

# **SME Performance Review**

## **2017/2018**

*Methodological note on WP 3*

Berlin, 27 July 2018

**DIW Econ GmbH**

Mohrenstraße 58

10117 Berlin

Contact:

Dr. Anselm Mattes

Tel. +49.30.20 60 972 - 0

Fax +49.30.20 60 972 - 99

[amattes@diw-econ.de](mailto:amattes@diw-econ.de)

[www.diw-econ.com](http://www.diw-econ.com)

## Table of contents

1.	Introduction .....	1
2.	Historical data .....	2
2.1	Scope of historical data delivered.....	2
2.1.1	Sources.....	2
2.1.2	Estimation procedure .....	3
2.2	Detailed Country Sources .....	7
2.2.1	EU28.....	7
2.2.2	Non- EU countries.....	9
3.	Now-casts .....	13
4.	Forecasts .....	15

## 1. Introduction

This document is intended to give an overview of the data sources and methodologies used to compile the dataset on structural SME indicators in WP3 of the SME Performance Review 2017/2018 commissioned by the European Commission (DG GROW).

The key objective of Work Package 3 is to compile a historical database from 2008 to 2015/2016, derive now-casts for the years 2016 and 2017 (if official data are not yet available for 2016), and generate forecasts for the years 2018 and 2019 of the following core indicators by enterprise size class (Eurostat variable codes in parenthesis<sup>1</sup>):

- Number of enterprises (V11110)
- Number of persons employed (V16110)
- Value added at factor costs (V12150)

Furthermore, WP3 delivers historical data for a number of non-core indicators (only EU28 member states):

- Turnover or gross premiums written (V12110)
- Gross investment in tangible goods (V15110)
- Production value (V12120)
- Gross operating surplus (V12170)
- Total purchases of goods and services (V13110)
- Personnel costs (V13310)
- Number of employees (V16130)

The industrial classification used is NACE rev. 2. The data cover the non-financial business economy (sections B-J and L-N). The data are broken down by the following firm size classes<sup>2</sup>:

- Micro enterprises (0 to 9 persons employed)
- Small enterprises (10 to 49 persons employed)
- Medium enterprises (50 to 249 persons employed)
- Large enterprises (250 and more persons employed)

---

<sup>1</sup> A detailed description of all indicators can be accessed via <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:086:0001:0169:EN:PDF>

<sup>2</sup> The size-class definition follows the standards of the Eurostat SBS database and is based on the number of persons employed only. The official EU SME definition additionally takes into account turnover and balance sheet data.

In geographical terms, the data cover the EU28 member states as well as Albania, Brazil, China, Former Yugoslavian Republic of Macedonia (FYROM), Iceland, Japan, Moldova, Montenegro, Russia, Serbia, Turkey, and USA. For EU28 member states, the database provides information at the 1-digit (sections) and 2-digit (divisions) level. For non-EU countries the database provides information on the core indicators at the 1-digit level only.

The database spans a multi-year period ranging from 2008 to the most recent year for which data was available, which varies by country. In most cases this was 2015, although for some countries and indicators 2016 data was available.

The data are primarily sourced from Eurostat and the respective national statistical organisations (NSOs). In case of missing or confidential data, DIW Econ derived estimates of the historical data using different methodologies. Now-casts for the years 2016 and 2017 and forecasts for the years 2018 and 2019 were generated for the core indicators (number of enterprises, number of persons employed and valued added at factor costs) for firm size classes for all EU28 member states.

The second section provides detailed information on the processes and methodologies used to generate the historical database while the third section provides detailed information on the now-casting methodology and the fourth section discusses the generation of the forecasts.

## 2. Historical data

### 2.1 Scope of historical data delivered

#### 2.1.1 Sources

The primary source of the database is the data provided by the Structural Business Statistics (SBS) database available on the Eurostat website. Wherever possible, the data was taken from the *SMEs - annual enterprise statistics by size class*.<sup>3</sup> The *SMEs – annual enterprise statistic by size class* contains all core and non-core indicators for the industrial and construction sections.

