

OECD Science, Technology and Industry Working Papers
2018/14

A taxonomy of digital intensive sectors

Flavio Calvino,
Chiara Criscuolo,
Luca Marcolin,
Mariagrazia Squicciarini

https://dx.doi.org/10.1787/f404736a-en

OECD

OECD Science, Technology and Industry Working Papers

The release of this working paper has been authorised by Andrew Wyckoff, OECD Director for Science, Technology and Innovation and by Martine Durand, OECD Director for Statistics Directorate.

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the authors.

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to Directorate for Science, Technology and Innovation, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

This publication is a contribution to the OECD Going Digital project, which aims to provide policymakers with the tools they need to help their economies and societies prosper in an increasingly digital and data-driven world.

For more information, visit www.oecd.org/going-digital.

#GoingDigital

This document, as well as any data and any map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Note by Turkey

The information in this document with reference to “Cyprus” relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Turkey recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the “Cyprus issue”.

Note by all the European Union Member States of the OECD and the European Union

The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

© OECD 2018

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for commercial use and translation rights should be submitted to rights@oecd.org.

A TAXONOMY OF DIGITAL INTENSIVE SECTORS

Flavio Calvino, Chiara Criscuolo, Luca Marcolin, and Mariagrazia Squicciarini[1]

Abstract

This study proposes a taxonomy of sectors according to the extent to which they have gone digital. The taxonomy accounts for some of the key facets of the digital transformation, and recognises that sectors differ in their development and adoption of the most advanced "digital" technologies, in the human capital needed to embed them in production and in the extent to which digital tools are used to deal with clients and suppliers. The indicators used to classify 36 ISIC revision 4 sectors over the period 2001-15 are: share of ICT tangible and intangible (i.e. software) investment; share of purchases of intermediate ICT goods and services; stock of robots per hundreds of employees; share of ICT specialists in total employment; and the share of turnover from online sales.

Some sectors (e.g. telecom and IT services) are positioned consistently at the top of the distribution of sectors, independently of the indicator used, while others are positioned in the bottom quartile of digital intensity across all available indicators (e.g. agriculture, mining and real estate). This is true for the classification of sectors based on both 2001-03 and 2013-15 data. Sectors positioned in the middle, instead, differ across indicators, suggesting that they are engaged in the digital transformation at different rates, depending on the dimensions considered.

The study further proposes an overall summary indicator of the digital transformation in sectors which encompasses all the considered dimensions. Preliminary results suggest that this summary indicator performs reasonably well in approximating sectors' intensities in the different dimensions of the digital transformation.

Table of contents

A TAXONOMY OF DIGITAL INTENSIVE SECTORS 3

Abstract 3

1. The digital transformation: a tale of many tales 5

2. Measuring the digital transformation of sectors: rationale and main challenges 8

3. Digitalisation: key indicators 10

3.1. ICT investment 10
3.2. Purchases of ICT intermediates 14
3.3. Robot use 19
3.4. ICT specialists 22
3.5. Online sales 24

4. Digital and less digital intensive sectors 27

4.1. Key stylised facts 27
4.2. A “global” taxonomy 30

5. Preliminary conclusions and next steps 34

Appendix 41

1. The digital transformation: a tale of many tales

The digital transformation is changing the way production is organised and managed and is affecting economic performance. However, thirty years after Solow stating that "you can see the computer age everywhere but in the productivity statistics" (Solow, 1987: 36), assessing to what extent going digital translates into broad-based economic growth and benefits society as a whole remains challenging. Inherent difficulties exist in measuring the scope and pace of such a multifaceted phenomenon. Traditional metrics, which only partially capture the degree of penetration of (certain) digital technologies struggle to mirror the fast pace at which the digital transformation unfolds. Also, the usefulness of well-established measures of technology diffusion and adoption is at times reduced by the low granularity or inadequate cross-country coverage of the data that are available.

Evidence-based policy making calls for a better understanding of what the digital transformation is and entails for economies and societies. The present study therefore proposes a taxonomy of sectors according to the extent to which they have gone digital. The taxonomy accounts for some of the key facets that the digital transformation may take, and recognises that sectors differ in their development and adoption of new technologies (see e.g. Comin and Mestrieri, 2013), their human capital endowment, industrial structure or organisation of the production. The digital transformation in fact does not only entail investing in and developing the most advanced "digital" tools, but also embedding them in production with the appropriate human capital, and using them when dealing with clients and suppliers.

This report presents a number of indicators and uses them to classify sectors (aggregated into 36 ISIC revision 4 sectors) based on country-sector information for the period 2001-15. The different indicators relate to technology, market and human capital-related features, namely: share of ICT tangible and intangible (i.e. software) investment; share of intermediate purchases of ICT goods and services; stock of robots per hundreds of employees; share of ICT specialists in total employment (here often referred to as "ICT specialists intensity"); and the share of turnover from online sales. The relevance of each of these dimensions, as well as the sources of information and the assumptions, calculations or estimations made to obtain a meaningful time series and to overcome data availability constraints are discussed in detail in the rest of the report.

To help reduce the multidimensionality of the measures considered, while accounting for the complexity characterising the digital transformation, an overall summary indicator is proposed, called "global", constructed using all other indicators. This overall summary indicator should nevertheless be considered as a general pointer, as analysis of the different components shows that the dimensions considered are at times uncorrelated or even negatively correlated. This is true especially with respect to ICT specialists and online sales.

While distinguishing sectors which are digital intensive from those that are less digital intensive, the proposed "global" taxonomy by construction masks some of the heterogeneity displayed by sectors across the different underlying indicators. To the best of the authors' knowledge, McKinsey (2015) is the only other study attempting to propose a similar taxonomy of digital *sectors*. McKinsey's work however differs from the one proposed here along a number of dimensions: it focuses on the United States only; it displays lower granularity, by relying on more aggregate sectors; and it exploits official

but also self-gathered data. In addition, and perhaps more importantly, it substantially differs in the dimensions of the digital transformation that are considered, and the methodological choices and assumptions underlying the indicators used.

The aim of the present taxonomy is to provide an operational tool helping analysts and policy makers to better understand the sectoral heterogeneity of the digital transformation and its relationship with various aspects of production. It benchmarks sectors rather than attempting to measure the size of the digital economy. By doing so it helps revealing e.g. which parts of the economy are embracing the digital transformation to a greater extent and along which dimension.

This taxonomy is meant to complement other well-established typologies of sectors based on R&D or knowledge intensity of production (OECD, 2003; Hirsch-Kreinsen et al., 2008; Galindo-Rueda and Verger, 2016), the innovative performance of firms in sectors (OECD, 2011a), or the fact that sectors are producers or users of ICT or of media content (OECD, 2002; 2011b). However, by taking into account the encompassing nature of the digital transformation, the present taxonomy is able to provide a more accurate picture of the extent to which sectors are going digital, and to account for the different dimensions of this transformation.

As mentioned, while relevant, the taxonomy proposed here provides an imperfect picture of the digitalisation of production, and cannot address some of its aspects.

First, the taxonomy does not account for the development and adoption of frontier technologies such as machine learning or 3D printing (as portrayed, for instance in OECD, 2017a). This is mainly due to the current lack of access to comparable and representative data on the deployment of such technologies at the country-sector-year level. However if, as it is reasonable to expect, the unobserved adoption of these frontier technologies is (strongly) correlated with any of the dimensions included here, the present taxonomy remains informative also in this respect.

Second, due to data constraints as the data is at the sector and not at the firm-level, the taxonomy does not capture the within-sector heterogeneity likely to characterise the development and adoption of digital technologies. OECD work (Andrews et al., 2016), for instance, provides evidence of a widening gap between firms at the frontier and all other firms, even within narrowly-defined sectors. It further highlights that frontier firms (top 2%) in ICT services experienced a significantly higher growth in productivity and in market share not only vis-à-vis non-frontier firms, but even within the group of best performing firms (top 10%) in the ICT services sector. This evidence suggests that, by exploiting sector-level averages or aggregate information about technology production and use, the taxonomy does not accurately mirror within-sector heterogeneity, and may hide the fact that scarcely digitalised firms may still operate in digital intensive sectors. In fact, the taxonomy can only rely on firm-level surveys of ICT use in businesses to measure online sales. Since such surveys do not contain information on some of the dimensions highlighted above (e.g. ICT investment, purchases of ICT intermediates) they cannot be the only source of information for the taxonomy. Previous OECD work (OECD, 2009a; Spiezia, 2011) explored the within-sector heterogeneity of technology adoption by accessing micro-level survey data on firms' innovative efforts and ICT investment across multiple OECD countries with a uniform methodology. It is not clear at this stage, however, whether a similar approach could be pursued in the future to inform the digital transformation discussion, given constraints to cross-country data availability.

Third, the use of the taxonomy for analytical purposes should carefully consider the purpose of the analysis itself. A study may benefit more from a taxonomy based on only one of the indicators, rather than the "global" classification. In addition, the taxonomy reflects the technological trajectories of a sample of developed OECD countries with similar GDP per capita, which are likely to be different from the trajectories of developing or emerging countries. This is even more true when considering that diffusion and adoption of technologies differ across countries even within the sub-sample considered.

The remainder of this paper is structured as follows: section 2 discusses the rationale and key challenges associated with measuring the digital transformation of sectors. Section 3 introduces the indicators underlying the taxonomy, their importance, the methodologies devised to make them comparable and meaningful, as well as some descriptive statistics. Section 4 introduces the "global" indicator and the classification of sectors which can be derived from it. Some first conclusions and a technical data Appendix follow.

2. Measuring the digital transformation of sectors: rationale and main challenges

Our approach to measure the digital transformation starts from two considerations.

The first one is that digitalisation is a complex phenomenon that is hardly captured by a single indicator. As a matter of fact, digital technologies affect different dimensions of economic activities, including capital, labour and markets.

The second consideration is that different sectors of activity are affected by digital technologies in heterogeneous ways. This heterogeneity across sectors will likely depend by the dimension of digitalisation considered. For instance, a sector where a high number of digital transactions occur might not necessarily be characterised by a very high presence of ICT specialists, and vice versa.

In this framework, the indicators presented here aim at representing digitalisation in:

- its ***technological component***, by looking at ICT investment, ICT intermediate consumption and robots
- the ***human capital*** required to embed technology in production, by looking at ICT specialists intensity
- the way it changes the way firms behave on the output ***market***, by looking at online sales.

Considerable efforts were devoted to sourcing the data and ensuring that the country, sector and time coverage of the resulting dataset are uniform across indicators, so as to enhance comparability.[2]

To this end, five main concerns have been addressed:

- *Sectoral coverage*. Constructing the different indicators is complicated by the fact that available data are often classified according to national sectoral classifications and/or aggregated at different levels. Moreover, sometimes coverage of selected sectors changes from country to country for the same indicator. In addition, even for those countries presenting data according to international standardised classification systems, time series of most indicators are affected by the 2008 ISIC classification revision (from revision 3 to revision 4).[3] To address these challenges, a number of conversions between classifications have been made. In some cases where information is available at a higher level of aggregation for a few sectors, the same intensity of the more aggregated sector is attributed to all underlying subsectors. The final sectoral classification used for the present taxonomy follows the 38-category breakdown proposed in the System of National Accounts and relies on the ISIC revision 4 sectoral classification. Two of the 38 sectors are excluded: household production and production by international organisations. The complete list of sectors considered is reported in Table A1 in the Appendix. For two indicators (i.e. robots and online sales) sectoral coverage is limited, as discussed more in detail below.
- *Country coverage*. Geographical coverage differs across indicators: some indicators are not available for a sub-set of countries; others suffer from restrictions stemming from the existing sectoral breakdown and / or from

appropriately converting relevant figures from national sectoral classifications to ISIC. For others, the nature of the indicator requires working at the cross-country level only. As a consequence, the taxonomy and the descriptive statistics shown rely on the largest "common denominator" of countries for which all sector-specific indicators can be constructed or exist, namely: Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.[4] The extent to which the measures computed generalise to other countries remains an empirical question that deserves further investigation.

- *Time coverage*. While for some of the indicators considered the available time series at the country-sector level is long (e.g. ICT investment), for others it is short or incomplete (e.g. ICT use survey). Also, in some cases, the time series differ across countries or sectors for the same dimension. Interpolation and extrapolation have been used, although to a minimal extent, to address these issues, with a view to construct a time series for the years 2000-15. Also, to try and reduce measurement error, and avoid that statistics mirror possible time-specific features or shocks, the descriptive statistics below as well as the final classifications of sectors are calculated over 3-year windows. In the interest of brevity, we report figures for the beginning and the end of the period, i.e. for the periods 2001-03 and 2013-15.
- *Monetary values*. Many of the indicators used in the taxonomy are measured in monetary terms. Comparisons of their change over time across indicator, sector or country would be misleading if price dynamics differed along these dimensions. All monetary values are therefore deflated with sector or country-level price series.
- *Outliers*. To avoid that the presence of outliers may bias the statistics, extreme values, both in levels and in changes of each indicator are removed where necessary, and the time and country coverage adjusted as a consequence.

3. Digitalisation: key indicators

The indicators used in the taxonomy are presented below, together with a discussion of their rationale and purpose, i.e. what they try and capture, the assumptions made and how they are developed, to ensure that they provide an objective and coherent picture of the underlying economic phenomenon across countries.

3.1. ICT investment

Building on a long-standing literature (e.g. Colecchia and Schreyer, 2002; Pilat and Lee, 2001), this study considers investment in information and communication technologies (ICT) a fundamental part of the digital transformation of sectors and countries. Data about sectors' investment intensities in both ICT equipment and in software and databases are separately used, to account for the tangible and intangible component of ICT. Data are mainly sourced from the System of National Accounts (SNA) tables of gross fixed capital formation (GFCF) by assets.[5]

"Investment in ICT equipment" is constructed as investment in computer hardware and telecommunication equipment as a share of non-residential gross fixed capital formation in the sector.[6] "Investment in software and databases" (hereafter: "Investment in software") is constructed as investments in software and databases divided by the same denominator used in the case of ICT equipment, i.e. non-residential gross fixed capital formation in the sector.[7]

Software, ICT tangible and non-residential GFCF are separately deflated using different country-wide deflators, given that different price dynamics characterise these assets. More detailed information on the deflation, on the data cleaning procedures, and on possible alternative indicators is available in the Appendix (methodological annex).

Both measures of ICT investment do not account for ICT investment which is "embodied" in other forms of capital. For example, investment in software which is purchased with computers may be recorded as ICT tangible investment rather than software investment. Similarly, computing equipment or software investment embodied in machinery is likely to accrue to non-ICT rather than ICT capital investment. Lastly, investment in machinery used to produce ICT investment goods is usually recorded as non-ICT investment.[8] These forms of mismeasurement are likely to bias downwards the "digital" content of some sectors, which rely on non-purely-ICT forms of investment for production. Current data availability limits the scope for a more precise measurement of such "embodied" ICT investment at the country-sector level, which is therefore left to future research.

3.1.1. Investment in ICT equipment

Figure 1 shows the sector-specific cross-country averages of "Investment in ICT equipment" for the periods 2001-03 and 2013-15.

Focusing on the 2013-15 period, "Telecommunications" and "IT" services emerge as top investors, followed by other services sectors including "Marketing and other", "Finance", or "Legal and accounting".[9] Conversely, most sectors in manufacturing tend to show relatively lower shares of investment in ICT equipment. This may be partially due to the relatively greater size of the denominator, i.e. the amount of non-residential GFCF, of

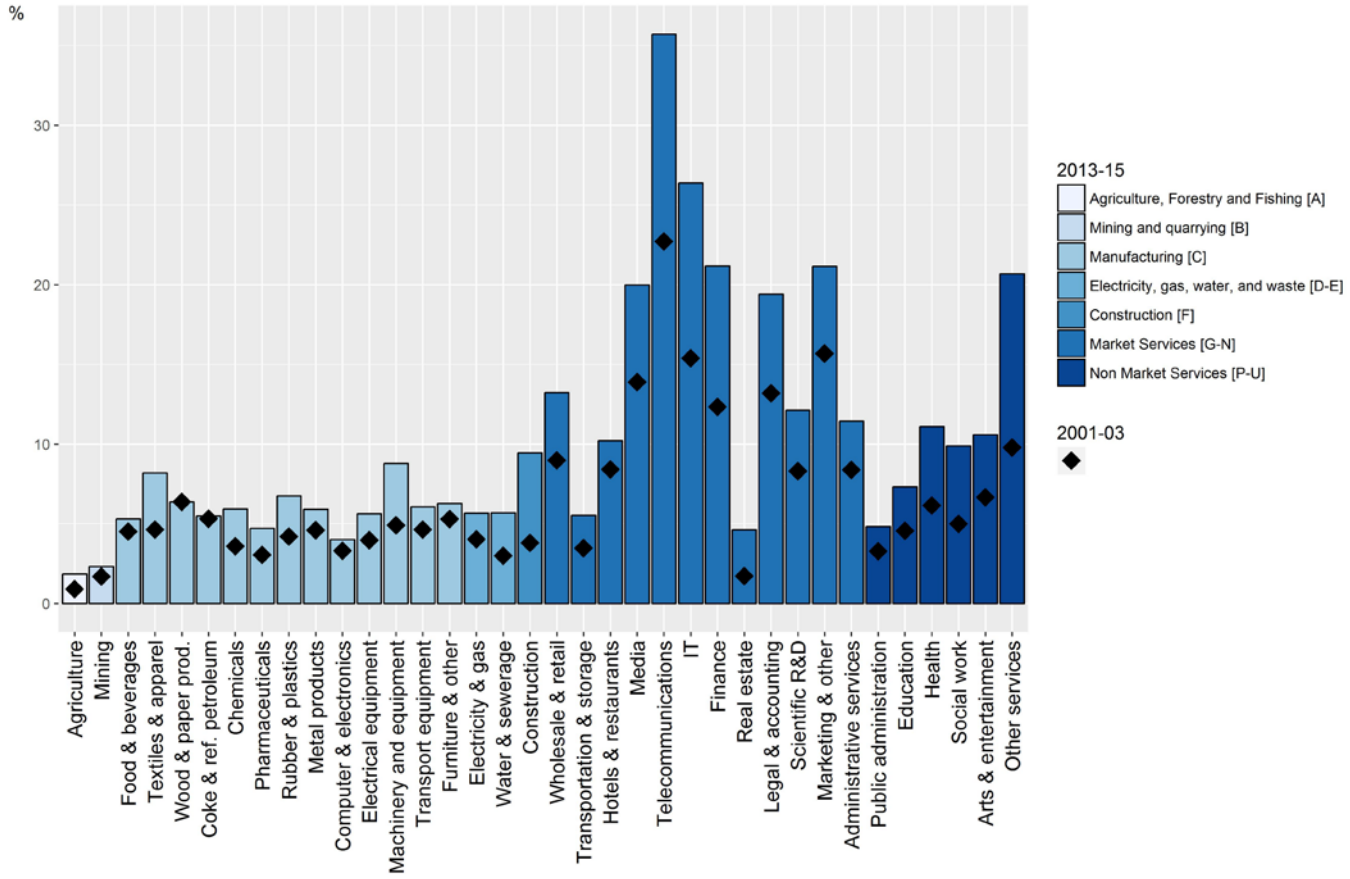
these sectors, which tend to be relatively more capital intensive and often make significant investments in assets other than ICT.

