

CompNet 8th Vintage (20e): A primer on its quality and most appropriate use

1 April 2022

By Filippo di Mauro and Javi Miranda (*)

(*) Inputs by Duy Hoang, Sergio Inferrera, Alessandro Zona Mattioli as well as comments by Eric Bartelsman, Matthias Mertens and Peter Haug are gratefully acknowledged

Motivation

- The Users of the CompNet dataset are widely heterogeneous. Broadly, we recognize two types:
 - Academics, interested in the granularity of the information provided and who want to independently decide how and which part of the dataset to use
 - Policy/economic analysts, interested in comparisons across countries and concerned about macro economic trends, who prefer to use - and quickly - just a subset of the dataset ('Reduced dataset')
- Different users call for different presentation of the same data
- In this presentation, we address the second type of users and show some comparability exercises done with official sources (Eurostat SBS) and across countries in the CompNet dataset
- We focus our analysis on key variables in the CompNet Reduced Dataset (CRD) including turnover, employment, and its derivative labor productivity.
- We want to: 1) show that – overall – the CompNet dataset fits well the official sources, and 2) encourage its use by these type of users, though with the specific provisos we will indicate.
- **Note 1: in this analysis we use the 20e sample for comparability purposes. However, we plan to extend this analysis to the full sample in due course.**
- **Note 2: here we focus on labor productivity. We will be extending a similar analysis to other crucial variables of the dataset, too.**

Main points

1. The CompNet dataset is broadly consistent with Eurostat aggregates trends.
 1. Starting with labor productivity, we report on the specific actions taken when (quantitatively relevant) consistency could not be ensured. This included suppressing certain years or specific sectors from the CRD
 2. We are working to expand the same analysis and data cleaning to the (15 or so) variables included in the CRD
 2. Users are encouraged to fully acknowledge that CompNet is a firm-level based dataset. While consistency with macro trends is important, its value-added strength lies on the analysis of distributions. E.g. how skewed or concentrated are productivity developments (firm distribution analysis)? how does this vary across sectors within one or across countries? How is this skewness related to labor costs and capital? How these relationships change overtime?...
- Even after the cleaning, Users are thus encouraged to use it for that purpose and sort of questions.
- If used for more aggregated – sectoral time-series - analyses, one should avoid comparison to published levels, but rather concentrate on trends (year-on-year dynamics).

Objectives

- The ultimate objective of this presentation is to reassure members that CompNet is a high standard and solid dataset, also when taking a macro perspective
- We are fully aware that a broad match of macro trends is critical for the credibility of our dataset and for its use for policy
 - We work constantly on such broad consistency!
- Of course, we do not expect a perfect match with macro data, e.g. Eurostat. The most obvious reason is that the latter are subject to a number of statistical adjustments standard in the production of official statistics.
- Should important mismatches – although overall limited – take place, it is our objective to:
 - Highlight and comprehend the underlying reasons
 - When possible, come up with and implement actual corrections
 - Encourage Users to also engage their respective NSIs - together with us - to come up with explanations/solutions

Tools

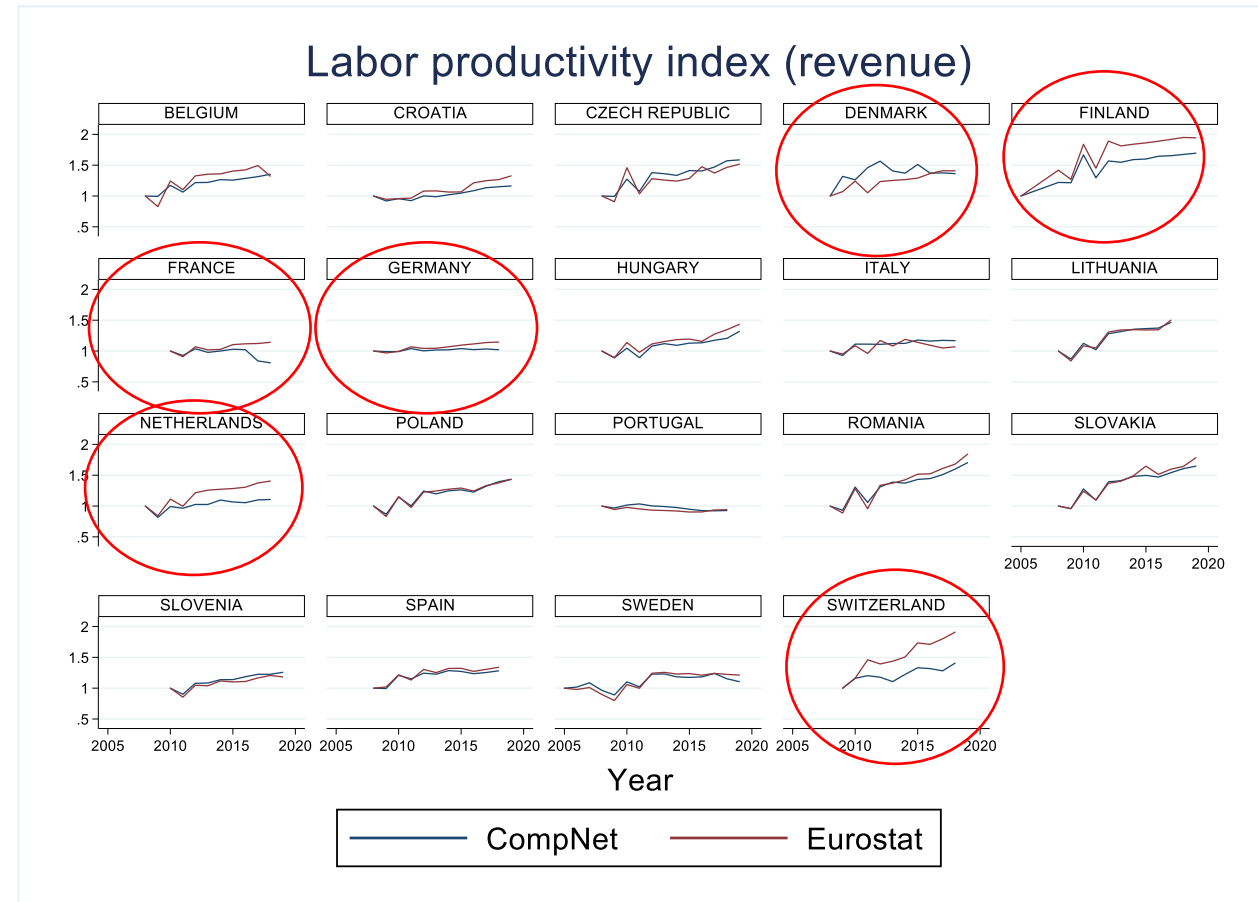
- We use Eurostat published statistics as benchmark for the corresponding variables in our Dataset, across the 19 countries included in the 8° vintage
- We concentrate here on labor productivity. Aside being one of the most utilized indicator, it is also a ratio of revenues and employment, whose data we will also separately analyze to provide further insights on our dataset
- We look at both (i) the time series of the above variables as well as (ii) the evolution of the respective distributions to unveil possible 'suspicious' dynamics
- The mismatches that we are going to underline are either
 1. directly corrected when easily identifiable and large in magnitude, or
 2. Highlighted and sent for a deeper analysis to our 'Data quality Team', who is preparing 'Data quality country fiches', to guide detailed and ongoing dialogue with individual data providers.
- We will provide the code to replicate all the results in this PowerPoint. We encourage Users to adopt the code and incorporate suggested best practices also in their own analysis.

Labor productivity index (before correction)

- We start by studying labor productivity, computed as a ratio between total revenues and total employment at the sector level.
- Although informative, Labor productivity level is not something we aim at reproducing*; we are more concerned with the trend and the underlying distribution → **Results on the level in the appendix**
- To better understand whether we broadly match macro information, we look therefore at developments in the labor productivity index, computed as follows
 - $$\text{Index} = \frac{\text{Labor productivity}_{c,t}}{\text{Labor productivity}_{c,\text{first year}}}$$
- There are some differences between CompNet and Eurostat in the growth rates, mostly coming from latter years. See red circles in the charts

Results

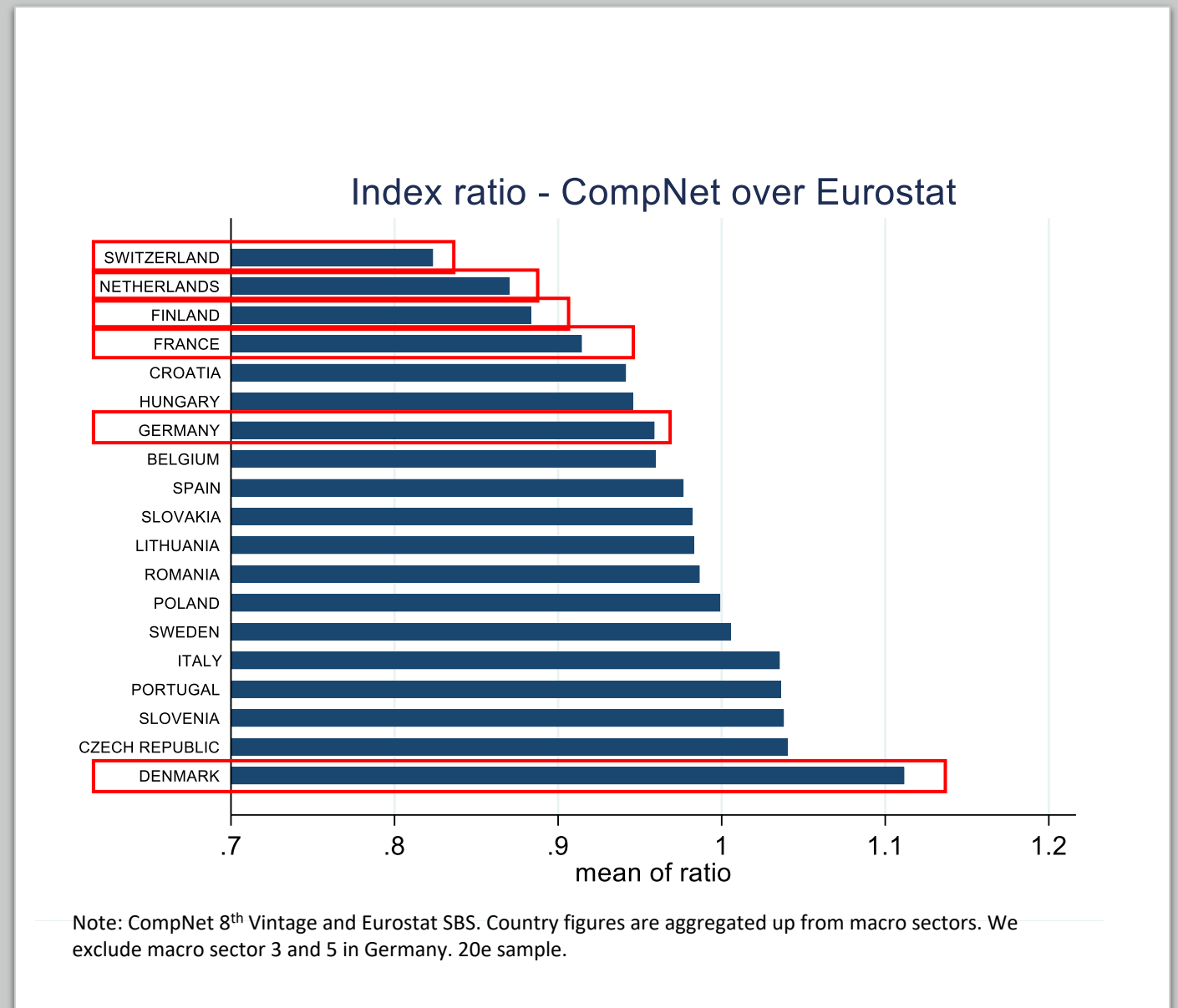
- Six countries have problematic and quantitatively relevant departures from the macro trends (see red circles), namely :
 - Denmark, Finland, France, Germany, Netherlands, Switzerland.



Note: CompNet 8th Vintage and Eurostat SBS. Country figures are aggregated up from macro sectors. We exclude macro sector 3 and 5 in Germany. 20e sample.

Index ratio – comparison with Eurostat

- To summarise the extent of the matching (CompNet-Eurostat) we create an index of adherence to macro trends, computed as the average (over years) ratio between CompNet and Eurostat in terms of labor productivity
- All in all, there are only FOUR countries that depart the most (over 10%) from official aggregate statistics, namely: **Switzerland, Netherlands and Finland** (for which we have lower values), and **Denmark** (for which we have larger values)
- We will inspect such countries (plus Germany and France) in greater detail in the next part of this presentation, looking at the elements of labor productivity, namely labor and total revenues.



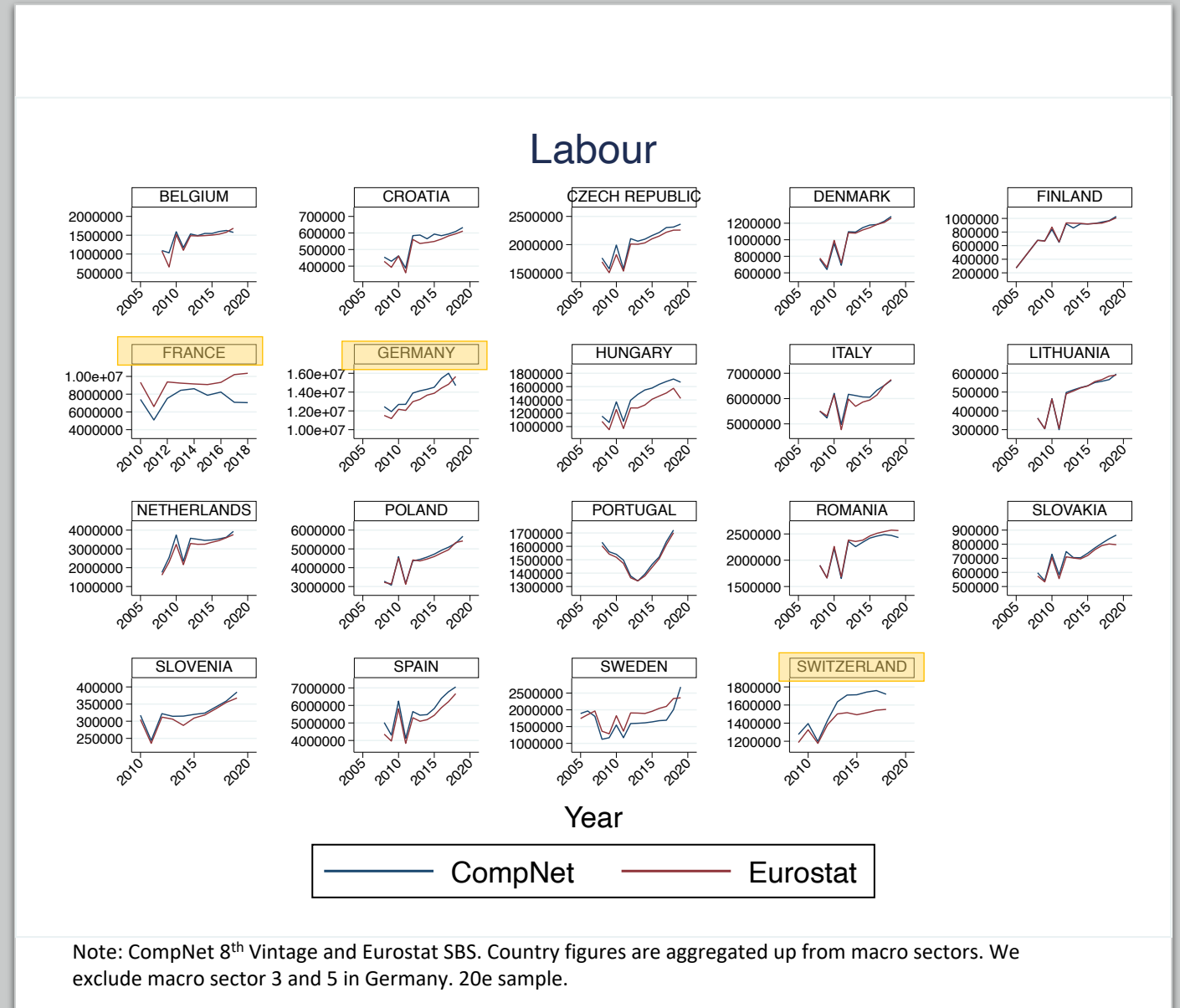
Labor

Results

- Employment moves in the same direction in most cases.
- There is no significant difference between CompNet and Eurostat: the two series are virtually identical
- Divergence (among the 6 countries considered) included:
 - Lower level in France
 - Larger level in Switzerland and Germany
 - Break in the latest available years in Germany and France

Actions

- Exploring reasons and possible corrective actions specially when there is divergence.
- The difference in levels can be addressed through easy to implement level shifts if matching this series is necessary. As before we worry mostly about diverging trends not levels. For this reason working with indexes might be more useful.



