# 9th Vintage CompNet Data Quality Report

First insights

# April 2023

# Contents

1	Introduction	3
<b>2</b>	Comparison methodology	4
3	France	4
	3.1 Data Sources	4
	3.1.1 Structural Business Statistics (SBS) from OECD	4
	3.1.2 French CompNet $\ldots$	5
	3.2 Coverage rates	6
	3.3 Indices and time series	7
	3.4 Summary	11
<b>4</b>	Portugal	12
	4.1 Data sources	12
	4.1.1 Structural Business Statistics (SBS) from OECD	12
	4.1.2 Portuguese CompNet	12
	4.2 Coverage rates	13
	4.3 Indices and time series	14
	4.4 Summary	20
<b>5</b>	Slovakia	20
	5.1 Data sources	20
	5.1.1 Structural Business Statistics (SBS) from OECD	20
	5.1.2 Slovakian CompNet	21
	5.2 Coverage rates $\ldots$	22
	5.3 Indices and time series	22
	5.4 Summary $\ldots$	24
6	Latvia	25
-	6.1 Data sources	25
	6.1.1 Structural Business Statistics (SBS) from OECD	25
	6.1.2 Latvian CompNet	$\overline{25}$
	6.2 Coverage rates	26

	6.3	Indices and times series	27
	6.4	Summary	30
7	Slov	renia	<b>31</b>
	7.1	Data sources	31
		7.1.1 Structural Business Statistics (SBS) from OECD	31
		7.1.2 Slovenian CompNet	32
	7.2	Coverage rates	32
	7.3	Indices and time series	33
	7.4	Summary	36
8	Ger	many	36
0	8 1	Data sources	36
	0.1	8.1.1 Structural Business Statistics (SBS) from OECD	36
		8.1.2 Corman CompNot	37
	<b>8</b> 9	Coverage rates	20
	0.2 0.2	Indices and time series	39
	0.0		40
	8.4	Summary	43
9	Use	r Guidance	44
10	App	oendix	46
	10.1	France	46
	10.2	Portugal	47
	10.3	Slovakia	49
	10.4	Slovenia	50
	10.5	Germany	50

### 1 Introduction

The CompNet data set contains a unique and vast compilation of firm-related indicators and variables. The goal of this country report is to provide a first comparison of CompNet aggregates for the TSI participating countries against "official" statistics of the Structural Business Statistics published by Eurostat (as well as the OECD) and to highlight features of the CompNet data users should be aware of.<sup>1</sup>

We focus on comparison of aggregate levels and trends between statistics derived from CompNet and the OECD-SBS data set. When possible, we will offer explanations for the differences and suggestions for what to do when this happens. In principle, levels and trends should line up well but in practice they do not always. One reason for this is that the underlying data sources for the construction of the CompNet and SBS data set differ. Second, coverage changes in the data underlying the construction of the SBS statistics generate well known breaks in the official statistics. We try to avoid these breaks in the CompNet data. This makes comparison of levels between SBS and the CompNet data extremely difficult.

With regards to CompNet, one should also note that the number of observations on which the indicators are derived varies by sector and years. For example, across all years manufacturing and the retail sector have very good coverage and indicators are computed on thousands of observations. For other sectors, specifically real estate activity, and accommodation, indicators rely on fewer observations. Users should keep in mind and interpret indicators and variables of retail and accommodation sector with caution.

In the next section we will focus on a subset of key variables which we compare across the two data sources. We will focus on:

- Number of firms
- Number of employees (headcounts) or alternatively the number of persons employed
- Labour costs (payroll) defined as gross wages before taxes + employer's contribution to social security
- Sales (turnover) or alternatively the gross output
- Value-added at factor cost
- Labour productivity as turnover per employee or value-added per employee

The structure of the report is as follows: In section 2 the comparison methodology between both data sets is introduced and the results for the six TSI countries are presented in sections 3 to 8. Finally, section 9 contains general comments on user guidance.

<sup>&</sup>lt;sup>1</sup>The OECD reproduces indicators from Eurostat although sometimes for different years so we use them here interchangeably. We treat Eurostat's and the OECD's structural business statistics data base under the same heading because the OECD publishes Eurostat data in the OECD structural and demographic business statistics data base. The contents of both data sets are more or less identical and sourced from the same data.

### 2 Comparison methodology

We compare the following variables from CompNet to the SBS of the OECD: number of firms, turnover, value added, employment and labor productivity. Comparisons are made for the years 2010-2019 mainly (taking into account both OECD-SBS and CompNet data availability) both at the country and at the 1 digit sector level. For each variable we compute the coverage rate  $(c_{it})$  as follows:

$$c_{it} = \frac{aggregate \ SBS \ statistic_{it}}{Compnet \ statistic_{it}} \tag{1}$$

where coverage rates >1 indicate that the "official" aggregates are higher than the CompNet aggregates and vice versa for coverage rates <1. A second measure of comparison is given by the trend analysis. In this case we fix a base year for both datasets and compare how a certain variable has evolved over time. This allows us to compare the two databases in terms of evolution over time. The index is computed as:

$$I_{it} = \frac{var_{it}}{var_{base}} \tag{2}$$

The following section presents first insights into the data quality of the data for the six participating countries in the TSI.

#### **3** France

#### 3.1 Data Sources

#### 3.1.1 Structural Business Statistics (SBS) from OECD

Source Data: Elaboration of Annual Statistics of Companies (ESANE).

#### Main Variables:

- Number of Enterprises: Includes firms with 0 employees.
- Turnover: Gross sales million EUR or NAC.
- Value Added: Value added at factor cost.
- Employment: Number of persons employed. This represents the number of employees plus the number of unpaid persons employed.
- Payroll: Remuneration of employees million EUR or NAC. Defined as wages/salaries + social security costs.
- Labour productivity: defined as TUTT/EMPN gross output per person employed and as VAFC/EMPN value added at factor cost per person employed.

**Changes in Coverage:** The SBS industry coverage has expanded over time. Starting from the reference year 2008 data is available for Sections B (mining and quarrying) to N (administrative and support service activities) and Division S95 of NACE Rev.2. This

means that the total or macro sector aggregates for the years before 2008 (if they are available) tend to be lower.

**Changes in Statistical Units:** The statistical units are "enterprises". France changed their enterprise concept in 2017 – switching from a 'legal unit' concept to a true 'enterprise' concept. This leads to known breaks after 2016 in the SBS data.

#### Time Coverage: 1996-2019.

**Size Classification:** For both Eurostat and OECD-SBS data are broken down into firm size classes according to the number of persons employed.

**Country-specific issues France:** According to Eurostat, there are several issues to keep in mind when using the French SBS data which are relevant for our comparisons. These issues are responsible for breaks in official statistics particularly acute in Construction, Wholesale and Retail, and Professional Services. Bottom line is, the official statistics have notable breaks in its official series and cannot be considered to be an internally consistent time series. According to Eurostat comparability over time is only possible for the French SBS statistics between 1996-2007.