Interestingly, the same sectors appearing as the top investors over the period 2013-15 are also the ones having experienced the highest increase in intensity over time. This can be seen focusing on the diamonds in Figure 1, which compares the 2013-15 sector-specific averages with the same statistics computed for the 2001-03 period, i.e. the beginning of the available time series.

Figure 2 further presents the cross-country within-sector dispersion of the indicator, computed for the 2013-15 period. A significant degree of cross-country heterogeneity emerges, especially in sectors featuring high shares of investments in ICT equipment, but also in a few other manufacturing sectors.

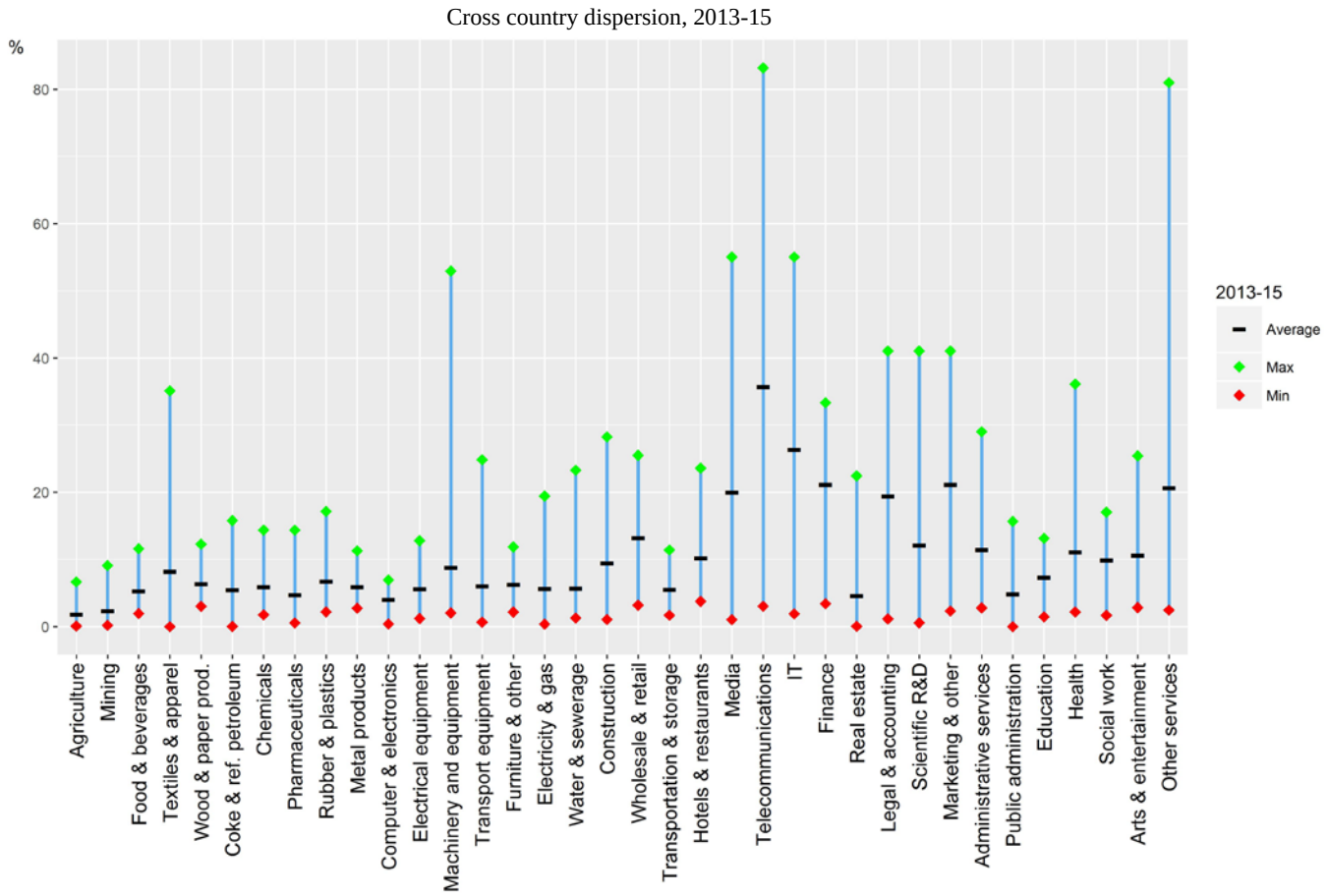
Figure 1. Investment in ICT equipment as a % of non-residential GFCF, by sector

Unweighted averages across countries, 2001-03 and 2013-15



Note: Sector-specific values are unweighted averages over the years 2001-03 and 2013-15 and across countries. Countries covered are: Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.
Source: OECD Annual National Accounts; EU-KLEMS.

Figure 2. Investment in ICT equipment as a % of non-residential GFCF, by sector



Note: Sector-specific values are unweighted averages over the years 2013-15, for each country ("Average"). "Max" ("Min") represent the country with highest (lowest) value of the indicator in the sector and period considered.
Source: OECD Annual National Accounts; EU-KLEMS.

3.1.2. Investment in software and databases

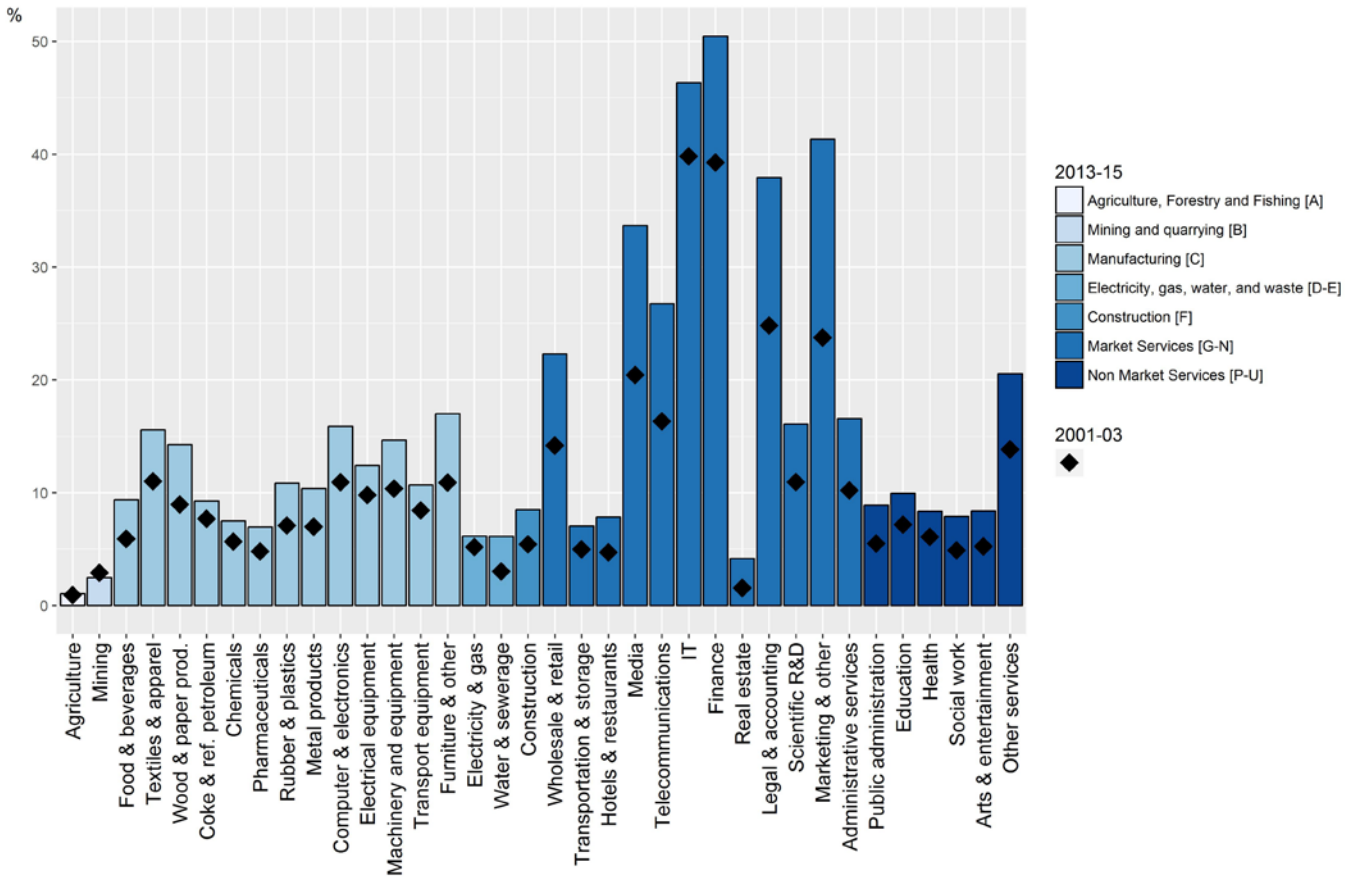
Figure 3 illustrates the "Investments in software and databases" indicator focusing again on sector-specific cross-country averages in two periods (2013-15 and 2001-03).

Focusing on the 2013-15 period, "Finance" and "IT" services emerge as the sectors with the highest shares of investment in software and databases over total investment. "Marketing and other", "Legal and accounting" and, to some extent, "Media" also score relatively high in terms of software investment intensity. A possible explanation may be that these sectors need relatively more ICT assets to perform their core activities, as compared to manufacturing sectors.

These market services sectors also experience the highest variation, when comparing the 2013-15 values with the 2001-03 ones.

Figure 3. Investment in software and databases as a % of non-residential GFCF, by sector

Unweighted averages across countries, 2001-03 and 2013-15

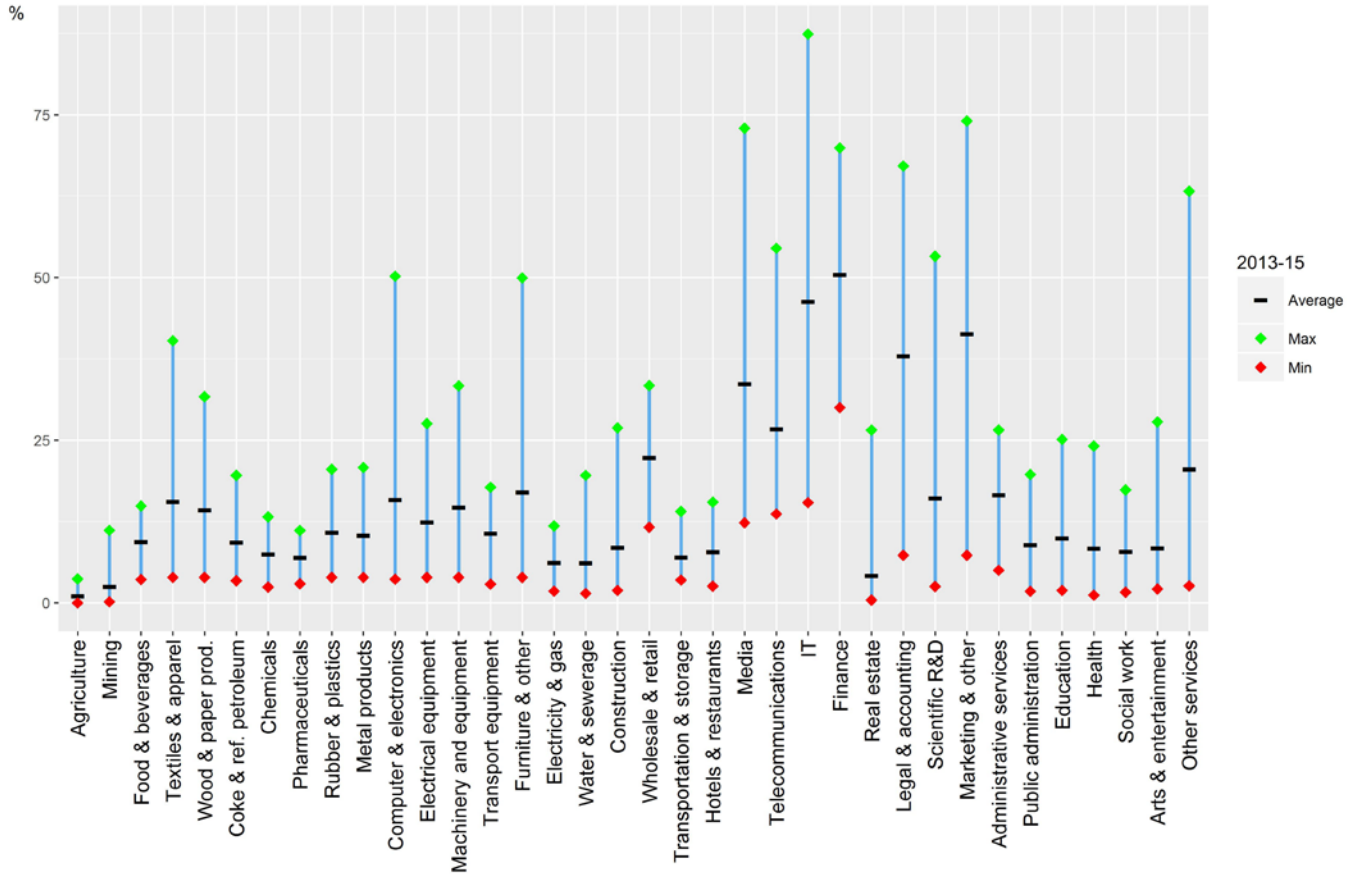


Note: Sector-specific values are unweighted averages over the years 2001-03 and 2013-15 and across countries. Countries covered are: Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.
Source: OECD Annual National Accounts; EU-KLEMS; Corrado et al. (2012).

Similarly to the case of investment in ICT equipment, there appears to be a significant degree of cross-country heterogeneity, especially in sectors with high shares of investments in software and databases, but also in few other manufacturing sectors, as presented in Figure 4.

Figure 4. Investment in software and databases as a % of non-residential GFCF, by sector

Cross country dispersion, 2013-15



Note: Sector-specific values are unweighted averages over the years 2013-15, for each country ("Average"). "Max" ("Min") represent the country with highest (lowest) value of the indicator in the sector and period considered.
Source: OECD Annual National Accounts; EU-KLEMS; Corrado et al. (2012).

3.2. Purchases of ICT intermediates

This study is one of the first to rely on the composition of the consumption of intermediate materials to proxy the digital intensity of sectors, in addition to using data about investment in ICT capital or software.[10] The SNA and OECD (2009b) recommend the capitalisation of expenditure if a purchase has a "useful life of more than one year", i.e. it is not entirely consumed in the annual process of production. However, this only paints a limited view of overall ICT use. For example a firm purchasing software with one-year licences will record the transaction as intermediate consumption but if the same software was purchased with a two year license it would be recorded as investment. Moreover, firms nowadays can extract productivity gains from expenses in technological items which have a shorter "useful life", such as IT consulting or data processing. These expenses may be the result of a strategy which substitutes what used to be an investment in fixed assets with the purchase of intermediates from a different company or sector. Companies may, for instance, forgo investing in servers for data storage and processing,

and purchase access to servers which are owned by an external provider, as it happens in the case of e.g. cloud computing. In addition, some purchases of technological goods and services may be acquired for the production of, de facto, assets that have a "useful life" of longer than one year, but may not yet be included in the fixed asset boundary of the SNA (e.g. network construction, although in practice related expenditures will be capitalised, for example in machinery or software used in producing these networks). Lastly, there may be differences across countries in the propensity to record expenditures in the national accounts as investment or intermediate consumption (see e.g. Ahmad, 2003).