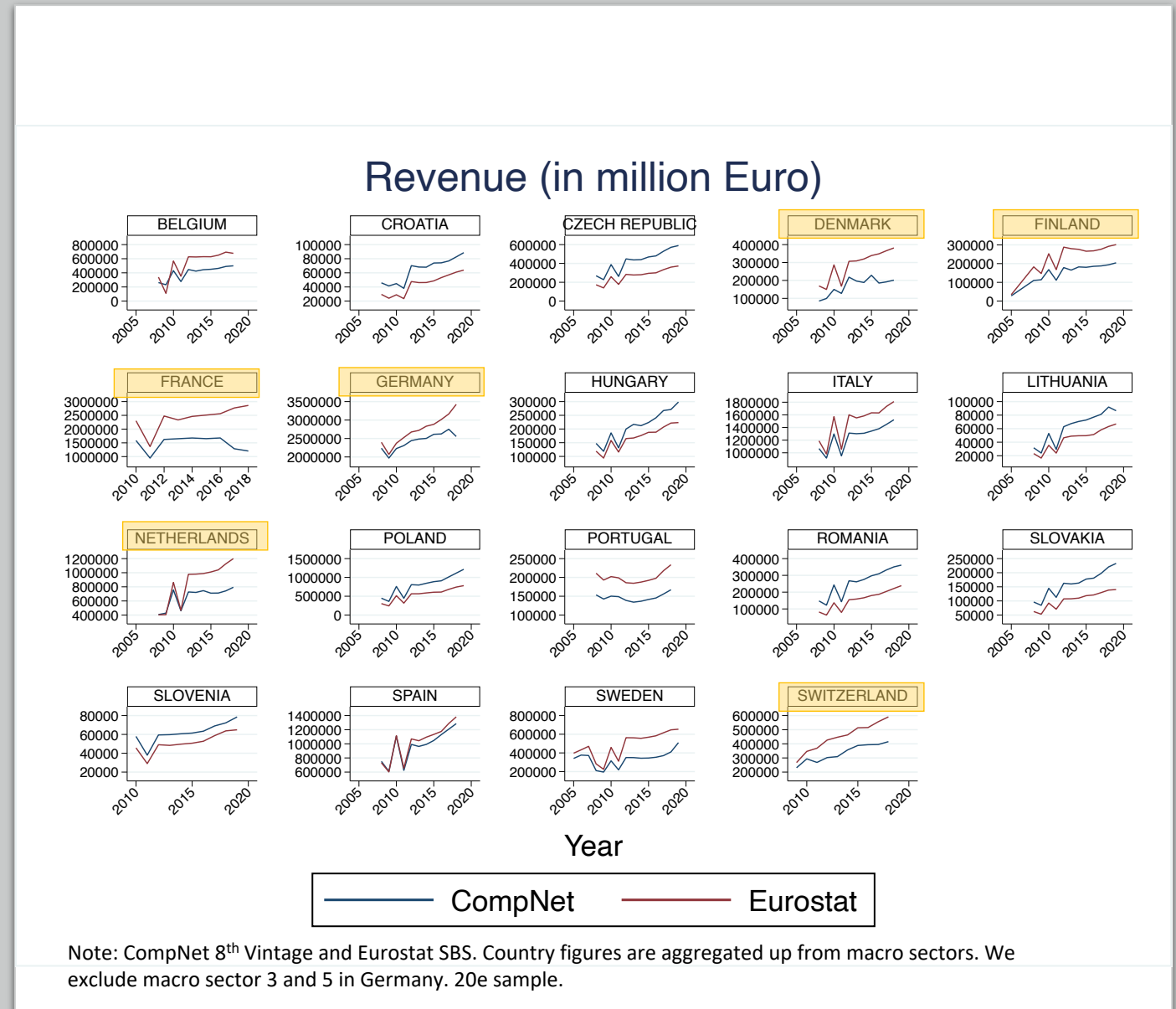
Revenue

Results

- Revenue numbers are different in CompNet and Eurostat, but the series typically comove.
- All the six “problematic” countries have some issues related to revenues
- Several things may explain these jumps: compositional shifts, firms' coverage, divergent patterns within a size-class or a sector, or poor performance of the weighting scheme

Actions

- The small discrepancies in levels can be addressed through weights, while more problems may arise when the trends are divergent (countries highlighted)
- To assess the divergent trends, we use below three strategies:
 1. Using demeaned series, to understand whether we have compositional shifts
 2. Analysis of the productivity distribution and of size class coverage
 3. Analysis of the productivity distribution and macro-sector coverage



(1) Demeaning

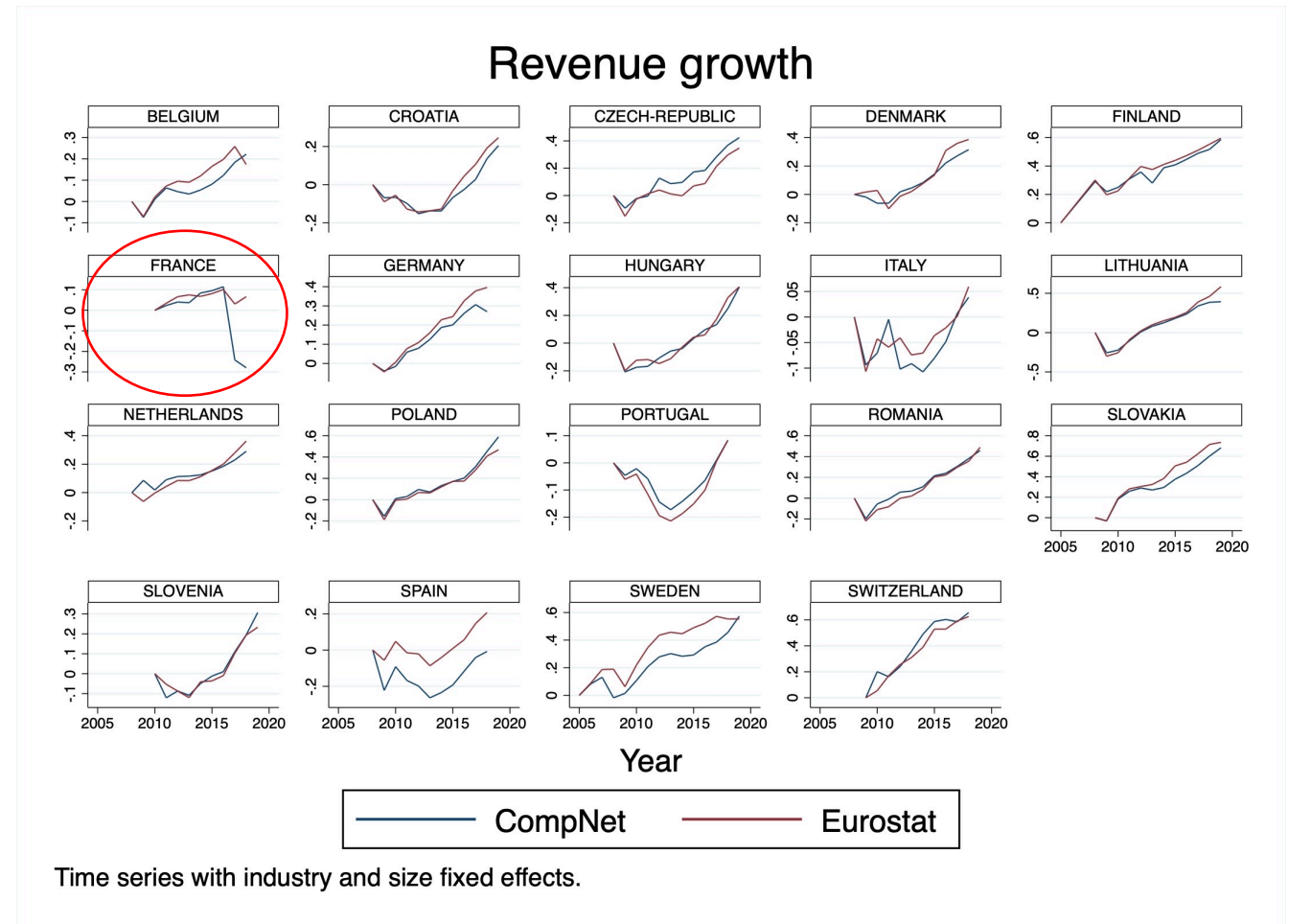
→ Demeaning the time series by industry and size class averages can solve the issue if compositional shifts are causing it

Results

- The series cleaned from these fixed effects is much cleaner! No issues in compositional shifts, except for France and Germany in latest years
- Another small divergence is spotted in Spain. However, it does not raise any issue: aggregate numbers match pretty well and macro-sector analysis (in the appendix) does not uncover any problem

Actions

- We advise users to utilize such series when looking at growth trends in output
- Pairwise, we advise **not to use** the last two years of observations for France



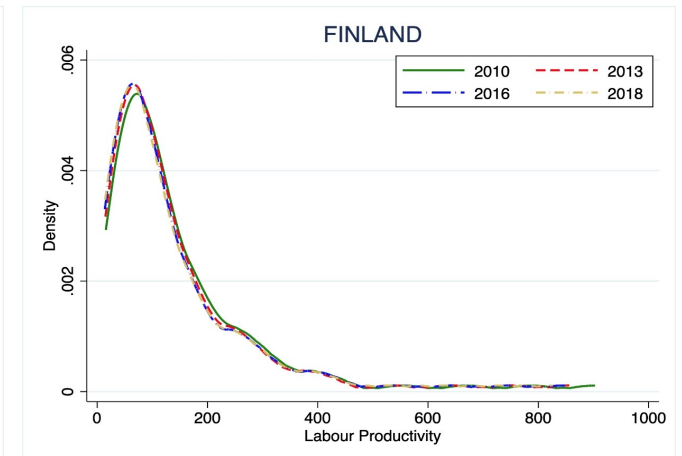
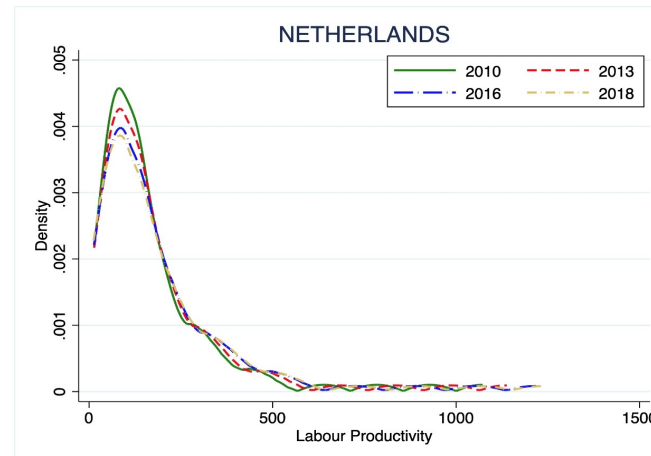
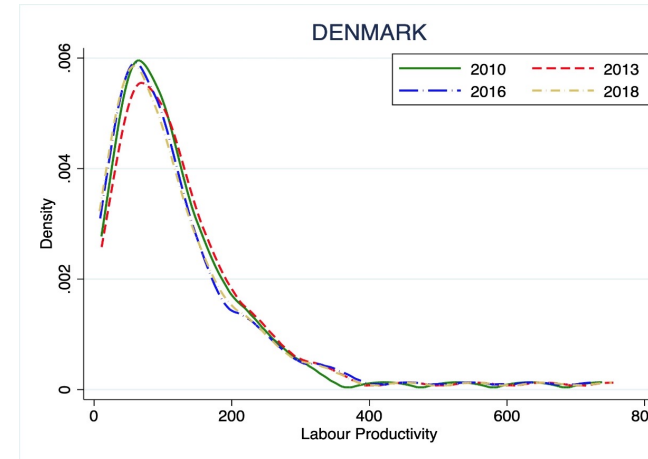
Note: CompNet 8th Vintage and Eurostat SBS. Country figures are aggregated up from macro sectors. We exclude macro sector 3 and 5 in Germany. 20e sample.

(2) Labor productivity distribution

- One possible concern with industry/size compositional shifts is that they may disturb the underlying productivity distributions and thus pose a risk to the inferences obtained from the data. We expect stable distributions over time with fairly small shifts as a result of economy wide shocks (e.g. business cycle, COVID)
- Hence, we plot the distribution of labor productivity for few of the problematic countries

Results 1

- For Denmark, Finland, Netherlands the change in the productivity distribution is virtually null.
 - ➔ Great sign upon first approach on the quality of the dataset but we will keep exploring



Note: CompNet 8th Vintage and Eurostat SBS. Country figures. 20e sample.

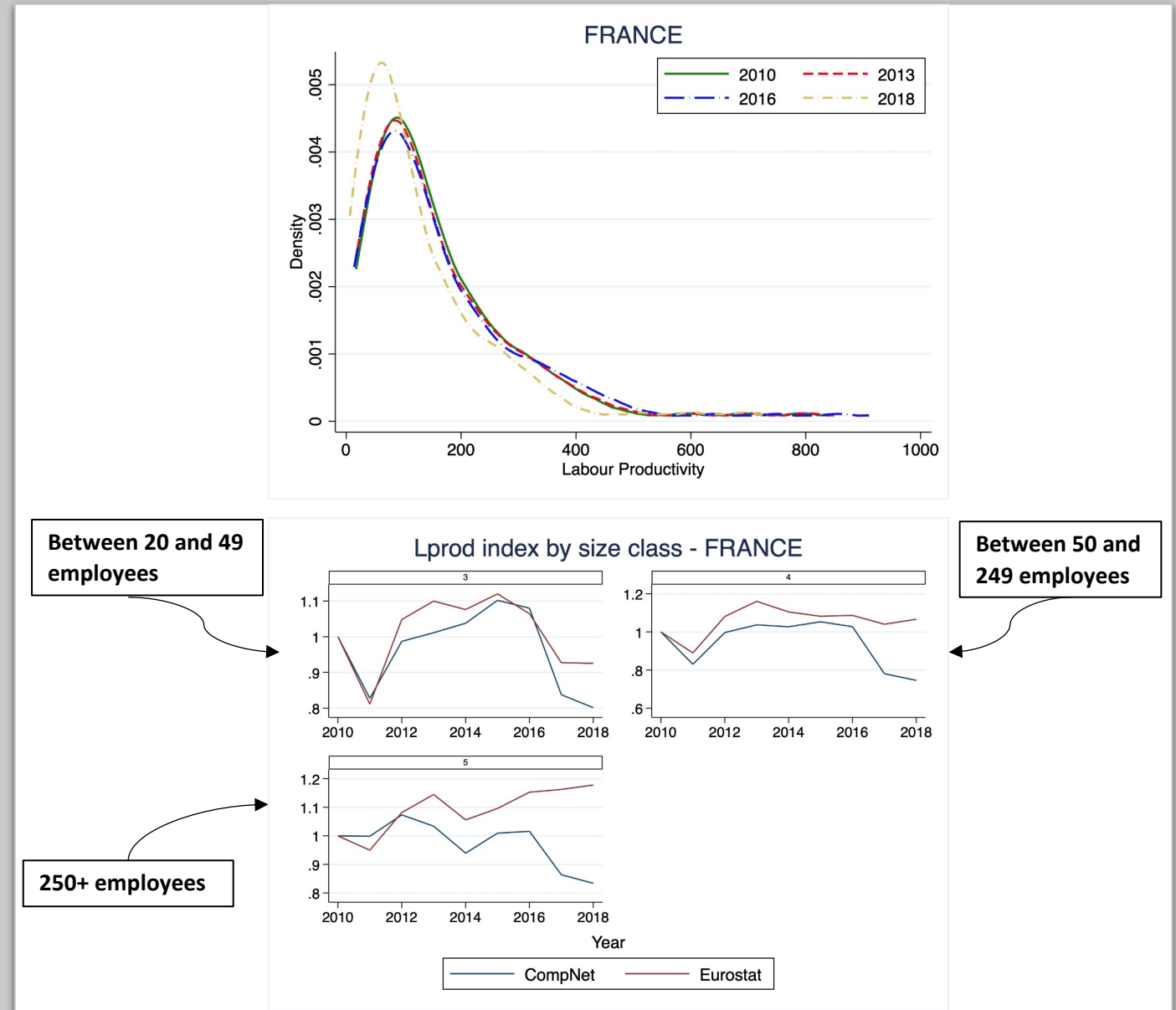
(2) Labor productivity distribution - France

Results 2

- On the other hand, in France the productivity distribution shifts abruptly to the left in 2018 (kernel chart)
- In this year, the coverage of large productive firms decreases! (See charts by size class)
 - In the appendix, you can find a general decrease in the number of firms in France. We argue that these are the largest and most productive ones
- The issue is clearly a sample bias

Action:

- Users are advised to eliminate 2017 and 2018 from their analysis when using France
- The CompNet Data quality group will engage INSEE to explain the nature of such departures



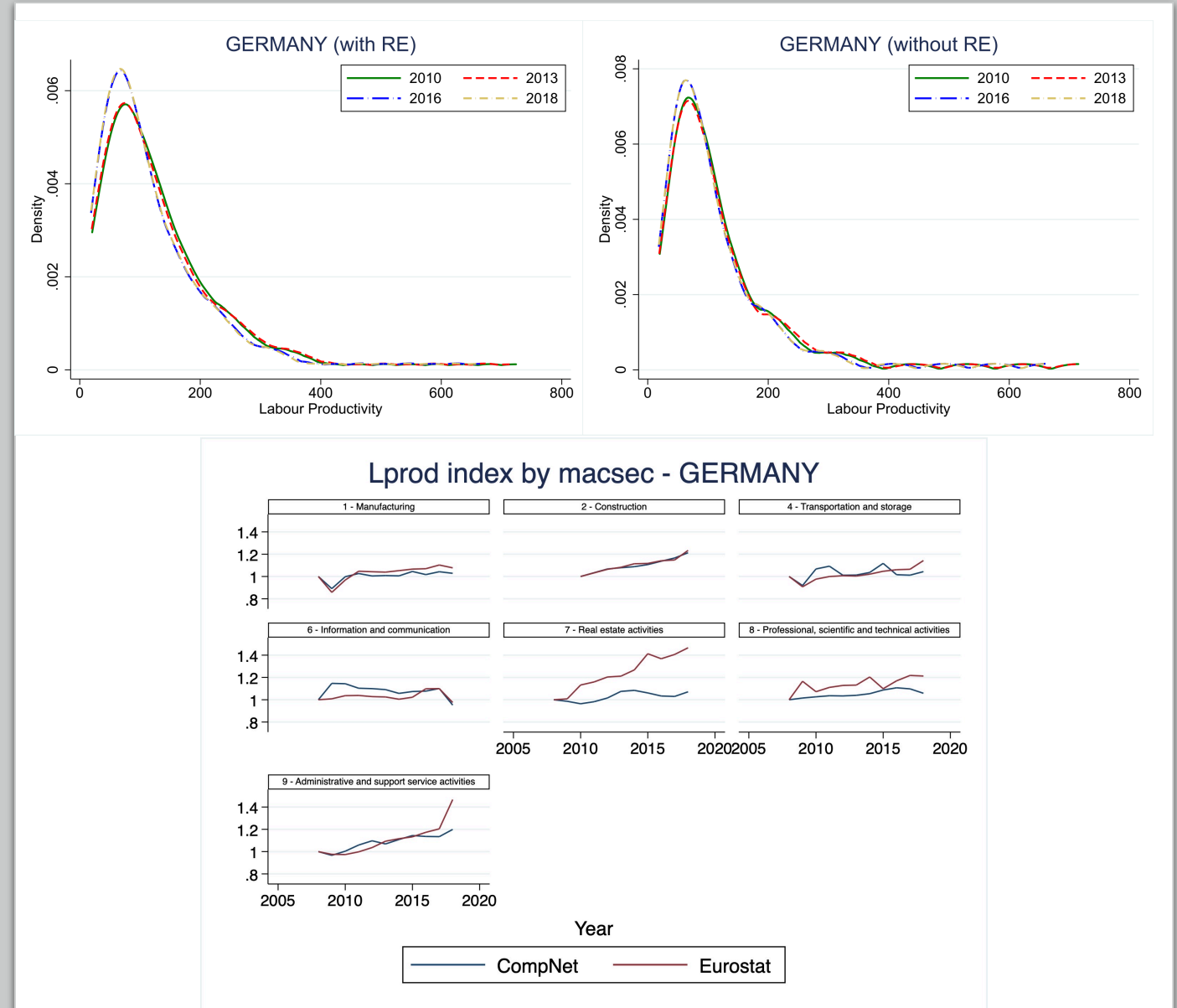
(2) Labor productivity distribution - Germany

