This version: 21 December 2018

CompNet's 6th vintage of data: Novelties and main stylised facts

Report coordinated and drafted by Paloma Lopez-Garcia. The report includes the valuable contributions of Daniele Aglio, Richard Bräuer, Peter Haug, Jan Paul van der Kerke, Matthias Mertens, Matteo Sartori, Roberta Serafini, Ana Cristina Soares and Alessandro Zona Mattioli, all from ECB/IWH. Research assistance by Daan Beijersbergen and Marta Colombo and comments from Ettore Dorrucci and Filippo di Mauro are gratefully acknowledged.

Table of contents

Table of Figures

1. Introduction

The economic literature has since long recognised that firm-level data delivers crucial information for understanding the drivers of competitiveness, as aggregate performance depends strongly on firm-level decisions (on labour and capital markets as well as innovation and technological capacity). Moreover, widespread heterogeneity in firm's behaviour has been well documented (Caves 1998, Bartelsman and Doms 2000), thus highlighting the limits of models based on the representative agent hypothesis. These findings also suggest that better knowledge of the underlying distribution of size and productivity might be required in order to assess aggregate productivity growth, and thus competitiveness.

In policy terms, an important implication of the existing firm heterogeneity is that a similar policy shock might yield different results on (aggregate) competitiveness measures across countries or industries, with important consequences for welfare and distribution, depending upon the specific firm configuration prevailing at any moment in time.[1] It then follows that we need not only to improve on firm-level indicators, moving from averages to the knowledge of the entire distribution of firm performance, but also to incorporate in a more systematic way the impact of firm heterogeneity on 'standard' assessments of competitiveness.

Despite the increasing demand for micro-founded analysis by policy makers, firm-level cross-country studies are scarce. The main reason is that firm-level data collected for administrative purposes are confidential in nature and lack comparability across countries. In this context, the commercial databases like the ones compiled by Bureau van Dijk (BvD) have been widely used for research and policy analyses. The drawback of these databases is, however, that the collection of variables such as employment or exports, which are crucial for the analysis of trade and productivity, is not compulsory in all countries. As a result, the coverage of firms providing the required information for competitiveness analysis is drastically reduced and/or biased.[2] Moreover,

---

[1] Melitz and Redding (2014) provide a comprehensive summary of the different channels through which a trade shock interacts with firm heterogeneity in driving aggregate productivity.

[2] The coverage of Amadeus varies across countries, depending on each country's specific data collection rules and criteria. For example, only 11% of firms listed in Amadeus/Belgium have information on employment and turnover, while that percentage goes up to 66% in Spain. In Italy only about one-third of firms provide sufficient information to compute labour productivity and Total Factor Productivity. The value of exports is reported by a negligible number of firms in most countries.

country sources are greatly heterogeneous in nature – as detailed in the accompanying cross-country comparability report[3]– and commercial databases can do little to tackle the issue.

The "raison d'être" of the Competitiveness Research Network (CompNet), founded by the EU System of Central Banks in 2012, is to fill this gap and produce comparable cross-country firm-based information that suits the analyses conducted by the participating institutions as well as the research community. Beyond the dataset per se, the value added of the network is to bring together researchers working on competitiveness issues in different institutions. The network also counts on the collaboration of both the providers and users of data, which facilitates the production of the indicators demanded at each point of time by the participating policy institutions.

The aim of this report is to present the new vintage of the cross-country dataset, as well as to introduce the reader to the main novelties since the last data release in terms of governance and underlying methodology. Hence this report should be complemented with other documentation to get a complete picture of the Network. First, newcomers to the network might want to revise previous CompNet documentation to get a better sense of its methodology and achievements[4]; second, data users are encouraged to read the recently produced cross-country comparability report (Altomonte et al. 2018) and user guide (Aglio et al. 2018). These two reports refer to the current data vintage.

Over time, the CompNet dataset has increasingly been utilised for research and policy analysis, in line with a strong interest within European institutions and National Central Banks to further expand the use of firm-level information to analyse the micro-foundation of aggregate developments[5]. As a result of such demand, the number of CompNet's stakeholders has increased since the last release of the dataset (5th vintage, in 2016). Besides the European Central Bank and a number of National

---

[3] Altomonte et al. (2018): "Assessing the reliability of the CompNet micro-aggregated dataset for policy analysis and research: Coverage, representativeness and cross-EU countries comparability". This report will become available on line on CompNet website (see www.comp-net.org).

[4] For more information on the initial stages of the Network and fundamentals of the micro-aggregated dataset, please refer to the papers documenting previous vintages, in particular the horizontal paper Lopez-Garcia and di Mauro (2015) and the module-specific papers: Berthou et al. (2015) (trade); Ferrando et al. (2015) (financial) and Fernández et al. (2017) (labour).

[5] See for example Mario Draghi's speech" Moving to the frontier: Promoting diffusion of innovations", to be found here: https://www.ecb.europa.eu/press/key/date/2017/html/sp170313_1.en.html, the article on "Firm heterogeneity and competitiveness in the European Union" published in the ECB EB of March 2017 (https://www.ecb.europa.eu/pub/pdf/other/eb201702_article02.en.pdf), or the Chapter. 2 of the EBRD Transition Report 2017-2018. Besides, a number of papers shedding light on exchange rate pass-through or input misallocation, among others, have already been published. For more details, please visit CompNet's webpage: www.comp-net.org.

Central Banks of the ESCB, the European Commission, the European Investment Bank and the European Bank for Reconstruction and Development have become members of the network. The Halle Institute for Economic Research and the Tinbergen Institute, two European research institutes renowned for their expertise in productivity analysis, also actively participate in the project, with the Halle Institute also collaborating in a number of important tasks like the preparation of the statistical code, the collection of data, and the provision of other services. Another important development has been the involvement as data providers of a number of National Statistical Institutes and other national research centres with access to firm-level data (see next section for greater detail). Following these efforts, the country coverage of this vintage has grown from 15 (in the 5th vintage) to 18 countries, including the six biggest EU economies (DE, FR, IT, ES, NL and PL).

A key feature of the 6th CompNet data vintage is the considerable progress made to further enhance cross-country comparability. If CompNet data are to be used for policy analyses of competitiveness, trade and productivity, maximum comparability has to be ensured *conditional* on the fact that the database relies on administrative sources, i.e. it was originally not tailored for cross-country comparisons. Indeed, the fact that CompNet relies on existing data rather than on a new survey is cost-efficient, but also implies that sampling criteria, variable definition or cleaning of data, to name a few key issues, are decided by other parties and for other purposes. This is a common problem faced by all data sources based on administrative data. In this context, it is key that all data comparability issues are: (i) addressed with state-of-the-art statistical methods; and (ii) properly documented, so that researchers and policy makers can account for them in their analyses. For that purpose, this report is accompanied by a cross-country comparability report (Altomonte et al. 2018) which shows in great detail both the country-specific characteristics of the metadata and of the indicators being eventually produced. Most importantly, the cross-country comparability report highlights the remaining biases as well as the strengths and weaknesses of the CompNet dataset relative to other similar granular data sources, and provides clear guidelines to CompNet data providers in order to consistently improve the quality of the dataset.

The current data vintage incorporates fundamental improvements to both the content and the procedure of data collection. Regarding the content, we have updated the methodology to estimate some of our main indicators, such as Total Factor Productivity (henceforth, TFP). We have also included new indicators that contribute to the ongoing policy debate. Amongst them, three should be highlighted from the outset. First, the new dataset includes a new module on *distressed firms* still operating in the market, as well as their characteristics. Second, *job creation and destruction* rates have been computed for narrowly defined sectors or across-size classes in order to single out such

components of gross job flows. Third, for the first time information has been collected at the *regional level* (i.e. NUTS2 level). The latter will make it possible to explore, for instance, the productivity distributions across different regions within a given country.

Regarding the improvements in data procedures, the CompNet code has been totally rewritten to increase its efficiency and robustness, correct small glitches, include a better weighting system and incorporate confidentiality checks which the data providers can adjust to their respective country-specific requirements.

The presentation of the new dataset is done in two steps. *First*, the report documents the 6th data vintage of CompNet in terms of coverage, representativeness, and indicators. This is done in Sections 2 and 3. For completeness, the information therein should be complemented with the cross-country report on data comparability (Altomonte et al. 2018) and the updated user guide (Aglio et al. 2018). Both documents are available on CompNet's webpage. *Second*, the report highlights selected stylised facts of potential high value for research and policy, emerging from the new dataset. The discussion, in Section 4, is structured in the same fashion across main indicators: we first show how the relevant CompNet indicator behave vis-à-vis other (aggregate) sources, then we illustrate its granular features, and finally we use the data to uncover new facts or to shed light on a number of issues of interest to policy makers. This preliminary analysis intends to be solely suggestive and aims to "raise the appetite" of researchers in the various topics covered by the report.

2. Data sources, coverage and representativeness

This section provides an overview of the methodology and raw data used to construct the CompNet dataset, with special emphasis on the representativeness and cross-country differences. For more information please refer to the cross-country comparability report (Altomonte et al. 2018).

2.1 The micro-distributed approach

In order to preserve confidentiality of firm-level sources and improve cross-country comparability, CompNet adopted from the beginning the so-called "micro-distributed approach", developed by Bartelsman, Haltiwanger and Scarpetta (2004) and described in Lopez-Garcia, di Mauro et al (2015). In this approach, a common protocol is used to extract relevant information – not only the mean, but also other moments of the distribution of each variable – from existing firm-level datasets available

within each of our data providers. The code is sent to the data providers who run it on their firm-level data and send in return only aggregated information to preserve confidentiality.

The common methodology ensures the harmonisation of industry coverage, variable definitions and units, target population of firms, deflators, treatment of outliers, and estimation methodologies. Regarding cross-country comparability of the raw data, the use of indicator-specific population weights ensures that the firm distribution by macro-sector and size class resembles the one of the population of firms (according to Eurostat Structural Business Statistics). The use of PPPs improves the comparison of nominal variables across countries. Moreover, given that some countries do not sample firms below a given size threshold, all indicators are computed both for the sample of firms with 20 and more employees as well as for the sample of firms with at least one employee.

2.2 Basic country information

Table 2.1 shows the sources of raw data in each of the countries included in this report. In most cases, data providers rely on business registers, complemented with other firm-level sources either to enrich firm coverage or to include information on value of exports. Despite the diverse sources of raw data, in all countries the target population of the dataset is narrowed down and the data consistently includes non-financial corporations with employees.

Table 2.1: Sources of raw data

Country	Data source name	Acronym	Institution responsible for source	Data provider	firms included in dataset*	Source specific Information
Belgium	-	-	-	National Bank of Belgium (BACH)	-	Micro-information underlying the financial ratio's in the BACH database*
Croatia	Yearly financial statements of firms	FINA	Financial Agency Croatia	Croatian National Bank	NFC drawn from total economy	
Czech Republic	Annual report of economic units in selected production in	P501	Statistics Czech Republic	Czech National Bank	NFC drawn from total economy	full coverage for firms with >50 employees, stratified survey for smaller firms
	Extrastat/Intrastat foreign trade transaction data	TRADE	Statistics Czech Republic		NFC drawn from total economy	
	Business Register	RES	Statistics Czech Republic		NFC drawn from total economy	
Denmark	Accounts statistics - non-agricultural industries	Acc. Stat.	Statistics Denmark	Central Bank of Denmark	NFC	
	General enterprise statistics	Gen. Stat.	Statistics Denmark		NFC drawn from total economy	
Finland	Structural business and financial statement statistics data	SBS	Statistics Finland	Statistics Finland	NFC drawn from total economy	
	International trade statistics data	ITS	Finnish Customs		NFC drawn from total economy	
France	Regime of Normal Real Profits	BRN	Statistics France	Statistics France	NFC drawn from total economy	Complementing sources with RSI. BRN covers large firms +788K
	Simplified Regime for the Self-Employed	RSI	Statistics France		NFC drawn from total economy	Complementing sources with BRN. RSI covers small firms below 788K
Germany	administrative firm-level data	AFiD	German Statistics	German Statistics	Manufacturing	German sample only covers the manufacturing sector, subsequently only firms with more than 20 employees are included
Hungary	Tax registry database of National Tax and Customs Administration	NAV	National Tax and Customs Administration	Central Bank Hungary	NFC drawn from total economy	Non-mandatory variables for tax-records are underreported. E.g. 30% of Firms do not report the number of employees.
	Business Registry	VR	Statistics Hungary and Central Bank of Hungary		NFC drawn from total economy	
	Export-Import data of Hungarian Enterprises	Külker	Statistics Hungary		NFC drawn from total economy	
Italy	Statistical Business Register	ASIA	Statistics Italy	Statistics Italy	NFC drawn from total economy	
	Balance Sheets of non-financial companies	BIL	Statistics Italy		NFC drawn from total economy	
	Large enterprise survey	SCI	Statistics Italy		NFC drawn from total economy	
	Foreign Trade Statistics based on custom data	COE	Statistics Italy		NFC drawn from total economy	Complementary source which is targeted at large firms (+ 100 employees)
Lithuania	Statistical Survey on the Business Structure (Annual ques	F01	Statistics Lithuania	Central Bank Lithuania	NFC drawn from total economy	
	Business Register	BR	Centre of Registers		NFC drawn from total economy	
	Customs, Customs declarations	CU	Customs of the Republic of Lithuania		NFC drawn from total economy	
Netherlands	Statistics finances of non-financial enterprises	SFO	Statistics Netherlands	Statistics Netherlands	NFC drawn from total economy	Full coverage for small firms (< 40 Mln balance sheet total); Large firms (> 40 Mln balance sheet totals) are surveyed
	Business register	ABR	Statistics Netherlands		NFC drawn from total economy	
Poland	Reports on revenues, costs, profit and outlays on fixed as	F01	Statistics Poland	Central Bank Poland	NFC	Exclusion of firms with less than 10 employees
	Stat. financial report	F02	Statistics Poland		NFC	Exclusion of firms with less than 10 employees
Portugal	Central balance sheet database, annual survey	CBSD	Central Bank of Portugal	Banco de Portugal (BACH)	NFC	Micro-information underlying the financial ratio's in the BACH database*
	Simplified corporate information	IES	Statistics Portugal and Central Bank of Portugal		NFC drawn from total economy	Micro-information underlying the financial ratio's in the BACH database*
Romania	Balance sheet information on non-financial enterprises	Bal. Sheet	Ministry of Public finances	National Bank Romania	NFC drawn from total economy	
	Exports and imports of goods, firm-level data	TRADE	Statistics Romania		NFC drawn from total economy	
Slovakia	Annual report on production industries	Reports	Statistics Slovakia	National Bank of Slovakia	NFC drawn from total economy	Exclusion of firms with less than 20 employees
	statistical register of organizations	Register	Statistics Slovakia		NFC drawn from total economy	Exclusion of firms with less than 20 employees
	foreign trade statistics	Customs	Statistics Slovakia		NFC drawn from total economy	Exclusion of firms with less than 20 employees
Slovenia	Slovenia Public and Legal Records and Related Services	AJPES	Agency for Public Legal Records and Related Services	Univ. of Ljubljana	NFC drawn from total economy	
Spain	CBSO voluntary survey	CBA	Central Bank of Spain	Banco de España (BACH)	NFC	Micro-information underlying the financial ratio's in the BACH database*
	Spanish mercantile register	CBB	Mercantile registry		NFC	Micro-information underlying the financial ratio's in the BACH database*
Sweden	Structured business statistics	SBS	Statistics Sweden	Statistics Sweden	NFC drawn from total economy	
	International trade in goods	ITG	Statistics Sweden		NFC drawn from total economy	
	Business register	BR	Statistics Sweden		NFC drawn from total economy	

Table 2.2 shows basic country-specific information. The second column provides information on the period covered by each country, mostly from early 2000s to 2015, the third column specifies if there are data available for firms with less than 20 employees (what we call the "full sample", which covers all firms with at least 1 employee) or only for firms with 20 employees or more, the "20E sample". This is the case in some countries due to sampling exclusions of the underlying data sources, like Poland, in which the sources exclude firms with less than 10 employees, or Slovakia which samples firms with more than 20 employees or above a certain turnover threshold. Note that data for Germany, provided for the first time by the Federal Statistical Office (destatis), refers to firms with 20 or more employees operating in the manufacturing sector.[6]

Regarding sector coverage, CompNet includes firms operating in all sectors of the business economy with the exception of mining and agriculture, utilities and the financial sector (see Annex 1). However, due to sampling choices of the underlying sources or confidentiality reasons some countries do not cover some other sectors, which are detailed in the fourth column of **Table 2.2**. The fifth column specifies if the country provides export information. Exporting activity of firms is provided so far by 13 of 18 countries in CompNet, including now Germany, which makes this dataset unique in that it includes harmonized joint information on share and characteristics of exporters in each manufacturing 2-digit NACE Rev. 2 industry. The last column shows which countries provide information aggregated to the regional level.

[6] In previous vintages of CompNet, the German data were provided by the Bundesbank. The data were very rich in terms of availability of financial information but lacked information on exports and, above all, referred to a sample of large firms demanding ratings or with large FDI or export activity. Hence, the sample was fairly biased towards well-performing firms. In this vintage the Statistical Institute draws from a representative sample of manufacturing firms with at least 20 employees, for which there is also export information.

Table 2.2: Country-specific information

Country	Time span	Sample available	Excluded sectors (besides the sectors excluded by default)*	Export information	Regional information
BE	2004-2015	Full and 20e		No	Yes
CZ	2003-2015	20e		Yes	Yes
DE	1999-2014	20e	Construction and Services	Yes	No
DK	2000-2015	Full and 20e		No	Yes
ES	2009-2015	Full and 20e		No	Yes
FI	1999-2015	Full and 20e	Real estate activities	Yes	Yes
FR	2004-2014	Full and 20e		Yes	No
HR	2008-2015	Full and 20e		Yes	No
HU	1999-2015	Full and 20e		Yes	No
IT	2001-2014	Full and 20e		Yes	Yes
LT	2000-2015	Full and 20e		Yes	Yes
NL	2000-2014	Full and 20e		No	No
PL	2005-2015	20e		Yes	No
PT	2006-2015	Full and 20e	Manufacture of tobacco products, Manufacture of coke and refined petroleum products, Postal and courier activities, Real estate activities	No	No
RO	2005-2015	Full and 20e	Manufacture of coke and refined petroleum products, Air transport, Postal and courier activities, Real estate activities	Yes	Yes
SI	2005-2016	Full and 20e		Yes	Yes
SE	2003-2015	Full and 20e		Yes	No
SK	2000-2015	20e		Yes	Yes

*The sectors excluded in CompNet are: mining and agriculture, utilities, financial sector and public administration.

2.3 Coverage and representativeness of the data

Table 2.3 shows the coverage of employment and firms in CompNet relative to the population of non-financial corporations with employees operating in the sectors covered by CompNet (which are detailed in Annex 1). The population data is extracted from the Structural Business Statistics of Eurostat and refers to firms with at least 1 employee (with at least 20 employees in Poland, Slovakia, the Czech Republic and Germany) in 2011. On average, CompNet covers about 60% of the corresponding population of firms and employment although with large cross-country variation, ranging from 11% of firms in Italy to about 80% in Romania or 90% in Slovakia. Note that the number of firms shown for CompNet excludes observations dropped after the outlier cleaning procedure.

Table 2.3: Coverage of the population, 2011.

Country	Coverage vs. population	
	Employment	Number of firms
Belgium	44%	19%
Croatia	52%	38%
Czech Republic*	72%	72%
Denmark	53%	87%
Finland	50%	45%
France	57%	41%
Germany*	51%	42%
Hungary	57%	44%
Italy	39%	11%
Lithuania	69%	37%
Netherlands	35%	18%
Poland*	75%	74%
Portugal	56%	31%
Romania	68%	76%
Slovakia*	86%	90%
Slovenia	50%	28%
Spain	25%	15%
Sweden	40%	32%

Sources: CompNet 6th vintage and Eurostat, 2011.

Notes: due to data availability, representativeness for Spain is measured in 2012 and Slovakia in 2014. Germany only consists of the manufacturing sector.* 20E sample.

Even if coverage is not complete, what matters is that available firms are representative of the underlying population. In this respect, CompNet applies a reweighting scheme based on inverse probability weights to the raw data so that the share of firms by macro-sector and size class is the same as in the population.[7] Note that this year, for the first time, reweighting is dependent on the specific indicator and therefore provides more accurate results, and it applies to both the full sample (firms with at least 1 employee) and the 20E samples (firms with at least 20 employees). More details about the exact reweighting procedure can also be found in Annex 2. **Table 2.4** and

[7] Note that only few countries participating in CompNet have access to the firm census, which would be required to reweight the samples to ensure they are representative of the population of firms at the 2-digit industry (or finer) and size-class. As a consequence, we chose to go for the common denominator across all countries which was the population figures given by the Structural Business Statistics of Eurostat and available only at the macro-sector and size-class level. This weighting does not ensure, however, full representativeness of the sample at the 2-digit industry level as long as there are large differences in the share of employment of each detailed industry within a given macro-sector.

Table 2.5 below show respectively the share of total employment in each of the main macro-sectors (manufacturing, construction and services)[8] and across size classes[9] in CompNet and Eurostat (the numbers of Eurostat are reported in brackets).

Table 2.4: Share of employment by main sector in CompNet and population (in brackets), 2011

Country	Manufacturing	Construction	Services
Belgium	29.7% (27.5%)	12.3% (17.7%)	57.9% (54.7%)
Croatia	34.3% (43.4%)	12.3% (25.4%)	53.3% (31.1%)
Czech Republic*	52.3% (56.2%)	6.88% (9.34%)	40.8% (34.3%)
Denmark	23.3% (31.3%)	12.8% (15.3%)	63.7% (53.3%)
Finland	30.6% (31.0%)	13.7% (16.9%)	55.6% (51.9%)
France	22.4% (24.9%)	14.3% (16.3%)	63.2% (58.6%)
Germany*	100%	-	-
Hungary	32.9% (34.9%)	8.90% (11.6%)	58.1% (53.3%)
Italy	37.8% (33.1%)	10.3% (15.5%)	51.8% (51.3%)
Lithuania	27.2% (30.2%)	12.7% (18.1%)	59.9% (51.5%)
Netherlands	19.5% (17.1%)	10.7% (13.7%)	69.6% (69.1%)
Poland*	44.9% (54.7%)	8.21% (11.5%)	46.8% (33.6%)
Portugal	28.4% (24.1%)	13.7% (15.6%)	57.8% (60.2%)
Romania	35.7% (40.1%)	12.4% (15.2%)	51.8% (44.6%)
Slovakia*	50.2% (64.3%)	5.95% (9.43%)	43.7% (26.1%)
Slovenia	45.9% (43.6%)	9.23% (18.5%)	44.7% (37.8%)
Spain	24.3% (21.8%)	12.9% (16.3%)	62.6% (61.7%)
Sweden	21.2% (28.0%)	13.9% (17.1%)	64.8% (54.8%)

Sources: CompNet 6th vintage and Eurostat, 2011.