Purchasing ICT goods or services grants companies the possibility to (i) use the technology at exactly the capacity required and, possibly, in a more customised fashion; (ii) adjust the use of the technology fast and frequently; (iii) avoid maintenance costs; and (iv) the possibility to increase in scale even without access to important financial resources (which might be required to make an investment). The contribution of ICT technologies therefore may be under-estimated if one were to look at investment data only and not also at services-related expenses.[11]

This study measures ICT intermediate consumption on the basis of the OECD Inter-Country Input-Output (ICIO) database, as well as national Input-Output tables. In the so-called intermediate matrix, these databases record the purchases of intermediate goods and services carried out by a given sector and country from any other sector in the same country or abroad (for the ICIO, in particular). The proposed indicator of "purchases of ICT services" measures the purchases carried out by a given sector for services produced by the ISIC revision 3 sector number 72 "Computer and related activities". These services include hardware and software consultancy, maintenance of computing equipment and data processing.[12] The proposed indicator of "purchases of ICT goods" measures the purchases carried out by a given sector for intermediate goods produced by the ISIC revision 3 sectors 30, 32 and 33 "Computer and electronics".[13] This measure somewhat overstates the actual "digital intensity" of sectors, given that we are mostly interested in the purchases of computing equipment intermediates, but lack of data at the needed level of disaggregation impinges upon the possibility to rely on more accurate estimates.[14]

For the machinery production sectors (ISIC revision 3 sectors 29 to 35), intermediate materials from the "Computer and electronics" sector are likely to be microchips or electronic components, used in the production of goods that are subsequently sold-on to other consumers, and so are not 'used' by the producing firms as "substitute" or 'complementary' to investment. For this reason, the purchases of ICT intermediate goods are put to missing for these sectors.[15]

Lastly, purchases of IT intermediates are deflated using the price of output of ISIC revision 4 sectors 26 ("Computer and electronics") and 62-63 ("Telecommunications" and "IT" services) in the country which is purchasing the intermediates.[16] The denominator, i.e., sectoral output, is deflated by the price of output in the purchasing sector. A measure of intensity is then obtained by dividing the two deflated figures.

The sectoral coverage is adjusted to match the A38 ISIC revision 4 sectoral classification in National Accounts, resulting in equal intensities for two or more sectors where a more disaggregated sectoral classification is not available in the ICIO (e.g. electricity, gas and water supply, or business services). More information on the deflation and cleaning of the data series is provided in the Appendix (methodological annex).

As for ICT investment here above, the proposed indicators only measure direct consumption of ICT intermediates, as opposed to intermediate consumption of ICT goods

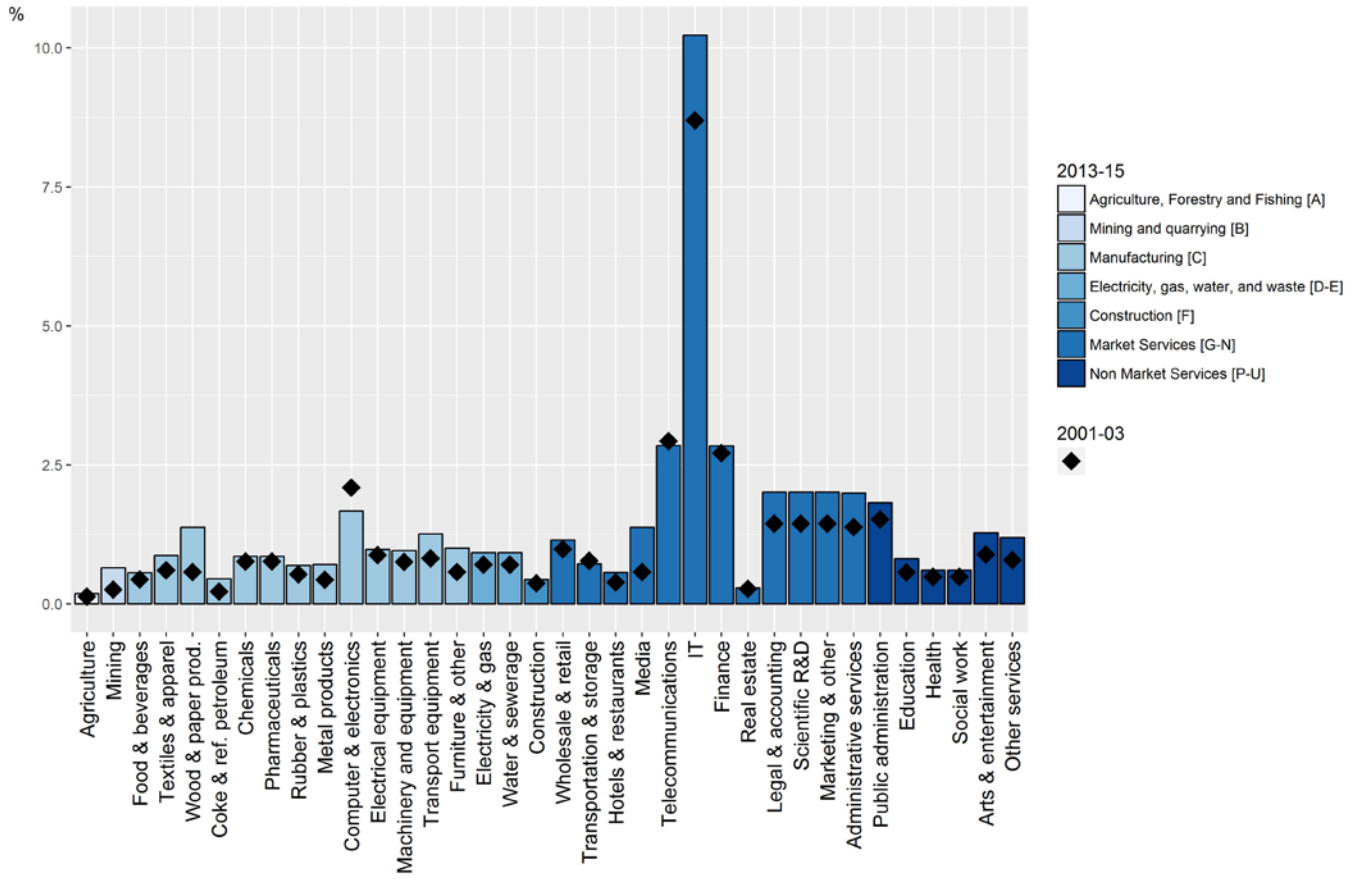
and services embodied in other intermediates. In the case, for instance, of a manufacturing company purchasing professional or business services which are strongly relying on ICT services, the constructed indicator does not consider the ICT services as consumed by the manufacturing company. A new indicator of "embodied" ICT intermediate consumption can be constructed in the future on the basis of the intermediate matrix of the ICIO in ISIC revision 4.

3.2.1. Purchases of ICT services

The statistics presented in Figure 5 below show the sector-specific purchases of ICT services as an average for the 2013-15 and 2001-03 periods.

Figure 5. Purchases of intermediate ICT services as a % of output, by sector

Unweighted averages across countries, 2001-03 and 2013-15



Note: The sector's intensity is calculated as sectoral deflated purchases from the ICT service or good production sector over deflated sectoral output. Sector-specific values are unweighted averages over the years 2001-03 and 2013-15 and across country. Countries covered are: Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.
Source: OECD Inter-Country Input-Output matrices; OECD, Annual National Accounts Database; and OECD Structural Analysis (STAN) Database.

The ICT producing sector ("IT" services) displays the highest intensity in ICT services in 2013-15, likely because it produces both datasets and analysis using them, and has seen the highest increase in intensity over time. All other sectors display significantly lower

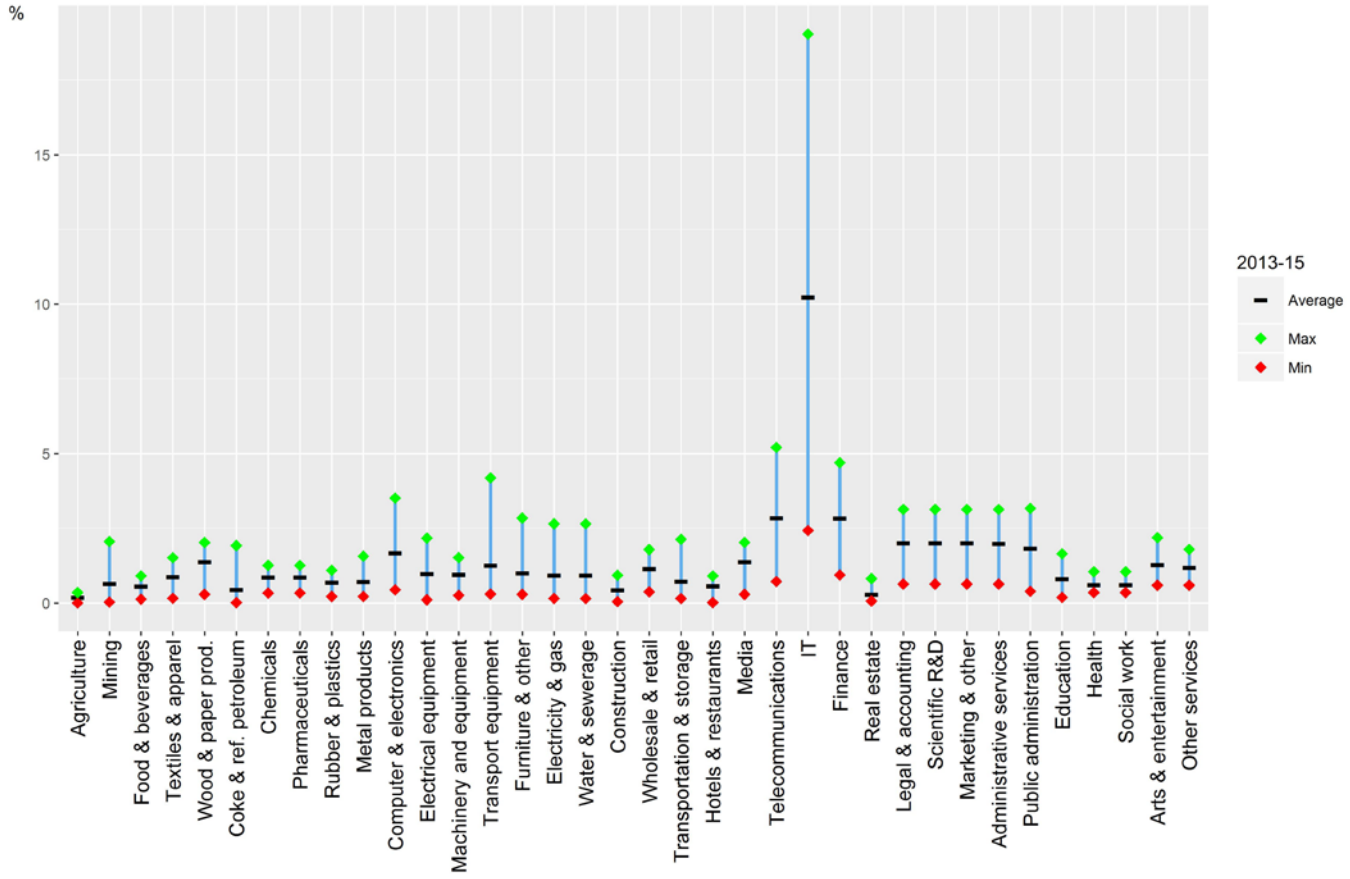
intensities, and manufacturing sectors in particular, while "Telecommunications", "Finance" and business services display the second highest intensities after the ICT service producing sector. This result is coherent with other indicators, where these sectors are positioned at the top of the distribution of digital intensive sectors.

The comparison of the two different periods reported in Figure 5 in addition shows that some manufacturing sectors ("Transport equipment", "Furniture and other" manufacturing, "Wood and paper products") may be "catching up" by increasing their intensity in the use of ICT intermediates at a faster pace than the currently already digitally-intensive sectors.

Figure 6 reports the dispersion of intensities across countries. Generally, sectors with the highest dispersion in intensity across countries also display higher average intensities for 2013-15. This is particularly evident when focusing on the "IT" sector.

Figure 6. Purchases of intermediate ICT services as a % of output, by sector

Cross country dispersion, 2013-15



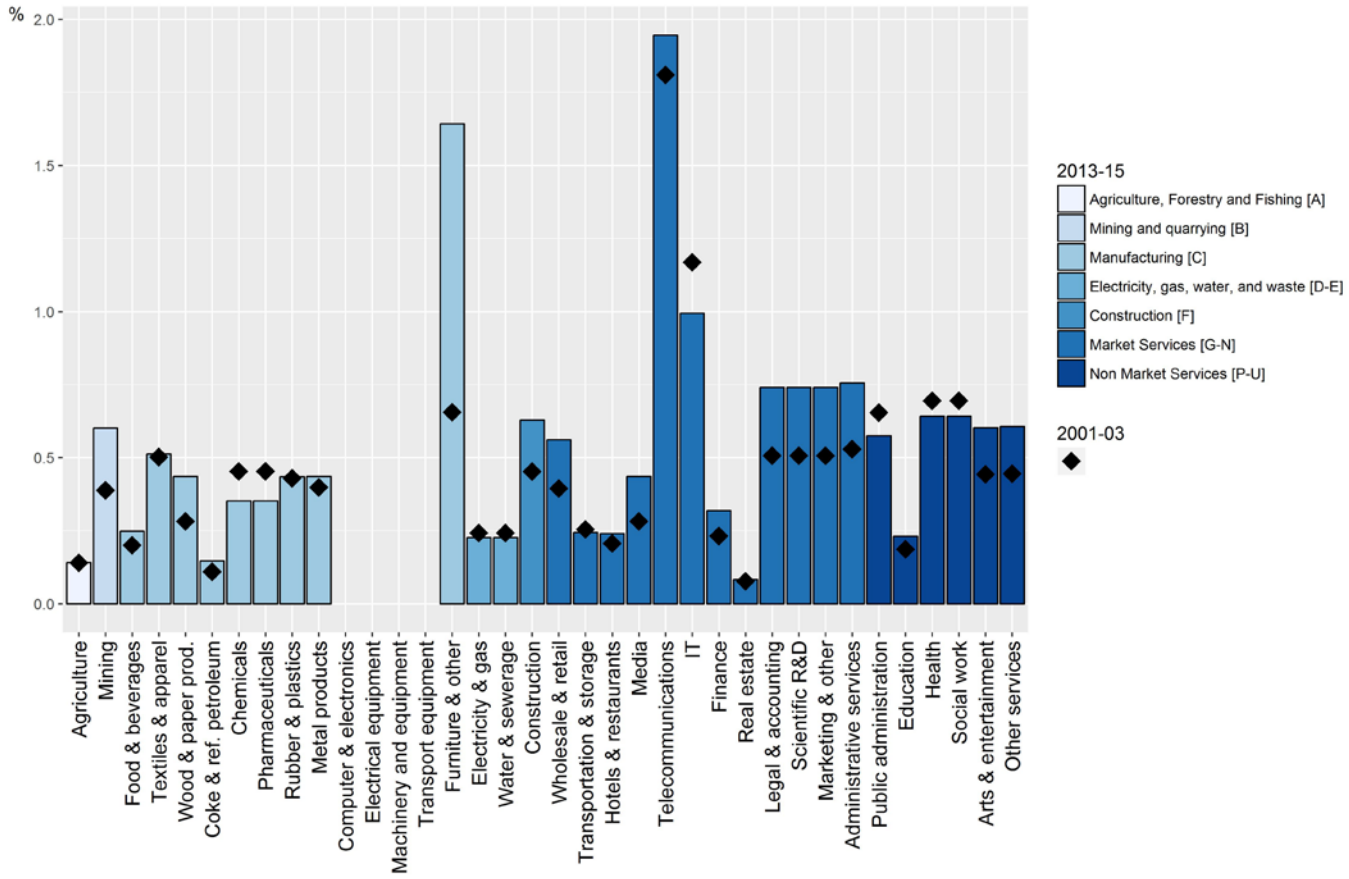
Note: Sector-specific values are unweighted averages over the years 2013-15, for each country ("Average"). "Max" ("Min") represent the country with highest (lowest) value of the indicator in the sector and period considered.
Source: OECD Inter-Country Input-Output matrices; OECD, Annual National Accounts Database; and OECD Structural Analysis (STAN) Database.

3.2.2. Purchases of ICT goods

Figures 7 and 8 present the same descriptive statistics related to the intermediate consumption of ICT goods. By construction, the purchases made by machinery-producing sectors are set to missing, including those by the "Computer and electronics" production sector. The highest intensity is therefore concentrated in the "Telecommunications" service sector, followed by "Furniture and other" manufacturing (which includes repairs and installation of machinery) and "IT" services. The non-machinery manufacturing production displays once again a relatively lower average intensity in ICT intermediates than the one observed in the case of services - whether public services are considered or excluded - and lower average growth rates between the two periods.

Figure 7. Purchases of intermediate ICT goods as a % of output, by sector

Unweighted averages across countries, 2001-03 and 2013-15



Note: The sector's intensity is calculated as sectoral deflated purchases from the ICT good production sector over deflated sectoral output. Sector-specific values are unweighted averages over the years 2001-03 and 2013-15 and across country. Countries covered are: Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.
Source: OECD Inter-Country Input-Output matrices; OECD, Annual National Accounts Database; and OECD Structural Analysis (STAN) Database.