#### 3.1.2 French CompNet

- Source Data and Coverage: The French CompNet raw data have been drawn as a (near full-)sample from all of the firms' tax declarations. Not included in the raw data set seem to be micro firms subject to micro-BIC.
- Source Data: Administrative data from tax declarations including firms that are taxed according to the Regime of Normal Real Profits (BRN or Bénéfice Réel Normal), and firms taxed according to the Simplified Tax Regime (RSI or Régime Sim-plifié d'Imposition).
- Main Variables:
  - Number of Enterprises: including only firms with at least 1 employee.
  - Employment: headcounts of the number of employees at a certain date with employed shareholders/owners excluded.
  - Turnover: Turnover at factor cost + Increase in the stock/inventory of manufactured finished - or semi-finished products + capitalized internal activities.
  - Payroll: Gross wages and salaries paid to employees + other monetary or nonmonetary expenses for employee benefits (social security contributions, payroll taxes, benefits).
  - Value added: defined as turnover expenses for purchased goods and services.
  - Labour productivity: aggregated nominal value added and/or aggregated nominal turnover divided by aggregated number of employees.

- Methods: The initial sample is reweighted to approximate the full population of firms using firm population numbers taken from Eurostat SBS at the two-digit level for 5 size classes.
- Changes in Statistical Units: Composition of the sample has changed in 2017, 2018, and 2019 due to the change from legal unit to an enterprise.
- **Time Coverage:** 2004-2019
- Industry Coverage: In contrast to the OECD/Eurostat SBS data as well as the EUKLEMS data set, CompNet data are generally not collected for all NACE 2 sections (CompNet: macro sectors) and divisions (CompNet: two-digit sector). Sections A (agriculture, forestry, and fishing), B (mining and quarrying), D (electricity, gas, steam, and air conditioning supply), E (water supply; sewerage, waste management, and remediation activities), K (financial and insurance activities) and also the sections O to U (all divisions > 82) are not included.
- Samples: Full sample and the sample of firms with at least 20 employees ("20e sample"). Weighted and unweighted versions of the CompNet data are available. Comparisons should be based on weighted versions. Weights are designed to approximate the full population using population numbers taken from SBS at the two-digit level for 5 size classes.
- **Discrepancies:** CompNet statistics exclude firms with no employees. However, the CompNet weighting procedure using firm numbers from the SBS weights is based on the count of firms including those with zero employees. This means that the 1-9 size class in the CompNet data receives a disproportionately high weight. As such statistics from size class 1-9 might overshoot the true numbers in the population, we recommend using the 20e sample.

#### 3.2 Coverage rates

It should be noted that while the OECD-SBS database includes all 0 employees firms (self-employed individuals), the CompNet data set does not. For this reason, interested users should mainly use statistics from the 20e sample (includes only firms with at least 20 employees) and preferably the version which is weighted to match population totals.

Table 1 shows the coverage rate, as defined above, for selected years, for the 20e weighted sample. Other variables like personnel compensation, intermediate inputs and value-added are missing for a few years and tend to have low coverage rates for the years that data is available. Since these variables are primarily only available for the Manufacturing and Construction sectors, making economy-wide comparisons is impossible. The coverage rates for the 20e weighted sample for the remaining variables however are above or below 1, indicating differences in the degree of comparability between the OECD-SBS statistics and the CompNet statistics.

Figure 36 in the Appendix shows the coverage rate for the number of firms. The coverage rates remain close to 0.9 over the years until 2016, thereafter we observe a

Variable	2012	2013	2014	2015	2016	2017	2018	2019
number of firms	0.90	0.91	0.88	0.86	0.94	0.71	0.76	0.81
revenues	1.36	1.44	1.60	1.24	1.34	1.32	1.31	1.27
value added	0.55	0.57	0.67	0.54	0.57	0.60	0.61	0.59
workers employed	1.17	1.12	1.09	1.10	1.20	1.18	1.32	1.20

Table 1: Coverage rate table for the 20e sample weighted - France

significant decrease. In Figures 37 and 38, the revenue and number of persons employed coverage rates start from 2010. In the case of Figure 37, the coverage rate for revenue remains above 1 suggesting that the CompNet statistics systematically underestimate revenue relative to the OECD-SBS data. The same is also true for the number of persons employed in Figure 38.

#### 3.3 Indices and time series

Next, we show how indices vary over time. Figure 1 shows the evolution of the index variable, constructed as mentioned above, for the entire economy over time for the 20e sample. In general, the indices created using the CompNet data follow those using the OECD-SBS data for nearly all the variables of interest. However, there exist some exceptions.



Figure 1: Evolution of indexes over time for the entire economy (20e weighted sample).

Figures 2-6 show the sector-wise differences in the time-series evolution of the key indices. In the case of the number of firms in Figure 2 the CompNet time series pattern follows the OECD-SBS one closely even if the latter trend is always below, especially after 2016. This may be explained by the change in the definition of Enterprise and it is true

across all sectors. Furthermore, the comparison between CompNet and OECD-SBS data in the number of persons employed in Figure 3 shows more discrepancy after 2016.



Figure 2: Number of firms by sector

Figure 3: Number of persons employed by sector



Figure 4 shows the sector-wise time-series comparison of personnel compensation using the CompNet data and the OECD-SBS data. As noted earlier, since the OECD-SBS data on personnel compensation is unavailable for all sectors except the Construction and Manufacturing sectors, we observe only one time-series plot for the CompNet data in the panels for the remaining sectors.



In Figure 5, we see a sharp dip in the revenue index in the years 2013 for the Transportation and Storage sector and 2014 for the Real Estate Activities sector. For the remaining sectors, the revenue index using the CompNet data closely follows that from the OECD-SBS data with some exceptions. Figure 6 shows the same sharp dip seen in the panel for the Transportation and Storage, and Real Estate Activities sectors on the revenue index since value-added is measured by gross revenue minus expenses. However, in this case, the OECD-SBS trend is always below the CompNet data.



Figure 5: Revenue index across sectors



Figure 7 shows the sector-wise evolution of the labour productivity index (in valueadded terms). The panels in this figure reflect the same sector-wise patterns as that of the previous two figures since labour productivity here is measured using value-added output.



Figure 7: Labour Productivity index (value added) by sector

#### 3.4 Summary

- CompNet coverage rates against OECD-SBS are quite different. Notably, coverage rates for the number of firms are all close to one until 2016, while it significantly decreases thereafter. On the other hand, the coverage rates for revenue and persons employed remains above 1 for all the period considered.
- CompNet over-performs the benchmark notably for the number of firms index especially after 2016. This is probably due to the different enterprise concept France adopted starting from 2017.
- There is large heterogeneity when looking at 1-digit sectors. Value added growth is the furthest above the OECD-SBS data in almost all of them, except "Administrative and support service activities".
- When it comes to the revenue benchmark, significant differences between CompNet and the OECD-SBS data are observed in all sectors, except Manufacturing and Administrative and Support Service activities. Considering the OECD-SBS statistics we see a sharp dip in the revenue index in the years 2013 for the "Transformation and Storage" sector and 2014 for the "Real Estate Activities" sector.

• Users should take into account these discrepancies presented in previous figures. Relying on time series for these sectors in particular should be careful when formulating conclusions.

### 4 Portugal

#### 4.1 Data sources

#### 4.1.1 Structural Business Statistics (SBS) from OECD

#### Main Variables:

- Number of Enterprises: Includes firms with 0 employees.
- Turnover: Gross sales million EUR or NAC.
- Value Added (at factor cost).
- Employment: Number of persons employed. This represents the number of employees plus the number of unpaid persons employed.
- Payroll: Remuneration of employees million EUR or NAC. Defined as wages/salaries + social security costs.
- Labour productivity: defined as TUTT/EMPN gross output per person employed and as VAFC/EMPN value added at factor cost per person employed.