[8] The 7 service macro-sectors are aggregated to a single “service sector” using sector weights.

[9] CompNet considers five size classes: 1-9, 10-19 employees, 20 to 49 employees, 50 to 249 employees and more than 250 employees.

Notes: Number in parenthesis refers to the figures in Eurostat Structural Business Statistics. * refers to the 20E sample. In the case of Germany only manufacturing firms are included.

As a result of the reweighting scheme, CompNet replicates well the distribution of employment across the main macro-sectors. Turning to the employment share of the different size classes, according to

Table 2.5 in Italy and, to a lesser extent the Netherlands, micro-firms are slightly under-represented. Note that very large firms are under-represented across all countries. This is quite extreme in some cases, which results in an over-representation of micro firms, as the percentages sum up to 100 in each country. This is a result of the, perhaps too strict, outlier treatment which drops firms at the very extreme of the size distribution.

Table 2.5: Share of employment by size class in CompNet and population (in brackets), 2011

Country \ Size Classes	1 - 9 Employees	10 - 19 Employees	20 - 49 Employees	50 - 249 Employees	> 250 Employees
Belgium	21.5% (26.3%)	12.8% (7.78%)	20.3% (12.4%)	24.4% (16.8%)	20.8% (36.5%)
Croatia	27.2% (9.29%)	13.9% (13.0%)	17.4% (15.1%)	26.4% (27.8%)	14.9% (34.6%)
Czech Republic*	-	-	16.2% (16.5%)	38.3% (32.8%)	45.3% (50.5%)
Denmark	41.1% (23.0%)	14.0% (9.55%)	17.8% (12.6%)	19.3% (21.6%)	7.57% (33.1%)
Finland	28.7% (28.2%)	14.0% (8.73%)	18.5% (11.2%)	24.9% (18.4%)	13.7% (33.3%)
France	30.4% (25.7%)	14.5% (8.05%)	19.2% (11.3%)	24.9% (15.9%)	10.8% (38.9%)
Germany*	-	-	5.06% (7.33%)	27.5% (24.7%)	67.2% (53.4%)
Hungary	37.2% (37.1%)	15.1% (8.58%)	15.4% (9.34%)	20.5% (16.7%)	11.6% (28.1%)
Italy	23.0% (41.0%)	18.3% (11.8%)	21.2% (10.8%)	25.7% (14.2%)	11.6% (21.8%)
Lithuania	23.4% (28.8%)	13.9% (11.1%)	20.2% (15.7%)	29.4% (23.0%)	12.8% (21.1%)
Netherlands	16.9% (26.2%)	13.4% (8.50%)	20.1% (11.5%)	30.1% (20.9%)	19.2% (32.8%)
Poland*	-	-	13.4% (13.6%)	40.2% (34.4%)	46.2% (51.9%)
Portugal	36.5% (32.1%)	16.4% (11.8%)	19.4% (13.7%)	19.8% (18.4%)	7.66% (23.8%)
Romania	29.3% (21.7%)	13.3% (8.17%)	18.4% (12.4%)	28.2% (23.3%)	10.5% (34.2%)
Slovakia*	-	-	13.9% (14.6%)	34.3% (32.9%)	51.7% (52.4%)
Slovenia	24.3% (36.9%)	11.7% (9.96%)	16.3% (8.53%)	28.6% (22.6%)	18.8% (21.8%)
Spain	33.3% (37.7%)	17.2% (9.54%)	20.2% (11.4%)	17.2% (14.6%)	11.8% (26.6%)
Sweden	39.8% (21.9%)	17.7% (9.72%)	22.3% (13.4%)	17.9% (20.0%)	2.01% (34.8%)

Sources: CompNet 6th vintage and Eurostat, 2011.

Notes: Number in parenthesis refers to the figures in Eurostat Structural Business Statistics. * refers to the 20E sample. In the case of Germany only manufacturing firms are included.

2.4 Sample biases

Even when firms in the raw data are weighted to replicate the distribution by size and sector in the population, there could be within-cell biases. That is, even if the share of micro-firms in manufacturing, for example, is the exact same as in the population, CompNet could be sampling only the very best firms within that cell so they are not representative of the average manufacturing-micro firm in the country. To assess the presence of this type of bias, **Table 2.6** compares the average firm size in a given macro-sector and size class in CompNet and in the population of firms (numbers of Eurostat reported in brackets).

According to the numbers shown in the table, there is no evidence of within-cell biases in the CompNet data. The only exception is the smaller-than-in-the-population size of the very large firms operating in any of the main macro-sectors, which is consistent with the findings in

Table 2.5 and might be a by-product of the very strict cleaning protocol.

Table 2.6: Average firm size within each cell, in CompNet and the population, 2011

Belgium				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.64 (13.24)	31.23 (30.90)	104.9 (103.8)	598.4 (736.3)
Construction	13.42 (13.23)	30.31 (29.94)	82.15 (97.46)	312 (514)
Services	13.48 (13.27)	30.37 (30.06)	90.19 (98.68)	718.2 (1301.9)

Croatia				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.44 (13.22)	30.22 (29.88)	106.1 (102.2)	541.6 (548.5)
Construction	13.56 (13.35)	30.21 (29.71)	90.44 (99.08)	361.5 (587)
Services	13.25 (12.59)	29.84 (26.34)	91.45 (84.95)	387.3 (463.0)

Denmark				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	14.09 (13.29)	31.34 (30.87)	98.45 (97.56)	585.6 (779.0)
Construction	13.80 (13.42)	29.83 (29.42)	80.15 (89.66)	n.a.
Services	13.83 (13.44)	29.90 (26.58)	87.52 (97.87)	410.0 (720.3)

Finland				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	14.08 (13.28)	31.20 (30.29)	99.61 (102.1)	501.1 (790.4)
Construction	13.78 (13.77)	29.15 (29.20)	78.66 (90.14)	n.a.
Services	13.84 (14.98)	30.16 (31.54)	90.61 (116.1)	391.4 (886.1)

Hungary				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.82 (13.75)	31.82 (31.59)	103.2 (106.3)	582.5 (769.3)
Construction	13.51 (13.45)	28.36 (29.54)	92.19 (92.90)	n.a.
Services	13.46 (13.41)	29.67 (30.04)	83.31 (96.65)	413.2 (940.48)

France				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.51 (13.98)	31.16 (34.57)	103.2 (111.3)	511.8 (846.6)
Construction	13.35 (14.80)	29.94 (32.36)	80.32 (100.9)	n.a.
Services	13.30 (16.31)	30.60 (35.93)	95.09 (122.9)	410.7 (1763.)

Italy				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	14.07 (13.33)	31.13 (30.08)	95.51 (96.78)	435.0 (722.3)
Construction	13.66 (12.90)	29.73 (28.81)	75.62 (85.76)	n.a.
Services	13.68 (12.94)	30.41 (29.69)	92.41 (97.66)	525.3 (1167.)

Lithuania				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.65 (13.41)	30.98 (29.91)	98.80 (80.12)	436.9 (441.2)
Construction	13.52 (13.50)	29.97 (30.02)	89.33 (96.26)	374.2 (409.4)
Services	13.23 (13.18)	29.65 (32.31)	86.93 (78.39)	522.9 (784)

Table continued

Netherlands				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.70 (14.92)	30.17 (34.59)	95.09 (108.4)	566.3 (606.6)
Construction	13.54 (14.67)	29.82 (32.19)	89.94 (96.75)	n.a.
Services	13.51 (18.20)	30.03 (39.19)	93.92 (130.9)	n.a.

Portugal				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.56 (13.50)	30.44 (30.43)	90.04 (95.89)	433.1 (500.2)
Construction	13.05 (13.01)	29.52 (29.46)	79.06 (91.88)	376.7 (703.5)
Services	13.11 (13.08)	29.61 (29.68)	82.72 (93.70)	685.3 (997.6)

Romania				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.73 (13.73)	31.10 (31.08)	106.0 (106.1)	537.2 (698.6)
Construction	13.44 (13.44)	30.27 (30.16)	89.91 (98.29)	n.a.
Services	13.20 (13.32)	29.65 (29.86)	93.98 (101.1)	n.a.

Slovenia				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	14.12 (13.38)	31.69 (27.19)	109.4 (104.3)	658.5 (569.8)
Construction	13.89 (13.12)	30.32 (29.29)	83.58 (93.38)	n.a.
Services	13.69 (13.20)	30.11 (16.35)	92.07 (85.63)	409.3 (706.0)

Spain				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.53 (13.38)	29.81 (29.89)	94.26 (99.93)	641.7 (685.0)
Construction	13.19 (13.71)	29.06 (29.82)	76.53 (94.31)	538.6 (856.5)
Services	13.33 (13.19)	29.33 (29.43)	91.86 (100.0)	735.3 (1153.)

Sweden				
Sector	**10-19 Employees**	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	13.49 (15.21)	30.65 (33.77)	92.07 (112.1)	326.5 (811.0)
Construction	13.25 (15.09)	29.04 (32.86)	66.48 (95.86)	n.a.
Services	13.31 (14.49)	29.41 (34.27)	72.69 (111.3)	495.6 (997.6)

Poland*			
Sector	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	32.96 (30.34)	111.4 (108.4)	536.3 (652.5)
Construction	31.45 (29.08)	99.30 (97.49)	415.8 (644.0)
Services	30.93 (29.19)	101.6 (102.4)	732.0 (909.5)

Slovakia*			
Sector	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	32.34 (30.04)	108.5 (108.4)	678.6 (742.9)
Construction	31.21 (30.86)	92.77 (93.00)	512.8 (597.2)
Services	31.41 (29.75)	97.76 (96.06)	724.9 (785.3)

Germany*			
Sector	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	34.35 (34.14)	114.1 (106.4)	685.2 (909.1)

Czech Republic*			
Sector	**20-49 Employees**	**50-249 Employees**	**>250 Employees**
Manufacturing	31.42 (30.62)	109.5 (106.1)	564.8 (663.6)
Construction	29.61 (29.23)	95.36 (92.53)	478.3 (681.6)
Services	29.79 (29.25)	100.7 (98.12)	634.7 (1053.)

Sources: 6th vintage of CompNet and Eurostat.

Note: Numbers in parenthesis refer to the figures in Eurostat Structural Business Statistics. * Figures rely on the 20E sample.

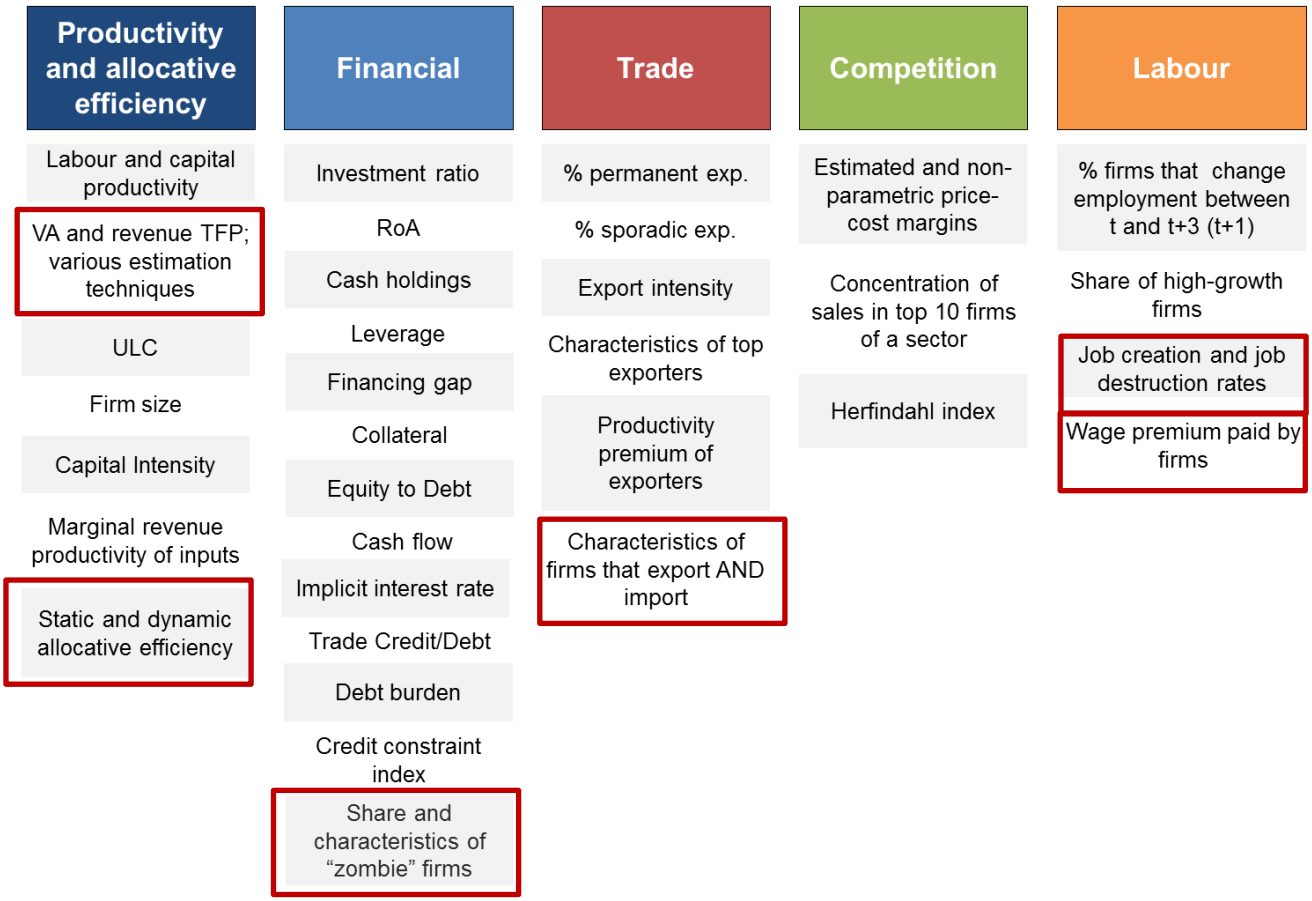
3. Overview of indicators

3.1 Modules

The range of indicators collected by CompNet is very broad in order to meet the needs of policy-makers and researchers participating in the network. **Figure 3.1** provides an overview of the main indicators organized by topic (the new indicators are highlighted). The productivity and allocative efficiency module covers TFP, capital and labour productivity as well as several measures of input

misallocation. The financial module covers a wide range of indicators of the financial position of the firm, including our own credit constrained indicator developed using firm-level data from the ECB-SAFE survey (see Ferrando et al. 2015 and the user guide for more details). The trade module is quite unique in that it computes export intensity and productivity and financial characteristics of exporters in each 2-digit industry of manufacturing in a comparable way across all countries with export information (see Berthou et al. 2015 and the user guide for more details). The competition module computes parametric and non-parametric mark-ups as well as other indicators of sector competition like the Herfindahl index or the share in sector sales of the top 10 firms in terms of size (see the user guide for more details). Finally, the labour module includes transition matrices which record the share of firms moving from any quantile of the size distribution to another one within a sector, and job flows (see Fernandez et al. 2017 and the user guide for more details).

Figure 3.1: Main indicators collected, by topic



The new vintage of data includes some novelties which will be briefly reviewed here. For more details, please refer to the updated user guide of CompNet's dataset (Aglio et al. 2018).

Regarding productivity and allocative efficiency, the innovations are twofold. First, Total Factor Productivity is now estimated according to different methodologies, both parametric and non-parametric, letting the researcher choose which one to use. The non-parametric measures include real value added of the firm divided by a weighted average of inputs used (with weights 1/3 for capital and 2/3 for labour) while the parametric ones are based on both revenue and value added production function estimations assuming either a Cobb-Douglas or a Translog production function. Further, we now compute the Petrin-Sivadasan indicator of labour misallocation (see Petrin and Sivadasan 2013). According to this methodology, the mean absolute gap between marginal productivity and cost of a given input across firms can be interpreted as an approximation to the aggregate gain that would occur if every firm had a one-unit change in that input in the efficient direction, holding everything else constant. Compared to the Hsieh-Klenow indicator of misallocation, which is also computed, the Petrin-Sivadasan indicator has the advantage of not assuming equal marginal costs across all firms in a given industry. The disadvantage, however, is that the marginal cost of labour has to be approximated by the average wage at the firm. Note that besides the Petrin-Sivadasan and the Hsieh-Klenow misallocation indicators, the code also computes the Olley-Pakes gap, i.e. the within-sector covariance between the relative productivity and size of a firm, which proxies to what extent resources are allocated to the most productive firms in the sector.

The financial module has also been expanded to include new indicators on "distressed firms". These are non-viable firms that are still operating in the market. The definition of "non-viable firm" varies in the literature so we provide different indicators for the researcher to choose: first, we flag firms with positive profits but below interest payments over three consecutive years; second, firms with persistent (3 consecutive years) negative profits which are not high-growth firms; and, finally, firms that price below marginal costs. The dataset includes the share of each of these firms in any level of aggregation as well as a complete set of their characteristics and joint distributions. For more information, refer to Section 4.3.

Regarding trade, we have added information on the import intensity (besides the export intensity) of firms and also of two-way traders, that is, of firms both exporting and importing. Regarding competition measures and to improve upon previous vintages, we added a new state-of-the-art firm-level mark-up estimation routine based on the approach by De Loecker and Warzynski (2012). Finally, the labour module includes now job flows following Davis et al (1996), i.e. job creation and destruction rates, of firms operating in a given level of aggregation, and also by size class. The labour module includes as

well a new variable aimed at proxying firm's human capital: the wage premium. The wage premium is computed as the ratio of the average wage in a given firm to the median wage paid by firms operating in the same 2-digit industry. The idea behind this indicator is that if a firm pays more than its competitors in a narrowly defined sector, it could be because its workers are more skilled.

3.2 Levels of aggregation and statistics

In CompNet we consider 5 different levels of aggregation: the country level, the NUTS2 regional level, the macro-sector level (sectors at 1-digit corresponding to the NACE REV.2 sections), the macro-sector and size class[10] and, finally, the sector level (at 2-digit in NACE REV.2). CompNet constructs one dataset for each level of aggregation including all distributions and joint distributions referring only to firms operating at that particular level of aggregation.

As mentioned in the introduction, the beauty of the CompNet dataset is that it does not only collect the mean of each of the indicators for each level of aggregation but also further moments/properties of the distribution. See **Table 3.1** for a list of the summary statistics collected for each indicator in CompNet.

Table 3.1: Statistics of each indicator included in the CompNet database

Moment in CompNet	Definition
p1, p5, p10, p25, p50, p75, p90, p95, p99	Population percentiles of the variable
mean	Mean of the variable
sd	Standard deviation of the variable
skew	Skewness of the variable
kurt	Kurtosis of the variable's distribution
sum_weights	Number of firms in the population
tot_mark, count, obs	Number of observations in the sample

Notes: the moments, apart from tot_mark, of the distribution are estimated for the population in CompNet using indicator specific weights.

3.3 Joint distributions

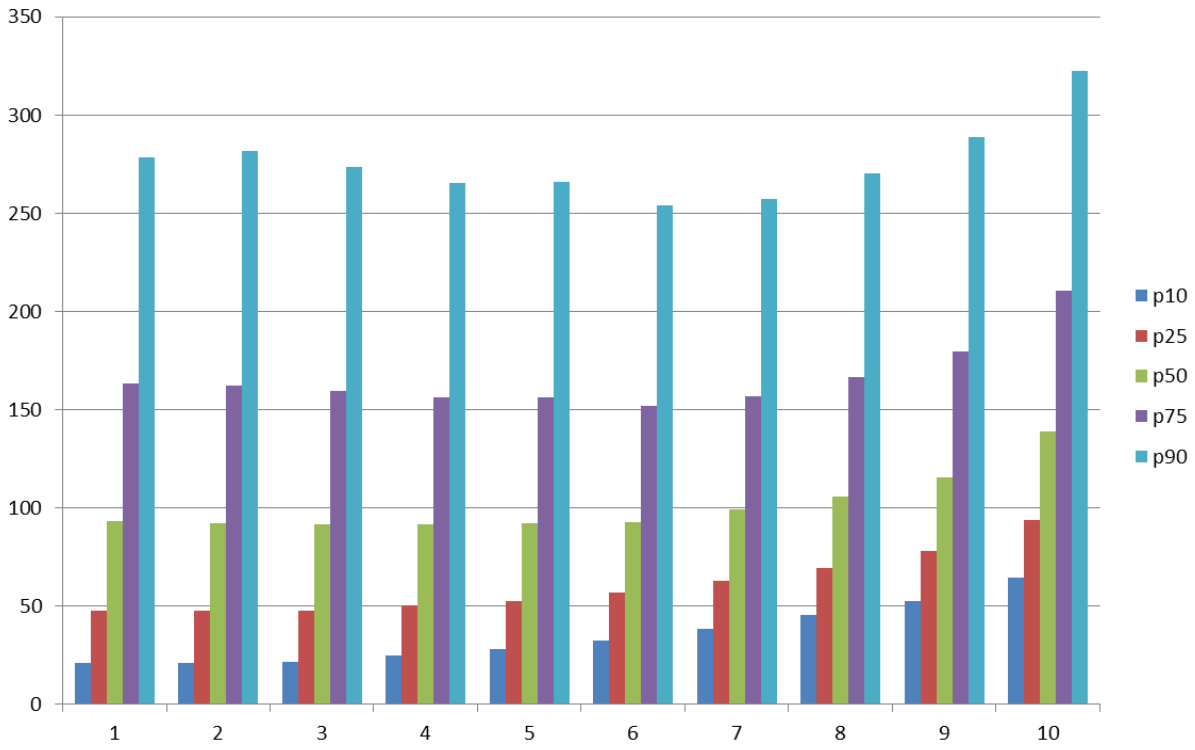
Finally, CompNet computes the so-called "joint distributions". For a given level of aggregation, we consider firms in each decile of the distribution of one of our main (continuous) variables and compute their characteristics. This is a very powerful tool to understand the differences between, for example,

[10] CompNet considers five size classes: 1-9, 10-19 employees, 20 to 49 employees, 50 to 249 employees and 250 employees or more.

firms at the top and bottom of the size or productivity distribution. For variables which take discrete values, for example a dummy taking the value 1 if a firm is classified as “distressed” and 0 if classified as a “healthy” firm, the joint distribution database provides complete information on firms in each of those splits, within a given level of aggregation. This enables the researcher to explore, following that example, different characteristics of “distressed” and “healthy” firms in a given 2-digit industry, across countries and years.