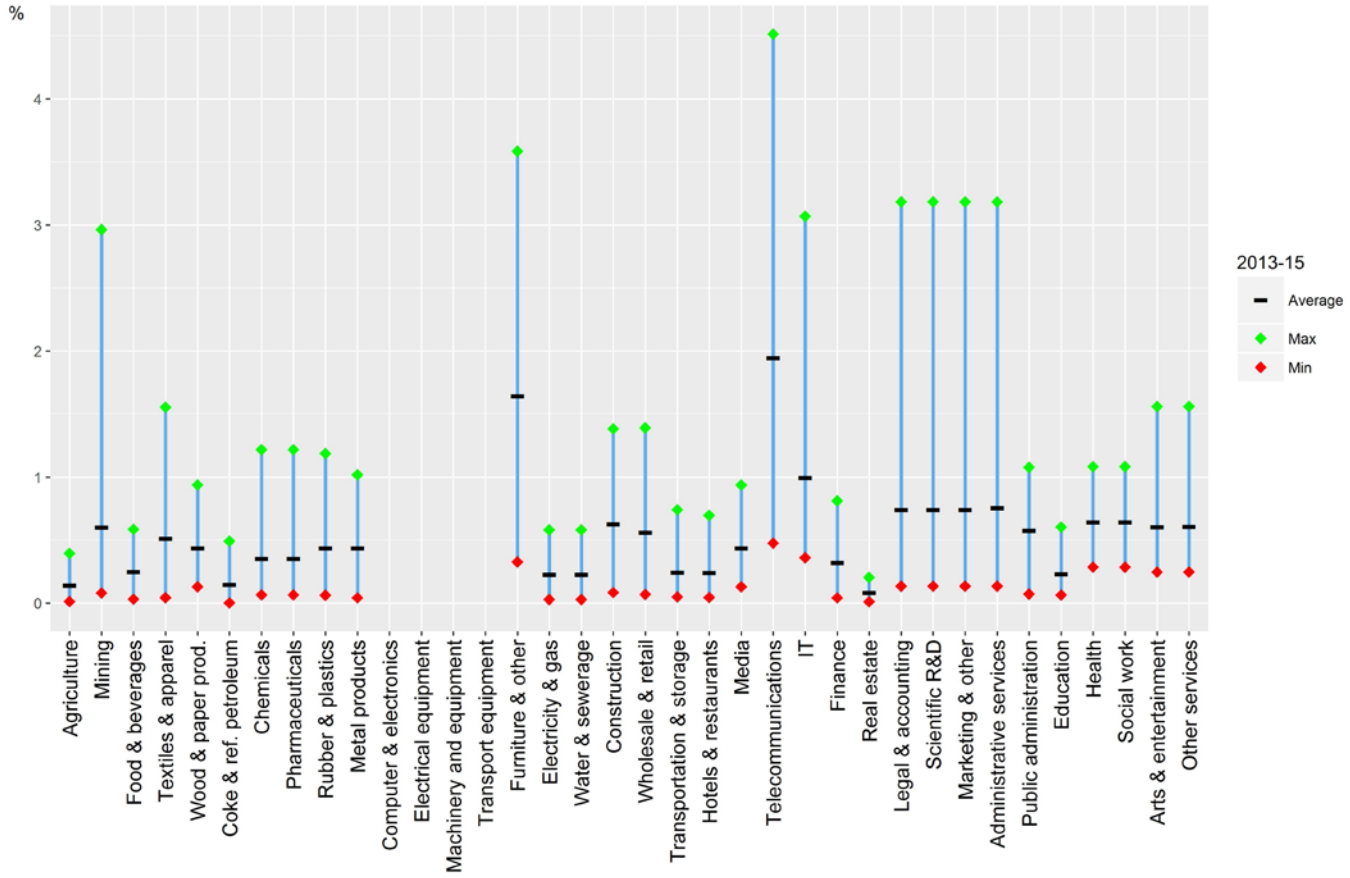
One important driver of the alleged reduction in corporate reliance on ICT investment in favour of consumption of ICT intermediates may be product differentiation and customisation. If this is true, one would expect the phenomenon to be more evident for

services rather than manufacturing sectors: high levels of product differentiation and customisation in services may indeed require great(er) flexibility in production.

Figure 8 shows however that this phenomenon is not equally widespread across countries, with dispersion being higher in services than in manufacturing sectors.

Figure 8. Purchases of intermediate ICT goods as a % of output, by sector

Cross country dispersion, 2013-15



Note: Sector-specific values are unweighted averages over the years 2013-15, for each country ("Average"). "Max" ("Min") represent the country with highest (lowest) value of the indicator in the sector and period considered.
Source: OECD Inter-Country Input-Output matrices; OECD, Annual National Accounts Database; and OECD Structural Analysis (STAN) Database.

3.3. Robot use

Robots and robotics represent one of the key elements contributing to the digital transformation of sectors and countries. Robot-related data have been recently used to assess the consequences of automation on employment, wages and productivity (e.g. Graetz and Michaels, 2015; Acemoglu and Restrepo, 2016; Dauth et al., 2017).[17] In line with these studies and building on OECD work (OECD, 2017b; De Backer et al., 2018), robot use is here measured as the stock of industrial robots owned by a given sector,

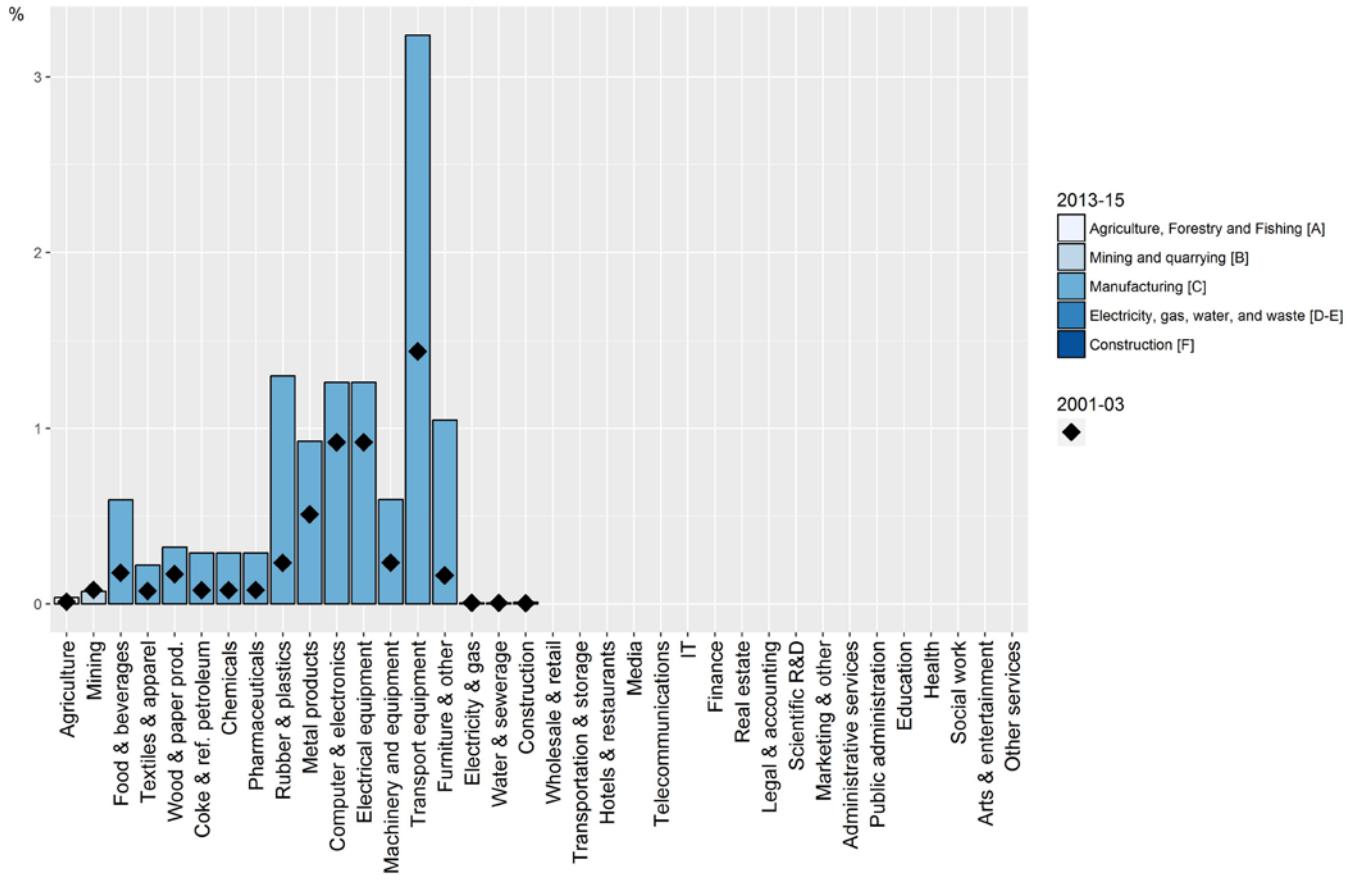
normalised by sector-level number of employees (hundreds). Data on shipments of robots are sourced from the International Federation of Robotics (IFR), which gathers information from national robotics associations and suppliers. Employment data are instead compiled combining a wide range of sources, as detailed in the Appendix (methodological annex).

Robot stock per hundreds of employees is considered at the country-sector-year level, although mainly for manufacturing sectors. The IFR in fact does not provide data for most other sectors, with the exception of construction, utilities, and the R&D producing sector. Additional details, including the definition of industrial robots, the construction and depreciation of the stock measure and the methodology used to impute missing values, are discussed in the Appendix (methodological annex).

Figure 9 provides a graphical visualisation of the robot use intensity indicator, as a sector-specific cross-country average for the periods 2001-03 and 2013-15.

Figure 9. Robots stock per hundreds of employees, by sector

Unweighted averages across countries, 2001-03 and 2013-15



Note: Sector-specific values are unweighted averages over the years 2001-03 and 2013-15 and across countries. Countries covered are: Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.
Source: IFR, OECD STAN, OECD Annual National Accounts, WIOD.

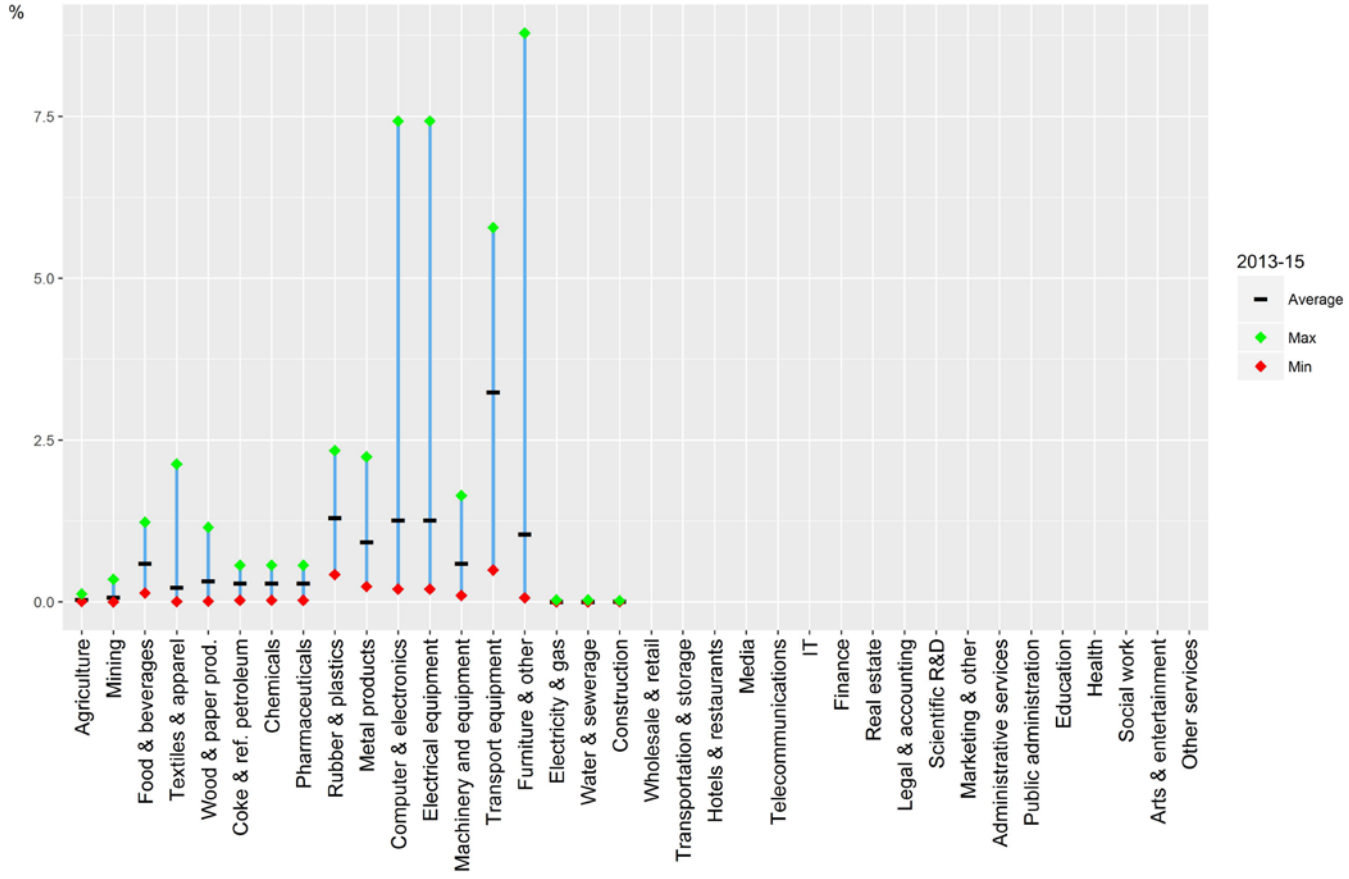
Focusing on the period 2013-15, the stock of robots per hundreds of employees results highest in the manufacturing of "Transport equipment", where there are more than 3 robots for every 100 employees on average. This corroborates anecdotal evidence suggesting an increase in the degree of robotisation of the sector. All other sectors considered exhibit less than half of the number of robots per hundreds of employees of the "Transport equipment" sector.

Focusing instead on the time variation in robots use intensity, automation, as measured by this indicator, has experienced significant increases over time in the "Transport equipment" sector, as well as "Rubber and plastics", and the "Furniture and other" manufacturing sector.

High cross-country heterogeneity in robot use intensity is especially evident in "Computer and electronics", "Electrical equipment", and "Furniture and other" manufacturing sectors (Figure 10). These sectors show a maximum value that is even higher than the one reported in the "Transport equipment" sector. Japan is the country that scores highest in these sectors.

Figure 10. Robots stock per hundreds of employees, by sector

Cross country dispersion, 2013-15



Note: Sector-specific values are unweighted averages over the years 2013-15, for each country ("Average"). "Max" ("Min") represent the country with highest (lowest) value of the indicator in the sector and period considered.
Source: IFR, OECD STAN, OECD Annual National Accounts, WIOD.

3.4. ICT specialists

The main measure of digital-related human capital embedded in production is the number of ICT specialists working in the different sectors. This ICT specialists indicator is measured as the proportion of workers in the sector who are employed in the following ISCO 2008 3-digit occupations: 251 (*Software and applications developers and analysts*), 252 (*Database and network professionals*), 133 (*Information and communications technology service managers*) and 351 (*Information and communications technology operations and user support*).[18] The selection of these occupations is the result of a thorough analysis based on five different methodologies, detailed in the Appendix (methodological annex).

Given the four ICT specialists occupations above, labour force surveys with occupational and sectoral information are used to compute the share of workers employed in ICT or digital intensive occupations per country - sector - year. This requires substantial work to (i) clean the underlying employment data, and (ii) homogenise sectoral and occupational classifications (ISIC and ISCO), especially as these classifications have changed during the period considered (i.e. from ISIC revision 3 to ISIC revision 4 in 2008; and from ISCO88 to ISCO08 in 2011). As the number of surveyed individuals in the occupation-sector cells can be small in some countries, data are pooled across countries and a single value calculated per each sector and year.

The descriptive evidence in Figure 11 aligns with expectations: in the period 2013-15 ICT specialists are mostly occupied in the ICT-producing sectors (both services and goods). Machinery manufacturing sectors and selected services sectors such as “Finance” and “Public administration” follow. This might reflect the fact that it is difficult to increase the number of ICT specialists quickly, given supply constraints, despite the dramatic speed of the digital transformation. Interestingly, the proportions of ICT specialists over total sectoral employment appear to have remained essentially unchanged over the two periods considered, except for “IT” services.

Box 1 below provides additional insights for an alternative measure that can be used to assess the human capital required to embed technology in production, focusing rather than on quantity – as the present indicator does – on the quality of labour.

Box 1 - ICT task intensity

Another way to assess the digital penetration of sectors relies on measuring how much of the labour force has the skills needed for the transformation to happen. Skills are needed to develop new technologies as well as to embed them in production. While fully defining which skills are most demanded in a digital world goes well beyond the scope of this study, it is possible to identify a set of skills which are related to the use of ICT at work and measure which occupations and/or sectors in the economy are most intensive in them.

Grundke et al. (2017) exploits the information reported in the Programme for the International Assessment of Adult Competencies (PIAAC) dataset to measure the frequency with which surveyed individuals carry out tasks which are related to the use of ICT on the job. This frequency is measured at the individual level, on the basis of the joint answer to the questions reported in Table A3 (see the Appendix). Each question is given a different weight based on state-of-the-art factor analysis, so that the retrieved frequency is a weighted average of the individual’s answers to these questions. The weights used correspond to the sampling weights reported for each individual in PIAAC. The indicator ranges between 0 and 100 per cent.