**Changes in Coverage:** The SBS industry coverage has expanded over time. Starting from the reference year 2008 data is available for Sections B (mining and quarrying) to N (administrative and support service activities) and Division S95 of NACE Rev.2. This means that the total or macro sector aggregates for the years before 2008 (if they are available) tend to be lower.

**Time Coverage:** SBS data for Portugal is available for the years 1996-2019. Break in data series between 2007 and 2008.

#### 4.1.2 Portuguese CompNet

#### • Main Variables:

- Number of Enterprises: including only firms with at least 1 employee.
- Employment: headcounts of the number of employees at a certain date with employed shareholders/owners excluded.
- Turnover: Turnover at factor cost + Increase in the stock/inventory of manufactured finished or semi-finished products + capitalized internal activities.
- Payroll: Gross wages and salaries paid to employees + other monetary or nonmonetary expenses for employee benefits (social security contributions, payroll taxes, benefits).

- Value added: defined as turnover expenses for purchased goods and services.
- Labour productivity: aggregated nominal value added and/or aggregated nominal turnover divided by aggregated number of employees.
- Methods: The initial sample is reweighted to approximate the full population of firms using firm population numbers taken from Eurostat SBS at the two-digit level for 5 size classes.
- Time Coverage: 2010-2020.
- Samples: Full sample and the sample of firms with at least 20 employees ("20e sample"). Weighted and unweighted versions of the CompNet data are available. Comparisons should be based on weighted versions. Weights are designed to approximate the full population using population numbers taken from SBS at the two-digit level for 5 size classes.
- **Discrepancies:** CompNet statistics exclude firms with no employees. However, the CompNet weighting procedure using firm numbers from the SBS weights is based on the count of firms including those with zero employees. This means that the 1-9 size class in the CompNet data receives a disproportionately high weight. As such statistics from size class 1-9 might overshoot the true numbers in the population, we recommend using the 20e sample.

#### 4.2 Coverage rates

Table 2 shows the coverage rate, as defined above, for selected years, for the 20e weighted sample. Other variables like personnel compensation, intermediate inputs and value-added tend to have low coverage rates for the years that data is available. Since these variables are primarily only available for the Manufacturing and Construction sectors, making economy-wide comparisons is impossible. The coverage rates for the 20e weighted sample for the remaining variables however are close to 1, indicating the high degree of comparability between the OECD-SBS statistics and the CompNet statistics.

Variable	2012	2013	2014	2015	2016	2017	2018	2019
number of firms	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
revenues	1.59	1.63	1.63	1.61	1.61	1.63	1.63	1.62
value added	0.99	1.01	1.01	1.01	1.02	1.03	1.02	1.01
workers employed	1.00	1.01	1.00	0.99	0.99	0.99	0.99	0.99

Table 2: Coverage rate table for the 20e sample weighted - Portugal

Figure 39 in the Appendix shows the coverage rate for the number of firms. The coverage rates remain close to 1 over the years. In Figures 40 and 41 are shown the revenue and number of persons employed coverage rates. In the case of Figure 40, the coverage rate for revenue remains above 1 suggesting that the CompNet statistics systematically underestimate revenue relative to the OECD-SBS data. Instead, the number of persons employed in Figure 41 is always very close to 1.

#### 4.3 Indices and time series

Next, we show how indices vary over time. Figure 8 shows the evolution of the index variable, constructed as mentioned above, for the entire economy over time for the 20e sample. In general, the indices created using the CompNet data follow those using the OECD-SBS data for nearly all the variables of interest, the most evident exception is labor productivity.



Figure 8: Evolution of indexes over time for the entire economy (20e weighted sample).

Figures 9-13 show the sector-wise differences in the time-series evolution of the key indices. In the case of the number of firms in Figure 9 the CompNet time series pattern follows the OECD-SBS one very closely. Furthermore, the comparison between CompNet and OECD-SBS data in the number of persons employed in Figure 10 shows some discrepancy in the Real Estate sector.



Figure 9: Number of firms by sector



15

Figure 11 shows the sector-wise time-series comparison of personnel compensation using the CompNet data and the OECD-SBS data. As noted earlier, since the OECD-SBS data on personnel compensation is unavailable for all sectors except the Construction and Manufacturing sectors, we observe only one time-series plot for the CompNet data in the panels for the remaining sectors.



Figure 11: Personnel compensation by sector

In Figure 12, we see some discrepancies in the revenue index for the Information and Communication, and Real Estate sectors. For the remaining sectors, the revenue index using the CompNet data closely follows that from the OECD-SBS data. Figure 13 shows the same discrepancies seen in the panel for the Information and Communication, and Real Estate sectors on the revenue index since value-added is measured by gross revenue minus expenses.



Figure 12: Revenue index by sector



Figure 14 shows the sector-wise evolution of the labour productivity index (in valueadded terms). The panels in this figure reflect the same sector-wise patterns as that of the previous two figures since labour productivity here is measured using value-added output.



Figure 14: Labour Productivity index (value added) by sector

Figure 15 displays the kernel density for the distribution of the sectoral labour productivity for Portugal's 20e weighted sample. Mean productivity has remained quite stable from the years 2010 and 2012 to 2019. The shape of the distribution has remained more or less the same over the years.

Figure 16 shows a positive correlation between labour productivity and firm size across industries for the 20e weighted sample. That is, industries with large firms tend to be more productive. The regression line is flatter in 2019 suggesting that the positive correlation is weaker relative to that in 2010 and 2012.



Figure 15: Distribution of sectoral labour productivity

Figure 16: Correlation between average firm size and labor productivity



### 4.4 Summary

- CompNet coverage rates against OECD-SBS are overall satisfactory. Notably, coverage rates for the number of firms are all equal to one. On the contrary, the coverage rate for revenue remains above 1 suggesting that the CompNet statistics systematically underestimate revenue relative to the OECD-SBS data.
- In general, the indices created using CompNet and OECD-SBS data are very similar, the most important exception is labour productivity. For this variable, CompNet statistics over-performs the benchmark until 2016.
- The evolution of the CompNet time series indices follows the OECD-SBS one very closely at 1-digit sectors, only small discrepancies in specific sectors exist. For example, the number of persons employed index is above the benchmark only in Real Estate Activities.
- When it comes to the revenue benchmark, significant differences between CompNet and the respective benchmark are observed in sectors such as Information and Communication and Real Estate Activities.
- Users should take into account these discrepancies presented in previous figures. Relying on time series for these sectors in particular should be careful when formulating conclusions.

# 5 Slovakia

#### 5.1 Data sources

#### Data source: Statistical Office of the Slovak Republic

#### 5.1.1 Structural Business Statistics (SBS) from OECD

#### Main Variables:

- Number of Enterprises: Includes firms with 0 employees. The average and median firm sizes have fallen by 55% and 61% respectively between the years 2000 and 2020.
- Turnover: Sales
- Value Added: Value added at factor cost.
- Employment: Number of persons employed. This represents the number of employees plus the number of unpaid persons employed.
- Payroll: Remuneration of employees. Defined as wages/salaries + social security costs
- Labour productivity: defined as revenues per workers employed and as value added per workers employed
- Time Coverage: 2005-2019.

#### 5.1.2 Slovakian CompNet

#### Source Data: National Bank of Slovakia, Statistics Slovakia,

The current sample for Slovenia is available for firms with 20e.