Figure 3.2 shows an example of a (continuous) joint distribution. In particular, it shows the productivity distribution of firms in different deciles of the firm size distribution in Sweden in 2011. It becomes clear that there is a discontinuity in the productivity of firms by size. Up to the median of the size distribution, the productivity of firms does not change significantly with size. However, productivity increases continuously with size when considering firms in the upper half of the size distribution. The interesting part is that not only the mean or median productivity of firms increases with size, but the full distribution of productivity shifts to the right.

Figure 3.2: Labour productivity distribution by size decile in Sweden



Note: Full sample of Sweden, year 2011. Labour productivity computed as total turnover divided by total number of workers.

Table 3.2 shows the set of main indicators, both continuous and discrete, which serve as the basis for the joint distributions, while **Table 3.3** shows the set of characteristics computed for firms in each decile of the distribution of the main indicators in **Table 3.2**.

Table 3.2: Main indicators used for the joint distributions

Main variables		
Topic	**Name**	**Definition**
Financial	Credit constraint index	Data for firms that display financial restrictions when planning investments or credit constrained firms, based on the Ferrando-Ruggieri (2015) approach.
	Share and characteristics of distressed firms	Distributions conditional on firms being in a distressed status, using various methodologies based on markups, interest payment to operating profits ratio, negative profits and negative profits conditional on not displaying high growth
	Interest rate	Interest over operating profits
	Investment ratio	Growth rate of capital plus depreciation, divided by Capital
Trade	% permanent exporters	Data for exporting and non exporting firms
	top 10 exporters	Information related to top 10 exporters by export revenues
Labour	Employment	Data conditional on deciles of the size distribution, according to employment
Productivity and allocative efficiency	Labour productivity	Data conditional on deciles of log of ratio of real value added over employees
	Solow Residual	Data conditional on deciles of non-parametric Solow's residuals
	TFP	Data conditional on deciles of TFP from a Cobb-Douglas estimation at different levels of aggregation

Table 3.3: Characteristics of firms in each decile of the main indicators

Characteristics		
Topic	**Name**	**Definition**
Financial	Credit constraint index	Data for firms that display financial restrictions when planning investments or credit constrained firms, based on the Ferrando-Ruggieri (2015) approach.
	Share and characteristics of distressed firms	Firms being in a distressed status, using various methodologies based on markups, interest payment to operating profits ratio, negative profits and negative profits conditional on not displaying high growth
	Cash holdings	Cash divided by total assets
	Collateral	Capital divided by total assets
	Debt burden	Interest paid over total assets
	Equity to Debt	Equity divided by debt
	Implicit interest rate	Interest over operating profits
	Investment ratio	Growth rate of capital plus depreciation, divided by Capital
	Leverage	Debt divided by total assets
	RoA & Profit Margin	Operating profit-loss divided by total assets & EBIT over turnover
	Capital	Fixed assets deflated with GDP deflator
	Real Turnover & Value Added	Turnover and Value added deflated with the GDP or sector specific deflator
Trade	% exporters	Data for exporting and non exporting firms
	% export status	Share of firms exporting since more than 3 years, sporadic exporters, firms exporting and importing
	Import and Export Intensity	Import/Export value over turnover
	Trade credit & Trade debt	Accounts payable/receivable over total assets
Labour	% firms that change employment	Annual employment growth rate bwtween t and t+1 (t+3)
	Labour	Average of the number of workers in FTE over the year
	Labour costs	Labour costs divided by value added, average labour cost, wage premium
	Job creation rate and Job destruction rate	Weighted average of positive/negative growth rates of number of employees
Competition	Herfindahl index	Hirschman-Herfindahl index of market concentration
	Markups	Markup à la DeLoecker-Warzynski (2012) with materials as input, from a Cobb-Douglas estimation at various aggregation levels and using value added or revenues as inputs
	Price-cost margins	Estimated and non estimated: revenues minus materials minus labour cost, divided by revenues
	Top 10 firms' market power	Concentration of turnover in top 10 firms of a given sector
Productivity and allocative efficiency	Capital growth rate	Annual real capital growth rate
	Indicator of market imperfection	Dobbelaere-Mairesse (2013) indicator of market imperfection, using different production function specifications
	Labour and Capital productivity	Real value added divided by capital or turnover/real value added divided by employees
	Marginal revenue product of inputs	Marginal revenue product of capital and labour, estimated according to different specifications
	Misallocation	Petrin-Sivadasan gap (2013) measure of misallocation, using various specifications at different levels of aggregation
	Capital intensity	Real capital divided by the number of employees
	SR	Non-parametric Solow's residuals, from the equation: logSR = log(real value added) - 1/3*log(real capital) - 2/3*log(number of employees)
	TFP	TFP from a Cobb-Douglas estimation at various levels of aggregation, adjusted or non adjusted
	Unit Labour Cost	Nominal labour costs divided by real value added

4. New evidence from CompNet

This section validates, describes and "puts to work" selected indicators of the 6th vintage of the CompNet dataset. The aim is threefold. *First*, each indicator is contrasted with similar aggregated data from other sources to ensure we find common dynamics. It should be kept in mind, however, that the underlying data of CompNet refers to non-financial corporations with employees operating in the business economy. This means that it collects information of only a part of the economy, excluding agriculture, utilities, the public and financial sector, as well as self-employed workers. Consequently, one cannot expect an exact match of CompNet data with aggregate sources, although they should certainly show consistent developments over time and across countries.[11] *Second*, we aim at showing some of the granularity available for each indicator, in terms of size, sector or productivity details to name but a few. *Third*, and most importantly, we provide examples of how this dataset can help shedding light on several issues currently debated in the policy or academic fora. To show the potential of the data, each section indicates possible ways to put the indicators to work and displays some new evidence on specific topical issues. The purpose is not to carry out an in-depth analysis, but to suggest how the new CompNet dataset may support further work in certain areas.

4.1 Productivity

Productivity is the key determinant of welfare in most economic models and the core driver of competitiveness. It is, therefore, one of the core indicators of CompNet. However, the concept is elusive. For this reason, CompNet offers to data users a wide variety of productivity estimators and indicators to serve the needs of diverse research agendas. We provide a non-parametric labour productivity and Solow residual using fixed weights. In addition to such productivity indicator derived from macroeconomic concepts, we also estimate one- and two-digit sectoral level production functions to determine firm-level TFP. Specifically, we follow the Wooldridge (2009) methodology to control for the well-known simultaneity bias, which combines computational simplicity and wide adoption. Using this technique, we estimate value added and gross output production functions at the sector and macro-sector level and store all resulting firm-level TFP estimates.

[11] See also the accompanying cross-country comparability report (Altomonte et al. 2018).

4.1.1 Definition, validation and granularity

In more detail, our battery of productivity indicators includes[12]:

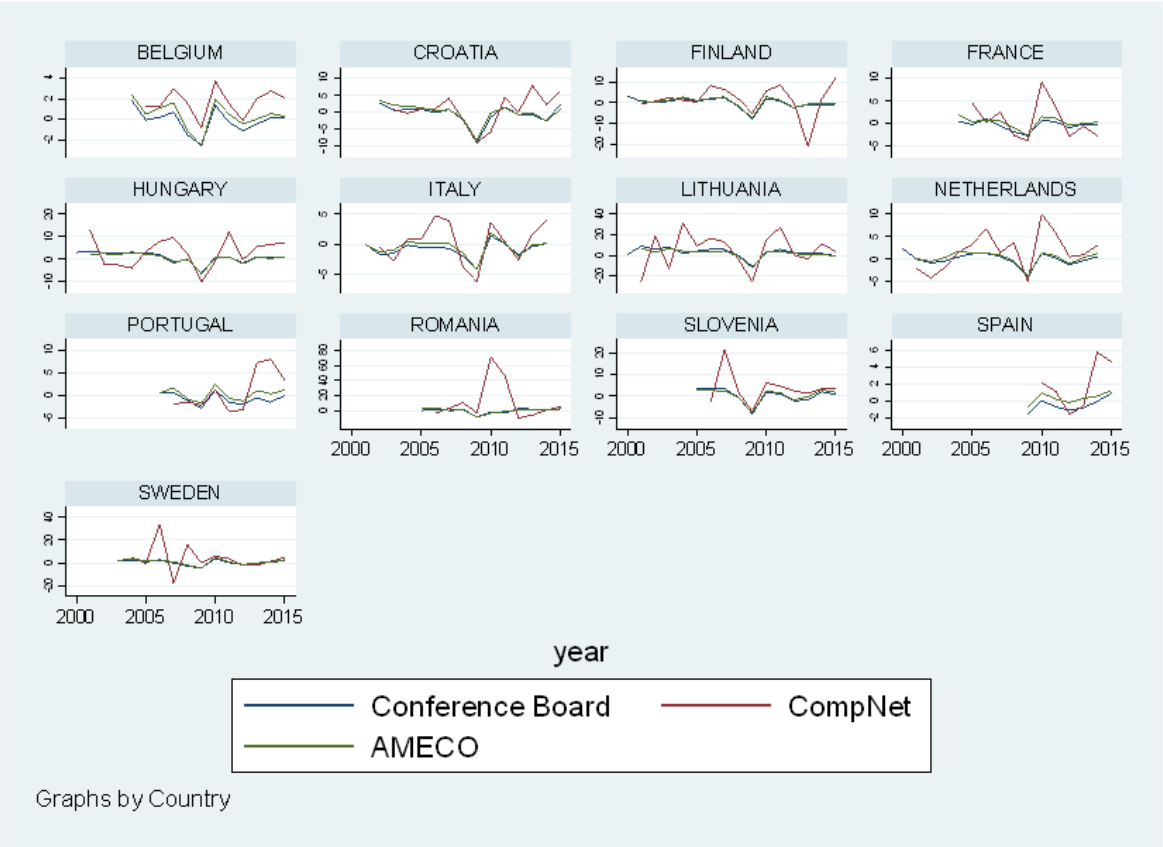
- Non-parametric measures of labour productivity: labour productivity is computed as (i) real value added per worker or (ii) real turnover per worker.
- Parametric TFP and marginal revenue productivity of inputs: we estimate eight different production functions, including value added and gross output production Cobb Douglas and Translog functions at the 2-digit (sector) and 1-digit (macro-sector) NACE level. We provide estimates for the country-specific elasticities of inputs, TFP estimates and, uniquely, derived marginal productivities of labour and capital.
- Solow residual: as an "hybrid" between micro-level production function estimation and macro-level growth accounting, we also provide a Solow residual - i.e. as the ratio of a firm's real value added to its inputs labour and capital, weighted 2/3 and 1/3 respectively.

To validate the data we compare TFP growth computed in CompNet with that from other aggregate sources like AMECO or the Conference Board. Note, however, that both aggregate sources derive their statistics from national accounts using different techniques to estimate the contribution of certain segments of the economy, like self-employed workers, and calculate TFP as a Solow residual. CompNet, differently, computes TFP at the firm level and tracks the mean (or median) in a sector or in the overall business sector over time. For the purpose of this comparison exercise, we consider the CompNet TFP indicator computed as the Solow residual defined above and use the full sample of firms.

Figure 4.1 shows that despite the arguably large differences in terms of concept and measurement, the dynamics in TFP computed by CompNet track reasonably well those of aggregate sources, although volatility is higher in CompNet, maybe due to the fact that CompNet includes only the business sector.

[12] For further technical details we refer the reader to the updated user's guide, Aglio, D . et al. (2018).

Figure 4.1: TFP growth: Conference Board and AMECO vs. CompNet

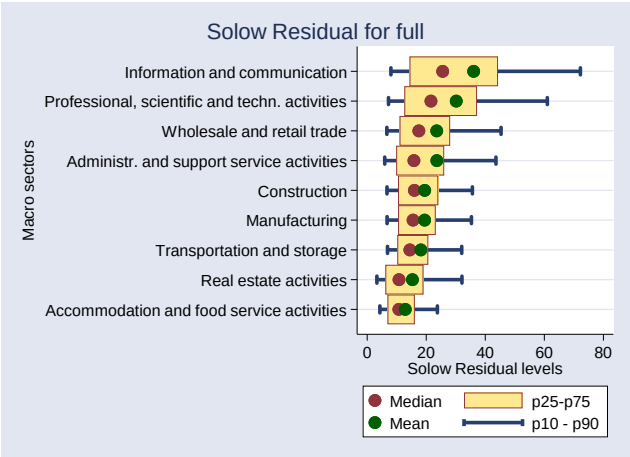


Sources: 6th vintage of CompNet, full sample, AMECO and Conference Board.
Note: CompNet TFP data refer to firms with employees operating in the non-financial business sector. TFP in CompNet is calculated as the ratio of a firm's real value added to its inputs labour and capital, weighted 2/3 and 1/3 respectively. We show the growth rate of the mean in the business sector. Conference Board and AMECO refer to the whole economy and calculate TFP as difference between aggregate real output growth and aggregate input growth, as derived from the national accounts (Solow residual). No data from Denmark in Eurostat.

Given that our data closely follow aggregate performance, one can use the granularity of CompNet to drill deeper and try to understand which firms contribute the most to aggregate performance. Most modern macroeconomic models keep tractability by featuring many (often even homogenous) firms, each of which is too small to affect the aggregate (e.g. Smets and Wouters 2007). However, as it is evident from CompNet's more granular data, this approach is not plausible. Top productive firms are responsible for a large part of the productivity dynamics of different sectors. We illustrate this using macro-sector level TFP data from 14 European countries with full sample. We show separately countries in Western and Eastern Europe as productivity levels are still quite different in both regions.

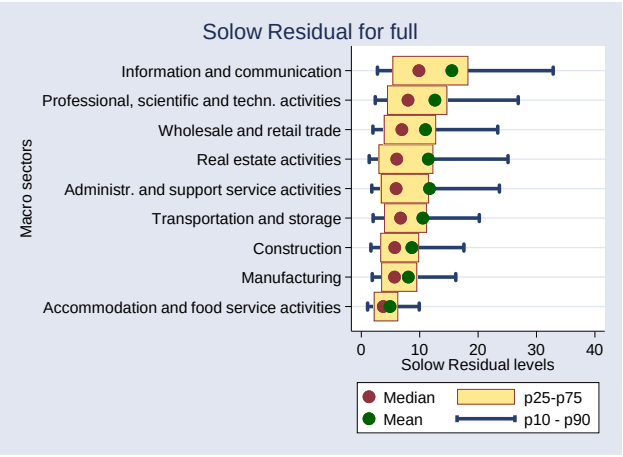
The boxplots below show different deciles of the distribution of TFP in each macro-sector (unweighted average across countries), as well as the mean of the distribution.[13] It is clear that sector differences in mean TFP are driven by top firms while TFP at the bottom of the distribution is comparable across all sectors.

Figure 4.2: Productivity distribution by macro-sector: Western European countries, 2010



Sources: 6th vintage of CompNet, full sample, year 2010.
Notes: the countries included are BE, DK, FI, FR, IT, NL, PT, ES and SE.

Figure 4.3: Productivity distribution by macro-sector: Eastern European countries, 2010



Sources: 6th vintage of CompNet, full sample, year 2010.
Notes: the countries included are HR, HU, LT, RO and SI.

4.1.2 Putting the indicators to work: Growth in the age of superstar firms

Firm size and market concentration have been steadily increasing over the last decades. Recently, a number of papers have explored the negative consequences of this development (e.g. Autor et al. 2018, Eeckhout and Kircher 2018, De Loecker and Eeckhout 2017). However, the top productive firms, which have a large influence on mean productivity in each sector as shown above, tend also to be the largest ones, as well-established evidence (including from CompNet) suggests.

In the following, we explore the connection between superstar firms and sector productivity growth. To do so, cross-country 2-digit sectors are grouped into 3 categories according to their mean TFP growth:

[13] TFP computed as a residual from macro-level production functions, where labour has a coefficient of 2/3 and capital of 1/3

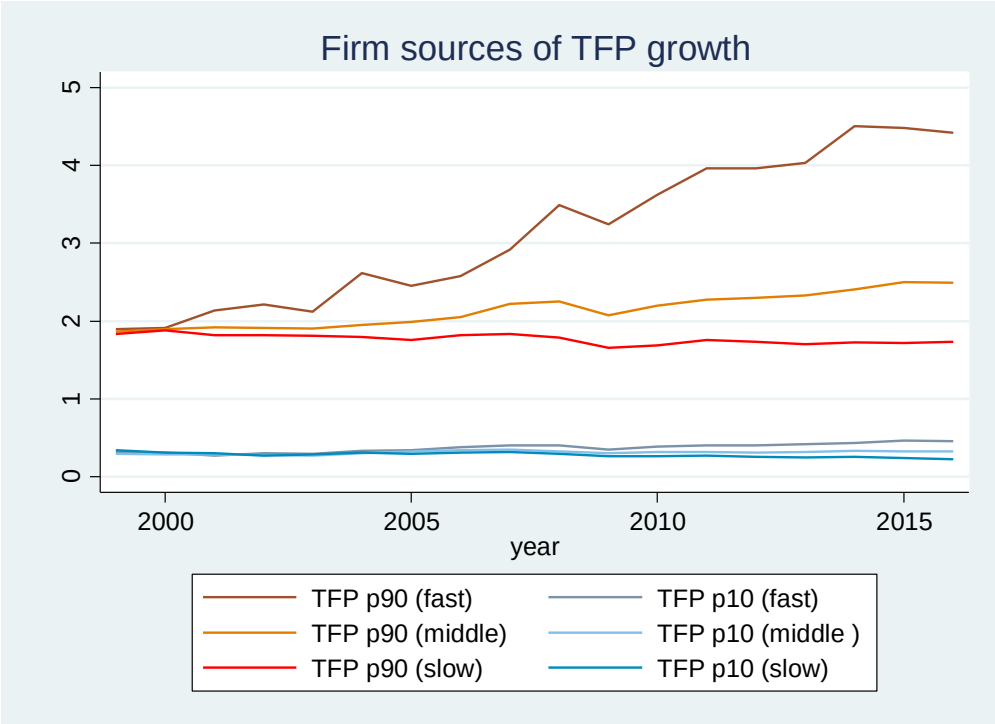
(1) *fast* growth sectors belong to the upper third of the yearly distribution[14], (2) *moderately* growing sectors belong to the middle third of the distribution; and (3) *slow* growing sectors are those in the bottom third of the distribution. **Figure 4.4** shows the dynamics of the top and bottom 10 percentile firms in terms of TFP in each of the 3 groups of sectors described above. We find that the most important factor that allows differentiating across sectors' performance is the TFP performance of firms at the top 10% of the distribution, while the performance of lower tails is similar in the 3 groups of sectors[15]. In other words, in sectors with relatively high productivity growth, the distribution of productivity has gotten more skewed. Obviously, this is not the only source of aggregate productivity gains: entry and exit and reallocation of resources also affect aggregate performance. Nonetheless, without movement in the underlying distribution of firm efficiency, long-term growth cannot materialize. It is thus worrying that over the last 15 years and over a wide variety of sectors and countries, movement of the productivity distribution has been associated with a widening gap between the best performers and the rest.

This finding is a challenge for economic policy aimed at reigning in top/largest firms: hindering their expansion might actually slow down aggregate sector productivity growth, since the other segments of the distribution show little dynamism. Hence, as long as this pattern remains, there may be a trade-off between fostering productivity growth and reducing firm concentration. Policies aimed at speeding up the catching up of less productive firms with high potential could alleviate this trade-off and increase the productivity potential.

[14] There are 550 sectors in the distribution. This is done for every year.

[15] Note that this is a repeated cross-section so we do not follow the same firm over time.

Figure 4.4: TFP cumulative growth of tail firms in sectors with different dynamics



Sources: 6th vintage of CompNet data, full sample.
Notes: The chart includes the 14 countries with full sample information. TFP is indexed to average productivity in 1999, which is the first year.

4.2 Unit Labour Costs

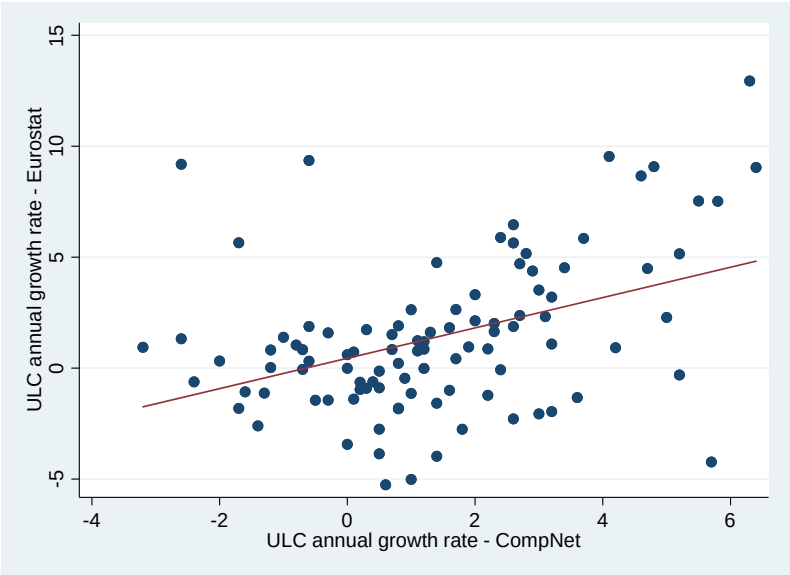
A measure widely used to assess the development of cost-competitiveness of countries is the Unit Labour Cost (ULC) — the ratio between labour compensation per employee and labour productivity. It measures the extent to which labour costs raise in line with productivity gains, going up if compensation rises faster than productivity — a possible indication of falling cost-competitiveness.

4.2.1 Definition, validation and granularity

ULC is computed in CompNet as average nominal labour cost per employee divided by real labour productivity at the firm level. This definition is also consistent with the one used by Eurostat, the ECB, the European Commission or the OECD.[16]

To get a sense of the plausibility of the data gathered under CompNet, **Figure 4.5** shows the ULC growth rate of the median firm in the business sector[17] in each country-year in CompNet versus the annual ULC growth rate published by Eurostat and computed as average compensation per employee divided by GDP per person employed. Despite the different definitions and coverage of the two datasets, the correlation is about 0.4.

Figure 4.5: Correlation between the annual growth rate of ULC in CompNet and Eurostat



Sources: Eurostat and 6th vintage of CompNet, 20E sample.
Notes: countries included are BE, HR, DE, CZ, FI, FR, HU, IT, LT, PL, NL, PT, ES, SK, SI, and SE over the period 2006-2015.