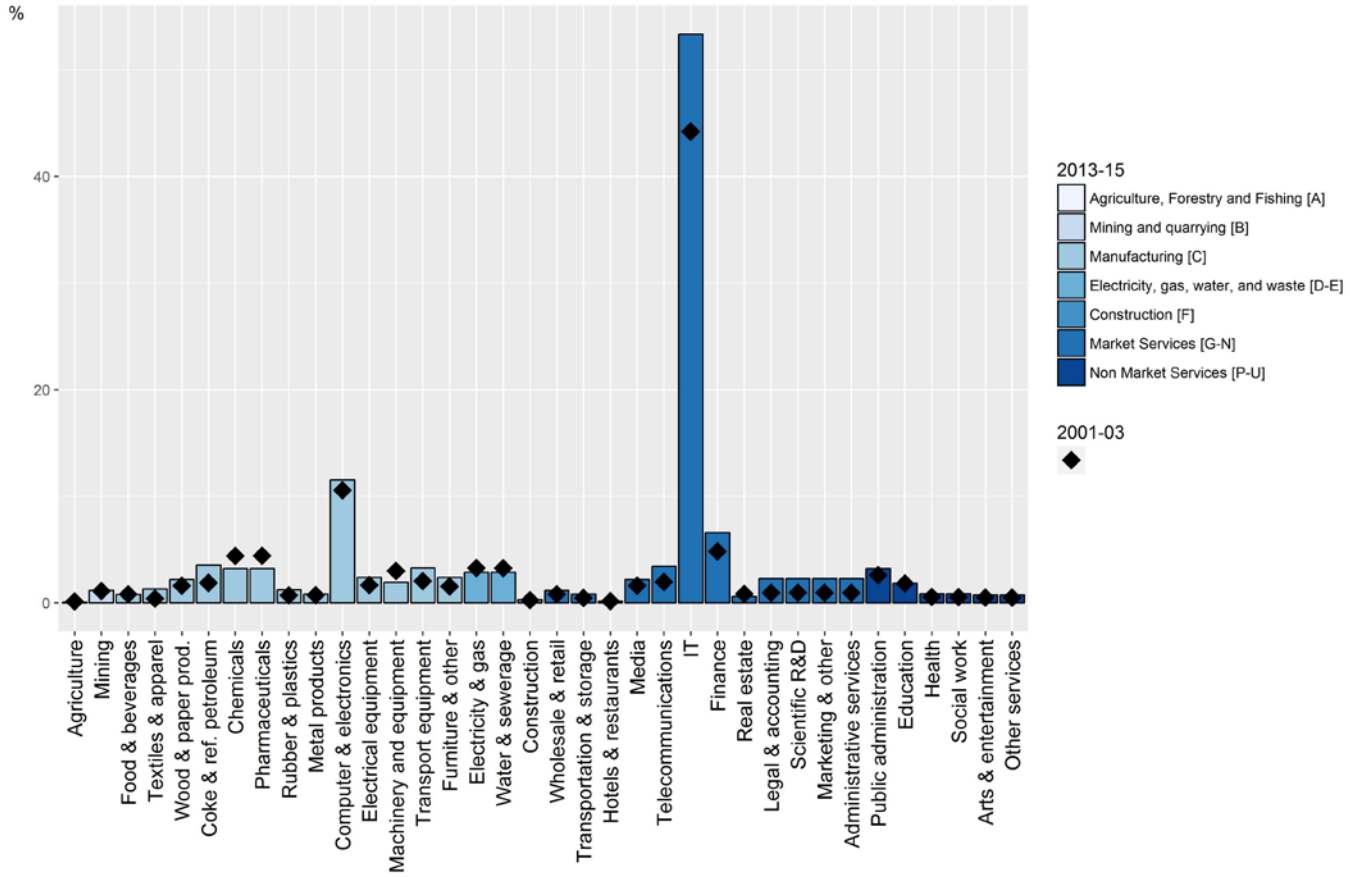
In light of the restricted number of individuals per country-sector reporting the complete set of information needed to estimate the task frequency, such a measure is not constructed at the country level. Furthermore, as only one wave of PIAAC has been implemented to date, this indicator is available for one cross-section only.

As a consequence, Figure 12 only shows the cross-country sector-specific values of this indicator for the year 2012. While the "IT" services sector is still positioned at the top, differences between manufacturing and services sectors appear less marked than for other indicators. In line with what could be expected, some services sectors are positioned at the bottom of the distribution, as is the case of "Hotels and restaurants" services, or "Transportation and storage" services. "Construction" also shows a very low value.

The absence of a time series constrains the use of this indicator for the purposes of the present study. In particular, the indicator could be used to enrich the taxonomy for the 2013-15 period, but not the one for 2001-03, thus hindering comparability over time.

Figure 11. ICT specialists as a % of total employment, by sector

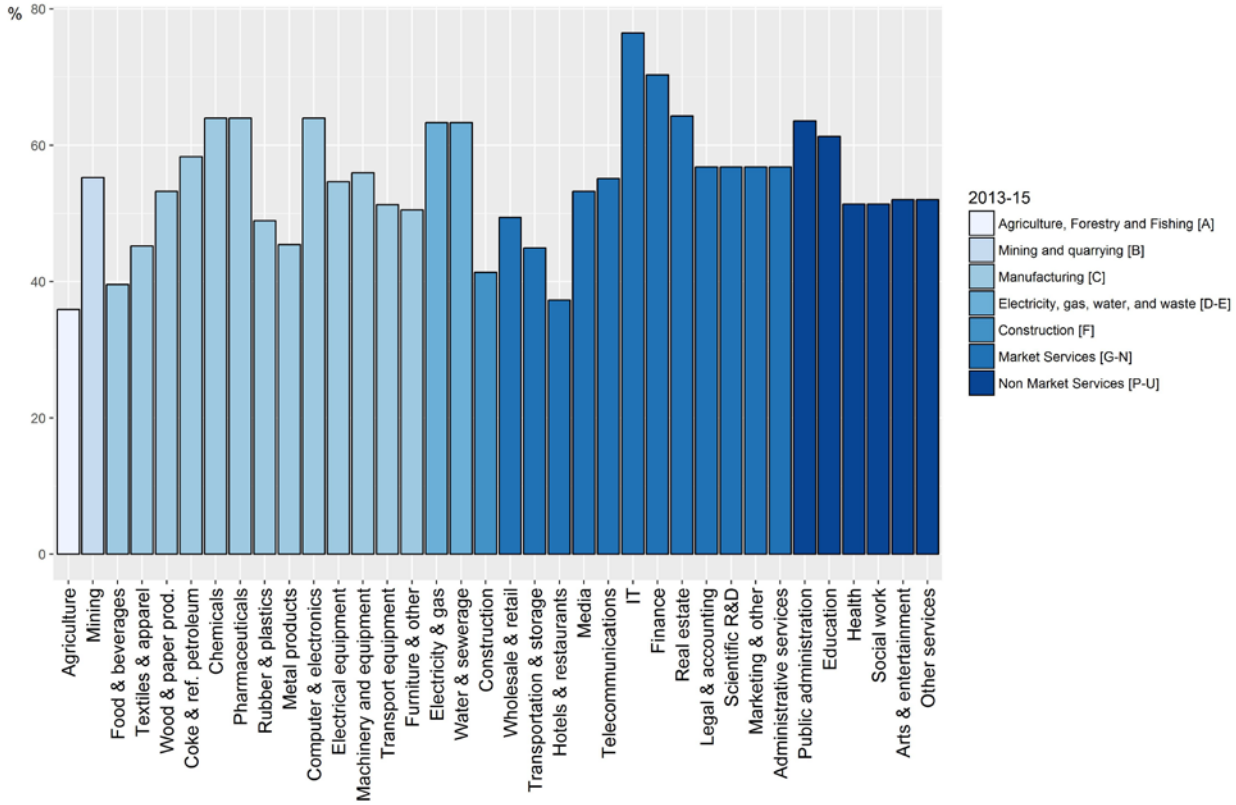
Aggregate across the considered countries, 2001-03 and 2013-15



Note: The sector's intensity is calculated as the number of workers employed in ICT specialist occupations, over total employment in the sector. Sector-specific values are weighted averages over the considered countries, where weights are equivalent to the country's total employment in the sector. The values are then averages over the years 2001-03 and 2013-15. Countries covered are: Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.
Source: Australian, Canadian, European, and Japanese Labour Force surveys; Japanese population census; U.S. Current Population Census.

Figure 12. ICT task intensity (%), by sector

Weighted averages across countries, 2012



Note: These values are calculated as a weighted average over individuals working in the sector across the following countries: Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom, and the United States.
Source: OECD Programme for the International Assessment of Adult Competencies (PIAAC) dataset.

3.5. Online sales

Nowadays a significant number of commercial transactions occur in electronic marketplaces (OECD, 2014), which creates opportunities for new types of consumer experiences and business models, and influences the dynamics of market competition.

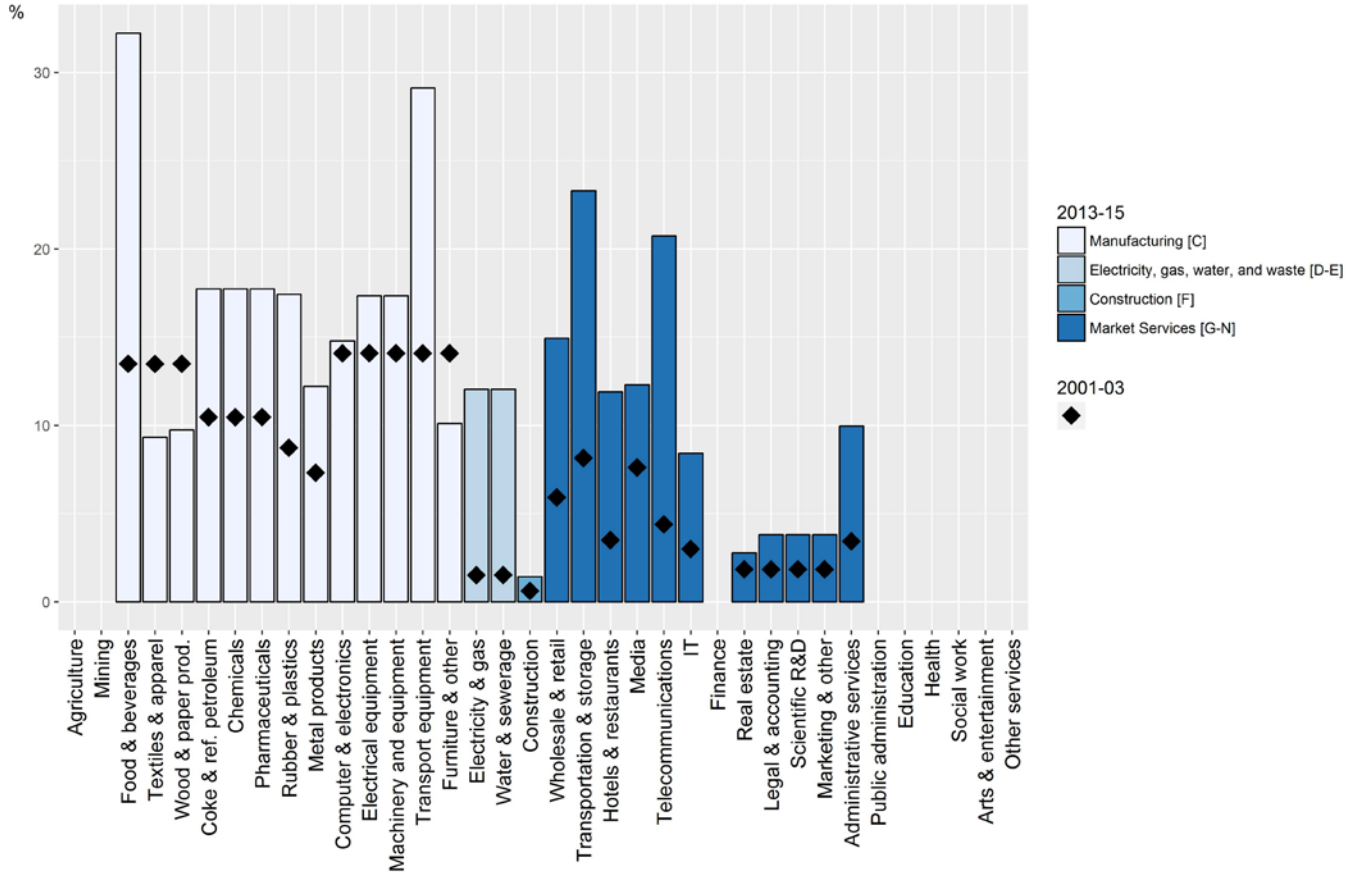
The diffusion of e-commerce as complement (or substitute) to more conventional commercial activities becomes thus an important aspect of the digital transformation.

To proxy such a feature of the digital transformation, data about the share of turnover from online sales complement other indicators in the taxonomy. This e-sales indicator captures the extent to which businesses operating in different sectors are confronted with more digital marketplaces.[19]

Data are sourced at the country-sector-year level from the Eurostat "Digital Economy and Society Statistics, Comprehensive Database" in both NACE revision 2 and NACE revision 1.1, then adapted to ISIC revision 4 classes for the entire period. Equivalent data for non-European countries at a sufficient level of sectoral disaggregation are not available at present.

Figure 13. Share (%) of turnover from online sales, by sector

Unweighted averages across countries, 2001-03 and 2013-15



Note: Sector-specific values are unweighted averages over the years 2001-03 and 2013-15 and across countries. Countries covered are: Austria, Denmark, Finland, France, Italy, the Netherlands, Norway, Sweden, and the United Kingdom.
Source: Eurostat.

Even within the European sample, data availability is heterogeneous across sectors and countries.[20] By sampling design, the indicator is not available for the following ISIC revision 4 sectors: sectors 1 to 9 ("Agriculture", "Mining"), 64 to 66 ("Finance"), 84 and above (public, social and personal services).

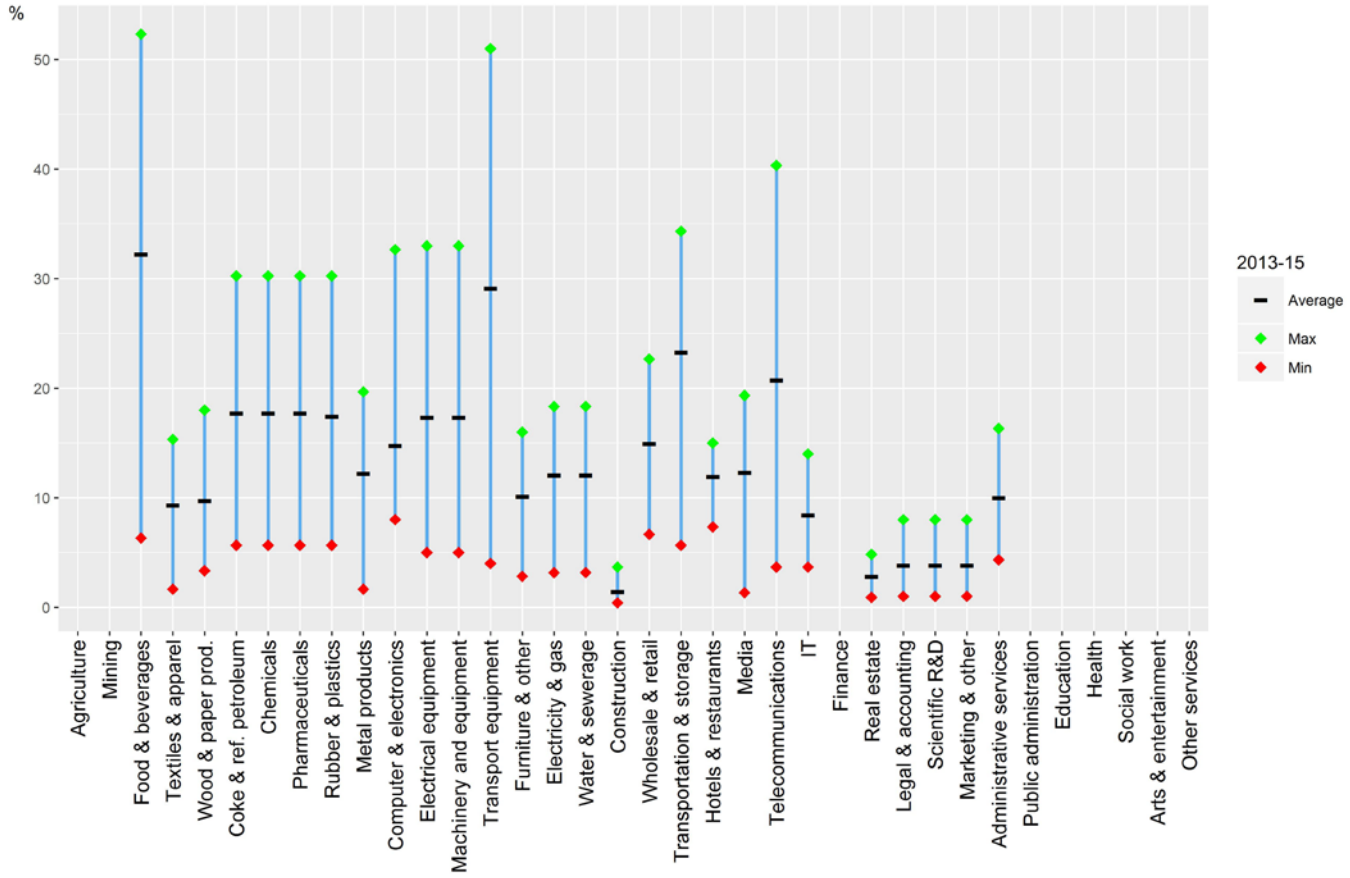
The statistics shown in Figure 13 provide a graphical visualisation of the share of turnover from online sales based on the subset of available countries that best matches the set of countries included for the computation of the other indicators in the taxonomy. The figure shows that, in the period 2013-15, the "Food and beverages" and the "Transport equipment" manufacturing sectors exhibit the highest shares of turnover from online sales, followed by the "Transportation and storage" and the "Telecommunications" service sectors. A lower value in "Wholesale and retail" seems indicative of the fact that a significant portion of turnover in this sector is still generated in non-electronic form.

Comparing the shares of turnover from online sales in the periods 2013-15 and 2001-03 suggests that most sectors witnessed an increase in the share of turnover from e-commerce. Similarly to other indicators, the largest increases seem to occur in sectors that reach the top of the distribution in 2013-15. Inference from the time variation in this indicator, however, should be derived with caution, in light of the data manipulations required to adjust sectoral classifications and aggregation levels as available in the original data.

Lastly, Figure 14 focuses on the cross-country variation underlying the 2013-15 averages. Sectors with the highest average intensity also display more significant dispersion in the share of revenues from online sales.

