups, only the group head is registered at the Swiss Federal Tax Administration and has a VAT number. The group head declares the turnover of all the members in his group. This total turnover must be redistributed between the active members according to employment and economic activity using a mathematical model.

As described in the previous chapter, there is a lag between the reference year and the insertion of annual turnover in the business register. Therefore the universe of the reference year must be created using the following criteria: includes all enterprises that were active during the reference year as well as those that had an active VAT number during the reference year. This universe is created using historical files of the business register. The annual turnover data inserted in the business register takes into account the historical variables (enterprise structure, demography, economic activity, type, etc.) of the enterprises. The annual turnover data inserted in the business register represent a snapshot of the “historical” annual turnover data which provide an accurate record of the economic variables during the reference year.

Turnover is relevant for all active and market units. The only source of turnover we have at the moment is the VAT tax returns. On one side, its perimeter is wider than the business statistical universe, because the Swiss Federal Tax Administration collects taxes from certain types of businesses that are not recognized by the business statistics, such as fictive companies, or consortia. Also, although they are considered inactive in the statistical system, certain entities of the administrative system generated turnover, for example, a company that was active during the reference year and generated a turnover, but was shut down the following year. On the other hand, the coverage of VAT is smaller than the business statistical universe, particularly for certain economic activities not submitted to VAT.
The business register is also used as a sampling frame. The business survey frame takes a snapshot of the register, a frozen state for a given date. Once a year a frame is set and kept in this frozen state for all business survey samples of the Swiss Federal Statistical Office. The annual turnover was inserted in the business register for the first time this year. Therefore this new variable will be included in the next survey frame. The difficulty is to combine the needs of the business register and the needs of the survey users. The former needs a record of the “historical” annual turnover data which provides an accurate record of the economic variables of an enterprise as well as a great potential for temporal data analysis. The latter require “current” turnover data that take into account the enterprise demography, including employment changes.

However in the laps of time between the end of the fiscal year and the insertion of the turnover data in the business register, structural enterprise changes such as births, merges, splits and deaths can occur. Therefore this definition of the “historical” annual turnover does not meet the needs of the statistical universe or sampling frame, which requires a “current” turnover taking into account the current variables of the enterprises.

The turnover statistical universe is composed of all business units that could potentially be used in business statistics, for which the turnover variable must be available. The statistical universe only includes active enterprises, including the newly created ones since the reference year. As mentioned previously, over 60% of the universe is covered by the VAT administrative data. When turnover data is not available from the VAT administrative source, the remaining missing units are estimated using the mathematical model described above.

A method to estimate this “current” turnover is currently being developed. The “historical” turnover value will be recorded once a year in the business register. In addition to these data, an estimation of the “current” turnover value will be provided. A recalculation of this value will be triggered every time a sampling frame is created.
Conclusion

The use of administrative data from the VAT in the field of statistics is a real opportunity, as well as a major challenge for all statistical actors. On the one hand, collaboration between data providers and the register enables us to reduce the burden on respondents as well as to provide the turnover to the business register. But on the other hand, the use of VAT data raises many questions and forces statisticians to change their habits. Indeed, the statistical community needs to adapt to this new source of information as well as to its content.

Despite these difficulties, VAT turnover has a great potential for statistical work, particularly for when estimating, plausibilising or imputing missing values. Moreover, turnover as a new stratification and weighting variable serves not only to measure the quality of the register, but also to reduce the size of samples and the frequency of surveys. It certainly brings a new perspective to the definitions, but it also increases the effectiveness of results and data processing capabilities.

The main difficulties in the treatment and use of the turnover data was to combined with the needs of the various users. In order to comply with all their needs, two different sets of annual turnover data must be produced every year: a “historical” and a “current” annual turnover. The annual turnover was inserted in the business register for the first time this year and this new variable will also be included in the next survey frame for the first time. If this new methodology is validated in the following year, this dual dataset will be used from then on.