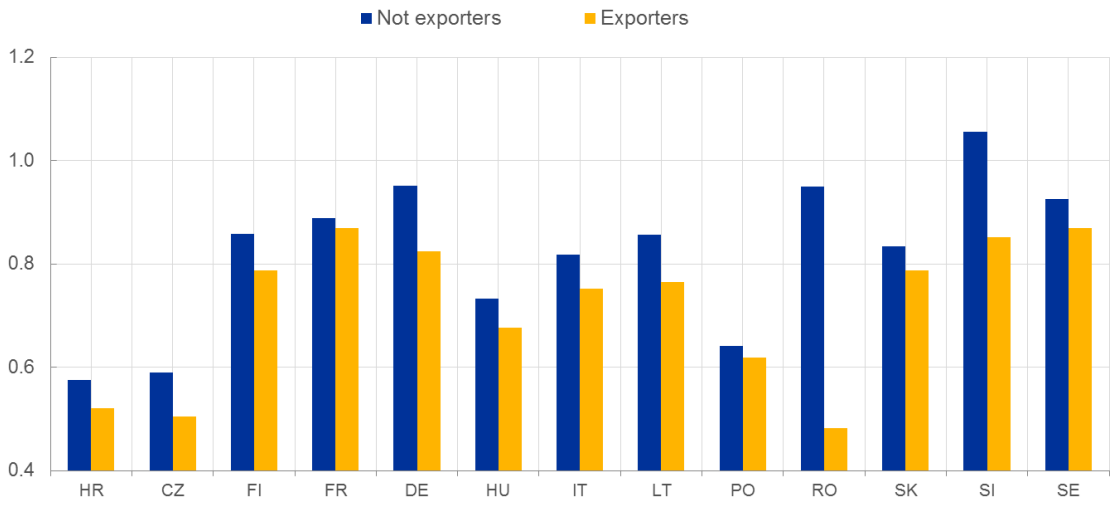
Aggregate figures can hide substantial differences across sectors, or types of firms within a given sector. Particularly, we expect that internationally active firms display higher cost-competitiveness than

[16] Alternative definitions of ULC are nominal labour costs divided by nominal value added or real labour costs (deflated by CPI or similar deflators) divided by real value added.

[17] The median value is used because it is more robust to outliers than the mean value.

firms not subject to international competition. **Figure 4.6** shows median ULC of exporting and non-exporting firms operating in the same 2-digit manufacturing industry (according to the NACE REV.2 classification), that is, operating within narrowly defined sectors, aggregated up to the country level using sector weights. Although exporting firms are more competitive than non-exporting firms across all countries, the difference is largest in Eastern countries, where ULC of non-exporting firms is about 12% higher (50% if we include Romania) than that of exporting firms in the same sector vs. 8% in Western countries. France, on the other hand, displays the smallest difference.

Figure 4.6: ULC of exporting and non-exporting firms operating in the same sector



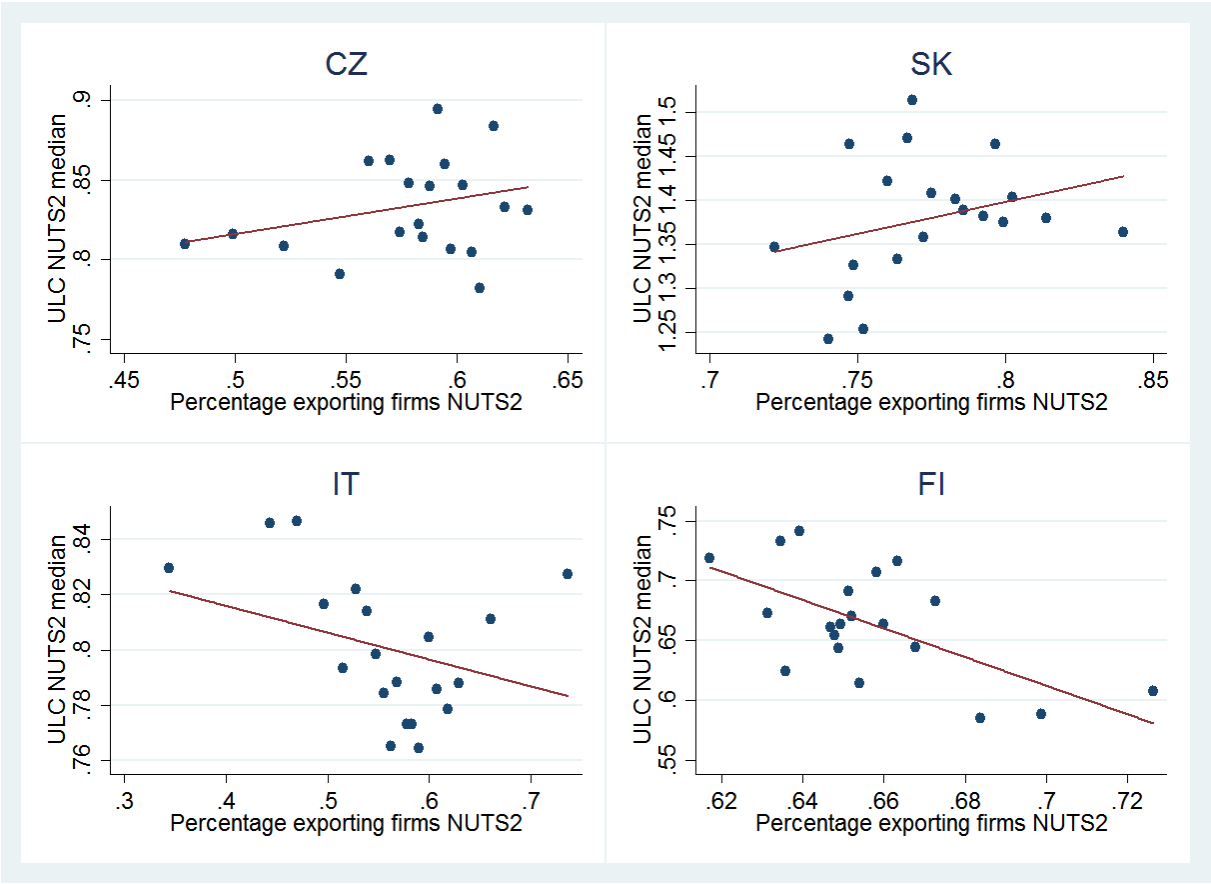
Sources: Own calculations on 6th vintage of CompNet, sample of firms with at least 20 employees.

4.2.2 Putting the indicators to work: ULC and the competitiveness of export-oriented regions

As a new feature of CompNet, we have collected data at the NUTS2 regional level within each country. We use this information to explore the question of whether regions more exposed to international competition show higher competitiveness (or lower ULC). For the purpose of the section, international exposure is measured as the share of exporting firms in each region. Given that we only have export information for manufacturing firms, the charts refer to the manufacturing sector only. **Figure 4.7** shows results for two Western (Italy and Finland) and two Eastern (Czech Republic and Slovakia) economies. The reason is that dynamics are quite different between both regions. In particular, for the two Western economies included we find a negative relation between the share of exporting firms and the median ULC in the region, particularly strong for Finland and somewhat

weaker for Italy. The situation in the two Eastern countries is radically different though: there is a positive correlation between the median ULC and the share of exporting firms in the region.[18] These findings could be related to the fact that eastern European countries are deeply involved in global value chains and, therefore, might be able to incorporate into the production process high-quality inputs which require the employment of skilled, more expensive labour (see Chiacchio et al. 2018). However, this is only one hypothesis. Understanding the factors behind these interesting differences deserves further research.

Figure 4.7: Median ULC and regional export orientation



Sources: 6th vintage of CompNet, sample of firms with at least 20 employees.
Notes: 38 NUTS2 regions; the period covered is 2005-2015. Export orientation is measured by share of exporting firms in a given sector (only sectors 10-33 manufacturing business included).

[18] The figure is a bin-scatter plot controlling for NUTS2-region fixed effects.

4.3 Distressed firms

By distressed firms we refer to firms active in the market despite displaying characteristics that would be expected to force them to exit. Indeed, in a well-functioning economy the creative-destruction process, fuelled by competition, should be able to urge poorly performing firms to improve their efficiency, shrink or shut down. Therefore, the documented rise of distressed firms (see for example McGowan et al., 2017) can be interpreted as a sign that this process is either distorted or slowing down, which could have important consequences on the overall economy, decreasing aggregate efficiency and productivity.

There is a growing literature concerned with the impact of distressed firms on the overall economy, and particularly on healthy firms.[19] This literature focuses on three channels: 1) distressed firms exhibit lower levels of productivity and hold back innovation; 2) distressed firms hold resources (including credit) that could be more efficiently assigned to better performing firms; 3) distressed firms exhibit lower turnover and investment. Concerning the drivers, a number of recent papers explore the link between this phenomenon of distressed firms and the quality of the banking sector. Schivardi et al. (2017), in particular, show that in Italy under-capitalized banks tend to prolong credit to distressed firms, helping their survival, to a greater extent than stronger banks. It is in this context that distressed firms are labelled as “zombie” firms.

4.3.1 Definition, validation and granularity

The economic literature has used different definitions of distressed firms. Among those, the OECD (McGowan et al., 2017) defines distressed firms as firms having an interest coverage ratio below 1 over three consecutive years, conditional on having operated at least 10 years.[20] The reason for the latter is that young firms, particularly if they are high-growth firms, could record very low profits the first years of operations but, still, be viable in the medium term. Other authors, instead, define non-viable firms as firms with persistent negative profits (Bank of England 2013).

Given that the characteristics of firms with an interest coverage ratio below 1 *and positive profits* are radically different from those with an interest coverage ratio below 1 *and negative profits*, CompNet

[19] See Cooper, Haltiwanger and Power (1999); Bank of England (2013); Bank of Korea (2013); Acharya et al. (2017); Deutsche Bank (2018); Andrews et al., (2016); McGowan et al., (2017).

[20] The interest coverage ratio is defined as the ratio of operating income to interest expenses.

collects information on both types of distressed firms separately. Therefore, distressed firms are defined as follows in CompNet:

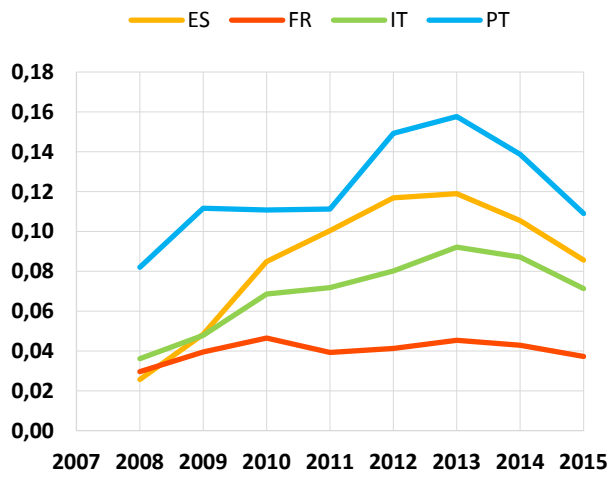
- Interest coverage definition: we flag firms facing an interest coverage ratio below 1 for three consecutive years, *conditional on the firm having positive profits*.
- Negative profits: firms with negative operating profits for three consecutive years.
- Not high-growth: firms with negative profits for three consecutive years, excluding firms that have experienced high growth over the same period.[21] By doing so, this definition refrains from flagging firms which frontload a lot of investment to facilitate growth. It is therefore a similar rationale as the one applied by the OECD when choosing firms with more than 10 years of activity.
- Below 1 mark-ups: it flags firms pricing below marginal costs.

To validate the data, we plot the evolution over time of the CompNet share of *not high-growth* distressed firms (shown in **Panel C of Figure 4.8**) and compare it with the shares of distressed firms from two other sources: 1) the paper by Storz et al. (2017), which uses the ORBIS database and defines distressed firms as firms with negative investment, negative return on assets and the ratio of EBITDA to financial debt of less than 5% for 2 consecutive years (**Panel A of Figure 4.8**); 2) the ECB survey on the access to finance of enterprises (SAFE), where distressed firms are defined as firms with lower turnover, lower profits and higher interest expenses compared to the previous six months (**Panel B of Figure 4.8**). Despite the different definitions, the dynamics are roughly similar, in all instances: the share of distressed firms steadily increases over time up to 2013, and starts declining thereafter.

[21] High growth is defined as firms undergoing average annual employment growth of more than 20% over the t-3/t period.

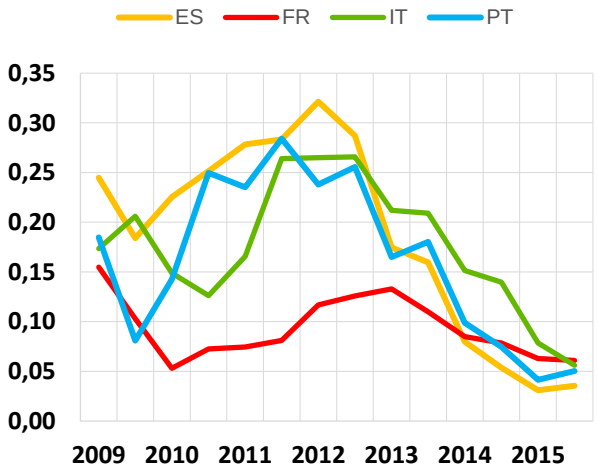
Figure 4.8: Share of distressed firms over time, different sources

Panel A: ORBIS



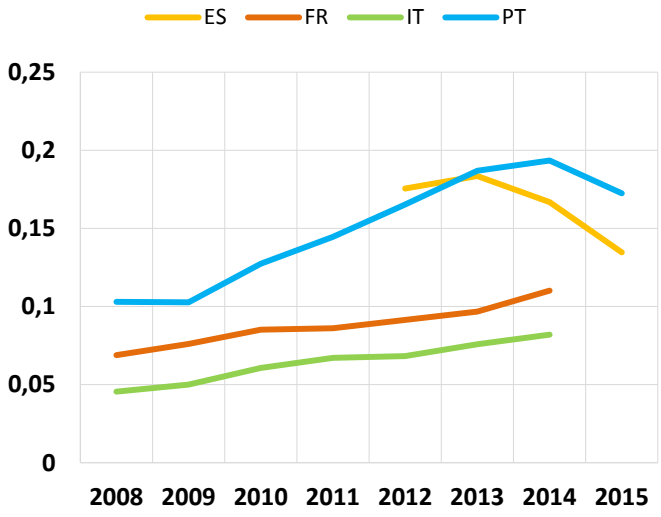
Source: ORBIS.
Note: Distressed firms are defined according to Storz et al. (2017): Non-financial firms with negative investment, negative return on assets and EBITDA to financial debt of less than 5% for two consecutive years.

Panel B: SAFE-survey



Sources: SAFE survey
Notes: countries included are ES, FR, IT and PT, over the period 2009 to 2015, Distressed firms are defined as firms experiencing lower turnover, lower profits and higher interest expenses compared to the previous six months

Panel C: CompNet *Not high growth* definition.

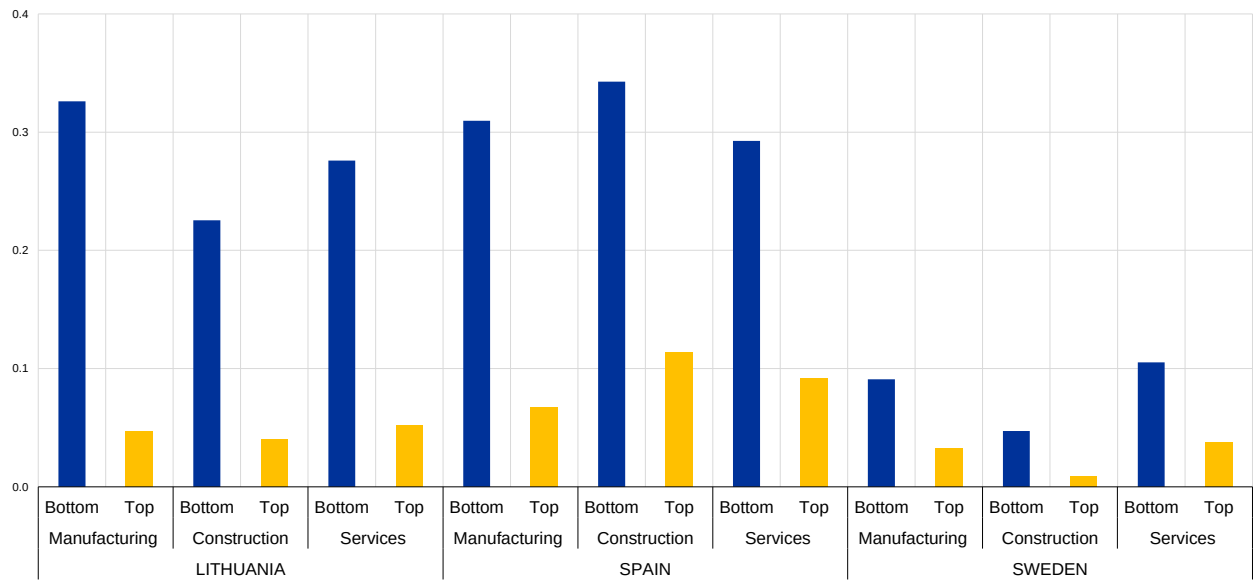


Sources: 6th vintage of CompNet, drawing from the full sample.
Notes: Not high growth captures firms with negative operating profits for three consecutive years, excluding the firms that experienced high growth in employment during that period. Countries included are ES, FR, IT, PT over the period 2008 to 2015.

CompNet collects information not only on the aggregate share of distressed firms, but also on their characteristics or prevalence across sectors, size classes and productivity deciles. To get a sense of the granular richness of the information present in the 6th data vintage, **Figure 4.9** shows the share of distressed firms at the top and bottom of the productivity distribution in different macro-sectors in 2013. We show the results for 3 countries: one Euro-area Eastern country (Lithuania), one Euro-area Western economy (Spain) and one non-Euro area Western country (Sweden). In all countries, we observe a significantly higher concentration of distressed firms among the bottom performing firms in each sector, which is what we should expect.

The granularity of CompNet is not only exploited by looking at differences along the productivity distribution, but also in terms of sectoral differences. For instance we see that in Lithuania the concentration of distressed firms in the high productivity group is similar across sectors, while the share of distressed firms amongst the low productive firms is highest in the manufacturing sector. In Spain, on the other hand, the share of distressed firms, both at the top and bottom of the productivity distribution is highest in construction while in Sweden they can be found in services.

Figure 4.9: Share of distressed firms at the top and bottom of the sector productivity distribution



Sources: 6th vintage of CompNet, full sample.
Notes: Not high definition of zombie firms is used. Low (high) productive defined as firm belonging to the 20th (80th) centile of the productivity distribution using revenue based Cobb-Douglas productivity function.

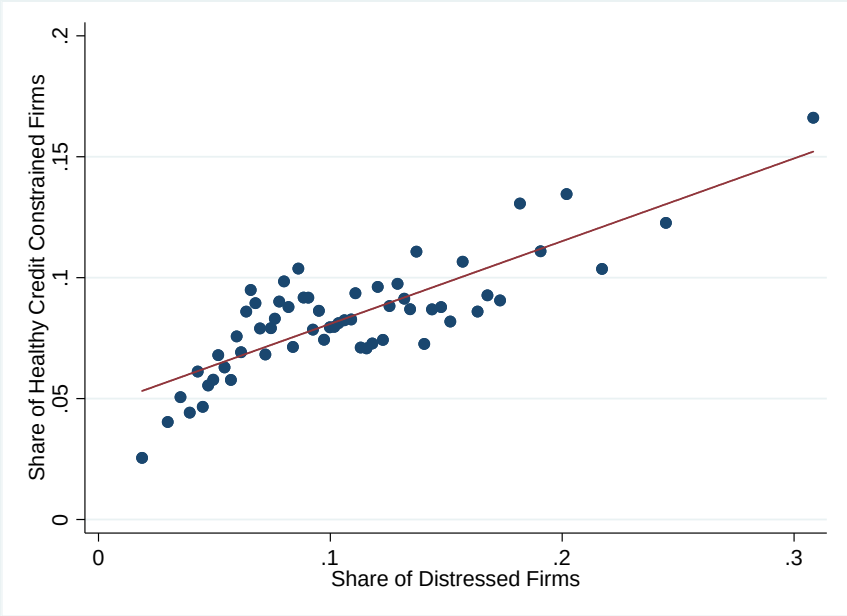
4.3.2 Putting the indicators to work: "Zombie" firm congestion

As mentioned above, the existence of distressed firms represents a potential source of inefficiency since these firms can hold back resources that could otherwise be reallocated to more productive entities. We follow McGowan et al. (2017) to explore in detail the nature and consequences of the phenomenon. In their paper, McGowan et al. (2017) show that a higher share of industry capital owned by distressed firms (defined as firms that have persistent difficulties in serving their interest payments) is associated with lower capital investment and lower employment growth of the typical "healthy" firm and also to less productivity-enhancing capital reallocation. This, in turn, increases the negative impact of distressed firms on economy-wide investment and employment.

Along these lines, we start by showing the sway of distressed firms on other (healthy) firms in the same sector and then move to their overall impact on sector employment and investment. As it was mentioned above, distressed firms can hold productive resources, for example credit, which could otherwise be reallocated to more productive firms (Gopinath et al. 2017, Schivardi et al. 2017). We show that this is indeed the case in **Figure 4.10** where the share of distressed firms – defined as firms with negative profits excluding high-growth firms – in a given country-sector-year is plotted against the share of credit constrained healthy firms in the same country-sector-year.[22]

[22] The share of credit constrained firms is computed according to the CompNet ICC indicator. This indicator ranks firms according to their probability of being credit constrained, which depends on their financial position and size, and then computes the share above a certain threshold provided by SAFE. For more information please refer to Ferrando et al. (2015).

Figure 4.10: Share of distressed firms and credit constrained healthy firms in the same sector



Sources: 6th vintage of CompNet, full sample.
Notes: Bin-scatter controlling for country FE. The share of credit constrained firms is measured according to the CompNet ICC indicator of credit constraints (see Ferrando et al. 2015). On the x axis we have the share of distressed firms according to the not high growth definition. Both variables are measured at the 2-digit industry for a given year and country. The countries included are: BE, HR, DK, FI, FR, HU, IT, LT, NL, PL, PT, RO, ES, SI and SE.