---

<sup>3</sup> See <http://ec.europa.eu/eurostat/web/structural-business-statistics/structural-business-statistics/sme>

Additional data sources used:

- *Annual detailed enterprise statistics*: Where the primary source did not provide any data, division and section totals (i.e. the sum of all enterprise size classes of a given section or division) were taken from the *Annual detailed enterprise statistics* on Eurostat. This source was primarily used for the *gross investment in tangible good* indicator as well as for several of the non-core indicators in the trade and services sections which were not included in the *SMEs – annual enterprise statistic*.<sup>4</sup>
- National statistical offices (NSO): Data not included in the SBS database were partially sourced from the websites of the respective NSO. Additionally, all national statistical offices were contacted and asked to provide the missing data within the bounds of their respective confidentiality regulations.

To visualise the source of a specific value in the dataset a colour coding was applied in the country spreadsheets:

- Green: values sourced from Eurostat.
- Yellow: values provided by the national statistical offices upon request or sourced from the NSO web site.
- Blue: all estimated values, as well as the now- and forecasts.

### 2.1.2 Estimation procedure

Whenever neither Eurostat nor NSOs provided data due to confidentiality or because data was not available, the missing values were estimated. To ensure consistency among estimations, a number of generic estimation methods were established.

Table 1 lists the methods most commonly applied to estimate missing data.

---

<sup>4</sup> The data used can be accessed via [http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=sbs\\_na\\_ind\\_r2&lang=en](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=sbs_na_ind_r2&lang=en)

**Table 1: Estimation methods**

Method	Description
Only one value missing in a given section/division in a given year	
Simple difference	The missing size class was estimated as the simple difference between the total and the sum of all other size classes. A missing section/division total was calculated as the sum of all size classes.
More than one value missing in a given section/division in a given year	
Used average size-class distribution of previous and following year	Condition: no missing values in this section/division for the previous and following year Assumption: distribution of the missing size classes is the same as the average distribution of the previous and following years
Use size-class distribution of the previous year	Condition: no missing values in this section/division for the previous year Assumption: distribution of the missing size classes is the same as in the previous year
Use size-class distribution of the following year	Condition: no missing values in this section/division for the following year Assumption: distribution of the missing size classes is the same as in the following year
Interpolation	Condition: value of the missing size class is available for this section/division for the previous and following year Assumption: missing value is the average of the value in this size class for the previous and following year
Extrapolation	Condition: value of the missing size class is available for this section/division for the previous or following year Assumption: missing value is equal to the respective value in the previous or following year
Used information from 1-digit level	Condition: All values for a given section are given/were estimated If the complete data for all but one division belonging to this section was available/already estimated, the missing data of the remaining division was calculated by simple difference with the section level data.
Estimating total (1)	Condition: total value in section/division is available for two previous or following years Assumption: The growth rate of the section/division total is equal to the growth rate between the two previous/following years.
Estimating total (2)	Condition: Section <b>and</b> division total for the previous or following year are given. Assumption: share of division of corresponding section is the same as the previous/following year
<i>Persons employed inferred from number of enterprises</i>	Condition: values for <i>number of enterprises</i> are provided for the respective section/division Assumption: Average number of persons employed per enterprise is equal to the average number of persons employed per enterprise in this size class in the previous/following year (if both previous and following year were available, the average was taken). If no information on average number of persons employed was available at the division level (previous and following year missing as well) average employment of the respective size class

	was taken from the section level.
<i>Value added inferred from the number of persons employed</i>	Condition: values for <i>number of persons employed</i> are provided for the respective section/division Assumption: Average value added per person employed is equal to the average value added per person employed in this size class in the previous/following year (if both previous and following year were available, the average was taken). If no information on average value added per person employed was available at the division level (previous and following year missing as well), average value added per person employed of the respective size class was taken from the section level.
Adjusting national data to be consistent with Eurostat SBS data	In the case where Eurostat SBS totals were provided and data provided by NSOs was not consistent with this, the size-class distribution from the NSO data was applied to the SBS total.

Source: DIW Econ.

Regarding the core indicators, a full dataset was created for the period of 2008 until the most recent year for which at least section (or respectively division) total values were available - in most cases 2015, although 2016 data was available for some countries. All missing values were estimated using the estimation methods described above. For the non-core indicators estimations were only made if a section/division total was available. If a total value was available, the size-class distribution for non-core indicators was estimated using the distribution of *value added at factor cost*. Results were then adjusted to ensure consistency between sections and their corresponding divisions. An exception to this approach is the indicator *number of employees* which was estimated by applying the ratio of *number of persons employed* to *number of employees* from the section level to each division.