- Main Variables:
  - Number of Enterprises: including only firms with at least 20 employees. The total number of observations tripled over time.
  - Employment: headcounts of the number of employees at a certain date with employed shareholders/owners excluded. The number of employees has remained constant over time despite a decrease in firm size between the years 2000-2019.
  - Turnover: Gross output that includes 1) Turnover at factor cost + 2) Increase in the stock/ inventory of manufactured finished- or semi-finished products + 3) Capitalized internal activities or increase in the value of total assets by the construction of own machinery, self-constructed buildings or other selfconstructed investment goods. Capitalized internal activities do not include other non-financial revenues.
  - Payroll: includes all the gross wages and salaries paid to employees, payments in the form of other monetary or non-monetary expenses for employee benefits (social security contributions, payroll taxes, benefits) are included.
  - Value added: defined as turnover-expenses for purchased goods and services
  - Labour productivity: aggregated nominal value added and/ or aggregated nominal turnover divided by the aggregated number of employees
- **Time Coverage:** The time coverage for Slovakia in Compnet is the years 2000-2019.
- Industry Coverage: The share of 5 size classes has remained constant for the last 10 years. Sector 75 is only available for years before 2007 and sector 12 only for the years 2000-2006 and from 2013-2015
- Sample Composition: The 20e sample is available in weighted and unweighted structure. It is worth mentioning that the number of observations in the 9th vintage has tripled over time. The average median firm size has reduced by 55% and 61% respectively between the years 2000 and 2020. Note that the share of size class 3 firms has increased at the expense of the decrease in size class 4 and 5. It is also important to note that The median and average age of firms nearly tripled during the same period as expected for post-socialist economies. Comparisons should be based on weighted versions. Weights are designed to approximate the full population using population numbers taken from SBS at the two-digit level for 5 size classes.
- Discrepancies: However, some variables are available for certain years. For example, Capital stock is only available for 2000-2019, debtors, intangible fixed assets, and investments are available for the years 2000-2019. Therefore, some of the capital-related variables could not be calculated for 2020. The values for export, imports, and debt are available only after 2003. The effective tax rate is also available for the years 2012 -2020.

#### 5.2 Coverage rates

Table 3 shows the coverage rate, as defined above, for selected years, for the 20e weighted sample. Other variables, like compensation to employees and intermediate inputs, have very little coverage in the OECD-SBS database across sectors. They are mainly only available for the manufacturing and retail sector, making comparisons at the economy-wide dimension impossible. Coverage rates of the economy-wide number of firms and employees align well between both statistics. However, value added seems to be systematically overestimated in the CompNet data while the aggregates on revenue by the OECD-SBS seem to be somewhat higher than in the CompNet data. Discrepancies between both data sources might arise from the way stocks are accounted for or by differences in the estimation of intermediate inputs. The level differences should be taken into account when working wih the CompNet data.

Variable	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Number of firms	0.96	0.95	0.97	0.94	0.98	0.97	0.99	0.99	0.96	0.97
Revenues	1.14	1.14	1.18	1.20	1.21	1.19	1.20	1.17	1.13	1.12
Value Added	0.45	0.46	0.46	0.46	0.49	0.51	0.50	0.50	0.49	0.49
Employees	0.96	0.96	0.96	1.00	0.99	0.97	0.98	0.97	0.96	0.95

Table 3: Coverage rate table for the 20e sample weighted - Slovakia

Figures 42 and 43 show that the differences in the coverage rates for revenue and value added are driven by sectors 5, 7 and 8 in particular. Level comparisons for those sectors should therefore be avoided or treated with caution.

#### 5.3 Indices and time series

Next, we show how indices vary over time. Figure 17 shows the evolution of the index variable, constructed as mentioned above, for the entire economy over time for the 20e sample. We see that despite the level differences identified above all variables of interest follow a similar growth pattern over time. The largest differences arise in labor productivity and value-added for which the OECD-SBS index lies above the CompNet index.



Figure 17: Evolution of indexes over time for the entire economy (20e weighted sample).

The following Figures 18 and 19 show the labor productivity and value-added index in more detail by sector. The larger growth trend in the OECD-SBS statistics in labor productivity is especially evident in sectors 6,7,8 and 9 for which the CompNet is systematically smaller. In the case of value-added, differences seem to arise from the sectors 5 and 7.







Figure 19: Value added index by sector

#### 5.4 Summary

- CompNet coverage rates against OECD-SBS are overall satisfactory. Notably, coverage rates for the number of firms are all equal to one.
- Coverage rates for value-added stand out; CompNet underestimates value-added relative to the OECD-SBS data because the latter source has very little coverage for this variable across sectors (mainly available for the Transport and storage sector).
- In general, indices created using CompNet and OECD-SBS data are similar until 2015 for Value-Added, and Labour Productivity, after which we see a dip in the CompNet data. For indices like the Number of firms, Labour, Wages, and Revenue CompNet and OECD SBS data are similar until 2020.
- The dip in the labor productivity index is explained by the dip in the real estate sector, professional services sector, information and communication, and wholesale and retail sectors.
- Concerning the labor productivity index, OECD-SBS reports a spike in 2012 which is not apparent in the CompNet data. This spike can be explained by spikes in Real Estate activities, professional services, administrative services, and wholesale and retail sectors.
- Users should take into account these discrepancies and use CompNet mainly for gathering micro-aggregated moments across and within sectors, countries, and years in order to estimate structural models of firm dynamics and to investigate correlations between variables.

### 6 Latvia

#### 6.1 Data sources

#### 6.1.1 Structural Business Statistics (SBS) from OECD

Source Data: Central Statistical Bureau of Latvia

#### Main Variables:

- Number of Enterprises: Includes firms with 0 employees.
- Turnover: Sales
- Value Added: Value added at factor cost.
- Employment: Number of persons employed. This represents the number of employees plus the number of unpaid persons employed.
- Payroll: Remuneration of employees. Defined as wages/salaries + social security costs
- Labour productivity: defined as revenues per workers employed and as value added per workers employed

**Changes in Coverage:** The SBS for Latvia has good coverage at the 1-digit sector level for the following variables: number of firms, turnover, value added and number of persons employed. However, for both turnover and value added, there is no data for sectors 7 (real estate activities) and 8 (professional, scientific and technical activities) in the year 2012. Coverage of remaining variables vary depending on sector and year. Remuneration of employees and intermediate inputs (total purchases of goods and services), and value added as turnover minus purchases of goods and services have very little coverage and are essentially only available for the manufacturing and construction sector (sectors 1 and 2 respectively). Moreover, they are only available for years 2008-2019.

Time coverage: 2005-2019.