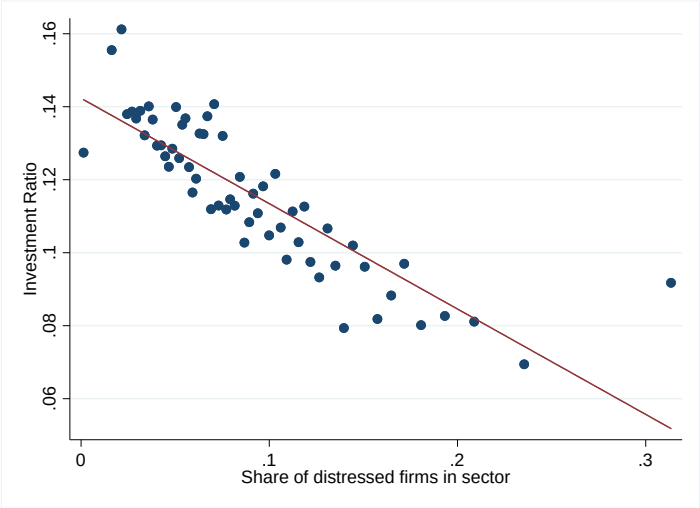
This positive correlation is consistent with the reference literature (Acharya et al., 2017; Andrews and Petroulakis, 2017) and serves as motivation for further studies along this line. In particular, Acharya et al. (2017) showed how the favourable post-crisis monetary conditions caused credit to flow more abundantly to low-productivity firms with pre-existing lending relationships with banks, rather than to high productivity firms. Therefore, healthy firms in industries with a prevalence of distressed firms suffered significantly from credit misallocation, which slowed down the economic recovery.

We turn next to the impact of distressed firms on average sector investment and job creation. **Figure 4.11** shows the correlation between the share of distressed firms in a given country-sector-year and the corresponding investment ratio[23] in 14 EU countries with full sample while **Figure 4.12** reports the correlation with the job creation rate. The definition of "distressed firms" used in the chart is based on

[23] Defined as growth rate of capital plus depreciation, divided by capital.

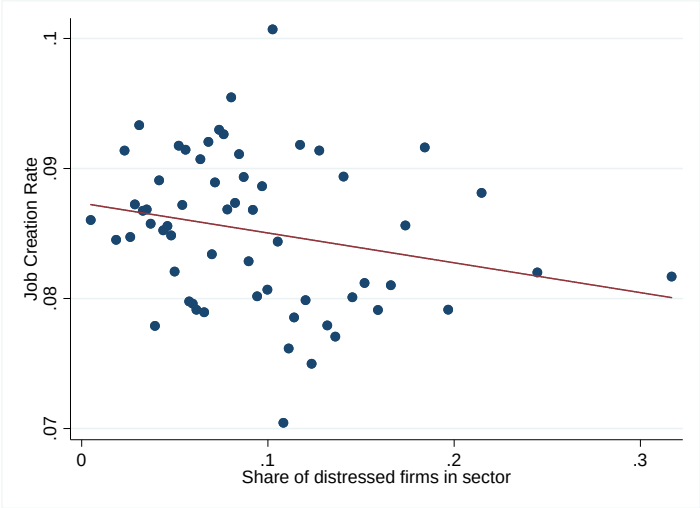
the *interest coverage* definition, i.e. firms with operating profits-interest payments ratio lower than 1 for 3 consecutive years (conditional on positive profits), in order to better compare with the results in McGowan et al. (2017).

Figure 4.11: Correlation between sector share of distressed firms and investment



Source: 6th vintage of CompNet, full sample.
Notes: On the y axis the median investment ratio is reported, measured as growth rate of capital plus depreciation, divided by capital. Zombie firms are defined as firms with interest payments higher than operating profits for 3 consecutive years, conditional on positive profits. Both variables are measured at the 2-digit industry for a given year and country, aggregated in a small number of bins for visualization purposes. The graph is based on the full sample of BE, HR, DK, FI, FR, HU, IT, LT, NL, PT, RO, ES, SI, and SE.

Figure 4.12: Correlation between sector share of distressed firms and job creation rate



Source: 6th vintage of CompNet, full sample.
Notes: On the y axis the median job creation rate is reported, measured as weighted average of positive growth rates of number of employees. Zombie firms are defined as firms with interest payments higher than operating profits for 3 consecutive years, conditional on positive profits. Both variables are measured at the 2-digit industry for a given year and country, aggregated in a small number of bins for visualization purposes. The graph is based on the full sample of BE, HR, DK, FI, FR, HU, IT, LT, NL, PT, RO, ES, SI, and SE.

In line with the existing literature, in country-sectors-years where the concentration of distressed firms is more prominent, the overall levels of both investments and creation of new jobs is significantly lower. Hence, the macroeconomic implications of the prevalence of these firms can be sizeable.

4.4 Trade

Availability of comparable *cross-country* data on export and import activity is highly appreciated by policy makers and researchers but scarce, particularly when considering granular information. In order to fill this gap, data providers were asked to merge firms' balance sheet data with information on their

export activities whenever available. The provision of cross-country indicators of export and import intensity, as well as information on characteristics of exporters (and of different type of exporters) within each of the 2-digit manufacturing industries constitutes one of the main comparative advantages of CompNet. More specifically, within a three-year window, the dataset allows distinguishing between new (not exporting at *t-1* and exporting at *t* and *t+1*) and established exporters (exporting at *t-1*, *t* and *t+1*), switchers (exporting at *t-1*, not exporting at *t* and exporting at *t+1*), two-way traders (firms that both import and export over the full period), and firms that stop exporting (exporting at *t-1* and not exporting at *t* and *t+1*).[24]

4.4.1 Definition, validation and granularity

CompNet collects annual data on exports and imports, in nominal terms, for *manufacturing sector firms*. In the following, we focus on exports. A firm is considered an exporter if in any given year its export value is above 1,000 EUR and exports represent at least 0.5% of its total turnover.[25]

Table 4.1 below provides a full overview of the sources of data, time coverage and thresholds' levels by country. Note that reporting thresholds vary across countries and time. Time coverage also varies across countries and spans over at least a 10-year period. For most countries where export data are available, information on imports is also provided.

[24] We have also collected the prevalence rate and characteristics of different types of exporters defined over a two-year window. The reason is that requiring a firm to stay in the sample for three consecutive years, in order to classify it as a given type of exporter, reduces sometimes sensibly the sample size.

[25] As observed total exports in the custom databases and alike can be larger than the total turnover recorded in the annual accounts, values of exports exceeding 150% of total turnover have been considered to be misreported and omitted. Note that for countries that use custom or intrastat / extrastat declarations on exports at the firm level, the minimum amount of exports may be much larger (for instance, in Belgium for the 2006-2010 period, intra EU trade is observed for firms exporting to the EU 27 at least 600,000 EUR in a given year).

Table 4.1: Country-specific information on exports

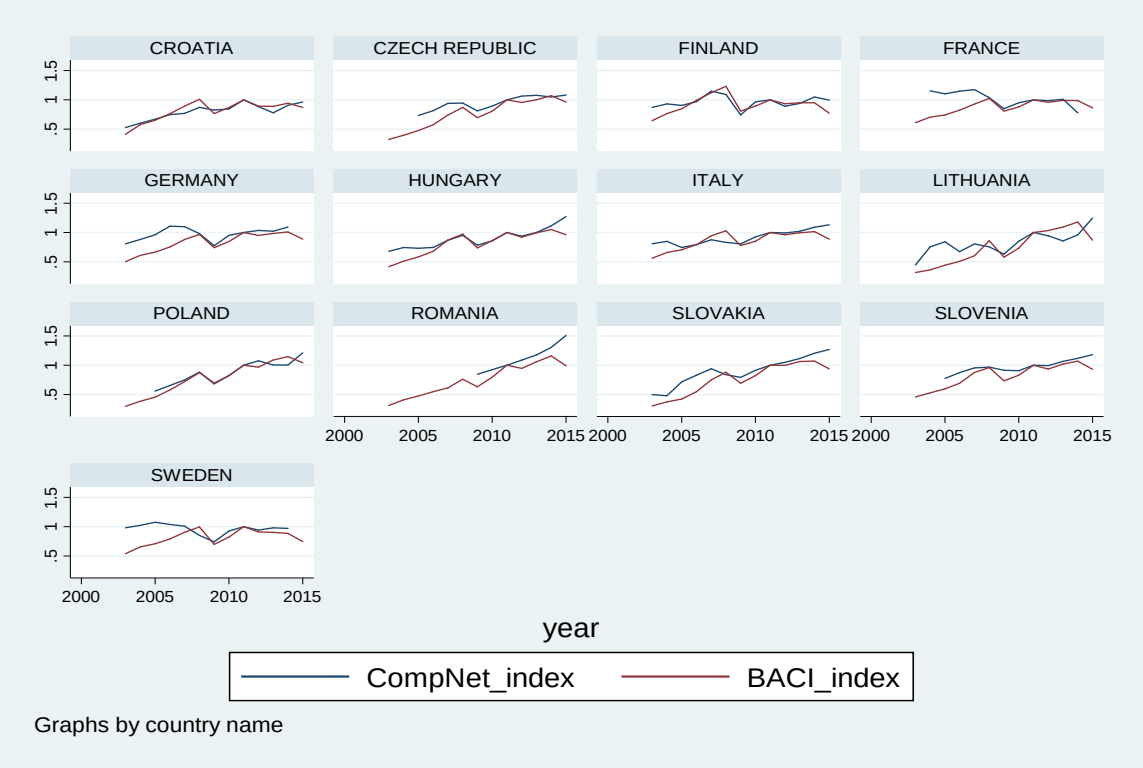
Country	Export Data	Import Data	Coverage	Data Source	Reporting Threshold
BE	No	No			
CZ*	Yes	Yes	2005-2015	Foreign trade transaction data (Statistics Czech Republic)	Intra EU transaction thresholds: 4 mil. CZK before 2008, 8 mil. CZK after.
DK	No	No			
DE*	Yes	No	1999-2014	Administrative firm-level data (Statistics Germany)	Intra EU transaction thresholds: 500,000 euros (until 2004, 200,000 euros). Extra EU transaction threshold: 1,000 euros.
ES	No	No			
FI	Yes	Yes	1999-2015	International trade statistics data (Finnish Costums)	Intra-EU imports / exports in euros : 100,913 /100,913 (2000-2001); 100,000 / 100,000 (2002-2005) ; 100 000 / 200 000 (2006-2007) ; 200,000/ 300,000 (2008-2010) ; 275 000 / 500 000 (2011-2013); 500 000/500 000 (2014-2015). Extra-EU: 1,000 euros until 2008 and no threshold 2009-2015.
FR*	Yes	No	2004-2014	Statistical Office of France	Intra-EU: threshold based on total intra-EU exports for the calendar year 38,100 euros (1998) ; 99,100 (2001); 100,000 (2002); 150,000 (2006); 460,000 (2011-2014). Extra-EU: 1,000 euros per transaction.
HR	Yes	Yes	2002-2016	Financial Agency Croatia	Intra EU transaction threshold: 1,000,000 kn (exports), 1,900,000 kn (imports), in 2018.
HU*	Yes	Yes	1999-2015	Export-Import data of Hungarian Enterprises (Statistics Hungary)	Intra EU transaction threshold: exports threshold in Million HUFs 25 for 2004 and 100 since, for imports 25 in 2004, 40 in 2005, 60 in 2006-2007, 100 million since 2008.
IT	Yes	Yes	2001-2014	Foreign trade statistics data, based on customs data (Statistics Italy)	Annual threshold of 1000 euros.
LT	Yes	Yes		Customs, customs declarations (Customs of the Republic of Lithuania)	Reporting thresholds based on import and export values. Threshold only make up 5% of total trade. Trade data is biased towards larger firms.
NL	No	No			
PL*	Yes	No	2005-2015	Statistical Office of Poland	Only firms with more than 10 employees are included.
PT	No	No			
RO*	Yes	Yes	2005-2015	Exports and imports of goods, firm-level data (Statistics Romania)	Intra EU transaction threshold: 900,000 lei (2018).
SI	Yes	No	2005-2016	Agency for Public Legal Records and Related Services	Intra EU transaction threshold: 220,000 euros (exports), 140,000 euros (imports), in 2018.
SK*	Yes	Yes	2004-2015	foreign trade statistics (Statistics Slovakia)	Intra EU transaction threshold: 400,000 euros (exports), 200,000 euros (imports), in 2018.
SE	Yes	Yes	2003-2014	International trade in goods (Statistics Sweden)	Trade data contains exports above the thresholds of 4.5 mil. SEK, and imports above 9 mil. SEK. Trade below the threshold is determined via VAT declarations.

*CompNet data available for the 20E sample.

In order to validate the CompNet export data we compare the evolution of aggregate exports for the manufacturing sector in the CompNet and the CEPII-BACI dataset for the overlapping countries[26]. As can be seen in **Figure 4.13** both series show very similar dynamics for all countries in the sample.

[26] We consider the trade data provided by the CEPII-BACI because it includes information on export values and quantities by country pairs and 6-digit products of the Harmonized Commodity Description dataset over an extended period of time. Differently, data available in Eurostat used to compare aggregate levels in 2012 is only available for a short time frame. Note that the levels in both datasets are not exactly comparable since some of these goods could be exported by non-manufacturing firms.

Figure 4.13: Export dynamics in CompNet and CEPII-BACI, 2011=1



Sources: 6th vintage of CompNet 20E sample and CEPII-BACI dataset.
Note: The index is equal to 1 in 2011.

4.4.2 Putting the indicators to work: The Happy Few

Over the last years, studies on international trade have moved the focus away from industries towards firms and products (see Melitz 2003, Bernard et al. 2003). The increasing availability of granular data has indeed shed additional light on the role of firm heterogeneity and uncovered a number of important facts. First of all, a large share of exports and employment is generated by a very small number of exporting firms, called the “happy few” following the work of Mayer and Ottaviano (2008).In other words, although exporters represent a very small share of the total number of active firms, they account for large share of economic activity. This is confirmed with CompNet data and shown in **Table 4.2** below. The first column of the Table shows the (very low) percentage of firms with employees operating in international markets, which ranges from 4% in Italy to 46% in Slovenia. The prevalence rate of exporters is much higher if one considers only firms with at least 20 employees, given the concentration of exporting firms in large size classes (**Table 4.2, second column**). But even in the

latter case, exporters account for a disproportionate share of sector employment, labour costs, real value added or turnover (**Table 4.2, columns 3-6**).

Table 4.2: Relative importance of exporting firms. Share of total, 2013

Country	Number of firms with at least one employee	Number of firms with at least 20 employees	Employment (20E)	Labour Costs (20E)	Real Value Added (20E)	Real Turnover (20E)
CZ		24%	37%	36%	34%	33%
DE		25%	65%	72%	72%	79%
FI	14%	60%	74%	76%	80%	82%
FR		63%	74%	76%	78%	78%
HR	26%	65%	71%	76%	77%	86%
HU		48%	64%	73%	70%	83%
IT	4%	24%	29%	33%	35%	34%
LT	17%	61%	72%	77%	81%	92%
PL		64%	76%	77%	74%	73%
RO		33%	53%	62%	69%	73%
SE	20%	72%	78%	79%	80%	85%
SI	46%	88%	93%	94%	96%	96%
SK		80%	85%	86%	85%	89%
Average	21%	54%	67%	71%	72%	76%

Source: 6th vintage of CompNet, full and 20E sample.

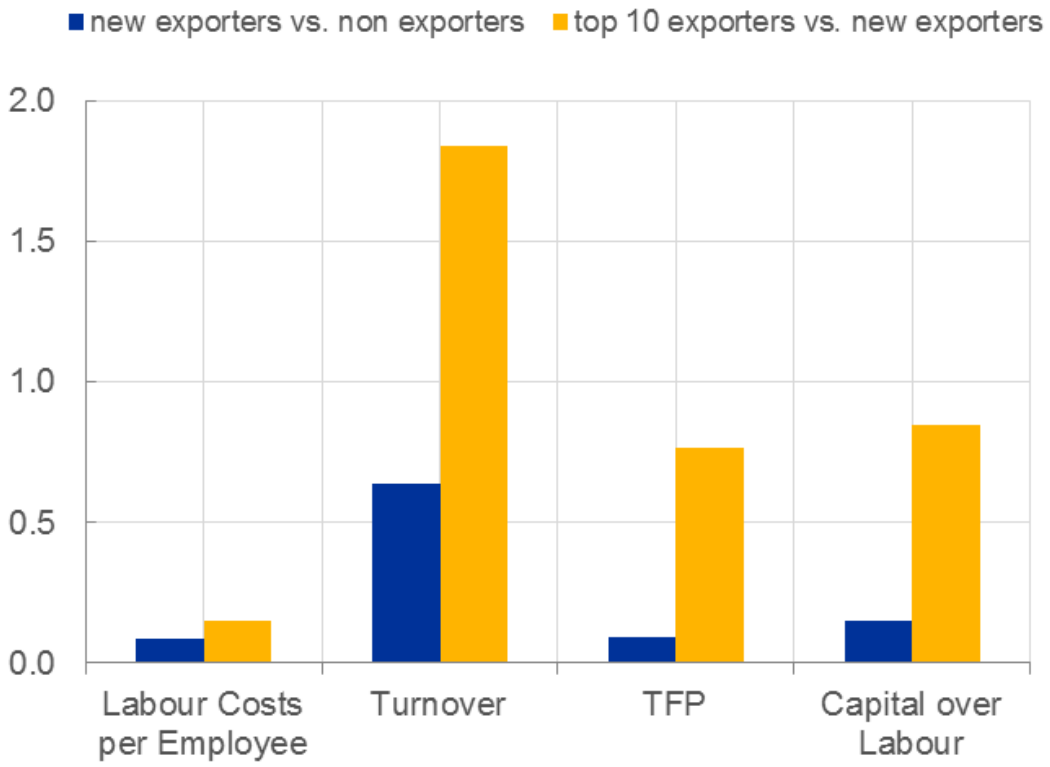
Secondly, exporters are substantially different from domestic firms: they are larger, more productive, more capital-intensive and pay higher wages. One important question is whether this is the result of *self-selection into exporting*, or whether there is *learning by exporting*, i.e. the exporting activity itself is a driver of a firm's productivity increase, for example due to exposure to best practices and new technologies. It could be also a combination of both, i.e. there can be a "two-way link" between trade and productivity.[27]

We explore this issue with CompNet data by estimating the premia of new exporting firms relative to non-exporting firms operating in the same 2-digit industry, after controlling for country, sector and time fixed-effects. If new exporting firms show higher productivity or size than domestic in their same sector, it would be evidence of *self-selection* of the best firms into exporting. To explore if there is also *learning by exporting*, we estimate the premia of top exporters in the sector, that is, of firms that have

[27] See ECB (2017).

been exporting for a prolonged period of time, relative to new exporters in the same sector. Consistently with the empirical literature, we find that new exporters are significantly larger, in terms of turnover, pay higher wages, are more efficient and more capital intensive than domestic firms operating in their same sector. The premia is even larger in all instances if one compares the performance of top exporters with those of firms which just started to export. These results highlight the existence of a two-way link between productivity and trade. [28]

Figure 4.14: Premia of new and top exporting firms relative to other firms in the same sector



Source: 6th vintage of CompNet, 20E sample.
Note: The chart shows the coefficients of the export dummy, indicating whether the firm is exporter or not, from OLS regressions where the dependent variable is the log of the performance indicators, controlling for country, time and sector dummies. Countries included are FI, HR, IT, LT, SE and SI.

[28] Note that this exercise was done with the 20E sample. That is, we compare exporting and domestic firms with at least 20 employees. The estimated premia using the full sample, that is, comparing exporting firms with domestic firms with at least 1 employee, are in all cases larger.

4.5 Mark-ups

The presence of market power influences economic activity: it determines innovation incentives, the allocation of resources and market shares across firms, and, thereby, long run aggregate production and welfare. This implies a need to understand the dynamics and determinants of market power.

The 6th vintage of the CompNet dataset sheds light on how market power has evolved across European countries through a number of indicators: price-cost margins, price over marginal cost mark-ups, Hirschman-Herfindahl indices, and profit margins. Since we cannot cover all indicators in this report, we focus in this section on mark-ups. For a comprehensive definition of other market power indicators we refer the interested reader to the user's guide (Aglio et al. 2018).

4.5.1 Definition, validation and granularity

CompNet calculates firm-specific mark-ups based on different gross output production function specifications by using the framework of De Loecker and Warzynski (2012). Specifically, mark-ups are computed as:

$$\boldsymbol{\mu_{it} = \alpha_{it}^{M} * \frac{P_{it}Q_{it}}{P_{it}^{M}M_{it}}}, \qquad \textbf{(1)}$$

where μ_{it} denotes the mark-up, α_{it}^{M} is the output elasticity of intermediate inputs, and $\frac{P_{it}Q_{it}}{P_{it}^{M}M_{it}}$ is the inverse of the share of intermediate input expenditures in revenues. We recover α_{it}^{M} from estimating a production function at different levels of aggregation (macro-sectors and NACE Rev. 2 two-digits sectors) and assuming different functional forms (Cobb-Douglas and Translog). From equation (1) it becomes clear that mark-up levels can be very sensitive to different intermediate input variable definitions. Therefore, it is important to note that potential differences in country-specific variable definitions can be absorbed by using appropriate normalizing techniques when comparing mark-up evolutions across different countries (see below).

Alternatively, mark-ups can be calculated using firms' labour input decision instead of using firms' intermediate input decision. However, we believe that the later generates a more reliable and comparable mark-up estimate because a necessary condition for an unbiased mark-up estimate is that the production factor based on which mark-ups are calculated is a flexible input (e.g. De Loecker and Warzynski 2012; De Loecker et al. 2016). Changing flexibility in labour markets (e.g. due to labour market reforms) might introduce variation in estimated mark-ups that reflects changes in input rather

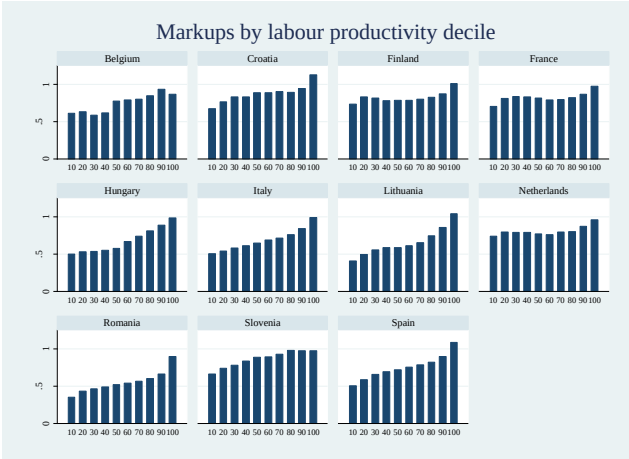
than output market environments. We believe that such concerns are especially relevant in a cross-country context and, therefore, prefer to use firms' intermediate input decision to calculate mark-ups.

Validating our mark-up measures against other aggregated sources is difficult as no other database offers micro-aggregated mark-up estimates. However, we can check whether our mark-up estimates display correlations with other key variables that are in line with the literature.

Most existing studies report a positive correlation between productivity (profitability), size, and mark-ups (e.g. De Loecker and Warzynski 2012; De Loecker et al. 2016; Dhyne et al. 2011). Note however, that theoretically such correlation depends on the underlying preference structures (Dhingra and Morrow 2016). As can be seen from **Figure 4.15** for most countries in our sample group, we find that mark-ups and firm size are positively correlated. For Finland, the Netherlands, and Hungary we find, however, a negative correlation between mark-ups and firm size (except for the very large firms in Hungary and the Netherlands). Similarly, as shown by **Figure 4.16**, across all countries, firms' mark-ups and labour productivity are positively correlated.