One overall exception in comparison to the treatment of the other countries, as well as in comparison to last year's approach, is France. Data for the years 2008 and 2009 was not taken from Eurostat and processed as described above, but back-casted instead. To this effect, the methodological approach for the years 2008 and 2009 follows the now-casting approach (see section 3 for a detailed description).

If no data was available for the section level, no estimates were provided. When estimating missing values, multiple constraints on the data had to be taken into account. As not all constraints could be respected at all times, an ex-ante ranking of constraints was established and applied to all estimated data. These constraints are listed below in decreasing order of importance:

1. All data from the Eurostat SBS database – including both data from the SME statistics and totals from the *Annual detailed enterprise statistics* – were taken 'as is'. No adjust-

ments were made.<sup>5</sup> Data sourced from the national statistical offices – including both data sourced from the NSOs websites and data provided upon request – were taken ‘as is’ to the maximum extent permissible. Only where the data conflicted with Eurostat values, the NSO data was adjusted to ensure consistency.

2. When estimating missing values, the most important constraint was ensuring that the size classes in a given section/division added up to the section/division total.
3. Wherever possible, consistency between sections and corresponding divisions was ensured in the sense that the values of all divisions add up to the values of the corresponding section. This was ensured not only for the division totals but also for each size class.
4. Whenever indicators could be placed in a direct relationship to each other, care was taken to ensure consistency across indicators. Two cases of cross-indicator consistency were handled most often:
  - 4.1. Consistency between the number of enterprises and all other indicators: Whenever there were no enterprises in a size class for a given section/division, the value for all other indicators in the same size class should also be zero.
  - 4.2. Consistency between *number of persons employed* and *number of enterprises*: The average number of persons employed per enterprise should be within the range given by the relevant size class. For example: the average number of persons employed in the size class 10-49 should lie between 10 and 49 persons employed.
5. When estimating an indicator, care was taken to ensure an economically plausible development of this indicator for the relevant size class over time. For example, a high volatility in the number of enterprises over time in a given size class was considered implausible.

Even when taking utmost care to produce consistent and sensible results, certain pitfalls could not be entirely avoided.

- For some countries, the range of data available was very low. Thus, the data quality is limited due to the large amount of estimations required. This is especially the case for the mining section, where missing values due to confidentiality are very prominent.

---

<sup>5</sup> There were only very rare exceptions to this rule when the Eurostat SBS data were obviously wrong or inconsistent.

- The estimation methods described above all have a ‘smoothing’ effect on the data. Thus, high volatility over time in the actual data may have been underestimated substantially. This is especially true for the indicator value added at factor cost.

All country datasets were reviewed by the responsible country experts. The country experts checked whether the data and estimations were plausible from an economic point of view and statistically consistent. For this purpose they sourced additional data (if available) and applied their specific country and industry expertise. The country experts’ comments were implemented to improve data quality.

## 2.2 Detailed Country Sources

### 2.2.1 EU28

The EU28 member states can be sorted into two groups based on the sources used to compile the respective country dataset. For countries in *Group 1* only data from the SBS database was used whereas for countries in *Group 2* additional national data could be applied. Table 2 shows a listing of the two groups and the respective countries:

**Table 2: EU 28 sources used**

<b>Group 1</b>		
Sources: <i>Eurostat SMEs - annual enterprise statistics by size class</i>		
<i>Eurostat Annual detailed enterprise statistics</i>		
<input type="checkbox"/> Austria	<input type="checkbox"/> Finland	<input type="checkbox"/> Poland
<input type="checkbox"/> Belgium	<input type="checkbox"/> Italy	<input type="checkbox"/> Romania
<input type="checkbox"/> Cyprus	<input type="checkbox"/> Luxembourg	
<input type="checkbox"/> Denmark	<input type="checkbox"/> Malta	
<input type="checkbox"/> Estonia	<input type="checkbox"/> Netherlands	
<b>Group 2</b>		
Sources: <i>Eurostat SMEs - annual enterprise statistics by size class</i>		
<i>Eurostat Annual detailed enterprise statistics</i>		
<i>National data provided by the NSO or sourced from the NSO website</i>		
<input type="checkbox"/> Bulgaria	<input type="checkbox"/> Hungary	<input type="checkbox"/> Slovenia
<input type="checkbox"/> Croatia	<input type="checkbox"/> Ireland	<input type="checkbox"/> Spain
<input type="checkbox"/> Czech Republic	<input type="checkbox"/> Latvia	<input type="checkbox"/> Sweden
<input type="checkbox"/> France	<input type="checkbox"/> Lithuania	<input type="checkbox"/> United Kingdom
<input type="checkbox"/> Germany	<input type="checkbox"/> Portugal	
<input type="checkbox"/> Greece	<input type="checkbox"/> Slovakia	

Source: DIW Econ.