#### 6.1.2 Latvian CompNet

#### Source Data: Central Statistical Bureau of Latvia

#### Main Variables:

- Number of Enterprises : including only firms with at least 1 employee.
- Employment: headcounts of the number of employees at a certain date with employed shareholders/owners excluded
- Turnover: Turnover at factor cost + Increase in the stock/inventory of manufactured finished or semi-finished products + capitalized internal activities

- Payroll: Gross wages and salaries paid to employees + other monetary or nonmonetary expenses for employee benefits (social security contributions, payroll taxes, benefits).
- Value added: defined as turnover-expenses for purchased goods and services
- Labour productivity: aggregated nominal value added and/ or aggregated nominal turnover divided by aggregated number of employees
- Time Coverage: 2007-2017
- Industry Coverage: All sectors at the 1-digit level
- Samples: Full sample and the sample of firms with at least 20 employees ("20e sample"). Weighted and unweighted versions of the CompNet data are available. Comparisons should be based on weighted versions. Weights are designed to approximate the full population using population numbers taken from SBS at the two-digit level for 5 size-classes.
- Discrepancies: CompNet statistics exclude firms with no employees. However, the CompNet weighting procedure using firm numbers from the SBS weights is based on the count of firms including those with zero employees. This means that the 1-9 size class in the CompNet data receives a disproportionate high weight. As such statistics from size class 1-9 might overshoot the true numbers in the population, we recommend using the 20e sample.
- Changes in coverage: Focusing on the 20e weighted sample, Latvian CompNet has good coverage for all aforementioned variables, except for payroll. This data, while available for all sectors, is only available from 2011 onwards (years 2011-2017).

#### 6.2 Coverage rates

Table 4 shows the coverage rate, as defined above, for selected years, for the 20e weighted sample. While the coverage for number of firms persons employees is 1, SBS OECD database has slightly higher coverage for revenues. Coverage for value added, however, is much lower at around 0.4, meaning that CompNet coverage rates are overestimated. This is due mostly due to CompNet slightly overestimating all sectors, and sectors 7 and 8, (Real estate activities, and Professional, scientific and technical activities) having especially high coverage rates (around 0.3). Other variables, like compensation to employees, intermediate inputs and value added as turnover minus purchases of goods and services, have very little coverage in the SBS database across sectors. They are mainly only available for the manufacturing and construction sectors, making comparisons at the economy wide dimension not possible. However, coverage in those sectors in which the data is not missing is good.

Variable	2010	2011	2012	2013	2014	2015	2016	2017
number of firms	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
revenues	1.06	1.09	1.09	1.09	1.08	1.05	1.15	1.18
value added	0.38	0.38	0.38	0.37	0.38	0.37	0.39	0.41
workers employed	0.98	0.97	0.95	0.98	0.97	0.98	0.99	0.99

Table 4: Coverage rate table for the 20e sample weighted - Latvia

#### 6.3 Indices and times series

Next, we show how indices vary over time. Figure 20 shows the evolution of the index variable, constructed as mentioned above, for the entire economy over time for the 20e sample. The number of firms, number of persons employed and wages have very similar patterns and the indices tightly follow each other. With wages, however, the Compnet line is flat in 2010, which can be explained by the fact that CompNet reports wages only from 2011 onwards. Also to note for wages, CompNet does report wages for all sectors while OECD only reports wages for the Manufacturing and Construction sectors (sectors 1 and 2 respectively), as seen in Figure 21.





For revenues and value added, the indices also follow similar trends, however they exhibit some discrepancies after 2015; specifically, there is a dip in 2016 in the CompNet which does not appear in the OECD data. This can be explained by a dip in the Wholesale and retail trade (sector 3) and Administrative and support sector activities (sector 9) compared to OECD as evident from Figure 22. Worth to note in Figure 22 the gap in 2012 for sectors 7 and 8, as previously explained. In addition to missing data, the OECD data and CompNet data differ vastly for these two sectors, despite following similar trends

after 2013. Lastly, Labor productivity index in Figure 20 stands out the most. OECD's labor productivity spikes in 2012, which seems to be due to increases in the values of Wholesale and retail trade and Information and Communication sectors, sectors 3 and 6 respectively, as seen in Figure 23.



Figure 22:	Revenue	index	by	sector
------------	---------	-------	----	--------





In Figures 24 and 25 we plot the evolution of the main variables of interest first aggregated at the economy level over time, and then for the Manufacturing (sector 4) and Wholesale and retail trade sector (sector 3) using CompNet data.



Figure 24: Time series of selected variables for the entire economy



Figure 25: Time series of selected variables for Manufacturing and Wholesale and retail trade sectors

It becomes evident that in the whole economy, all variables were affected by the 2008 financial crisis, however that labor productivity, revenue and value added have recovered since quickly, while number of firms and number of persons employed has not recovered since. Similarly, in Figure 25 we can see the similar trends in the Wholesale and retail trade sector. However, the Manufacturing sector has shown slower recovery overall, which may be due to the fact that even pre-2008 financial crisis, this sector contributed lower levels to the country's productivity.

#### 6.4 Summary

- CompNet coverage rates against OECD-SBS are overall satisfactory. Notably, coverage rates for the number of firms are all equal to one. However, the coverage rate for revenue remains above 1 suggesting that the CompNet statistics underestimate revenue relative to the OECD-SBS data.
- Coverage rates for value-added stand-out; CompNet overestimates value-added relative to the OECD-SBS data because the latter source has very little coverage for this variable across sectors (mainly available for the Manufacturing and Construction sectors).
- In general, indices created using CompNet and OECD-SBS data are similar until 2015, after which there are small divergences in the labor productivity index, value added index, and revenue index with the CompNet showing a small dip in 2016 compared to the OECD. This dip can be explained by a dip in sectors 3 (Wholesale and retail trade) and 9 (Administrative and support activities).

- Concerning the labor productivity index, OECD-SBS reports a spike in 2012 which is not apparent in the CompNet data. This spike can be explained by spikes in Wholesale and retail trade, and Information and Communication sectors (sectors 3 and 6 respectively).
- Concerning wages, while CompNet reports wages for all sectors from 2011 onwards, OECD-SBS only reports wages for Manufacturing and Construction sectors (sectors 1 and 2 respectively).
- Users should take into account these discrepancies and use CompNet mainly for gathering micro-aggregated moments across and within sectors, countries, and years in order to estimate structural models of firm dynamics and to investigate correlations between variables.

### 7 Slovenia

#### 7.1 Data sources

#### 7.1.1 Structural Business Statistics (SBS) from OECD

**Source Data:** Agency of the Republic of Slovenia for Public Legal Records and Related Services (AJPES)

#### Main Variables:

- Number of Enterprises: Includes firms with 0 employees.
- Turnover: Gross sales million EUR or NAC.
- Value Added (at factor cost).
- Employment: Number of persons employed. This represents the number of employees plus the number of unpaid persons employed.
- Payroll: Remuneration of employees million EUR or NAC. Defined as wages/salaries + social security costs.
- Labour productivity: defined as TUTT/EMPN gross output per person employed and as VAFC/EMPN value added at factor cost per person employed.

Time and changes in coverage: SBS data for Slovenia is available for the years 2005-2019. The coverage is relatively good for most of the variables in the analysis. However, some information is missing before 2008 in "Manufacturing" and "Construction" regarding total purchases of goods and services, value-added as turnover minus purchases of goods and services, and wages. In the latter sectors, information is also missing between 2005-2019 for the 1-digit sectors wholesale and retail trade, transportation and storage, accommodation and food service, information and communication, real estate activities, professional, scientific and administrative, and support services.