Results 3

- Productivity distribution in Germany shifts inwards throughout time
 - This explains the low productivity growth observed throughout time
- In this case, specific sectors may be driving such divergence with Eurostat
 - Real Estate is problematic in Eurostat, in view of its huge productivity growth

Action:

- Users are advised to eliminate Real Estate sector from their analysis when using Germany
- The CompNet Data quality group will engage Destatis to explain the nature of such departures



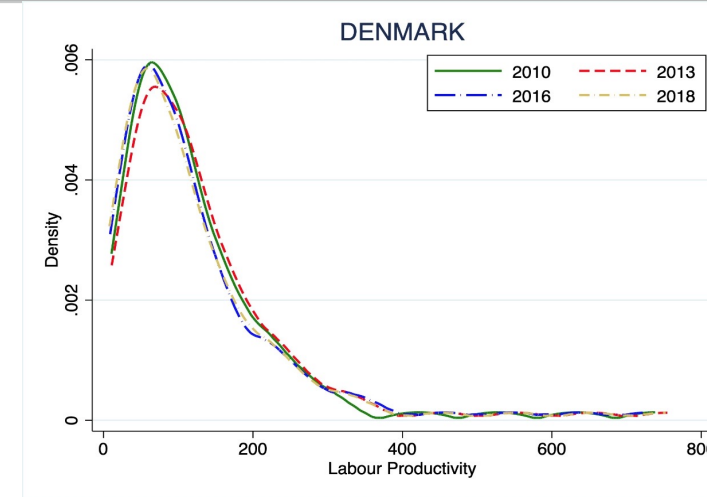
(2) Labor productivity distribution - Denmark

Results 4

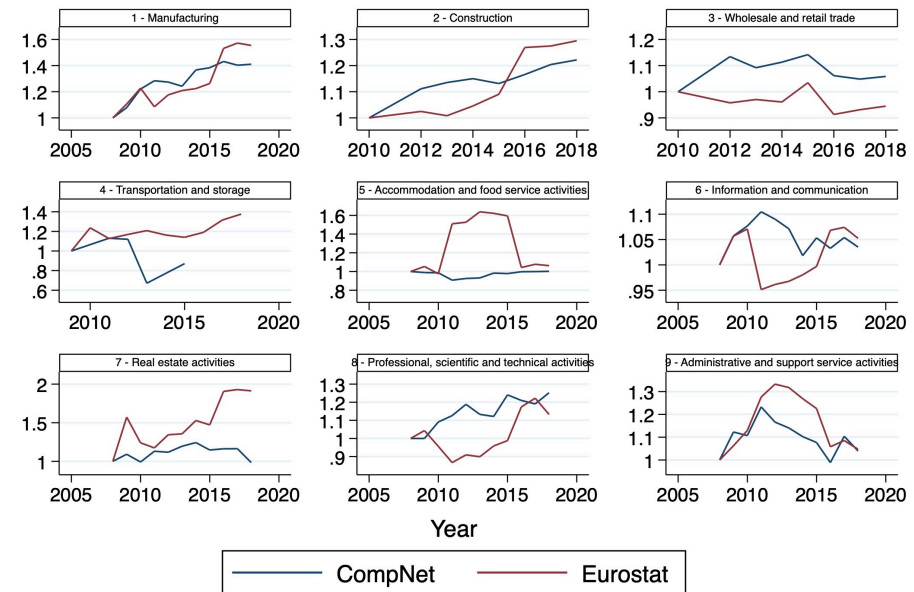
- Productivity distribution dynamics in Denmark appear less stable than in other countries
- Specific sectors may be driving this as shown by divergences with Eurostat
 - Transportation and Storage, and Accommodation and Food Service Activities

Action:

- Users are advised to eliminate such sectors and exercise caution when making inferences based on time trends
- The CompNet Data quality group will engage Statistics Denmark to explain the nature of such departures



Lprod index by macsec - DENMARK



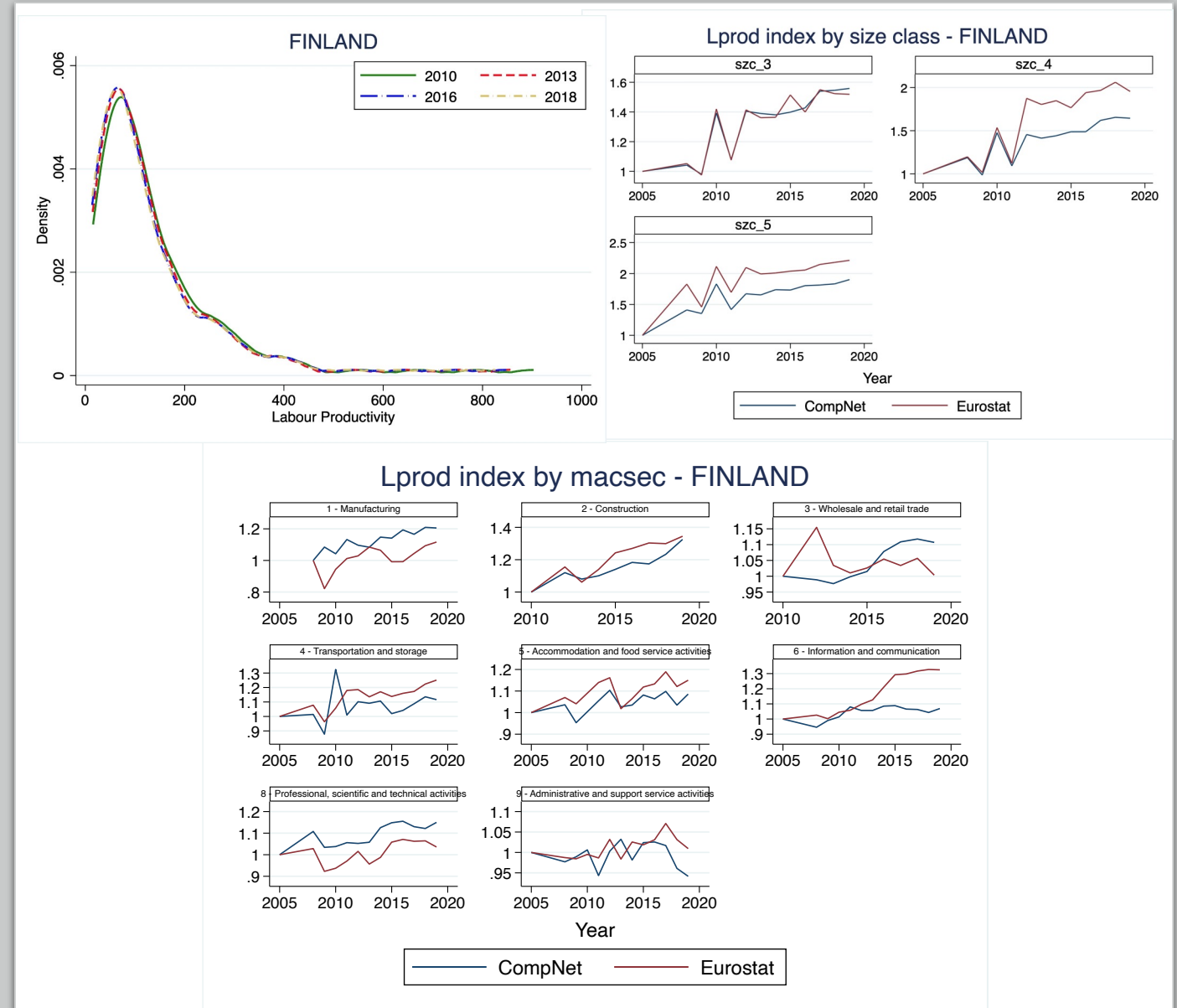
(2) Labor productivity distribution - Finland

Results 5

- In Finland, the productivity distribution is rather stable over time
- However, in aggregate Eurostat shows larger productivity growth
- The sectorial analysis shows that this comes from all sectors, except Professional, scientific and technical Activities & Manufacturing
- Furthermore, analysis of size-class evolution of productivity shows that the larger productivity growth is due to higher growth in Eurostat among the largest firms (firms with more than 50 employees)

Action:

- The CompNet Data quality group will engage Statistics Finland to explain the nature of such departures



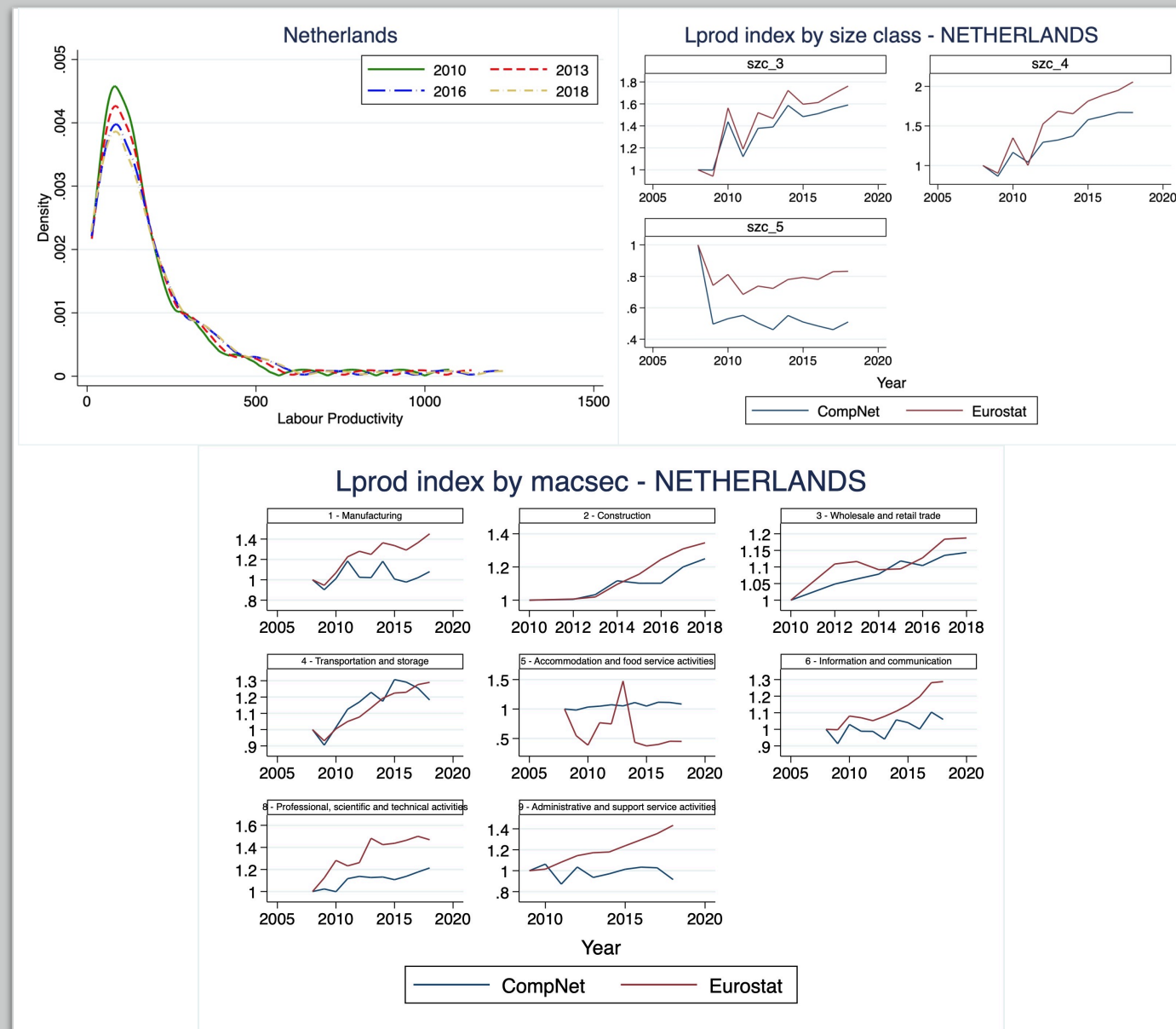
(2) Labor productivity distribution - Netherlands

Results 6

- In the Netherlands, the productivity distribution is rather stable over time, but Eurostat shows larger productivity growth
- The sectorial analysis shows that this issue is shared across sectors, with the notable exception of the Accommodation and food service activities & Transportation and Storage
- Size-class analysis shows that the discrepancy is relative to the whole size distribution and not to particular size-classes

Action:

- The CompNet Data quality group will engage Statistical Netherlands to explain the nature of such departures



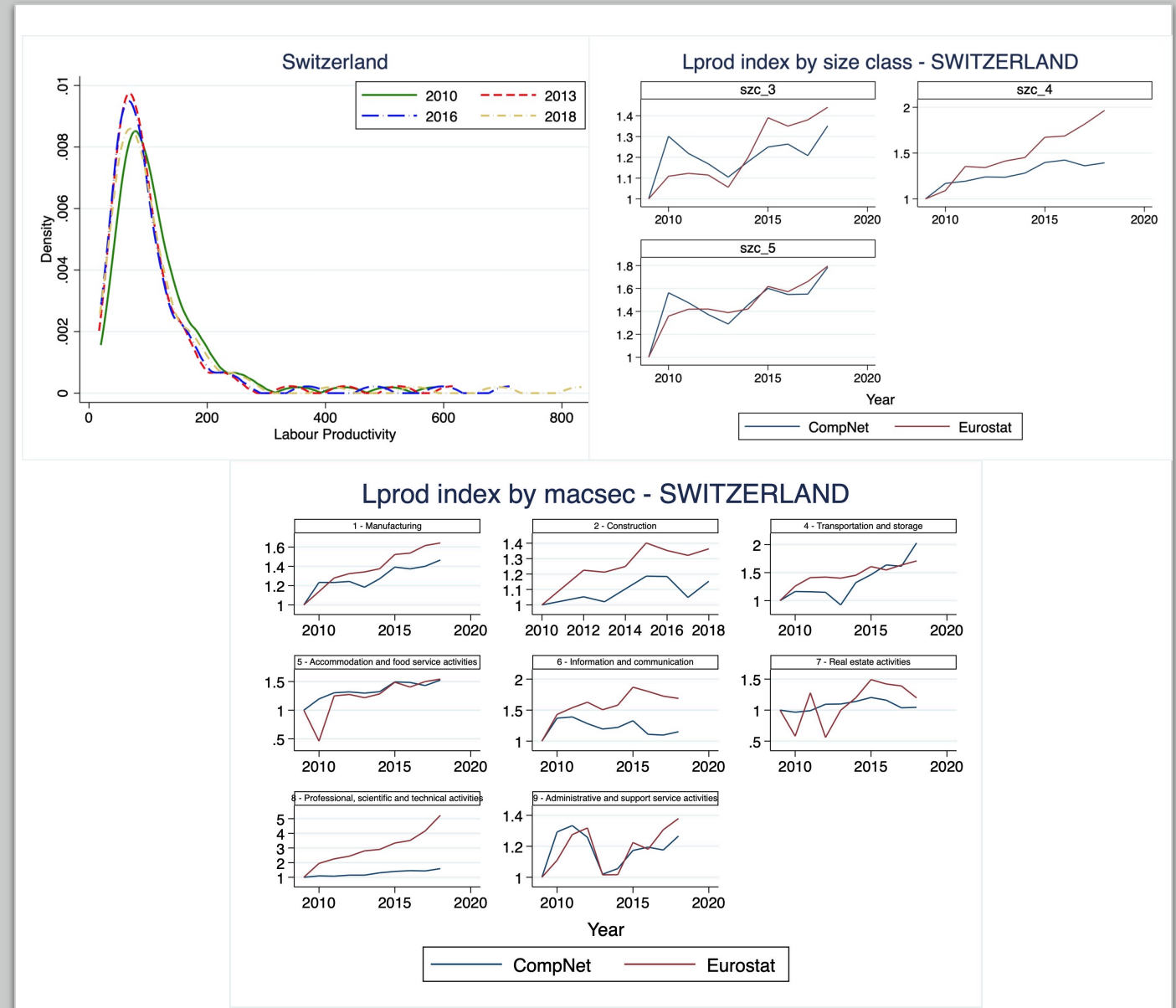
(2) Labor productivity distribution - Switzerland

Results 7:

- Productivity in Switzerland shifted throughout time
- The source of such abrupt movements can be found in specific sectors, namely Professional and Scientific Activities (8) and Information and Communication (6)

Action:

- The CompNet Data quality group will engage Statistics Switzerland to explain the nature of such departures



Labor productivity index (after correction)