For some countries data comparability is not given over time due to structural breaks in the data collection. This is caused by methodological changes in the data collection process.

**Table 3: EU 28 data issues**

country	Data issue
Finland	There is a structural break between the years 2012 and 2013. This was caused by a classification change, in which SBS and business register data was combined in a common production/data collection process.
France	We observe a structural break between the years 2011 and 2012, due to increased data coverage of the self-employed workforce (via ESANE).
Germany	There is a structural break in sections G and I between the years 2008 and 2009, due to methodological changes in the data collection process. Additionally, there is another structural break between 2013 and 2014 affecting the values for micro enterprises, due to methodological improvements. Methodological changes concerning section H might affect the intertemporal comparability. In section L there is a structural break from 2014 to 2015 as private landlords are no longer considered, resulting in a strong decline in the micro size class for purely methodological reasons.
Greece	Structural change in statistics from 2014 onwards due to improvements in the applied methodology as well as the alignment of SBS data with tax data.
Netherlands	There is a structural break affecting all sections between the years 2009 and 2010. This was caused by a change in data sources for compiling the database. From 2010 onwards there is no firm-size threshold anymore. In addition, we observe a structural break between the years 2012 and 2013. A change in the data collection process is likely but has not yet been confirmed by the Dutch NSO.
Ireland	There was a significant upward revision of value added from 2015 onwards affecting national accounts data that are used for now-casts. This is mainly reflected in manufacturing.
Slovakia	From 2010 on data refers to the entire population of enterprises including sole proprietors. This effects the values regarding micro enterprises.
United Kingdom	Value added development expressed in euros appears very volatile between 2014 and 2016 due to substantial exchange rate fluctuations.

Source: DIW Econ.

### 2.2.2 Non-EU countries

For Non-EU countries, a wide selection of national sources was employed. Table 4 below lists the data sources used for each country. Wherever the structure of the data differs from the general structure laid-out above, country-specific issues are listed in Table 5.

**Table 4: Non-EU countries sources used**

country	sources
Albania	Data provided by the NSO of Albania (Instituti i Statistikave) upon request
Brazil	All data was sourced from the IBGE (Instituto Brasileiro de Geografia e Estatística) Website: <a href="http://www.ibge.gov.br/english/estatistica/economia/cadastroempresa/2015/default_tab.shtm">http://www.ibge.gov.br/english/estatistica/economia/cadastroempresa/2015/default_tab.shtm</a>  <a href="https://www.ibge.gov.br/english/estatistica/economia/contasnacionais/2014/defaulttab_sinoticas_ods.shtm">https://www.ibge.gov.br/english/estatistica/economia/contasnacionais/2014/defaulttab_sinoticas_ods.shtm</a>
China	All data was sourced from the National Bureau of Statistics of China. Website: <a href="http://www.stats.gov.cn/tjsj/ndsj/2016/indexeh.htm">http://www.stats.gov.cn/tjsj/ndsj/2016/indexeh.htm</a>
FYR of Macedonia	All data was provided by the NSO of Macedonia (State Statistical Office of the Republic of Macedonia) upon request
Iceland	All data was provided by the NSO of Iceland upon request.
Japan	Data was retrieved from Statistics Japan – Statistics Bureau, Ministry of Internal Affairs and Communications Website: <a href="http://www.stat.go.jp/english/index.htm">http://www.stat.go.jp/english/index.htm</a> <a href="http://www.stat.go.jp/english/data/roudou/index.html">http://www.stat.go.jp/english/data/roudou/index.html</a>
Moldova	All data was searched from Statistica Moldova - National Bureau of Statistics of the Republic of Moldova Website: <a href="http://statbank.statistica.md/pxweb/pxweb/en/40%20Statistica%20economica/40%20Statistica%20economica_24%20ANT_ANT020/ANT020100.px/?rxid=b2ff27d7-0b96-43c9-934b-42e1a2a9a774">http://statbank.statistica.md/pxweb/pxweb/en/40%20Statistica%20economica/40%20Statistica%20economica_24%20ANT_ANT020/ANT020100.px/?rxid=b2ff27d7-0b96-43c9-934b-42e1a2a9a774</a>
Montenegro	Data on Number of enterprises and Value added was provided by the NSO of Montenegro upon request. Number of persons employed were retrieved from the website of the NSO (Monstat). <a href="https://www.monstat.org/enq/page.php?id=1148&amp;pageid=23">https://www.monstat.org/enq/page.php?id=1148&amp;pageid=23</a>
Russia	All data was sourced from the Federal State Statistical Service of the Russian Federation Website: <a href="http://www.gks.ru">http://www.gks.ru</a>
Serbia	All data was provided by the NSO (Statistical Office of the Republic of Serbia, Department for Structural Business Statistics) upon request
Turkey	All data was provided by the NSO (Turkstat – Turkish Statistical Institute) upon request