#### 7.1.2 Slovenian CompNet

#### • Main Variables:

- Number of Enterprises: including only firms with at least 1 employee.
- Employment: headcounts of the number of employees at a certain date with employed shareholders/owners excluded.
- Turnover: Turnover at factor cost + Increase in the stock/inventory of manufactured finished or semi-finished products + capitalized internal activities.
- Payroll: Gross wages and salaries paid to employees + other monetary or nonmonetary expenses for employee benefits (social security contributions, payroll taxes, benefits).
- Value added: defined as turnover expenses for purchased goods and services.
- Labour productivity: aggregated nominal value added and/or aggregated nominal turnover divided by the aggregated number of employees.
- Methods: The initial sample is reweighted to approximate the full population of firms using firm population numbers taken from Eurostat SBS at the two-digit level for 5 size classes.
- Time Coverage: 2002-2021.
- Industry Coverage: In contrast to the OECD/Eurostat SBS data as well as the EUKLEMS data set, CompNet data are generally not collected for all NACE 2 sections (CompNet: macro sectors) and divisions (CompNet: two-digit sector). Sections A (agriculture, forestry, and fishing), B (mining and quarrying), D (electricity, gas, steam, and air conditioning supply), E (water supply; sewerage, waste management, and remediation activities), K (financial and insurance activities) and also the sections O to U (all divisions > 82) are not included.
- Samples: Full sample and the sample of firms with at least 20 employees ("20e sample"). Weighted and unweighted versions of the CompNet data are available. Comparisons should be based on weighted versions. Weights are designed to approximate the full population using population numbers taken from SBS at the two-digit level for 5 size classes.
- **Discrepancies:** CompNet statistics exclude firms with no employees. However, the CompNet weighting procedure using firm numbers from the SBS weights is based on the count of firms including those with zero employees. This means that the 1-9 size class in the CompNet data receives a disproportionately high weight. As such statistics from size class 1-9 might overshoot the true numbers in the population, we recommend using the 20e sample.

#### 7.2 Coverage rates

The following Table 5 shows coverage rates for selected years and the 20e weighted sample against comparable information in OECD-SBS. In general, CompNet ensures satisfactory coverage for Slovenia.

Variable	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
number of firms	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
revenues	0.81	0.82	0.81	0.80	0.81	0.80	0.80	0.81	0.82	0.80
value added	0.69	0.71	0.67	0.67	0.69	0.70	0.71	0.72	0.72	0.71
workers employed	0.94	0.96	0.97	0.97	0.94	0.97	0.99	0.99	0.99	0.99

Table 5: Coverage rate table for the 20e sample weighted - Slovenia

One can see that both CompNet and OECD-SBS data are comparable for the above key variables. However, revenues and value-added seem to be somewhat underestimated in CompNet data. This holds across all sectors with the only exceptions being the coverage rate in sector 8 ("Professional, scientific and technical activities") that shows a large spike around the year 2010 and sector 7 ("Real estate activities") for which data is missing (see Figure 44 in the appendix).

#### 7.3 Indices and time series

Next, we show how indices vary over time. Figure 26 shows the evolution of the index variable, constructed as mentioned above, for the entire economy over time for the 20e sample. The indices created using the CompNet data almost perfectly follow the ones arising from the OECD-SBS statistics.



Figure 26: Evolution of indexes over time for the entire economy (20e weighted sample).

The following figures show the sector-wise index developments for the key variables of interest in the 20e weighted sample. In the number of firms almost no differences arise across sectors. Figure 28 depicts the labor index and indicates that CompNet and OECD-SBS statistics differ more strongly for certain sectors (namely sectors 3, 7 and 8).



Figure 27: Number of firms by sector

Figure 28: Number of persons employed by sector



Figures 29 and 30 show the revenue and value added by sector respectively. The revenue index between CompNet data and OECD-SBS is well aligned for sectors 1,2,3,4 and 5. For the remaining sectors either one of the data shows anomalies, such as large

spikes. For the value added index a similar picture arises with differences between both data sources specifically arising in sectors 6 and 8.



Figure 29: Revenue index by sector

#### Figure 30: Value added by sector



### 7.4 Summary

- CompNet coverage rates against OECD-SBS are overall satisfactory. Notably, coverage rates for firms' numerosity are all equal to one.
- CompNet underperforms the benchmark notably for labour productivity, valueadded, and revenue (see Figure 27). For these variables, the explanation likely lies in underlying definitions not entirely converging.
- There is a large heterogeneity when looking at 1-digit sectors. Labor productivity growth is the furthest above the benchmark in almost all of them, notably "Manufacturing", "Transportation", "Accommodation and food services activities" and "Administrative and support service activities".
- When it comes to the revenue benchmark, significant differences between CompNet and the respective benchmark are observed in sectors such as "Wholesale and retail trade", "Manufacturing" after 2015, "Transportation and storage" and, notably "Professional, scientific and technical activities".
- Users should take into account these discrepancies presented in previous figures. Relying on time series for these sectors in particular should be careful when formulating conclusions.

### 8 Germany

#### 8.1 Data sources

#### 8.1.1 Structural Business Statistics (SBS) from OECD

#### Main Variables:

- Number of Enterprises: Includes firms with 0 employees.
- Turnover: Gross sales million EUR or NAC.
- Value Added (at factor cost).
- Employment: Number of persons employed. This represents the number of employees plus the number of unpaid persons employed.
- Payroll: Remuneration of employees million EUR or NAC. Defined as wages/salaries + social security costs.
- Labour productivity: defined as TUTT/EMPN gross output per person employed and as VAFC/EMPN value added at factor cost per person employed.

**Changes in coverage:** Changes in Coverage: Since we will use the weighted sample of firms with at least 20 employees ("20e weighted sample") for the German CompNet, the relevant benchmark within the OECD SBS dataset for Germany are the Business Statistics by Employment Size Class (BSC). The latter source, indeed, breaks down information by size classes which allows selecting groups of firms for proper comparison.

When considering relatable samples also in terms of overlapping 1-digit sectors, from 2005 until 2020 the SBS data for Germany has good coverage for most of the variables in the analysis. Two exceptions are intermediate inputs and remuneration of employees whose information is missing before 2008 in "Manufacturing" and "Construction" and entirely missing for all the other sectors.<sup>2</sup>

All SBS data series exhibit a break in 2018 induced by Eurostat starting to report results for enterprises rather than legal units in SBS in accordance with the EU definition of enterprises as stipulated in the Council Regulation (EEC) No  $696/93^3$ .

#### 8.1.2 German CompNet

Source data: Federal Statistical Office (Destatis), AFID Register

#### Main Variables:

- Number of Enterprises: For manufacturing, only firms with more than 19 employees. For all the other 1-digit sectors, firms with annual turnover of at least 17,500 EURO.
- Employment: headcounts of the number of employees at a certain date with employed shareholders/owners excluded.
- Turnover: Gross output, which includes: 1) Turnover at factor cost: gross sales revenues minus customer discounts, returns and allowances; excluding indirect taxes but including subsidies on products and production (sales include: revenues from selling manufactured finished-or semi-finished goods, revenues from selling goods bought for resale, and revenues from services offered.) + 2) Increase in the stock/ inventory of manufactured finished- or semi-finished products + 3) Capitalized internal activities, i.e., increase in the value of total assets by the construction of own machinery, self-constructed buildings or other self-constructed investment goods (excluding software, licenses, patents, copyrights developed in-house). This definition does not include other non-financial revenues (e.g. revenues from liquidating reserves, unexpected payments of demands that have been already written off, etc. or revenues from selling tangible or intangible non-financial fixed assets). Furthermore, financial revenues are also excluded.
- Payroll: Gross wages and salaries paid to employees + other monetary or nonmonetary expenses for employee benefits (social security contributions, payroll taxes, benefits).
- Value added: defined as turnover expenses for purchased goods and services.