Figure 14. Share (%) of turnover from online sales, by sector

Cross country dispersion, 2013-15



Note: Sector-specific values are unweighted averages over the years 2013-15, for each country ("Average"). "Max" ("Min") represent the country with highest (lowest) value of the indicator in the sector and period considered..
Source: Eurostat.

4. Digital and less digital intensive sectors

4.1. Key stylised facts

A few stylised facts emerge from the descriptive statistics shown above.

First, and not surprisingly, ICT-intensive sectors emerge as the most digital intensive ones across most dimensions. This is true for "IT" services and mostly for "Telecommunication" services, while it is less so for their manufacturing counterpart, i.e. the computer producing sector. The sectors appearing amongst the least digital intensive, across all indicators considered, also align with anecdotal evidence. These are "Agriculture", "Real estate" and "Mining".

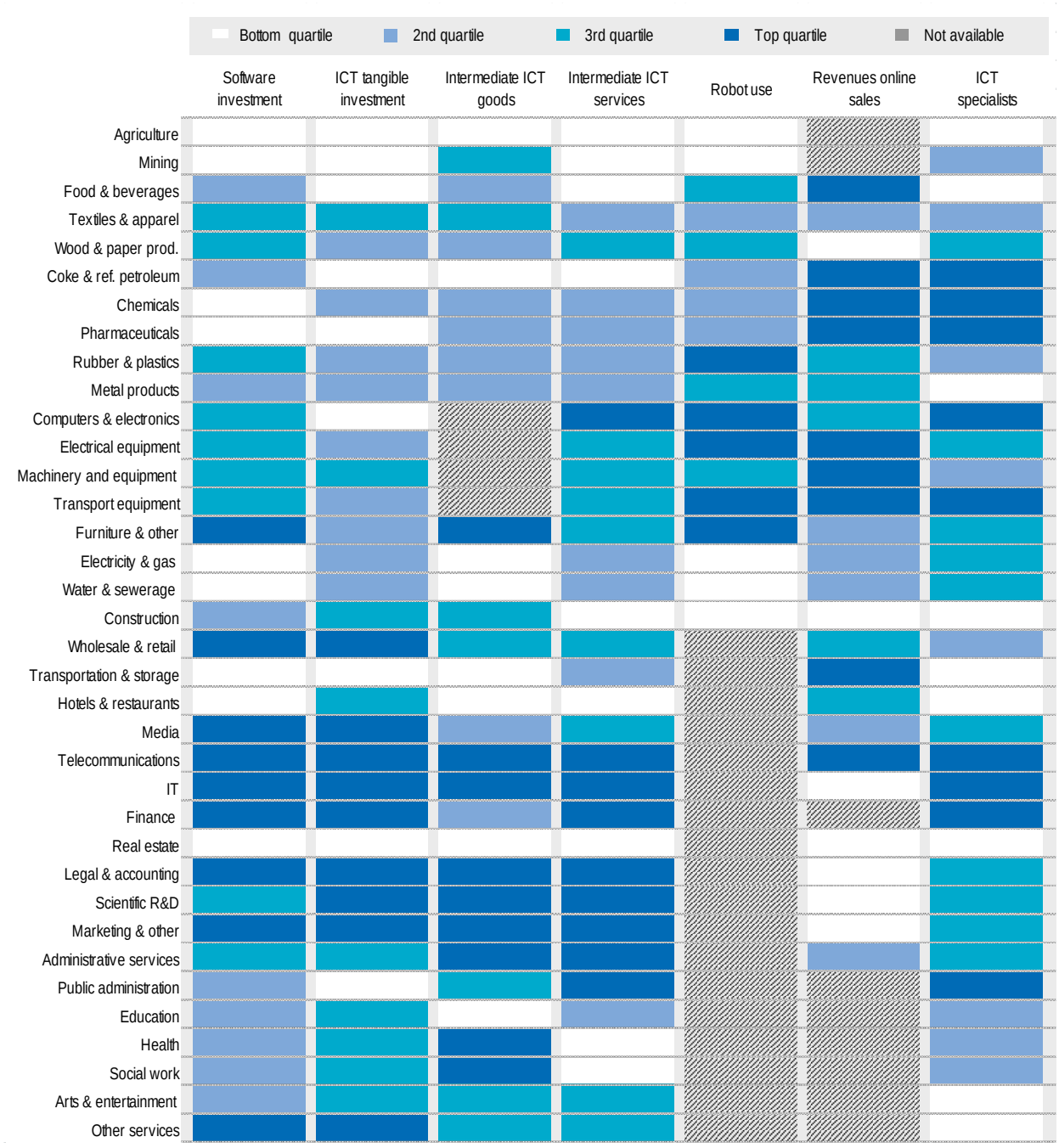
Table 1 summarises the classification of digital intensity according to all indicators considered. For each indicator the table reports whether a particular sector is reported as being relatively low (or high) in that particular dimension. In the table, if a sector is associated with a white cell in an indicator's column it will have the lowest level of digital intensity (bottom quartile) and the darker the colour, the more digital intensive the sector in that particular dimension (2nd and 3rd quartile) with the "blue" sectors being the most digital intensive (top quartile).

Second, sectors which are at the top (resp., bottom) of the classification in 2001-03 tend to remain at the top (resp., bottom) also at the end of the period considered. There are exceptions, of course, such as in the case of software investment intensity (the top sector in 2001-03 is "IT" services" while in 2013-15 it is "Finance).

These differences may reflect the possible "catching up" of some sectors which are accelerating their digital transformation, at least in some dimensions. In general, though, it remains true that displaying a higher digital intensity at the beginning of the period positively correlates with increasing intensity over the considered years. A simple correlation of the change in sectoral intensity and the intensity at the beginning of the period is reported in Figure 15. These positive associations remain positive and robust when controlling for sector and country fixed effects.[21]

Third, large heterogeneity across all indicators (and over time) emerges for sectors that are close to the median digital intensity in a particular indicator, i.e. sectors change their position significantly across indicators. Different sectors appear very much engaged in the digital transformation at different rates, depending on the considered dimension of digitalisation. This strengthens the importance of a multidimensional approach when classifying sectors according to their digital intensity, as sectors may be advancing in the digital transition according to one or more dimensions but not others. Table 2 shows that the rankings of the sectors obtained with different indicators are not necessarily correlated among each other. Table A2 in the Appendix instead shows the association among the intensities themselves, conditional on country and sector fixed effects, as obtained by estimating an OLS regression.[22]

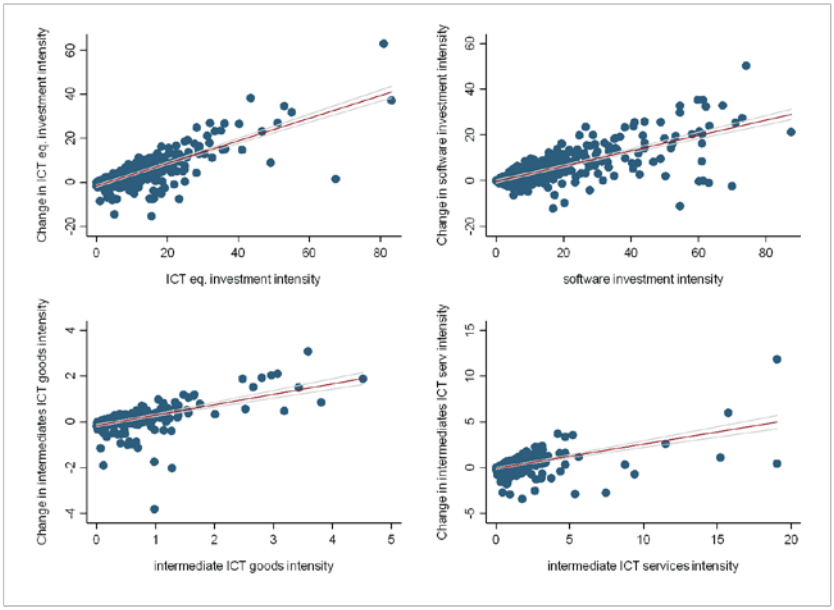
Table 1. Sectoral taxonomy of digital intensity, by indicator, 2013-15



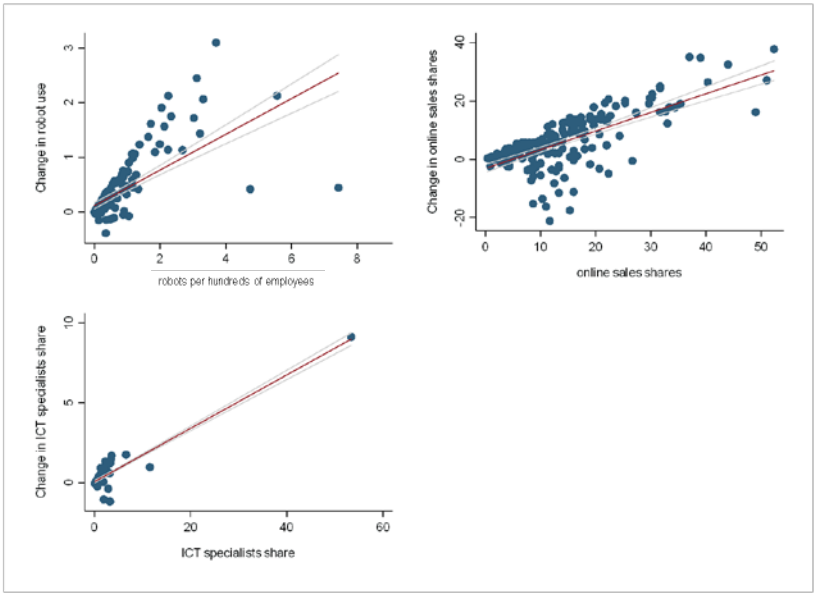
Note: All underlying indicators are expressed as sectoral intensities. For each indicator, sectoral values represent averages across countries and years (for the period 2013-15 only). The colour of the cells in the table corresponds to the quartile of the sectoral distribution to which the sector belongs.
Source: OECD Science, Technology and Industry Scoreboard 2017 (OECD, 2017c), based on based on Annual National Accounts, STAN, ICIO, PIAAC, International Federation of Robotics, World Bank, Eurostat Digital Economy and Society Statistics, national Labour Force Surveys, US CPS, INTAN-Invest and other national sources.

Figure 15. Correlation of initial intensity level and its change
Pooled sample of sector-country values

(a) Investment and Purchases of ICT intermediates



(b) Robot use, turnover from online sales and ICT specialists



Note: Pooled sample of sector-country values for the two periods considered in the taxonomy. Sector-country specific values are simple averages over the 3 years considered in each period. The confidence interval for the interpolation line is set to 95%. The "extreme" value of the "ICT specialists" sector is the "IT" services sector, and is not driven by an obvious outlier in the underlying sector-country distribution. Removing this data point, however, does not change the sign of the plotted correlation.
Source: Annual National Accounts, STAN, ICIO, PIAAC, International Federation of Robotics, World Bank, Eurostat Digital Economy and Society Statistics, national Labour Force Surveys, US CPS, INTAN-Invest and other national sources.

Table 2. Rank correlations across indicators

	ICT investment	Software investment	ICT intermediate goods	ICT intermediate services	Robot use	Online sales	ICT specialists
ICT investment	1						
Software investment	0.0463	1					
ICT intermediate goods	-0.0706	0.1151	1				
ICT intermediate services	0.1812*	-0.0338	0.0703	1			
Robot use	0.1535*	0.3227*	0.0052	0.0523	1		
Online sales	0.0728	0.2673*	-0.0853	0.1807*	0.2639*	1	
ICT specialists	-0.064	-0.1152	-0.105	0.3013*	-0.1635*	0.2113*	1

Note: Pooled sample of sector-country values for the two periods considered in the taxonomy. Sector-country specific values are simple averages over the 3 years considered in each period. Starred cells highlight rank correlations which are significant at least at 10%.
Source: Annual National Accounts, STAN, ICIO, PIAAC, International Federation of Robotics, World Bank, Eurostat Digital Economy and Society Statistics, national Labour Force Surveys, US CPS, INTAN-Invest and other national sources.

4.2. A "global" taxonomy

As previously stated, the taxonomy proposed here is meant to be an operational tool for analysis and policy design, as having different classifications of sectors depending on the dimension considered complicates matters. This so-called "global" taxonomy of sectors attempts to summarise the different classifications and exploits information from each and every one of the indicators considered thus far.

In particular, the position of a given sector *j* in the summary (hereafter "global") classification is computed as an average of sector *j*'s position in each considered dimension of the digital transformation, where the weight is 1/36, i.e. the total number of sectors included in the taxonomy. These sector and indicator-specific scores (i.e. a sector's position divided by 36) are then averaged across indicators, to yield one value per sector.[23] The classification of sectors by these values is the mentioned "global" taxonomy. It should be noted that the values themselves (i.e. the average of the scores across indicators) have no interpretation by themselves, and are only useful for classification purposes.[24]

Table 3 shows which sectors fall into which quartile of the sectoral distribution, based on the "global" taxonomy for the periods 2001-03 and 2013-15. There are of course important similarities with the classification based on the underlying indicators, as reported in Table 1. "Agriculture", "Mining" and "Real estate" are positioned in the bottom quartile, while "IT" services, "Telecommunications" and "Finance" appear at the top. Few sectors change quartile from the initial to the final period, but the exact positions display greater dynamism: differences in the two periods are more prominent when looking at the positions themselves, rather than "only" at the quartiles.

The "global" taxonomy underpinning Table 3 is obtained using one of many possible methodologies to aggregate the information provided by different indicators of digitalisation. While useful to propose a synthetic view of the multifaceted phenomenon at hand, this "global" taxonomy, being at the sectoral level, possibly hides the within-sector heterogeneity that may exist. What is more, it falls short of accounting for differences in magnitudes between sectors in the underlying indicators. For example,

while "Finance" is always positioned above "Agriculture", the "global" taxonomy does not reflect the fact that the value for "Finance" is 10 times that of "Agriculture" for one indicator, but only 5 times for another indicator.[25]

Table 3. Sectoral taxonomy of digital intensity: "global" indicator

Sector denomination	ISIC rev.4	Quartile of digital intensity: 2001-03	Quartile of digital intensity: 2013-15
Agriculture, forestry, fishing	01-03	Low	Low
Mining and quarrying	05-09	Low	Low
Food products, beverages and tobacco	10-12	Low	Low
Textiles, wearing apparel, leather	13-15	Medium-low	Medium-low
Wood and paper products, and printing	16-18	Medium-high	Medium-high
Coke and refined petroleum products	19	Medium-low	Medium-low
Chemicals and chemical products	20	Medium-low	Medium-low
Pharmaceutical products	21	Medium-low	Medium-low
Rubber and plastics products	22-23	Medium-low	Medium-low
Basic metals and fabricated metal products	24-25	Medium-low	Medium-low
Computer, electronic and optical products	26	High	Medium-high
Electrical equipment	27	Medium-high	Medium-high
Machinery and equipment n.e.c.	28	High	Medium-high
Transport equipment	29-30	High	High
Furniture; other manufacturing; repairs of computers	31-33	Medium-high	Medium-high
Electricity, gas, steam and air cond.	35	Low	Low
Water supply; sewerage, waste management	36-39	Low	Low
Construction	41-43	Low	Low
Wholesale and retail trade, repair	45-47	Medium-high	Medium-high
Transportation and storage	49-53	Low	Low
Accommodation and food service activities	55-56	Low	Low
Publishing, audiovisual and broadcasting	58-60	Medium-high	Medium-high
Telecommunications	61	High	High
IT and other information services	62-63	High	High
Finance and insurance	64-66	High	High
Real estate	68	Low	Low
Legal and accounting activities, etc.	69-71	High	High
Scientific research and development	72	Medium-high	High
Advertising and market research; other business services	73-75	High	High
Administrative and support service activities	77-82	High	High
Public administration and defence	84	Medium-high	Medium-high
Education	85	Medium-low	Medium-low
Human health activities	86	Medium-high	Medium-low
Residential care and social work activities	87-88	Medium-low	Medium-low
Arts, entertainment and recreation	90-93	Medium-low	Medium-high
Other service activities	94-96	Medium-high	High

Note: "High" identifies sectors in the top quartile of the distribution of the values underpinning the "global" taxonomy, "medium-high" the second highest quartile, "medium-low" the second lowest, and "low" the bottom quartile.

A first test aimed at assessing the robustness of the "global" taxonomy relative to its underlying indicators is thus proposed. Table 4 reports the results of an econometric estimation where the position of a sector in the "global" taxonomy is correlated to the sector's intensity in each underlying dimension of digitalisation. The tests considers whether the sector is above or below the "global" median (columns "i"), in the top quartile of the distribution ("high digital intensive", columns "ii"), or in the bottom quartile ("low digital intensive", columns "iii").