USA	Number of persons employed and number of enterprises taken from the U.S. Census Bureau <a href="http://www.census.gov/econ">Website: http://www.census.gov/econ</a> Section totals for value added were taken from the Department of Commerce, Bureau of Economic Analysis <a href="http://www.bea.gov/industry/index.htm">Website: http://www.bea.gov/industry/index.htm</a>
-----	--

Source: DIW Econ.

Because of data limitations and differing national industrial classification there are certain aspects to consider.

**Table 5: Non-EU countries data issues**

Country	Data issue
Albania	<ul style="list-style-type: none"> <li><input type="checkbox"/> Data for 2008-2012 only available in NACE Rev. 1.1</li> <li><input type="checkbox"/> Data from 2012 onwards is in NACE Rev. 2</li> <li><input type="checkbox"/> Value added is in basic prices instead of factor costs</li> </ul>
Brazil	<ul style="list-style-type: none"> <li><input type="checkbox"/> Value Added is only given for sector totals, class size break downs were calculated by assuming the same economies of scale as in EU-average.</li> </ul>
China	<ul style="list-style-type: none"> <li><input type="checkbox"/> Size class breakdown for enterprises (20-300, 300+, Total) differs from SBS typology</li> <li><input type="checkbox"/> The data is originally provided in the national industrial classification and therefore the data is only a best fit to NACE Rev. 2 classification</li> <li><input type="checkbox"/> Only few data with a size-class breakdown could be provided (only number of enterprises). Apart from that it was only possible to provide section totals.</li> <li><input type="checkbox"/> The statistical yearbook provided only for some sections data for 2016.</li> <li><input type="checkbox"/> Two different sources for value added: Statistical Yearbook 2016 (covering 2011-2014) and Third Economic Census (published every fifth year) until 2013.</li> <li><input type="checkbox"/> Differences in value added depending on source</li> </ul>
FYR of Macedonia	<ul style="list-style-type: none"> <li><input type="checkbox"/> Data for 2008 is only available in NACE 1.1 and were therefore estimated</li> </ul>

Japan	<ul style="list-style-type: none"> <li>❑ Data is based on a “best fit” mapping from the Japan Standard Industrial Classification to NACE Rev. 2.</li> <li>❑ Size-class breakdown (0-9, 10-49, 50-299, 300+) differs from SBS typology</li> <li>❑ No data provided on sector N</li> <li>❑ As the Economic Census is only conducted every second year, values in between had to be calculated by average.</li> <li>❑ Value added: Calculation of average productivity per section (Total VA per section/Total number of persons employed in section). Value added of class sizes was calculated by assuming the same economies of scale as in Germany.</li> <li>❑ Changes in class size break up between 2014 and 2015 due to recourse to different data bases (2012-2014 Labour Force Survey, after 2014 Economic Census)</li> </ul>
Moldova	<ul style="list-style-type: none"> <li>❑ Data on enterprises is only available in NACE Rev. 1.1 until 2014.</li> <li>❑ Enterprise size-class definition as well as the size-class breakdown for small enterprises (1-10, 11-50) deviates from the SBS typology</li> </ul>
Montenegro	<ul style="list-style-type: none"> <li>❑ Data on enterprises is only available from 2011 to 2014</li> <li>❑ Data on number of persons employed is only available for sector totals</li> </ul>
Russia	<ul style="list-style-type: none"> <li>❑ Data is only available in NACE Rev. 1.1</li> <li>❑ Enterprise size-class definition as well as the size-class breakdown deviates from the SBS typology</li> <li>❑ Information on Number of Persons Employed in SMEs given by the Federal State Statistical Service likely to be incorrect.</li> <li>❑ Data on value added is only available for section totals.</li> </ul>
Turkey	<ul style="list-style-type: none"> <li>❑ Data for 2008 is only available in NACE Rev. 1.1</li> <li>❑ Data from 2009 onwards is provided in NACE Rev. 2</li> <li>❑ Size-class breakdown (1-19, 20-49, 50-249, 250+) differs from SBS typology</li> </ul>