<sup>&</sup>lt;sup>2</sup>Notice that, while for value added comparison between CompNet and OECD SBS is still possible given that the latter directly contains measures for it, information on remuneration of employees for the missing sectors in OECD SBS cannot be reverted.

<sup>&</sup>lt;sup>3</sup>From Eurostat: "According to the EU enterprise definition, an enterprise is the smallest combination of legal units that is an organisational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its resources. An enterprise carries out one or more activities at one or more locations. An enterprise may correspond to a single legal unit ("simple enterprise") or consist of several legal units ("complex enterprise")".

- Labour productivity: aggregated nominal value added and/or aggregated nominal turnover divided by the aggregated number of employees.
- Time Coverage: 2001-2002 (only "Manufacturing"), 2003-2005 (all 1-digit sectors except "Wholesale and retail trade; repair of motor vehicles and motorcycles" and "Accommodation and food service activities"), 2006-2018 (all CompNet 1-digit sectors).
- Industry Coverage: All CompNet 1-digit sectors, that is, the following NACE Rev.2 sections: "Manufacturing", "Construction", "Wholesale and retail trade; repair of motor vehicles and motorcycles", "Transportation and storage", "Accommodation and food service activities", "Information and communication", "Real Estate activities", "Professional scientific and technical activities", and "Administrative and support service activities".
- Samples: Only the 20e weighted sample is available. Weights are designed to approximate the full population using population numbers taken from SBS at the two-digit level for 5 size classes.
- **Discrepancies:** Despite CompNet statistics generally excluding firms with zero employees, the weighting procedure uses SBS numerosity figures that include them. This means that the 1-9 size class in CompNet receives a disproportionately high weight. Therefore, usage of the 20e weighted sample is usually recommended in CompNet. Nevertheless, for Germany this is not necessarily the case.

In both SBS and CompNet the number of enterprises accrues from the same threshold, i.e., an annual turnover of less than 17,500 EURO. Moreover, both databases exclude firms with zero employees in a calendar year. However, in CompNet only for the "*Manufacturing*" 1-digit sector it is sufficient having more than 19 employees for a given firm to be included, which also explains Germany being solely available in the 20e weighted sample.

Employment comparison between OECD SBS and CompNet is respectively between persons employed and number of employees. Furthermore, sales are similarly defined across the two databases (indirect taxes like VAT, customer discounts, returns, and allowances are all excluded) apart from OECD SBS diverging for not including subsidies. Additionally to sales, CompNet includes inventory increases and capitalized internal activities within total turnover. Instead, other non-financial income, financial income, and extra-ordinary income are equally excluded by both databases.

These discrepancies together with those for intermediate inputs (at market prices in CompNet and at purchasers' prices in OECD SBS) can be expected to induce differences in the indicators of value added at factor costs as well. They would seemingly reflect on measures of labor productivity computed as the ratio between turnover and employment (both taken as aggregates at the 1-digit sector level).

Finally, when confronting wages it has be borne in mind that pension payments to retired staff are taken into account in CompNet. But, consistently with the definition set out in paragraph 7.21 of SNA 1993, this is not the case in OECD SBS.

#### 8.2 Coverage rates

Table 8.2 below shows country-level coverage rates for the 20e weighted sample against comparable information in OECD-SBS. Overall, CompNet guarantees a satisfactory coverage for Germany. Until 2017, totals for turnover are larger than the benchmark and this is even more the case for value added. This was predictable considering that CompNet, differently from OECD-SBS, defines turnover including increases in the inventory and capitalized internal activities on top of sales. The wider distance between CompNet and OECD-SBS aggregate value-added figures can be likely traced back to the intermediate inputs definition of the former (at market prices) further inflating measures with respect to the definition of the same item in the "official" statistics (at purchasers' prices)<sup>4</sup>.

Quite surprisingly, CompNet overperforms also for employment. This is difficult to reconcile with concepts of employment in CompNet and OECD-SBS respectively as employees and persons employed. One possible explanation is the fact that, rather than being computed as a yearly average<sup>5</sup>, CompNet employment for Germany is observed at a certain date with possible seasonal effect.

Interestingly, in all 1-digit sectors CompNet time series for the number of firms are almost perfectly aligned with the benchmark (see Figure 45). Only "Manufacturing" reports a coverage rate lower than 1 by more than 0.01. The same sector is also the unique to which a different inclusion parameter for firms applies: while all other sectors in CompNet follow the same OECD-SBS criteria of including enterprises with at least 17,500 EURO turnover and one employee in a calendar year, CompNet "Manufacturing" comprises all firms with at least twenty employees. Hence, for "Manufacturing" CompNet covers additional units than OECD-SBS coinciding with firms with at least twenty employees and annual turnover below 17,500 EURO. Despite being an understandably residual category of the firms population, it may help to explain the leading role of "Manufacturing" in CompNet outnumbering the benchmark. Moreover, "Manufacturing" often drives CompNet excessive values against "official" aggregates for most of the variables, together with "Construction", "Wholesale and retail trade; repair of motor vehicles and motorcycles", and "Accommodation and food service activities".

Coverage rates rise abruptly above one in 2018, corresponding to Eurostat switching from legal units to enterprises as presentation unit of SBS. Meanwhile, such change did not take place on underlying data from Destatis and, consequently, in CompNet<sup>6</sup>. Therefore, the comparability in this year between the two datasets is questionable.

All these considerations reflect into dynamics for labor productivity computed as the ratio of turnover over employment at the 1-digit sectoral level. The country-level measure is reverted as unweighted mean over all 1-digit sectors for each year. The more

<sup>&</sup>lt;sup>4</sup>According to Eurostat Glossary, indeed, "The purchaser price of goods includes any transport charges paid separately by the purchaser to take delivery at the required time and place".

<sup>&</sup>lt;sup>5</sup>Germany diverges from the standard definition of employment in CompNet which is: "Headcounts of the number of employees (yearly average) with employed shareholders/owners excluded".

<sup>&</sup>lt;sup>6</sup>From Eurostat: "The enterprise according to the EU definition is the central presentation unit of business sector-specific SBS. The legal unit will still work as survey unit for SBS".

comprehensive definition of turnover in CompNet prevails on the higher employment measurements, which boils down to OECD-SBS trailing also in terms of labor productivity even though catching up over the most recent years. The country level, however, veils a certain degree of heterogeneity among 1-digit sectors according to whether each sector has a stronger advantage relating to turnover or employment in CompNet with respect to OECD-SBS. Therefore, for "Manufacturing"<sup>7</sup> and "Transportation and storage" CompNet labor productivity underperforms against the benchmark, while for "Real Estate activities", "Professional scientific and technical activities", and "Administrative and support service activities" this is only the case starting from 2017. On the other hand, "Accommodation and food service activities" drives the result of CompNet displaying more pronounced cross-sector average labor productivity than OECD-SBS.

Variable	2010	2011	2012	2013	2014	2015	2016	2017	2018
Number of firms	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Turnover	0.92	0.91	0.92	0.93	0.93	0.90	0.92	0.92	1.04
Value added	0.86	0.85	0.83	0.84	0.85	0.85	0.87	0.87	1.02
Employment	0.95	0.93	0.92	0.93	0.93	0.93	0.91	0.91	1.03

Table 6: Coverage rate table for the 20e sample weighted - Germany

#### 8.3 Indices and time series

When looking at the country-level evolution of indices (constructed like in equation 1) over time in Figure 31, trends for the number of firms, value added, and turnover (or revenues) appear almost entirely overlapping between CompNet and OECD SBS.