Figure 4.15: Median manufacturing sector mark-ups across firm size deciles

Figure 4.16: Median manufacturing sector mark-ups across firm productivity deciles

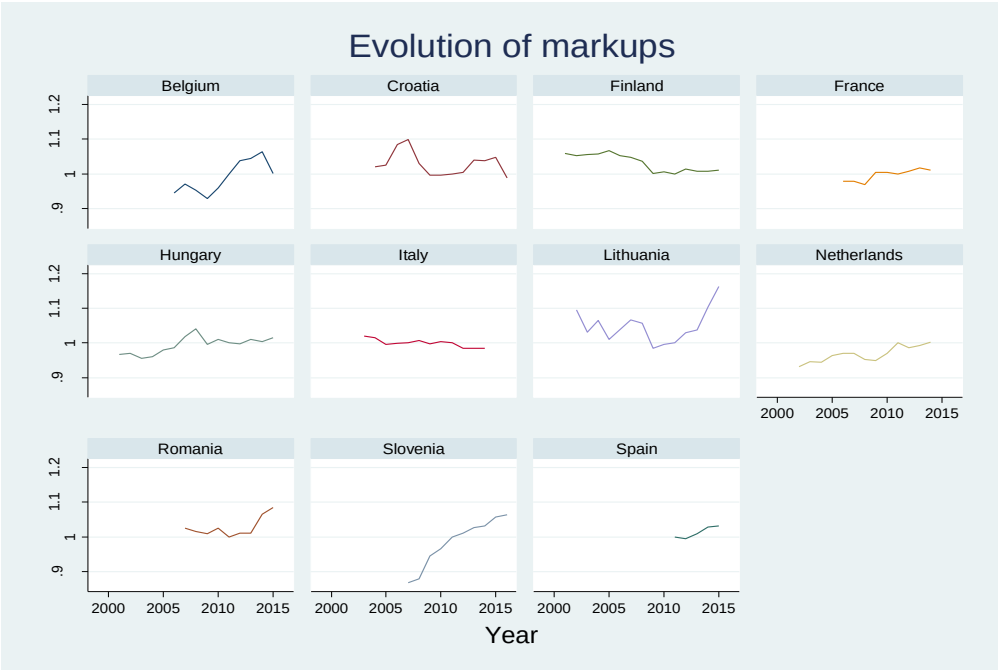


Source: 6th vintage of CompNet data, full sample.
Notes: Median mark-ups are normalized by country averages.

Next, we investigate how mark-ups evolved over time. A recent study by De Loecker and Eeckhout (2017) reports an increase in mark-ups within the last 50 years in the U.S. Such a "rise of market power" has important implications for macroeconomic dynamics and could explain several secular trends observed in many developed economies, including the fall of labour shares and the slowdown

of aggregate productivity. **Figure 4.17** reports the evolution of median mark-ups for the manufacturing sector across different European countries using the CompNet dataset. We absorb potential level differences due to country-specific differences in variable definitions by normalizing median mark-ups to 1 in the year 2011. The Figure shows that in most countries we indeed observe an increase in mark-ups, which generally is in line with the evidence reported in De Loecker and Eeckhout (2017). However, in Finland's, Croatia's, and Italy's manufacturing sector mark-ups tend to fall in the observation period. Although for convenience, we restrict ourselves here to the manufacturing sector, CompNet also estimates mark-ups and production functions for other macro-sectors.[29] In principle, this allows exploring how structural changes (i.e. the reallocation of economic activity across macro-sectors) affect aggregate market power. Generally speaking, linking mark-up dynamics to country-specific characteristics (i.e. competition policies or the degree of firm concentration) and showing what the evolution of mark-ups implies for country-specific macroeconomic outcomes could be a valuable topic for future research based on the CompNet dataset.

Figure 4.17: Time evolution of manufacturing sector mark-ups across countries, 2011=1



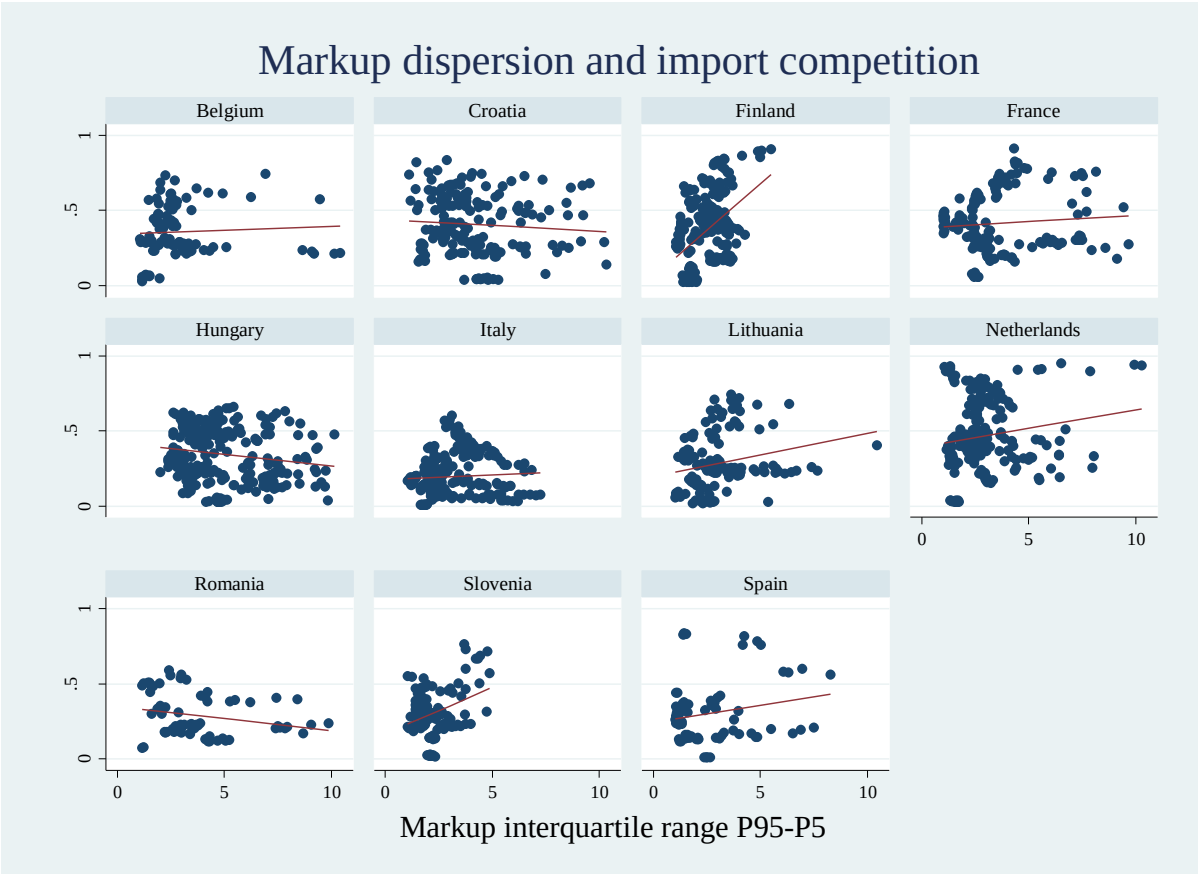
Source: Own calculations based on the 6th vintage of CompNet full sample (manufacturing sector).
Notes: Values for the year 2011 are normalized to 1.

[29] The time evolution of mark-ups across all macro-sectors is very similar.

4.5.2 Putting the indicators to work: Imports and mark-up dispersion

Mark-ups dispersion might indicate market power that is typically associated with a misallocation of resources (Epifani and Gancia 2011; Dhingra and Morrow 2016). By exposing producers to more competition, international trade may reduce market dispersion, improve allocative efficiency and increase welfare (Edmond et al. 2015). On the other hand, an incomplete pass-through of cost savings due to cheaper intermediate input imports may allow globally sourcing firms to increase their mark-ups compared to non-globally sourcing firms. To date, there is only limited evidence on how trade and mark-up dispersion are empirically linked to each other (e.g. Lu and Yu 2015).

Figure 4.18: Within-sector interquartile range of mark-ups and sector import shares



Sources: 6th vintage of CompNet full sample (manufacturing sector) and United Nations Comtrade Database
Notes: Based on the full sample. We dropped outliers with respect the interquartile ranges.

To motivate this topic a bit more, **Figure 4.18** plots import shares defined as imports over domestic production plus imports and calculated from the United Nations COMTRADE Database against two-digit sector specific interquartile ranges of mark-ups. We see that in most cases import shares and mark-up dispersion are positively correlated. This relationship is especially strong for Finland and Slovenia. A negative correlation between mark-up dispersion and import shares can only be found in three out of eleven cases (Hungary, Croatia and Romania) and even there, this relationship is comparably weak. Does this imply that import competition tends to widen the mark-up dispersion? We do not know. At least we find that across most countries in our sample, sectors with higher import shares display a wider mark-up distribution. We believe that the more in-depth and causal investigation of that or similar topics constitute relevant fields for future research based on our dataset.

4.6 Wages

Wage formation, existing feedback loops between inflation expectations and wage growth, and firms' pass-through of increasing labour costs to prices are fundamental for the inflation outlook. Hence, the inclusion of this variable in the CompNet dataset is fundamental for some stakeholders like Central Banks.

4.6.1 Definition, validation and granularity

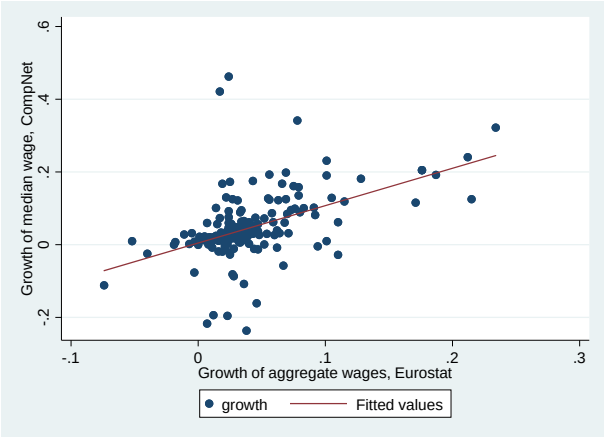
CompNet collects data on firm‘s average labour cost per employee. This indicator is computed as gross wages and salaries plus overtime payments and employers' social security contributions, divided by the total number of employees in the firm (average of the year, normally in FTE).

To get a sense of the plausibility of the data, **Figure 4.19** shows the median wage growth rate in each country-year in CompNet versus the annual wage growth rate provided for the whole economy by Eurostat. Despite the different target populations of both sources, the coefficient of correlation is reasonable, around 0.5.

Besides average wage growth in a given country, CompNet collects information on the distribution of wages, by sector, as well as wage levels in different splits of firms. This granularity helps understanding aggregate trends. One example is provided in **Figure 4.20** where average nominal

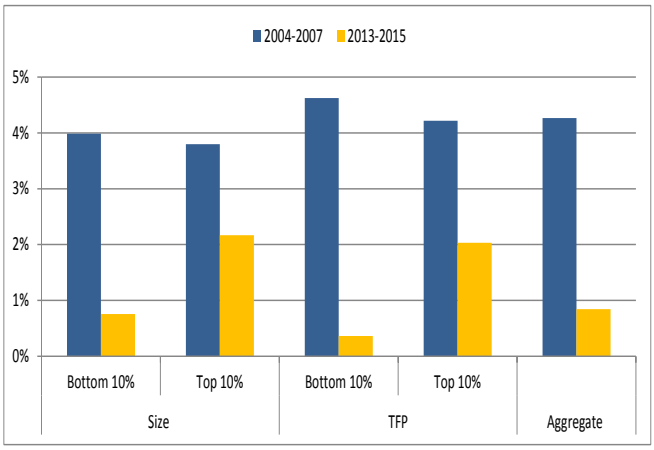
wage growth of manufacturing firms in western economies[30] is shown for two periods: the pre-crisis period, defined as 2004-2007, and the post-crisis one, 2013-2015. Looking first at the wage growth considering all firms in the sector (far-right bars), the post-crisis slow-down in wage growth becomes evident: average annual nominal wage growth in the post-crisis period is about half the one in the pre-crisis years, confirming the aggregate numbers provided by other institutions. However, this decline has not been the same across all firms operating in manufacturing, but rather concentrated in small and low productive firms.

Figure 4.19: Wage growth in Eurostat and CompNet



Growth of median wage Sources: Eurostat and 6th vintage of CompNet, 20E sample.
Notes: countries included are BE, HR, CZ, DK, FI, FR, DE, HU, IT, LT, NL, PL, PT, RO, SK, SI, ES and SE over the period 2000-2015.

Figure 4.20: Average wage growth before and after the crisis of different firms in Western Europe, manufacturing sector



Sources: Eurostat and 6th vintage of CompNet, 20E sample.
Notes: countries included are BE, DE, DK, FI, PT, SE, FR, NL and IT. The pre-crisis period is 2004-2007 while the post-crisis is defined as 2013-2015.

4.6.2 Putting the indicators to work: Subdued wage growth in the post-crisis period

Nominal wage growth in Western European countries has been relatively subdued in recent years, despite increasingly tight labour markets. The factors behind this common development are not entirely clear although policy-makers consider that a number of concurrent factors have been at play.

[30] Wage developments in Western and Eastern EU countries have been very different over the recent period. Hence this sub-section will concentrate in Western EU countries: Belgium, Germany, Denmark, Finland, Portugal, Sweden, France, Netherlands, and Italy.

Amongst them: (i) weak productivity growth; (ii) weak inflation developments affecting real wages; (iii) compositional effects related to the return to the labour force of workers with lower than average wages; (iv) underemployment engendering "hidden" labour market slack; (v) welfare and labour market reforms increasing not only flexibility, but also labour supply; and (vi) one-off, idiosyncratic factors (e.g. Brexit in the UK).

Indeed, notwithstanding some rebound after 2013, labour productivity remains weaker than before the crisis across western economies. The direction of causality between productivity and wage growth, however, probably goes both ways: low productivity growth has capped wage growth, but low wage growth relative to productivity may have lifted labour demand and, therefore, increased employment. However, while weak productivity growth has contributed to the slowdown in wage growth, it cannot fully explain it. This is clearly seen in **Figure 4.21**, showing the correlation between median *nominal* wage and productivity growth in each 2-digit industry and year of western countries included in CompNet, distinguishing between the pre-crisis (2004-2007) and the post-crisis period (2013-2015).[31] Note that the figure is a "bin-scatter", that is, it groups observations in bins, both in the x- and y-axis, to show a cleaner picture. Moreover, country-specific fixed effects are controlled for to take away country-invariant factors which could affect the correlation.[32] Given any productivity growth rate, nominal wage growth in the post-crisis is significantly lower than in the pre-crisis period. In this line, a study conducted by the Bank of England for the UK shows that the fall in productivity growth accounts for less than half of the slowdown in UK pay growth (see Bank of England, 2017).

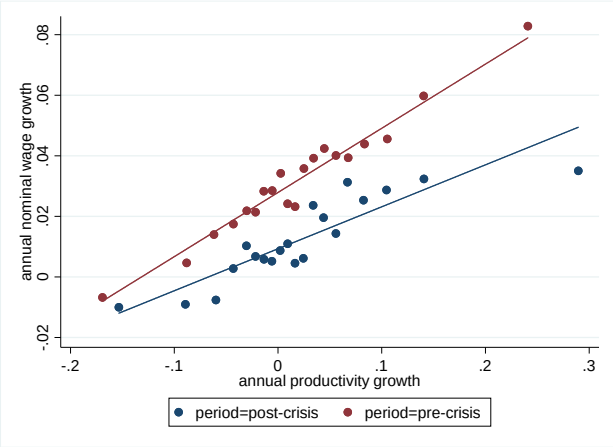
One possible reason behind this fact, also listed above, is that inflation has been subdued in the post-crisis period, standing on average at 0.7% relative to 2.2% in the pre-crisis period.[33]However, even when controlling for weaker inflation we find that *real* wage growth in the post-crisis period is lower for any given productivity growth rate (**Figure 4.22**).

[31] The countries included are Belgium, Denmark, Finland, Germany, Portugal, Sweden, France and Netherlands.

[32] That is, the figure shows the residual after regressing wage and productivity growth on a set of country dummies.

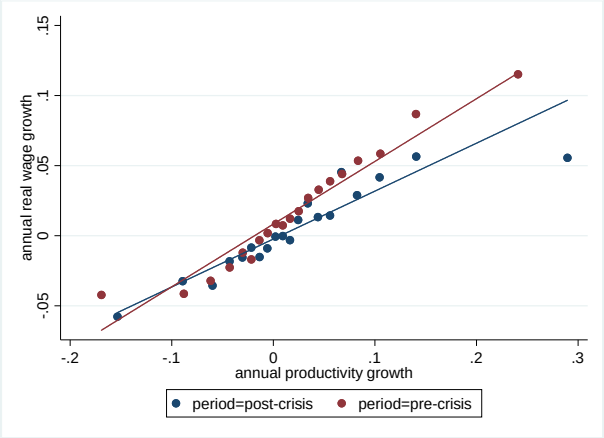
[33] Data from Eurostat. The numbers correspond to EA average annual HICP growth in the pre-crisis and post-crisis period, between 2004 and 2007 and 2013 and 2015 respectively.

Figure 4.21: Nominal wage and productivity annual growth in 2-digit industries, western economies



Sources: 6th vintage of CompNet, sample of firms with at least 20 employees
Notes: countries included are BE, DK, FI, FR, DE, NL, PT and SE over the period 2000-2015.

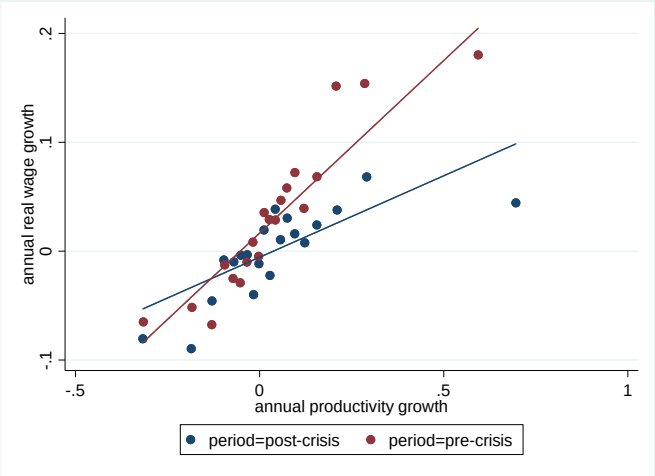
Figure 4.22: Real wage and productivity annual growth in 2-digit industries, western economies



Sources: 6th vintage of CompNet, sample of firms with at least 20 employees
Notes: countries included are BE, DK, FI, FR, DE, NL, PT and SE over the period 2000-2015.

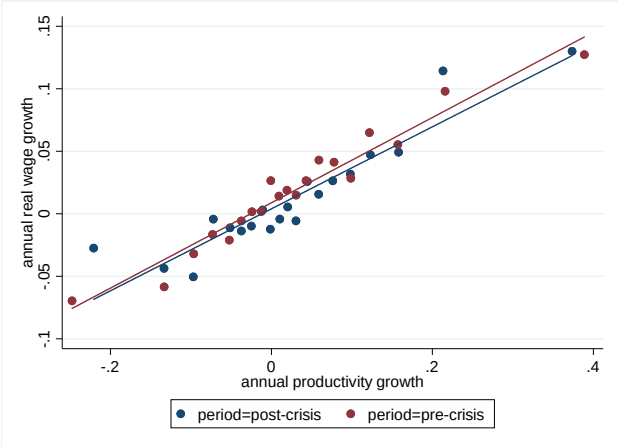
One obvious question is whether this disconnection between wages and productivity growth in the post-crisis period has taken place across all firm types. To explore this issue, we replicate **Figure 4.22** but considering only firms at the bottom and top 10% of the TFP distribution in any 2-digit sector. **Figure 4.23** and **Figure 4.24** show that the increasing disconnection between wages and productivity is concentrated in the lower tail of the productivity distribution; the correlation between real wages and productivity growth amongst the most productive firms in each sector has not varied between the pre and post-crisis period.

Figure 4.23: Real wage and productivity growth in bottom 10% productive firms within each 2-digit industry, western economies



Sources: 6th vintage of CompNet, sample of firms with at least 20 employees
Notes: countries included are BE, DK, DE, FI, FR, IT, NL, PT and SE over the period 2000-2015.

Figure 4.24: Real wage and productivity growth in top 10% productive firms within each 2-digit industry, western economies



Sources: 6th vintage of CompNet, sample of firms with at least 20 employees
Notes: countries included are BE, DK, DE, FI, FR, IT, NL, PT and SE over the period 2000-2015.

This interesting finding, confirmed after controlling for sector, year and country fixed effects, could be driven by a composition effect, that is, by a shift towards lower payed jobs in some countries during the crisis and post-crisis, particularly in low productive firms.

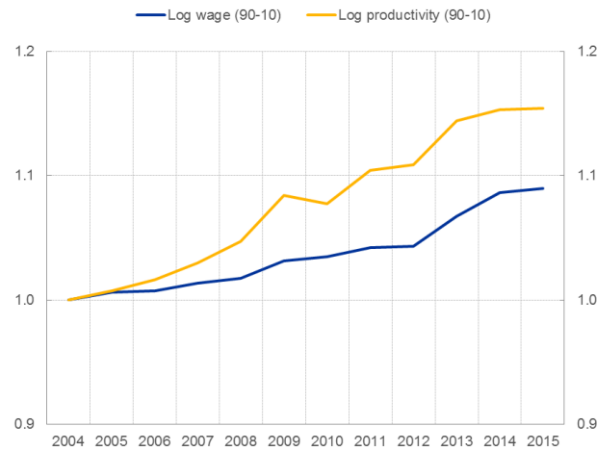
To explore further this possibility, we plot the log difference between the top and bottom decile of the wage distribution in each 2-digit sector, aggregated to the country level using sector employment weights, as well as the log difference between the top and bottom productivity levels. We focus for this exercise on two countries, Belgium and Spain, given that aggregating across countries would mask interesting heterogeneity.[34] In both countries (**Figure 4.25**) the within-sector productivity *and* wage dispersion are increasing over time, with a cumulative growth of about 15% for labour productivity and 10% for wages. These are very similar numbers to those reported in OECD (2017). Note that during the sovereign debt crisis, productivity dispersion fell in Spain due probably to the cleansing effect of

[34] These are two countries similar in terms of good coverage of micro-firms but distinct in terms of the impact of the crisis.