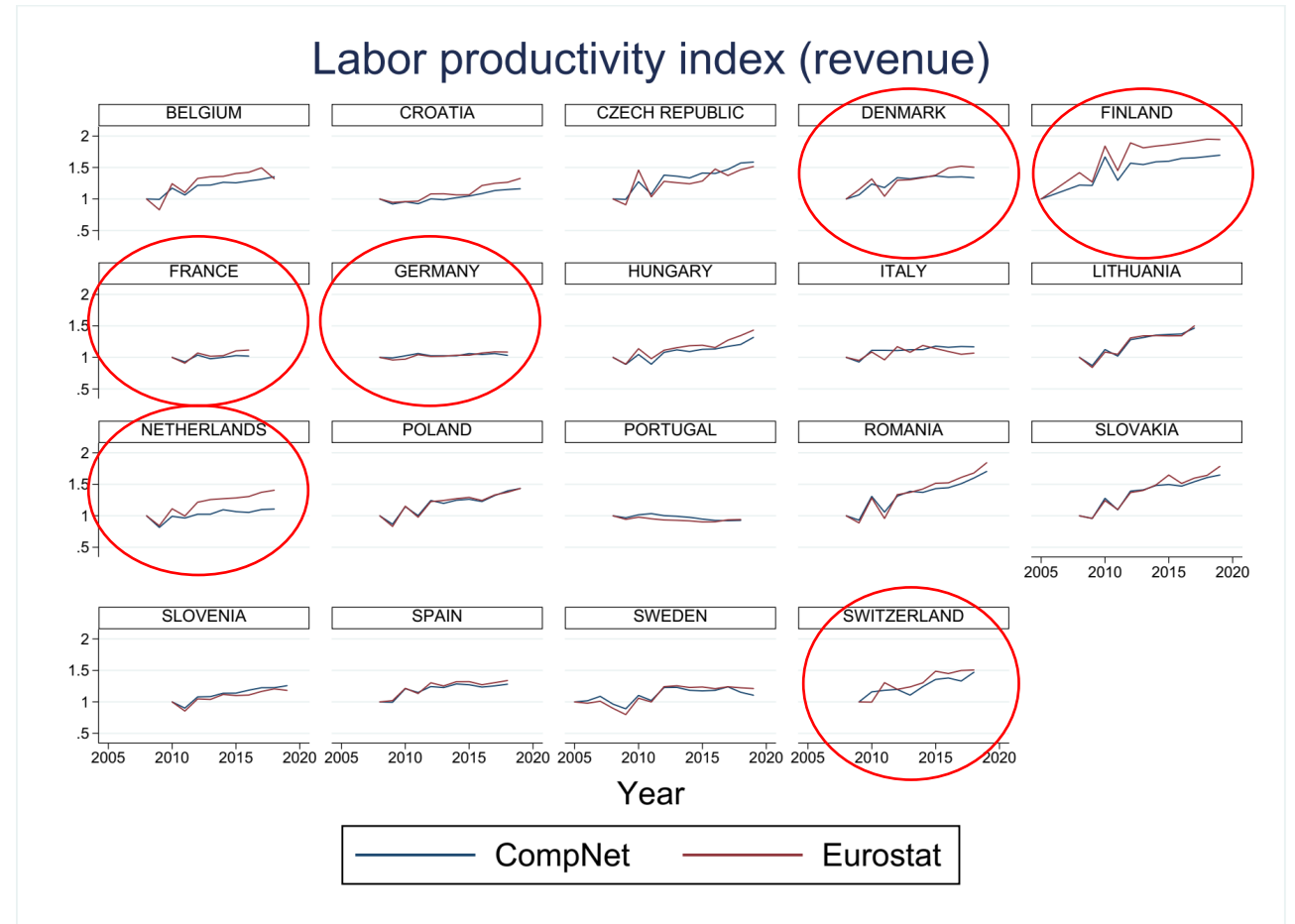
- The comparability improves tremendously when few problematic data points are removed

Results

1. In most of the cases we have strikingly similar movements
2. Furthermore, when trends do not match, the divergence is quantitatively not important.
 - We are working to address such divergent trends, identified in the earlier charts

Actions

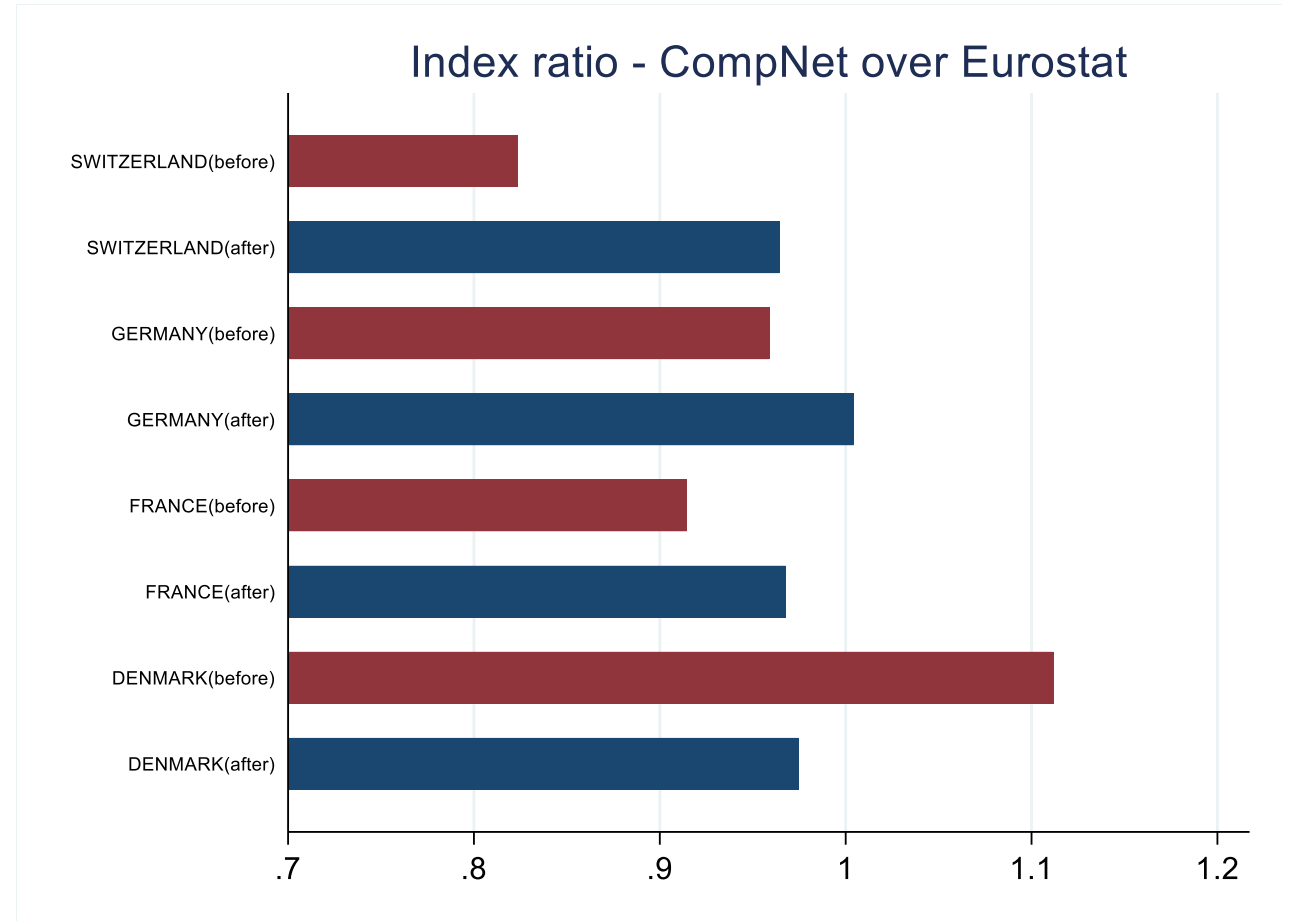
- Finland and Netherlands still show suspicious growth but after our check, these are not problems from specific sectors or years. These countries require further investigation.



Note: CompNet 8th Vintage and Eurostat SBS. Country figures are aggregated up from macro sectors following the recommendations provided in this presentation. 20e sample.

Country specific index ratio (before- and after-correction)

- The above results are also confirmed by comparing the above computed **index of adherence to macro trends** (slide 7) before and after the proposed corrections.
- All in all, the following procedures are advised to Users
- **Denmark**
 - Drop the Transportation and Storage, and Accommodation and Food Service Activities sectors
- **France**
 - Drop year 2017 and 2018
 - **Germany**
 - Drop Real Estate sector
- **Switzerland**
 - Drop Professional and Scientific Activities, and Information and Communication sectors



Note: CompNet 8th Vintage and Eurostat SBS. Country figures are aggregated up from macro sectors following the recommendations provided in this presentation. 20e sample.

Conclusion

- In this presentation we have highlighted some of the departures from the Eurostat values for labor productivity, identified possible causes, implemented a small set of adjustments (edits to the data), and proposed analytical methods that yield robust inferences when looking at trends
- We have shown that the productivity distributions are relatively stable across time, indicating that there are no systematic biases
 - France is an exception in the last years
- Both of these things imply that the dataset is very reliable when used for its most important features (analysis of the distributions) but also for the analysis of trends when using robust methods
 - We encourage member institutions to use the data in their analysis with confidence but to be aware that strict level comparisons with Eurostat should be avoided.
 - Furthermore, we encourage you to provide us with your feedback as you become more comfortable using the data so that we can continuously improve the dataset (e.g. checking odd movements as well as adding new variables, modules, etc.)
- The Network is now extending this analysis to all the variables included into the 'reduced dataset', and release additional corrected series in the next few weeks
- We will pursue further training on the best use of the data

Complete chartbook

You can take a look at the complete chartbook at this [link](#).

Charts include figures on growth and levels (divided by macro-sector) on:

- Revenues
- Employment
- Labor productivity

The title of the file specifies the variable plotted, whether it is plotted the index or the level, and the country used to plot the data

Appendix

Labor productivity level

- We start by studying labor productivity **level**. Although informative, Labor productivity **level** is not something we aim at reproducing; we are more concerned with the trend and the underlying distribution
- Labor productivity = $\frac{\text{Total revenues}_{s,t}}{\text{Total employment}_{s,t}}$

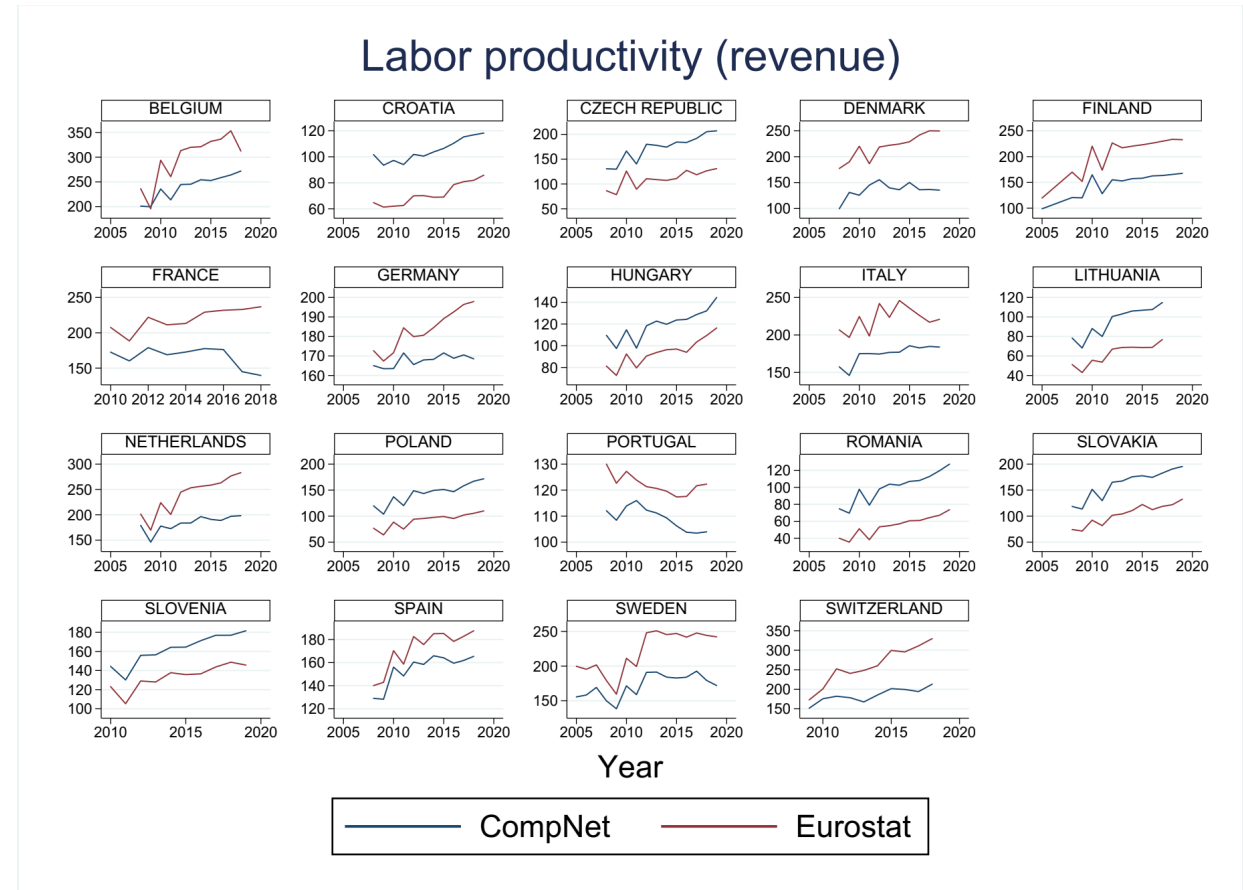
Results:

1. There are differences in levels for all countries,
2. The time series however tend to follow the same trend.
3. No specific pattern is detected on whether CompNet is larger or smaller than Eurostat

Actions

- Labor productivity levels are not fully informative: data aggregated blindly from firm-level information can reasonably diverge from official statistics, that undergo extensive post-production cleaning
- Better therefore to look at an aggregate index (see next slide) to look at **respective trends**.

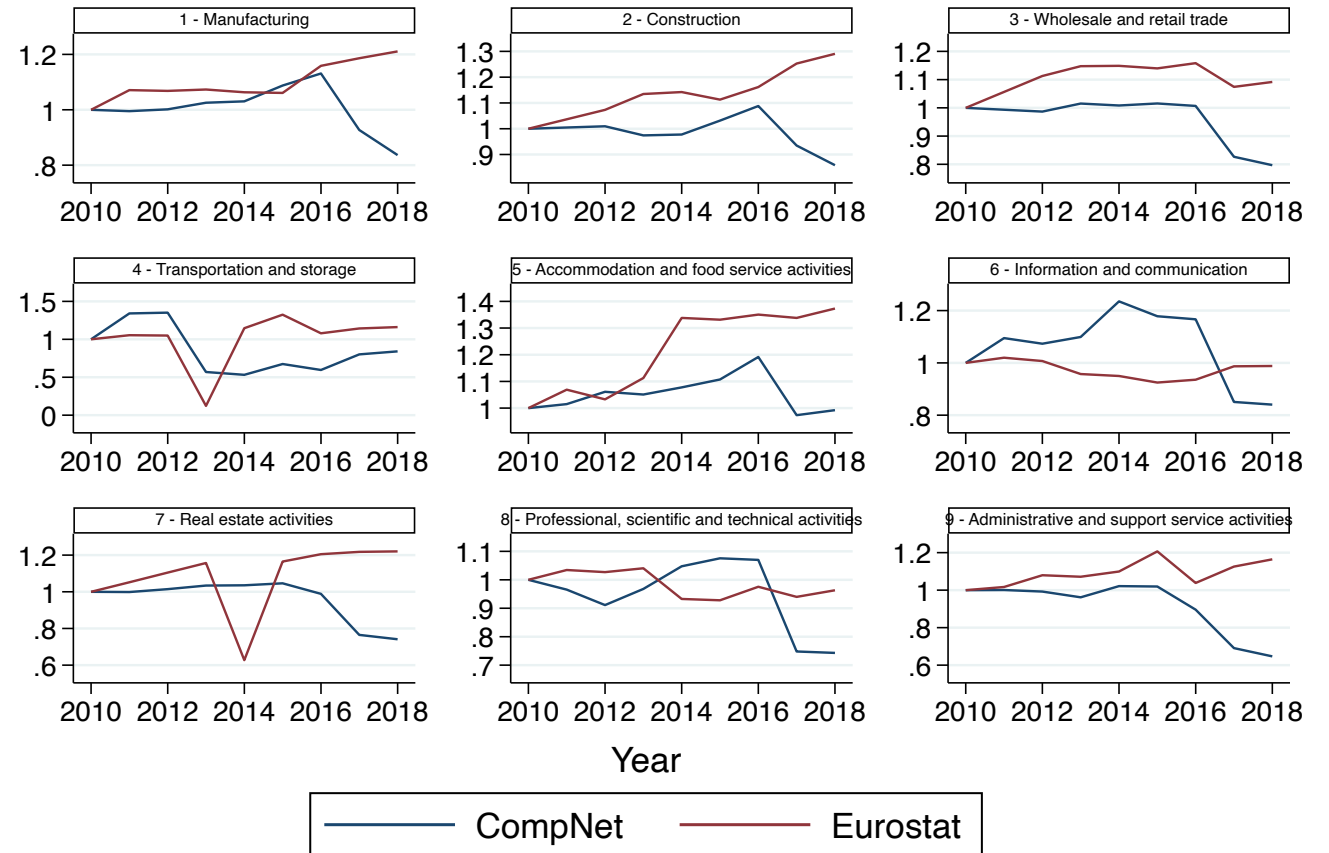
*We use aggregate nominal revenues at the macro-sector level



France

- France has identical patterns, until 2017 when CompNet shows a significant drop across sectors.