Table 4. Correlations between the "global" taxonomy and single dimensions of the digital transformation

	ICT equipment investment			Software investment			ICT intermediate goods			ICT intermediate services		
	i	ii	iii	i	ii	iii	i	ii	iii	i	ii	iii
1 if digital intensive	0.660***			0.923***			0.612***			0.687***		
	(0.0598)			(0.0570)			(0.0591)			(0.0640)		
1 if high digital intensive		0.782***			1.127***			0.727***			1.036***	
		(0.0690)			(0.0647)			(0.0760)			(0.0704)	
1 if low digital intensive			-0.530***			-0.773***			-0.535***			-0.519***
			(0.0715)			(0.0704)			(0.0669)			(0.0766)
Observations	860	860	860	860	860	860	762	762	762	860	860	860
R-squared	0.252	0.257	0.196	0.321	0.346	0.220	0.366	0.354	0.331	0.144	0.227	0.076

	Robot use			Online sales			ICT specialists		
	i	ii	Iii	i	ii	iii	i	ii	iii
1 if digital intensive	0.709***			-0.0952			0.469***		
	(0.0903)			(0.0823)			(0.0668)		
1 if high digital intensive		0.921***			-0.0714			0.906***	
		(0.114)			(0.0892)			(0.0731)	
1 if low digital intensive			-0.558***			-0.145			-0.368***
			(0.0927)			(0.0970)			(0.0782)
Observations	428	428	428	469	469	469	860	860	860
R-squared	0.271	0.277	0.229	0.242	0.241	0.243	0.055	0.154	0.026

Note: Estimates obtained with OLS regressions where the dependent variable is reported as a column header, and explanatory variables are the variable in the row header, and country, year and country-year dummies. "1 if digital intensive" is a dummy variable with value 1 if the sector is above the median of the distribution of sectors for the indicator in the column header. "1 if high digital intensive" is the same, but for sectors positioned in the top quartile by indicator intensity; "1 if low digital intensive" is the same, but for sectors positioned in the bottom quartile by indicator intensity. The classification of sectors by "global" digital intensity here used is defined on initial period information (2001-03). All dependent variables are standardised. Estimates are based on the pooled sample of sector-country values for the two periods considered in the taxonomy. Sector-country specific values are simple averages over the 3 years considered in each period. ***p<0.01,**p<0.05,*p<0.1

The estimates suggest that the "global" taxonomy is performing reasonably well in approximating the sectors' intensities in the different dimensions of the digital transformation. Conditional on other country and period correlates, switching from a less

digital intensive to a digital intensive sector is associated to an increase in the sector's intensity of ICT investment, intermediate consumption of ICT products, robot use and ICT specialists. Associations are stronger in the case of the sectors at the top quartile of the "global" intensity distribution, relative to other quartiles. Conversely, sectors at the bottom of the distribution display lower intensities in all indicators, relative to sectors in other quartiles. Associations with a sector's intensity in e-sales are not statistically different from zero.

5. Preliminary conclusions and next steps

This paper presents a taxonomy mirroring the extent to which sectors have gone digital, based on intensities in various dimensions of the digital transformation. It takes into account how companies in the digital era invest in the most advanced digital assets, whether they embed such digital technologies in production by means of relying on the necessary human capital, and whether they use them to enhance their businesses.

For each of the seven dimensions of the digital transformation considered, the study discusses the rationale for considering them, the methodological challenges and data constraints faced, and the heterogeneity in the relevance of these dimensions across sectors and countries. While the different dimensions are found to be only partially correlated among each other, some sectors are positioned consistently as being the most digital intensive, across different indicators. Those sectors that are close to the median level if digital intensity display significant heterogeneity across indicators. Both aspects also characterise the "global" digital intensity taxonomy, which averages the classifications along the different dimensions of digitalisation.

Far from providing a full picture of such a complex phenomenon and how it is reflected in production, the taxonomy is meant as an operational tool to assess how sectors characterised by different digital intensities are evolving, and whether the extent to which a sector is digital intensive is correlated with any economic dimension of interest. As such, the taxonomy can complement other established taxonomies, such as those based on ICT use or R&D intensities.

Future work could increase the number of indicators underpinning the taxonomy, conditional on data availability, in particular by including: (i) a measure of intensity in ICT-related innovative output, as proxied by ICT-related patent families filed by companies operating in the sector; (ii) a measure of ICT "embodied" investment and/or value added, as already discussed in this study; or (iii) a measure of robots or automation in the services sectors.

Lastly, the taxonomy would benefit of an indicator of the value of data used by companies in production. Collection and analysis of data can represent an increasingly important source of value for digital companies. While the number of firms using big data can be sourced for European countries for 2016 in the Eurostat Digital Economy and Society Statistics, Comprehensive Database, a complete time series across sectors and countries is not currently available, to the best of the authors' knowledge. Similar data constraints, as well as definitional and conceptual issues, characterise the measurement of *value* of data produced or used by companies, even at the aggregate and sector-country level. These constraints are the object of prospective work at the OECD, which aims at the development of new measures of the digital economy.

Notes

[1] OECD STI/PBD and STI/EAS. In the data collection and elaboration process, this study benefitted from the contributions of: Alessandra Colecchia, Timothy DeStefano, Hélène Dernis, Robert Grundke, William Haines, Peter Horvát, Giorgio Presidente, Jan Tscheke, Norihiko Yamano, Colin Webb, and Belen Zinni. We are also grateful to Nadim Ahmad, Carol Corrado, Elif Koksal-Oudot, Molly Lesher, Pierre Montagnier, Paul Schreyer, Vincenzo Spiezia, and participants to the ESCoE conference for helpful feedback and suggestions.

[2] This would not have been possible without the contributions of the large number of OECD colleagues mentioned in the acknowledgements.

[3] See UN Stats: https://unstats.un.org/unsd/cr/registry/isic-4.asp.

[4] The indicator for Revenues of online sales is an exception. A sufficiently populated time series by country and sector (at the required sectoral disaggregation level) for online sales is available for European countries only at this stage.

[5] These data are complemented with EU-KLEMS data (see Jäger, 2016) for two countries.

[6] "Non-residential" refers to the fact that dwellings are excluded from the computation of gross fixed capital formation. For robustness, an alternative version of these indicators was calculated using a different denominator, the sum of investment in machineries and equipment and investment in software and databases.

[7] The choice of non-residential GFCF as denominator is not void of consequences. For example a manufacturing firm with higher overall investment but the same level of investment in ICT as a services firm of equal output size will display much lower intensity, despite the same overall ICT investment. An argument in favour of non-residential GFCF is developed in the Appendix (methodological annex).

[8] In Korea in 2014, for instance, GFCF in "semiconductor manufacturing machinery" and "display board manufacturing machinery", all non-ICT equipment, exceeded GFCF in computers and telecommunication equipment.

[9] Note that the "Other services" sector, another sector that scores relatively high, includes activities associated with the repair of computers.

[10] The only other study which exploits a similar measure of digitalisation is, to our knowledge, van Ark et al. (2016). Their study, however, covers only 3 countries (Germany, United Kingdom, United States).

[11] Note that this does not exclude possible complementarities between ICT investment (either tangible or software) and ICT intermediate purchases. Assessing the existence of such complementarities, however, falls outside the scope of the present paper.

[12] A second indicator, which is not exploited in this study, considers also the purchases from ISIC revision 3 sector 64 "Post and Telecommunication". Ideally, only telecom services would fit the purpose of this study. However distinguishing postal from telecom services is not easily done for data expressed in ISIC revision 3. So as to avoid overestimating purchases of ICT services, a lower-bound figure including only services from sector 72 is preferred here. More precise figures could be sources from Supply and Use Tables, as pointed out below.

[13] This sectoral category was created ad-hoc for the ICIO and Trade in Value Added database, so as to smooth the conversion from ISIC revision 3 to ISIC revision 4. Isolating intermediates produced by sector 30 ("Manufacture of office, accounting and computing machinery") from those of the combined 30, 32 and 33 sector is left as possible future refinement of the presented methodology, and in particular to the use of Supply and Use Tables (see below).

[14] For purchases of both ICT services and goods, Supply and Use Tables (SUTs) would have provided a more precise source of information than the ICIO, as (i) they do not convert supplied products into sectors but rather report sales of precisely defined ICT intermediate goods and services, and (ii) they report purchases by sectors which are expressed in the latest sectoral classification, i.e. ISIC revision 4. However, a complete time series of purchases in ISIC revision 4 covering all years and countries of interest is not available in the SUTs at the time of writing, which limits the use of this source for the present study. Nevertheless, a robustness exercise was conducted, whereby purchases from the SUTs were compared to purchases as sourced from the ICIO, in volumes, for the available countries and the years where no data estimation was necessary (2010 and 2011). Correlations reach 88% (pairwise) and 81% (rank) for purchases of ICT goods, and 94% (pairwise) and 75% (rank) for purchases of ICT services.

[15] A similar assumption could be made for ICT services, too, and especially for the computer producing sector, which may purchase software to then sell it in a bundle with computing equipment.

[16] The price series used here should deflate only the domestic purchases, while imported intermediates should be deflated by the price of imports of ICT goods and services in the purchasing country. This is especially important for products, such as computers, where domestic companies supply only a small share of intermediates. However, at the time of writing, import prices were not available for a sufficiently large number of countries and years, hence this venue could not be pursued. The current deflation therefore assumes that import prices of ICT intermediates and domestic ICT output prices are the same.

[17] Ongoing work is exploring the dynamics of a different indicator of automation, i.e. the share of enterprises whose business processes are automatically linked to suppliers or customers. This can be sourced from the Eurostat ICT usage and e-commerce in enterprises surveys, and is also available for services sectors, showing rather good correlation with the robot indicator in manufacturing.

[18] The occupation 352 (*Telecommunications and broadcasting technicians*), which is additionally included in the OECD-Eurostat definition of ICT specialists (OECD 2015a), was not included here, because workers in this occupation are mostly working with moving pictures technology and are not usually involved in the production of ICT goods and services. Investigating the existence of a natural threshold among the top ICT specialist occupations also excludes occupation 352 (see here further below).

[19] Some caution should be applied when using online sales as a proxy of digital intensity of sectors. Firstly, higher levels of fragmentation of production or outsourcing in the sector could

result in higher intensities, to the extent that these business-to-business transactions are carried out through the internet. Secondly, some countries may display higher values of the indicator, since cultural factors may still influence the propensity of economic agents to sell and purchase online. However, this would affect the taxonomy only if cultural factors influenced the propensity to transact online differently across sectors within the same country. Lastly, this indicator of e-commerce does not measure online transactions of data or their value.

[20] The methodological annex in the Appendix discusses how this issue is addressed, and provides other information about the data source.

[21] These associations are estimated as OLS regressions of the change over time in the indicator (for each country-sector pair) on the initial level of the indicator. The correlation remains positive and statistically significant even when considering percentage changes. The regression for ICT specialists only controls for sector fixed effects, as one single value for all countries is available. In the ICT specialist's graph, the top values are for the "IT" services sector.

[22] The estimation by OLS regressions allows the correlations to have magnitude higher than one, contrary to Table 2.

[23] As a consequence, being positioned as the 4th top sector by intensity in a given dimension of digitalisation always yields a value equal to 4/36 = 0.11. In computing the cross-indicator average, however, we cannot correct for missing values in a particular indicator, therefore sector-year cells for which a particular indicator is missing translate in missing values in the indicator specific taxonomy and are reported as grey (not available) in Table 1. In the calculation of the global taxonomy indicator these observations are not taken into account when deriving the average.

[24] Sectors with the same value for an indicator are attributed the same position and have the same score as a consequence.

[25] This becomes more important when the relative position of two sectors is inverted across two indicators. A different aggregating methodology could take into account the fact that e.g. sector (a) has 10 times the value of sector (b) for one indicator, while sector (b) has 5 times the value of sector (a) for another indicator.

References

Acemoglu, D. and P. Restrepo (2016), "The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment", *NBER Working Paper*, No. 22252, National Bureau of Economic Research, Cambridge, MA, http://dx.doi.org/10.3386/w22252.

Ahmad, N. (2003), "Measuring Investment in Software", *OECD Science, Technology and Industry Working Papers*, No. 2003/06, OECD Publishing, Paris, http://dx.doi.org/10.1787/335303788330.

Andrews, D., C. Criscuolo and P. Gal (2016), "The best vs the rest: the global productivity slowdown, divergence across firms and the role of public policy", *OECD Productivity Working Papers*, No. 5/2016, OECD Publishing, Paris, http://dx.doi.org/10.1787/63629cc9-en.

Colecchia, A. and P. Schreyer (2002), "ICT Investment and Economic Growth in the 1990s: Is the United States a Unique Case? A Comparative Study of Nine OECD Countries", *Review of Economic Dynamics*, Vol. 5/2, Elsevier Science, USA, pp. 408-442, https://doi.org/10.1006/redy.2002.0170.

Comin, D and M. Mestieri (2013), "If Technology Has Arrived Everywhere, Why Has Income Diverged?" *NBER Working Paper*, No. 19010. National Bureau of Economic Research, Cambridge, MA, http://dx.doi.org/10.3386/w19010.

Corrado, C., Haskel, J., Jona-Lasinio, C. and Iommi, M. (2012), "Intangible Capital and Growth in Advanced Economies: Measurement Methods and Comparative Results", *IZA Discussion Paper*, No. 6733, IZA, Bonn.

Dauth, W., S. Findeisen, J. Suedekum and N. Woessner (2017), "German robots – The impact of industrial robots on workers", *CEPR Discussion Paper*, No. 12306, Centre for Economic Policy Research, London.

De Backer, K. *et al.* (2018), "Industrial robotics and the global organisation of production", *OECD Science, Technology and Industry Working Papers*, No. 2018/03, OECD Publishing, Paris, http://dx.doi.org/10.1787/dd98ff58-en.

Galindo-Rueda, F. and F. Verger (2016), "OECD Taxonomy of Economic Activities Based on R&D Intensity", *OECD Science, Technology and Industry Working Papers*, No. 2016/04, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jlv73sqqp8r-en.

Graetz, G. and G. Michaels (2015), "Robots at Work", *CEP Discussion Paper*, No. 1335, Centre for Economic Performance, London School of Economics, London.

Grundke, R., S. Jamet, M. Kalamova, F. Keslair and M. Squicciarini (2017), "Skills and Global Value Chains: A Characterisation", *OECD Science, Technology and Industry Working Papers*, No. 2017/05, OECD Publishing, Paris, http://dx.doi.org/10.1787/cdb5de9b-en.

Hirsh-Kreinsen, H., K. Hahn and D. Jacobsen (2008), "The Low-tech Issue", in H. Hirsch-Kreinsen, D. Jacobsen (eds.), *Innovation in Low-Tech Firms and Industries*, Edward Elgar, Cheltenham, pp. 3-24.

Jäger, K. (2016), "EU KLEMS Growth and Productivity Accounts 2016 release - Description of Methodology and General Notes", available at http://www.euklems.net/.

McKinsey Global Institute (2015), "Digital America: A tale of the haves and have-mores", *McKinsey Global Institute Report*, McKinsey & Company.

OECD (2018), "ICT investment" (indicator), http://dx.doi.org/10.1787/b23ec1da-en (accessed on 06 March 2018).

OECD (2017a), *The Next Production Revolution: Implications for Governments and Business*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264271036-en.

OECD (2017b), "Determinants and impact of automation: an analysis of robots' adoption in OECD countries", internal document.

OECD (2017c), *OECD Science, Technology and Industry Scoreboard 2017: The digital transformation*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264268821-en.

OECD (2015a), "Proposal for a EUROSTAT-OECD Definition of ICT Specialists", internal document.

OECD (2015b), "New skills for the digital economy: Measuring the demand and supply of ICT skills at work", *OECD Digital Economy Papers*, No.258, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jlwnkm2fc9x-en.

OECD (2014), *Measuring the Digital Economy: A New Perspective*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264221796-en.

OECD (2011a), "Innovative sectors", in *OECD Science, Technology and Industry Scoreboard 2011*, OECD Publishing, Paris, http://dx.doi.org/10.1787/sti_scoreboard-2011-66-en.

OECD (2011b), *OECD Guide to Measuring the Information Society 2011*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264113541-en.

OECD (2009a), *Innovation in Firms: A Microeconomic Perspective*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264056213-en.

OECD (2009b), *Handbook on Deriving Capital Measures of Intellectual Property Products*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264079205-en.

OECD (2004), *The Economic Impact of ICT. Measurement, Evidence and Implications*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264026780-en.

OECD (2003), *OECD Science, Technology and Industry Scoreboard 2003*, OECD Publishing, Paris, http://dx.doi.org/10.1787/sti_scoreboard-2003-en.

OECD (2002), *Measuring the Information Economy 2002*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264099012-en.

Pilat, D. and F. Lee (2001), "Productivity Growth in ICT-producing and ICT-using Industries: A Source of Growth Differentials in the OECD?", *OECD Science, Technology and Industry Working Papers*, No. 2001/04, OECD Publishing, Paris, http://dx.doi.org/10.1787/774576300121.

Solow, R. (1987), "We'd better watch out", *New York Times Book Review*, New York Times, New York, July 1987.

Spiezia, V. (2011), "Are ICT Users More Innovative? An Analysis of ICT-Enabled Innovation in OECD Firms", *OECD Journal: Economic Studies*, Vol. 2011/1, http://dx.doi.org/10.1787/eco_studies-2011-5kg2d2hkn6vg.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer and G. J. de Vries (2015), "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production", *Review of International Economics*, Vol. 23/3, John Wiley & Sons Ltd, pp. 575–605, https://doi.org/10.1111/roie.12178.

Tukey, J. W. (1977), *Exploratory Data Analysis*, Addison-Wesley, Reading, MA.

van Ark, B., A. Erumban, C. Corrado and G. Levanon (2016), "Navigating the New Digital Economy: Driving Digital Growth and Productivity from Installation to Deployment", *Conference Board Report*, The Conference Board, New York.