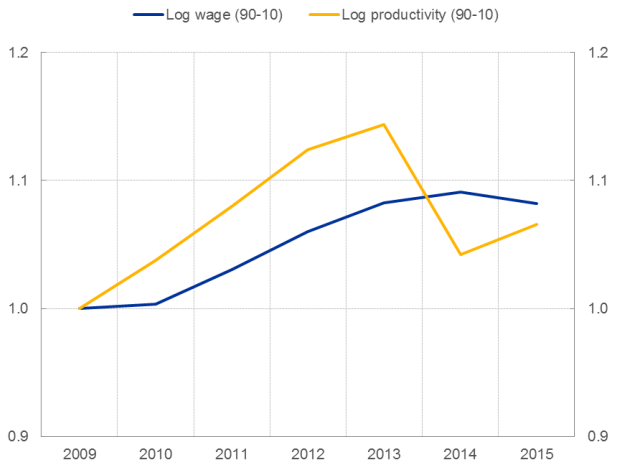
the crisis, whereby low productive firms exited the market. **Figure 4.26** splits the dynamics in wage dispersion in two halves: dispersion in the lower part of the wage distribution (log difference between the 50th and the 10th decile of wages, in a given 2-digit industry) and dispersion in the upper part of the wage distribution (log difference between the 90th and the 50th decile of wages). While in Belgium the dispersion in both halves of the wage distribution increased similarly over time, with the exception of the last year, in Spain dispersion increased mostly at the bottom of the distribution. This result deserves further research but lends some support to the possibility that muted wage growth in the post-crisis period is driven by a composition effect.

Figure 4.25: Wage and productivity dispersion

Belgium



Spain

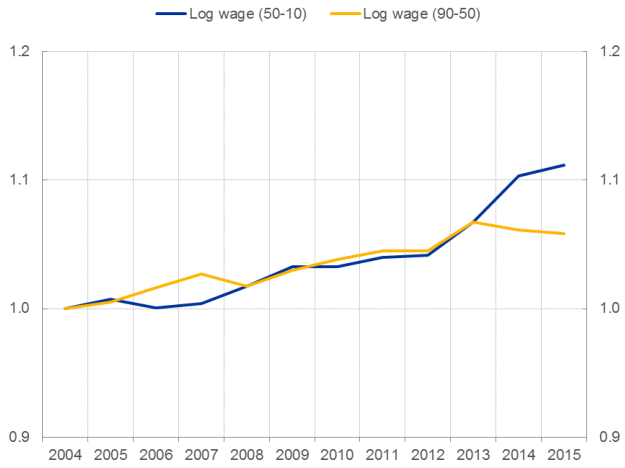


Sources: 6th vintage of CompNet, sample of firms with at least 1 employee

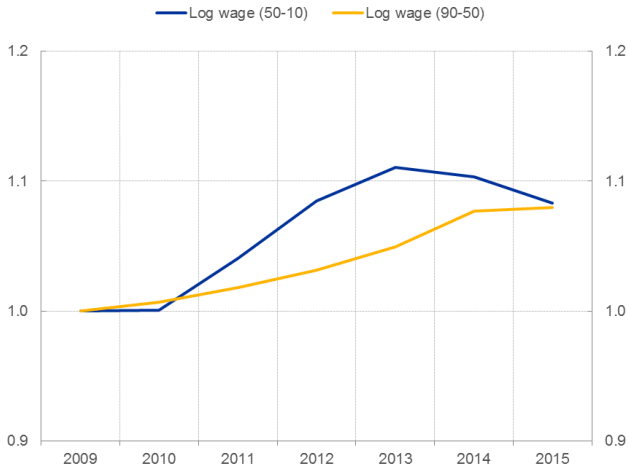
Notes: Computed within each 2-digit sector and then aggregated to the country level using sector employment weights.

Figure 4.26: Dispersion below and above median wages

Belgium



Spain



Sources: 6th vintage of CompNet, sample of firms with at least 1 employee

Notes: Computed within each 2-digit sector and then aggregated to the country level using sector employment weights.

4.7 Job flows and high-growth firms

There is broad consensus about the fact that relatively few firms, be them small or large, known as high-growth firms, are responsible for the majority of jobs created. The early works of Birch and Medoff (1994) for the US estimated that during the 1988-1992 period, 4% of active firms created about 60% of the jobs. These firms, however, also destroyed disproportionately more jobs. Hence, access to data on gross job flows of various types of firms, including high-growth firms, can help policy makers to better understand existing bottlenecks in the labour market.

4.7.1 Definition, validation and granularity

For the first time, CompNet computes job creation and destruction rates within each level of aggregation. We follow Davis and Haltiwanger (1996) and compute job flows as the weighted employment growth rate of firms in a given level of aggregation with positive (job creation rates) or negative (job destruction rates) growth. Hence:

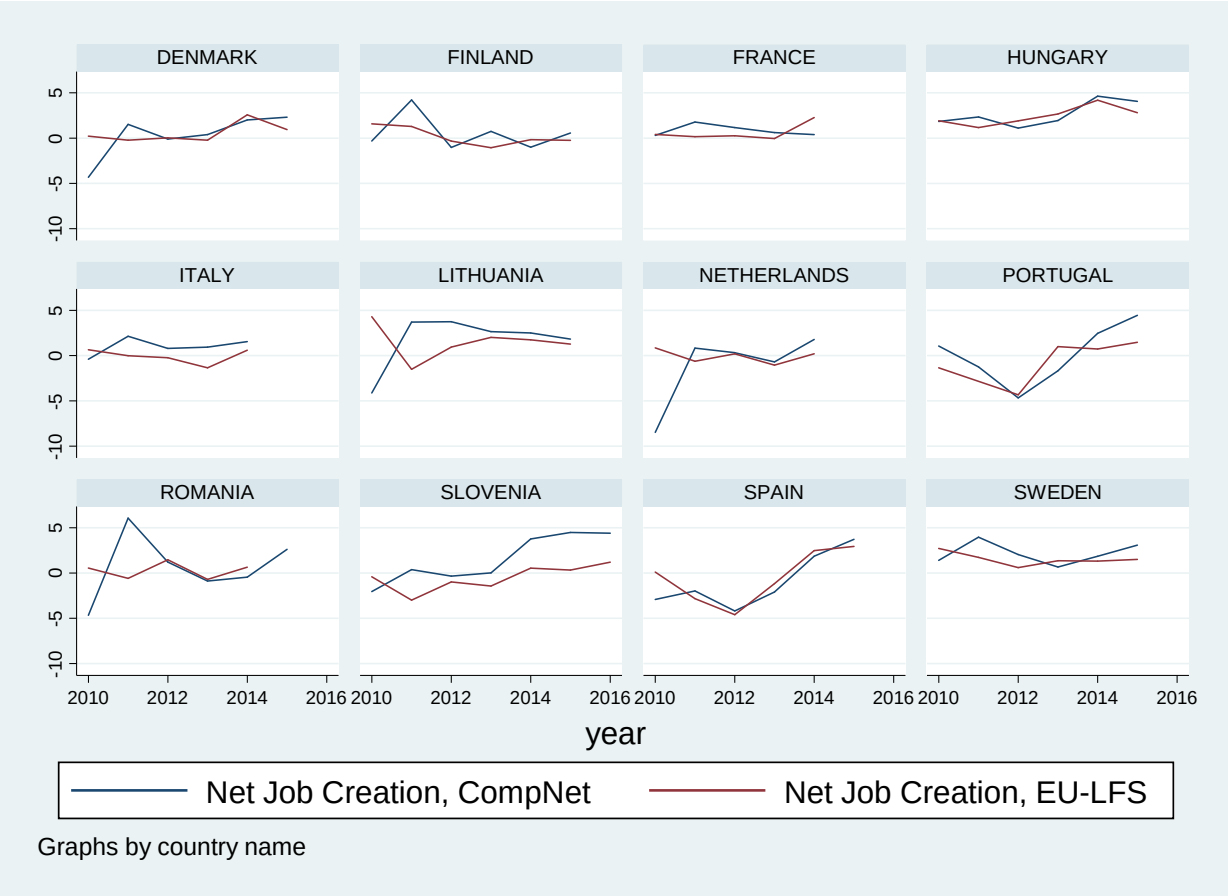
$$\boldsymbol{Job\ Creation\ (Destruction)\ Rate_{st}} = \sum_{i \in s} \boldsymbol{g_{it}} \times \boldsymbol{w_{it}}\ \boldsymbol{if\ g_{it}} > \mathbf{0}\ (\boldsymbol{g_{it}} \leq \mathbf{0}) \tag{2}$$

with $$\boldsymbol{g_{it}} = \frac{(\boldsymbol{employment_{it}} - \boldsymbol{employment_{it-1}})}{((\boldsymbol{employment_{it}} + \boldsymbol{employment_{it-1}})/\mathbf{2})} \tag{3}$$

and $$\boldsymbol{w_{it}} = \frac{\boldsymbol{employment_{it}}}{\sum_{i \in s} \boldsymbol{employment_{it}}} \tag{4}$$

Figure 4.27 shows for each country with information on micro-firms (i.e. with the full sample) net job creation -the difference between job creation and destruction rates-, computed by CompNet and by the Labour Force Survey Longitudinal dataset (Eurostat). The Labour Force Survey computes labour market transitions defined as the number of persons transiting between different labour market statuses. We define job creation rates as the share in total employment of persons transiting from unemployment or inactivity to employment, and job destruction rates as the share of employed transiting from employment to inactivity or unemployment. Transitions from employment to employment are not taken into account since they also include workers who stay in their current job, not only workers who switch to a different job. Despite the (admittedly large) differences in definitions, net job creation rates from CompNet are of the same order of magnitude, and show similar dynamics as the labour transition flows provided by Eurostat.

Figure 4.27: Net Job Creation Rates in CompNet and Eurostat



Source: Own calculations from the 6th vintage of CompNet data, full sample and Labour market transitions from the EU-LFS

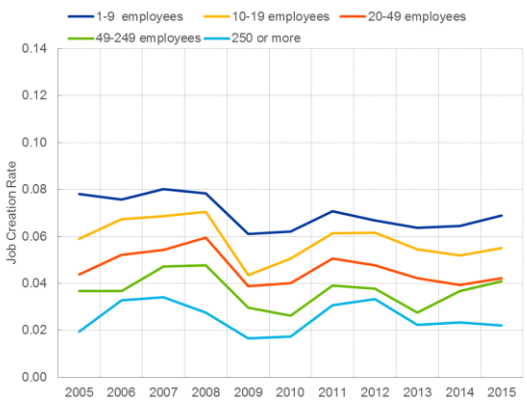
Aggregate job flows like the ones shown above hide, however, large heterogeneity across sectors and, within each sector, across types of firms. This is clearly shown below where job creation (**Figure 4.28**) and destruction (**Figure 4.29**) rates are depicted for firms in different size classes within the manufacturing sector in two countries with very different developments: Belgium and Italy. Looking at the levels, rather than trends, in both countries job flows decrease with size; that is, they are largest in the smallest size class. This is particularly the case for job destruction rates resulting in negative net job creation of micro-firms in both countries, which is in line with the literature.[35] Looking at the trends, it is noteworthy that job creation rates, and to a lesser extent job destruction rates, decrease in Italy

[35] See Haltiwanger et al. (2013) for a detailed study of job creation and destruction by size and age for the US

over the whole period of analysis. In Belgium, on the other hand, only job creation rates of micro-firms show this declining trend whereas, in the rest of the size classes, job flows are roughly similar at the end and at the beginning of the period. Both the larger dynamism of small firms and, in particular, the declining trends in job flows have been found in other countries like the US (see Decker et al. 2016) and it has largely been attributed to the decrease in dynamism of the very young firms. Although we do not have information on the age of firms, we also find that the declining trend is most pronounced amongst the smallest size class, where the young firms normally belong. Finally, the severity of the sovereign debt crisis in Italy can be easily traced by the large increase in job destruction rates, particularly for very small firms.

Figure 4.28: Job creation rates by size class, manufacturing sector.

Belgium



Italy

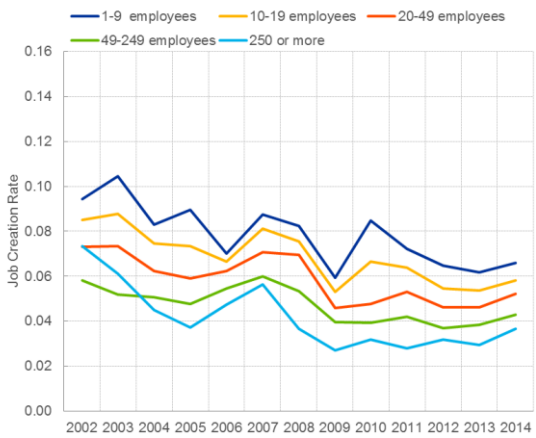
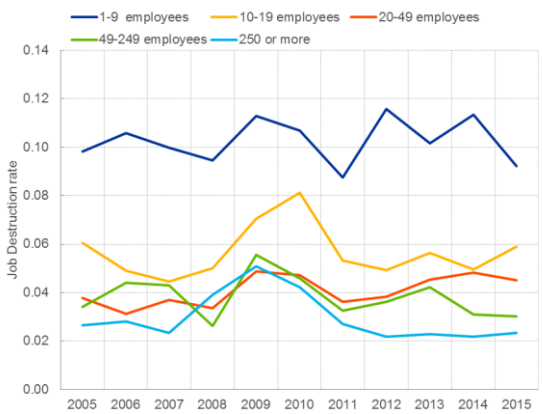
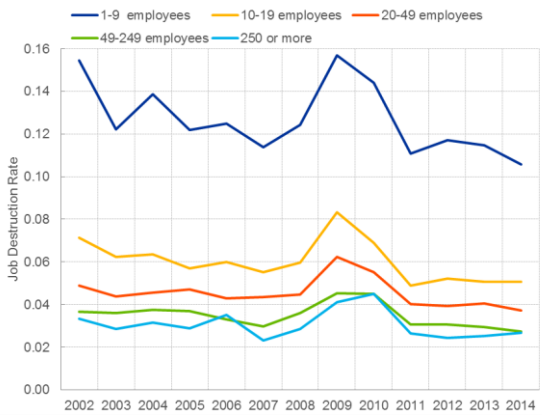


Figure 4.29: Job destruction rates by size class, manufacturing sector.

Belgium



Italy

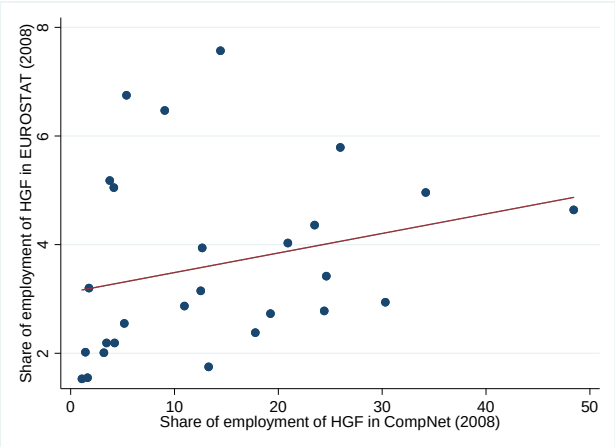


Sources: Own calculations on the 6th vintage of CompNet data, full sample.

Turning to *high-growth firms*, these are firms experiencing extraordinary high employment growth for a number of consecutive years. Both their share and characteristics can be approximated using the transition matrices available in CompNet. Transition matrices classify firms in a given sector according to their transitions from one quintile of the size distribution to another in 3-year windows. We flag as high-growth firms those moving from the first or second quintile to the fifth quintile, and between the first and the fourth quintile of the size distribution, within their sector of activity. Cumulative employment growth of these firms is, on average, around 70% which corresponds to around 20% of employment growth per year, matching the standard classification of high-growth firms in the literature (see for example Schreyer 2000).

In order to get a sense of the soundness of our definition, we compare in **Figure 4.30** the total share of employment in high-growth firms in CompNet with that provided by the structural business statistics of Eurostat. Eurostat provides employment in firms with three consecutive years of employment growth above 20%.The comparison is done for 2008 and includes only overlapping countries, which in this case are only Hungary, Sweden and Italy. The coefficient of correlation is around 0.25. [36]

Figure 4.30: Employment in high-growth firms in CompNet and Eurostat



Sources: Own calculations on 6th vintage of CompNet full sample and Eurostat
Notes: HGF in CompNet calculated as the share of firms transiting from the first and second quintile to the fifth quantile and from the first to the fourth quintile of the size distribution in 3 years in a given sector out of the total number of firms in the sector. In Eurostat, HGF defined as firms with 3 consecutive years of employment growth above 20%.Countries included are the only overlapping countries between both databases: SE, HU and IT.

[36] The data from Eurostat is sourced from the sub-section "Business demography" of the Structural Business Statistics and is very incomplete. Although in theory it is available from 2007 to 2013, only in 2008 there are data for at least 3 countries also included in CompNet.

4.7.2 Putting the indicators to work: High-growth firms and job creation

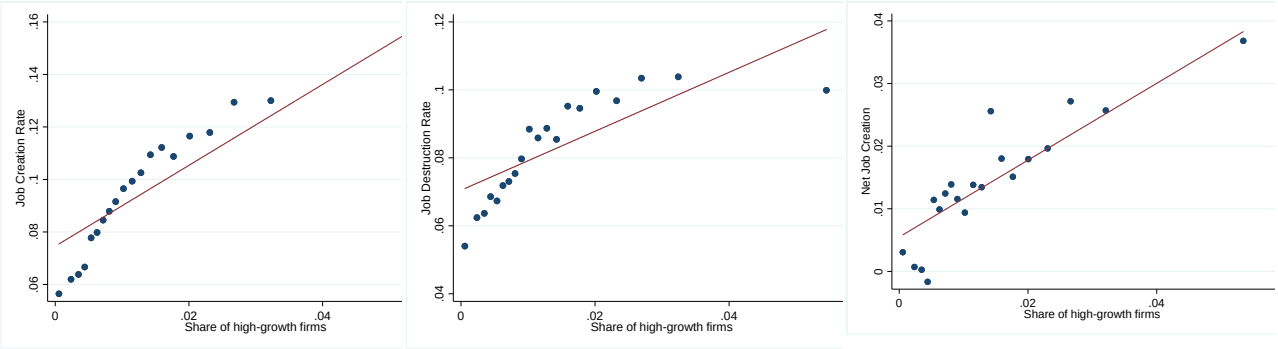
In the mid-1970s David Birch created for the first time a longitudinal database able to follow firms over time using the Dun & Bradstreet records of firms attempting to establish credit with other firms or seeking credit information. Using this new database, Birch found that around 80% of net new jobs were created by firms with 100 employees or less (Birch 1979: The Job Generation Process. The main results are also in Birch 1981). Birch's claim that 8 out of 10 net new jobs were created by small firms prompted the interest of policy-makers who, in the high unemployment days of the early 80s, were interested in policies able to reduce unemployment. In the academic world, Birch' findings where revolutionary at the time. They implied that inter-class movements (small firms growing until they are classified as large firms) were a major factor in determining overall employment growth. Birch also discovered that the rates of job losses across regions were pretty similar. Differences in net employment growth were due to differences in the job gain rates. In other words, rapidly growing areas were replacing lost jobs at 2 or 3 times the rate of the declining ones. More importantly, about 80% of the replacement jobs were created by establishments that were four years or younger:

“not all small businesses are job creators. The job creators are the relatively few younger ones that start-up and expand rapidly in their youth, outgrowing the ``small'' designation in the process” (Birch 1981, page 8).

Using CompNet data we can confirm that in sectors with a high share of high-growth firms, lost jobs are indeed replaced by a disproportional number of new jobs. **Figure 4.31** shows in its Panel A the correlation between the share of high-growth firms in a given 2-digit industry and the sector-specific job creation rate; in the middle panel, the correlation between the share of high-growth firms and job destruction rates; and in Panel C the correlation between net job creation and the share of high-growth firms in the sector. Note that the Figures are bin-scatter, that is, they group observations in different bins in the x- and y-axis to get a cleaner picture. Moreover, country-specific effects are controlled for to account for the fact that we pool together countries in different stages of development.

Figure 4.31: Job flows and high-growth firms

A. Job Creation Rates and HGF B. Job Destruction Rates and HGF C. Net Job Creation and HGF



Sources: 6th vintage of CompNet data, full sample.
Notes: Full sample. The countries included are BE, ES, FI, HU, IT, LT, PT, RO and SE.

It becomes clear than in line with the literature high-growth firms contribute greatly to both job creation and destruction. However, as Panel C of the figure above shows, the rate of replacement of lost jobs is larger than one so this type of firms contributes positively to net employment creation. Perhaps one of the most important challenges currently in the EU is related to the creation of jobs. Given the reported importance of high-growth firms for net job creation, cross-country and cross-sector information on these firms can help policy makers designing the appropriate framework to nurture their creation, survival and growth.

5. Concluding remarks

Policy institutions are increasingly demanding more granular data to complement their analysis of aggregate indicators and understand underlying trends and their micro foundation. These data should be as comparable as possible across countries, and cover a sufficient number of EU countries and years.

The first best in this regard would be that an international body, such as Eurostat, responds to this policy need and compiles, checks, harmonises and publishes firm-level data or micro-aggregated data. However, this scenario is not likely to materialise in the short-to-medium term. Hence, other avenues are being undertaken.

One possible avenue is to use commercial databases. This avenue has the advantage of providing large amount of firm-level data for several countries, but - as examined in the accompanying cross-country comparability report - also clear drawbacks when it comes to the analysis of competitiveness, productivity and exports. The main loopholes are the lack of appropriate coverage and the presence of representativeness biases in the country samples of firms reporting basic variables such as employment or turnover, which prevents a reliable analysis of productivity.

Given these constraints, CompNet's dataset remains in our view the "first best" for benchmark analysis of competitiveness. CompNet consists of a network of researchers and policy makers with access to firm-level data in a number of countries. Micro-aggregated indicators are collected, which preserves confidentiality but enables benchmarking analysis. In this set-up, the data users keep also control over the production of the data, which has clear net benefits, given that researchers are in the best position to decide what indicators are needed.

Over the recent period, we have undertaken a major revision of the procedures and indicators in order to improve cross-country comparability and include new, relevant information. There is still room for improvement but we hope that the current vintage of data will be useful to inform and complement the policy analysis undertaken by European institutions.