Lprod index by macsec - FRANCE

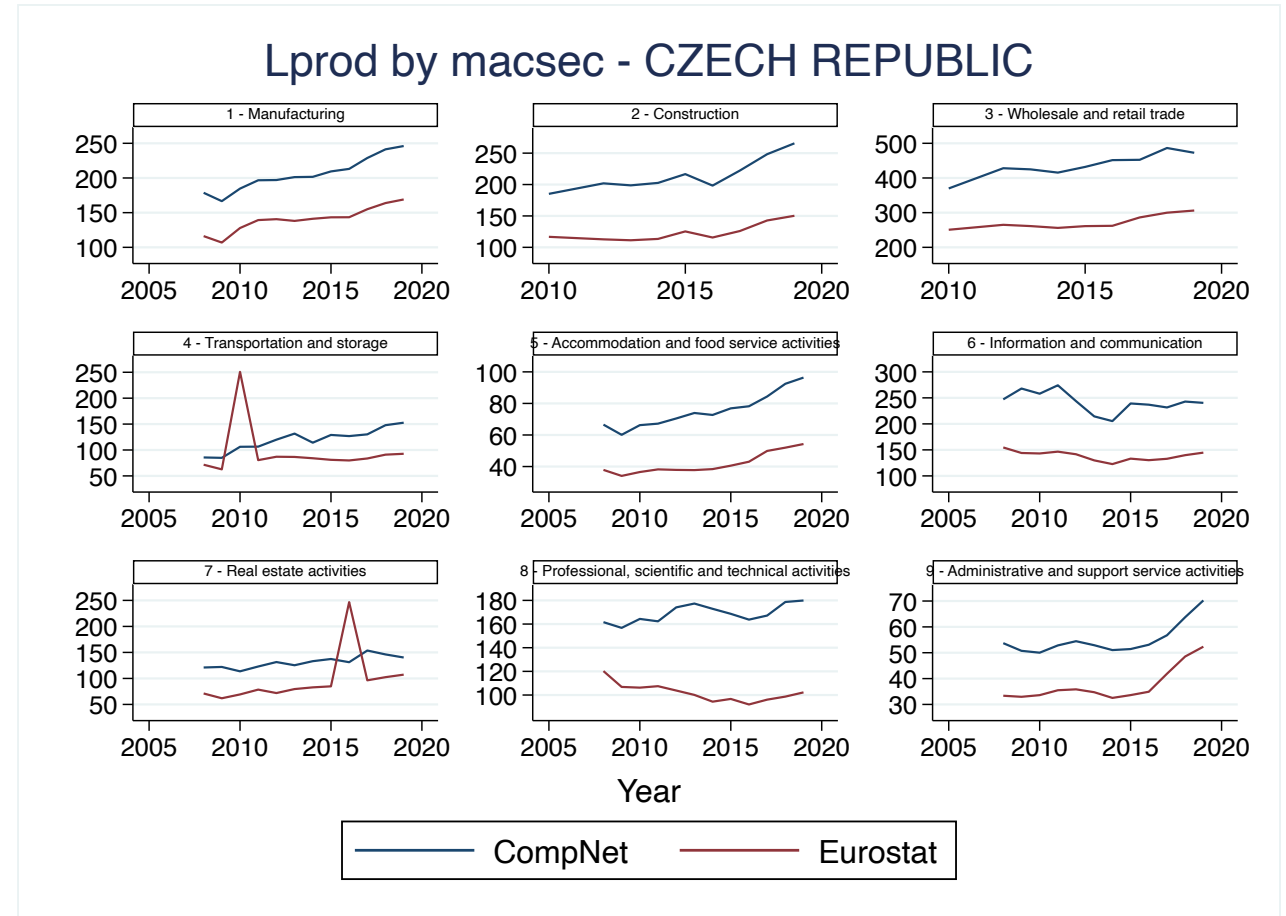


- France has a problematic definition of firm. This is reflected in the distribution of age across time (look at changes in 2008 and in 2017)
- The jumps in the p1, p50, and p99 of the age distribution are highly suspicious

	country	year	OV00_firm_age_p1	OV00_firm_age_p50	OV00_firm_age_p99
1	FRANCE	2004	0	7	34
2	FRANCE	2005	0	7	35
3	FRANCE	2006	0	7	36
4	FRANCE	2007	0	7	36
5	FRANCE	2008	0	10	53
6	FRANCE	2009	0	10	53
7	FRANCE	2010	0	10	54
8	FRANCE	2011	0	10	54
9	FRANCE	2012	0	10	55
0	FRANCE	2013	0	10	56
1	FRANCE	2014	0	9	57
2	FRANCE	2015	1	11	58
3	FRANCE	2016	2	12	59
4	FRANCE	2017	9	18	61
5	FRANCE	2018	9	18	62

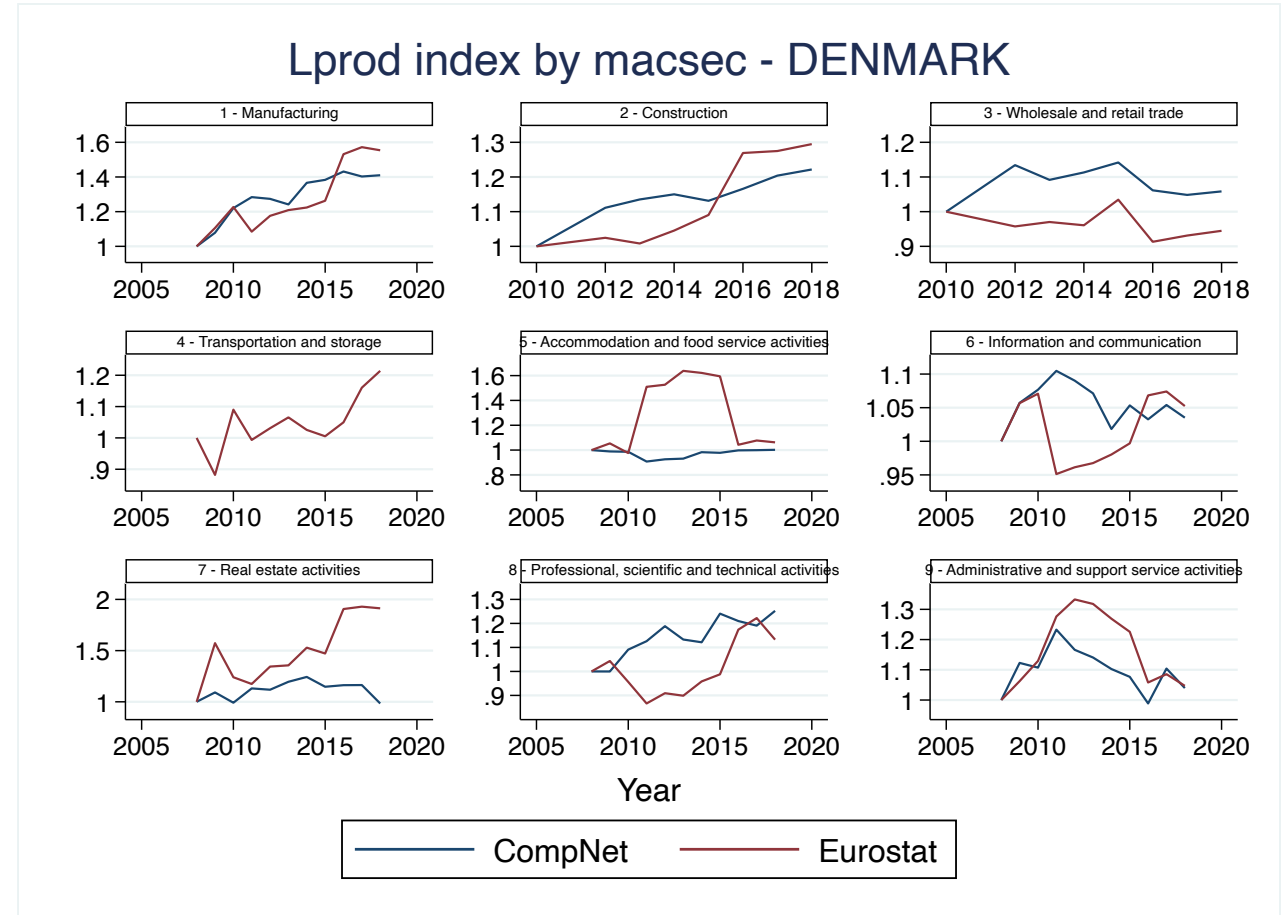
Czech Republic

- CompNet is always larger than Eurostat
- Does this come from the smaller number of firms in the sample?
- Despite the jump, patterns still move in the same direction, but it's still a notable difference.



Denmark

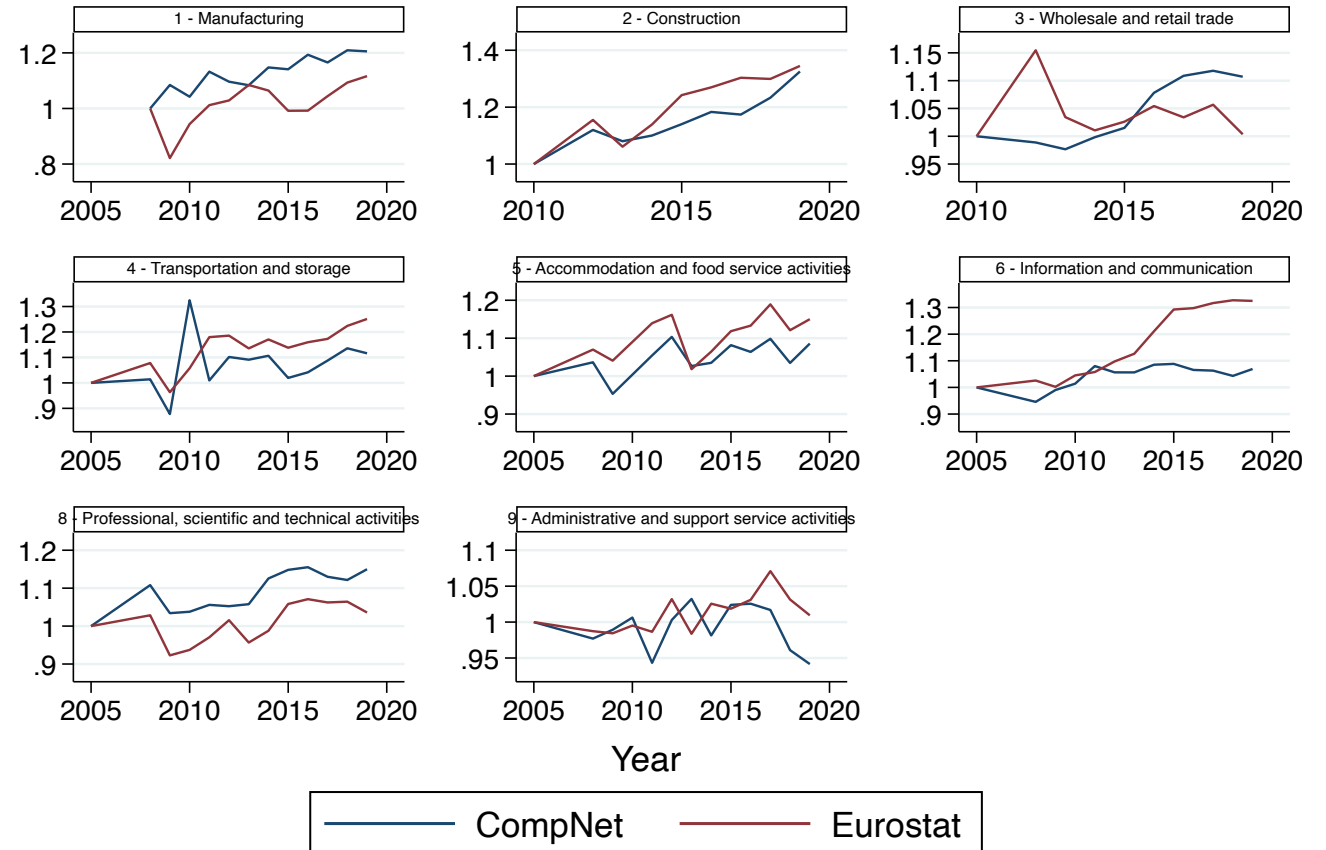
- Jump in Denmark particularly odd in Eurostat in Accommodation and Real estate
- Manufacturing and Construction behave well
- Positive trend in CompNet in wholesale, whereas it is negative in Eurostat



Finland

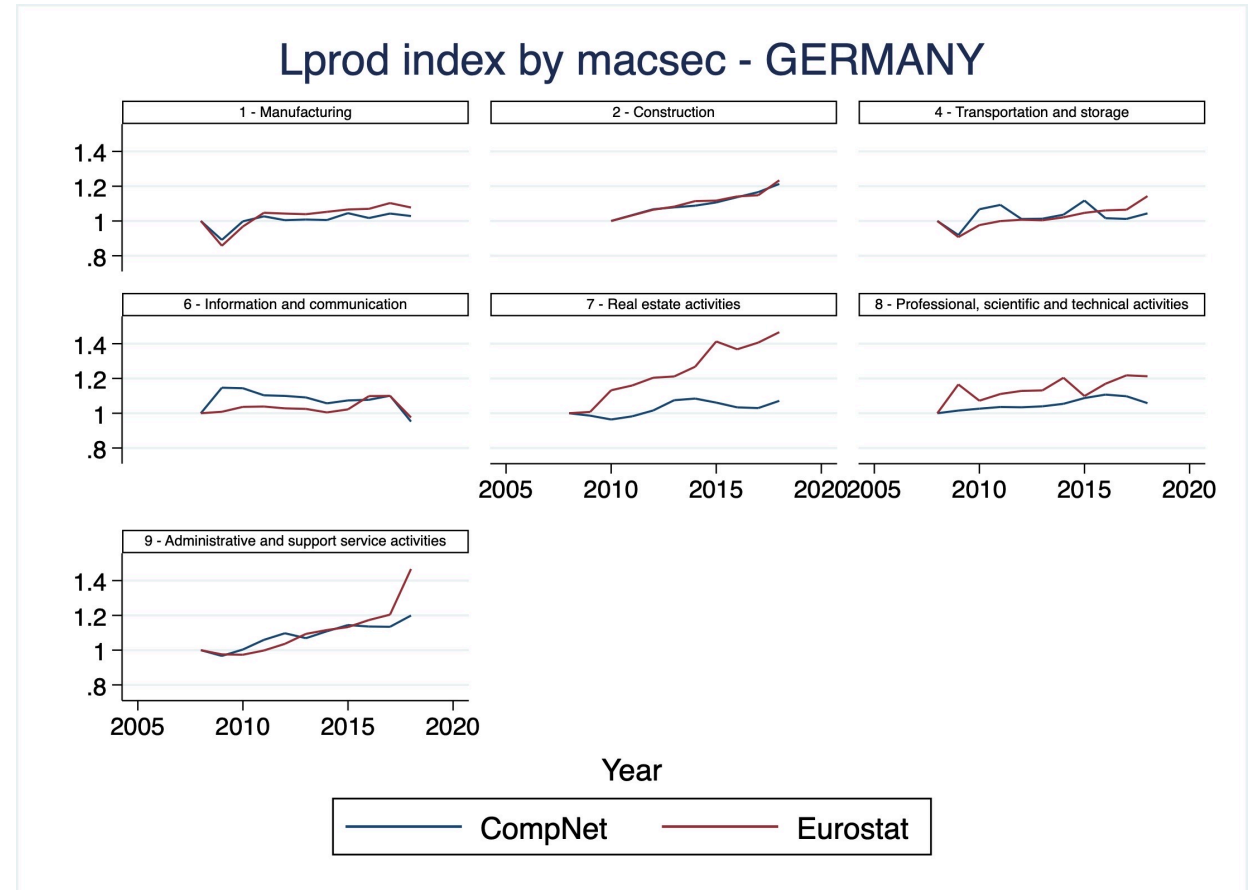
- Eurostat has higher growth (but almost identical pattern at national level)
 - Manufacturing and Prof, Sci activities are exceptions
- This is clearly not a sector-specific problem.

Lprod index by macsec - FINLAND



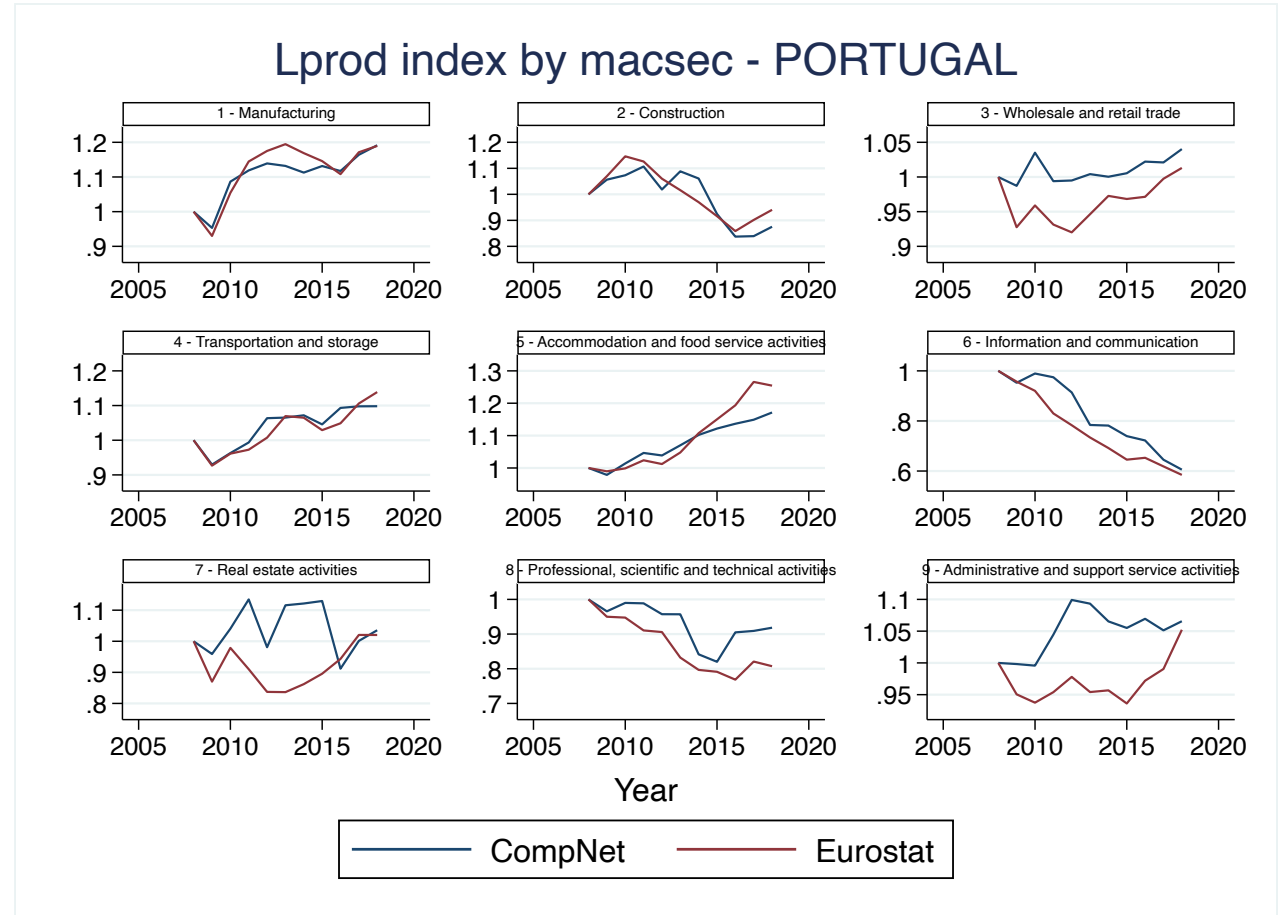
Germany

- There is a clear contrast in trend between CompNet and Eurostat in Real Estate
- Manufacturing and Construction behave extremely well