Appendix

Additional tables

Table A1. List of considered sectoral aggregations

Sector denomination	*Short sector label (in Figures)*	*ISIC rev.4 code*
Agriculture, forestry, fishing	Agriculture	01-03
Mining and quarrying	Mining	05-09
Food products, beverages and tobacco	Food & beverages	10-12
Textiles, wearing apparel, leather	Textiles & apparel	13-15
Wood and paper products, and printing	Wood & paper prod.	16-18
Coke and refined petroleum products	Coke & ref. petroleum	19
Chemicals and chemical products	Chemicals	20
Pharmaceutical products	Pharmaceuticals	21
Rubber and plastics products	Rubber & plastics	22-23
Basic metals and fabricated metal products	Metal products	24-25
Computer, electronic and optical products	Computer & electronics	26
Electrical equipment	Electrical equipment	27
Machinery and equipment n.e.c.	Machinery and equipment	28
Transport equipment	Transport equipment	29-30
Furniture; other manufacturing; repairs of computers	Furniture & other	31-33
Electricity, gas, steam and air cond.	Electricity & gas	35
Water supply; sewerage, waste management	Water & sewerage	36-39
Construction	Construction	41-43
Wholesale and retail trade, repair	Wholesale & retail	45-47
Transportation and storage	Transportation & storage	49-53
Accommodation and food service activities	Hotels & restaurants	55-56
Publishing, audiovisual and broadcasting	Media	58-60
Telecommunications	Telecommunications	61
IT and other information services	IT	62-63
Finance and insurance	Finance	64-66
Real estate	Real estate	68
Legal and accounting activities, etc.	Legal & accounting	69-71
Scientific research and development	Scientific R&D	72
Advertising and market research; other business services	Marketing & other	73-75
Administrative and support service activities	Administrative services	77-82
Public administration and defence	Public administration	84
Education	Education	85
Human health activities	Health	86
Residential care and social work activities	Social work	87-88
Arts, entertainment and recreation	Arts & entertainment	90-93
Other service activities	Other services	94-96

Table A2. Conditional correlations across indicators

	ICT investment	Software investment	ICT intermediate goods	ICT intermediate services	Robot use	Online sales	ICT specialists
ICT investment	1	0.0659**	-0.0140	-0.001	0.0433	0.0934*	0.0287***
Software investment	0.0942**	1	-0.0123	0.0523	0.0682	0.189***	0.0306***
ICT intermediate goods	-0.0173	-0.0105	1	0.0805**	0.008	-0.0439	0.0048
ICT intermediate services	-0.001	0.0629	0.109**	1	0.005	0.105*	0.0294***
Robot use	0.0133	0.0169	0.0104	0.000464	1	0.690***	0.010***
Online sales	0.0921*	0.140***	-0.0617	0.0670*	0.261***	1	0.0268***
ICT specialists	1.275***	0.951***	0.168	0.760***	2.085***	0.861***	1

Note: Estimates obtained with OLS regressions where the dependent variable is reported as a column header, and explanatory variables are the variable in the row header, and sector and country dummies. Each cell signals a different regression, pooling the sample of sector-country values for the two periods considered in the taxonomy. Sector-country specific values are simple averages over the 3 years considered in each period. *** p<0.01, ** p<0.05, * p<0.1.

Methodological Annex

ICT investment

The main challenges faced when constructing ICT investment-related indicators depend on i) the availability of time series at the required level of sectoral aggregation, and ii) the deflation of ICT assets.

Investment data are sourced at the ISIC revision 4 SNA A38 disaggregation level (excluding "Private households with employed persons" and "Extraterritorial organisations"), and are available in the detailed table 8A of the OECD Annual National Accounts Database. Missing values of each indicator are updated with values computed at the STAN a21 level of aggregation, for the relevant A38 sectors. Information for Germany and Spain is complemented with data from EU KLEMS (Jäger, 2016).

Extensive discussions about deflation challenges are not the purpose of this work (see for instance Colecchia and Schreyer, 2002; or OECD, 2004). Some key methodological choices are made here, keeping in mind the main purpose of this study, i.e. comparing the digital intensity of sectors.

Country-specific implicit deflators are retrieved for ICT tangible and non-residential gross fixed capital formation series from time series in current and previous year prices. When the latter information is missing, this is backed up from investment in volumes and current prices, if available. Implicit deflators may in some case underestimate the price decline of ICT equipment as they may not necessarily be constructed with a hedonic methodology. However, unreported analysis shows good correlations with alternative harmonised deflator series produced by the OECD. Country-specific deflators are applied to the software and databases investment series, based on Corrado et al. (2012), and extrapolated using information provided by the OECD Statistics Directorate.

Some interpolations/extrapolations are used to get the maximum possible coverage for the deflators. A basic data cleaning puts few implausible values to missing. The final coverage for this indicator includes at most 36 2-digit sectors for the following countries:

Australia, Austria, Belgium, Canada, Czech Republic, Germany, Denmark, Estonia, Finland, France, Greece, Hungary, Ireland, Israel, Italy, Japan, Luxembourg, Latvia, Mexico, Netherlands, Norway, New Zealand, Portugal, Slovak Republic, Slovenia, Spain, Sweden, United Kingdom, and the United States.

The construction of investment intensities implied a last methodological choice, i.e. the denominator to be used. Alternative indicators could have been constructed dividing ICT investments by output or by value added, rather than by non-residential gross fixed capital formation. The choice of non-residential capital formation, while common (e.g. OECD, 2018), implies that sectors displaying high overall investments (and therefore a high denominator) will score lower than sectors that have low overall investments (and therefore a low denominator). As pointed out in the main body of the study, this may downplay the intensity of manufacturing sectors more than services.

On the other hand, unreported analysis shows that services sectors (such as "Telecommunications", "IT" services, "Finance" or "Wholesale and retail") get on average the highest shares of total economy-wide ICT (tangible and intangible) investment.[26] As a consequence, a downward bias of manufacturing sectors in ICT intensities when these are constructed using non-residential GFCF at the denominator is in fact consistent with the relative importance of manufacturing sectors in overall ICT capital investment.

Purchases of ICT intermediates

This indicator measures the "direct" intermediate consumption of a sector *j* and country *c* of intermediates sold by the ICT-producing sectors in any of the 62 countries contained in the OECD Inter-Country Input-Output (ICIO) database, in deflated USD and as a percentage of sector *j*'s output.

The body of the study does not provide details on the deflation procedure. At present, the ICIO series are expressed in current prices. The purchases of intermediate ICT goods are here deflated with the price of output of the computer-producing sector (ISIC revision 4 sector 26) in the purchasing country, independently on the purchasing sector, as found in the OECD Structural Analysis (STAN) dataset. When a country-specific deflator is not available, the deflator is constructed as a simple average of the available G7 countries in the sample. For purchases of intermediate ICT services, the price of output of the "IT and other information services" producing sector (ISIC revision 4 sector 62 and 63) is chosen instead.

The sector-specific output deflator from STAN is used to deflate the output of the purchasing sector. When this is not available, the economy-level GDP deflator is used instead, as sourced from the OECD Annual National Accounts database. Countries for which this is not available either are dropped from the sample.

The resulting series are cleaned from outliers in the growth rate of both levels and intensities. The cleaning procedure targets year-on-year changes of the country-sector values of more than 50%. The value at time t giving origin to a 50% (or greater) change is replaced by the interpolation of values in t-1 and t+1, and treated iteratively. The resulting data series are finally extrapolated to populate the year 2015, and the years 2012-2014 for a subset of countries.

The final data covers at most 34 sectors for the following countries: Argentina, Australia, Austria, Belgium, Bulgaria, Brazil, Canada, Chile, Colombia, Costa Rica, Cyprus, Czech Republic, Germany, Denmark, Estonia, Finland, France, Greece, Croatia, Hungary,

Indonesia, Ireland, Iceland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Malta, Mexico, Netherlands, Norway, New Zealand, the People's Republic of China, Poland, Portugal, Romania, Russian Federation, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States.

As in the case of investment, an alternative indicator could be constructed using a different denominator, *i.e.* dividing purchases of intermediates by value added, rather than output. This is not a neutral choice, as it results in higher intensities for sectors which classify as low in value added but high in output (relative to other sectors), such as "Coke and refined petroleum products" or "Metal products". In analogy to what was proposed for investment intensities, a preference was given to a denominator (output) which is gross of the numerator, so that a higher numerator in a sector does not also mechanically translate in a lower denominator. A robustness specification using value added as denominator can be pursued in future research.

Robot use

An industrial robot is defined as "an automatically controlled, reprogrammable, multipurpose manipulator programmable on three or more axes, which can be either fixed in place or mobile for use in industrial automation applications" (definition ISO:8373, as reported by the IFR).

The baseline stock data reported by the IFR assumes a 100% depreciation rate at the end of the 12th year of life of the robot, a somewhat problematic hypothesis. A different measure of robots stock in the sector is thus constructed as follows. First, the initial stock is taken as reported in the IFR source data. Second, sales from subsequent years are added to the first-year stock. Third, an annual depreciation rate of 10% is applied. This is in line with the approach in Graetz and Michaels (2015).

For some countries, robot stock and sales are only available at the country-level for the first years of the sample period. In these cases, the shares of robot sales and stock across sectors in later years are used to allocate robots to these sectors in earlier years. A number of robustness tests are carried out, calculating the shares with all available years and with the last year of the sample period. The two methods yield minor differences in the time series. Therefore only shares from the last sample period are used, which improves the overall sectoral coverage of the data.

Employees data are primarily sourced from the OECD STAN database. If these are not available, information is supplemented with data from the Annual National Accounts Database, and with data from the WIOD 2013 Social Accounts (Timmer et al., 2015).

The dataset reports robot stocks from 1993, although for the purposes of this study, the time series is restricted to 2001-15. The final coverage for robot stock per employee includes at most 18 2-digit sectors for the following countries: Australia, Austria, Belgium, Brazil, Canada, Czech Republic, Germany, Denmark, Estonia, Finland, France, Greece, Hungary, Indonesia, India, Ireland, Italy, Japan, Korea, Latvia, Lithuania, Mexico, Netherlands, Norway, Poland, Portugal, Russian Federation, Slovak Republic, Slovenia, Spain, Sweden, Chinese Taipei, Turkey, United Kingdom, United States.

ICT specialists

A first definition of ICT specialist occupations results from the joint effort of the OECD and Eurostat (OECD, 2015a) and includes the five three-digit ISCO08 occupations mentioned above. The guiding principle for selecting these five occupations is their

primary involvement in the production of ICT goods and services (OECD, 2015a). Occupations which require specialised skills in using ICT only as a tool, not involved in the production of ICT goods and services, are thus excluded in this definition.

Greater accuracy in the definition of ICT occupations can be obtained on the basis of the tasks carried out by workers on the job, information which is gathered by the PIAAC survey. OECD (2015b) proposes three different methods to achieve this, to which the present study adds the measure of ICT task intensity of occupations from Grundke et al. (2017).

OECD (2015b) selects three distinct sets of tasks that are related to ICT usage on the job. The first set is called "use of communication and information search" (CIS) and includes the tasks "send/receive email" and "find work-related information on the Internet". The second set "use of office productivity software" (OPS) includes the tasks "use word processors" and "use spreadsheets". Both CIS and OPS require ICT-generic skills, but OPS involves a more sophisticated use of ICT and a higher level of ICT skills. The third set of tasks contains the "use of programming languages" which is a genuine ICT specialist task (OECD, 2015b). To select the top 20 occupations for each of the three sets of ICT related tasks (CIS, OPS and ICT specialists) across countries, the following procedure for each of the three sets of tasks is implemented. For each country and ISCO 2008 3-digit occupation, the share of workers conducting these tasks on a daily basis is calculated and occupations are classified according to this share. To select the top 20 occupations across countries, OECD (2015b) computes for each occupation the share of countries where this occupation is among the top 20 occupations. A similar intuition is used in this study to define how many occupations should be selected to estimate the sector employment intensity in ICT specialists (see here below).

The question nevertheless remains about the extent to which the sectoral intensities calculated over each of the 5 different methodologies exploited to define the ICT specialists' occupations are correlated. The answer to this question is a function of how many occupations are considered. When taking the top 5 occupations by ICT intensity as emerging from each of the considered methodologies, the correlations are high between OECD (2015a), the ICT specialist definition in OECD (2015b), and the ICT task intensity indicator from Grundke et al. (2017). The other indicators for a wider use of basic ICT technology in OECD (2015b) are also highly correlated between each other. When taking the top 5% occupations per each methodology, instead, correlations are less robust.

This study therefore investigates, for each methodology, whether "natural" cut off points exists, to identify the "ideal" number of occupations to be considered. This is achieved by computing the proportion of countries where a given occupation is within the first 20 ICT occupations, and then considering whether including the *n*-th occupation among the ICT ones changes this proportion significantly. For example, for the ICT specialist definition in OECD (2015b), the natural cut-off point appears after the first five occupations, whereby the percentage of countries where the occupation is among the first 20 ICT occupations drops from 78% to 67% when considering the sixth top ICT occupation. By considering these thresholds, the methodologies in OECD (2015a), OECD (2015b) for specialists, and Grundke et al. (2017) deliver the same 4 occupations, which therefore are the ones considered in this study to define the sector employment intensity in ICT specialists.

The top 5 occupations identified by each methodology proposed above are:

- Eurostat/OECD classification of ICT experts (OECD, 2015a):

  - 133- Information and Communications Technology Services Managers
  - 251- Software and Applications Developers and Analysts
  - 252- Database and Network Professionals
  - 351- Information and Communications Technology Operations and User Support Technicians
  - 352- Telecommunications and Broadcasting Technicians.
- OPS (OECD, 2015b) - Top 5 occupations based on proportion of countries where the occupation is among the top 20 occupations for daily users of Office Productivity Software (OPS):
  - 121- Business Services and Administration Managers
  - 122- Sales, Marketing and Development Managers
  - 241- Finance Professionals
  - 242- Administration Professionals
  - 261- Legal Professionals.
- CIS (OECD, 2015b) - Top 5 occupations based on proportion of countries where the occupation is among the top 20 occupations for daily users of Communication and Information Search (CIS):
  - 121- Business Services and Administration Managers
  - 122- Sales, Marketing and Development Managers
  - 133- Information and Communications Technology Services Managers
  - 241- Finance Professionals
  - 252- Database and Network Professionals.
- ICT specialists (OECD, 2015b) - Top 5 occupations based on proportion of countries where the occupation is among the top 20 occupations for daily users of programing languages:
  - 133- Information and Communications Technology Services Managers
  - 231- University and Higher Education Teachers
  - 251- Software and Applications Developers and Analysts
  - 252- Database and Network Professionals
  - 351- Information and Communications Technology Operations and User Support Technicians.
- ICT skill indicator (Grundke et al., 2017) - Top 5 occupations according to the ICT skill indicator computed by occupation across PIAAC countries (round 1 and round 2):
  - 133- Information and Communications Technology Services Managers
  - 242- Administration Professionals
  - 251- Software and Applications Developers and Analysts
  - 252- Database and Network Professionals
  - 351- Information and Communications Technology Operations and User Support Technicians.

ICT task intensity

The measure of ICT task intensity is based on the answers to the questions here below, as provided by workers surveyed in the OECD Programme for the International Assessment of Adult Skills (PIAAC). Data are available for 31 OECD and non-OECD countries, and more precisely: Australia, Austria, Belgium (Flanders), Canada, Chile, the Czech Republic, Germany, Denmark, Estonia, Finland, France, Greece, Ireland, Italy, Israel, Japan, Korea, Lithuania, the Netherlands, New Zealand, Norway, Poland, the Russian Federation (excluding Moscow), Singapore, the Slovak Republic, Slovenia, Spain,

Sweden, Turkey, the United Kingdom (England and Northern Ireland), the United States. Only a few of these are effectively used in the taxonomy, for coherence with other indicators. While individuals in PIAAC report their sector of occupation to the 4 digit ISIC revision 4 disaggregation level, the indicator is calculated at the 2-digit level instead, to be consistent with the other indicators in the taxonomy. The countries included in the taxonomy were all surveyed in 2012.

Table A3. Items included in the construction of the ICT task intensity indicator

Items included in the construction of the ICT skills indicator
G_Q05e Frequency of excel use
G_Q05g Frequency of programming language use
G_Q05d Frequency of transactions through internet (banking, selling/buying)
G_Q05a Frequency of email use
G_Q05c Frequency of simple internet use
G_Q05f Frequency of word use
G_Q05h Frequency of real-time discussions through ICT Computer
G_Q01b Frequency of Reading letters, emails, memosG_Q02a Frequency of Writing letters, emails, memos
G_Q06 Level of Computer Use required for the job
F_Q06b Frequency of working physically over long periods

Source: Grundke et al. (2017) compilation based on PIAAC data.

Online sales

Data are sourced from the Eurostat "Community survey on ICT usage and e-commerce in enterprises", in the Digital Economy and Society Statistics, Comprehensive Database by Eurostat.[27] It provides annual data on the use of ICTs and e-commerce in enterprises with 10 or more persons employed. A detailed definition of e-commerce transactions is available in OECD (2011b).

In order to be consistent with the ISIC revision 4 A38 sectoral classification adopted in this work, a correspondence table is applied to convert NACE revision 1.1 into NACE revision 2 sectors. Furthermore, values for aggregations of multiple 2-digit sectors are attributed to all relevant sub-sectors, where needed, which results in multiple A38 sectors displaying the same intensity. For instance, sectors 10, 13, and 16 show the same average value for 2001-03 (Figure 14), while sectors 69, 72, and 73 do so for both 2001-03 and 2013-15.

Lastly, data are cleaned from outliers based on Tukey's method (1.5 times the interquartile range, see Tukey, 1977), as well as interpolated or extrapolated whenever some years were missing for a given country-sector pair, but at least eight years are available.

Appendix Notes

[26] Relative low shares of ICT investment in manufacturing over total ICT investment can relate to (i) the small size of manufacturing output relative to services output in the considered economies; (ii) mismeasurement of ICT investment of manufacturing sectors, where ICT capabilities are embodied in machineries (e.g. robots) and hence recorded as non-ICT tangible investment; (iii) the propensity of sectors to outsource IT functions in the form of purchased services rather than investment; (iv) the high ICT capital investment effort of some services sectors, such as logistics, telecommunications, ICT services.

[27] See http://ec.europa.eu/eurostat/web/digital-economy-and-society/data/comprehensive-database, accessed in July 2017.