USA	<ul style="list-style-type: none"> <li>❑ Data is based on a “best fit” mapping from the NAICS 2002/NAICS 2007 to NACE Rev. 2</li> <li>❑ Size-class breakdown (0-9, 10-49, 50-299, 300+) differs from SBS typology</li> <li>❑ Number of enterprises and number of persons employed of 2016 were estimated by assuming the same development as for establishments</li> <li>❑ Size-class breakdown for value added was calculated by assuming the same economies of scale effects as in the UK.</li> </ul>
-----	---

Source: DIW Econ.

### 3. Now-casts

The now-casts are based on data from Eurostat national accounts, more precisely the Quarterly National Accounts detailed breakdowns by industry and product.<sup>6</sup> Two of the core indicators, i.e. persons employed (V16110) and value added (V12150), are estimated based on the national accounts data. The third indicator, i.e. number of enterprises, is estimated based on a bridge equation from persons employed to number of enterprises.

The now-casts for persons employed (V16110) are based on the Quarterly National Account employment data by branch (10 branches). The now-casts for value added (V12150) are based on the Quarterly National Accounts value added aggregates at current prices (10 branches). The quarterly national accounts data are available until the fourth quarter of 2017 and on a NACE rev.2 section level.<sup>7</sup>

The now-casting procedure follows a top-down approach including the following steps:

- ❑ First, to obtain now-casts at section level for the size class “total”, we applied – for each NACE rev. 2 section – the growth rates of employment and value added in a given year of the national accounts data to the corresponding year and level of V16110 and V12150 in our data-set.
- ❑ Second, to now-cast the size-class breakdown at section level, we run a multivariate regression to estimate the elasticity of growth (of persons employed and value added) of a given size class with regard to the growth of the respective section total. This estimated relationship is then used to determine the development of a particular size class within a

<sup>6</sup> Last updated: 02/05/2018.

<sup>7</sup> Some of the sections are aggregated in the national accounts data such as for example B, D and E or G-I.

section. Predicted annual growth in sections B and D is limited to the maximum change in historical growth.

□ Third, to now-cast the size classes at division level, we run a similar regression to assess the elasticity of growth (of persons employed and value added) of a given size class at division level with regard to the respective size class at section level. This estimated elasticity is then used to determine the growth rate of persons employed and value added of the respective size classes at division level. Again, predicted annual growth rates in sections B and D are capped at the maximum change in historical growth.

□ In order to obtain now-casts for the number of enterprises (V11110), a bridge equation is used. Without loss of generality, the following methodology is used to derive now-casts for V11110:

i) The ratio of  $V16110_{p,t,j}$  with respect to  $V11110_{p,t,j}$  is calculated for the latest year available in the data set (2015 or 2016):

$$\frac{V16110_{p,t,j}}{V11110_{p,t,j}}$$

where  $V16110_{p,t,j}$  is the number of persons employed at time  $t$  in section  $j$  for size class  $p$ .

ii) In a second step, the corresponding section, division and size-class now-casts for V16110 are divided by this ratio (i.e. persons employed per enterprise) in order to obtain the now-casts for V11110 for the years 2016 and 2017 respectively:

$$V11110_{p,t,j} = \frac{V16110_{p,t,j}}{\left(\frac{V16110_{p,t-1,j}}{V11110_{p,t-1,j}}\right)}$$

for  $t=2016, 2017$ .