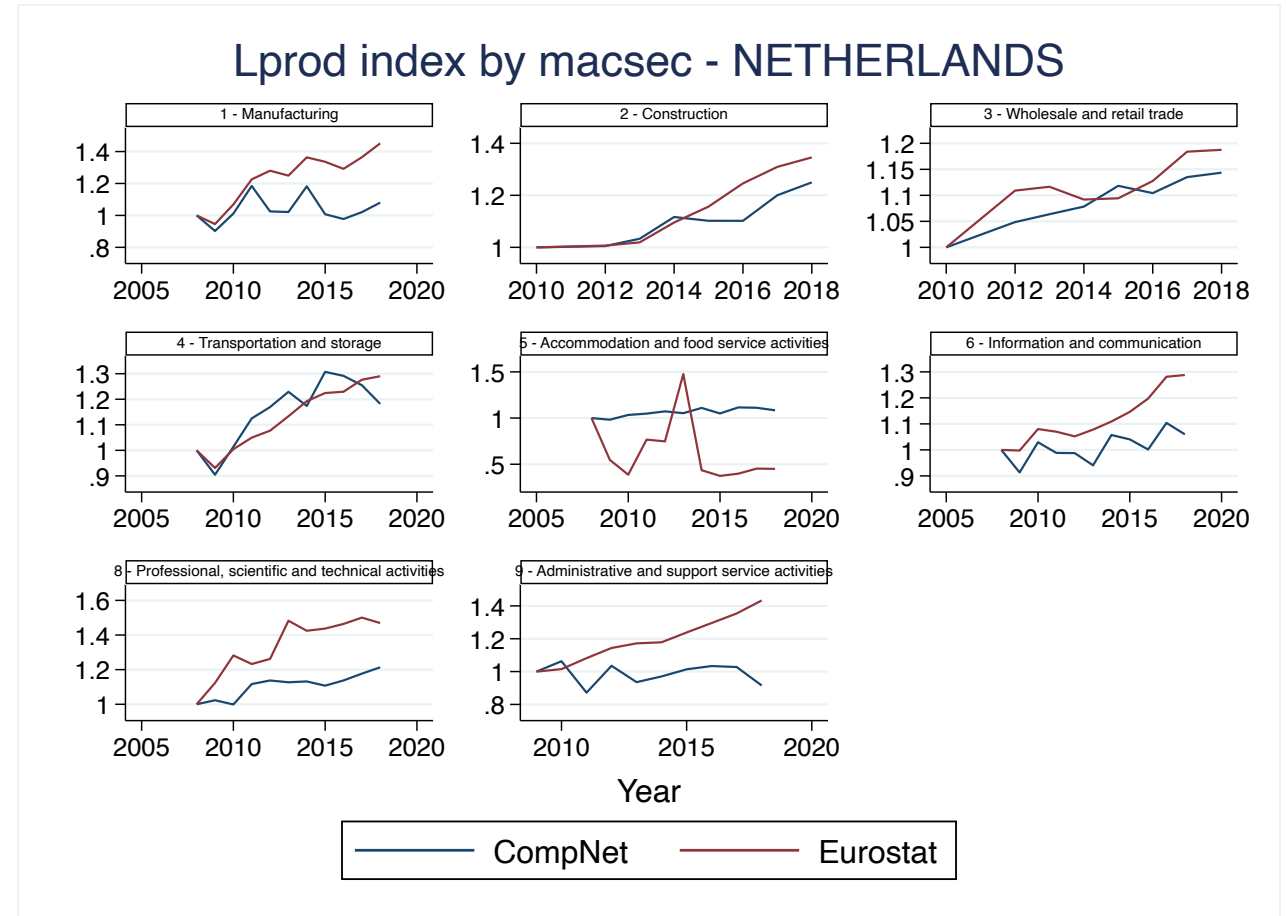
Portugal

- There is a strange break in 2011 when trends of CompNet and Eurostat reverse.
- Most likely coming from sector 7 and 9.
- The break is not quantitatively relevant



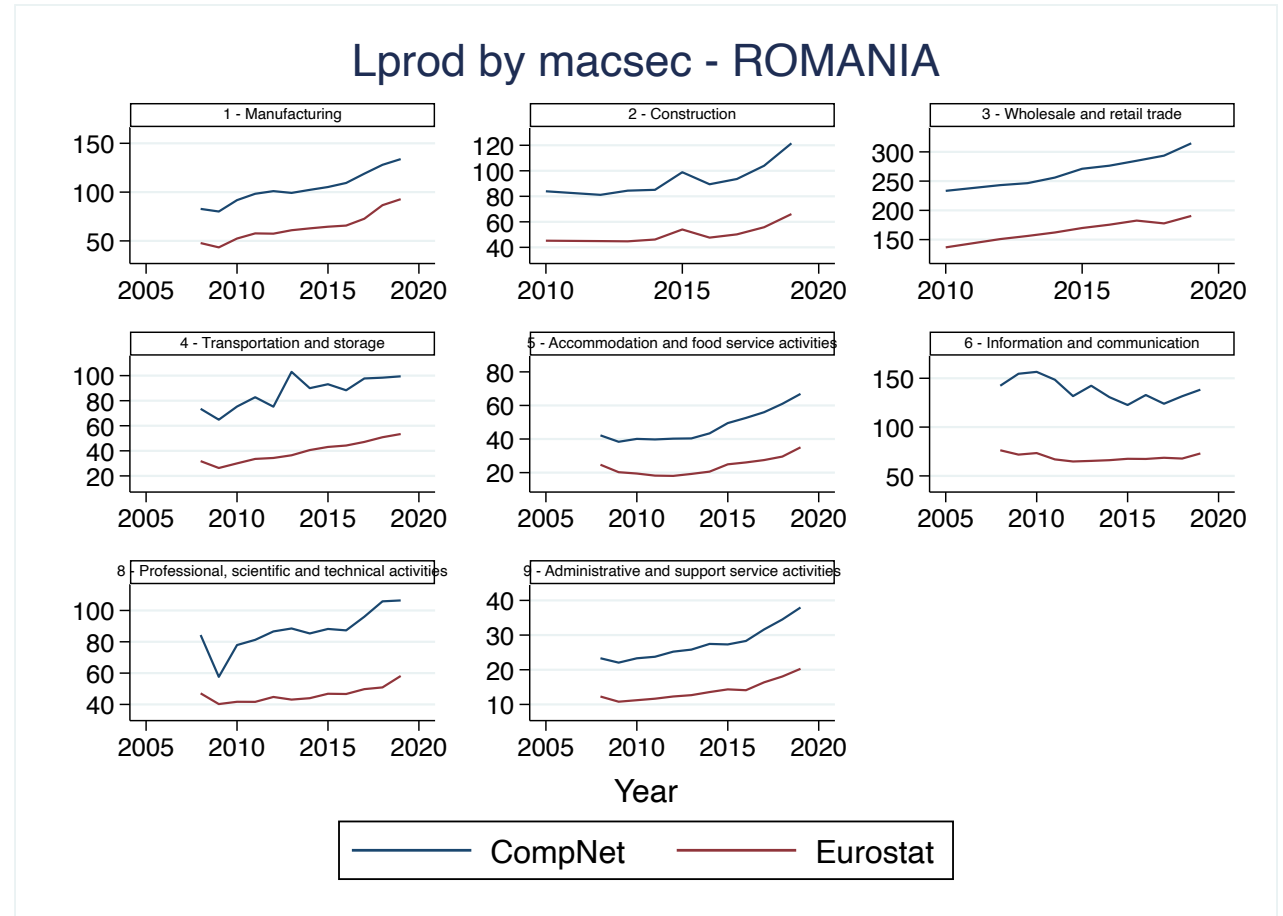
Netherlands

- Stronger growth in Manufacturing might be driving the overall trend
- Other divergences include Info and communication, Professional and SciTech activities, and Administrative services
- Odd movements in sector 5 in Eurostat

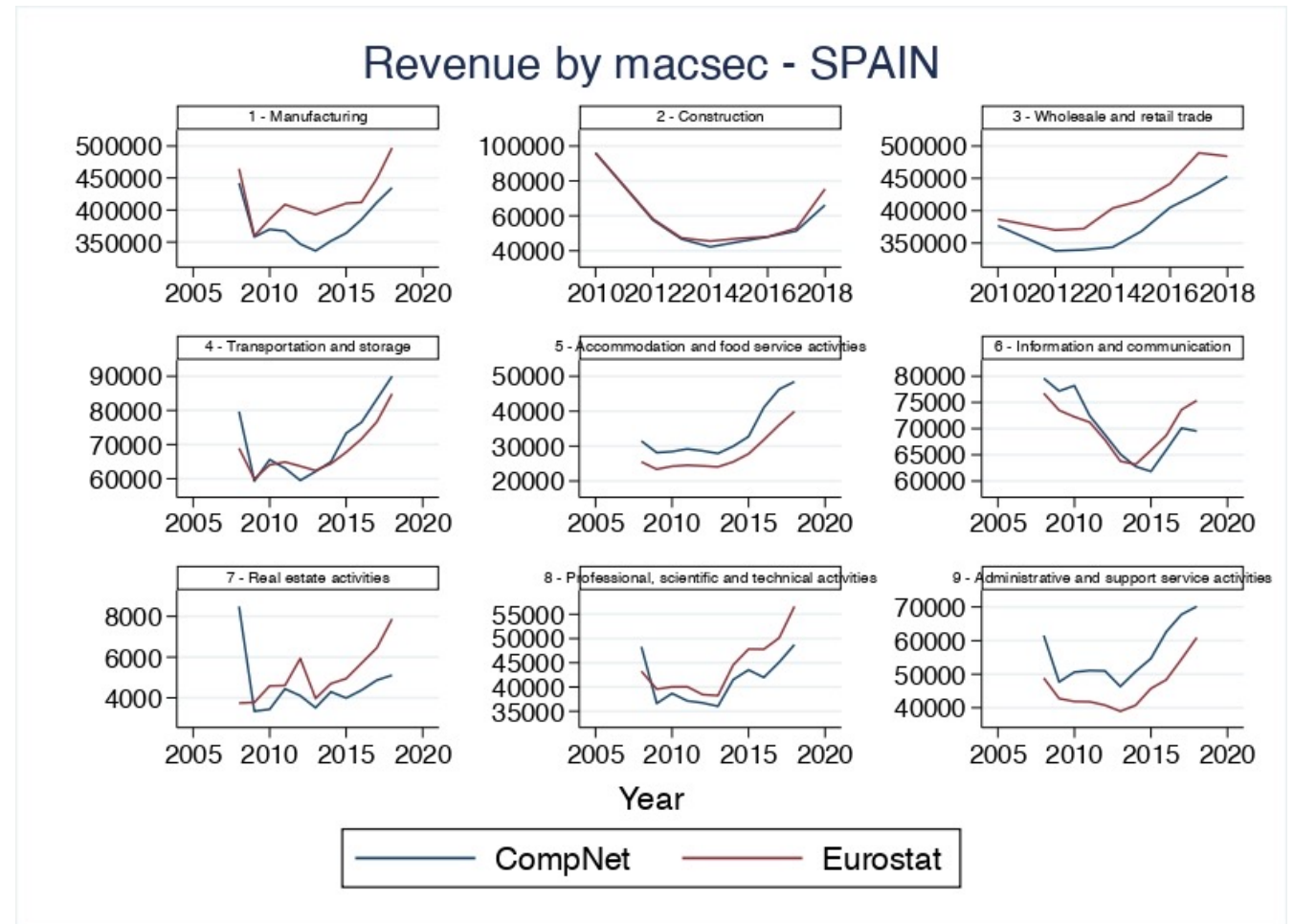


Romania

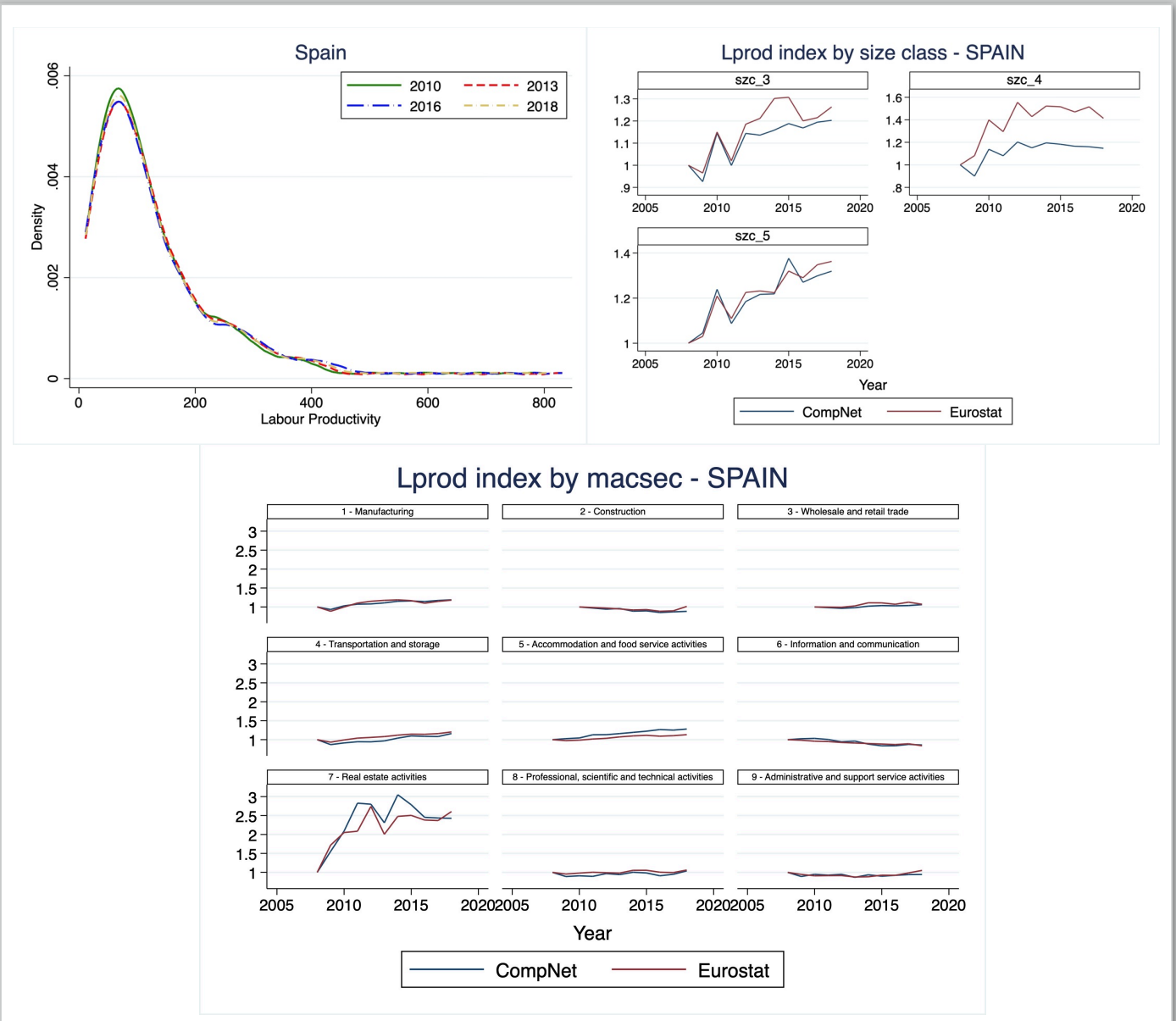
- We have larger labor productivity in ALL sectors in Romania