<sup>&</sup>lt;sup>7</sup>For "*Manufacturing*", the edge of CompNet over OECD-SBS being starker for employment rather than turnover is consistent with the hypothesis that additional firms in CompNet have at least twenty employees but extremely low revenues.





Source: CompNet 9th Vintage unconditional\_mac\_sector\_20e\_weighted.dta and OECD-SBS SSIS\_BSC\_ISIC4.

For labor productivity, instead, CompNet delivers a more subdued growth rate by circa 0.05 (i.e., five percentage points) in 2017. Besides reflecting the benchmark narrowing the gap for this variable, it is also clear that the differential observable between the two datasets in labor productivity developments stems from a divergence of the same magnitude, but of inverse direction, between indices for employment in the same year. More sustained gains in employment for CompNet seem to lead more rapidly rising wages as well. Nevertheless, CompNet wages outperform OECD-SBS by a larger factor (circa 0.1, i.e., ten percentage points) than the one for employment. This suggests that faster employment increases compound with the deviating underlying definition of wages (from remuneration of employees in SBS) for determining a higher pace in the evolution of the latter variable for CompNet<sup>8</sup>.

Also for indices, when shifting to the 1-digit sectoral level differentiated patterns emerge. Hastened CompNet employment growth is led by "Manufacturing", "Construction", "Wholesale and retail trade", "Accommodation and food service activities", and, over the last few years, by "Professional, scientific and technical activities" and "Administrative and support service activities" too (see Figure 32).

<sup>&</sup>lt;sup>8</sup>Since CompNet, on the contrary of OECD-SBS, also accounts for pension payments to retired staff, the accelerated wage growth in CompNet is consistent with ageing trends in the German population. According to the World Bank, the share of population aged 65 or above rose by two percentage points between 2010 and 2020 reaching 22%.



Figure 32: Number of persons employed by sector

Figure 33: Number of persons employed by sector



As shown in Figure 33 CompNet turnover rose ahead of OECD-SBS in most of the years only for "Manufacturing", "Construction", and "Professional, scientific and technical activities", but just for "Manufacturing" by a measure sufficient to lift the labor productivity trend consistently above the benchmark from 2013 onward. Reversely, for

"Transportation and storage", "Accommodation and food service activities", "Real estate activities", and "Administrative and support service activities" labor productivity indices are the furthest below the "official" statistics, the result of "Transportation and storage" and "Real estate activities" being mostly driven by sluggish turnover growth at least from 2015.

Like it was noticed for time series, for indices there is still an evident break in 2018 which again originates from Eurostat SBS revising its "presentation unit".

#### 8.4 Summary

- Comparable OECD-SBS-BSC data series exhibit a break in 2018 which thwarts comparison with CompNet for that year. Moreover, the benchmark dataset only provides information on remuneration of employees for "*Manufacturing*" and "*Construction*".
- Given overlapping data sources, also thresholds are almost the same across the two databases. That is, to be included are only firms with a turnover of at least 17,500 EURO and at least one employee during a calendar year. Just CompNet "Manufacturing" differs since it covers all firms with more than 19 employees. CompNet users should be aware of this specificity for the "Manufacturing" sector.
- CompNet coverage rates against OECD-SBS are overall satisfactory. Totals for turnover and value added are larger in CompNet than the benchmark as a consequence of differing underlying definitions for the variables.
- Coverage rates for firms numerosity are all extremely close to one, except for "Manufacturing". This is likely due to CompNet containing additional units in that sector with a remarkably low turnover in comparison with the size of their employment.
- CompNet outnumbers the benchmark also for employment and wages. For these variables as well, the explanation likely lies in underlying definitions not entirely converging.
- The 1-digit sectors leading CompNet excessive figures are "Manufacturing", "Construction", "Wholesale and retail trade; repair of motor vehicles and motorcycles", and "Accommodation and food service activities". Users relying on aggregates for these sectors in particular should be careful when formulating conclusions.
- When looking at labor productivity, CompNet outperforms the benchmark at the country level while the reverse holds for "Manufacturing". This is consistent with the sector having additional firms in CompNet with at least twenty employees and extremely low revenues. CompNet labor productivity is also lower in "Transportation and storage" and, only more recently, in "Real Estate activities", "Professional scientific and technical activities", and "Administrative and support service activities".
- Time dynamics for number of firms, turnover, and value added are almost perfectly aligned between CompNet and the benchmark. The rise in employment instead

is faster in CompNet, which in turn hinders growth for labor productivity and accelerates that for wages.

- For "Transportation and storage", "Accommodation and food service activities", "Real estate activities", and "Administrative and support service activities" labor productivity growth is the furthest below the benchmark. Users relying on time series for these sectors in particular should be careful when formulating conclusions.
- Academics, policy analysts, and researchers in general should use Compnet mainly for gathering micro-aggregated moments across and within sectors, countries, and years in order to estimate structural models of firm dynamics and to investigate correlations between variables.

### 9 User Guidance

- The 20e weighted sample should be the preferred sample for analysis since it is most representative of the population and comparable with the OECD-SBS data.
- Comparison between the OECD-SBS data and CompNet data should not be based on variable levels, but rather on trends of the indices constructed.
- Indices and variables in the CompNet dataset are derived using different underlying numbers of observations depending on the sector. The interested user should keep this in mind. Manufacturing and the Retail sector tend to have the highest number of observations, while other sectors like the Accommodation and Food Services sector tend to be missing quite a lot of observations on certain variables.
- Users should use CompNet mainly for analyzing productivity trends across sectors, countries, and years or to gather moments aggregated at the sector level to use as targets for structural models of firm dynamics. Other possible applications of CompNet are correlational studies on factor allocation at the industry level.
- Users interested in time-series analysis or the behavior of individual firms may not find the CompNet data fully suited to their needs.

CompNet was designed to allow users to explore moments of the data distributions as well as the correlations between key variables. For example, one can study the distributions different variables, or relationship between two or more key variables, such as the relationship between productivity and firm size. This is where its strength lies. In this regard it is important to note that these correlations appear to be robust to the data issues discussed in this document.

Below we provide two examples of potential types of analyses. As you will see the inferences from these types of analyses are not significantly affected by any of the underlying data considerations described above.

Figure 34 displays the kernel density for the distribution of the sectoral labour productivity for France's 20e weighted sample. Mean productivity has risen from the years 2007 and 2012 to 2019. The shape of the distribution has remained more or less the same over the years with only a slight change in the spread: the distribution in 2019 seems to have slightly less spread than the ones in 2007 and 2012.



Figure 35 shows a positive correlation between labour productivity and firm size across industries for the 20e weighted sample. That is, industries with large firms tend to be more productive. The regression lines are flatter over time suggesting that the positive correlation is weaker in 2019 relative to that in 2007.

Figure 35: Correlation between average firm size and labor productivity



# 10 Appendix

### 10.1 France









Figure 38: Coverage: Number of persons employed, 20e sample

### 10.2 Portugal



Figure 39: Coverage: Number of firms, 20e sample



Figure 40: Coverage: Revenue, 20e sample

Figure 41: Coverage: Number of persons employed, 20e sample





Figure 42: Value added coverage rate by sector

Figure 43: Revenue coverage rate by sector



### 10.4 Slovenia



Figure 44: Revenue coverage rate by sector

### 10.5 Germany



Figure 45: Number of firms by sector