The rationale behind this approach is that the historic relationship between persons employed and enterprises in a given size class is typically stable over time and can be extrapolated into the future. Exploitation of the bridge equation is favoured over the alternative of applying econometric estimations because this approach ensures consistency between the indicators. As mentioned before, the same approach has been used to estimate (back-cast) missing 2008 and 2009 data for the three main indicators for France.

## 4. Forecasts

The starting dataset for the forecasts is a balanced country and division level panel data set, spanning from 2008 to 2017 as obtained from the historical data compilation and now-casting process. The forecasts for the three core indicators “number of persons employed”, “number of enterprises” and “value added at factor costs” are generated using a top-down approach similar to the one adopted for the generation of the now-casts. The forecasting process requires the use of a predictor which is available over the whole forecasting period. The main predictor variable chosen for this purpose is employment of the total economy taken from the annual macro-economic (AMECO) database maintained by DG Economic and Financial Affairs.<sup>8</sup> The variable used is labelled “NETN - Employment, persons: total economy (national accounts)” and the unit is 1000 persons. It is available up to 2019.

These employment forecasts are used as predictor variables in regression models to forecast the number of persons employed (V16110) for the years 2018 and 2019. The forecasting procedure follows a top-down approach including the following steps:

- First, to forecast persons employed at section level and for size-class “total”, we run an OLS regression to assess the (historical) elasticity between employment growth in the total economy (AMECO) and the growth of persons employed in a given section.<sup>9</sup> This estimated elasticity is then used to determine the growth rate of persons employed for the years 2018 and 2019 at section level which then in turn is applied to determine the number of persons employed for the years 2018 and 2019.
- Second, to forecast the size-class breakdown at section level, we run an OLS regression to assess the elasticity of growth of persons employed of a given size class with regard to the growth of the respective section total. This estimated elasticity is then used to determine the development of a particular size class within a section.
- Third, to forecast persons employed at division level, we run an OLS regression to assess the elasticity of growth of persons employed of a given-size class at division level with regard to the respective size class at section level. This estimated elasticity is then used to determine the growth rate of persons employed of the respective size classes at the division level.

---

<sup>8</sup> Data sourced from the AMECO database was last updated on 03/05/2018.

<sup>9</sup> Based on the data set obtained from the now-casting process.

Subsequently, the section, division and size-class specific forecasts of “number of enterprises” and “value added” are derived using a bridge equation approach. The ratios of interest are “persons employed per enterprise” to forecast the number of enterprises and “value added per person employed” to forecast value added.

The bridge equation formulas for the number of enterprises are nearly identical to the one described above. The bridge equation formulas for value added are derived analogously:

- i) Labour productivity (i.e. the ratio of  $V12150_{p,t,j}$  with respect to  $V16110_{p,t,j}$ ) is calculated for the year  $t$ :

$$Labour\_Productivity_{p,t,j} = \frac{V12150_{p,t,j}}{V16110_{p,t,j}}$$

where  $V12150_{p,t,j}$  is value added at factor costs at time  $t$  in section  $j$  for size class  $p$ .

- ii) Since the AMECO forecasts of value added for 2018 and 2019 are much more optimistic than the respective employment forecasts (which implies a growth in nominal labour productivity), a constant bridging equation results in an underestimation of value added (in contrast to the ratio of persons employed per enterprise in a given size class which is highly stable over time). Hence, the ratio of value added per person employed (labour productivity) was adjusted by the difference between the AMECO forecasts of the growth rate of value added and the AMENCO forecasts of the growth rate of employment:

$$\begin{aligned} Labour\_Productivity_{p,t,j} &= Labour\_Productivity_{p,t-1,j} \\ &* [1 + (Growth\_Rate\_VA_{AMECO,t} \\ &- Growth\_Rate\_Employment_{AMECO,t})] \end{aligned}$$

for  $t=2018, 2019$ .

- iii) Estimated labour productivity is then multiplied by corresponding section, division and size-class forecasts for the number of persons employed for 2018 and 2019 to obtain forecasts for value added for the years 2018 and 2019 respectively:

$$V12150_{p,t,j} = Labour\_Productivity_{p,t,j} * V16110_{p,t,j}$$

for  $t=2018, 2019$ .