Spain - revenue

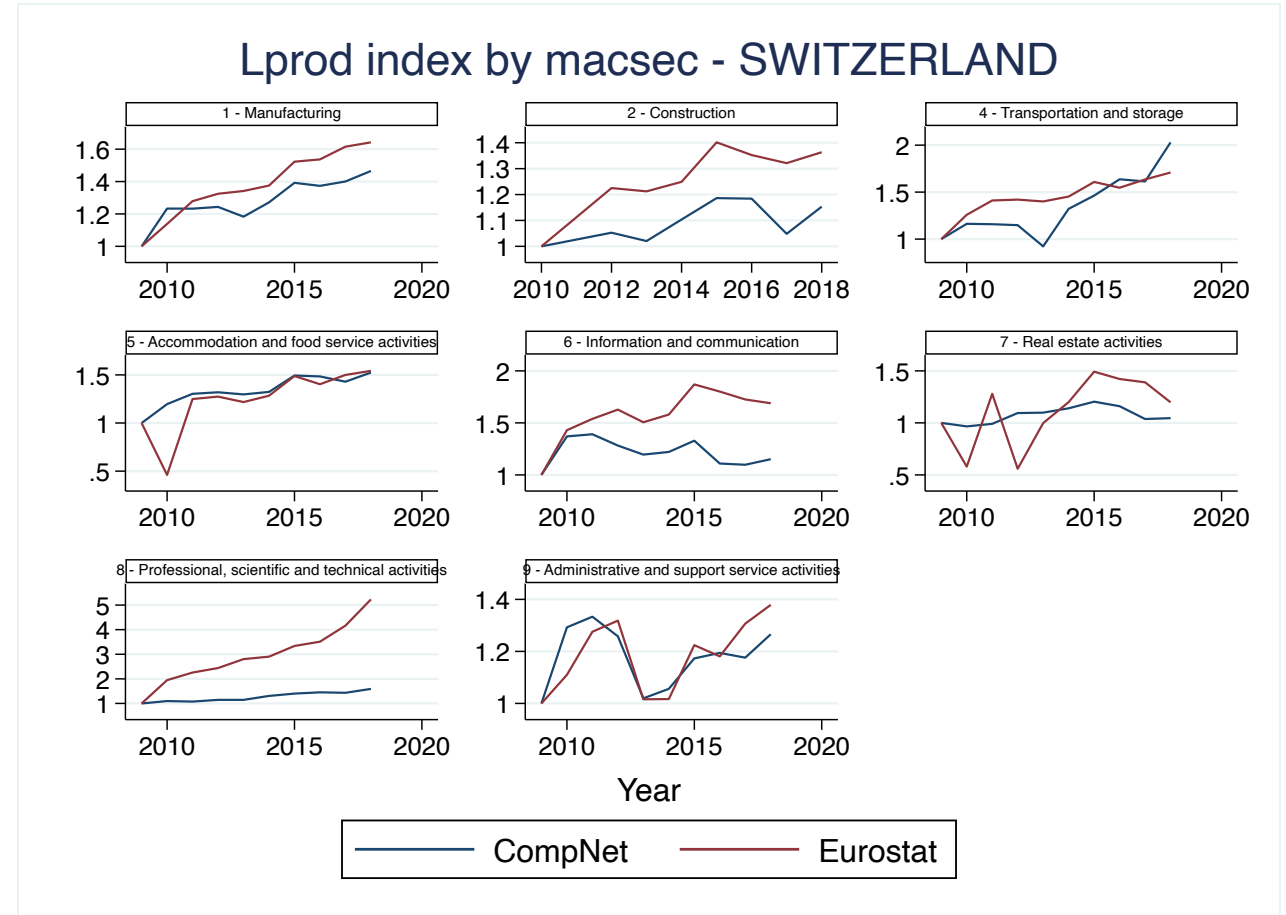


Spain – productivity distribution

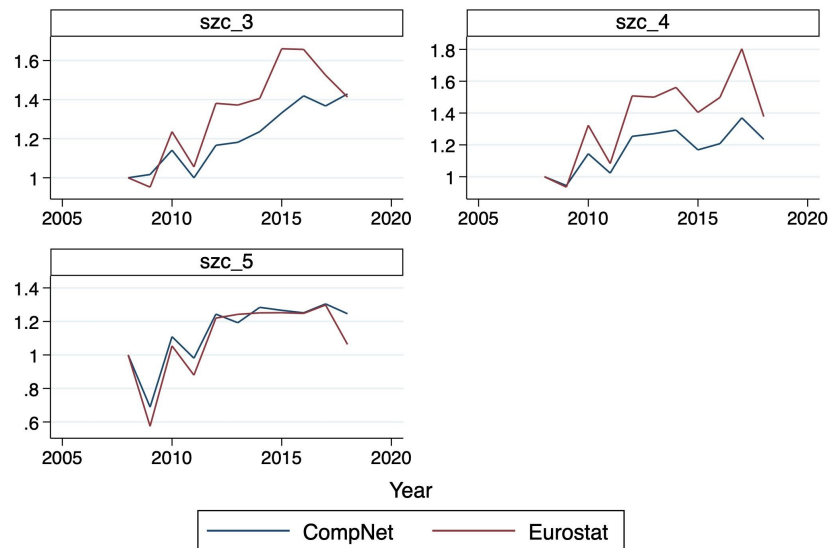


Switzerland

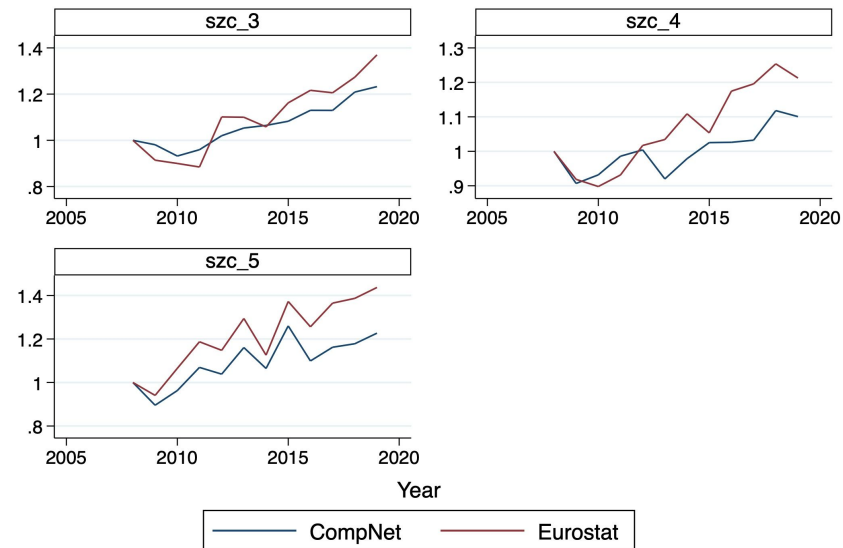
- Significant better growth in Eurostat.
- This is true for most sector except 9.



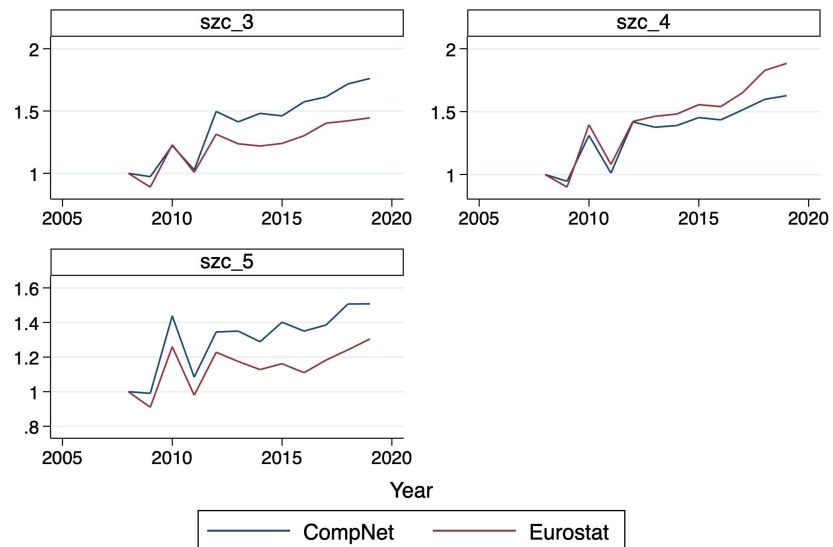
Lprod index by size class - BELGIUM



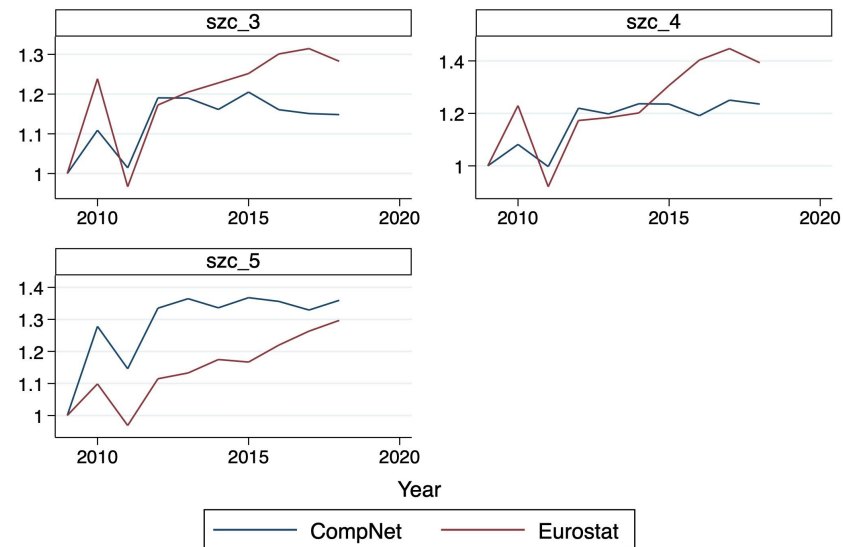
Lprod index by size class - CROATIA



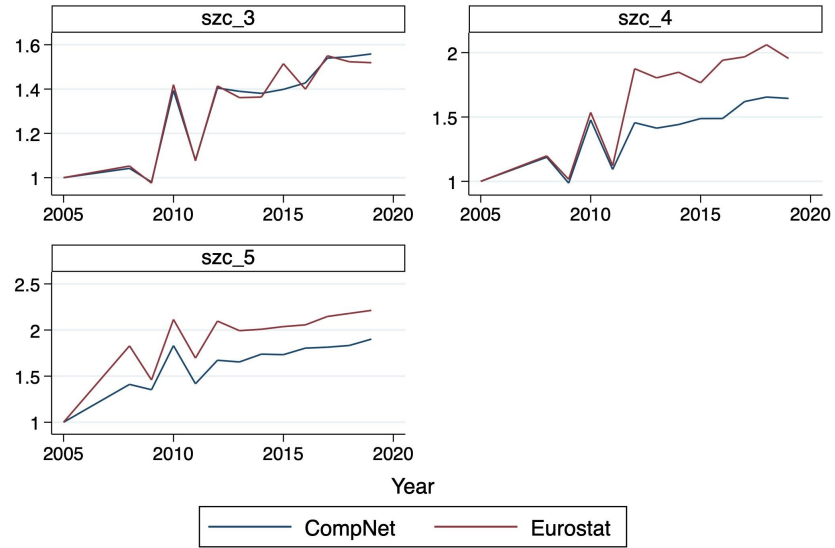
Lprod index by size class - CZECH REPUBLIC