6. Annexes

6.1 Annex 1: Raw variable definitions and sector details

Table 6.1: Raw variable definition (most common across countries)

Raw variable	Definition
Capital (Tangible fixed assets)	Tangible fixed assets
Raw materials (intermediate inputs)	Use of materials + energy + services
Labour cost	Compensation of employees (wages and salaries plus social contributions)
Value added	Turnover - Raw materials
Number of employees, headcount	number of workers (please indicate persons employed or persons engaged (includes proprietors), and FTE or headcount)
Turnover	Total sales net of VAT at basic prices
Unadjusted export value	Total exports by the firm, not adjusted for reporting threshold
Threshold adjusted export value	Export value + reporting thresholds
Import value	Total imports by the firm, as reported
Total assets	Total assets
Cash and cash equivalents	Cash and balances at banks
Cash flow (from profit/loss statement)	Net income + depreciation+ extraordinary income
Profit/loss	EBIT
Interest paid (or financial charges)	Interest on financial debts + other financial expenses
Long term debt	Loans due in more than 1 year
Short-term debt	Loans due within 1 year
Total inventories	Inventories and consumable biological assets
Depreciation	Depreciation on tangible assets
Trade credit (accounts payable)	Trade credit or Accounts payable (Liabilities related to purchased goods and services)
Trade debt (accounts receivable)	Trade debt or Accounts receivable
Trade credit (accounts payable)	Short term debt + trade credit + other current liabilities
Non-current liabilities	Liabilities - (Short term debt + trade credit + other current liabilities)
Shareholder funds (equity)	Equity
Profits and losses before taxes	Earnings before taxes (EBT)

Other current assets	Current assets – Trade debtors – Total inventories
Other non-current liabilities	Provisions
Other fixed assets	Total fixed assets - tangible fixed assets - intangible fixed assets - financial assets
Intangible fixed assets	Total intangible fixed assets
Current assets	Cash and other assets expected to turn in cash within a year
Other current liabilities	Current liabilities - Short term debt - trade credit
Total fixed assets	Total Fixed Assets
Dividends	Dividends
Firm’s birth year	Actual year of birth

Table 6.2: Macro-Sectors and Two-Digit NACE Rev. 2 Sectors Covered by CompNet

NACE Rev. 2 Section	Macro-sector in CompNet	Description	Sector in CompNet	Description
C	1	Manufacturing	10	Manufacture of food products
			11	Manufacture of beverages
			12	Manufacture of tobacco products
			13	Manufacture of textiles
			14	Manufacture of wearing apparel
			15	Manufacture of leather and related products
			16	Manufacture of wood and of products of wood and cork, except furniture
			17	Manufacture of paper and paper products
			18	Printing and reproduction of recorded media
			19	Manufacture of coke and refined petroleum products
			20	Manufacture of chemicals and chemical products
			21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
			22	Manufacture of rubber and plastic products
			23	Manufacture of other non-metallic mineral products
			24	Manufacture of basic metals
			25	Manufacture of fabricated metal products, except machinery and equipment
			26	Manufacture of computer, electronic and optical products
			27	Manufacture of electrical equipment
			28	Manufacture of machinery and equipment n
			29	Manufacture of motor vehicles, trailers and semitrailers
			30	Manufacture of other transport equipment
			31	Manufacture of furniture
			32	Other manufacturing
			33	Repair and installation of machinery and equipment

NACE Rev. 2 Section	Macro-sector in CompNet	Description	Sector in CompNet	Description
F	2	Construction	41	Construction of buildings
			42	Civil engineering
			43	Specialised construction activities
G	3	Wholesale and retail trade; repair of motor vehicles and motorcycles	45	Wholesale and retail trade and repair of motor vehicles and motorcycles
			46	Wholesale trade, except of motor vehicles and motorcycles
			47	Retail trade, except of motor vehicles and motorcycles
H	4	Transportation and storage	49	Land transport and transport via pipelines
			50	Water transport
			51	Air transport
			52	Warehousing and support activities for transportation
			53	Postal and courier activities
I	5	Accommodation and food service activities	55	Accommodation
			56	Food and beverage service activities
J	6	Information and communication	58	Publishing activities
			59	Motion picture, video and television program production, sound recording and music publishing
			60	Programming and broadcasting activities
			61	Telecommunications
			62	Computer programming, consultancy and related activities
			63	Information service activities
L	7	Real Estate activities	68	Real estate activities
M	8	Professional scientific and technical activities	69	Legal and accounting activities
			70	Activities of head offices; management consultancy activities
			71	Architectural and engineering activities; technical testing and analysis
			72	Scientific research and development
			73	Advertising and market research
			74	Other professional, scientific and technical activities
			75	Veterinary activities
N	9	Administrative and support service activities	77	Rental and leasing activities
			78	Employment activities
			79	Travel agency, tour operator and other reservation service and related activities
			80	Security and investigation activities
			81	Services to buildings and landscape activities
			82	Office administrative, office support and other business support activities

6.2 Annex 2: Data cleaning, weighting and deflating

6.2.1 Data cleaning

The CompNet code applies two routines that affect the raw variables before being fed into the actual indicator computation. The first routine loops through the main raw variables eliminating impossible values. The second routine focusses on assessing implausible values along a few criteria, and deletes them if the criteria do not hold. We discuss each routine in more depth.

Impossible values

The ratio behind the impossible values routine is focused on preserving as much useful information as possible. Hence, small violations in certain accounting identities are not judged and treated as being data inconsistencies. We could test whether the difference between turnover and intermediate inputs, which should be equal to value added, holds in our datasets. However we observe small violations of this identity. This can be explained by the plurality of data providers and heterogeneous underlying data sources. Instead of applying invasive accounting routines, we rely on our outlier treatment to filter out miss-measured values.

Therefore, the first routine investigates the raw variables provided by the national counterparts on the basis of accounting identities. The content of the following variables is deleted if they show negative values: turnover, capital, labour, totals assets, cash holdings, long term debt holdings, trade credit, trade debt, interest payments, other fixed assets, current assets, dividend payments and depreciations. The single observation is therefore treated as missing by the code. In a similar fashion the interest payments and debts are checked. If interest payments are smaller than the debt value both observations are turned to missing values.

Outlier dropping

The second routine focusses on measurement errors and flags the respective variables as outliers. Previous vintages of data collection indicated a trade-off. The outlier procedure must not affect or distort aggregate results by limiting the number of observations used for the indicator calculation. But still, it must be strict enough to correctly filter out values that can be identified as outliers. A factor that further complicates the creation of this code is that the routine is written without fine-tuning it to an individual data source.

Before the routine starts, the data is split into bins according to the two digit sector and year. Within these bins consequently, three checks are applied.

1. Is a value more than three standard deviations away from the median?
2. Is a value in the top or bottom 1 percentile?
3. Is the growth of a value with respect to the previous year in the top or bottom 1 percentile?

If all of these conditions are fulfilled, the value is set to missing. Literature labels this as a *lenient routine*. Given the quality of the data sources and the institutions behind them this lenient routine can be justified. The outlier procedure is assessed after each round of data collection and will possibly be strengthened in future vintages.

6.2.2 Data aggregation, weighting and deflation

The CompNet dataset aims to enable researchers to look beyond simple aggregations in the *firm population*. Thus, all indicators should be understood as attributes of the underlying firm population and not of a sample, unless explicitly stated otherwise. To achieve this, CompNet's output is weighted with inverse probability weights: using Eurostat, we gather the number of firms in a given size-class and NACE Rev. 2 macro-sectors. From this, we compute the probability of a firm to end up in this sector- size class bin. Firms for which we are unable to compute this weight are discarded. Moreover, where we want to express indicators in real terms, we use publicly available deflators from Eurostat.

This form of reweighting gives a correct picture of the underlying firm distribution if and only if there is no selection into reporting within these bins. Since we deal with administrative data, non-reporting should be a function of survey design, not of endogenous firms' decisions.

We construct weights separately for every indicator. This is because some of them require many variables and/or lagged values. For these specific ones, the probability to be in the sample might be indeed radically different than for more standard variables. Since descriptive statistics make up the bulk of the running time of the code, we use the inbuilt summarize command of Stata to speed up computations. While this does not allow inverse probability weighting, using the same weights as variance weights leads to the same output.

All regressions are computed using standard inverse probability weights. A potential problem of this approach is that the macro-sector-size class level might not be the level at which survey probabilities vary: If e.g. the sample is drawn to represent all regions of a country, our method will not be able to correct for this. However, adjusting to the individual survey schemes of the different countries would create additional work for the data providers and introduce another breaking point into the code. It would arguably also disrupt the homogeneity of data treatment, as each country would be weighted differently.

6.2.3 Purchasing Power Adjustment

To improve the cross-country comparability of the CompNet data, we must eliminate systematic price differences between countries that are not caused by differences in exchange rates or inflation and reflect differences in purchasing power. To this end, the concept of purchasing power parity (PPP) is applied to the CompNet data.

The PPP and exchange rate statistics we use are based on OECD annual data[37]. In the OECD database, PPP's are expressed in terms of national currency of a particular country per US dollar[38].

In the CompNet dataset all variables measured in monetary units are expressed in Euros, even for countries outside the Euro zone. Consequently, PPPs between CompNet countries should be in the Euro/Euro dimension. Therefore, the original OECD PPP values have been multiplied by the inverse of the exchange rate (in national currency per USD)[39]. This makes the resulting PPP values (only available at country level) dimensionless. According to the base year of the time varying deflators we used to purge price variation from monetary variables (turnover, value added, capital stock), we applied PPP values of 2005 for all years.

The actual PPP-adjustment procedure was of a post-collection nature, i.e. only the resulting descriptive statistics (and not the underlying original variables) have been adjusted[40]:

- For non-parametric variables the percentiles and moments have been simply multiplied by the country-specific PPP or its inverse 1/PPP, if necessary.
- For monetary variables based on parametric estimations (e.g. TFP measures), PPP or PPP^{-1} had to be multiplied with a weighting factor (if necessary). This has to be done because not all input factors used in a production function estimation are measured in monetary units (e.g. labour). However, this simple multiplication technique could only be applied to indicators derived from Cobb-Douglas type production functions. As the underlying production functions have only been estimated for the (one-digit) macro-sector and the two-digit sector (but not at the country, macro-sector size class and NUTS2 level), a weighted one-digit macro-sector factor according to the number of firms has been

[37] See https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm#indicator-chart.

[38] For more information regarding how OECD calculates PPP, see the manual, link: https://www.oecd-ilibrary.org/economics/eurostat-oecd-methodological-manual-on-purchasing-power-parities_9789264189232-en.

[39] Source: https://data.oecd.org/conversion/exchange-rates.htm#indicator-chart.

[40] See e.g. Gal (2013), appendix B3. As can be easily seen, this method can also be applied to the descriptive statistics of the underlying variable.

used for the country level and the macro-sector weighting factor for the macro-sector size class level. A correction at the NUTS2 level was not possible.

6.3 Annex 3: Confidentiality procedure

Although the literature has long recognized that firm-level data delivers crucial information about a wide range of phenomena, economic research based on these data has been hampered so far by issues of confidentiality and comparability. As a result, the CompNet data collection and indicator construction process has been designed in such a way that both issues are resolved. We describe the CompNet confidentiality procedure in two parts, one part focussing on the raw firm-level data and another part covering the eventual output of the code, the output indicators. Both parts contribute to the fact that the user of the final data will not be able to uniquely identify individual firms based on the aggregated data.

6.3.1 Raw variables

The conditions of dealing with firm-level information and the obligations surrounding confidentiality differ across countries and across member institutions. The CompNet secretariat and the individual data providers work together intensively in compiling the dataset, but the code is ran in a decentralized way in each of the respective institutions. This means that no individual firm-level data is made available to the secretariat of CompNet. In this way, each member institution can satisfy their individual confidentiality constraints.

6.3.2 Output Indicators

The second aspect of the confidentiality procedure is ensuring that the eventual output indicators leave no room for identifying individual firms. Also in this regard, each member institution can have individually-specified conditions to satisfy. The CompNet code includes a specific routine, which is run in the final stage of the computation that checks the eventual output cells. This routine includes thresholds for the minimum number of observations to guarantee that no individual firm can be identified and tests for statistical dominance. If a cell is based on a limited amount of underlying micro-observations, making the identification of individual firms possible, the cell is dropped. This information is not eliminated from the total distribution; it is only left out of the specific cell. The second test is the test for statistical dominance. It includes thresholds for the largest permissible size share a single observation takes on in a given cell.

These thresholds can be set a-priori by the data providers to satisfy their country or institution specific conditions. These are the parameters which can be chosen:

1. The minimal number of observations for the 1% and 99% percentiles can be adjusted.
2. The minimal number of observations for the 5% and 95% percentiles can be adjusted.
3. The parameter for statistical dominance can be adjusted. This is the largest permissible share an observation takes on in a cell.

7. References

Acharya, V., T. Eisert, C. Eufinger and C. W. Hirsch (2017): “Whatever it takes: The real effects of unconventional monetary policy”, *CEPR Discussion Paper*, (No. DP12005).

Aglio, D., M. Christophori, J.P. van de Kerke, F. Marder, M. Sartori, V. Schramm, D. Trofimenkov and A. Zona Mattioli. It has been commented by R. Bräuer, M. Mertens, F. di Mauro, P. López-Garcia, revised and approved by P. Haug (2018): “User Guide for the 6th Vintage of the CompNet Dataset”, *The Competitiveness Research Network* (available on line at www.comp-net.org).

Altomonte, C., E. Bartelsman, J.P. van de Kerke, P. López-Garcia, F. di Mauro, M. Melitz, M. Polder and S. Roux: "Assessing the reliability of the CompNet micro-aggregated dataset for policy analysis and research: Coverage, representativeness and cross-EU countries comparability", *The Competitiveness Research Network* (available on line at www.comp-net.org).

Andrews, D. and F. Petroulakis (2017): “Breaking the Shackles: Zombie Firms, Weak Banks and Depressed Restructuring in Europe”, (No. 1433) *Economic Department Working Papers*, OECD.

Arkolakis, C., A. Costinot, D. Donaldson and A. Rodríguez-Clare (2018): “The elusive pro-competitive effects of trade”, *The Review of Economic Studies (forthcoming)*.

Asker, J., A. Collard-Wexler and J. De Loecker (2014): "Dynamic Inputs and Resource (Mis)Allocation", *Journal of Political Economy,* Vol. 122, No. 5, pp. 1013-1063.

Autor, D., Dorn, D. Katz, L., Patterson, C. and Van Reenen, J. (2017): “The Fall of the Labor Share and the Rise of Superstar Firms”, NBER Working Paper No. 23396.

Bank of England (2013): *Inflation report*. Retrieved from: https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2013/august-2013.pdf?la=en&hash=FE0AAD6C68F5E20AB94ACD521CE25FDA0CB9ED0E

Bank of England (2017): *Inflation report.* Retrieved from: https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2017/february-2017.pdf

Bank of Korea (2013): *Financial Stability Report*. Retrieved from: http://centerforfinancialstability.com/fsr/kor_fsr_201304.pdf

Bartelsman, E. J., J. Haltiwanger and S. Scarpetta (2004): “Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries”, (No. 1374) *IZA Discussion Paper Series*.

Bartelsman, E., J. Haltiwanger and S. Scarpetta (2013): “Cross-Country Differences in Productivity: The Role of Allocation and Selection”, *American Economic Review*, Vol. 103, No. 1, pp. 305-334.

Bartelsman, E.J. and M. Doms (2000): “Understanding Productivity: Lessons from Longitudinal Microdata”, *Journal of Economic Literature*, Vol. 38, 2000, pp. 569-594.

Bernard, A.B., J. Eaton, J.B. Jensen, S. Kortum (2003): “Plants and Productivity in International Trade”, *American Economic Review,* Vol. 93, No. 4, pp. 1268-1290.

Berthou, A., E. Dhyne, M. Bugamelli, A.M. Cazacu, C.V. Demian, P. Harasztosi, T. Lalinsky, J. Meriküll, F. Oropallo, and A.C. Soares (2015): “Assessing European firms’ exports and productivity distributions: the CompNet trade module”, *ECB WP 1788*.

Birch, D. (1981). Who Creates the Jobs?, The Public Interest 65: 3-14.

Birch, D. L., and J. Medoff (1994): “Gazelles”, in L. C. Solmon and A. R. Levenson (ed.), *Labor markets, employment policy and job creation*, (Boulder: Westview Press) 159167.

Caves, R.E. (1998): “Industrial Organization and New Findings on the Turnover and Mobility of Firms”, *Journal of Economic Literature*, Vol. 36, pp.1947-1982.

Chiacchio, F., K. Gradeva and P. Lopez-Garcia (2018): “The post-crisis TFP growth slowdown in CEE countries: exploring the role of Global Value Chains”, (No. 2143) *ECB Working Paper*.

Cooper, R., J. Haltiwanger and L. Power (1999): “Machine replacement and the business cycle: lumps and bumps”, *American Economic Review*, Vol. 89, No. 4, pp. 921-946.

Davis, S.J. and J. Haltiwanger (1996): *Job Creation and Destruction*, (Cambridge, MA: The MIT Press).

De Loecker, J. and J. Eeckhout (2017): “The rise of market power and the macroeconomic implications”, (No. w23687) *NBER Working Paper*.

De Loecker, J. and F. Warzynski (2012): “Mark-ups and firm-level export status”, *American Economic Review*, Vol. 102, No. 6, pp. 2437-2471.

De Loecker, J., P.K. Goldberg, A.K. Khandelwal and N. Pavcnik (2016): “Prices, mark-ups, and trade reform”, *Econometrica*, Vol. 84, No. 2, pp. 445-510.

Decker, R. A., J. Haltiwanger, R.S. Jarmin and J. Miranda (2016): “Where has all the skewness gone? The decline in high-growth (young) firms in the US”, *European Economic Review*, Vol. 86, pp. 4-23.

Deutsche Bank (2018): The persistence of zombie firms in a low yield world, *Policy Note*.

Dhingra, S. and J. Morrow (2016): “Monopolistic competition and optimum product diversity under firm heterogeneity”, forthcoming in: *Journal of Political Economy*.

Dhyne, E., A. Petrin and F. Warzynski (2011): “Prices, Mark-ups and Quality at the Firm-Product Level”, *Mimeo*, University of Minnesota.

ECB (2017): “Firm heterogeneity and competitiveness in the European Union”, ECB Economic Bulletin 2017(2).

Edmond, C., V. Midrigan and D.Y. Xu (2015): “Competition, mark-ups, and the gains from international trade”, *American Economic Review*, Vol. 105, No. 10, pp. 3183-3221.

Eeckhout, J. and P. Kircher (2018): “Assortative Matching With Large Firms”, *Econometrica*, Vol.86, No. 1, pp. 85-132.

Epifani, P. and G. Gancia (2011): “Trade, mark-up heterogeneity and misallocations”, *Journal of International Economics*, Vol. 83, No. 1, pp. 1-13.

European Central Bank (2017): Survey on the access to finance for SMEs in the euro area (SAFE), *Statistical Data Warehouse*. Retrieved from: https://www.ecb.europa.eu/pub/pdf/other/ecb.accesstofinancesmallmediumsizedenterprises201705.en.pdf?17da4ff2a730b7ababea4037e4ce8cae

Fernández, C., R. García, P. Lopez-Garcia, B. Marzinotto, R. Serafini, J. Vanhala and L. Wintr (2017): “Firm growth in Europe: an overview based on the CompNet labour module”, *ECB WP 2048*.

Ferrando, A. and A. Ruggieri (2015): “Financial Constraints and Productivity: Evidence from Euro Area Companies”, (No. 1823) *ECB Working Paper*.

Ferrando, A., M. Iudice, C. Altomonte, S. Blank, M.H. Felt, F. Meinen, K. Neugebauer and I. Siedschlag (2015): “Assessing the financial and financing conditions of firms in Europe: the financial module in CompNet”, *ECB WP 1836*.

Gal, P. (2013): "Measuring Total Factor Productivity at the Firm Level using OECD-ORBIS", *OECD Economics Department Working Papers*, No. 1049, OECD Publishing, Paris, https://doi.org/10.1787/5k46dsb25ls6-en.

Gopinath, G., Ş. Kalemli-Özcan, L. Karabarbounis and C. Villegas-Sanchez (2017): “Capital allocation and productivity in South Europe”, *The Quarterly Journal of Economics*, Vol. 132, No. 4, pp.1915-1967.

Haltiwanger, J., R.S. Jarmin and J. Miranda, J. (2013): “Who creates jobs? Small versus large versus young”, *Review of Economics and Statistics*, Vol. 95, No. 2, pp. 347-361.

Kaldor, N. (1978): “The effect of devaluations on trade in manufacturers”, in N. Kaldor (ed.), *Further Essays on Applied Economics* (London: Duckworth), pp. 99-118.

Lopez-Garcia, P. and F. Di Mauro (2015): “Assessing European competitiveness: the new CompNet microbased database”, (No. 1764) *ECB Working Paper*.

Lu, Y. and L. Yu (2015): “Trade Liberalization and Mark-up Dispersion: Evidence from China's WTO Accession”, *American Economic Journal: Applied Economics*, Vol. 7, No. 4, pp. 221-53.

Mayer, T. and G.I.P. Ottaviano (2008): “The Happy Few: The Internationalisation of European Firms”, *Intereconomics,* Vo. 43, No. 3, pp 135–148.

McGowan, M.A., D. Andrews and V. Millot (2017): “The Walking Dead? Zombie Firms and Productivity Performance in OECD Countries”, (No. 1372) *Economics Department Working Paper*.

Melitz, M.J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”, *Econometrica*, Vol. 71, No. 6, pp. 1695-1725.

Melitz, M. J, and S.J. Redding (2014): “Heterogeneous Firms and Trade”, *Handbook of International Economics*, 4th ed, 4: 1-54. Elsevier.

OECD (2017): “The Great Divergence(s): The link between growing productivity dispersion and wage inequality”, Directorate for Science, Technology and Innovation *Policy Note*.

Petrin, A. and J. Sivadasan (2013): “Estimating lost output from allocative inefficiency, with an application to Chile and firing costs”, *Review of Economics and Statistics*, Vol. 95, No. 1, pp. 286-301.

Schivardi, F., E. Sette and G. Tabellini (2017): “Credit Misallocation During the European Financial Crisis”, (No. 3/2017) *Bank of Italy Working Paper*.

Schreyer, P. (2000): “High-growth firms and employment” (No. 2000/3) *OECD Science, Technology and Industry Working papers*.

Smets, F. and R. Wouters (2007): “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach”, *American Economic Review,* Vol. 97, No.3, pp.586-606.

Storz, M., M. Koetter, R. Setzer, A. Westphal (2017): “Do we want these two to tango? On zombie firms and stressed banks in Europe”, (No. 2104) *ECB Working Paper*.

Syverson, C. (2011): “What Determines Productivity?”, *Journal of Economic Literature*, Vol. 49, No 2, 2011, pp. 326-365.

Wooldridge, J. M. (2009): “On estimating firm-level production functions using proxy variables to control for unobservables”, *Economics Letters*, Vol. 104, No. 3, pp. 112-114.