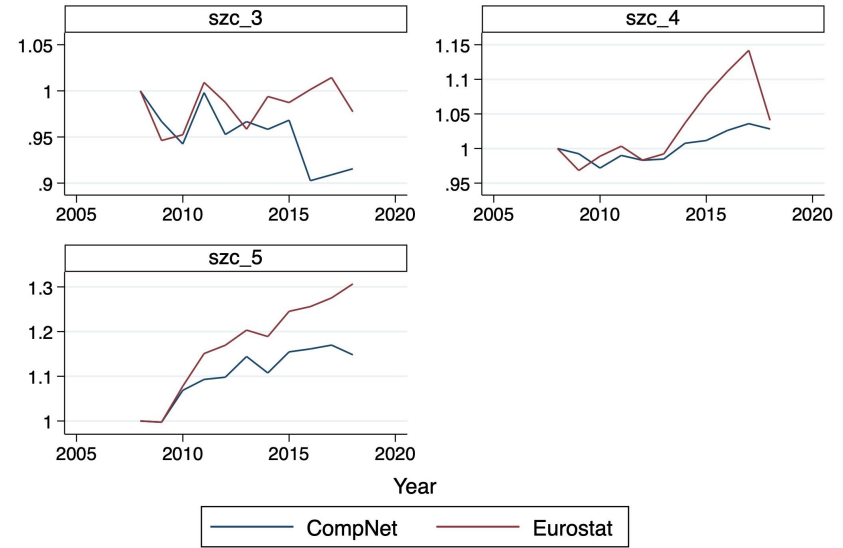
Lprod index by size class - DENMARK



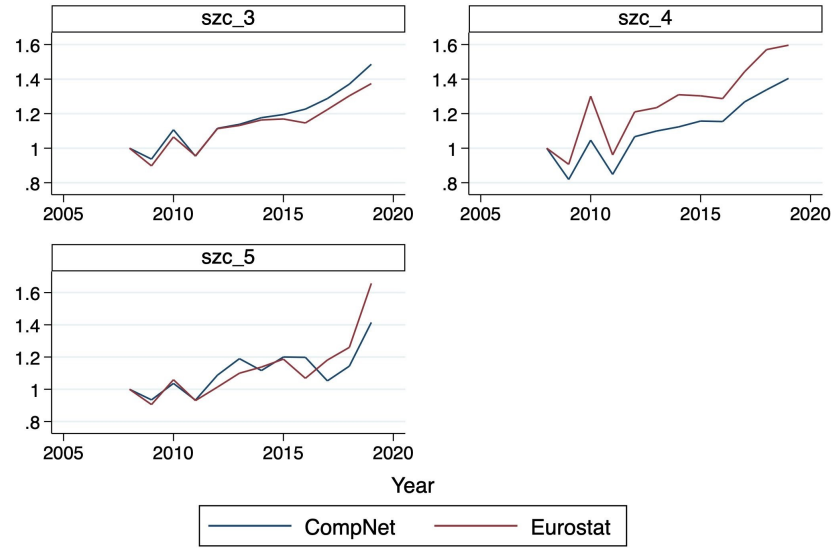
Lprod index by size class - FINLAND



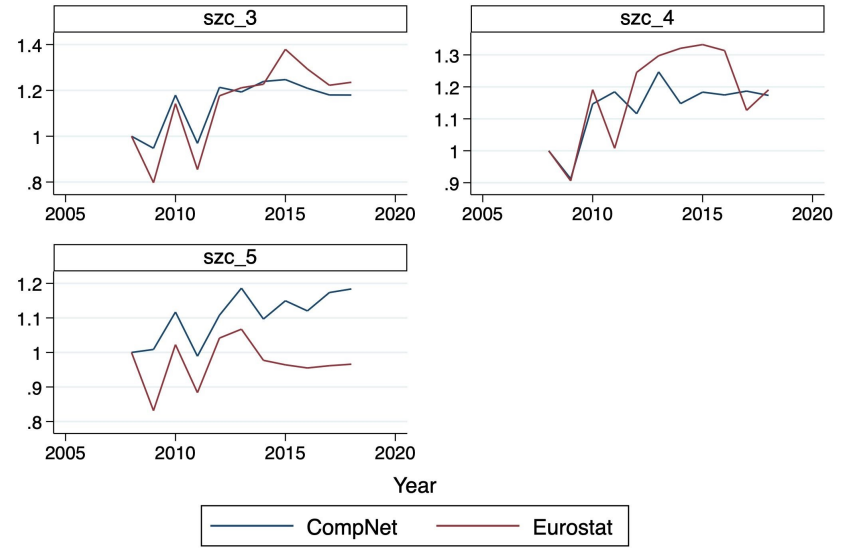
Lprod index by size class - GERMANY



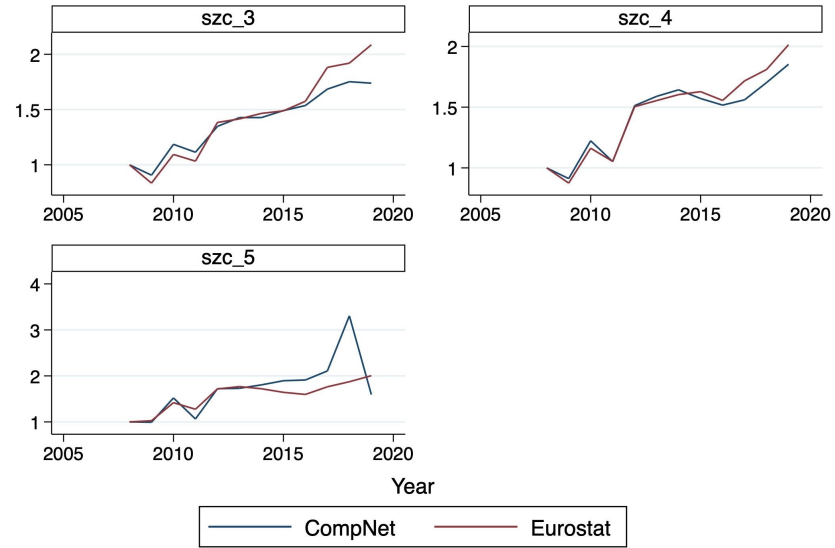
Lprod index by size class - HUNGARY



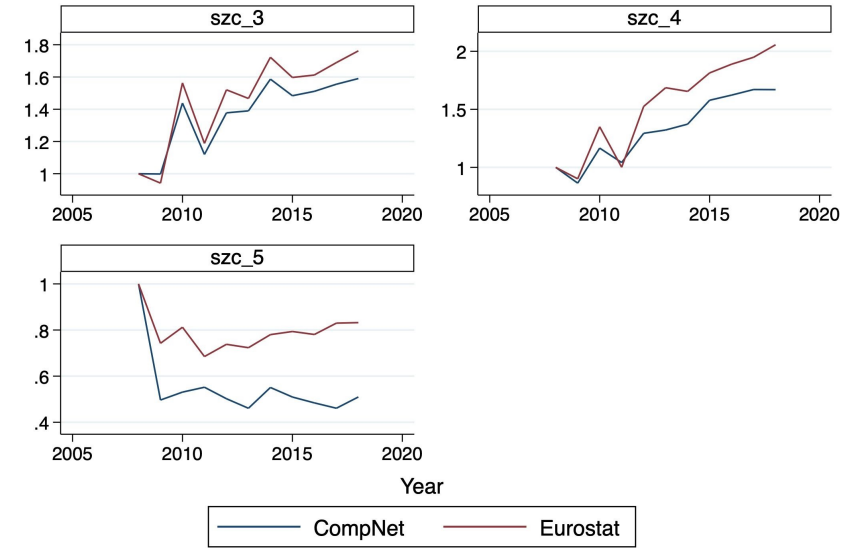
Lprod index by size class - ITALY



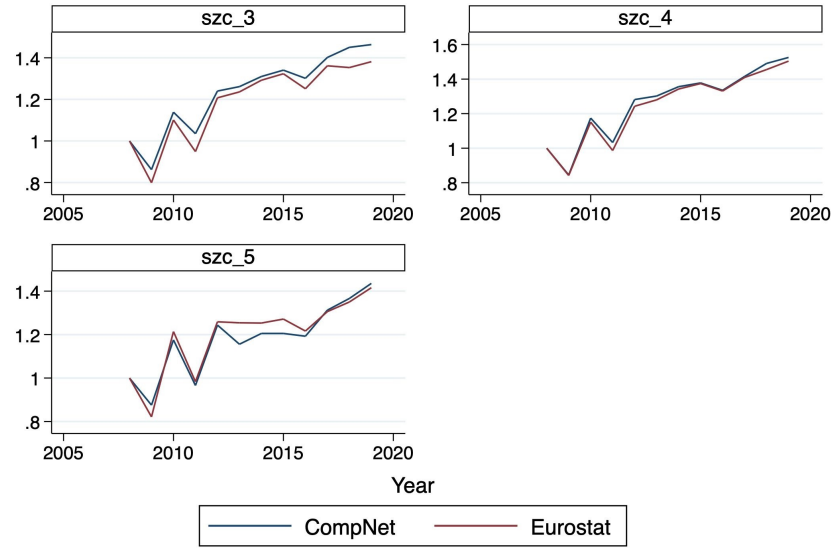
Lprod index by size class - LITHUANIA



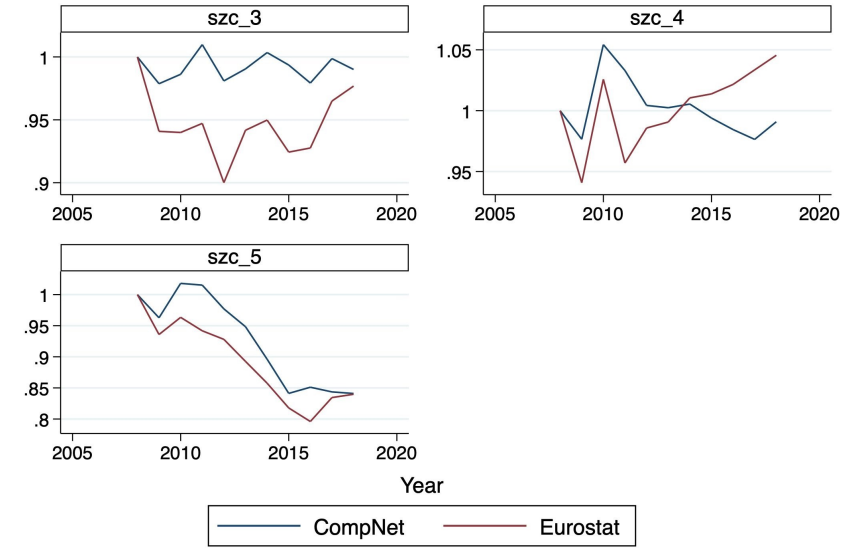
Lprod index by size class - NETHERLANDS



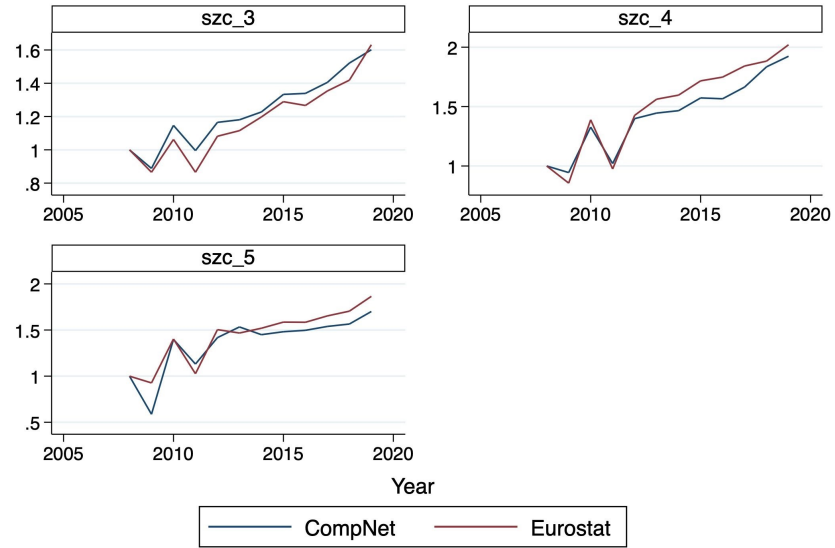
Lprod index by size class - POLAND



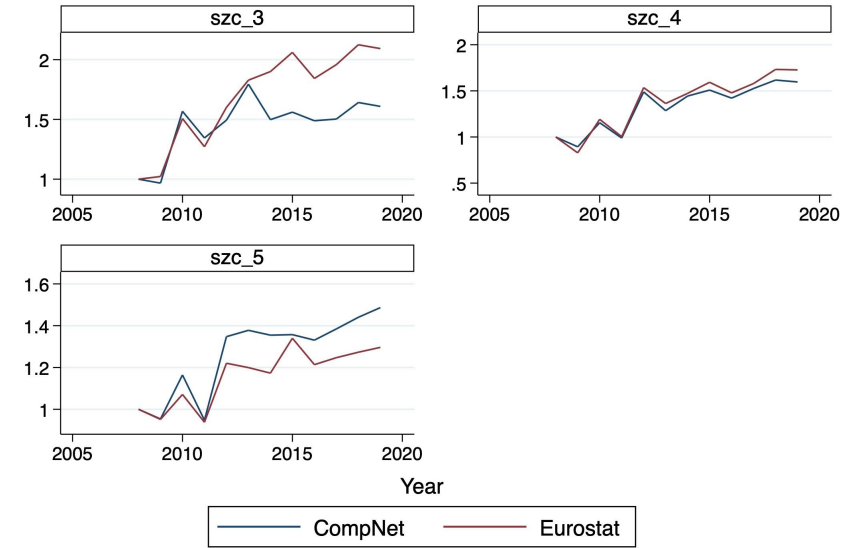
Lprod index by size class - PORTUGAL



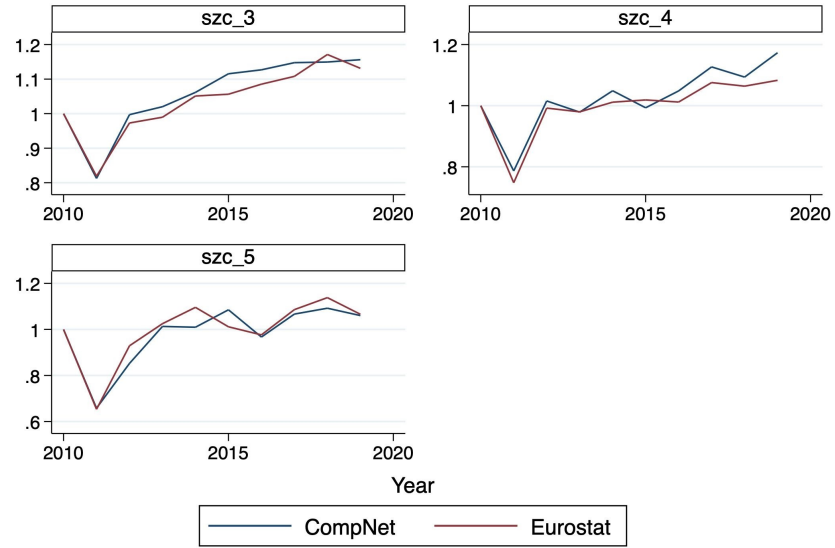
Lprod index by size class - ROMANIA



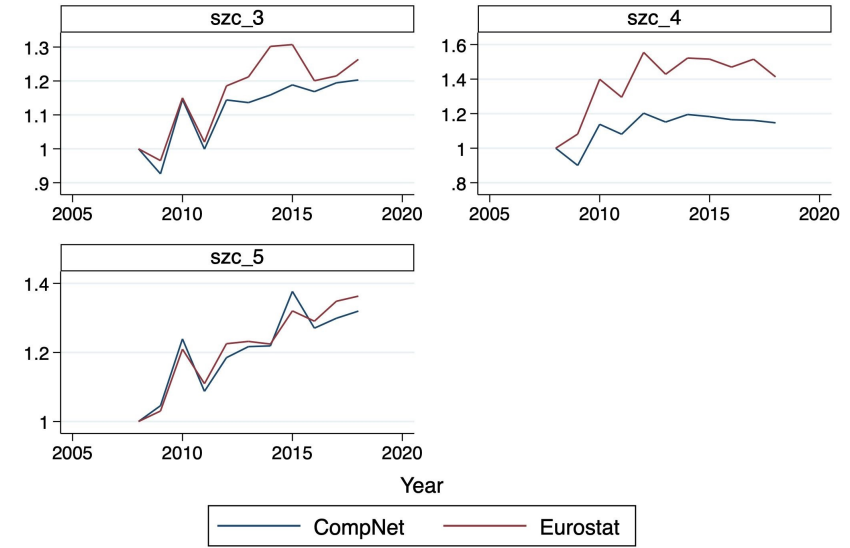
Lprod index by size class - SLOVAKIA



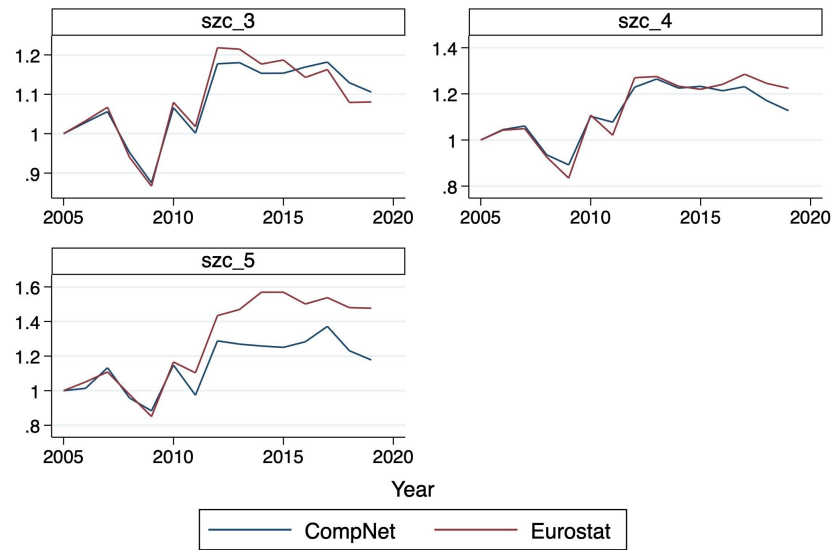
Lprod index by size class - SLOVENIA



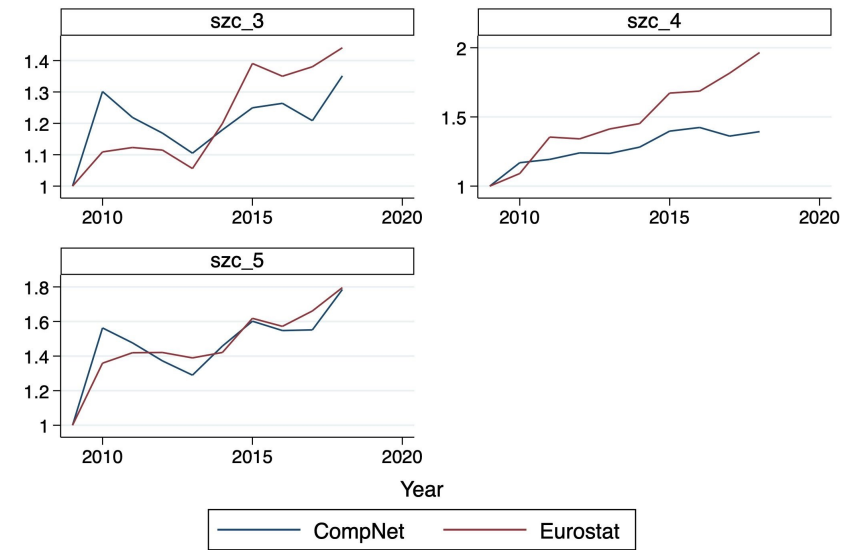
Lprod index by size class - SPAIN



Lprod index by size class - SWEDEN



Lprod index by size class - SWITZERLAND



Number of firms growth

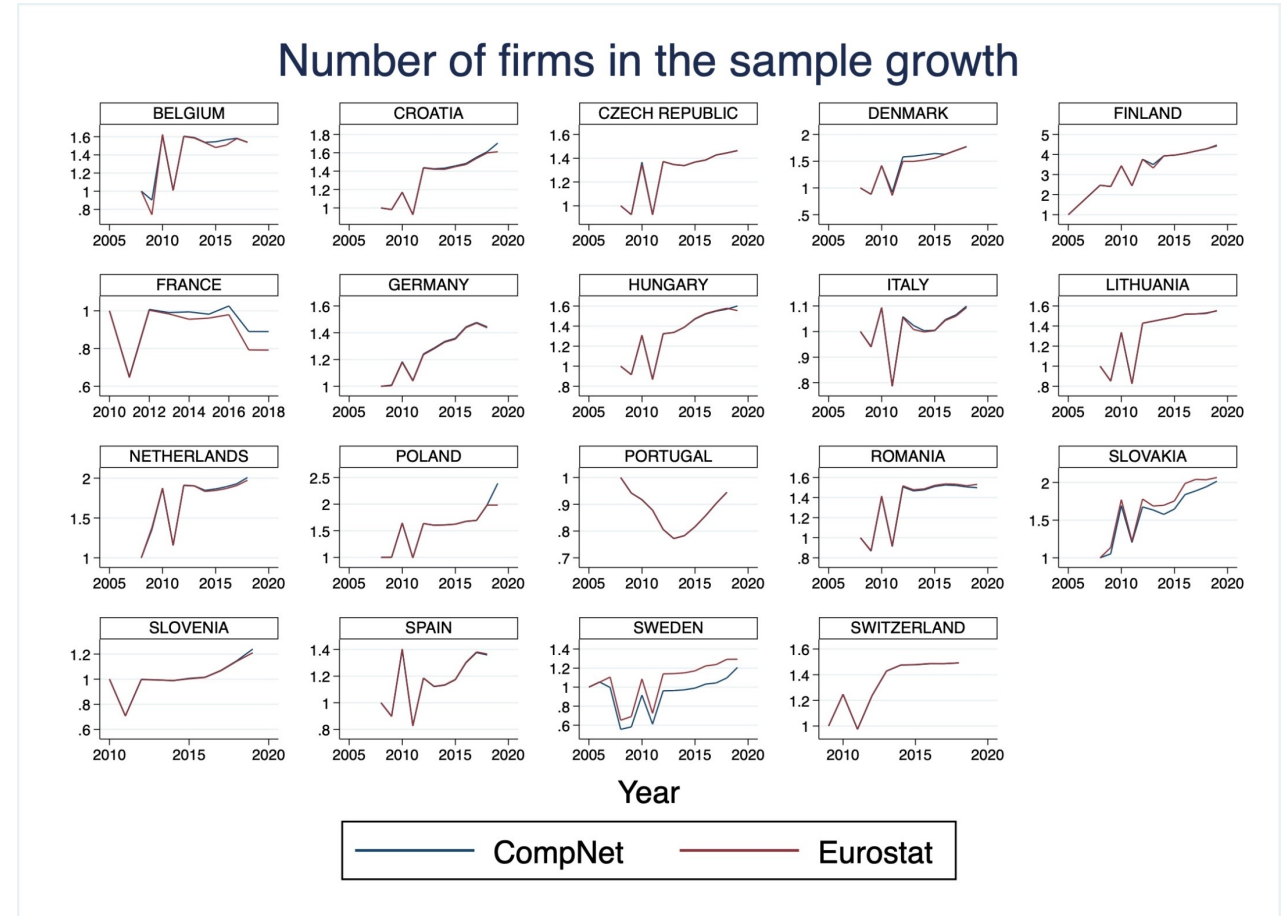
- We compare the time series related to the number of firms in CompNet with the one of Eurostat

Results

- CompNet follows pretty closely Eurostat dynamics in terms of firms in the sample
- The spike present **both** in Eurostat and in CompNet call for a cooperation with NSIs to understand underlying dynamics

Actions

- We advise to use such series (LV21_I_sw in the dataset) when studying the number of firms in the economy



COUNTRY AND SECTOR COVERAGE, 20E SAMPLE

Panel A: Country Coverage					
Country	Years (1)	Employment unweighted (2)	Employment weighted (3)	Number of firms unweighted (4)	Number of firms weighted (5)
Belgium	2000-2018	0.76	1.05	0.74	1.03
Croatia	2002-2019	0.86	1.03	0.84	1.01
Czech Republic	2005-2019	0.71	1.04	0.49	1.00
Denmark	2001-2018	0.80	1.00	0.86	1.03
Finland	1999-2019	0.89	0.99	0.88	1.00
France	2009-2015	0.70	0.81	1.01	1.07
Germany*	2005-2018	-	1.05	-	1.00
Hungary	2003-2019	0.86	1.12	0.83	1.02
Italy	2006-2018	0.75	1.02	0.70	1.00
Lithuania	2000-2019	0.83	1.00	0.81	1.00
Netherlands	2007-2018	0.85	1.06	0.81	1.03
Poland	2002-2019	0.79	1.02	0.62	1.02
Portugal	2005-2018	0.91	1.01	0.90	1.00
Romania	2007-2019	0.85	0.98	0.86	1.00
Slovenia	2002-2019	0.89	1.03	0.82	1.03
Slovakia	2000-2019	0.88	1.04	0.79	1.01
Spain	2008-2018	0.68	1.08	0.62	1.00
Sweden	2008-2018	0.61	0.90	0.78	1.04
Switzerland	2009-2018	0.67	1.11	0.33	1.00
TOTAL	2009-2016	0.58	1.01	0.59	1.01
Cross-country simple average	2009-2016	0.75	1.02	0.72	1.01

Panel B: Macro – Sector Coverage (balanced sample excluding France)				
Macro-sector	Employment unweighted (1)	Employment weighted (2)	Number of firms unweighted (3)	Number of firms weighted (4)
Manufacturing	0.53	1,03	0.56	1,00
Construction	0.57	1,03	0.51	1,00
Wholesale and retail trade	0.78	1,01	0.73	1,00
Transportation and storage	0.49	1,05	0.42	1,00
Accommodation and food service activities ICT	0.76	1,05	0.70	1,04
ICTAccommodation and food service activities	0.55	1,01	0.50	1,01
Professional Activities	0.42	1,01	0.40	1,01
Administrative and service	0.49	1,06	0.38	1,00

[Back](#)